THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 56, No. 1, January 1977
Printed in U.S.A.

Design of Quantizers for Real-Time Hadamard-Transform Coding of Pictures

By F. W. MOUNTS, A. N. NETRAVALI, and B. PRASADA

(Manuscript received June 29, 1976)

A methodology is developed to obtain subjectively optimum quantizers for Hadamard-transformed still pictures. To exploit the perceptual redundancies that depend upon the local properties of the picture, a small block (2 × 2 × 2, horizontal-vertical-temporal) is used. A series of subjective tests was carried out to determine the visibility of impairment in the reconstructed picture when noise, which simulated the quantization noise, was added to the Hadamard coefficients in the transform domain (H-noise). A design procedure for quantizers was developed using these visibility functions. These quantizers minimize the "mean-square subjective distortion" (MMSSD) due to quantization noise. The resulting picture quality and entropy were compared with that of Max-type quantizers which minimize the "mean-square error" (MMSE). This comparison indicates that the MMSSD quantizers based on subjective visibility of the quantization noise are less companded than the MMSE quantizers. Also for the same number of quantization levels, pictures coded with MMSSD quantizers have better quality and less entropy than the pictures coded with minimum mean-square quantizers.

I. INTRODUCTION

A methodology is developed in this paper to establish fidelity criteria that characterize human observers' perception of noisy transform-coded pictures and to obtain optimum quantizers for the transform coefficients based on these fidelity criteria. The perceptual effects of impairments introduced in the transform domain are, in general, quite different from the impairments introduced in the picture domain. Our experiments, which are performed with the Hadamard transform of a stationary picture, determine the visibility of impairments in the reconstructed picture when noise (H-noise), which simulates the quantization noise, is added to a Hadamard coefficient. Functions that give the appropriate

subjective weighting of the quantization noise as a function of the quantity to be quantized are derived from these experiments. These functions, called visibility functions, are used in a systematic way to design quantizers for PCM and DPCM coding of the transform coefficients. These quantizers are compared with minimum mean-square error (MMSE) quantizers, both in terms of picture quality and bit rates.

1.1 Relationship to previous work

Considerable attention has been paid to the transform domain in the recent work on picture coding.[1–14] Transform domain processing has several potential advantages. It produces less correlated (but not necessarily independent) coefficients. It redistributes the image energy so that a large amount of energy is packed in a few of the coefficients. Moreover, on inverse transformation at the receiver, both noise from quantization of coefficients and the channel errors get distributed over the block in a manner given by the inverse transform of a particular coefficient.

A number of different transforms have been investigated; among them are: Karhunen-Loeve, Fourier, Hadamard, Haar, cosine, and slant transform. There have been several attempts[5,15,16] to compare the various transforms and to find their relative merits for coding of pictures. Almost all of these comparisons have been with respect to the following three criteria: (*i*) the correlation between the coefficients, (*ii*) the mean-square approximation error caused by setting some of the coefficients to zero, and (*iii*) the computational complexity in obtaining transform coefficients from picture elements (pels) and vice versa. Perceptual factors and the dependence of the picture quality on the particular transform and the block size have not received the attention they deserve.

The irreversible processing of the transform coefficients, which determines the trade-off between picture quality and bit rate, has been performed in a number of ways; for example, (*i*) zonal sampling or masking, which drops some predetermined higher-order coefficients; (*ii*) threshold sampling, which drops those coefficients whose values are below a predetermined threshold (a certain amount of addressing information must be sent in using this technique); (*iii*) quantization of the coefficients—both amplitude (PCM) and differential amplitude (DPCM) quantization[9,14] have been considered. Most of the work on quantization of the coefficients has centered around minimization of mean-square error as a criterion in designing the quantization characteristics. Several assumptions on the probability of the coefficients have been made, including the familiar gaussian case,[15] to carry out this minimization. Exploitation of the psycho-visual properties of the viewer and the optimization of the quantizers for the best subjective quality of the picture

has often been mentioned in the literature; however, no systematic methods are available for achieving these.

Our work on obtaining the quantization characteristics may be compared with that reported by Landau and Slepian[1] and Tasto and Wintz.[4] For this reason, we give a brief review of their reports. In both only a single frame of picture data is used despite the fact that the quantization noise is more visible when a sequence of frames of the same scene is coded.

Landau and Slepian considered both Karhunen-Loeve and Hadamard basis vectors for the linear transformation and found that the Karhunen-Loeve transformation required solution of an almost degenerate eigenvalue problem. They then used Hadamard transformation with a 4×4 block. The number of quantization levels given to each of the first ten coefficients was approximately proportional to the variance of that coefficient, and the last six coefficients (H_{11} to H_{16}) were dropped. The first coefficient was quantized by a 64-level uniform quantizer. Coefficients H_2 through H_{10} were quantized with quantizers having a companding characteristic given by a function of the form $y = k\sqrt{x}$.

Two arguments led Landau and Slepian to this quantization strategy. Firstly, since the variances of the lower coefficients are in general larger, coding them more accurately reduced the mean-square error. Secondly, the higher coefficients tend to be large in the busy regions of the pictures, where the viewers have more tolerance to amplitude errors. Thus, they used in an empirical way the consideration of a characteristic of the viewer as well as the statistics in the design of quantizers. They carried out over 100 experiments in which the decision levels and the representative levels of the quantizers were changed. However, since the number of choices is so large, their search could not be exhaustive and, therefore, their quantizers are the best only among those that they investigated.

Tasto and Wintz proposed an encoder using a 6×6 adaptive Karhunen-Loeve transform whose coefficients are quantized by what the authors call a "subjectively" optimized system of quantizers. This is done by first starting with a quantizer that minimizes the mean-square quantization error and then changing it by a trial-and-error procedure to obtain the "best" picture quality in the authors' judgment. The "best" is again from among those encountered in the trial-and-error procedure. They also conducted subjective rating experiments to compare the performance of the minimum mean-square quantizers with the "best" quantizers.

1.2 Basic objectives and approach

Our basic objectives are to obtain fidelity criteria in the transform domain which incorporate psycho-visual properties, and to develop

systematic methods for the optimum design of coders based on these fidelity criteria. As mentioned above, perceptual properties of the human viewer have not been given sufficient importance in the transform-coding literature and, consequently, good models do not exist to explain the subjective effects of the quantization errors in the coefficients when the coefficients are inverse transformed to obtain the picture element (pel).

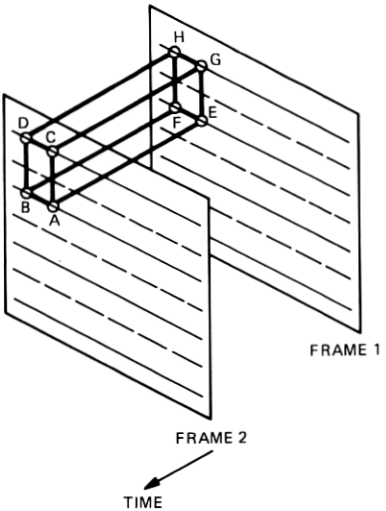
In the pel domain, some efforts[17–19] have been made to measure properties of human vision in psychophysical experiments and then utilize these to design coders. It is not easy to extend or utilize these techniques for the transform domain where we deal with blocks of pels instead of one pel at a time. Imperfect reproduction of coefficients of the block distributes distortion over the entire block upon inverse transformation.

To take advantage of both the perceptual and statistical properties, some of the factors one has to study are:

(*i*) Spread of the quantization error by inverse transformation.
(*ii*) Visibility of the quantization error in different coefficients.
(*iii*) Statistical decorrelation.
(*iv*) Probability distributions of the coefficients.

In this paper, we do not attempt to solve this general problem but restrict ourselves to nonadaptive coding of stationary pictures using a $2 \times 2 \times 2$ (horizontal-vertical-temporal) Hadamard transform. Although a temporal structure of the block is not relevant for still pictures, it will be used in the next phase of our work which will treat coding of a sequence of pictures. The Hadamard transformation has been chosen for its simplicity in implementation. The objective in choosing a small block is to exploit the perceptual redundancies which depend on local properties of the picture. The small block ensures that the quantization noise can be placed in parts of the picture where it is least visible. However, it does result in some loss of coding efficiency on statistical grounds. To compensate at least partially for this, we also discuss the differential coding of the first transform coefficient H_1.

In Fig. 1 the definition of the Hadamard coefficients for the block size $2 \times 2 \times 2$ is given. H_1 is the sum of the element brightnesses within the block. H_2 is the sum of the line differences within the block. H_3 is the planar difference. H_4 is the sum of the element differences within the block. It may be noted that for stationary pictures, H_5, H_6, H_7, and H_8 are all zero; further, any noise added to the first four coefficients gets repeated in the reconstructed signal at half the frame rate due to the block structure. As mentioned earlier, in coding frames of a single picture scene, the "nonmoving" noise patterns are, in general, less annoying than the moving noise patterns normally encountered in a television system



	A	B	C	D	E	F	G	H
H_1	+	+	+	+	+	+	+	+
H_2	+	+	−	−	+	+	−	−
H_3	+	−	−	+	+	−	−	+
H_4	+	−	+	−	+	−	+	−
H_5	+	+	+	+	−	−	−	−
H_6	+	+	−	−	−	−	+	+
H_7	+	−	−	+	−	+	+	−
H_8	+	−	+	−	−	+	−	+

Fig. 1—Definition of Hadamard coefficients. The pel positions are A, B, C, D, E, F, G, H. The Hadamard coefficients are H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8.

and, for this reason, a system was built to give a more realistic representation of television coding impairment. This system is described in Section II.

Our method for determining the visibility functions of the noise in H_2, H_3, and H_4 consists of the following. We add H-noise (which simulates the quantization noise) to a coefficient whenever its magnitude exceeds a threshold. This is done because each of these coefficients consists of difference quantities of pels and, therefore, may be expected to mask the noise as some function of their amplitude. For the DPCM coding of H_1, we add H-noise whenever the magnitude of the difference of H_1 from its previous block value is higher than a threshold. Again, this difference of H_1 can be taken as a measure of signal busyness. The effect of this H-noise impairment on the picture is then compared by the subject in an A-B test with simple additive white noise impairment of the picture. This method of judging pictures is similar to the one used by Candy and Bosworth.[20] The experimental method is discussed in detail in Section II.

1.3 Summary of results

The visibility functions for the following conditions have been measured: H_2-noise as a function of $|H_2|$; H_3-noise as a function $|H_3|$; H_4-noise as a function of $|H_4|$; and H_1-noise as a function of $|\Delta H_1|$, where ΔH_1 is the adjacent block difference in the horizontal direction.

The study of these visibility functions indicates that H_3 is the least

important coefficient and can be dropped entirely with little impairment. PCM quantizers with minimum mean-square subjective distortion (MMSSD) have been designed for H_2 and H_4 coefficients, and an MMSSD DPCM quantizer has been designed for H_1 using the corresponding visibility functions as the fidelity criteria. These quantizers have been implemented and have been compared in subjective tests with the corresponding quantizers optimized with respect to the minimum mean-square error criteria. Details of this approach and the results are given in the subsequent sections.

II. EXPERIMENTAL SYSTEM

The experimental system described in this section has been designed with considerable flexibility as a vehicle for future research. The system has real-time capabilities for adaptive and nonadaptive Hadamard transform coding of a $2 \times 2 \times 2$ block of pels.

2.1 System block diagram

A block diagram of the experimental system is shown in Fig. 2. The video signal is generated by a vidicon camera scanned with 271 lines interlaced 2:1. The video signal has a bandwidth of 1 MHz and is sampled at the Nyquist rate. Each picture sample is PCM encoded with amplitude accuracy of 8 bits per pel.

A frame memory is incorporated in the system to accommodate the transform block. Alternate frames of the digitized pictures, say the odd frames, are stored in the frame memory via data select switch 1. Memory 1 consists of two line delays and four small delays for linking the data from the present and previous frames. It ensures, during even frames, simultaneous presentation of all the elements from the two frames that comprise the data block to the Hadamard transform logic. It may be noted that the system is designed for spatially overlapped block processing. The output corresponding to nonoverlapping blocks is selected by memory 2 for decoding and experimentation. The spatially overlapping blocks are suitable for the study of various kinds of predictive encoders. This facility is also very useful for a flicker-free display of the Hadamard coefficients on a television screen.

The Hadamard transform circuit is a serial and parallel combination of adders and subtracters to implement the canonic forms shown in Fig. 1. In the processor circuit, the magnitudes of all the coefficients are rounded off to the eight most significant bits, which are used for further processing. This rounding off does not produce any visible impairment on inverse transformation. Capabilities exist in the processor circuit to insert eight independent quantizers, one for each coefficient.

For the subjective experiments, the processor circuit permits two

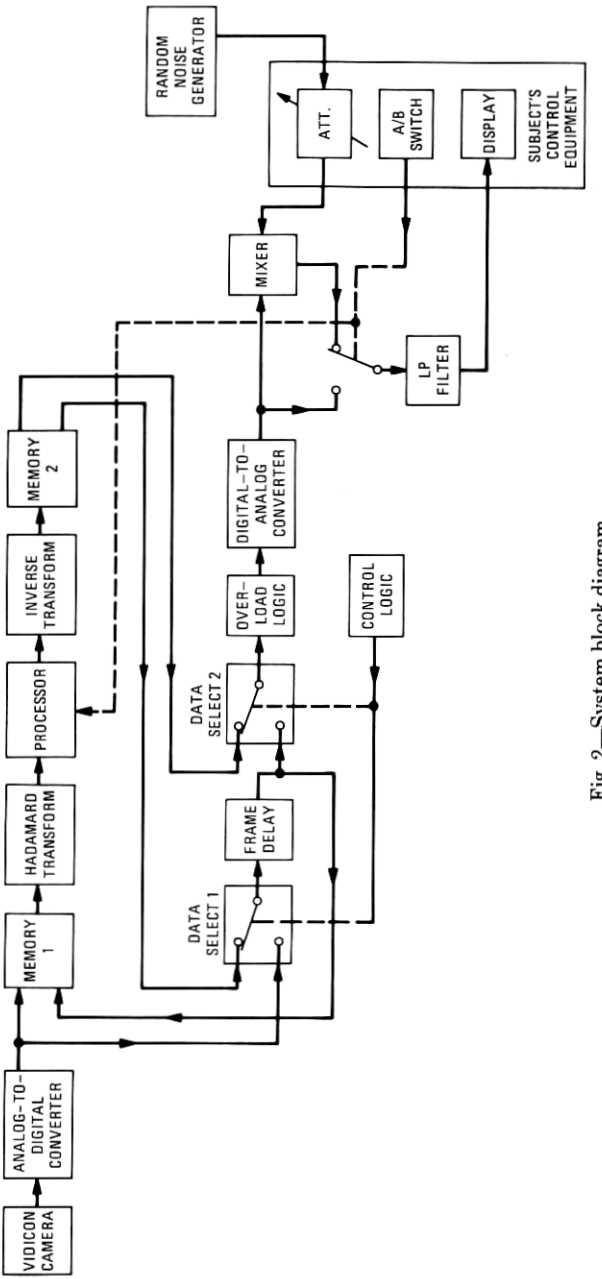


Fig. 2—System block diagram.

modes of operation controlled by the A/B switch. The details of the subsystem are shown in Fig. 3. The A/B switch is under the control of the subject. In the A mode, unimpaired coefficients are fed to the inverse-transform circuit. This provides the original picture in the reconstructed signal domain. In the B mode, a controlled amount of pseudo-random noise is added to one of the coefficients only when the magnitude of a control signal (which is the coefficient itself in this diagram) exceeds some reference threshold. This noise, which we call the H-noise, is generated at the sample rate by an 8-bit pseudo-random generator having a period of 2^{15} words and which is not synchronized with the line or the frame rate of the picture. An amplitude limiter controls the magnitude of the noise to the level set by the experimenter. The sign bit for the noise word is obtained from the output of a white noise source, and has equal probability of being a "0" or a "1".

Since the addition of pseudo-random noise results in doubling of the maximum amplitude of the noisy coefficient, the sum of the coefficient and noise, and the other coefficients, are divided by 2 prior to inverse transformation to prevent overload.

The inverse transformation network is similar to the transformation network and is used to reconstruct simultaneously all of the pels of the block.

It may be recalled that the alternate frames (odd frames) of the input are stored in the frame memory via data select switch 1. The recon-

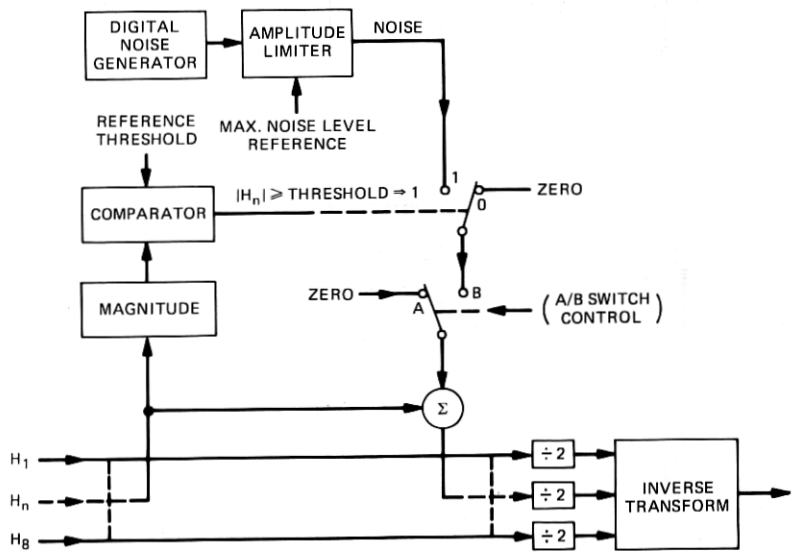


Fig. 3—Noise adding circuits. Coefficients are divided by 2 to prevent overload in the inverse transform function after noise addition.

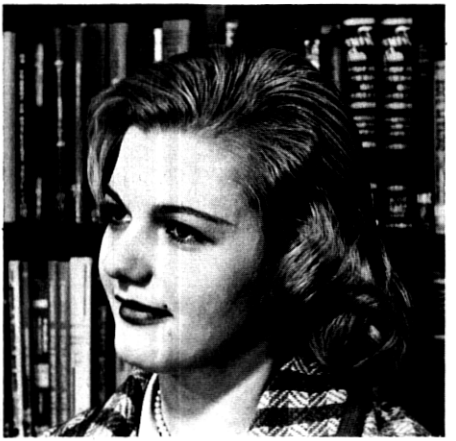


Fig. 4—Original picture used for subjective tests.

structed pels corresponding to the even frames are stored in the proper time slots in the frame memory by data select switch 1. Thus, the frame memory contains both processed and unprocessed data and is utilized fully. Data select switch 2 ensures that the reconstructed pels corresponding to the even frames that are stored in the frame memory are fed to the digital-to-analog converter in the proper time sequence.

The original picture used for the subjective tests is shown in Fig. 4. The scanned and filtered version (by a 1-MHz *Picturephone*® filter) is shown in Fig. 5a. Figure 6 shows the picture of the coefficients using overlapping blocks. Figure 6a shows coefficient H_1, which is essentially a "block-low-pass-filtered" version of the picture and preserves much of the picture information. On the other hand, Figs. 6b (H_2 coefficient), 6c (H_3 coefficient), and 6d (H_4 coefficient) show a variety of edge information.

2.2 Experimental details

The experimental setup for determining a visibility function is shown in the simplified block diagram in Fig. 7. The experimenter adds H-noise to a selected coefficient whenever the absolute value of the coefficient exceeds a threshold. The amount of noise and the threshold are varied. This is presented as condition B to the subject. Condition A is the unimpaired picture plus white noise. By turning an attenuator knob, the subject can control the amount of white noise added to the unimpaired picture. He can switch between conditions A and B by the A/B switch provided. An experiment consists of the subject changing the attenuator until he finds the pictures in the switch positions A and B to be subjec-

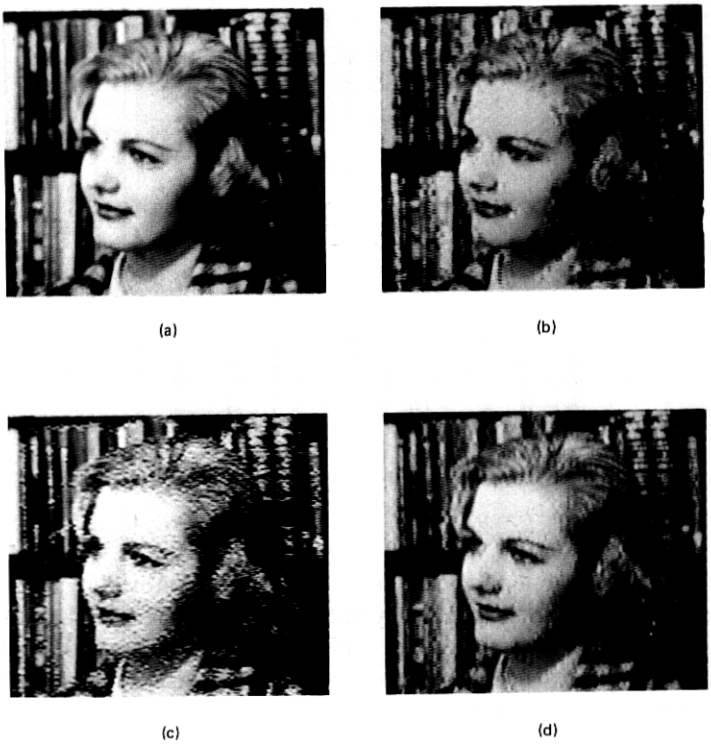


Fig. 5—Filtered test picture and H-noise added pictures. (a) Filtered test picture (1-MHz *Picturephone*® filter). (b) Picture with noise added to H_1. (c) Picture with noise added to H_2. (d) Picture with noise added to H_4.

tively equivalent. The subject can switch between A and B conditions as often as he likes and can look at the test conditions as long as he likes. When he arrives at the subjective equivalence, he gives the attenuator reading to the experimenters on an intercom. He is then given the next test condition. In one sitting, a subject makes 28 judgments of which the first four are considered as training. The remaining 24 are recorded as data. The experiment is also characterized by the following:

(*i*) The picture has 271 lines, interlaced 2:1 at 30 frames per second.
(*ii*) The visible portion of the picture is about 13 cm × 12 cm.
(*iii*) High light brightness is 74 foot-lamberts.
(*iv*) Low light brightness is 4 foot-lamberts.
(*v*) Room illumination is 57 foot-candles.

The scan lines of the Conrac monitor were broadened to correspond to the *Picturephone* display tube. Subjects were seated at a distance of

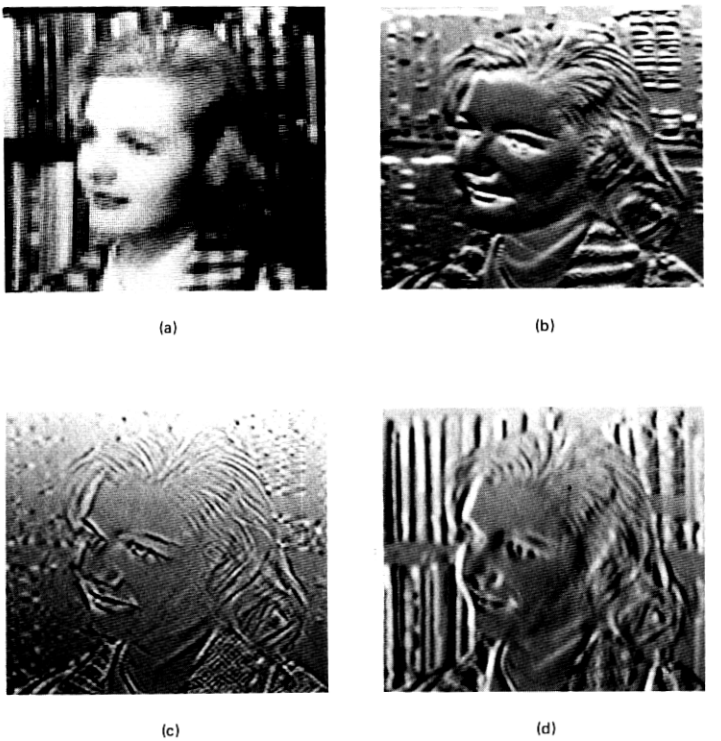


Fig. 6—Pictures of coefficients. (a) Picture of H_1 coefficient (with no output filter). (b) Picture of H_2 coefficient (with output filter). (c) Picture of H_3 coefficient (with output filter). (d) Picture of H_4 coefficient (with output filter).

about 80 cm from the monitor. All of the six subjects used had experience in judging coded television pictures.

III. TEST DATA AND ANALYSIS

Results of a typical subjective test are shown in Fig. 8. In this case, the absolute value of H_4 was compared to a threshold and the noise (H-noise) was added to H_4. In this figure, H_4-noise is plotted in dB on the X-axis and the "equivalent white noise" is plotted on the Y-axis. Each datapoint is an average of the readings obtained from six subjects. Under the assumption that the equivalent white noise (V_W) is proportional to the H-noise (V_H), the results for each threshold should fall on a 45-degree straight line. The lines drawn in Fig. 8 are the best unity-slope straight lines obtained by the least square fitting to the datapoints. Figures 9 and 10 show similar data for H_2, and H_1, respectively. In the case of H_1, $|\Delta H_1|$ is compared to a threshold. Notice that in each case the quantity that

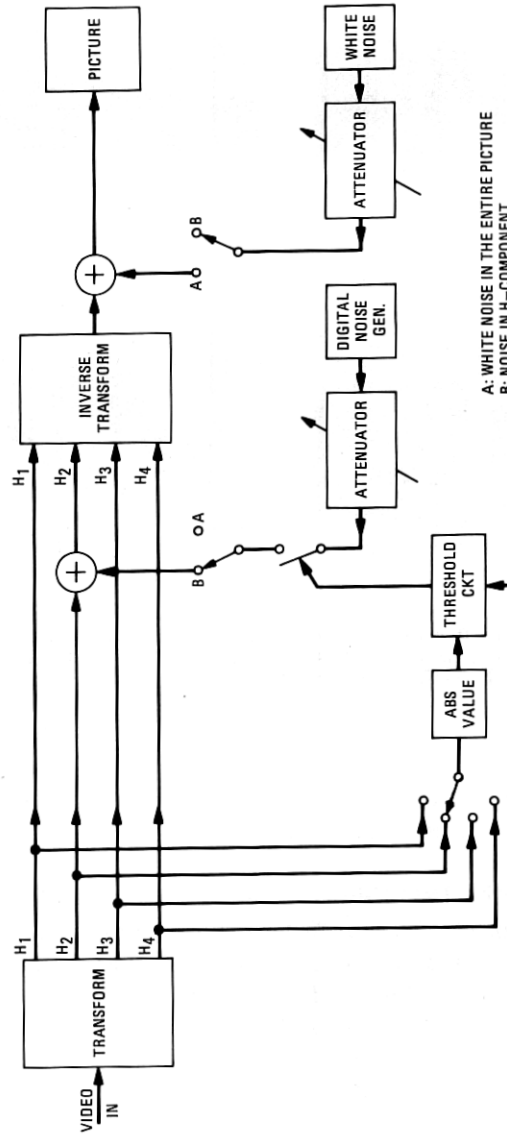


Fig. 7—Experimental setup. Noise is added to H_2 coefficient as a function of its magnitude.

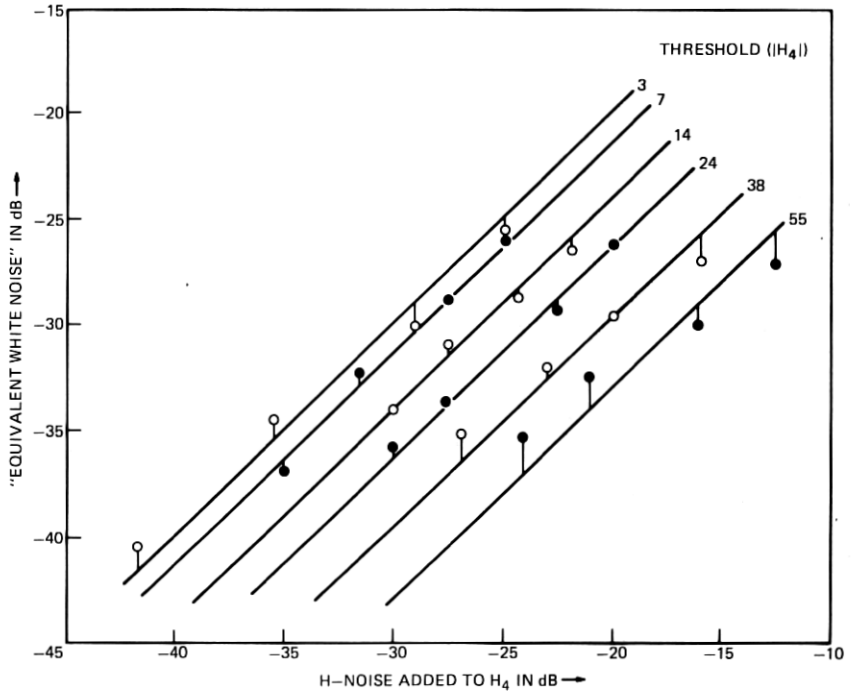


Fig. 8—Plot of "equivalent white noise" vs H_4 noise for different thresholds on $|H_4|$.

is compared to a threshold is a measure of busyness of the picture in a local area. The locations in the picture where noise is added and its appearance are dependent upon the quantity that is compared to a threshold and the coefficient to which the noise is added.

The pictures with H-noise impairments are shown in Fig. 5. The H-noise added in each of the three pictures has a peak value of 100 units* (signal range is 0 to 255 units). In Fig. 5c, noise is added to H_2 in all blocks which have $|H_2|$ more than five units, whereas in Fig. 5d, noise is added to all blocks in which $|H_4|$ is more than five units. While H_2 is the line difference, Fig. 5c has noise whenever an edge has a sufficiently large component along the horizontal direction, whereas Fig. 5d has noise whenever an edge has a sufficiently large component along the vertical direction. Also notice the difference in the appearance of the noise. H_2-noise is much more noticeable than H_4-noise. Figure 5b shows noise

* This is much more than the noise used for any test condition, but has been used to demonstrate the effects in a photograph.

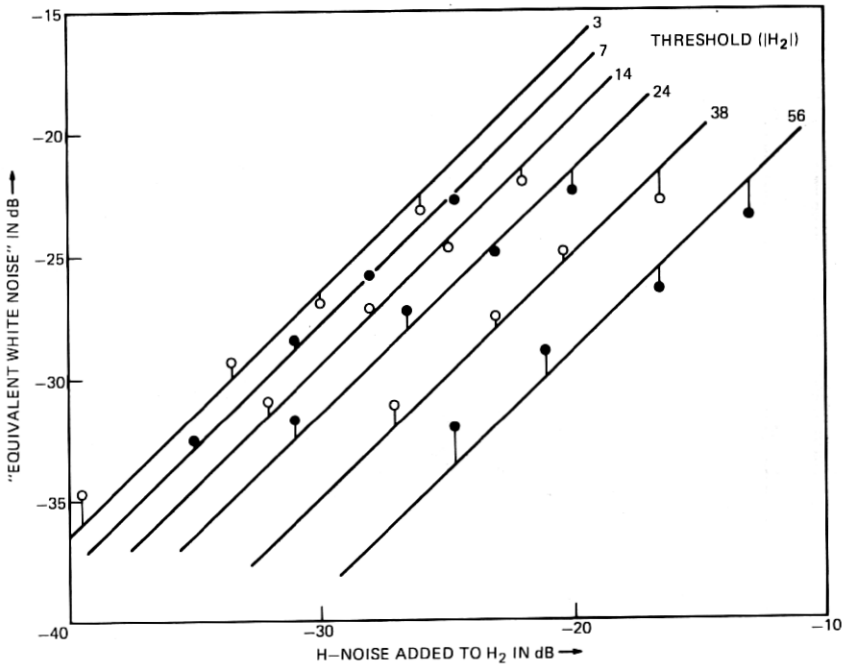


Fig. 9—Plot of "equivalent white noise" vs H_2 noise for different thresholds on $|H_2|$.

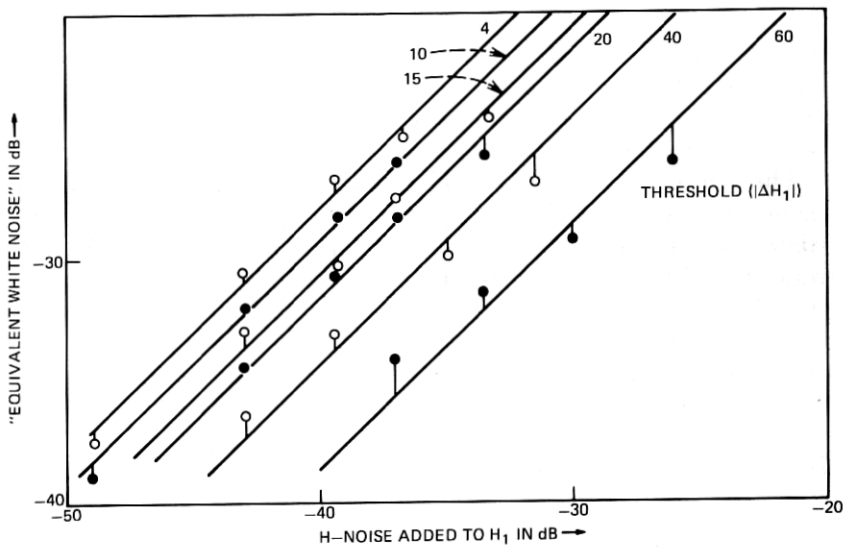


Fig. 10—Plot of "equivalent white noise" vs H_1 noise for different thresholds on $|\Delta H_1|$.

of 100 units added to H_1 whenever $|\Delta H_1|$ is more than 10 units. Here again, noise gets added to all blocks having horizontal interblock edges. Also the pattern generated by H_1 noise is much more objectionable than the pattern generated either by H_2-noise or H_4-noise.

The relationship of proportionality between the equivalent white noise and the H-noise for a threshold z is written in the form:

$$V_w = F(z)V_H, \tag{1}$$

where z can take on the value of $|H_2|$, $|H_3|$, $|H_4|$, or $|\Delta H_1|$. The constant of proportionality $F(z)$ is the equivalent white noise power when a unit H-noise is added to the particular coefficient for all blocks of the picture, where the magnitude of the corresponding coefficient ($|\Delta H_1|$ in the case of H_1) is greater than or equal to the threshold z. We next assume the additivity of the equivalent white noise power with respect to the coefficient value; i.e., if the equivalent white noise power when a unit of H-noise is added to H_2 and $T_1 \leqq |H_2| < T_2$ is V_{w1}, and the equivalent white noise power when a unit of H-noise is added to H_2 and $T_2 \leqq |H_2| < T_3$ ($T_1 < T_2 < T_3$) is V_{w2}, then the equivalent white noise power when a unit of H-noise is added to H_2 and $T_1 \leqq |H_2| < T_3$ is $(V_{w1} + V_{w2})$. Under this assumption, $F(z)$ can be written as an elemental sum of the equivalent white noise powers. Thus,

$$F(z) = \int_z^\infty f(x)dx, \tag{2}$$

where $f(z)$ is called the visibility function.

Using this procedure, visibility functions were computed. They are shown in Fig. 11. Notice that we have assumed that the occurrences of positive and negative coefficients (ΔH_1, H_2, H_3, H_4) are similar, and the noise visibility does not depend upon the sign of the coefficient. This results in the visibility functions being symmetrical about zero. The value of the visibility function shows the relative importance of the various transform coefficients. The larger the value, the more important is the coefficient. In general, the visibility functions decrease as a function of their arguments. This is a combined effect of several factors, such as (*i*) the decrease in the number of blocks having large coefficient values ($|\Delta H_1|$ in the case of H_1), (*ii*) the dependence of the perception of noise on the magnitudes of the coefficients (which correspond to the sharpness of the boundary in the pel domain), and (*iii*) the contextual importance of the specific regions of the picture.

Psycho-visual techniques which measure the detectability of perturbations in the neighborhood of edges,[18,21–23] and the just noticeable differences in the amplitudes of edges have been widely applied to DPCM coding. Since these deal with over-simplified stimuli and surround and are almost always detection experiments, their use in picture coding may

not always result in better coders. In any case, these techniques cannot be easily applied in the transform domain because we are not dealing with the single pels but with blocks comprising pels from more than a single line and frame. Also, the perturbations must be introduced in the transform coefficient, whereas the annoyance to the perturbations must be judged in the pel domain.

Our approach, which obtains the visibility functions as outlined above, has the following limitations:

(*i*) Since the visibility functions are tied to the picture content, they admittedly vary from picture to picture, especially if the picture content is changed significantly. They also depend upon the class of viewers and the viewing conditions. Thus, any optimization based on the visibility functions is strictly applicable to a restricted situation. This is under-

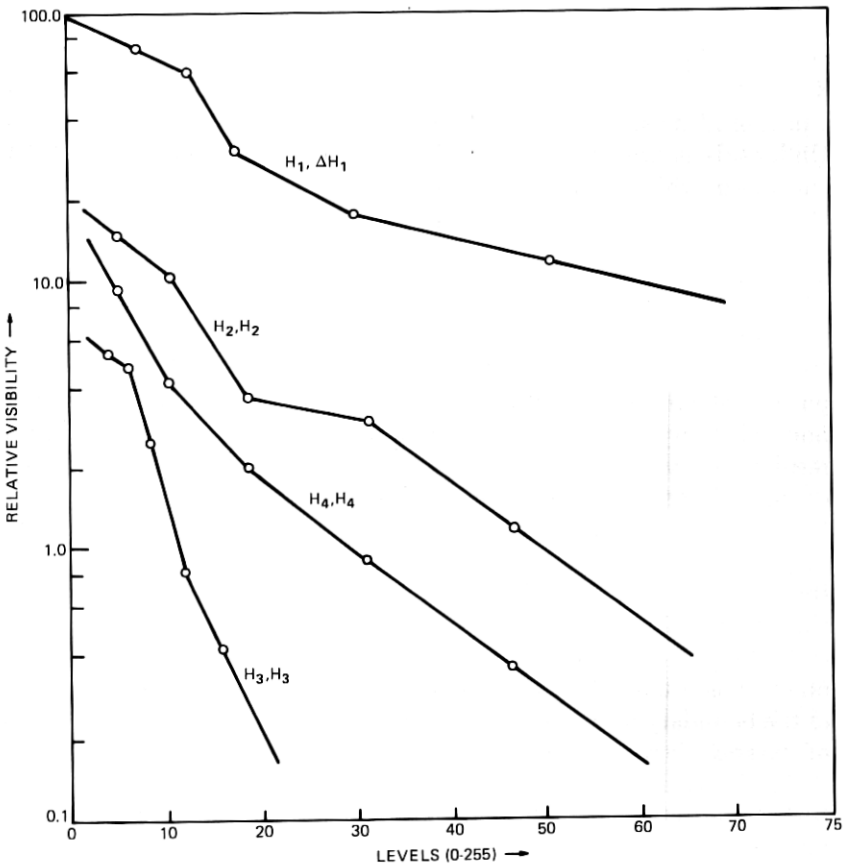


Fig. 11—Plot of visibility functions. Notation H_i, H_j indicates noise added to H_i, when H_j is thresholded. The H_j threshold level is shown on the X-axis.

standable since the human perception does indeed vary with the picture content, the viewing conditions, and the particular viewer. We demonstrate in Section IV that the results we obtained using these visibility functions are not overly sensitive to the picture content and are reasonable for a class of pictures rather than a particular picture.

(*ii*) The simulation of the quantization noise by the H-noise is fairly accurate for H_2, H_3, and H_4. However, in the case of H_1, DPCM techniques are used to code ΔH_1. The changes in the appearance of noise as a function of threshold are not completely reflected in the measurements; i.e., while the noise that is added to H_1 does look like granular noise at low thresholds, it does not look like slope overload at high thresholds. Also, at high thresholds, the noise is added to fewer blocks in the picture and the appearance of such an impaired picture is different from the appearance of a white noise impaired picture. Therefore, in some cases subjective equivalence is hard to achieve.

(*iii*) It would be better if the perceptual, statistical, and contextual effects were explicit in the visibility function and could be controlled separately. Unfortunately, such is not the case.

(*iv*) It is seen from eq. (2) that the process of obtaining the visibility function involves differentiation of the data, which is known to introduce some noise. By adding H-noise to a coefficient when the quantity to be compared to a threshold is within a small range of values, it is possible to avoid this differentiation.

IV. RESULTS

In this section, we present certain conclusions drawn from the visibility functions and then describe their application to the design of quantizers for the coefficients. Visibility functions shown in Fig. 11 clearly show the relative importance of various coefficients. H_1 is the most important, H_2 is the next, followed by H_4, and H_3 is the least important. The visibility of H noise depends upon the patterns associated with a particular coefficient. These patterns depend upon the inverse transform and are shown in Fig. 12 for H_2, H_3, and H_4, respectively. In each case, noise of a given amount is added to one of the coefficients, and the background is assumed to be flat. The higher the spatial frequency of the pattern, the lower the visibility of the noise. Thus, H_2 noise is more visible than H_4 noise because the interlace gives the H_2 noise pattern lower spatial frequency than the H_4 noise pattern.

4.1 Visibility of H_1 noise

An experiment was performed to utilize the well-known property of the human eye that the brightness discrimination decreases as the brightness level increases, called Weber's law in the psycho-visual lit-

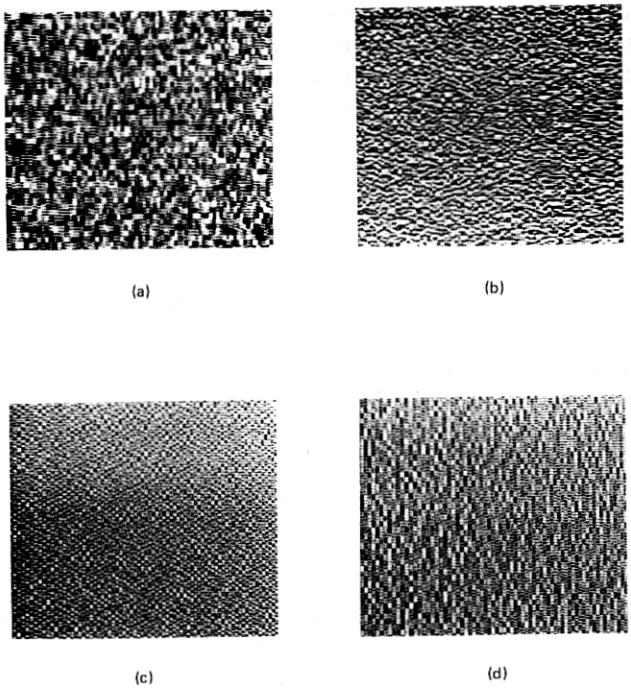


Fig. 12—Pictures of noise patterns for coefficients. (a) Picture of noise added to H_1 on a flat background. (b) Picture of noise added to H_2 on a flat background. (c) Picture of noise added to H_3 on a flat background. (d) Picture of noise added to H_4 on a flat background.

erature[24–26] (see Ref. 27 for a recent application of Weber's law for picture coding in the pel domain).

In this experiment, noise was added to H_1 as a function of H_1, since H_1 corresponds to the average brightness in the block. The results of this subjective experiment showed large variations from observer to observer. When the data for the observers was averaged, there was no significant variation in the visibility of noise as a function of H_1. This could be due to the following: (*i*) If the gamma of the display tube used was not unity, it would have partially compensated for Weber's law effects. (*ii*) We were working with the head and shoulders view of a person. In general, for such a picture, the highlights are on the forehead or the cheek of the person. These regions are contextually very important causing the visibility of noise to be high. (*iii*) The picture we used was such that the low-light areas had more spatial detail than the highlight areas; thus, the latter two effects may have compensated for the Weber's law.

Measurements of the gamma of the monitor indicate that the visibility function in this case cannot be fully explained on the basis of the compensation of Weber's law by the gamma of the display tube. It seems that at least for this class of pictures, namely the head and shoulders view of a person, the advantage that could be gained by the utilization of Weber's law is compensated for by the other effects.

4.2 "Frozen" vs "unfrozen" noise visibility

It may be recalled that the experiment on visibility was done using a block size of $2 \times 2 \times 2$. In this case, any noise added to the first four coefficients remained unchanged for two frames. The noise in the coefficients in this case may be called the "frozen" noise because it remains unchanged for two frame periods. An experiment was performed to determine the visibility functions for H_4 coefficient for a block size of 2×2 (horizontal-vertical). Since all the experiments have been carried out with a stationary picture, the only difference between the experiment with the block size of 2×2 and a block size of $2 \times 2 \times 2$ is the coefficient noise. For the block size of 2×2, the coefficient noise changes from frame to frame and is called "unfrozen" noise. Figure 13 shows the visibility functions for H_4 with "frozen" and "unfrozen" noises. Although the visibility functions of "unfrozen" noise are generally a little lower than that of "frozen" noise, due to lower temporal frequency, the differences are small.

In Section III, it was mentioned that the visibility functions can be used as fidelity criteria for the design of quantizers. We describe below how these results are used to design quantizers.

4.3 PCM coding of H_2, H_3, and H_4

It is assumed that little interaction exists between the Hadamard coefficients, so that the quantization transfer characteristics for the coefficients can be obtained independent of each other. It is recalled from Fig. 11 that H_3 was the least important coefficient and, therefore, it was decided to drop the transmission of H_3 altogether.

Minimum mean-square error quantizers are obtained by minimizing the mean-squared quantization error. If N is the number of levels, and $P_{H_k}(\cdot)$ is the histogram for $|H_k|$, then we minimize the distortion D given by

$$D = \sum_{j=1}^{N} \int_{X_j}^{X_{j+1}} (|H_k| - Y_j)^2 \, P_{H_k}(|H_k|) \, d(|H_k|), \qquad k = 2,3,4, \quad (3)$$

with respect to $\{X_j\}$, $j = 2, \ldots, N$ and $\{Y_j\}$, $j = 1, \ldots, N$. This gives us the well-known Max quantizer.[28] MMSE quantizers are obtained by weighting the quantization error according to the frequency of its oc-

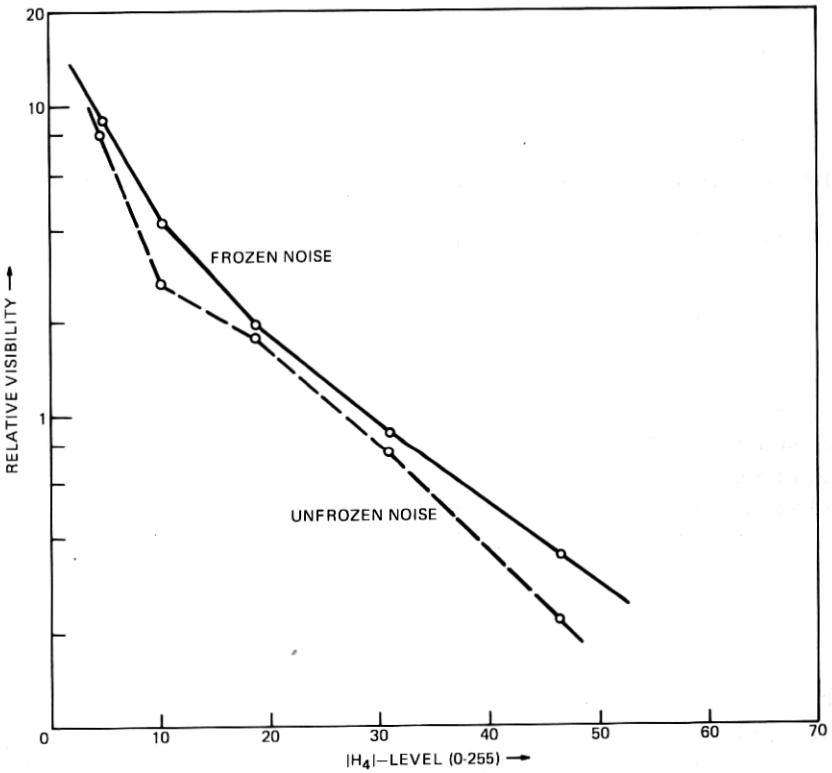


Fig. 13—Visibility function for "frozen" and "unfrozen" H_4 noise.

currence. Minimum mean-square subjective distortion quantizers, on the other hand, weight the quantization error according to its subjective visibility. This can be achieved by substituting $f_{H_k}(\cdot)$ for $P_{H_k}(\cdot)$ in the expression for the distortion. The term $f_{H_k}(\cdot)$ is the visibility function for the coefficient H_k. Standard programming techniques were used to minimize the distortion D in both cases.

The histograms for $|\Delta H_1|$, $|H_2|$, $|H_3|$, and $|H_4|$ are shown in Fig. 14. In general, these decrease faster than the visibility functions. This is exemplified in Fig. 15 in which the histogram and the visibility function for ΔH_1 are plotted with the same scale on the X-axis. We shall see later that this fact results in larger companding of the MMSE quantizers than the MMSSD quantizers and, consequently, poor reproduction of busy areas of a picture.

Typical quantizer characteristics are shown in Fig. 16. The MMSE quantizer is more companded than the MMSD quantizer. Note also that the dynamic range of the MMSE quantizer is smaller. The performance

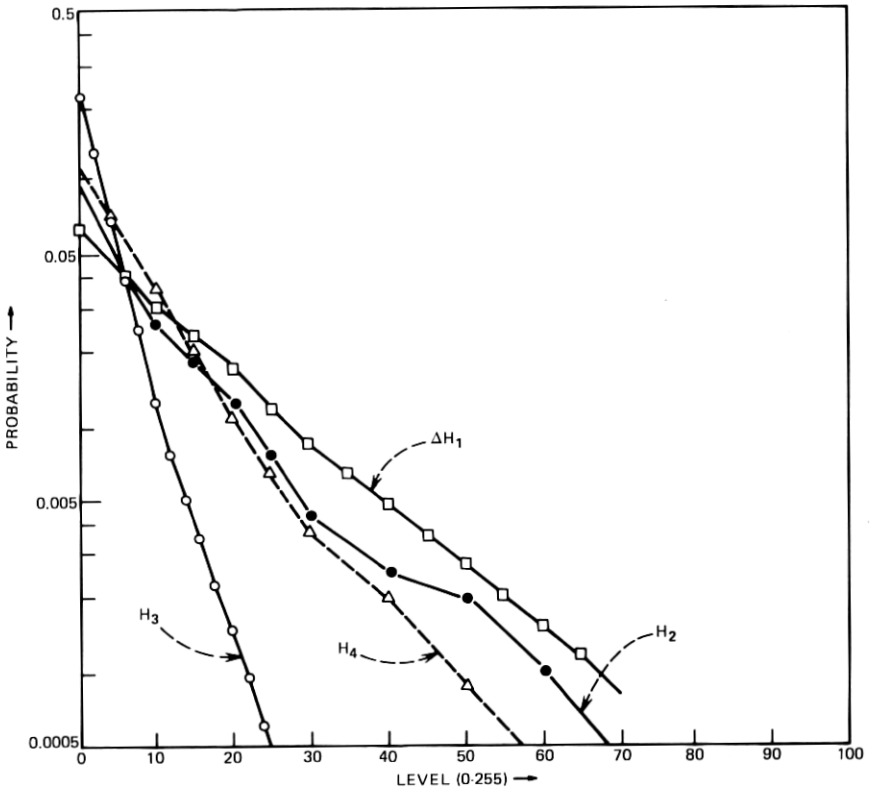


Fig. 14—Histograms of coefficients. All the histograms are assumed to be symmetric about zero.

of these two types of quantizers (MMSSD and MMSE) was compared in an A-B test with different numbers of levels. Figure 17 shows the results of such a test for the coefficient H_4. In this test, MMSE quantizers with levels 5 to 8 were compared in terms of picture quality with MMSSD quantizers with levels 3 to 9 using a random pairing by six skilled subjects. The numbers in the table indicate the percentage of observers who preferred the MMSSD quantizers over the MMSE quantizers. The picture coded with the 5-level MMSSD quantizers was preferred by 100 percent of the subjects over the 6-level MMSE quantizer. Figure 18 shows similar comparisons for the quantization of H_2. Here again, for the same number of levels, the picture quality using the MMSSD quantizers is always better than using MMSE quantizers. Moreover, picture quality using the 6-level MMSSD quantizer and 7-level MMSE quantizer is equivalent. Figures 19 and 20 show the entropy of the quantized output using both the MMSSD and the MMSE quantizers having levels 3 to 8 for H_4 and H_2, respectively. In the case of H_4, the difference between the entropies of

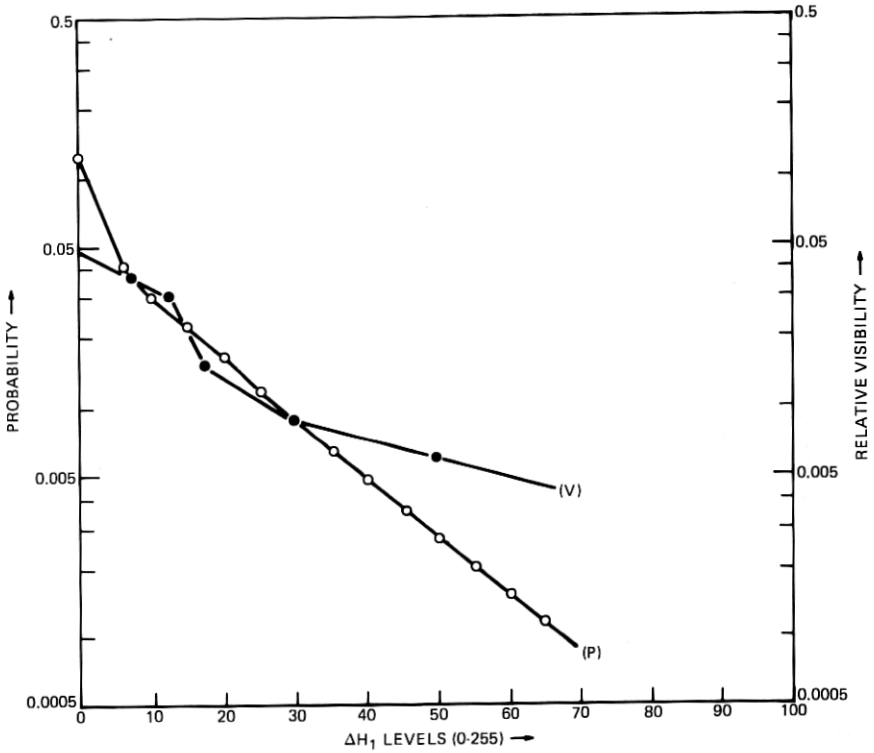


Fig. 15—Comparison of probability density (P) and visibility (V) for H_1 noise.

the output of the MMSE and the MMSSD quantizer for the same number of levels is about 0.2 bit. Since the picture quality with the 7-level visibility quantizer was better than with the 8-level MMSE quantizer, the gain by the use of the MMSSD quantizer is of the order of 0.5 to 0.6 bit for the transmission of H_4. Similar remarks can be made about the quantization of H_2.

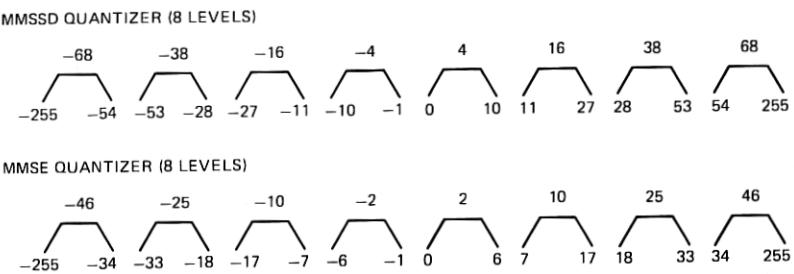


Fig. 16—Typical (MMSSD and MMSE) quantizer characteristics for H_4 coefficient. Notation $\overset{z}{\frown}$ implies that all input levels between x and y (including x,y) are represented as z.

MMSE QUANTIZERS Number of Levels	MMSSD QUANTIZERS Number of Levels 3	4	5	6	7	8	9
5	50	33	83	100	100	100	100
6	50	33	100	100	100	100	100
7	33	33	33	87	100	100	100
8	17	17	50	33	100	100	100

Fig. 17—Comparison of picture quality of MMSSD and MMSE quantizers of different levels for H_4.

4.4 Coding of H_1

Unlike the H_2, H_3, and H_4 coefficients, the H_1 coefficient is not a difference signal. It represents the average brightness within the block and thus carries the low-frequency information which should be coded relatively precisely. Uniform PCM coding of H_1 requires 7 to 8 bits for good picture quality. As mentioned before, efforts to compand the PCM quantizer by using the Weber's law effect were not very successful. Therefore, it was decided to DPCM encode H_1. Since the block size used is small, there is substantial correlation between the H_1 values of adjacent blocks. This was exploited by using a DPCM coding of H_1 with horizontally adjacent blocks for prediction. The quantizers for such a DPCM coder are obtained from the visibility function of H_1 under the control of $|\Delta H_1|$ in a manner similar to the above by minimizing the mean-square subjective distortion due to the quantization noise. The resulting quantizer scales are companded due to the monotonic decrease of the visibility function with respect to ΔH_1, as shown in Fig. 11. Quantizer scales have also been obtained by minimizing the mean-square quantization error.* As noted before, MMSE quantizer scales are more companded and have less dynamic range compared to the MMSSD quantized scales. Using these two types of scales, experiments have been performed

* Although the visibility function and the histograms are obtained from the difference signal $|\Delta H_1|$, and the quantity that is quantized is the differential signal (i.e., the difference between the present H_1 and the coded value of H_1 from the previous block), it is expected that the quantizer characteristics will not change appreciably by using difference instead of the differential signal in eq. (3).

MMSE QUANTIZERS: Number of Levels	MMSSD QUANTIZERS: Number of Levels 3	4	5	6	7	8	9
5	0	17	100	100	100	100	100
6	0	0	0	100	100	100	87
7	0	0	0	50	100	100	100
8	0	17	0	17	33	100	87

Fig. 18—Comparison of picture quality of MMSSD and MMSE quantizers of different levels for H_2.

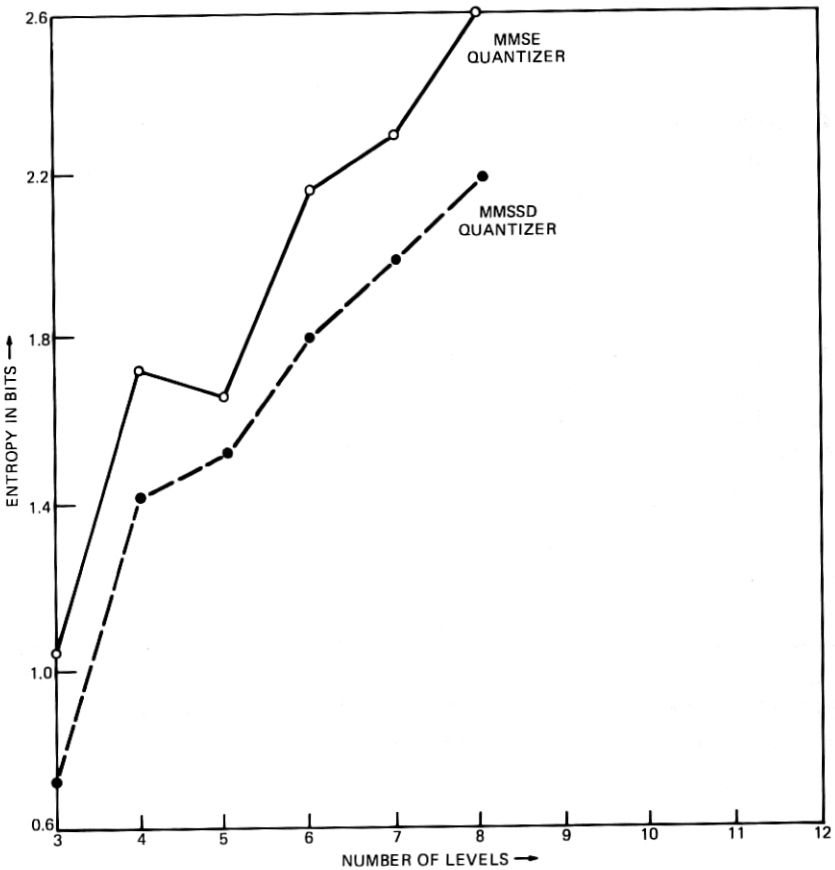


Fig. 19—Plots of entropy of outputs of MMSSD and MMSE quantizers of different levels for H_4.

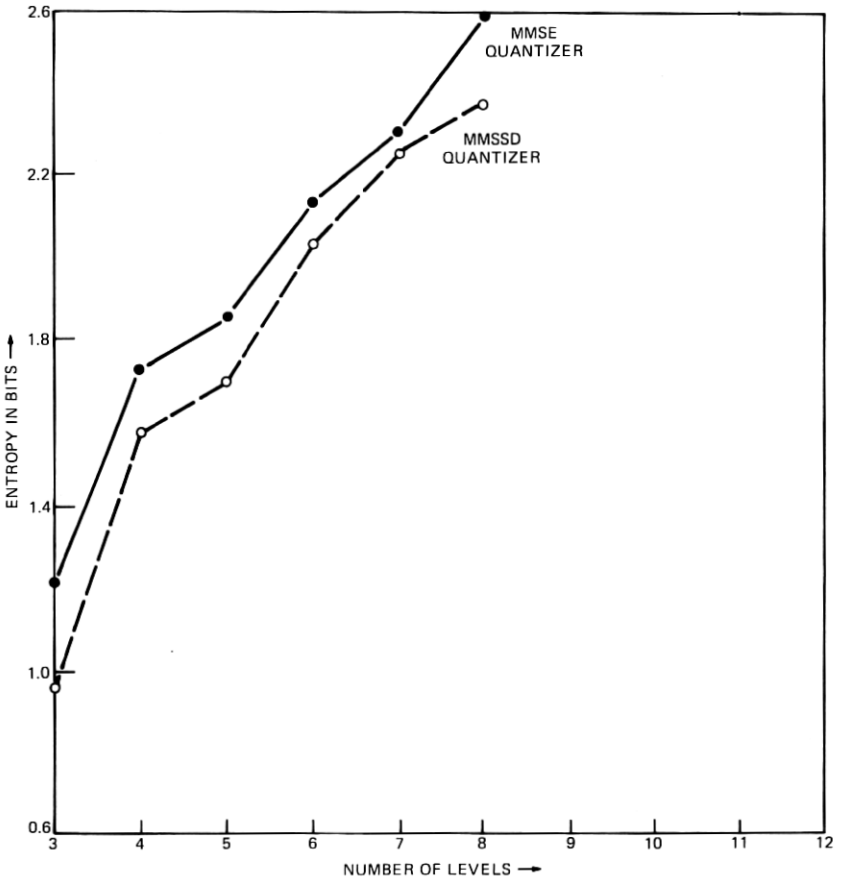


Fig. 20—Plots of entropy of outputs of MMSSD and MMSE quantizers of different levels for H_2.

to compare the picture quality for the same number of levels. The results of such a comparison are shown in Fig. 21. It is seen that, for the same number of levels, all the subjects preferred the picture coded with the MMSSD quantizers over the picture coded with MMSE quantizers. Moreover, picture quality using a 24-level MMSSD quantizer is equivalent to the picture quality using a 30-level MMSE quantizer. Entropies of the quantized signal with MMSSD and MMSE quantizers of different levels are shown in Fig. 22. Here again, visibility quantizers perform better than MMSE quantizers by about 0.35 bit per block for the same number of levels. Picture quality using a 24-level MMSSD quantizer can be produced by an MMSE quantizer with an increase in entropy of 0.6 bit per block. It is worth noting tht, due to the DPCM coding of H_1, the bits required for H_1 could be almost halved. However, H_1 still remains more important than H_2 and H_4 and requires more bits for satisfactory transmission.

MMSE QUANTIZERS Number of Levels	MMSSD QUANTIZERS Number of Levels 20	22	24	26	28	30	32
26	50	100	100	100	100	100	100
28	17	100	100	83	100	100	100
30	17	50	50	67	50	100	100
32	0	50	67	67	67	100	100

Fig. 21—Comparison of picture quality of MMSSD and MMSE quantizers of different levels for DPCM coding of H_1.

4.5 Combined quantization of all coefficients

Combined quantization of all the coefficients requires investigation of the optimal number of quantizer levels to be given to each one of them. In the case of gaussian random vectors using Karhunen-Loeve transformation and mean-square error criterion, optimal bit allocation for the various transform coefficients is well known.[29,30] However, in our case, none of these assumptions are strictly valid. In fact, our assumption that the optimum quantizer characteristics for different coefficients can be obtained independently is not strictly true and, for this reason, we tried to evaluate the picture quality by quantizing all the coefficients. By trial and error, a near-perfect picture was produced by using 36, 13, and 7 quantization levels for ΔH_1, H_2, and H_4 respectively, and by dropping H_3. This resulted in a total entropy of about 2.17 bits per pel. In single-frame photographic reproduction, no difference could be observed between the coded picture and the low-resolution original shown in Fig. 5a. Several other "head and shoulders" type of pictures were coded using the same combination of levels. Although, in each case the picture appeared to have a reasonable quality, the visibility of the quantization and the resulting picture quality varied slightly. This implies that the quantizers we obtained by optimizing the visibility of the quantization noise for one particular picture were not overly sensitive to variation in picture content.

V. CONCLUSIONS

A systematic method for quantizing Hadamard coefficients has been given. This method gives the best quantizers in a subjective and probabilistic sense. We have compared the resulting quantizers with MMSE quantizers and found the MMSSD quantizers to be better both in terms

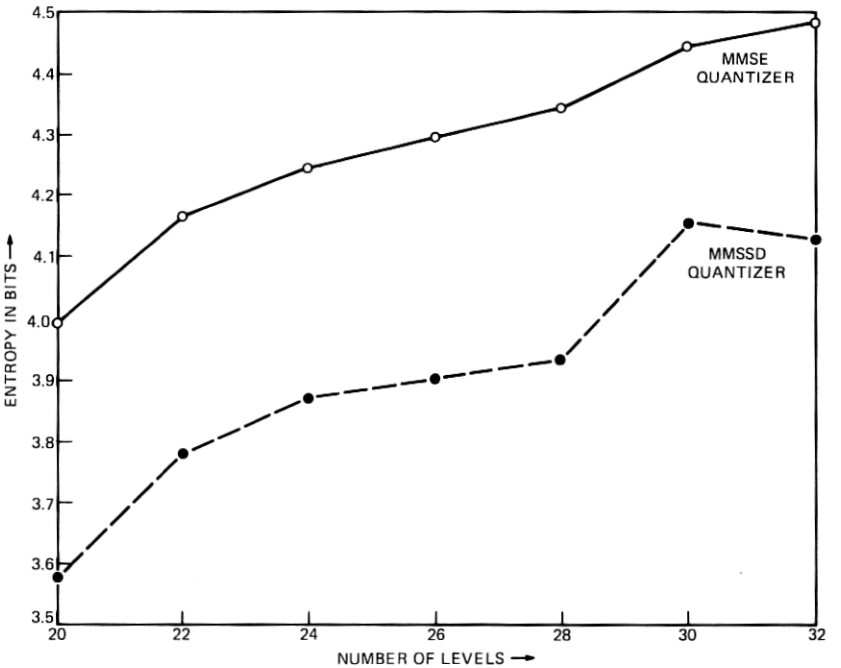


Fig. 22—Plots of entropy of outputs of MMSSD and MMSE quantizers of different levels for DPCM coding of H_1.

of the subjective picture quality and entropy. We do not imply that there are no better quantizer than the MMSSD quantizer, since by taking many other factors into consideration, one could come up with a better quantizer. We do find that the minimum visibility quantizers are optimum with respect to our model and the approach used for weighting the quantization noise.

Investigations are in progress for adaptive and predictive coding of the coefficients; our findings will be reported in a future paper.[31]

VI. ACKNOWLEDGMENT

We would like to thank J. C. Candy, B. G. Haskell, and K. A. Walsh for help in various phases of this work. Thanks are also due to E. G. Bowen, R. H. Bosworth, R. C. Brainard, E. F. Brown, J. E. Berrang, and R. L. Schmidt for generously giving their time as subjects.

REFERENCES

1. H. J. Landau and D. Slepian, "Some Computer Experiments in Picture Processing for Bandwidth Reduction," B.S.T.J., *50,* No. 5 (May–June 1971), pp. 1525–1540.
2. W. K. Pratt and H. C. Andrews, "Fourier Transform Coding of Images," Proc. Hawaii Intern. Conf. System Sciences, January 1968.
3. W. K. Pratt, J. Kane, and H. C. Andrews, "Hadamard Transform Image Coding," Proc. IEEE, *57,* No. 1 (January 1969), pp. 58–68.

4. M. Tasto and P. A. Wintz, *Picture Bandwidth Compression by Adaptive Block Quantization,* Technical Report TR-EE-70-14, July 1970, Purdue University, Lafayette, Indiana.
5. P. A. Wintz, "Transform Picture Coding," Proc. IEEE, *60,* No. 7 (July 1972), pp. 809–820.
6. J. W. Woods and T. S. Huang, *Picture Bandwidth Compression by Linear Transformation and Block Quantization,* 1969 Symposium on Picture Bandwidth Compression, Massachusetts Institute of Technology, New York: Gordon and Breach, 1972.
7. G. B. Anderson and T. S. Huang, "Piecewise Fourier Transformation for Picture Bandwidth Compression," IEEE Trans. Commun. Technol., *COM-19* (April 1971), pp. 133–140.
8. A. Habibi and P. A. Wintz, "Image Coding by Linear Transformation and Block Quantization," IEEE Trans. Commun. Technol., *COM-19* (February 1971), pp. 50–60.
9. A. Habibi, "Hybrid Coding of Pictorial Data," IEEE Trans. Commun. Technol., *COM-22,* No. 5 (May 1974), pp. 614–620.
10. S. C. Knauer, "Video Compression Algorithm for Hadamard Transform Processing," IEEE Trans. Electromagnet. Compat., *EMC-18,* No. 1 (February 1976), pp. 28–36.
11. *Applications of Walsh Functions,* Conference Proceedings and Symposium, 1971 through 1974, IEEE Electromagnetic and Compatibility Group.
12. B. J. L. Fino, "Recursive Generation and Computation of Fast Unitary Transforms," Memo No. ERL-M-514, Electronics Research Laboratory, University of California, Berkeley, January 1974.
13. R. M. Haralick and K. Shanmugam, "Comparative Study of a Discrete Linear Basis for Image Data Compression," IEEE Trans. Systems, Man and Cybernetics, *SMC-4,* No. 1 (January 1974), p. 16.
14. M. Ishii, "Picture Bandwidth Compression by DPCM in the Hadamard Transform Domain," Fujitsu Sci. & Tech. J., *10,* No. 3 (September 1974), pp. 51–65.
15. H. C. Andrews, *Computer Techniques in Image Processing,* New York: Academic Press, 1970, Chapter 7.
16. J. Pearl and H. C. Andrews, "Performance Measures for Transform Data Coding," IEEE Trans. Commun. Technol., *COM-20* (June 1972), pp. 411–415.
17. W. F. Schreiber, "Picture Coding," Proc. IEEE, *55* (March 1967), pp. 320–330.
18. J. O. Limb, Ref. 17, pp. 364–379.
19. T. L. Mannos and D. J. Sakrison, "The Effects of Visual Fidelity Criterion on the Encoding of Images," IEEE Trans. Inform. Theory, *IT-20,* No. 4 (July 1974), pp. 525–536.
20. J. C. Candy and R. H. Bosworth, "Methods for Designing Differential Quantizers Based on Subjective Evaluations of Edge Busyness," B.S.T.J., *51,* No. 7 (September 1972), pp. 1495–1516.
21. G. F. Newell and W. K. E. Geddes, "The Visibility of Small Perturbations in Television Displays," BBC Engineering Division, Research Department, Report T-106, 1963.
22. H. Harms and E. Aulhorn, "Studien Über den Grenzkontrast. I Mitteilung. Ein neus Grenzphänomen.," Grafes Arch. Opthal., *157,* 1955, pp. 3–23.
23. D. Teller, "The Influence of Borders on Increment Thresholds," Ph.D. Thesis, Department of Psychology, University of California, Berkeley, 1965.
24. S. S. Stevens, "Handbook of Experimental Psychology," New York: Wiley, 1951.
25. K. Hacking, "The Relative Visibility of Random Noise Over the Grey-Scale," J. Brit. Inst. Radio Eng., *23,* No. 4 (April 1962), pp. 307–310.
26. T. G. Stockham, "Image Processing in the Context of a Visual Model," Proc. IEEE, *60,* No. 7 (July 1972), pp. 828–842.
27. F. Kretz, "Subjectively Optimal Quantization of Pictures," IEEE Trans. Commun., *COM-23* (November 1975), pp. 1288–1292.
28. J. Max, "Quantizing for Minimum Distortion," IEEE Trans. Inform. Theory, *IT-6* (March 1960), pp. 7–12.
29. J. J. Y. Huang and P. M. Schultheiss, "Block Quantization of Correlated Gaussian Random Variables," IEEE Trans. Commun. Syst., *CS-11* (September 1963), pp. 289–296.
30. A. Segall, "Bit Allocation and Encoding for Vector Sources," IEEE Trans. Inform. Theory, *IT-22* (March 1976), pp. 162–169.
31. F. W. Mounts, A. N. Netravali, and B. Prasada, "Some Experiments in Adaptive and Predictive Hadamard Transform Coding of Pictures," unpublished work.