A Subjective Evaluation of Noise-Shaping Quantization for Adaptive Intra-/Interframe DPCM Coding of Color Television Signals

BERND GIROD, HÅKAN ALMER, LEIF BENGTSSON, BJÖRN CHRISTENSSON, AND PETER WEISS

Abstract—**Nonuniform quantizers for just not visible reconstruction errors in an adaptive intra-/interframe DPCM scheme for component-coded color television signals are presented, both for conventional DPCM and for noise-shaping DPCM. Noise feedback filters that minimize the visibility of reconstruction errors by spectral shaping are designed for Y, $R-Y$, and $B-Y$. A closed-form description of the "masking function" is derived which leads to the one-parameter "b quantizer" characteristic. Subjective tests that were carried out to determine visibility thresholds for reconstruction errors for conventional DPCM and for noise shaping DPCM show significant gains by noise shaping. For a transmission rate of around 30 Mbits/s, reconstruction errors are almost always below the visibility threshold if variable length encoding of the prediction error is combined with noise shaping within a 3:1:1 system.**

I. Introduction

FOR a transmission of color television signals at broadcast quality, a digital broad-band channel with a rate of around 30 Mbits/s is likely to be internationally standardized soon. In order to reduce the rate of 216 Mbits/s of the digital studio 4:2:2 TV signal [1] to the desired transmission rate, data compression is required which, however, should result in a picture quality that is superior to the quality provided by today's PAL, SECAM, or NTSC systems. A first step of data compression can be a sampling rate conversion to 10.125 MHz for the luminance signal Y and to 3.375 MHz for the color difference signals $R-Y$ and $B-Y$, resulting in a 3:1:1 system, as it has similarly been proposed, e.g., in [2]–[5] and investigated in [6], [7]. Furthermore, the horizontal and vertical blanking intervals need not be transmitted, which leaves approximately 2 bits for each sample in the average.

Adaptive intra-/interframe DPCM with variable-length coding of the prediction error signal is a source coding scheme that has been favored for 30 Mbit/s transmission of broadcast quality TV signals in several publications, e.g., [2]–[5], [8]–[10]. Additionally, a combination with noise shaping has been suggested to be an efficient means of reducing the visibility of reconstruction noise [11]. What picture quality results from a combination of the proposed algorithms at the given transmission rate, however, has not been investigated yet. In order to answer this question, quantization characteristics that lead to just not visible reconstruction errors in conjunction with the coding scheme envisaged have to be determined in subjective tests.

For intraframe DPCM systems without noise shaping, quantizers for just not visible errors have been presented in [12]–[14] for the luminance and in [15], [16] for the color difference signals. An investigation concerning the visibility of quantization errors for an adaptive intra-/interframe luminance DPCM coder has been reported by Westerkamp [17].

In this paper, we present quantization characteristics for the luminance and the color difference signals that lead to just not visible reconstruction errors in an adaptive intra-/interframe DPCM scheme. These quantizers have been determined by means of subjective tests. In order to evaluate the improvements that can be achieved by reconstruction noise shaping [11], we compare the quantization characteristics for just not visible errors for systems with and without noise shaping.

Section II very briefly reviews adaptive intra-/interframe DPCM and describes the details of the specific coding algorithm used in the subjective tests. In Section III, the fundamentals of noise shaping are summarized and noise feedback filters are designed for Y, $R-Y$, and $B-Y$ based on results from psychophysics literature. Section IV describes the subjective testing methodology that we used to determine quantizers for just not visible distortions. Sections V and VI state the results of our subjective tests, and give quantizers for just not visible quantization errors for the luminance and the color difference signals with conventional DPCM and with noise shaping DPCM. Finally, Section VII relates the subjective test results to bit rate for both fixed and variable wordlength encoding of the prediction error.

II. The Adaptive Intra-/Interframe DPCM Algorithm

In the following section, we will very briefly review adaptive intra-/interframe DPCM and describe the details of the coding scheme that has been used in our subjective tests.

Fig. 1 shows an adaptive intra-/interframe DPCM system. It is based on the idea that a television signal can be predicted efficiently if the predictor is switched between two modes.

- *Interframe Prediction:* In regions of the picture where the signal contains only small changes from frame to frame, the amplitude of the current sample can be predicted accurately from the corresponding sample in the previous frame [Fig. 2(a)], according to

$$\hat{S}_{\text{inter}} = S'_{20}. \tag{1}$$

- *Intraframe Prediction:* In regions of the picture where the signal contains rapid motion, the amplitude of the interframe prediction error is large, and a better prediction is obtained by a linear combination of the amplitudes of adjacent signal samples in the same field [Fig. 2(b)], according to

$$\hat{S}_{\text{intra}} = a_1 S'_1 + a_2 S'_2 + a_3 S'_3 + a_4 S'_4. \tag{2}$$

Adaptive intra-/interframe prediction has been discussed by many authors, e.g., [2], [4], [8], [9], [17], [18], [19]. Switching between intraframe prediction and interframe prediction can be controlled in two different ways.

1) With *feedback adaptation*, the predictor is switched

Paper approved by the Editor for Image Processing of the IEEE Communications Society. Manuscript received December 1, 1986. This paper was presented in part at the International Picture Coding Symposium, Stockholm, Sweden, June 1987.

B. Girod is with the Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, Universität Hannover, Germany.

H. Almer, L. Bengtsson, B. Christensson, and P. Weiss are with the Research Department, Swedish Telecommunications Administration, Stockholm, Sweden.

IEEE Log Number 8718639.

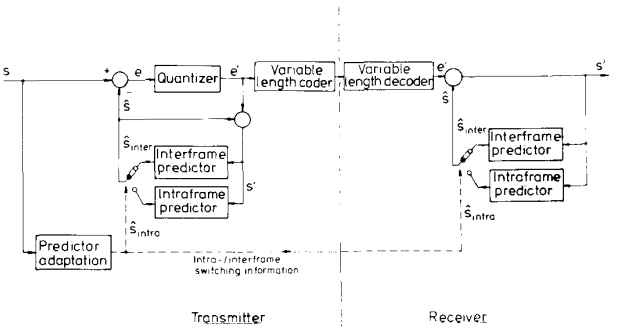
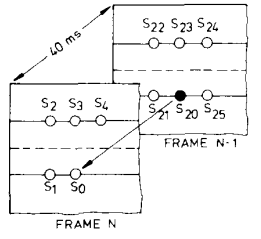
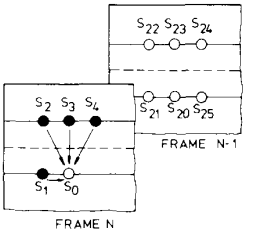


Fig. 1. Block diagram of an adaptive intra-/interframe DPCM system with feedforward adaptation.



$\hat{S}_{inter} = S'_{20}$

(a)



$\hat{S}_{intra} = a_1 S'_1 + a_2 S'_2 + a_3 S'_3 + a_4 S'_4$

(b)

Fig. 2. Samples used for interframe prediction (a) or for intraframe prediction (b) of the current sample S_0.

based on previously transmitted information only. As this information is equally available at transmitter and receiver, no additional adaptation information need be transmitted. The adaptation algorithm has to be implemented at both transmitter and receiver.

2) With *feedforward adaptation*, the predictor is switched based on data that are not available at the receiver. The adaptation algorithm is implemented at the transmitter only and the adaptation state has to be transmitted additionally to the prediction error signal. Different adaptation algorithms can be used with the same receiver as long as they obey an agreed format to signal the prediction mode. The system shown in Fig. 1 utilizes feedforward adaptation.

The coding efficiencies obtainable with either approach are approximately equivalent [19]. For feedforward adaptation, the prediction error entropy is smaller, which makes up for the additional adaptation information to be transmitted. In general, feedforward prediction adaptation is preferable since the receiver is less complex than for feedback adaptation, and the robustness against transmission errors is usually much better.

In our subjective tests, we used a feedforward adaptation scheme that signals the adaptation state on the basis of a fixed block structure [5], [9], [19]. Each block consists of nine horizontally adjacent luminance samples within one scan line and the corresponding three $R-Y$ samples and three $B-Y$ samples (Fig. 3). Switching is encompassed by a comparison of the accumulated absolute values of the intraframe and the interframe luminance prediction errors e [9], [19], i.e.,

$$\text{If } \sum_{\text{block}} |e_{\text{inter}}| < \sum_{\text{block}} |e_{\text{intra}}|, \text{ then } \hat{S} = \hat{S}_{\text{inter}}$$

$$\text{else} \qquad \hat{S} = \hat{S}_{\text{intra}}. \tag{3}$$

The adaptation mode is determined from the luminance signal only and then adopted for the color difference signals.

The quantization characteristic for just not visible quantization errors is quite dependent on the choice of prediction coefficients a_1, a_2, a_3, and a_4 (2) [12]. *Isotropic intraframe prediction* yields the coarsest quantizer for just not visible reconstruction errors [20], [21]. In order to keep reconstruction errors below the visibility threshold not in the average, but everywhere in the picture, the isotropic predictor minimizes the prediction error power at edges of most unfavorable orientation. In terms of prediction error entropy, the isotropic predictor performs reasonably close to the optimum for typical images, such that it is equally useful for fixed wordlength encoding and for variable wordlength encoding of the prediction error.

Table I lists the coefficients of isotropic intraframe predictors that are used in our codec for the luminance and for the color difference signals. It also states the horizontal and vertical bandwidths of the signals that have been assumed in the optimization of the isotropic predictor. The vertical bandwidth is characterized by the equivalent horizontal bandwidth in MHz for an interlaced 625 line/50 Hz system. Prediction coefficients were restricted to be multiples of 1/4. For coefficients that are multiples of 1/8, only a small improvement can be observed.

III. Design of a Noise-Shaping DPCM System

A. Reconstruction Noise Shaping by Additional Quantization Error Feedback

A standard DPCM system with memoryless quantizer usually produces white reconstruction noise [11]

$$n = S' - S. \tag{4}$$

Exceptions from the flat power spectrum are limit cycles of the DPCM coder that can occur for a constant input signal S or slope overload that can occur at large signal discontinuities.

The reconstruction noise is perceived by the human visual system with its specific transfer function that depends on both spatial and temporal frequency. Noise shaping fits the reconstruction noise spectrum to the frequency characteristic of the human visual system such that larger quantization errors are allowed for just not visible distortions [22], [23].

The most advantageous solution to modify the standard DPCM system towards a noise-shaping system is an additional quantization error feedback (Fig. 4) [11]. The quantization error q is calculated by subtraction of input and output of the quantizer, and is fed back to the input of the DPCM loop through a linear filter with transfer function $H(\omega_x, \omega_y, \omega_t)$. If the quantization error can be considered a statistically independent, additive noise source with power spectrum $Q(\omega_x, \omega_y, \omega_t)$, the power spectrum of the reconstruction error is

$$N(\omega_x, \omega_y, \omega_t) = Q(\omega_x, \omega_y, \omega_t) \cdot |1 - H(\omega_x, \omega_y, \omega_t)|^2. \tag{5}$$

$H(\omega_x, \omega_y, \omega_t)$ allows us to shape the reconstruction noise spectrum as desired.

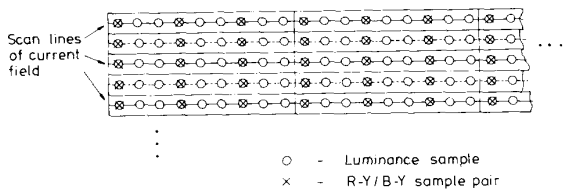


Fig. 3. Illustration of the block structure for predictor adaptation. For each block of nine horizontally adjacent luminance samples and the corresponding chrominance samples, a common prediction mode is used.

TABLE I
ISOTROPIC INTRAFRAME PREDICTORS FOR LUMINANCE AND COLOR DIFFERENCE SIGNALS OF A 3:1:1 SYSTEM

Signal	sampling frequency	Horizontal bandwidth assumed	Vertical bandwidth assumed	Prediction coefficients a_1	a_2	a_3	a_4
Y	10.125 MHz	5.0 MHz	4.5 MHz	1/2	0	1/4	1/4
R-Y	3.375 MHz	1.6 MHz	4.5 MHz	1/4	0	1/2	1/4
B-Y	3.375 MHz	1.6 MHz	4.5 MHz	1/4	0	1/2	1/4

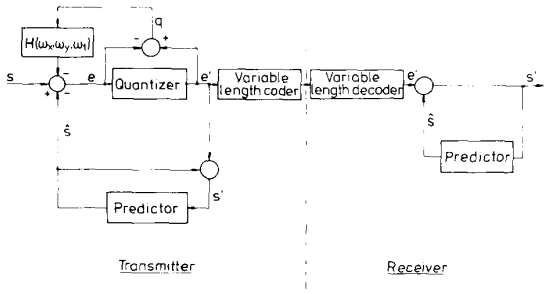


Fig. 4. Noise-shaping DPCM system with additional quantization error feedback.

The receiver of the DPCM system with additional quantization error feedback does not differ from the standard receiver (Fig. 4). This feature is highly desirable for future communication systems:

1) For program distribution, the receiver should be simple, while the transmitter may be complex.

2) Both standard DPCM transmitters and those with noise shaping can be operated in the same network.

B. Stability Considerations

Systems with quantization error feedback can oscillate under certain conditions even though their input signal is constant. For a DPCM system, these limit cycles can have consequences for the entropy of the prediction error. While the influence of noise shaping on the variance or on the entropy of the quantized prediction error e' usually is negligibly small [11], this is no longer true if the prediction error standard deviation is small compared to the distance between the inner quantizer representative levels.

A limit cycle that involves the N innermost representative levels of a quantizer could theoretically lead to a (memoryless) entropy of the quantized prediction error of $\log_2(N)$ bit/sample if all N levels occur equally frequent. The limit cycle behavior of the coder puts a lower limit on the data rate that can reliably be reached for very "easy" pictures or for very coarse quantizers. Aiming at a data compression to around 2 bits/sample, we can accept limit cycles involving up to three levels.

The following stability condition for a DPCM system with additional quantization error feedback is derived in [11] based on [24].

The maximum (minimum) value of a limit cycle oscillation in terms of the prediction error e is less (greater) than or equal to the maximum (minimum) value of e in the range where the quantization error characteristic q(e) = e' − e is greater than (less than) or equal to the stability border

$$q = (\overset{+}{-})m \cdot |e| \tag{6}$$

with

$$m = \frac{1}{\sum_i |a_i + h_i|}.$$

The a_i are the prediction coefficients and the h_i are the noise-shaping filter coefficients at corresponding locations i. This stability condition can be applied to the b quantization error characteristic that will be introduced in Section IV-A. The b quantizer is the coarsest quantizer with quantization errors

$$-b\sqrt{|e|} \le q(e) \le b\sqrt{|e|}. \tag{7}$$

The five inner representative levels of the b quantizer are 0, $\pm 2b^2$, $\pm 6b^2$ if the finite-wordlength signal representation in the DPCM loop is neglected. From (6), we conclude that for $m = 1$, which corresponds to conventional DPCM with prediction according to (1) or (2) with coefficients from Table I, a degenerate one-level limit cycle, i.e., a constant offset between S' and S, can occur. Limit cycles can involve only the inner three representative levels if $m > 1/2$, which holds for

$$\sum_i |h_i| < 1. \tag{8}$$

Besides small amplitude limit cycles, large amplitude oscillations can occur if the quantization error characteristic leaves the stability sector in the quantizer overload range. This problem can be overcome by very simple means. We do not feed back the quantization error through the noise feedback filter $H(\omega_x, \omega_y, \omega_t)$ whenever the prediction error is above (below) the maximum (minimum) representative level of the quantizer. This can be incorporated into the noise-shaping DPCM coder by substituting the explicit calculation of the quantization error (Fig. 4) by a look-up table (Fig. 5). The look-up table contains a modified quantization error characteristic as shown in Fig. 6.

C. Design of a Noise Feedback Filter for the Luminance Signal

How should $H(\omega_x, \omega_y, \omega_t)$ be chosen in order to minimize the visibility of the reconstruction noise n? This question cannot be answered without a visibility measure that is realistic on one hand and mathematically tractable on the other hand. $H(\omega_x, \omega_y, \omega_t)$ has an impact mainly on the spectral shape of the reconstruction noise, and thus a frequency weighted mean-square error measure

$$\begin{aligned} E &= \int_{\omega_x}\int_{\omega_y}\int_{\omega_t} |M(\omega_x, \omega_y, \omega_t)|^2 \\ &\quad \cdot N(\omega_x, \omega_y, \omega_t)\, d\omega_x\, d\omega_y\, d\omega_t \\ &= \int_{\omega_x}\int_{\omega_y}\int_{\omega_t} |M(\omega_x, \omega_y, \omega_t)|^2 Q(\omega_x, \omega_y, \omega_t) \\ &\quad \cdot |1 - H(\omega_x, \omega_y, \omega_t)|^2\, d\omega_x\, d\omega_y\, d\omega_t \end{aligned} \tag{9}$$

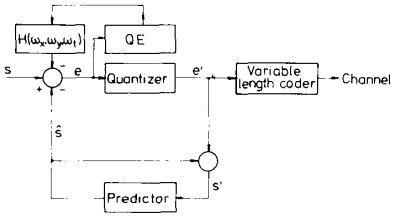


Fig. 5. Noise-shaping DPCM coder with additional quantization error feedback that is realized as a modified quantization error table *QE* (Fig. 6).

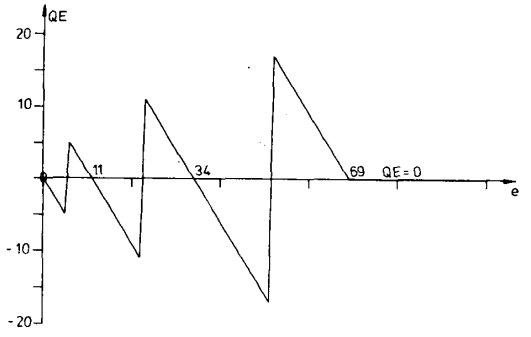


Fig. 6. Modified quantization error table *QE* in the noise-shaping system according to Fig. 5. Positive representative levels in this example are 0, 11, 34, and 69. *QE* suppresses the additional feedback of quantizer overload errors.

seems to be adequate. If we assume $Q(\omega_x, \omega_y, \omega_t)$ to be constant, visibility measure (9) can be substituted by

$$E' = \int_{\omega_x}\int_{\omega_y}\int_{\omega_t} |M(\omega_x, \omega_y, \omega_t)|^2 \cdot |1-H(\omega_x, \omega_y, \omega_t)|^2 \, d\omega_x \, d\omega_y \, d\omega_t. \quad (10)$$

$M(\omega_x, \omega_y, \omega_t)$ is the frequency response of the human visual system.

Measurements on the human visual system, as they have been conducted by Robson [25], Kelly [26], or Van Nes *et al.* [27], show that its modulation transfer function (MTF) for luminance sinewave stimuli possesses a spatiotemporal bandpass characteristic that is separable in space and time for high frequencies only. For an evaluation of (10), closed-form representations of the spatiotemporal MTF are useful. Based on the measurements of Van Nes *et al.* [27], Koenderink *et al.* [28] suggest

$$M(\omega_x, \omega_y, \omega_t) = \exp\left(\frac{-(\omega_x+\omega_y)^2\sigma_1^2}{2}\right) \cdot \sqrt{\frac{\left[\exp\left(-\frac{1}{2}(\omega_x+\omega_y)^2\sigma_2^2\right)-1\right]^2+\omega_t^2\tau_1^2}{1+\omega_t^2\tau_1^2}} \cdot \frac{|\omega_t(\tau_3-\tau_2)|}{\sqrt{(1+\omega_t^2\tau_2^2)(1+\omega_t^2\tau_3^2)}} \cdot \left|\frac{\sin\left(\frac{\pi\omega_t}{\omega_0}\right)}{\frac{\pi\omega_t}{\omega_0}}\right|$$

with

$$\sigma_1 = 1.0 \text{ min}$$
$$\sigma_2 = 8.5 \text{ min}$$
$$\tau_1 = 30 \text{ ms}$$
$$\tau_2 = 50 \text{ ms}$$
$$\tau_3 = 500 \text{ ms}$$
$$\omega_0 = 2\pi \cdot 40 \text{ Hz}$$

(Fig. 7). From the nonlinear "visual filter" proposed by Lukas and Budrikis [29] to fit Robson's data [25], the following transfer function can be derived by a linearization for small stimuli:

$$M(\omega_x, \omega_y, \omega_t) = \frac{1}{k+L_B}\exp\left[-\frac{1}{2}(\omega_x^2+\omega_y^2)\sigma_e^2\right] \cdot \frac{1}{1+\omega_t^2\tau_e^2} - \frac{L_B}{(k+L_B)^2}\exp\left[-\frac{1}{2}(\omega_x^2+\omega_y^2)\sigma_i^2\right] \cdot \frac{1}{1+\omega_t^2\tau_i^2} \quad (12)$$

with

$$\sigma_e = 1.02 \text{ min}$$
$$\sigma_i = 8.04 \text{ min}$$
$$\tau_e = 21 \text{ ms}$$
$$\tau_i = 48 \text{ ms}$$
$$k = 0.3 \text{ cd/m}^2$$
$$L_B = \text{background luminance.}$$

Both transfer functions (11) and (12) are spatially isotropic. The design of noise feedback filters with either model yields almost identical results, such that we can restrict ourselves to the transfer function (11) in the following.

Fig. 8 shows the gain by noise shaping for a one-tap horizontal noise feedback filter

$$H(\omega_x, \omega_y, \omega_t) = H(\omega_x) = h_x \cdot \exp(-j\omega_x\Delta_x) \quad (13)$$

as a function of the noise feedback coefficient h_x for different viewing distances. Δ_x in (13) is the viewing angle between horizontally adjacent samples. The gain that can be achieved by spatial noise shaping and the optimum noise feedback coefficient depends on the viewing distance. A viewing distance of six times screen height is recommended [30]. The corresponding curve in Fig. 8 possesses a flat minimum of approximately -2 dB at around $h_x = 5/8$. The gain that can be achieved by any other 2D intrafield noise feedback filter over the horizontal one-tap filter is marginal, as long as stability condition (8) is obeyed.

A larger gain of 4.4 dB in terms of E' (10) is theoretically obtained for a spatiotemporal filter

$$H(\omega_x, \omega_y, \omega_t) = \frac{3}{8}\exp[-j\omega_x\Delta_x] + \frac{1}{4}\exp[-j\omega_y\Delta_y - j\omega_t\Delta_t] + \frac{1}{4}\exp[j\omega_y\Delta_y - j\omega_t\Delta_t] \quad (14)$$

where Δ_y is the viewing angle between vertically adjacent interlaced scan lines in a frame and Δ_t is the time interval

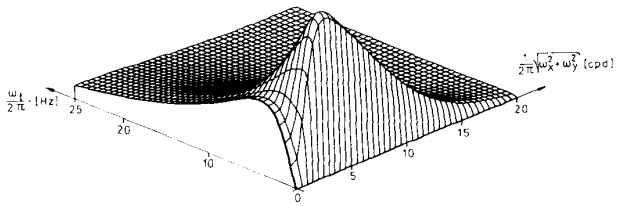


Fig. 7. Spatiotemporal modulation transfer function of the human visual system according to (11).

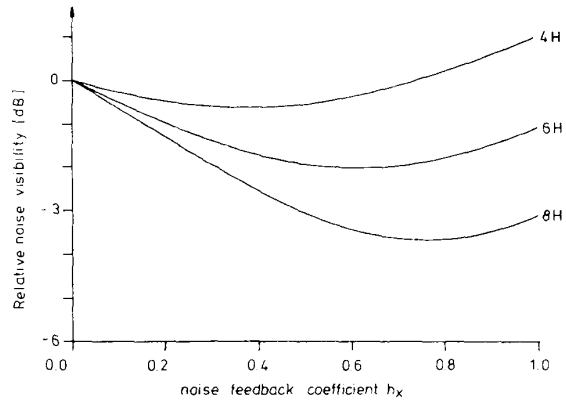


Fig. 8. Relative noise visibility according to (10) with a horizontal one-tap noise feedback filter as a function of noise feedback coefficient h_x (13). Viewing distances are $4H$, $6H$, and $8H$. Data for luminance signal Y, sampling frequency 10.125 MHz.

between fields (Δ_t = 20 ms). There exists, however, a general problem in the perception of temporal high-frequency noise, which can be interpreted as a Doppler effect by eye movements. If the human eye moves with a velocity of (v_x, v_y) relative to the television screen, visibility criterion (10) should be modified to its motion-compensated version

$$E_{mc} = \int\int\int_{\omega_x\omega_y\omega_t} |M(\omega_x, \omega_y, \omega_t - v_x\omega_x - v_y\omega_y)|^2 \cdot |1 - H(\omega_x, \omega_y, \omega_t)|^2 \, d\omega_x \, d\omega_y \, d\omega_t. \quad (15)$$

For certain eye velocities, E_{mc} indicates a deterioration of the picture quality by 3D noise shaping. Subjective tests that are reported in [31] have shown that even for still picture contents, saccadic eye movements suffice to cancel the additional gain by the 3D filter (14) over the best 1D filter (13).

In conclusion, the best noise feedback filter for the luminance signal Y sampled at 10.125 MHz and viewed from $6H$ is a simple horizontal tap with a noise feedback coefficient h_x = 5/8 that can easily be incorporated into the quantization error table QE (Fig. 5). The theoretical gains that we expect for such a noise feedback filter are listed in Table II for viewing distances $4H$ and $6H$.

Typical reconstruction error patterns for a conventional DPCM coder and for the corresponding noise-shaping DPCM coder are shown in Fig. 9 for a small window of the original picture shown in Fig. 16(a). The fine structure of reconstruction errors with noise shaping is clearly less annoying than conventional DPCM quantization noise. In the background, it can additionally be observed that noise shaping breaks up annoying low-frequency limit-cycle patterns. When viewed as a picture sequence with dynamic reconstruction noise, the differences between conventional DPCM and noise-shaping DPCM are even more striking than with frozen noise in Fig. 9.

TABLE II
THEORETICAL GAINS BY NOISE SHAPING

signal	noise feedback filter	gain 6H in dB	gain 4H in dB
Y	eq. (13) $h_x = 5/8$	2.0	0.3
R-Y	eq. (18) $h_y = 1/2$	2.2	0.6
B-Y	eq. (18) $h_y = 3/4$	4.8	2.2

D. Design of Noise Feedback Filters for the Color Difference Signals

Measurements of the spatiotemporal modulation transfer of the wide-band chromaticity channel (*RG* channel) and the narrow-band chromaticity channel (*YB* channel) of the human visual system indicate that both possess spatiotemporal low-pass characteristics and both are separable in space and time [32]. We can roughly approximate the spatial characteristic of the transfer function by

$$M(\omega_x, \omega_y) = \exp\left[-\frac{\sigma^2}{2}(\omega_x^2 + \omega_y^2)\right]. \quad (16)$$

Fitting this approximation to data published in [32] for the luminance range that corresponds to television viewing conditions, we found spread constants of approximately σ = 1.5 min for the *RG* channel and σ = 2.25 min for the *YB* channel. The resulting model transfer functions for both channels are shown in Fig. 10. It should be noted that the data on the spatiotemporal chromaticity MTF given in the literature are not completely consistent. For example, the spatial bandwidths of both chromaticity channels given in [33] are significantly smaller than in [32].

In order to determine the optimum noise feedback filter characteristics for the color difference signals, we assume in the following that the human eye's transfer characteristic for $R-Y$ reconstruction noise is identical to the transfer function of the RG channel, and that the $B-Y$ transfer function is determined by the transfer function of the YB channel. As for the luminance noise feedback filter, a frequency-weighted mean-squared error criterion (10) is used. First-order noise-feedback filters, using only the previous quantization error in the current line (sample S_1 in Fig. 2) or using only the quantization error at the same horizontal position of the previous line of the current field (sample S_3 in Fig. 2), have been considered. Their transfer functions are

$$H(\omega_x, \omega_y, \omega_t) = H(\omega_x) = h_x \cdot \exp[-j\omega_x\Delta_{xc}] \quad (17)$$

for horizontal noise shaping or

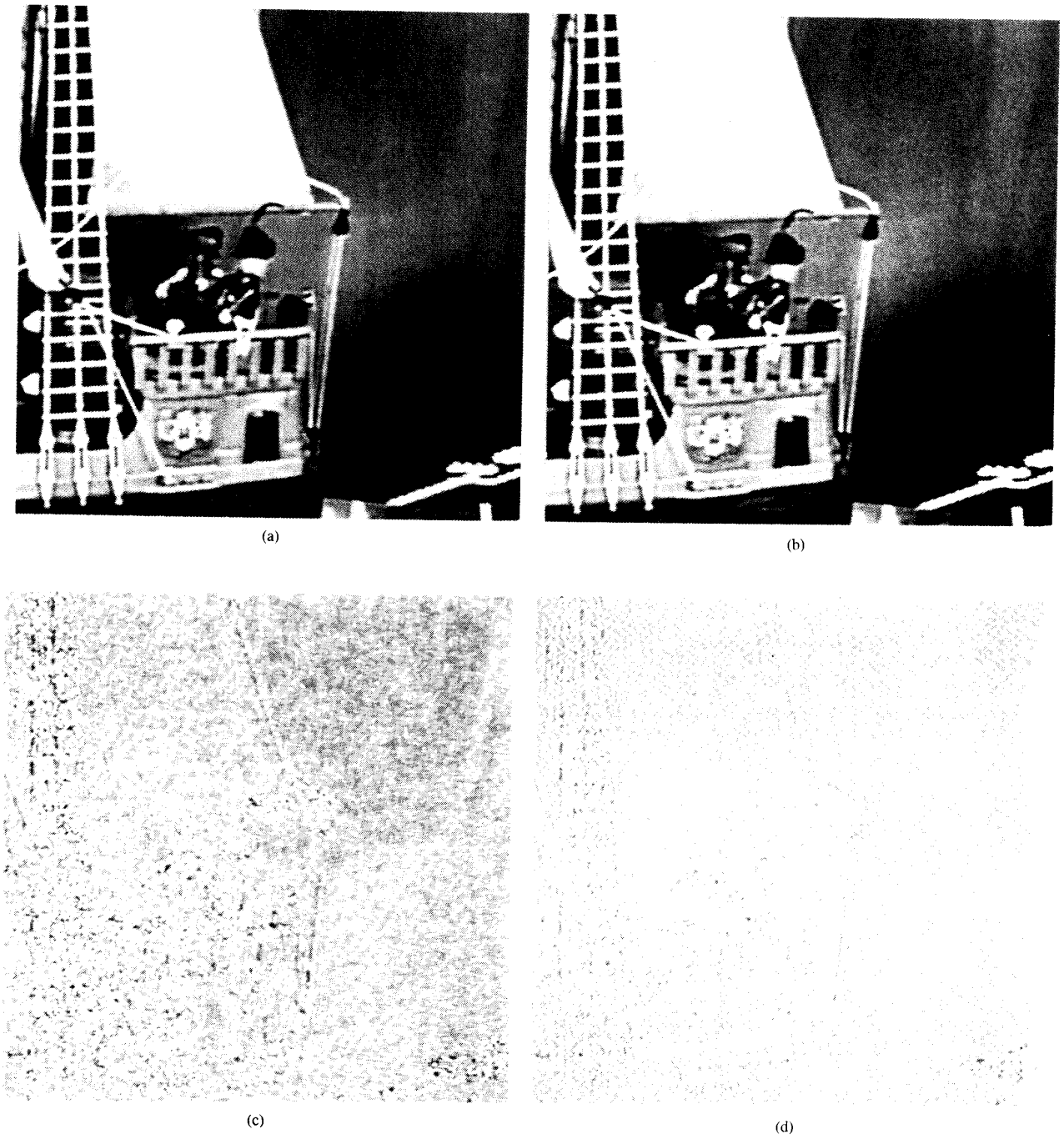
$$H(\omega_x, \omega_y, \omega_t) = H(\omega_y) = h_y \cdot \exp[-j\omega_y \cdot 2\Delta_y] \quad (18)$$

for vertical noise shaping where Δ_{xc} is the viewing angle between horizontally adjacent color difference samples and Δ_y again is the distance between vertically adjacent interlaced scan lines within one frame. For a sampling frequency of 3.375 MHz and a viewing distance of $6H$, we obtain

$$\Delta_{xc} = 4.2 \text{ min} \qquad 2\Delta_y = 2.0 \text{ min}. \quad (19)$$

The noise visibility E' according to (10) is shown in Figs. 11 and 12 as a function of the noise feedback coefficients for both horizontal and vertical noise shaping for $R-Y$ and $B-Y$. Vertical noise shaping is clearly superior, which is due to the fact that $2\cdot\Delta_y$ is significantly smaller than Δ_{xc} (19).

The gains that can be achieved by vertical noise shaping of the color difference signals for a viewing distance of $6H$ are



(a) (b) (c) (d)

Fig. 9. Illustration of typical reconstruction error patterns for a conventional DPCM coder and for a noise-shaping DPCM coder. The predictor is fixed to the intraframe mode. The quantizer corresponds to a square-root envelope function according to (21) with $b = 2.6$. The window displayed covers approximately 290 pels $\times$ 320 lines. The error signal n is multiplied by 4 and superimposed to a medium grey value. (a) Reconstructed signal S', conventional DPCM. (b) Reconstructed signal S', noise-shaping DPCM. (c) Reconstruction error n, conventional DPCM. (d) Reconstruction error n, noise-shaping DPCM.

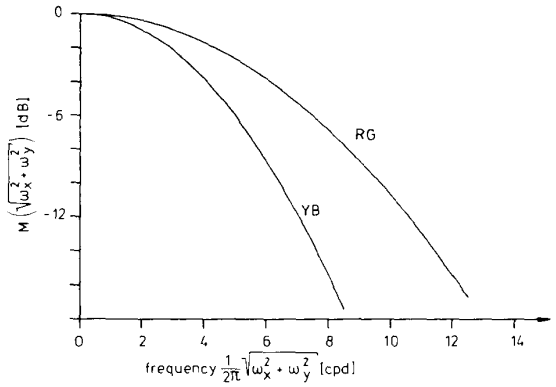


Fig. 10. Spatial transfer function of the *RG* and the *YB* channel of the human visual system according to (16).

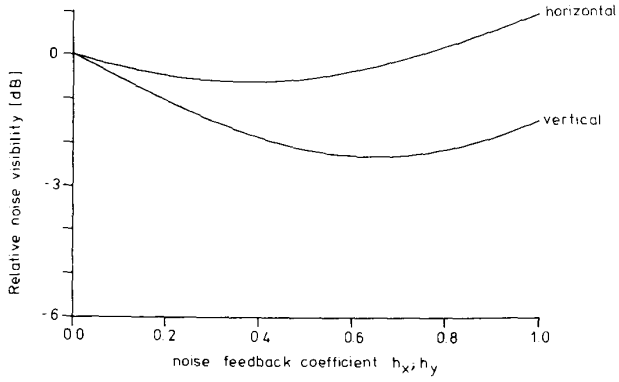


Fig. 11. Relative noise visibility according to (10) versus noise feedback coefficient for horizontal (17) and vertical (18) noise shaping of *R-Y*.

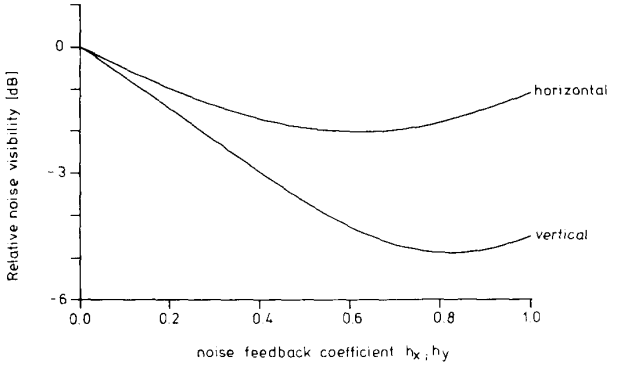


Fig. 12. Relative noise visibility according to (10) versus noise feedback coefficient for horizontal (17) and vertical (18) noise shaping of *B-Y*.

listed in Table II. Table II also includes the gains that are achieved with the optimum coefficients for $6H$ if the picture is viewed from $4H$.

IV. Measurement of Quantizer Parameters for Just Not Visible Reconstruction Errors

In this section, we describe the methodology that we have used for our subjective tests. Our aim is nonuniform "threshold quantizer" characteristics that produce just not visible reconstruction errors in natural scenes. In Section IV-A, we try to motivate the use of a specific type of quantizer that possesses only one free parameter. Section IV-B describes our experimental setup for the subjective tests, while Section IV-C describes the subjective test procedure.

A. Parametrization of the "Masking Function" by a Square-Root Quantization Error Characteristic

For large magnitudes of the prediction error in a DPCM system, coarser quantization is allowed than for small prediction errors. In order to determine a nonuniform quantization characteristic for just not visible reconstruction errors, several researchers have conducted subjective tests to determine the "masking function" [34], [12]–[17]. The masking function reflects a relationship between prediction error and the allowed amplitude of the quantization error that was assumed to be governed by the masking effect at spatial luminance discontinuities, investigated originally by Fiorentini *et al.* [35]. Once the masking function is known, it is straightforward to construct the coarsest nonuniform quantizer for just not visible reconstruction errors [36]: the masking function simply serves as the envelope of the quantization error characteristic

$$q(e) = e' - e. \tag{20}$$

In order to measure the masking function, an approximation by a staircase function or by a piecewise linear function has been proposed [13]. The masking function is accordingly characterized by a limited number of parameters; e.g., for the measurement of a staircase masking function, typically more than five parameters have to be varied. Unfortunately, a tradeoff exists between the parameters, and as a consequence, not just one, but an infinite number of masking functions exists [37]. In the case of many parameters, their tradeoff can only be investigated with tremendous effort. Additionally, the masking function will more likely contain artifacts that are due to a specific picture material. It is desirable to parameterize the masking function with as few parameters as possible.

In the Appendix, we derive a masking function that yields the coarsest quantizer for just not visible quantization errors at edges of most unfavorable orientation, velocity, height, and edge spread:

$$m(e) = b \cdot \sqrt{|e|}. \tag{21}$$

As (21) contains only one free parameter b, we denote a quantizer corresponding to this envelope function as "b quantizer." The derivation of the "masking function" (21) does not involve any spatial masking effect [35], but simply uses a *local* mean-squared error visibility criterion.

A comparison of masking functions given in the literature [13], [17] to (21) shows a close resemblence (Figs. 13 and 14). In our subjective tests, we mainly have considered b quantizers.

B. The Experimental Setup

A block diagram of the experimental setup used for the subjective tests is shown in Fig. 15. The test scene is picked up by a camera and coded by a flexible testbed that is able to perform a variety of adaptive intra-/interframe DPCM algorithms in realtime. Details of the testbed have been described in [5]. The reconstructed signal S' (Fig. 1) is displayed on a color monitor and judged by the subjects.

Two different scenes have been used for the experiments. Both have been carefully constructed to contain the most critical cases with respect to the visibility of reconstruction errors. The scene "ship," that was used for the evaluation of luminance reconstruction errors, contains a ship rotating slowing on a turntable (approximately 3 revolutions/min), circled by a toy train (approximately 6 rounds/min) (Fig. 16). The background contains slight, slowly moving shadows, which turned out to produce the most visible reconstruction

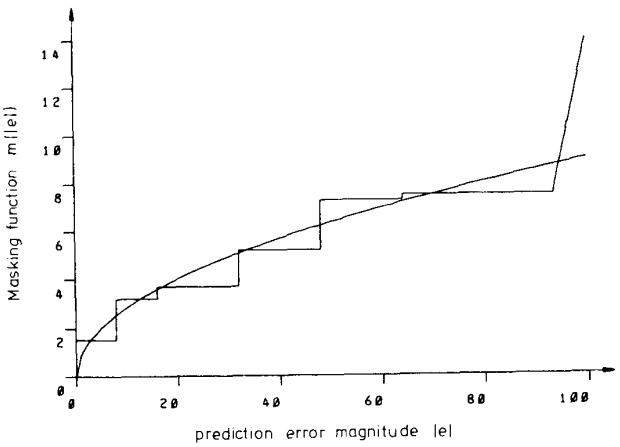


Fig. 13. Masking function measured by Pirsch for a fixed intraframe predictor at a sampling frequency of 8.867 MHz with a quincunx sampling structure [13]; approximation according to (21) with $b = 0.9$.

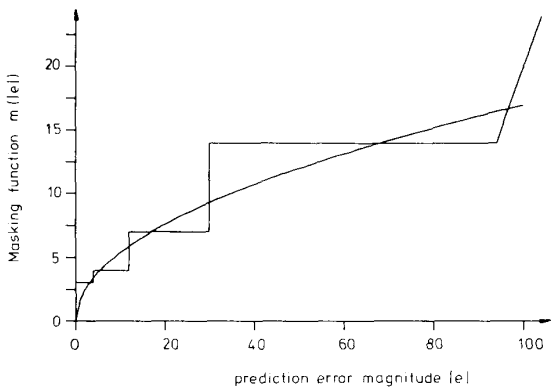


Fig. 14. Masking function measured by Westerkamp for an adaptive intra-/interframe predictor at a sampling frequency of 10 MHz [17]; approximation according to (21) with $b = 1.7$.

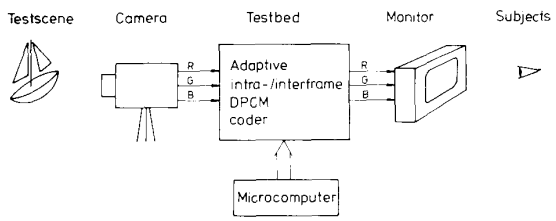
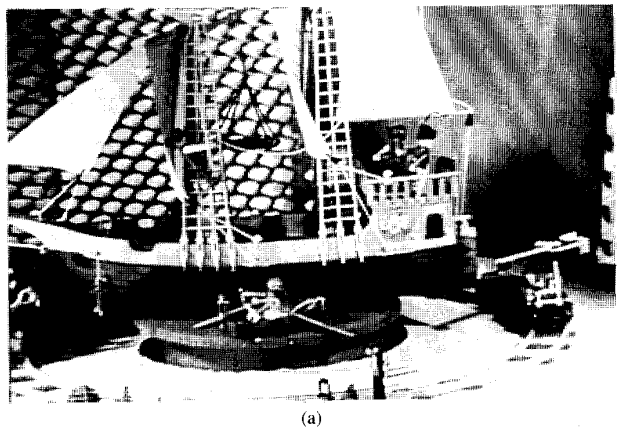
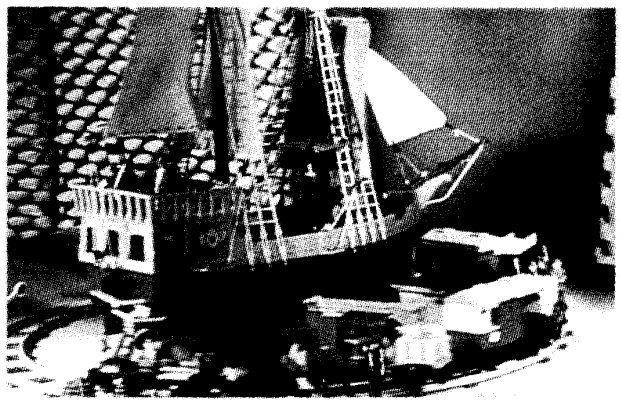


Fig. 15. Experimental setup for the subjective tests.

errors. The scene was displayed without color difference signals. The scene "toys" for the investigation of chrominance reconstruction errors is shown in Fig. 17. It contains toys with saturated colors that are partly rotating on a turntable (16 revolutions/min). A deflated plastic ball is rotated and lifted up and down by a hidden robot arm. A distant lamp that is periodically dimmed (0.533 Hz) produces some pulsating reflexes and a slightly changing overall illumination. The automatic iris of the camera closes briefly once every turntable rotation and causes the entire scene contents to change. The background contains a large uniformly dark area in which chrominance noise can be perceived especially well. Although both scenes never exactly repeat themselves, the same critical



(a)



(b)

Fig. 16. Two time instances of test scene "ship" used for the subjective testing of luminance reconstruction errors.

situations for perception of reconstruction errors happen quite regularly.

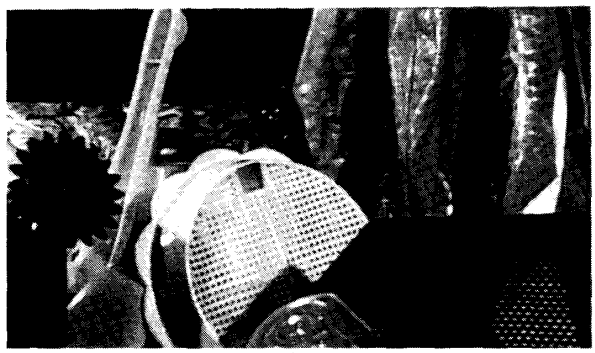
The scenes were picked up by RGB mobile 3/4 in tube broadcasting cameras with 2D aperture correction. The unweighted signal-to-noise ratio of the overall analog input was measured for the scene "toys" to be SNR = 47 dB/55 dB/52 dB for $Y/R-Y/B-Y$, respectively.

The flexible testbed digitizes the components $Y/R-Y/B-Y$ at sampling rates of 10.125/3.375/3.375 MHz. For the three components, isotropic intraframe prediction values (2) with coefficients according to Table I and interframe prediction values (1) are calculated. A change detector (3) switches the prediction mode based on the luminance signal only. A common fixed quantizer is used for both prediction modes.

For reconstruction noise shaping, a horizontal one-tap noise feedback filter is used in the luminance channel and vertical one-tap feedback filters are used in the color difference channels, as summarized in Table II. The output of the coder was displayed on the RGB color monitor of a subjective testing facility that corresponds to CCIR Recommendation 500 [30]. The gamma of the monitor was 2.2. For the luminance tests, the monitor was carefully adjusted such that the displayed brightness on the screen was in accordance with CCIR Recommendation 500. For the chrominance tests, the display was calibrated with respect to color as well. The monitor's white value was adjusted to $D65$. The chromaticity coordinates of the phosphors have been measured and are given in Table III.

C. The Subjective Test Procedure

In order to determine the visibility of a specific impairment, we use a forced-choice comparison test [13]. The coder



(a)

(b)

Fig. 17. Two time instances of test scene "toys" used for the subjective testing of chrominance reconstruction errors.

TABLE III
CHROMATICITY COORDINATES (CIE 1931) OF THE PHOSPHORS MEASURED FOR THE MONITOR USED IN THE CHROMINANCE TESTS

	x	y
Red (R)	0.68	0.29
Green (G)	0.30	0.60
Blue (B)	0.14	0.06

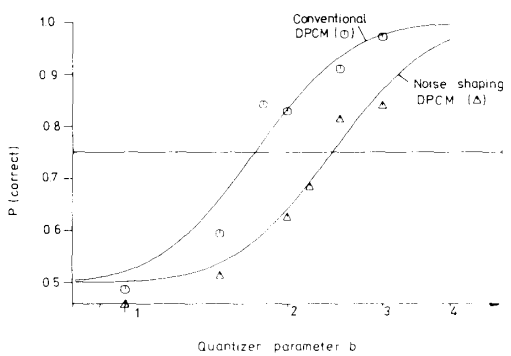


Fig. 18. Probability of correct answers versus coarseness of a b quantizer for luminance reconstruction errors with conventional DPCM and with noise-shaping DPCM; viewing distance is $6H$.

switches between two signals randomly. One of the signals is the DPCM-coded signal S' (Fig. 1) that contains the reconstruction errors to be judged by the subjects. The other signal, which has been coded by an 8 bit PCM only and thus does not contain DPCM reconstruction errors, serves as a reference. The subjects have to decide whether the randomly presented signal is the PCM or the DPCM signal by pressing one out of two possible buttons. If they clearly can distinguish both modes, they give 100 percent correct answers. If they cannot see a difference, they always have to guess, and the probability of correct answers is 50 percent. If we consider the probability of correct answers as a function of the strength of an impairment, we find a transition range between "seen" and "not seen" (Fig. 18). We define a probability of 75 percent correct answers as the value that corresponds to the "visibility threshold."

The subjective tests were carried out with subjects that had experience in the television field. The specific impairment was demonstrated before each test session. Thirteen male subjects participated in the luminance tests, and 13 males and 1 female in the chrominance tests. All subjects had normal or properly corrected acuity. For the chrominance tests, all subjects were checked for color deficiencies by pseudoisochromatic test-charts. The luminance tests and the chrominance tests were conducted within three days each. Each subject participated in two sessions a day. After some training, all subjects showed a stable performance in their responses. For the rare case when a subject obviously failed to concentrate, his responses were excluded from further evaluation.

Four or five subjects participated simultaneously in a test session. The visibility of reconstruction errors for a certain fixed quantizer has been tested as follows. In a training phase of approximately 1 min, the PCM signal and the DPCM signal were introduced. After that, in the judgment phase, either signal was presented randomly, and the subjects gave their judgments for each presentation. Each judgment phase consisted of 10–12 presentations of 8 s for scene "ship" or 6 s for scene "toys." A test session, which lasted 20 to 30 min, consisted of 8–11 training and judgment phases. The type of impairment was similar throughout each test session. For a fair comparison, standard DPCM and noise-shaping DPCM were mixed within each test session. Each sesssion started with impairments that are clearly visible, such that the subjects became acquainted with the type of impairment. Then impairments were presented in the order of decreasing visibility. Switching between coder modes (PCM and DPCM) was never done directly, but a uniform field was displayed in between, such that the subjects did not have the chance to detect an impairment more easily in the instant of switching.

V. Threshold Quantizers for the Luminance Signal

For our adaptive intra-/interframe DPCM coder, we have measured the visibility of reconstruction noise with a b quantizer according to (21) in a conventional DPCM coder and a noise-shaping DPCM coder by the forced-choice test described in the previous section. The estimated probability of correct answers is shown as a function of the parameter b in Fig. 18. Each measurement point in Fig. 18 represents typically 140 responses, which indicates a standard deviation always below 5 percent for statistically independent responses [38]. We have fitted curves

$$P_{\text{correct}}(b) = \frac{1}{2} + \frac{1}{2}\,\text{erf}\left[\frac{\log_{10}\left(\frac{b}{\theta}\right)}{\sigma_{\text{ens}}}\right]$$

with

$$\text{erf}\,(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-t^2}{2}\right) dt \qquad (22)$$

to the data. θ is the visibility threshold according to our definition in Section IV-C. σ_{ens} is the "ensemble standard deviation," which reflects the uncertainty of the single subject in the perception of near-threshold stimuli and the distribution of individual visibility thresholds within the group of subjects. The ensemble standard deviation may not be confused with the standard deviation of the measurement of the visibility threshold. For the results presented in this paper, the standard deviation of the visibility threshold measurement is typically 20 percent of the ensemble standard deviation.

Table IV states the visibility thresholds corresponding to Fig. 18 together with upper and lower deviation values of b corresponding to σ_{ens}. The notation used for the threshold values is illustrated in Fig. 19. Threshold quantizers for a viewing distance of $6H$ are listed in Table V. The visibility threshold of $b = 1.7$ for conventional DPCM viewed at $6H$ agrees with Westerkamp's result for adaptive intra-/interframe DPCM [17] (Fig. 14). Westerkamp obtained his results with similar predictors, but with a different adaptation strategy.

With noise shaping, there is a gain which we can express is decibels according to

$$G_{ns} = 20 \cdot \log_{10} \left[\frac{b_{ns}}{b_{st}} \right] \text{dB} \tag{23}$$

where b_{ns} is the visibility threshold with noise shaping and b_{st} is the visibility threshold for a conventional DPCM system. If we compare this gain to the numbers in Table II based on the design model, we find that the actual gain is slightly larger than predicted for a viewing distance of $6H$. For $4H$, the model predicts an almost complete loss of noise-shaping gain. The subjective tests that we carried out for a viewing distance of $4H$ do not confirm this prediction (Table IV). The noise-shaping gain does not drop significantly for the shorter viewing distance of $4H$. An explanation for this finding might be an additional low-pass characteristic of the display device, e.g., as a result of the unavoidable electron beam aperture.

For comparison, we also determined uniform threshold quantizers for our coding scheme. Similar to the definition of the b quantizer (21), a uniform quantizer can be defined by a constant envelope function

$$m(e) = A. \tag{24}$$

Equation (24) defines a uniform quantizer uniquely if additionally one representative level of the quantizer is known. We have used the representative level "0" in our measurements.

The results of the subjective tests for uniform quantization are shown in Fig. 20 and in Table VI. Assuming that it is the distance between the three innermost levels of a uniform quantizer that determines the visibility of reconstruction errors, a uniform quantizer and a b quantizer should be judged similarly in the subjective tests if

$$A \approx b^2. \tag{25}$$

Equation (25) helps to explain the visibility thresholds for uniform quantization as well as the relatively larger ensemble standard deviation (Table VI). Accordingly, the noise-shaping gain

$$G_{ns} = 20 \log_{10} \left[\frac{a_{ns}}{a_{st}} \right] \text{dB} \tag{26}$$

is almost twice as large as for b quantization. The model calculations presented in Section III-C do not apply. For uniform quantization, the visibility of reconstruction errors is determined almost exclusively by the limit cycle behavior of the coder. Noise shaping breaks up annoying low-frequency limit cycles, and thus results in an even larger gain of G_{ns} = 5 dB.

TABLE IV
VISIBILITY THRESHOLDS FOR b QUANTIZATION OF THE LUMINANCE SIGNAL TOGETHER WITH THE ENSEMBLE STANDARD DEVIATION AS SHOWN IN FIG. 18. THE NOTATION OF THE THRESHOLD IS ILLUSTRATED IN FIG. 19. THE STANDARD DEVIATION OF THE THRESHOLD MEASUREMENT IS TYPICALLY 20 PERCENT OF THE ENSEMBLE STANDARD DEVIATION

viewing distance	conventional DPCM	noise shaping DPCM	noise shaping gain G_{ns} [dB]
6H	$b = 1.7^{+0.7}_{-0.5}$	$b = 2.4^{+0.9}_{-0.7}$	3.0
4H	$b = 1.3^{+0.4}_{-0.3}$	$b = 1.8^{+0.3}_{-0.3}$	2.8

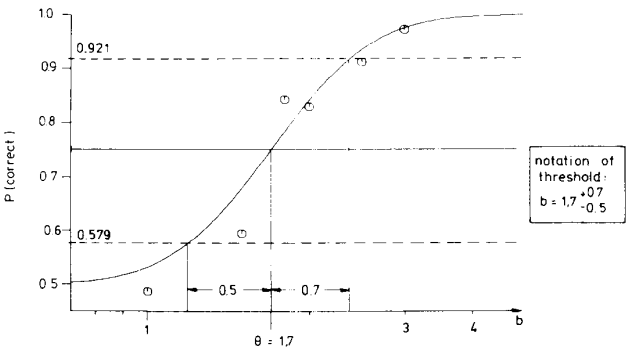


Fig. 19. Threshold notation in Tables IV, VI, and VII illustrated for the measurements with a conventional DPCM coder as presented in Fig. 18.

TABLE V
THRESHOLD QUANTIZERS FOR THE LUMINANCE SIGNAL WITH CONVENTIONAL DPCM AND WITH NOISE SHAPING DPCM; VIEWING DISTANCE $6H$

conventional DPCM (b=1.7)		noise shaping DPCM (b=2.4)	
prediction error	representative level	prediction error	representative level
...-103	-120		
-102 ... -71	-85		
-70 ... -45	-56	... -93	-116
-44 ... -25	-33	-92 ... -52	-69
-24 ... -11	-16	-51 ... -22	-34
-10 ... -3	-5	-21 ... -6	-11
-2 ... 2	0	-5 ... 5	0
3 ... 10	5	6 ... 21	11
11 ... 24	16	22 ... 51	34
25 ... 44	33	52 ... 92	69
45 ... 70	56	93 ...	116
71 ... 102	85		
103 ...	120		

The threshold b quantizers measured are coarser than the corresponding uniform quantizers (Tables IV–VI). Thus, b quantization results in both a smaller number of quantizer representative levels and a smaller entropy of the prediction error for just not visible reconstruction errors.

VI. THRESHOLD QUANTIZERS FOR THE COLOR DIFFERENCE SIGNALS

In order to determine quantizers for just not visible reconstruction errors in the color difference signals, we have

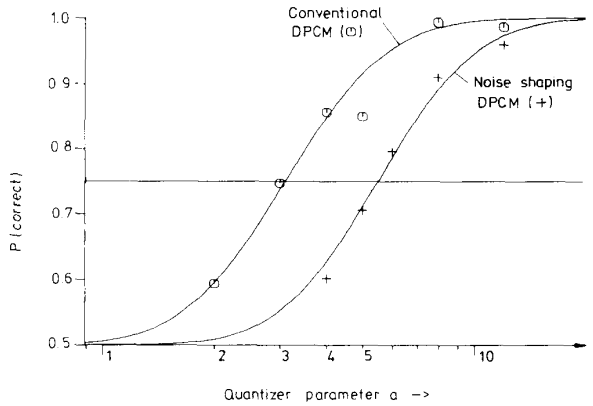


Fig. 20. Probability of correct answers versus coarseness of a uniform quantizer for luminance reconstruction errors with conventional DPCM and with noise-shaping DPCM; viewing distance is $6H$.

TABLE VI
VISIBILITY THRESHOLDS FOR UNIFORM QUANTIZATION OF THE LUMINANCE SIGNAL TOGETHER WITH THE ENSEMBLE STANDARD DEVIATION AS SHOWN IN FIG. 20; THE STANDARD DEVIATION OF THE THRESHOLD MEASUREMENT IS TYPICALLY 20 PERCENT OF THE ENSEMBLE STANDARD DEVIATION

viewing distance	conventional DPCM	noise shaping DPCM	noise shaping gain G_{ns} [dB]
6H	$A = 3.1^{+2.0}_{-1.2}$	$A = 5.5^{+3.4}_{-2.1}$	5.0

measured the visibility of reconstruction noise with b quantizers according to (21) for $R-Y$ and for $B-Y$. Similarly as for the luminance, we carried out forced-choice subjective tests (Section IV) both without and with vertical noise shaping (Table II).

During the first test series, we measured visibility thresholds for reconstruction errors in one color difference signal only. This was done by introducing quantization only in the component of interest and displaying it together with the reference PCM signal of the other components. The single component visibility thresholds for $6H$ are listed in Table VII in the lines "$R-Y$" and "$B-Y$", respectively, together with their ensemble standard deviations. The noise-shaping gains are very close to the theoretical values in Table II for both $R-Y$ and $B-Y$. $B-Y$ errors are generally less visible than $R-Y$ errors, as also has been reported by other authors [15], [16].

If both $R-Y$ and $B-Y$ contain reconstruction errors from quantizers according to the single component visibility thresholds, the combined impairment exceeds the visibility threshold. We thus have to lower the b parameters found in the single component measurements for just not visible chrominance reconstruction errors. Subjective tests with combined $R-Y/B-Y$ impairments were carried out by using a combination of quantizers with

$$b_{R-Y} = c \cdot 2.0; \quad b_{B-Y} = c \cdot 2.4 \tag{27}$$

for conventional DPCM, and with

$$b_{R-Y} = c \cdot 2.6; \quad b_{B-Y} = c \cdot 4.0 \tag{28}$$

for noise-shaping DPCM. The correction factor c defines a combination of quantizers that produce a certain fraction of the single component threshold impairment. For $c = 1$, both $R-Y$ and $B-Y$ reconstruction errors are at their single component thresholds. We have measured visibility as a function of c for both conventional DPCM and noise-shaping DPCM. In both cases, the visibility threshold for combined $R-Y/B-Y$ impairments is approximately 80 percent of the single component threshold (Table VII). The noise-shaping gain determined in the single component tests is confirmed by this new, independent measurement up to 0.1 dB. The resulting threshold quantizers for just not visible chrominance errors are listed in Tables VIII and IX.

TABLE VII
VISIBILITY THRESHOLDS FOR b QUANTIZATION OF THE COLOR DIFFERENCE SIGNALS TOGETHER WITH THE ENSEMBLE STANDARD DEVIATION; THE STANDARD DEVIATION OF THE THRESHOLD MEASUREMENT IS TYPICALLY 20 PERCENT OF THE ENSEMBLE STANDARD DEVIATION

Signal	viewing distance	conventional DPCM	noise shaping DPCM	noise shaping gain G_{ns} [dB]
R-Y	6H	$b = 2.0^{+0.6}_{-0.4}$	$b = 2.6^{+0.7}_{-0.6}$	2.3
B-Y	6H	$b = 2.4^{+0.6}_{-0.5}$	$b = 4.0^{+0.8}_{-0.6}$	4.4
R-Y B-Y	6H	$c = 0.81^{+0.20}_{-0.16}$	$c = 0.82^{+0.14}_{-0.12}$	2.4 4.5
R-Y B-Y	4H	$c = 0.71^{+0.23}_{-0.17}$	$c = 0.55^{+0.24}_{-0.17}$	0.1 2.2

TABLE VIII
THRESHOLD $R-Y$ QUANTIZERS FOR JUST NOT VISIBLE CHROMINANCE RECONSTRUCTION ERRORS WHEN COMBINED WITH THE QUANTIZERS IN TABLE IX; VIEWING DISTANCE $6H$

conventional DPCM (b=1.62)			noise shaping DPCM (b=2.13)		
prediction error		representative level	prediction error		representative level
		.			.
		.			.
		.			.
...	-95	-110			.
-94 ...	-66	-79			
-65 ...	-42	-52	...	-74	-92
-41 ...	-24	-31	-73 ...	-42	-55
-23 ...	-11	-16	-41 ...	-19	-28
-10 ...	-3	-5	-18 ...	-5	-9
-2 ...	2	0	-4 ...	4	0
3 ...	10	5	5 ...	18	9
11 ...	23	16	19 ...	41	28
24 ...	41	31	42 ...	73	55
42 ...	65	52	74 ...		92
66 ...	94	79			
95 ...		110			.
		.			.
		.			.
		.			
		.			

For conventional DPCM, the visibility threshold is only slightly lower when combined $R-Y/B-Y$ reconstruction noise is viewed from a distance of $4H$ (Table VII). This is due to a color shift within a large area of the dark background in the test scene "toys" (Fig. 17), which is perceived rather independently from the viewing distance. With noise shaping, we observe an expectedly large drop of the visibility threshold. The theoretical noise shaping gains for $4H$ (Table II), which consider single component impairments, and the subjective test results for combined $R-Y/B-Y$ impairments correspond quite well, but we have to be careful comparing them. It is not clear how the human visual system combines $R-Y$ and $B-Y$ reconstruction errors.

VII. b Quantization and Bit Rate

Throughout this paper, we have characterized quantizers with a square-root envelope function (21) by the parameter b. Visibility improvements are stated in decibels according to (23). In this section, we relate the quantizer parameter b to bit rate and to quantization noise power.

For a DPCM system with *fixed wordlength* encoding of the quantized prediction error, the number of quantizer represent-

TABLE IX
THRESHOLD *B–Y* QUANTIZERS FOR JUST NOT VISIBLE CHROMINANCE RECONSTRUCTION ERRORS WHEN COMBINED WITH THE QUANTIZERS IN TABLE VIII; VIEWING DISTANCE *6H*

conventional DPCM (b=1.94)		noise shaping DPCM (b=3.28)	
prediction error	representative level	prediction error	representative level
... -95	-113	... -43	-64
-94 ... -61	-76	-42 ... -11	-21
-60 ... -34	-45	-10 ... 10	0
-33 ... -15	-22	11 ... 42	21
-14 ... -4	-7	43 ...	64
-3 ... 3	0		
4 ... 14	7		
15 ... 33	22		
34 ... 60	45		
61 ... 94	76		
95 ...	113		

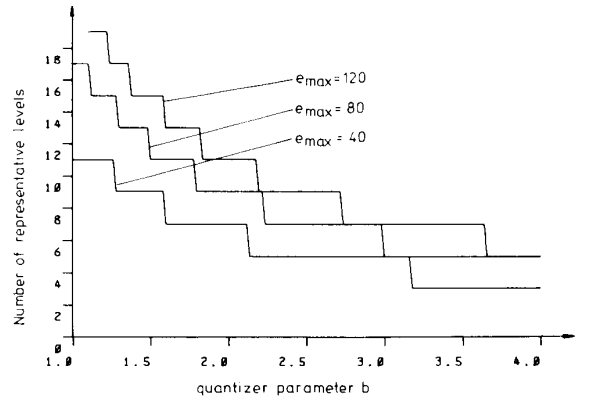


Fig. 21. Number of representative levels of a b quantizer for different maximum prediction error magnitudes e_{max}.

ative levels determines the required transmission bit rate. Other than for variable wordlength encoding of e', we must allow the quantization error characteristic $q(e)$ to exceed the envelope function (21) if $|e| > e_{max}$. The number of representative levels for the b quantizer is plotted in Fig. 21 for different e_{max}. $e_{max} = 120$ is sufficient for the luminance signal in almost all situations, while $e_{max} = 40$ is adequate for the color difference signals. Assuming these maximum prediction error magnitudes, Table X summarizes the number of quantizer representative levels required for just not visible reconstruction errors. From Table X, we conclude that even with noise shaping for fixed wordlength encoding of the prediction error, a 30 Mbit/s transmission is not possible without visible reconstruction errors.

With *variable wordlength* encoding, the average wordlength of the quantized prediction error can approach its entropy. The distribution of probabilities of the prediction error in a DPCM system Pr(e) often times resembles a discrete Laplacian [39], i.e.,

$$\Pr(e) = \begin{cases} A \cdot \exp\left(\dfrac{-\sqrt{2}\,|e|}{\sigma_e}\right) & \text{for integer } e \\ 0 & \text{else} \end{cases} \qquad (29)$$

where σ_e^2 is the prediction error variance. Fig. 22 shows the

TABLE X
NUMBER OF QUANTIZER REPRESENTATIVE LEVELS REQUIRED FOR JUST NOT VISIBLE RECONSTRUCTION ERRORS

Signal	viewing distance	conventional DPCM	noise shaping DPCM
Y	6H	13	9
R-Y	6H	7	5
B-Y	6H	7	3
Y	4H	17	13
R-Y	4H	9	9
B-Y	4H	7	5

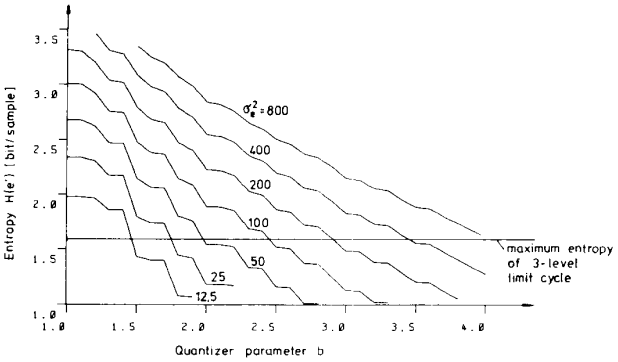


Fig. 22. Entropy of the b quantized prediction error $H(e')$ as a function of the quantizer parameter b for a Laplacian probability distribution of the prediction error with variance σ_e^2.

entropy of the quantized prediction e',

$$H(e') = -\sum_{e'} \Pr(e') \cdot \log_2 [\Pr(e')] \text{ bit} \qquad (30)$$

if a b quantizer is applied to the pdf (29) for different prediction error variances σ_e^2. The corresponding power of the quantization noise

$$P_q = 10 \cdot \log_{10}\left(\frac{E[q^2]}{255^2}\right) \text{ dB} \qquad (31)$$

is shown in Fig. 23. The ripples in Figs. 22 and 23 are due to the finite precision of the number representation in the DPCM loop according to (29) which has been taken into account for the construction of the b quantizer. From Fig. 23, we conclude that the quantization noise power increases by approximately 6 dB when b is increased by 6 dB (23). Indeed, this relation motivates definition (23). Furthermore, in Fig. 22, an increase of b by 6 dB yields a saving in terms of $H(e')$ (30) of approximately 1.3 bits/sample. For a DPCM system with optimum variable length encoding of the quantized prediction error, we can translate the noise shaping gains G_{ns} in Tables IV and VII into approximate bit rate gains by multiplication with a factor of 0.22 bit/dB. This has been done in Table XI.

Note, however, that these bit rate gains are valid only as long as the entropies involved are significantly larger than the maximum entropy for a three-level limit cycle (see Section III-B). This maximum limit-cycle entropy of 1.58 bits/sample is marked in Fig. 22. For transmission rates around 30 Mbits/s, we have measured in a practical system that the bit rate gains in Table XI are usually valid for the luminance signal. For the color difference signals, limit cycles tend to set a lower bound for the bit rate with noise shaping. As a consequence, the bit rate gains measured for the color difference signals are often times smaller than those given in Table XI.

We cannot report bit rates for a variable length encoding of the prediction error for just not visible reconstruction errors

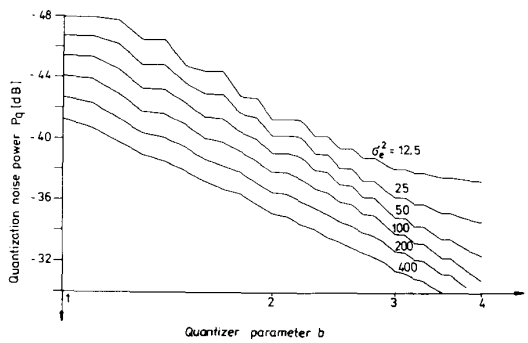


Fig. 23. Quantization error power P_q after b quantization of a Laplacian probability distribution with variance σ_e^2.

TABLE XI
NOISE-SHAPING GAINS TRANSLATED INTO APPROXIMATE BIT RATE SAVINGS

Signal	viewing distance	noise shaping gain [bit/sample]	[Mbit/s]
Y	6H	0.7	5.0
R-Y	6H	0.5	1.3
B-Y	6H	1.0	2.5
Y/R-Y/B-Y	6H		8.8
Y	4H	0.6	4.7
R-Y	4H	0.0	0.0
B-Y	4H	0.5	1.2
Y/R-Y/B-Y	4H		5.9

that are generally valid as the exact figures very much depend on the scene contents. A realistic DPCM coder has to deliver a constant bit rate to the transmission channel despite variable wordlength encoding of the prediction error. It has to control the coarseness of the quantizer characteristic adequately, and the degree of impairment will vary with scene contents rather than the bit rate. The prediction error variance of the luminance signal Y exceeds a value of $\sigma_e^2 = 250$ only for extremely critical scenes, while for $R-Y$ or $B-Y$, $\sigma_e^2 = 100$ is already a very large value. For a transmission rate of around 30 Mbits/s, we can conclude from Fig. 22 that we hardly ever will have to exceed the visibility threshold for a viewing distance of $6H$ if we utilize reconstruction noise shaping for Y, $R-Y$, and $B-Y$.

VIII. Conclusion

In this paper, we have presented nonuniform quantizers for just not visible reconstruction errors in an adaptive intra-/interframe DPCM scheme for component-coded color television signals, both for conventional DPCM and for noise-shaping DPCM. Noise feedback filters that minimize the visibility of reconstruction errors by spectral shaping have been designed for Y, $R-Y$, and $B-Y$ based on results from psychophysics literature. A closed-form description of the "masking function" has been derived, which leads to the one-parameter b quantizer characteristic. Subjective tests were carried out to determine visibility thresholds for reconstruction errors for conventional DPCM and for noise-shaping DPCM. The subjective test results show significant gains by noise shaping. They have been related to the transmission bit rate that is required for just not visible reconstruction errors.

For adaptive intra-/interframe DPCM of television signals within a 3:1:1 system for a bit rate of around 30 Mbits/s, we came to the following conclusions.

- For a fixed length of the prediction error codewords, visible reconstruction errors cannot be avoided.
- For variable wordlength encoding of the prediction error, we hardly ever exceed the visibility threshold for a viewing distance of $6H$ if we utilize reconstruction noise shaping for Y, $R-Y$, and $B-Y$.
- The gain of noise-shaping DPCM over conventional DPCM corresponds to about 5 Mbits/s for the luminance signal and up to 4 Mbits/s for the color difference signals.

Appendix

Derivation of the b Quantizer

In the following, we will motivate a masking function with only one free parameter. Our approach describes natural pictures as consisting of edges of different orientation φ, velocity V, and height H. Edges can be abrupt luminance discontinuities with a small edge spread σ_{edge}, as well as smooth luminance transitions with large σ_{edge}. The basic idea of our approach is that a representative scene contents will contain edges with all possible combinations of φ, V, H, and σ_{edge}, and that the most unfavorable combination of these edge parameters decides about the visibility of reconstruction errors in the picture signal.

Consider a luminance edge profile

$$s(\xi) = H \cdot f\left(\frac{\xi}{\sigma_{\text{edge}}}\right) \tag{A1}$$

where H is the height of the edge, σ_{edge} is the edge spread, ξ is a spatial coordinate orthogonal to the edge, and $f(\xi)$ is a prototype edge profile of unit height (Fig. 24). The prediction error across the edge profile $e(\xi)$ is obtained with a convolution of the reconstructed signal $s'(\xi)$ and a projection $p(\xi)$ of the impulse response of the linear intra-/or interframe predictor into the ξ axis:

$$e(\xi) = s(\xi) - s'(\xi) * p(\xi) \tag{A2}$$

(Fig. 25). With the approximation

$$s \approx s' \tag{A3}$$

we obtain

$$e(\xi) = s(\xi) * (\sigma(\xi) - p(\xi)). \tag{A4}$$

For a predictor with a coefficient sum equal to "1," the prediction error is approximately proportional to the local edge slope, i.e.,

$$e(\xi) \approx C_{\varphi V} \frac{ds(\xi)}{d\xi} = \frac{H \cdot C_{\varphi V}}{\sigma_{\text{edge}}} \cdot \frac{df\left(\frac{\xi}{\sigma_{\text{edge}}}\right)}{d\xi} \tag{A5}$$

where $C_{\varphi V}$ is a factor that depends on velocity and orientation of the edge only. We describe the quantization noise q by an additive noise source uniformly distributed between $-m$ and $+m$ where m is controlled by the prediction error e, and accordingly the mean-squared quantization error at locus ξ is

$$E[q^2(\xi)] = \frac{1}{3}[m(e(\xi))]^2. \tag{A6}$$

$m(e)$ is the envelope of the quantization error characteristic (20) which, for a threshold quantizer, is equal to a masking function. The approximation of the quantizer by an additive independent noise source is justified as long as the number of quantizer levels is sufficiently large.

As a visibility criterion, we use a simplification of the sophisticated threshold model proposed in [40] which leads to the local mean-squared quantization error Q_{edge} that can be

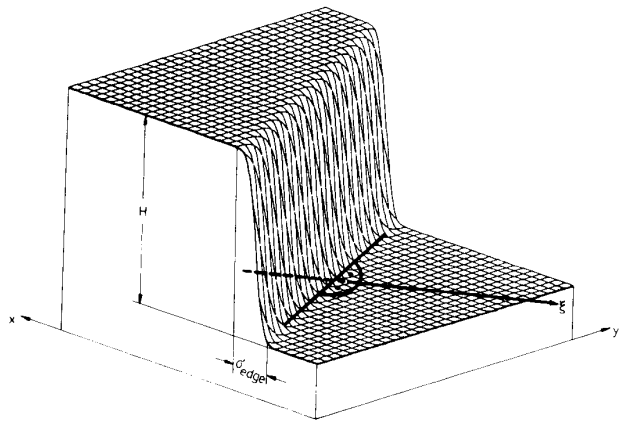


Fig. 24. Luminance edge profile.

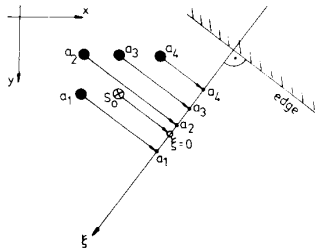


Fig. 25. Projection of an intraframe predictor impulse response (2) onto the ξ axis.

computed by an accumulation across the edge profile, i.e.,

$$Q_{\text{edge}} = \int_{\xi} E[q^2(\xi)]\, d\xi. \tag{A7}$$

With (A6) and (A5), we obtain

$$Q_{\text{edge}} = \int_{v} \frac{1}{3}\,\sigma_{\text{edge}} \cdot \left(m\left(\frac{H \cdot C_{\varphi v}}{\sigma_{\text{edge}}} \cdot \frac{df(v)}{dv} \right) \right)^2 dv. \tag{A8}$$

The coarsest quantizer will result if the local mean-squared quantization error Q_{edge} is as small as possible for the most unfavorable combination of H, $C_{\varphi V}$, and σ_{edge}. As $m(e)$ usually has its minimum at $e = 0$ and is monotonically increasing with increasing $|e|$, Q_{edge} is largest when H and $C_{\varphi V}$ take on their maximum values. σ_{edge} can vary in a very large range. Thus, the coarsest quantizer results if Q_{edge} is independent of σ_{edge}. This is achieved only by choosing the envelope function

$$m(e) = b \cdot \sqrt{|e|}. \tag{A9}$$

As (A9) contains only one free parameter b, we denote a quantizer corresponding to this envelope function as a "b quantizer."

ACKNOWLEDGMENT

The joint effort of the authors that is reported in this paper evolved out of common work done in the European COST 211 bis project. One of the authors' work has been supported generously by an Ernst von Siemens fellowship. The help of Norddeutscher Rundfunk and of the Swedish Broadcasting Company in the experimental setup is gratefully acknowledged. Especially, the authors wish to thank all their colleagues who participated in the subjective tests.

REFERENCES

[1] CCIR Recommendation 601, "Encoding parameters of digital television for studies," in *CCIR Recommendations and Reports,* vol. XI, ITU, Geneva, Switzerland, 1982.

[2] D. Westerkamp, "Adaptive intra-/interframe DPCM-coding for transmission of colour TV-signals with 34 Mbit/s," in *Proc. 1984 Int. Zurich Seminar Digital Commun.,* 1984, pp. C 2.1-C 2.7.

[3] D. Mischler, "TV signals transmission with 34 Mbit/s," presented at the Int. Picture Coding Symp., Rennes, France, 1984.

[4] H. Buley and L. Stenger, "Inter/Intraframe coding of color TV signals for transmission at the third level of the digital hierarchy," *Proc. IEEE,* vol. 73, pp. 765-772, Apr. 1985.

[5] P. Weiss, L. Bengtsson, and B. Christensson, "Adaptive DPCM-coding of broadcast TV-signals at 34 Mbit/s," presented at the Int. Picture Coding Symp., Tokyo, Japan, 1986.

[6] P. Pirsch and M. Bierling, "Changing the sampling rate of video signals by rational factors," in *Signal Processing II: Theories and Applications,* H. W. Schüssler, Ed., *Proc. EUSIPCO 1983,* pp. 171-174.

[7] B. Girod and W. Geuen, "Vertical sampling rate decimation and line-offset decimation of the color difference signals," submitted to *Signal Processing*.

[8] T. Koga, Y. Iijima, K. Iinuma, and T. Ishiguro, "Statistical performance analysis of an interframe encoder for broadcast television signals," *IEEE Trans. Commun.,* vol. COM-29, pp. 1868-1876, Dec. 1981.

[9] H. Yamamoto, Y. Hatori, and H. Murakami, "30 Mbit/s codec for the NTSC color TV signal using an interfield-intrafield adaptive prediction," *IEEE Trans. Commun.,* vol. COM-29, pp. 1859-1867, Dec. 1981.

[10] H. J. Grallert and A. Starck, "Component encoding of color television signals in 34, 70, and 140 Mbit/s channels," in *Proc. 1984 Int. Zurich Seminar Digital Commun.,* pp. C1.1-C1.6.

[11] B. Girod, "Reconstruction noise shaping in the context of predictive TV signal coding," in *Proc. Int. Conf. Commun.,* Amsterdam, The Netherlands, 1984, pp. 711-717.

[12] P. Pirsch, "A new predictor design for DPCM coding of TV signals," in *Proc. Int. Conf. Commun.,* Seattle, WA, 1980, pp. 31.2.1-31.2.5.

[13] ——, "Design of quantizers for video signals using subjective tests," *IEEE Trans. Commun.,* vol. COM-29, pp. 990-1000, July 1981.

[14] R. Schäfer, "Design of adaptive and nonadaptive quantizers using subjective criteria," *Signal Processing,* vol. 5, pp. 333-345, July 1983.

[15] F. X. J. Lukas and F. Kretz, "DPCM quantization of color television signals," *IEEE Trans. Commun.,* vol. COM-31, pp. 927-932, July 1983.

[16] R. Schäfer, "DPCM coding of the chrominance signals for the transmission of color TV signals at 34 Mbit/s," *Signal Processing,* vol. 6, pp. 187-199, 1984.

[17] D. Westerkamp, "Optimierung einer adaptiven intraframe/interframe-DPCM unter Berücksichtigung der Wahrnehmbarkeit von Quantisierungsfehlern" (in German), Ph.D. dissertation, Univ. Hannover, Germany, 1985.

[18] P. Pirsch, "Adaptive intra-/interframe DPCM coder," *Bell Syst. Tech. J.,* vol. 61, pp. 747-764, May 1982.

[19] B. Girod, "Design of switched predictors for interframe DPCM coding of television signals," in *Proc. Int. Conf. Digital Signal Processing,* 1984, pp. 570-576.

[20] ——, "Prädiktoroptimierung unter Berücksichtigung multiplikativen Quantisierungsrauschens" (in German), *Kleinheubacher Berichte,* vol. 27, Fernmeldetechnisches Zentralamt, Ed., Darmstadt, 1983, pp. 393-403.

[21] ——, "Isotropic intraframe prediction," presented at the Int. Picture Coding Symp., Rennes, France, 1984.

[22] P. Noll, "On predictive quantizing schemes," *Bell Syst. Tech. J.,* vol. 57, pp. 1499-1532, May-June 1978.

[23] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Englewood Cliffs, NJ: Prentice-Hall, 1984, pp. 362-371.

[24] P. Pirsch, "Stability conditions for DPCM coders," *IEEE Trans. Commun.,* vol. COM-30, pp. 1174-1184, May 1982.

[25] J. G. Robson, "Spatial and temporal contrast sensitivity functions of the visual system," *J. Opt. Soc. Amer.,* vol. 56, pp. 1141-1142, 1966.

[26] D. H. Kelly, "Visual contrast sensitivity," *Opt. Acta,* vol. 24, pp. 107-129, 1977.

[27] F. L. van Nes, J. J. Koenderink, H. Nas, and M. A. Bouman,

"Spatiotemporal modulation transfer in the human eye," *J. Opt. Soc. Amer.*, vol. 57, pp. 1082–1088, Sept. 1967.

[28] J. J. Koenderink, W. A. van de Grind, and M. A. Bouman, "Foveal information processing at photopic luminances," *Kybernetik*, vol. 8, no. 4, pp. 128–144, 1971.

[29] F. X. J. Lukas and Z. L. Budrikis, "Picture quality prediction based on a visual model," *IEEE Trans. Commun.*, vol. COM-30, pp. 1679–1692, July 1982.

[30] "Method for the subjective assessment of the quality of television pictures," CCIR Rec. 500-2, 1982.

[31] B. Friedrich, "Spektrale Formung des Quantisierungsgeräusches bei der DPCM-Codierung von Fernsehsignalen" (in German), *Diploma* thesis, Univ. Hannover, Germany, 1985.

[32] G. J. C. van der Horst and M. A. Bouman, "Spatiotemporal chromaticity discrimination," *J. Opt. Soc. Amer.*, vol. 59, pp. 1482–1488, Nov. 1969.

[33] A. Watanabe, H. Sakata, and H. Isono, "Chromatic spatial sine-wave responses of the human visual system," NHK Lab. Note 198, Tokyo, Japan, Mar. 1976.

[34] J. O. Limb and C. B. Rubinstein, "On the design of quantizers for DPCM coders: A functional relationship between visibility, probability and masking," *IEEE Trans. Commun.*, vol. COM-26, pp. 573–578, May 1978.

[35] A. Fiorentini, M. Jeanne, and G. Franchi, "Measurements of differential threshold in the presence of a spatial illumination gradient," *Atti Ford. Ronchi*, vol. 10, pp. 371–379, 1955.

[36] D. K. Sharma, "Design of absolutely optimal quantizers for a wide class of distortion measures," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 693–702, Nov. 1978.

[37] D. Bultmann, "Subjektive Tests zur Bestimmung der gerade nicht mehr wahrnehmbaren Quantisierungsfehler bei der DPCM-Codierung von Farbfernsehsignalen" (in German), Diploma thesis, Univ. Hannover, Germany, May 1979.

[38] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1965.

[39] J. B. O'Neal, Jr., "Predictive quantizing systems (differential pulse code modulation) for the transmission of television signals," *Bell Syst. Tech. J.*, vol. 45, pp. 689–721, May–June 1966.

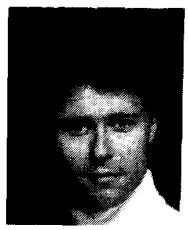
[40] B. Girod, "Ein Modell der menschlichen visuellen Wahrnehmung zur Irrelevanzreduktion von Fernsehluminanzsignalen" (in German), Ph.D. dissertation, Univ. Hannover, Germany, 1987.

★

Bernd Girod was born in Bielefeld, Germany, on December 1, 1957. He received the M.S. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1980, and the Doctoral degree from the Universität Hannover, Hannover, Germany, in 1987.

Temporarily, he worked at Siemens Central Research Laboratories, Munich, Germany. Since 1981 he has been with the Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, Universität Hannover. His work has been supported by an Ernst von Siemens-fellowship since 1985. Currently, he is a Visiting Scientist at Massachusetts Institute of Technology, Cambridge, MA. His research interests include coding of television signals and modeling of the human visual system.

Dr. Girod is a member of the German Informationstechnische Gesellschaft (ITG) des VDE.



Håkan Almer was born in Stockholm, Sweden, in 1959. He received the M.S. degree in electrical engineering from the Royal Institute of Technology, Sweden, in 1983.

In 1983 he joined Ericsson where he developed systems for digital television transmission using data compression. Since 1985 he has worked at the Research Department of Swedish Telecom where his main interest is algorithm development for data compression of moving pictures.

★

Leif Bengtsson was born in Stockholm, Sweden, in 1957. He received the M.S. degree in electrical engineering from the Royal Institute of Technology, Sweden, in 1981.

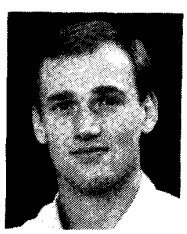
Since 1981 he has worked at the Research Department of Swedish Telecom where his main interest is algorithm research for visual communication systems.

★

Björn Christensson was born in Stockholm, Sweden, in 1957. He received the M.S. degree in electrical engineering from the Royal Institute of Technology, Sweden, in 1983.

In 1983 he joined the Swedish Telecom Radio Services where he worked with digital television coding and transmission. Since 1986 he has worked at the Research Department of Swedish Telecom with algorithm development for both high and low bit rates.

★



Peter Weiss was born in Stockholm, Sweden, in 1958. He received the M.S. degree in electrical engineering from the Royal Institute of Technology, Sweden, in 1981.

Since 1981 he has worked at the Research Department of Swedish Telecom where he has developed systems and algorithms for visual communications.