STATISTICS FOR EFFICIENT LINEAR AND NON-LINEAR PICTURE ENCODING

THOMAS KUMMEROW
Institut fur Nachrichtentechnik
Technische Universitat Braunschweig

Summary

Differential pulse code modulation (DPCM) of video signals can be improved by assuming the model of the general Markov Source, which has not only linear but also nonlinear statistical dependencies. The proposed type of coder will have a fixed code word length and synchronous bit-rate. It consists of a differential coder with linear prediction and a quantizer with an additional nonlinear coder. The nonlinear coder can be considered as changing the quantizing characteristic according to the past of the signal. The improvement over PCM and over DPCM is measured by the signal to quantizing noise ratio. Picture material under investigation consists of 4 different single frames of Standard CCIR 625-line television format with relatively high detail. Two types of linear prediction have been investigated : one dimensional (previous element) and two dimensional spatial prediction. The nonlinear coder is controlled by up to three previous elements of the same and the previous line.

The measurements show an improvement over previous element DPCM of up to 9 dB S/N which would result in additional saving of 1.5 bits per coded element.

I. Introduction

Differential pulse code modulation systems have been widely investigated for efficient transmission of television signals [1][2][3]. O'Neal [2] has analyzed a DPCM System with linear prediction as shown in Fig. 1. The system consists of a linear predictor and a quantizer. The prediction coefficients are determined from the covariances of the signal samples. The prediction error e is then quantized with a quantization characteristic matched to the probability distribution of the error sequence, in order to achieve the minimum mean square quantizing noise. Such a system makes use only of the linear statistical dependencies between adjacent picture elements. Higher order redundancy can be removed by incorporating an additional nonlinear stage into the coder.

Several proposals are known, which suggest a variable quantizing range [5], depending on the past of the signal. Musmann [6][7] has proposed a coder which combines linear prediction with a variable quantizing characteristic), that is switched depending on the previous element amplitude. This type of coder has been investigated by Cohen [8] for video signals and Schlink [9] for speech signals.

In extension of the original concept of Musmann our approach will be to choose a fixed and rather simple prediction algorithm but to change the quantizing characteristic depending on various different combinations of previously transmitted values. With the perfomance criterion of a minimum mean square quantizing noise the quantizing characteristic is always closely matched to the conditional probability density of the differences which must be quantized. It is not the purpose of this paper to describe an implementation of the coder but to show the maximum gain that can be achieved and to find the most favourable source model. All results are based on statistical measurements of real picture material.

II. Signal Source and Statistical Properties

Picture material for the statistical measurements consists of 4 selected scenes as shown by Fig. 2. These photographs were scanned by a plumbicon tube scanner of European TV-standard using line interlace. The sync-less video signal was quantized linearly into 256 steps. The analog signal to noise ratio was better than 50 dB. For each photograph 512 by 512 picture elements were taken from the active part of the frame. The horizontal sampling distance and the line to line distance have a ratio of 3:2 due to the sampling rate of 10 MHz (Fig. 3). After taking the samples into the computer (DEC PDP 15) they were recorded on magnetic tape for further processing. Fig. 4 shows the amplitude distribution and the element difference distribution of the signal (4 frames). The ratio of the peak to peak composite signal to the rms value of video was 17 dB, the ratio of the rms video to the rms value of the element differences was 14 dB.

III. Linear Coding

The design of DPCM systems using linear prediction is described in the above-mentioned paper of O’Neal from 1966. Although accurate values for the prediction coefficients are given, most practical coders use only approximated values. For the previous element coder a prediction coefficient of 1 has been used by Limb and Mounts [10]. For spatial prediction the coefficients of 0.5 for elements A and D of Fig. 3 have shown profitable. Connor, Pease & Schooles [11] have subjectively tested several spatial prediction methods and have found that a prediction with (A+D)/2 is superior to other algorithms.

With respect to a simple implementation we have chosen these two prediction rules for our further investigations. For the previous element coder the prediction error to be transmitted is e = X-A, and for the spatial coder the transmitted prediction error is e = X-(A+D)/2.

The minimal noise which is introduced into the signal by quantizing these difference signals is given by the formula of Panther and Dite [12]

$$\sigma^2_{qmin} = \frac{1}{12\ K^2} \left(\int_V p^{1/3}(e)\ de \right)^3 \qquad (1)$$

where p(e) is the probability density of the quantizer input and V its range. K represents the number of quantizing levels we choose.

With e equal to the video amplitude x and a constant probability density p(x) = 1/V over the range V the optimum quantization is uniform. This is the case of straight PCM and formula (1) results in

$$\sigma_{qPCM} = \frac{V}{\sqrt{12}\ K} \qquad (2)$$

If we denote the ratio of the peak to peak composite video signal (which is 1.43 V) to rms quantizing noise a σ_q by S/N, we have (with $n = \log_2 K$)

$$S/N = 20 \log_{10} (1.43 \sqrt{12}\ K) = 13.9 + 6n \text{ (in dB)} \qquad (3)$$

for PCM video signals. For a quantization with n = 8 bits (K = 256) equ.(3) gives S/N = 61.9 dB.

For our actual computer measurements of S/N formula (1) had to be modified as the prediction error e is already in a quantized form. So the probability density function p(e) was substituted by a probability distribution P(e) and the integral by a sum. With the summation range ± M, the number of original quantizing levels, we get

$$\sigma^2_q \min = \frac{1}{12\ K^2} \left(\sum_{e=-M}^{+M} P^{1/3}(e) \right)^3 \qquad (4)$$

Evaluating eq. (4) for our picture material gave an improvement of 10.1 dB over PCM for previous element prediction and 11.0 dB for spatial prediction. This gain of DPCM over PCM is 2 dB less than the values measured by O'Neal. The difference compared to O'Neal may be explained by the non-optimum prediction ru e used here, and by variation of the picture material. Since 6 dB of quantizing noise is equivalent to 1 bit per sample, the theoretical advantage of the resulting DPCM systems would be less than 2 bit per sample. Nevertheless, reasonable or good picture quality can be achieved by linear coding with three bits or four bits respectively. This is mainly caused by the properties of the human visual perception, which are matched very well by differential coding systems. So it is reasonable to relate our following results of the nonlinear coder to the values we measured for linear DPCM to show the advantage of nonlinear over linear coding.

IV. Non-Linear Coding

With linear coding, the quantizing characteristic is fixed and depends on the amplitude distribution of the quantizer input signal. In a more general view, this distribution is variable and depends on the past of the signal. If the video process is considered as a Markov source, for each state of the source given by a certain set of previously transmitted samples, the conditional probability density of the following sample is defined. In the same way the conditional probability density of the prediction error is defined.

According to this source model, the quantizer should no longer have a fixed quantizing characteristic, but one depending on the signal past. In contrast to a normal DPCM System a combined linear and nonlinear coder would have the structure shown by Fig. 5. Here the variable quantizing characteristic is replaced by a fixed quantizer with sufficiently fine level spacing, followed by a digital coder. This coder is controlled by selected values of the preceding source output. It minimizes the noise inherent in each transmitted value, while producing always the same output bit rate. Since the quantizer control is derived from the reconstructed signal and not from the coder input, the same information is at disposal of the receiver for controlling the dequantizer, assuming the transmission has been error-free.

To determine the average quantizing noise of the nonlinear coder, we must consider the properties of the Markov source. For an N^{th}-order Markov process any state S_j of the source is defined by N-1 previously quantized samples X_n: $S_j = S(X_1,...,X^{N-1})$. The conditional probability density for the N^{th}-unqunatized sample is then:

$$p(X_N/S_j) = p(X_N/X_1,...,X_{N-1})$$

Similarly the conditional probability density of the prediction error e_N can be defined as

$$p(e_N/S_j) = p(e_N/X_1,...., X_{N-1})$$

For our measurements the value e_N is also in a quantized form, so we have a discrete conditional probability distribution for e_N:

$$P(e_N/S_j) = P(e_N/X_1,...., X_{N-1})$$

If the quantization characteristic for the state S_j is matched properly to this distribution the average quantizing noise of the state S_j follows from eq. (4)

$$\sigma_q^2 \min (S_j) = \frac{1}{12\ K^2}\left(\sum_{e_N=-M}^{+M} P^{1/3}(e_N/S_j) \right)^3 \tag{5}$$

We get the minimum quantizing noise of the whole system by averaging over all states S_j or all combinations of $X_1,..., X_{N-1}$

$$\sigma_q^2 \min = \frac{1}{12K^2} \sum_{S_j} P(S_j)\left(\sum_{-M}^{M} P^{1/3}(e_N/S_j) \right)^3 \tag{6}$$

With $P(e_N, S_j) = P(S_j) \cdot P(e_N/S_j)$ eq. (6) is reduced to

$$\sigma_q^2 \min = \frac{1}{12K^2} \sum_{S_j} \left(\sum_{-M}^{M} P^{1/3}(e_N, S_j) \right)^3 \tag{7}$$

Eq. (7) has been used for computing the minimum mean square noise of the nonlinear coding system from the measured joint probability distributions. This formula always assumes an optimal quantization characteristic which is derived according to the rules of Panther and Dite from the conditional probability distribution of the value to be coded, given the previous state S_j. The value e_N we want to quantize will be either the element difference (X-A) or the spatial difference $(X-\frac{A+D}{2})$. The states S_j will here be defined by combinations of the surrounding picture elements A,B,C,D,E as shown in Fig. 3.

V. S/N Measurements

To illustrate this further, we will look at some results we obtained for the quantizing noise σ_q. As mentioned before, previous element DPCM gives an improvement over PCM of 10.1 dB. With nonlinear coding and the knowledge of the amplitude of the element A the improvement is 12.6 dB showing a gain of 2.5 dB. To achieve this gain, the quantizer must adapt its characteristic to the value A, which means it must switch between 256 different characteristics corresponding to the 256 different levels A can assume.

When the number of characteristics is reduced to 64 or 16 by utilizing only the most significant 6 or 4 bits of A, the above-mentioned gain is reduced to 1.8 dB and 1.3 dB respectively. When instead of the amplitude A the previous difference (B-A) is known, the gain is 3.2 dB or about half a bit. If the difference (B-A) is not known, but instead the amplitudes of both B and A, then using the most significant 6 bits of both elements, the S/N improvement will be 16.5 dB over PCM or 6.4 dB over DPCM, corresponding to 1 bit per element saving, but at the expense of theoretically 4096 different quantization characteristics.

Apparently a lot of information about the coded difference is found in the preceding horizontal difference (B-A) and in the difference (A-D) between the preceding element A and element D in the line before, so the following experiment seemed reasonable : The mentioned differences were quantized logarithmically into 15 levels with the innermost three levels each equal to one original 8 bit step, and the adjacent levels doubling their step size successivly (Fig. 6). Both differences, coded with 4 bits each, together with the most significant 4 bits of the amplitude of A were used to control the quantizer characteristic. For a quantizer that is optimally matched to this signal past, described by the above 12 bits, a gain of 9.6 dB over normal DPCM can be achieved, which is a saving of 1.6bits/sample.

So far only the previous-element prediction has been incorporated in the nonlinear coder system. In a similar way we have explored the prediction of (A+D)/2, which gives a significant improvement of picture quality and has the other advantage of reducing the visibility and propagation of eventual transmission errors. This prediction seems to be at least equivalent to various other more complicated spatial predictions. This improvement in picture quality is not so evident from our S/N measurements. Here the gain over PCM of a normal DPCM coder with prediction of (A+D)/2 is only 11 dB in comparison with 10 dB for a previous-element coder. Nevertheless, the measurements were carried on to see if the difference will increase or decrease with nonlinear coding. We will find only very little change in the additional gain of nonlinear coding, but the initial difference of 1 dB remains visible in most cases.

The element A alone when used to control the quantization characteristic adds only 2 dB to the initial 11 dB, significant improvement again comes from two separate elements. Here the elements A and E have proven to be the best choice of any two of the surrounding elements. With both values coded by 6 bits the gain is 6.4 dB and identical to previous element prediction. This choice is also better than that of the 3 separate elements A,C and D each coded with 4 bits, which gives the same total number of bits, but a gain of only 6.1 dB. This again leads to the supposition that not alone are the absolute values of the elements X significant, but also their differences, especially when these are small. To utilize this, the measurements were continued in the following way. The differences of (C-A) and (C-D) were formed and requantized logarithmically in the same way as described before (Fig.6), each producing 4 bits of information. Together with the most significant 4 bits of the amplitude of A they served as a combination of the signal past for controlling the quantizer. The S/N improvement over PCM was 19 dB for this case or 8 dB over the linear prediction alone. When (C-A) and (C-D) are replaced by (B-A) and (A-D) we get an improvement of 19.4 dB which is very close to the corresponding result of the previous element prediction.

At this point the results of our hitherto performed measurements have been presented. Table I shows a summary of all numerical results. We have seen that considerable improvement of the S/N ratio results from adapting the quantizer characteristic to the probability distribution of the input value which has to be processed. The amplitude distribution forming the characteristic of a DPCM quantizer is well known and correponds to the difference probability of Fig. 4. But what happens to the probabilities if they are conditioned by special states of the signalpast? Although these question has not yet been investigated quantitatively, our measurements have shown clearly different effects. Most obvious is the sharpening or flattening of the distribution, while still centered around zero (Fig. 7a). Use of a variable preamplifier at the quantizer input has been proposed to match the input distribution to the quantizer characteristic [5]. Another effect is the displacement of the distribution, caused by a non-sufficient prediction rule (Fig. 7b). In these cases the conditional expectation differs from the predicted value. These errors are impossible when using an optimal nonlinear prediction. The third important case is given by asymmetries in the distribution (Fig. 7c). A simple reason for these can result from the fact that the previous value is already near to the limit of the amplituderange. In practice, all these effects and perhaps numerous others overlay each other, making a quantitative analysis for the purpose of simplifying the nonlinear coder design rather difficult.

VI. Conclusions

The measurements described above have been based on a special system consisting of linear prediction and nonlinear quantization. In fact they have a more general meaning. The S/N values derived by formula (7) do include all the gain achievable by optimum nonlinear prediction. On the other hand, the linear prediction part of the system becomes less useful, as more information is processed in the nonlinear coder, especially if the same information is used by both parts.

Thus in this work the gain of optimum nonlinear quantization as well as optimum prediction has been,determined. With the restriction of using not more than 12 bits of information about the signal past, combinations of past values were found that seem near to optimum. If a nonlinear quantizer is controlled by this informations, it gives an improvement of 9 dB over a fixed quantizer. To test the significance of this statement, we are planning to do a simulation of the nonlinear coder with actual quantization of 1,2 or 3 bits for single frames. With the result of this simulation and a quantitative analysis of the conditional probability distributions a realtime-coder shall be built which will be able to process broadband television signals.

Table I - Values of the improvement of the S/N ratio of the nonlinear coder as compared to straight PCM.

Z (Combination of previous elements used for controlling the quantizer, see Fig. 3)	$\sigma_{qmin}(e\|Z)$ $e = X - A$	$\sigma_{qmin}(e\|Z)$ $e = X - \frac{A + D}{2}$
None (Linear DPCM)	10.1 dB	11.0 dB
A_8	12.6	13.0
$e'_9 = A - B$ (previous prediction error)	13.3	-
$e'_9 = A - \frac{B + E}{2}$ (previous prediction error)	-	15.2
A_6, B_6	16.5	-
A_6, E_6	-	17.4
A_4, C_4, D_4	-	17.1
$A_4, (C - A)_{4L}, (C - D)_{4L}$	-	19.0
$A_4, (B - A)_{4L}, (A - D)_{4L}$	19.6	19.4

Subscripts show number of most significant bits of the element values used for controlling the quantizer. Subscript L means logarithmic quantization of the element value according to Fig. 6.

References

1. 1C. C. Cutler, Differential Quantization of Communication Signals, Patent No. 2 605 361, July 1952 (applied for June 1950).

2. J. B. O'Neal, "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals, Bell Syst. Techn. J., Vol. 45, pp 689-721, May 1966.

3. J. B. Millard and H. I. Maunsell, "Digital Encoding of the Video Signal", BSTJ 50, Feb. 1971, pp 459 - 497.

4. A. V. Balakrishnan, R. L. Kutz and R. A. Stampfl, "Adaptive Data Compression for Video Signals", Washington, D.C., NASA Technical Note TN - D 3395, April 1966.

5. L. S. Golding and P. M. Schultheiss, "Study of an Adaptive Quantizer", Proc. IEEE, Vol. 55, pp 293 - 297, March 1967.

6. H. G. Musmann, "Codierung von Videosignalen", NTZ, Vol. 24, pp 114 - 116, February 1971.

7. H. G. Musmann, “Theoretical Aspects of Intraframe Coding”, Deutsche Luft- und Raumfahrt, Forschungsbericht 1972.

8. F. Cohen, “A Nonlinear Coding System Using the Principle of a Switched Quantizer for the Transmission of Television Signals”, to be published in NTZ, Germany, 1972.

9. W. Schlink, “A Redundancy Reducing PCM System for Speech Signals”, Proc. Intl. Zurich Seminar on integrated Systems for Speech, Video and Data Communications, March 1972, pp F4/1 - 4.

10. J. O. Limb and F. W. Mounts, “A Digital Differential Quantizer for Television”, BSTJ 48, Sept. 1969, pp 2583 - 2599.

11. D. J. Connor, R. F. W. Pease and W. G. Scholes, “Television Coding Using Two Dimensional Spatial Prediction”, BSTJ 50, March 1971, pp 1049 - 1061.

12. P. F. Panter and W. Dite, “Quantizing Distortion in Pulse Count Modulation with Nonuniform Spacing of Levels, Proc. IEEE, Vol. 39, pp 44 - 48, January 1951.

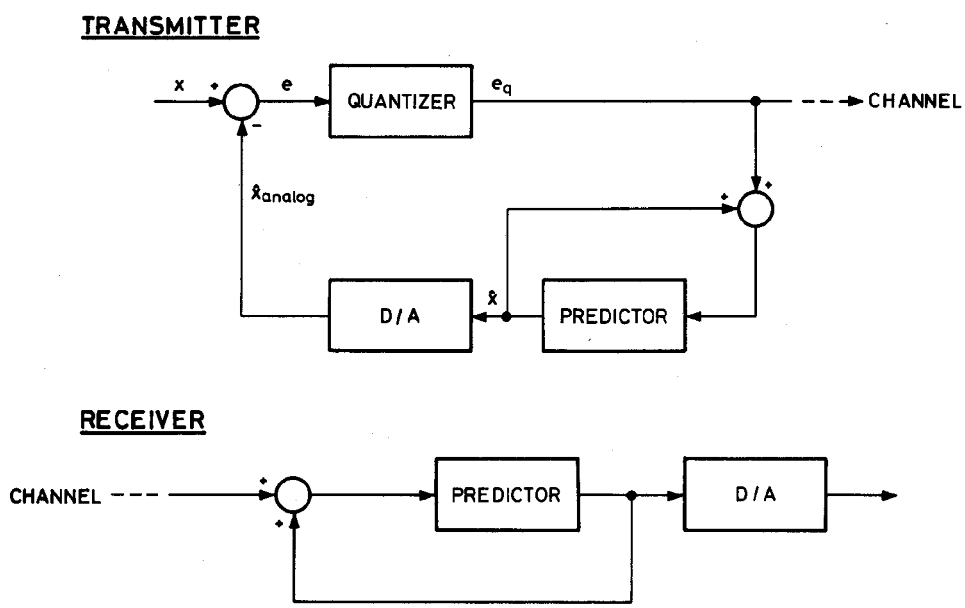

Fig. 1 - Block diagram of linear DPCM encoder and decoder

Fig. 2 - Pictures of four slides scanned to obtain television signals

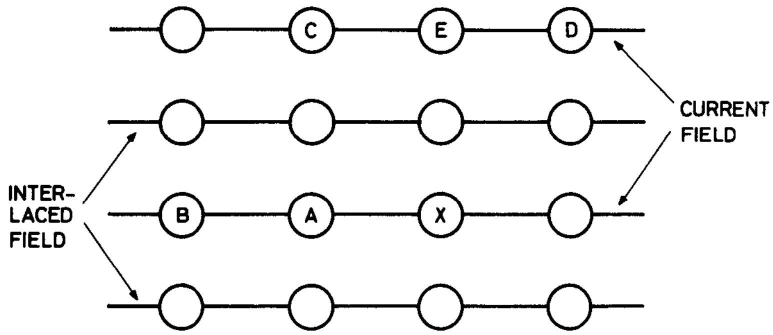

Fig. 3 - Sampling pattern and denotation of elements used for prediction and coder control

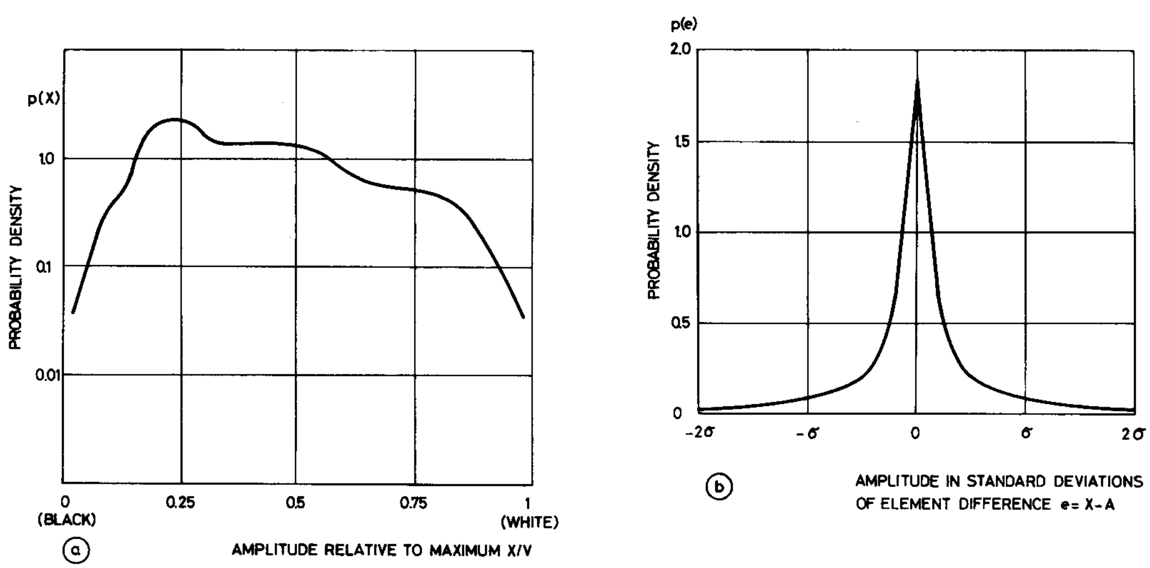

Fig. 4 - Probability distribution of video amplitude and element difference obtained from the four scenes of Fig. 2

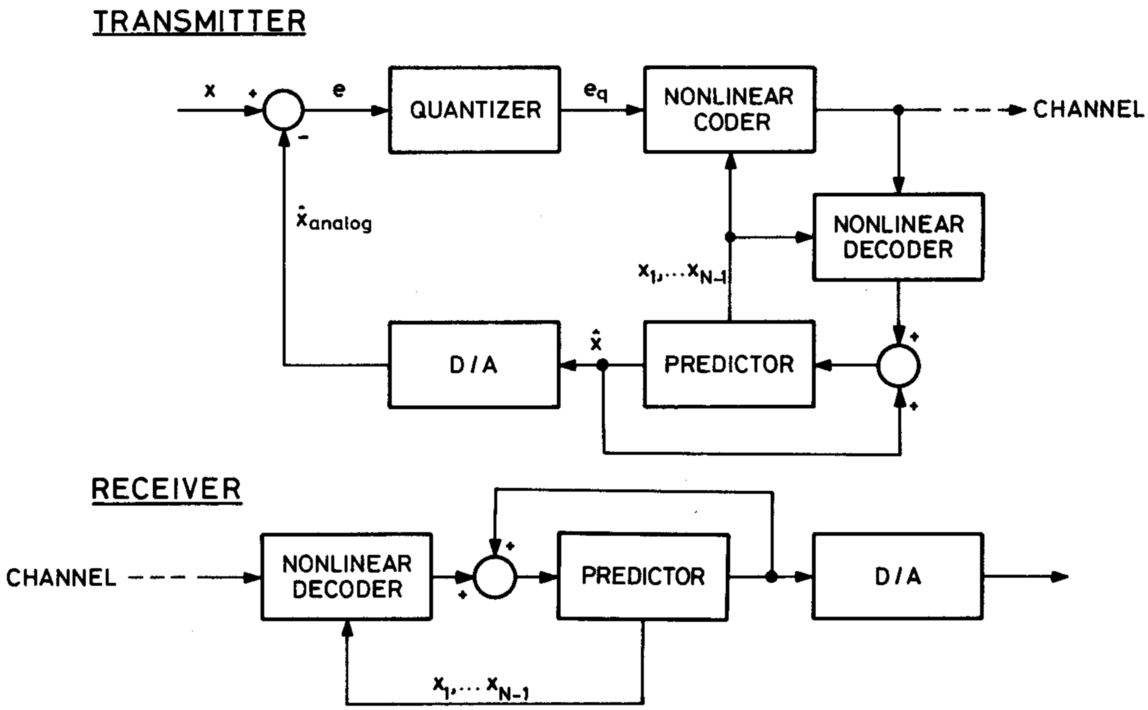

Fig. 5 - Block diagram of a combined system of linear and non-linear DPCM encoder and decoder

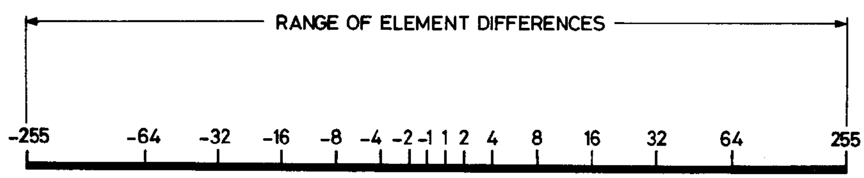

Fig. 6 - Quantization characteristic of element differences as used for nonlinear coder control

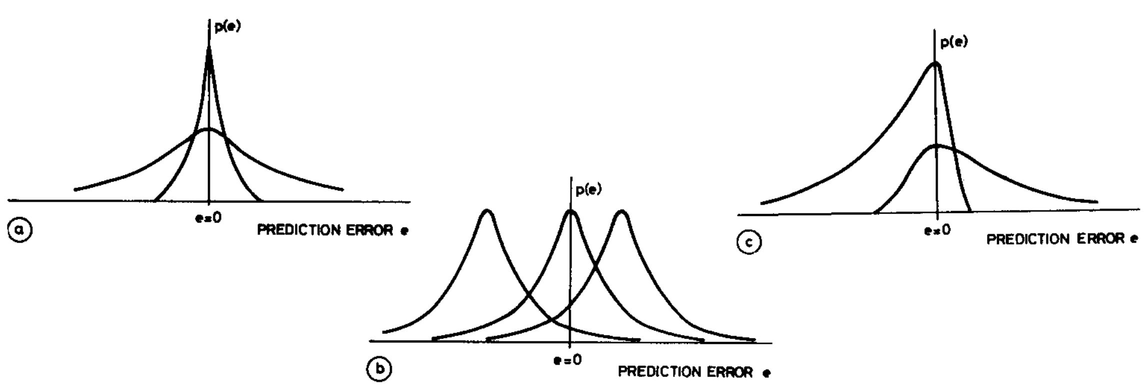

Fig. 7 - Different types of conditional probability density functions of the prediction error e