APPLICATION OF RATE DISTORTION IN IMAGE CODING

By

K. FAZEKAS

Department of Microwave Telecommunication,
Technical University, Budapest

Received October 5, 1979
Presented by Dr. L. PÁSZTORNICZKY

In most of different digital intraframe image coding systems, developed in the recent decade, the linear predictive coding and transforming procedures are applied. In realizing these systems, the minimum ms (mean square) error is used as quality criterion or its term expressed in decibels as an equivalent signal-to-noise ratio respectively, because of its relative simple analitical handling. Thus defined quality measure is applied to compare systems.

The signal-to-noise ratio is known not to be an adequate measure of image quality, in final account it is defined by subjective tests. The rate distortion theory is applicable to give the bound of the subjectively determined quality measure in the first place for communication purposes.

The rate-distortion function $R(D)$ gives the information rate R required for the transmission of a signal, when the distortion introduced by the digitizing process is measured by the quantizing noise power D. The rate-distortion function can be defined from the minimization of mutual information function $I(x, y)$ (Fig. 1) for a maximum acceptable distorion D'. In general, however, the minimization of mutual information function with the constraint, that the average distortion must not exceed an acceptable value D', is a fairly difficult problem. In this paper only a bound applicable for image sources will be discussed.

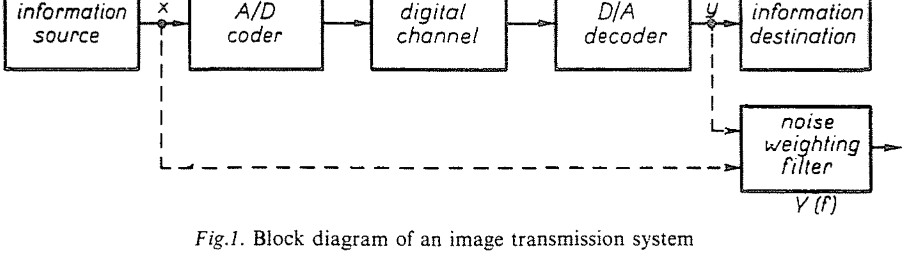

Fig.1. Block diagram of an image transmission system

One of the significant application areas of digital image coding systems is the transmission of television signals. For television signals, noise at low frequencies is more damaging than noise at higher frequencies. Frequency-weighting networks have been designed on the basis of subjective tests, that can be used to accurately measure the damaging effect of noise. In this manner the ratio of signal power to frequency-weighted noise power can be defined, which is a much better measure, than the usual signal-to-noise ratio.

The ratio of signal power to frequency-weighted noise power by O'Neal can be expressed as the sum in decibels:

$$S/N_y = T_b + T_p + T_s \tag{1}$$

where T_b is the bit rate of system, T_p is the predictability of signal, and T_s is a quantity determined by subjective nature of the user. The simplified block diagram of the image transmission system and of the frequency-weighting filter are shown in Fig. 1. The information source emits a band-limited signal, with a power spectrum $S(f)$. The power spectrum of quantizing noise caused by digitization operation, is $N(f)$. The damaging effect of noise depends on its power spectrum. This effect can be measured by means of frequency-weighting filter. The upper bound of the ratio of signal power to frequency-weighted noise power is for this system

$$S/N_y \leqq 6\frac{R}{2B} - 10\lg K_p - \frac{1}{B}\int_0^B 10\lg Y(f)\,\mathrm{d}f. \tag{2}$$

The first term T_b is as six times the number of bits. K_p is the predictability of signal and it is determined by the redundancy of the signal. K_p is a constant less than 1, therefore T_p is positive. The third term T_s is positive, since $Y(f)$ is less than 1. The determination of T_p and of K_p is fundamental in practical predictive communication systems.

1. Determination of the constant K_p for predictive system

The predictability constant is the ratio of entropy power of the signal to signal power

$$K_p = Q_s/P_s$$

the quota how the signal power can be reduced by linear prediction. For Gaussian signals the constant K_p can be computed according to the theory of linear prediction in time domain, or the power spectrum of the signal in frequency domain.

The Figure 2 shows the simplified block diagram of predictive communication system. The sampling of stationary signal $S(t)$ at twice its bandwidth produces the sequence of sample values $\{S_i\} = S_0, S_1, \ldots$. At the same time, the predictor composes the sequence $\{\hat{S}_i\} = \hat{S}_0, \hat{S}_1, \ldots$ based on a linear estimate of each sample value. The linear estimate of the next sample value S_j is based on the previous sample values $S_{j-1}, S_{j-2}, S_{j-n}$ according to the next relation

$$\hat{S}_j = a_1 S_{j-1} + a_2 S_{j-2} + \ldots + a_n S_{j-n}. \tag{3}$$

By substracting $e_j = S_j - \hat{S}_j$ the transmitted error sequence $\{e_i\} = e_0, e_1, \ldots$, is derived and the receiver uses this sequence to reconstruct the original analog signal $S(t)$. The constants $a_1, a_2, \ldots a_n$ are determined so, that each $\hat{S}_j$ is the best linear ms estimate of S_j. This minimizes the expected ms value of the sequence $\{e_i\}$ and causes the members of this sequence to be independent. The constants $a_1, a_2, \ldots a_n$ are determined from the autocorrelation function of the signal $S(t)$.

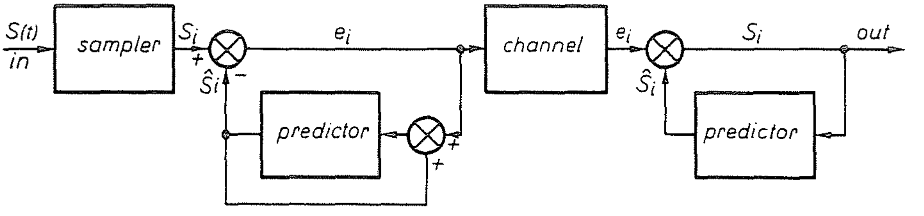

Fig. 2. Model of the predictive system

If $S(t)$ is a kth-order Markov process then n, the number of constants required to achieve the optimum estimate $\hat{S}_j$ is equal to k. The entropy and the entropy power of this sequence $\{e_i\}$ will be equal to that of the signal $S\{t\}$, since the transformation from set $\{S_i\}$ to the set $\{e_i\}$ is measure preserving. Is the signal $S(t)$ a Gaussian one, then so are the sequencies $\{S_i\}$ and $\{e_i\}$. Since the prediction process causes the members of $\{e_i\}$ to be independent, their entropy power equals that of the signal, so $Q_s = E\{e_i^2\}$, the mean square value of the error sequence.

Considering in the frequency domain a Gaussian signal with power spectral density $S(f)$, we can write

$$\lg K_p = \frac{1}{B}\int_0^B \lg S(f)\mathrm{d}f - \lg\frac{1}{B}\int_0^B S(f)\mathrm{d}f. \tag{4}$$

Noll obtained similar results in his comparative study of predictive quantizing systems.

2. Some application of rate distortion function

In designing a practical system, the rate distortion function determined on a given information source, is an upper bound on the ratio of signal power to frequency-weighted noise power of the system.

The image sources often are modelled as Markov process, so in this case the obtained results are specially interesting.

For coding a one-dimensional Markov source the rate distortion function is

$$R(D') = \frac{1}{2}\, ld \left[\frac{\sigma^2(1-\rho_h^2)}{D'} \right] \tag{5}$$

and for two-dimensional coding

$$R(D') = \frac{1}{2}\, ld \left[\frac{\sigma^2(1-\rho_h^2)(1-\rho_v^2)}{D'} \right] \tag{6}$$

where σ^2 is the variance, ρ_h and ρ_v are the horizontal and vertical correlation factors.

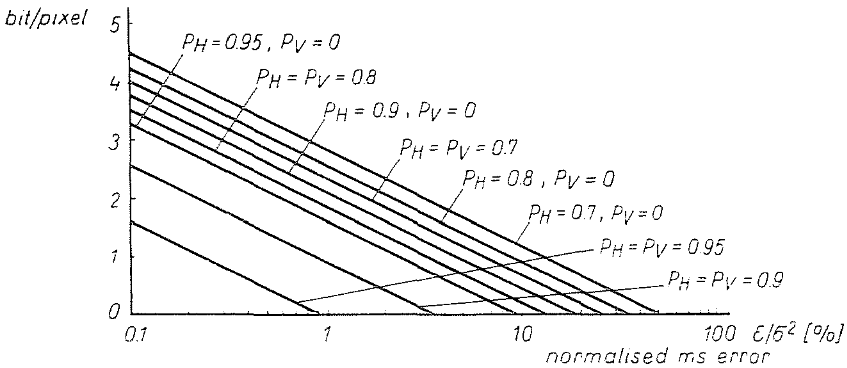

Fig. 3. The rate-distortion function in case of signals with different correlation coefficients

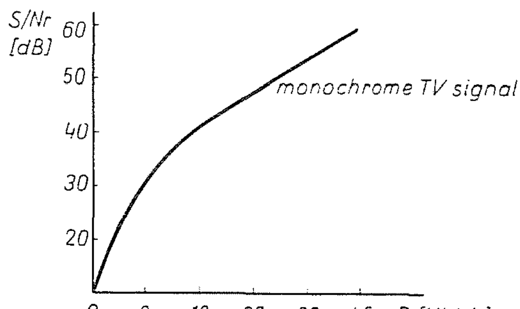

Fig. 4. The rate-distortion function in case of monochromatic TV signals

These relations give accurate results at low distortion. The Fig. 3 gives the rate distortion function for various correlation factors. For monochrome television signals, the value computed at $\rho_h = 0.96$ is in good agreement with experimental results. The function for TV signals has been plotted separately in Fig. 4. The S/N_y plotted in Fig. 4 are upper bounds, consequently the ratio of signal power to frequency-weighted noise power of a given coder falls always below the bound.

For colour images, the appropriate frequency-weighting and the rate distortion function $R(D)$ can be determined on the basis of a colour perception model, of course now parametrically corresponding to the three components of colour signal. For colour television signals the realization problems of coder are simpler to solve relying on the intensity and the colour-difference signals.

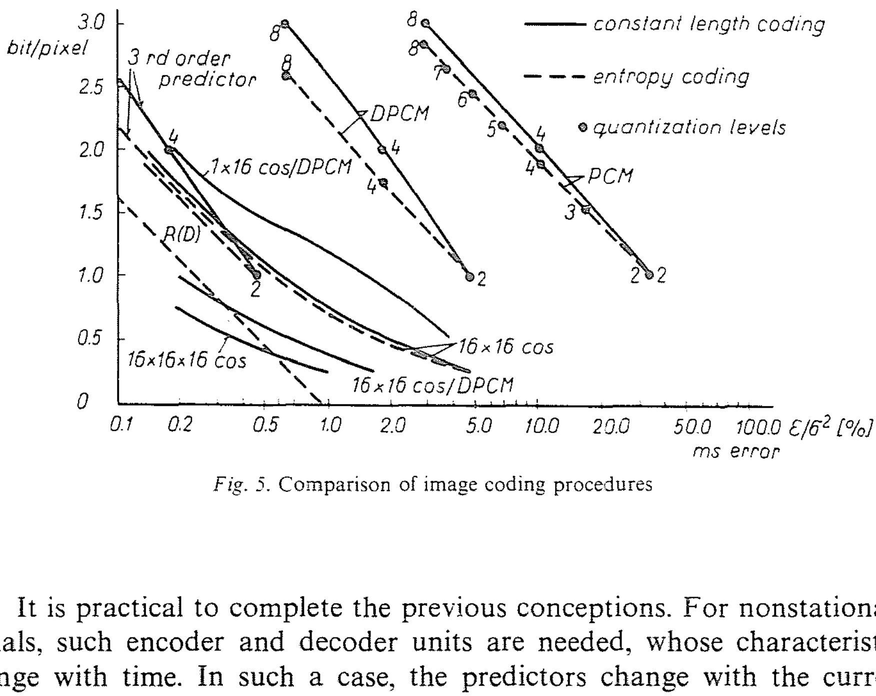

Fig. 5. Comparison of image coding procedures

It is practical to complete the previous conceptions. For nonstationary signals, such encoder and decoder units are needed, whose characteristics change with time. In such a case, the predictors change with the current statistical character of signal in encoder and decoder (adaptive coding), and information on this change must be transmitted to the receiver. Thus, some fraction γ of the bit rate R is used to specify the nature of redundancy and T_p becomes a function of γR. Taking this into account, it can be written:

$$S/N_y \leqq 6\,\frac{(1-\gamma)R}{2B} + T_p'(\gamma R) + T_s\,. \tag{7}$$

In designing practical systems, the value γ should be chosen to maximize S/N_y. In every case, $T_p' \leqq T_p$ and the relation (2) is applicable for nonstationary signals as well as for stationary ones.

Finally, Fig. 5 it is a comparison between various types of image coding systems and gives the realization bounds.

Summary

The rate-distortion function $R(D)$ gives the information rate R required for the transmission of a signal, when the distortion introduced by the digitizing process is measured by quantizing noise power D. In general, the evaluation of rate-distortion function is a difficult problem. In this paper only a procedure adaptable in designing of predictive encoders will be discussed.

Kálmán FAZEKAS, H-1521 Budapest