NEUVIEME COLLOQUE SUR LE TRAITEMENT DU
SIGNAL ET SES APPLICATIONS

NICE du 16 au 20 MAI 1983

Waveform Coding with Multipath Search Techniques

Heinz Georg Fehn

University of Bremen, Department of Electrical Engineering, Universitätsallee, 28 Bremen 33

RESUME

Pour des vitesses de transmission réduites (R = 1-2 bit/sample) la qualité de toute méthode de codage dépend principalement des propriétés du quantifieur. L'erreur de codage peut être réduite en utilisant les "multipath search techniques". Cette publication démontre quelles améliorations on peut obtenir grace aux "multipath search techniques": 1) quantification du vecteur, 2) codage Tree, 3) codage Trellis pour une vitesse de transmission R = 1 bit/sample. On y compare les résultats obtenus par les "multipath search techniques" avec les méthode traditionelles utilisées pour le codage des sources stationnaires.

SUMMARY

In the low bitrate region the performance of any coding scheme is crucially dependent on the properties of the quantizer. The coding error can be reduced by using multipath search techniques. This paper demonstrates the performance of three multipath search techniques: 1) vector quantization (codebook coding), 2) tree coding and 3) trellis coding working on a bitrate R = 1 bit/sample. The improvements in multipath search techniques are explained in relation to conventional coding schemes using examples of stationary sources.

Waveform Coding with Multipath Search Techniques

H. G. Fehn

1. Introduction

At low bit-rates the performance of any coding scheme is crucially dependent on the properities of the quantizer. The poor performance of any coding scheme when operating at a rate R of 1 bit/sample is the reason to reconsider the possibilities of using multipath search techniques which, in contrast to conventional strategies, are based on a delayed decision about binary data representing waveform samples. Conventional coding schemes are based on instantaneous decision: the encoder converts an input sample x_n at time n into a channel codeword c_n which contains information about x_n (as in Puls-Code-Modulation (PCM)) or on x_n and its predecessors x_{n-1}, x_{n-2}, ... (as in Differenz-PCM (DPCM))/1/. The decoder converts the received channel codeword c_n into an output sample y_n. In multipath search techniques schemes (MSC), on the other hand, future values x_{n+1}, x_{n+2}, ... are considered as well, before a (delayed) decision is made about the optimum c_n to be released.

Multipath search techniques schemes can be divided into three classes i) vector quantization /2/ (codebook coding), ii) tree coding and iii) trellis coding. The classes of MSC schemes are given by the arrangement of the output sequences. In vector quantization schemes the set of output sequences is arranged in a codebook whose elements are not restricted in any way. In tree and trellis coding schemes the output sequences are arranged in the form of a tree or trellis. The arrangement of output sequences in MSC schemes working at a rate R of 1 bit/sample is shown in Fig. 1.

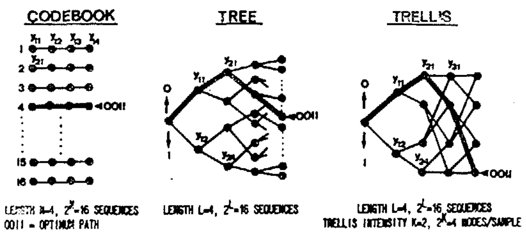

Fig. 1. Classes of MSC schemes.

The heavy line indicates a codebook sequence $\underline{y}$ of length N and a path through the tree and trellis of depth L described by the binary channel sequence OOII.

The main structure of MSC schemes shows Fig. 2. Samples x_n of the analog input signal are fed into the buffer of length N. Let vector $\underline{x}$ describe the bufferd input sequence. The encoder compares $\underline{x}$ with a collection of possible output sequences $\underline{y}_k$, $k=1,2,\dots,2^N$ where $\underline{y}_k^T=(y_{k1}, y_{k2}, \dots, y_{kN})$. The collection of these sequences which are either stored or deterministically

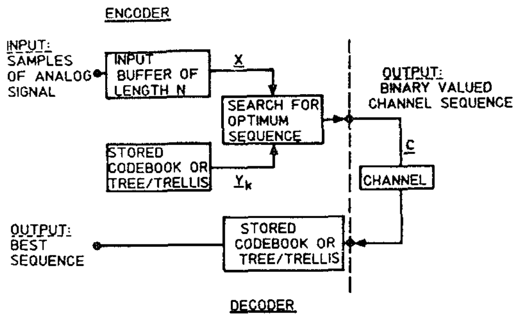

Fig. 2. Main structure of MSC schemes.

generated when needed must be available at both the transmitter and receiver. The optimum output sequence is the nearest neighbor sequence, i.e. the sequence with the smallest squared error

$$E_k = (\underline{x}_k - \underline{y}_k)^T(\underline{x}_k - \underline{y}_k) \qquad (1)$$

The decoder is informed about the chosen output sequence (i.e. the sequence that best describe the bufferd input) by a binary channel sequence $\underline{c}$. On the basis of this sequence the decoder outputs the corresponding output sequence or at least portions of it /3,4/.

2. Conventional Source Coding

At this point it is useful to review the main results on source coding of stationary zero-mean random sequences, either with Gaussian (G) or gamma (Γ) shaped probability density function (pdf), and to compare them with the information-theoretical bounds. Conventional source coding gives us the upper-bound performance of multipath search techniques, whereas the information-theoretical results lower-bounds the performance of any coding scheme.

As a criterion of performance the mean-squared value of the reconstruction error $r_n=x_n-y_n$, i.e. the mean-squared error (MSE)

$$\sigma_r^2 = E[r_n^2] = E[(x_n-y_n)^2] \qquad (2)$$

will be used throughout this paper.

2.1 Memoryless Sources

For a given average rate R (bits/sample) the lower bound D(R) of the mean-squared error is given by the rate-distortion theory as /5/

$$D_G(R) = 2^{-2R}\sigma_x^2 \qquad (3a)$$

$$D_\Gamma(R) \leq D_G(R) \qquad (3b)$$

These values lower-bound the MSE performance of any coding scheme operating on the appropriate i.i.d. source samples of variance σ_x^2. The MSE produced by a memoryless quantizer is

$$\sigma_q^2(R) = \epsilon_q^2(R)\,\sigma_x^2 \qquad (4)$$

where $\epsilon_q^2(R)$, the quantizer performance factor, depends on the pdf of the input samples and the number 2^R of quantizer steps. Tab. 1 shows the

Waveform Coding with Multipath Search Techniques

H. G. Fehn

rate-distortion bounds D(1) and the quantizer performance factor $\epsilon_q^2(1)$ for a memoryless source with Gaussian- and gamma-pdf, resp.

pdf	$\epsilon_q^2(1)$	$D(1)/\sigma_x^2$
Gaussian source $\frac{1}{\sqrt{2\pi}\sigma_x}\exp(-x^2/2\sigma_x^2)$	0.363	0.25
Gamma source $\frac{\sqrt[4]{3}}{\sqrt{8\pi\lvert x\rvert\sigma_x}}\exp(-2\lvert x\rvert/\sigma_x)$	0.667	0.139

Tab. 1. MSE performance for memoryless sources; Bitrate R=1 bit/sample /5,6/.

2.2 Sources with Memory

As a second step consider sources with memory. Only in the case of Gaussian autoregressive sources the rate-distortion bound is known and given by /5/

$$D_G(R) = 2^{-2R}\sigma_i^2 \qquad (5)$$

The variance $\sigma_i^2=\gamma_i^2\sigma_x^2$ is given by the source variance and the spectral flatness measure γ_i^2 /7/. For example consider a first-order autoregressive source (ar(1)-source) with an autoregressive constant ρ which defines the normalized correlation between neighboring samples. In this case the spectral flatness measure yields $\gamma_i^2=1-\rho^2$ and the rate-distortion bound reads

$$D_G(R) = 2^{-2R}(1-\rho^2)\sigma_x^2 \qquad (6)$$

if $R\geq \log_2(1+\rho)$. For non-Gaussian sources $D_G(R)$ serves again as an upper bound, i.e. smaller distortions can be expected for non-Gaussian sources of same variance.

In a DPCM-coding scheme the difference d_n between input x_n and its predicted value $\hat{x}_n$ is coded and transmitted. Assume an error-free channel, then the reconstruction error r_n equals the quantization error q_n between quantizer input and output. Therefore the MSE reads

$$\sigma_q^2(R) = \epsilon_q^2(R)\,\sigma_d^2 \qquad (7)$$

where $\sigma_d^2=E[(x_n-\hat{x}_n)^2]$ is the variance of the DPCM difference signal. Note that the quantizer performance factor will depend on the pdf of the difference signal. In the case of a Gaussian ar(1)-source and optimum first-order prediction with h=ρ /8/, i.e. the estimation of x_n is $\hat{x}_n=h\,y_{n-1}$, the mean-squared error is

$$\sigma_{r,G}^2(R) = \epsilon_q^2(R)\,\frac{\gamma_x^2}{1-\epsilon_q^2(R)\rho^2}\,\sigma_x^2 \qquad (8)$$

which can be considerably higher than $D_G(R)$ of (5) if the correlation ρ between neighboring samples is high and if R is small.

The difference between the rate-distorion bound and DPCM performance can be explained by the fact that DPCM has contraints of the quantizer performance factor $\epsilon_q^2(R)\geq 2^{-2R}$ and of the difference signal variance $\sigma_d^2\geq\gamma_x^2\sigma_x^2$ due to quantization noise feedback inherent in DPCM loops. The estimate $\hat{x}_n$ of x_n is based on their quantized version, i.e.

$$\hat{x}_n = f(y_{n-1}, y_{n-2}, \ldots) \qquad (9)$$

instead of

$$\hat{x}_n = f(x_{n-1}, x_{n-2}, \ldots) \qquad (10)$$

This in fact yields a loss in prediction gain described by the denominator of (8). With multipath search techniques it is possible to reduce the loss in prediction gain such that

$$\sigma_r^2(R) = \epsilon_q^2(R)\,\gamma_x^2\sigma_x^2 \qquad (11)$$

and it is also possible to narrow the gap between $\epsilon_q^2(R)$ and information-theoretical limit D(R) such that

$$\sigma_q^2(R) \rightarrow D(R) \qquad (12)$$

3. Multipath Search Techniques

3.1 Vector Quantization

The vector quantization (codebook coding) follows Fig. 2. In a codebook 2^N typical output sequences $\underline{y}_k$, $k=1,2,\ldots,2^N$ are stored. Each of these sequences is indicated by a N-bit index. The N-length buffered input sequence $\underline{x}$ is compared with these sequences on the basis of the squared error defined in (1). An exhaustive search leads to the optimum output sequence defined by $\min_k\{E_k\}=E_j$. Its index j is transmitted as an N-bit channel sequence and the decoder can output the optimum sequence $\underline{y}_j$ if it has access to an identical codebook. The most important step is obviously to find the appropriate codebook. With restricted output values $y_{kl}\in\{\pm\Delta\}$, $k=1,2,\ldots,2^N$; $l=1,2,\ldots,N$ the codebook coding procedure is that of PCM. By choosing $\Delta=E[|x_n|]$ we would obtain the quantizer performance factor given in Tab. 1. However, the codebook should contain a set of typical, i.e. highly probable sequences. Based on a generalization of iterative procedures given by LLoyd and Max /9/,/10/ codebooks can be produced which are either based on known probabilistic models or on a training sequence of data /2/. It is also possible to pick the elements of the codebook as generated by a pseudo-random-noise generator and to find a good codebook by a trial-and-error method. Hints about the appropriate pdf can be drawn from the results obtained from the Blahut algorithm /11,12/.

3.2. Memoryless Sources

The simulation of MSC schemes was first done with a memoryless gamma source. Ten input waveform, each consisting of 32 000 samples, were used and the results were averaged. Fig. 3.

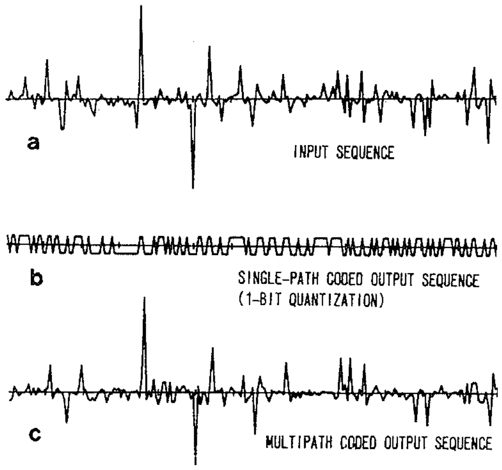

Fig. 3. Comparision of waveforms; memoryless gamma source.

shows a segment of the waveform of an input sequence with its corresponding output sequence of a memoryless quantizer and vector quantization (R=1 bit/sample, N=8). Fig. 4 compares the normalized error variance $\epsilon_q^2(1)$ of vector quantization a memoryless gamma source with i) iterative codebooks and ii) randomly generated (stochastic) codebooks. In both cases

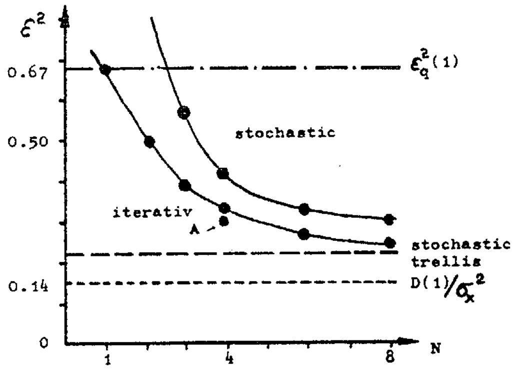

Fig. 4. Normalized performance factor $\epsilon_q^2(1)$ vs. codebook length N; memoryless gamma source, Bitrate R= 1 bit/sample.

of codebook construction the normalized error variance $\epsilon_q^2(1)$ decreases with increasing codebook length N. For example with N=8 the normalized error variance $\epsilon_q^2(1)$ yields 0.26 instead of $\epsilon_{q,\eta}^2(1)=0.667$ in the case of conventional quantization. The decreasing error variance $\epsilon_q^2(1)$ for gamma sources are mainly due to the fact that infrequently occuring samples of high amplitude find the appropriate reconstruction value (compare Fig. 3.). Examination of the codebook sequences have shown that in very many cases each sequence consists of one high amplitude value and N-1 values of typical small amplitudes most of them close to zero. For large N the distribution tends to be close to that given by the Blahut algorithm /12-14/.

A scheme has been proposed where N-1 elements of $\underline{y}$ are set to zero. For example in a codebook of length N=4 the first eight sequences have only one nonzero value Δ ; the second eight sequences are filled similary with one nonzero value Δ^*. This five level codebook with elements $0, \pm\Delta, \pm\Delta^*$ is easy to implement. No time consuming iterative procedures or trial-and-error methods (in stochastic codebook generation) are nescessary to find an appropriate codebook. The optimum

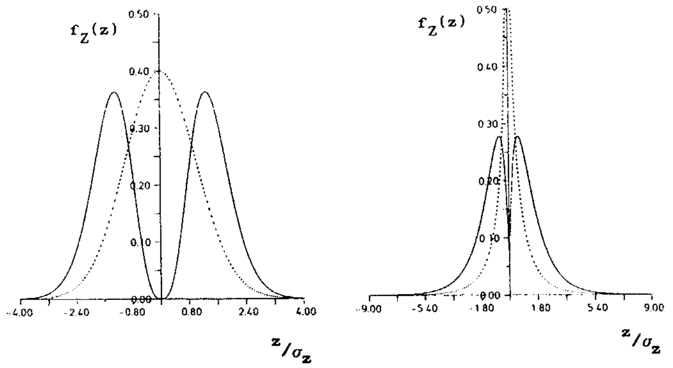

Fig. 5. Resulting pdf for optimization the nonzero values in codebook coding; memoryless sources with a) Gaussian and b) gamma pdf (the input pdf is shown by the dotted line).

nonzero values can be found by optimization of the reconstruction values of a N-level quantizer for a memoryless source with samples z_j. The samples z_j can be determine from the input source samples x_i as

$$|z_j| = \max\{|x_i|\} \quad i=1,2,\ldots,N \quad j=1,2,\ldots \qquad (13)$$

The optimization of Δ, Δ^* must be done with respect to the pdf of the quantizer input samples z_j (Max-quantizer /10/). The pdf of the samples z_j for Gaussian and gamma input sources are shown in Fig. 5. With optimized values $\Delta=3\sigma_x$ and $\Delta^*=1\sigma_x$ a normalized error variance $\epsilon_q^2(1) =$ 0.3 was realized for a memoryless gamma source (Point A in Fig.4). Evaluation on the pdf of the reconstruction sequences shows, that the pdf tends to optimum pdf obtained by the Blahut algorithm.

3.3 Tree and Trellis Coding

Tree and trellis coding algorithms are most easily understanding by considering the generalized DPCM decoder structure of Fig. 6.
At the decoder any binary channel sequence of channel codewords $c_n \epsilon \{0, I\}$ is D/A converted to an innovation sequence i_n of analog values taken from a 2^R-ary alphabet. If R=1 bit/sample we have $i_n \epsilon \{\pm\Delta\}$; in adaptive controlled systems these values may differ from sample to sample. The innovation sequence is fed into a linear filter (time-invariant or adaptive) whose output is the reconstruction sequence y_n. The set of all sequences can be arranged in a tree or

Waveform Coding with Multipath Search Techniques

H. G. Fehn

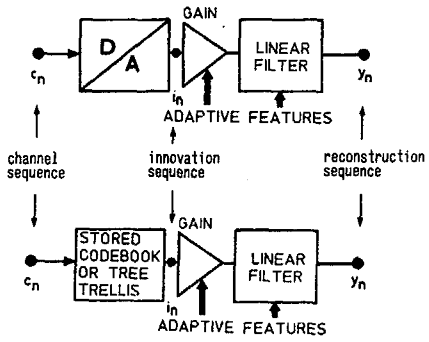

Fig. 6. Generalized DPCM structure.

trellis structure having two branches per node if the channel codewords are binary. Each channel sequence defines a unique path through the innovation tree (trellis) and the corresponding - generated via the linear filter - reconstruction tree. In a tree coding scheme the number of nodes increases exponentially with the tree depth L. In order to limit the number we can make use of a specific structure called trellis (Fig. 1). Here the number of nodes is limited to 2^K per sample (where K is called the intensity of a trellis). During the initial fan-out, i.e. as long as the trellis depth L is less than K, the number is smaller and given by 2^L. In other words, the trellis starts as a tree which then collapses to a specific structure of the trellis if $L > K$.

Using the channel sequence c_n calculating the innovation sequence on a basis of a given algorithm known at the encoder and decoder a deterministic tree (trellis) coding scheme results. on the other hand, in a stochastic scheme the channel sequence c_n is used to determine innovation values out of a set of random variates either stored or generated when needed by a pseudo-random-noise generator. The statistics of the random variates, populating the nodes of the tree (trellis), should by determined beforehand with respects to the statistics of the source/12, 15,16/. An example for the potential of stochastic trellis coding memoryless gamma sources is given in Fig. 3 (stochastic trellis).

3.4. Sources with Memory

A few examples may show the potential of tree and trellis coding. Fig. 7 shows the results to encode ar(1)-sources with Gaussian and gamma innovation, resp. The improvements in signal-to-noise ratio ($SNR = 10\log_{10} \sigma_x^2/\sigma_r^2$) in tree coding schemes over conventional DPCM (equivalent to tree depth L=1) yields with increasing depth L up to 1.3 dB and 2.9 dB, resp. These improvements are in agreement with the loss in prediction gain in dB ($-10\log_{10}(1-\epsilon_q^2(1)\rho^2)$). The comparision of the prediction gains (GPR) of Gaussian ar(1)-sources with various correlation

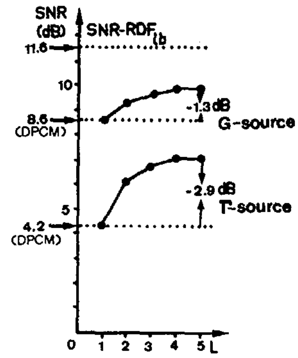

Fig. 7. SNR in tree coding schemes vs. tree depth L; ar(1)-source (Gaussian- and gamma innovation with autoregressive constant $\rho = 0.85$), Bitrate R=1 bit/sample.

coefficients ρ in DPCM-systems and tree coding schemes together with the theoretical prediction gains without loss (GPR_{OF}) and with loss (GPR_{MF}) shows Fig. 8. The gains GPR_{OF} and GPR_{MF} are adequate to $-10\log_{10}\gamma_x^2$ and $-10\log_{10}(\gamma_x^2/(1-\epsilon_q^2(1)\rho^2))$, resp. In all cases the loss of prediction gain is negligible with tree depth $L \geq 5$.

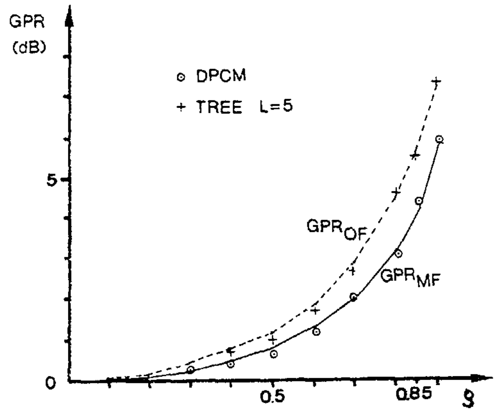

Fig. 8. Comparision of prediction gains in DPCM in tree coding schemes; ar(1)-source (Gaussian innovation) with various autoregressive constants ρ, Bitrate R=1 bit/sample.

Various authors /3,4,17-19/ have found similar improvements in the range of 1-3 dB. In all these cases the improvements can be explained as being due to the compensation of feedback quantization noise and therefore they can be estimated to be upper-bounded by the loss in prediction gain.

It is obvious that the performance of the MSC coders is still restricted by the quantizer performance factor $\epsilon_q^2(R)$. Further improvements can be obtained by applying a stochastic tree- or trellis coding scheme. As an example Fig. 9 compares the signal-to-noise ratio of deterministic- and stochastic trellis coded gamma source (ar(1)-source with $\rho = 0.85$). The results of Fig. 9 are based on a trellis depth L=1024 which implies a coding delay of 1024 samples (intolerable in speech coding). We note, that the

Waveform Coding with Multipath Search Techniques

H. G. Fehn

improvement in SNR over DPCM is significantly higher if the trellis intensity K is large enough. With K=8 an inprovement of nearly 7 dB (10.8 dB instead of 4.2 dB in the DPCM scheme) is obtained. The comparison of waveforms demonstrates the potential of stochastic trellis coding (Fig. 10). The stochastic scheme (c) is able to follow the

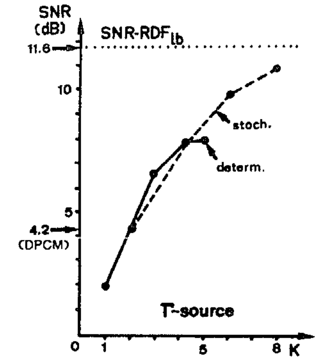

Fig. 9. SNR in trellis coding schemes vs. trellis intensity K; ar(1)-source (gamma innovation with autoregressive constant ϱ =0.85), Bitrate R=1 bit/sample.

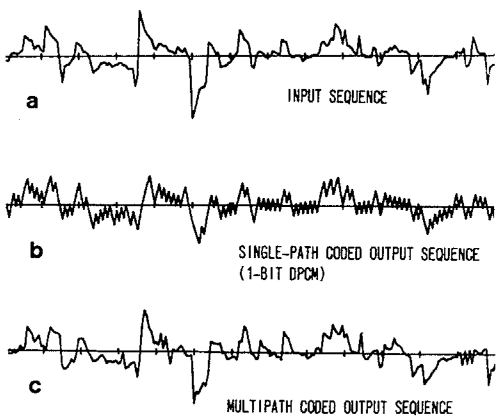

Fig.10. Comparision of waveforms; ar(1)-source (gamma innovation with autoregressive constant ϱ =0.85).

input signal (a), whereas the slope of the output sequence in DPCM scheme is bounded by the quantizer stepsize $q_n \in \{\pm\Delta\}$. To preserve the gain over conventional DPCM the trellis depth L can be reduced by a factor of 1/4 to 1/16, i.e. a delay of 256 and 64 samples, resp./13 /. In this case the random variates populating the nodes of the trellis during initial fanout should be redefined to $E[|x_n|]$ to prevent insufficient decisions during the initial fanout.

Conclusions

In this paper it is shown, that MSC schemes are able to improve the performance of waveform coders up to 7 dB on expense of increasing complexity (in contrast to conventional coding schemes). An exhaustive search algorithm was used to obtain these gains whereas better algorithms such as M-L-algorithm /16/ and Viterbi algorithm /20/ are known. It is also shown, that the performance in deterministic schemes encoding sources with memory are upper bounded by the loss in prediction gain and thus in the range of 1-3 dB (or an equivalent gain if the source is not of first order). Stochastic schemes do not know such restrictions; they are upper-bounded by the information-theoretical limits. At this point it must be mentioned that this bound implies an infinite delay: such MSC schemes are not realizable. However, simple codebooks and tolerable delays in trellis coding up to 128-256 samples allow gains in the midrange of 4 dB in comparision with conventional coding schemes even if real sources (such as speech) /12/ are encoded.

References

1. N. S. Jayant, Ed.,"Waveform Quantization and Coding," New York, IEEE Press, 1976.
2. Y. Linde, A. Buzo, R. M. Gray, "An Algorithm for Vector Quantizer Design," IEEE Trans. Commun., vol. COM-28, pp. 84-95, Jan. 1980.
3. J. Uddenfeldt, L. H. Zetterberg, "Algorithms for Delayed Encoding in Delta Modulation with Speech-Like Signals," IEEE Trans. Commun., vol. COM-24, pp. 652-658, June 1976.
4. A. C. Gorris, "Tree Encoding of Speech Using Variable Length Symbol Release Techniques," Master's Thesis, Dep. Elec. Eng. Texas A&M Univ., College Station, Aug. 1979.
5. T. Berger, "Rate Distortion Theory," Englewood Cliffs, NJ., Prentice Hall, 1971.
6. P. Noll, R. Zelinski, "Bounds on Quantizer Performance in the Low Bit Rate Region," IEEE Trans. Commun., vol. Com-26, pp. 300-304, Feb. 1978.
7. J. D. Markel, A. H. Gray, " Linear Prediction of Speech," New York, Springer Verlag, 1976.
8. R. W. Stroh, R. R. Boorstyn, " Optimum and Adaptive Differential Puls Code Modulation," Polytech. Inst. Brooklyn, Brooklyn, NY. Tech. Rep. NASA Grant NGR 33-006-040, 1970.
9. S. P. LLoyd, "Least Squares Quantization in PCM," Bell Labs., 1957.
10. J. Max, "Quantization for Minimum Distortion," IRE Trans. Inform. Theory, vol. IT-6, pp. 7-12, Mar. 1960.
11. R. E. Blahut, "Computation of Channel Capacity and Rate-Distortion Functions," IEEE Trans. Inform. Theory, vol. IT-18, pp. 460-473, July 1972.
12. H. G. Fehn, P. Noll, "Multipath Search Coding of Stationary Signals with Applications to Speech," IEEE Trans. Commun., vol. COM-30, pp. 687-701, 1982.
13. H. G. Fehn, "Untersuchungen von Mehrwegesuchverfahren zur Codierung von Modell- und Sprachquellen," (in German), Dissertation, Uni. Bremen, 1981.
14. S. G. Wilson, D. W. Lytle, "Trellis Coding of Continuous Amplitude Memoryless Sources," IEEE Trans. Inform. Theory, vol. IT-23, pp. 404-409, Mar. 1977.
15. P. Noll, "Quellencodierung mittels Verfahren der vektoriellen Quantisierung," (in German), Frequenz 36, pp. 187-193, 1982.
16. J. B. Bodie, "Multi-Path Tree-Encoding for Analog Data Sources," Commun. Res. Lab., Fac. Eng., McMasters Univ., Hamilton, Ont., Canada, CRL Int. Rep., Series CRL-20, 1974.
17. Y. Linde, R. M. Gray, " A Fake Process to Data Compression," IEEE Trans. Commun., vol. COM-26, pp. 840 - 846, June 1978.
18. L. H. Zetterberg, J. Uddenfeldt,"Adaptive Delta Modulation with Delayed Decision," IEEE Trans. Commun., vol. COM-22, pp. 1195-1198, Sept. 1974.
19. H. G. Fehn, P. Noll. G. Wessels, "Treecodierung von stationären Quellen und Sprachsignalen," (in German), NTG-Fachberichte, Band 65, (Informations- und System theorie in der digitalen Nachrichtentechnik), pp. 73-80, 1978.
20. C. D. Forney, Jr., "The Viterbi Algorithm," Proc. IEEE vol. 61, pp.268-278, Mar. 1973.