European Conference on
Speech Technology
Edinburgh, Scotland, UK
September 1987

BIT ALLOCATION BASED IN PARAMETRIC RESIDUAL ENVELOPE FOR ADAPTIVE PREDICTIVE CODERS

A. MORENO* and A. MOLINER*

ABSTRACT

This work describes a coder that works at bit rates of 9.6Kbits/s to 32 Kb/s. Basically the system consist on a waveform DPCM coder with an improvement in the quantizer. This new feature consist on quantizing the prediction error taking into account its waveform characteristics. The prediction residual of a voiced signal is characterized by an energy sychronous with pitch. The envelope of this signal holds this information and can be used to quantize properly the prediction error. In this work a parametric version of the residual envelope is used in two ways: Dynamic bit assignment in time domain and adaptive control of the dynamic range of the quantizer.

INTRODUCTION

Adaptive predictive coding is a usefull way to encode speech waveforms. SNR can be increased by using a quantizer based in the waveform of the prediction residual. In this paper, authors present an adaptive predictive coder whose quantifier works w ith an estimation of the residual waveform. This information holds in a parametric version of its envelope and is used to control the range of the quantifier and to assign quantification bits dynamically.

This paper is organized in this manner: In the next section we describe the system used to codify. Next we study the method to obtain the optimal bit asignment according to minimize the mean square error between the input signal and the recovered one. We also discuss the theoretic gain introduced by this system. The number of parameters to transmit and its codification are studied next. Finally, the proposed system is compared with others in terms of segmented SNR and subjective quality tests.

10.21437/ECST.1987-175

SYSTEM DESCRIPTION

The basic configuration of the transmission system is depicted in Fig. 1. The input signal x(n) is applied to an adaptive DPCM /2/,/3/ coder where the input to the quantizer Q is the difference between the input signal non quantized x(n) and a prediction of it $\tilde{x}(n)$.

The quantizer used in Fig. 1 works in an adaptive way. The step size depends on two variables: the variance of the input signal and the number of bits assigned.

For each sample of the input signal to the quantizer the step size is given by the expression

$$\Delta(n) = G(R(n))\sqrt{E(n)} \qquad (1)$$

where R(n) is the number of bits/sample, E(n) is the estimation of the energy of the n^{th} sample and G is the optimal step size for a Gaussian uniform quantizer with unit variance, which is evaluated as a function of the number of bits. To obtain E(n) and R(n) we will use the parametrized residual envelope. In the next section we discuss how to obtain a parametric version of the envelope.

------

* ETSI Telecomunicacion C/ Jorge girona Salgado,s/n 08034 Barcelona Spain
This work is supported by the Generalitat de Catalunya DOG 294/1984

ENVELOPE PARAMETRIZATION

The analytic signal associated with a real signal x(n) is defined as:

$$a_x(n) = x(n)+j\ h_x(n) \qquad (2)$$

Where $h_x(n)$ is the Hilbert Transform of x(n). The envelope of the signal x(n) is the module of its analytic signal. This envelope can be parametrized in the following way. We can compute samples of the analytic signal Fourier Transform $A_x(k)$ $(0<k\leqslant N-1)$ obtained from a windowed signal x(n). $(0<n\leqslant N-1)$ Applying linear prediction over the sequence $A_x(k)$ we estimate the predictor coefficients c(q) (q=1,..,Q) according to minimize the mean square error.

$$E = \sum_{k=0}^{N-1}\left[A_x(k)-\sum_{q=1}^{Q} c(q)\ A_x(k-q)\right]^2 \qquad (3)$$

Solving this problem, as it was previously reported by the authors /1/ by correlation method,a smoothed version $\hat{e}_x(n)$ of the original envelope is obtained

$$\hat{e}_x^{\,2}(n) = k_o / \left|N(1+ \sum_{q=1}^{Q} c(q)\exp(j(2\ /N)nq))\right|^2 \qquad (4)$$

Applying this method over the residual we can see that this parametric envelope gives an estimation of the residual energy.

Fig. 2.a) shows a residual signal, b) its envelope obtained with 9 parameters over 256 samples.

BIT ASSIGNMENT IN TIME DOMAIN

In this section the waveform distortion will be represented as a function of the residual energy. Waveform distortion is defined as the mean square value of the difference between the input and the output of the coder. Let us suppose that there are no transmission errors. Then, as the signal is coded with a system like Fig. 1 the error after decoding the signal is equal to the quantization error.

The variance of the quantization error is proportional to the one of the input signal of the quantizer and inversely to the square of the number of quantization levels.

Distortion can be expressed as:

$$D = \frac{1}{N}\sum_{n=1}^{N} D(n) = \frac{1}{2N}\sum_{n=1}^{N} K2^{-2R(n)}\ \hat{e}_e^{\,2}(n) \qquad (5)$$

Where D(n), R(n) and $\hat{e}_e^{\,2}(n)$ are, respectively, the distorsion, the number of quantization bits and the square of the parametrized envelope of the n^{th} sample, ($\hat{e}_e^{\,2}(n)$ is an estimation of the energy) and K is a term which depends on the probability density function of the residual. For simplicity, we will suppose K equals to a constant.

The average number of bits/sample for the input signal is given by

$$R = \frac{1}{N}\sum_{n=1}^{N} R(n) \qquad (6)$$

We want, for a given rate R, to find the optimal bit assignment for each sample R(n) according to minimize the distortion.

Therefore the problem is to minimize (5) with the restriction (6).The solution is:

$$R(m)_{opt} = R + \frac{1}{2} \log_2 \frac{\hat{e}_e^{\,2}(n)}{\prod_{m=1}^{N} (\hat{e}_e^{\,2}(m))^{1/N}} \qquad (7)$$

This means that $R(n)_{opt}$ is proportional to the log value of the residual envelope. In order to get R(n) an integer value, an iterative algorithm is used. Using the parametric envelope for each sample is calculated the number of bits R(n) and its step size from (1)

THEORETICAL GAIN

For an optimal bit assignment, distortion is obtained by substituting $R(n)_{opt}$ in the expression (5)

$$Dmin = \frac{K}{2}\, 2^{-2R} \prod_{m=1}^{N} (\hat{e}_e^{\,2}(n))^{1/N} \qquad (8)$$

We can see that D_{min} is proportional to the geometric mean of the residual energy estimation.

Theoretical gain SNR is defined as the increase in signal to noise ratio in relation to the Adaptive PCM coder (APCM) which uses feed-forward adaptation on the step size.

Suposing the distortion gain K the same for the residual and the signal, we obtain

$$D_{APCM} = E\; K\; 2^{-2R} \qquad (9)$$

Where E is the energy of the input signal.

Denoting by G the whole SNR gain we have:

$$G = 10 \log_{10} \frac{2E}{\frac{1}{N}\sum_{m=1}^{N} e_e^{\,2}(m)} + 10 \log_{10} \frac{\frac{1}{N}\sum_{m=1}^{N} e_e^{\,2}(m)}{\prod_{m=1}^{N} (\hat{e}_e^{\,2}(m))^{1/N}} \qquad (10)$$

The first part of this equation is the gain apported by the spectral flatness measure /5/ and the second one gives a time flatness measure. Arithmetic mean always is bigger than geometric mean and this second term is positive.

SYSTEM PARAMETERS

TABLE 1. shows the system parameters at bit rates from 9.6 Kb/s to 32 Kb/s. This bit assignement has been chosen based in SNRseg measurements and informal subjetive quality test. Short delay filter LAR parameters are quantized uniformly. Previously mean and variance had been calculated. The mean value of each parameter is substracted at the transmitter and is added in the receiver. After an envelope sensibility study, envelope parcors have been chosen to be quantized in a similar way. Real and imaginary part of each parcor is quantized independly.

RESULTS

At bit rates lower than 24Kb/s this coder was compared with ATC /4/. ATC uses a cosinus transform and bit assignement based in a 12 poles LPC model. Quality is compared in SNRseg meausures and informal subjective test. At 9.6 Kb/s ATC gives 1.8 dB over this coder. At 12 Kb/s the SNRseg is the same with both coders and at 16Kb/s the coder performance is 2 dB better than ATC. At 32 Kb/s this coder was compared with others ADPCM coders whith better results.

REFERENCES

/1/ A.Moreno, M.A. Lagunas.'Envelope/phase representation in signal modeling'.Proc. EUSIPCO-86

/2/ N.S.Jayant.'Digital Coding of Speech Waveforms:PCM,DPCM and DM Quantizers' Proc IEEE vol 62 May(1974)

/3/ B.S. Atal 'Predictive Coding of Speech at Low Bit Rates' IEEE Trans. on Communications Vol com 30,N.4 April(1982).

/4/ J.M. Tribolet, R.E. Crochiere.'Frequency Domain Coding of Speech'IEEE Trans. on ASSP vol ASSP-27, n.5 October(1979).

/5/ N.S.Jayant,P.Noll.'digital Coding of Waveform. Principles and Applications to Speech and Video' Prentice-Hall, Inc. (1984)

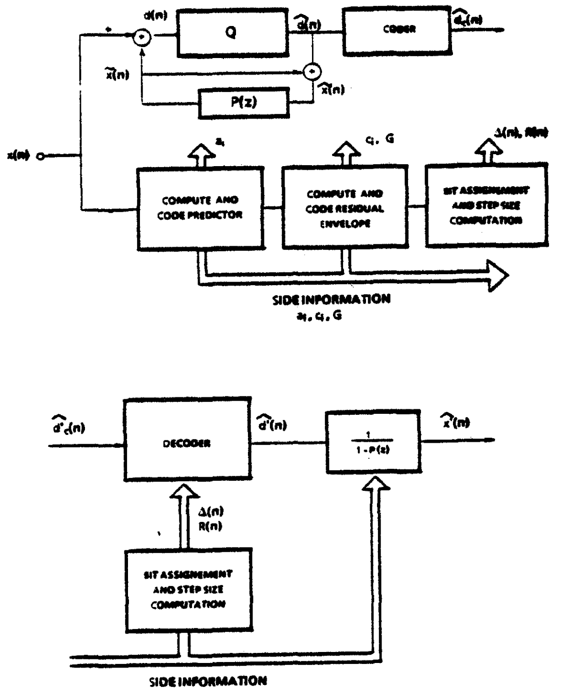

Fig 1. Top Coder.Bottom Decoder

Fig 2. a)Residual.b)Parametric envelope

Total bit rate (Kbits)	9.6	12	16	24	32
Frame length (samples)	256				
Residual data (bit/frame)	232	308	436	672	928
Residual data (bit/sample)	0.9	1.18	1.7	2.62	3.62
Side information (bit/frame)	76			96	
Side information (Kbits/s)	2.5			3	
LPC Parcors (bit/frame)	37			45	
1	5			6	
2	4			6	
3	4			5	
4	3			5	
5-6	3			4	
7,8,9	3			3	
10,11,12	2			2	
Envelope parcors (bit/frame)	34			46	
1	3x2			4x2	
2,3,4	2x2			3x2	
5,6,7,8,9	2x2			2x2	
Envelope Gain	5				

TABLE 1