L. Torres, E. Masgrau, and M.A. Lagunas (eds.)

ROBUST LPC VECTOR QUANTIZATION BASED ON KOHONEN'S DESIGN ALGORITHM

José A. Rodríguez-Fonollosa, Enrique Masgrau, A.Moreno

Departament de Teoria del Senyal i Comunicacions
Universitat Politècnica de Catalunya
Apdo. 30.002, 08080 Barcelona, Spain.

This paper describes a Multistage Vector Quantization scheme in which the codewords are adapted to follow the input statistics. This adaptation is computationally very simple and requires no additional bit transmission. The adaptation algorithm is shown to be closely related with the vector quantizer design techniques known as LBG and Kohonen's. We have studied the application of the developed scheme to quantize the LPC parameters and some results are presented in which the resulting adaptive structure is shown to outperform not only the non-adaptive multistage vector quantizer but also the conventional full-search vector quantizer.

1. INTRODUCTION

Vector quantization (VQ) is a simultaneous quantization of a sequence of samples or vector. This process allows to make effective use of the interrelations among the different vector components and performance arbitrarily close to the ultimate rate-distortion can be achieved by VQ if the vector dimension is high enough [1].

Nevertheless, the exponential growth in complexity forces the use of low-dimensionality VQ in practical systems. In this case some kind of adaptation is necessary to obtain adequate performance, specially when we deal with non-stationary inputs as in speech coding.

Several kinds of adaptive VQ-based coding schemes can be found in the literature. For example, the popular CELP [2] can be seen as an adaptive vector quantizer whose codewords are frame to frame adapted to the input statistics (autocorrelation) by linear filtering.

In the work we describe here we have followed a different approach in which the adaptation algorithm is directly derived from the VQ design techniques. The resulting scheme will be useful when some local stationarities are expected in the input vector statistics. In this communication we explore its application to quantize the LPC parameters, Video coding is another area in which this adaptive quantizer has been shown to be useful [10].

It is well known that, for a given bit rate, a speaker dependent codebook, i.e. designed for the specific speaker whose speech is being coded, would work better than a speaker-independent codebook. Therefore, and adaptive quantizer is also expected to outperform a non-adaptive one when it is used to quantize the LPC parameters. This scheme would also be able to adapt to other variations in the speech spectrum as those due to the acoustic environment of the speaker and to the recording conditions, and the result will be an increase both in performance and in robustness.

The adaptive vector quantization of the LPC parameters have been also studied by other authors. In [3] a system that changed the codewords in time is described. Nevertheless this change creates the necessity of transmitting the new vectors to the receiver with a significant increase in bit-rate. In the scheme that we present the multistage VQ structure [4] is shown to allow the redesign of the quantizer of one stage using the information given by the rest of the stages, therefore, no additional bit transmission is needed.

In [5] vector linear prediction is used to make effective use of the considerable redundancy between different speech frames within one phoneme. Our approach can also be viewed as a vector predictor of reduced complexity. Furthermore the quantizer is expected to follow not only the local stationarities but also the long term (speaker) statistics of the input vectors.

In the following section of the paper we will first present the adaptive multistage vector quantization (AMSVQ) scheme. This scheme was previously reported in [6] where the adaptation algorithm was derived as an LMS type minimization. In this paper, we extend our preliminary report and show the close relation between the proposed algorithm and the VQ design techniques known as LBG [7] and Kohonen's [8].

In the last part of the paper the developed scheme is applied to the quantize the LPC parameters. We will first compare the performance of the adaptive with the non-adaptive multistage structure and the full-search VQ. Then this quantized LPC parameters are used in the CELP and the Multipulse coder and both the SNR results and the subjective quality is discussed.

2. DESCRIPTION

Multistage vector quantization (MSVQ) has always been seen as a suboptimal VQ scheme with reduced complexity and storage. It consists of successively approximating the input vector in several cascaded VQ stages, where the input vector from each stage is the quantization error from the preceding stage. In [4] the MSVQ is applied to the quantization of the LPC parameters, and it is shown that, in the case of Euclidean distance measures such as the log-area ratio, that quantizer is very close to a theoretically predicted asymptotically optimal rate distortion relationship.

In the scheme we present the multistage structure has been used to develop a continuously adaptive VQ. The objective of the adaptive algorithm is to update the last used codevector $\mathbf{c}_i$ in order to minimize the exponentially weighted mse E_i defined as

$$E = \sum_{j=0}^{\infty} \beta^j \, ||\mathbf{x}_i(n-j) - \mathbf{c}_i(n+1)||^2 \qquad (1)$$

where $\mathbf{x}_i(n-j)$ are the input vectors previously quantized by $\mathbf{c}_i$. The minimization of (1) gives

$$\mathbf{c}_i(n+1) = \frac{\sum_{j=0}^{\infty} \beta^j \, \mathbf{x}_i(n-j)}{\sum_{j=0}^{\infty} \beta^j}$$

$$\mathbf{c}_i(n+1) = (1-\beta) \left[\sum_{j=1}^{\infty} \beta^j \, \mathbf{x}_i(n-j) + \mathbf{x}_i(n) \right] \qquad (2)$$

Note that $\mathbf{x}_i(n)$ is the last input vector $\mathbf{x}(n)$ and the infinite sum can be identified as the previous $\mathbf{c}_i$, so we can express the above equation as

$$\mathbf{c}_i(n+1) = \beta \, \mathbf{c}_i(n) + (1-\beta) \, \mathbf{x}(n)$$

$$\mathbf{c}_i(n+1) = \mathbf{c}_i(n) + (1-\beta) \, (\mathbf{x}(n) - \mathbf{c}_i(n)) \qquad (3)$$

or

$$\mathbf{c}_i(n+1) = \mathbf{c}_i(n) + (1-\beta) \, \mathbf{e}(n) \qquad (4)$$

where $\mathbf{e}(n)$ is the quantization error.

This update equation is the well known Kohonen's algorithm, used to design what he calls self-organizing feature maps [8], i. e., vector quantizers. This algorithm, as it appears in (4), is not useful for adapting the quantizer as it is used because the receiver does not known the value of the quantization error. Nevertheless in a MSVQ the error of one stage is quantized by the following stages, therefore, this quantized error is available at the decoder and is used instead of the actual error. The resulting updating equation is then

$$\mathbf{c}_i(n+1) = \mathbf{c}_i(n) + \mu \mathbf{e}_q(n) \qquad (5)$$

where $\mu=(1-\beta)$ is the step size and $\mathbf{e}_q(n)$ is, in general, the sum of the outputs of the following stages. In the most simple case, (2 stages), the VQ equations are

$$\mathbf{c}_i = q_1(\mathbf{x}) \qquad (6)$$

$$\mathbf{e}_q = q_2(\mathbf{x} - \mathbf{c}_i) \qquad (7)$$

$$\mathbf{z} = \mathbf{c}_i + \mathbf{e}_q \qquad (8)$$

where $\mathbf{x}$ is the input vector, $\mathbf{z}$ the quantized vector, $\mathbf{c}_i$ the output of the first quantizer and $\mathbf{e}_q$ the contribution of the second codebook (Fig. 1).

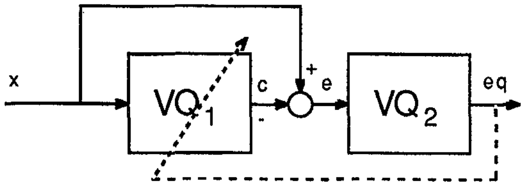

Figure 1. AMSVQ system with 2 stages

Then $\mathbf{e}_q$, that is an estimation of the quantization error of the first codebook, is used to adapt the first quantizer as equation (5) indicates.

As it is shown in [6] this algorithm can also be obtained applying a LMS-type minimization algorithm to the error at the output of the first codebook. In our simulations we only adapted the first stage that was a full-search codebook. Nevertheless, a similar algorithm can be derived for gain-shape vector quantizers and other kind of structures as the tree-searched vector quantizer [9].

Although robust to variations in the signal statistics the above adaptation make the quantizer more sensitive to channel errors that the conventional multistage structure. This problem can be reduced using a leakage factor similar to that used in the

LMS-type algorithms of practical ADPCM schemes. Then (5) is modified to give:

$$\mathbf{c}_i(n+1) = (1-\gamma)\, \mathbf{c}_i(n) + \mu \mathbf{e}_q(n) + \gamma\, \mathbf{c}_{i0}(n) \qquad (10)$$

where γ is a small value that control the memory of the system and $\mathbf{c}_{i0}$ is the initial (design) value of the codevector $\mathbf{c}_i$.

3. CODEBOOK DESIGN

The codebook design approach is similar to that used in conventional multistage VQ [1]. We need a representative speech training set from different speakers, but thanks to the adaptive nature of the AMSVQ system the results are not expected to be very conditioned by the election of the training sequence.

Starting with the first codebook and using the LBG clustering algorithm [7], the codebooks are constructed in succession. The training sequence for the second codebook is obtained as the first codebook quantization error. The problem is that , to obtain this quantization error, we need the output of the second codebook to apply the adaptation algorithm (5) to the first codebook. As $\mathbf{e}_q$ is not available to obtain this training sequence, the actual quantization error $\mathbf{e}$ is used to adapt the first codebook (equation (4)). Further details can be found in [6].

4. RESULTS

The AMSVQ system has been applied to the quantization of LPC parameters with important improvements respect to the conventional multistage structure. Quantization of the LPC coefficients has been extensively studied. Generally, the inverse sines of the reflection coefficients or the log-area ratio values are quantized. We chose the log-area ratio values with the mse distortion for our experiments. Therefore, we will define the LAR-SNR as

$$\text{LAR-SNR} = \frac{\sum_{m=1}^{M} \sum_{i=1}^{N} v_i^2(m)}{\sum_{m=1}^{M} \sum_{i=1}^{N} (v_i(m)-v_{qi}(m))^2}$$

where v_i is the i-th log-area ratio.

Vector quantizers have proven to be very efficient in encoding the predictor parameters. Nevertheless, for high quality coding, 24-30 bits are required and a full search codebook become impractical. Structured codebooks, as the multistage, must be used to reduce the complexity. The developed adaptation algorithm makes the suboptimal multistage structure efficient and robust against different speakers, languages and environments.

To design the quantizers a training speech sequence formed by 80 sentences of 7 female and 7 male speakers was used. And to test the quantizers we chose 24 sentences of a different male speaker (not included in the training sequence). The silent segments (background noise) were not processed.

The speech signal was not preemphasized and the correlation method with a Hamming window of 200 samples (25 ms) was used to obtain ten log area parameters every 160 samples (20 ms).

4.1. High-rate quantization

The first results we present have been obtained using a codebook of 8 codewords (3 bits/frame) in the first stage and 3 codebooks of 256 codewords (8 bits/frame) in the second, third and fourth stage. The number of bits for each frame is thus 27 (3+8+8+8) and at 50 frames/s, the bit rate equals 1350 bps. To obtain a fast adaptation and robustness against channel errors, only the 8 codewords of the first quantizer were adapted.

First of all, we studied the performance of a single codebook of 8 codewords (3 bits) when it was adapted using the actual error $\mathbf{e}$ (forward adaptation). The results showed an increase in signal to noise ratio (LAR-SNR) of more of 2 dB for values of the step size between 0.1 and 1, and local maximum were observed near μ=0.1 and μ=0.6, so these values and μ=0 were used to design three different sets of codebooks.

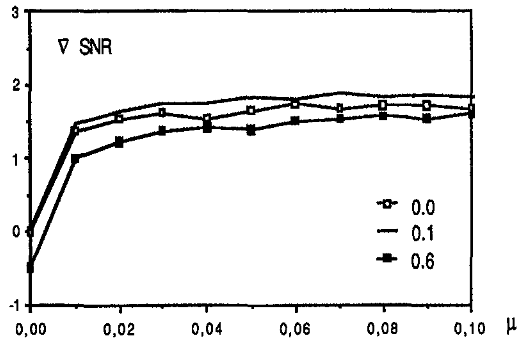

Figure 2. Increase in SNR for the three designed AMSVQ

Then, the designed codebooks have been used to quantize the test sequence with different values of the step size μ. Fig. 2 illustrates the increment in LAR-SNR respect to the 19 dB of the conventional (μ=0), multistage vector quantizer. The best results are now obtained with μ=0.4 and the AMSVQ system designed for μ=0.1. Nevertheless smaller values of the step size as μ=0.1 or μ=0.01 give also a significant increase of the performance and can be a good choice for providing robustness against channel errors.

4.2. Low-rate quantization

The aim of this experiment was to compare the full-search VQ with the AMSVQ. Due to the complexity of the full-search scheme only 10 bits per frame were used. We tried different distributions of bits between the first and the second stage. Fig. 3 compares the LAR-SNR obtained with each scheme and in table I the complexity of the full-search and the best AMSVQ schemes is also shown .

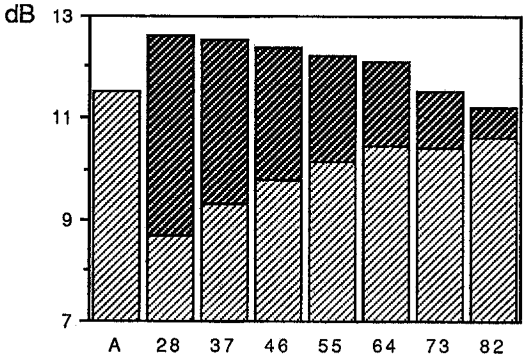

Figure 3. LAR-SNR for some AMSVQ schemes with different bit distributions and the Full-search VQ (A).

Scheme	bits/frame	Comp. Cost	LAR-SNR
VQ	10	1024	11,52
AMSVQ	3+7	136	12,53
AMSVQ	2+8	260	12,62

Table I. Computational cost and LAR-SNR for some 10 bits VQ schemes.

4.3. Speech coding

The developed AMSVQ of the LPC parameters has been integrated in a CELP and a MultiPulse coder. As, it was expected, the results show that an increase in LAR-SNR gives also an increase in the global performance of the coder and both speech quality and SNR is improved (Table II).

bits/frame	CELP	Multipulse
oo	10,57	14,50
AMSVQ-27	10,00	13,30
MSVQ-27	9,65	12,90
AMSVQ-19	9,38	12,80
MSVQ-19	9,15	12,20

Table II. SNR for different coders and some LPC VQ schemes

5. CONCLUSIONS

The application of the developed adaptation algorithm to the multiple stage vector quantizer allows to increase the performance and reduce the complexity of previous VQ-based schemes. This algorithm is very simple and requires no additional bits. The quantizer is continuously redesigned and it is less sensible to the chosen training sequence. Nevertheless, one of the most important results is the increase in robustness across different speakers, languages and environments that the adaptation algorithm provides.

The results show that the AMSVQ increases the performance of conventional multistage quantizers and can be a good choice for coding the LPC coefficients with high quality.

REFERENCES

[1] J. Makhoul, S. Roucos, and H. Gish. "Vector Quantization in Speech Coding". Proc IEEE, vol. 73, November 1985.

[2] M.R. Schroeder, and B.S. Atal. "Code-excited linear prediction (CELP): High-quality Speech at very low bit rates". Proc. ICASSP, Vol. 3, pp 941-944, March 1985.

[3] D.B. Paul. "An 800 bps adaptive vector quantization vocoder using a perceptual distance measure". IEEE Proc. ICASSP, Boston, April 1983.

[4] B.H. Juang, A.H. Gray Jr. "Multiple stage Vector Quantization for Speech Coding". IEEE Proc. ICASSP, pp 597-600, November 1982.

[5] M. Young, G. Davidson, and A. Gersho. "Encoding of LPC Spectral Parameters using Switched-Adaptive Interframe Vector Predict.". IEEE Proc. ICASSP, pp 402-405, April 1988.

[6] J.A. Rodríguez-Fonollosa. "Adaptive Multistage Vector Quantization". Proc. MELECON, pp 225-228, Lisboa, April 1989.

[7] Y. Linde, A. Buzo, and R.M. Gray. "An Algorithm for Vector Quantizer Design". IEEE Trans on ASSP, vol. 34, no. 4, August 1986.

[8] T. Kohonen. *Self-Organization and Associative Memory*. Springer-Verlag, Berlin 1988.

[9] J.A. Rodríguez-Fonollosa. "Cuantificación Vectorial Adaptativa aplicada a la codificación de voz". PhD dissertation, DTSC-UPC, Barcelona, Spain, July 1989.

[10] J.R. Fonollosa, and J.A. Rodríguez-Fonollosa. "Parallel Adaptive Multistage Vector Quantization for digital Video compression". Proc. EUSIPCO'90.