VARIABLE MODEL ORDER LPC QUANTIZATION

Pasi Ojala, Ari Lakaniemi

Nokia Research Center
Speech and Audio Systems Laboratory
Tampere, Finland

ABSTRACT

This paper presents a new method to apply variable bit-rate predictive quantization of the variable model order LPC parameters. In addition, the method is employed to interpolate the parameters within the analysis frame. The LPC model order selection algorithm of this work is based on the characteristics of the input signal and on the performance of the LPC model. Hence, the variable bit-rate LPC quantization is source controlled. The number of quantized parameters needs to be identical in successive frames to be able to apply the predictive quantization and to interpolate parameters inside the frame. Therefore, the order of the LPC model of the previous frame needs to be expanded or reduced to be the same as the current frame LPC model. The advantage of variable model order LPC quantization is the lowered average bit-rate compared to fixed rate while the speech quality remains the same.

1. INTRODUCTION

Speech encoders produce for each analysis frame a set of coding parameters which the decoders use to construct the synthesized speech. Most commonly, codecs have a fixed bit-rate meaning that the number of bits used to represent each speech frame is fixed. Variable bit-rate operation, on the other hand, means that the number of bits may vary from frame to frame. Examples of variable rate codecs can be found e.g. in [4][6].

In the presented encoding algorithm the variability is due to varying the number of parameters as well as the quantization bit-rate of each parameter. This paper introduces a method to select the order of the LPC model so that a smallest possible number of parameters is used for each speech frame while maintaining high quality of reconstructed speech.

Since the characteristics of speech are partly voiced, partly unvoiced and partly noisy, the speech encoder can change the order of the LPC model based on the knowledge about the spectral content of the speech frame and the performance of the linear prediction. Typically, voiced speech with multiple formants in the spectral envelope needs a high model order to reconstruct the spectrum and to maintain good speech quality. On the other hand, speech frames with flat spectrum do not contain any formants, and hence, a lower order model is good enough to maintain the high quality of decoded speech. The benefit of the variable model order LPC is the low bit-rate of the low model order LPC parameter quantization. With a reliable model order selection the quality of decoded speech will remain the same despite of the lowered average bit-rate of the LPC quantization.

Several AR model order selection criteria can be found in literature [2]. Most frequently used linear model order criteria are Akaike Information Criterion (AIC) and Rissanen's Minimum Description Length (MDL). Both criteria are based on the cost function with the variance of prediction error and weighted model order. The cost function finds the minimum model order that satisfies the conditions.

Typically, modern codecs apply predictive line spectrum pair (LSP) quantization to minimize the bit-rate while maintaining high perceptual speech quality [3]. The predicted LSP coefficients are subtracted from the LSP parameters of the current frame resulting in residual which is typically quantized applying vector quantization [5]. Predictive quantization has been proved to be efficient in terms of bit-rate and quality. The variable model order LPC, however, causes difficulties in quantization of the parameters. Since the number of LSP parameters in successive frames can vary, the direct prediction is impossible. In this paper, the predictive quantization of the variable model order LPC parameters is solved by using forced model order expansion or reduction to incorporate the different number of parameters. Hence, variable model order LPC quantization in LSP domain can be applied. The performance of the quantization is measured using the spectral distortion as discussed in [5]. In addition, subjective listening tests are carried out. The performance tests show no significant difference between fixed and variable bit-rate quantization.

This paper first presents the LPC model order criterion used in the work. Then the method to apply predictive LSP quantization and interpolation to variable order LPC model is discussed in Section 2. The performance of variable model order quantizers is evaluated in Section 3 with objective and subjective measures. The conventional fixed rate LPC quantizer is applied as a comparison. Finally, Section 4 summarizes the paper.

2. VARIABLE ORDER LPC

2.1 LPC model order selection

The prediction error of AR model decreases monotonically with increasing model order. Typically, an increase in LPC model order will improve the prediction. However, in the case of speech signal, the prediction error decreases slower, when the spectral envelope of the signal is relatively flat, e.g. when signal contains unvoiced speech or stationary noise. The model order criterion is applied to select a lower order when the reduction in prediction error with increasing order is below a predetermined threshold. The proposed model order criterion determines the reduction in prediction error and compares the difference to a threshold as follows.

$$\hat{\delta}^2{}_k - \hat{\delta}^2{}_{k-1} \le t_k, \quad k \in \{10, 9, \ldots, 1\} \tag{1}$$

where $\hat{\delta}^2{}_k$ is the variance of the prediction error of the k :th order LPC model, and t_k is the applied threshold. The condition in Equation 1 is started with the highest model order. If the condition holds, and the difference in variances is lower than the applied threshold, i.e., the reduction in prediction error with increasing model order is not significant, the determination will be repeated with lower model orders until the lowest order is reached or the condition fails.

In this work three model orders of 10, 8 and 6 are applied. The applied threshold is a compromise between reduced model order and model fit. With a high threshold, the model order decreases too much, and the reduced AR-model does not necessarily explain all the information the signal contains. On the other hand, a too low threshold introduces unnecessary parameters into the model and the prediction is only marginally improved with the cost of extra coefficients and high bit-rate. The threshold is selected based on listening tests. Optimal value minimizes the order rate without introducing distortion to the sound. Figure 1 presents the model order selection for a short speech sample.

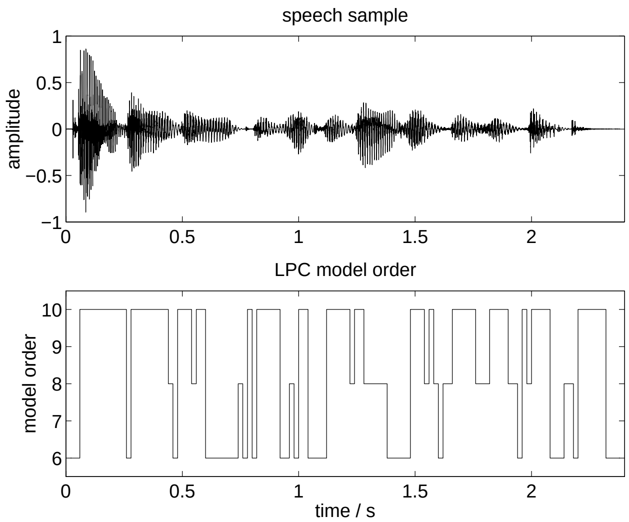

Figure 1. Results of LPC model order selection.

As Figure 1 shows, LPC model order 10 dominates the sample containing male speech. The highest order is selected for clearly voiced periods, whereas the lowest order is selected for silence and unvoiced speech.

2.2 Quantization of LPC parameters

Since the predictive LSP quantization has been proved to be the most efficient method to encode the LPC parameters, it is also applied in the variable model order LPC quantization of this work. When the order of the previous LPC model is different to the current one, the predicted LSP vector has a different length compared to that of the current frame LSP vector. This makes the predictive quantization impossible. In addition, the average values of LSP parameters are different in the case the model order changes. Therefore, the distribution of the prediction residual would be unnecessary large for the quantization purposes, thus lowering the quality.

The solution to the problem is to modify the LPC model of the previous analysis frame. First, the order of the current LPC model is selected in the LPC analysis using the criteria described earlier. The order of the LPC model of the previous frame is then expanded or shortened to be identical to that of the current frame. The modification is done by first converting the parameters into reflection coefficients using the equations described for example in [1]. If the previous model order was lower than the current one, zeros will be added to the previous reflection coefficient vector so that the model orders are same. In the case the previous model order was higher, the last reflection coefficients are removed. After these operations, the modified reflection coefficient vector is transformed back to LPC parameters and further into the LSP domain. Hence, the number of parameters of the previous and current LPC model are now identical, and the conventional predictive LSP quantization can be applied.

Adding zeros to the reflection coefficient vector does not affect the performance of the LPC model. On the other hand, when the reflection coefficient vector is shortened by few coefficients, the LPC spectrum will lose some of its fine structure. Nevertheless, the disadvantage of the lost fine structure is negligible compared to the advantages of predictive vector quantization.

Naturally, the LSP predictor and quantizer need to be separate for each model order. The predictor and the quantization codebooks are selected based on the knowledge about the selected model order as presented in Figure 2.

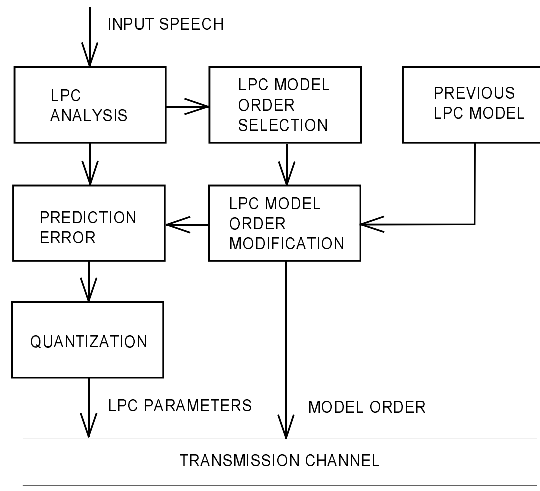

Figure 2. The predictive quantization of the variable model order LPC parameters.

2.3 Interpolation of LPC parameters

The LPC parameters need to be interpolated within the speech frame to assure a smooth transition from frame to frame. Typically, a linear interpolation is applied in each subframe. The parameters in first subframe(s) are interpolated between the current and previous LPC parameters. Now, there exists the same problem as in the case of predictive quantization. When interpolating the parameters, the orders of the successive models need to be the same. Therefore, a similar modification of the LPC model order as described in Section 2.2 needs to be done. The only difference to the previously described method is that the lower order model (current or previous) is always expanded to be identical with the higher model order. The model order expansion is done because it does not result in any loss of information of the spectral fine structure as does the model order shortening. Figure 3 presents the interpolation procedure. The orders of the current and previous LPC model are given to the modification algorithm in which the previous model will be expanded if the order of the current LPC model is higher. Similarly, the algorithm will expand the current LPC model as described earlier, if the order is lower than that of the previous LPC model. Then, the identical order LPC models are interpolated for each subframe.

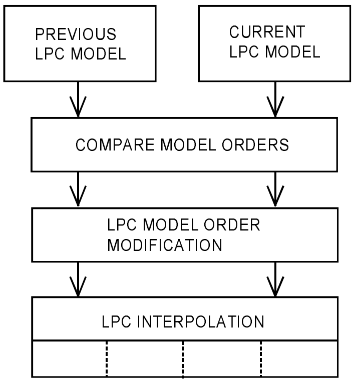

Figure 3. The variable order LPC model interpolation using the model order modification.

2.4 Variable model order LPC decoder

The index of selected LPC model order is transmitted to the decoder among the quantized LPC parameters as a side information. The LPC parameters with variable model order are decoded and further interpolated for successive subframes using the method described above in Sections 2.2 and 2.3 by modifying the LPC model parameters so that the previous frame model order matches to the model order of the current frame.

3. VARIABLE BIT-RATE

3.1 Quantizer training

The variable model order LPC quantizer training material consisted of nine different languages. The total amount of material was about 240,000 frames of length 20 ms resulting in a data base of about one hour and 20 minutes. The speech material was preprocessed by using modified IRS filter.

The training material was first classified using the same model order criterion as that of the variable rate encoder. The classified material was then used for the corresponding model order quantizer training. To compare the quantizer quality, a fixed rate LPC quantizer was trained using the same data base.

3.2 LPC codebook bit allocation

Table 1 presents the bit allocation of the various variable model order codebooks. Two separate bit allocations were designed for the 10th and 6th order models to evaluate the quality degradation with lowering bit-rate. Huffman coding is applied to encode the LPC model order information to minimize the average bit-rate.

Model order	Length of VQ split	Bit allocation	Total bits
10	3 + 3 + 4	7 + 9 + 9 7 + 8 + 9	25 24
8	4 + 4	10 + 10	20
6	3 + 3	8 + 8 8 + 7	16 15

Table 1. The bit allocations of the designed LPC codebooks for the variable model order quantizers.

3.3 Quantization performance

The LPC quantizers were evaluated with objective spectral distortion (SD) measure. The evaluation material was different from the training material consisting of two languages. The total length of the material was about 20,000 frames.

LPC bit allocation	Average SD (dB) total	10	8	6	Outliers (%) 2-4 dB	> 4 dB
25, 20, 16	1.01	1.06	1.03	0.89	1.80	0
25, 20, 15	1.03	1.06	1.04	0.98	2.06	0.01
24, 20, 16	1.04	1.13	1.03	0.88	2.11	0
24, 20, 15	1.06	1.13	1.04	0.98	2.31	0.01
25 (Fixed)	1.05	1.06	1.09	0.99	1.45	0

Table 2. Spectral distortion values for the variable rate and fixed rate LPC model quantization at different bit configurations. Average SD values are presented for the whole material and for each model order.

The spectral distortion values in Table 2 indicate that the performance of the variable rate bit allocation of 25, 20 and 16 bits for the 10th, 8th and 6th order LPC model, respectively, is close to that of the fixed rate LPC quantizer with 25 bits. The average spectral distortion of the fixed rate LPC is also presented for frames corresponding to the model orders of variable model order LPC quantizers. With the given results, both variable and fixed rate quantizers can be considered as transparent. However, the variable rate quantizer operates at significantly lower bit-rate.

The results in Table 2 also imply that changing the bit allocation of the codebook of the 10th or 6th order LPC model does not change the average distortion of the other model order quantizers. The average SD values decrease gracefully with decreasing quantizer bit-rate. However, in transient frames, in which the model order is changing, the average spectral distortion is slightly higher. Table 3 presents the average SD values for transient frames for the variable model order quantization as well as the average SD of the corresponding frames of the fixed rate LPC quantization.

Model order change	Variable rate LPC Average SD (dB)	Fixed rate LPC Average SD (dB)
up	1.13	1.10
down	0.99	1.06

Table 3. Spectral distortion values for the variable model order quantization in transient frames compared with the values of the fixed rate quantization within the corresponding frames.

Results in Table 3 show that the transient in model order and the change of the quantizer does not cause significant difference in the average spectral distortion compared to that of the fixed rate quantization. However, the average SD with increasing model order is higher. Typically, the LPC analysis with increasing model order faces a transient frame containing complex spectrum which is difficult to model using linear prediction.

The distribution of model orders with selected bit configurations for active speech is presented in Table 4. The Huffman coding of the mode information was designed based on this distribution.

Model order	Probability (%)	LPC bits	Mode bits	Average bits
10	41.6	25	1	21.1 + 1.6
8	33.3	20	2	
6	25.1	16	2	
10 (Fixed)	100	25	-	25

Table 4. Bit-rate comparison between variable rate and fixed rate LPC quantizers having the same performance. A variable rate Huffman coding is applied in mode information.

The overall reduction in average bit allocation is 3.9 bits per frame when the presented variable rate LPC quantizer is used. The mode information adds 1.6 bits to the average bits.

3.4 Subjective speech quality

The variable rate LPC quantization with two different bit allocations (Table 2) were also compared with fixed rate using a subjective listening test. The quantizers were embedded into a ACELP type speech codec of 7.35 kbit/s with 25 bit fixed rate LPC quantization. The tested LPC quantizer bit allocations consisting of the conventional fixed rate and two variable rate quantizers are presented in Table 5.

LPC mode	LPC bit allocation	Average LPC bits	Average mode bits	Total bits
Fixed	25	25	-	25
Variable	25, 20, 16	21.1	1.6	22.7
Variable	24, 20, 15	20.4	1.6	22.0

Table 5. Bit allocation for the tested LPC quantizers. The mode information (Huffman coded) is included into the variable rate LPC quantization bit-rate.

The subjective listening tests were performed using a Paired Comparison test. Four female and four male speech samples consisting of sentence pairs were compared. 12 subjects performed the tests. In total, 192 votes were given in both comparisons. Table 6 presents the subjective test results.

LPC bit-rates in comparison	Preference of fixed rate
25 (fixed) vs. 21.1 (variable)	49.5 %
25 (fixed) vs. 20.4 (variable)	51.0 %

Table 6. Paired comparison test results.

The listening test results in Table 6 show that the higher bit-rate variable and fixed rate LPC quantizations do not have any significant difference. The lower bit-rate variable model order quantization performs only slightly worse, as expected according the objective measures in Table 2.

4. SUMMARY

The object of this work was to create a source controlled variable model order LPC quantization and interpolation algorithm for a variable bit-rate speech codec. In this paper a method to search the optimal LPC model order based on the performance of the LPC analysis and the characteristics of the input signal was provided. The algorithms to apply predictive quantization and interpolation of variable model order LPC parameters were presented. When the order of the LPC model changes between successive frames, the LPC model of the previous frame is expanded or reduced so that the model order is the same than that of the current frame. The advantage of the variable model order LPC is the lowered average bit-rate of the parameter quantization. A transparent quantizer was constructed with an average bit-rate of 21.1 bits per frame, while the corresponding fixed rate quantizer with the same quantization performance consumes 25 bits per frame. Evaluation of the variable and fixed rate quantizers with spectral distortion and subjective listening test showed no significant difference in performance.

5. REFERENCES

[1] Kondoz A. M. *Digital Speech (Coding for Low Bit Rate Communication Systems).* John Wiley & Sons, New York. 1994.

[2] Dickie, J. R. and Nandi A. K. *A comparative study of AR order selection methods*. Signal Processing 40. 1994.

[3] Soong F. K. and Juang B. H., *Line Spectrum Pair (LSP) and Speech Data Compression*. Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, Albuquerque, NM, Apr. 1994.

[4] Ojala P. *Toll Quality Variable-Rate Speech Codec*. Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, Munich, Germany, Apr. 21-24 1997.

[5] Paliwal K. K. and Atal B. S., *Efficient Vector Quantization of LPC Parameters at 24 Bits/Frame*. IEEE Transactions on Speech and Audio Processing, Vol. 1, No. 1, Jan. 1993.

[6] Unno T., Barnwell T. and Clements M. *The Multimodal Multipulse Excitation Vocoder*. Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing, Munich, Germany, Apr. 21-24 1997.