Zlatan ESKIC, Zoran PERIC, Jelena NIKOLIC

Faculty of Electronic Engineering, Nis

A method of designing an adaptive uniform quantizer for LPC coefficients quantization

***Abstract**. This paper proposes a method of designing the adaptive uniform quantizer for frame by frame LPC coefficients quantization. The method firstly determines the support region thresholds of two uniform quantizers designated to quantize the minimal and the maximal value of LPC coefficients of each frame. Based on this, the uniform quantizer thresholds estimation for LPC coefficients quantization are provided. The results obtained by testing the proposed method in processing the speech signal from the TIMIT data base are presented and disscused in the paper.*

***Streszczenie.** W artykule zaproponowano metodę zuniformowane go adaptacyjne kwantowania współczynnika LPC (linear prediction coders). Początkowo obliczana jest minimalna i maksymalna wartość LPC dla każdej ramki. Następnie zuniformowany współczynnik jest określany. Zaprezentowano test metody na przykładzie przetwarzania sygnału mowy z bazy TIMIT. (**Metoda projektowania adaptacyjnego kwantyzera LPC**)*

Keywords: Linear predictive coding, LPC coefficients, Adaptive scalar quantization
Słowa kluczowe: LPC, kwantowanie adaptacyjne.

Introduction

Linear prediction coding is one of the widest used speech processing methods. Many of the standardized speech coders at low bit rates are the Linear Prediction Coders (LPC) [1, 2, 3]. LPC coders are determined by the LPC coefficients (LPCs), which are, in general, inappropriate for quantization due to their wide dynamic range and concomitant instability issues in the synthesis filter [4]. On the other hand, it is shown that the quality of a quantized signal is generally influenced by the width of a quantizer's support region and the number of quantization levels [5]. The importance of a suitable support region choice has been pointed out in [6], where the logarithmic decrease of the support region threshold with the number of quantization levels has been ascertained.

The LPC coefficients are usually transformed, before vector or scalar quantization, to more convenient forms (e.g. reflection coefficients and Line Spectrum Pairs) [7, 8, 9, 10]. The fact that with the transformation of the LPC coefficients the computational complexity may be significantly incresead, leave us the space to research a less complex solutions that consider direct quantization of LPC coefficients. Observe that the novel model of quantizer that achieves signal-to-noise ratio values near the optimal one with a little increase in complexity when compared to uniform quantizers has been introduced in [11]. Generally, the optimization of a quantizer design is based on the exact statistical knowledge of the signal to be quantized [9, 5]. However, if the statistics of the signal to be quantized is not completely known, or it varies with time, it is of practical interest to consider application of a kind of adaptation technique. Accordingly, in this paper we propose an adaptation method which is applied to adjust the uniform quantizer support region regarding the dynamic of the LPC coefficients. In other words, by processing the LPC coefficients in frame by frame manner, as in [12], and by applying the novel adaptation method, we actually have a goal to shrink a wide dinamic range of the LPC coefficients, and accordingly the support region of the quantizer, that further may result in the improvement of the signal-to-noise ratio, i.e. of the objective quality of the quantized signal. Particularly, in this paper we propose a simple and an efficient method of determining the quantizer's support region thresholds, which is described in detail in the third section of the paper. The results obtained by testing the proposed method in processing the speech signal from the widely used TIMIT data base are presented in the fourth section of the paper along with the conclusion that sumarises the gain achieved in the paper.

Linear Predicitive Coding

The method of predicting a sample of an input signal based on several previous samples is denoted in literature as Linear Predictive Coding (LPC). Although the possible application fields of this method are vast (e.g. fields of economy [13], geology [14] and medicine [15]), its basic application is in speech processing, speech recognition and speaker recognition [14, 16, 17, 18]. According to the LPC method, the prediction of the current n-th speech sample is carried out by using a linear combination of p previous samples [14, 17, 18]

$$\hat{x}_n = -\sum_{k=1}^{p} a_k x_{n-k} \ , \tag{1}$$

where a_k, $k = 1,...p$, are model parameters or LPC coefficients and p is the number of samples that defines the order of LPC analysis. By approaching the infinite value of p, the prediction of the n-th sample approaches the actual sample value. Nevertheless, p is usually of the order ten to twenty, which implies almost accurate prediction with the limited cost of computation. Accordingly, the prediction error is always introduced. This error, defined by the difference between the estimated sample $\hat{x}_n$ and the original sample x_n,

$$e_n = x_n - \hat{x}_n = x_n + \sum_{k=1}^{p} a_k x_{n-k} = \sum_{i=0}^{p} a_k x_{n-k} \ , \tag{2}$$

is also named the LPC residual. Very low short-term correlations between the samples of residul signal results in approximately flat envelope of its power spectrum. By taking z transformation of (2), it follows that

$$E(z) = A(z) \cdot X(z) \ , \tag{3}$$

where $X(z)$ and $E(z)$ are z transformations of the speech signal and the residual signal, respectively. In the last equation we denoted by $A(z)$ the transfer function of the predictor

$$A(z) = 1 + \sum_{k=1}^{p} a_k z^{-k} \ , \tag{4}$$

that named the predictor filter the whitening filter since it removes the short-term correlation present in the speech signal and, therefore, it flattens the spectrum.

A New Method for LPC Coefficients Quantization

The method we propose in this paper uses a uniform scalar quantization to quantize each LPC coefficient independently. Generally, a uniform scalar quantizer having the support region [a, b] and N quantization levels is defined by the decision thresholds x_i

(5) $$x_i = a + i\Delta,\ i = 0,...,N,$$

and the representation levels y_i

(6) $$y_i = a + \frac{(2i-1)\Delta}{2},\ i = 1,...,N,$$

(7) $$\Delta = \frac{b-a}{N},$$

where Δ is the uniform quantizer step-size [7], [8]. What we actually propose in this paper is the new method for determining the support region tresholds that define the uniform quantizer support region. The quantization is performed in frame-by-frame manner where the LPC coefficients of each frame are firstly determined by applying the Dublin-Levinson-Itakura method [14], afterwhich they are used for the support region adaptation of the uniform quantizer. Namely, the method we propose firstly determines the minimal and the maximal value of LPC coefficients for each frame providing in such a manner the set of minimal and the set of maximal values, denoted as a set of extreme values. Let us denote by $a_{\min}$ and $a_{\max}$ the minimal and the maximal value of the set of minimal LPC coefficient values, and similarly, by $b_{\min}$ and $b_{\max}$ the minimal and the maximal value of the set of maximal LPC coefficient values. Setting these extreme values of both sets as the support region thresholds of the corresponding uniform quantizers, i.e. of the quantizer of minimal LPC coefficient values and the quantizer of maximal LPC coefficient values, one can determine the step-sizes of the appropriate disposable uniform quantizers:

(8) $$\Delta_a = \frac{a_{\max} - a_{\min}}{N_{\text{ex}}},$$

(9) $$\Delta_b = \frac{b_{\max} - b_{\min}}{N_{\text{ex}}}.$$

Accordingly, one can resume that we actually have two uniform scalar quantizers at disposal having the same number of quantization levels N_{ex}, but the different support regions. These two uniform quantizers are designated to quantize the minimal and the maximal value of LPC coefficients of each frame. By the appropriate uniform scalar quantization of the minimal and the maximal value of LPC coefficients of each frame, the thresholds estimation of the quantizer for uniform quantization of the LPC coefficients are provided. Since such determined thresholds may differ for each frame, regarding the fact that the minimum and the maximum coefficient of the frames might be different, in the Fig.1, the uniform quantizer for LPC coefficient quantization is denoted as adaptive uniform quantizer.

In other words, the method we propose consists of the following steps:

Step 1. Determining the minimal and the maximal value of the LPC coefficients for each frame, assuming a prediction order of p.

Step 2. Putting all the minimal and all the maximal values in the appropriate set of minimal values and the set of maximal values, respectively, both called the set of extreme values.

Step 3. Determining the minimal and the maximal value of the LPC coefficients in both sets, i.e., $a_{\min}$ and $a_{\max}$, for the set of minimal values and $b_{\min}$ and $b_{\max}$, for the set of maximal values.

Step 4. Setting $a_{\min}$ and $a_{\max}$, and similarly, $b_{\min}$ and $b_{\max}$, as the quantizer thresholds of the uniform quantizer of minimal values and the uniform quantizer of maximal values, respectively, i.e. of the quantizers of extreme values.

Step 5. Designing the uniform quantizers of extreme values having the same number of the quantization levels N_{ex}, but the different step sizes Δ_a and Δ_b, respectively.

Step 6. Determining the thesholds of the adaptive uniform quantizer by quantizing the minimal and the maximal LPC coefficient of the current frame using uniform quantizers designed in the previous step.

Step 7. Designing the adaptive uniform scalar quantizer having N_p quantization levels, for the current frame LPC coefficient quantization.

Step 8. Uniform quantization of the current frame LPC coefficients using the adaptive uniform quantizer defined in the Step 7.

Step 9. Processing of the next frame by returning to the Step 6.

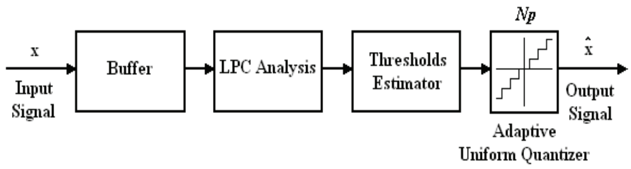

Fig.1. Procedure for LPC coefficients quantization

Numerical Results and Conclusions

For an estimate of the quantizer performance, a signal-to-noise ratio (SNR) is usually used along with the quantization rate, so that the problem of the quantizer design is of the rate - SNR optimization type [14]. In this paper, as in [19], we consider segmental signal-to-noise ratio (SNR_{seg}), i.e. signal-to-noise ratio averaged over the segments or frames:

(10) $$\mathrm{SNR}_{\text{seg}} = \frac{1}{M}\sum_{j=0}^{M-1} 10\log_{10}\left[\sum_{n=m_j-L+1}^{m_j} \frac{x_n^2}{(x_n - \hat{x}_n)^2}\right],$$

where m_0, m_1, ..., $m_{\text{M-1}}$ are the end-times for m frames, each of which is of length L. It is important to mention that problems might occure with SNR_{seg} measure, if frames of silence are included, since large negative SNR values bias the overall measure of SNR_{seg}. Therefore, silent frames are not usually considered in SNR analysis.

The quantization rate (or just rate) is an indicator of the quantizer complexity, which for the proposed quantization scheme is determined as:

(11) $$R = \frac{p \cdot n_p + 2n_{\text{ex}}}{L},$$

where the prediction order is denoted by p, the rate of the quantizer for the uniform quantization of LPC coefficients by n_p, the rate of both, quanitzer of minimal and quantizer of maximal values, by n_{ex}, and finally, the number of samples per frame, or frame lenght, by L.

We have conducted an experiment on the speech signal produced by a male speaker from the widely used TIMIT data base. The original speech signal taken from the TIMIT data base has firstly been sampled by the frequency of 16 kHz and has been represented by 16 bits. Then, the speech signal is down-sampled to the sampling frequency of 8 kHz. Such obtained speech signal we have segmented into frames having lenghts L = 240, i.e. having duration of 30 ms. The total processed number of frames is 400, where the frames of silence has not been considered.

In our analysis we have assumed that the order of linear prediction is $p = 10$. Numerical values of the SNR_{seg} which has been determined for the assumed prediction order and a different quantization levels $N_p = N_{ex}$ (16 to 256), are presented in Table 1 along with the numerical values of the corresponding quantization rate R. Moreover, in the Fig.2, SNR_{seg} dependance on the quantization rate is presented. It is obvious that the increase of rate R cause a slower increase of SNR_{seg}. To compare how well the quantization is performed, one can compare SNR_{seg} values ascertained for the proposed adaptive uniform quantizer with the optimal SNR_{seg} value, which we have calculated for a signal with unquantized LPCs values i.e. for the predictor order of $p = 10$. The optimal SNR_{seg} value we have actually deteremined from the relation between the original and the predicted signal where we have assumed the prediction order of $p = 10$. For the available speech signal we have ascertained the optimal value of SNR_{seg} of 10.8729 dB (see Table 2.). By comparing the SNR_{seg} values, provided in the Table 1, with the optimal value given in the Table 2, it can be noticed that SNR_{seg} of the proposed quantizer apporaches to the optimal SNR_{seg} value with the increase of the quantization levels. It can be easily noticed from Table 1 and Fig.2 that rate - SNR optimization is satisfied for $R = 0.30$ since further increase of rate R results in negligible increase of SNR_{seg}. Corresponding number of rate – SNR optimal quantization levels for this case are $N_p = N_{ex} = 64$ ($n_p = n_{ex} = 6$). This actually means that $2n_{ex} + p\, n_p = 72$ bits are used for quantization of each frame having the lenght $L = 240$.

Table 1. Numerical values of the SNR_{seg} and the quantization rate R of the proposed quantizer for a different quantization levels $N_p = N_{ex}$ and the prediction order of $p = 10$

N_p	SNR_{seg}	R
16	9.0321	0.20
32	9.9973	0.25
64	10.5705	0.30
128	10.7899	0.35
256	10.8426	0.40

Table 2. The optimal SNR_{seg} value for the unquantized LPCs values and the prediction order of $p = 10$

	$p = 10$
Optimal SNR_{seg}	10.8729

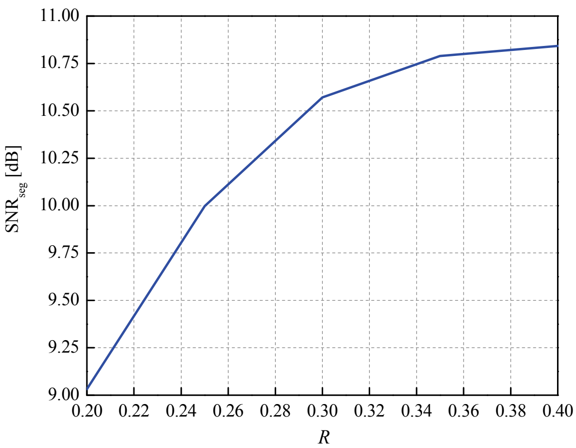

Fig.2. Dependance of SNR_{seg} on a quantization rate R for the prediction order of $p = 10$

Additionaly, one can easily calculate the corresponding bit rate by multiplying the sampling frequency of 8000 Hz with the quantization rate $R = 0.30$ bit/sample which results in 2400 bit/s.

In order to ascertain how the order of predictor affects to the quantization performance, the values of SNR_{seg} that we have determined for a different values of predictor order $p = \{8, 9, 10, 11, 12\}$, which are near the initial assumed value of $p = 10$ for $N_p = N_{ex} = 64$ are presented in Table 3. By comparing these values it can be noticed that an increase of prediction order above value of 10 results in a negligible increase of SNR_{seg}, while increase of R is not negligible. Therefore, in this paper we assume $N_p = N_{ex} = 64$ quanitization levels and the prediction order of $p = 10$. For such assumed solution we have determined both the original and the estimated signal for 40 samples of the 59'th frame (see Fig.3). From the Fig.3 one can notice that the proposed method provides very good aproximation of the original signal.

Table 3. Numerical values of the SNR_{seg} determined for the proposed quantizer having $N_p = N_{ex} = 64$ quantization levels and a different prediction order $p = \{8, 9, 10, 11, 12\}$

	SNR_{seg}
$p = 8$	10.3451
$p = 9$	10.5024
$p = 10$	10.5705
$p = 11$	10.6457
$p = 12$	10.6508

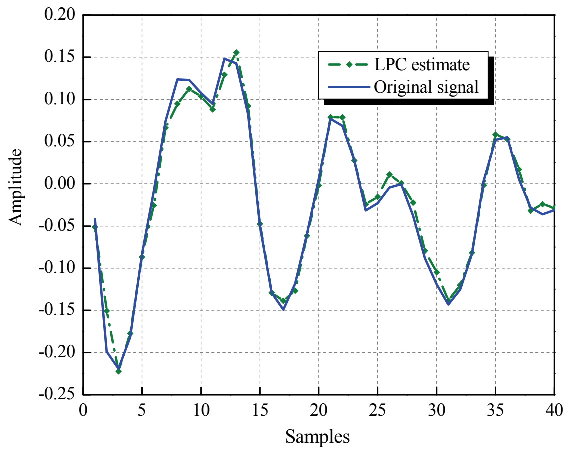

Fig.3. Original signal versus LPC estimate for $N_p = N_{ex} = 64$ and the prediction order of $p = 10$

Based on the obtained results one can conclude that presented adaptive quantizer can be used for direct quantization of LPC coefficients, which has advantage in less processing comlexity than in the cases where transformation of LPC coefficients are used and then the transformed coefficient are quantized. We have shown that in the case of quantizer having $N_p = N_{ex} = 64$ quantization levels (bit rate of 2400 b/s), the proposed adaptive quantizer has the best performance when the predictor is of order $p = 10$. We believe that this simple solution for an adaptive quantizer designing will influence further developement of methods for designing a quantizers with less complexity and possibly better performance.

REFERENCES

[1] Digital Cellular Telecommunications System (Phase 2); *Enhanced Full Rate (EFR) Speech Transcoding, GSM 06.60 Version 4.0.0*, ETSI, (1997).

[2] Recommendation G.729, *Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)*, ITU-T, (1996).

[3] Recommendation GSM 06.10, *GSM Full Rate Speech Transcoding*, ETSI, (1992).
[4] Sugamura N., Itakura F., "Speech Analysis and Synthesis Methods Developed at ECL in NTT – From LPC to LSP", *Speech Communication*, 5 (1986), No. 2
[5] Jayant N. S., Noll P., *Digital Coding of Waveforms, Principles and Applications to Speech and Video*, Prentice Hall, New Jersey, (1984), pp. 115-251.
[6] Na S., "On the Support of Fixed-Rate Minimum Mean-Squared Error Scalar Quantizers for a Laplacian Source", *IEEE Trans. Inf. Theory*, 50 (2004), No. 5, pp. 937-944.
[7] Paliwal K. K., Atal B. S., "Efficient Vector Quantization of LPC Parameters at 24 bits/frame", *IEEE Trans. Acoust., Speech, Signal Process.*, 1 (1993), pp. 3–14
[8] Paliwal K. K., Kleijn W. B., *Quantization of LPC Parameters*, Speech Coding and Synthesis, Elsevier, Amsterdam (1995), pp. 433–466.
[9] Hanzo L., Somerville F. C. A., Woodard J. P., *Voice Compression and Communications: Principles and Applications for Fixed and Wireless Channels*, New York: IEEE Press, (2001).
[10] W. F. LeBlanc, B. Bhattacharya, S. A. Mahmoud, and V. Cuperman, "Efficient Search and Design Procedures for Robust Multi-Stage VQ of LPC Parameters for 4 Kb/s Speech Coding", *IEEE Trans. Speech Audio Process*, 1 (1993), No. 4, pp. 373–385
[11] Peric Zoran, Nikolic Jelena, Eskic Zlatan, Krstic Srbislava, Markovic Nebojsa, "Design of Novel Scalar Quantizer Model for Gaussian Source", *Information Technology and Control*, 37 (2008), No. 4, pp. 321-325
[12] Peric Zoran, Nikolic Jelena, Mosic Aleksandar, Panic Stefan, "A Switched-Adaptive Quantization Technique Using μ-Law Quantizers", *Information Technology and Control*, 39 (2010), No. 4, pp. 317-320
[13] Pindryck S. P. , and D. L. Rubinfeld, *Econometric Models and Econometric Forecasts*, McGraw-Hill, New York, (1998).
[14] Makhoul J., *Linear Prediction: A Tutorial Review*, Proc. IEEE, 63(4): (1975), 561–580
[15] Sharifzadeh H. R., McLoughlin I.V., Ahamdi F., "Voiced Speech from Whispers for Post-Laryngectomised Patients", *IAENG International Journal of Computer Science*, 34 (2009), Issue 4
[16] Lipeika A., "Optimization of Formant Feature Based Speech Recognition", *Informatica*, 21 (2010), No. 3, 361-374
[17] Markel J. D., Gray A. H., *Linear Prediction of Speech*, Springer, Berlin, 2nd edition, (1980).
[18] Hayes M. H., *Statistical Digital Signal Processing and Modeling*, J.Wiley & Sons, Inc., New York, (1996).
[19] Zhijian Yuan, *The Weighted Sum of the Line Spectrum Pair for Noisy Speech, (Master Thesis: Helsinki University of Technology)*, (2003).

Authors*: dipl. Ing. Zlatan Eskic, E-mail: zlataneskic@yahoo.com;, prof. dr Zoran Peric, E-mail: zoran.peric@elfak.ni.ac.rs; MSc Jelena Nikolic, E-mail: jelena.nikolic@elfak.ni.ac.rs; Faculty of Electronic Engineering, University of Nis, Aleksandra Medvedeva 14, 18000 Nis, Serbia.*