Zoran PERIĆ[1], Jelena NIKOLIĆ[1], Jelena LUKIĆ[1], Dragan DENIĆ[1]

Faculty of Electronic Engineering, University of Niš

Two-stage Quantizer with Huffman Coding Based on G.711 Standard

***Abstract**. This paper considers the implementation of Huffman coder for coding of segments of the quantizer based on the G.711 standard. The signal quality obtained with the two-stage quantizer satisfies and overreaches the quality value defined by the G.712 standard, while the implementation of Huffman coder leads to reduction of the average bit rate for 1.25 bit/sample compared to the fixed rate G.711 quantizer. This enables a high quality speech transmission over the channel with less occupancy of its bandwidth.*

***Streszczenie.** W artykule analizowano zastosowanie kodera Huffmana do kodowania kwantyzera bazującego na standardzie G.711. Jakość sygnału otrzymana z dwustopniowego kwantyzera jest satysfakcjonująca przewyższając wymagania zdefiniowane przez standard G.712. (**Dwustopniowy kwantyzer bazujący na kodowaniu Huffmana pracujący w standardzie G.711**).*

Keywords: G.711 standard, lossless Huffman coding, speech signal compression.
Słowa kluczowe: standard G.711, kodowanie Huffmana

Introduction

Lossy speech signal quantizer defined by the G.711 standard provides high quality of the reconstructed signal for the fixed value of bit rate [1]. After G.711 standard was adopted in 1972., a novel G.711.0 standard, which defines lossless compression of speech signal previously processed by the lossy G.711 quantizer, was proposed in 2009. [2]. Lossless signal compression, proposed by the G.711.0 standard, implies the implementation of one of the several different coding techniques where the choice depends on the input signal characteristics [2]. G.711.0 standard, as well as G.711 standard, are widely applied in PSTN networks and in the telephone networks with the packet switched data transmission, i.e. in the VoIP.

In this paper, we propose a two-stage quantizer based on the G.711 standard with the lossless Huffman coding [3, 4, 5]. Specifically, the Huffman coder is utilized in the first quantization stage, and it is applied for coding of segments of the G.711 quantizer, while the coding method in the second quantization stage employs the code words of constant length to code the cells within the segments. Huffman coding, as well as Golomb-Rice coding, employed in [2, 6], belong to a group of VLC (Variable Length Code) entropy techniques for coding with variable length code words [3, 4]. Although the Golomb-Rice coding is simpler than Huffman coding, in this paper Huffman coding is utilized because the number of input symbols, or the number of quantizer segments is low. The procedure of Huffman coding includes determining the optimal length of code words for a given probability of symbols and it is suitable for application for a small number of symbols [3, 4, 5]. The goal of this paper is to show how the application of lossless compression technique, in particular Huffman coding technique, can reduce the average bit rate of the G.711 quantizer, which provides high-quality coded speech signal.

The rest of the paper is organized as follows. After the introductory section, the two-stage quantizer properties are presented and the application of Huffman coder on segments of the G.711 quantizer is explained. Then, the numerical results are provided and discussed in detail. At the end of paper the conclusions are derived to emphasize the importance of the proposed two-stage quantizer with Huffman coding.

Two-stage quantizer and Huffman coder properties

The quantizer based on the G.711 standard is a companding system which compression characteristic is a piecewise linear approximation to the μ-law characteristic ($\mu = 255$) [1]. The support region [$-x_{\max}$, $x_{\max}$] of this symmetric quantizer is divided into $2L = 16$ segments, while each segment is further divided into $m = 16$ cells. The cell or step size widths Δ_i of the G.711 quantizer, defined by the slope of the compression characteristic, in consecutive segments is related by a power of 2, and remain constant within the thresholds of the segment:

$$\Delta_i = 2^i \frac{x_{\max}}{255m}, \quad \frac{\Delta_{i+1}}{\Delta_i} = 2, \quad i = 0,1,2,\ldots,L-1. \tag{1}$$

Parameter $x_{\max}$ from the previous expression represents the support region threshold of the G.711 quantizer.

For a more precise insight into the two-stage quantizer performance, we provide the definition of the following parameters:

$$x_i = \frac{\left(2^i - 1\right)x_{\max}}{255}, \quad i = 0,1,2,\ldots,L, \tag{2}$$

$$x_{ij} = x_i + j\Delta_i, \quad i = 0,1,2,\ldots,L-1, \quad j = 1,2,\ldots,m-1, \tag{3}$$

$$y_{ij} = x_i + \frac{(2j-1)}{2}\Delta_i, \quad i = 0,1,2,\ldots,L-1, \quad j = 1,2,\ldots,m, \tag{4}$$

where x_i are the segments thresholds, x_{ij}, y_{ij} are the cells thresholds and quantizer reproduction levels within the i-th segment, respectively. By assuming that the speech signal, at the quantizer input, can be modeled with the Laplacian probability density function (pdf) [3, 4]:

$$p(x,\sigma) = \frac{1}{\sqrt{2}\sigma} e^{-\frac{|x|\sqrt{2}}{\sigma}}, \tag{5}$$

the following explicit expressions for granular D_g and overload D_{ov} distortion of the G.711 quantizer can be derived [3]:

$$D_g = \frac{1}{12}\sum_{i=0}^{L-1} \Delta_i^2 \left(\exp\left(-\frac{\sqrt{2}x_i}{\sigma}\right) - \exp\left(-\frac{\sqrt{2}x_{i+1}}{\sigma}\right) \right), \tag{6}$$

$$D_{ov} = e^{-\frac{\sqrt{2}x_{\max}}{\sigma}} \left(\left(x_{\max} - y_{L-1,m} + \frac{\sigma}{\sqrt{2}} \right)^2 + \left(\frac{\sigma}{\sqrt{2}} \right)^2 \right). \tag{7}$$

$y_{L-1,m}$ is the reproduction level from the last cell of the last quantizer segment. By summing the granular and overload distortion the total distortion D can be derived and used for signal to quantization noise ratio SQNR determination:

(8) $$\mathrm{SQNR}(\sigma)=10\log\left(\sigma^2/D\right),$$

when the signal variance σ^2 is known. SQNR value is the measure of the reconstructed signal quality.

The G.711 standard defines coding of speech signal samples with code words of constant length equal to 8 bits, so the average bit rate is 8 bit/sample ($R=\log_2(2Lm)$) [3, 4]. In accordance with the standard, these eight bits are used as follows: one bit is used for the sample sign coding, the next three bits are used to define the segment to which the current sample belongs and the last four bits are used to define the cell to which the current sample belongs within the particular segment.

In this paper the novel two-stage coding method is proposed. In the first stage, the Huffman coder [5] for the G.711 quantizer segments coding with code words of variable length is used, while in the second stage, where the cell position within the segment is coded, coding method with code words of constant length is used. In accordance with the VLC Huffman coder properties, the average bit rate required for the coding of segments at the end of the first stage, is less than 4 bit/sample and is equal to:

(9) $$\overline{R}_I(\sigma)=\sum_{i=1}^{2L} l_i P_i(\sigma),$$

where l_i is the length of code word used for coding of the i-th segment and P_i is the i-th segment probability which is, for the assumed pdf function, defined with the thresholds of the i-th segment:

(10) $$P_i(\sigma)=\int_{x_i}^{x_{i+1}} p(x,\sigma)dx,\ i=0,1,2,...,L-1.$$

In the second stage, the uniform cells from the quantizer segments are coded with code words of constant length:

(11) $$R_{II}=\log_2(m),$$

therefore, the second coding stage generates the last four bits of the complete code word from the output of the two-stage quantizer.

Numerical results

This part of the paper is dedicated to the presentation and analysis of numerical results obtained by using the proposed two-stage speech signal processing system. First, one can verify the correctness of Huffman coding method, considered in this paper, by using the code word lengths and segment probabilities given in Table 1. for the proposed two-stage quantizer designed for $x_{\max}$ = 70 and standard deviation σ = 1. For this particular example the obtained average bit rate of Huffman coder is 2.92 bit/sample which is less than 4 bit/sample.

Fig. 1 shows the dependence of SQNR on the speech signal relative input variance for different values of $x_{\max}$. From the Fig. 1, one can notice that the quantizer designed for $x_{\max}$ = 40 achieves the highest average value of $\mathrm{SQNR_{av}}$:

Table 1. Code word lengths l_i and segment probabilities P_i for the two-stage quantizer designed for $x_{\max}$ = 70 and $\sigma=\sigma_{ref}$ =1

(l_1,l_{16})	(l_2,l_{15})	(l_3,l_{14})	(l_4,l_{13})	(l_5,l_{12})	(l_6,l_{11})	(l_7,l_{10})	(l_8,l_9)
(12,12)	(11,10)	(9,8)	(7,6)	(5,4)	(3,3)	(3,2)	(3,3)
P_1=P_{16}	P_2=P_{15}	P_3=P_{14}	P_4=P_{13}	P_5=P_{12}	P_6=P_{11}	P_7=P_{10}	P_8=P_9
0	0	0	0.001	0.032	0.123	0.183	0.161

(12) $$\mathrm{SQNR_{av}}=\frac{1}{k}\sum_{i=1}^{k}\mathrm{SQNR}(\sigma_i),$$

where k represents the number of standard deviation values σ_i included in the calculation. Therefore, the previously mentioned quantizer represents the best solution when the highest average SQNR is required. In addition, one can observe that the signal quality obtained by the two-stage quantizers designed for $x_{\max}$ = 60, $x_{\max}$ = 70 and $x_{\max}$ = 80 meets and exceeds the G.712 standard [7] in the whole range of relative input signal variances.

The difference between average bit rates of the proposed two-stage quantizer and the fixed rate G.711 quantizer, is caused by the application of VLC Huffman coder in the first coding stage. Therefore, the Fig. 2 shows only the dependence of the average bit rate of Huffman coder on the relative input signal variance. From Fig. 2, one can observe that the increase of $x_{\max}$ leads to the decrease of mean value of the average bit rate of Huffman coder $\overline{R}_I$ defined as follows:

(13) $$\overline{R}_I=\frac{1}{k}\sum_{i=1}^{k}\overline{R}_I(\sigma_i),$$

where k represents the number of standard deviation values σ_i included in the calculation. From the Fig. 2 one can conclude that the lowest mean value of the average bit rate of Huffman coder, or the highest compression ratio, achieves the proposed quantizer designed for $x_{\max}$ = 80.

The total average bit rate of the two-stage quantizer is defined with:

(14) $$\overline{R}=\overline{R}_I+R_{II}.$$

Obviously, it is always less than 8 bit/sample. The two-stage quantizer with Huffman coding, designed for $x_{\max}$ = 70, reaches the total average bit rate of 6.75 bit/sample (2.75 bits are used for the sample sign and segment position coding and 4 bits are used for coding of the cell position within the segment). This is 1.25 bit/sample less than the bit rate obtained by coding with constant length code words, i.e. by using only the G.711 quantizer. Observe that a recently proposed lossless coding method has been applied on the nonadaptive and forward adaptive G.711 quantizers [8] and the obtained results have shown that the nonadaptive G.711 quantizer, designed for $x_{\max}$ = 70, provides an average bit rate value of 7.18 bit/sample, which is for 0.43 bit/sample higher than the total average bit rate corresponding to the proposed two-stage quantizer. In addition, the proposed two-stage quantizer provides for about 0.25 bit/sample compression compared to the low complex lossy forward adaptive compandor model that uses approximately 7 bit/sample for the information about the input signal presentation [9].

Most often, the quantizer designing need to be based on simultaneous fulfillment of two opposite requirements, maximal signal quality and minimal value of bit rate [10]. As it can be seen from Figs. 1 and 2, with the increase of $x_{\max}$ over $x_{\max}$=40 (the average SQNR has the highest value, therefore this case is taken for benchmark) signal quality decreases, and the total average bit rate of the two-stage quantizer also decreases. This situation requires definition of a special criterion for selecting the best quantizer solution. The criterion proposed in this paper relies on the G.711 standard requirement that with increase of average bit rate for 1 bit/sample, signal quality increases for approximately 6 dB [3, 4]:

$$\text{(15)} \qquad \mathrm{SQNR}_{\mathrm{av1}} - \mathrm{SQNR}_{\mathrm{av2}} < \left(\overline{R}_{I1} - \overline{R}_{I2}\right) \cdot 6\,\frac{dB}{bit}\,.$$

The left side of the previous inequality represents a signal quality decrease due to $x_{\max}$ increase over the value of $x_{\max} = 40$, while the right side of the inequality is theoretical, i.e. expected quality decrease. Therefore, the parameters indexed with 1 refer to the quantizer designed for $x_{\max} = 40$, while the parameters indexed with 2 refer to quantizers designed for other $x_{\max}$. As long as the decrease in signal quality is less than expected, i.e. value on the left side of

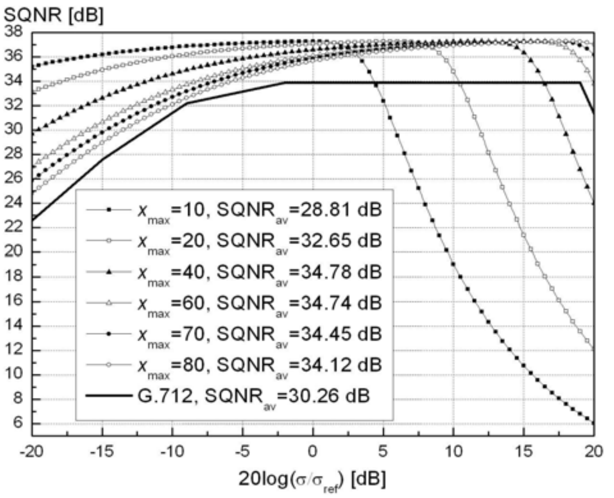

Fig.1. Dependence of SQNR on relative input signal variance for different values of $x_{\max}$ and $\sigma_{ref} = 1$

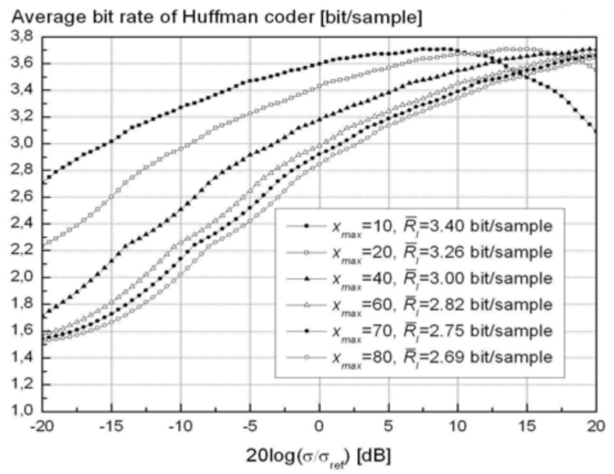

Fig.2. Dependence of the average bit rate of Huffman coder on the relative input signal variance for different values of $x_{\max}$

the inequality is less than the value on the right side, the quantizer indexed with 2 represents a better solution. The upper limit of the $x_{\max}$ value, for which it makes sense to design quantizer, is determined once the sign of inequality is changed. In accordance with the established criterion, the best solution is the quantizer that gives the maximal value of difference between expected and real quality decrease. We have ascertained that the best solution represents the quantizer designed for $x_{\max} = 80$ and the total average bit rate of 6.69 bit/sample.

Conclusion

In this paper we have proposed the two-stage quantizer with lossless Huffman coding employed in the first stage for coding of segments of the G.711 quantizer, while the coding method in the second quantization stage employs the code words of constant length to code the cells within the segments. It has been shown that the proposed quantizer designed for the support region threshold of $x_{\max} = 40$ gives the highest average SQNR. In addition, we have ascertained that with the proposed quantizer, designed for $x_{\max} = 80$, the average bit rate can be reduced by 1.25 bit/sample compared to the widely used fixed rate G.711 quantizer. Therefore, with the proposed two-stage quantizer the transmission of high-quality speech signal over the channel with less occupancy of its bandwidth can be provided.

REFERENCES

[1] Recommendation G.711, *Pulse Code Modulation of Voice Frequencies*, ITU-T, (1972).

[2] Recommendation G.711.0, *Lossless Compression of G.711 Pulse Code Modulation*, ITU-T, (2009).

[3] Jayant N. S., Noll P., *Digital Coding of Waveforms, Principles and Applications to Speech and Video*, Prentice Hall, New Jersey, (1984), pp. 115-251.

[4] Chu W. C., *Speech Coding Algorithms, Foundation and Evolution of Standardized Coders*, John Wiley & Sons, New Jersey, (2003), pp. 143-183.

[5] Salomon D., *A Concise Introduction to Data Compression*, Springer, London, (2008), pp. 61-91.

[6] Perić Z., Dinčić M., Petković M., Design of a Hybrid Quantizer with Variable Length Code, *Fundamenta Informaticae*, 98 (2010), No. 2, 233-256.

[7] Recommendation G.712, *Transmission Performance characteristics of Pulse Code Modulation*, ITU-T, (1992).

[8] Perić Z., Dinčić M., Denić D., Jocić A., Forward Adaptive Logarithmic Quantizer with New Lossless Coding Method for Laplacian source, *Wireless Personal Communications*, 59 (2011), No. 4, 625-641.

[9] Nikolić J., Perić Z., Antić D., Jovanović A., Denić D., Low Complex Forward Adaptive Loss Compression Algorithm and Its Application in Speech Coding, *Journal of Electrical Engineering*, 62 (2011), No. 1, 19-24.

[10] Eskić Z., Perić Z., Nikolić J., Method of Designing Adaptive Uniform Quantizer for LPC Coefficient Quantization, Przeglad Elektrotechniczny, 87 (2011), No. 7, 245-248.

***Authors**: prof. dr Zoran Perić, Serbia, E-mail: zoran.peric@elfak.ni.ac.rs; dr Jelena Nikolić, E-mail: jelena.nikolic@elfak.ni.ac.rs; Jelena Lukić, E-mail: jellaprok@yahoo.com; prof. dr Dragan Denić, E-mail: dragan.denic@elfak.ni.ac.rs, University of Niš, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš.*