FORWARD MEAN-ADAPTIVE QUASILOGARITHMIC QUANTIZER FOR CODING OF CORRELATED SOURCES

UDC (621.391:004):004.415.3

Stefan Tomić, Zoran Perić, Milan Tančić

University of Niš, Faculty of Electronic Engineering, Department of Telecommunications, Niš, Republic of Serbia

Abstract. *This paper presents a forward mean-adaptive quasilogarithmic quantizer for correlated sources. This quantizer exploits the correlation between adjacent input signal samples, while keeping the low computational complexity. Quasilogarithmic quantizer is highly applicable, due to its robustness, providing approximately constant objective output signal quality, when applied to signals with different statistical characteristics. The impact of the application of various compression factors is also analyzed and described. The results obtained indicate that the forward mean-adaptive quasilogarithmic quantizer is suitable for the application in coding of highly correlated sources, as high quality speech and music signals.*

Key words: *adaptive coding, audio compression, correlation, quantization*

1. INTRODUCTION

Constant increase of available data transmission bandwidth in telecommunications increases a demand for high quality digital audio. Digitalization is one of the most important processes in signal processing and it consists of three phases: sampling, quantization and coding. Sampling frequency determines how often we sample the input signal and record it current value. Higher sampling frequency is mostly used when we demand a high quality digital representation of the signal. High quality audio (speech or music) signals are mostly sampled with sampling frequencies equal to 16 KHz or 44 KHz [1]. Higher sampling frequency outputs more samples per second, which are highly correlated in comparison to the signal sampled at smaller sampling frequency.

Quantization is second and very important phase in digitalization process. By applying quantization we determine how many different representation levels digital signal will have. Quantization can be divided based on few criteria, as: (i) memoryless or with

memory, which depends on whether quantizer encoding rules consider previous samples, (ii) uniform or non-uniform, based on the representational levels step size and (iii) scalar or vector, depending upon are samples quantized individually or as blocks [1]. Choice of quantizer which is used in a coding scheme is often based on the statistical characteristics of the signal which will be processed.

This paper presents a forward mean-adaptive (FMA) quasilogarithmic quantizer, designed for the application in quantization of aforementioned highly correlated input audio signals. It was shown in [2] and [3] how the correlation of the input signal samples can be utilized in speech signal coding. In [4] the signal correlation was exploited in ECG signal compression. In this paper, we utilize the correlation which exists between the adjacent audio signal samples for more efficient coding. Quasilogarithmic quantizer is known for its robustness, which makes it applicable to sources with undetermined statistical parameters [5]. FMA quantizer relies on the robustness of the quasilogarithmic quantizer and on the high correlation of the input. Input signal is being processed in frame by frame manner, with the utilization of the mean value of the frame sample amplitudes. The correlation is further increased by subtracting the quantized mean value of the frame from the frame sample amplitudes. This process reduces the amplitude dynamics of the frame samples, which makes the frame more suitable for coding. We showed in [6] that the statistical parameters of the input signal could be successfully utilized for the purpose of reducing the number of bits per sample required for coding, while obtaining high quality output. Unlike the coding scheme presented in [6], which uses information about two statistical parameters of the input signal (mean value and variance), the FMA quasilogarithmic quantizer utilizes only the mean value of the frame, reducing the overall number of computations performed.

The performance of the FMA quasilogarithmic quantizer is evaluated through an objective quality measure called signal to quantization noise ratio (SQNR). Additionally, the results are compared to the quantization schemes with the similar complexity level, such as pulse code modulation (PCM) and the original quasilogarithmic quantizer. The rest of the paper is organized as follows. Section 2 describes the basics of companding quantization, with the focus on quasilogarithmic quantization. In Section 3, the design of the FMA quasilogarithmic quantizer is described in detail. The objective quality assessment is presented in Section 4, while Section 5 is devoted to conclusions and future research directions.

2. QUASILOGARITHMIC QUANTIZATION

Most real signals (speech, audio, image) have non-uniform distribution, with the characteristics that the probability of the signal taking low amplitude values is much greater than the probability of taking larger ones [1]. This implies that it is more suitable to use non-uniform quantizers whose quantization levels are more densely distributed in the area of low input signal amplitudes (on a smaller mutual distance) and widely in the area of greater input signal amplitudes. One way of designing non-uniform quantizers is the application of companding technique, which is mainly used in case when larger number of quantization levels are required [1]. According to companding technique, non-uniform quantizer can be realized as serial connection of compressor, uniform quantizer and expandor.
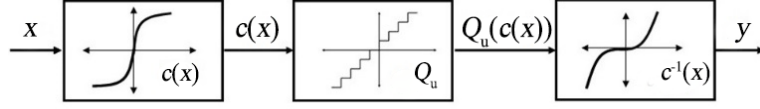
Fig. 1 Block diagram of non-uniform companding technique

Companding technique is performed as follows: input signal (denoted by x) is firstly compressed by compressor function $c(x)$. Afterwards it is quantized with the application of a uniform quantizer (Q_u) and finally expanded by using a non-uniform inverse compressor function $c^{-1}(x)$, which characteristics is inverse to the compressor. The overall structure, which consists of the compressor, uniform quantizer and the expandor is called a compandor. Companding quantizers are divided to optimal and logarithmic, while there are two types of logarithmic quantizers: A law and μ-logarithmic quantizers. Quasilogarithmic quantizer represents a logarithmic quantizer defined with the μ compression law. Quasilogarithmic quantizer performs a signal compression by applying a compressor function defined as in [1], [5].[7]:

$$c_\mu(x) = \frac{x_{max}}{\ln(1+\mu)}\ln\left(1+\mu\frac{|x|}{x_{max}}\right)\mathrm{sgn}(x), \quad |x| \le x_{max},\tag{1}$$

where μ represents the compression factor and x_{max} represents the support limit of the quantizer. Since a coding scheme needs to be theoretically designed and observed, we must introduce a model of an input signal. It has been shown that the Laplacian source can be successfully used to model speech or audio signal [1]. Laplacian probability density function of the original random variable X can be expressed by:

$$p(x) = \frac{1}{\sqrt{2}\sigma}\exp\left(-\frac{|x|\sqrt{2}}{\sigma}\right).\tag{2}$$

By applying quantization, we make an inevitable error in representing the original signal. This error is summed up in a measure named signal distortion, which can be divided to: granular and overload distortion. Granular distortion for companding quantizers can be defined by means of Bennett integral [1], [8] and [9]:

$$D_g(Q) = \frac{x_{max}^2}{3N^2}\int_{-x_{max}}^{x_{max}}\frac{p(x)}{[c'(x)]^2}\,dx.\tag{3}$$

where $c'(x)$ denotes the first derivate of compressor function. By replacing the probability density function $p(x)$ in Eq. 3 and substituting $c(x)$ with $c_\mu(x)$, given by Eq. 1, we obtain an explicit expression for granular distortion of quasilogarithmic quantizer for Laplacian source:

$$D_g(Q) = \frac{\ln^2(1+\mu)\ \sigma^2}{3N^2}\left[\frac{1}{\mu^2}\frac{x_{max}^2}{\sigma^2} + \frac{x_{max}}{\sigma}\frac{\sqrt{2}}{\mu} + 1\right].\tag{4}$$

Overload distortion for the assumed sequence can be calculated as:

$$D_o(Q) = \sigma^2 \exp\left(-\frac{\sqrt{2}x_{max}}{\sigma}\right).$$ (5)

Finally, total distortion can be calculated as a sum of granular and overload distortion: $D(Q) = D_g(Q) + D_o(Q)$ and it is given by:

$$D(Q) = \frac{\ln^2(1+\mu)}{3N^2}\sigma^2\left[\frac{1}{\mu^2}\frac{x_{max}^2}{\sigma^2} + \frac{x_{max}}{\sigma}\frac{\sqrt{2}}{\mu} + 1\right] + \sigma^2 \exp\left(-\frac{\sqrt{2}x_{max}}{\sigma}\right).$$ (6)

By defining distortion we are able to inspect the signal degradation which occurs in the quantization process. Thus defined distortion is used for calculating SQNR, which also depends on the input signal variance (σ^2) and it is given by:

$$\text{SQNR[dB]} = 10\log_{10}\left(\frac{\sigma^2}{D(Q)}\right).$$ (7)

SQNR can be also theoretically obtained, which could be useful when designing a quantizer. This procedure allow us to inspect the performance of the quantizer in wide range of input variances, and also to determine which parameters suit the best to the possible application of the quantizer. The following figure presents the theoretically obtained SQNR for quasilogarithmic quantizer when using 8 bits per sample and various compression factors (20, 50, 100 and 255). One can observe that the highest value of SQNR is obtained when compression factor is equal to 20. The downside of using a small compression factor value is that the highest SQNR is obtained for a narrow range of input signal variance. By using higher compression factor value, we obtain lower maximal SQNR, but the curve is more constant for various input signal variances. This implies that the use of higher compression factor values is a good choice for the coding of high dynamics input signal. The figure shows the importance of choosing appropriate compression factor, for the purpose of obtaining high quality output signal.
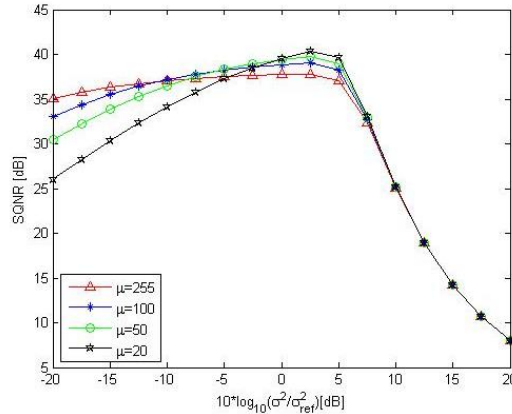

Fig. 2 Theoretical SQNR for quasilogarithmic quantizer when using
8 bits per sample and different compression factors

3. FORWARD MEAN-ADAPTIVE QUASILOGARITHMIC QUANTIZER DESIGN

Logarithmic quantizers with higher compression factor value are often used when it is necessary to provide approximately constant SQNR in wide range of input signal variances. This makes it suitable for the application to non-stationary signals as speech and audio signal. The choice of support limit also highly impacts to the performance of the quantizer [10]. The support limit of the quantizers implemented in the FMA quantization scheme is calculated as [8]:

$$t_{max} = \frac{1}{\sqrt{2}} \ln\left(\frac{3\mu N^2}{\ln^2(\mu+1)}\right),$$

(8)

where N denotes the number of representational levels that quantizer implements, calculated as 2^r, where r represents the bit rate of the quantizer.

The basic goal of the FMA quasilogarithmic quantizer is to achieve compression of a input signal, by using smaller bit rate to represent the side information, while providing the high quality of the quantized signal. Block schemes of encoder and decoder of the proposed quasilogarithmic quantizer are presented in Fig. 3.
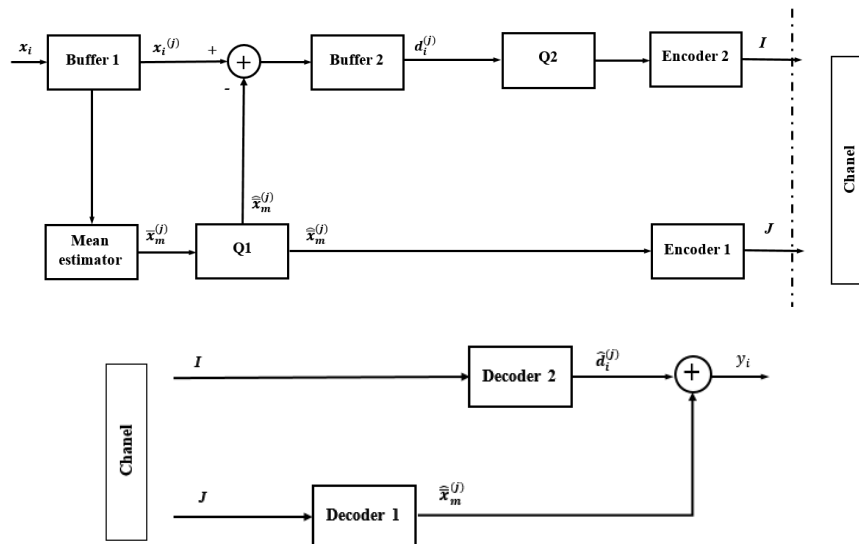

Fig. 3 Block scheme of encoder (above) and decoder (bellow)
of the FMA quasilogarithmic quantizer

The first step in the quantization process is decomposition of the input signal into frames. The input signal $x = x_i(j)$, $i = 1, 2,…,M$, $j = 1,2,…,L$ is divided into L frames of M samples that are brought to the Buffer 1, one frame at a time. Afterwards, the mean value of the input signal samples is estimated for each frame and quantized by the compandor defined with the μ compression law (Encoder 1 and Decoder 1, $\mu=255$). The mean value is quantized with 4 bits per frame, which implies that the quantizer Q1 uses 16 representational

levels (2^4). The quantized mean of the current frame is used for lowering the amplitude dynamics of the frame, for the purpose of better representation. This is performed by subtracting the quantized value of the frames mean from the each sample amplitude in the frame. The obtained difference signal is defined as:

$$d_i^{(j)} = x_i^{(j)} - \hat{\bar{x}}_m^{(j)}, i = 1, 2, ..., M, j = 1, 2, ..., L,$$ (9)

where $\hat{\bar{x}}_m^{(j)}$ represents the quantized value of the mean of the frame, M represents the frame size, while L denotes the total number of frames. This process makes an adaptation of the input signal to the mean value. Thus defined difference signal frame is then fed to the Buffer 2. Buffered difference signal frames are than quantized with the implementation of another μ compression law compandor (Encoder 2 and Decoder 2, μ=255). The encoder of the FMA quantizer sends the indexes (I and J) representing the encoded difference frame samples and the mean value of the original input frame. The quantized output signal samples $y = y_i(j)$, $i = 1, 2,...,M, j = 1,2,...,L$ are reconstructed by adding the decoded represents to the decoded quantized mean value of the current frame:

$$y_i^{(j)} = \hat{d}_i^{(j)} + \hat{\bar{x}}_m^{(j)}, \ i = 1, 2, ..., M, j = 1, 2, ..., L,$$ (10)

As FMA quasilogarithmic quantizer uses side information the total bit rate also depends on the frame size and the bits per sample used for representing the side information. The side information introduces additional bits used for coding in total. The increment in bit rate depends on the frame size. Since we use 4 bits for representing the mean, the total bit rate will be lower if the frame size is greater. This implies that, in most cases, the total bit rate that FMA quantizer implements will not be a round number. The total bit rate of the FMA quasilogarithmic quantizer can be calculated as:

$$R = r_1 + \frac{r_2}{M},$$ (11)

where r_1 denotes the bit rate used for coding the difference signal samples, while r_2 denotes the bit rate used for representing the side information. We also calculated the amplitude dynamics of the input signal, for the purpose of additionally investigating the input signal statistics. Amplitude dynamics was calculated as:

$$Ad = 20 \log_{10} \left(2 \frac{x_{max}^+}{x_{min}^+} \right),$$ (12)

where x_{max}^+ denotes the maximal positive input signal sample, while x_{min}^+ represents the minimal positive input signal sample.

At this point, we are able to access the original and the reconstructed value of the input audio signal, so that we can conduct the objective quality measure and evaluate the performance of the designed quantizer. By applying quantization, we make an inevitable error in representing the original signal. This error is summed up in signal distortion, which is experimentally calculated as:

$$D = \frac{1}{S} \sum_{n=1}^{S} (x_n - y_n)^2,$$ (13)

where $S = M \times L$ denotes the total number of the input signal samples, while x_n and y_n represent the original and quantized input signal samples values, respectively. Thus defined distortion is used in the experiment for calculating SQNR, which also depends on the input signal variance (σ^2) and it is given by:

$$\text{SQNR[dB]} = 10\log_{10}\left(\frac{\sigma^2}{D}\right). \tag{14}$$

3. EXPERIMENTAL RESULTS AND ANALYSIS

This section presents the experimental results collected in the research. The experiment was conducted for the purpose of investigating the objective quality of the designed quantizer output. As mentioned above, the performance of a quantizer is usually expressed through SQNR. This measure quantifies the distortion of the output signal along with the quantization noise and express it in decibels. In the experiment, we applied the FMA quasilogarithmic quantizer to the two types of correlated sources, high quality music and speech audio signals. Both signals are sampled at 44.1 KHz, which indicates high correlation factor. The correlation factor was calculated from:

$$\rho = \frac{\sum_{k=1}^{S} x_k x_{k+1}}{\sum_{k=1}^{S} x_k^2}, \tag{15}$$

where x_k represents the current input signal sample. The relevant statistical characteristics of the test signals are shown in Table 1. One can observe that both signals are characterized by high correlation and amplitude dynamics, while the variance of the signals are low.

Table 1 Statistical characteristics of the input signals

Input signal	Signal variance	Samples number	Correlation factor	Amplitude dynamics [dB]
Classical music	0.0089	832531	0.9938	84.6772
Speech signal	0.0021	2914083	0.9917	86.73

The performance overview of the FMA quasilogarithmic quantizer consists of applying the quantizer to the test signals and collecting the results. Additionally, we applied the original quasilogarithmic quantizer and PCM coding scheme to the same input signals, for the purpose of better understanding the performance of the proposed quantizer. This comparison should confirm the high objective quality of the output signal obtained by applying the proposed quantizer and show the possible gain in SQNR obtained in comparison to the similar quantization techniques.

Table 2 presents the results from applying the FMA quasilogarithmic quantizer to classical music signal, with statistics presented in Table 1. Additionally, the SQNR results for the quasilogarithmic quantizer (SQNR_Q) and for the PCM system (SQNR_{PCM}) are

presented. SQNR for the PCM coding scheme was calculated as described in [6], while SQNR for quasilogarithmic quantizer was calculated by applying Eq. 7. We must underline that the bit rate and frame size are related just to FMA quantizer, as the quasilogarithmic quantizer and PCM do not implement division of the input signal into frames and do not utilize the side information. Compared quantization techniques implement round number of bits per sample, which in this experiment amounts to 8 bits per sample. Frame sizes are varying from 10 to 40 samples, so that the mean value corresponds more accurately to each sample in the frame. High correlation factor implies that the adjacent samples in smaller frame would also have low variance. This implies that process of deduction the quantized mean value of the frame from the frame samples makes the signal more suitable for coding. In this experiment, we implement compression factor of 511, for the purpose of obtaining more robust coding scheme.

Table 2 SQNR for different frame sizes obtained for classical music signal

Bit rate	Frame size	$SQNR_{FMAQ}$	$SQNR_Q$	$SQNR_{PCM}$
8.4	10.0	43.1987	35.6995	34.8305
8.2667	15.0	42.4821	35.6995	34.8305
8.2	20.0	41.8439	35.6995	34.8305
8.16	25.0	41.2283	35.6995	34.8305
8.1333	30.0	40.6835	35.6995	34.8305
8.1143	35.0	40.19	35.6995	34.8305
8.1	40.0	39.7789	35.6995	34.8305

One can notice that the FMA quasilogarithmic quantizer provides gain in SQNR, compared to other considered techniques, for all frame sizes used in the experiment. Since we use different total bit rate for each frame size, this must be considered when the gain is calculated. For example, when the frame size is equal to 20, the total bit rate amounts to 8.2. This means that we add a 0.2 bits per input signal sample, for representing the mean. It is known that for PCM coding scheme SQNR increases about 6 dB when bit rate is increased by 1 bits per sample [1]. This implies that, in this case, SQNR should increase 0.2 times 6, which equals to 1.2 dB. Table 2 shows that, in the case of frame size is equal to 20, the actual gain amounts to about 6.4 dB, when compared to quasilogarithmic quantizer and 7 dB, in comparison to the PCM system. By excluding the 1.2 dB gain, caused by the bit rate increase, the real gain of the FMA quasilogarithmic quantizer still amounts to about 5.2 dB and 5.8 dB, when compared to quasilogarithmic quantizer and PCM coding scheme, respectively. This shows a great improvement in objective quality of the output audio signal, when FMA quantizer is applied.

The results from the application of the FMA quasilogarithmic quantizer to speech signal are presented in Table 3. Additionally, the results are also compared to the original quasilogarithmic quantizer and PCM coding scheme. One can observe that FMA quasilogarithmic quantizer provides gain in SQNR, in comparison to the observed coding schemes. By observing the same frame size of 20 samples, we can see that FMA quasilogarithmic quantizer provides gain in SQNR of about 3.5 dB when compared to the original quasilogarithmic quantizer, while the gain amounts to about 2.6 dB in comparison to the PCM. By deducting the gain which is a product of higher bit rate (8.2

bits per sample) which amounts to 1.2 dB, actual gain is equal to 2.3 dB and 1.4 dB, when compared to quasilogarithmic quantizer and PCM, respectively.

Table 3 SQNR for different frame sizes obtained for male speech signal

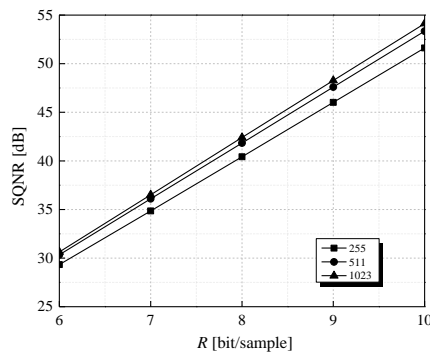| Bit rate	Frame size	$SQNR_{FMAQ}$	$SQNR_Q$	$SQNR_{PCM}$
8.4	10.0	38.9964	34.7252	35.5742
8.2667	15.0	38.5789	34.7252	35.5742
8.2	20.0	38.1938	34.7252	35.5742
8.16	25.0	37.8021	34.7252	35.5742
8.1333	30.0	37.4254	34.7252	35.5742
8.1143	35.0	37.0922	34.7252	35.5742
8.1	40.0	36.7044	34.7252	35.5742

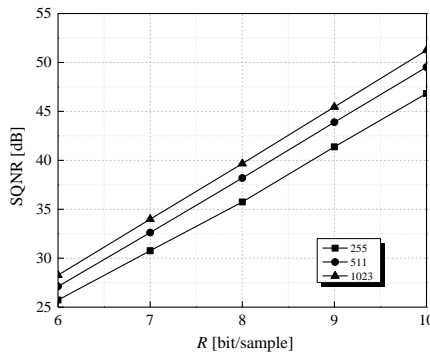**Fig. 4** SQNR for different bit rates and compression factors for classical music signal

Fig. 5 SQNR for different bit rates and compression factors for speech signal

As we showed that the proposed FMA quasilogarithmic quantizer provides gain in SQNR in comparison to quasilogarithmic quantizer and PCM, we shortly present the impact of compression factor choice on the obtained SQNR. For the convenience of presenting the results, we present integer values of the bit rate. The actual bit rate is 0.2 bits per sample larger due to the side information added, for the frame size equal to 20. The previous figures (Fig. 4 and Fig. 5) show SQNR obtained for classical music and speech signal, for the case of using different bit rates and compression factors. One can notice that the higher compression factor results in higher SQNR values. Fig. 4 shows that the smaller value of the compression factor outputs higher theoretical SQNR values. As mentioned, when using a higher compression factor, the SQNR curve is more constant in the wider range of variances. This makes the proposed quantizer suitable for the application to the input signals characterized by high amplitude dynamics. This is confirmed by obtaining higher SQNR values for all bit rates and input signals in the experiment.

6. CONCLUSION

This paper has presented forward mean-adaptive quasilogarithmic quantizer and its possible application in audio signal coding. It was demonstrated how the correlation between adjacent samples in the input signal frames could be utilized for achieving higher signal compression, while preserving high output signal quality. Also, the mean adaptive coding scheme was presented and it has been shown how introducing a low computational complexity adaptation could significantly increase objective quality of the coded output signal. The performance of the proposed FMA quasilogarithmic quantizer was evaluated through SQNR for different bit rates and frame sizes. Additionally, the numerical results have been obtained for the original quasilogarithmic quantizer and for PCM coding scheme, for the purpose of performance comparison. By comparing results, we showed that proposed FMA quasilogarithmic quantizer provides a significant gain in SQNR, while preserving low computational complexity. Moreover, an impact of implementing different compression factors on the objective quality of the output signal has been presented. The successful implementation of the FMA quasilogarithmic quantizer with smaller frame sizes indicate that the proposed quantizer could be utilized in coding schemes which require low delay, as in real time. Finally we can conclude that the presented FMA quasilogarithmic quantizer can be successfully implemented in the coding of high amplitude dynamics audio signals.

REFERENCES

[1] N. S. Jayant, P. Noll, Digital coding of waveforms. New Jersey, Prentice Hall, Chapter 5, pp. 221-251, 1984.

[2] V. Despotović, Z. Perić, L. Velimirović, V. Delić, "DPCM with forward gain-adaptive quantizer and simple switched predictor for high quality speech signals," *Advances in Electrical and Computer Engineering*, vol.10, no.4, pp.95-98, 2010.

[3] V. Despotović, Z. Perić, "ADPCM using a second-order switched predictor and adaptive quantizer," *Advances in Electrical and Computer Engineering*, vol.11, no.3, pp.61-64, 2011. doi:10.4316/AECE.2010.04015

[4] Z. Perić, D. Denić, J. Nikolić, A. Jocić, A. Jovanović, "DPCM quantizer adaptation method for efficient ECG signal compression", *Journal of Communications Technology and Electronics*, vol. 58, no. 12, pp. 1241-1250, 2013.

[5] ITU-T, Recommendation G.711: Pulse Code Modulation (PCM) of Voice Frequencies, International Telecommunication Union, 1972.

[6] Z. Perić, S. Tomić, M. Tančić, "High quality speech signal coding with the application of BTC algorithm", International Conference on Advanced Technologies, Systems and Services in Telecommunications - TELSIKS, Nis, Serbia, October 2015, pp. 23-26

[7] J. Nikolić, Z. Perić, D. Aleksić, D. Antić, "Linearization of optimal compressor function and design of piecewise linear compandor for Gaussian source," *Advances in Electrical and Computer Engineering*, vol.13, no.4, pp.73-78, 2013. doi:10.4316/AECE.2013.04013

[8] J. Nikolić, Z. Perić, D. Aleksić, "Optimization of μ-law companding quantizer for Laplacian source using Muller's method", Przegląd Elektrotechniczny, vol. 89, no. 3A,pp. 206-208, 2013.

[9] S. Na and D. L. Neuhoff, "On the support of MSE-optimal, fixed-rate, scalar quantizers", *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2972–2982, Nov. 2001.

[10] Z. Perić, M. Tančić, A. Jocić , B. Došović, "The Analysis Support Region Influence on Performance of Differential Pulse Code Modulation System with Quasi-logarithmic Quantizer", XIII International Conference SAUM 2016, Nis, Serbia, pp. 316-319.