INFORMATICA, 2010, Vol. 21, No. 3, 375–391

Forward Adaptation of Novel Semilogarithmic Quantizer and Lossless Coder for Speech Signals Compression

Zoran H. PERIC[1], Milan S. SAVIC[1], Milan R. DINCIC[1],
Dragan B. DENIC[1], Momir R. PRASCEVIC[2]
[1]*Faculty of Electronic Engineering, University of Nis*
Aleksandra Medvedeva 14, 18000 Nis, Serbia
[2]*Faculty of Occupational Safety, University of Nis*
Carnojevica 10 A, 18000 Nis, Serbia
e-mail: zoran.peric@elfak.ni.ac.rs

Received: October 2009; accepted: April 2010

Abstract. In this paper new semilogarithmic quantizer for Laplacian distribution is presented. It is simpler than classic A-law semilogarithmic quantizer since it has unit gain around zero. Also, it gives for 2.97 dB higher signal-to-quantization noise-ratio ($SQNR$) for referent variance in relation to A-law, and therefore it is more suitable for adaptation. Forward adaptation of this quantizer is done on frame-by-frame basis. In this way G.712 standard is satisfied with 7 bits/sample, which is not possible with classic A-law. Inside each frame subframes are formed and lossless encoder is applied on subframes. In that way, double adaptation is done: adaptation on variance within frames and adaptation on amplitude within subframes. Joined design of quantizer and lossless encoder is done, which gives better performances. As a result, standard G.712 is satisfied with only 6.43 bits/sample. Experimental results, obtained by applying this model on speech signal, are presented. It is shown that experimental and theoretical results are matched very well (difference is less than 1.5%). Models presented in this paper can be applied for speech signal and any other signal with Laplacian distribution.

Keywords: semilogarithmic quantizer, forward adaptive quantizer, lossless coder.

1. Introduction

Quantizers play an important role in theory and practice of modern day signal processing. Quantizers are mainly design for one variance of input signal and for in advance known distribution. We use Lloyd's–Max's quantizer or optimal companding quantizer (Jayant and Noll, 1984; Peric and Nikolic, 2007; Peric *et al.*, 2009) for constant input variance or very narrow range of variance. Lloyd–Max's quantizer has negligible better performance in relation to optimal companding quantizer (Peric *et al.*, 2009; Sakran *et al.*, 2009). Companding technique is simpler than Loyd–Max's quantization and therefore it has wider application (Peric *et al.*, 2009; Sakran *et al.*, 2009). Logarithmic quantizers are designed for wide range of input signal variance (Jayant and Noll, 1984). Due to its robustness and simplicity, logarithmic quantizers have a wide application (Jayant and Noll, 1984; Aldajani, 2008; Peric *et al.*, 2008; Lyon, 2008; Zavadskas, 2008). European ITU-T G.711 standard (ITU-T, Recommendation G.711, 1972; Sakran *et al.*, 2009) is based upon semilogarithmic compression law. In the early applications, nonadaptive quantizers were dominant, due to simplicity. Today, adaptive quantizers are wide used, since they have higher average $SQNR$, which is almost constant in wide range of variance (Nikolic and Peric, 2008). Forward adaptive quantizers have higher $SQNR$ for about 1 dB in relation to backward adaptive quantizers (Nikolic and Peric, 2008; Chu, 2003; Kondoz, 2004). In many modern applications, combination of quantizer and lossless coder is used. The most often, quantizer and lossless coder are design separately, due to simplicity, but obtained performances are not optimal. Optimal performances can be obtained only with joined design of quantizer and lossless coder, which is done in this paper. There are many coders with higher complexity and higher coding delay. These coders can use linear prediction. For example, linear prediction can be used for speech recognition, as it was described in Bastys *et al.* (2010). But, we do not use linear prediction since our aim is to design simple coder based on logarithmic compression law.

The first aim of this paper is introduction of novel semilogarithmic compression law. As distinct from the classic semilogarithmic A-law, new semilogarithmic law has unit gain in area around zero, and therefore it is simpler for realization. Expressions for distortion for this new law for Laplacian distribution are given in closed form. Optimization of parameters $x_{\min}$ (border between uniform and logarithmic part), and $x_{\max}$ (maximal amplitude of quantizer) is done with a view to maximizing $SQNR$ for referent variance. Also, optimization of N_1 (number of levels in uniform region) for fixed total number of levels N is done. It is shown that for $N = 128$ (i.e., for bit-rate of 7 bits/sample), new semilogarithmic law gives for 2.97 dB higher $SQNR$ for referent variance, in relation to A-law. On the other hand, A-law has constant $SQNR$ in wide range of input variance. Therefore, for nonadaptive quantizers, A-law is better. But, for adaptive quantizers, new semilogarithmic law is much more suitable. To prove that, forward adaptation of new semilogarithmic quantizer is done on frame-by-frame basis. It is shown that standard G.712 can be satisfied with 7 bits/sample. On the other hand, classic A-law for $N = 128$ levels has maximal $SQNR$ of 32.142 dB (recall that G.712 standard requires $SQNR$ of minimum 34 dB), and therefore A-law quantizer cannot satisfy standard G.712 with bit-rate of 7 bits/sample, neither in nonadaptive case nor with adaptation. Beside simplicity, this is the second big advantage of novel semilogarithmic law in relation to A-law.

With a view to further decreasing of bit-rate, model which consists of forward adaptive new semilogarithmic quantizer and lossless encoder is used. Lossless encoder is modification of lossless coder presented in Peric *et al.* (2009). On the output of adaptive quantizer indexes are obtained, whereby each index corresponds to one of $N = 2^r$ representation levels. First $N_1 = 2^{r_1}$ indexes correspond to levels in uniform region. To apply lossless encoder, subframes of M_1 indexes are formed. Lossless encoder works on the following way: (i) if all indexes inside subframe are less or equal to N_1, then all indexes in that subframe are coded with r_1 bits; (ii) if at least one index in subframe is greater than N_1, then all indexes in that subframe are coded with r bits. In that way, double adaptation is done: on variance within frame and on amplitude within subframe.

One of the main contributions of this paper is joined design of forward adaptive quantizer and lossless coder, which gives optimal performances. Optimal values of parameters $x_{\min}$ and $x_{\max}$ are obtained by minimization of average bit-rate R_{av}, with condition that $SQNR$ stays beyond 34 dB. Also, optimizations of parameters N_1 and M_1 are done. It is shown that this model can satisfy G.712 standard with bit-rate of only 6.43 bits/sample, which is significant result. Experiment on speech signal is done, using models presented above. Obtained experimental results for R_{av} and $SQNR$ match with theoretical results very well (the difference is less than 1.5%).

Based on facts presented above, we can say that models presented in this paper can be very successfully applied for compression of speech and other signals with Laplacian distribution.

This paper is organized in the following way. In Section 2, A-law is described, novel semilogarithmic law is presented and their comparison is done. In Section 3, adaptation of new semilogarithmic quantizer is done and after that model with adaptive quantizer and lossless encoder is presented. Experimental results are given in Section 4. Section 5 concludes the paper.

2. Novel Semilogarithmic Companding Law

In this section, classic semilogarithmic A-law will be described first, and after that novel semilogarithmic law will be presented.

2.1. *A-Law Companding*

Assume that an input signal is characterized by continuous random variable X with probability density function (pdf) denoted by $p(x)$. In the rest of the paper we assume that information source is Laplacian source with memoryless property and zero mean value. The pdf of Laplacian source is given by

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

A scalar quantizer with N levels is characterized by the set of real numbers $t_1, t_2, \dots, t_{N-1}$, called *decision thresholds*, satisfying $-\infty = t_0 < t_1 < \cdots < t_{N-1} < t_N = +\infty$ and set of numbers $y_1, \dots, y_N$, called *representation levels*, satisfying $y_j \in \alpha_j = [t_{j-1}, t_j)$ for $j = 1, \dots, N$. Sets $\alpha_1, \alpha_2, \dots, \alpha_N$ form the partition of the set of real numbers R and are called *quantization cells*. The quantizer is defined as many-to-one mapping Q: $R \to \{y_1, y_2, \dots, y_N\}$ defined by $Q(x) = y_j$ where $x \in \alpha_j$. Cells $\alpha_2, \dots, \alpha_{N-1}$ are called *inner* (or *granular*) *cells* while α_1 and α_N are called *outer* (or *overload*) *cells*. In such way, cells $\alpha_2, \dots, \alpha_{N-1}$ form granular while cells α_1 and α_N form an overload region. Quantizers can be uniform or nonuniform. Nonuniform quantizers can be designed using iterative Lloyd–Max algorithm or companding technique. Companding technique consists of the following steps: (i) compress the input signal x by applying the compressor function $c(x)$; (ii) apply the uniform quantizer on the compressed signal; (iii) expand the quantized version of the compressed signal using an inverse compressor function $c^{-1}(x)$. The most often used compressor functions are optimal and logarithmic. Optimal compressor function gives maximal $SQNR$ for referent variance and it is used when input variance is constant or in very narrow range around referent variance. Logarithmic compressor function is robust which means that it gives almost constant $SQNR$ in wide range of variance and it is used when input variance changes with time in wide range. There are two widely used logarithmic functions: A-law and μ-law. In this paper A-law is considered.

Compressor function for A-law is given with:

$$c_1(x) = \begin{cases} \dfrac{Ax}{1+\log A}, & \text{for } |x| \leqslant x_{\min}, \\ \dfrac{x_{\max}(1+\log \frac{A|x|}{x_{\max}})}{1+\log A}\operatorname{sgn}x, & \text{for } x_{\min} < |x| \leqslant x_{\max}. \end{cases} \tag{2}$$

It consists of two parts: linear and logarithmic. $x_{\min}$ is border between those two parts where $x_{\max}$ is maximal amplitude of quantizer. Parameter A denotes ratio $A = x_{\max}/x_{\min}$. A-law is used in many systems, especially in PCM telephone systems in Europe, where value $A = 87.6$ is used. With N_1 and N_2 are denoted numbers of levels in linear and logarithmic part, respectively. N_1 and N_2 can be expressed over total number of levels N as

$$\begin{aligned} &N : 2x_{\max} = N_1 : \frac{2x_{\max}}{1+\log A} \quad \Rightarrow \quad N_1 = \frac{N}{1+\log A}; \\ &N_2 = N - N_1 = \frac{N\log A}{1+\log A}. \end{aligned} \tag{3}$$

During quantization an irreversible error is made, which is expressed by distortion. Distortion is most commonly defined as mean-squared difference between original and quantized signal. Total distortion D_t consists of granular distortion D_g in granular region and overload distortion D_o in overload region, i.e,.

$$D_t = D_g + D_o. \tag{4}$$

Granular distortion D_g consists of two parts: distortion in linear part D_{g1} and distortion in logarithmic part $D_{g2} \cdot D_{g1}$ is calculated as

$$D_{g1} = \frac{\Delta_1^2}{12}P_1, \tag{5}$$

where $\Delta_1 = \frac{2x_{\min}}{N_1}$ and $P_1 = \int_{-x_{\min}}^{x_{\min}} p(x)\,\mathrm{d}x$ is probability that input signal belongs to linear part. For Laplacian distribution $P_1 = 1 - e^{-\frac{\sqrt{2}x_{\min}}{\sigma}}$. D_{g2} can be calculated using Bennett integral as

$$D_{g2} = \frac{\Delta_2^2}{6} \int_{x_{\min}}^{x_{\max}} \frac{p(x)}{[c_1'(x)]^2}\,\mathrm{d}x, \tag{6}$$

where $\Delta_2 = 2(x_{\max} - \frac{x_{\max}}{1+\log A})/N_2$.

Lemma 1. *For semilogarithmic A-law it is valid that Δ_2 is equal to $\Delta = 2x_{\max}/N$.*

Proof.

$$\Delta_2 = \frac{2(x_{\max} - \frac{x_{\max}}{1+\log A})}{N_2} = \frac{2x_{\max}\log A}{(1+\log A)N_2}. \tag{7}$$

Using expression for N_2, given with (3), we obtain

$$\Delta_2 = \frac{2x_{\max}\log A}{(1+\log A)\frac{N\log A}{(1+\log A)}} = \frac{2x_{\max}}{N} = \Delta. \tag{8}$$

This concludes the proof. Using Lemma 1, expression (6) becomes

$$D_{g2} = \frac{\Delta^2}{6} \int_{x_{\min}}^{x_{\max}} \frac{p(x)}{[c_1'(x)]^2}\,\mathrm{d}x. \tag{9}$$

Overload distortion is defined as

$$D_o = 2 \int_{x_{\max}}^{\infty} (x - y_N)^2 p(x)\,\mathrm{d}x. \tag{10}$$

For Laplacian distribution, applying simple mathematic calculation, we can obtain the following expressions:

$$D_{g1} = \frac{x_{\min}^2}{3N_1^2}\big(1 - e^{-\frac{\sqrt{2}x_{\min}}{\sigma}}\big), \tag{11}$$

$$D_{g2} = \frac{(1+\log A)^2}{3N^2}\big[\mathrm{e}^{-\frac{\sqrt{2}}{\sigma}x_{\min}}\big((x_{\min})^2 + \sqrt{2}\sigma x_{\min} + \sigma^2\big) - e^{-\frac{\sqrt{2}}{\sigma}x_{\max}}\big(x_{\max}^2 + \sqrt{2}\sigma x_{\max} + \sigma^2\big)\big], \tag{12}$$

$$D_o = e^{-\frac{\sqrt{2}}{\sigma}x_{\max}}\left(x_{\max}^2 - 2\left(x_{\max} + \frac{\sigma}{\sqrt{2}}\right)\left(y_N - \frac{\sigma}{\sqrt{2}}\right) + y_N^2\right). \tag{13}$$

Signal-to-quantization noise ratio ($SQNR$) is given with:

$$SQNR[\mathrm{dB}] = 10\log_{10}\frac{\sigma^2}{D_t} \tag{14}$$

2.2. *New Semilogarithm Compression Function*

In this section we propose new semilogarithm quantizer which consists of uniform and logarithmic part and which is defined with compression function:

$$c_2(x) = \begin{cases} x, & \text{for } |x| \leqslant x_{\min}, \\ \dfrac{x_{\max}}{B}\left(1+\log\left(\dfrac{B}{x_{\max}}|x|\right)\right)\mathrm{sgn}x, & \text{for } x_{\min} < |x| \leqslant x_{\max}. \end{cases} \tag{15}$$

$x_{\min}$ separates uniform and logarithmic part. One very important fact is that $c_2(x)$ and its first derivation $c_2'(x)$ are continuous in point $x_{\min} \cdot x_{\max}$ is maximal input amplitude of quantizer. With parameter B is denoted ratio $B = x_{\max}/x_{\min}$. $N = N_1 + N_2$ is total number of levels, whereas N_1 and N_2 are numbers of levels in uniform and logarithmic part, respectively. It is valid that

$$\begin{aligned} &N : 2x_{\min}(1+\log B) = N_1 : 2x_{\min} \quad \Rightarrow \quad N_1 = \frac{N}{1+\log B}; \\ &N_2 = N - N_1 = \frac{N\log B}{1+\log B}. \end{aligned} \tag{16}$$

Similarlly to classic A-law, with D_{g1}, D_{g2} and D_o are denoted distortions in uniform, logarithmic and overload region. Similarly to Lemma 1, using expression for N_2 in (16), we can prove that $\Delta_2 = \frac{2x_{\min}(1+\log B)-2x_{\min}}{N_2} = \frac{2x_{\min}\log B}{N_2}$ is equal to $\Delta = \frac{2x_{\min}(1+\log B)}{N}$. Starting from (5), (9) and (10), using, putting $c_2'(x)$ instead $c_1'(x)$ in (9) and approximating y_N with $x_{\max}$, following expressions are obtained:

$$D_{g1} = \frac{x_{\min}^2}{3N_1^2}\left(1-e^{-\frac{\sqrt{2}x_{\min}}{\sigma}}\right), \tag{17}$$

$$\begin{aligned} D_{g2} = \frac{(\log B)^2}{3N_2^2}\Big(&e^{-\frac{\sqrt{2}}{\sigma}x_{\min}}\left(x_{\min}^2+\sqrt{2}\sigma x_{\min}+\sigma^2\right) \\ &- e^{-\frac{\sqrt{2}}{\sigma}x_{\max}}\left(x_{\max}^2+\sqrt{2}\sigma x_{\max}+\sigma^2\right)\Big), \end{aligned} \tag{18}$$

$$D_o = \sigma^2 e^{\frac{-\sqrt{2}}{\sigma}x_{\max}}. \tag{19}$$

Total distortion and $SQNR$ are calculated using expressions (4) and (14).

N and N_1 are chosen to be power of 2, i.e., $N = 2^r$ and $N_1 = 2^{r_1}, r$ and r_1 are integers. In this paper we consider only the case with $N = 128 = 2^7$, because we want that average bit-rate be $R_{\text{av}} \leqslant 7\text{bits/sample}$. In Fig. 1 dependences of $SQNR$ on input variance for different combinations (N_1, N_2) are given. For all combinations, maximization of $SQNR$ is done and optimal values for $x_{\min}$ and $x_{\max}$ are calculated

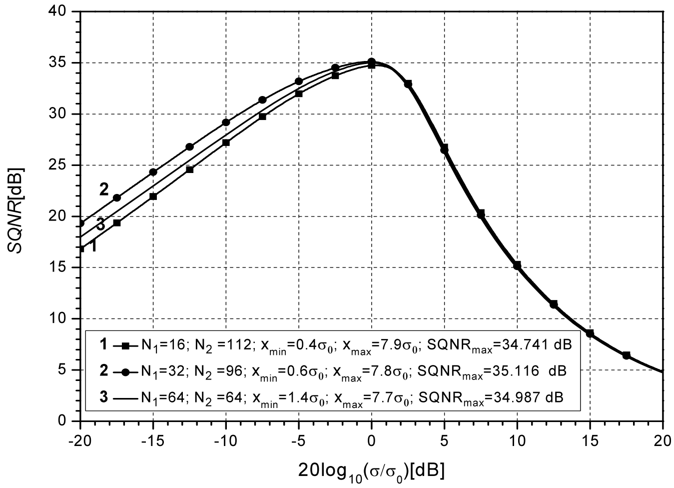

Fig. 1. Comparation of $SQNR$ for novel semilogarithmic law, for different values of $N_1 and N_2$.

such that $SQNR$ has maximal value for referent variance σ_0^2. From this figure we can see that the highest $SQNR$ is achieved for combination $(N_1, N_2) = (32, 96)$. Also, the curve for this combination is widest which is very important for adaptation. Therefore, we can conclude that this combination is optimal. For this combination, optimal values of parameters are $x_{\min} = 0.6\sigma_0$ and $x_{\max} = 7.8\sigma_0$, and maximal $SQNR$ is 35.116 dB.

In the Fig. 2 dependences of $SQNR$ on input variance are given for A-law quantizer with $N = 128$ levels, and for new semilogarithm quantizer given with $c_2(x)$ for combination $(N_1, N_2) = (32, 96)$. We can conclude that:

(i) A-law has constant $SQNR$ for wide range of input variance and it is very suitable for nonadaptive quantizers. But, maximal $SQNR$ for A-law for $N = 128$ levels is 32.142 dB, which means that A-law cannot satisfy G.712 standard (which requires 34 dB) with bit-rate of 7 bits/sample, neither with nor without adaptation.

(ii) New semilogarithmic law has maximal $SQNR$ (for referent variance) 35.116 dB, which is over 2.97 dB higher than maximal $SQNR$ for A-law. Therefore, new semilogarithmic law is very suitable for adaptation. This is its main advantage. Since maximal $SQNR$ is higher than 34 dB, this means that if adaptation is applied on this semilogarithmic quantizer, G.712 standard will be satisfied with 7 bits/sample. We will show this in the next section.

3. Forward Adaptation of Novel Semiogarithmic Quantizer and Lossless Encoder

3.1. *Forward Adaptation*

In this section forward adaptation on novel semilogarithm quantizer, given with $c_2(x)$ is done. Given analysis is similar to analysis in Chu (2003), Kondoz (2004), Hersent *et al.* (2005), Nikolic and Peric (2008).

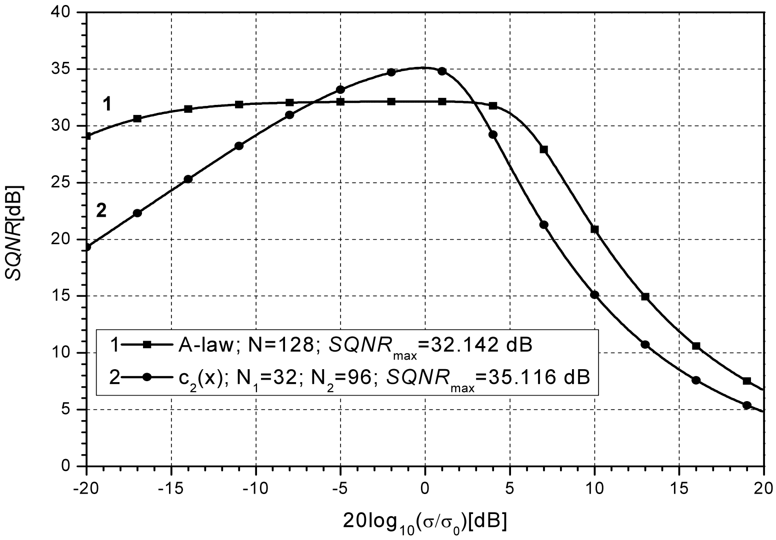

Fig. 2. Comparation of $SQNR$ between A-law and novel semilogarithmic law, in wide range of input variance.

Forward adaptive lossy encoder consists of: buffer with M samples, gain estimator, log-uniform scalar quantizer with K levels for gain quantization, divider and fixed semilogarithm quantizer given with $c_2(x)$, designed for referent variance σ_0^2. M input samples are loaded into input buffer and average variance of these samples, denoted with σ^2, is calculated in 'variance estimator'. σ^2 is quantized in log-uniform scalar quantizer which is designed so that logarithmic variance $(10\log(\sigma^2/\sigma_0^2))$, in range $(-20\,\mathrm{dB}, 20\,\mathrm{dB})$ in relation to referent variance σ_0^2, is divided on K uniform intervals. In logarithmic domain, thresholds are

$$\alpha_i[\mathrm{dB}] = -20 + i\Delta, \quad i = 0, 1, \ldots, K \tag{20}$$

and representation levels are

$$\hat{\alpha}_i[\mathrm{dB}] = -20 + \Big(i - \frac{1}{2}\Big)\Delta, \quad i = 1, \ldots, K, \tag{21}$$

where $\Delta[\mathrm{dB}] = (20 - (-20))/K = 40/K$. In linear domain, thresholds are

$$\sigma_i = 10^{\alpha_i/20}, \quad i = 0, \ldots, K \tag{22}$$

and representation levels are

$$s_i = 10^{\hat{\alpha}_i/20}, \quad i = 1, \ldots, K. \tag{23}$$

So, if $\sigma \in (\sigma_{i-1}, \sigma_i)$ then σ is quantized to s_i. Gain is defined as

$$g_i = \frac{s_i}{\sigma_0}, \quad i = 1, \ldots, K. \tag{24}$$

Therefore, we have K discrete levels of gain. Input samples from buffer are divided with g_i and guided to fixed semilogarithm quantizer with N levels, designed for σ_0^2. If thresholds for fixed quantizer are denoted with $t_j^f,\ j = 0,\dots,N$ and representation levels with $y_j^f,\ j = 1,\dots,N$, then thresholds for adaptive quantizer, for $\sigma \in (\sigma_{i-1}, \sigma_i)$, are $t_j^a = g_i t_j^f,\ j = 0,\dots,N$ and representation levels are $y_j^a = g_i y_j^f,\ j = 1,\dots,N$. If border between uniform and logarithmic part of fixed semilogarithmic quantizer is denoted with $x_{\min}^f$ and maximal amplitude of this fixed quantizer is denoted with $x_{\max}^f$, then border between uniform and logarithmic part, and maximal amplitude of adaptive quantizer, for $\sigma \in (\sigma_{i-1}, \sigma_i)$, are $x_{\min}^a = g_i x_{\min}^f$ and $x_{\max}^a = g_i x_{\max}^f$.

Additional (side) information that determine which gain level (from K levels) is used, should be sent to receiver. Therefore, we need $\log_2 K$ extra bits for every frame of M samples. Hence, bit-rate for forward adaptive quantizer is

$$R = \log_2 N + \frac{\log_2 K}{M}. \tag{25}$$

Dependence of $SQNR$ on input variance is presented in Fig. 6, for $M = 200$, $K = 32$, $N_1 = 32$, $N_2 = 96$ and $x_{\min}^f = 0.6\sigma_0$. For large values of $K (K \geqslant 32)$, $SQNR$ is almost constant and average $SQNR$ is very close to maximal $SQNR$, which is 35.116 dB. This proves that G.712 standard can be satisfied, with bit-rate of 7.025 bits/sample, by adaptation of novel semilogarithmic quantizer. This is not possible with A-law.

3.2. *Lossless Coder*

With a view to further decreasing of bit-rate, in this section we consider model which consists of forward adaptive new semilogarithmic quantizer and lossless encoder. Models of encoder and decoder are presented in Fig. 3. Lossless encoder is very simple, and it is modification of lossless encoder given in Peric *et al.* (2009).

On the output of the forward adaptive encoder, indexes are obtained. There are N different indexes $I = \{1,\dots,N\}$, whereby each index corresponds to one of $N = 2^r$ representation levels of quantizer. First $N_1 = 2^{r_1}$ indexes, $I_1 = \{1,\dots,N_1\}$, correspond

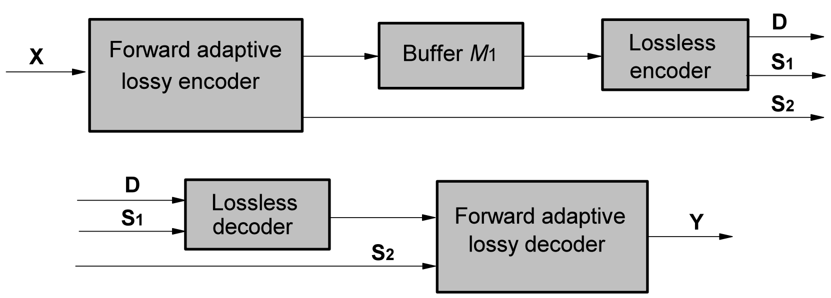

Fig. 3. Models of encoder and decoder (D – bitstream of encoded input samples; S_1 – additional bit in every subframe, for lossless encoder; S_2 – additional information in frame, for forward adaptation).

to representation levels in uniform part. Lossless encoder works in the following way. Firstly, subframes of M_1 indexes are formed. If all indexes in subframe belong to I_1, then each index in that subframe is coded with r_1 bits. If at least one index in subframe is higher than N_1, then all indexes in that subframe are coded with r bits. One additional bit for each subframe should be sent to receiver, to determine whether r_1 or r bits are used for indexes in that subframe. This additional bit increases bit-rate for $1/M_1$ bits/sample. In this model two adaptations are applied: on variance inside frame and on maximal amplitude inside subframe.

Index belongs to I_1 if input sample belongs to uniform part. For input standard deviation $\sigma \in (\sigma_{i-1}, \sigma_i)$, input sample belongs to uniform part if it is smaller than $x^a_{\min} = g_i x^f_{\min}$. Probability that input sample is smaller than $x^a_{\min}$, denoted with p_1, for Laplacian distribution is:

$$p_1(\sigma, s_i) = 1 - e^{\frac{-\sqrt{2}x^a_{\min}(\sigma_i)}{\sigma}} = 1 - e^{-\frac{\sqrt{2}g_i x^f_{\min}}{\sigma}} = 1 - e^{-\frac{\sqrt{2}s_i x^f_{\min}}{\sigma\sigma_0}}; \\ \sigma \in (\sigma_{i-1}, \sigma_i). \tag{26}$$

Probability that M_1 samples belong to uniform part is $p_1^{M_1}$. For $\sigma \in (\sigma_{i-1}, \sigma_i)$, bit-rate $R(\sigma, s_i)$ is

$$R(\sigma) = \big(p_1(\sigma, s_i)\big)^{M_1} r_1 + \big(1 - \big(p_1(\sigma, s_i)\big)^{M_1}\big) r + \frac{1}{M_1} + \frac{\log_2 K}{M}. \tag{27}$$

Average bit-rate in the range (σ_{i-1}, σ_i), denoted with $R^{(i)}_{\text{av}}$, is

$$R^{(i)}_{\text{av}} = \frac{1}{\sigma_i - \sigma_{i-1}} \int_{\sigma_{i-1}}^{\sigma_i} R(\sigma, s_i)\,\mathrm{d}\sigma. \tag{28}$$

Average bit-rate is

$$R_{\text{av}} = \frac{1}{K}\sum_{i=1}^{K} R^{(i)}_{\text{av}}. \tag{29}$$

Integral in (28) cannot be solved in closed form. But, for large K $(K \geqslant 32)$, interval (σ_{i-1}, σ_i) is vary narrow and therefore σ from (σ_{i-1}, σ_i) can be approximated with s_i, i.e., $\sigma \approx s_i$. Using this approximation, expression (26) becomes

$$p_1 \equiv p_1(\sigma, s_i)\big|_{\sigma = s_i} = 1 - e^{-\frac{\sqrt{2}x^f_{\min}}{\sigma_0}}. \tag{30}$$

In that way, $R(\sigma)$ for σ in (σ_{i-1}, σ_i) is approximated with $R(s_i)$, which is constant and therefore integral in (28) can be solved very easily. Since $R^{(i)}_{\text{av}},\ i = 1, \ldots, K$, are equal to each other, there is no need for averaging in (29). Therefore, average bit-rate, R_{av}, is given with

$$R_{\text{av}} = \Bigl(1 - e^{-\frac{\sqrt{2}x^f_{\min}}{\sigma_0}}\Bigr)^{M_1} r_1 + \Bigl(1 - \Bigl(1 - e^{-\frac{\sqrt{2}x^f_{\min}}{\sigma_0}}\Bigr)^{M_1}\Bigr) r + \frac{1}{M_1} + \frac{\log_2 K}{M}. \tag{31}$$

We take that frame for adaptation has length of $M = 200$ samples. Also, we take that $N = 128$ and $K = 32$. Now, we should find optimal values for $x^f_{\min}, x^f_{\max}, N_1$ and M_1. Dependence of average bit-rate R_{av} on M_1 is given in Fig. 4. (This figure is drown for $x^f_{\min} = 0.95\sigma_0$, but dependence is very similar for all other values of interest). We can see that optimal integer value is $M_1 = 2$, and this value will be used in further analysis.

$SQNR$ depends on both $x^f_{\min}$ and $x^f_{\max}$ whereas R_{av} depends only on $x^f_{\min}$. Firstly, we consider one fixed value for $x^f_{\max}$, and for that value we find value for $x^f_{\min}$, denoted with $x^f_{\min}(x^f_{\max})$, in the following way. Since R_{av} is decreasing function of $x^f_{\min}$, we should find maximal value of $x^f_{\min}$ so that $SQNR \geqslant 34$ dB, and this value is $x^f_{\min}(x^f_{\max})$. Now, changing $x^f_{\max}$ in some narrow range, we should find $x^f_{\min}(x^f_{\max})$ for each $x^f_{\max}$ in that range. Optimal $x^f_{\max}$ is one that gives the highest $x^f_{\min}(x^f_{\max})$. This highest value is optimal value of $x^f_{\min}$.

We consider three combinations for (N_1, N_2): $(16, 112), (32, 96)$ and (64, 64). For each of these combinations, parameters $x^f_{\min}$ and $x^f_{\max}$ are obtained in the way described above, and dependences of average bit-rate R_{av} on $SQNR$ are given in Fig. 5. We can see that for required $SQNR$ of 34 dB, combination $(N_1, N_2) = (32, 96)$ gives the smallest $R_{\text{av}} = 6.432$ bits/sample. Therefore, this is optimal combination. For this combination, optimal values of parameters are: $x^f_{\min} = 0.95\sigma_0$ and $x^f_{\max} = 8\sigma_0$.

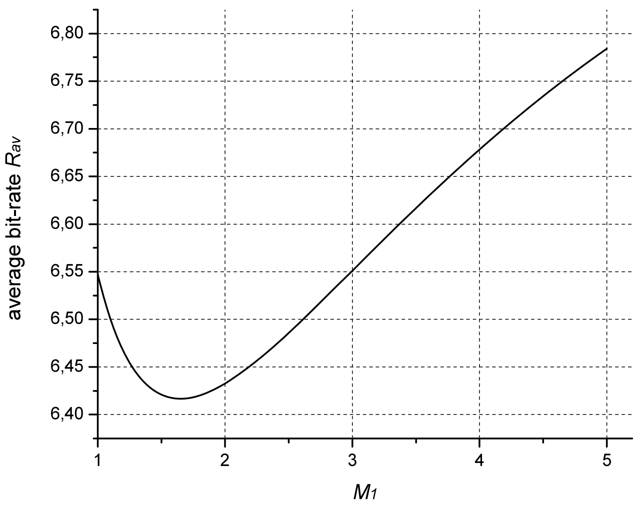

Fig. 4. Dependence of average bit-rate R_{av} on subframe length M_1.

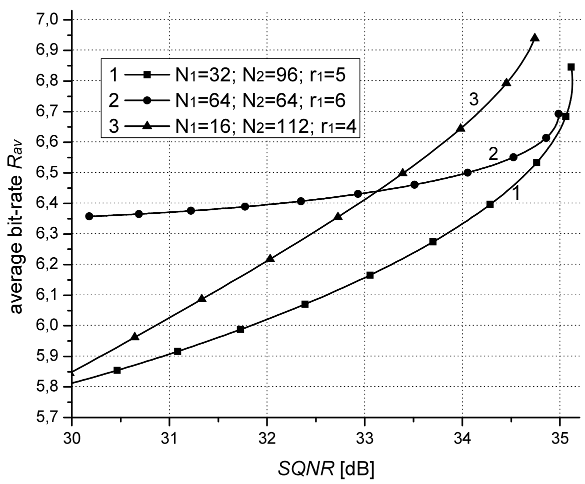

Fig. 5. Dependence of average bit-rate R_{av} on $SQNR$ for different values of N_1 and N_2.

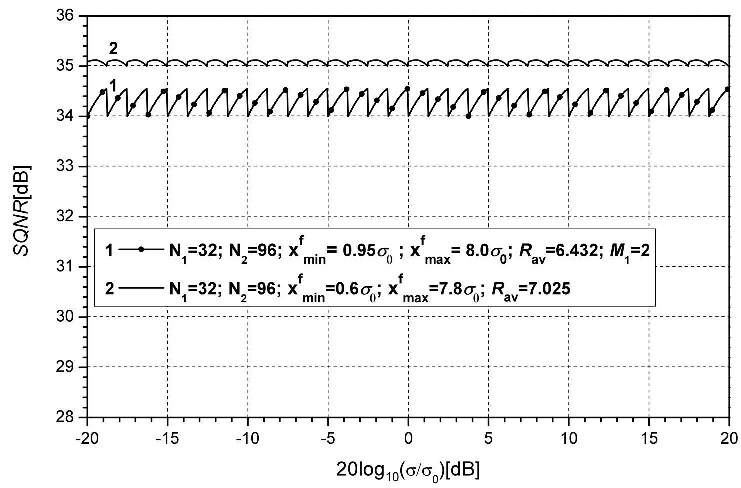

Fig. 6. Comparation of $SQNR$ for models described in Sections 3.1. and 3.2., for $K = 32$.

As it can be seen from the above procedure, R_{av} (this is parameter of lossless encoder) has influence on $x^f_{\min}$ (this is parameter of quantizer). The fact that parameters of lossless encoder have influence on parameters of quantizer, and reverse, represents joined design of quantizer and lossless encoder. Joined design gives optimal values of parameters, as opposed to separate design of quantizer and lossless encoder. Therefore, joined design gives better performances in relation to separate design. Using this joined design, standard G.712 is satisfied with average bit-rate R_{av} of only 6.432 bits/sample.

In the Fig. 6, $SQNR$ in dependence on input variance for $K = 32$ is given, for two cases. One case is forward adaptive quantizer (described in Section 3.1) and the second

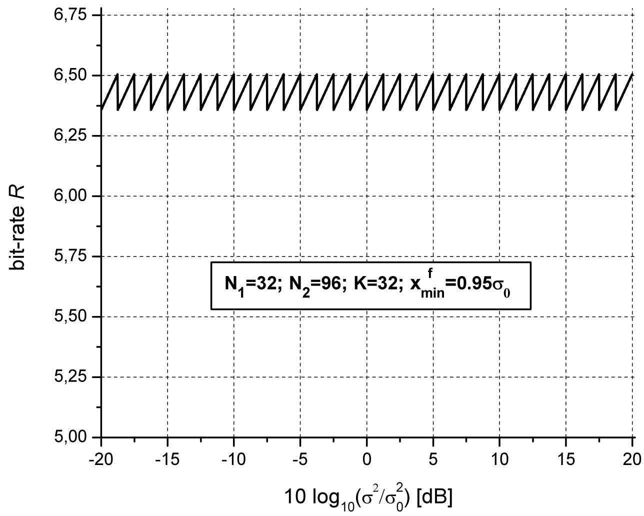

Fig. 7. Bit-rate R in dependence on input variance.

case is model with forward adaptive quantizer and lossless encoder (described in this section). For the second case joined design is done. Optimal values of parameters for both cases are given in the figure.

Dependence of bit-rate R on input variance (given with (27)) is presented in Fig. 7, for $K = 32$. We can see that inside one interval $\sigma \in (\sigma_{i-1}, \sigma_i)$, bit-rate R changes slightly, from minimal value $R_{\min} = R(\sigma_{i-1}) = 6.358$ bits/sample, to maximal value $R_{\max} = R(\sigma_i) = 6.507$ bits/sample. This means that small error will be done if $R(\sigma)$, for $\sigma \in (\sigma_{i-1}, \sigma_i)$, approximates with $R(s_i) = 6.432$ bits/sample. For larger K, this error will be even smaller. For example, for $K = 64$ we have that $R_{\min} = R(\sigma_{i-1}) = 6.40$ bits/sample, $R_{\max} = R(\sigma_i) = 6.475$ bits/sample and $R(s_i) = 6.438$ bits/sample. In this way, validity of approximation, given with (30), is proved.

4. Experimental Results for Speech Coding

In this section we present experimental results, obtained by applying models from Section 3 on speech signal. In our experiment we consider F frames with M speech samples. In order to provide the experimental values of the average signal to quantization noise ratio within the each of F frames we define the following relation:

$$SQNR_p^{ex} = 10\log_{10}\frac{(\sigma_p^{ex})^2}{D_p^{ex}}, \quad p = 1, \ldots, F, \tag{32}$$

where $(\sigma_p^{ex})^2$ denotes the variance of the input speech samples within the pth frame, $p = 1, \dots, F$:

$$\left(\sigma_p^{ex}\right)^2 = \frac{1}{M}\sum_{q=1}^{M} x_{pq}^2, \quad p = 1, \dots, F \tag{33}$$

and D_p^{ex} denotes the average distortion for the pth frame, $p = 1, \dots, F$:

$$D_p^{ex} = \frac{1}{M}\sum_{q=1}^{M} \left(x_{pq} - y_{pq}^a\right)^2, \quad p = 1, \dots, F. \tag{34}$$

With x_{pq} and y_{pq}^a are denoted the input speech samples and the outputs of the adaptive semilogarithmic quantizer, respectively. By averaging the signal to quantization noise ratios within the each of F frames (32), we can obtain experimental results (average values of $SQNR$, denoted with $SQNR_a^{ex}$).

$$SQNR_a^{ex} = \frac{1}{F}\sum_{p=1}^{F} SQNR_p^{ex}. \tag{35}$$

Experiments are done for two cases. In the first case, forward adaptive semilogarithmic quantizer (described in Section 3.1) is used. Experimental value of average $SQNR$ is 35.45 dB. Theoretical value, obtained in Section 3.1, is 35.116 dB. We can see that matching of theoretical and experimental results is very good, since difference is less than 1.5%.

In the second case, experiment is done with model described in Section 3.2, which consists of forward adaptive quantizer and lossless encoder. Experimental value for average bit-rate R_{av} is 6.33 bits/sample, whereas theoretical value (from Section 3.2) is 6.432 bits/sample. The difference is less than 1.5%. Therefore, for this case we also have very well matching between theoretical and experimental results.

Experimental results for both described cases are presented in Fig. 8.

5. Conclusion

In this paper the novel semilogarithmic quantizer was presented. It was shown that it is much more suitable for adaptation than A-law quantizer. Forward adaptation of that novel quantizer was done, and it was shown that G.712 standard can be satisfied with bit-rate of 7.02 bits/sample, which is not possible with A-law. After that, lossless encoder was applied, with a view to further decreasing bit-rate, and G.712 standard was satisfied with bit-rate of 6.43 bits/sample. Joined design of adaptive quantizer and lossless encoder was done, and therefore optimal performances were obtained. Theoretical results were verified by experiment on speech signal. It was shown that matching of theoretical and experimental results was very well, since the difference was less than 1.5%. We can

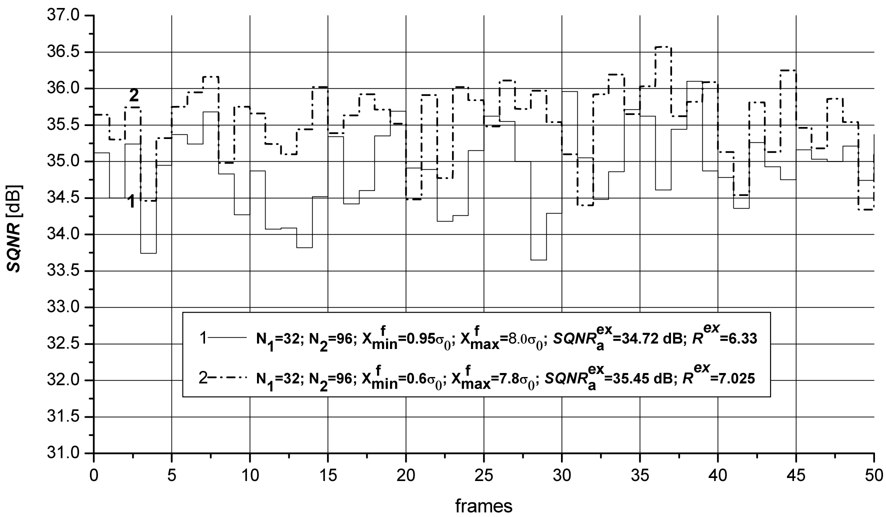

Fig. 8. Experimental results for $SQNR$ for models described in Sections 3.1 and 3.2.

conclude that models presented in this paper are good solution for compression of speech signal and other signals with Laplacian distribution. Future work and further improvements of this model can be done using perceptual measures and voice detection, at it was done in Kajackas and Anskaitis (2009).

References

Aldajani, M.A. (2008). Logarithmic quantization in the least mean squares algorithm. *Digital Signal Processing*, 18, 321–333.

Bastys, A., Kisel, A., Salna, B. (2010). The use of group delay features of linear prediction model for speaker recognition. *Informatica*, 21(1), 1–12.

Chu, C.W. (2003). *Speech Coding Algorithms*. Wiley, New York, 143–183.

Hankerson, D., Harris, G.A., Johnson Jr., P.D. (2004). *Introduction Information Theory and Data Compression*. 2nd ed. Chapman & Hall, London.

Hersent, J.P., Gurle, D. (2005). *Beyond VoIP Protocols, Understanding Voice Technology and Networking Techniques for IP Telephony*. Wiley, New York.

Hiwasaki, Y., Ohmuro, H., Mori, T., Kurihara, S., Kataoka, A. (2006). A G.711 embedded wideband speech coding for VoIP conferences. In: *IEICE Transactions on Info and Systems,* E89-D(9), pp. 2542–2552.

ITU-T, Recommendation G.711. (1972). Pulse Code Modulation (PCM) of Voice Frequencies.

Jayant, N.S., Noll, P. (1984). *Digital Coding of Waveforms*. Prentice-Hall, New York.

Kajackas, A., Anskaitis, A. (2009). An investigation of the perceptual value of voice frames. *Informatica*, 20(4), 487–498.

Kondoz, A. (2004). *Digital Speech, Coding for Low Bit Rate Communication Systems*. Wiley, New York.

Lyon, A.D. (2008). The μ-law CODEC. *Journal of Object Technology*, 7(8), 17–31.

Nikolić, J., Perić, Z. (2008). Lloyd–Max's algorithm implementation in speech coding based on forward adaptive technique. *Informatica*, 19(2), 255–270.

Peric, Z., Nikolic, J. (2007). An effective method for initialization of Lloyd–Max's algorithm of optimal scalar quantization for Laplacian source. *Informatica*, 18(2), 279–288.

Peric, Z., Mosic, A., Panic, S. (2008). Robust and switched nonuniform scalar quantization of Gaussian source in a wide dinamic range of power. *Automatic Control and Computer Science*, 42(6), 334–341.

Peric, Z., Petkovic, M., Dincic, M. (2009). Simple compression algorithm for memoryless laplacian source based on the optimal companding technique. *Informatica*, 20(1), 99–114.

Sakran, H., Shokair, M., Elazm, A.A. (2009). Combined interleaving and companding for PAPR reduction in OFDM systems. *Progress in Electromagnetics Research C*, 6, 67–78.

Sayood, K. (2006). *Introduction to Data Compression*. Elsevier, Amsterdam.

Zavadskas, E.K., Turskis, Z. (2008). A new logarithmic normalization method in games theory. *Informatica*, 19(2), 303–314.

Z.H. Perić was born in Nis, Serbia, in 1964. He received the BSc degree in electronics and telecommunications from the Faculty of Electronic Engineering, Nis, Serbia, Yugoslavia, in 1989, and MSc degree in telecommunications from the University of Nis, in 1994. He received the PhD degree from the University of Nis, also, in 1999. He is currently professor at the Department of Telecommunications and vicedean of the Faculty of Electronic Engineering, University of Nis, Serbia. His current research interests include the information theory, source and channel coding and signal processing. He is particularly working on scalar and vector quantization techniques in speech and image coding. He was author and coauthor in over 150 papers in digital communications. Dr Peric has been a reviewer for: IEEE Transactions on Information Theory, Informatica, COMPEL. He is member of Editorial Board of Journal "Electronics and Electrical Engineering".

M.S. Savić was born in Vranje, Serbia, in 1984. He received the BSc degree in computer science from the Faculty of Electronic Engineering, Nis, in 2008. He is currently on doctoral studies in telecommunications on the same faculty, and he is scholar of Ministry of Science Republic of Serbia. His current research interests include source coding and quantization of speech signals and images.

M.R. Dincic was born in Niš, Serbia, in 1983. He received the BSc degree in telecommunications from the Faculty of Electronic Engineering, Nis, in 2007. He is currently on doctoral studies on the same faculty, and he is scholar of Ministry of Science Republic of Serbia. His current research interests include source coding and quantization of speech signals and images.

D.B. Denic was born in Vlase, township Vranje, Serbia, in 1963. He graduated on Faculty of Electronic Engineering in May 1988, with theme in field of measurement automatization. He received the MSc degree in electrical science from the University of Niš, in 1992. He received the PhD degree from the University of Niš, in 1996. He is currently a professor at the Department of Measurement, University of Niš, Serbia. He actively participated in many projects of science and research for Ministry of Science and Technological Development and in many significant projects for industry. He is the author or coauthor of more than 50 papers published in international journals or conference proceedings. As expert he was involved in period of more than 3 years as independent consultant of GPI (Gurley Precision Instruments), leading company in his field of research. His main fields of interest are the measurement science, position measurement, telemetry, computer based measurement systems, scalar quantization and A/D conversion.

M.R. Prascevic is an assistant professor of Noise and Vibration at the University of Nis, Faculty of Occupational Safety of Nis. He was born on May 15, 1965. He received the

BS degree and the MS degree in electronic engineering from University of Nis, Faculty of Electronic Engineering, Serbia, in 1990 and 1995, respectively, and the PhD degree in environmental engineering from University of Nis, Faculty of Occupational Safety of Nis, Serbia, in 2004. His research areas include environmental noise, measurement and control of noise and vibration, speech signal processing and machine condition monitoring. He is a member of the International Institute of Acoustics and Vibration and Serbian Acoustical Society.

Originalaus logaritminio tipo kvantuoklio ir koderio be nuostolių tiesioginė adaptacija kalbos signalų suglaudinimui

Zoran H. PERIC, Milan S. SAVIC, Milan R. DINCIC, Dragan B. DENIC, Momir R. PRASCEVIC

Šiame straipsnyje yra pateiktas naujas logaritminio tipo kvantuoklis Laplaso skirstiniui. Jis yra paprastesnis negu klasikinis A-dėsnio logaritminio tipo kvantuoklis, kadangi jis turi vienetinį stiprinimą nulio aplinkoje. Taip pat juo gaunamas 2.97 dB aukštesnis signalo/kvantavimo triukšmo santykis *(SQNR)* etaloninei dispersijai atžvilgiu A-dėsnio kvantuoklio, todėl jis labiau tinkamas adaptacijai. Kvantuoklio tiesioginė adaptacija yra atliekama kalbos signalo kadrams. Šiuo būdu G.712 standartas yra išpildomas naudojant 7 bitus/imčiai, kas yra neįmanoma klasikiniam A-dėsnio kvantuokliui. Kiekvieno kadro viduje yra formuojami pokadriai ir jiems naudojamas koderis be nuostolių. Tokiu atveju yra atliekama dviguba adaptacija: adaptacija pagal dispersiją kadrų viduje ir adaptacija pagal amplitudę pokadrių viduje. Yra atliekamas koderio be nuostolių ir kvantuoklio sujungimas, kuris užtikrina geresnį darbingumą. Naudojant pasiūlytą kvantuoklį G.712 standartas yra patenkinamas su tiktai 6.43 bitų/imčiai. Pateikti eksperimentų rezultatai, gauti pritaikant šį modelį kalbos signalui. Parodyta, kad eksperimentiniai ir teoriniai rezultatai sutampa labai gerai (skirtumas yra mažesnis nei 1.5%). Šiame straipsnyje pateikti modeliai gali būti pritaikyti kalbos signalui ir bet kokiam kitam signalui, turinčiam Laplaso skirstinį.