Rev. Roum. Sci. Techn.– Électrotechn. et Énerg.
Vol. **62**, *3*, pp. 299–304, Bucarest, 2017

MODIFIED WIDEBAND SPEECH CODING SYSTEM WITH EMBEDDED G.711 CODERS

MILAN TANČIĆ, ZORAN PERIĆ, NIKOLA SIMIĆ

Key words: G.711 coders, G.711.1 standard, Segmental uniform quantization, Hadamard transform coding.

In this paper, we have proposed a novel model for wideband speech signal coding, structurally similar to G.711.1 coding system. Unlike G.711.1 coding system, the proposed model has lower complexity that is provided by employing G.711 coders in both branches after applying transformation coding. This way, signal is not being divided into frames, which leads to substantially reduced processing delay. Furthermore, special attention is paid to the determination of optimal values of support range for G.711 coders. System performance is measured by using signal-to-quantization-noise ratio which represents a standard measure of reconstructed signal quality. Experiments are done for male and female speakers, due to difference in speech signal dynamics. The obtained results show that the proposed model ensures experimental gain up to 5 dB compared to G. 711 model, whereas signal processing delay is decreased from 5 ms to 0.25 ms, compared to G. 711.1 standard.

1. INTRODUCTION

The basic method for speech signal coding, the first one standardized by ITU-T, is recommendation G.711 [1, 2]. This method covers speech signal sampling at 8 kHz, encoding with 8 bits/sample and provides 64 kb/s bit-rate transmission by using log-compressed pulse-code modulation and it presents the standard for narrowband speech signal coding. In order to provide wideband speech signal transmission in commercial services, firstly it has been proposed standardization of wideband speech coding scalable with G.711 at the ITU-T meeting in 2008. This method, standardized by ITU-T, is entitled as G.711.1 standard [3, 4]. The purpose of this extension is to provide wideband scalability for G.711. The main feature of G.711.1 is that signal is divided into two bands using low-pass and high-pass frequency filters. In the branch after low-pass filter, it is used G.711 coder for encoding, whereas in the branch after high-pass filter, it is used interleaved conjugate-structured vector quantization. Signal filtering is commonly used for processing other types of signals, as images or audio, too [5, 6]. Unlike G.711 standard which is suitable for speech signals, current audio and image processing schemes incorporate some kind of transform coding, such as discrete cosine transform (DCT) [6], to reduce dynamic range, while quantizer optimization could be considered to enhance quality of reconstructed signal [7].

This paper proposes modification of G.711.1 standard which employs G.711 coders in both branches for encoding (that are applied after transforming signal instead of application of low-pass and high-pass filters). This way, system complexity is simplified, since that vector quantization is not implemented. The main motivation for this research was additional reducing of coding delay compared to G.711.1 standard in order to obtain more reliable real-time transmission. According to G.711.1 standard, speech signal is divided into frames of 40 samples which causes 5 ms delay (0.125 ms per sample), whereas the proposed modifications of wideband speech coding system, provide delay reducing to 0.25 ms, which is caused by coding only the first two samples (one sample per branch).

Digitized wideband speech signal, sampled at 16 kHz, firstly is processed through transform mapping scheme which decomposes signal into two independent sequences. The next step of the proposed method is encoding of such decomposed signal. For this task, G.711 coders are used in both branches, after transformations t_1 and t_2, which represent modified Hadamard transformations. In ITU-T G.711 recommendation, speech signal coding is done using a bit-rate of $R = 8$ bits/sample whereas compression factor value is $\mu = 255$. The coder in the branch after transformation t_1 is designed for $R_1 = 8$ bits/sample and $\mu = 255$, according to the G.711 recommendation. In order to reduce the average bit-rate required for coding, the coder we propose in the branch after transformation t_2 is designed for both $R_2 = 4$ and $R_2 = 6$ bits/sample. However, the main designing is the determination of the optimal support range value, thus special attention is paid to finding the optimal values of quantizer's support range in the branches, $x_{\max 1}$ for quantizer Q_1 and $x_{\max 2}$ for quantizer Q_2. Experiments are done for both male and female speech signals sampled at 16 kHz, because of their different dynamics and in order to discuss the optimal support range values.

The support range values for G.711 coders in both branches are obtained as multiplication of standard deviation of adequate branch signal and the optimal support range value of quasi-logarithmic quantizer designed for speech signal modeled with Laplacian probability density function (PDF) of the unit variance ($\sigma^2 = 1$). The values of the standard deviations σ_1 and σ_2, where σ_1 represents the standard deviation for signal y_1 (in the branch after application of transformation t_1), whereas σ_2 describes standard deviation of signal y_2 (in the branch after application of transformation t_2), depend on the input signal's standard deviation σ and its correlation coefficient ρ [8]. However, the input signal can have different variance σ and correlation coefficient ρ depending on the nature of the signal (male and female), thus it is very important to discuss optimal support range values for quantizers that provide the highest signal-to-quantization-noise ratio (SQNR), which is used as a standard objective coding quality measure.

This paper is organized as follows. In Section 2, description of the proposed coding scheme for wideband speech signal with all implementations described in details is presented. In Section 3, numerical results are shown and the analysis of these results is given. Finally, contributions of the paper are summarized in Section 4.

University of Nis, Faculty of Electronic Engineering, Department of Telecommunications, Aleksandra Medvedeva 14, 18000 Nis, Serbia, tel. (+381)18/529214, E-mail: milan_tancic@yahoo.com, zoran.peric@elfak.ni.ac.rs,simicnikola90@gmail.com

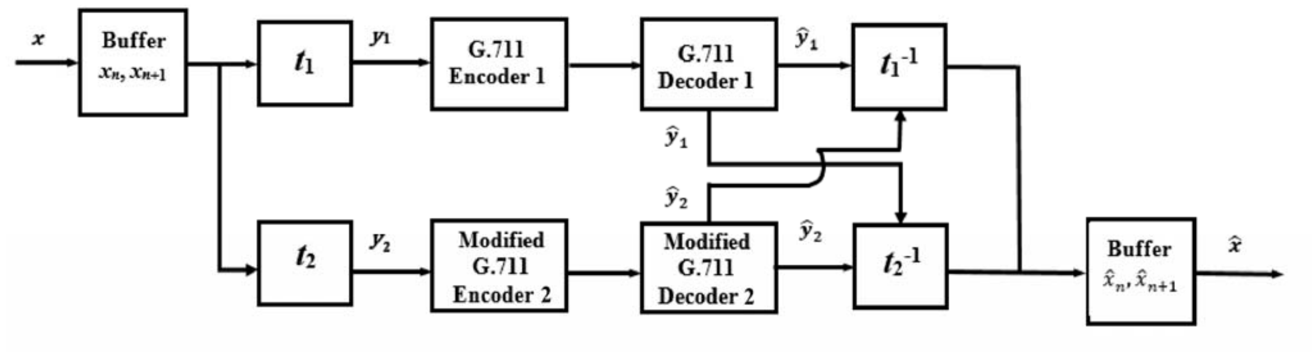

Fig. 1 – Modified wideband speech coding system.

2. WIDEBAND SPEECH SIGNAL CODING SYSTEM WITH EMBEDDED G.711 CODERS

Nowadays, the most common method for speech signal coding sampled at 16 kHz (wideband extension) is G.711.1 standard. It provides high quality of reconstructed signal, but processing delay is still relatively high and it can produce problems in unreliable networks. As a result, the main idea of this paper is to provide lower processing delay, ensuring high quality of reconstructed signal. Speech signal coding in this paper is designed by modifying G.711.1 standard [3, 4]. Wideband speech signal is firstly decomposed into two independent sequences (two branches) using transformations t_1 and t_2, while in the second step, it is encoded using G.711 coders. The main purpose of transformations is to exploit speech signal correlation [1, 9]. In Fig. 1, the proposed scheme for coding of the wideband speech signal is shown [1, 8, 10–13].

After decomposition and transformation, signal is fed to G.711 encoders designed for different bit-rates. Transformations that are used in this coding scheme are defined using the following expressions [8]:

$$t_1 = \frac{x_k + x_{k+1}}{2}, \tag{1}$$

$$t_2 = \frac{x_k - x_{k+1}}{2}, \tag{2}$$

where x_k and x_{k+1} are samples of the information signal, which are transformed to a new signal sequences y_1 and y_2. Transformations defined in this way represent the modification of the original Hadamard H_1 transform [8]. Modification is reflected in the usage of different coefficient which multiplies 2×2 matrix, since it is used $1/2$ instead of $1/\sqrt{2}$. This way, narrower dynamics range is obtained for the signals sequences y_1 and y_2 compared to the input signal x dynamics range. On the other hand, signal reconstruction using inverse transformations is simple as the form of transform and inverse matrices is the same. After transformation, these independent sequences are coded separately by using G.711 coders. G.711 coders at transmitter and receiver present G.711 quantizers, which are implemented using segmental uniform quantization. This kind of quantization works as follows.

Firstly, N representational levels of quantizer are divided into L segments, wheareas each segment consists of n levels. In order to achieve higher precision, we have decided to use $L = 16$ segments, where the number of levels in each segment $n = N / L$ depends on the total number of quantization levels N, i.e. the total average bit-rate of a branch ($R_1 = 2^{N(Q1)}$, $R_2 = 2^{N(Q2)}$). Signal sequence y_1 is fed at the entrance of quantizer Q_1 (G.711 Encoder1/ Decoder1 in the Fig. 1), while signal sequence y_2 is fed at the entrance of quantizer Q_2 (Modified G.711 Encoder2/ Decoder2 in the Fig. 1). The quantizer Q_1 is designed for bit-rate R_1=8 bits/sample, according to G.711 standard, whereas modification, compared to G.711 standard, is reflected in the fact that quantizer Q_2 is designed for lower bit-rates, $R_2 = 4$ and $R_2 = 6$ bits/sample. The boundaries between these segments are decision thresholds x_i, and they are defined as:

$$x_i = \frac{\left(2^i - 1\right)}{255} \cdot x_{\max}, \quad i = 1,\ 2,\ ...,\ L, \tag{3}$$

where $x_{\max}$ represents quantizers' support range whereas i is a counter of quantizers' segments that goes from 1st to Lth segment. In each segment L_i, n representational levels are allocated and its width is defined as:

$$\Delta_i = \frac{x_{i+1} - x_i}{n}, \tag{4}$$

where x_i and x_{i+1} are decision thresholds, while n is a number of representational levels placed into a segment. Decision thresholds x'_j and representational levels y'_j for quantization level placed in each segment ($j = 1, 2, \ldots, n$) are defined as:

$$x'_j = j \cdot \Delta,\ j = 1, 2, \ldots, n, \tag{5}$$

$$y'_j = (j - 0.5) \cdot \Delta,\ j = 1, 2, \ldots, n, \tag{6}$$

where Δ presents quantization level width, j is a counter of representational levels inside of each segment L_i that goes from 1st to nth representational level.

At the receiver, quantized sequences $\hat{y}_1$ and $\hat{y}_2$ form quantized signal $\hat{x}$, using inverse transformations t_1^{-1} and t_2^{-1}:

$$t_1^{-1} = \hat{y}_1 + \hat{y}_2 = \hat{x}_k, \tag{7}$$

$$t_2^{-1} = \hat{y}_1 - \hat{y}_2 = \hat{x}_{k+1}, \tag{8}$$

where $\hat{x}_k$ and $\hat{x}_{k+1}$ represent samples of the quantized signal $\hat{x}$.

One of the main tasks in this paper is the determination of the optimal values for the support range of quantizers Q_1 and Q_2, since those values have great influence on reconstructed signal quality, measured with SQNR. The support range values $x_{\max 1}$ and $x_{\max 2}$ (for quantizers Q_1 and Q_2 respectively), depend on the signal variances σ_1^2 and σ_2^2

of sequences y_1 and y_2, obtained after applying transformations. These variances depend on the input signal variance σ_x^2 and correlation coefficient ρ [14]:

$$\sigma_1^2 = \frac{\sigma_x^2}{2}(1+\rho), \tag{9}$$

$$\sigma_2^2 = \frac{\sigma_x^2}{2}(1-\rho). \tag{10}$$

The values of support ranges $x_{\max1}$ and $x_{\max2}$ are obtained as multiple values of σ_1 and σ_2, respectively, and they can be defined as:

$$x_{\max 1} = l\cdot\sigma_1\cdot x_{\max}^{\sigma=1}, \tag{11}$$

$$x_{\max 2} = l\cdot\sigma_2\cdot x_{\max}^{\sigma=1}. \tag{12}$$

The parameters σ_1 and σ_2 in equations (11−12) are standard deviations (square-roots of variances σ_1^2 and σ_2^2) of signals y_1 and y_2, respectively, l is a coefficient of multiplication whereas $x_{\max}^{\sigma=1}$ is an optimal support range value for Laplacian source of the unit variance, given by [15–18]:

$$x_{\max}^{\sigma=1} = \frac{1}{\sqrt{2}}\ln\left(\frac{3\mu N^2}{\ln^2(\mu+1)}\right). \tag{13}$$

Furthermore, the parameter μ presents compression factor while N is a number of representational levels and it represents the average number of representational levels N_1 and N_2 for which are designed quantizers Q_1 and Q_2, respectively.

Due to the importance of bit allocation, the optimal values of R_1 and R_2 can be obtained using the following equations [8]:

$$R_m = R + \frac{1}{2}\log_2\frac{\sigma_m^2}{\prod_{m=1}^{M}\left(\sigma_m^2\right)^{\frac{1}{M}}}. \tag{14}$$

The equation (14) represents general expression for obtaining optimal bit-rate where M is the total number of branches whereas m can have values 1 or 2 for the proposed model. Consequently, the optimal values of R_1 and R_2 for the proposed coding scheme could be defined as:

$$R_1 = R + \frac{1}{4}\log_2\frac{1+\rho}{1-\rho}, \tag{15}$$

$$R_2 = R + \frac{1}{4}\log_2\frac{1-\rho}{1+\rho}. \tag{16}$$

Theoretical background is given through calculating SQNR and SQNR gain in the case when the signal modeled with Laplacian source is present at the entrance of the system. Laplacian probability density function for certain quantization level is defined as [1]:

$$p_i(x_{\max},\sigma) = \int_{x_i'}^{x_{i+1}'}\frac{1}{\sqrt{2}\sigma}\exp\left(-\frac{|x|\sqrt{2}}{\sigma}\right)\mathrm{d}x\,. \tag{17}$$

The granular distortion depends on the number of representational levels, quantizer's support range and the input signal standard deviation and it is defined with:

$$D_g(N, x_{\max},\sigma) = \frac{1}{6}\sum_{i=0}^{L-1}\Delta_i^2\cdot p_i(x_{\max},\sigma)\,. \tag{18}$$

On the other hand, the overload distortion depends on the quantizer's support range and the input signal variance:

$$D_{ov}(x_{\max},\sigma) = \sigma^2\cdot\exp\left(-\frac{\sqrt{2}x_{\max}}{\sigma}\right). \tag{19}$$

In the end, the total distortion represents the sum of the granular and overload distortion:

$$D = D_g + D_{ov}. \tag{20}$$

As it is presented in (18) and (19), the total distortion is highly dependent on the quantizer's support range. Thus, it is very important to design quantizers for optimal values of $x_{\max}$.

The quality of reconstructed signal is measured using signal-to-quantization-noise ratio (SQNR) which is a standard measure and it is defined as [1, 7, 8] and [14–16]:

$$\mathrm{SQNR[dB]} = 10\cdot\log_{10}\left(\frac{\sigma^2}{D}\right). \tag{21}$$

In addition, quality of the proposed coding scheme can also be measured by calculating the gain of coded signal over G.711 standard [1, 8]:

$$G[\mathrm{dB}] = 10\cdot\log_{10}\left(\frac{D_{G.711}(R,\sigma)}{D(R_1,R_2,\sigma)}\right). \tag{22}$$

3. EXPERIMENTAL RESULTS

In this section, numerical results obtained by theoretical calculations and by performing experiments are presented. Theoretical calculations are done for the case of the Laplacian source. The calculations are based on both granular and overload distortion estimation, using expressions (18) and (19), respectively, while SQNR is obtained using expression (21). The results are shown for medium bit-rates of G.711 coder, *i.e.* $R = 6$ bits/sample and $R = 7$ bits/sample, whereas the proposed scheme is designed so that the total average bit-rate remains equal.

Firstly, in Figs. 2 and 3 it is presented the theoretical comparison of SQNR for a wide range of the input signal variances [15, 16].

Figure 2 shows the SQNR for the proposed coding scheme over G.711 standard for the case when the bit-rates of the quantizers Q_1 and Q_2 have values $R_1 = 8$ bits/sample and $R_2 = 4$ bits/sample, respectively, whereas the quantizer in G.711 standard is modified and designed for $R = 6$ bits/sample, in order to provide corresponding comparison. By observing Fig. 2 in the range of input signal variances [–20 dB, 5 dB], it can be seen that theoretically calculated SQNR of the proposed coding scheme achieves about 30 dB, that is about 5.5 dB better performance comparing to G.711 standard. However, the exact gain, which represents the difference between achieved SQNR and the one obtained by using G.711 standard according to (22), depends on the input signal variance. By observing Fig. 2, it can be concluded that the proposed coding scheme achieves gain between 3.5 dB and 7 dB over G.711 standard in the range of input signal variance range [–15 dB, 20 dB].

To confirm the suitability of the proposed coding scheme, theoretical calculations are done also for the bit-rate cases $R_1 = 8$ bits/sample and $R_2 = 6$ bits/sample, while quantizer from G.711 standard is modified and designed for

$R = 7$ bits/sample. This way, the total average bit-rates of the proposed and compared system are equal.

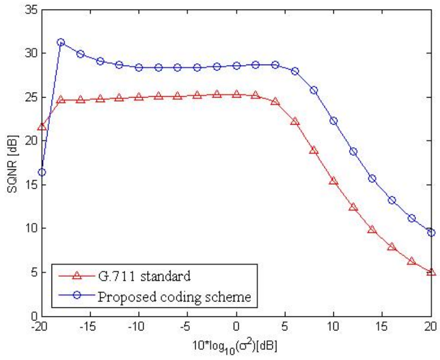

Fig. 2 – SQNR for G.711 standard ($R = 6$ bits/sample) and the proposed coding scheme ($R_1 = 8$ bits/sample, $R_2 = 4$ bits/sample).

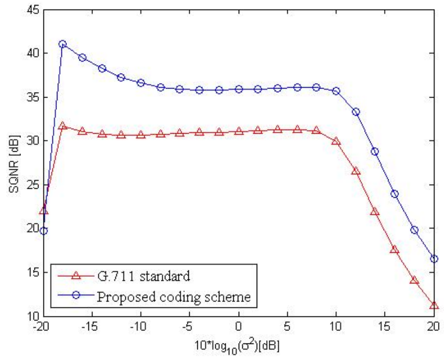

Fig. 3 – SQNR in wide range of input signal variances for G.711 standard ($R = 7$ bits/sample) and the proposed coding scheme ($R_1 = 8$ bits/sample, $R_2 = 6$ bits/sample).

In Fig. 3 it is shown SQNR for the proposed coding scheme and G.711 standard, depending on the input signal variance for $R_1 = 8$ bits/sample and $R_2 = 6$ bits/sample.

Figure 3 shows that theoretically calculated SQNR of the proposed coding scheme achieves about 35 dB in the range of input signal variances [–20 dB, 10 dB], with the peak of 41 dB in the case when the input signal variance is –17.5 dB, while it approximately linearly decreases for higher variances. It can be noticed that SQNR for G.711 standard is about 5.5 dB lower comparing to the proposed scheme. The corresponding gain for the proposed coding scheme over G.711 standard, in the case of $R_1 = 8$ bits/sample and $R_2 = 6$ bits/sample, for the variance range [–20 dB, –17.5 dB] is varying from negative (–2 dB) to positive (9 dB). This means that for extremely low values of variance and for the observed bit-rates, the proposed scheme does not provide better performance. However, this is valid only for the input signal variance range around –20 dB. Moreover, it can be seen that the proposed coding scheme achieves about 6 dB gain over G.711 standard in the range of input signal variances [–15 dB, 20 dB].

This way, it can be concluded that the proposed coding scheme is suitable for wideband speech signal coding, since that high quality (4÷7 dB higher SQNR over G.711 standard for the first bit-rate combination and 5÷9 dB for the second bit-rate combination) is achieved, providing lower complexity at the same time.

After theoretical comparison, in the rest of this section will be discussed experimental results. As it was already said, the experiments are done for both male and female speech signal, recorded in Laboratory of Acoustics, Faculty of Electronic Engineering, University of Nis. As it could be expected, the difference in signals' dynamics causes different coding quality. The compression factor μ by G.711 recommendation has value $\mu = 255$ and it is the same for all observed experimental and theoretical cases. Moreover, the experiment is done for two sets of quantizers' support range values. Firstly, support range values of quantizers are obtained using (11) and (12), whereas in the second case, these values are twice larger.

Table 1 shows the results for male speech signal processing. The input signal is sampled at 16 kHz (wideband extension), since that sampling at this frequency provides less correlation coefficient of the input signal than in the case when it is sampled at 44.1 kHz (standard for audio CD and MPEG-1). The correlation coefficient in this case is $\rho = 0.9672$. Next, the input signal variance of recorded male signal is $\sigma^2 = 0.0021$. The experiment is performed for two sets of quantizers' Q_1 and Q_2 support range values. Firstly, the results are provided for $l = 1$ and the set is obtained using (11) and (12). This way obtained values are: $x_{\max 1} = 0.3733$ (the support range for quantizer Q_1) and $x_{\max 2} = 0.0482$ (the support range for quantizer Q_2). The support range for corresponding G.711 model has the same value as the support range of quantizer in the branch after applying transformation t_1 in the proposed system. The bit-rates for Q_1 and Q_2 are R_1=8 bits/sample (G.711 standard) and $R_2 = 4$ bits/sample (the modification of G.711 standard), while for the purpose of comparison, the bit-rate for G.711 standard is also modified and it amounts $R = 6$ bits/sample, which is the average value of bit-rates for quantizers Q_1 and Q_2. This way is provided appropriate comparison of the proposed scheme and G.711 standard.

From Table 1, it can easily be seen that proposed system for wideband speech signal coding provides the gain of 3.8951 dB over G.711 standard. However, the advantage of the proposed coding scheme is even better for higher values of the support ranges for quantizers Q_1 and Q_2 and optimal values are obtained using twice larger support region ($l = 2$). The aforementioned values are: $x_{\max 1} = 0.7466$ and $x_{\max 2} = 0.0964$. The support range value for the quantizer in G.711 standard has the same value as quantizer Q_1 of the proposed coding scheme and it amounts $x_{\max} = 0.7466$. In this case, the proposed system provides 4.383 dB higher SQNR comparing to G.711 standard. It has been experimentally shown that this set of support range values for quantizers Q_1 and Q_2 is optimal because every other set causes decreasing in the sense of SQNR.

In the case of female speech signal, which has greater dynamics, the results are similar. In Table 2, the results of female speech signal coding using the proposed system and G.711 standard are presented. The sampling frequency is the same as in the case of male speech signal processing $f = 16$ kHz, correlation coefficient is $\rho = 0.9586$, whereas the signal variance amounts $\sigma^2 = 0.0062$. The first set of support range values is obtained using (11) and (12) and the obtained values are: $x_{\max 1} = 0.6367$ and $x_{\max 2} = 0.0925$. The support range value for quantizer in G.711 standard has value $x_{\max} = 0.6367$, like support range value for quantizer Q_1 of the proposed coding scheme. The bit-rate for quantizer in G.711 standard amounts $R = 6$ bits/sample due to the adequate comparison with the proposed coding, since that its average bit-rate amounts $R = 6$ bit/sample.

Table 1

SQNR - the proposed coding scheme and G.711 in the case of male speech signal for the average bit-rate R = 6 bits/sample

		First set of support range values		Second set of support range values	
R	μ	SQNR[dB]	$SQNR_{G.711}$ [dB]	SQNR[dB]	$SQNR_{G.711}$ [dB]
$R_1 = 8$ $R_2 = 4$	255	28.6658	24.7077	29.5098	25.1268

Table 2

SQNR – the proposed coding scheme and G.711 in the case of female speech signal for the average bit-rate R = 6 bit/sample

		First set of support range values		Second set of support range values	
R	μ	SQNR[dB]	$SQNR_{G.711}$ [dB]	SQNR[dB]	$SQNR_{G.711}$ [dB]
$R_1 = 8$ $R_2 = 4$	255	25.9416	23.5089	28.6113	25.0890

The greater dynamics of female speech signal causes lower gain and SQNR, comparing to the processed male speech signal, for 2.7242 dB using the proposed system for coding, and 1.1988 dB lower SQNR using G.711 standard. However, SQNR obtained by using the proposed scheme is still 2.4327 dB higher than classic G.711 standard. The advantage of the proposed system is higher for the second set of support range values, which is twice larger than the first one. The obtained support range values are $x_{max1} = 1.2734$ and $x_{max2} = 0.185$ for Q_1 and Q_2, respectively. For this set of support range values ($l = 2$), the proposed coding scheme provides 3.5223 dB higher SQNR in comparison to G.711 standard.

Table 3 shows SQNR for the same male speech signal as Table 1, but for different values of bit-rates for which the quantizers are designed: $R_1 = 8$ bits/sample (G.711 standard) and $R_2 = 6$ bits/sample (modified G.711 standard) for the proposed coding scheme, whereas bit-rate for G.711 standard is also modified and amounts $R = 7$ bits/sample which presents the average value of bit-rates for quantizers Q_1 and Q_2. The sampling frequency, correlation coefficient and signal variance remains the same, as in the case presented in Table 1, but the support range values are different because quantizers are designed for different bit-rates. The first set of support range values ($l = 1$) amounts $x_{max1} = 0.4128$ for quantizer Q_1 and $x_{max2} = 0.0540$ for quantizer Q_2. The support range value for G.711 standard is the same as the support range value for quantizer Q_1 and amounts $x_{max1} = 0.4128$.

From Table 3 it is evident that the proposed scheme for wideband speech signal coding provides the gain of 3.5893 dB over G.711 standard, for the first set of the support range values. Also, for the second set of support range values ($l = 2$), that are twice larger and equal to $x_{max1} = 0.8355$ and $x_{max2} = 0.1081$, for the proposed coding scheme and $x_{max} = 0.8355$ for G.711 standard. For this set of support range values, the proposed coding scheme provides 5.0292 dB higher SQNR comparing to G.711 standard.

Finally, Table 4 shows the results of female speech signal processing, the same as in Table 2, for different system parameters: $R_1 = 8$ bits/sample (according to the G.711 standard) and $R_2 = 6$ bits/sample (modification of the G.711 standard) for the proposed coding scheme, whereas G.711 standard is also modified and its quantizer is designed for $R = 7$ bits/sample. The sampling frequency remains the same and it is equal $f = 16$ kHz, correlation coefficient is $\rho = 0.9586$ while the signal variance amounts $\sigma^2 = 0.0062$. The first set of support range values ($l = 1$) is obtained using (11) and (12) and the obtained values are: $x_{max1} = 0.7132$ and $x_{max2} = 0.1036$. The support range value for quantizer in G.711 standard has value $x_{max} = 0.7132$, like support range value for quantizer Q_1 of the proposed coding scheme. The bit-rate for quantizer in G.711 standard amounts $R = 7$ bits/sample due to the adequate comparison. In the case of the second set for support range values ($l = 2$), when these values are twice larger and equal to $x_{max1} = 1.4264$ and $x_{max2} = 0.2072$ for the proposed coding scheme and $x_{max} = 1.4264$ for G.711 standard, the proposed coding scheme provides the gain of 4.9335 dB over G.711 standard.

Table 3

SQNR - the proposed coding scheme and G.711 in the case of male speech signal for the average bit-rate R = 7 bits/sample

		First set of support range values		Second set of support range values	
R	μ	SQNR[dB]	$SQNR_{G.711}$ [dB]	SQNR[dB]	$SQNR_{G.711}$ [dB]
R_1=8 R_2=6	255	34.7762	31.1869	36.1886	31.1594

Table 4

SQNR - the proposed coding scheme and G.711 in the case of female speech signal for the average bit-rate R = 7 bits/sample

		First set of support range values		Second set of support range values	
R	μ	SQNR[dB]	$SQNR_{G.711}$ [dB]	SQNR[dB]	$SQNR_{G.711}$ [dB]
R_1=8 R_2=6	255	30.9070	28.8145	36.005	31.0715

By observing Table 3 and Table 4, it can be concluded that female speech signal processing reaches 3.8692 dB lower SQNR, comparing to male speech signal using the proposed coding scheme for $R = 7$ bits/sample, while female speech signal achieves 2.3724 dB lower SQNR using G.711 standard. However, SQNR obtained by using the proposed scheme is 2.0925 dB higher than the one achieved using G.711 standard, for the first set of support range values whereas for the second set of support range values proposed scheme provides 4.9335 dB of gain over G.711 standard. In the case of mail speech signal, the proposed coding scheme provides 3.5893 dB of gain over G.711 standard for the first set of support range values, whereas for the second set of support range values the gain amounts 5.0292 dB. This means that the advantage of the proposed coding scheme is increased when the support range values for quantizers are twice larger, which is considered as the second set of these values. The second set of these values present optimal values because for any other set of support range values decreases the coding quality.

4. CONCLUSIONS

In this paper it is proposed a new scheme for wideband speech signal coding, structurally similar to G.711.1 standard with difference in the fact that the proposed scheme applies G.711 coders in both branches unlike G.711.1 standard which exploits vector quantization in one branch. This fact decreases signal delay from 5 ms to 0.25 ms, because signal is not divided into frames and only first two samples affect on the delay (0.125 ms per sample). Theoretically, the proposed scheme provides averagely 5.5 dB higher SQNR for speech signal processing, modeled with Laplacian source, comparing to G.711 standard, in the case when the average bit-rate is R = 6 bits/sample, and 6.5 dB, in the case when the average bit-rate is R=7 bits/sample.

The experiment has been done for both male and female speech signal for two sets of bit-rates as well as two sets of support range values. The results have shown that the proposed scheme for male speech signal processing provides gain about 3.5 ÷ 4 dB over G.711 standard for the first discussed set of the support range values, whereas the second set provides 4.5 ÷ 5 dB higher SQNR than G.711 standard, depending of quantizers' bit-rate, and it provides optimal values of support range values. Discussion of different bit-rate values is provided, too. It has been shown that female speech signal processing is followed with lower SQNR performance comparing to male speech signal, for the same system parameters, due to the fact that it has greater dynamics. However, female speech signal processing using the proposed scheme still provides gain of 2.5 dB and 3.5 dB averagely, over G.711 standard for both sets of support range values, respectively. The adequate comparison for different bit-rates is provided, too.

To sum up the proposed scheme provides higher SQNR for both male and female signal processing, comparing to G.711 standard. Furthermore, the complexity of the proposed system provides 20 times lower delay comparing to G.711.1 standard, and 2.5 ÷ 5 dB higher SQNR comparing to G.711 standard. The main idea of this paper was to propose a coding scheme for wideband speech signal processing (sampled at 16 kHz) which provides more reliable real-time transmission by reducing processing delay that is generated with G.711.1 standard, but on the other hand, to achieve great quality of reconstructed signal, which is compared to G.711 standard. In the future, we will intend to provide even higher reconstructed signal quality by applying forward or backward adaptation technique. Moreover, we will provide experiments for different specific applications and research possible difficulties of processing a signal of various variances.

AKNOWLEDGEMENTS

This work is supported by Serbian Ministry of Education, Science and Technologic Development (Project TR 32035).

Received on February 11, 2016

REFERENCES

1. N.S. Jayant, P. Noll, *Digital Coding of Waveforms*, Prentice Hall Pb, 1984.
2. *** *Pulse code modulation (PCM) of voice frequencies*, ITU-T Recommendation G.711, 1972.
3. Hiwasaki, Yusuke, Hitoshi Ohmuro, ITU-T G. 711.1: *Extending G. 711 to higher-quality wideband speech*, Communications Magazine, IEEE 47.10 (2009): 110–116.
4. *** *Wideband embedded extension for G.711 pulse code modulation*, ITU-T Recommendation G.711.1, 2008.
5. A. Ciobanu, V. Nita, C. Negrescu, D. Stanomir, *Improved Audio Edit Detection Using the Phase Analysis of the Harmonics of the Electric Network Signal*, Rev. Roum. Sci. Techn. – Électrotechn. et Énerg. **61**, *4*, pp. 394–397, 2016.
6. R.O. Preda, D.N. Vizireanu, S. Halunga, *Active Image Forgery Detection Scheme Based on Semi-Fragile Watermarking*, Rev. Roum. Sci. Techn. – Électrotechn. et Énerg., **61**, *1*, pp. 58–62, Bucarest, 2016.
7. Z. Peric, L. Velimirovic, M. Stankovic, A. Jovanovic, *A Comprehensive Analysis of the Scalar Compandor Model Designed Using Spline Functions*, Rev. Roum. Sci. Techn. – Électrotechn. et Énerg., **60**, *3*, pp. 283–292, 2015.
8. K. Sayood, *Introduction in data compression*, 3rd ed, 2005.
9. R.J.P. de Figueiredo, *A neural-network-based approach to speech signal prediction*, Rev. Roum. Sci. Techn. – Électrotechn. et Énerg.,**55**, *1*, pp. 42–48, 2010.
10. Y. Hiwasaki, T. Mori, S. Sasaki, H. Ohmuro, A. Kataoka, *A wideband speech and audio coding candidate for ITU-T G.711 WBE standardization*, Proc. of the IEEE ICASSP, 2008, pp. 4017–4020.
11. Y. Hiwasaki, H. Ohmuro, T. Mori, S. Kurihara, A. Kataoka, *A G.711 Embedded Wideband Speech Coding for VoIP Conferences*, IEICE Transactions on Information and Systems, **9**, pp. 2542–2552, 2006.
12. Jerry D.Gibson, *Speech Compression*, Information, *7*, *2*, pp. 2078–2489, 2016.
13. Z. Peric, J. Nikolic, J. Lukic, D. Denic, *Analysis of Two-stage Quantizer with Embedded G.711 Quantizer and Segmental Uniform Quantizer*, Electronics & Electrical Engineering, **19**, *2*, pp. 88–91, 2013.
14. Z. Peric, M. Tancic, S. Tomic, D. Ciric, *Subband Coding of Audio Signal with Logarithmic Compandors*, Proceedings of Telsiks 2015, Serbia, Nis, October 14–17, pp. 19–22, 2015.
15. Z. Perić, D. Aleksić, M. Stefanović, J. Nikolić, *New approach to support region determination of the μ-law quantizer*, Elektronika Ir Elektrotechnika, Electronics & Electrical Engineerings, **19**, *8*, pp. 111–114, 2013.
16. D. Aleksić, Z. Perić, J. Nikolić, *Support region determination of quasi-logarithmic quanitzer for Laplacian source*, Przeglad Elektro-tevhniczny, **88**, *7A*, pp. 130–132, 2012.
17. S. Na, *On the Support of Fixed-Rate Minimum Mean- Squared Error Scalar Quantizers for a Laplacian Source*, IEEE Trans. Inf. Theory, **50**, *5*, pp. 937–944, 2004.
18. S. Na, D.L. Neuhoff, *On the Support of MSE-optimal, Fixed-rate, Scalar Quantizers*, IEEE Trans. Inf. Theory, **47**, *7*, pp. 2972–2982, 2001.