An Algorithm for Simple Differential Speech Coding Based on Backward Adaptation Technique

Stefan S. TOMIĆ*, Zoran H. PERIĆ, Jelena R. NIKOLIĆ

Faculty of Electronic Engineering, University of Nis
Aleksandra Medvedeva 14, Nis, Serbia
e-mail: stefan@elfak.rs, zoran.peric@elfak.ni.ac.rs, jelena.nikolic@elfak.ni.ac.rs

Abstract. This paper presents a simple differential speech signal coding algorithm, based on backward adaptation. The considered algorithm is executed in frame by frame manner, by implementing predictive and adaptive quantization techniques. Both prediction and adaptation are performed backward, based on the previously quantized input signal frame. This enables us to obtain high quality output signal, without increasing the bit rate. This research puts emphasis on the quantizer design, with the optimal support limit determination, and theoretical performance evaluation. Objective quality of the output signal is evaluated through signal to quantization noise ratio (SQNR). We perform theoretical and experimental analysis of the algorithm performance and provide comparative results of implementing speech signal coding techniques with similar complexity. Experimental results show that our simple differential speech coding algorithm satisfies the G.712 Recommendation for high-quality speech coding at the bit rate of 6 bits per sample. This indicates that the algorithm can be successfully implemented in high quality speech signal coding.

Key words: speech coding algorithm, backward adaptive technique, prediction coding, differential coding.

1. Introduction

It is well known that due to living in the digital era, we meet digitized signals daily, regardless of our occupation and personal interest. As capability to speak represents a crucial human characteristic, which separates humans from other living beings, speech signal takes high place in digital signal processing. Therefore, digital processing of speech signal has always occupied interest and a lot of research was conducted, for the purpose of representing speech signal with the smaller bit rate, while preserving sufficient quality (Jayant and Noll, 1984; Kondoz, 2004; Chu, 2003; Sayood, 2017; Rabiner and Schafer, 1978). Speech signal coding implies the application of a data compression technique to digitized speech signal samples, by using certain amount of bits per sample (Gibson, 2016). The goal is to achieve high quality, by using fewer bits per sample possible. Speech signal coding plays a crucial role in any technology used for speech transmission, as: voice over Internet protocol (VOIP), digital cellular communications, and any

---

*Corresponding author.

system working with digitized speech (Chu, 2003). Latest expansion in software applications, which offer functionality of voice and video communication, further emphasizes on speech coding, for the purpose of more efficient usage of the available data transmission bandwidth. By lowering the bit rate, we also lower the objective quality of the signal, so that signal compression always presents a compromise between the desired signal quality and the available bandwidth for transmission, or memory for storing the output signal. Different speech coding standards require fulfilling certain conditions regarding bit rate and signal quality (Chu, 2003).

This paper presents a novel waveform speech signal coding algorithm, based on differential coding scheme. Waveform coding is suitable for application in speech signal coding, and it is known that this type of coding can provide the highest level of speech signal quality (Jayant and Noll, 1984; Kondoz, 2004; Chu, 2003). One of the most familiar and commonly used, simple waveform coder is defined by the G.711 standard (ITU-T, 1972). G.711 coder is known as Pulse Code Modulation (PCM), with two basic types, μ-law and A-law. As PCM utilizes 8 bits per sample, a lot of research was performed for reducing the bit rate, while preserving original quality of the input signal (Kondoz, 2004; Gibson, 2016). Along with the bit rate, modifications and improvements have been done in the usage of available frequency bands (ITU-T, 2008). For instance, ITU-T, Recommendation G.711.1 allows implementation in narrowband and wideband speech signal coding, with sample frequency up to 16 kHz and bit rate ranging from 64 to 90 kbit/s. Since the used bit rate has significant impact on overall algorithm performance, certain techniques have been introduced to lower the bit rate, like linear prediction technique (Jayant and Noll, 1984; Kondoz, 2004; Chu, 2003). Linear prediction implies that the value of the current input signal sample can be represented as a linear combination of the previous samples. The earliest system which implemented this technique was Differential Pulse Code Modulation (DPCM), where prediction has been implemented for calculating the error (difference) signal (Jayant and Noll, 1984; Suma, 2012). Difference signal is obtained by subtracting the input signal sample from its predicted value. The main benefit of implementing differential coding is that the obtained difference signal is characterized by smaller variance and lower dynamic amplitude range. This means that the difference signal is more convenient for coding than the original input signal, regardless of the applied coding scheme type. Simplicity and benefits from implementing differential coding maintains this and related techniques interesting for exploitation and motivate researchers to consider it when designing new solutions and algorithms. For instance, in Uddin *et al.* (2016), a low bit rate speech signal coding algorithm with the implementation of DPCM has been analysed, and it has been shown that differential coding is suitable for application in coding of correlated speech signals. Predictive coding, which utilizes correlation as well, has been recently implemented in designing a LPC-Based Fronthaul Compression Scheme (Ramalho *et al.*, 2017, 2018), designed to compress LTE signal, by means of predictive and Huffman coding. The compression system from Ramalho *et al.* (2017, 2018) is characterized by low computational complexity and low latency, while preserving high output signal quality. Microcontroller implementation of DPCM and ADPCM in speech signal compression has also been recently presented

in Sarade (2017), and it has been shown how ADPCM can take advantage of correlated signals to achieve more efficient signal compression compared to DPCM technique.

Further implementing the adaptation techniques into DPCM resulted in ADPCM (Adaptive DPCM) algorithm, which exploits correlation between the input signal samples and reduces the bit rate required for coding (Jayant and Noll, 1984). Adaptation is most commonly implemented as forward or backward adaptation (Jayant and Noll, 1984; Kondoz, 2004; Chu, 2003; Sayood, 2017). Adaptive quantization has wide area of implementation, as it represents a simple method for increasing the objective quality of the output signal. It is successfully applied in speech coding (Suma, 2012; Dinčić *et al.*, 2016), image and video coding (Ortega and Ramchandran, 1995), as well as in wireless sensor networks (Fang and Li, 2008). In this paper, we combine adaptive quantization with simple differential coding to perform high quality speech coding. The algorithm is convenient for improvements by implementing more complex differential coding and quantization techniques. As we design a low complexity algorithm, this is left to future research.

The coding algorithm developed in this research is based on the backward adaptive technique, so it does not require transmission of the side information. This allows the developed coding algorithm to be used in the cases when small delay is required, without increasing the bit rate, as is the case of using forward adaptation, where transmission and coding of the side information is required. The goal is to achieve high quality output speech signal, which would meet the G.712 Recommendation for a high-quality speech coding (ITU-T, 2001), by using a low bit rate. Speech signal can be modelled with Gaussian or Laplacian probability density function (p. d. f.) (Kondoz, 2004). The algorithm we propose considers both models of the input signal p. d. f. and as will be shown, in the case of the Gaussian distribution it provides higher quality of the output signal for each value of the compression factor μ. Eventually, for achieving high performance of the proposed algorithm, it is important to determine the support limit of the quantizer in the specific manner to minimize the signal distortion (Na and Neuhoff, 2001), which we also take into account in designing our quantizer.

The rest of the paper is organized as follows. Section 2 describes the support limit determination for the Gaussian source model. In Section 3, the basic principles of differential coding are described. The description of the novel differential speech signal coding algorithm is presented in Section 4, while the results of its application are presented and discussed in Section 5. Finally, Section 6 is devoted to conclusions.

2. Support Limit Determination for the Quasilogarithmic Quantizer of Gaussian Source

In designing speech signal coding algorithm, input signal has to be modelled by some probability distribution function (PDF). It is well known that speech signal can be successfully modelled by Gaussian (normal) distribution (Jayant and Noll, 1984; Kondoz, 2004; Chu, 2003; Sayood, 2017). PDF for a random variable with the Gaussian distribution with

mean value denoted by α and variance by σ^2 is defined as Jayant and Noll (1984):

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

Without loss of generality, we can assume that information source is memoryless Gaussian source with mean value equal to zero. PDF of thus defined Gaussian source is given by:

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

Quantization is a significant part in the digitalization process. A quantizer is defined as the structure of encoder and decoder (Jayant and Noll, 1984). Role of a quantizer is to perform mapping of the input signal amplitudes into the group of permitted amplitudes. An N-level scalar quantizer Q is defined by mapping $Q: R \in Y$, where R represents a set of real numbers, while $Y \equiv y_1, y_2, y_3, \ldots, y_N \subset R$, is a set of representation levels that makes the code book of size $|Y| = N$ (Jayant and Noll, 1984). Scalar quantizer divides a set of real numbers into N cells $R_i = (t_{i-1}, t_i]$, $i = 1, \ldots, N$, where t_i, $i = 0, 1, \ldots, N$ are decision thresholds of the quantizer, defined as $Q(x) = y_i$, $x \in R_i$. In practice, during the process of scalar quantization, the region of an input signal is divided into granular and overload regions, which are separated by the support region thresholds $-t_{N-1}$ and t_{N-1}. Minimum and maximum support region thresholds define the support region of the quantizer $[-t_{N-1}, t_{N-1}]$, which has a high influence on the total distortion of the quantizer (Jayant and Noll, 1984; Na and Neuhoff, 2001; Na, 2004). In order to determine the optimal support region, we implement an iterative numerical method for determining the support region thresholds. Optimal support limit can be found by minimizing the total distortion. In the case of quasilogarithmic quantizer and an input signal modelled with Gaussian source, with mean value equal to zero and variance $\sigma^2 = 1$, granular distortion is defined by Jayant and Noll (1984):

$$D_g = \frac{\ln^2(\mu + 1)}{3N^2} \left[\frac{t_{N-1}^2}{\mu^2} + t_{N-1} \frac{2}{\mu} \sqrt{\frac{2}{\pi}} + 1 \right], \tag{3}$$

where μ represents the compression factor. In the case of implementing larger compression factors, as is the case with the quantizer defined with the G.711 standard, where $\mu = 255$, $(\mu \gg 1)$ the term of the equation $\frac{1}{\mu^2}$ tends towards zero, so that Eq. (3) can be approximated as:

$$D_g = \frac{\ln^2(\mu + 1)}{3N^2} \left[\frac{2}{\mu} \sqrt{\frac{2}{\pi}} t_{N-1} + 1 \right]. \tag{4}$$

For a quantizer with a support region threshold t_{N-1} and input signal modelled by Gaussian PDF of zero mean and unit variance, overload distortion can be calculated as Na and

Table 1
Support limit convergation for different starting values chosen.

R bit rate	Compression factor μ	Starting value t_{N-1}	1st iteration	2nd iteration	3rd iteration
		3	4.3869	4.1760	4.2028
		3.5	4.3001	4.1869	4.2014
6	255	4	4.2264	4.1963	4.2002
		4.5	4.1622	4.2046	4.1991
		5	4.1052	4.2122	4.1981

Neuhoff (2001):

$$D_o = \sqrt{\frac{2}{\pi}} \frac{2}{t_{N-1}^3} \exp\left\{\frac{-t_{N-1}^2}{2}\right\}. \tag{5}$$

The total distortion represents the sum of the granular and the overload distortion $D = D_g + D_o$. Minimizing the distortion in relation to the t_{N-1} is performed as:

$$\frac{\partial D}{\partial t_{N-1}} = 0. \tag{6}$$

From Eq. (6) we can obtain the optimal support limit, which is iteratively calculated by:

$$t_{N-1}^{(i)} = \sqrt{\left|2\ln\left(\frac{1}{\mu}\frac{\ln^2(\mu+1)}{3N^2}\frac{t_{N-1}^{4(i-1)}}{t_{N-1}^{2(i-1)}+3}\right)\right|}. \tag{7}$$

For the purpose of increasing the algorithm speed, we use only two iterations and obtain the result near the optimal one, where optimality refers to the solution of Eq. (6). The initial values are chosen intuitively, as it is known that the iterative procedure converges to the same output for different initial values. Table 1 shows the convergation of support limit values, when different starting values are chosen.

As already stated, in this research we have chosen to implement quasilogarithmic quantizer to perform input signal quantization. It represents a good choice, especially in the case of adaptive quantization, where we can easily adjust the quantizer by changing the compression factor. Also, a quasilogarithmic quantizer can be defined by the closed form formulas, which makes the design and analysis simple. Quasilogarithmic quantizer performs signal compression by applying a compressor function defined by Jayant and Noll (1984):

$$c_\mu(t) = \frac{t_{N-1}}{\ln(1+\mu)}\ln\left(1+\mu\frac{|t|}{t_{N-1}}\right)\text{sgn}(t), \quad |t| \leqslant t_{N-1}. \tag{8}$$

By finding and implementing the optimal support limit and choosing the appropriate compression factor, we obtain a simple and efficient quantizer, which combined with the differential coding and backward adaptation technique is capable to provide a high quality output signal.

3. Theoretical Background of Differential Coding

Signals in nature are often correlated, meaning that the successive samples have values similar to a certain extent (Jayant and Noll, 1984; Chu, 2003). Wideband speech signal, sampled at 16 kHz, can be considered to be a correlated source, which can be practically examined by calculating the correlation coefficient of a real speech signal. For an input signal $x[n]$, $n = 1, \ldots, M$, with the length of M samples, the correlation coefficient can be calculated by:

$$\rho = \frac{\sum_{n=1}^{M-1} x[n] \times x[n+1]}{\sum_{n=1}^{M} (x[n])^2}. \tag{9}$$

Differential coding (DPCM) exploits the correlation of the successive samples of the input signal to reduce the redundancy of the input. It is based on quantizing the difference (prediction-error) signal $e[n]$, which is obtained by subtracting the input signal $x[n]$ from the signal prediction $x_p[n]$ formed between the actual value of the current sample and predicted value of the previous sample (Jayant and Noll, 1984; Chu, 2003). Prediction is commonly performed by implementing linear prediction, in which current sample is predicted as a linear combination of the previously quantized samples (Jayant and Noll, 1984; Makhoul, 1975; Markel and Gray, 2013). The number of previous samples used in prediction represents the order of a predictor. Function of the k-th order linear predictor is defined as:

$$x_p[n] = \sum_{l=1}^{k} a_l \hat{x}[n-1], \tag{10}$$

where a_l, $l = 1, 2, \ldots, k$ are the prediction coefficients, while $\hat{x}[n-1]$ represents the previously quantized input signal samples. The main benefit of differential approach is that prediction-error signal has smaller variance and dynamic amplitude range, which makes it more suitable for quantization (Suma, 2012). Quantizing prediction-error signal enables us to achieve higher SQNR for a given resolution, or equally given number of quantization levels N.

DPCM encoder and decoder are presented in Fig. 1. Output of the encoder are indices $i[n]$, by which the decoder obtains quantized prediction error, which is used in forming the quantized input. The thus defined DPCM system has been successfully implemented in speech (Jayant and Noll, 1984; Ortega and Ramchandran, 1995), image coding (Zschunke, 1977), as well as in ECG signal coding (Peric *et al.*, 2013a). We implement DPCM principles in designing a speech coding algorithm presented in this paper. The combination of differential coding and backward adaptation provides a simple and efficient coding scheme without increasing the bit rate used. The contribution of this research lies also in the optimal quantizer design with utilization of correlation in speech signal coding. This enables us to obtain high quality output signal, while implementing the simplest differential coding. The following section describes the design and implementation of the algorithm we propose.
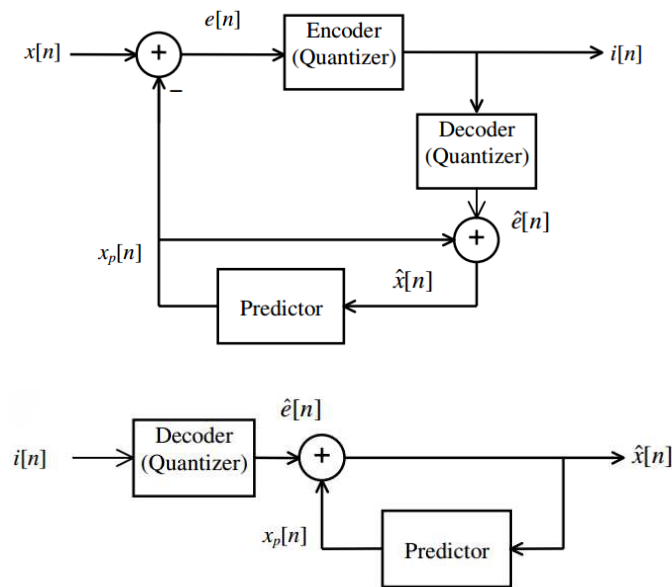
Fig. 1. DPCM encoder (top) and decoder (bottom).

4. Algorithm for Differential Adaptive Coding of Speech Signal

Designing a speech coding algorithm is not a simple task, since a speech coding algorithm has to comply with the following Chu (2003):

1. Use low bit rate;
2. Provide high output signal quality;
3. Be robust in a wide area of possible speakers and languages;
4. Have low computational complexity and low coding delay.

We have chosen to implement low complexity coding techniques with the possibility of using small frame sizes. This enables the algorithm to be implemented when small delay is required, as is the case in real time speech communication or package transmission of data. Since the coding algorithm includes encoder and decoder, both blocks will be described in steps separately. The purpose of the encoder is to convert the input signal into a form convenient for transmission and to send the encoded bit-stream to transmission channel as indices (Jayant and Noll, 1984). The designed algorithm encoder is described in the following.

As we implement backward adaptation technique, the first frame of the input signal does not have the required information for this process, as there was no information before. This is overcome by implementing PCM encoder (ITU-T, 1972) to the first frame of the input signal. As the same bit rate was implemented for all frames, the usage of PCM encoder for the first frame cannot significantly degrade the overall quality of the output, while it provides the initialization of the algorithm. The first encoded frame is used for accessing the statistical information required for the backward adaptation process, which is

based on the previous, encoded frame. In the following steps, we calculate and implement the statistical data from the previous frame for coding the current one. Firstly, we calculate the correlation coefficient, starting from the second frame. The correlation coefficient used for coding the current jth frame can be determined with the implementation of Eq. (9) as:

$$\rho^{(j-1)} = \frac{\sum_{n=1}^{M-1} \hat{x}^{(j-1)}[n] \times \hat{x}^{(j-1)}[n+1]}{\sum_{n=1}^{M} (\hat{x}^{(j-1)}[n])^2}, \quad j = 2, \ldots, L, \tag{11}$$

where L represents the total number of the input signal frames, M is the total number of samples in the frame, while $x[n]$ and $x[n+1]$ denote the current and the next sample of the current input signal frame, respectively. The next parameter used is the variance of the previously quantized frame, defined by:

$$\big[\sigma_{\hat{x}}^{(j-1)}\big]^2 = \frac{1}{M} \sum_{n=1}^{M} \big(\hat{x}^{(j-1)}[n] - \eta^{(j-1)}\big)^2, \tag{12}$$

where $\eta^{(j-1)} = \frac{1}{M} \sum_{n=1}^{M} \hat{x}^{(j-1)}[n]$, $j = 2, \ldots, L$ represents the mean value of the quantized $(j-1)$th signal frame.

Difference signal frame is formed in sample by sample manner, as:

$$d^{(j)}[n] = x^{(j)}[n] - \rho^{(j-1)}\hat{x}^{(j)}[n-1]. \tag{13}$$

The next step is to calculate difference frame variance, which is defined by:

$$\big[\sigma_d^{(j)}\big]^2 = \big[\sigma_{\hat{x}}^{(j-1)}\big]^2 \big(1 - \big[\rho^{(j-1)}\big]^2\big). \tag{14}$$

The last step of encoding is the application of the designed adaptive quasilogarithmic quantizer (Q_{ad}) to the difference signal sample:

$$\hat{d}^{(j)}[n] = Q_{ad}\big[d^{(j)}[n]\big], \tag{15}$$

where the adaptive quantizer has the support region threshold defined by:

$$t_{N-1}^{(j)} = t_{N-1} \times \sigma_d^{(j)}, \tag{16}$$

where t_{N-1} is obtained from Eq. (7). Thus defined adaptive support region thresholds contain information about the input signal, as in the case of implementing backward adaptation and differential coding represents the information about the previously quantized difference signal frame. In this manner, by implementing few simple calculations we adapt the quantization process to the input signal and its statistics and obtain better signal compression. Finally, the output of the described encoder is sent through the channel as binary information (I). The encoded values are received and decoded for the purpose of reconstructing the input signal.

Decoding represents an inverse process to encoding, with the purpose of extracting the input signal from its encoded representation. The first frame of the input signal is decoded by applying the PCM decoder (ITU-T, 1972). By reconstructing the first frame, we are able to extract the data required for decoding the following one. As in the encoding procedure, firstly we extract the correlation coefficient of the previous frame, by implementing Eq. (11). The next statistical parameter from the previous frame used is the variance, calculated by implementing Eq. (12). When we have the information about the previous frame variance, we can apply the inverse quantization to obtain the difference signal sample:

$$\hat{d}^{(j)}[n] = Q_{ad}^{-1}[I].\tag{17}$$

Actual input signal sample can be reconstructed as:

$$\hat{x}^{(j)}[n] = \hat{d}^{(j)}[n] + \rho^{(j-1)}\hat{x}^{(j)}[n-1].\tag{18}$$

This step completes the decoding procedure, as we obtain the values of the quantized input signal samples.

The encoding algorithm can be presented as a series of the following steps:

Step 1. The first frame ($j = 1$) is fed to the buffer.
Step 2. The samples of the first frame are successively encoded by PCM encoder and decoded by local PCM decoder for the purpose of initialization.

From the second to the last frame (for $j = 2, \ldots, L$) the following steps repeat:

Step 3. Each frame is fed to the frame buffer.
Step 4. From the previously quantized frame, the correlation coefficient $\rho^{(j-1)}$ and variance $[\sigma_{\hat{x}}^{(j-1)}]^2$ are calculated (Eqs. (11) and (12)).
Step 5. Difference signal is formed in sample by sample manner, where the correlation of the previously quantized frame and the previously quantized signal sample are used (Eq. (13)).
Step 6. Difference signal is fed to the input of variance-adaptive quasilogarithmic quantizer, which is adapted to the variance $[\sigma_d^{(j)}]^2$ given by Eq. (14).
Step 7. Encoded difference signal frames are sent through the channel as binary information.

The decoding procedure is performed by applying the following steps:

Step 1. The first frame is decoded with PCM decoder.

From the second to the last frame the following is repeated:

Step 2. Output frames are fed to the frame buffer.
Step 3. Information about the previous frame correlation coefficient and variance is extracted.
Step 4. Output signal sample is obtained from the quantized difference signal, quantized previous sample and from the correlation coefficient of the previously quantized frame (Eq. (18)).

Step 5. Samples are fed to the output frame buffer, where we have reconstructed signal samples.

By performing quantization, we introduce an irreversible error due to rounding the current values of the input signal to the representation levels. This error is named quantization error and it can be expressed by distortion, which is commonly defined as the average value of a mean-squared error (Jayant and Noll, 1984). Eqs. (3)–(5) define theoretical distortion for the quasilogarithmic quantization of Gaussian source. In the case of applying algorithm in real input signal coding, distortion can be calculated by:

$$D = \frac{1}{S} \sum_{n=1}^{S} \left(x[n] - \hat{x}[n] \right)^2,$$ (19)

where S represents the total number of the input signal samples. Thus defined distortion with input signal power determines signal to quantization noise ratio, an objective signal quality measure, used in this paper (Jayant and Noll, 1984):

$$\mathrm{SQNR[dB]} = 10 \log_{10} \left(\frac{\frac{1}{S} \sum_{n=1}^{S} (x[n])^2}{D} \right).$$ (20)

5. Numerical Results and Analysis

This section presents the numerical results of applying the proposed algorithm in speech signal coding. In addition, we observe the theoretical characteristics of the basic coding schemes implemented in differential speech signal coding, to show the improvements which are results of the algorithm design. As we implement quasilogarithmic quantization, firstly we analyse the theoretical performance of the quasilogarithmic quantizer. This represents the case when quantization is performed only by applying quasilogarithmic quantization, with 6 bits per sample and the usage of different compression factors. Speech signal is modelled with the Gaussian distribution with zero mean and unit variance, as it has been assumed in the proposed algorithm design. Additionally, we show the theoretical performance of the quasilogarithmic quantizer in the case of using Laplacian distribution with zero mean and unit variance for modelling a speech signal. In the case of using Laplacian distribution, optimal support limit of quasilogarithmic quantizer can be defined as Jayant and Noll (1984):

$$t_{N-1}^{Lap} = \begin{cases} \frac{1}{\sqrt{2}} \ln \left(\frac{3N^2 \mu}{\ln^2(1+\mu)} \right), & \mu \geqslant 255 \\[2em] \frac{1}{\sqrt{2}} \ln \left(\frac{\frac{3N^2 \mu}{\ln^2(1+\mu)}}{1 + \frac{1}{\mu} \ln(\frac{3N^2 \mu}{\ln^2(1+\mu)})} \right), & \mu < 255. \end{cases}$$ (21)
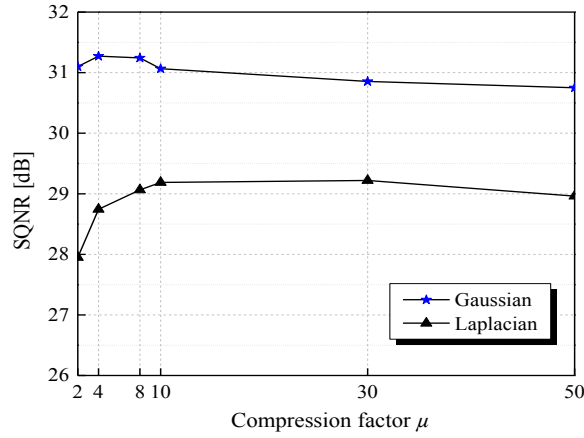
Fig. 2. Theoretical dependence of SQNR on compression factor for the quasilogarithmic quantizer that is designed for Gaussian vs. Laplacian source models for speech signal.

The total distortion of the quasilogarithmic quantizer designed for the Laplacian source can be calculated by Kondoz (2004), Peric *et al.* (2013b):

$$D^{Lap} = \frac{\ln^2(\mu+1)}{3N^2}\left[\frac{(t_{N-1}^{Lap})^2}{\mu^2} + \frac{\sqrt{2}t_{N-1}^{Lap}}{\mu} + 1\right] + \exp\left\{-\sqrt{2}t_{N-1}^{Lap}\right\}. \tag{22}$$

We show the results for the case of using smaller compression factor values, ranging from 2 to 50, as our algorithm implements adaptive quantization and it is expected that it will have the best performance in this range of compression factors. Figure 2 shows the benefits of modelling speech signal with Gaussian source, as the quasilogarithmic quantizer designed for Gaussian source provides better theoretical SQNR characteristics.

The figure shows that the highest theoretical SQNR value in the case of using Gaussian model is obtained for the compression factor equal to 4 and it amounts to around 31.3 dB. When using Laplacian model, maximum is obtained in the case when compression factor is equal to 10, and it amounts to around 29.2 dB. One can notice that the theoretical maximum in SQNR is greater when using Gaussian source model, and the gain amounts to 2.1 dB. For higher compression factor values SQNR characteristics are in steady decline, while the gain is obtained for all compression factor values. This confirms the benefits of modelling speech signal with the Gaussian distribution.

In order to evaluate the performance of the proposed coding algorithm, an experiment was also conducted. The input signal has been a 15 seconds long sequence of a male speech signal, sampled at 16 kHz. The objective quality measure used in the experiment is SQNR, expressed in dB. Along the performance of the proposed algorithm, we analyse the performance of the algorithm without using adaptive quantization and PCM scheme. The advantage of implementing adaptive quantization is observed through the obtained gain in SQNR, especially for lower compression factor values. Comparison with PCM indicates significantly higher objective output signal quality obtained with the implementation of the proposed algorithm.

Table 2
Experimentally obtained SQNR for the proposed algorithm with comparison.

R bit rate	Frame size	Compression factor μ	SQNRP$_{CM}$ [dB]	SQNR$_{NON-AD}$ [dB]	SQNR$_{AD}$ [dB]
		10	24.0995	3.7499	35.7590
		20	24.0995	7.4873	36.3638
		30	24.0995	9.7586	36.4615
6	**40**	**40**	24.0995	11.3872	36.4439
		50	24.0995	12.6633	36.4267
		100	24.0995	16.5379	36.0445
		255	24.0995	21.2701	35.2108

By observing Table 2, one can notice that the proposed algorithm (SQNR$_{AD}$) satisfies the G.712 Recommendation, by using 6 bits per sample for coding, as it provides SQNR greater than 34 dB (ITU-T, 2001), for all values of compression factor implemented. Additionally, the proposed algorithm provides gain in SQNR from 11.1 to about 12.36 dB, when compared to PCM. We should point out that PCM uses fixed compression factor, equal to 255, so SQNR in the column has a single value. The constant gain in SQNR of more than 10 dB, compared to the widely used PCM, confirms that the proposed coding algorithm provides high quality output speech signal. The version of the differential algorithm without implementing adaptive quantization provides lower objective quality of the output signal, for all values of the compression factor used (SQNR$_{NON-AD}$). When we use smaller compression factors, the non-adaptive quantizer is not suitable for application, while for the adaptive version of the algorithm it provides approximately constant performance for different compression factor values. In the case of using compression factor equal to 255, the gain in SQNR is close to 14 dB in favour of the proposed differential speech coding algorithm with adaptive quantization, when compared to non-adaptive version, while the gain amounts to about 11 dB, in comparison to the PCM.

Furthermore, we can compare our results with the results obtained in earlier researches, by using similar or greater complexity coding schemes, where some of them also comply with the G.712 Recommendation. Table 3 shows the comparative results obtained by implementing DPCM with uniform quantization and the second order predictor (Suma, 2012), providing SQNR of about 30 dB, while using 6 bits per sample. One can notice that the proposed algorithm provides gain in SQNR of about 6.5 dB, compared to the case of implementing the basic DPCM coding scheme for the same bit rate. This gain is somewhat justified by the more suitable design of our quantizer. It is also justified by the fact that we have processed the wideband speech signal, which is more correlated compared to the narrowband speech signal processed in Suma (2012). Additionally, Table 3 shows the performance of the fixed and adaptive companding quantizer with variable-length codeword (Perić *et al.*, 2013c) and transform coding with forward adaptive quantization (Tancic *et al.*, 2016). In both cases, SQNR obtained is greater than 34 dB, while it is achieved by using around 6.5 bits per sample. The algorithm proposed in this paper overreaches SQNR of the compared coding schemes by using around 0.5 bits per sample less for coding. We can easily estimate the performance of the proposed algorithm for the case of using 6.5 bits per sample. This could be done by obtaining the SQNR performance for the case of using

Table 3
Comparative SQNR performance of coding schemes which comply with the G. 712 Recommendation.

Bit rate[a]	SQNR[a] [dB]	Bit rate[b]	SQNR[b] [dB]	Bit rate[c]	SQNR[c] [dB]	Bit rate (proposed)	SQNR$_{AD}$ [dB] (proposed)
6	30	6.5	35.143	6.52	34.603	6	36.46

[a] Suma (2012), [b] Perić *et al.* (2013c), [b] Tancic *et al.* (2016).

7 bits per sample and by finding the mean value of this value and the one obtained in the case of using 6 bits per sample from Table 3. The estimated SQNR performance of the proposed algorithm for 6.5 bits per sample amounts to about 39.51 dB. This shows that in the case of using the same bit rate, the proposed algorithm provides gain in SQNR from 4.3 dB to 4.9 dB, when compared to the aforementioned coding schemes.

The experimental results of applying the proposed algorithm to a real speech signal and comparative results shown in Tables 2 and 3, confirm the suitability of the proposed algorithm to be implemented in speech signal coding.

6. Conclusion

This paper has presented a simple, differential wideband speech signal coding algorithm, with the implementation of backward adaptation technique. The algorithm implements low complexity signal coding techniques and provides high quality output speech signal, while using low bit rate. By implementing differential coding, instead of common waveform coding of input signal samples, we perform quantization on the difference signal, which has lower amplitude dynamics and, as shown, is more suitable for performing quantization. The difference signal has been quantized by the backward adaptive quasilogarithmic quantizer, whose design has been presented in this paper. By inspecting the theoretical performance of the quasilogarithmic quantizer, we have demonstrated the advantages of modelling the input signal with Gaussian distribution, since in that case SQNR overreaches the results of using Laplacian distribution to model the input signal. The experimental results have shown the benefits of implementing adaptive quantization and differential speech coding. Moreover, the experimental results have shown that the proposed algorithm can satisfy the G. 712 Recommendation for high-quality speech coding, in the case of using 6 bits per sample for coding. Comparative results of the proposed algorithm with similar complexity speech coding schemes have confirmed that the proposed algorithm can be successfully implemented in speech signal coding.

References

Chu, W.C. (2003). *Speech Coding Algorithms*. John Wiley & Sons, New Jersey.

Dinčić, M., Perić, Z., Jovanović, A. (2016). New coding algorithm based on variable-length codewords for piecewise uniform quantizers. *Informatica*, 27(3), 527–548.

Fang, J., Li, H. (2008). Distributed adaptive quantization for wireless sensor networks: From delta modulation to maximum likelihood. *IEEE Transactions on Signal Processing*, 56(10), 5246–5257.

Gibson, J.D. (2016). Speech compression. *Information*, 7(2), 7–23.

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

ITU-T, Recommendation G.712 (2001). *Transmission Performance Characteristics of Pulse Code Modulation Channels*.

ITU-T, Recommendation G.711.1 (2008). *Wideband Embedded Extension for ITU-T G.711 Pulse Code Modulation*.

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

Kondoz, A.M. (2004). *Digital Speech: Coding for Low Bit Rate Communication Systems*. John Wiley & Sons, New Jersey.

Makhoul, J. (1975). Linear prediction: a tutorial review. *Proceedings of the IEEE*, 63(4), 561–580.

Markel, J.D., Gray, A.J. (2013). *Linear Prediction of Speech*. Springer Science & Business Media, New York.

Na, S. (2004). On the support of fixed-rate minimum mean-squared error scalar quantizers for a Laplacian source. *IEEE Transactions on Information Theory*, 50(5), 937–944.

Na, S., Neuhoff, D.L. (2001). On the support of MSE-optimal, fixed-rate, scalar quantizers. *IEEE Transactions on Information Theory*, 47(7), 2972–2982.

Ortega, A., Ramchandran, K. (1995). Forward-adaptive quantization with optimal overhead cost for image and video coding with applications to MPEG video coders. In *IS&T/SPIE's Symposium on Electronic Imaging: Science & Technology*. International Society for Optics and Photonics, pp. 129–138.

Peric, Z., Denic, D., Nikolic, J., Jocic, A., Jovanovic, A. (2013a). DPCM quantizer adaptation method for efficient ECG signal compression. *Journal of Communications Technology and Electronics*, 58(12), 1241–1250.

Peric, Z., Aleksic, D., Stefanovic, M., Nikolic, J. (2013b). New approach to support region determination of the μ-law quantizer. *Elektronika ir Elektrotechnika*, 19(8), 111–114.

Perić, Z.H., Nikolić, J.R., Mosić, A.V., Petković, M.D. (2013c). Design of fixed and adaptive companding quantizer with variable-length codeword for memoryless Gaussian source. *Informatica*, 24(1), 71–86.

Rabiner, L.R., Schafer, R.W. (1978). *Digital Processing of Speech Signals*. Prentice-Hall, New Jersey.

Ramalho, L., Fonseca, M.N., Klautau, A., Lu, C., Berg, M., Trojer. E., Höst, S. (2017). An LPC-based fronthaul compression scheme. *IEEE Communications Letters*, 21(2), 318–321.

Ramalho, L., Freire, I., Lu, C., Berg, M., Klautau, A. (2018). Improved LPC-based fronthaul compression with high rate adaptation resolution. *IEEE Communications Letters* (accepted for publication).

Sarade, S. (2017). Speech compression by using adaptive differential pulse code modulation (ADPCM) technique with microcontroller. *Journal of Electronics and Communication Systems*, 2(3), 1–9.

Sayood, K. (2017). *Introduction to Data Compression*. Morgan Kaufmann.

Suma, M.N. (2012). Performance analysis of DPCM and ADPCM. In *International Conference on Electronic Design and Signal Processing (ICEDSP)*, pp. 19–23.

Tancic, M., Peric, Z., Tomic, S., Simic, N. (2016). Speech signal coding using forward adaptive quantization and simple transform coding. *Elektronika ir Elektrotechnika*, 22(3), 74–77.

Uddin, S., Ansari, I.R., Naaz, S. (2016). Low bit rate speech coding using differential pulse code modulation. *Advances in Research, SCIENCEDOMAIN International*, 8(3), 2348–2394.

Zschunke, W. (1977). DPCM picture coding with adaptive prediction. *IEEE Transactions on Communications*, 25(11), 1295–1302.

S.S. Tomić was born in Vranje, Serbia, in 1989. He received MS degree in electrical engineering from the Faculty of Electronic Engineering, Nis, Serbia, in 2013. He is currently a student of doctoral studies at the same faculty. His current research interests include audio and speech signal coding. He has published over 10 papers, from which 6 are in peer-reviewed international journals.

Z.H. Perić was born in Niš, Serbia, in 1964. He received the BS, MS and PhD degrees from the Faculty of Electronic Engineering, University of Niš, Serbia, in 1989, 1994 and 1999, respectively. He is a full-time professor at Department of Telecommunications, Faculty of Electronic Engineering, University of Niš. His current research interests include the information theory and signal processing. He is an author and co-author of over 240 papers. Dr. Perić has been a reviewer in a number of journals, including *IEEE Transactions on Information Theory*, *IEEE Transactions on Signal Processing*, *IEEE Transactions on Communications*, *Compel*, *Informatica*, *Information Technology and Control*, *Expert Systems with Applications and Digital Signal Processing*.

J.R. Nikolić was born in Prokuplje, Serbia, in 1978. She received the BS, MS and PhD degrees in electrical engineering from the Faculty of Electronic Engineering, University of Niš, Serbia in 2003, 2006 and 2011, respectively. She is currently an assistant professor at the Department of Telecommunications at the same Faculty. Her current research interests include audio, speech and ECG signal compression and approximation methods with application in source coding. She has published over 100 papers, from which 41 in SCI journals. She is a member of the Editorial Board of the Journal of Advanced Computer Science & Technology.