ANALYSIS OF DIFFERENTIAL PULSE CODE MODULATION WITH FORWARD ADAPTIVE LLOYD-MAX'S QUANTIZER FOR LOW BIT-RATE SPEECH CODING

ZORAN PERIĆ [1], ALEKSANDAR JOCIĆ [1], JELENA NIKOLIĆ [1],
LAZAR VELIMIROVIĆ [2], DRAGAN DENIĆ [1]

Key words: Differential pulse code modulation (DPCM), Lloyd-Max's quantizer, Forward adaptation, Predictor, Correlation coefficient, Prediction gain, Signal to quantization noise ratio.

In this paper, a DPCM (Differential Pulse Code Modulation) system with forward adaptive Lloyd-Max's quantizer is presented. This quantizer is designed for low bit rate, where the first and the second order linear predictors are used in the proposed DPCM system solution. It is shown how SQNR (Signal to Quantization Noise Ratio) and G_p (Prediction Gain) depend on the correlation coefficients of the predictor of the first and the second order. The obtained experimental values of parameters SQNR and G_p are presented and compared with the corresponding theoretical values for the given system. In this manner the selection of optimal DPCM system correlation coefficients values is performed and the possibilities of this DPCM system application in the speech coding are indicated.

1. INTRODUCTION

Speech coding is the process of obtaining a compact representation of speech signals for efficient storage and transmission over band-limited wired and wireless channels [1, 2]. Today, speech coders have become essential components in telecommunications and in multimedia infrastructure [3, 4]. A speech coder converts a digitised speech signal into a coded representation, which is usually transmitted in frames. A speech decoder receives the coded frames and provides the reconstructed speech signal [1–4]. In speech and image coding systems, the ability to intelligently adapt a quantizer in order to best match the varying input signal characteristics is essential for achieving high performance. In forward

[1] University of Nis, Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia; E-mail: zoran.peric@elfak.ni.ac.rs, aleksandar.jocic@elfak.ni.ac.rs, jelena.nikolic@elfak.ni.ac.rs, dragan.denic@elfak.ni.ac.rs.

[2] Mathematical Institute of the Serbian Academy of Sciences and Arts, Kneza Mihaila 36, 11001 Belgrade, Serbia; E-mail: velimirovic.lazar@gmail.com

Rev. Roum. Sci. Techn.– Électrotechn. et Énerg., **58**, *4*, p. 424–434, Bucarest, 2013

adaptive systems, the coder does all the work by determining the optimal adaptation rule and by broadcasting it periodically to the decoder, which can therefore be kept simple. The price paid for this decoder simplicity is the increased complexity of the coder as well as the increased cost due to the transmission of the update quantizer parameters via side-information to the decoder [3, 4].

Realization of Differential Pulse Code Modulation (DPCM) is based on a technique that firstly predicts current sample value based upon previous samples and then performs coding of the difference between actual value of sample and its predicted value [1, 2]. This difference is called prediction error. Since it is necessary to predict sample values, DPCM is a form of predictive coding [1, 2]. In DPCM, code words represent differences between actual values of samples and their predicted values, unlike PCM where code words represent sample values. This is the reason of conducting the analysis presented in this paper for low bit rates. The goal of the paper is to obtain a solution that can be effectively applied for coding of speech signals.

A speech coding algorithm based on forward adaptive technique in which adaptive Lloyd-Max’s quantizer is implemented is described in [5, 6]. Although Lloyd and Max developed an algorithm for designing an optimal quantizer having a minimal possible distortion, this algorithm is too time consuming for the large number of quantization levels [2]. In particular, utilization of Lloyd-Max’s algorithm is demanding from the aspect of arithmetic complexity and memory resources required, which increase with the bit rate. Precisely for this reason, the implementation of Lloyd-Max’s quantizer is mostly limited to lower bit rates, *i.e.* to a smaller number of quantization levels, as we consider in this paper. Unlike [5, 6] where a PCM system is considered, in this paper we consider DPCM system. In addition, unlike [7] and [8], where in DPCM system the optimal companding quantizer is designed and used for high bit-rate speech coding, in this paper we utilize Lloyd-Max's quantizer for low bit-rate speech coding. Lloyd-Max's quantizer is an optimal quantizer for the given probability density function and is simple for design and implementation only for low bit-rates. As highlighted in [5, 6] for high bit-rates Lloyd-Max's quantizer is complex to design and, accordingly, an optimal companding quantizer, which is simpler to design and which has performance close to the optimal one, is more preferable. However, for low bit-rates, we consider in this paper, performances of optimal companding quantizers are far from the optimal. As in [5], in this paper forward adaptive technique is utilized. Specifically, forward adaptive technique is applied to DPCM system in which a linear predictor is designed so that its coefficients are determined from the correlation coefficients of an input signal. The simple DPCM speech coding scheme with the first order switched predictor and forward adaptive quantizer is presented in [7]. The quantizer is realized using the companding model and it is adapted to a short-term estimate of the input signal standard deviation on each frame. An adaptive speech coding scheme based on an optimal companding quantizer and the utilization of correlation is presented in [8]. The main idea in both papers is to provide coding of speech signals at high bit rates with the performance improvement over the PCM by exploiting a correlation between samples within frames, where a higher correlation, a higher improvement. The idea with the utilization of the speech signal prediction has also been observed in [9], where linear predictive coding has been observed. Specifically, different from the quantizer presented in this paper, belonging to the class of waveform coders, in [9] a quantizer from the class of parametric coders has been observed, where only similarity in approaches is in the utilization of the prediction. Due to the significance of correlation in DPCM system, in this paper we study the influence of correlation coefficients on SQNR. The rest of the paper is organized as follows: In section 2 a detailed description of Lloyd–Max's quantizer design is given. Theoretical background of DPCM system is provided in section 3. Section 4 presents and discusses numerical results. Finally, section 5 is devoted to the conclusions which summarise the contribution achieved in the paper.

2. **LLOYD-MAX'S QUANTIZER**

Lloyd and Max proposed an algorithm for designing optimal quantizers using mean-square error distortion measure [1, 2, 10–12]. The most commonly used criterion by which the set of the optimal parameters of a quantizer is determined is the criterion of minimum distortion. Necessary and sufficient conditions for a quantizer optimality for a mean-square distortion measure are described in [1, 2, 10–13]. In accordance with the above mentioned criterion, for a fixed number of quantization levels N, necessary conditions for optimality of the decision thresholds $t_1, t_2, ..., t_{N-1}$ and the representation levels $y_1, y_2, ..., y_N$, are derived by differentiating (finding the first derivative) the following expression for distortion:

$$D = \sum_{i=1}^{N} \int_{t_{i-1}}^{t_i} \left(x - y_i\right)^2 p\left(x\right) \mathrm{d}x \,, \tag{1}$$

per each of these parameters and by equating the obtained expressions to zero:

$$\frac{\partial D}{\partial y_i} = 0 \,, \quad i = 1,..., N \,, \tag{2}$$

$$\frac{\partial D}{\partial t_i} = 0 \,, \quad i = 1,..., N-1 \,, \tag{3}$$

where $t_0 = -\infty$ and $t_N = \infty$. These conditions yield:

$$t_i^{(k+1)} = \frac{y_i^{(k)} + y_{i+1}^{(k)}}{2}, \; i = 1,...,N-1, \tag{4}$$

$$y_i^{(k+1)} = \frac{\int_{t_{i-1}^{(k+1)}}^{t_i^{(k+1)}} xp(x)\,\mathrm{d}\,x}{\int_{t_{i-1}^{(k+1)}}^{t_i^{(k+1)}} p(x)\,\mathrm{d}\,x}, \; i = 1,...,N, \tag{5}$$

where k is a number of iterations ($k = 0,1,...$). The procedure for determining the thresholds and the representation levels is an iterative procedure. Condition for the termination of the procedure is that the relative change of distortion in two successive iterations is less or equal than 0.005 [6].

In this paper, the input signal is modelled by Laplacian probability density function (PDF) [2], where the design of the Lloyd-Max's quantizer is performed for the low bit rates by using the optimal design parameters specified for these bit rates in.

3. THEORETICAL BACKGROUND OF DPCM SYSTEM

DPCM is a technique of converting an analog into a digital signal in which an analog signal is sampled and then the difference between the actual sample value and its predicted value is quantized. A predicted value of the actual sample is based on the value of the previous sample or the values of the previous samples [1, 2]. Basic concept of DPCM is based on the fact that most source signals show a significant correlation between successive samples so that quantizer uses redundancy in sample values which provides lowering bit rate [1, 2]. Fig. 1 illustrates our DPCM system with a Lloyd-Max's quantizer in the adaptive scheme.

The input signal $x[n]$ is processed in a frame by frame manner. The first, samples in a current frame pass through the buffer, and then a gain g is calculated and quantized. Based on the so obtained quantized gain, Lloyd-Max's quantizer adaptation is performed. In the feedback loop of DPCM system there is a fixed predictor and at its output a predictive signal $\hat{x}[n]$ is obtained. The samples values of the signal difference $d\,[n]$ is equal to the difference between the current values of the input samples $x\,[n]$ and their predictive values $\hat{x}[n]$. Quantizer of the signal difference $d[n]$ is a Lloyd-Max's quantizer designed for low bit rate, which is applied in the adaptive scheme. The process of adaptation involves normalization of samples of the signal difference $d[n]$ by using quantized gain $\hat{g}$. Gain g is defined as the square root of the estimated variance of the signal difference σ_d^2 (6), for each frame separately [2]:

$$\sigma_d^2 = \sigma_x^2\left(1-\sum_{i=1}^{P} a_i\rho_i\right),\quad \sigma_x^2 = \frac{1}{M}\sum_{n=1}^{M} x^2(n), \tag{6}$$

where σ_x^2 is the variance of an input signal within a given frame, M is the frame lenght, P is the order of the predictor, ρ_i and a_i, $i = 1, ..., P$ are correlation coefficients of the input signal and coefficients of the P^{th} order predictor.

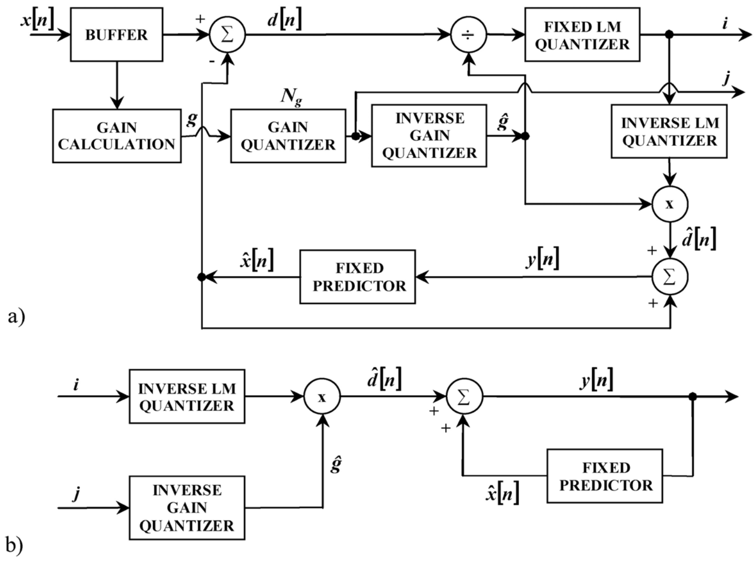


Fig. 1 – DPCM system: a) coder; b) decoder.

For gain quantization ($g = \sigma_d$) we use the log-uniform quantizer, since in [5] it is shown that the log-uniform quantizer has a better performance than the uniform quantizer. Representation levels of the considered gain quantizer are determined by $20\log_{10}(\hat{g}_i) = 20\log_{10}(\sigma_{\min}) + (2i-1)\Delta/2$, $i = 1, ..., N_g$, where $\Delta = 20\log_{10}(\sigma_{\max}/\sigma_{\min})/N_g$ and where dynamic range of the signal difference variance is defined by $[20\log_{10}(\sigma_{\min}), 20\log_{10}(\sigma_{\max})]$ and N_g is the number of the quantization levels of the gain quantizer. Accordingly, it holds $\hat{g} \in \{\hat{g}_1, ..., \hat{g}_{N_g}\}$. Each sample of the signal difference from a specific frame is normalized by the quantized value of the gain $\hat{g}$, thus adapting to the fixed Lloyd-Max's quantizer, which is designed for $\sigma_d = 1$. Due to the quantization error $e_g = g - \hat{g}$, adaptation of samples of the signal difference $d[n]$ is not perfect. In forward adaptive schemes an information of the quantized gain value (index j in Fig. 1) is transferred to the receiving part of the DPCM system (decoder, Fig. 1b), in order to enable denormalization of samples of the signal difference and reconstruction of the output signal samples $y[n]$.

Functional dependency of coefficients of the P^{th} order predictor a_i on correlation coefficients of the input signal ρ_i , $i = 1, ..., P$ is well known [2]. For the cases considered in this paper, when the order of predictor are $P = 1$ and $P = 2$, the coefficients of the fixed predictor are defined as follows:

$$a_1 = \rho_1 \, , \, P = 1, \tag{7}$$

$$a_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2} \, , \; a_2 = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} \, , P = 2. \tag{8}$$

Prediction gain G_p of DPCM system is defined by [2]:

$$G_p = 10\log\left(\frac{\sigma_x^2}{\sigma_d^2}\right) = 10\log\left(\frac{\sum_{n=1}^{M} x^2[n]}{\sum_{n=1}^{M} d^2[n]}\right). \tag{9}$$

By combining (9) with (6) for G_p one obtains:

$$G_p = 10\log\left(\frac{\sigma_x^2}{\sigma_d^2}\right) = 10\log\left(\frac{1}{1-\sum_{i=1}^{P} a_i \rho_i}\right). \tag{10}$$

It can be observed that for a given the P^{th} order predictor, G_p depends only on correlation coefficients ρ_i and on coefficients of predictor a_i, $i = 1, ..., P$.

For evaluation of the quality of the quantized signal, signal to quantization noise ratio is usually used SQNR = 10 log (σ_x^2 / D) [1, 2], where D is a distortion inserted by quantization procedure. Since in this paper we assume frame by frame procession, SQNR for a signal of L frames, each of length M, is then determined by [1, 2]:

$$\text{SQNR}_{\text{DPCM}} = 10\log\left(\frac{\sum_{F=1}^{L}\sum_{n=1}^{M} x_F^2[n]}{\sum_{F=1}^{L}\sum_{n=1}^{M} (x_F[n] - y_F[n])^2}\right), \tag{11}$$

where subscript F here only denotes the ordinal number of frame and where it holds $x_F[n] - y_F[n] = d_F[n] - \hat{d}_F[n]$. Considering (9) and (11) one can rewrite:

$$\text{SQNR}_{\text{DPCM}} = \text{SQNR}_{\text{LM}} + G_{\text{p}}, \tag{12}$$

where $SQNR_{LM}$ is SQNR of the applied Lloyd-Max's quantizer:

$$\mathrm{SQNR}_{\mathrm{LM}} = 10\log\left(\frac{\sigma_d^2}{D}\right). \tag{13}$$

From equations (7)–(10) one can notice that G_p depends on the correlation coefficients ρ_i, $i = 1, ..., P$ of an input signal $x[n]$. This observation allows analysis of the whole DPCM system, i.e. analysis of $SQNR_{DPCM}$ (12), depending on the assumed values of correlation coefficients ρ_i, $i = 1, ..., P$ of an input signal $x[n]$. The obtained results are presented in the following section.

4. NUMERICAL RESULTS

By equations (8) and (10), for $P = 2$ is defined the dependence of prediction gain G_p on correlation coefficients ρ_1 and ρ_2. Fig. 2 shows the dependence of G_p on ρ_2 for the following ρ_1 values: $\rho_1 = 0.70$, $\rho_1 = 0.80$ and $\rho_1 = 0.85$.

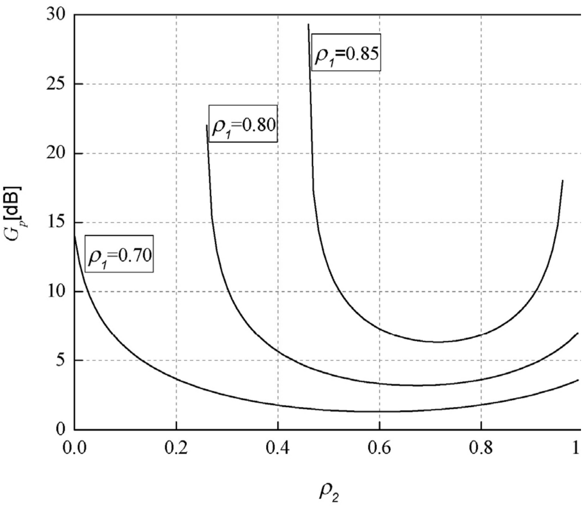


Fig. 2 – Theoretical dependence of prediction gain G_p on correlation coefficients ρ_1 and ρ_2.

These graphs represent the theoretical dependence of G_p on ρ_1 and ρ_2, where a wider range of ρ_2 values is presented than the range from which it takes ρ_2 in realistic calculations. Namely, the coefficient ρ_2 always takes the lower value of coefficient ρ_1 [1, 2, 8]. By appropriate choice of coefficients ρ_1 and ρ_2 one can drastically increase the value of G_p. Accordingly, as will be seen in the following analysis, we propose the choice of coefficients ρ_1 and ρ_2 to provide the maximum of $SQNR_{DPCM}$.

In Figs. 3, 4 and 5 the experimental results are presented as functional dependencies of $SQNR_{DPCM}$ on ρ_1 and ρ_2. The analysis was done on a recorded speech signal of 12000 samples, which was sampled with the frequency of 8 kHz, for the case where the frame length amounts to $M = 200$. The results were obtained by applying (11) on the whole available speech signal for all assumed the values of coefficients ρ_1 and ρ_2. Fig. 3 shows the dependence of $SQNR_{DPCM}$ on ρ_1 for the first-order predictor, and two values of bit rate of quantizer $R = 2$ bit/sample and $R = 3$ bit/sample. The maximum values of $SQNR_{DPCM}$ are obtained for the values of $\rho_1 = 0.81$ and $\rho_1 = 0.87$ for $R = 2$ bit/sample and $R = 3$ bit/sample, respectively (see Table 1). Thus, the maximum value of $SQNR_{DPCM}$ is not achieved for the same value of ρ_1 at different bit rates R, and accordingly, predictor needs special design for different values of R.

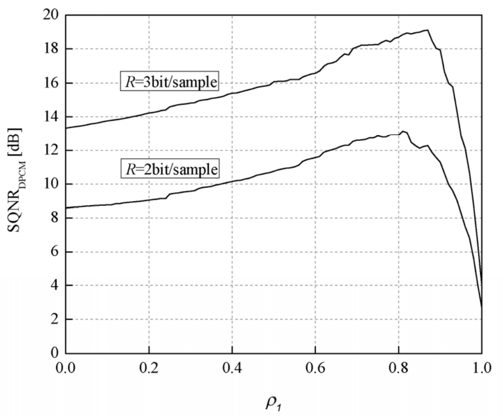


Fig. 3 – Dependency of $SQNR_{DPCM}$ on ρ_1 for $P = 1$, $R = 2$ bit/sample and $R = 3$ bit/sample.

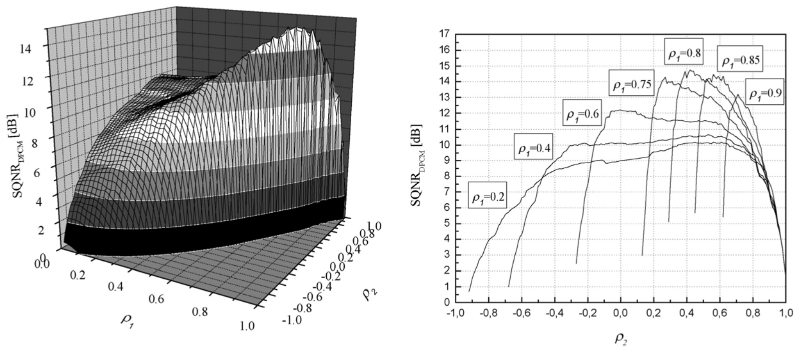
Table 1

The maximum value of $SQNR_{DPCM}$ obtained for the specific values of ρ_1 and ρ_2

R [bit/sample]	$P = 1$		$P = 2$		
	ρ_1	$SQNR_{DPCM}$ [dB]	ρ_1	ρ_2	$SQNR_{DPCM}$ [dB]
2	0.81	13.14	0.82	0.51	14.79
3	0.87	19.11	0.86	0.57	21.47

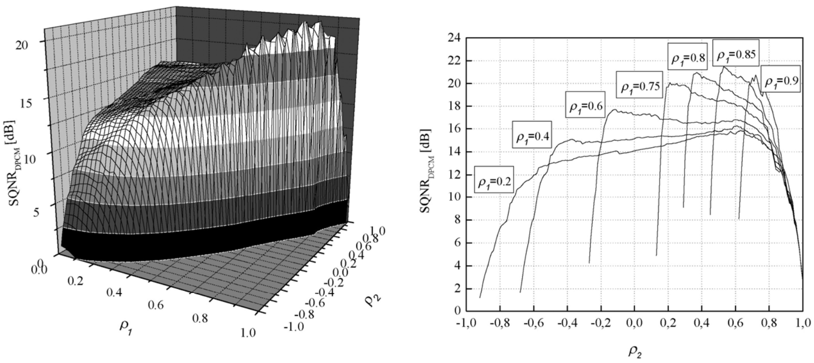
Figs. 4 and 5 show the dependence of $SQNR_{DPCM}$ on ρ_1 and ρ_2 for the second-order predictor. In Fig. 4, this dependence is shown for $R = 2$ bit/sample, and in Fig. 5 for $R = 3$ bit/sample. Figs. 4 a and 5 a show the 3D dependencies of $SQNR_{DPCM}$ on ρ_1 and ρ_2, whereas in Figs. 4 b and 5 b the dependencies of $SQNR_{DPCM}$ on ρ_2 are shown for several fixed values of ρ_1. The maximum value of $SQNR_{DPCM}$ is obtained for $\rho_1 = 0.82$ and $\rho_2 = 0.51$ for $R = 2$ bit/sample whereas for $R = 3$ bit/sample the maximum value of $SQNR_{DPCM}$ is obtained for $\rho_1 = 0.86$ and

$\rho_2 = 0.57$ (Table 1), *i.e.* for the highly correlated cases. In the case of the second-order predictor, the maximum value of $SQNR_{DPCM}$ can not be achieved for the same value of pairs ρ_1 and ρ_2 for a different R so that predictor needs special design for different R.



a) b)

Fig. 4 – Dependency of $SQNR_{DPCM}$ on ρ_1 and ρ_2 for $R = 2$ bit/sample and $P = 2$: a) 3D; b) 2D.



a) b)

Fig. 5 – Dependency of $SQNR_{DPCM}$ on ρ_1 and ρ_2 for $R = 3$ bit/sample and $P = 2$: a) 3D; b) 2D.

Based on the equations (1), (10), (12) and (13), the functional dependence of $SQNR_{DPCM}$ on variance of frame of an input signal σ_x^2 is determined, in the assumed dynamic range of σ_x^2 in dB scale of [–20dB, 20dB] [2]. This dependence is shown in Fig. 6. The analysis is performed for the second order predictor, and the two values of bit rate of quantizer $R = 2$ bit/sample and $R = 3$ bit/sample. For the values of coefficients ρ_1 and ρ_2, the values determined by the experimental analysis were used (Table 1), which achieves maximum value of $SQNR_{DPCM}$. From the Fig. 6 one can observe that $SQNR_{DPCM}$ has approximately constant value in the whole range σ_x^2 in dB scale. The reason for that behavior of $SQNR_{DPCM}$ lies in application of forward adaptation. Obviously the $SQNR_{DPCM}$ characteristic is periodical where the number of periodical intervals equals to the number of quantization levels of the log-uniform gain quantizer (N_g = 16). For R = 2 bit/sample the maximum value of $SQNR_{DPCM}$ amounts to 13.61 dB, where $SQNR_{LM}$ = 7.54 dB and G_p = 6.07 dB, whereas for R = 3 bit/sample the maximum value of $SQNR_{DPCM}$ amounts to 21.06 dB, where $SQNR_{LM}$ = 12.64 dB and G_p = 8.42 dB. These values are noted in Table 2 by index "*theor*", because its represent the theoretically obtained results. In the same table, there are the experimentally obtained values $SQNR_{DPCM}^{exp}$, $SQNR_{LM}^{exp}$ and G_p^{exp} calculated for the available speech signal. One can notice that the experimentally obtained results are better than the theoretically obtained results. The reason is that the experimental results were obtained for the values of ρ_1 and ρ_2 that are optimal for the available speech signal.

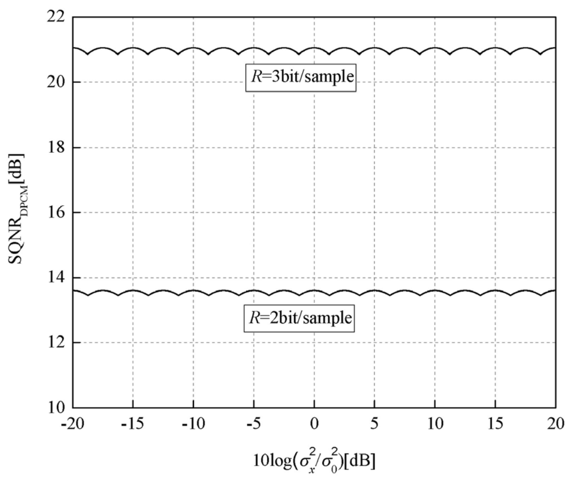


Fig. 6 – Theoretical dependence of $SQNR_{DPCM}$ on the variance of an input signal σ_x^2.

Table 2

Theoretical and experimental values of SQNR and G_p for the selected values of ρ_1 and ρ_2 and P = 2

R [bit/sample]	$SQNR_{LM}^{theor}$ [dB]	$SQNR_{LM}^{exp}$ [dB]	Gp^{theor} [dB]	Gp^{exp} [dB]	$SQNR_{DPCM}^{exp}$ [dB]
2	7.54	8.60	6.07	7.48	14.79
3	12.64	13.33	8.42	8.96	21.47

5. **CONCLUSION**

In this paper, an analysis of DPCM system with forward adaptive Lloyd-Max's quantizer designed for low bit rate and with a linear predictor of the first and the second order is carried out. It is shown that the values of correlation coefficients in the design of the linear predictor of DPCM system have a large impact on the $SQNR_{DPCM}$. By applying the proposed DPCM system on the available speech signal, the conclusion has been derived that for different bit rates of Lloyd-Max's quantizer a separate choice of correlation coefficients should be done to obtain the highest value of $SQNR_{DPCM}$. The presented features of our DPCM system solution indicate that the obtained solution should be of practical significance for quantization of highly correlated signals that as well as speech signals have Laplacian probability density function.

Received on October 29, 2012

ACKNOWLEDGMENTS

This work is supported by Serbian Ministry of Education and Science through Mathematical Institute of Serbian Academy of Sciences and Arts (Project III44006) and by Serbian Ministry of Education and Science (Project TR32045).

REFERENCES

1. A. Gersho, R. M. Gray, *Vector Quantization and Signal Compression*, Norwell, Kluwer Academic Publishers, 1992.
2. N. S. Jayant, P. Noll, *Digital Coding of Waveforms*, New Jersey, Prentice-Hall, 1984.
3. D. Minoli, *Voice over MPLS – Planning and Designing Networks*, New York, McGraw-Hill, 2002.
4. O. Hersent, J. Petit, D. Gurle, *Beyond VoIP Protocols – Understanding Voice Technology and Networking Techniques for IP Telephony*, New Jersey, John Wiley & Sons, 2005.
5. J. Nikolic, Z. Peric, *Lloyd–Max's algorithm implementation in speech coding algorithm based on forward adaptive technique*, Informatica, **19**, *2*, pp. 255–270, 2008.
6. Z. Peric, J. Nikolic, *An effective method for initialization of Lloyd–Max's algorithm of optimal scalar quantization for Laplacian source*, Informatica, **18**, *2*, pp. 279–288, 2007.
7. V. Despotovic, Z. Peric, L. Velimirovic, V. Delic, *DPCM with forward gain-adaptive quantizer and switched first order predictor for high quality speech signals*, Advances in Electrical and Computer Engineering, **10**, *4*, pp. 95–98, 2010.
8. V. Despotovic, Z. Peric, L. Velimirovic, *Switched adaptive quantizer for speech compression based on optimal companding and correlation*, IET Signal Processing, **5**, *7*, pp. 701–707, 2011.
9. R. J. P. De Figueiredo, *A neural-network-based approach to speech signal prediction*, Rev. Roum. Sci. Techn. – Électrotechn. et Énerg., **55**, *1*, p. 42–48, 2010.
10. S. P. Lloyd, *Least squares quantization in PCM*, IEEE Trans. Inform. Theory, **28**, pp. 129–137, 1982.
11. J. Max, *Quantization for minimum distortion*, IEEE Trans. Inform. Theory, **6**, pp. 7–12, 1960.
12. P. E. Fleischer, *Sufficient conditions for achieving minimum distortion in a quantizer*, IEEE Znt. Conc. Rec., Part 1, pp. 104–111, 1964.
13. Y. Koren, I. Yavneh, Alon Spira, *A multigrid approach to the scalar quantization problem*, IEEE Trans. Inform. Theory, **51**, *8*, pp. 2993–2998, 2005.