Optimization of Quantizer's Segment Threshold Using Spline Approximations for Optimal Compressor Function

Lazar Velimirović[1*], Zoran Perić[2], Miomir Stanković[3], Jelena Nikolić[2]

[1]Mathematical Institute of the Serbian Academy of Sciences and Arts, Belgrade, Serbia
[2]Department of Telecommunications, Faculty of Electronic Engineering, University of Niš, Niš, Serbia
[3]Mathematical Department, Faculty of Occupational Safety, University of Niš, Niš, Serbia
Email: *velimirovic.lazar@gmail.com, zoran.peric@elfak.ni.ac.rs

ABSTRACT

In this paper, the optimization of quantizer's segment threshold is done. The quantizer is designed on the basis of approximative spline functions. Coefficients on which we form approximative spline functions are calculated by minimization mean square error (MSE). For coefficients determined in this way, spline functions by which optimal compressor function is approximated are obtained. For the quantizer designed on the basis of approximative spline functions, segment threshold is numerically determined depending on maximal value of the signal to quantization noise ratio (SQNR). Thus, quantizer with optimized segment threshold is achieved. It is shown that by quantizer model designed in this way and proposed in this paper, the SQNR that is very close to SQNR of nonlinear optimal companding quantizer is achieved.

Keywords: Optimization of Quantizer's Segment Threshold; Mean Square Error; Second-Degree Spline Functions; Compressor Function

1. Introduction

Quantization, in mathematics and digital signal processing, is the process of mapping a large set of input values to a smaller set—such as rounding values to some unit of precision. The most common type of quantization is known as scalar quantization. A device or algorithmic function that performs quantization is called a quantizer [1-3].

Quantizers can be uniform and nonuniform. Uniform quantizers are suitable for signals that have approximately uniform distribution. How most of the signals do not have a uniform distribution there is a need for using nonuniform quantizer [1-3]. One of the most used methods for the realization of the nonuniform quantizer is companding technique, in which a specific compressor function is applied on an input signal. The most often used compressor functions are optimal compressor function [1-3]. The optimal compressor function gives the maximum signal to quantization noise ratio (SQNR) for the reference variance of the input signal. By knowing compressor function, model quantizer is completely defined. However, although easy for analysis, it is shown that this model is very difficult to realize [1-3]. Therefore, in order to achieve easier practical realization, approxi-

mation of the optimal compressor function and linearization of the optimal compressor function is performed.

A comprehensive analysis of SQNR behavior in a wide range of variances for the piecewise uniform scalar quantizer PUSQ designed for a Laplacian source according to the piecewise linear approximation to the optimal compressor law is reported in [4]. The linearization of the optimal compressor function is done in [5]. The demerit of method described by [5] is a high complexity of quantizer and high complexity of coding and decoding. The analysis of compressor function for Laplacian source is shown in [6]. Quantizer designed on the basis of approximative spline functions, whose support region is divided on segments of equal size is described in [7].

In this paper, in order to reduce complexity and maintenance of a reasonably good performance of quantizer, we develop a new method of construction quantizer which introduces optimization of the segment threshold and different number of cells per segments. The number of cells per segments is determined depending on value of optimized segment threshold. Also, depending on value of optimized segment threshold, approximate spline functions, by which the optimal compressor function is approximated, are determined. Unlike with the quantizer described in [7], the support region of proposed quantizer model is not divided on segments of equal size. Segments'

---

*Corresponding author.

size at the proposed model depends on segment threshold that is being optimized. The value of segment threshold depends on the size of the support region. In papers [8-10] detailed analyses are performed on which the importance of support region choice could be well seen. In the papers mentioned above, the influence of support region choice on scalar quantizer performances that are designed for Laplacian source is analysed. By designing the proposed quantizer based on approximate spline functions and optimized threshold segment, SQNR that is close to that of the nonlinear optimal companding quantizer is obtained.

The rest of the paper is organized as follows: In Section 2 the detailed description of spline functions of the second-degree and optimal compressor function is given. Design of quantizer based on spline functions of the second-degree is described in Section 3. The procedure for optimization of the segment threshold is described in Section 4. Finally, Section 5 presents numerical results and discusses their implications.

2. Approximate Second-Degree Spline Functions and Optimal Compressor Function

The theory of approximation is the area of numerical mathematics that deals with problems of replacing one function with other. It is shown that the spline of degree 2 better approximates the optimal compressor function than the spline of degree 1 [7]. Therefore, in this paper, the approximation of the optimal compressor function using spline function of the second-degree is done.

A function Q is called a spline of degree 2 if [11]:

1) The domain of Q is an interval $[a, b]$.

2) Q and Q' are continuous on $[a, b]$.

3) There are points x_i (called knots) such that $a = x_0 < x_1 < \cdots < x_L = b$ and Q is a polynomial of degree at most 2 on each subinterval $[x_i, x_{i+1}]$.

In brief, a quadratic spline is a continuously differentiable piecewise quadratic function, where quadratic includes all linear combinations of the basic functions $x \to 1$, x, x^2. The second-degree spline, also called a quadratic spline, is consisted of parabola parts between two consecutive knotes, but elected to have the same tangent at knote [11]. The approximate quadratic spline function $g(x)$, which approximates a nonlinear compressed function $c(x)$, has the following form [11]:

$$g(x) = \begin{cases} r_1 + p_1 x + q_1 x^2, & x \in [0, x_1] \\ r_2 + p_2 x + q_2 x^2, & x \in [x_1, x_2] \\ \vdots \\ r_L + p_L x + q_L x^2, & x \in [x_{L-1}, x_L] \end{cases} \quad . \quad (1)$$

In this paper, the coefficients of the second-degree spline, r_i, p_i and q_i, $i = 1, \cdots, L$, are determined by

minimizing the mean-square error (MSE) as follows:

$$F(x) = \sum_{i=1}^{L} \frac{1}{x_i - x_{i-1}} \int_{x_{i-1}}^{x_i} \left(c_i(x) - g_i(x) \right)^2 dx, \quad (2)$$

$$\frac{\partial F(x)}{\partial r_i} = 0, \frac{\partial F(x)}{\partial p_i} = 0, \frac{\partial F(x)}{\partial q_i} = 0, i = 1, \cdots, L. \quad (3)$$

The optimal compressor function $c(x)$ by which the maximum SQNR is achieved for the reference variance of an input signal is defined as [1-3]:

$$c(x) = \begin{cases} x_{max} \dfrac{\displaystyle\int_0^x p^{1/3}(x)dx}{\displaystyle\int_0^{x_{max}} p^{1/3}(x)dx}, & 0 \le x \le x_{max} \\[6mm] -x_{max} \dfrac{\displaystyle\int_x^0 p^{1/3}(x)dx}{\displaystyle\int_{-x_{max}}^0 p^{1/3}(x)dx}, & -x_{max} \le x \le 0 \end{cases} \quad (4)$$

Without diminishing the generality, the quantizer design will be done for the reference input variance of $\sigma_{ref}^2 = 1$ and $0 \le x \le x_{max}$.

3. Design of Quantizer Based on Approximate Second-Degree Spline Functions

The performances of a quantizer are usually specified in terms of SQNR [1-3]:

$$SQNR = 10\log\left(\frac{\sigma^2}{D_g + D_o}\right) = 10\log\left(\frac{\sigma^2}{D}\right), \quad (5)$$

measured in decibels [dB], with σ^2 denoting the variance of an input signal x and with D denoting the total distortion. The total distortion D is equal to the sum of the granular distortion D_g and the overload distortion D_o. The overload distortion D_o is determined as follows [1-3]:

$$D_o = 2 \int_{x_{max}}^{\infty} (x - y_{max})^2 p(x)dx, \quad (6)$$

where $p(x)$ is Laplacian probability density function (PDF) which is defined as follows [1]:

$$p(x) = \frac{1}{\sqrt{2}\sigma} \exp\left(-\frac{|x|\sqrt{2}}{\sigma}\right). \quad (7)$$

In the rest of the paper we assume symmetry about zero in the proposed quantizer model design. The Laplacian PDF is symmetrical about zero. Namely, without loss of generality, we assume that information source is Laplacian source with memoryless property, the unit

variance and zero mean value. The reproduction level $y_{\max}$ is determined from the condition:

$$y_{\max} = g_L^{-1}\left(x_{\max} - \frac{\Delta}{2} \right). \qquad (8)$$

The g_L presents the approximate spline function of the last segment (see **Figure 1**). The step size is equal to:

$$\Delta = \frac{2x_{\max}}{N}. \qquad (9)$$

The optimal support region value of the proposed quantizer is as follows [5]:

$$x_{\max} = \frac{3}{\sqrt{2}} \ln\left(N + 1 \right). \qquad (10)$$

The total number of the reproduction levels per segments in the first quadrant is:

$$\sum_{i=1}^{L} N_i = \frac{N}{2}, \qquad (11)$$

where the optimal number of reproduction levels per segments, N_i, is determined from the following condition:

$$N_i = \frac{N}{2} \frac{c_i\left(x_i \right) - c_{i-1}\left(x_{i-1} \right)}{c_L\left(x_L \right)}, \quad i = 1, \cdots, L, \qquad (12)$$

Reproduction levels are determined as the solution of the approximate spline function as follows:

$$y_{i,j} = g_i^{-1}\left(\left(\frac{2j-1}{2} \right) \Delta \right), \quad i = 1, \; j = 1, \cdots, N_i \qquad (13)$$

$$y_{i,j} = g_i^{-1}\left(\sum_{k=1}^{i-1} N_k + \left(\frac{2j-1}{2} \right) \Delta \right), \qquad (14)$$
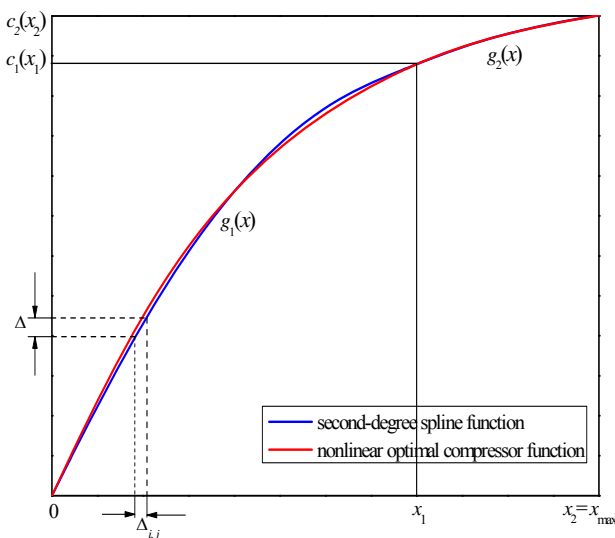

Figure 1. Second-degree spline function and nonlinear optimal compressor function for the number of segments $2L = 4$.

where $i = 2, \cdots, L, \; j = 1, \cdots, N_i$. For the $y_{i,j}$ is taken the solution that belongs to the spline function domain. Observe that indexes i and j indicate on the j-th reproduction levels within the i-th segment. Cells lengths per segments of the considered quantizer is equal to:

$$\Delta_{i,j} = \frac{\Delta}{g_i'\left(y_{i,j} \right)}, \quad i = 1, \cdots, L, \; j = 1, \cdots, N_i. \qquad (15)$$

Denoted by $\Delta_{i,j}$ the j-th cells lengths within the i-th segment.

The granular distortion is determined by Bennett's integral [1-3]:

$$D_g = \frac{x_{\max}^2}{3N^2} \int_{-x_{\max}}^{x_{\max}} \frac{p\left(x \right)}{\left[c'\left(x \right) \right]^2} \, dx. \qquad (16)$$

The granular distortion for the proposed model, based on formula (16), is equal to:

$$D_g = 2\frac{x_{\max}^2}{3N^2} \sum \frac{p\left(y_{i,j} \right)}{\left[g_i'\left(y_{i,j} \right) \right]^2} \Delta_{i,j}, \qquad (17)$$

where $i = 1, \cdots, L, \; j = 1, \cdots, N_i$.

4. Optimization of Quantizer's Segment Threshold

In this Section, the segment threshold optimization is described. Solving the system of equations defined by Equation (3), where is the optimal value of the support region, $x_L = x_{\max}$, defined by Equation (10), the expressions for determining coefficents $r_i, p_i, q_i, \; i = 1, \cdots, L$ are obtained. The values of coefficents $r_i, p_i, q_i, \; i = 1, \cdots, L$ depend on the value of the segment threshold x_i. The segment threshold x_i is determined in the following:

Step 1. Based on the coefficents $r_i, p_i, q_i, \; i = 1, \cdots, L$ expressed depending on the segment threshold x_i, the second-degree spline functions, $g_i(x)$, defined by Equation (1) is desgined.

Step 2. On the basis of the second-degree spline functions obtained in this way, $g_i(x)$, the proposed quantizer as described in Section 3 is designed. This way, an expression for SQNR of proposed quantizer depending on the segment threshold x_i is achieved.

Step 3. The segment threshold x_i is numerically determined so that maximal SQNR value of proposed quantizer is achieved (see **Figures 2** and **3**).

5. Numerical Results and Conclusions

Numerical results presented in this section are obtained for the case when the number of segments is equal to $2L = 4$ and for number of levels $N = 16$ and $N = 32$. For the number of segments $2L = 4$, the approximate quadratic spline function $g(x)$ defined by Equation (1) is equal to:
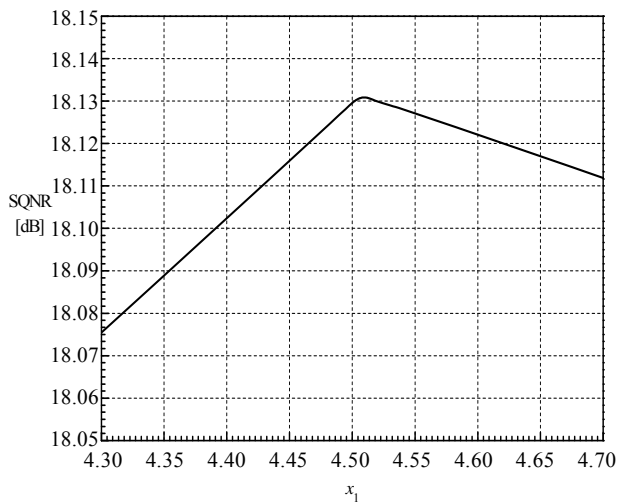
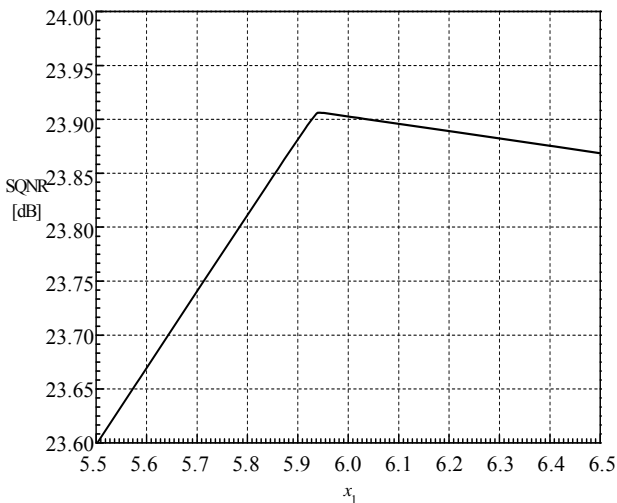**Figure 2. Numerical determination of the segment threshold x_1 for the number of levels $N = 16$.**

Figure 3. Numerical determination of the segment threshold x_1 for the number of levels $N = 32$.

$$g(x) = \begin{cases} r_1 + p_1 x + q_1 x^2, & x \in [0, x_1] \\ r_2 + p_2 x + q_2 x^2, & x \in [x_1, x_2] \end{cases}. \qquad (18)$$

The number of conditions set is equal to the number of coefficients that should be determined. In fact, as we have three knotes and two subintervals (see **Figure 1**), and each second-degree polynomial has three coefficients, means to determinate the six coefficients in total [11]. Solving the system of equations defined by Equation (3), the coefficents r_1, p_1, q_1, r_2, p_2, q_2 are determined. The values of coefficents r_1, p_1, q_1, r_2, p_2, q_2 depend on the value of the segment threshold x_1 which is numerically determined as described in Section 4.

In **Figures 2** and **3** the dependence of SQNR of the proposed quantizer model on the segment threshold x_1 for the number of levels $N = 16$ and $N = 32$ is shown. Based

on **Figure 2** it can be concluded that the maximal value of SQNR of the proposed quantizer is achieved for the value of the segment threshold $x_1 = 4.51$. Also, based on **Figure 3** it can be concluded that the maximal value of SQNR of the proposed quantizer is achieved for the value of the segment threshold $x_1 = 5.94$.

Table 1 shows the values of SQNR of the proposed quantizer which segment threshold x_1 numerically determined, (SQNR^N), the values of SQNR of the proposed quantizer having equidistant segment thresholds (the segment threshold x_1 is at the middle of support region) [7], (SQNR^S), and the values of SQNR of nonlinear optimal companding quantizer, $c(x)$:

$$[-x_{max}, x_{max}] \rightarrow [-x_{max}, x_{max}],$$

(SQNR^O), for the number of levels $N = 16$ and $N = 32$. The granular distortion D_g and the overload distortion D_o of nonlinear optimal companding quantizer defined as in [1-3]:

$$D_g^O = \frac{9}{2N^2}\left(1 - e^{-\frac{\sqrt{2}x_{max}}{3}}\right), \qquad (19)$$

$$D_o^O = \frac{e^{-\sqrt{2}x_{max}}\left(N^2 + \sqrt{2}Nx_{max} + x_{max}^2\right)}{N^2}. \qquad (20)$$

In **Figure 4** dependency of SQNR^N, SQNR^S and SQNR^O on the number bits per sample (R = 4 bit/sample and R = 5 bit/sample) is shown. Analyzing the results shown in **Table 1** and **Figure 4**, one can notice that design of quantizer based on approximate spline of degree 2, with the optimized segment threshold x_1, achieved higher SQNR than design of quantizer based on approximate spline of degree 2 with the segment threshold x_1 that is it the middle of the support region. Also, analyzing the results shown in **Table 1** and **Figure 4**, one can conclude that design of quantizer based on approximate spline of degree 2, with the optimized segment threshold x_1, achieved SQNR very close to that of nonlinear optimal companding quantizer.

Comparing the performance of the proposed quantizer model with the optimized segment threshold x_1, with the quantizer model having equidistant segment thresholds [7], and with the nonlinear optimal companding quantizer [1-3], it can be concluded that the proposed model is a very effective solution because a simple quantizer model achieves a high quality signal. For the case when the number of levels is lower than 16, the complexity of the

Table 1. The values of SQNR^N, SQNR^S and SQNR^O for the number of levels $N = 16$ and $N = 32$.

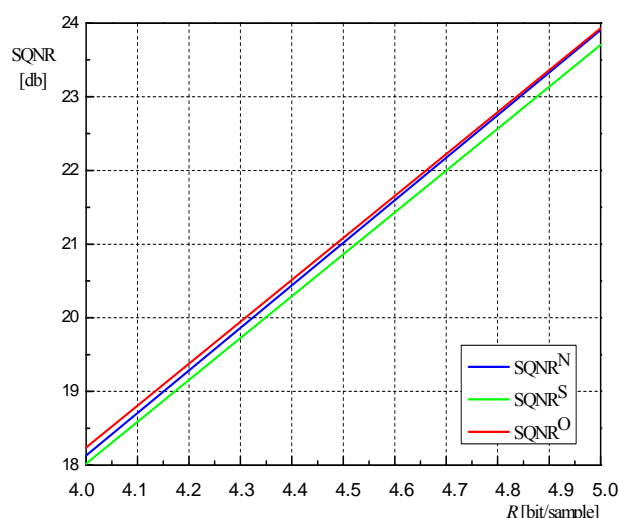| N	SQNR^N [dB]	SQNR^S [dB]	SQNR^O [dB]
16	18.1314	18.0205	18.2404
32	23.9065	23.7028	23.9303

Figure 4. Dependency of SQNR^N, SQNR^S and SQNR^O on the number bits per sample.

proposed quantizer model may be larger than the optimal Lloyd-Max's scalar quantizer model [1-3]. For the case when the number of levels is more than 32, the complexity of the proposed quantizer model becomes high. Due to these limitations, this paper analyzes the quantizer model for the number of levels $N = 16$ and $N = 32$, *i.e.* for the number bits per sample $R = 4$ and $R = 5$. The proposed quantizer model besides Laplacian PDF, is convenient for other probability density functions.

6. Acknowledgements

REFERENCES

[1]   N. S. Jayant and P. Noll, "Digital Coding of Waveforms, Principles and Applications to Speech and Video," Prentice Hall Secondary Education Division, Upper Saddle River, 1984.

[2]   A. Gersho and R. M. Gray, "Vector Quantization and Signal Compression," Kluwer Academic Publishers, Boston, Dordrecht, London, 1992. doi:10.1007/978-1-4615-3626-0

[3]   L. R. Rabiner and R. W. Schafer, "Introduction to Digital Speech Processing," Foundations and Trends in Signal Processing, Hanover, 2007.

[4]   J. Nikolić, Z. Perić, D. Antić, A. Jovanović and D. Denić, "Low Complex Forward Adaptive Loss Compression Algorithm and ITS Aplication in Speech Coding," *Journal of Electrical Engineering*, Vol. 62, No. 1, 2011, pp. 19-24. doi:10.2478/v10187-011-0003-5

[5]   Z. Perić, M. Petković and M. Dinčić, "Simple Compression Algorithm for Memoryless Laplacian Source Based on the Optimal Companding Technique," *Informatica*, Vol. 20, No.1, 2009, pp. 99-114.

[6]   Z. Perić and J. Nikolić, "Analysis of Compressor Function for Laplacian Source's Scalar Compandor Construction," *Data Recording*, *Storage and Processing*, Vol. 8, No. 2, 2006, pp. 15-24.

[7]   L. Velimirović, Z. Perić, J. Nikolić and M. Stanković, "Design of Compandor Quantizer for Laplacian Source for Medium Bit Rate Using Spline Approximations," *Facta Universitatis*, Vol. 25, No. 1, 2012, pp. 90-102.

[8]   S. Na, "On the Support of Fixed-Rate Minimum Mean-Squared Error Scalar Quantizers for a Laplacian Source," *IEEE Transactions on Information Theory*, Vol. 50, No. 5, 2004, pp. 937-944. doi:10.1109/TIT.2004.826686

[9]   Z. Perić, J. Nikolić and D. Pokrajac, "Estimation of the Support Region for Laplacian Source Scalar Quantizers," *Journal of Electrical Engineering*, Vol. 58, No. 1, 2007, pp. 47-51.

[10] Z. Perić, J. Nikolić and D. Pokrajac, "Analysis of Support Region for Laplacian Source's Scalar Quantizers," *Proceedings of 7th IEEE Conference on Telecommunications in Modern Satelite, Cable and Broadcasting Services TELSIKS* 2005, Niš, 28-30 September 2005, Vol. 2, pp. 491-494.

[11] W. Cheney and D. Kincaid, "Numerical Mathematics and Computing," 6th Edition, Thomson Higher Education, Belmont, 2008.