



NEW METHOD FOR CONSTRUCTION OF OPTIMAL SCALAR QUANTIZERS FOR LAPLACIAN SOURCE

Zoran Peric ¹⁾, Jelena Nikolic ¹⁾, Dragoljub Pokrajac ²⁾

¹⁾ Faculty of Electronic Engineering, University of Nis, 18000 Nis, Aleksandra Medvedeva 14, Serbia,
peric@elfak.ni.ac.yu, njelena@elfak.ni.ac.yu

²⁾ Computer and Information Sciences Department and Applied Mathematics Research Center, Delaware State University, 1200 N DuPont Hwy, Dover, Delaware 19901, USA, dpokraja@desu.edu

Abstract: In this paper we consider methods for computing the necessary parameters when constructing the optimal scalar quantizers for Laplacian source. We investigate two approaches to the problem of finding the sets of optimal parameters. The first approach requires solving the transcendental equations, but provides nearly optimal values of the scalar quantizers' parameters on successive manner. The proposed approach is an approximation method that linearizes transcendental equations providing simple and fast computing of scalar quantizers' parameters. We demonstrate that the proposed technique provides parameters values that are very close to the optimal ones.

Keywords: Scalar quantization, Laplacian source, quantizer parameters computing.

1. INTRODUCTION

Quantizers play an important role in theory and practice of contemporary signal processing. A vast amount of research has been performed in the area of quantization. One of the most important issues from the engineers' point of view is the design and implementation of quantizers to meet performance goals. Considerable work has been focused on the design of optimal quantizers for source coding of images, speech, etc. [1,2]. Lloyd [3] and Max [4] independently proposed an algorithm to compute optimal quantizers using mean-square error distortion measure. They provided the nonlinear quantization procedure in order to minimize the quantization noise. The primary goal when designing an optimal Lloyd-Max's quantizer is to select the representation levels and the decision thresholds in order to provide the minimum possible average distortion for a fixed number of quantization levels N . Particularly, Lloyd-Max's algorithm is an iterative algorithm, which in each iteration performs calculation of all representation levels and decision thresholds of the N levels scalar quantizer. The amount of necessary calculations and the number of iterations are the deficiencies of this algorithm, and they typically increase with the number of quantization levels N .

Consequently, this paper considers methods for finding the necessary parameters for construction of optimal scalar quantizers for Laplacian source.

Namely, in order to determine the scalar quantizers' parameters for Laplacian source it is necessary to solve the transcendental equations. This problem was considered in [5] and was resolved by introducing the Lambert W function and some approximations. We suggest one very fast and simple approximation that comprises new efficient method for solving transcendental equations. By linearizing the transcendental equations, the proposed method provides nearly optimal parameters' values, [1,2] that are necessary for designing scalar quantizers.

2. SCALAR QUANTIZATION

Scalar quantizers are most commonly defined in terms of their decision thresholds $\{t_0, t_1, \dots, t_N\}$ and real-valued quantization points $\{y_1, y_2, \dots, y_N\}$. We define the N -level quantizer in terms of its cells widths $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ and the distances from the representative levels to the nether decision thresholds, i.e., reconstruction offsets $\{\delta_1, \delta_2, \dots, \delta_N\}$. The negative thresholds and the quantization points are symmetric to their nonnegative counterparts. Symbolically, the decision thresholds are:

$$t_0 = -\infty < t_1 < \dots < t_{N-1} < t_N = \infty \quad (1)$$

and the cells widths are given by:

$$\alpha_j = (t_{j-1}, t_j] \quad j = 1, 2, \dots, N. \quad (2)$$

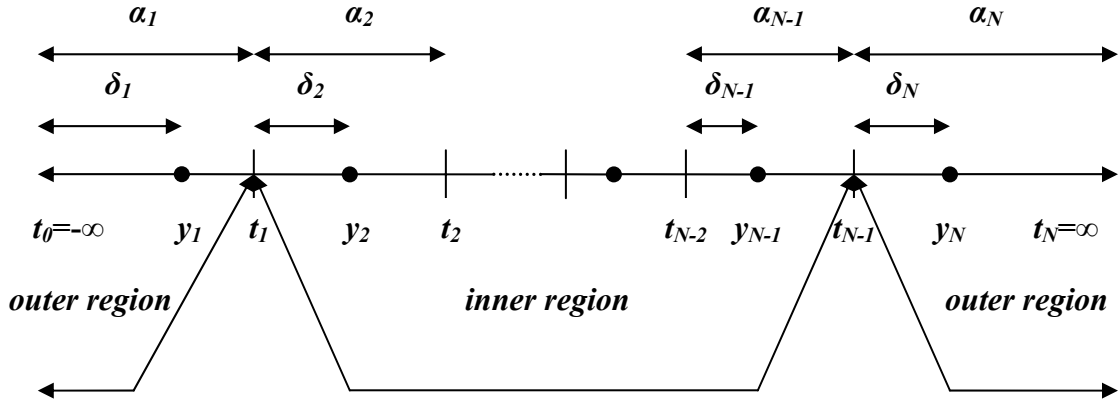


Fig. 1 - Illustration of the inner region and the outer region of the scalar quantizer

The cells $\alpha_2, \dots, \alpha_{N-1}$ are referred to as the inner cells, while α_1 and α_N are referred to as the outer cells. The outer cells are unbounded ($t_0 = -\infty$ and $t_N = \infty$) and are also called overload cells, while the inner cells are bounded and are called granular cells. The overload cells (α_1 and α_N) comprise the overload region, (outer region), while the set of all granular cells ($\alpha_2, \dots, \alpha_{N-1}$), is called granular region, (inner region), as depicted in Fig.1. A quantized signal has value y_j when the original signal belongs to the quantization cell α_j . Therefore, N -level scalar quantizer, Q , can be defined as a functional mapping of an input value x onto an output representation, such that:

$$Q(x) = y_j, x \in \alpha_j. \quad (3)$$

Formally, the reconstruction offsets are defined as:

$$\delta_j = y_j - t_{j-1}. \quad (4)$$

The quality of a quantizer can be measured by the distortion of the output representation y_j when compared to the original signal x . The primary goal of quantizer design is to select the reconstruction offsets and cells widths so as to provide the minimum possible average distortion for a fixed number of levels N . The quantizer distortion can be defined as [5]:

$$D = \sum_{j=1}^N \int_{t_{j-1}}^{t_j} d(x - y_j) p(x) dx, \quad (5)$$

where $p(x)$ is probability density function of continuous random variable (corresponding to the source signal) and $d(x - y_j)$ is the measure of distortion between an input signal x and quantized signal y_j . The most convenient and widely used measure of

distortion is the average mean-squared error (d_{mse}), i.e., quantization noise:

$$d_{mse}(\Delta) = |\Delta|^2, \quad (6)$$

where $\Delta = x - y_j$.

The discrete probability of each quantization point y_j is given by [5]:

$$p_j = \int_{t_{j-1}}^{t_j} p(x) dx. \quad (7)$$

These probabilities determine the output entropy of the quantizer as follows:

$$H = - \sum_{j=1}^N p_j \log_2 p_j. \quad (8)$$

A quantizer is optimal when no other N -level scalar quantizer can provide lower distortion with equal or lower output entropy. Quantizer optimization can be achieved by using unconditional optimization of extended function, known as objective function J , given by [5]:

$$J = D + \lambda H, \quad (9)$$

where λ is the Lagrange multiplier.

When the distortion measure defined in Eq. (6) is used, optimal values of the reconstruction offsets are given by [5]:

$$\delta_j = \frac{\int_{t_{j-1}}^{\alpha_j} x p(x + t_{j-1}) dx}{\int_0^{\alpha_j} p(x + t_{j-1}) dx}. \quad (10)$$

The second necessary condition for optimality [5] considers the optimality of the objective function Eq. (9). According to this condition, the decision thresholds t_j that define the cells widths α_{j-1} and α_j can be achieved when the two nearest reconstruction levels have the same rate-distortion penalties:

$$d(\alpha_{j-1} - \delta_{j-1}) - \lambda \log_2(p_{j-1}) = d(-\delta_j) - \lambda \log_2(p_j). \quad (11)$$

In this paper we consider case $\lambda = 0$, thereby Eq. (11) becomes:

$$d(\alpha_{j-1} - \delta_{j-1}) = d(-\delta_j). \quad (12)$$

If we use mean-square error distortion measure, Eq. (6), then the condition for optimality becomes:

$$(\alpha_{j-1} - \delta_{j-1})^2 = (-\delta_j)^2. \quad (13)$$

3. NEW METHOD

In this paper we combine two conditions for optimality. Namely, for particular probability density function (pdf) we derive the expression for the reconstruction offsets and substitute it in the second condition for optimality, Eq. (11). The result is an equation that enables numerical computation of the cells widths which are necessary for construction of scalar quantizers.

In this section the aforementioned idea is demonstrated in case of the Laplacian source. Particularly, we consider the Laplacian source with memoryless property. In such a case, assuming unit variance, the probability density function of continuous random variable can be expressed by the Laplacian distribution defined as [5]:

$$p(x) = \frac{\sqrt{2}}{2} e^{-\sqrt{2}|x|}. \quad (14)$$

Without loss of generality, due to symmetry of scalar quantizer's parameters that are necessary for construction of scalar quantizers, we may observe a one-sided pdf, where $x \geq 0$. Using Eq. (14), the memoryless property of the Laplacian pdf allows the substitution:

$$p(x + t_{j-1}) = e^{-t_{j-1}\sqrt{2}} p(x). \quad (15)$$

By substituting Eq. (15) into Eq. (10), the expression for determining optimal values of the reconstruction offsets δ_j as a function of the cells width α_j is derived as:

$$\delta_j(\alpha_j) = \frac{\sqrt{2}}{2} \left(1 - \frac{\sqrt{2}\alpha_j e^{-\alpha_j\sqrt{2}}}{1 - e^{-\alpha_j\sqrt{2}}} \right). \quad (16)$$

Substitution of Eq. (16) into Eq. (13) through simple mathematical transformations yields the following transcendental equation:

$$\alpha_{j-1} + e^{-\alpha_{j-1}\sqrt{2}} \left[\frac{\sqrt{2}}{2} + \sqrt{\left(\frac{\sqrt{2}}{2} \left(\frac{\sqrt{2}\alpha_j e^{-\alpha_j\sqrt{2}}}{1 - e^{-\alpha_j\sqrt{2}}} - 1 \right) \right)^2} \right] - \left[\frac{\sqrt{2}}{2} + \sqrt{\left(\frac{\sqrt{2}}{2} \left(\frac{\sqrt{2}\alpha_j e^{-\alpha_j\sqrt{2}}}{1 - e^{-\alpha_j\sqrt{2}}} - 1 \right) \right)^2} \right] = 0. \quad (17)$$

The problem of solving transcendental equations was considered in [5] and was resolved by introducing the Lambert W function and some approximations. In this paper we suggest more efficient solution to this problem.

In order to simplify the Eq. (17) we can introduce coefficient c_j that depends on the cell width α_j , such that:

$$c_j = \left[\frac{\sqrt{2}}{2} + \sqrt{\left(\frac{\sqrt{2}}{2} \left(\frac{\sqrt{2}\alpha_j e^{-\alpha_j\sqrt{2}}}{1 - e^{-\alpha_j\sqrt{2}}} - 1 \right) \right)^2} \right]. \quad (18)$$

Now, the Eq. (17) can be given by:

$$\alpha_{j-1} + c_j \cdot e^{-\alpha_{j-1}\sqrt{2}} - c_j = 0. \quad (19)$$

From this equation it is possible to resolve α_{j-1} , on successive manner, using α_j obtained in the previous computation step. In order to find nearly optimal solutions of the transcendental equations, we suggest an approximation by the first term of Taylor expression as follows:

$$e^{-\alpha_{j-1}\sqrt{2}} \approx e^{-\alpha_{j-1}^{(i)}\sqrt{2}} - \sqrt{2}(\alpha_{j-1} - \alpha_{j-1}^{(i)})e^{-\sqrt{2}\alpha_{j-1}^{(i)}}. \quad (20)$$

Combining this approximation with Eq. (19), it is we can compute the cell widths using the following iterative equation:

$$\alpha_{j-1}^{(i+1)} = \alpha_j \cdot \frac{1 - e^{-\sqrt{2}\alpha_{j-1}^{(i)}} (1 + \sqrt{2}\alpha_{j-1}^{(i)})}{1 - \alpha_j \sqrt{2} e^{-\sqrt{2}\alpha_{j-1}^{(i)}}}, \quad (21)$$

where i is counter of iterations.
 Let us summarize the proposed new method:

Step 1. Initialization: $\alpha_{N=\infty} \delta_N = \sqrt{2} / 2$ $\alpha_{N-1}^{(0)} = \sqrt{2}$

Step 2. Computation of $\alpha_{j-1}^{(i+1)}$ for $i=0,1$ and $j=N, \dots, 2$ from Eq. (21).

Step 3. Stopping criterion check:

$$|\alpha_{j-1}^{(i+1)} - \alpha_{j-1}^{(i)}| \leq \varepsilon \quad (22)$$

where $\varepsilon = 5 \cdot 10^{-4}$ [9].

Step 4. If the algorithm break is satisfied then $\alpha_j^{new} = \alpha_j^{(i+1)}$ and $\delta_j^{new} = \delta_j(\alpha_j)$, Eq. (16), for $j=N-1, \dots, 1$, otherwise go to Step 2.

4. NUMERICAL RESULTS

Considering the facts that $\alpha_{N=\infty}$ and analyzing an Eq. (16) we can make two following assumptions $\delta(\alpha_{N-1}) < \sqrt{2} / 2$ and $\alpha_{N-1} - \delta(\alpha_{N-1})$ is slightly greater than $\delta(\alpha_{N-1})$. Therefore, we can assume that the initial value of the cell width α_{N-1} is $\alpha_{N-1}^{(0)} \approx \sqrt{2}$. Our experiments demonstrate that such choice of initial values lead to fast convergence. It is very important to point out that the quantization points y_j are *not* obtained as an arithmetical mean of the decision thresholds that determine the quantization cells, as in the case of Lloyd-Max quantizers [3,4].

Let us describe the first computation step in more details. Considering the facts that $\alpha_{N=\infty}$, $\delta(\alpha_N) = \sqrt{2} / 2$ and $\alpha_{N-1}^{(0)} \approx \sqrt{2}$, and using Eq. (21), it is easy to calculate α_{N-1} and this value is marked here as $\alpha_{N-1}^{(1)}$. Substituting $\alpha_{N-1}^{(0)}$ with $\alpha_{N-1}^{(1)}$ in Eq. (21) the new value for α_{N-1} can be obtained and that value is the final value of α_{N-1}^{new} . The procedure of determining subsequent values of the step sizes α_j^{new} , $j=N-2, \dots, 1$ identically repeats as shown for α_{N-1}^{new} . It is very important to point out that for calculation of the cell width, α_{j-1}^{new} , we use, as an initial value, the value of the previously calculated cell width α_j^{new} .

Table 1. for $N=8$ compares the values of the cell widths α_j (due to the symmetry shown only for $j=N/2+1, \dots, N$) calculated from the transcendental equations (here referred to as Method I) and widths α_j^{new} obtained by using the proposed new method (the Method II), with the optimal values α_j^{opt} [1,2]. From the Table 1., it is obvious that the values of the cells width obtained by using both methods are nearly optimal. Table 2. compares the values of the reconstruction offsets that are calculated by using both methods δ_j and δ_j^{new} , for $j=N/2+1, \dots, N$ ($N=8$) with the optimal values δ_j^{opt} , [1,2]. Considering the Table 2., one can notice that appropriate values of the reconstruction offsets are almost identical.

Table 1. The cells widths, for scalar quantizers with $N=8$ levels

Method I	Method II	Optimal values [1,2]
$\alpha_8 = \infty$	$\alpha_8^{new} = \infty$	$\alpha_8^{opt} = \infty$
$\alpha_7 = 1.1269$	$\alpha_7^{new} = 1.1271$	$\alpha_7^{opt} = 1.127$
$\alpha_6 = 0.7196$	$\alpha_6^{new} = 0.7197$	$\alpha_6^{opt} = 0.720$
$\alpha_5 = 0.5332$	$\alpha_5^{new} = 0.5332$	$\alpha_5^{opt} = 0.533$

Table 2. The values of the reconstruction offsets, for scalar quantizers with $N=8$ levels

Method I	Method II	Optimal values [1,2]
$\delta_8 = 0.7071$	$\delta_8^{new} = 0.7071$	$\delta_8^{opt} = 0.707$
$\delta_7 = 0.4197$	$\delta_7^{new} = 0.4198$	$\delta_7^{opt} = 0.420$
$\delta_6 = 0.2998$	$\delta_6^{new} = 0.2998$	$\delta_6^{opt} = 0.300$
$\delta_5 = 0.2334$	$\delta_5^{new} = 0.2334$	$\delta_5^{opt} = 0.233$

5. CONCLUSION

The approach presented in this paper makes possible obviating very important problem of determining maximal amplitude of the input signal i.e., problem of determining granular region [6,7]. The proposed method makes possible simpler construction of optimal scalar quantizers demanding less memory space to store parameter values. Namely, the exact solutions of the transcendental equations can be successively obtained (as in Method I discussed here). However, using the linearization of the transcendental equation we could, as in the proposed novel method (Method II), get the solutions very close to the exact solutions but much faster. Analysis presented in this paper has the practical importance since it could be of great help to engineers. Particularly, it provides fast and efficient design of scalar quantizers that are used for source coding of images [8] and speech [2]. In comparison to the Lloyd-Max's algorithm that requires the existence of $2N+1$ initial values (where N is the number of quantization levels) the proposed approach requires only one initial value.

6. ACKNOWLEDGEMENTS

D. Pokrajac has been partially supported by NIH-funded Delaware IDeA Network of Biomedical Research Excellence (INBRE) Grant, DoD HBCU/MI Infrastructure Support Program (45395-MA-ISP Department of Army), National Science Foundation (NSF) Infrastructure Grant (award # 0320991) and NSF grant "Seeds of Success: A Comprehensive Program for the Retention, Quality Training, and Advancement of STEM Student" (award #HRD-0310163). We would like to thank the

anonymous reviewers for suggestions that significantly improved our manuscript.

7. REFERENCES

- [1] N.S. Jayant, Peter Noll, *Digital coding of waveforms*, Prentice-Hall, New Jersey, 1984., Chapter 4, pp. 129-139.
- [2] Wai C. Chu, *Speech coding algorithms*, John Wiley & Sons, New Jersey, 2003., Chapter 5, pp. 151-155.
- [3] S. P. Lloyd, Least squares quantization in PCM, unpublished memo., Bell Lab., 1957; *IEEE Transactions on Information Theory*, vol. IT-28, Mar. 1982, pp. 129-137.
- [4] J. Max, Quantizing for minimum distortion, IRE, *Transactions on Information Theory*, vol. IT-6, Mar. 1960., pp. 7-12.
- [5] Gary J. Sullivan, Efficient Scalar Quantization of Exponential and Laplacian Random Variables, *IEEE Transactions on Information Theory*, vol. 42, No. 5, September 1996., pp. 1365-1374.
- [6] Sangsin 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, May 2004., pp. 937-944.
- [7] Sangsin Na and David L. Neuhoff, On the Support of MSE-Optimal, Fixed-Rate, Scalar Quantizers, *IEEE Transactions on Information Theory*, Vol. 47, No. 7, November 2001., pp. 2972-2982.
- [8] Allen Gersho and Robert M. Gray, *Vector Quantization and Signal Compression*, Kluwer, Academ. Pub, 1992.
- [9] Robert Gray, Quantization and data compression, Lecture notes, Stanford University, 2004.



Zoran H. Peric was born in Nis, Yugoslavia, in 1964. He received the B. Sc. degree in electronics and telecommunications from the Faculty of Electronic Engineering, Nis, Serbia, Yugoslavia, in 1989, and M.Sc. degree in telecommunications from the University of Nis, in 1994. He received

the Ph. D. degree from the University of Nis, also, in 1999.

He is currently Professor at the Department of Telecommunications and vicedean of the Faculty of

Electronic Engineering, University of Nis, Serbia. His current research interests include the information theory, source and channel coding and signal processing. He is particularly working on scalar and vector quantization techniques in speech and image coding. He is author and coauthor in over 100 papers in digital communications. Dr Peric has been a Reviewer for *IEEE Transactions on Information Theory*. He is member Editorial Board of Journal "Electronics and Electrical Engineering"



Jelena R. Nikolic was born in Prokuplje, Yugoslavia, in 1978. She received the B. Sc. degree in telecommunications from the Faculty of Electronic Engineering, Nis, Serbia in 2003. She is currently Research Assistant on a Science Research Program, supported by Serbian government, on Faculty of

Electronic Engineering in the laboratory for telecommunications under supervision of prof. Zoran Peric, since 2003.

Her current research interests include the information theory, source and channel coding and signal processing. She is particularly working on scalar quantization techniques in speech and image coding. She is author and coauthor in over 15 papers in digital communications.

Dragoljub M. Pokrajac was born in Sibenik, Dalmatia, Croatia, in 1970. He received the B. Sc. degree in electronics and telecommunications from the Faculty of Electronic Engineering, Nis, Serbia, Yugoslavia, in 1993, M. Sc. degree in telecommunications from the University of Nis, 1997. He received the Ph.D. degree at Temple University, USA, 2002.



He is currently Assistant Professor at the Computer and Information Sciences Department and Graduate Department of Applied Mathematics and Theoretical Physics Delaware State University, Dover, DE, USA.

He is co-principal investigator of DoD/DoA Applied Mathematics Research Center Grant, and co-principal investigator of NIH-funded INBRE grant. His current research interests include video-surveillance, data bases and data mining, applied mathematics, spatial-temporal statistics, computational geometry and image processing. He is author and coauthor in over 80 papers and one patent. Dr Pokrajac has been a Reviewer for *IEEE Transactions on Education*.