ISSN 1392 – 124X INFORMATION TECHNOLOGY AND CONTROL, 2006, Vol. 35, No. 4

ENCODER DESIGN FOR SWITCHED PIECEWISE UNIFORM VECTOR QUANTIZATION OF THE MEMORYLESS TWO-DIMENSIONAL LAPLACIAN SOURCE

Zoran H. Peric, Vladimir M. Despotovic

Faculty of Electronic Engineerin
University of Nis, Serbia

Abstract. In this paper switched piecewise uniform vector quantization of two-dimensional memoryless Laplacian source is asymptotically analyzed for the case where the power of input signal varies in a wide range. One possible solution for encoder design is given for the same quantizer. The presented encoder is a compromise between memory space and number of logical circuits. Uniform quantizer optimality conditions and all main equations for the optimal number of levels and constant (nonoptimal) number of output points for each partition are presented (using rectangular cells). Switched quantization is used in order to give a higher quality in a wide range of signal volumes (variances). These systems, although not optimal, may have asymptotic performance arbitrarily close to the optimum. Furthermore, their analysis and implementation can be simpler than those of optimal systems.

Keywords: Piecewise Uniform Vector Quantization, Switched Quantization, Encoder design.

1. Introduction

The quantizers play an important role in the theory and practice of modern signal processing. The most basic of the approaches to source coding is a uniform scalar quantization [1, 2], which is commonly used for analog-to-digital conversion. Extensive results have been obtained on scalar quantization but more on vector quantization. The simplest vector quantization is two-dimensional vector quantization.

During the two dimensional vector quantization, vector obtained by sampling of input signal in two adjacent points is replaced with vector from an allowed set of vectors in such way that the quantization error is the smallest. Successful vector quantization depends on an appropriate choice of allowed vector set (codebook). Quantization is a necessary step in the digitalization process, but there are difficulties that cause the quantization error, which is unavoidable during this process. Hence, the quantization should be performed in such way that the quantization error does not reflect on signal reconstruction. This means that the allowed vector set should be chosen in a manner that the mean squared error (distortion) would be minimal.

The analysis of vector quantizer for arbitrary distribution of the source signal is given in paper [3]. The authors derived the expression for the optimum granular distortion and optimum number of output points. However, they did not prove the optimality of the proposed solutions. Also, they did not define the partition of the multidimensional space into subregions. In paper [4], the expressions for the optimum number of output points are derived, however the proposed partitioning of the multidimensional space for memoryless Laplacian source does not consider the geometry of the multidimensional source. In paper [5], vector quantizers of Laplacian and Gaussian sources are analyzed. The proposed solution for the quantization of memoryless Laplacian source, unlike in [4], takes into consideration the geometry of the source, however, the proposed vector quantizer design procedure is too complicated and unpractical.

The goal of this paper is to solve a quantization problem in a case of switched piecewise uniform vector quantizer (PUQ) for Laplacian memoryless source, for a case of constant number of output points on each domain. Indeed, we want to suggest the quantizer that, compared to the one with optimal number of output points [10], may have asymptotic performance arbitrarily close to the optimum. We will give a general and simple way to design a switched piecewise uniform vector quantizer. We will derive the optimal number of levels for each partition and the optimality of the proposed solution is proved. At the end, we will propose one solution for encoder design for this quantizer.

A block diagram of the encoder and decoder is shown in Figure 1. The encoder performs a mapping from the k-dimensional space R^k into the index set I, and the decoder maps the index set I into the finite subset C, which is the codebook. The codebook has a positive integer number of code vectors which define the codebook size *N*. The bit rate *R* depends on *N* and the vector dimension *k*. Since the bit rate is the number of bits per sample,

$$R = (\log_2 N)/k \,. \tag{1}$$

The decoding process is very simple and requires only a table (codebook) lookup, but the encoding procedure is complex and involves finding a best matching code vector, using a distortion measure as a criterion.

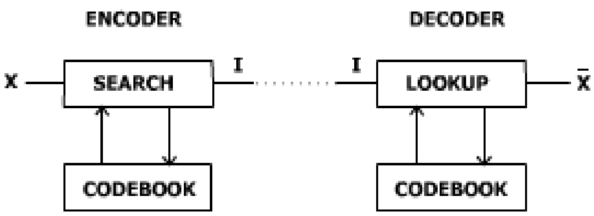

Figure 1. Block diagram of encoder/decoder

Let $\boldsymbol{X}$ be a two-dimensional random vector with joint density $f(\mathrm{x}) = f(x_1, x_2)$. *N* points vector quantizer is a function $Q(\boldsymbol{x})$ which maps $\boldsymbol{x}$ in R^2 into one of *N* output vectors. The quantizer is specified by the values of the output points and by a partition of the space R^2 into *N* cells $S_1, S_2, ..., S_N$, where $S_i = Q^{-1}(\mathrm{y}_i) \subset R^2$. Then we can write $Q(\mathrm{x}) = \mathrm{y}_i$, *if* $\mathrm{x} \in S_i$ *for* $i = 1, 2, ..., N$. A cell that is unbounded is called an overload cell. Each bounded cell is called a granular cell. Together all of overload (granular) cells are called the overload (granular) region.

The quality of a quantizer can be measured by the distortion of the resulting reproduction in comparison to the original. The total distortion *D* is a combination of the granular (D_g) and overload (D_o) distortions, $D = D_g + D_o$. The most convenient and widely used measure of distortion between an input vector $x = (x_1, x_2)$ and quantized vector y_i (y_{i1}, y_{i2}) is the average mean-squared error per dimension, i.e. quantization noise [7]

$$D = \frac{1}{2}\sum_{i=1}^{N}\int_{S_i} \left|\mathrm{x} - \mathrm{y}_i\right|^2 f(\mathrm{x})d\mathrm{x} \tag{2}$$

The mean-squared error (MSE) is used as the criterion for optimization.

The MSE of a two-dimensional vector source $x = (x_1, x_2)$, where x_i are zero-mean statistically independent Laplacian random variables of variance σ^2, is commonly used for the transform coefficients of speech or imagery. The first approximation to the long-time-averaged probability density function (pdf) of amplitudes is provided by the Laplacian model [6, p.32]. The waveforms are sometimes represented in terms of adjacent-sample differences. The pdf of the difference signal for an image waveform follows the Laplacian function [6, p.33]. The Laplace source is a model for speech [7, p.384]. We will consider two independent identically distributed Laplace random variables (x_1, x_2) with the zero mean. To simplify the vector quantizer, the Helmert transformation is applied to the source vector giving contours with constant probability densities. The transformation is defined as:

$$r = \frac{1}{\sqrt{2}}\left(|x_1| + |x_2|\right), \; u = \frac{1}{\sqrt{2}}\left(|x_1| - |x_2|\right). \tag{3}$$

In this paper, quantizers are designed and analysed under an additional constraint — each scalar quantizer is a uniform one.

PUQ consists of *L* optimal uniform vector quantizers. More precisely, our quantizer divides the input plane into *L* partitions and every partition is further subdivided into L_i ($1 \le i \le L$) subpartitions. Every concentric subpartition can be subdivided in four equivalent regions, i.e. the *j*-th subpartition in signal plane is allowed to have p_{ij} ($1 \le i \le L,\; 1 \le j \le L_i$) cells. We perform distortion optimization (D_i) in every partition under the constraint:

$$4\sum_{j=1}^{L_i} p_{ij} = N_i \,. \tag{4}$$

In this work, we design a piecewise uniform vector quantizer for optimal compression function. We perform analytical optimisation of the granular distortion and numerical optimization of the total distortion using rectangular cells.

2. Switched vector quantization model

During the two dimensional vector quantization, the vector obtained by sampling of the input signal in two adjacent points is replaced with vector from allowed set of vectors in such way that the quantization error is the smallest. Successful vector quantization depends on an appropriate choice of allowed vector set (codebook).

The nearest neighbor quantizer completely searches codebook. If the codebook is of size *N* (N=2^{16} for *R*=8 bit/sample, N=2^{15} for *R*=7.5 bit/sample), then *N* distortion estimations would be needed.

The switching quantization aims are to improve the quality of the signal-to-noise ratio in a wide range of the signal average power (i.e. variance) or to decrease the sample rate. The switching quantization is adaptive quantization for memoryless sources and it is aplicable only if adaptation is performed on the basis of the signal average power, what was done in this paper. As an input source, we will consider memoryless Laplacian source.

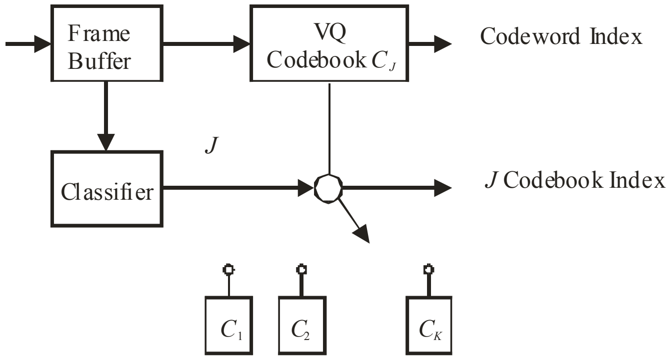

Figure 2. Switched codebook adaptive vector quantization

The basic scheme of switched codebook adaptation is shown in Figure 2. One simple technique is switched codebook adaptive vector quantization. This technique uses a classifier that looks at the contents of the input frame buffer and decides that the next block of vectors belongs to a particular statistical class of vectors from a finite set of K possible classes. Namely, the index specifying the class is used to select a particular codebook from a predesigned set of K codebooks. This index is also transmitted as side information to the receiver. Then each vector in the block is encoded by the vector quantizer which performs a search through the selected codebook.

One block is made of M vectors. The index to identify the class is sent at the end of the block, while the index to identify the codebook is sent with each vector. If each of the K codebooks has N code vectors, then the bit rate per sample is:

$$R = \frac{1}{n}\log_2 N + \frac{1}{n}\frac{\log_2 K}{M}, \tag{5}$$

where n is quantizer dimension. The second term in equation (5) is due to the side information [7].

We will use this technique for solving our problem. In our case, the quantizer dimension is $n = 2$. We have K codebooks, i. e. K piecewise uniform quantizers designed for particular value σ_{0j} and for the cover of particular input power range $\sigma^2 \in [\sigma_{1j}^2, \sigma_{2j}^2]$, where $\sigma_{1j} < \sigma_{0j} < \sigma_{2j}$. Also all codebooks are of the same size N. For selected value σ_{0j}, $j = 1, 2, ..., K$, we will design piecewise uniform vector quantizers. We should determine distortion for input Laplacian source whose power is $\sigma^2 \in [\sigma_{1j}^2, \sigma_{2j}^2]$, while the quantizer is designed in such way that it has the smallest distortion for the power σ_{0j}^2.

3. Piecewise uniform two-dimensional quantization

The initial expression for granular distortion is [8-10]:

$$D_g = 4\sum_{i=1}^{L}\sum_{j=1}^{L_i}\sum_{k=1}^{p_{i,j}} \int_{r_{i,j}}^{r_{i,j+1}} \int_{u_{i,j,k}}^{u_{i,j,k+1}} \left[\left(r - m_{i,j}\right)^2 + \left(u - \hat{u}_{i,j,k}\right)^2\right] \cdot \frac{1}{2\sigma^2} e^{-\frac{2r}{\sigma}} dr du . \tag{6}$$

The output point coordinates are given by the equations

$$m_{i,j} = \frac{r_{i,j+1} + r_{i,j}}{2}, \quad \hat{u}_{i,j,k} = \frac{u_{i,j,k} + u_{i,j,k+1}}{2}. \tag{7}$$

Rectangular cell dimensions are:

$$\Delta = r_{i+1} - r_i = r_{\max}/L, \quad \Delta_i = \frac{\Delta}{L_i},$$

$$\Delta_{ij} = \frac{r_{i,j} + r_{i,j+1}}{p_{i,j}} = \frac{2m_{i,j}}{p_{i,j}}, \quad i = 0,\ldots,L,$$

$$j = 0,\ldots,L_i . \tag{8}$$

The range of the quantizer is $r_{\max}$. The total number of output points is

$$N = \sum_{i=1}^{L} N_i , \tag{9}$$

where N_i is the number of output points in the i-th domain. We can also write:

$$\sum_{j=1}^{L_i} p_{i,j} = \frac{N_i}{4} . \tag{10}$$

We will now assume that the number of output points in the i-th domain N_i is constant under the constraint $N = \sum_{i=1}^{L} N_i = 2^{2R}$, where R is bit rate (number of bits per sample). If we assume that the bit rate is $R = R_1 + R_2$, where $2R_1$ bits are used for representing the number of partitions L, we can find R_1 from $L = 2^{2R_1}$. The number of output points in the ith domain N_i is:

$$N_i = 2^{2R_2} = 2^{2(R-R_1)} = 2^{2R - \frac{\log L}{\log 2}} = const. \tag{11}$$

Equation (6) can be written as:

$$D_g = \sum_{i=1}^{L} D_g(i) . \tag{12}$$

After integration over u and reordering, $D_g(i)$ becomes

$$D_g(i)=\sum_{j=1}^{L_i}[\int_{r_{i,j}}^{r_{i,j+1}}(r-m_{i,j})^2 4m_{i,j}\frac{1}{\sigma^2}e^{-\frac{2r}{\sigma}}dr+ \int_{r_{i,j}}^{r_{i,j+1}}\frac{m_{i,j}{}^3}{3p_{i,j}{}^2}4m_{i,j}\frac{1}{\sigma^2}e^{-\frac{2r}{\sigma}}dr]. \tag{13}$$

We will now assume that $\exp(-2r/\sigma)$ is constant over Δ_i. In that case, we can substitute $\exp(-2r/\sigma)$ with $\exp(-2m_{i,j}/\sigma)$. Equation (13) can be now written as:

$$D_g(i)=\sum_{j=1}^{L_i}4m_{i,j}\frac{1}{\sigma^2}e^{-\frac{2m_{i,j}}{\sigma}} \left[\int_{r_{i,j}}^{r_{i,j+1}}(r-m_{i,j})^2dr+\int_{r_{i,j}}^{r_{i,j+1}}\frac{m_{i,j}^2}{3p_{i,j}^2}dr\right]= \sum_{j=1}^{L_i}\left[2m_{i,j}\frac{\Delta_i^2}{12}+\frac{2m_{i,j}^3}{3p_{i,j}^2}\right]\cdot P(m_{i,j}), \tag{14}$$

where $P(m_{i,j})$ denotes the probability

$$P(m_{i,j})=\Delta_i f(m_{i,j})=\Delta_i\frac{2}{\sigma^2}e^{-\frac{2m_{i,j}}{\sigma}}. \tag{15}$$

By using the Langrangian multipliers, we can obtain the optimum number of cells in one region $p_{i,j}$, which yields the minimum granular distortion defined by the equation (14). Because we are designing an optimal quantizer for one value of variance σ_0, in calculating $p_{i,j}$ we will use σ_0 instead of σ. We will start from the following equation:

$$J=D_g(i)+\lambda\sum_{j=1}^{L_i}p_{i,j}. \tag{16}$$

After differentiating J with respect to $p_{i,j}$ and equalizing the derivate with zero we get

$$p_{i,j}=\frac{N_i}{4}\frac{m_{i,j}\sqrt[3]{g_0(m_{i,j})}}{\sum_{k=1}^{L_i}m_{i,k}\sqrt[3]{g_0(m_{i,k})}}, \tag{17}$$

where $g_0(m_{i,j})=\frac{1}{\sigma_0^2}e^{-\frac{2m_{i,j}}{\sigma_0}}$. If we multiply numerator and denominator with Δ_i, we can approximate the sum by the integral

$$p_{i,j}=\frac{N_i}{4}\frac{m_{i,j}\sqrt[3]{g_0(m_{i,j})}\cdot\Delta_i}{\int_{r_i}^{r_{i+1}}r\cdot\sqrt[3]{g_0(r)}\,dr}. \tag{18}$$

By substituting $p_{i,j}$ from equation (18) in equation (14) we get

$$D_g(i)=\frac{\Delta_i^2}{3}\sum_{j=1}^{L_i}m_{i,j}g_0(m_{i,j})\Delta_i+ \frac{64}{3N_i^2\Delta_i{}^2}I_0'(i)^2\sum_{j=1}^{L_i}m_{i,j}\frac{g(m_{i,j})}{g_0(m_{i,j})^{2/3}}\Delta_i. \tag{19}$$

After approximating the sum by the integral, we can rewrite (19) as

$$D_g(i)=\frac{\Delta^2}{3L_i^2}I(i)+\frac{64L_i^2}{3N_i^2\Delta^2}I_0'(i)^2I'(i). \tag{20}$$

The functions $I_0'(i)$, $I(i)$ and $I'(i)$ are defined as follows:

$$I_0'(i)=\int_{r_i}^{r_{i+1}}r\cdot\sqrt[3]{g_0(r)}dr=-\frac{9\sigma_0^{4/3}}{4}-e^{-\frac{2i\Delta}{3\sigma_0}} \left[\left(1-e^{\frac{2\Delta}{3\sigma_0}}\right)\left(1+\frac{2i\Delta}{3\sigma_0}\right)+\frac{2\Delta}{3\sigma_0}e^{\frac{2\Delta}{3\sigma_0}}\right];$$
$$I(i)=\int_{r_i}^{r_{i+1}}r\cdot g(r)dr=-\frac{1}{4}e^{-2i\Delta/\sigma} \left[\left(1-e^{2\Delta/\sigma}\right)\left(1+2i\Delta/\sigma\right)+2\Delta/\sigma\, e^{2\Delta/\sigma}\right];$$
$$I'(i)=\int_{r_i}^{r_{i+1}}r\cdot\frac{g(r)}{g_0(r)^{2/3}}dr=-\frac{9}{4}\left(\frac{\sigma_0{}^{5/3}}{3\sigma_0-2\sigma}\right)^2-e^{-\frac{2i\Delta}{3\sigma'}}\cdot \left[\left(1-e^{\frac{2\Delta}{3\sigma'}}\right)\left(1+\frac{2i\Delta}{3\sigma'}\right)+\frac{2\Delta}{3\sigma'}e^{\frac{2\Delta}{3\sigma'}}\right], \tag{21}$$

where $\sigma'=\frac{\sigma_0\sigma}{3\sigma_0-2\sigma}$.

After differentiating D_g from equation (20) with respect to L_i, and for σ_0 and equalizing the derivate with zero $\frac{\partial D_g(i)}{\partial L_i}=0$, we obtain the optimum number of subdomains in the i-th domain

$$L_{iopt}=\Delta\sqrt[4]{\frac{I_0(i)N_i^2}{64I_0'(i)^3}}, \tag{22}$$

where $I_0(i)$ is defined as

$$I_0(i)=\int_{r_i}^{r_{i+1}}r\cdot g_0(r)dr= -\frac{1}{4}e^{-2i\Delta/\sigma_0}\left[\left(1-e^{2\Delta/\sigma_0}\right)\left(1+2i\Delta/\sigma_0\right)+ 2\Delta/\sigma_0\,e^{2\Delta/\sigma_0}\right]. \tag{23}$$

4. Encoder design

Encoder design has to be a compromise between memory space and number of logical circuits (adders, comparators, counters etc.). The solution given in

Figure 3 is exactly the compromise between these demands. Encoder uniquely determines index I for every cell. Based on this parameter, the represent of the cell is obtained at the decoder side.

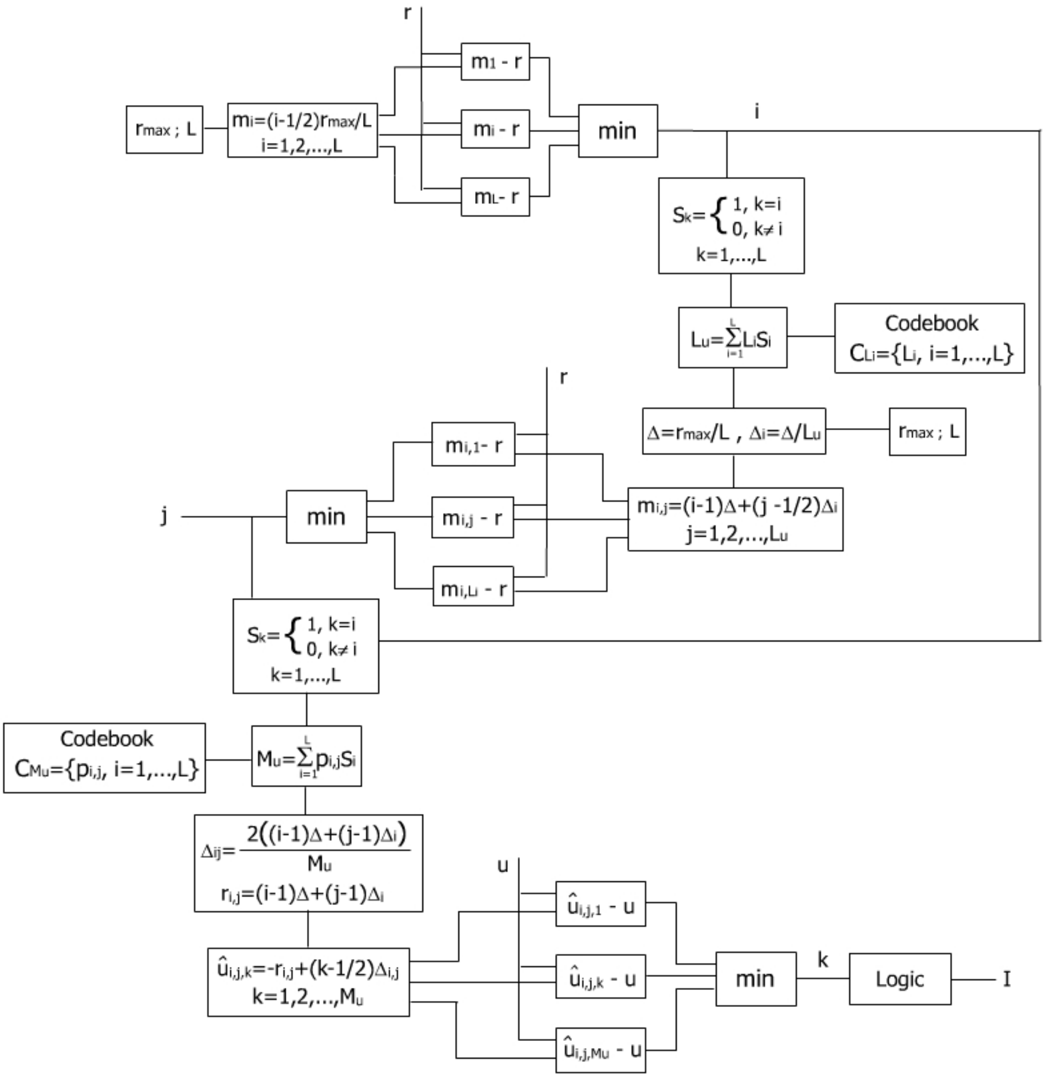

Figure 3. Encoder design

Using Helmert transformation, the input signal X is represented in *ru* coordinate system with r and u values. The encoder can be divided into 3 parts. The first part of the encoder is used for determination of the partition number where input signal belongs (index i). Beside input signal r, as input values, dynamic range of a two-dimensional quantizer r_{max} and the total number of partitions L are used for calculating of partition represents. We are looking for minimum subtraction between amplitudes of partition represents and input signal r.

In similar way, by subtracting amplitudes of subpartition represents of particular partition and input signal r, number of subpartition (index j) can be found. The input parameters are input signal r, dynamic range of quantizer r_{max}, the number of partitions L and the number of subpartitions for each partition L_i.

The third part of the encoder determines the position (number) of a cell in a subpartition of a particular partition, as a minimum subtraction between amplitudes of cell's represents and input signal u (index k). Every subpartition is divided into 4 equivalent regions, each having the same number of cells. Beside input signal u, as input values, dynamic range of a quantizer r_{max} and the number of cells for each region of a subpartition p_{ij} are used.

Indices i,j,k, the total number of cells for each partition N_i and the number of cells for each subpartition $4p_{ij}$ are used as parameters for determination of the unique index I of any cell where input signal belongs. One of possible ways of calculation of this index is:

$$I = (i-1)N_i + \sum_{q=1}^{j-1} 4p_{iq} + I_k, \tag{24}$$

where:

$$I_k = \begin{cases} k & \text{if the cell is in the } 1^{st} \text{ quadrant} \\ p_{i,j} + k & \text{if the cell is in the } 2^{nd} \text{ quadrant} \\ 2p_{i,j} + k & \text{if the cell is in the } 3^{rd} \text{ quadrant} \\ 3p_{i,j} + k & \text{if the cell is in the } 4^{th} \text{ quadrant}. \end{cases}$$

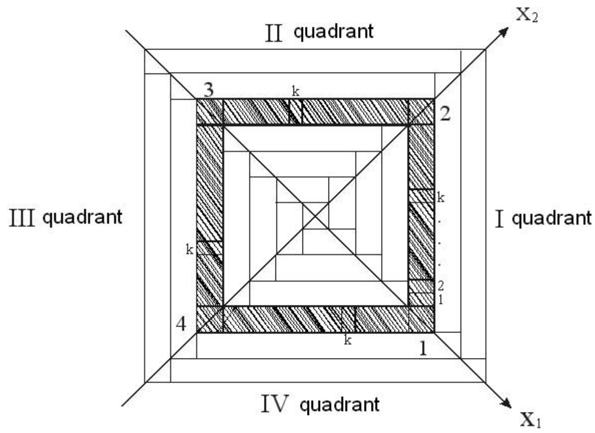

Figure 4. Cell position in two-dimensional space

Equation (24) is used for realization of logic block in Figure 3. The index I_k depends on the quadrant where particular cell is placed. Space partitioning for this case is shown in Figure 4.

Cell position depends on a sign of variables in input vector (x_1, x_2). If positive values are encoded with 0 and negative with 1, Table 1 shows a position of a cell in particular region of a subpartition.

Table 1

	x_1	x_2
I quadrant	0	0
II quadrant	1	0
III quadrant	1	1
IV quadrant	0	1

This encoder uses significantly less memory compared to the one that stores limits of every cell in a memory. Also, because the number of output points in each partition N_i is constant, the logic block is much simpler than in the case of the optimum number of output points, and less adders will be used in our case.

5. Conclusion

We suggested a model of switched piecewise vector quantizer which solves problem of variable input power in a wide range. It is shown that this switched quantizer can be applied for speech signals that have not only the random nature of instantaneous signal values, but also the random nature of the average power. During the switched quantizer design, the particular memory is needed. In cases where the memory resources are limited, it is possible to decrease the vector classes number *K,* but SNRQ will have larger variation due to input power changes.

We also proposed possible encoder design. The given solution is a compromise between memory space and number of logical circuits. Calculations for a special case of constant number of output points in each partition N_i were carried out. This significantly simplifies encoder compared to the one designed for optimal systems, and the performance is arbitrarily close to the optimum.

A simple expression for granular distortion, the number of subdomains and the number of output points in closed form is obtained. Memoryless Laplacian source is used, considering the possible application of this quantizer.

References

[1] **D. Hui, D.L. Neuhoff.** Asymptotic Analysis of Optimal Fixed-Rate Uniform Scalar Quantization. *IEEE Trans*, IT-47(3), 2001, 957–977.

[2] **S. Na, D.L. Neuhoff.** On the Support of MSE-Optimal, Fixed-Rate Scalar Quantizers. *IEEE Trans. Inform. Theory*, IT-47(6), 2001, 2972–2982.

[3] **F. Kuhlmann, J.A. Bucklew.** Piecewise Uniform Vector Quantizers. *IEEE Trans*., IT-34(5), 1988, 1259–1263.

[4] **P.F. Swaszek.** Unrestricted Multistage Vector Quantizers. *IEEE Trans*.,IT-38(3), 1992, 1169–1174.

[5] **D.G. Jeong, J. Gibson.** Uniform and Piecewise Uniform Lattice Vector Quantization for Memoryless Gaussian and Laplacian Sources. *IEEE Trans*., T-39(3), 1993, 786–804.

[6] **N.S. Jayant, P. Noll.** Digital Coding of Waveforms Principles and Applications to Speech and Video. *New Jersey: Prentice-Hall*, 1984.

[7] **A. Gersho, R.M. Gray.** Vector Quantization and Signal Compression. *Kluwer: Academ.Pub.*, 1992.

[8] **Z.H. Peric, J.D. Jovkovic, Z.J. Nikolic.** Two-Dimensional Laplas Source Quantization. *IEEE Conf. TELSIKS*-2001. *Nis, Serbia and Montenegro*, 33–37.

[9] **Z.H. Peric, A. Jovanovic, S. Bogosavljevic.** Optimal Piecewise Uniform Vector Quantization of the Memoryless Two Dimensional Laplacian source. *Third IEEE Workshop on "Intelegent Data Acquisition and Advanced Computing Systems: Technology and Applications, Sofia, Bulgaria* 5-7 *September,* 2005, 540-546.

[10] **Z.H. Peric, I.Lj. Tosic.** Piecewise uniform switched vector quantization of the memoryless two-dimensional Laplacian source. *International journal "Data Recording Storage and Processing, Vol*.6, *No*.1, 2004, 20-34.

Received April 2006.