Robust Companders

Demetrios Kazakos, Kami S. Makki
College of Science and Technology, Texas Southern University
3100 Cleburne Str. Houston, Texas, 77004, U.S.A
Department of Electrical Engineering & Computer Science, University of Toledo
2801 W. Bancroft, Toledo, Ohio 43606, U.S.A

Abstract—This paper considers the design of a robust quantizer for the class of input signal distributions having given quantiles and otherwise arbitrary shape. The quantizer model that consists of a compander and a uniform quantizer is utilized. The case of large number of quantization points is considered, and we use Bennett's and Gersho's approximation to the mean rth power distortion measure. We demonstrate that the piecewise linear compander provides robust quantization for the class of all input probability distributions having only their quantiles specified. The optimum robust solution is provided through the determination of all the required parameters. The problem is resolved for the case of block quantizers as well, and we show that the robust solution corresponds to a piecewise constant output point density function. The least favorable input multivariable density function is the piecewise uniform one.

I. Introduction

The design of quantizers has received much attention in the engineering literature [1, 2]. The majority of the work done on the subject of quantizer design deals with the case of perfectly known probability distribution function of the quantized signal. In reality, one has only partial knowledge of the distribution, and the quantizer must be designed for a class of distributions. It is desirable to design a quantizer that will minimize the average distortion for the worst case distribution of the specified class [3, 4]. Such a quantizer if it exists, is termed robust [1, 2]. In this paper we consider the case of large number of quantization levels, and we use the companding approximation to the average distortion due to Bennett [5]. We also utilize the quantizer model of Bennett [5], in which a Compressor is connected in series to a uniform quantizer and an expander, which is called the "companding" model. The idea of using Bennett's companding approximation for robust design of one dimensional quantizers originated in [3]. Also the problem of mismatch between the actual input statistics and the statistics assumed for the design of one dimensional quantizer have been studied in [7], [8].

We consider the class of signal probability distributions that have fixed quantiles. For this class, we seek the compander function that minimizes the maximum average distortion and maximizes the minimum average distortion. We show that the solution consists of a piecewise linear compander of an explicit form. Thus, we establish a robustness property for piecewise linear companders, which have been analyzed and designed by Dobrogowski [6]. In section III we develop a robust solution for the case of vector quantization and the rth power distortion measure. As performance criterion we utilize Gersho's formula for the rth power distortion, and the class of input signal distributions is characterized again in terms of fixed quantiles. It is shown that a robust quantizer exists, corresponding to a piecewise linear output point distribution.

II. Problem Statement and Solution for One Dimension

A quantizer is modeled as a compressor G with characteristic G(x), followed by a uniform N level quantizer on [a, b] and an expander G^{-1} at the output of the uniform quantizer. The expander is the inverse function of the compressor. It is assumed that G(x) is a monotone increasing function, and its derivative is denoted by g(x). The input signal x has a distribution F(x) and density function f(x), with support on [a, b]. Our partial knowledge of f(x) is characterized by the fact that only a number of quantiles of f(x) are known.

$$H = \{f(x); F(x_{k+1}) - F(x_k) = q_k, k = 0,1,\dots, M-1,$$
$$a = x_0 < x_1 < \dots < x_{M-1} < x_M = b\}$$

As performance measure of the quantizer we will use Bennett's [5] approximate expression for the mean square distortion for very large number N of quantizer output levels:

$$E(x - Q_N(x))^2 \cong \frac{1}{12N^2}\int_a^b \frac{f(x)}{[g(x)]^2}dx \tag{1}$$

where $Q_N(x)$ is a quantizer mapping.
The function g(x) is also the asymptotic output level density, i.e.: [9],

$$g(x) = \lim_{N\to\infty}\frac{N(x)}{N}$$

where N(x)dx is the number of quantization levels in the interval [x, x+dx]. Hence, we have the conditions:

$$\int_a^b g(x)dx = 1, g(x) > 0$$

We define the functionals:

$$D(f,g) \overset{\Delta}{=} \int_a^b \frac{f(x)}{[g(x)]^2}dx = \sum_{k=0}^{M-1} D_k(f,g) \tag{2}$$

$$D_k(f,g) \overset{\Delta}{=} \int_{x_k}^{x_{k+1}} \frac{f(x)}{[g(x)]^2}dx = k = 0,1,\dots, M-1 \tag{3}$$

The problems now are:

(A) $\sup_{f\varepsilon H} \inf_{g} D(f,g)$

(B) $\inf_{g} \sup_{f\varepsilon H} D(f,g)$

under the conditions that g(x) > 0, and $\int_a^b g(x)dx = 1$. In general, we have:

$$\inf_{g} \sup_{f\varepsilon H} D(f,g) \geq \sup_{f\varepsilon H} \inf_{g} D(f,g) \quad (4)$$

which means that the order of minimizing and maximizing is essential to the result. If (4) is an equality, then we say that a saddle point or robust solution exists.

Assuming g(x)>0 on [a, b], and using Hölder's inequality, [8], we find:

$$\int_{x_k}^{x_{k+1}} f^{1/3}(x)dx \leq [\int_{x_k}^{x_{k+1}} \frac{f(x)}{[g(x)]^2}dx]^{1/3}[\int_{x_k}^{x_{k+1}} g(x)dx]^{2/3} \quad (5)$$

or

$$[\int_{x_k}^{x_{k+1}} f^{1/3}(x)dx]^3 \leq D_k(f,g) \cdot [G(x_{k+1}) - G(x_k)]^2 \quad (6)$$

where:

$$G(x) = \int_a^x g(z)dz\,,$$

with equality if and only if:

$$g(x) = C \cdot [f(x)]^{1/3}, x_k \leq x \leq x_{k+1}\text{; C = Constant} \quad (7)$$

We now observe from (2), (6) that the solution of problems (A), (B) can be achieved in two steps. We fix at first the set of constants

$$\bar{G} = \{G(x_k); k = 1,..., M-1\}$$

and we find the constrained minimum over $g\varepsilon\bar{G}$. At the second step, we optimize by varying the M-1 parameters $\bar{G}$. The constrained minimization of D(f, g) is achieved by the M independent constrained minimizations of $\{D_k(f, g)\}$. Hence

$$\inf_{g\varepsilon G} D_k(f,g) = d_k^{-2}[\int_{x_k}^{x_{k+1}} f^{1/3}(x)dx]^3 \quad (8)$$

where:

$$d_k \stackrel{\Delta}{=} G(x_{k+1}) - G(x_k) > 0; \sum_{k=0}^{M-1} d_k = 1 \quad (9)$$

and the minimizing g is given by (7). If we apply Hölder's inequality again, we find:

$$[\int_{x_k}^{x_{k+1}} [f(x)]^{1/3}dx]^3 \leq (x_{k+1} - x_k)^2 [\int_{x_k}^{x_{k+1}} f(x)dx = (x_{k+1} - x_k)^2 q_k \quad (10)$$

Due to $g\varepsilon\bar{G}$, we have to find the extremum. Thus, the maximum is achieved for

$$g(x) = \text{Const} = (x_{k+1} - x_k)^{-1} \cdot d_k$$

and we find:

$$\inf_{g\varepsilon G} \sup_{f\varepsilon H} D_k(f,g) = q_k \cdot d_k^{-2}(x_{k+1} - x_k)^2 = \sup_{f\varepsilon H} \inf_{g\varepsilon G} D_k(f,g) \quad (11)$$

Summing (11) for k=0,1,…M-1 we find:

$$\sup_{f\varepsilon H} \inf_{g\varepsilon G} D(f,g) = \inf_{g\varepsilon G} \sup_{f\varepsilon H} D(f,g) = D(f^*, g^*) =$$

$$= \sum_{k=0}^{M-1} q_k (x_{k+1} - x_k)^2 d_k^{-2} \quad (12)$$

where f* is the uniform density and g* is the robust compander with piecewise constant slope.

As a final step, we will minimize (12) over the choice of the set $\bar{G}$. This amounts to minimizing (12) over $d_0 \dots d_{M-1}$, under the constraints

$$d_k > 0, \sum_{k=0}^{M-1} d_k = 1 \quad (13)$$

For this step, we use a single Lagrange multiplier z, and the problem is:

Minimize over $d_0, \dots d_{M-1}$, the quantity

$$J = \sum_{k=0}^{M-1} q_k (x_{k+1} - x_k)^2 d_k^{-2} + 2zd_k$$

Setting the partial derivative of J with respect to d_k equal to zero, we find:

$$z^{1/3} \cdot d_k = [q_k (x_{k+1} - x_k)^2]^{1/3} \quad (14)$$

III. MULTIDIMENSIONAL QUANTIZERS

In this section we generalize the previous results in two directions. We consider the rth power distortion measure and block quantization. Let x = (x1….xk) be a k-dimensional random vector with density f(x), with support on a k-dimensional bounded region V. An N-point block quantizer is a function y = Q(x); $V \to V$ which maps $x\varepsilon V$ into one of N output vectors {y1...yn}cV. The points {yj} and a partition of N disjoint sets Sj, such that $\bigcup_{j=1}^{N} S_j = V$, specify the quantizer Q. Thus: $Q(x) = y_j \text{ if } x\varepsilon S_j, j = 1,..., N$. A nonuniform block quantizer can be modeled by a block compressor followed by a uniform block quantizer and then a block expander. The block compressor is a nonlinear continuous mapping from V into itself, and the block expander is the exact inverse function of the compressor. The design of the compressor characteristic is, for large N, equivalent to the choice of the probability density function of the quantizer output y. This is because for large N the uniform quantizer approaches the behavior of an identity operator.

Let g(y), $y\varepsilon V$ denote the output point density function, which is to be chosen optimally. We measure the performance of such a quantizer by the distortion:

$$\Delta \stackrel{\Delta}{=} k^{-1} \cdot E \| x - Q(x) \|^r \quad (15)$$

where $r \geq 0$ and $\| \cdot \|$ denotes the usual l_2 norm. For r = 2, Δ is the mean square "per letter" distortion measure. Gersho [1] has developed heuristically the following approximate expression for large N,

$$\Delta = N^{-m} C(k,r) \int_V \frac{f(x)}{[g(x)]^m} dx \tag{16}$$

where $m = r \cdot k^{-1}$, and C(k, r) is called “coefficient of quantization.” Some values of C(k, r) are tabulated or bounded in **[1]**. The output point density function satisfies the constraints:

$$g(x) > 0, \int_V g(x) dx = 1 \tag{17}$$

Using Hölder’s inequality and constraint (17), Gersho showed that;

$$\int_V f(x) \cdot [g(x)]^{-m} dx \geq [\int_V f(x)]^{1/1+m \, 1+m} = \| f(x) \|_{1/1+m} \tag{18}$$

with equality if and only if

$$g(x) = C \cdot [f(x)]^{1/1+m} \tag{19}$$

We now define the functional

$$D(f,g) \stackrel{\Delta}{=} \int_V \frac{f(x)}{[g(x)]^m} dx \tag{20}$$

and are ready to pursue the robust design of g(x). Define the family of densities with known quantiles:

$$H=\{f(x); \int_{V_k} f(x) dx = q_k, k = 1,...M\} \tag{21}$$

where:

$$\bigcup_{k=1}^{M} V_k = V, V_j \Omega V_i = \emptyset \text{ for } i \neq j \tag{22}$$

We also define by v_k the volume of V_k, and we assume that all $v_k \neq 0$.

Define, now, the partial distortions:

$$D_k(f,g) \stackrel{\Delta}{=} \int_{V_k} \frac{f(x)}{[g(x)]^m} dx \tag{23}$$

and the set of output densities:

$$\bar{G} \stackrel{\Delta}{=} G(d) \stackrel{\Delta}{=} g(x); \{ \int_{V_k} g(x) dx = d_k > 0, k = 1,..., M; \sum_{k=1}^{M} d_k = 1\} \tag{24}$$

$$d \stackrel{\Delta}{=} (d_1 \cdots d_M)$$

Using Hölder’s inequality, and in a manner analogous to the derivation of (5), (6), we have the inequality:

$$\int_{V_k} [f(x)]^{1/1+m} dx \leq [D_k(f,g)]^{1/1+m} \cdot [\int_{V_k} g(x) dx]^{m/1+m} \tag{25}$$

or:

$$D_k(f,g) \geq d_k^{-m} \cdot [\int_{V_k} [f(x)]^{1/1+m}]^{1+m} \tag{26}$$

with equality if

$$g(x) = C \cdot [f(x)]^{1/1+m} \quad \text{for } x \varepsilon V_k ; C = \text{constant} \tag{27}$$

In a manner analogous to the derivation of (10), we find from Hölder’s inequality:

$$[\int_{V_k} [f(x)]^{1/1+m} dx]^{1+m} \leq [\int_{V_k} f(x) dx] \cdot [\int_{V_k} dx]^m \tag{28}$$

or

$$[\int_{V_k} [f(x)]^{1/1+m} dx]^{1+m} \leq q_k \cdot v_k^m \tag{29}$$

with equality if f(x) = constant for $x \varepsilon V_k$.

From (28), (29) we conclude:

$$\sup_{f \varepsilon H} \inf_{g \varepsilon \bar{G}} D_k(f,g) = v_k^m q_k \cdot d_k^{-m} \tag{30}$$

With respect to eq. (12), we find:

$$\sup_{f \varepsilon H} D_k(f,g) = q_k \cdot \sup_{x \varepsilon V_k} [g(x)]^{-m}$$

and:

$$\inf_{g \varepsilon \bar{G}} \sup_{f \varepsilon H} D_k(f,g) = q_k \cdot [\sup_{g \varepsilon \bar{G}} \inf_{x \varepsilon V_k} g(x)]^{-m}$$

Because of the constraint $g \varepsilon \bar{G}$, the max-min value is achieved by the uniform on V_k distribution g, i.e.:

$$g(x) = d_k v_k^{-1} \text{ for } x \varepsilon V_k$$

Hence:

$$\inf_{g \varepsilon G} \sup_{f \varepsilon H} D_k(f,g) = q_k d_k^{-m} v_k^m = \sup_{g \varepsilon H} \inf_{g \varepsilon \bar{G}} D_k(f,g) \tag{31}$$

and we have shown that the piecewise constant f* is the least favorable input density, and the piecewise constant g* is the minimax design solution, under the game constraints $f \varepsilon H$, $g \varepsilon \bar{G}$:

$$\inf_{g \varepsilon \bar{G}} \sup_{f \varepsilon H} D_k(f,g) = D(f^*, g^*) = \sup_{f \varepsilon H} \inf_{g \varepsilon \bar{G}} D(f,g) = \sum_{k=1}^{M} q_k v_k^m d_k^{-m} \tag{32}$$

As a final step, we minimize D(f*, g*) over the positive parameters $\{d_1 \ldots d_M\}$, under the constraint $\sum_{k=1}^{M} d_k = 1$. This last optimization in effect removes the constraint $g \varepsilon \bar{G}$, by scanning the whole parameter space through a variation of the parameters $d_1 \ldots d_M$. The minimization is performed using a Lagrange multiplier z. The problem now is:

Minimize over $d_1 \ldots . d_M$ the quantity

$$J = \sum_{k=1}^{M} q_k \cdot v_k^m d_k^{-m} + mzd_k$$

Setting the partial derivative of J with respect to d_k equal to zero, we find:

$$z^{1/m+1} \cdot d_k = (q_k v_k^m)^{1/1+m} \tag{33}$$

Summing (33) over k, and using the previous constraint we find z:

$$z^{1/m+1} = \sum_{k=1}^{M} (q_k v_k^m)^{1/1+m}$$

Thus, the optimum d_k is:

$$d_k^* = (q_k v_k^m)^{1/1+m} [\sum_{j=1}^{M} (q_j v_j^m)^{1/1+m}]^{-1} \tag{34}$$

The corresponding minimum distortion is:

$$D_0(f^*, g^*) \stackrel{\Delta}{=} \min_{\bar{G}} D(f^*, g^*) = [\sum_{k=1}^{M} (q_k v_k^m)^{1/1+m}]^{1+m} \tag{35}$$

In conclusion, the robust design of the quantizer amounts to a piecewise constant output density g(y), having the form:

$$g^*(y) = d_k^* v_k^{-1} \text{ for } y \varepsilon V_k,\ k = 1,...M$$

with d_k^* given by eq. (34). The least favorable input density function is the piecewise uniform one of the class H. Hence, the robust compander is the one that transforms f* into g*, and therefore is piecewise linear.

IV. Final Comments

We have developed a solution to the design of robust quantizers for the class of input signal distributions with known quantiles and arbitrary shape. The case of large number of quantization points was considered, and use of the Bennett-Gersho distortion approximation was made. The robust solution amounted to a piecewise constant output point density function, for the general block quantization and rth power distortion measure. Thus, the corresponding robust compander is piecewise linear, which makes it very easy to implement and of practical utility. The results of the present paper are limited to high quality quantization, which amounts to low distortion and large number of output quantization levels.

V. References

[1] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer , 1991.

[2] A. Rao, D. Miller, K. Rose, and A. Gersho, "Generalized Vector Quantization: Jointly Optimal Quantization and Estimation," IEEE International Symposium on Information Theory, Whistler B.C., Canada, September 1995.

[3] R. Hagen and P. Hedelin, “Robust vector quantization by a linear mapping of a block code,” *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 200–218, Jan. 1999.

[4] B. Farber and K. Zeger, “Quantizers with uniform encoders and channel optimized decoders,” *IEEE Transactions on Information Theory*, vol. 50, no. 1, pp. 62–77, Jan. 2004.

[5] W.R. Bennett, “Spectra of Quantized Signals”, Bell System Technical Journal, vol. 27, pp.446-472, July 1948.

[6] A. Dobrogowski, “Optimal Segmented Compression Characteristic for Speech”, *IEEE Transactions on Communications*, vol. COM-28, no. 2, February 1980, pp. 301-302.

[7] R.M. Gray and L.D. Davisson, “Quantizer Mismatch”, IEEE Transactions on Comm., vol. COM-23, April 1974, pp. 439-443.

[8] W. Mauersberger, “Experimental results on the performance of mismatched quantizers”, *IEEE Transactions on 1T*, Vol. 25, no. 3, May 1979, pp. 654-655.

[9] R.M. Gray and A.H. Gray, “Asymptotically optimal quantizers”, IEEE *Transactions on Information Theory*, vol. 23, no. 1, Jan. 1977, pp. 143-144.