ON THE SIZES OF VORONOI CELLS IN ENTROPY-CONSTRAINED VECTOR QUANTIZATION

Stephan F. Simon
Institut für Elektrische Nachrichtentechnik
Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen, 52056 Aachen, Germany
Tel:+49–241–807677, Fax:+49–241–8888196
E-mail: simon@ient.rwth-aachen.de

ABSTRACT

Voronoi cells for vector quantization subject to an entropy constraint are considered. It is shown that the constraint on the output entropy leads to a weaker and even vanishing dependency of the Voronoi cell's volume on the probability density function. Using some simplifying assumptions like linearization of a small part of the n–dimensional input space and modeling of the cell shapes as hyperspheres leads to an analytic expression of the quotient of the volumes of two neighboring Voronoi cells.

The results confirm the use of entropy coded lattice vector quantizers with optimized reproduction vectors in cases of vanishing dependency and may in other cases be exploited for the design of vector companders to be used in conjunction with lattice vector quantization.

1 INTRODUCTION

Chou *et al.* [1] provided a practical design algorithm for entropy-constrained vector quantizers which incorporates an entropy constraint directly into the iterative design process by extending the nearest neighbor quantization rule by weighted addition of codeword length or self-information.

In contrast to the high-resolution case (with constrained resolution) [2], where the so-called quantization- or Voronoi cells are small in regions with large probability density function (pdf) and vice versa, a different observation can be made in the entropy constrained case for sources with smooth pdf: the design algorithm which partitions the input space into a set of Voronoi cells yields, for a wide range of the wordlength weighting factor λ, partitions with approximately constant Voronoi cell volumes [3]. This observation was the incentive for this work in which the sizes of Voronoi cells in entropy-constrained vector quantization (ECVQ) are investigated theoretically.

After introducing some simplifying assumptions like linearization of the pdf and modeling of cell shapes as spheres, the proportion of the volumes of two neighboring cells is derived as a function of λ, of the gradient of the pdf, and of the vector dimension. The results reveal a better insight into the characteristics of ECVQ and provide some interesting conclusions on the design of vector quantizers.

2 BACKGROUND

The entropy-constrained minimization problem can be stated as follows: Given an n–dimensional random vector $\mathbf{X} = (X_1, ..., X_n)^T$, choose a quantizer $\mathbf{Q}$ and a variable-length mapping γ that minimize the average distortion

$$D = \mathrm{E}\{\rho(\mathbf{X}, \mathbf{Q}(\mathbf{X}))\}$$

subject to

$$\mathrm{E}\{|\gamma(\mathbf{Q}(\mathbf{X}))|\} \leq R.$$

This constrained minimization problem is usually replaced by a simpler unconstrained problem, the minimization of the Lagrangian

$$J_\lambda = \mathrm{E}\{\rho(\mathbf{X}, \mathbf{Q}(\mathbf{X})) + \lambda|\gamma(\mathbf{Q}(\mathbf{X}))|\}. \quad (1)$$

The n–dimensional N–point vector quantizer (VQ) $\mathbf{Q} : \mathbb{R}^n \to \mathbb{R}^n$ is characterized by a partition $P = \{\mathcal{P}_1, ..., \mathcal{P}_N\}$ of the n–dimensional Euclidean space $\mathbb{R}^n$ into N quantization cells and a codebook $C = \{\mathbf{y}_1, ..., \mathbf{y}_N\}$ consisting of N n–dimensional reproduction vectors. An n–dimensional source vector $\mathbf{x} = (x_1, ..., x_n)^T$ is mapped onto one of the codebook vectors, $\mathbf{Q}(\mathbf{x}) = \mathbf{y}_i$, if it belongs to the corresponding partition cell, $\mathbf{x} \in \mathcal{P}_i$. The partition cell can be defined by the biased nearest neighbor rule

$$\mathcal{P}_i = \{\mathbf{x} \in \mathbb{R}^n : \rho(\mathbf{x}, \mathbf{y}_i) + \lambda l_i \leq \rho(\mathbf{x}, \mathbf{y}_j) + \lambda l_j,\ \forall j \neq i\}. \quad (2)$$

$\rho(\mathbf{x}, \mathbf{y})$ measures the distortion between input and reproduction; in this work we consider the squared error $\rho(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y})$. $l_i = |\gamma(\mathbf{Q}(\mathbf{x}))|,\ \forall \mathbf{x} \in \mathcal{P}_i$ is the length (or alternatively the self-information) of the codeword $\gamma(\mathbf{y}_i)$ which is used to store or transmit the index i, corresponding to the codebook vector $\mathbf{y}_i$. The Lagrange multiplier λ reflects the relative importance of codeword length versus distortion.

Given the n–dimensional random vector $\mathbf{X}$ with joint probability density function $p_{\mathbf{X}}(\mathbf{x})$ and fixed λ, the ECVQ design algorithm [1] aims at minimizing the functional (1) by iteratively performing the following three steps on a training set which should be large enough to represent $p_{\mathbf{X}}(\mathbf{x})$ with sufficient accuracy.

i) For codebook C and codeword lengths $L = \{l_1, ..., l_N\}$ fixed, each element of the training set is assigned its appropriate codebook vector or index. This corresponds to fixing the partition P. Empty cells are eliminated.

ii) For P and C fixed, an entropy code with wordlengths L is generated.

iii) For P and L fixed, the new codebook vectors are obtained by $\mathbf{y}_i = \int\limits_{\mathcal{P}_i} \mathbf{x}\, p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$.

The algorithm is stopped when no further reduction of J_λ is obtained.

3 SIZES OF VORONOI CELLS IN ECVQ

3.1 Observations

The ECVQ design algorithm was applied to a large number of different sources, including the dimensions $n = 1, 2, 3, 4, 8$. Under consideration were correlated and uncorrelated sources as well as sources with nonlinear statistical dependencies. An observation which could be made in many of these experiments was that the volumes of bounded Voronoi cells tend to approximate a constant volume which depends on the Lagrange multiplier λ. Fig. 1 shows a representative example. In order to facilitate the visualization of the resulting Voronoi cells, the dimension $n = 2$ was chosen. The 10^5 2–dimensional training vectors were taken from a bivariate, correlated Gaussian distribution with variance $\sigma^2 = 1$. The codebook was initialized with $N = 800$ codebook vectors which also were realizations of the same distribution. The optimization with $\lambda = 0.01$ stopped after 18 iterations with $N = 139$ codebook vectors remaining. The resulting partition of $\mathbb{R}^2$ is depicted in Fig. 1 (left). The scatter diagram in Fig. 1 (right) shows the resulting volumes[1] v_i of the bounded cells versus their probabilities P_i. For cells with $P_i \gtrapprox \frac{1}{10} \max\limits_j P_j$, it seems at first glance that the cell volume v_i is independent of P_i in this example.

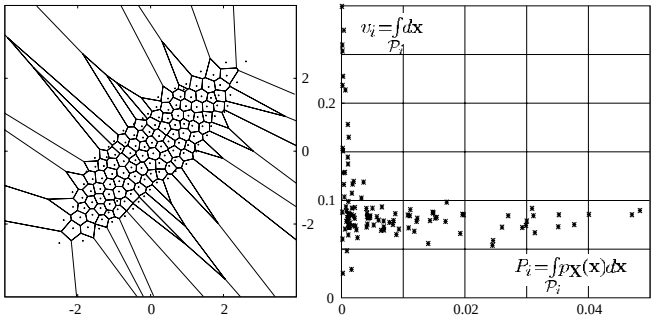

Fig. 1: Final partition after ECVQ optimization with random initialization (left) and resulting volumes of bounded quantization cells versus their probabilities (right).

For comparison, an example for VQ without entropy constraint — optimized for the same source — is also provided. In this special case where $\lambda = 0$ the ECVQ design algorithm equals the generalized Lloyd algorithm by Linde *et al.* [4]. Figure 2 shows the resulting partition and the scatter diagram

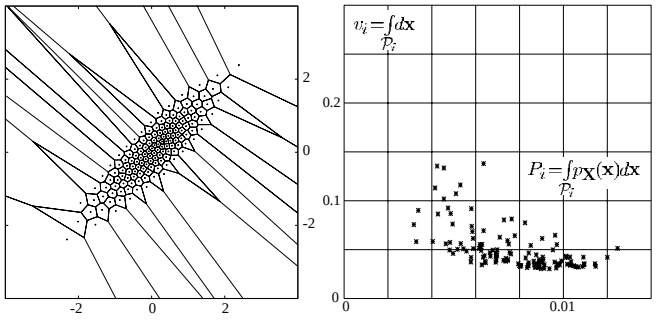

Fig. 2: Final partition after optimization according to [4] with random initialization (left) and resulting volumes of bounded quantization cells versus their probabilities (right).

of the cell volumes versus cell probabilities. This time, the design algorithm was initialized with $N = 150$ randomly chosen codebook vectors as aforementioned. During the optimization the codebook size N remained unchanged which is typical for the case without entropy constraint and also justifies the term "resolution-constrained VQ". The scatter diagram (Fig. 2, right) confirms our expectation of small Voronoi cells where the pdf is large and vice versa.

In the following approach the dependency of cell volumes on the factor λ, the distribution, and the dimension n is analyzed theoretically.

3.2 Theoretical Approach

In high-rate quantization theory the pdf is assumed to be constant inside a partition cell $\mathcal{P}_i$. The reproduction vector minimizing the distortion D_i of this cell is its center $\mathbf{c}_i$. In ECVQ, especially at moderate and lower rates, this assumption must be abandoned in order to explain the dependency of cell volumes on λ, the pdf, and the dimension n. We make the simplest extension of the constant pdf assumption and use a univariate linear approximation of the pdf $p_{\mathbf{X}}(\mathbf{x}) = p(x_1) = k(1 + sx_1)$ for a small portion of $\mathbb{R}^n$ which contains two Voronoi cells. The two cells are approximated by n–dimensional spheres which touch each other (see Fig. 3).

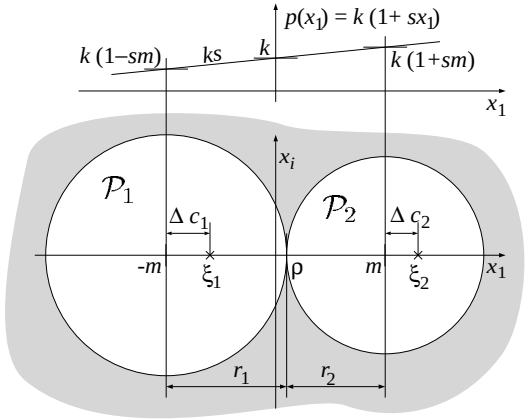

Fig. 3: Model assumptions: two spherical n–dimensional cells touching each other; univariately distributed input, linearly approximated.

[1] For convenience, we use the term "volumes" for all dimensions.

For the distortion of cell i we have

$$D_i = \int_{\mathcal{P}_i} p_{\mathbf{X}}(\mathbf{x}) \left((x_1 - (c_i + \Delta c_i))^2 + \sum_{j=2}^{n-1} x_j^2 \right) d\mathbf{x},$$

assuming that $\mathbf{c}_i = (c_i, 0, ..., 0)^T$ and $\mathbf{y}_i = (\xi_i, 0, ..., 0)^T$ (with $\xi_i = c_i + \Delta c_i$) are its center and reproduction vector, respectively.

We will first solve for the vector $\mathbf{y}_i$ which minimizes D_i. After substitution of $x_1 - c_i$ and some algebraic steps the distortion can be expressed as

$$D_i(\Delta c_i, c_i, r_i) = k \int_{x_1=-r_i}^{r_i} (1 + s(x_1 + c_i)) \cdot \left((x_1 - \Delta c_i)^2 \cdot V_{n-1}\left(\sqrt{r_i^2 - x_1^2}\right) + g_i(x_1, c_i, r_i) \right) dx_1$$

where $V_{n-1}\left(\sqrt{r_i^2 - x_1^2}\right)$ is the volume of an $(n-1)$–dimensional sphere with radius $\sqrt{r_i^2 - x_1^2}$ and

$$V_n(r) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} r^n. \tag{3}$$

The term $g_i(x_1, c_i, r_i)$ is independent of Δc_i. Setting $\frac{\partial}{\partial \Delta c_i} D_i(\Delta c_i, c_i, r_i)$ to zero and solving for Δc_i yields

$$\begin{aligned} \Delta c_i(c_i, r_i) &= \\ &= \frac{\int_{x_1=-r_i}^{r_i} x_1 (1 + sc_i + sx_1) \cdot V_{n-1}\left(\sqrt{r_i^2 - x_1^2}\right) dx_1}{\int_{x_1=-r_i}^{r_i} (1 + sc_i + sx_1) \cdot V_{n-1}\left(\sqrt{r_i^2 - x_1^2}\right) dx_1} \\ &= \frac{1}{n+2} \frac{s}{1 + sc_i} r_i^2. \end{aligned} \tag{4}$$

The probability for cell i is given by

$$P_i = \int_{\mathcal{P}_i} p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} = k(1 + sc_i)\, V_n(r_i). \tag{5}$$

In order to compute the proportion of the volumes $V_n(r_1)$ and $V_n(r_2)$, the point $\boldsymbol{\rho} = (\rho, 0, ..., 0)^T$ with $r_1 = m + \rho$ and $r_2 = m - \rho$ is considered (see Fig. 3). At this point both spheres with centers $c_1 = -m$ and $c_2 = m$ (without loss of generality) touch. After ECVQ-optimization, the equation

$$(m + \rho - \Delta c_1)^2 + \lambda l_1 = (m - \rho + \Delta c_2)^2 + \lambda l_2 \tag{6}$$

with $l_i = -\log_2 P_i$ holds.

Now the task is to find ρ. (For $\rho = 0$ both spheres would have equal volumes). From (6), (5) and (3) we have

$$\begin{aligned} 0 = 4 &\left(\frac{\rho}{m} - \frac{\Delta c_1 + \Delta c_2}{2m} \right) \left(1 - \frac{\Delta c_1 - \Delta c_2}{2m} \right) \\ &+ \frac{\lambda}{m^2} \log_2 \left(\frac{1 + sm}{1 - sm} \cdot \left(\frac{1 - \frac{\rho}{m}}{1 + \frac{\rho}{m}} \right)^n \right) \end{aligned} \tag{7}$$

with

$$\frac{\Delta c_1 + \Delta c_2}{2m} = \frac{sm}{n+2} \cdot \frac{\left(\frac{\rho}{m}\right)^2 + 2\frac{\rho}{m} sm + 1}{1 - (sm)^2}$$

and

$$\frac{\Delta c_1 - \Delta c_2}{2m} = \frac{sm}{n+2} \cdot \frac{\left(\frac{\rho}{m}\right)^2 sm + 2\frac{\rho}{m} + sm}{1 - (sm)^2}.$$

We use the substitutions $\tilde{\rho} = \frac{\rho}{m}$, $\tilde{\lambda} = \frac{\lambda}{m^2}$ and $\tilde{s} = sm$ to eliminate m and obtain non-dimensional quantities.

Now we search for the function $\tilde{\rho}(\tilde{\lambda}, \tilde{s}, n)$ which depends solely on the Lagrangian factor $\tilde{\lambda}$, on $\tilde{s}$ which represents the gradient of the pdf, and on the dimension n. The solutions of (7) in terms of $\tilde{\rho}$ can be computed numerically. The quotient of the volumes is then given by

$$\frac{V_n(r_1)}{V_n(r_2)} = \left(\frac{1 + \tilde{\rho}}{1 - \tilde{\rho}} \right)^n \tag{8}$$

4 RESULTS

4.1 Theoretical Results

Fig. 4 depicts $\tilde{\rho}(\tilde{\lambda}, \tilde{s}, n)$ for the two-dimensional case $n = 2$. In the special case without entropy constraint ($\tilde{\lambda} = 0$) which

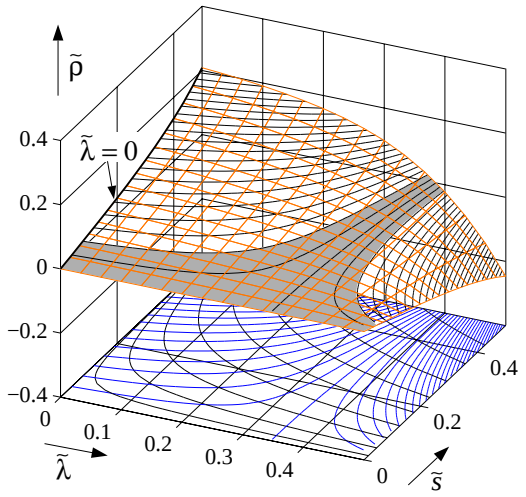

Fig. 4: Displacement of contact point $\tilde{\rho}(\tilde{\lambda}, \tilde{s}, n)$ for $n = 2$, (mesh and contour plot). Gray region: $-0.02 \le \tilde{\rho}(\tilde{\lambda}, \tilde{s}, n) \le 0.02$. Bottom: contour lines and $\lambda \cdot s^2 = \tilde{\lambda} \cdot \tilde{s}^2 = \text{const}$.

is represented by the fat line on the left, a strong dependency exists between the gradient of the pdf and the shift of the contact point. This dependency of $\tilde{\rho}$ on $\tilde{s}$ gets weaker as $\tilde{\lambda}$ increases. For $\tilde{\lambda} \approx 0.375$ (and $n = 2$) it almost vanishes. For larger values of $\tilde{\lambda}$ it even changes its sign[2]. In the example of Fig. 1 we had $\tilde{\lambda} \approx 0.4$. The contour lines which are plotted in steps of 0.01 are also projected on the bottom. The other curves on the bottom indicate where $\lambda \cdot s^2 = \frac{\lambda}{m^2} \cdot (sm)^2 = \tilde{\lambda} \cdot \tilde{s}^2 = \text{const}$. Usually, λ is a fixed parameter of the entropy-constrained VQ and s is fixed at a given location of the input space, so $\lambda \cdot s^2$ is fixed, too. Hence, a variation of $\tilde{\rho}$ can only be achieved by a variation of m. That means the quotient of the volumes

[2] For very large values of $\tilde{\lambda}$ the dependency of $\tilde{\rho}$ on $\tilde{s}$ may even be stronger than in the case $\tilde{\lambda} = 0$, but has negative sign. This is a pathological case which is automatically corrected by the ECVQ design algorithm by preferring cells with short words and dropping smaller cells and hereby increasing m.

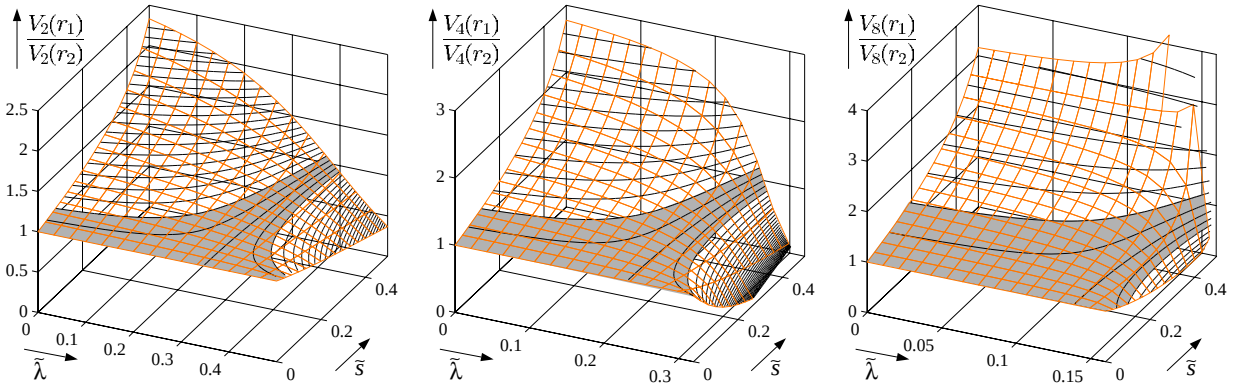

Fig. 5: Proportion of cell volumes for dimensions $n = 2,\ 4,\ 8$ (from left to right). Dark lines: contour lines.

of two neighboring cells at a given location is affected by enlarging or condensing them.

In the gray marked region we have $-\varepsilon \le \tilde{\rho}(\tilde{\lambda}, \tilde{s}, n) \le \varepsilon$ with $\varepsilon = 0.02$. Here, two neighboring cells have approximately equal size. We note that this region gets wide in terms of $\tilde{s}$ for $\tilde{\lambda}$ around 0.375.

By using (8), the quotient of volumes is obtained from $\tilde{\rho}$. Figure 5 depicts $\frac{V_n(r_1)}{V_n(r_2)}$ for the dimensions $n = 2,\ 4,\ 8$.

4.2 Comparison of Theory and Practice

Figure 6 provides a comparison of measured and predicted quotients $\frac{V_n(r_1)}{V_n(r_2)}$. Two-dimensional entropy constrained VQs were considered, optimized for a Gaussian i.i.d. source with variance $\sigma^2 = 1$ and parameters $\lambda = 0$ (marked by crosses) and $\lambda = 0.04$ (marked by circles), respectively. All pairs of adjacent bounded Voronoi cells were considered, regardless of their orientation.

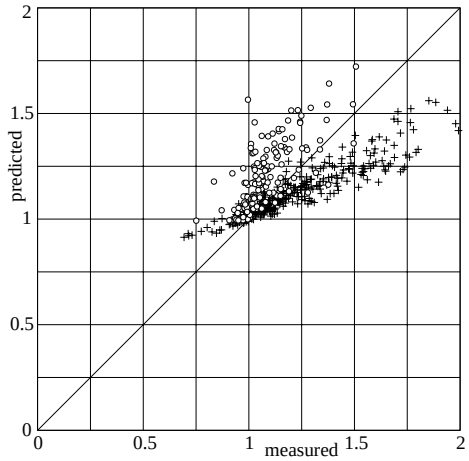

Fig. 6: A comparison of practical results and predicted values of the quotient $\frac{V_n(r_1)}{V_n(r_2)}$ for two-dimensional entropy constrained VQs designed for unit variance Gaussian input and $\lambda = 0$ (crosses) and $\lambda = 0.04$ (circles), respectively.

The predicted and measured values do not match exactly which is primarily due to the required coarse approximation of the cell shapes. Nevertheless, the comparison shows that the method presented here provides a way to approximately predict the resulting partition of vector quantizer design algorithms subject to an entropy constraint in terms of cell sizes.

5 CONCLUSION

Perhaps the most interesting region for practical coding applications is the one which satisfies

$$-\varepsilon \le \tilde{\rho}(\tilde{\lambda}, \tilde{s}, n) \le \varepsilon \tag{9}$$

for a small value of ε (see contour lines in Figures 4 and 5). In this region it is consequent to initialize an entropy-constrained VQ with a lattice partition. Besides speeding up the design algorithm, this initialization leads to considerable gains in rate-distortion performance if the training set is small [3]. If a very small performance loss can be tolerated, a significant reduction in complexity is obtained by replacing the entropy-constrained VQ by an appropriately scaled lattice VQ with optimized reproduction vectors and followed by entropy coding.

Now consider vector quantizers having numerous cell pairs for which (9) is not met. Replacing such a VQ by a lattice vector quantizer is not a good choice. But replacing it by a combination of a lattice vector quantizer and a suitable compander can lead to a simple but very efficient VQ system which possesses the desired quotients of the volumes of neighboring cells. The results presented here can then be exploited to adjust the parameters of such a compander.

REFERENCES

[1] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Trans. ASSP*, vol. 37, pp. 31–42, Jan. 1989.

[2] S. Na and D. L. Neuhoff, "Benett's integral for vector quantizers," *IEEE Trans. Inform. Theory*, vol. IT-41, pp. 886–900, July 1995.

[3] S. F. Simon and W. Niehsen, "Initialization of entropy-constrained vector quantizers with lattices," in *Proc. International Picture Coding Symposium PCS '96*, vol. 1, (Melbourne, Australia), pp. 219–223, Mar. 1996.

[4] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84–95, Jan. 1980.