Proc. Conf. Information Sciences, Systems (Princeton, NJ), pp. 677-682, March 1992

Lattice-based Nonuniform Vector Quantization

A. K. Khandani[1] and P. Kabal[1,2]

[1]McGill University, 3480 University, Montreal, Canada, H3A 2A7
[2]INRS–Telecommunications, 3 Place du Commerce, Verdun, Canada, H3E 1H6

Abstract: We propose some practical methods for applying a lattice-based uniform vector quantizer to a nonuniform source. The first method, denoted as cluster quantization, is based on using the k-fold cartesian product of a one-dimensional compander in conjunction with a lattice quantizer. This scheme has an asymptotic gain of 1.53 dB with respect to the optimum one-dimensional quantizer. The complexity is essentially the complexity of decoding of a lattice. The second method, denoted as quantizer shaping, is based on selecting an appropriate boundary for a lattice quantizer. By increasing the space dimensionality, this scheme becomes asymptotically optimum. As a practical shaping method, we use the Voronoi region around the origin of a lattice to shape the quantizer. By using binary lattices, we can construct quantizers with an integral bit rate. In an extension of this scheme, we use a lattice partition chain $\Lambda_q^0/\ldots/\Lambda_q^m/\Lambda_q^{m+1}$ to provide a set of $m+1$ Voronoi constellations, $C(\Lambda_q^i/\Lambda_q^{i+1})$, $i=0,\ldots,m$. A copy of the Voronoi region of Λ_q^i is centered around each point of $C(\Lambda_q^i/\Lambda_q^{i+1})$. This results in higher resolution for the partitions around the origin. This is denoted as a nonuniform Voronoi quantizer. The group property of the Voronoi constellations is used to decrease the complexity of the operations invovled in the quantization. These are the operations of shaping, encoding, addressing and reconstruction. The overall complexity is in the order of the linear mappings. By using binary lattices, we construct quantizers with a rate very close to an integer number. This reduces the redundancy associated with a binary indexing of the quantizer output.

1 Introduction

In quantizing a source $s\in\mathcal{S}\subset\mathbf{R}$, $\mathbf{R}$ is the set of the real numbers, the objective is to represent s using a discrete number of values, say $\hat{s}\in\hat{\mathcal{S}}\subset\mathbf{R}$, which are close to s. We assume that the source symbols are processed on a k-dimensional basis, $\mathbf{s}\in\mathcal{S}_k\subset\mathbf{R}^k$, $\{\,.\,\}^k$ denotes the k-fold cartesian product, and are quantized to $\hat{\mathbf{s}}\in\hat{\mathcal{S}}_k\subset\mathbf{R}^k$. Obviously $\mathcal{S}_k\subset\mathcal{S}^k$ and $\hat{\mathcal{S}}_k\subset\hat{\mathcal{S}}^k$. The probability density function of $\mathbf{s}$ is denoted by $P_k(\mathbf{s})$.

The similarity between s and $\hat{s}$ is measured in terms of a distortion measure which is a function defined on $\mathcal{S}\times\hat{\mathcal{S}}$, say $d(s,\hat{s})$, and extended to $\mathcal{S}_k\times\hat{\mathcal{S}}_k$, by,

$$d(\mathbf{s},\hat{\mathbf{s}})=\frac{1}{k}\sum_{i=0}^{k-1}d(s_i,\hat{s}_i)\,. \tag{1}$$

We use the r'th power of the l_2 norm, i.e.,

$$d(\mathbf{s},\hat{\mathbf{s}})=\frac{1}{k}\|\mathbf{s}-\hat{\mathbf{s}}\|_2^r=\frac{1}{k}\left(\sum_{i=0}^{k-1}|s_i-\hat{s}_i|^2\right)^{r/2}, \tag{2}$$

as the distortion measure. This is the same distortion measured as used in [1].

Quantization is achieved by partitioning part of the source space $\mathcal{S}_k$ bounded within the region $\mathcal{R}_q$ into N disjoint and exhaustive subregions $\mathcal{Q}_i$, $i=0,\ldots,N-1$, i.e.,

$$\bigcup_i \mathcal{Q}_i=\mathcal{R}_q\cap\mathcal{S}_k \quad\text{and}\quad \mathcal{Q}_i\cap\mathcal{Q}_j=\emptyset \quad\text{if}\quad i\neq j\,. \tag{3}$$

The region $\mathcal{R}_q$ is denoted as the quantizer shaping region.

Each $\mathcal{Q}_i$ has a unique reproduction symbol $\hat{\mathbf{s}}_i$. All the source vectors $\mathbf{s}\in\mathcal{Q}_i$ are quantized to $\hat{\mathbf{s}}_i$. This is called a k-dimensional vector quantizer (VQ).

The objective in the design of a VQ is to minimize the average distortion, namely,

$$D=\sum_{i=0}^{N-1}\mathbf{E}\left[d(\mathbf{s},\hat{\mathbf{s}}_i)\,|\,\mathbf{s}\in\mathcal{Q}_i\right]\int_{\mathcal{Q}_i}P_k(\mathbf{s})\,d\mathbf{s}\,, \tag{4}$$

for fixed N where $\mathbf{E}[\,.\,|\,.\,]$ denotes the conditional expectation. To achieve this objective, the reproduction symbols $\hat{\mathbf{s}}_i$'s are selected to minimize $\mathbf{E}\left[d(\mathbf{s},\hat{\mathbf{s}}_i)\,|\,\mathbf{s}\in\mathcal{Q}_i\right]$. For the distortion measure under consideration, this results in,

$$\hat{\mathbf{s}}_i=\mathbf{E}\left[\mathbf{s}\,|\,\mathbf{s}\in\mathcal{Q}_i\right]=\int_{\mathcal{Q}_i}\mathbf{s}\,P_k(\mathbf{s})\,d\mathbf{s}\,. \tag{5}$$

The indexing (labeling) process is the assignment of the indices $i=0,\ldots,N-1$ to the partitions. In using a VQ, a source vector $\mathbf{s}\in\mathcal{Q}_i$ is specified by the index i. For example, in a transmission system the indices are transmitted to the receiver and in a digital storage media the indices are stored.

The procedure of vector quantization is composed of shaping, encoding, addressing and reconstruction. For each source vector $\mathbf{s}$, shaping is to check if $\mathbf{s}\in\mathcal{R}_q$ or not, encoding is to find the region $\mathcal{Q}_i$ such that $\mathbf{s}\in\mathcal{Q}_i$, addressing is to produce the index i if $\mathbf{s}\in\mathcal{Q}_i$ and reconstruction is the production of $\hat{\mathbf{s}}_i$ from the index i. In general, shaping has no simple rule, encoding is achieved by an exhaustive search and addressing, reconstruction are achieved by the use of a lookup table. For a VQ with a large number of partitions all these operations have high complexity.

In a VQ with a large number of partitions, the major complexity is that of the encoding. In general, this is achieved by an exhaustive search. On the other hand, partitioning the space by the Voronoi region of a lattice has the same properties as in (3). This is known as a lattice quantizer. The complexity of quantization is that of decoding of the corresponding lattice. This is much easier than the exhaustive search associated with the general VQ. The only problem is that the partitions of a lattice quantizer have the same volume. This is a natural quantizer for a uniform source.

We study three methods for the lattice quantization of a nonuniform source. The first method is based on cascading a one-dimensional compander with a multi-dimensional lattice quantizer. We propose a practical method, denoted as cluster quantization, to implement this scheme. The second method is based on selecting an appropriate boundary (shaping region) for a lattice quantizer. As a practical shaping method, we use the Voronoi region around the origin of a lattice as the boundary of the quantizer. This is denoted as a Voronoi quantizer. The group property of the Voronoi constellations is used to decrease the complexity of the shaping and encoding operations. We propose an indexing method which results in low addressing complexity. By using this indexing method in conjunction with a set of suboptimum reconstruction vectors, we obtain a system with a low reconstruction complexity. The overall complexity is in the order of the linear mappings plus the decoding of a lattice. By using binary lattices, we construct quantizers with an integral bit rate. In an extension of the Voronoi quantizer, we use a chain of the lattices to partition the space into a set of concentric subregions. The quantizer has nonequal resolutions in different subregions.

Our analysis is based on continuous approximation. This is a usual approximation in calculating the performance measures associated with a discrete set of points. This approximation is based on assuming a continuous density of points within the shaping region. Assuming a continuous approximation, the performance measures which are expressed in terms of a multiple summation, are approximated by a multiple integral over the shaping region. In general, such an integral is easier to calculate.

The performance of the proposed schemes is measured in terms of the degradation in the quantization Signal-to-Noise-Ratio (SNR) with respect to the optimum VQ. This is denoted as the performance loss (P_l).

2 Optimum Vector Quantization

Define a baseline source as a uniform source with the support $[-1/2, 1/2]$. This source has a power of $1/12$ per dimension. This is denoted as the baseline power. All the sources under consideration are normalized to have the same power.

Assuming continuous approximation, the distortion of an optimum vector quantizer for a k-dimensional source with probability P_k is equal to, [1], [2], [3],

$$D = N^{-r/k} C(k,r) \left[\int_{S_k} P_k^{k/(k+r)} ds \right]^{(k+r)/k}, \qquad (6)$$

where $N^{-r/k} C(k,r)$ is the distortion of the optimum N-point uniform quantizer applied to the baseline source. This is denoted by $D_b(N,k,r)$. The value of $C(k,r)$ in dimensionality one and two are known, [4]. The corresponding quantizers are lattice quantizer. Lower and upper bounds on $C(k,r)$ are given in [2]. In the case of a lattice quantizer and $r = 2$ (mean square distortion measure), a conjecture for the upper bound of $C(k,2)$ is given in [5]. The asymptotic value of $C(k,2)$, $k \to \infty$, is equal to $1/2\pi e$, [2].

We assume that the source symbols are independent, i.e., $P_k(\mathbf{s}) = \prod_{i=0}^{k-1} p(s_i)$ and $S_k = S^k$. Substituting in (6), we obtain,

$$D = D_b(N,k,r) \left[\int_S p^{k/(k+r)} ds \right]^{k+r}. \qquad (7)$$

It is seen that the distortion is the composed of the product of the distortion of the optimum uniform quantizer and the factor,

$$L(p,k,r) = \left[\int_S p^{k/(k+r)} ds \right]^{k+r}. \qquad (8)$$

For a uniform source $L(p,k,r) = 1$. It is easy to show that, for a Gaussian source,

$$L(p,k,r) = \left(\frac{\pi}{6}\right)^{r/2} \times \left[\frac{k+r}{k}\right]^{(k+r)/2}, \qquad (9)$$

with the limiting value of $L(p,\infty,r) = (\pi e/6)^{r/2}$. Similarly, for a Laplacian source,

$$L(p,k,r) = \left(\frac{1}{6}\right)^{r/2} \times \left[\frac{k+r}{k}\right]^{k+r}, \qquad (10)$$

with the limiting value of $L(p,\infty,r) = (e^2/6)^{r/2}$. Figure (1) shows the value of $L(p,k,2)$ for the Gaussian and Laplacian sources as a function of k. This is the degradation in the optimum quantization SNR of these sources comparing to the optimum quantization of a uniform source (for the same rate).

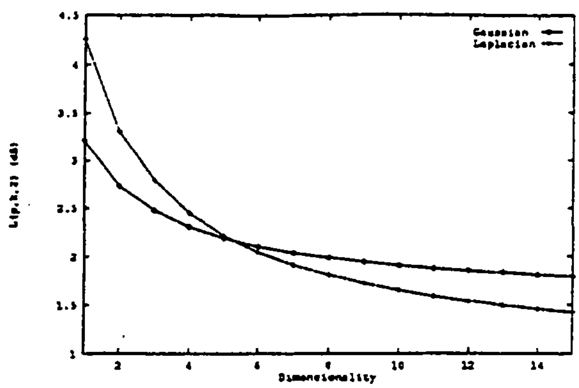

Fig. 1 $L(p,k,2)$ for the Gaussian and Laplacian sources.

In the following we discuss some properties of the function $L(p,k,r)$. We assume that the probability density p is a non-increasing function. This is the case for most of the interesting densities.

Theorem: $L(p,k,r)$ is a non-increasing function of k with the limit e^{rh_p} where h_p is the differential entropy of the source.

Proof: Using Holder's inequality [6], it can be shown that for any positive function f, $(\int f^{1/t})^t$ is a non-increasing function with respect to t. Substituting $f = p^{1/k}$ and $t = k + r$, proves the non-increasing property of the $L(p,k,r)$. To calculate the limit, from the fundamental theorems of source coding, we know that in the limit of $k \to \infty$, the density P_k tends to the constant value e^{-kh_p} over a region of volume e^{kh_p} and tends to zero outside of this region. Substituting in (6), proves the desired result.

This theorem shows that increasing the dimensionality results in better performance. This theorem also means that in the limit, a Gaussian source is the most difficult source to quantize (a Gaussian source maximizes the h_p).

The proof of the second theorem is based on the following lemma.

Lemma: The differential entropy of a monotonically decreasing density is positive.

Proof: It is easy to show that the set of the monotonically decreasing probability densities with constant average power constitute a convex region. The baseline source corresponds to a point on the boundary of this region. Also, differential entropy is a convex cap function of the density function and has a unique maximum point over the convex region (the maximum point corresponds to a Gaussian density). Consequently, the entropy increases as we move from the point corresponding to the uniform density toward the interior part of the region. Considering that the differential entropy of the uniform (baseline) source is zero, the desired result is proved.

Theorem: $L(p,k,r)$ is greater than or equal to one where the equality is satisfied for a uniform source.

Proof: Combining the lemma and the first theorem means that $L(p,k,r)$ is a decreasing function with a positive limit and consequently is positive for the whole range of k. By the direct substitution in (8), it is easy to verify that for a uniform source $L(p,k,r) = 1$.

This theorem means that a uniform source is the easiest source to quantize.

The following theorem gives a lowerbound to the performance of a uniform quantizer in conjunction with a nonuniform source.

Theorem: Let's $\mathcal{Q}_i$'s be a set of the congruent partitions of the region $\mathcal{A}$. Let's G, H denote the set of density functions which are nonincreasing, nondecreasing function of the l_2 norm and have a constant measure, M, over $\mathcal{A}$. The distortion associated with the density function $P_k \in G \cup H$, given in (4), is maximized if P_k is constant over $\mathcal{A}$.

Proof: Consider a density $P_k \in G \cup H$. Assume that the measure of P_k over $\mathcal{Q}_i$, $i = 0, \ldots, N-1$ is equal to M_i where $\sum_i M_i = M$. The total distortion of P_k is equal to, $D = \sum_i D_i$ where,

$$D_i = \int_{\mathcal{Q}_i} \|\mathbf{s} - \mathbf{E}[\mathbf{s} \mid \mathbf{s} \in \mathcal{Q}_i]\|_2^r P_k(\mathbf{s})\, d\mathbf{s}. \tag{11}$$

Let's $G' \subset G$ and $H' \subset H$ denotes the set of the nonincreasing, nondecreasing densities with the measure M_i, $i = 0, \ldots, N-1$ over $\mathcal{Q}_i$. We first show that for a density $P_k' \in G' \cup H'$, the distortion D_i, $i = 0, \ldots, N-1$ in (11) is maximized if P_k' is constant over $\mathcal{Q}_i$. It is easy to show that: (i) The sets G' and H' are convex sets with a boundary corresponding to a constant density. (ii) The set of the densities with the measure M_i, $i = 0, \ldots, N-1$ over $\mathcal{Q}_i$ is a convex set. This is denoted as F. The distortion given in (11) is a convex $\cap$ function of P_k and has a unique maximum over each of the convex sets F, G' and H'. The maximum points over G' and H' can not be located inside of the regions because in this case, as $G' \subset F$ and $H' \subset F$ we should have two maximum over the region F. This means that the maximum points over G' and H' are located on their boundary (corresponding to a uniform density). On the other hand, summing the D_i's in (11) and considering that $\mathcal{Q}_i$'s are congruent, it is easy to show that the distortion of a density which is constant over each $\mathcal{Q}_i$ and has a constant total measure over $\mathcal{A}$ is a constant value independent of the individual measures. This means that such a density has the same distortion as a density which is constant over the whole $\mathcal{A}$. This completes the proof.

3 One-dimensional Companding

Companding is a method of implementing a nonuniform quantization scheme using a uniform quantizer. This is based on the cascade of a zero memory nonlinearity, denoted as the compander, followed by a uniform quantizer and then followed by the inverse of the first nonlinearity. The optimum compander for dimensionality one is known, [7]. The conditions for the existence of the optimum compander in higher dimensionalities are hard to satisfy, [8].

We assume that the compander is equal to the k-fold cartesian product of a one-dimensional compander. In this case, it can be shown the average distortion is equal to,

$$D = N^{-r/k} C(k, r) L(p, 1, r). \tag{12}$$

The performance loss with respect to the optimum quantizer is equal to,

$$P_l = \frac{L(p, 1, r)}{L(p, k, r)}. \tag{13}$$

Figure (2) shows the P_l for the Gaussian and Laplacian sources as a function of k, $r = 2$. Using (9) and (10), it is easy to show that the asymptotic value of P_l for the Gaussian source is equal to, 2.81 dB and for the Laplacian source is equal to, 5.63 dB.

The gain with respect to the optimum one-dimensional quantizer is, $C(1, r)/C(k, r)$. For example, for $r = 2$ ($C(1, 2) = 1/12$, [4]) in dimensionality $k = 8$ using lattice E_8, $C(8, 2) = 0.071682$, [4], results in a gain of 0.65 dB and in dimensionality $k = 24$ using lattice Λ_{24}, $C(24, 2) = 0.065771$, [4], results in a gain of 1.0 dB. The asymptotic gain for $r = 2$ and $k \to \infty$ is equal to 1.53 dB, $C(\infty, 2) = 1/2\pi e$, [2].

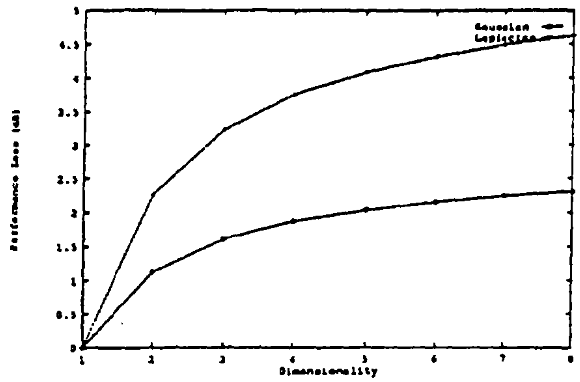

Fig. 2 The P_l of the cluster quantizer for the Gaussian and Laplacian sources, $r = 2$.

Later, we will discuss a practical method denoted as the Cluster Quantization to implement this scheme. Before starting the main discussion, we first talk about the concept of the Voronoi constellations, [9], [10].

4 Lattices, Voronoi constellation

A k-dimensional lattice Λ is a subset of points of $\mathbf{R}^k$ which form a group under ordinary vector addition. The points of a lattice can be written in form,

$$\mathbf{c} = \sum_{i=0}^{p-1} a_i \mathbf{u}_i, \tag{14}$$

where $p \geq k$, a_i's are integer numbers and the set of the k-dimensional vectors $\mathbf{u}_i$'s are a set of the generators for Λ. A subgroup Λ_s of Λ, denoted as Λ/Λ_s, is called a sublattice. A sublattice Λ_s partitions Λ into $|\Lambda/\Lambda_s|$ cosets of Λ_s. The set of the cosets form a group under addition modulo Λ_s. This is denoted as the quotient group. A lattice Λ is called binary if $Z^k/\Lambda/2^l Z^k$ is a valid partition chain for some integer l. For binary lattices, $|\Lambda/\Lambda_s|$ is an integral power of two.

Let's $\mathcal{V}(\Lambda)$ denotes the Voronoi region around the origin of the lattice Λ. A Voronoi constellation based on the partition Λ/Λ_s, denoted as $C(\Lambda/\Lambda_s)$, is composed of the points of Λ located inside of $\mathcal{V}(\Lambda_s)$. The number of such points is equal to $|\Lambda/\Lambda_s|$. The points of $C(\Lambda/\Lambda_s)$ are the coset leaders of the partition and form a group under vector addition modulo Λ_s. To uniquely specify the Voronoi constellation $C(\Lambda/\Lambda_s)$, one should also solve the problem of ties. This occurs when some of the points of Λ are located on the boundary of Λ_s, [10].

In the case of binary lattices, we have,

$$\Lambda = \Lambda_s + \mathbf{G}\mathbf{a}, \tag{15}$$

where $\mathbf{a}$ is a binary m-tuple ($|\Lambda/\Lambda_s| = 2^m$) and $\mathbf{G}$ is an $k \times m$ binary matrix. The points of the Voronoi constellation $C(\Lambda/\Lambda_s)$ can be written in the form, [10],

$$\mathbf{c} = [\mathbf{G}\mathbf{a}]_{\mathrm{mod}(\Lambda_s)}. \tag{16}$$

The binary vector $\mathbf{a}$ can be recovered from $\mathbf{c}$ using,

$$\mathbf{a} = [\mathbf{G}^{-1}\mathbf{c}]_{\mathrm{mod}(2)}. \tag{17}$$

4.1 Cluster Quantization

We employ a fine nonuniform quantizer with 2^m points along each dimension. Then the partitions of this quantizer are mapped to the points of the Voronoi constellation $C(Z/2^m Z)$. This is the set of points from the integer lattice bounded within the region $[-2^{m-1}, 2^{m-1}]$. The k-fold cartesian product space will be mapped to the points of the cubic constellation $C(Z^k/2^m Z^k)$. The total rate of the resulting quantizer is equal to km bits. Now, a lattice Λ_q, where $Z^k/\Lambda_q/2^m Z^k$ is a valid partition chain, is used to partition the cubic constellation into $2^{m_1} = |Z^k/\Lambda_q|$ clusters each containing $2^{m_2} = |\Lambda_q/2^m Z^k|$ points, $m_1 + m_2 = km$. Each cluster determines one of the final partitions. As a result of this grouping, the rate decreases by m_2 bits resulting in a rate of m_1 bits for the final quantizer. To have consistency with the continuous approximation, the lattice Λ_q is selected such that m_2 is a large number.

Cluster Quantization can be also considered as a method for nonuniform partitioning of the space while the encoding has a low complexity. Each partition is the union of some hypercubes. For larger values of m_2, the hypercubes are smaller which results in partitions with a smoother boundary.

5 Quantizer Shaping

In this method, region $\mathcal{R}_q$ is selected to act as an interface matching the probability density of the source to the (desired) uniform density in a lattice quantizer. In general, a uniform density of points within a region $\mathcal{R}_q$ induces certain marginal density along the space dimensions. Define a symmetrical region as a region for which the marginal densities are the same. It can be shown that a convex region results in a monotonically decreasing marginal density. We assume symmetrical convex shaping regions. For a region which is not a hypercube, the marginal densities are dependent. For any region, as the dimensionality tends to infinity, the marginal densities become independent of each other.

From the fundamental theorems of the information theory, we know that for any probability density p there exist an infinite dimensional region $\mathcal{R}^\infty$ such that a uniform density within that region results in an independent density equal to p along each dimension. It can be shown that a convex region $\mathcal{R}^\infty$ corresponds to a monotonically decreasing p and vice versa. For a Gaussian source, $\mathcal{R}^\infty$ is a sphere and for a Laplacian source $\mathcal{R}^\infty$ is a pyramid.

Consider a source s with the baseline power and with a non-increasing probability density function p of support $\mathcal{R}$. Obviously, $\mathcal{R} \geq 1$, with equality if p is uniform. The k-fold cartesian product of the source s, has the support $\mathcal{R}_k = \mathcal{R}^k$, namely a hypercube of edge length $\mathcal{R}$.

If the whole region $\mathcal{R}_k$ is quantized by a uniform quantizer, it is easy to show that the performance loss with respect to the optimum scheme is equal to,

$$P_l = \frac{[V(\mathcal{R}_k)]^{r/k}}{L(p,k,r)}, \tag{18}$$

where $V(\ .\ )$ denotes the volume. The density P_k has support on the region, $\mathcal{R}_k = \mathcal{R}^k$, $V(\mathcal{R}_k) = [V(\mathcal{R})]^k$. Substituting in (18) results in,

$$P_l = \frac{[V(\mathcal{R})]^r}{L(p,k,r)}. \tag{19}$$

As $k \to \infty$, $\mathcal{R}_k$ tends to a subset of $\mathcal{R}^k$ and $[V(\mathcal{R}_k)]^{r/k} \to e^{rh_r}$. We also showed earlier that as $k \to \infty$, $L(p,k,r) \to e^{rh_r}$. This means that the asymptotic value of P_l in (18) is equal to zero dB. This reflects the fact that in an infinite dimensional space, the optimum quantizer for any source is uniform. In extending this idea to a finite dimensional space, we use a region $\mathcal{R}_q \subset \mathcal{R}^k$ to select the source samples to be quantized.

It should be mentioned that in an infinite-dimensional space the volume of a solid is concentrated on its surface. This fact provides a connection between our way of looking at geometrical source coding and the one in [11], [12] and [13].

In the following, we study two methods for the geometrical source coding in a finite dimensional space.

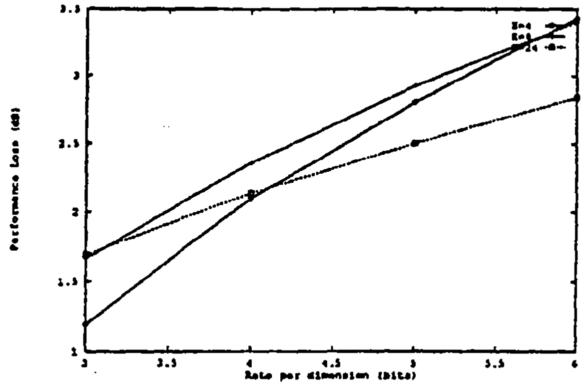

Fig. 3 The P_l of the shaping by rounding to the nearest point for a Gaussian source, $r = 2$.

5.1 Shaping by Rounding to the Nearest Point

In this method, a source vector $s \notin \mathcal{R}_q$ is quantized to the vector $A(s) \in \mathcal{R}_q$ which minimizes $\|s - A(s)\|_2^r$. It is easy to show that the total distortion is equal to,

$$D = N^{-r/k} C(k,r)\,[V(\mathcal{R}_q)]^{r/k} \int_{\mathcal{R}_q} P_k\, ds + \int_{s \notin \mathcal{R}_q} \|s - A(s)\|_2^r P_k\, ds. \tag{20}$$

The region $\mathcal{R}_q$ is selected to minimize the total distortion.

For each source vector s, we first check if s belongs to $\mathcal{R}_q$ or not. If $s \notin \mathcal{R}_q$, it is rounded to the nearest point on the surface of $\mathcal{R}_q$. After that the quantizer lattice is decoded.

It is easy to verify that for a Gaussian source, the optimum region is a hypersphere. Denoting the radius of the hypersphere by a, the total distortion is found as,

$$D(a) = N^{-r/k} C(k,r)(\pi a^2)^{r/2} \{\Gamma[(k/2)+1]\}^{-r/k}\, Y(a) + X(a), \tag{21}$$

where $\Gamma(\ .\ )$ denotes the gamma function,

$$Y(a) = \frac{2\,(6)^{k/2}}{\Gamma(k/2)} \int_0^a x^{k-1} \exp(-6x^2)\, dx, \tag{22}$$

and,

$$X(a) = \frac{2\,(6)^{k/2}}{\Gamma(k/2)} \int_a^\infty (x-a)^r x^{k-1} \exp(-6x^2)\, dx. \tag{23}$$

The parameter a is selected to minimize the total distortion. The corresponding P_l for $r = 2$ and for dimensionality $k = 4, 8, 24$ as a function of the rate per dimension, $(1/k)\log N$, is shown in Fig. (3).

It is seen that the geometrical source coding is more effective for a lower number of quantizer partitions, N. However, as we are using continuous approximation, by decreasing N, the validity of the results decreases.

5.2 Shaping by Clipping

In this case a source vector $s \notin \mathcal{R}_q$ is clipped to zero. This scheme results in some degradation with respect to the previous method but it is easier to implement. The total distortion is equal to,

$$D_t = N^{-r/k} C(k,r)\,[V(\mathcal{R}_q)]^{r/k} \int_{\mathcal{R}_q} P_k\, ds + \int_{s \notin \mathcal{R}_q} \|s\|_2^r P_k\, ds. \tag{24}$$

It is easy to verify that the surface of the optimum $\mathcal{R}_q$ is composed of the points with $P_k(\mathbf{s})\|\mathbf{s}\|_2^2 = c$. For a fixed number of quantizer regions, changing the constant c determines a scale factor to be applied to the quantization lattice. This provides a tradeoff between the approximation and the clipping errors. The optimum value of c is selected to minimize the total distortion given in (24). In this method, the test for $\mathbf{s} \in \mathcal{R}_q$ is easily achieved by checking if $P_k(\mathbf{s})\|\mathbf{s}\|_2^2 \le c$ or not.

6 Voronoi Quantizer

A set of lattice points $\mathbf{c}_i$ partition the space into a set of the congruent Voronoi regions. This set of the partitions in conjunction with a uniform source results in the reproduction vectors, $\hat{\mathbf{s}}_i = \mathbf{c}_i$, $\forall i$ corresponding to a local optimum VQ. A partition chain Λ_q/Λ_s results in a VQ with $N = |\Lambda_q/\Lambda_s|$, $\mathcal{R}_q = \Lambda_s$, $Q_i = \mathcal{V}(\Lambda_q)$, $\forall i$ and $\mathbf{c}_i \in C(\Lambda_q/\Lambda_s)$. This is denoted as a Voronoi quantizer. In this case, shaping is achieved by the decoding of Λ_s and encoding is achieved by the decoding of Λ_q.

In a Voronoi quantizer, the effectiveness of a shaping lattice depends on the statistic of the source. As there is no cost (except for the decoding complexity) associated with shaping, we always use the best shaping lattice. In general, the Voronoi region of a lattice with lower $C(k, 2)$ is more circular. Such a lattice is appropriate in quantizing a Gaussian source. It seems that for a Laplacian source, the Voronoi region of the lattice D_N^* is more similar to a pyramid and achieves better performance.

In a Voronoi quantizer, the indexing has an important effect on the addressing and reconstruction complexities. Considering (16), if the partition centered at $\mathbf{c}$ is indexed by $\mathbf{a}$, (17) provides an easy way for the addressing. If instead of the optimum $\hat{\mathbf{s}}_i$'s in (5), we use $\hat{\mathbf{s}}_i = \mathbf{c}_i$, (16) provides an easy way for the decoding. Obviously, this results in some degradation.

7 Nonuniform Voronoi Quantizer

In the following, we use a chain of the lattices to partition the space into a set of concentric partitions of different resolutions. Recently, Joeng and Gibson in [14] and [15] have proposed the structure of a multi-dimension compressor. They use a set of the concentric radial bands (with respect to l_1 or l_2 norm), to partition the space. The integer lattice (Z^k) is used for the quantization. The density of the lattice points is increased by a constant multiplicative factor as we pass from one radial band to the next one closer to the origin. The main difference between their work and our method is that here, instead of the concentric radial bands, we use a set of the nested lattice Voronoi regions to partition the space. Their method makes use of the optimum shaping region (for a given source statistics). Obviously, from the shaping point of view, their method is superior. However, as we will see later, in our method, the group property of the lattices is the source of a number of useful properties which decrease the complexity of the quantization.

Consider the lattice partition chain, $\Lambda_q^0/\Lambda_q^1, \ldots \Lambda_q^m/\Lambda_q^{m+1}$. For each partition $\Lambda_q^{m-i}/\Lambda_q^{m-i+1}$, $i = 0, \ldots, m$, a copy of $\mathcal{V}(\Lambda_q^{m-i})$ is centered around each point of $C(\Lambda_q^{m-i}/\Lambda_q^{m-i+1})$. This results in a set of partitions such that the resolution increases in m steps as we get closer to the origin.

To calculate the number of the total partitions, N, we know that in the i'th step, the number of the partitions increases by the factor, $|\Lambda_q^{m-i}/\Lambda_q^{m-i+1}| - 1$. Consequently,

$$N = \sum_{i=0}^{m} |\Lambda_q^i/\Lambda_q^{i+1}| - m. \tag{25}$$

From now on, we assume the chain of the binary lattices, $\Lambda_q/2\Lambda_q, \ldots, 2^m\Lambda_q/2^{m+1}\Lambda_q$. For this chain we have $|2^i\Lambda_q/2^{i+1}\Lambda_q| = 2^k$, k is the space dimensionality, and (25) reduces to,

$$N = (m+1)2^k - m. \tag{26}$$

In this case, to increase the quantization resolution, we can use the lattice $2^{i-p}\Lambda_q$ to partition the Voronoi regions of $2^i\Lambda_q$ into 2^{pk} congruent subregions. As a result of this subpartitioning, the quantizer bit rate increases by p bits per dimension.

7.1 Indexing

We assume that the partitions of the nonuniform Voronoi quantizer are indexed by $R_t = \lceil \log N \rceil$ binary digits. The redundancy is equal to $r = \lceil \log N \rceil - \log_2 N$. In the case of the binary lattices, if we select $m + 1 = 2^u$, using (26) and the inequality $\log_e z \ge 1 - (1/z)$, it can be shown that,

$$R_t = k + u, \tag{27}$$

and the redundancy satisfies,

$$r \le \left(\frac{m+1}{m} 2^k - 1\right)^{-1} \log_2(e). \tag{28}$$

This usually results in a negligible redundancy. The rate per dimension is equal to $R = 1 + (u/k)$.

For the chain of binary lattices, $\Lambda_q/2\Lambda_q/\ldots/2^m\Lambda_q/2^{m+1}\Lambda_q$, define the region A_i, $i = 0, \ldots, m$, as the set of the $\mathcal{V}(2^i\Lambda_q)$ centered around the points,

$$A_i = \left[\mathbf{c} : \mathbf{c} \notin C(2^{i-1}\Lambda_q/2^i\Lambda_q), \text{ and } \mathbf{c} \in C(2^i\Lambda_q/2^{i+1}\Lambda_q)\right]. \tag{29}$$

Using (16), the points of A_i can be written as,

$$\mathbf{c} = 2^i\,[\mathbf{G}\mathbf{a}]_{\mathrm{mod}(2\Lambda_q)}. \tag{30}$$

Using (30), the points of A_i are indexed by concatenating the binary representation of i (with $\lceil \log_2(m+1) \rceil$ binary digits) and the k-dimensional binary vector $\mathbf{a}$. This method of indexing, denoted as the natural indexing, will be very efficient if $m+1$ is an integral power of two.

7.2 Shaping, Encoding

Given a source vector $\mathbf{s}$, shaping and encoding is achieved by decoding the set of lattices $\Lambda_q, \Lambda_q^2, \ldots, 2^{m+1}\Lambda_q$. This is used to find the set A_i defined in (29) and the nearest point $\mathbf{c} \in A_i$ to $\mathbf{s}$. A source vector $\mathbf{s} \notin \mathcal{V}(2^{m+1}\Lambda_q)$ is truncated to zero.

7.3 Addressing

For a given set A_i and point $\mathbf{c} \in A_i$, the binary vector $\mathbf{a}$ is calculated by substituting $\mathbf{c}$ in,

$$\mathbf{a} = [2^{-i}\mathbf{G}^{-1}\mathbf{c}]_{\mathrm{mod}(2)}. \tag{31}$$

This is concatenated with the binary representation of the index i to achieve the addressing.

7.4 Decoding

If the optimum reconstruction vectors given in (5) are used, the decoding is achieved by a lookup table. However, this may not be practical for large values of N.

In a suboptimum method, the reconstruction vectors, $\hat{\mathbf{s}}_i$'s, are selected as the centers of the Voronoi regions. This is useful in conjunction with the natural indexing. In this case, the two parts of the label are used in (30) to produce the reconstruction vector.

8 Examples

An example of a nonuniform Voronoi quantizer based on the lattice partition chain $Z/2Z/2^2Z/2^3Z/2^4Z$ is shown in Fig. (4). Similar examples based on the partition chain $Z^2/2Z^2/2^2Z^2/2^3Z^2/2^4Z^2$ and $\Re Z^2/2\Re Z^2/2^2\Re Z^2/2^3\Re Z^2$ are shown in Figs. (5) and (6), $\Re$ is the rotational operator, [4].

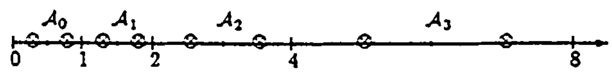

Fig. 4 The positive part of the one-dimensional nonuniform Voronoi quantizer based on the integer lattice, $u = 2$ ($m = 3$) and $p = 1$. The reconstruction levels are shown by the ⊗ sign.

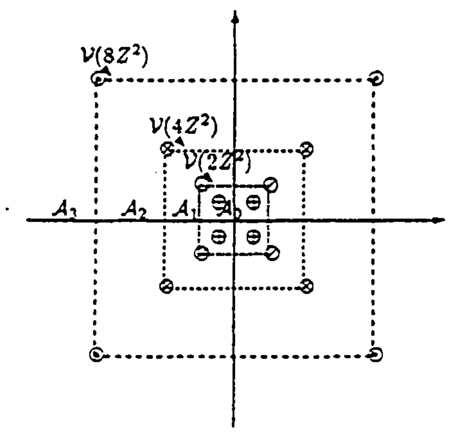

Fig. 5 The regions $\mathcal{A}_i$ for a two-dimensional nonuniform Voronoi quantizer based on the integer lattice, $u = 2$ ($m = 3$) and $p = 0$. The region $\mathcal{V}(16Z^2)$ is not shown.

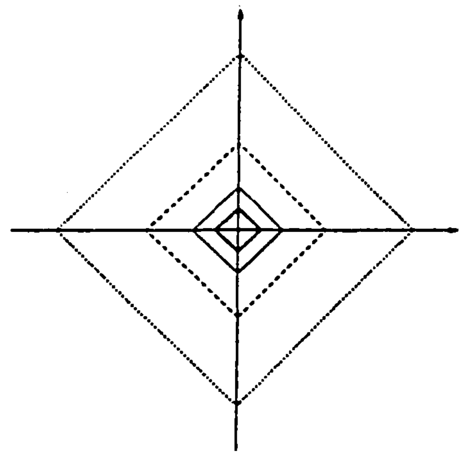

Fig. 6 The regions $\mathcal{A}_i$ for a two-dimensional nonuniform Voronoi quantizer based on the partition chain $\Re Z^2/\ldots/2^3\Re Z^2$. This structure is specially useful in conjunction with a Laplacian source.

9 Summary and Conclusions:

We have proposed some practical methods to apply a lattice-based *VQ* to a nonuniform source. The cluster quantization, uses the k-fold cartesian product of a one-dimensional compander together with a lattice quantizer. This scheme has an asymptotic gain of 1.53 dB with respect to the optimum one-dimensional quantizer. The quantizer shaping, is based on selecting an appropriate boundary for the quantizer. This scheme is asymptotically optimum. The Voronoi quantizer is based on using the points of a Voronoi constellation to partition the space. The group property of the lattice points is used to decrease the complexity of the operations. In an externsion of this scheme, we used a chain of binary lattices to partition the space into a set of concentric partitions such that the quantization resolution increases in m steps as we get closer to the origin.

References

[1] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 373–380, July 1979.

[2] P. L. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Inform. Theory*, vol. IT-29, p. 279, March 1983.

[3] J. A. Bucklew, "Properties of minimum mean squared error block quantizers," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 105–107, January 1982.

[4] J. H. Conway and N. J. A. Sloane, "Sphere packings, Lattices and groups." Springer–Verlag, 1988.

[5] J. H. Conway and N. J. A. Sloane, "A lower bound on the average error of vector quantizers," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 106–109, January 1985.

[6] R. G. Gallager, *Information theory and reliable communication*, John Wiley & Sons, New York, NY, 1968.

[7] B. Smith, "Instantaneous companding of quantized signals," *Bell Syst. Tech. J.*, vol. 52, pp. 653–709, May 1957.

[8] J. A. Bucklew, "A note on optimal multidimensional companders," *IEEE Trans. Inform. Theory*, vol. IT-29, p. 279, March 1983.

[9] J. H. Conway and N. J. A. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 820–824, November 1983.

[10] G. D. Forney, Jr., "Multidimensional constellations –Part II: Voronoi constellations," *IEEE J. Select. Areas Commun.*, vol. SAC-7, pp. 941-958, August 1989.

[11] D. J. Sakrison, "A geometrical treatment of the source encoding of a Gaussian random variable," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 481–486, May 1968.

[12] T. R. Fischer, "Geometric source coding of vector quantization," variable," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 137–145, January 1989.

[13] T. R. Fischer, "A pyramid vector quantizer," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 568–583, July 1986.

[14] D. G. Jeong and J. D. Gibson, "Image coding with uniform and piecewise uniform vector quantizers," *IEEE Global Telecomm. Conf.* pp. 81–84, Phoenix, Arizona, December 1991.

[15] D. G. Jeong and J. D. Gibson, "Piecewise linear vector quantizers," *IEEE Int. Symp. on Inform. Theory*, p. 314, Budapest, Hungary, June 24–28 1991. pp. 81–84, December 1991.