EMBEDDED VORONOI CODES FOR SUCCESSIVE REFINEMENT LATTICE VECTOR QUANTIZATION

Guillaume Fuchs

Fraunhofer Institut Integrierte Schaltungen
Am Wolfsmantel 33, 91058 Erlangen, Germany
email: amm-info@iis.fraunhofer.de

ABSTRACT

Lattice Vector Quantization (LVQ) is an interesting tool in source coding which can take advantage of a higher dimension than the scalar case while overcoming complexity limitations of conventional vector quantization. However, the high dimension and the relatively complex indexing of the codebooks make LVQ often unsuitable for getting a successive refinement of the source. For addressing this problem, the paper proposes a new class of LVQ called the embedded Voronoi codes. The new codes can gradually describe the source with a granularity of 1 bit/dimension by properly combining differently scaled Voronoi codes. A rate-distortion evaluation for a Gaussian source shows that the embedding of the codes comes at a minimal cost at low bit-rates while preserving LVQ advantages over scalar quantization.

Index Terms— Lattice Vector Quantization, Embedded Quantization

1. INTRODUCTION

Lattice Vector Quantization (LVQ) is a powerful tool in source coding as it inherits advantages over scalar Quantization (SQ) from Vector Quantization (VQ) while being less complex than conventional stochastic VQ. In LVQ, the codebook is defined as a finite subset of a regular point lattice. Due to the highly regular structure of lattices, the search and indexing of the nearest neighbor within the codebook can be algebraic. In this way, codebooks require no or almost no training and storage, and the computational complexity is limited. LVQ can therefore exploit higher dimensions and codebook sizes than the stochastic VQ allows in practical usages. LVQ was successfully adopted in image coding [1, 2, 3, 4] as well as in audio coding [5, 6].

Embedded Quantization is used for obtaining a successive refinement of the source information and finds applications in progressive transmission of the coded data. In the scalar case, embedded quantizers can be obtained easily by using the bit-plane coding principle. Such a coding is used in the image coding schemes EZW [7] and SPIHT [8], but also in audio coding [9, 10]. In the vector case, and moreover for LVQ, the problem is more complex. A solution for designing a successive refinement LVQ was already proposed in [11]. The method defines a constraint in the design of a LVQ which ensures to obtain an optimal successive refinement. However, the method requires the computation of specific geometric shapes and the design of singular entropy codes for each desired codebook. In that sense, the method is complex and cannot be easily applied for any lattices and dimensions.

Part of the work was performed at the University of Sherbrooke, Canada

The aim of this work is to define and study the properties of a new class of successive refinement LVQ, which is generic to any lattices. The new embedded quantization called embedded Voronoi codes exploits the remarkable properties of the Voronoi codes by decomposing a point of the lattice in a multitude of Voronoi codes scaled at the powers of 2. The bitstream is then suitable for a progressive transmission with a granularity as fine as 1 bit/dimension. Similar codes were used in a scalable audio coding scheme [12]. The present paper defines the codes more accurately and introduces their properties. Further, a theoretical evaluation in terms of rate-distortion is presented for a memoryless Gaussian source. The performance of the embedded Voronoi codes is compared to conventional bit-plane coding and Voronoi codes.

The paper is organized as follows: in the two first sections, some background about LVQ and embedded quantization is given. The new embedded quantization is introduced in section 4, followed by a results section evaluating its performance. The paper ends with a conclusion and a section dedicated to positioning the presented study in relation to prior work.

2. LATTICE VQ

2.1. Definition

A lattice Λ in N-dimensional space $\Re^N$, is defined as a set of points obtained by an integer combination of a linearly independent set of vectors $\{v_1, v_2, \ldots, v_n\}$:

$$\Lambda = \{y | y = \sum_{i=1}^{N} k_i v_i | k_i \in \mathbb{Z}, v_i \in \Re^N, \forall i = 1, 2, \ldots, N\} \quad (1)$$

Each point y of the lattice Λ is associated with a Voronoi region. The Voronoi region gathers all the points in $\Re^N$ closer to y than to any other points of the lattice.

$$\Omega(\Lambda, y) = \{x \in \Re^N | \|x - y\|^2 \leq \|x - z\|^2, \forall z \in \Lambda\}, y \in \Lambda \quad (2)$$

Due to the regular structure of lattices, all Voronoi regions in a lattice are simple translations of the zero-centered Voronoi region $\Omega_0(\Lambda)$, which can be expressed by:

$$\Omega_0(\Lambda) = \Omega(\Lambda, y) - y, \forall y \in \Lambda \quad (3)$$

This means that the Voronoi regions around lattice points are congruent polytopes, and form a regular tessellation partitioning of the infinite space $\Re^N$. The lattices can be ranked by their space filling advantage over the lattice $\mathbb{Z}^N$ representing the partition of a uniform scalar quantization [13]. The space filling is given by the normalized second moment of Voronoi region. According to this criterion, the hexagonal lattice A_2 is optimal for dimension 2 and will be used in the following for illustration purposes. For dimension 8, the Gosset lattice RE_8 is optimal and is adopted for evaluating the performance of the proposed successive refinement LVQ in the results section.

2.2. Voronoi codes

The Lattice Vector Quantization (LVQ) is a vector quantization whose codebook is a subset of an infinite lattice Λ. A codebook C of a LVQ can be defined as an intersection of a lattice Λ and a shape of support S.

$$C = \Lambda \cap S \tag{4}$$

The points outside the shape of support are rejected from the codebook. The lattice shaping is of major importance for the performance of the resulting LVQ. To get an optimal quantization the lattice should be shaped by a contour of constant probability density of the source [14]. For memoryless uniform, Gaussian, and Laplacian distributions the iso-probability contours are an N-cube, an N-sphere, and an N-octahedron, respectively.

The Voronoi codes introduced by Conway and Sloane [15] are a convenient way of shaping and at the same time indexing an LVQ codebook. The Voronoi codes are generated by shaping the lattice Λ by an integer multiple m of the Voronoi region of the same lattice Λ translated by a fixed vector a:

$$V(\Lambda, m\Lambda, a) = \Lambda \cap (\Omega_0(m\Lambda) + a) \tag{5}$$

a is a translation offset in $\Re^N$ that ensures that no point of Λ falls on the boundary of $\Omega_0(m\Lambda) + a$. An illustration of the Voronoi code is in A_2 given in Fig. 1. The code size of $V(\Lambda, m\Lambda, a)$ is m^N and can be indexed by efficient algorithms [15]. Usually, m is a power of 2 which leads to codebooks needing a rate of integer bits. However, the Voronoi codes' truncation is not always optimized for the source distribution. The Voronoi region is close to an N-sphere, especially for high dimensions, and can be considered quasi-optimal for a memoryless Gaussian source.

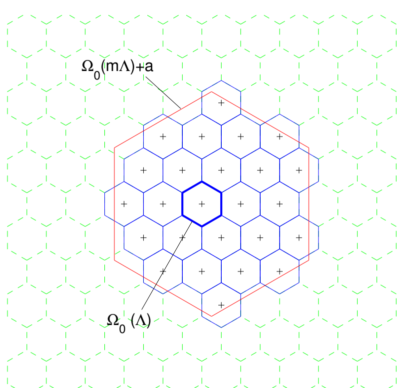


Fig. 1. Codebook defined by the Voronoi code $V(A_2, 5A_2, 0.25)$.

3. EMBEDDED QUANTIZATION

3.1. Definition

An embedded quantization can be defined as a set of K approximating functions $x \to Q_k[x], k \in \{0, .., K-1\}$, where each $Q_k[x]$ refines the representation of x given by previous approximation $Q_{k-1}[x]$. Embedded quantization allows to produce a scalable bitstream, in which decoding a subset of the coded data is sufficient for getting an approximation of x from $Q_k[x]$. This approximation is coarser than the approximation given by Q_{K-1} which can be obtained by decoding the entire coded data. The representation of the source is then sent in multiple stages, where in the k-th stage the information difference between $Q_k[x]$ and $Q_{k-1}[x]$ is transmitted.

3.2. Bit-plane coding

Uniform scalar quantization is well suited for getting an embedded quantization. It is usually achieved by decomposing N successive scalar quantized values into binary format. A bit-plane is formed by gathering the N bits having the same weight in the binary representations. The number K of bit-planes needed for describing the N-dimensional vector y of the quantized values is given by:

$$K = \log_2(\max_{n \in \{0,\ldots,N\}}(abs(y_n)) + 1) \tag{6}$$

The number K also needs to be transmitted for every vector and can be entropy coded. In case $K > 0$, the sign of each element of y is transmitted followed by the K bit planes sent sequentially from the most significant to the least significant bit-plane. Starting from the least significant bits, the bitstream can be truncated with a granularity of 1 bit and decoded with a coarser approximation.

4. EMBEDDED VORONOI CODES

In this study we define a new embedded quantization called embedded Voronoi codes based on LVQ and particularly the Voronoi codes. This section gives its definition and properties.

4.1. Definition

The embedded Voronoi codes $C^{(r)}$ of order r in the lattice Λ are defined as a sum of r Voronoi codes $V(\Lambda, 2\Lambda, a_i), i \in \{0 \ldots r\}$ scaled by factors 2^i:

$$C^{(r)} = 2^r V(\Lambda, 2\Lambda, a_r) + \ldots + 2V(\Lambda, 2\Lambda, a_1) + V(\Lambda, 2\Lambda, a_0) \tag{7}$$

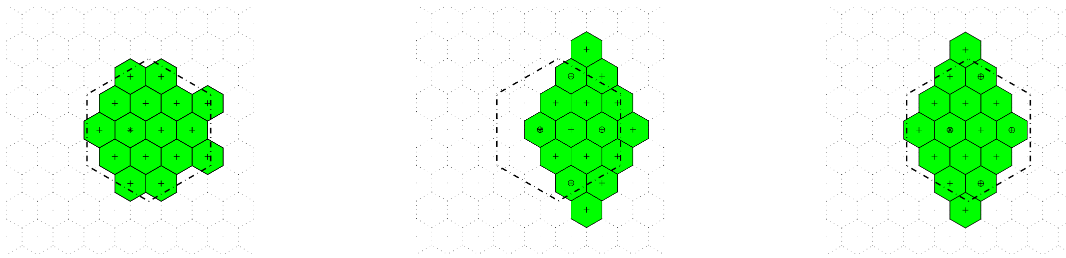
$C^{(r)}$ forms a family of codebooks in Λ of size $2^{N(r+1)}$. It can be easily shown that $C^{(r)}$ are subsets of Λ. This is deduced from the self-similarity property of lattices which can be summarized by the following formula [16]:

$$\Lambda = m\Lambda + V(\Lambda, m\Lambda, a) = \bigcup_{c \in \Lambda, v \in V(\Lambda, m\Lambda, a)} mc + v \tag{8}$$

$C^{(r)}$ requires a bit-rate of $r + 1$ bits per dimension and can achieve a multi-rate quantization with a resolution of 1 bit per dimension. The minimal order r to use for coding a point y of the lattice Λ can be found by iteratively searching for $y \in C^{(r)}$. Moreover, the embedded Voronoi codes are suitable for a successive refinement as it decomposes y into $r + 1$ vectors:

$$y = \sum_{i=0}^{r} 2^i v_i \tag{9}$$

where v_i are codevectors issued from the Voronoi codes $V(\Lambda, 2\Lambda, a_i)$ and each of them can be coded on N bits. The first codevector $2^r v_r$ represents the coarsest approximation of y and can be obtained by decoding the first N bits. The following codevectors



(a) Voronoi codes (b) Embedded Voronoi codes ($a_i = a, \forall i$) (c) Embedded Voronoi codes such as (11)

Fig. 3. 4 bits codebooks issued from Voronoi codes and embedded Voronoi codes .

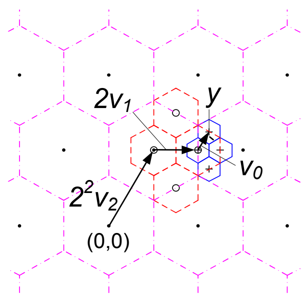


Fig. 2. Decomposition of a lattice point y from A_2 by an embedded Voronoi code of order 2.

$2^i v_i, \forall i \in \{r-1 \dots 0\}$ are the successive refinements. The approximation of y can be refined r times by decoding each time N additional bits, i.e. with a decoding granularity of 1 bit per dimension. Fig. 2 illustrates the decomposition in the lattice A_2 obtained by an embedded Voronoi code of order 2.

In the next subsections we will underline some important properties of the embedded Voronoi codes which are often inherited or induced from the Voronoi codes.

4.2. Source mismatch

The embedded Voronoi code of order r is not systematically contained in the Voronoi region scaled by the factor $2^{(r)}$. This property illustrated for the case of lattice A_2 in Fig. 3 is an important aspect regarding the performance of the new embedded quantization. Indeed, the scaled Voronoi region is a quasi-optimal shape of support for LVQ in case of a memoryless Gaussian, especially at higher dimension. The divergence of the codebook shape of $C^{(r)}$ from the Voronoi region will consequently penalize the performance of the quantization for such a source. This is the price that must be paid for embedding the code compared to a single stage Voronoi code $V(\Lambda, 2^r\Lambda, a)$.

4.3. Centroid of the code

Given that for most of the source distributions the origin is the optimal centroid, it is important to know and control the centroid of $C^{(r)}$. In the specific case where the different Voronoi codes $V(\Lambda, 2\Lambda, a_i)$ involved in $C^{(r)}$ use the same offset a, the centroid $E(C^r)$ will drift away from the origin as it can be observed in Fig. 3 (b). The centroid of such codebooks can then be expressed as:

$$\begin{aligned} E(C^{(r)}) &= (2^{(r)} + 2^{(r-1)} + \ldots + 1))E(V(\Lambda, 2\Lambda, a)) \\ &= (2^{r+1} - 1)E(V(\Lambda, 2\Lambda, a)) \end{aligned} \tag{10}$$

where $E(V(\Lambda, 2\Lambda, a))$ is a non-null centroid of $V(\Lambda, 2\Lambda, a)$. For avoiding such a drift, we propose to constrain the offsets a_i $V_i(\Lambda, 2\Lambda)$ such that:

$$\begin{cases} a_r = a \\ a_{r-1} = \cdots = a_0 = -a \end{cases} \tag{11}$$

Under such a condition the centroid of C^r comes $(2^{r+1} - 1)$ times closer to the origin and is equal to $E(V(\Lambda, 2\Lambda), a)$. In this manner the centroid of the embedded Voronoi code is of the same order as the Voronoi codes as shown in Fig. 3 (c).

4.4. Decoding of an approximation

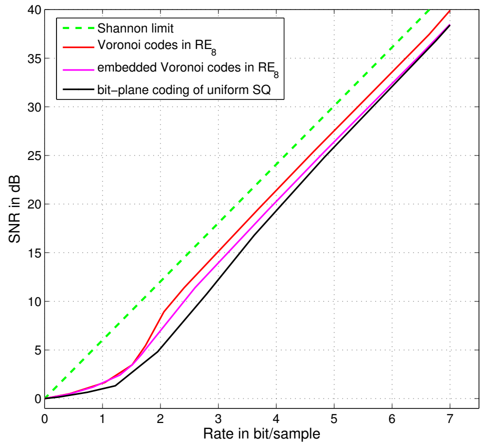
The distortion caused by the embedded Voronoi codes is equal to the granular distortion of lattice Λ if the bitstream is fully decoded. In case of truncation of the bitstream, the distortion increases but can be limited at the decoder side by adjusting the decoded approximation. The necessity of such an adjustment comes from the fact that the scaled Voronoi region used for generating the Voronoi codes in (5) is not entirely covered and defined by the different Voronoi regions of the generated codes. That means that most of the time we have the following inequality:

$$\Omega_0(m\Lambda) \neq \bigcup_{c \in V(\Lambda, m\Lambda)} \Omega_0(\Lambda) + c \tag{12}$$

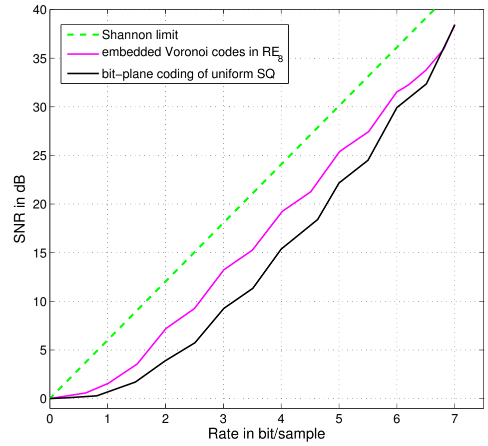
As a direct consequence for the embedded Voronoi codes, a subset of the decomposition $y' = v_r + \ldots + v_l$ is necessarily the best representative of the reachable codes in C^r. For getting an optimal approximation, an adjustment has to be computed at the decoder side. In case C^r fulfilled the condition (11), the decoded approximation then can be expressed as follows:

$$\begin{cases} y & \text{if } R = N(r+1) \\ \sum_{i=r}^{l} 2^i v_i - (2^l - 1)E(V(\Lambda, 2\Lambda), a) & \text{if } R < N(r+1) \end{cases} \tag{13}$$

where $R \in \{0, \ldots, N(r+1)\}$ is the number of bits considered for decoding an approximation of y and l is the number of subvectors dropped from the decomposition of y by the decoder and is given by $l = r + 1 - \lfloor \frac{R}{N} \rfloor$.



(a) Performance for fixed bit-rates



(b) Performance for scalable bit-rates

Fig. 4. Rate-distortion evaluation of the embedded Voronoi codes.

5. RESULTS

To evaluate the coding performance of the embedded Voronoi codes, we consider a zero-mean and unity variance i.i.d. Gaussian random process. The embedded Voronoi codes are applied at dimension 8 in the lattice RE_8. The Signal to Noise Ratio (SNR) is computed at different rates. The different rates are obtained by tuning a global gain g which is applied to the signal before quantization. The inversion gain $1/g$ is applied at the end of the decoding process. The higher g is, the lower is the distortion and the higher is the rate R. The performance of the embedded Voronoi codes is compared to:

- The theoretical limit of Shannon of SNR: $20R\log_{10}2 \approx 6.02R$
- The Voronoi codes in RE_8
- The bit plane coding used after a uniform scalar quantization.

An ideal entropy coding was used to transmit the order r of the embedded Voronoi codes, the factor m of Voronoi codes and the number K of bit-planes.

In Fig. 4 (a), the SNR is measured at different rates when encoding and decoding are done at the same fixed-rates. When comparing to the conventional Voronoi codes, it can be seen that the embedding property of the embedded Voronoi codes comes at a small cost of performance at low bit-rates. This penalty was explained in subsection 4.2. However, the embedded Voronoi codes still keep the advantages of LVQ over a uniform SQ. At high bitrates, e.g. 7 bits per dimension and more, the embedded Voromoi codes show higher performance degradation and the advantage over the uniform SQ vanishes.

In Fig. 4 (b), the performance in case of progressive transmission is evaluated. This was achieved by generating a single scalable bitstream at a rate of 7 bits/sample and by truncating it in order to reach different decoded bit-rates. Only the bit-plane coding and the embedded Voronoi codes can produce a scalable bitstream. It can be seen that the embedded Voronoi codes perform significantly better than the bit-plane coding. This is partly due to the adjustment defined in subsection 4.4 and used at the decoder side by the embedded Voronoi codes. One can also observe an oscillation of the SNR curves. The period of the oscillation corresponds to 1 bit/sample, which shows that an optimal truncation is obtained when an integer number of bits per sample is removed from the bitstream. In this case, the same number of bits is dropped from all transmitted values, which induces a better quantization of the signal.

6. CONCLUSION

Embedded Voronoi codes allow a successive refinement of the source while maintaining the advantages of the LVQ over the uniform SQ. The definition of the embedded Voronoi codes is generic and can be applied to any lattices. The cost of the embedding property is a suboptimal codebook shape which can lead to a certain mismatch with the source distribution especially at high bit-rates. Finally, the embedded Voronoi codes can also be applied for the scalar case, and can be used as a bit-plane coding extension for higher dimension.

As a future work, one can study the performance of the embedded Voronoi codes in conjuction with an entropy coding of the successive refinements. Such an association has been already done for the bit-plane coding [17]. It can also be interesting to study further the combination of the embedded Voronoi codes with advanced multi-rate LVQ techniques like the one described in [5] for exploring the reachable tradeoffs between adaptation to the source and bitrate scalability.

7. RELATION TO PRIOR WORK

The presented embedded Voronoi codes are mainly based on the Voronoi codes introduced in [15]. A previous work addressed the same problem in [11], but used a different approach. The authors defined a hard constraint for obtaining a successive refinement without or with minimal loss of optimality in terms of rate-distortion. The obtained solution requires designing each embedded quantizer individually. The solution cannot be applied easily to any dimensions and lattices. In this work, this constraint was relaxed, allowing some loss of efficiency, but making the embedded Voronoi codes more universal and generic. Similar codes were used in a scalable audio coding scheme presented by the author of this paper in [12]. The present work builds on those results by providing accurate definitions and a theoretical analysis.

8. REFERENCES

[1] M. Antonini, M. Barlaud, and P. Mathieu, "Image coding using lattice vector quantization of wavelet coefficients," *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP*, pp. 2273–2276, May 1991.

[2] M. Barlaud, P. Sole, T. Gaidon, M. Antonini, and P. Mathieu, "Pyramidal lattice vector quantization for multiscale image coding," *IEEE Trans. Image Processing*, pp. 367–381, July 1994.

[3] A. Woolf and G. Rogers, "Lattice vector quantization of image wavelet coefficient vectors using a simplified form of entropy coding," *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP*, vol. 5, pp. 269–272, Apr. 1994.

[4] W.-H. Kim, Y.-H Hu, and T.Q. Nguyen, "Wavelet-based image coder with entropy-constrained lattice vector quantizer (ECLVQ)," *IEEE Trans. on Circuits ans Systems*, vol. 45, no. 8, pp. 1015–1030, Aug. 1998.

[5] S. Ragot, B. Bessette, and R. Lefebvre, "Low-Complexity Multi-Rate Lattice Vector Quantization with Application to Wideband TCX Speech Coding at 32 kit/s," *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP*, 2004.

[6] B. Bessette, R. Lefebvre, and R. Salami, "Universal speech/audio codng using hybrid ACELP/TCX techniques," *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP*, pp. 301–304, Mar. 2005.

[7] J.M. shapiro, "Embedded image coding using zerotrees of wavelet coefficients," Dec. 1993.

[8] A. Said and W.A. Pealman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits Systems Video Technol.*, vol. 6, pp. 243–250, June 1996.

[9] J. Li, "Embedded Audio Coding (EAC) with Implicit Psychoacoustic Masking," *ACM Multimedia*, pp. 592–601, Dec. 2002.

[10] Ralf Geiger, Jürgen Herre, Gerald Schuller, and Thomas Sporer, "Fine grain scalable perceptual and lossless audio coding based on intmdct," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, ICASSP*, 2003, pp. 445–448.

[11] D. Mukherjee and S.K. Mitra, "Successive refinement lattice vector quantization," *IEEE Trans. on Signal Processing*, vol. 11, no. 12, pp. 1337–1348, Dec. 2002.

[12] G. Fuchs and R. Lefebvre, "A Scalable CELP/Transform Coder for Low Bit Rate Speech and Audio Coding," *AES 120th convention at Paris*, May 2006.

[13] E. Agrell and T. Eriksson, "Optimization of lattices for quantization," *IEEE Trans. on Information Theory*, , no. 5, pp. 1814–1828, 1998.

[14] A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Plublishers, 1992.

[15] J.H. Conway and N.J.A. Sloane, "A Fast Encoding Method for Lattice Codes and Quantizers," *IEEE Trans. on Information Theory*, vol. 29, no. 6, pp. 820–824, November 1983.

[16] G.D. Forney, "Coset codes. I. Introduction and geometrical classification," *IEEE Trans. on Information Theory*, vol. 34, no. 5, pp. 1123–1151, Sept. 1988.

[17] Sung-Hee Park, Yeon-Bae Kim, and Yang-Scock Seo, "Multilayer bit-sliced bit-rate scalable audio coding," in *Audio Engineering Society Convention 103*, 9 1997.