PREDICTIVE MULTIPLE-SCALE LATTICE VQ FOR LSF QUANTIZATION

*Adriana Vasilache, Marcel Vasilache * and Ioan Tăbuş*

Signal Processing Lab., Tampere Univ. of Technology, Finland
* Nokia Research Center, Tampere, Finland

ABSTRACT

This paper introduces a new lattice quantization scheme, the multiple-scale lattice vector quantization (MSLVQ), based on the truncation of the D_{10}^{+} lattice. The codebook is composed of several copies of the truncated lattice scaled with different scaling factors. A fast nearest neighbor search is introduced. We compare the performance of predictive MSLVQ for quantization of LSF coefficients with the quantization technique used in the codec G.729 and show the better performance of our method in terms of spectral distortion. The MSLVQ scheme achieves the transparent quality at 21 bits/frame.

1. INTRODUCTION

Linear predictive coding (LPC) method is one of the most popular approaches used for describing the short-term spectrum of speech signal. In many speech coding systems, LPC coefficients are transformed in the line spectrum frequencies (LSF) representation in view of quantization.

During the last years, many vector quantization schemes have been proposed for the transmission of the LSF parameters. One way to rank various quantization methods is to compare the bit rates at which is achieved the transparent quality [6], which usually are in the range 20-30 bits/frame. For practical applications the transparent quality is too costly and very recent standards have settled the tradeoff cost-performance at lower than transparent quality requirements, one typical example being the standard G.729, where a non-transparent, but still good quality, is achieved with 18 bits/frame by employing a low complexity multistage combined with split VQ, in a predictive scheme.

It was observed that exploiting the interframe correlation the transparent quality is obtained at lower bit-rates, as proven by the better results of predictive VQ when compared to memoryless VQ [10]. Principal component analysis can provide another effective method of taking advantage of LSF coefficients redundancy, as shown in [13], where the Karhunen Loeve transform is applied to LSF coefficients and only 7 out of 10 of the transformed coefficients are transmitted, resulting in transparent quality at a rate of 22 bits/frame. In [12] a dynamic codebook ordering followed by entropy coding is shown also to use very efficiently the redundancy of LSF parameters, reducing from 24 to 20 the number of bits/frame. Another method that uses the memory in the process of transmission and reception of the quantization index is presented in [11]. In [16] successive LSF coefficients are partitioned into variable length segments of strings that are classified in a finite number of classes.

Some critical factors in wireless applications are the amount of memory required to store the codebooks and the complexity of computation used for comparing the input vector to each of the codevectors. One efficient method to address these issues is the use of lattice VQ. Results on this line have been reported in [14] where by means of a split VQ, three low-dimensional (2,2,3) lattice codebooks and a two dimensional stochastic vector quantizer are employed. With this scheme the transparency is obtained at 28 bits/frame. A two stage tree-structured VQ with a pyramidal truncated lattice in the second stage is proposed by Pan [8].

We present in this work the use of a 10-dimensional lattice VQ that has better performance than G.729 at 18 bits/frame and is "transparent" at 21 bits/frame. We present a very fast method for searching the codebook for the nearest neighbor. The scheme has very low memory requirements, since there is no need to store the codebook.

First we review the spectral quantization problem, the lattice vector quantization method, and pay special attention to the lattice D_{10}^{+} which will be used in this paper. The new predictive multiple-scale lattice VQ will be further introduced, followed by experimental results and conclusions.

2. LPC QUANTIZATION

The LPC parameters are extracted from the input signal and transmitted in order to reconstruct at the decoder the short-term spectral envelope of the decoded signal. The LPC parameters are the coefficients of the p-th order prediction polynomial $A(z)$, obtained by applying the Levinson algorithm to a frame of speech signal. At the decoder, the decoded excitation is filtered by the all-pole filter $H(z) = 1/A(z)$. The order of $A(z)$ is commonly taken $p = 10$. The LPC parameters are transformed into an equivalent set, the line spectrum pairs (LSF) and their quantized values are transmitted to the receiver. The process of quantizing the filter parameters to a finite number of bits/frame is referred to as LPC quantization.

The spectral distortion is often used as an objective measure of the encoding performance:

$$SD = \left\{ \frac{10}{\pi} \int_0^{\pi} \left[\log_{10} |A_n(e^{j\omega})|^2 - \log_{10} |\hat{A}_n(e^{j\omega})|^2 \right]^2 d\omega \right\}^{1/2} \quad (1)$$

where SD is given in dB, $A_n(e^{j\omega})$ and $\hat{A}_n(e^{j\omega})$ are the spectra of the $n-th$ speech frame without and with quantization respectively.

3. LATTICE VQ

A lattice can be defined as all the linear combinations with integer coefficients of the basis vectors, $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_d$ [1]. Other definitions also exist, but in the context of vector quantization we prefer this one. Lattice points are equally distributed in space and every point in the lattice has the same spatial distribution of its neighbors. A lattice can be described by means of the generator matrix (formed by the basis vectors) [1] or by properties associated with the lattice point coordinates [2]. All the lattice points have integer components but they can be scaled by a positive real number such that they suit the application demands.

A truncated and properly scaled lattice can be used as a codebook for a vector quantizer. The advantages of this VQ method are the fast nearest neighbor search algorithm and the reduced memory demands. However, reaching with a lattice VQ a good coding efficiency depends on how close to the uniform distribution is the distribution of data, and, as observed in [1], the higher the number of dimensions, the closer the distribution is to a uniform one.

A truncated lattice can be defined as the set of lattice points having the norm less than a given value K:

$$T = \{x \in \Lambda | N(x) \leq K\} \tag{2}$$

where Λ is the lattice and $N(x)$ is a norm of x. If $N(x)$ is the Euclidean norm the truncation will be spherical [1] and in the case of the l_1 norm the truncation will be pyramidal [8]. The lattice points will be grouped on spherical or pyramidal shells. The number of points on each shell can be calculated by means of the θ series for spherical shells [2] or for the pyramidal shells by means of the $\tilde{\theta}$ series that will be further introduced.

3.1. The lattice D_{10}^+

Recently it has been found out that D_{10}^+ is the lattice with the lowest second moment in 10 dimensions [1], which makes it the best candidate for a lattice quantizer in 10 dimensions. This lattice is defined as:

$$D_{10}^+ = D_{10} \cup (D_{10} + \underbrace{[1/2 \ldots 1/2]}_{10}) \tag{3}$$

where the lattice D_{10} is:

$$D_{10} = \{(x_i) \in \mathbb{Z}^{10} | \sum x_i = even\} \tag{4}$$

3.2. Truncation of the lattice

Let $N(x)$ be the norm of the vector x. As stated before, the lattice can be spherically or pyramidally truncated. In both cases the number of lattice points in the truncation must be determined. For the first case this problem is solved in [2] and for the second case (pyramidal truncation i.e. $N(x) = \sum_{i=1}^{n} |x_i|$) we shall present a solution in what follows.

We introduce the following $\tilde{\theta}$ series:

$$\tilde{\theta}_2(z) = \sum_{m=-\infty}^{\infty} q^{|m+1/2|},$$

$$\tilde{\theta}_3(z) = \sum_{m=-\infty}^{\infty} q^{|m|},$$

$$\text{and } \tilde{\theta}_4(z) = \sum_{m=-\infty}^{\infty} (-1)^m q^{|m|}$$

where $q = e^{\pi i z}$.

With these new defined series we follow a similar methodology to the one presented at page 45 in [2] (which uses θ series) for computing the number of points on each shell.

The θ series for the lattice Λ in the case of l_1 norm are:

$$\tilde{\theta}_\Lambda(z) = \sum_{\mathbf{x} \in \Lambda} q^{N(\mathbf{x})} = \sum_{m=0}^{\infty} N_m q^m \tag{5}$$

where N_m is the number of points of the lattice that have the norm m.

The $\tilde{\theta}$ series of the integers $\mathbb{Z}$ are $\tilde{\theta}_3(z)$ while for its translate $\mathbb{Z} + 1/2$ the series are $\tilde{\theta}_2(z)$.

The l_1 norm has the property $N((x_1, x_2, \ldots, x_n)) = N(x_1) + N(x_2) + \ldots + N(x_n)$ (like the Euclidean norm). Thus we have:

$$\tilde{\theta}_{\mathbb{Z}^n}(z) = \tilde{\theta}_{\mathbb{Z}}(z)^n = \tilde{\theta}_3(z)^n. \tag{6}$$

The lattice D_n is the sub-ensemble of the Z_n lattice for which the sum of the components of a lattice vector is even. Its $\tilde{\theta}$ series can be obtained like in the Euclidean norm case:

$$\tilde{\theta}_{D_n}(z) = \frac{1}{2}(\tilde{\theta}_3(z)^n + \tilde{\theta}_4(z)^n). \tag{7}$$

And finally for the lattice D_n^+ the $\tilde{\theta}$ series are

$$\tilde{\theta}_{D_n^+}(z) = \frac{1}{2}(\tilde{\theta}_2(z)^n + \tilde{\theta}_3(z)^n + \tilde{\theta}_4(z)^n). \tag{8}$$

For $n = 10$ the distribution of points on the first 6 pyramidal shells of norm m is as follows:

m	0	2	4	5	6	7
N_m	1	200	6800	512	103000	28160

Table 1. Distribution of lattice points on pyramids

and on spheres of radius m is as presented in [2]:

m^2	0	2	2.5	4	4.5	6	6.5
N_m	1	180	512	3380	5120	16320	23040

Table 2. Distribution of lattice points on spheres

3.3. A fast algorithm for nearest neighbor search in D_{10}^+

A fast nearest neighbor (NN) algorithm in D_{10}^+ can be obtained by first searching in the D_{10} lattice (as presented in [2]), then in the translated D_{10} lattice (see (3)) and taking the best result. As the search has to be performed on a truncated lattice we have derived a suboptimal algorithm that uses binary search. The algorithm is based on the fact that in the truncated lattice an approximative NN of a point x can be found by checking the NN's in D_{10}^+ for a set of points having similar orientations as x but different norms. If NN of x is inside the truncation limit nothing else has to be done, otherwise the search procedure is used. The initial range of norms is set to be from 0 to $N(x)$ and the test point is placed in the middle. At each step the NN of the test point is obtained and if it is within the truncation limit the lower search limit is increased to the test position otherwise the upper limit is similarly decreased. The search stops when the search range is smaller than a 'significant' distance in the lattice or if the same NN's are obtained at both ends of the search region.

This is a general algorithm that works for both spherical and pyramidal truncation. In the case of pyramidal truncation the l_1 norm is used when making the decision inside/outside.

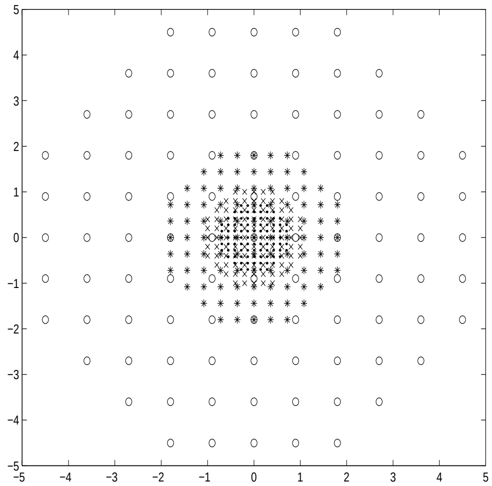

Figure 1: MSLVQ for lattice $\mathbb{Z}^2$ with the scales 0.7(·), 1.0(x), 1.8(*) and 4.5(o)

3.4. Enumeration of the codevectors

Several methods for the enumeration of the points of a pyramidal truncated lattice [3] or spherical and pyramidal [7] exist in the literature. We have utilized an algorithm that demands memory only for the table (2) and is suited (without further adaptation) to the spherical truncation of the lattice D_{10}^+. It uses a procedure of enumerating combinations and it is based on the special properties of the lattice vectors on each layer (e.g. on the first layer there can be vectors that have only two components of unitary absolute values, the others being zero).

4. PREDICTIVE MULTIPLE-SCALE LATTICE VQ

A predictive VQ [10] scheme exploits the memory of the input vectors by computing first the predicted vector and then quantizing the prediction error. The prediction is based only on *previous quantized* vectors. The predictor may be of AR or MA type and it has been shown experimentally [10] that it is sufficient to use only the diagonal elements of the prediction matrices. The AR predictor used in this study has the form: $\omega_t^p = \sum_{i=1}^N p_i \hat{\omega}_{t-i}$, where ω_t^p is the predicted value, $\hat{\omega}_{t-i}$ the previous quantized values, N the predictor order and p_i the predictor coefficients.

In this work the truncated lattice is used for the quantization of the prediction errors. It is well known that in the scalar case the distribution of the prediction error is close to a Laplacian distribution. From Figure 1 one can see that a multiple-scale truncated lattice is closer to Lloyd Max optimum quantizer when compared to an uniform lattice, for a given size of the codebook.

The constant K defining the truncation of the lattice in (2) is selected such that T contains all lattice points on a given number of shells. Now introducing also the multiple scaling factors we will define our codebook as follows:

$$\left\{ sy \quad | \quad y \in D_{10}^+, \;\; N(y) < K, \;\; s \in \{s_1, \ldots, s_S\} \right\} \qquad (9)$$

where s is the scaling factor and S is the number of scaling factors.

The index of the codevector to be transmitted should be formed by the index of the lattice point and the index of the scaling factor.

Some related methods used to obtain a lattice quantizer for a non-uniform source have already been presented in the literature as the multidimensional companding [9] and the embedded algebraic VQ [15] or piecewise uniform lattice VQ [5]. We have tested the companding by transforming the distribution of the input vectors into a uniform one using spherical companding. However, in the truncated lattice there is a small number of norm levels (e.g. 15 for 2^{21} codevectors) that might induce large errors in the norm at the uncompanding procedure.

5. EXPERIMENTAL RESULTS

We experimented the predictive multiple-scale lattice VQ for the quantization of LPC parameters, by utilizing the same LPC computation procedure used in G.729 codec. There, a 10-th order LPC analysis with 10 ms analysis frame is employed, based on the autocorrelation method and the use of a 30 ms Hamming window.

One part from the training set of the TIMIT speech database, consisting of 115006 frames, has been used for the design of the AR predictor of order 2 and the optimization of scales. The results were not significantly different when we used 4 scales or 8 scales. The testing data have been the remaining part of the training set, T_{trn}, (up to 1400000 frames) and the test set ,T_{tst}, (about 500000 frames) of the TIMIT speech database. The experiments have shown that a spherical truncated lattice performs better than a pyramidal truncated one. This is in part due to the repartition on pyramidal shells of the lattice points given in subsection 3.2. It should be noted that due to the structure of the lattice, the number of available codevectors at a given bit-rate (power of 2) is greater than the number of used codevectors. This results in different bit utilization efficiencies for the different truncations, number of scales or bit-rates. However the unused codevector indexes can be used for increasing the resilience of coding.

$\overline{SD}$ (dB)	> 1 (%)	> 4 (%)	> 6 (%)	$[2, 4]$ (%)	Data
1.356	56.17	1.03	0.19	16.93	T_{tst}

Table 3. Spectral distortion for a pyramidal truncated lattice (110513 codevectors, K = 6), 2 scales (1.0 3.0), (18 bits=17+1):

$\overline{SD}$ (dB)	> 1 (%)	> 4 (%)	> 6 (%)	$[2, 4]$ (%)	Data
1.199	47.97	0.73	0.02	10.57	T_{trn}
1.201	48.07	0.74	0.02	10.61	T_{tst}

Table 4. Spectral distortion for a spherical truncated lattice (25513 codevectors, K = 2.45), 8 scales (0.7 1.0 1.2 1.5 1.8 2.0 4.0 6.0), (18 bits=15+3):

$\overline{SD}$ (dB)	> 1 (%)	> 4 (%)	> 6 (%)	$[2, 4]$ (%)	Data
1.156	43.42	0.75	0.02	9.45	T_{trn}
1.157	43.51	0.76	0.02	9.49	T_{tst}

Table 5. Spectral distortion for a spherical truncated lattice (48553 codevectors, K = $\sqrt{6.5}$), 4 scales (0.7 1.0 1.8 4.5), (18 bits = 16+2)

The predictive MSLVQ scheme is compared with the codec G.729 in terms of spectral distortion of the quantized LSF coefficients (1). For the same bit-rate the average spectral distortion with our quantization scheme is smaller than that of G.729 and the number of outliers is also smaller. Informal listening tests have shown good performance of the MSLVQ scheme.

$\overline{SD}$ (dB)	> 1 (%)	> 4 (%)	> 6 (%)	$[2; 4]$ (%)
1.347	67.74	0.46	0.02	12.00

Table 6. Spectral distortion for G.729

The results for the spherical companding from Table 7 show that, as expected, the violation of the high rate assumption leads to worse quantization performance.

bits	$\overline{SD}$ (dB)	> 1 (%)	> 4 (%)	> 6 (%)	$[2; 4]$ (%)
19	1.496	58.95	5.32	1.78	10.79
20	1.404	50.08	4.68	1.51	10.14
21	1.348	44.37	4.43	1.36	9.97
22	1.308	40.55	4.29	1.24	9.88

Table 7. Spectral distortion using companding for spherical truncated lattice

It is important to note that at a bit-rate of 21 bits/frame the MSLVQ scheme can give an average spectral distortion of 0.93 dB.

bits	$\overline{SD}$ (dB)	> 1 (%)	> 4 (%)	> 6 (%)	$[2; 4]$ (%)
19	1.081	37.53	0.61	0.01	7.79
20	1.015	32.45	0.47	0.00	6.42
21	0.927	25.74	0.33	0.00	4.59
22	0.866	21.14	0.23	0.00	3.44

Table 8. Spectral distortion for a spherical truncated lattice, 4 scales

bits	$\overline{SD}$ (dB)	> 1 (%)	> 4 (%)	> 6 (%)	$[2; 4]$ (%)
19	1.094	39.70	0.43	0.01	8.21
20	1.022	33.90	0.31	0.00	6.61
21	0.961	28.86	0.22	0.00	5.38
22	0.880	22.46	0.10	0.00	3.79

Table 9. Spectral distortion spherical truncated lattice, 8 scales

6. CONCLUSION

In this work we have introduced a new lattice quantization scheme. It uses a truncation of the D_{10}^{+} lattice. The codebook is composed of several copies of this truncation, scaled with different scaling factors. This way one has to encode the index of the lattice codevector and the corresponding scale. A method for the calculation of the number of lattice points on each pyramidal shell has been presented as well as a fast algorithm for the nearest neighbor search in the truncated lattice (pyramidal or spherical). The resulting quantization scheme has been applied to LPC quantization and was shown to perform better (for the same bit rate - 18 bits/frame) than the quantization scheme from the codec G.729. Besides this performance in terms of spectral distortion, it should be noted that no memory is necessary for the codebook and a fast search algorithm is available. The MSLVQ scheme achieves an average spectral distortion of 0.93 dB at 21 bits/frame.

7. REFERENCES

[1] E. Agrell. *Voronoi-based coding*. PhD thesis, School of Electrical and Computer Engineering, Chalmers Universoty of Technology, Göteborg, Sweden, 1997.

[2] J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups*. Springer-Verlag,New York, 1992.

[3] T. R. Fischer and J. Pan. Enumeration encoding and decoding algorithms for pyramyd cubic lattice and trellis codes. *IEEE Trans. on Information Theory*, 41(6):2056–2061, November 1995.

[4] Coding of speech at 8kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP). ITU Telecommunications Standardization Sector, December 1995. Draft Recommandation.

[5] D. G. Jeong and J. D. Gibson. Uniform and piecewise uniform lattice vector quantization for memoryless gaussian and laplacian sources. *IEEE Trans. on Information Theory*, 39(3):786–803, May 1993.

[6] W. Kleijn and K. Paliwal, editors. *Speech coding and Synthesis*. Elsevier Science, 1995.

[7] J. Moureaux, P. Loyer, and M. Antonini. Efficient indexing method for lattice quantization applications. In *ICIP'96*, volume 3, pages 447–450, 1996.

[8] J. Pan. Two-stage vector quantization-pyramidal lattice vector quantization and application to speech LSP coding. In *ICASSP'96*, volume 2, pages 737–740, 1996.

[9] S. F. Simon. On suboptimal companding. In *Proc. Data Compression Conference*, pages 438–447, Snowbird, Utah, March 1998.

[10] J. Skoglund and J. Lindén. Predictive VQ for noisy channel spectrum coding: AR or MA? In *Proceedings of ICASSP'97*, volume 2, pages 1351–1354, Munich, Germany, 1997.

[11] M. Skoglund and J. Skoglund. On non-linear utilization of intervector dependency in vector quantization. In *Proceedings of ICASSP'98*, Seattle, USA, May 1998. IEEE.

[12] S. Sridharan and J. Leis. Two novel lossless algorithms to exploit index redundancy in VQ speech compression. In *Proceedings of ICASSP'98*, Seattle, USA, May 1998.

[13] H. L. Vu and L. Lois. Optimal transformation of LSP parameters using neural network. In *Proceedings of ICASSP'97*, volume 2, pages 1339–1342, Munich, Germany, 1997. IEEE.

[14] M. Xie and J.-P. Adoul. Fast and low-complexity LSF quantization using algebraic vector quantizer. In *ICASSP'95*, volume 1, pages 716–719, May 1995.

[15] M. Xie and J.-P. Adoul. Embedded algebraic vector quantizers (EAVQ) with application to wideband speech coding. In *ICASSP'96*, volume 1, pages 240–243, May 1996.

[16] C. S. Xydeas and T. M. Chapman. Multi codebook vector quantization of LPC parameters. In *Proceedings of ICASSP'98*, Seattle, USA, May 1998. IEEE.