MODELING AND ANALYSIS OF VECTOR-QUANTIZED M-CHANNEL SUBBAND CODECS

Innho Jee and R. A. Haddad
Department of Electrical Engineering
Polytechnic University
Five Metrotech Center, Brooklyn, NY 11201
E-mail : ijee@photon.poly.edu

ABSTRACT

This paper demonstrates that the scalar non-linear gain-plus-additive noise quantization model can be used to represent each vector quantizer in an M-band subband codec. The validity and accuracy of this analytic model is confirmed by comparing the calculated model quantization errors with actual simulation of the optimum LBG vector quantizer. We compute the mean squared reconstruction error(MSE) which depends on N the number of entries in each codebook, k the length of each codeword, and on the filter bank coefficients. We form this MSE measure in terms of the equivalent scalar quantization model and find the optimum FIR filter coefficients for each channel in the M-band structure for a given bit rate, given filter length, and given input signal correlation model. Specific design examples are worked out for a 4-tap filter in a two-band paraunitary structure. Theoretical results are confirmed by extensive Monte Carlo simulation.

1. INTRODUCTION

Subband coding and vector quantization have been shown to be effective methods for low rate coding of speech, still image, video, and HDTV signals[1]. The idea of subband coding is to split the frequency band of the signal into a number of subands and then to encode each subband separately using a bit allocation algorithm which reflects the energy in each subband. The system under study is the critically sampled filter bank shown in Fig.1(a). The codebook for the vector quantizer for each channel is constructed from the Linde-Buzo-Gray(LBG) algorithm[8] using 500,000 samples of an AR(1) signal $x(n)$ passed through a bank of FIR filters designed under PR constraints. The overall constraint for the structure is B bit/sec which is to be allocated among these M channels,

$$\frac{1}{M}\sum_{j=0}^{M-1} B_j = B \qquad (1)$$

where B_j is the number of bits allocated to quantizer Q_j. Each VQ has a codebook of N_i entries of length k_i, and therefore B_j satisfies

$$B_j = \frac{log_2 N_j}{k_j}. \qquad (2)$$

Fig.1(b) results from a polyphase transformation of the filter bank and the scalar quantization models described in the sequel. These can be analyzed and optimized using scalar optimization method in Ref[2],[3].

2. MODELING THE VECTOR QUANTIZER

2.1 VECTOR QUANTIZER

As shown in Fig.2(a) an N-level k-dimensional quantizer is a mapping, Q, that assigns to each input vector, $\underline{v} = (v_0, v_1, .., v_{k-1})$, a reproduction vector, $\underline{\hat{v}} = Q(\underline{v})$, drawn from a finite reproduction alphabet, $\hat{A} = \{\underline{\hat{v}}_i; i = 1, 2, ..., N\}$. The quantizer Q is completely described by the reproduction codebook $\hat{A}$ together with the partition, $S = \{S_i; i = 1, 2, .., N\}$, of the input vector space into the sets $S_i = \{\underline{v}; Q(\underline{v}) = \underline{\hat{v}}_i\}$ of input vectors mapping into the i-th reproduction codeword. The quantizer performance can be measured by the distortion, $D = \frac{1}{k}E||\underline{v} - Q(\underline{v})||^m$, where $||.||$ denotes the usual l_2 norm. We wish to choose $\hat{v}_1, ..., \hat{v}_N$ to minimize D. The k-dimensional m^{th} power distortion-rate function of an optimal vector quantizer in high resolution is given by

$$D^k_{VQ}(B) = C(k,m)2^{-(m/k)B}[\int [p(\underline{v})]^{k/(m+k)} d\underline{v}]^{(m+k)/k} \qquad (3)$$

in Ref[4],[6]. The constant $C(k, m)$ is a function of the vector dimension k and of m and represents how well cells can be packed in k-dimensional space. The density function $p(\underline{v})$ is the k-dimensional joint pdf of the vector process. The properties of an optimized VQ for mean squared error distortion over a frame are[9]

$$E\{\underline{\tilde{v}}\} = \underline{0}, \quad E\{\underline{\hat{v}}^t \underline{\tilde{v}}\} = 0. \qquad (4)$$

2.2 APPROXIMATE OPTMIZED VECTOR QUANTIZER MODEL

According to Jayant & Noll[5], the short-time pdf of a speech segment can be approximated by a Gaussian pdf. The mean squared quantization error averaged over a frame in optimized vector quantizer coding can be computed approximately using the asymptotic distortion-rate function derived for a Gaussian random signal[7],

$$D_{VQ}^{k} \approx \tau 2^{-2B/k} (det\Gamma)^{1/k} \triangleq \sigma_{\underline{\tilde{v}}}^{2} \tag{5}$$

where k, B and Γ denote respectively the vector dimension, the number of bits allocated to the quantizer and the covariance matrix of the input signal, and τ is a correction factor

$$\tau = 2\pi ck(1+\frac{2}{k})^{k/2+1} \tag{6}$$

where c is the quantization coefficient for the VQ. The coefficients of quantization values are unknown except for $k = 1$ and 2. However, there are a number of approximations based on lower or upper bounds. The results in this paper are based on using the values given by the Voronoi lattice upper bound[6]. It is computationally burdensome to directly estimate $det\Gamma$. However, using the Toeplitz distribution theorem[5],

$$\lim_{k\to\infty} det\Gamma^{1/k} = exp[\frac{1}{2\pi}\int_{-\pi}^{\pi} log_e S_{xx}(e^{j\omega})d\omega] = \sigma_{e,min}^{2} \tag{7}$$

where $S_{xx}(e^{j\omega})$ is the power spectral density of the random signal $\{X(n)\}$ and $\sigma_{e,min}^2$ the energy of the minimum prediction error. When the vector dimension k and the predictor order are reasonably large, the quantization error in Eq.(5) can be further simplified to

$$D_{VQ}^{k} \approx \tau 2^{-2B/k} \sigma_e^2 \tag{8}$$

where σ_e^2 is the variance of the prediction error sequence using a finite memory optimal predictor in mean square sense[5].

For a scalar pdf-optimized quantizer, the quantization error variance in each channel is

$$\sigma_{b_j}^2 = \beta(B_j) 2^{-2B_j} \sigma_j^2 \tag{9}$$

where σ_j^2 is the variance of the signal input to the quantizer and $\beta(B_j)$ depends on the pdf of the input signal v and on B_j. Then the optimum allocation of bits is known to be

$$B_j = B + \frac{1}{2} log_2 \frac{\sigma_j^2}{(\prod_{i=0}^{M-1}\sigma_i^2)^{1/M}}. \tag{10}$$

This result implies that the relation between the distortion and bit-rate used for coding each speech vector (having a resonably large dimension) in VQ coding reduces to the same form as that used in the conventional memoryless scalar quantizer, except that the scalar signal variance is replaced by the prediction error variance.

3. GAIN-PLUS-ADDITIVE NOISE MODEL FOR VQ

The gain-plus-additive model for the pdf optimized scalar quantizer is shown in Fig.2(b). In that model, we know[3],[5]

$$E\{\tilde{v}\} = 0, \quad E\{\tilde{v}\hat{v}\} = 0 \tag{11}$$

$$\alpha = 1 - \frac{\sigma_{\tilde{v}}^2}{\sigma_v^2}, \quad \sigma_r^2 = \alpha(1-\alpha)\sigma_v^2 = \alpha\sigma_{\tilde{v}}^2. \tag{12}$$

We show that this represention can also be used for an optimized VQ. The distortion per frame in the LBG algorithm is

$$\frac{1}{k}\sum_{i=n-(k-1)}^{n} |v(i) - \hat{v}(i)|^2. \tag{13}$$

We show that this distortion measure equals D_{VQ}^k of Eq.(8).

- Assume $E\{|v(n-i) - \hat{v}(n-i)|^2\}$ is same for all i in that block. Can we use D_{VQ}^k of Eq.(8) as this measure ?
- Is it true that $\tilde{v}(n-i)$ is orthogonal to $\hat{v}(n-i)$ as required by Eq.(4) ?, where $\tilde{v}(i) = v(i) - \hat{v}(i)$.

Thus, we calculate $E\{\tilde{v}(i)\}$ and $E\{\hat{v}(i)\tilde{v}(i)\}$ as follows.

- $E\{\tilde{v}(i)\} \to \frac{1}{k}\sum_{i=0}^{k-1}\tilde{v}(i)$ for each block; then sum over blocks.
- $E\{\hat{v}(i)\tilde{v}(i)\} \to \frac{1}{k}\sum_{i=0}^{k-1}\hat{v}(i)\tilde{v}(i)$ for each block; then sum over all blocks.

From these simulations we will show that $E\{\tilde{v}(i)\} \simeq 0$, $E\{\hat{v}(i)\tilde{v}(i)\} \simeq 0$. So, we can use $D_{VQ}^k = \tau 2^{-2B/k}\sigma_e^2 = \frac{1}{k}\sum_{i=n-(k-1)}^{n}|v(i)-\hat{v}(i)|^2$ in the pdf-optimized vector quantizer. Comparing Eq.(11) and (12) for the scalar quantizer with Eq.(8) for VQ, we see that if $\sigma_{\underline{\tilde{v}}}^2$ of VQ obtained from $\sigma_{\underline{\tilde{v}}}^2 = \frac{1}{k}\sum_{i=0}^{k-1}\sigma_{\tilde{v}_i}^2$ per block and averaged over all blocks equals $\sigma_{\tilde{v}}^2$ of the scalar quantizer, we can say

$$\alpha = 1 - \frac{\sigma_{\tilde{v}}^2}{\sigma_v^2} = 1 - \frac{\tau 2^{-2B/k}\sigma_e^2}{\sigma_v^2} \tag{14}$$

τ, which depends on k, the vector dimension, is given in a table in Ref[6]. Also, from the theory of linear optimum prediction[5] $\sigma_e^2 = E\{(\hat{v}-v)^2\}$ and the optimal prediction error is represented as

$$\sigma_e^2 = \gamma_v^2\sigma_v^2. \tag{15}$$

Thus

$$\alpha = 1 - \tau 2^{-2B/k} \gamma_v^2 \quad (16)$$

where γ_v^2 is the spectral flatness measure which is the reciprocal of the maximum prediction gain

$$\gamma_v^2 = min\{^\infty \sigma_e^2\}/\sigma_v^2 = [max\{^\infty G_p\}]^{-1} \quad (17)$$

where G_p is the prediction gain of predictor[5]. Note that γ_v^2 is independent of the quantizer. We calculate γ_v^2 in the following way. Consider a zero-mean process {X(n)} with power spectral density $S_{xx}(e^{j\omega})$. This signal is filtered by $H(e^{j\omega})$. Its filtered signal spectral density is $S_{vv}(e^{j\omega}) = |H(e^{j\omega})|^2 S_{xx}(e^{j\omega})$ and γ_v^2 is

$$\gamma_v^2 = \frac{exp[\frac{1}{2\pi}\int_{-\pi}^{\pi} log_e S_{vv}(e^{j\omega})d\omega]}{\frac{1}{2\pi}\int_{-\pi}^{\pi} S_{vv}(e^{j\omega})d\omega}. \quad (18)$$

Eq.(16) gives us a theoretical value for α. Then use of Eq.(14) gives $\sigma_{\tilde{v}}^2 = (1-\alpha)\sigma_v^2$. This theoretical value is compared with simulated quantization error variance. In our case, B is small and $k = 8$ is also small. Therefore to improve the accuracy of the model we introduce an empirically determined correction factor δ which depends on B and k[7]. Hence

$$\alpha = 1 - \tau 2^{-2(B/k-\delta)} \gamma_v^2. \quad (19)$$

The optimized VQ mean squared error is now

$$\sigma_{\underline{\tilde{v}}}^2 = \tau 2^{-2(B/k-\delta)} \gamma_v^2 \sigma_v^2. \quad (20)$$

4. SIMULATIONS

1. An input AR(1) ($\rho = 0.95$, mean=0, var=1.0) signal is passed through a 4-tap Binomial QMF[10]. This filtered signal is used as a training signal for codebook design using the LBG algorithm. We choose k=8 for vector dimension, N=32, 64 for codebook addresses, and n=500,000 samples for training sequences. The number of training vectors used in codebook generation is $\geq 100N$. The average distortion in this algorithm is the mean-square error distortion. We show simulation results in Table 1. From these simulations we see that $E\{\tilde{v}(i)\} \simeq 0$, $E\{\hat{v}(i)\tilde{v}(i)\} \simeq 0$. The requirements for the pdf optimized quantizer of Eq.(11) are satisfied. So, we can use $D_{VQ}^k = \tau 2^{-2B/k}\sigma_e^2 = \frac{1}{k}\sum_{i=n-(k-1)}^{n} |v(i) - \hat{v}(i)|^2$ to measure distortion in the optimized vector quantizer.

2. We compare $E\{|\tilde{v}(i)|^2\}$ from test on VQ experimentally with $\sigma_{\tilde{v}}^2$ from theoretical scalar gain-plus-additive noise model Eq.(18) and (19). These results are shown in Table 2 with the correction factor δ equal to zero. An even closer match can be found by selecting, δ, from the empirically obtained universal table, as shown in Table 3. From these simulations we conclude that optimum vector quantizer in an M-channel subband coder can be modeled by the scalar gain-plus-additive noise scalar model.

3. We design specific example for the paraunitary, two band 4-tap case. The optimization algorithm is based on the exhaustive search of all possible bit allocations constrained by the total number of bits with Monte Carlo simulation. And we choose the one with minimum MSE among them. Considering the complexity of vector quantizer, we choose bit rate from 0.5 - 1.0 bit/sample. We assume that each quantizers are to allocated only integer bits and the high frequency components of the subband signal gets at least 1 bit and the low frequency compoments of the signal gets maximum 11 bits for a codebook with 2048 entries(or 11/8 bit/sample). Also, we choose test sequence with 64,000 samples to validate the theory.

4. Calculation Procedure
   (a) Fix B, k, ρ, and codebook.
   (b) Choose $h_0(n)$.
   (c) Calculate τ, γ_v^2, α, and σ_r^2.
   (d) Calculate optimum $h_0(n)$ using approach in [2] and MSE (Mean Squared Error).
   (e) Is $(\text{MSE})^i \leq (\text{MSE})^{i-1}$? If yes, go to step (c) and if no, stop.

5. CONCLUSIONS

We actually used the scalar model for the VQ's and formulated MS reconstruction error as in Ref[2]. The optimal filter coefficients for the paraunitary, two band 4-tap case are shown in Table 4 and Table 5, along with comparision of the MS reconstruction error as obtained by Monte-Carlo simulation and as calculated using our model. We conclude that the scalar gain-plus-additive noise model provides an accurate represention of the optimum VQ in a subband codec and can be used as the basis for the design of optimum filter banks in presence of VQ's.

References

[1] H. Gharavi, "Subband coding algorithm for video application: Videophone to HDTV-conferencing," *IEEE Trans. Circuit Syst. Video Techno.*, vol. 1, No. 2, pp. 174-183, June 1991.

[2] R. A. Haddad and Kyusik Park, "Modeling, analysis and optimum design of quantized M-band filter banks," submitted to *IEEE Trans. Signal Processing*, Nov. 1993.

[3] Kyusik Park and R. A. Haddad, "Optimum subband filter bank design and compensation in presence of quantizers," in *Proc. 27th Asilomar Conf. on Sign. Syst. and Comp.*, Pacific Grove, CA, Nov. 1993.

[4] A. Gersho, "Asymptotically optimal block quantization," in *IEEE Trans. on Inf. Theory*, vol. IT-25, No. 4, pp. 373-380, July 1979.

[5] N. S. Jayant and P. Noll, *Digital Coding of Waveforms, Englewood Cliffs, N.J., Prentice-Hall*, 1984.

[6] J. H. Conway and N. J. A. Sloane "A lower bound on the average error of vector quantizer" in *IEEE Trans. on Inf. Theory*, vol. IT-31, No. 1, 1985

[7] Vladimir Cuperman, "Joint bit allocation and dimension optimization for vector transform quantization" in *IEEE Trans. on Inf. Theory*,vol. No. 1, pp. 302-305,Jan. 1993.

[8] Y. Linde, A. Buzo and R.M. Gray, Vector Quantization and Signal Compression, Kluwer Academic "An algorithm for vector quantizer design," in *IEEE Trans. on Comm.*, vol. COM-28, No. 1, pp 84-95, Jan. 1980.

[9] A. Gersho and R. Gray, *Vector Quantization and Signal Compression, Kluwer Academic Publishers*, 1992.

[10] A. N. Akansu, R. A. Haddad and H. Calgar, "The binomial QMF-wavelet transform for multiresolution signal decomposition," in *IEEE Trans. Signal Processing*, vol. 41 No. 1, pp13-19, Jan. 1993.

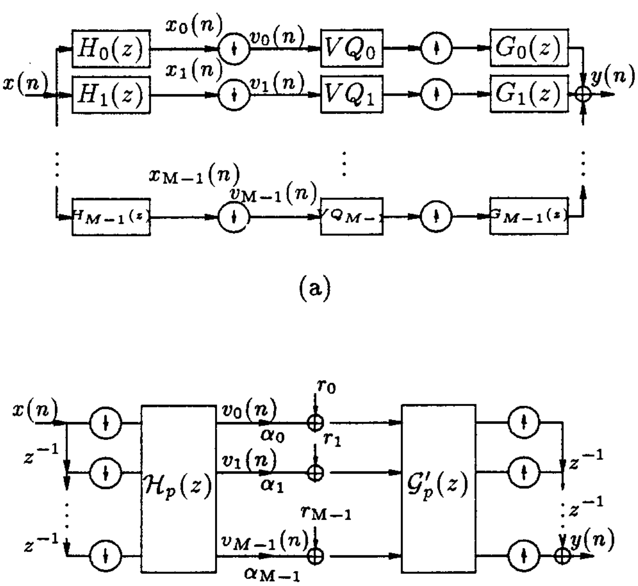

Figure 1: (a) M-band filter bank structure with vector quantizers, (b) polyphase equivalent structure.

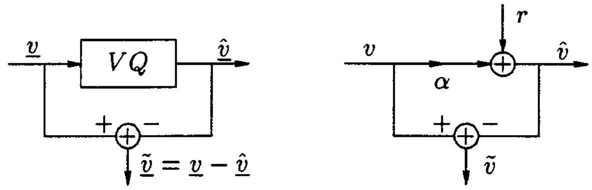

Figure 2: (a) vector quantizer, (b) equivalent scalar gain-plus-additive noise model.

Codebook	$E\{\tilde{v}(i)\}$	$E\{\tilde{v}(i)\hat{v}(i)\}$	$E\{v^2(i)\}$	$E\{\|\tilde{v}(i)\|^2\}$
N=32,k=8	-3.57E-4	8.97E-4	1.9651	0.1187
N=64,k=8	-2.25E-4	2.32E-3	1.9651	0.0861

Table 1. Simulation results using AR(1) signal (n=500,000 samples, ρ=0.95) for LBG Vector Quantizer.

Bit rate (B)	$E\{\|\tilde{v}(i)\|^2\}_{sim}$	$\sigma_{\tilde{v}}^2$
0.625	0.1187	0.1152
0.75	0.0861	0.0969

Table 2. Comparision $E\{|\tilde{v}(i)|^2\}_{sim}$ from test on VQ experimentally with $\sigma_{\tilde{v}}^2$ from equivalent scalar gain-plus-additive noise model theoretically.

B	0.25	0.5	0.75	1.0
k=8	0.5450	0.1499	-0.0853	-0.2434
k=12	0.1323	-0.2855	-0.5371	
k=16	-0.1476	-0.5780		

Table 3. Values of δ for AR(1) Gaussian input. B is the VQ rate in bit/sample, k is the VQ dimension.

B	B_0	B_1	MSE	MSE_{sim}
0.50	7	1	0.087293	0.088310
0.625	9	1	0.064425	0.064764
0.75	11	1	0.048300	0.049312
1.0	11	5	0.041292	0.042670

Table 4. Optimal Bit Allocation and MSE_{sim} and MSE_{model}. Inside training sequence n=500,000 and test sequence n=64,000 samples.

B	$h_0(0)$	$h_0(1)$	$h_0(2)$	$h_0(3)$
0.5	0.488488	0.832218	0.226195	-0.132770
0.625	0.488485	0.832219	0.226199	-0.132771
0.75	0.488486	0.832219	0.226198	-0.132771
1.0	0.488478	0.832221	0.226204	-0.132772

Table 5. Optimal filter coefficients for Paraunitary 4-tap 2-band Filter Bank.