Optimal Distortion Measures for the High Rate Vector Quantization of LPC Parameters

William R. Gardner[1,2*]

Bhaskar D. Rao[1]

[1]Department of ECE
University of California, San Diego
La Jolla, CA 92093-0407

[2]QUALCOMM Inc.
6455 Lusk Blvd.
San Diego, CA 92121

Abstract

This paper presents a class of quadratically weighted distortion measures which provide optimal performance for the high rate vector quantization (VQ) of linear predictive coding (LPC) parameters. It is shown that the quantization distortion of a high rate VQ converges to a quadratically weighted measure, where the quadratic weighting matrix is a "sensitivity" matrix, which is a generalization of the scalar sensitivity concept to the vector case. The sensitivity matrix is the second order term of the Taylor series expansion of the original distortion measure. Closed form expressions and computationally efficient methods for computing the sensitivity matrices of the different LPC parameterizations are given, which involve no numerical integration and can be implemented in real-time on modern DSP chips. In the general case, the "sum of sensitivity weighted scalar errors" is not equivalent to the original distortion measure. However the sensitivity matrix of the line spectral pair (LSP) frequencies is exactly diagonal, demonstrating that for LSPs only a "sum of sensitivity weighted scalar errors" will result in optimal performance.

Introduction

The linear predictive coding (LPC) model for speech is often used in modern mid-to-low rate speech compression systems, and much work has been done on quantization of the LPC filter coefficients. Recent work on high quality speech coding systems has focussed on determining LPC quantization schemes which require a minimal number of bits per vector of LPC coefficients, while achieving a low Log Spectral Distortion (LSD) between the unquantized LPC vectors and the quantized LPC vectors. Designing a quantizer which directly minimizes the overall LSD is difficult due to the complexity of the LSD measure, and typically quantizers are designed to minimize simpler distortion measures, such as mean squared (MSE) or weighted mean squared error (WMSE) between the original and quantized LPC coefficients, reflection coefficients, log area ratios (LARs), arcsine parameters, or LSP frequencies.

This paper introduces a class of quadratically weighted distortion measures which are optimal for high rate quantizers. It is shown that, at high rate, the distortion incurred by quantizing a vector approaches a simple, quadratically weighted error, where the quadratic weighting matrix is called the "sensitivity matrix." The diagonal terms of the sensitivity matrix are related to the scalar sensitivities of the parameters, and the off-diagonal terms are related to the interactions which occur when multiple parameters are quantized simultaneously. Closed form expressions and computationally efficient algorithms are given for computing the sensitivity matrices of LPC coefficients, reflection coefficients, LAR parameters, arcsine parameters, and LSP frequencies. Importantly, the sensitivity matrix for LSP frequencies is shown to be diagonal, implying that for LSP frequencies only a vector quantizer trained by minimizing an appropriate weighted WMSE measure (i.e. the "sum of sensitivity weighted scalar errors") will result in optimal performance.

Distortion Measures at High Rate

This section describes the properties of the quantization distortion in high rate vector quantization systems, and uses these properties to derive simple distortion measures which provide optimal performance at high rate.

Let $d(\boldsymbol{x}, \bar{\boldsymbol{x}})$ be a continuously differentiable distortion function which measures the distortion in quantizing the vector $\boldsymbol{x}$ to the vector $\bar{\boldsymbol{x}}$. For example, the LSD measure satisfies these conditions, where the LSD in dB^2 is given by

$$LSD(\boldsymbol{a}, \bar{\boldsymbol{a}}) = \frac{\beta}{2\pi}\int_{-\pi}^{\pi} (\ln(|A(\omega)|^2) - \ln(|\bar{A}(\omega)|^2))^2 d\omega.$$

Here $\beta = (10/\ln(10))^2$, $A(\omega) = 1 - \sum_{i=1}^{v} a_i e^{j\omega i}$, and $\bar{A}(\omega) = 1 - \sum_{i=1}^{v} \bar{a}_i e^{j\omega i}$. Performing a multidimensional

*This material is based upon work supported under a National Science Foundation Graduate Fellowship.

Taylor series expansion of $d(\boldsymbol{x}, \bar{\boldsymbol{x}})$ about $\boldsymbol{x} = \bar{\boldsymbol{x}}$, holding the second vector in $d(\boldsymbol{x}, \bar{\boldsymbol{x}})$ constant, and noting that the first and second terms of the Taylor expansion are zero, results in

$$d(\boldsymbol{x}, \bar{\boldsymbol{x}}) = \frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}})^T \boldsymbol{D}(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}) + O(||\boldsymbol{x} - \bar{\boldsymbol{x}}||^3).$$

where $\boldsymbol{D}(\bar{\boldsymbol{x}})$ is an n by n dimensional matrix with j, kth element defined by

$$D_{j,k}(\bar{\boldsymbol{x}}) = \left. \frac{\partial^2 d(\boldsymbol{x}, \bar{\boldsymbol{x}})}{\partial x_j \partial x_k} \right|_{\boldsymbol{x}=\bar{\boldsymbol{x}}}.$$

For most common distortion measures, including the LSD, $\boldsymbol{D}(\boldsymbol{x})$ is positive definite. As the rate of a quantizer increases, the distance between a vector and its closest quantization vector, as measured by $d(\boldsymbol{x}, \bar{\boldsymbol{x}})$, approaches zero, i.e. $||\boldsymbol{x} - \bar{\boldsymbol{x}}||$ gets small $\forall \boldsymbol{x}$ quantized to $\bar{\boldsymbol{x}}$, so

$$d(\boldsymbol{x}, \bar{\boldsymbol{x}}) \rightarrow \frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}})^T \boldsymbol{D}(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}); \quad \forall \boldsymbol{x} \text{ quantized to } \bar{\boldsymbol{x}}$$

$\boldsymbol{D}(\boldsymbol{x})$ is here termed the "sensitivity matrix," since its elements represent the relative sensitivities of quantizing the various parameters. The diagonal elements of the sensitivity matrix are the scalar sensitivities which represent the degree to which quantization error in a particular scalar parameter increases the overall distortion. The off-diagonal elements represent cross-sensitivity terms, which relate to the interactions which occur when multiple parameters are quantized simultaneously. In general, the sensitivity matrix is not diagonal, and the "sum of sensitivity weighted scalar errors" will not converge to the true distortion measure.

The above discussion shows that, in high rate vector quantizers, only the second order term in the Taylor series expansion of the distortion measure is relevant. If a high rate vector quantizer is trained by minimizing the expected value of the quadratic distortion measure

$$\hat{d}(\boldsymbol{x}, \bar{\boldsymbol{x}})) = \frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}})^T \boldsymbol{D}(\boldsymbol{x})(\boldsymbol{x} - \bar{\boldsymbol{x}}), \qquad (1)$$

the resulting quantizer will approach the optimal vector quantizer which has been trained by minimizing the original distortion measure, $d(\boldsymbol{x}, \bar{\boldsymbol{x}})$. There are several advantages which can be gained by using $\hat{d}$ rather than d in training and quantizing. First, it is easy to compute the centroid of a quadratic measure such as in equation 1, whereas it may be impossible to compute the centroid of the original error measure. This allows the quantizer to be built using the generalized Lloyd or LBG algorithms. Without the existence of an expression for the centroid of a region, constructing a quantizer becomes difficult or impossible. Second, the complexity incurred in computing the distortion measure for each vector in $\bar{\boldsymbol{X}}$ during quantization may be much less for the quadratic error measure than for the true error measure. This quantization scheme involves computing the actual $\boldsymbol{D}(\boldsymbol{x})$ during quantization only once per input vector. In the next section it is shown that this computation is simple for many parameter sets used in LPC parameter quantization schemes.

LPC Sensitivity Matrices

Denote the set of LPC filter coefficients corresponding to a vth order LPC filter by the vector

$$\boldsymbol{a} = [a_1 a_2 \ldots a_v]^T.$$

Most systems for compressing telephone bandwidth speech use a linear prediction filter of 10th order, i.e. $v = 10$. By taking derivatives, it can be shown that [1, 2]

$$\left. \frac{\partial^2 LSD(\boldsymbol{a}, \bar{\boldsymbol{a}})}{\partial \bar{a}_k \partial \bar{a}_l} \right|_{\bar{\boldsymbol{a}}=\boldsymbol{a}} = 4\beta R_A(k - l),$$

where $R_A(k)$ is the autocorrelation of the LPC filter impulse response, $h(n)$, i.e.,

$$R_A(k) = \sum_{n=0}^{\infty} h(n)h(n + k).$$

Hence, the sensitivity matrix with respect to the LPC coefficients, is

$$\boldsymbol{D}_A(\boldsymbol{a}) = 4\beta \boldsymbol{R}_A,$$

where

$$\boldsymbol{R}_A = \begin{bmatrix} R_A(0) & R_A(1) & \cdots & R_A(v-1) \\ R_A(1) & R_A(0) & & R_A(v-2) \\ \vdots & & \ddots & \vdots \\ R_A(v-1) & R_A(v-2) & \cdots & R_A(0) \end{bmatrix}$$

is the standard symmetric Toeplitz autocorrelation matrix. For the "autocorrelation" method of traditional LPC analysis, the elements of $\boldsymbol{R}_A$ are the computed autocorrelation values of the input frame of speech divided by the vth order prediction error. This means that, for LPC coefficients, the sensitivity matrix is easy to compute in real-time, and the computation involves no numerical integration techniques. Training a high rate VQ by minimizing this quadratically weighted error measure can be shown to be equivalent to training a VQ by minimizing the standard linear prediction error, and the Itakura-Saito measure. VQs trained by minimizing these measures, e.g. [3], will approach the optimal performance in LSD as the rate gets large.

There are several one-to-one vector functions which transform the vector of LPC coefficients, $\boldsymbol{a}$, into another length v vector, $\boldsymbol{p}$, e.g. reflection coefficients, LSP frequencies, etc. Denote the transformation from LPC vector $\boldsymbol{a}$ to the parameter set $\boldsymbol{p}$ by the function $\boldsymbol{p}(\boldsymbol{a})$ and the reverse transformation by $\boldsymbol{a}(\boldsymbol{p})$. Then,

$$\left. \frac{\partial^2 d(\boldsymbol{a}(\boldsymbol{p}), \boldsymbol{a}(\bar{\boldsymbol{p}}))}{\partial \bar{p}_k \partial \bar{p}_l} \right|_{\bar{\boldsymbol{p}}=\boldsymbol{p}}$$

$$= \sum_{m=1}^{v} \sum_{n=1}^{v} \frac{\partial a_m(\bar{p})}{\partial \bar{p}_k} \frac{\partial^2 d(a(p), \bar{a})}{\partial \bar{a}_m \partial \bar{a}_n} \frac{\partial a_n(\bar{p})}{\partial \bar{p}_l} \bigg|_{\substack{\bar{p}=p \\ \bar{a}=a(p)}} .$$

In more revealing matrix form, this says that the second derivative matrix with respect to p is

$$D_p(p) = J_p^T(p) D_A(a(p)) J_p(p) \qquad (2)$$

where $J(p)$ is the n by n Jacobian matrix of the transform $a(p)$, which has its j, kth element defined by

$$(J_p(p))_{j,k} = \frac{\partial a_j(\bar{p})}{\partial \bar{p}_k} \bigg|_{\bar{p}=p} \qquad (3)$$

Then, the distortion incurred by quantizing the vector p to $\bar{p}$ is, at high rate,

$$d(a(p), a(\bar{p})) \approx \frac{1}{2}(p - \bar{p})^T D_p(p)(p - \bar{p})$$
$$= \frac{1}{2}(p - \bar{p})^T J_p^T(p) D_A(a(p)) J_p(p)(p - \bar{p}).$$

This result holds for continuous one-to-one parameter mappings.

For the Log Spectral Distortion measure in particular,

$$D_A(a(p)) = 4\beta \hat{R}_A,$$

and the sensitivity matrices for reflection coefficients, LSPs, etc., can be found by multiplying the autocorrelation matrix on both sides by the appropriate Jacobian matrices.

Reflection Coefficient Sensitivities

The LPC coefficients can be transformed into the vector of reflection coefficients, k, and vice versa, using well known methods[4]. In order to determine the sensitivity matrix for the set of reflection coefficients, the Jacobian matrix of the mapping from the reflection coefficients to the LPC parameters, $J_k(k)$, must be computed. The matrix $J_k(k)$ can be partitioned as

$$J_k(k) = \left[j_{k_1} | j_{k_2} | \cdots | j_{k_v} \right] .$$

The "step-up" procedure [4] for converting the reflection coefficients to LPC coefficients can be written in matrix form as

$$\begin{bmatrix} 1 \\ -a \end{bmatrix} = K_v K_{v-1} \cdots K_1 \begin{bmatrix} 1 \\ o \end{bmatrix}, \qquad (4)$$

where

$$K_m = \left[\begin{array}{ccccc|c} 1 & & 0 & & k_m & \\ & 1 & & k_m & & \\ 0 & & \ddots & & 0 & 0 \\ & k_m & 0 & 1 & & \\ k_m & & & & 1 & \\ \hline & & 0 & & & 0 \end{array} \right] .$$

In the above expression, the matrix K_m is symmetric, the upper left matrix block is $m+1$ by $m+1$ and the lower right matrix block is $v - m$ by $v - m$. Taking derivatives of both sides of equation 4 with respect to k_m reveals that

$$\begin{bmatrix} 0 \\ -j_{k_m} \end{bmatrix} = K_v \cdots K_{m+1} L_m K_{m-1} \cdots K_1 \begin{bmatrix} 1 \\ o \end{bmatrix},$$

where

$$L_m = \left[\begin{array}{ccccc|c} & 0 & & & 1 & \\ & & & 1 & & 0 \\ & & \cdot & & & \\ & 1 & & 0 & & \\ 1 & & & & & \\ \hline & & 0 & & & 0 \end{array} \right] .$$

In this way, the vector j_{k_m} can be determined by performing the traditional step-up procedure, but replacing the iteration $a_i^{\{m\}} = a_i^{\{m-1\}} + k_m a_{m-i}^{\{m-1\}}$ with the iteration $a_i^{\{m\}} = a_{m-i}^{\{m-1\}}$ only at the mth stage. To determine the entire Jacobian matrix, this modified step-up procedure can be run for each m from 1 to v to determine each of the columns. Much of the computation at the mth step is the same as that at the $m-1$th step, and efficient algorithms for determining the entire Jacobian matrix can be developed.

Since the LAR and Arcsine parameters are one-to-one scalar functions of the reflection coefficients, i.e. $LAR_m = f(k_m)$ and $ARCSINE_m = g(k_m)$, the Jacobian matrices of the transformations from LARs and Arcsines to LPC parameters are simply the Jacobian matrix of the transformation from reflection coefficients to LPC parameters, with the columns of the Jacobian matrix scaled by $1/f'(k_m)$ and $1/g'(k_m)$ respectively [1, 2]. Thus, the complexity required to compute the sensitivity matrices of the LAR and Arcsine parameters is similar to that required to compute the sensitivity matrix of the reflection coefficients.

Line Spectral Pair Sensitivities

The LSP frequencies are another set of parameters often used for LPC quantization. The LSP frequencies are the angular frequencies of the roots of the polynomials

$$P(z) = A(z) + z^{-(v+1)} A(z^{-1})$$

and

$$Q(z) = A(z) - z^{-(v+1)} A(z^{-1}).$$

which exist on the unit circle in the first and second quadrant of the complex plane. Let the LSP frequencies be denoted by the vector ω with elements $\omega_1, \omega_2, \cdots, \omega_v$, so the roots of $P(z)$ correspond to the odd indices and the roots of $Q(z)$ correspond to the even indices. Then, the mapping from the set of LSP frequencies to $A(\omega)$ is given by

$$A(\omega) = \frac{1}{2}(P(\omega) + Q(\omega))$$

where

$$P(\omega) = (1+e^{-j\omega}) \prod_{i\,odd} \left(1 - 2\cos\omega_i e^{-j\omega} + e^{-2j\omega}\right)$$

and

$$Q(\omega) = (1-e^{-j\omega}) \prod_{i\,even} \left(1 - 2\cos\omega_i e^{-j\omega} + e^{-2j\omega}\right).$$

For notational purposes, define $\tilde{p}_0(n) = \delta(n)+\delta(n-1)$, $\tilde{p}_i(n) = \delta(n)-2\cos(\omega_{2i-1})\delta(n-1)+\delta(n-2)$, $\tilde{q}_0(n) = \delta(n)-\delta(n-1)$, and $\tilde{q}_i(n) = \delta(n)-2\cos(\omega_{2i})\delta(n-1)+\delta(n-2)$, and define $\tilde{p}_i(\omega)$ and $\tilde{q}_i(\omega)$ as the discrete time Fourier transforms of $\tilde{p}_i(n)$ and $\tilde{q}_i(n)$, so

$$P(\omega) = \prod_{i=0}^{v/2} \tilde{p}_i(\omega) \qquad \text{and} \qquad Q(\omega) = \prod_{i=0}^{v/2} \tilde{q}_i(\omega).$$

The sensitivity matrix for the LSP frequencies is given by

$$\boldsymbol{D}_\omega(\boldsymbol{\omega}) = 4\beta \boldsymbol{J}_\omega^T(\boldsymbol{\omega}) \boldsymbol{R}_A \boldsymbol{J}_\omega(\boldsymbol{\omega}), \tag{5}$$

where $\boldsymbol{J}_\omega(\boldsymbol{\omega})$ is the Jacobian matrix of the transformation from LSP frequencies, $\boldsymbol{\omega}$, to LPC coefficients. If the n, ith element of the Jacobian matrix is denoted by

$$j_i(n) = \left.\frac{\partial a_n(\bar{\boldsymbol{\omega}})}{\partial \bar{\omega}_i}\right|_{\bar{\boldsymbol{\omega}}=\boldsymbol{\omega}}, 1 \leq n, i \leq v$$

then simple differentiation reveals the discrete time Fourier pair

$$j_i(n) \Leftrightarrow J_i(\omega) = \frac{\partial A(\omega)}{\partial \omega_i} = \begin{cases} \frac{1}{2}\frac{\partial P(\omega)}{\partial \omega_i}; i\ odd \\ \frac{1}{2}\frac{\partial Q(\omega)}{\partial \omega_i}; i\ even \end{cases}$$

$$= \begin{cases} \sin(\omega_i) e^{-j\omega} \prod_{j=0:j\neq(i+1)/2}^{v/2} \tilde{p}_j(\omega); i\ odd \\ \sin(\omega_i) e^{-j\omega} \prod_{j=0:j\neq i/2}^{v/2} \tilde{q}_j(\omega); i\ even \end{cases} \tag{6}$$

Thus, the elements of the Jacobian matrix for the LSP to LPC transformation can be found by determining the coefficients in equation 6. For example, the coefficient corresponding to the $e^{-3j\omega}$ term in $J_1(\omega)$ is equal to $j_1(3)$. Given the Jacobian matrix and the autocorrelation matrix $\boldsymbol{R}_A$, the sensitivity matrix for the LSP frequencies can be evaluated. The overall computation of the optimal weightings for the LSP frequencies requires no divisions, square roots, or power computations, which are typically time consuming operations in real DSP implementations. The computation requires about 1000 multiply-accumulates per frame, which can be easily performed in real-time on modern DSP chips. Source code for computing these weightings can be found in [1, 2].

The following theorem states an important and surprising result.

Theorem
For $k \neq l$, $1 \leq k,l, \leq v$,

$$(\boldsymbol{D}_\omega(\boldsymbol{\omega}))_{k,l} = \left.\frac{\partial LSD(\boldsymbol{a}(\boldsymbol{\omega}), \boldsymbol{a}(\bar{\boldsymbol{\omega}}))}{\partial \bar{\omega}_k \partial \bar{\omega}_l}\right|_{\bar{\boldsymbol{\omega}}=\boldsymbol{\omega}} = 0,$$

where ω_i is the ith LSP frequency. This theorem states that the sensitivity matrix for the LSP frequencies is diagonal. The proof of this theorem is given in [1, 2].

An important consequence of this theorem is that, at high rate, *the sum of sensitivity weighted scalar errors in LSP frequencies is equivalent to LSD!* This implies that a high rate VQ constructed by minimizing this relatively simple measure will achieve the minimum possible LSD. Also, this makes the computation of the LSP sensitivity matrix simple, since only the diagonal elements need to be computed. The sensitivity matrices of the LPC coefficients, reflection coefficients, LAR parameters, and arcsine parameters are not diagonal, and thus the "sum of sensitivity weighted scalar errors" in these parameter sets are not equivalent to the LSD at high rate.

The optimal LSP weightings have been tested on split 3-3-4 VQs of LSP parameters, and the performance of the VQs trained using these weightings is superior to that of VQs trained by minimizing the distortion measures used in [5, 6]. The full results of these experiments are reported in [1, 2]. At high rates leading to average LSDs of 1-2 dB^2, VQs trained by minimizing the optimal measure presented here gave performances 0.5-1.0 bits better than VQs trained by minimizing the measures used in [5, 6], and 2.5 bits better that VQs trained by minimizing the unweighted sum of squared errors in the LSP frequencies. These results indicate that although the optimal measure produces results superior to those produced by the measures proposed in [5, 6], these other measures are not far from the optimal measure. A theoretical analysis of the performance of VQs which are trained by minimizing suboptimal distortion measures, such as in [5, 6], has also been developed, and can be found in [1, 2, 7].

References

[1] W. R. Gardner and B. D. Rao, "Theoretical Analysis of the High Rate Vector Quantization of LPC Parameters," *submitted to IEEE Trans. on Speech and Audio.*

[2] W. R. Gardner, *Modeling and Quantization Techniques for Speech Compression Systems.* PhD thesis, University of California, San Diego, 1994.

[3] A. Buzo, A. H. Gray, R. M. Gray, and J. D. Markel, "Speech Coding based upon Vector Quantization," *IEEE Trans. on ASSP,* Oct. 1980.

[4] J. D. Markel and A. H. Gray, *Linear Prediction of Speech.* New York: Springer-Verlag, 1976.

[5] R. Laroia, N. Phamdo, and N. Farvardin, "Robust and Efficient Quantization of Speech LSP Parameters using Structured Vector Quantizers," in *Proceedings of ICASSP,* 1991.

[6] K. K. Paliwal and B. Atal, "Efficient Vector Quantization of LPC Parameters at 24 bits/frame," *IEEE Trans. on SAP,* vol. 1, Jan. 1993.

[7] W. R. Gardner and B. D. Rao, "Analysis of High Rate LPC Vector Quantizers Designed by Minimizing Suboptimal Distortion Measures," in *Proceedings of the IEEE Asilomar Conference on Signals, Systems, and Computers,* 1994.