14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

LOW-COMPLEXITY WIDEBAND LSF QUANTIZATION BY PREDICTIVE KLT CODING AND GENERALIZED GAUSSIAN MODELING

Marie Oger[1], *Stéphane Ragot*[1] *and Marc Antonini*[2]

[1]France Télécom R&D/TECH/SSTP, Av. Pierre Marzin, 22307 Lannion Cedex
[2]Lab. I3S-UMR 6070 CNRS and Univ. of Nice Sophia Antipolis, rte des Lucioles, 06903 Sophia Antipolis
E-mail: {marie.oger, stephane.ragot}@francetelecom.com, am@i3s.unice.fr

ABSTRACT

In this paper we present a new model-based coding method to represent the linear-predictive coding (LPC) parameters of wideband speech signals (sampled at 16 kHz). The LPC coefficients are transformed into line spectrum frequencies (LSF) and quantized by switched AR(1)/MA(1) predictive Karhunen-Loeve transform (KLT) coding. Compared to previous work, the main novelty lies in the use of improved quantization to represent the (transformed) prediction error of LSF parameters. Generalized Gaussian modeling is applied for this purpose. We review existing methods to fit the free model parameter of a generalized Gaussian model to real data and show that that the distribution of the prediction error for LSF parameters is indeed very close to Laplacian. Experimental results show that the proposed LSF quantization method has a performance close to classical vector quantization (AMR-WB LPC quantization) at 36 and 46 bits per frame with a much lower complexity for both design and operation.

1. INTRODUCTION

A parametric approach based on Gaussian mixture models (GMM) has been developed for the vector quantization (VQ) of linear-predictive coding (LPC) parameters [1, 2]. This approach has brought interest in the design of model-based quantization methods as opposed to standard constrained VQ requiring stochastic codebook training based on a given source database. Several variants of GMM-based LPC quantization have been proposed including predictive methods [2, 3].

In this paper, we study a simplified framework of GMM-based VQ in which predictive Karhunen-Loeve transform (KLT) coding is used. This is equivalent to source coding with singular value decomposition (SVD) or principal component analysis (PCA). This also corresponds to the special case of predictive GMM-based VQ with a single Gaussian component in the GMM. The motivations for setting such a restriction are as follows:

- The complexity of GMM-based VQ is roughly linear in the GMM order. This is true for both storage requirement and computational cost. For instance, typical LSF quantization methods [2, 3] use a GMM of 4 or 8, which implies a significant complexity overhead compared to KLT coding with one Gaussian component.
- The bit mapping in GMM-based VQ is usually not optimized to be robust against bit errors. The related bit allocation methods [2] distribute a certain amount of codewords among Gaussian components, in such a way that a single bit error in the overall quantization index can have a dramatic impact on the reconstructed LSF parameters. KLT coding with one Gaussian component can avoid this problem.

Although KLT coding has some advantages in terms of complexity and robustness against bit errors, its performance is lower than that of GMM-based VQ [4]. To compensate for this limitation, we propose to refine the source modeling by using non-Gaussian models for transformed LSF prediction errors.

The main contribution of this work lies in the application of generalized Gaussian modeling to improve the quantization performance of predictive KLT coding. Generalized Gaussian modeling has been used extensively in image and video coding – see for instance [5, 6]. However, its application to speech coding is quite new. We review existing methods to fit the free model parameter of a generalized Gaussian model to real data and show that the prediction error density function for LSF parameters is very close to Laplacian. Futhermore we show that generalized Gaussian modeling brings a non-neglegible performance gain over simple Gaussian modeling for predictive KLT coding of LSF parameters.

This paper is organized as follows. We review existing LPC quantization methods based on KLT coding in Section 2. We distinguish memoryless and predictive variants and give an outline of the method developed in this work. The generalized Gaussian distribution is defined in Section 3, where we also review estimation methods for the free parameter of a generalized Gaussian model. The proposed predictive model-based quantization method is described in Section 4. Experimental results for wideband LSF quantization are presented and discussed in Section 5 before concluding in Section 6.

2. LPC QUANTIZATION BASED ON KLT CODING

We follow here the notations of [1]. The probability density function (pdf) of LSF vectors $\mathbf{x}$ in dimension n can be modeled [2] by a Gaussian mixture model of order M given by

$$f(\mathbf{x}|\Theta) = \sum_{i=1}^{M} \rho_i f_i(\mathbf{x}|\theta_i), \quad (1)$$

$$\text{where } f_i(\mathbf{x}|\theta_i) = \frac{1}{\sqrt{(2\pi)^n |\det(\mathbf{C}_i)|}} \, e^{-\frac{1}{2}(\mathbf{x}-\mu_i)^T \mathbf{C}_i^{-1}(\mathbf{x}-\mu_i)}, \quad (2)$$

with the following constraints: $\rho_i > 0$ and $\sum_{i=1}^{M} \rho_i = 1$. The dimension n is usually set to 16 or 18 for wideband speech signal (sampled at 16 kHz). The set of GMM parameters is given by

$$\Theta = \{\rho_1, \cdots, \rho_M, \mu_1, \cdots, \mu_M, \mathbf{C}_1, \cdots, \mathbf{C}_M\},$$

where ρ_i, μ_i and $\mathbf{C}_i$ are respectively the weight (a priori probability), the mean vector and the covariance matrix of the i-th GMM component. For a given source database, Θ is usually estimated using the E-M algorithm [7].

2.1 (Memoryless) mean-removed KLT coding

Memoryless GMM-based quantization is illustrated in Figure 1 (a). For an input LSF vector $\mathbf{x}$ of dimension n, the quantized LSF vector $\hat{\mathbf{x}}$ is selected among M candidates $\widehat{\mathbf{x}}^{(i)}$, with $i = 1, \cdots, M$, by minimizing a distortion criterion:

$$\hat{\mathbf{x}} = \hat{\mathbf{x}}^{(j)} \text{ where } j = \arg\min_{i=1,\cdots,M} d(\mathbf{x}, \widehat{\mathbf{x}}^{(i)}). \quad (3)$$

The selection criterion d can be the log-spectral distortion [2] or a simple weighted Euclidean distance [4]. The candidate $\hat{\mathbf{x}}^{(i)}$ is the representative of $\mathbf{x}$ in the i-th GMM component (or cluster). The

(a) Mean-removed KLT coding

(b) Predictive KLT coding

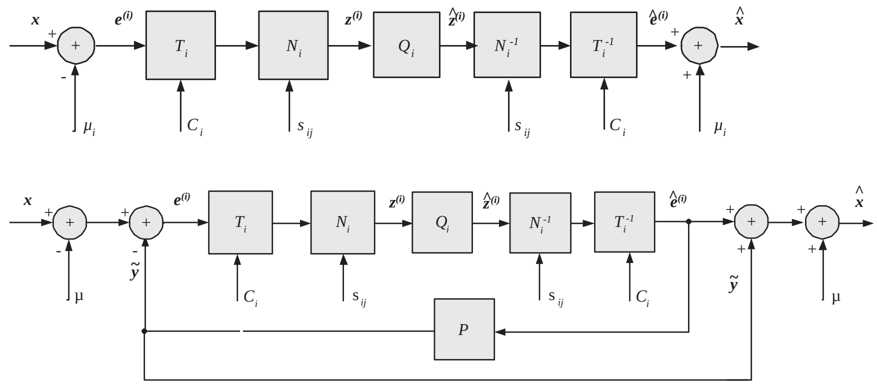

Figure 1: Principle of GMM based VQ: T_i and N_i corresponds respectively to KLT transform and normalization by standard deviation

candidates $\hat{\mathbf{x}}^{(i)}$ are computed in [2] by mean-removed Karhunen-Loeve transform (KLT) coding using the parameters μ_i and $\mathbf{C}_i$ of the i-th GMM component. The KLT matrix T_i and normalization factors σ_{ij} are derived from the eigenvalue decomposition of covariance matrices $\mathbf{C}_i$ [2]:

$$\mathbf{C}_i = T_i \; diag(\sigma_{i1}^2, \cdots, \sigma_{in}^2) \; T_i^T \tag{4}$$

where $\sigma_{i1}^2 \geq \cdots \geq \sigma_{in}^2$ are the eigenvalues of $\mathbf{C}_i$ and the matrix T_i comprises the eigenvectors of $\mathbf{C}_i$. In Figure 1 (a) the block N_i refers to the normalization by σ_{ij}. Usually the quantization Q_i of the source $\mathbf{z}^{(i)}$ relies on a Gaussian model assumption – for instance in [2] non-uniform scalar quantization with high-rate-optimized Gaussian companding is used.

2.2 Predictive KLT coding

Predictive KLT coding is illustrated in Figure 1 (b). In this case we define the mean-removed input LSF vector $\mathbf{y}$ of dimension n where $\mathbf{y} = \mathbf{x} - \mu$ and μ is the long-term mean of $\mathbf{x}$. The GMM model approximates the pdf of the prediction error $\mathbf{e} = \mathbf{y} - \tilde{\mathbf{y}}$. In general a moving-average (MA) or autoregressive (AR) linear predictor P is used to compute the prediction $\tilde{\mathbf{y}}$. The actual quantization of $\mathbf{e}$ follows the GMM-based VQ described previously for the memoryless case.

2.3 Outline of the proposed method

The quantization method developed in this work is similar to the predictive GMM-based VQ of Figure 1 (b). The main difference is that we choose to work with only one GMM component to reduce complexity and improve robustness against bit errors. Furthermore we apply generalized Gaussian modeling on each component z_j of the source $\mathbf{z}$ (which corresponds to the normalized KLT-transformed prediction error). We apply an efficient quantization (model-based Lloyd-Max scalar quantization) using this non-Gaussian model for $\mathbf{z}$.

3. GENERALIZED GAUSSIAN MODELING

3.1 Definition

The pdf of a zero-mean generalized Gaussian random variable z with standard deviation σ is given by:

$$g_\alpha(z) = \frac{A(\alpha)}{\sigma} e^{-\left|B(\alpha)\frac{z}{\sigma}\right|^\alpha}, \tag{5}$$

where α is a shape parameter describing the exponential rate of decay and the tail of the density function. The parameters $A(\alpha)$ and $B(\alpha)$ are given by:

$$A(\alpha) = \frac{\alpha B(\alpha)}{2\Gamma(1/\alpha)} \quad \text{and} \quad B(\alpha) = \sqrt{\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}}, \tag{6}$$

where $\Gamma(.)$ is the Gamma function defined as:

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha+1} \, dt. \tag{7}$$

The Laplacian and Gaussian distributions correspond to the special case $\alpha = 1$ and 2 respectively. The generalized Gaussian model is useful to approximate symetric unimodal distributions.

3.2 Estimation of the shape parameter α

Estimation methods of the shape parameter α are reviewed here. We classify estimation methods in "closed loop" and in "open loop". "Closed loop" methods estimate α by minimizing a distance criterion between data and model, while "open loop" methods provide an estimate of α without any distance criterion.

We assume that we are given samples $\{z_1, z_2, \ldots, z_N\}$ from a random variable z of pdf $g(z)$. We estimate the shape parameter α of a generalized Gaussian model $g_\alpha(z)$ approximating $g(z)$. To do so the normalized histogram $\tilde{p}(z)$ of $z_1, z_2, \ldots, z_N$ is compared with a sampled version of $g_\alpha(z)$. The sampling step size is defined as a constant q. The estimation procedure is illustrated in Figure 2. Without loss of generality the random variable z is supposed to be of unit variance and zero-mean.

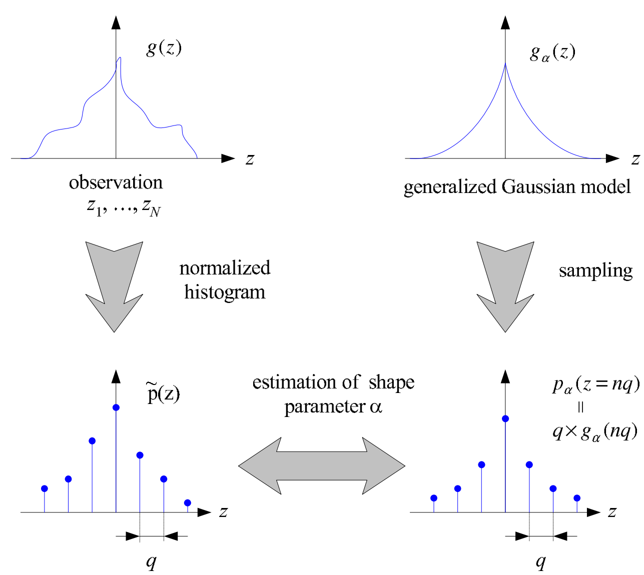

Figure 2: Estimation procedure for the shape parameter α.

3.3 Estimation methods in closed loop

3.3.1 χ^2 quantity [8]

The χ^2 quantity evaluates a kind of distance between two probability density functions. We use here a χ^2 distance given by:

$$\chi^2(\alpha) = \sum_{z=\dots,-q,0,q,\dots} \frac{(\tilde{p}(z) - p_\alpha(z))^2}{\tilde{p}(z) + p_\alpha(z)}, \tag{8}$$

where $p_\alpha(z) = g_\alpha(z) \times q$.
The estimated shape parameter is obtained by minimization :

$$\hat{\alpha} = \arg\min_\alpha \chi^2(\alpha) \tag{9}$$

3.3.2 Kolmogorov-Smirnov statistic [8]

The Kolmogorov-Smirnov statistic is defined as:

$$KS(\alpha) = \max_z |G(z) - G_\alpha(z)| \tag{10}$$

where $G(z)$ and $G_\alpha(z)$ are the distributions:

$$G(z) = \sum_{n=-\infty}^{z/q} \tilde{p}(nq) \quad \text{and} \quad G_\alpha(z) = \sum_{n=-\infty}^{z/q} p_\alpha(nq) \tag{11}$$

The estimated shape parameter $\hat{\alpha}$ is found by minimization:

$$\hat{\alpha} = \arg\min_\alpha KS(\alpha) \tag{12}$$

3.3.3 Kullback-Leibler divergence [9]

The Kullback-Leibler divergence is given by:

$$D(\tilde{p}||p_\alpha) = -\sum_{z=\dots,-q,0,q,\dots} p_\alpha(z) \log \frac{\tilde{p}(z)}{p_\alpha(z)} \tag{13}$$

The estimated shape parameter $\hat{\alpha}$ is obtained by minimizing this measure between the histogram $\tilde{p}$ and the generalized Gaussian model:

$$\hat{\alpha} = \arg\min_\alpha D(\tilde{p}||p_\alpha) \tag{14}$$

3.4 Estimation methods in open loop

Several methods are omitted here, e.g. ML estimation [5].

3.4.1 Estimation based on kurtosis (κ_α)

For a generalized gaussian source z of shape parameter α, the kurtosis κ_α is given by [8]:

$$\kappa_\alpha = \frac{E(z^4)}{E(z^2)^2} = \frac{\Gamma(5/\alpha)\Gamma(1/\alpha)}{\Gamma(3/\alpha)^2} \tag{15}$$

It can be verified that $\log \kappa_\alpha$ is approximatively a linear function of $1/\alpha$ [6]:

$$\log \kappa_\alpha \approx \frac{1.447}{\alpha} + 0.345 \tag{16}$$

Based on this approximation, the shape parameter α can be estimated as [6]:

$$\hat{\alpha} \approx \frac{1.447}{\ln \hat{\kappa} - 0.345}, \tag{17}$$

where $\hat{\kappa}$ is estimated from the data:

$$\hat{\kappa} = \frac{\frac{1}{n}\sum_{i=1}^n z_i^4}{\left(\frac{1}{n}\sum_{i=1}^n z_i^2\right)^2} \tag{18}$$

3.4.2 Method proposed by Mallat [10]

A relation between the variance $E(z^2)$, the mean of the absolute value $E(|z|)$ and the shape parameter α is given by [6]:

$$\frac{E(|z|)}{\sqrt{E(z^2)}} = \frac{\Gamma(2/\alpha)}{\sqrt{\Gamma(1/\alpha)\Gamma(3/\alpha)}} = F(\alpha) \tag{19}$$

The shape parameter α can be estimated as:

$$\hat{\alpha} = F^{-1}\left(\frac{\hat{m}_1}{\sqrt{\hat{m}_2}}\right) \tag{20}$$

where $\hat{m}_1 = \frac{1}{n}\sum_{i=1}^n z_i^2$ and $\hat{m}_2 = \frac{1}{n}\sum_{i=1}^n |z_i|$.

3.4.3 Estimation based on differential entropy (h_α)

The differential entropy of a generalized Gaussian distribution is given by [6]:

$$h_\alpha = \frac{1}{2}\log_2\left[\frac{4\Gamma(1/\alpha)^3}{\alpha^2\Gamma(3/\alpha)}\right] + \frac{1}{\alpha \log 2} \tag{21}$$

Based on high rate quantization theory, it can be shown that the entropy rate R of the quantized random variable z and the differential entropy $h(z)$ are related as follows :

$$R \approx h(z) - \log_2 q \tag{22}$$

The estimated shape parameter $\hat{\alpha}$ is given by [11]:

$$\hat{\alpha} = h_\alpha^{-1}\left[\hat{R} + \log_2 q\right] \tag{23}$$

where R is the estimated entropy rate of z:

$$\hat{R} = -\sum_{z=\dots,-q,0,q,\dots} \tilde{p}(z) \log_2 \tilde{p}(z) \tag{24}$$

4. PREDICTIVE KLT CODING WITH GENERALIZED GAUSSIAN MODELING

The coding method developed in this work is illustrated in Figure 3. It follows the principle of predictive KLT coding explained in Section 2. We compute the mean-removed source $\mathbf{y} = \mathbf{x} - \mu$ where μ is the long-term average of $\mathbf{x}$. The estimate $\tilde{\mathbf{y}}$ is given by the switched AR(1)/MA(1) predictor P, for which:

- The MA(1) prediction matrix is set to $diag(1/3, \cdots, 1/3)$ (as in AMR-WB LPC quantization [12]),
- The AR(1) (diagonal) prediction matrix is optimized in closed-loop.

This switched predictor, which is a kind of safety-net VQ, yields a performance close to AR(1) with improved robustness against channel impairments. Due to switching between AR(1) and MA(1), one bit is sent to indicate the selected prediction ; the selection is performed in open-loop by minimizing the energy of the prediction error $\mathbf{e} = \mathbf{y} - \tilde{\mathbf{y}}$. The KLT matrix T and normalization N by standard deviation σ_j are derived from the eigenvalue decomposition of the covariance of $\mathbf{e}$. Two different quantization methods are used: either non-uniform scalar quantization with high-rate Gaussian companding [2] or model-based Lloyd-Max quantization. In the latter case we optimize Lloyd-Max centroids *based on a generalized Gaussian model*, which is different from classical codebook training with a given source database.

4.1 Model-based scalar quantization

4.1.1 Companded scalar quantization for a Gaussian model [2]

As shown in Figure 3, the source $\mathbf{z} = (z_1, \cdots, z_n)$ can be encoded by companded scalar quantization. Under the high-rate and

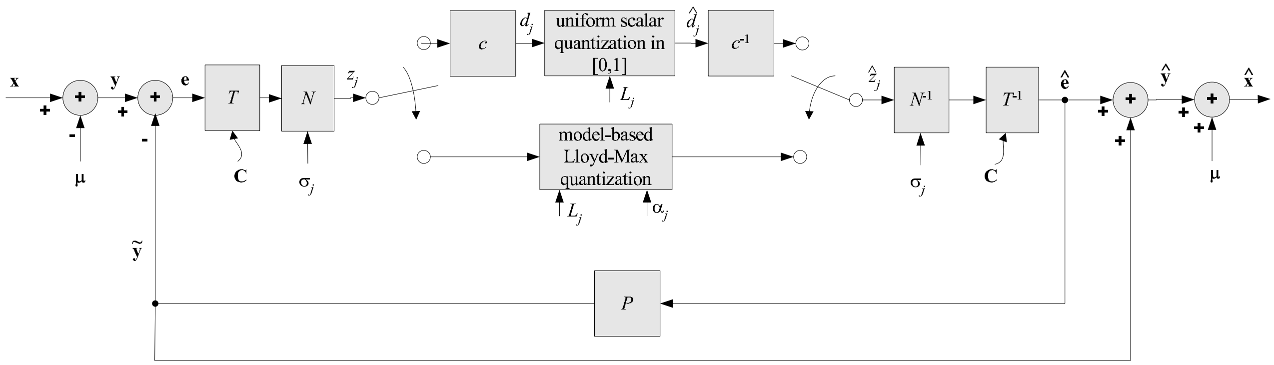

Figure 3: Predictive KLT coding with model-based non-uniform scalar quantization.

Gaussian assumptions, $\alpha_{j=1\dots n} = 2$, the optimal companding for a unit-variance random variable z is given by [3]:

$$c(z) = \frac{1}{2}(1 + erf(z/\sqrt{6})), \qquad (25)$$

where erf is the error function $erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2}\, dt$. The source $\mathbf{z} \in [-\infty, +\infty]^n$ is mapped into a source $\mathbf{d} \in [0,1]^n$ with $\mathbf{d} = (c(z_1), \cdots, c(z_N))$. Uniform scalar quantization in $[0,1]^n$ is applied to $\mathbf{d}$: if d_j is quantized with $L_j \geq 1$ scalar levels, the reconstruction $\hat{d}_j$ is given by $\hat{d}_j = ([d_j L_j - \frac{1}{2}] + \frac{1}{2})/L_j$ where $[.]$ denotes the rounding to the nearest integer.

4.1.2 Lloyd-Max quantization for a generalized Gaussian pdf

An alternative approach is shown in Figure 3, where the source $\mathbf{z} = (z_1, \cdots, z_n)$ is encoded by Lloyd-Max quantization [13]. We optimize here the Lloyd-Max quantizer using the theoretical pdf of the source model. This model-based approach allows to circumvent the costly training of stochastic codebooks using a database. This makes the quantization more versatile, as it can be easily reoptimized by updating the generalized Gaussian model parameters.

The decision thresholds t_i, and the reconstruction levels s_i of the quantizer are found by the following iterative process until convergence of s_i:

$$t_i = \frac{1}{2}(s_i + s_{i-1}) \qquad i = 2, \dots, L_j \qquad (26)$$

$$\text{with } t_0 = -\infty \text{ and } t_{L_j+1} = +\infty$$

$$s_i = \frac{\int_{t_i}^{t_{i+1}} z\, g_{\alpha_j}(z)\, dz}{\int_{t_i}^{t_{i+1}} g_{\alpha_j}(z)\, dz} \qquad i = 1, \dots, L_j \qquad (27)$$

where L_j is the allocated number of levels.

4.2 Bit allocation

The problem of bit allocation to several generalized Gaussian random variables has been studied in [6]. Given an allocation per sample $R_{tot} = \frac{1}{n}\sum_{j=1}^{n} R_j$, the scalar level quantization $L_{j=1\dots n}$ for a random variable $z_{j=1\dots n}$ of shape parameter $\alpha_{j=1\dots n}$ is defined as:

$$L_j = 2^{R_j} = 2^{-\frac{1}{2}\log_2\left(\frac{\lambda}{n}\right) + \frac{1}{2}\log_2(2\log(2)\mathscr{F}(\alpha_j))}, \qquad (28)$$

$$\text{where} \quad \mathscr{F}(\alpha_j) = \frac{\Gamma(1/\alpha_j)^3}{3\alpha_j^2 \Gamma(3/\alpha_j)} e^{\frac{2}{\alpha_j}} \qquad (29)$$

$$\text{and} \quad \lambda = 2^{-R_{tot}} 2\log(2) \prod_{i=1}^{n} \left(\frac{\mathscr{F}(\alpha_j)}{n}\right). \qquad (30)$$

The scalar quantization levels $L_{j=1\dots n}$ are rounded and further optimized by a greedy bit algorithm similar to [3]. Note that the optimization of L_j can be easily constrained to take into account the robustness against bit errors.

5. EXPERIMENTAL RESULTS FOR WIDEBAND LSF QUANTIZATION AND DISCUSSION

5.1 Experimental setup

The experimental setup is similar to that of [3]. The database used for this work is the NTT-AT wideband speech material (sampled at 16 kHz) without silence frames . The downsampling to 12.8 kHz and linear-predictive analysis of AMR-WB [12] is used to extract LSF vectors of dimension $n = 16$. A training database comprising 607,386 LSF vectors was extracted, and a separate test database of 202,112 LSF vectors was also built.

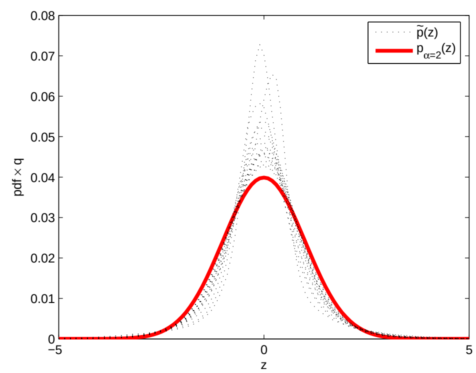

Figure 4: Normalized histograms $\tilde{p}(z_j)$ of data samples $z_{j=1\dots16}$ compared with a Gaussian model with $p_{\alpha=2}(z) = g_{\alpha=2}(z) \times q$.

5.2 Shape parameters of generalized Gaussian models

Figure 4 compares the normalized histograms $\tilde{p}(z_j)$ of samples $z_{j=1\dots16}$ with a Gaussian distribution, while Table 1 presents the values of the shape parameter α_j for the estimation methods using the same step size q. These results show that generalized Gaussian modeling provides a better approximation than a Gaussian model. Furthermore, it turns out that methods in "closed loop" give better results than the "open loop" ones. The best estimation in "open loop" is the method proposed by Mallat. For the next part, the estimation method of the shape parameter will use the method proposed by Mallat. This allows to optimize design parameters in a very efficient way. The shape parameters $\alpha_{j=1\dots16}$ are between 0.89 and 1.70 with the method proposed by Mallat.

5.3 Spectral distortion

The performance of LSF quantization is evaluated with the usual spectral distortion (SD) [14]. The SD statistics obtained for switch AR(1)/MA(1) predictive LSF quantization can be found in Table 2. The total bit allocation is 36 or 46 bits ; one bit is used to indicate the predictor switch, while the rest – 35 or 45 bits – is allocated to quantize the transformed prediction error $\mathbf{z}$.

Table 1: Shape parameters $\alpha_{j=1...16}$ estimated on the training database of vectors $\mathbf{z}$.

j	χ^2	KS	KL	Mallat	κ_α	h_α
1	1.04	1.11	1.03	0.98	0.93	0.91
2	1.14	1.31	1.14	1.11	1.06	1.05
3	1.15	1.22	1.15	1.12	1.05	1.08
4	1.18	1.28	1.17	1.14	1.05	1.09
5	1.19	1.16	1.19	1.15	1.06	1.11
6	1.35	1.42	1.35	1.32	1.20	1.24
7	1.39	1.32	1.39	1.37	1.26	1.31
8	1.44	1.44	1.43	1.40	1.25	1.32
9	1.50	1.50	1.50	1.49	1.37	1.42
10	1.48	1.52	1.47	1.46	1.33	1.39
11	1.47	1.46	1.47	1.46	1.33	1.38
12	1.52	1.54	1.52	1.51	1.40	1.45
13	1.63	1.63	1.62	1.62	1.49	1.54
14	1.66	1.68	1.66	1.65	1.51	1.57
15	1.71	1.73	1.71	1.70	1.56	1.61
16	0.95	0.99	0.94	0.89	0.86	0.87

The results show that model-based Lloyd-Max (LM) quantization improves the performance compared to companded scalar quantization (SQ). This is due to the fact that the related companding is optimized for a Gaussian source coded at high bit rates; yet, the high-bit rate assumption is not valid for practical LPC quantization. In the case of LM quantization with $\alpha = 2$ the gain in average SD over companded SQ is around 0.03-0.11 dB but the amount of outliers is slightly increased. In the case of LM quantization with optimal α the gain in average SD is more significant (around 0.05-0.13 dB) and the amount of outliers is significantly reduced.

If we compare the performance of Lloyd-Max quantization for $\alpha = 2$ and optimal α, it turns out that generalized Gaussian modeling still brings a non-neglegible improvement.

The performance of the LPC quantizer used in AMR-WB is also reported. It shows that the performance of the proposed model-based coding is close (but slightly inferior to) classical constrained VQ. The performance gap between AMR-WB and LM with optimized α is around 0.07-0.10 dB in average SD.

Table 2: Results for switched AR(1)/MA(1) predictive LSF quantization vs AMR-WB LPC quantization: comparison between companded SQ and model-based LM quantization.

(a) Results at 36 bits per frame

Quantization methods	avg. SD (dB)	$SD \geq 2$ dB (%)	$SD \geq 4$ dB (%)
Companded SQ	1.36	6.71	0.675
LM, $\alpha = 2$	1.25	9.43	1.480
LM, optimal α	1.23	6.51	0.293
AMR-WB [12]	1.13	3.01	0.015

(b) Results at 46 bits per frame

Quantization methods	avg. SD (dB)	$SD \geq 2$ dB (%)	$SD \geq 4$ dB (%)
Companded SQ	0.90	2.42	0.323
LM, $\alpha = 2$	0.87	4.20	0.710
LM, optimal α	0.85	1.71	0.057
AMR-WB [12]	0.78	0.45	0.003

5.4 Complexity

The memory requirement for fixed-point AMR-WB LPC quantization tables is 6.7 kword (16-bit words) [12]. For the predictive KLT coding with model-based Lloyd-Max quantization, the memory requirement is 0.8 kword assuming a fixed-point implementation. In particular we need to store for each bit allocation (36 or 46 bits) the KLT matrix T, the eigenvalues σ_j, the LM centroids as well and the number of quantizations. It would be possible to reduce even more the memory consumption, for instance by forcing a unique KLT matrix T independent of bit allocation. The memory cost of model-based Lloyd-Max quantization is very small, yet it depends slightly on bit allocation. On the other hand non-uniform scalar quantization with high-rate Gaussian companding has a complexity independent of bit allocation, yet this technique implies to store tables to implement the compander c.

Moreover, the total computational cost of predictive KLT coding with model-based Lloyd-Max quantization in fixed-point is estimated around 0.1 wMOPS (weighted Million Operations per Second) whereas it is around 1.9 wMOPS for AMR-WB LPC quantization.

6. CONCLUSIONS

In this paper we presented a predictive KLT quantization method using generalized Gaussian modeling for wideband LSF speech parameters. This method was compared to AMR-WB LPC quantization. The proposed method has much lower complexity (computation cost, storage requirement) and similar performance.

REFERENCES

[1] P. Hedelin and J. Skoglund, "Vector quantization based on Gaussian mixture models," *IEEE Trans. Speech and Audio Proc.*, vol. 8, no. 4, pp. 385–401, Jul. 2000.

[2] A. D. Subramanian and B. D. Rao, "Pdf optimized parametric vector quantization of speech line spectral frequencies," *IEEE Trans. Speech and Audio Proc.*, vol. 11, no. 2, pp. 130–142, Mar. 2003.

[3] M. Oger, S. Ragot, and R. Lefebvre, "Companded lattice VQ for efficient parametric LPC quantization," *Proc. Eusipco*, 2004.

[4] S. Ragot, H. Lahdili, and R. Lefebvre, "Wideband LSF quantization by generalized Voronoi codes," in *Proc. Eurospeech*, Sep. 2001, pp. 2319–2322.

[5] R. L. Joshi and T. R. Fisher, "Comparison of generalized Gaussian and Laplacian modeling in DCT image coding," *IEEE. Sign. Proc. Lett.*, vol. 2, no. 5, pp. 81–82, May 1995.

[6] C. Parisot, M. Antonini, and M. Barlaud, "3d scan based wavelet transform and quality control for video coding," *EURASIP*, vol. 2003, no. 1, pp. 521–528, Jan 2003.

[7] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *JRSSB*, vol. 39, no. 1, pp. 1–38, 1977.

[8] A. Papoulis, *Probability, random variables, and stochastic processes*, Mc Graw-Hill, 1965.

[9] S. Kullback, *Information Theory and Statistics*, New York: Wiley, 1959.

[10] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.

[11] K. Kokkinakis and A. K. Nandi, "Speech modelling based on generalized Gaussian probability density functions," *ICASSP*, pp. 381–384, 2005.

[12] ITU-T G.722.2, *Wideband Coding of Speech at Around 16 kbit/s using Adaptive Multirate Wideband*, 2002.

[13] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall,Inc, 1984.

[14] K.K. Paliwal and W.B. Kleijn, *Quantization of LPC Parameters*, pp. 433–466, in *Speech Coding and Synthesis*, W.B. Kleijn and K.K. Paliwal eds., Elsevier Science, 1995.