

## NONLINEAR ESTIMATION OF MISSING $\Delta$ LSF PARAMETERS BY A MIXTURE OF DIRICHLET DISTRIBUTIONS

Zhanyu Ma<sup>a</sup>, Rainer Martin<sup>b</sup>, Jun Guo<sup>a</sup>, and Honggang Zhang<sup>a</sup>

<sup>a</sup>Pattern Recognition and Intelligent System Lab.,

Beijing University of Posts and Telecommunications, Beijing, China.

<sup>b</sup>Institute of Communication Acoustics, Ruhr-University Bochum, Bochum, Germany

mazhanyu@bupt.edu.cn, rainer.martin@rub.de, guojun@bupt.edu.cn, zhhg@bupt.edu.cn

### ABSTRACT

In packet networks, a reliable scheme to handle packet loss during speech transmission is of great importance. As a common representation of the linear predictive coding (LPC) model, the line spectral frequency (LSF) parameters are widely used in speech quantization and transmission. In this paper, we propose a novel scheme to estimate the missing values occurring during LPC model transmission. In order to exploit the boundary and ordering properties of the LSF parameters, we utilize the  $\Delta$ LSF representation and apply the Dirichlet mixture model (DMM) to capture the correlations among the elements in the  $\Delta$ LSF vector. With the conditional distribution of the missing part given the received part, an optimal nonlinear minimum mean square error estimator for the missing values is proposed. Compared to the previously presented Gaussian mixture model based method, the proposed DMM based nonlinear estimator shows a convincing improvement.

**Index Terms**— Line spectral frequency, packet loss, Dirichlet distribution, mixture modeling, neutrality property

### 1. INTRODUCTION

Efficient quantization and transmission of the linear predictive coding (LPC) model plays an important role in parametric speech coding. Conventionally, the LPC coefficients are converted to line spectral frequency (LSF) parameters [1–3]. The LSF representation, among others, is the most stable and efficient one in LPC model quantization and transmission [4].

When transmitting speech over a packet network, the delayed or lost packets should be estimated from the available information so that additional latency is avoided. In a previous study [5], the joint distribution of the missing LSF elements and the receiving ones is modeled by a Gaussian mixture model. Based on the conditional distribution of the missing part given the received one, an optimal minimum mean square error (MMSE) estimator was derived by utilizing the intra-frame correlation of differentially-encoded LSF coefficients. Due to the nice properties of the Gaussian distribution,

a closed-form solution was obtained. This method enhanced the quality of the speech when frame losses occurred and consumed significantly less memory, compared to a histogram-based approach.

By considering the boundedness [6] and ordering properties [7, 8], the LSF parameters can be transformed into the  $\Delta$ LSF domain. The  $\Delta$ LSF parameters have less variability and the range is more limited compared to the absolute LSF value [4, 9]; therefore, schemes of quantizing the  $\Delta$ LSF, rather than the LSF, have been introduced in, *e.g.*, [7–11]. In [8], the underlying distribution of the  $\Delta$ LSF parameters was modeled by a Dirichlet mixture model (DMM). An efficient DMM-based vector quantization (VQ) scheme was proposed and shown to be superior to the conventional Gaussian mixture model (GMM)-based VQ. Hence, the  $\Delta$ LSF parameters, instead of the LSF parameters, were transmitted after quantization.

In this paper, we propose a DMM-based method to nonlinearly estimate the  $\Delta$ LSF values lost during transmission. Similar to [8], the underlying distribution of the  $\Delta$ LSF is modeled by a DMM. With the neutrality properties of the Dirichlet variable, the correlation between the received part and the missing part is studied and an optimal nonlinear estimator for the missing  $\Delta$ LSF values is derived. This estimator can be expressed in an analytically tractable form, and therefore, the calculation is facilitated. Previous work on Dirichlet/generalized Dirichlet distribution based estimator mainly focused on image processing, which can be found in, *e.g.*, [12].

In the remaining parts of this paper, we will introduce the basic idea of the  $\Delta$ LSF parameters, study the neural properties of the Dirichlet variable, derive the DMM-based nonlinear estimator, and show extensive experimental comparisons.

### 2. $\Delta$ LSF REPRESENTATIONS OF THE LSF PARAMETERS

The LSF parameters are widely used in speech transmission because of their relatively uniform spectral sensitivity [13].

The LSF parameters with dimensionality  $K$  are defined as

$$\mathbf{s} = [s_1, s_2, \dots, s_K]^T, \quad (1)$$

which are interleaved on the unit circle [1].

By recognizing that the LSF parameters are in the interval  $(0, \pi)$  and are strictly ordered, one representation, namely the  $\Delta$ LSF, was utilized in [7] for the purpose of LSF quantization [8]. It has been demonstrated that the DMM based VQ performs better than the conventional GMM based VQ [8]. With a transformation matrix  $A$ , the relation between the LSF parameters  $\mathbf{s}$  and the  $\Delta$ LSF parameters  $\mathbf{x}$  is [7]

$$\mathbf{x} = \varphi(\mathbf{s}) = A\mathbf{s} = [s_1, s_2 - s_1, \dots, s_K - s_{K-1}]^T, \quad (2)$$

where

$$A = \frac{1}{\pi} \begin{bmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -1 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{bmatrix}_{K \times K}. \quad (3)$$

According to [7, 8], the underlying distribution of the  $K$ -dimensional  $\Delta$ LSF parameters can be modeled efficiently with a  $(K + 1)$ -dimensional DMM.

### 3. DIRICHLET MIXTURE MODEL

By concatenating  $x_{K+1} = 1 - \sum_{k=1}^K x_k$  to the end of  $\mathbf{x}$ , we obtain a new  $(K + 1)$ -dimensional vector  $\tilde{\mathbf{x}} = [x_1, \dots, x_K, x_{K+1}]^T$ , which is named the *complete*  $\Delta$ LSF vector. From another point of view, this vector can be obtained by dividing the range  $[0, \pi]$  into  $K + 1$  intervals  $s_1 - 0, \dots, s_K - s_{K-1}, \pi - s_K$ , and then scaled by a factor  $\frac{1}{\pi}$ . Thus, the vector  $\tilde{\mathbf{x}}$  denotes the proportions of these  $K + 1$  intervals to the whole range  $[0, \pi]$ . Since the summation of the elements in  $\tilde{\mathbf{x}}$  is equal to one, we assume that  $\tilde{\mathbf{x}}$  is Dirichlet distributed, with  $K$  degrees of freedom, as [14]

$$f(\tilde{\mathbf{x}}) = \text{Dir}(\tilde{\mathbf{x}}; \boldsymbol{\alpha}) = \frac{\Gamma(\sum_{k=1}^{K+1} \alpha_k)}{\prod_{k=1}^{K+1} \Gamma(\alpha_k)} \prod_{k=1}^{K+1} x_k^{\alpha_k - 1},$$

where  $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_{K+1}]^T$  is the parameter vector. With a set of  $N$  *i.i.d.* observations  $\mathbf{X} = [\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_N]$ , we can denote the likelihood function for the observations by a mixture of Dirichlet densities with  $I$  components as

$$f(\mathbf{X}) = \prod_{n=1}^N \sum_{i=1}^I \pi_i \frac{\Gamma(\sum_{k=1}^{K+1} \alpha_{ki})}{\prod_{k=1}^{K+1} \Gamma(\alpha_{ki})} \prod_{k=1}^{K+1} (x_{kn})^{\alpha_{ki} - 1},$$

where  $\boldsymbol{\alpha}_i = [\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{K+1,i}]^T$  is the parameter vector for the  $i$ th mixture component,  $\pi_i$  is the nonnegative weighting factor for the  $i$ th component, and  $\sum_{i=1}^I \pi_i = 1$ . By applying the expectation-maximization (EM) algorithm, the parameters in the DMM can be estimated with the method introduced in [7]<sup>1</sup>.

<sup>1</sup>Bouguila et al. also proposed an EM algorithm in [15]. In this paper, we directly estimate  $\alpha_{ki}$  instead of estimating  $\ln \alpha_{ki}$  as in [15].

## 4. ESTIMATION OF THE MISSING VALUES

The Dirichlet vector  $\tilde{\mathbf{x}}$  is negatively correlated. When some elements in  $\tilde{\mathbf{x}}$  are missing, we can recover them by using the correlations between the missing part and the received part.

### 4.1. Properties of the Dirichlet Variable

The Dirichlet vector  $\tilde{\mathbf{x}}$  is a neutral vector [8, 16]. This neutrality concept has been proposed in [16] and the independence properties of the proportion ratios were further studied in [17–19]. Here, we review two important properties of the neutral vector.

#### Property 4.1 (Aggregation of Dirichlet Variable)

If a vector  $\tilde{\mathbf{x}} = [x_1, x_2, \dots, x_{K+1}]^T$  is Dirichlet distributed as  $\text{Dir}(\tilde{\mathbf{x}}; \boldsymbol{\alpha})$ , the new vector  $\tilde{\mathbf{x}}_{i \cap j} = [x_1, x_2, \dots, x_i + x_j, \dots, x_{K+1}]^T$  is also Dirichlet distributed as  $\text{Dir}(\tilde{\mathbf{x}}_{i \cap j}; \boldsymbol{\alpha}_{i \cap j})$ , where  $\boldsymbol{\alpha}_{i \cap j} = [\alpha_1, \alpha_2, \dots, \alpha_i + \alpha_j, \dots, \alpha_{K+1}]^T$ .

#### Property 4.2 (Neutrality Property)

For  $\tilde{\mathbf{x}} = [x_1, x_2, \dots, x_{K+1}]^T$  drawn from  $\text{Dir}(\tilde{\mathbf{x}}; \boldsymbol{\alpha})$ ,  $x_1$  is independent of  $\mathbf{x}_{\setminus 1} = [\frac{x_2}{1-x_1}, \dots, \frac{x_{K+1}}{1-x_1}]^T$ . More generally, for any  $k \in \{1, 2, \dots, K+1\}$ ,  $x_k$  is independent of  $\mathbf{x}_{\setminus k} = [\frac{x_1}{1-x_k}, \dots, \frac{x_{k-1}}{1-x_k}, \frac{x_{k+1}}{1-x_k}, \dots, \frac{x_{K+1}}{1-x_k}]^T$ . Furthermore,  $\tilde{\mathbf{x}}_{\setminus k}$  is distributed as  $\tilde{\mathbf{x}}_{\setminus k} \sim \text{Dir}(\tilde{\mathbf{x}}_{\setminus k}; \boldsymbol{\alpha}_{\setminus k})$ , where  $\boldsymbol{\alpha}_{\setminus k} = [\alpha_1, \dots, \alpha_{k-1}, \alpha_{k+1}, \dots, \alpha_{K+1}]^T$ .

### 4.2. Relating Distributions

Assuming that the *complete*  $\Delta$ LSF parameter vector  $\tilde{\mathbf{x}}$  is partitioned into two parts, *i.e.*, the missing part  $\tilde{\mathbf{x}}^M$  and the received part  $\tilde{\mathbf{x}}^R$ , as

$$\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{\mathbf{x}}^M \\ \tilde{\mathbf{x}}^R \end{bmatrix}. \quad (4)$$

Meanwhile, the parameter vector in the  $i$ th mixture component can also be partitioned correspondingly as

$$\boldsymbol{\alpha}_i = \begin{bmatrix} \boldsymbol{\alpha}_i^M \\ \boldsymbol{\alpha}_i^R \end{bmatrix}. \quad (5)$$

After normalizing the missing parts by itself, we get a normalized vector

$$\check{\mathbf{x}}^M = [\check{x}_1^M, \check{x}_2^M, \dots, \check{x}_M^M]^T. \quad (6)$$

Each element in  $\check{\mathbf{x}}^M$  is defined as

$$\check{x}_m^M = \frac{x_m^M}{\sum_{m=1}^M x_m^M} = \frac{x_m^M}{1 - \sum_{r=1}^R x_r^R}, \quad (7)$$

where  $x_m^M$  is the  $m$ th element in  $\tilde{\mathbf{x}}^M$  and  $M$  is the length of  $\tilde{\mathbf{x}}^M$ . According to property 4.1 and 4.2, it can be shown that the distribution of the normalized missing part  $\check{\mathbf{x}}^M$ , on a mixture component basis, can be written as

$$f_i(\check{\mathbf{x}}^M) = \text{Dir}(\check{\mathbf{x}}^M; \boldsymbol{\alpha}_i^M) = \frac{\Gamma(\sum_{m=1}^M \alpha_{mi}^M)}{\prod_{m=1}^M \Gamma(\alpha_{mi}^M)} \prod_{m=1}^M (\check{x}_m^M)^{\alpha_{mi}^M - 1}. \quad (8)$$

For the same reasoning, the normalized version of the received parts  $\tilde{\mathbf{x}}^R$  can be represented as

$$\tilde{\mathbf{x}}^R = [\tilde{x}_1^R, \tilde{x}_2^R, \dots, \tilde{x}_R^R]^T \text{ and } \tilde{x}_r^R = \frac{x_r^R}{\sum_{r=1}^R x_r^R}, \quad (9)$$

where  $x_r^R$  denotes the element in  $\tilde{\mathbf{x}}^R$  and  $R$  is the length of  $\tilde{\mathbf{x}}^R$ . The marginal PDF of  $\tilde{\mathbf{x}}^R$ , on a mixture component basis, is then denoted as

$$f_i(\tilde{\mathbf{x}}^R) = \text{Dir}(\tilde{\mathbf{x}}^R; \boldsymbol{\alpha}_i^R) = \pi_i \frac{\Gamma(\sum_{r=1}^R \alpha_{ri}^R)}{\prod_{r=1}^R \Gamma(\alpha_{ri}^R)} \prod_{r=1}^R (\tilde{x}_r^R)^{\alpha_{ri}^R - 1}. \quad (10)$$

### 4.3. Optimal MMSE Estimator

Based on the probability theory [20,21], the optimal estimator for the missing part  $\tilde{\mathbf{x}}^M$ , in terms of the MMSE criterion, is the conditional mean of the missing part  $\tilde{\mathbf{x}}^M$  given the received part  $\tilde{\mathbf{x}}^R$ , which can be calculated as

$$\mathbf{E}_{f(\tilde{\mathbf{x}}^M|\tilde{\mathbf{x}}^R)}[\tilde{\mathbf{x}}^M] = \int \tilde{\mathbf{x}}^M f(\tilde{\mathbf{x}}^M|\tilde{\mathbf{x}}^R) d\tilde{\mathbf{x}}^M. \quad (11)$$

The conditional PDF of  $\tilde{\mathbf{x}}^M$  given  $\tilde{\mathbf{x}}^R$  is

$$\begin{aligned} f(\tilde{\mathbf{x}}^M|\tilde{\mathbf{x}}^R) &= \frac{f(\tilde{\mathbf{x}})}{f(\tilde{\mathbf{x}}^R)} \\ &= \frac{\sum_{i=1}^I \pi_i f_i(\tilde{\mathbf{x}})}{\int \sum_{i=1}^I \pi_i f_i(\tilde{\mathbf{x}}) d\tilde{\mathbf{x}}^M} \\ &= \sum_{i=1}^I \left[ \frac{\pi_i f_i(\tilde{\mathbf{x}}^R)}{\sum_{i=1}^I \pi_i f_i(\tilde{\mathbf{x}}^R)} \cdot \frac{f_i(\tilde{\mathbf{x}})}{f_i(\tilde{\mathbf{x}}^R)} \right] \\ &= \sum_{i=1}^I \left[ \frac{\pi_i f_i(\tilde{\mathbf{x}}^R)}{\sum_{i=1}^I \pi_i f_i(\tilde{\mathbf{x}}^R)} \cdot f_i(\tilde{\mathbf{x}}^M|\tilde{\mathbf{x}}^R) \right]. \end{aligned} \quad (12)$$

In the last line of (12),

$$\frac{\pi_i f_i(\tilde{\mathbf{x}}^R)}{\sum_{i=1}^I \pi_i f_i(\tilde{\mathbf{x}}^R)} = \frac{\pi_i f_i(\tilde{\mathbf{x}}^R)}{\sum_{i=1}^I \pi_i f_i(\tilde{\mathbf{x}}^R)}, \quad (13)$$

since  $f_i(\tilde{\mathbf{x}}^R) \propto f_i(\tilde{\mathbf{x}}^R)$ . This quantity can be calculated explicitly with (10) and is independent of  $\tilde{\mathbf{x}}^M$ . Therefore, the conditional mean of  $\tilde{\mathbf{x}}^M$  given  $\tilde{\mathbf{x}}^R$  can be calculated as

$$\begin{aligned} \mathbf{E}_{f(\tilde{\mathbf{x}}^M|\tilde{\mathbf{x}}^R)}[\tilde{\mathbf{x}}^M] &= \sum_{i=1}^I \pi_i' \int \tilde{\mathbf{x}}^M f_i(\tilde{\mathbf{x}}^M|\tilde{\mathbf{x}}^R) d\tilde{\mathbf{x}}^M \\ &= (1 - \sum_{r=1}^R x_r^R) \cdot \sum_{i=1}^I \pi_i' \int \tilde{\mathbf{x}}^M f_i(\tilde{\mathbf{x}}^M) d\tilde{\mathbf{x}}^M \end{aligned} \quad (14)$$

$$\begin{aligned} &= (1 - \sum_{r=1}^R x_r^R) \cdot \sum_{i=1}^I \pi_i' \cdot \mathbf{E}_{f_i(\tilde{\mathbf{x}}^M)}[\tilde{\mathbf{x}}^M], \end{aligned} \quad (15)$$

where  $\pi_i' = \frac{\pi_i f_i(\tilde{\mathbf{x}}^R)}{\sum_{i=1}^I \pi_i f_i(\tilde{\mathbf{x}}^R)}$ . From (14) to (15) we used the fact that  $\tilde{\mathbf{x}}^M = (1 - \sum_{r=1}^R x_r^R) \tilde{\mathbf{x}}^M$  and applied the method of integration by substitution.

The above equation means that the optimal estimator of the missing part  $\tilde{\mathbf{x}}^M$ , in terms of MMSE sense, is the weighted sum of the means of the normalized missing part  $\tilde{\mathbf{x}}^M$  scaled by the received part  $\tilde{\mathbf{x}}^R$ .

## 5. EXPERIMENTAL RESULTS AND DISCUSSION

### 5.1. Experimental Setup

The DMM based VQ for the  $\Delta$ LSF representation was introduced in [8]. In stead of transmitting the  $K$ -dimensional LSF vector directly, the  $K$ -dimensional  $\Delta$ LSF vector was quantized and transmitted. At the receiver side, the LSF vector can be obtained from the received  $\Delta$ LSF vector. In case that some elements are lost during transmission, we can use the above proposed DMM based nonlinear optimal MMSE estimator to recover the missing part from the received elements. In order to make extensive comparisons, we test the proposed estimator with both narrow band (NB) and wide band (WB) data. For each type of data, we consider two transmission scenarios: a) the elements in the  $\Delta$ LSF vector are transmitted individually and only one element is lost during transmission. b) the  $\Delta$ LSF vector are partitioned into subsets and then transmitted. Only one subvector is lost during transmission. The partition strategy follows the commonly used method in the LSF transmission as:

1. The NB data case. According to the GSM AMR coder [5, 22], the 10-dimensional LSF/ $\Delta$ LSF vector are partitioned into three subsets as  $\{3, 3, 4\}$ .
2. The WB data case. When applying the split VQ [23, 24], the 16-dimensional LSF/ $\Delta$ LSF vector are usually partitioned into five subsets as  $\{3, 3, 3, 3, 4\}$ .

Please note, when recovering the missing elements from the *complete*  $(K + 1)$ -dimensional  $\Delta$ LSF vector, we need to estimate not only the missing elements but also the redundant element (*i.e.*,  $x_{K+1}$ ) in the *complete*  $\Delta$ LSF vector. This means that, if the elements with location index 1, 2, and 3 are missing in the NB data case, the missing part (that will be estimated) in (4) is  $\tilde{\mathbf{x}}^M = [x_1, x_2, x_3, x_{11}]^T$ .

### 5.2. Results and Discussion

To make fair and extensive comparisons, we conducted and compared three estimation methods:

1. The proposed DMM+ $\Delta$ LSF method, which models the  $\Delta$ LSF parameters by a DMM.
2. The GMM+LSF method introduced in [5], which models the LSF parameters by a GMM.
3. The GMM+ $\Delta$ LSF method, which models the  $\Delta$ LSF parameters by a GMM and follows the optimal estimation strategy in [5].

For method 1 and 3, the corresponding LSF parameters can be obtained from the estimated  $\Delta$ LSF parameters. The estimation error in the LSF domain, calculated in terms of mean square error (MSE), is used as the criterion for performance

**Table 1.** Performance comparisons with NB data.

Method	Model order	MSE ( $\times 10^{-3}$ )														
		Missing element										Avg.	Missing subvector			Avg.
		1	2	3	4	5	6	7	8	9	10		1	2	3	
DMM+ $\Delta$ LSF	16 mix.	3.5	3.8	4.7	6.9	8.5	7.6	7.8	6.5	8.4	8.3	6.6	18.8	51.5	61.5	43.9
	32 mix.	3.2	3.5	4.6	6.5	8.2	7.1	7.4	6.2	8.2	8.2	6.3	18.3	49.1	60.7	42.7
	64 mix.	3.1	3.3	4.2	6.2	8.1	6.5	7.2	6.0	8.1	8.1	6.2	17.8	48.9	60.2	42.3
GMM+LSF	16 mix.	3.6	4.4	6.5	13.9	17.8	14.9	14.7	13.9	12.2	10.4	11.2	23.8	67.4	61.7	60.0
	32 mix.	3.3	3.9	5.6	11.3	15.3	13.5	13.3	13.1	11.5	10.1	10.1	22.1	63.2	60.8	48.7
	64 mix.	2.9	3.5	5.0	10.3	13.3	11.5	12.5	12.5	11.2	9.9	9.3	20.5	64.5	60.6	48.5
GMM+ $\Delta$ LSF	16 mix.	4.4	5.1	6.9	13.5	15.6	15.2	17.6	14.4	18.9	12.2	12.4	24.7	72.4	93.0	63.4
	32 mix.	4.0	4.4	6.5	12.9	14.1	14.5	16.4	13.0	17.2	11.4	11.4	24.4	67.0	87.2	59.5
	64 mix.	3.9	4.2	6.2	12.0	12.3	12.9	14.6	11.6	15.0	10.5	10.3	21.6	64.7	81.4	55.9

**Table 2.** Performance comparisons with WB data.

Method	Model order	MSE ( $\times 10^{-3}$ )																						
		Missing element															Avg.	Missing subvector					Avg.	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		16	1	2	3	4		5
DMM+ $\Delta$ LSF	16 mix.	1.4	1.8	2.5	3.3	3.3	2.9	2.7	3.0	2.9	2.7	3.1	3.3	3.8	3.6	3.8	3.4	3.0	9.0	21.3	14.6	16.5	29.7	18.2
	32 mix.	1.3	1.5	2.5	3.1	3.0	2.8	2.7	2.7	2.8	2.6	2.9	3.1	3.6	3.5	3.5	3.4	2.8	8.6	20.8	14.3	16.0	29.3	17.8
	64 mix.	1.3	1.3	2.4	3.0	2.9	2.7	2.6	2.6	2.6	2.5	2.9	3.1	3.6	3.5	3.5	3.3	2.7	8.4	20.8	14.0	16.0	29.2	17.7
GMM+LSF	16 mix.	1.5	2.9	4.6	7.8	7.0	6.1	6.2	5.8	5.8	6.3	7.7	8.4	9.1	7.5	6.0	4.1	6.1	10.9	25.8	19.2	26.8	30.5	22.6
	32 mix.	1.4	2.4	4.0	6.8	6.0	5.5	5.6	5.4	5.2	5.7	7.0	8.2	8.5	6.5	5.5	3.9	5.5	10.0	23.9	17.7	25.9	29.8	21.5
	64 mix.	1.2	2.1	3.8	6.3	5.2	5.1	5.2	5.1	4.7	5.4	7.1	7.5	8.0	6.3	5.4	3.8	5.1	9.7	22.4	16.5	25.6	29.8	20.8
GMM+ $\Delta$ LSF	16 mix.	2.0	2.3	4.7	7.1	6.4	7.2	5.6	7.4	6.4	6.3	8.1	9.3	10.6	9.0	8.0	5.8	6.6	13.3	31.8	27.6	33.0	58.2	32.8
	32 mix.	1.8	2.4	4.4	6.6	5.8	6.8	5.6	6.7	6.0	6.1	7.5	9.3	10.0	8.6	7.3	5.9	6.3	11.8	29.3	26.1	31.4	55.8	30.9
	64 mix.	1.7	1.8	4.2	5.8	5.4	6.5	5.2	6.3	5.5	5.9	7.7	8.6	10.3	9.1	7.2	5.8	6.1	11.5	28.2	24.6	31.1	51.1	30.1

comparison. We ran 50 rounds of simulations and the mean values are reported.

The TIMIT database [25] was used to get a training and a test set of LSF/ $\Delta$ LSF parameters. TIMIT is a corpus of phonemically and lexically transcribed speech of female and male American English speakers of different dialects. To obtain the NB data, the 16 kHz speech signal was firstly downsampled by a factor of 2. With window length equal to 25 milliseconds and step size equal to 20 milliseconds, approximately 497,000 LSF/ $\Delta$ LSF vectors were selected for the training partition, and about 178,000 LSF/ $\Delta$ LSF vectors were obtained for the test partition. The Hann window was applied to each frame and no prefilter was used. All the silent frames were removed. Table 1 lists the comparison results when transmitting the LSF/ $\Delta$ LSF elements individually. When the LSF/ $\Delta$ LSF parameters are partitioned into subvectors and transmitted, the comparisons are also shown in Table 1. It can be observed that the proposed DMM+ $\Delta$ LSF method performs better (with smaller MSE) than the other two methods in both scenarios. As the model order (number of mixture components) increases, the estimation performance is also improved. We believe this is because the  $\Delta$ LSF representation of the LSF parameters can explicitly exploit the boundary and ordering properties. The  $\Delta$ LSF representation captures the correlation between the missing part and the received part more efficiently. Moreover, since the GMM+ $\Delta$ LSF method performs the worst among all the three methods, it suggests that DMM is an efficient statistical model in describing the underlying distribution of the  $\Delta$ LSF parameters.

Similar facts can also be observed from Table 2, where the WB data is used for evaluation. In the WB data case, frame extraction settings are the same as those in the NB data case. About 497,000 and 178,000 LSF/ $\Delta$ LSF vectors were obtained from the training and test partitions, respectively.

**Table 3.** Comparisons of model complexities.

Number of mixture components	Number of free parameters					
	DMM+ $\Delta$ LSF		GMM+LSF		GMM+ $\Delta$ LSF	
	NB	WB	NB	WB	NB	WB
16	191	287	335	527	367	559
32	383	575	671	1055	735	1119
64	767	1151	1343	2111	1471	2239

The model complexity comparisons, in terms of free parameters used in describing the model, are listed in Table 3. Compared to the other two methods, the DMM+ $\Delta$ LSF method performs better and has the lowest model complexity.

## 6. CONCLUSIONS

As an efficient representation of the line spectral frequency (LSF) parameters, the  $\Delta$ LSF parameters are used in linear predictive (LPC) model quantization and transmission. In order to deal with the packet loss occurring during the  $\Delta$ LSF parameter transmission, we proposed an optimal nonlinear estimator, in the sense of minimum mean square error, to recover the missing part from the received available information. The proposed method is based on a Dirichlet mixture model (DMM). With the neutrality property of the Dirichlet variable, an analytically tractable solution is derived. With both NB and WB data evaluations in all the transmission scenarios, we can conclude that transmitting the LPC model with the  $\Delta$ LSF representation and applying the proposed DMM+ $\Delta$ LSF method can significantly enhance the signal quality when packet loss occurred. Furthermore, the DMM+ $\Delta$ LSF method has requires less memory.

## Acknowledgement

This work is partly supported by the ‘‘Foundational Research Funds for the Central Universities’’ No. 2013XZ11, NSFC grant No. 61273217, Chinese 111 program of Advanced Intelligence and Network Service under grant No. B08004, and EU FP7 IRSES MobileCloud Project (Grant No. 612212).

## 7. REFERENCES

- [1] F. Itakura, "Line spectrum representation of linear predictive coefficients of speech signals," *Journal of the Acoustical Society of America*, vol. 57, pp. 535, 1975.
- [2] P. Vary and R. Martin, *Digital Speech Transmission: Enhancement, Coding and Error Concealment*, John Wiley & Sons, Ltd, Chichester, England, 2006.
- [3] J. Benesty, M. M. Sondhi, and Y. Huang, Eds., *Springer Handbook on Speech Processing*, Springer, 2008.
- [4] K. K. Paliwal and W. B. Kleijn, *Speech Coding and Synthesis*, chapter Quantization of LPC parameters, pp. 433–466, Amsterdam, The Netherlands: Elsevier, 1995.
- [5] R. Martin, C. Hoelper, and I. Wittke, "Estimation of missing LSF parameters using Gaussian mixture models," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001, vol. 2, pp. 729–732 vol.2.
- [6] Z. Ma and A. Leijon, "Bayesian estimation of beta mixture models with variational inference," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 11, pp. 2160–2173, Nov. 2011.
- [7] Z. Ma and A. Leijon, "Modeling speech line spectral frequencies with Dirichlet mixture models," in *Proceedings of INTERSPEECH*, 2010, pp. 2370–2373.
- [8] Z. Ma, A. Leijon, and W. B. Kleijn, "Vector quantization of LSF parameters with a mixture of Dirichlet distributions," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 1777 – 1790, Sep. 2013.
- [9] F. Soong and B. Juang, "Optimal quantization of LSP parameters using delayed decisions," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1990, pp. 185–188.
- [10] F. Soong and B. Juang, "Line spectrum pair (LSP) and speech data compression," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Mar. 1984, vol. 9, pp. 37–40.
- [11] F. Lahouti and A. K. Khandani, "Quantization of line spectral parameters using a trellis structure," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2000, vol. 5, pp. 2781–2784.
- [12] N. Bouguila, "Non-Gaussian mixture image models prediction," in *Proceedings of IEEE International Conference on Image Processing (ICIP)*, 2008, pp. 2580 – 2583.
- [13] J. Li, N. Chaddha, and R. M. Gray, "Asymptotic performance of vector quantizers with a perceptual distortion measure," *IEEE Transactions on Information Theory*, vol. 45, pp. 1082 – 1091, May 1999.
- [14] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- [15] N. Bouguila, D. Ziou, and J. Vaillancourt, "Unsupervised learning of a finite mixture model based on the Dirichlet distribution and its application," *IEEE Transactions on Image Processing*, vol. 13, no. 11, pp. 1533–1543, Nov. 2004.
- [16] R. J. Connor and J. E. Mosimann, "Concepts of independence for proportions with a generalization of the Dirichlet distribution," *Journal of American Statistical Association*, vol. 64, no. 325, pp. 194–206, 1969.
- [17] I. R. James and J. E. Mosimann, "A new characterization of the Dirichlet distribution through neutrality," *The Annals of Statistics*, vol. 8, no. 1, pp. 183–189, 1980.
- [18] S. Kotz, N. L. Johnson, and N. Balakrishnan, *Continuous Multivariate Distributions: Models and applications*, Wiley series in probability and statistics: Applied probability and statistics. Wiley, 2000.
- [19] B. A. Frigyik, A. Kapila, and M. R. Gupta, "Introduction to the Dirichlet distribution and related processes," Tech. Rep., Department of Electrical Engineering, University of Washington, 2010.
- [20] J. Zhang and D. Ma, "Nonlinear prediction for Gaussian mixture image models," *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 836–847, 2004.
- [21] V. Krishnan, *Probability and Random Processes*, Wiley Survival Guides in Engineering and Science. Wiley, 2006.
- [22] ESTI GSM 06.71, "Digital cellular telecommunications system (phase 2+); Adaptive Multi-Rate (AMR); Speech processing functions; General description," 1998.
- [23] S. So and K. K. Paliwal, "A comparative study of LPC parameter representations and quantisation schemes for wide-band speech coding," *Digital Signal Processing*, vol. 17, pp. 114–137, Jan. 2007.
- [24] S. Chatterjee and T.V. Sreenivas, "Analysis-by-synthesis based switched transform domain split VQ using Gaussian mixture model," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2009, pp. 4117–4120.
- [25] "DARPA-TIMIT," *Acoustic-phonetic continuous speech corpus, NIST Speech Disc 1.1-1*, 1990.