# AN ADAPTIVE, SCALABLE PACKET LOSS RECOVERY METHOD

*Christian Feldbauer and W. Bastiaan Kleijn*

KTH (Royal Institute of Technology) Stockholm, Sweden
Sound and Image Processing Lab, School of Electrical Engineering
feldbauer@tugraz.at, bastiaan.kleijn@ee.kth.se

**ABSTRACT**

We propose a packet loss recovery method that uses an incomplete secondary encoding as redundancy. The recovery is performed by minimum mean squared error estimation. The method adapts to the loss scenario and is rate scalable. It incorporates a statistical model for the quantizers to facilitate real-time adaptation. We apply the method to the encoding of line-spectral frequencies, which are commonly used in speech coding, illustrating the good performance of the method.

***Index Terms—*** Robust coding, Packet loss, Speech coding

## 1. INTRODUCTION

Audio-visual communication over packet networks has become commonplace. System cost is reduced if the encoding is robust to packet loss and can react to congestion by adaptation of the rate. In this paper, we propose a rate-scalable and robustness-scalable encoding method and describe its application to speech coding.

Methods to address packet loss can be divided into receiver-based and sender-based methods (e.g., [1]). Sender-based methods, which introduce redundancy in the transmitted bit stream, are generally more powerful but require changes in both encoder and decoder, whereas receiver-based methods require changes in the decoder only.

We propose a sender-based packet loss recovery method that uses an incomplete secondary encoding as redundancy. The method can easily be added to existing systems. In contrast to multiple-description coding (MDC), our secondary encoding is not a complete description of the signal. A legacy decoder simply ignores the bit-stream component corresponding to the secondary encoding. Matched decoders, however, use the secondary description to improve the quality of the reconstructed speech in case of packet losses.

The high performance of our system is partly due to statistical modeling at the receiver. Statistical signal models facilitate signal reconstruction when the bit stream is damaged. For instance, in [2] a-priori knowledge of speech parameters is used to conceal bit errors caused by noisy channels. In [3, 4] Gaussian mixture models (GMMs) are used to predict speech parameters of lost packets from previous packets. GMMs are also used in [5] to estimate missing parameters of incomplete descriptions. The fore-mentioned systems [2, 3, 5] do not rely on the transmission of redundancy and only exploit the dependency of parameters either between blocks or within a single block.

Our approach combines statistical signal modeling with the transmission of redundant information (cf. also [4, 6]). Because our approach is entirely based on an analytic, continuous signal model, our method facilitates real-time adaptation of the redundancy during transmission.

A unique aspect of our coding architecture is that it is loss-scenario flexible and rate-scalable: depending on short-term packet loss statistics and network constraints, the primary as well as the secondary encodings can adapt in rate. This ensures good reconstruction quality on the one hand and a constrained total rate on the other hand. Existing loss recovery methods based on vector quantizers (VQs) (e.g., [4, 6]) are not rate-scalable unless VQs for a finite selection of different redundancy rates are trained. Moreover, existing methods also have dependencies on the primary rate (such as with discrete statistical models, e.g., states of Markov models referring to VQ cells [6]) or on a particular packet loss scenario [4]. These dependencies require different estimators or the training of several statistical models for different rates and different scenarios. Our estimator's flexibility and the modeling of both the signal properties and the effect of quantization eliminate these problems.

This paper is structured as follows. In section 2 we present the basic structure of the packet-loss recovery system. In section 3, we develop the estimators needed to make the structure work in practice. Section 4 shows experimental results, and section 5 presents the conclusions.

## 2. SCALABLE ROBUST CODING

We design our coding architecture such that it can be used with legacy source coding systems and so that the redundancy rate can be adjusted in a continuous manner in real time. Further flexibility can be obtained by using a basic source coder that can be adapted in rate in real time. We assume that the coders operate on a blockwise basis, and that each signal block is transmitted in an independent packet.

The basic design of our system, which is outlined in Figure 1, consists of a primary source coder (which can be a legacy coder) that encodes a random parameter vector $\mathbf{X}$ once per coding block and a (relatively low-rate) secondary coder that transmits additional information about $\mathbf{X}$ that is mostly redundant. The secondary coder provides an estimate of the parameter vector $\mathbf{X}$ of the primary coder in the current block based on *i)* parameter vectors in the previous blocks (and/or future blocks, e.g., from a jitter buffer) and possibly the present block, *ii)* redundant information transmitted in other packets, and *iii)* prior information about the distribution of the source-coder parameters.

The secondary coder exploits the parameter vector of the previous block by means of a predictive coding structure. To facilitate scalability, we use scalar quantizers (SQs) for the residuals of the components of $\mathbf{X}$ in the predictive coder. To ensure availability of the output of the predictive coder when a packet is lost, the information of the secondary coder of a block must reside in a different packet than the information corresponding to the primary coder.

The authors thank the Austrian Science Fund (FWF) for grant J2541-N15 and the Swedish Research Council for grant 2005-4107.

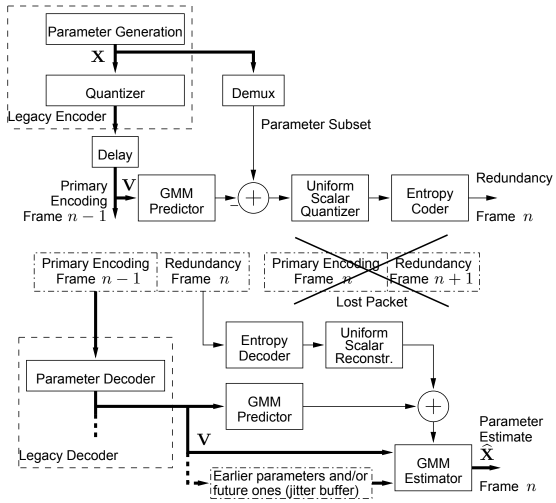


**Fig. 1**. Block diagram of the proposed redundancy layer encoder (top) and decoder for the scenario of recovering a lost parameter vector $\mathbf{X}$ (bottom).

The low rate of the secondary coder and the usage of SQs leads, in general, to the coding of only a subset of the primary-coder parameters in the secondary coder. Since the secondary coder uses the past-block information only to predict the value of a subset of the primary-coder parameters, a second stage is needed to obtain the estimates for the full set of primary-coder parameters. The need for only a subset of the parameters is a function of the dependencies between the parameters and of the distortion criterion that is used. This is most easily understood for parameters with identical marginal distribution and a uniform squared error criterion. Consider the limiting case where the parameters are identical. In this case, encoding one of the parameters at a rate of, say, $H$ results in significantly better performance than encoding both parameters at $H/2$. The validity of this argument decreases progressively with decreasing dependency between the parameters. If the parameters are independent, all must be encoded (not considering the effects of quantization at very low rates).

To ensure good performance, both the prediction stage and the full-parameter set estimation stage must be performed with general statistical estimation methods. When such statistical estimation methods are based on parameter distributions that are obtained from data based on a large number of relevant signals, they implicitly use prior knowledge of the signal. In our implementation we use GMMs to describe the joint parameter distributions that we use for our estimators as well as the scalar distributions of the prediction residuals to enable an adaptation of the entropy coder (arithmetic coder) in real time.

As will be seen in section 3, the scalability requirement makes the design of the estimators more challenging. For high performance and particularly, to obtain a relation between rate and estimation performance, the effect of quantization must be modeled and considered in the estimators.

For decoding we can initially distinguish four scenarios for each coding block: *i)* the bit sequence corresponding to both the primary and the secondary coder is received *ii)* the bit sequence of the primary coder is lost, but the bit sequence of the secondary coder is received *iii)* the bit sequence of the primary coder is received and that of the secondary coder is lost *iv)* the bit sequence of both coders is lost. It is straightforward to create an appropriate decoding structure for each situation. We assume the squared error criterion. When both bit sequences are lost, the conditional mean of the parameters given past (and/or future) parameters is the optimal reconstruction. If only the primary coder bit sequence is available, the primary coder output is used directly. If only the secondary coder bit sequence is available, the secondary coder functions normally and the estimator of the second decoding stage uses the output of the secondary coder and the output of previous blocks of the primary coder as input. If both bit sequences are available, the second stage estimator can additionally use the current block primary coder output as input. In practice, this improves the performance over having only the primary output available only when also the rate of the primary coder is low.

The four scenarios for the decoder can readily be generalized to the case that includes two or more sequential blocks with lost bit sequences and to the case that redundancy of more blocks is added to a packet. The estimators must be appropriately modified. Note that the information transmitted by the predictive quantizers is not adjusted.

## 3. ESTIMATION PROCEDURES

In this section, we discuss the statistical models and the estimation procedures that can be used to implement the proposed coding architecture. We base our estimation on a *single* Gaussian mixture model for the distribution of a random *supervector* $\mathbf{Z}$, which contains both the current block parameter vector and the parameter vectors of previous (and/or future) blocks:

$$p_{\mathbf{Z}}(\mathbf{z}) = \sum_{m=1}^{M} p_C(m)\,\mathcal{N}(\mathbf{z}|\boldsymbol{\mu}^{(m)}, \mathbf{C}^{(m)}), \tag{1}$$

where $\mathcal{N}(\mathbf{z}|\boldsymbol{\mu}^{(m)}, \mathbf{C}^{(m)})$ is a multivariate Gaussian density with index $m$, which has mean $\boldsymbol{\mu}^{(m)}$, covariance matrix $\mathbf{C}^{(m)}$, and prior probability $p_C(m)$. Importantly, our distribution model facilitates the modification of the quantizers in real time. The model is adjusted analytically for the effect of the quantizers.

### 3.1. Flexible Estimator using a single GMM

We denote the random parameter vector of the current block by $\mathbf{X}$ and the parameter vectors of previous (and/or future) blocks by $\mathbf{V}$. We further represent $\mathbf{V}$ and, if present, known parameters of the present block by a single vector $\mathbf{Y}$. A GMM-based minimum mean squared error (MMSE) estimator of $\mathbf{X}$ given $\mathbf{Y}$ is then [5]

$$\hat{\mathbf{x}} = \mathrm{E}\{\mathbf{X}|\mathbf{Y}=\mathbf{y}\} = \sum_{m=1}^{M} p_{C|\mathbf{Y}}(m|\mathbf{y})\,\boldsymbol{\mu}^{(m)}_{\mathbf{X}|\mathbf{Y}} \quad \text{with} \tag{2}$$

$$p_{C|\mathbf{Y}}(m|\mathbf{y}) = \frac{p_C(m)\mathcal{N}(\mathbf{y}|\boldsymbol{\mu}^{(m)}_{\mathbf{Y}}, \mathbf{C}^{(m)}_{\mathbf{Y}})}{p_{\mathbf{Y}}(\mathbf{y})}, \tag{3}$$

which requires

$$\boldsymbol{\mu}^{(m)}_{\mathbf{X}|\mathbf{Y}} = \boldsymbol{\mu}^{(m)}_{\mathbf{X}} + \mathbf{C}^{(m)}_{\mathbf{XY}}\left(\mathbf{C}^{(m)}_{\mathbf{Y}}\right)^{-1}\left(\mathbf{y} - \boldsymbol{\mu}^{(m)}_{\mathbf{Y}}\right) \tag{4}$$

$$\mathbf{C}^{(m)}_{\mathbf{X}|\mathbf{Y}} = \mathbf{C}^{(m)}_{\mathbf{X}} - \mathbf{C}^{(m)}_{\mathbf{XY}}\left(\mathbf{C}^{(m)}_{\mathbf{Y}}\right)^{-1}\mathbf{C}^{(m)}_{\mathbf{YX}}. \tag{5}$$

As was described in section 2, different scenarios may emerge in transmission over a packet network, and this can make the evaluation of equation (3) cumbersome. We prefer an estimation algorithm that is flexible and which facilitates estimation based on data vectors of different dimensions. This can be attained by training several statistical models and switching between them depending on the actual scenario. However, it is more convenient and elegant to use a single GMM and select the appropriate dimensions.

We train the GMM for the supervector $\mathbf{Z}$. The vectors $\mathbf{X}$ and $\mathbf{Y}$ are components of the the supervector $\mathbf{Z}$. That is, $\mathbf{X} = \mathbf{P_X Z}$ and $\mathbf{Y} = \mathbf{P_Y Z}$, where $\mathbf{P_X}$ and $\mathbf{P_Y}$ are permutation matrices that extract the appropriate elements of $\mathbf{Z}$. We note that not all elements of $\mathbf{Z}$ need to be used in a particular estimation task. We then use

$$\boldsymbol{\mu}_{\mathbf{X}}^{(m)} = \mathbf{P_X}\boldsymbol{\mu}^{(m)} \quad \text{and} \quad \boldsymbol{\mu}_{\mathbf{Y}}^{(m)} = \mathbf{P_Y}\boldsymbol{\mu}^{(m)} \tag{6}$$

for the mean vectors and

$$\mathbf{C}_{\mathbf{X}}^{(m)} = \mathbf{P_X C}^{(m)}\mathbf{P}_{\mathbf{X}}^{\mathsf{T}}, \quad \mathbf{C}_{\mathbf{XY}}^{(m)} = \mathbf{P_X C}^{(m)}\mathbf{P}_{\mathbf{Y}}^{\mathsf{T}},$$
$$\mathbf{C}_{\mathbf{YX}}^{(m)} = \mathbf{P_Y C}^{(m)}\mathbf{P}_{\mathbf{X}}^{\mathsf{T}}, \quad \text{and} \quad \mathbf{C}_{\mathbf{Y}}^{(m)} = \mathbf{P_Y C}^{(m)}\mathbf{P}_{\mathbf{Y}}^{\mathsf{T}} \tag{7}$$

for the covariance matrices. Instead of switching between different GMMs, we can now switch between different permutation matrices, which are specified by element indices only, resulting in a flexible and memory efficient implementation.

### 3.2. A Scalable Estimator

Suppose a GMM of the unquantized variables $\mathbf{X}, \mathbf{Y}$ is available, but that the estimation has to be performed given a scalar-quantized version $\check{\mathbf{Y}}$. The MMSE estimator requires the statistics of $\mathbf{X}, \check{\mathbf{Y}}$. It is possible to correct the covariance matrices of the unquantized variables to get the statistics of the quantized versions.

For each mixture component, we model the effect of the uniform SQs by additive, zero-mean, Gaussian noise independent of all unquantized variables

$$\begin{aligned} \check{Y}_i^{(m)} &= Y_i^{(m)} + N_i^{(m)}, \\ \mathrm{E}\left\{Y_j^{(m)} N_i^{(m)}\right\} &= 0, \ \forall i, j \in \{1, \ldots, d_y\} \end{aligned} \tag{8}$$

and also independent of the variables to be estimated. Given these assumptions, we only have to correct the diagonal of $\mathbf{C}_{\mathbf{Y}}^{(m)}$:

$$\mathbf{C}_{\check{\mathbf{Y}}}^{(m)} = \mathbf{C}_{\mathbf{Y}}^{(m)} + \mathrm{diag}\left[\sigma^2_{N_1^{(m)}}, \ldots, \sigma^2_{N_{d_y}^{(m)}}\right] \tag{9}$$
$$\mathbf{C}_{\mathbf{X}\check{\mathbf{Y}}}^{(m)} = \mathbf{C}_{\mathbf{XY}}^{(m)}. \tag{10}$$

The additive noise in our model does not represent the actual quantization noise. However, when we select the noise variance such that the mutual information between a Gaussian component and its noisy version is equal to the entropy of the quantized Gaussian, we obtain the same result as that obtained with the conventional backward-channel model of rate-distortion theory.

For an SQ $q_i(\cdot)$ that operates directly on $Y_i$ (i.e., without prediction), we compute the additive noise variance as

$$\sigma^2_{N_i^{(m)}} = \frac{\sigma^2_{Y_i^{(m)}}}{2^{2H(q_i(Y_i^{(m)}))} - 1}. \tag{11}$$

At sufficiently high resolutions, the entropy of the quantized Gaussian $H(q_i(Y_i^{(m)}))$ as a function of the step size $\Delta_i$ can be obtained by well-known high-rate approximations, and at vanishing rates, a look-up table is recommended.

### 3.3. Incomplete Description as Redundancy

As was stated in section 2, it is often optimal to have the secondary coder operate on an incomplete description of the parameter vector $\mathbf{X}$. In the simplest case, the subset contains only a single element $X_i$ of $\mathbf{X}$. To facilitate finding the best choice in terms of best estimation performance at a given rate, we can derive an approximation for the estimation error variance based on a simple Gaussian model (i.e., $M = 1$ in (1)).

For a Gaussian model, the MMSE estimate is $\hat{\mathbf{x}} = \boldsymbol{\mu}_{\mathbf{X|Y}}$ and the estimation error variance is $MSE = \mathrm{tr}\,\mathbf{C}_{\mathbf{X|Y}} = \mathrm{tr}\,\mathbf{C}_{\mathbf{X|V}} - MSE_\Delta$. $MSE_\Delta$ is the reduction in variance due to having $\check{X}_i = X_i + N_i$ in addition to the side information $\mathbf{V}$. Upon applying the inversion lemma for a partitioned matrix and simplification of the trace we obtain

$$MSE_\Delta = \frac{||\mathbf{C}_{\mathbf{X}X_i} - \mathbf{C}_{\mathbf{XV}}\mathbf{C}_{\mathbf{V}}^{-1}\mathbf{C}_{\mathbf{V}X_i}||^2}{\sigma^2_{N_i} + \sigma^2_{X_i} - \mathbf{C}_{X_i\mathbf{V}}\mathbf{C}_{\mathbf{V}}^{-1}\mathbf{C}_{\mathbf{V}X_i}}. \tag{12}$$

The variance $\sigma^2_{N_i}$ is a function of the rate spent on $X_i$ (cf. (11)). To account for the prediction prior to quantization (see Figure 1), the design of the SQ has to be based on the properties of the prediction residual $R_i$ instead of $X_i$. For the Gaussian model, the residual is also Gaussian with $\sigma^2_{R_i} = \sigma^2_{X_i} - \mathbf{C}_{X_i\mathbf{V}}\mathbf{C}_{\mathbf{V}}^{-1}\mathbf{C}_{\mathbf{V}X_i}$, and the additive noise variance in (12) is computed using $\sigma^2_{N_i} = \sigma^2_{R_i}/(2^{2H_i} - 1)$. Therefore, the rate saving due to predictive encoding is $H'_i = 1/2\log_2((\sigma^2_{X_i} + \sigma^2_{N_i})/(\sigma^2_{R_i} + \sigma^2_{N_i}))$. The optimal single variable $X_{i^*}$ is then easily found by $i^* = \arg\max_i MSE_\Delta$.

The form of (12) indicates a saturation in estimation performance at higher rates (or equivalently for $\sigma^2_{N_i} \ll \sigma^2_{X_i} - \mathbf{C}_{X_i\mathbf{V}}\mathbf{C}_{\mathbf{V}}^{-1}\mathbf{C}_{\mathbf{V}X_i}$). Therefore, for higher rates, a subset with more variables should be encoded. To find the optimal subset and bit allocation, (12) can be modified and used in an iterative manner. We replace $\mathbf{V}$ by $\mathbf{V}_i$ that combines side information and all variables of the current vector but $X_i$ and obtain $MSE_\Delta(H_i)$ as a function of the entropy $H_i$ while all other entropies $H_{j \neq i}$ are considered to be fixed. The form of (12) allows us to easily derive the slopes $\frac{\partial MSE}{\partial H_i}$ for all variables of the vector $\mathbf{X}$. We then iteratively redistribute rate from the variable with the flattest slope to the variable with the steepest slope under the constraints $\sum_i H_i = H$ and $H_i \geq 0$. The procedure converges quickly towards a rate distribution where the slopes of the variables with non-zero rate are equal.

Although the obtained rate distribution is strictly optimal only for Gaussian random variables and a linear estimator, we use it also for non-Gaussian data and the GMM MMSE estimator. Given a bit allocation, the actual SQs are designed using the properties of the actual residuals. Since our system works with relatively low rates, we use a look-up table for the quantizer step size versus residual entropy. Our second-stage estimator does not require a GMM that includes the residuals, but we calculate the additive noise variance to model $\check{X}_i^{(m)}$ based on a virtual step size $\Delta'_i$ such that $H(q_i(X_i)) = H_i + H'_i$.

## 4. APPLICATION TO LSF PARAMETERS

The line-spectral frequencies (LSF) are commonly used in state-of-the-art speech coders, e.g., [7]. We use these parameters (which are mutually dependent) to provide experimental evidence that the proposed procedure performs well on data encountered in practical systems.

The adaptive multi-rate (AMR) speech coder [7], operating at a rate of 10.2 kbit/s, serves as a base system. In this mode, the AMR codec spends 26 bits for the encoding of 10 LSFs using split vector quantization (SVQ) after first-order moving-average prediction. The LSF block rate is 50 Hz. The AMR codec is used to compute unquantized and quantized LSFs from the complete TIMIT database (after down-sampling to 8 kHz sampling rate), which provides us with more than 700,000 LSF vectors for GMM training and more than 200,000 vectors for testing.

The experiments simulate a scenario with a single lost LSF vector. Previously received vectors that are needed for the estimation of the lost one are available correctly at the decoder. One GMM

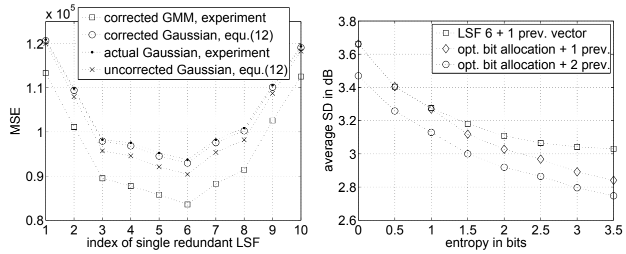


**Fig. 2**. Left: Estimation error variance when a single LSF (scalar-quantized with a 2 bit entropy constraint after prediction) of the current block is given in addition to the complete vector of the previous block. Right: Averaged log-spectral distortion over entropy.

**Table 1**. Optimal bit allocation for the secondary coder using a Gaussian model and assuming one previous LSF vector is given. All not shown numbers are 0.

| LSF | total entropy in bits | | | | | | |
|---|---|---|---|---|---|---|---|
| index | 0.50 | 1.00 | 1.50 | 2.00 | 2.50 | 3.00 | 3.50 |
| 1 | | | | | | | |
| 2 | | | | | | | |
| 3 | | | 0.46 | 0.58 | 0.68 | 0.77 | 0.78 |
| 4 | 0.16 | 0.36 | | | | | 0.20 |
| 5 | | | 0.51 | 0.58 | 0.60 | 0.67 | 0.65 |
| 6 | 0.34 | 0.47 | | 0.11 | 0.30 | 0.42 | 0.55 |
| 7 | | | 0.42 | 0.49 | 0.47 | 0.47 | 0.48 |
| 8 | | 0.17 | | | 0.11 | 0.24 | 0.35 |
| 9 | | | 0.11 | 0.24 | 0.34 | 0.43 | 0.49 |
| 10 | | | | | | | |

with $M = 32$ mixture components and full covariance matrices is trained using the EM algorithm for a supervector $\mathbf{Z}$ consisting of three consecutive LSF vectors (the first one unquantized and the latter two AMR-quantized). Thus, the model allows the second-stage estimator to take 0, 1, or 2 previous vectors as input. However, the first-stage estimator predicts always only from the first previous vector (as in Figure 1). The second-stage estimator corrects the GMM according to section 3.2 to achieve rate scalability. In addition to the GMM, a Gaussian model is estimated from the same data. This model will be used only in the bit allocation algorithm.

We show results from estimation experiments with different redundancy encodings. Unless otherwise noted, the final estimator uses one previous parameter vector as side information. The left plot of Figure 2 shows the estimation error variance (line with boxes) when a single LSF is transmitted at an entropy of 2 bits for its prediction residual. The best estimation performance is achieved by taking LSF 6. The plot also shows the predicted MSE (circles) based on the Gaussian model according to (12). To show that additive noise is a valid model for the SQs, in the left graph of Figure 2 also the performance of an estimator based on a Gaussian model for a supervector that includes scalar-quantized LSFs is plotted (points). The deviations from (12) are negligible. On the other hand, when the additive noise term in (12) is ignored, the MSE cannot be predicted accurately (crosses).

In the right plot of Figure 2, averaged log-spectral distortion as a function of the entropy is presented for three different configurations: *i)* only LSF 6 is encoded as redundancy (boxes), *ii)* the optimal LSF subset and bit allocation is used (diamonds), and *iii)* the same redundancy encoding as in *ii)* is used but the estimator takes two consecutive past LSF vectors as input (circles). The optimal bit allocation is determined iteratively as described in section 3.3 and is shown in Table 1. For a total entropy of 0.5 bit, the rate is allocated only among two LSFs. At higher rates more LSFs build the optimal subset. We note that LSFs 1, 2, and 10 are never selected, presumably because of strong dependencies on the selected LSFs.

Unlike the optimal encoding, the performance of the single LSF saturates around 2 bit. The gain by having the second previous vector available as side information is between 0.1 dB and 0.2 dB. A spectral distortion of the order of 3 dB with one previous vector and 2 bit redundancy makes the performance of our scalable method comparable to the VQ-based system in [4].

## 5. CONCLUSIONS

We propose a packet loss recovery strategy that *i)* is continuously rate-scalable, *ii)* can be applied to various loss scenarios, *iii)* can be implemented efficiently, *iv)* can easily be added to legacy systems, and *v)* can be used as an enhancement layer. It therefore constitutes an attractive solution to an existing problem.

The two core elements of the method are *i)* a GMM-based MMSE estimator that is loss-scenario adaptive and rate-scalable by the use of an analytical model for the quantizers, and *ii)* a redundancy layer consisting of an incomplete description. Individual parameters are efficiently encoded using simple scalar quantization and adaptive arithmetic coding, thus yielding continuous rate scalability. The proposed bit allocation algorithm generalizes 'reverse water-filling' to the problem of scalar encoding of dependent variables for the case of a final MMSE estimation stage and available side information. Our experiments using LSFs demonstrate that a small parameter subset as redundant data is the optimal choice at low rates typical for packet loss recovery. The spectral distortions achieved by our scalable method are comparable to those of existing, non-scalable methods.

## 6. REFERENCES

[1] M. Y. Kim and W. B. Kleijn, "Comparative rate-distortion performance of multiple description coding for real-time audiovisual communication over the internet," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 625–636, Feb. 2006.

[2] T. Fingscheidt and P. Vary, "Softbit speech decoding: A new approach to error concealment," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 3, pp. 240–251, Mar. 2001.

[3] J. Lindblom, J. Samuelson, and P. Hedelin, "Model based spectrum prediction," in *Proc. IEEE Workshop on Speech Coding*, Delavan, WI, Sept. 2000, pp. 117–119.

[4] Y. Agiomyrgiannakis and Y. Stylianou, "Coding with side information techniques for LSF reconstruction in voice over IP," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing*, vol. 1, Philadelphia, PA, Mar. 2005, pp. 141–144.

[5] R. Martin, C. Hoelper, and I. Wittke, "Estimation of missing LSF parameters using Gaussian mixture models," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing*, vol. 2, Salt Lake City, UT, May 2001, pp. 729–732.

[6] A. M. Peinado, Á. M. Gómez, V. Sánchez, J. L. Perez-Córdoba, and A. j. Rubio, "Packet loss concealment based on VQ replicas and MMSE estimation applied to distributed speech recognition," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing*, vol. 1, Philadelphia, PA, Mar. 2005, pp. 329–332.

[7] 3GPP TS 26.090, "Adaptive multi-rate (AMR) speech codec; transcoding functions," Technical Specification, V6.0.0, Dec. 2004.