Bit-Erasure Channel Decoding for GMM-based Multiple Description Coding

Yannis Agiomyrgiannakis, Yannis Stylianou

Department of Computer Science, University of Crete, Greece
jagiom@csd.uoc.gr, yannis@csd.uoc.gr

10.21437/Interspeech.2007-474

Abstract

Multiple Description Coding (MDC) is a plausible way to use the diversity of packet networks to increase the robustness of the transmission to packet losses. The redundancy that is introduced via MDC can also be used to increase the robustness of the transmission to bit-errors. This paper presents a novel decoding method for GMM (Gaussian Mixture Model)-based MDC in the presence of detected bit-errors. Particularly for speech transmission over bit-erasure channels, is is shown that the proposed method considerably improves the quality of the received speech spectral envelopes when one side-description is damaged. In highly correlated descriptions, for example, single and double bit-errors can almost be corrected.

Index Terms: Multiple Description Coding, Source/Channel Coding, bit-erasure channel decoding, Speech Coding, Voice-over-IP

1. Introduction

There is a growing need towards the transmission of multimedia over IP networks. Unfortunately, in IP networks there is no mechanism to assure the quality-of-service and the arrival of a packet may be delayed or the ordering of the packets may be altered. This is usually compensated by a buffer at the decoder, but the size of this buffer is restricted in real-time communications, resulting to increased packet losses. Multiple Description Coding (MDC) is a plausible framework to increase the robustness of communication upon packet losses. In MDC, the source is encoded in two side descriptions which are routed independently. If both descriptions arrive, we get a high-quality reconstruction of the source, referred to as *"central reconstruction"*. If only one description is received, an intermediate quality reconstruction is made, referred to as *"side reconstruction"*. The MDC framework can be seen as a mechanism to introduce redundancy into the bitstream by trading the distortion of the central reconstruction (referred to as *central distortion*) for the distortion of the side reconstructions (referred to as *side distortions*).

In the typical MDC framework, the network is modeled as a packet erasure channel (a description can be either received or lost); the possibility of introducing bit-errors is ignored. Introducing bit-error correction directly to MD bitstreams may be suboptimal [1]. However, this is not the case if the MD quantizer is optimized for both packet-erasure and bit-erasure channels [2]. The importance of error-resilient MDC becomes more apparent with the convergence between wireline and wireless networks due to the fact that wireless transmission suffers from bit-errors.

The first systematic approach on MDC was made by Vaishampayan in [3] with Multiple Description Scalar Quantization (MDSQ). This method is extended to transform coding by Vaishampayan and Batllo in [4]. A further extension to sources that can be modeled with a GMM is provided by Samuelsson and Plasberg in [5]. The latter method (referred to as GMM-MDSQ$_{TC}$ in the rest of the paper) can find applications in speech coding and particularly in the quantization of spectral envelopes is sinusoidal speech codecs [6],[7],[8],[9].

This paper extends GMM-MDSQ$_{TC}$ to the case of bit-erasure channels. It is shown that the redundancy introduced by GMM-MDSQ$_{TC}$ can be used to reduce the penalty of bit-errors that occur in one description. The alternative would be to disregard the damaged description and conduct a side-reconstruction using the undamaged description. The proposed method can be used to make a reconstruction that is significantly better than this side reconstruction. It is demonstrated that in a practical setting of MDC for speech spectral envelopes, single and double bit-errors can almost be corrected in terms of the mean-squared-error. However, correcting more bit-errors faces a significant computational barrier.

The proposed method is related to the bit-erasure channel decoding method initially introduced by Subramaniam in [10] (Chapter 6). When the number of bit-errors is known, the channel decoder provides a set of *candidate indices* that contains the transmitted indices. For example for a 30-bit encoding and 2 detected bit-errors there are 30*29=870 candidate indices and thus 870 candidate codevectors. In [10], a reconstruction is made by averaging these vectors according to the *conditional probability* of the current vector given the information of past vectors (speech spectral envelopes). In this paper, the correlated but undamaged side description is used to significantly reduce the number of candidate codevectors of the damaged side description while the averaging is made using the *probability* of the source. The extension of the proposed method to the case where both descriptions are damaged is also discussed.

The rest of the paper is organized as follows: Section 2 presents the necessary background for the paper, Section 3 describes the proposed bit-erasure channel decoding method for GMM-MDSQ$_{TC}$ and Section 4 presents an experimental evaluation. Finally, Section 5 concludes the paper.

2. Background

2.1. Multiple Description Scalar Quantization

In MDSQ, the correlations between the two descriptions are introduced via an *index-assignment* matrix that links the two side description codebooks with the central description codebook [3]. Let $\lambda \sim p(\lambda)$ be the scalar random variable which describes the encoded data and r be the total rate of the MDSQ system. Assuming balanced descriptions with equal rates, each description will then have a rate of $r_s = \frac{1}{2}r$ bits. Let $C_0 \equiv \{\hat{\lambda}^{(0)}_{i,j} : i = 1, ..., 2^{r_s}, j = 1, ..., 2^{r_s}\}$ be the codebook of the central description, $C_1 \equiv \{\hat{\lambda}^{(1)}_i : i = 1, ..., 2^{r_s}\}$ be the codebook of description 1 and $C_2 \equiv \{\hat{\lambda}^{(2)}_j : j = 1, ..., 2^{r_s}\}$ be the codebook of description 2.

The two side descriptions, say i and j, are routed independently to the receiver that has three decoders, a central decoder and two side decoders, one for each description. When both descriptions are received, the central decoder returns $\hat{\lambda}_{i,j}^{(0)} \in C_0$. When only the first description is received, side decoder 1 returns $\hat{\lambda}_i^{(1)} \in C_1$. Accordingly, when only the second description is received, side decoder 2 returns $\hat{\lambda}_j^{(2)} \in C_2$. The encoder finds the pair of indices (i, j) that minimizes the *total distortion*:

$$D_{tot} = D_0 + \frac{\rho}{1-\rho}(D_1 + D_2), \tag{1}$$

where ρ is the packet loss probability, D_0 is the mean central distortion and D_1, D_2 are the mean distortions for the two side descriptions, respectively. The total distortion measure in (1) is based on the assumption that each description is routed through a symmetric channel and that both channels have the same packet loss probability ρ. In practice, varying ρ provides a mechanism to exchange central distortion for side distortions. The index assignment mapping links a pair of side codebooks entries (i, j) to one central codebook entry. Note that at higher loss probabilities ρ, C_0 has less than 2^R codepoints. Therefore, not all possible pairs of indices (i, j) correspond to a central description codepoint. The encoder can produce only the pairs of indices (i, j) that are linked to a central description codepoint and those pairs will be referred to as *valid indices*. The corresponding central and side description codepoints will be referred to as *valid codepoints*.

Let $Q_{i,j}^{(0)}$ be the quantization cell associated with central codepoint $\hat{\lambda}_{i,j}^{(0)}$, $Q_i^{(1)}$ and $Q_j^{(2)}$ be the quantization cells associated with side codepoints $\hat{\lambda}_i^{(1)}$ and $\hat{\lambda}_j^{(2)}$, respectively. Let $I_j(i) \equiv \{j : \hat{\lambda}_{i,j}^{(0)} \in C_0\}$ be the set of valid indices j when i is known, and $I_i(j) \equiv \{i : \hat{\lambda}_{i,j}^{(0)} \in C_0\}$ be the set of valid indices i when j is known. Then, the side description quantization cells $Q_i^{(1)}$, $Q_j^{(2)}$ can be expressed as unions of the $Q_{i,j}^{(0)}$ cells:

$$Q_i^{(1)} = \bigcup_{j \in I_j(i)} Q_{i,j}^{(0)}, \quad Q_j^{(2)} = \bigcup_{i \in I_i(j)} Q_{i,j}^{(0)}, \tag{2}$$

When only the quantization cells $Q_{i,j}^{(0)}$ and the mappings $I_i(j)$, $I_j(i)$ are known, the optimal (in the mean-square-error sense) codepoints can be computed by taking expectations over the quantization cells using $p(\lambda)$:

$$\hat{\lambda}_{i,j}^{(0)} = \int_{Q_{i,j}^{(0)}} \lambda p(\lambda|i,j) d\lambda, \tag{3}$$

$$\hat{\lambda}_i^{(1)} = \sum_{j \in I_j(i)} \int_{Q_{i,j}^{(0)}} \lambda p(\lambda|i) d\lambda, \tag{4}$$

$$\hat{\lambda}_j^{(2)} = \sum_{i \in I_i(j)} \int_{Q_{i,j}^{(0)}} \lambda p(\lambda|j) d\lambda, \tag{5}$$

where $p(\lambda|i,j)$, $p(\lambda|i)$ and $p(\lambda|j)$ is the pdf of λ inside the quantization cells $Q_{i,j}^{(0)}$, $Q_i^{(1)}$ and $Q_j^{(2)}$, respectively.

Construction techniques for the design of $N(0,1)$ MDSQ codebooks and the corresponding index assignment tables can be found in [3] and [11].

2.2. Transform Coding based on MDSQ

The scalar MDSQ codebooks can be used to construct a multiple description transform coding system (MDSQ$_{TC}$). A high-rate analysis for such a system is provided by Batllo and Vaishampayan in [4]. According to this analysis, the optimal bit allocation for resolution-constrained quantization of the multivariate Gaussian case (with diagonal covariance matrix) is:

$$r_p = \frac{\frac{1}{2}R}{P} + \frac{1}{2}log_2\left(\frac{\sigma_p^2}{\left(\prod_{i=1}^{P}\sigma_p^2\right)^{\frac{1}{P}}}\right), \tag{6}$$

where P is the number of encoded Gaussians, R is the total rate assigned to the Gaussians, r_p is the rate assigned to the p-th Gaussian random variable, and σ_p^2 its variance.

2.3. GMM-based MDSQ$_{TC}$

Sources like the spectral envelopes have non-Gaussian distributions which are usually modeled via GMM. Efficient high-rate quantization for these sources can be made by GMM-based quantization, initially proposed by Subramaniam in [10]. In GMM-based quantization, the source is encoded with a set of transform coders and the “best" (in the mean-square-error sense) encoding is transmitted through the network together with the index of the corresponding encoder. Each of these transform coders is designed according to the statistics of a multivariate Gaussian component of a GMM. This idea is extended to multiple description coding by Samuelsson and Plasberg in [5]. We shall refer to this scheme as GMM-MDSQ$_{TC}$.

In GMM-MDSQ$_{TC}$, the data vector $\mathbf{x} \in \mathbb{R}^P$ is encoded with M MDSQ$_{TC}$ encoders and the “best" encoding is transmitted along with the index of the correspond encoder. Let $\alpha_{x,m}$ be the prior probability, $\mu_{x,m}$ the mean and $\Sigma_{x,m}$ the covariance matrix of the m-th Gaussian component of the GMM. An eigenvalue decomposition is made to each covariance matrix $\Sigma_{x,m}$:

$$\Sigma_{x,m} = V_{x,m}\Lambda_{x,m}V_{x,m}^T \tag{7}$$

where the columns of $V_{x,m}$ are the eigenvectors of $\Sigma_{x,m}$ and

$$\Lambda_{x,m} = diag(\sigma_{m,1}^2,\ \sigma_{m,2}^2,\ \ldots,\ \sigma_{m,P}^2)$$

is a diagonal matrix with the eigenvalues (variances) $\sigma_{m,p}^2$ on it's diagonal.

The m-th MDSQ$_{TC}$ encoder assumes that the statistics of $\mathbf{x}$ follow the statistics of the m-th Gaussian component of the GMM, namely $N(\mu_{x,m}, \Sigma_{x,m})$. Vector $\mathbf{x}$ is translated and rotated in order to obtain a zero mean vector $\mathbf{x}'_m$ with diagonal covariance matrix $\Lambda_{x,m}$:

$$\mathbf{x}'_m = V_{x,m}^T(\mathbf{x} - \mu_{x,m}). \tag{8}$$

The *uncorrelated* vector $\mathbf{x}'_m$ is then quantized with a series of MDSQ encoders to obtain M sets of central and side reconstructions $\{\hat{\mathbf{x}}'_{0,m}, \hat{\mathbf{x}}'_{1,m}, \hat{\mathbf{x}}'_{2,m}\}$, $m = 1, ..., M$. The encoding that minimizes the distortion function

$$D_m = \|\mathbf{x}'_m - \hat{\mathbf{x}}'_{0,m}\|_2^2 + \frac{\rho}{1-\rho}\left(\|\mathbf{x}'_m - \hat{\mathbf{x}}'_{1,m}\|_2^2 + \|\mathbf{x}'_m - \hat{\mathbf{x}}'_{2,m}\|_2^2\right), \tag{9}$$

where we recall that ρ is the packet loss probability, is selected for transmission. Note that D_m is based on the assumption that each description is routed through an independent symmetric channel with loss probability ρ. Furthermore, the distortion is efficiently calculated using $\mathbf{x}'_m$ instead of $\mathbf{x}$ because the distortion measure is not affected by $V_{x,m}^T$ which is a unitary transform. Let m' be the encoding that minimizes D_m. Then, the indices of the first and the second description are of the form:

$$I_1 = \{m', i_1, i_2, ..., i_P\}, \tag{10}$$

$$I_2 = \{m', j_1, j_2, ..., j_P\}, \tag{11}$$

where (i_p, j_p) refer to the index pairs of the MDSQ quantizer that encode the p-th dimension of $\mathbf{x}'_{m'}$. The m'-th MDSQ$_{TC}$ decoder performs the inverse operation. If both descriptions are received, the decoder returns the central reconstruction $\hat{\mathbf{x}}'_{0,m'}$. If only one of these descriptions is received, then the corresponding reconstruction $\hat{\mathbf{x}}'_{1,m'}$ or $\hat{\mathbf{x}}'_{2,m'}$ is obtained. Finally, the resulting codevector is obtained by rotating and translating the reconstruction $\hat{\mathbf{x}}'_{k,m'}$, $k = 0, 1, 2$ according to the equation:

$$\hat{\mathbf{x}} = \mu_{m'} + V_{x,m'}\hat{\mathbf{x}}'_{k,m'}. \tag{12}$$

Details regarding the bit-allocation for GMM-MDSQ$_{TC}$ can be found in [5].

3. Bit-Erasure Channel Decoding for GMM-based MDSQ

Assume that description 1 (index I_1) is received without bit-errors, while description 2 contains bit-errors. The channel decoder provides a set of L candidate indices $U_2 = \{I_{2,l} : l \in 1, \cdots, L\}$ for the second description. The indices I_1, $I_{2,l}$ are decomposed to the following set of component and scalar quantization indices:

$$I_1 = \{m', i_1, i_2, \cdots, i_P\}, \tag{13}$$

$$I_{2,l} = \{m_l, j_{1,l}, j_{2,l}, \cdots, j_{P,l}\},\ l = 1, \cdots, L \tag{14}$$

where P is the number of dimensions. Only a subset of the candidate indices U_2 are *valid*, in the sense that the pair $(I_1, I_{2,l})$ can be produced from the encoder. Let $U'_2 \subseteq U_2$ be the set of *valid candidate indices*:

$$U'_2 = \{I_{2,l} : m_l = m' \ \&\ (i_p, j_{p,l}) \text{ is valid } \forall p \in \{1, \cdots, P\}\}, \tag{15}$$

where we recall that a pair of indices $(i_p, j_{p,l})$ is considered to be valid if the MDSQ index assignment matrix maps the indices to a central codebook entry. Note that the component index m' is transmitted twice, one in each description. Therefore every valid candidate index must have the same component index. For convenience, let us renumber the elements in U'_2 so that $U'_2 = \{I_{2,l}, l = 1, \cdots, L'\}$, where L' is the size of the set U'_2.

Following the nomenclature of Section 2.3, let $\mathbf{x}'_{m'} = [x'_{m',1}, x'_{m',2}, \cdots, x'_{m',P}]^T$ be the P uncorrelated variables that are encoded with the m'-th MDSQ$_{TC}$ quantizer, and $\sigma^2_{m',1}$, $\sigma^2_{m',2}$, $\ldots$, $\sigma^2_{m',P}$ be the corresponding variances. Let $r_{m',p}$ be the number of bits allocated to each of the $x'_{m',p}$, $p = 1, ..., P$. A set of pre-trained MDSQ codebooks for the $N(0, 1)$ Gaussian is obtained using the methods in [3],[11]: one pack of MDSQ codebooks (central and side codebooks) for each loss probability ρ and each rate r_p. The GMM-MDSQ$_{TC}$ encoder/decoder operates with the subset of MDSQ codebooks that corresponds to the loss probability of the network. Let $\hat{\lambda}^{(0)}_{i,j}(r)$ be the (i, j)-th entry of the central description codebook trained for $N(0, 1)$ variables at a side description rate of r bits, and $S_{i,j}(r)$ the length of the corresponding quantization cell. Let $\mathbf{y}^{(l)}$, $l = 1, ..., L'$ be the *valid candidate codevectors*:

$$\mathbf{y}^{(l)} = \begin{bmatrix} \sigma_{m',1}\hat{\lambda}^{(0)}_{i_1,j_{1,l}}(r_{m',1}) \\ \sigma_{m',2}\hat{\lambda}^{(0)}_{i_2,j_{2,l}}(r_{m',2}) \\ \vdots \\ \sigma_{m',P}\hat{\lambda}^{(0)}_{i_P,j_{P,l}}(r_{m',P}) \end{bmatrix}. \tag{16}$$

Let $Q^{(l)}$ be the P-dimensional quantization cell associated with codevector $\mathbf{y}^{(l)}$:

$$Q^{(l)} = Q_{i_1,j_{1,l}} \times Q_{i_2,j_{2,l}} \times \cdots \times Q_{i_P,j_{P,l}}, \tag{17}$$

where $Q_{i_p,j_{p,l}}$ is the (scalar) central description quantization cell associated with the codepoint $\sigma_{m',p}\hat{\lambda}^{(0)}_{i_p,j_{p,l}}(r_{m',p})$. The cell $Q_{i_p,j_{p,l}}$ is an interval of length $S_{i_p,j_{p,l}}(r_{m',p})\sigma_{m',p}$. The optimal MSE (Mean Square Error) reconstruction is given by the following formula:

$$\tilde{\mathbf{x}}'_{1,m'} = \sum_{l=1}^{L'} \left(\frac{\int_{\mathbf{x}\in Q^{(l)}} \mathbf{x}p(\mathbf{x}; m')d\mathbf{x}}{\sum_{k=1}^{L'} \int_{\mathbf{x}\in Q^{(k)}} p(\mathbf{x}; m')d\mathbf{x}} \right), \tag{18}$$

where $p(\mathbf{x}; m')$ is the pdf of the m'-th Gaussian component, namely $N(0, \Lambda_{x,m'})$. A computationally attractive high-rate approximation can be made if we assume that $p(\mathbf{x}; m)$ is approximately constant inside each quantization cell $Q^{(l)}$, therefore:

$$\tilde{\mathbf{x}}'_{1,m'} \approx \sum_{l=1}^{L'} \frac{w_l}{\sum_{k=1}^{L'} w_k} \mathbf{y}^{(l)}, \text{ where}$$

$$w_l = \prod_{p=1}^{P} p(\hat{\lambda}^{(0)}_{i_p,j_{p,l}}(r_{m',p}))S_{i_p,j_{p,l}}(r_{m',p})\sigma_{m',p}, \tag{19}$$

where $p(\lambda)$ is the $N(0, 1)$ pdf.

The decoded GMM-MDSQ$_{TC}$ reconstruction can be obtained by rotating and translating $\tilde{\mathbf{x}}'_{1,m'}$ according to the statistics of the m'-th Gaussian component of the GMM (see equation (12)):

$$\tilde{\mathbf{x}}_1 = \mu_{m'} + V_{x,m'}\tilde{\mathbf{x}}'_{1,m'}. \tag{20}$$

The reconstruction formula (19) has a complexity that increases linearly with L', but the number L' of valid candidate codevectors increases rapidly with the number of detected but uncorrected bit-errors. Assume that each description is quantized with 30 bits, and that the channel code only detects bit-errors. The number of candidate codevectors when there are K bit-errors is $L = \begin{pmatrix} 30 \\ K \end{pmatrix}$. For $K = 1 \Rightarrow L = 30$, for $K = 2 \Rightarrow L = 30 * 29 = 870$, while for $K = 3 \Rightarrow L = 30 * 29 * 28 = 24360$. The number of valid candidate codevectors is smaller than L: $L' \leq L$, but the complexity remains overwhelming for more than 2 bit-errors because all candidate codevectors should be checked for validity. A way to tackle the complexity is to split the bits of each description into two parts and use a different error detection code for each part.

The case where both descriptions are damaged with bit-errors can be handled in a similar manner, using equation (19). The only difference in this case is that the candidate indices and the valid indices (those that could originate from the GMM-MDSQ$_{TC}$ encoder) should now constructed from both descriptions. Such a scheme provides a better reconstruction that the average of the source (the alternative solution when both descriptions are damaged), but even such a reconstruction cannot be considered as good for a speech coding application. However, if additional correlated information is available, the integration in equation (19) can be made using the conditional pdf of the source instead of the source pdf in order to provide an improved reconstruction.

4. Experiments and Results

An evaluation is made using 20-dimensional RCC (Real Cepstrum Coefficient) spectral envelopes derived from 20 ms speech frames. Each frame was analyzed with a harmonic sinusoidal model that estimates the amplitudes and the phases of the harmonics by solving a least squares problem. Harmonic amplitudes were fitted to a 20-th order regularized cepstral envelope in Bark scale [9],[12]. These features are suitable for sinusoidal speech coding [9], voice conversion [12] and they are similar to cepstral coefficients used in [13],[14] for speech coding. We trained a GMM with 16 components using 400.000 samples from the training set of the TIMIT database, while the test-set consisted of 10.000 samples from the corresponding test-set of TIMIT. The RCC source was encoded with 60 bits using two balanced 30 bit descriptions. Let D_0 be the distortion of the central reconstruction and D_1, D_2 be the distortions of the corresponding side reconstructions. A number of 1 and 2 bit-errors were introduced to description 2 and the proposed decoder was evaluated using the measured distortions $D_{ber,1}$, $D_{ber,2}$, respectively. The candidate indices for each vector were computed by introducing 1-2 random bit-errors to the 30 bit description and changing every possible set of two bits; a total of $\begin{pmatrix} 30 \\ 1 \end{pmatrix} = 30$ candidates for the 1 bit error case and a total of $\begin{pmatrix} 30 \\ 2 \end{pmatrix} = 30 * 29 = 870$ candidates for the 2 bits error case. The evaluation is made with the MSE (Mean-Square-Error) distortion measure using the raw RCC parameters. The results are depicted in Figure 1. The distortions were evaluated for several loss probabilities ρ ranging from 0 to 0.5. At $\rho = 0$, the central distortion D_0 is very low but the side distortions D_1, D_2 are very high. As the loss probability increases, the side distortions become lower at the cost of a higher central distortion. From $\rho = 0.2$ and above, the central/side descriptions almost converge to the state of highest redundancy. At $\rho = 0.5$, the side distortions D_1, D_2 are equal to the distortion provided by a GMM-based quantizer at a rate of 30 bits. When the second description has bit-errors, distortion $D_{ber,1}$ (or $D_{ber,2}$) provided by the proposed method is much lower than D_1 which corresponds to a complete loss of description 2. At the higher redundancy point ($\rho = 0.5$) the proposed method almost corrects the single/double bit errors, providing a reconstruction that is close to the one provided by the central description. Furthermore, the benefits of the proposed method appear even from low redundancies ($\rho \geq 0.03$).

A practical aspect of this work is that it makes possible to reallocate a portion of the redundancy provided for bit-error correction to the GMM-MDSQ$_{TC}$ quantizer. For example, the bit-error correction capability of the channel coding could be reduced by 1, allowing to reallocate the saved bits to the GMM-MDSQ$_{TC}$ quantizer in order to provide a better reconstruction when there are no bit-errors.

5. Discussion

A novel bit-erasure channel decoding method for GMM-based MDC is presented. It is shown that the correlations between the descriptions can be used to considerably reduce the impact of single and double bit-errors that occur in a single side description. However, the complexity increases rapidly for more than two bit-errors. A practical aspect of this work is that it justifies a design where a portion of the redundancy that is given for bit-erasure channel coding can be reallocated to the GMM-based MDC quantizer.

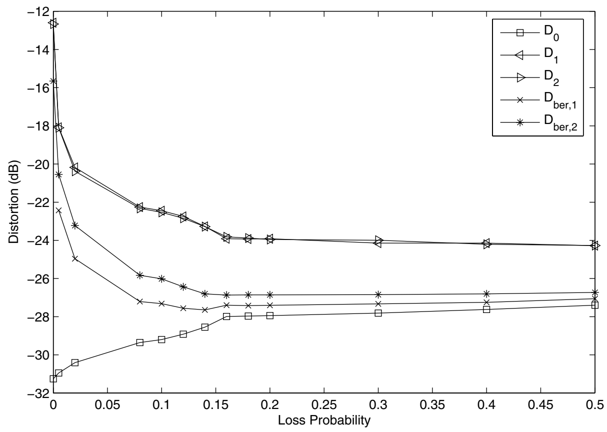

Figure 1: GMM-MDSQ$_{TC}$ for an 20-dimensional RCC source of speech spectral envelopes. Each description is encoded using 30 bits. D_0,D_1 and D_2 correspond to the average distortions of the central and the two side descriptions, respectively. $D_{ber,1}$ and $D_{ber,2}$ corresponds to the average distortion when description 2 contains 1 and 2 bit-errors. Distortions are depicted in dB scale.

6. References

[1] H. Coward, R. Knopp, and B. Servetto, "On the performance of multiple description codes over bit error channels." in *Proc. of IEEE Int. Symposium on Information Theory*, vol. 3, June 2001.

[2] Y. Zhou and W.-Y. Chan, "Multiple Description quantizer design using a channel optimized quantizer approach," 2004.

[3] V. Vaishampayan, "Design of Multiple Description Scalar Quantizers," *IEEE Transactions on Information Theory*, vol. 39, no. 3, p. 821, 1993.

[4] J. Batllo and V.A.Vaishampayan, "Asymptotic Performance of Multiple Description Transform Codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 703–707, 1997.

[5] J. Samuelsson and J. Plasberg, "Multiple Description Coding based on Gaussian Mixture Models," *IEEE Signal Processing Letters*, June 2005.

[6] R. J. McAulay and T. F. Quatieri, "Sinusoidal Transform Coding," in *Jet Propulsion Lab., Proceedings of the Mobile Satellite Conference p 503-508 (SEE N88-25680 19-32)*, May 1988, pp. 503–508.

[7] R. Rudolph and E. Yu, "IMBE and AMBE speech compression," in *International IC'99*, 1999, pp. 232–234.

[8] B. W. Kleijn and K. K. Paliwal, *Speech Coding and Synthesis*. Elsevier, 1995.

[9] Y. Agiomyrgiannakis, "Sinusoidal Coding of Speech for Voice-over-IP." Ph.D. dissertation, Department of Computer Science, University of Crete, http://www.csd.uoc.gr/˜jagiom/voip.htm, Jan. 2007.

[10] A. D. Subramaniam, "Gaussian Mixture Models in compression and communication," Ph.D. dissertation, University of California, San Diego, 2003.

[11] P. Koulgi, S. L. Regunathan, and K. Rose, "Multiple Description quantization by deterministic annealing," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, 2003.

[12] Y. Stylianou, "Harmonic-plus-noise Models for speech, combined with statistical methods for speech and speaker modification," Ph.D. dissertation, Ecole Nationale. Sup'erieure des Telecommunications, Paris, France, 1998.

[13] T. Eriksson, H.-G. Kang, and Y. Stylianou, "Quantization of the spectral envelope for sinusoidal coders," in *Proc. ICASSP*, 1998.

[14] T. F. Quatieri, *Discrete-Time Speech Signal Processing: Principles and Practice*. Prentice Hall, 2001.