Wyner-Ziv Encoded Predictive Multiple Descriptions

Ashish Jagmohan, Narendra Ahuja
Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL-61801
Contact Author: Ashish Jagmohan, email: jagmohan@uiuc.edu, phone: 217-333-1869

Abstract

The predictive Multiple Description coding problem can be posed as a variant of the well-known Wyner-Ziv side-information problem. Predictive MD coding in this framework (termed the WYZE-PMD framework) eliminates the problem of predictive mismatch without requiring restrictive channel assumptions or high latency. In the present paper we analyze the performance of two-channel one-step predictive MD coding within the WYZE-PMD framework. Specifically, we obtain achievable rate-distortion (R-D) regions for the problem of MD coding in the presence of correlated decoder side-information, and use these to obtain the operational R-D performance for predictive MD coding under certain restrictions. We propose practical code constructions within the WYZE-PMD framework, and compare the performance of the proposed codes with conventional approaches, for communication of a first-order Gauss-Markov source over two erasure channels with independent failure probabilities. Results indicate that the proposed approach significantly out-performs conventional approaches, in terms of R-D performance.

1 Introduction

Multiple Description (MD) coding involves coding source information into multiple descriptions for lossy transmission, such that the quality of reconstruction degrades gracefully with the number of descriptions lost. MD coding of predictively encoded sequences (termed predictive MD coding) is of practical interest in low-latency multimedia applications that involve communication of compressed video/audio data using error-prone channels. Examples of such applications include video-conferencing and Internet broadcast of video/audio streams, and robust storage of video/audio data on redundant disks. The key problem encountered in predictive MD coding is that of predictive mismatch, which refers to a scenario in which there is a mismatch between the predictor symbols at the encoder and the decoder. In the context of MD coding, this mismatch arises because the subset of predictor symbol descriptions received at the decoder is unknown to the encoder.

For the two-channel case, previous approaches for predictive MD coding include the techniques described in [1, 2, 3]. The main shortcoming of these approaches is the strong assumptions required on channel behaviour to eliminate predictive mismatch. In particular, it is required that the subset of channels received over time remains fixed. This assumption is rarely satisfied in applications involving multimedia transmission over lossy

Proceedings of the Data Compression Conference (DCC'03)
1068-0314/03 $17.00 © 2003 IEEE

networks. Other techniques, such as [4] and the video coding technique presented in [1], avoid these assumptions at the cost of sacrificing low latency.

We have recently shown, in [5], that the predictive MD coding problem can be posed as a variant of the well-known Wyner-Ziv decoder side-information problem [6]. We have also shown, in [7], that performing predictive MD coding by transmitting coset information eliminates the problem of predictive mismatch without requiring restrictive assumptions or high latency. We term this framework Wyner-Ziv Encoded Predictive Multiple Descriptions or WYZE-PMD.

The aim of the present paper is to analyze the performance of two-channel one-step predictive MD coding within this framework. Specifically, we present the following results:

- We obtain achievable rate-distortion (R-D) regions for the problem of MD coding with correlated side-information available a-priori at the decoder.
- We obtain operational R-D performance expressions (using the MSE distortion measure) for two-channel predictive MD coding of a first-order Gauss-Markov source, under certain restrictions.
- We propose code constructions, within the WYZE-PMD framework, based on MD scalar quantizers and turbo codes. We show that the proposed constructions significantly out-perform conventional predictive MD coding approaches, in terms of R-D performance.

2 Predictive MD Coding using Coset Codes

In this section we give an example of a coset code construction for predictive MD coding of a real-valued source. Our aim is to illustrate the principles involved in WYZE-PMD coding. Further details can be found in [7].

2.1 Problem Description

Consider the communication of a M-dimensional source with memory, $\{\mathbf{V}_i\}_{i=1}^{\infty}, \mathbf{V}_i \in \mathbb{R}^M$, across a lossy channel using one-step predictive coding. Given the decoder reconstruction of source symbol $\mathbf{V}_{k-1}$ (denoted $\widehat{\mathbf{V}}_{k-1}$) the encoder communicates $\mathbf{V}_k$ by generating the innovation $\mathbf{T}_k = \mathbf{V}_k - E[\mathbf{V}_k|\widehat{\mathbf{V}}_{k-1}]$ which is input to the channel, where $E[\cdot]$ represents the expectation operator.

For the case of N-channel MD coding, the decoder reconstruction $\widehat{\mathbf{V}}_{k-1}$ can take one of multiple values, depending on which of the 2^N possible subsets of descriptions of $\mathbf{V}_{k-1}$ is received i.e. $\widehat{\mathbf{V}}_{k-1} \in \mathcal{R}_{k-1}$ where $\mathcal{R}_i = \{\widehat{\mathbf{V}}_i^j\}$ denotes the reconstruction set for the i^{th} symbol. The number of possible predictors grows exponentially with time and, in general, $|\mathcal{R}_k| = 2^{kN}$.

The problem of predictive MD coding can be formulated as a variant of the WZ decoder side-information problem as follows. The decoder reconstruction of the predictor $\widehat{\mathbf{V}}_{k-1}$ takes values in the reconstruction set $\mathcal{R}_{k-1} = \{\widehat{\mathbf{V}}_{k-1}^j\}$ with a probability mass function determined by the channel failure probabilities, $P(\widehat{\mathbf{V}}_{k-1} = \widehat{\mathbf{V}}_{k-1}^j) = q(j), j \in \{1, \ldots, |\mathcal{R}_{k-1}|\}, \sum_j q(j) = 1$. Thus, the encoder is required to compress $\mathbf{V}_k$ in the presence of the correlated decoder side-information $\widehat{\mathbf{V}}_{k-1}$, when the only information the encoder has about $\widehat{\mathbf{V}}_{k-1}$ are it's statistics, i.e. $\mathcal{R}_{k-1}$ and $q(\cdot)$.

Proceedings of the Data Compression Conference (DCC'03)
1068-0314/03 $17.00 © 2003 **IEEE**

2.2 Coset Code Construction

Consider a M-dimensional lattice Λ, for which the lattice quantization distortion is the desired MD central distortion D_0. Consider a sub-lattice $\Lambda' \subset \Lambda$ which induces a partition of Λ into $|\Lambda|\Lambda'|$ cosets of Λ' [8].

For the N-channel case, the WYZE-PMD encoder consists of the nearest-neighbour lattice quantization function $\mathbf{Q}(\cdot) : \mathbb{R}^M \to \Lambda$, the coset index function $\mathbf{C}(\cdot) : \Lambda \to \Lambda|\Lambda'$ which finds the index of the coset to which a lattice point belongs, and the MD coding functions $\mathbf{a}_i(\cdot) : \Lambda|\Lambda' \to J$, $i \in \{1, \ldots, N\}$ where i indexes the channels and J is the set of description labels. To communicate source symbol $\mathbf{V}_k$ the encoder transmits the innovation description $\mathbf{T}_k^i = \mathbf{a}_i(\mathbf{C}(\mathbf{Q}(\mathbf{V}_k)))$ on the ith channel.

The WYZE-PMD decoder consists of the MD decoding functions $\mathbf{g}_j(\cdot) : J \to \Lambda|\Lambda'$, $j \in \{0,1\}^N$ and the coset decoding function $\widehat{\mathbf{C}}(\cdot,\cdot) : \Lambda|\Lambda' \times \Lambda \to \Lambda$. For a given set of received descriptions denoted by $j_{dec} \in \{0,1\}^N$, the decoder invokes the appropriate MD decoding function $\mathbf{g}_{j_{dec}}$ to reconstruct the transmitted coset index $\widehat{c}_k^{j_{dec}}$. The coset decoding function decodes $\widehat{c}_k^{j_{dec}}$ to the coset lattice point closest to the MMSE estimator of $\mathbf{V}_k$ based on the decoder predictor $\widehat{\mathbf{V}}_{k-1}$ (which acts as side-information). It can be shown (cf. [7]) that the decoder can correctly decode $\mathbf{V}_k$, with fidelity upto the channel loss in the transmission of $\{\mathbf{T}_k^i\}$ if the minimum distance of the sublattice Λ' is large enough compared to the expected distance between $\mathbf{V}_k$ and $\widehat{\mathbf{V}}_k$.

The construction described above does not require any assumptions on the set $\mathcal{R}_{k-1}$, does not increase the latency with respect to one-step predictive coding, and requires the generation of multiple descriptions of only one innovation, leading to efficient coding. Finally, we note that the order in which the coset indices and the MD indices are generated can be reversed by employing multiple $\{\Lambda_j, \Lambda_j'\}_{j=1}^N$ codebook pairs.

3 R-D Region for MD Coding with Side-information

In this section we present achievable R-D regions for two-channel MD coding with correlated side-information present a-priori at the decoder. The results obtained will be subsequently used to analyze predictive MD coding within the WYZE-PMD framework.

3.1 Achievable R-D Region for Discrete Alphabet Sources

Consider a sequence, $\{X_i, Y_i\}_{i=1}^n$, $X_i \in \mathcal{X}$, $Y_i \in \mathcal{Y}$, of independent copies of a pair of random variables with joint distribution $p(x,y)$. X denotes the source which is to be coded, and Y is the side-information available only at the decoder. Let $\mathcal{I}_m = \{1, 2, \ldots, m\}$.

An $\{n, 2^{nR_1}, 2^{nR_2}, D_0, D_1, D_2\}$ MD code, for this scenario, consists of two encoding functions $\mathbf{a}_j : \mathcal{X}^n \to \mathcal{I}_{2^{nR_j}}, j = 1, 2$ and three decoding functions

$$\mathbf{g}_j : \mathcal{I}_{2^{nR_j}} \times \mathcal{Y}^n \to \widehat{\mathcal{X}}_j^n, j = 1, 2, \quad \mathbf{g}_0 : \mathcal{I}_{2^{nR_1}} \times \mathcal{I}_{2^{nR_2}} \times \mathcal{Y}^n \to \widehat{\mathcal{X}}_0^n$$

where $\widehat{\mathcal{X}}_j$ denote the decoder reconstruction alphabets. The expected reconstruction distortion tuple (D_0, D_1, D_2) for the code is given by $D_j = \frac{1}{n}\sum_{k=1}^n E[d_j(X_k, \widehat{X}_{jk})]$, $j = 0, 1, 2$, where $d_j(\cdot\cdot) : \mathcal{X} \times \widehat{X} \to \mathbb{R}^+$ are appropriate distortion measures. A quintuple $(R_1, R_2, D_0, D_1, D_2)$ is said to be achievable if for every $\epsilon > 0$, there exists for some n an $(n, 2^{nR_1}, 2^{nR_2}, \tilde{D}_0, \tilde{D}_1, \tilde{D}_2)$ code with $\tilde{D}_j \leq D_j + \epsilon, j = 0, 1, 2$.

Theorem 3.1 *An achievable rate region for a given distortion tuple (D_0, D_1, D_2) is given by the union of (R_1, R_2) satisfying*

$$R_1 \geq I(X;U_1) - I(U_1;Y), \quad R_2 \geq I(X;U_2) - I(U_2;Y)$$
$$R_1 + R_2 \geq I(X;U_0,U_1,U_2) + I(U_1,U_2) - I(U_1;Y) - I(U_2;Y) - I(U_0;Y|U_1,U_2) \qquad (1)$$

for some p.m.f. $p(x,u_0,u_1,u_2) = p(x).p(u_0,u_1,u_2|x)$ *such that* $D_j \geq \frac{1}{n}\sum_{k=1}^{n} E[d_j(X_k, \widehat{X}_{jk})]$, $j = 0,1,2$ *for the decoding functions* $\mathbf{g}_j$.

Proof: The proof methodology is similar to that used in [9, 10]. We present a brief outline.
For $j = 1,2$, draw $2^{nR'_j}$ n-vectors $\{\mathbf{U}_j(i) : i = 1,2,\ldots,2^{nR'_j}\}$ independently and uniformly using the respective marginal density $p(u_j)$. Next, for each jointly typical $(\mathbf{u}_1, \mathbf{u}_2)$ in this list, draw $2^{nR'_0}$ vectors $\{\mathbf{U}_0(i) : i = 1,2,\ldots,2^{nR'_0}\}$ using the conditional marginal density $p(u_0|u_1,u_2)$. For encoding a given $\mathbf{x} \in \mathcal{X}^n$, find a triple (c_0, c_1, c_2) such that $(\mathbf{x}, \mathbf{u}_0(c_0), \mathbf{u}_1(c_1), \mathbf{u}_2(c_2))$ are jointly typical. This can be done if

$$R'_1 \geq I(X,U_1), \quad R'_2 \geq I(X,U_2), \quad R'_1 + R'_2 \geq I(X;U_1,U_2) + I(U_1;U_2) \qquad (2)$$

Next, each of the three codebooks is partitioned randomly into $2^{n\tilde{R}_j}, j = 0,1,2$ bins. Let b_j denote the index of the bin in which $\mathbf{u}_j(c_j)$ falls. Then b_1 is transmitted on Channel 1, b_2 is transmitted on Channel 2 and b_0 is split arbitrarily among the two channels. Thus the channel rates are given by $R_1 \geq \tilde{R}_1$, $R_2 \geq \tilde{R}_2$ and $R_1 + R_2 = \tilde{R}_1 + \tilde{R}_2 + \tilde{R}_0$.

Decoding is performed as follows. Say only Channel i $(i = 1,2)$ is received. The decoder then finds a codeword $\mathbf{u}_i$ in bin b_i such that $(\mathbf{y}, \mathbf{u}_i)$ are jointly typical. This can be uniquely done if

$$R'_i - \tilde{R}_i \leq I(U_i;Y), \quad i = 1,2 \qquad (3)$$

When both channels are received, the decoder additionally finds a codeword $\mathbf{u}_0$ in b_0 such that $(\mathbf{y}, \mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2)$ are jointly typical. This can be uniquely done if

$$R'_0 - \tilde{R}_0 \leq I(U_0;Y|U_1,U_2) \qquad (4)$$

Combining (2),(3),(4) results in the required rate region.□

3.2 Achievable Region for Gaussian Sources

The bounds presented above can be generalized to the special case where (X,Y) are jointly Gaussian random variables, using the approach presented in [11].

Theorem 3.2 *Let (X,Y) be zero-mean jointly Gaussian random variables, such that $X \sim \mathcal{N}(0,\sigma_x^2)$, and $Y = \alpha(X+N)$ w.l.o.g. where $N \sim \mathcal{N}(0,\sigma_n^2)$ is independent of X. An achievable distortion region for a given rate pair (R_1, R_2) is given by the union of (D_0, D_1, D_2) satisfying*

$$D_1 \geq \beta \cdot 2^{-2R_1}, \quad D_2 \geq \beta \cdot 2^{-2R_2} \qquad (5)$$

$$D_0 \geq \frac{\beta^3 \cdot e^{-2(R_1+R_2)}}{\beta^2\gamma^2 - 2D_1D_2 + \beta(D_1+D_2) + 2\sqrt{(D_1-\beta)(D_2-\beta)(D_1D_2-\beta^2\gamma^2)}} \qquad (6)$$

where $\beta = \frac{\sigma_x^2\sigma_n^2}{\sigma_x^2+\sigma_n^2}$, $\gamma^2 = e^{-2(R_1+R_2)}$, *$R_1, R_2$ are in bits, and the MSE distortion measure is used.*

Proceedings of the Data Compression Conference (DCC'03)
1068-0314/03 $17.00 © 2003 IEEE

Proof: We present a brief outline. Consider the encoder functions $U_1 = X + N_1$, $U_2 = X + N_2$ where N_1, N_2 are zero-mean Gaussian random variables, which are independent of X and have the following correlation matrix.

$$\Phi_{N_1N_2} = \begin{bmatrix} \sigma_{n_1}^2 & \rho\sigma_{n_1}\sigma_{n_2} \\ \rho\sigma_{n_1}\sigma_{n_2} & \sigma_{n_2}^2 \end{bmatrix}$$

Consider the decoder functions

$$\mathbf{g}_1(U_1, Y) = aU_1 + bY, \;\; \mathbf{g}_2(U_2, Y) = cU_2 + dY, \;\; \mathbf{g}_o(U_1, U_2, Y) = eU_1 + fU_2 + gY \tag{7}$$

where the the decoder reconstructions are the MMSE estimates of X, based on their respective arguments. Motivated by [12], we consider the R-D region characterized by (2) and (3). Then for a fixed tuple (R_1, R_2), the R-D expressions given by (2), (3) can be evaluated to yield (5). Further, for fixed (D_1, D_2), the central distortion is minimized by selecting $\rho = -\sqrt{\frac{D_1D_2-\beta^2\gamma^2}{D_1D_2}}$. Substituting this value in the expression for D_0, got from the rate region, leads to (6).□

We have not proved the converse for Theorem 3.2—thus the presented rate-distortion region is not proved to be optimal. However, (5),(6) do yield the *known optimal* R-D region in two special cases. Firstly, for Y independent of X, $\beta \to \sigma_x^2$ and the region in Theorem 3.2 yields the optimal two-channel Gaussian MD coding region, as proved in [12]. Secondly for $R_1 \to 0$ (or $R_2 \to 0$), Theorem 3.2 yields the optimal Wyner-Ziv coding region, as proved in [11].

4 R-D Performance for Predictive MD Coding

In this section we present results on the operational R-D performance for two-channel one-step predictive MD coding of a first-order Gauss-Markov source, in the WYZE-PMD framework. We will subsequently use the derived results to evaluate the performance of the codes proposed in Section 5.

Consider a first-order stationary Gauss-Markov source $X_n = \rho X_{n-1} + N_n$ where $X_i \sim \mathcal{N}(0, \sigma_x^2)$, $N_n \sim \mathcal{N}(0, (1-\rho^2)\sigma_x^2)$. We consider the following special case of the problem at hand. We consider the case of balanced two-channel predictive MD coding ($R_1 = R_2 = R$, $D_1 = D_2 = D$) with exactly one channel received at the decoder at each time instant. Note that *which* channel is received varies arbitrarily with time. We will also assume that the decoder performs one-step decoding—i.e. the decoder uses only the immediately previous reconstruction $\widehat{X}_{k-1}$ as side information.[1]

For the case of Gaussian side-information Y, considered in Section 3.2, solving for the side-decoder MMSE estimates in (7) yields:

$$\widehat{X_i} = \mathbf{g}_i(U_i, Y) = (1 - \frac{D}{\sigma_x^2}) \cdot X + \widehat{N}_i, \quad i = 1, 2 \tag{8}$$

[1]This mirrors common practice in many practical predictive coding applications, where the decoder buffers only the most recent source reconstruction.

where $\widehat{N}_i \sim \mathcal{N}(0, D - \frac{D^2}{\sigma_x^2})$ and are independent of X. Thus, for a stationary source X and fixed single-channel distortion D, the single-channel decoder reconstructions have the stationary distribution given by (8).

Thus, for the special case we are considering, the decoder side-information has a fixed distribution as above. The achievable operational R-D performance can now be determined by using the bounds derived in Section 3.2. In particular, the decoder side-information while decoding X_n is given by

$$Y = (1 - \frac{D}{\sigma_x^2}) \cdot X_{n-1} + \widehat{N} = (1 - \frac{D}{\sigma_x^2})\rho \cdot (X_n + \frac{\tilde{N}}{\rho} + \frac{\widehat{N}}{\rho(1 - \frac{D}{\sigma_x^2})})$$

where $\tilde{N} \sim \mathcal{N}(0, (1-\rho^2)\sigma_x^2)$, $\widehat{N} \sim \mathcal{N}(0, D - \frac{D^2}{\sigma_x^2})$, and $\tilde{N}, \widehat{N}$ are independent of X_n.[2]

The side information is thus $Y = \alpha(X_n + N)$ with $\alpha = (1 - \frac{D}{\sigma_x^2})\rho$ and $N \sim \mathcal{N}(0, \sigma_n^2)$ with $\sigma_n^2 = \frac{1-\rho^2}{\rho^2}\sigma_x^2 + \frac{D}{1-\frac{D}{\sigma_x^2}} \cdot \frac{1}{\rho^2}$. Inverting (5) gives the information rate required to achieve distortion D

$$R \geq \frac{1}{2}\log(\frac{1-\rho^2}{D}\sigma_x^2 + \rho^2)$$

Also of interest is the required information rate when the central distortion D_0 is additionally constrained. In this case the required information rate can be similarly obtained by inverting (6).

The scenario considered above provides a metric for evaluating the absolute performance of practical codes. We defer the analysis of the more general case, where the two channels have potentially independent failure probabilities that may vary with time, for future work.

5 Code Construction for WYZE-PMD

In this section we describe practical WYZE-PMD code constructions for the two-channel predictive MD coding problem. We will consider throughout, as an example, the M-dimensional first-order Gauss-Markov source $\mathbf{X}_n = \rho\mathbf{X}_{n-1} + \mathbf{N}$, $\mathbf{X}_n \in \mathbb{R}^m$, where $\mathbf{X}_n \sim \mathcal{N}(\underline{0}, \sigma_x^2 \mathrm{I}_m)$, $\mathbf{N} \sim \mathcal{N}(\underline{0}, (1-\rho^2)\sigma_x^2\mathrm{I}_m)$. We will consider predictive MD coding of this source over two erasure channels, with failure probabilities p_i $(i = 1, 2)$. No other assumptions on channel behaviour are made—in particular, the channel failure events are independent, and both channels may fail simultaneously at a given time instant.

For a given source vector $\mathbf{x}_n \in \mathbf{X}_n$, encoding is performed as follows. Multiple description scalar quantization [13] is used to generate two descriptions of the source vector. Fig. 1 shows a sample MDSQ index assignment with $2k + 1 = 3$ diagonals, and the associated scalar quantizers. The matrix entries represent the quantizer indices for the central scalar quantizer. The row and column labels represent the codebook indices for the first and second channel descriptions respectively. For a given codebook index c_i in Codebook i $(i = 1, 2)$, the reproduction level $u_i(c_i)$ is given by the centroid of the central quantizer intervals lying in row/column c_i in the index assignment matrix. Thus, the MDSQ code consists of the functions $\mathbf{m}_i : \mathbb{R}^m \to \mathcal{U}_i^m \;\; i = 1, 2$.

[2]Note that $X_n = \rho X_{n-1} + N_n \Rightarrow X_{n-1} = \rho X_n + \tilde{N}$, with $\tilde{N}$ independent of X_n.

c_1 \ c_2	1	2	3	4	5	6
1	1	2				
2	3	4	5			
3		6	7	8		
4			9	10	11	
5				12	13	14
6					15	16

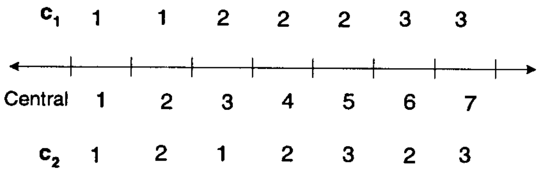

Figure 1: MDSQ index assignment with $k = 1$. $c_1(\cdot)$ is the Channel 1 codebook index and $c_2(\cdot)$ is the Channel 2 codebook index.

Next, the encoder partitions both codebooks into cosets, and transmits the index of the coset to which $\mathbf{m}_i(\mathbf{x}_n)$ belongs on Channel i. Approaching the information-theoretic bound requires the use of strong channel codes to partition the source code. It has been shown previously (cf. [14]) that the use of turbo codes yields performance close to the Wyner-Ziv bound. We briefly describe the procedure for generating the coset indices.[3]

The index vector $\mathbf{c}_i$ of the Channel i code-vector $\mathbf{u}_i(\mathbf{c}_i)$ is converted to a string of binary digits and is encoded using two systematic $\frac{k_n}{n_n}$ convolutional encoders. The second encoder is preceded by an interleaver which randomly interleaves the incoming bitstream. The resultant parity bitstreams, consisting of a total rate of $2(1-\frac{k_n}{n_n})m$ bits, are punctured so as to generate a reduced rate of mR_i bits, which represents the coset information for the channel code-vector. Thus the coset index functions $\mathbf{C}_i$ and the overall encoder functions $\mathbf{a}_i = \mathbf{C}_i(\mathbf{m}_i)$ are the following maps:

$$\mathbf{C}_i : \mathcal{U}_i^m \to \mathcal{I}_{2^{mR_i}}, \quad \mathbf{a}_i : \mathbb{R}^m \to \mathcal{I}_{2^{mR_i}}, \ i = 1, 2$$

Decoding is performed as follows. For each received channel i, the decoder uses the appropriate coset decoding function, $\widehat{\mathbf{C}}_i : \mathcal{I}_{2^{mR_i}} \times \mathbb{R}^m \to \mathcal{U}_i^m$, to reconstruct the channel code-vector from the the received coset information and the decoder side-information (i.e. the decoder reconstruction $\widehat{\mathbf{x}}_{n-1}$ which serves as a predictor). This involves iteratively decoding the channel code-vector from the received parity information, using the appropriate probability distributions induced on the coded symbols by the known side-information. If only one channel is received, the decoder then forms an MMSE estimate of the source-vector based on the decoded code-vector $\mathcal{U}_i^m$ and the side-information. When both channels are received, the decoder has unique knowledge of the central quantizer index from the decoded index vectors $(\mathbf{c}_1, \mathbf{c}_2)$. The decoder forms an MMSE estimate of the source-vector based on the central quantizer index and the side-information. If neither channel is received, the decoder's MMSE estimate of the source-vector is given by $\widehat{\mathbf{x}}_n = \rho \cdot \widehat{\mathbf{x}}_{n-1}$.

The key design issue in the above construction is the choice of the punctured transmission rate. The transmission rate required for the decoder to be able to correctly decode the received coset information is dependent on the closeness of the predictor $\widehat{\mathbf{x}}_{n-1}$ to the source-vector $\mathbf{x}_n$ (quantified, for example, by $I(\widehat{X}_{n-1}; X_n)$ for random variables, or the

[3] In the following description, we omit the optimizations that have been made to improve the performance of the turbo coder. Our aim in this discussion is to illustrate the basic principles of the code construction.

Euclidean norm for deterministic vectors). While encoding $\mathbf{x}_n$, the reconstruction set of decoder predictors $\mathcal{R}_{n-1}$ will, in general, contain $2^{2(n-1)}$ predictors corresponding to all possible channel behaviours at time instants $1, \ldots n-1$. Let $\mathcal{S}_i$ denote the set of predictors using which the decoder can correctly decode the Channel i code-vector from the received coset information. This set depends on the transmission rate R_i:

$$\mathcal{S}_i(R_i) = \{\widehat{\mathbf{x}}_{n-1} : \widehat{\mathbf{C}}_i(\mathbf{a}_i(\mathbf{x}_n), \widehat{\mathbf{x}}_{n-1}) = \mathbf{u}_i(\mathbf{c}_i),\ \mathbf{a}_i(\mathbf{x}_n) \in \mathcal{I}_{2^{mR_i}}\}$$

Then R_i should be selected such that the probability of decoder failure is negligible i.e. $P_{df} = P(\widehat{\mathbf{x}}_{n-1} \notin \mathcal{S}_i(R_i) | \widehat{\mathbf{x}}_{n-1} \in \mathcal{R}_{n-1}) < \epsilon$, $\epsilon > 0$. The choice of an appropriate R_i will, in general, depend upon ϵ, ρ and the channel failure probabilities p_i.

6 Results

Simulations were performed on a first-order Gauss-Markov source with CSNR = 17dB, with the two erasure channels having equal, independent failure probabilities $p_1 = p_2 = p$. The turbo coder consisted of two 16-state, rate $\frac{1}{2}$ systematic convolutional encoders. We compared the results of the proposed approach with those obtained using two alternative approaches.

The first alternative approach compared, was the low-latency MR-DPCM approach proposed in [1]. The MR-DPCM approach requires two predictors (one corresponding to each channel) to be maintained at the decoder. The encoder uses two independent DPCM loops to generate two predictively encoded sequences. As presented in [1], the low-latency MR-DPCM approach requires the same subset of channels to be received at all time instants. To take into account the independent channel failure probabilities, this approach was modified as follows. At each time instant the decoder reconstruction, used in computing distortion, corresponded to the received channels. When Channel j ($j = 1, 2$) was not received, the corresponding decoder predictor was estimated as $\widehat{\mathbf{x}}_k^j = \rho \widehat{\mathbf{x}}_{k-1}^j$. The second alternative approach considered was the independent coding approach, in which each source-vector was coded independently of all others. This approach completely avoids predictive mismatch, at the cost of lower compression efficiency.

Fig. 2(a) compares the operational D-R characteristics of the three approaches with channel failure probability $p = 0.1$. Fig. 2(b) compares the performance of the three approaches for fixed channel rate $R \approx 3$ bits per source-word, and varying channel failure probability.[4] In both cases, the transmission rate was selected such that $P_{df} \sim O(10^{-6})$. For the proposed approach and the MR-DPCM approach, one independently coded symbol was inserted after every 50 predictively coded symbols.

As can be seen, the low-latency MR-DPCM approach, which does not take predictive mismatch into account, performs the worst of the three approaches. This is illustrative of the performance loss caused by predictive mismatch, and motivates the need for avoiding mismatch. The proposed approach out-performs the conventional approaches by $7-10$ dB over a wide range of channel rates and channel failure probabilities. This demonstrates the efficacy of the proposed approach in exploiting the source correlation for efficient compression, while avoiding mismatch.

[4]The channel rate for the proposed approach was slightly *lower,* with $R = 2.91$ bits/source-word.

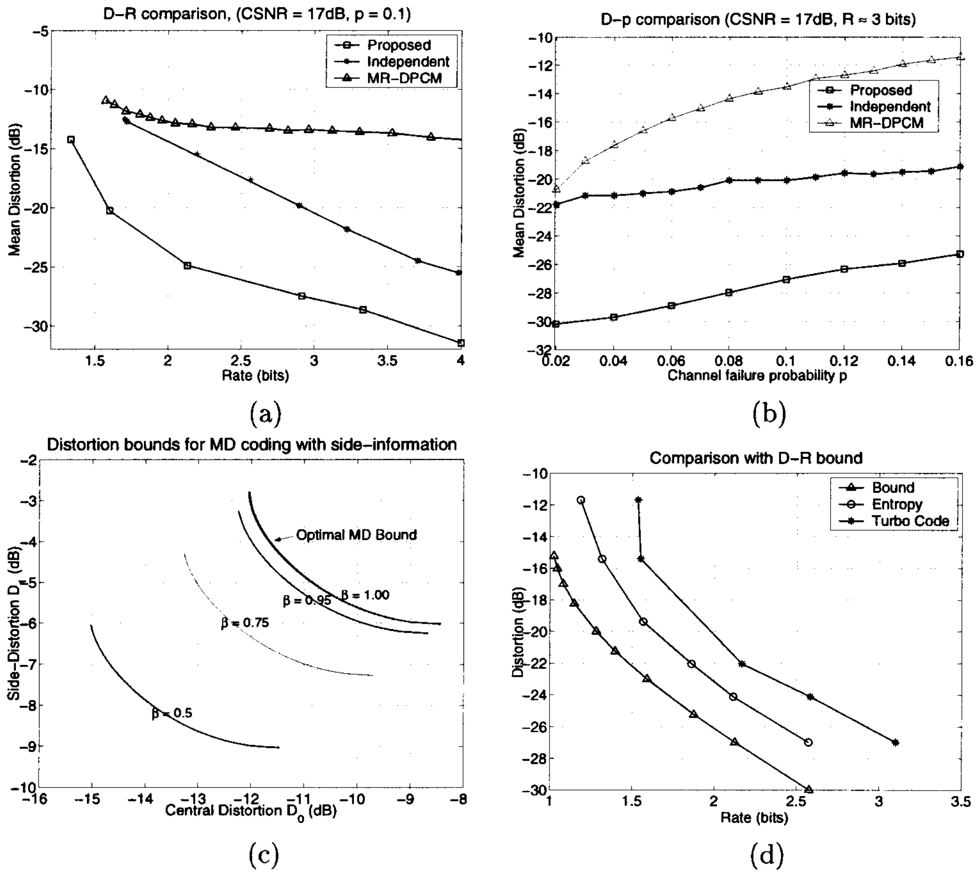

Figure 2: (a) Comparison of $D-R$ characteristics of the three approaches for channel failure probability $p_1 = p_2 = 0.1$. (b) Comparison of $D-p$ characteristics of the three approaches for channel rate $R_1 = R_2 = 3$ bits. (c) Achievable distortion bounds in Theorem 3.2 for $R_1 = R_2 = 1$ bit and $\sigma_x^2 = 1$. For $\beta = 1$, the bound is identical to the optimal distortion bound for conventional MD coding. (d) Comparison of $D-R$ performance of proposed approach with the achievable $D-R$ bound derived in Section 4. $\sigma_x^2 = 1$, $D_1 = D_2$, $D_0 = \frac{D_1}{10}$.

Fig. 2(c) plots the achievable distortion bound for two-channel Gaussian MD coding with side-information, as given by Theorem 3.2. The parameters for the plot are $R_1 = R_2 = 1$, $D_1 = D_2$ and $\sigma_x^2 = 1$. For $\beta = 1$, the side-information is independent of the source. As Fig. 2(c) shows, in this case the distortion bound becomes identical to the MD coding bound without side-information, derived in [12].

To evaluate the absolute performance of the proposed code, a set of simulations was performed under the restrictions specified in Section 4, with parameters $\sigma_x^2 = 1$, $D_1 = D_2$, $D_0 = \frac{D_1}{10}$. Fig. 2(d) compares the resulting performance of the proposed code with the achievable R-D bound (for constrained D_0) derived in Section 4. To help distinguish between the source coding loss and the channel coding loss, Fig. 2(d) also plots the channel code-vector entropy $H(\mathbf{u}_i|\widehat{\mathbf{x}}_{n-1})$. As can be seen the proposed approach incurs a loss of about 6 dB compared to the R-D bound. The key components of this loss are, (1) the source coding loss due to scalar MD quantization (1.5 dB for side reconstructions and 3 dB for the central reconstruction), and (2) the loss due to the non-Gaussianity of the channel code-vectors and the side-information, in practice. The use of dithered quantization in conjunction with higher-dimensional MD quantizers is expected to further improve the performance of the proposed approach.

References

[1] S. John and V.A. Vaishampayan, "Balanced interframe multiple description video compression," in *IEEE International Conference on Image Processing*, 1999, vol. 3, pp. 812–816.

[2] A.R. Reibman, H. Jafarkhani, Y. Wang, M.T. Orchard, and R. Puri, "Multiple-description video coding using motion-compensated temporal prediction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, pp. 193–204, 2002.

[3] A. Jagmohan and K. Ratakonda, "Multiple description coding of predictively encoded sequences," in *IEEE Data Compression Conference*, 2002, vol. 3, pp. 13–22.

[4] R. Puri and K. Ramchandran, "Multiple description source coding using forward error correction codes," in *Asilomar Conf. Sig. Sys. Computers*, 1999, pp. 342–346.

[5] A. Jagmohan, A. Sehgal, and N. Ahuja, "Predictive encoding using coset codes," in *IEEE International Conference on Image Processing*, 2002, vol. 2, pp. 29–32.

[6] A.D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 22, pp. 1–10, Jan. 1976.

[7] A. Jagmohan, A. Sehgal, and N. Ahuja, "Predictive multiple description coding using coset codes," To be presented at Int. Conf. Comm. Syst., Singapore, Nov. 2002.

[8] G.D. Forney Jr, "Coset codes. i. introduction and geometrical classification," *IEEE Transactions on Information Theory*, vol. 34, pp. 1123–1151, Sept. 1988.

[9] A.A.El Gamal and T.M. Cover, "Achievable rates for multiple descriptions," *IEEE Transactions on Information Theory*, vol. 28, pp. 851–857, Nov. 1982.

[10] C. Heegard and T. Berger, "Rate distortion when side information may be absent," *IEEE Transactions on Information Theory*, vol. 31, pp. 727–734, Nov. 1985.

[11] A.D. Wyner, "The rate-distortion function for source coding with side information at the decoder-ii: General sources," *Information and Control*, vol. 38, pp. 60–80, 1978.

[12] L. Ozarow, "On a source-coding problem with two channels and three receivers," *Bell Syst. Tech. Journ.*, vol. 59, pp. 1909–1921, Dec. 1980.

[13] V.A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 39, pp. 821–834, May 1993.

[14] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *IEEE Data Compression Conference*, 2002, pp. 252–261.

Proceedings of the Data Compression Conference (DCC'03)
1068-0314/03 $17.00 © 2003 IEEE