PREDICTIVE MULTIPLE DESCRIPTION CODING USING COSET CODES

A. Jagmohan, A. Sehgal, N. Ahuja

University of Illinois, Urbana
{jagmohan, asehgal, n-ahuja}@uiuc.edu

ABSTRACT

Multiple Description (MD) coding of predictively coded sources is of practical interest in several multimedia applications such as redundant storage of video/audio data, and real-time video/audio telephony. A key problem associated with predictive MD coding is the occurrence of predictive mismatch. In the present paper, we pose the problem of predictive MD coding as a variant of the Wyner-Ziv decoder side-information problem. We propose an approach based on the use of coset codes for predictive MD coding, which avoids predictive mismatch without requiring restrictive channel assumptions or high latency. We specifically consider two-channel predictive MD coding of a first-order Gauss-Markov process. Results indicate that the proposed approach significantly out-performs alternative approaches in terms of rate-distortion performance.

1. INTRODUCTION

Multiple Description (MD) coding involves coding source information into multiple descriptions for transmission over multiple lossy channels, such that the quality of reconstruction degrades gracefully with the number of descriptions lost. Predictive encoding is a well-known source coding technique used for low-latency compression of sources with memory. MD coding of predictively encoded sequences (termed predictive MD coding) is of practical interest in 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. Predictive mismatch refers to a scenario in which there is a mismatch between the predictor symbols at the encoder and the decoder, leading to an erroneous reconstruction of the predicted symbol at 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 (since feedback is assumed absent).

For the two-channel case, previous approaches for predictive MD coding include techniques based on multiple description scalar quantization with independent channel predictors [1], based on coding residual-of-residuals [2] and based on combining side residuals efficiently [3]. The main shortcoming of these approaches is the strong assumptions required on channel behaviour to eliminate predictive mismatch. Other techniques, such as [4] and the video coding technique presented in [1], avoid these assumptions at the cost of higher delay, thereby sacrificing the real-time nature of predictive coding.

In the present paper, we propose a technique for MD coding of predictively encoded sequences which eliminates predictive mismatch without requiring the restrictive assumptions made by previous approaches, and without requiring increased delay compared to conventional predictive coding. The key underlying concept is that predictive MD coding belongs to the class of predictive coding problems [5] which can be posed as variants of the well-known Wyner-Ziv decoder side-information problem [6]. Accordingly the proposed approach is based on the use of coset code constructions, which have been shown to achieve performance close to the information-theoretic bound for the Wyner-Ziv problem [7]. To illustrate the efficacy of the proposed approach, we consider predictive MD coding of a first-order Gauss-Markov process for transmission over two binary erasure channels using multiple scalar quantizers. The proposed approach significantly outperforms conventional approaches in terms of rate-distortion (R-D) performance, and can be extended to higher dimensional coset codes, such as lattice and trellis codes.

2. PROBLEM DESCRIPTION

In this section, we describe in detail the problem of predictive mismatch in the context of predictive MD coding. We then show how the problem under consideration can be formulated as a variant of the Wyner-Ziv side information problem.

2.1. Predictive Mismatch in Predictive MD Coding

Consider the communication of a M-dimensional source with memory, $\mathrm{S} = \{V_i\}_{i=1}^{\infty}, V_i \in \mathbb{R}^M$, across a lossy channel using one-step predictive coding. Given the decoder reconstruction of source symbol V_{k-1} (denoted $\widehat{V}_{k-1}$)

0-7803-7510-6/02/$17.00 ©2002 IEEE ICCS 2002

the encoder communicates V_k by generating the innovation $T_k = V_k - E[V_k|\widehat{V}_{k-1}]$ which is input to the channel, where $E[\cdot]$ represents the expectation operator. Denoting the channel output as $\widetilde{T}_k$, the decoder reconstructs V_k as $\widehat{V}_k = \widetilde{T}_k + E[V_k|\widehat{V}_{k-1}]$, such that for a given distortion metric $d(\cdot,\cdot)$, $d(V_k, \widehat{V}_k) = d(T_k, \widetilde{T}_k)$. Now, consider the case where the encoder does not have precise knowledge of $\widehat{V}_{k-1}$, and predictively encodes V_k using an incorrect predictor $\widetilde{V}_{k-1} \neq \widehat{V}_{k-1}$. Then the encoder transmits $\tau_k = V_k - E[V_k|\widetilde{V}_{k-1}]$ and the decoder reconstructs V_k as $\overline{V}_k = \widetilde{\tau}_k + E[V_k|\widehat{V}_{k-1}]$. This leads to predictive mismatch, which is manifested in increased decoder reconstruction distortion since $d(V_k, \overline{V}_k) > d(\tau_k, \widetilde{\tau}_k)$.

For the case of N-channel MD coding, the decoder reconstruction $\widehat{V}_{k-1}$ can take one of multiple values depending on the subset of descriptions of V_{k-1} received i.e. $\widehat{V}_{k-1} \in \mathbf{R}_{k-1}$ where $\mathbf{R}_i = \{\widehat{V}_i^j\}, j = 1 \ldots |\mathbf{R}_i|$ denotes the reconstruction set for the i^{th} symbol. Thus, assuming that the first source symbol V_1 is transmitted without predictive coding, $\mathbf{R}_1 = \{\widehat{V}_1^s; s \in \{0,1\}^N\}$, where s denotes the set of received descriptions ($s_i = 1$ if the i^{th} description is received) and $|\mathbf{R}_1| = 2^N$. In the absence of feedback (as is the case in real-time data transmission), the encoder does not know the set of descriptions $s_{dec} \in \{0,1\}^N$ received by the decoder, and hence does not know the predictor $\widehat{V}_1^{s_{dec}}$ to use for encoding V_2. The use of an incorrect predictor $\widehat{V}_{k-1}^{\bar{s}_{dec}} \in \mathbf{R}_1$ leads to predictive mismatch.

A naive solution, which eliminates predictive mismatch and ensures fidelity upto channel loss, is to encode V_k separately with respect to each possible predictor $\widehat{V}_1^s \in \mathbf{R}_1$, and transmit multiple descriptions of all resulting $|\mathbf{R}_1|$ innovations. However, the number of possible predictors grows exponentially with time and, in general, $|\mathbf{R}_k| = 2^{kN}$. Thus, the naive solution requires a large increase in rate of transmission over time, and is inefficient in a rate-distortion sense. Previous approaches [1, 2, 3] make the assumption that the subset of channels received over time stays the same. This assumption eliminates the exponential growth of $|\mathbf{R}_k|$, facilitating easier solutions to the problem of predictive mismatch. Such an assumption, however, is rarely satisfied in applications involving video/audio transmission over lossy networks. Other approaches (the video coding approach in [1], [4]) require a substantial increase in coding delay to circumvent this. These sacrifice low-latency, which is one of the key advantages of predictive coding, and are unsuitable for real-time applications such as video-conferencing and Internet telephony.

2.2. Predictive MD coding as the Wyner-Ziv problem

Consider two continuous-valued correlated Gaussian random variables $X, Y \in \mathbb{R}$. Consider the problem of compressing X with the correlated side information Y present only at the decoder. In [6], Wyner and Ziv showed that, if the encoder has knowledge of the statistics of Y, the R-D function for this problem is the same as in the case where the encoder has precise knowledge of Y [1] i.e.

$$\begin{aligned} R_{WZ}(D) = R_{X|Y}(D) &= \tfrac{1}{2}\log\frac{2\pi e\sigma^2_{X|Y}}{D} \quad && 0 \le D \le \sigma^2_{X|Y} \\ &= 0 && D \ge \sigma^2_{X|Y} \end{aligned}$$

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

In the sequel, we show how this formulation can be leveraged to design a practical solution to the predictive MD coding problem, which attains the desirable attributes outlined in Section 1.

3. PROPOSED APPROACH

In this section, we describe an approach based on coset codes, which solves the problem defined in Section 2. We then consider, in detail, MD coding of a real-valued first order Gauss-Markov process over two binary erasure channels.

3.1. 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 proposed 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 V_k the encoder transmits the innovation description $T_k^i = \mathbf{a}_i(\mathbf{C}(\mathbf{Q}(V_k)))$ on the ith channel. Thus, the encoder quantizes V_k onto the lattice Λ and generates N multiple descriptions of the coset index c_k (=$\mathbf{C}(\mathbf{Q}(V_k))$) of the resultant lattice point.

[1] If X, Y are not Gaussian, $R_{WZ}(D) \neq R_{X|Y}(D)$. However, knowledge of the statistics of Y can still be used to compress X more efficiently.

The proposed decoder consists of the MD decoding functions $\mathbf{g}_j(\cdot) : J \rightarrow \Lambda|\Lambda', j \in \{0,1\}^N$ and the coset decoding function $\widehat{\mathbf{C}}(\cdot,\cdot) : \Lambda|\Lambda' \times \Lambda \rightarrow \Lambda$. 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 $\widehat{c}_k^{j_{dec}}$. In general, if some descriptions are not received, the reconstructed coset index will not be identical to the transmitted coset index. This is the MD coding distortion due to channel loss. The coset decoding function decodes $\widehat{c}_k^{j_{dec}}$ to the coset lattice point closest to the MMSE estimator of V_k based on the decoder predictor $\widehat{V}_{k-1}$ (which acts as side-information), i.e.

$$\begin{aligned}\widehat{V}_k &= \widehat{\mathbf{C}}(\widehat{c}_k^{j_{dec}}, \widehat{V}_{k-1}) \\ &= \arg \min_{\lambda \in \Lambda, \mathbf{C}(\lambda) = \widehat{c}_k^{j_{dec}}} \left\| \lambda - E[V_k | \widehat{V}_{k-1}] \right\| \quad (1)\end{aligned}$$

It can be shown that the proposed decoder can correctly decode V_k, with fidelity upto the channel loss in the transmission of $\{T_k^i\}$, if the following condition (similar to [7] is satisfied.

$$e_{j_{dec}}(\widehat{V}_{k-1}) = \left\| \mathbf{Q}(V_k) - E[V_k | \widehat{V}_{k-1}] \right\| + \min_{\lambda \in \Lambda, \mathbf{C}(\lambda) = \widehat{c}_k^{j_{dec}}} \|\lambda - \mathbf{Q}(V_k)\| \leq \frac{d_{\Lambda'}}{2} \quad (2)$$

where $d_{\Lambda'}$ is the minimum distance of the sublattice Λ'. The event in which eqn. (2) is violated is termed a decoding failure event, and can be considered equivalent to the occurrence of predictive mismatch. Thus, predictive mismatch can be avoided if the sublattice Λ' is chosen with large enough $d_{\Lambda'}$ such that the probability of decoding failure is negligibly small, i.e.

$$E_j[\sum_{i:\widehat{V}_{k-1}^i \in \mathbf{R}_{k-1}} q(i) \cdot P(e_j(\widehat{V}_{k-1}^i) \geq \frac{d_{\Lambda'}}{2})] \approx 0 \quad (3)$$

where $E_j[\cdot]$ represents an expectation taken over the subscript index.

The benefits of the proposed approach are three-fold. Firstly, no assumptions on the set $\mathbf{R}_{k-1}$ (and thus on channel behaviour) are required–as long as $d_{\Lambda'}$ satisfies eqn. (3), the probability of predictive mismatch is negligibly small. Secondly, irrespective of the number of possible predictors $|\mathbf{R}_{k-1}|$, the proposed approach requires the generation of multiple descriptions of only one innovation, leading to efficient coding. Thirdly, the proposed approach does not lead to any increase in the coding delay with respect to baseline predictive coding, and can thus be used for real-time coding applications.

3.2. First-order Gauss-Markov Process

Consider a source whose output comprises a real-valued first-order Gauss-Markov process

$$x_n = \rho x_{n-1} + w_n \quad x_j, w_j \in \mathbb{R}, x_1 \sim \mathcal{N}(0,1)$$

where w_j is a sample of an i.i.d Gaussian noise process with mean = 0 and variance = $1-\rho^2$ and is independent of x_j. We consider predictive MD coding of this source process over two binary erasure channels, each of which has a failure probability of p. We will not make the assumption that at least one channel is always received–we consider the failure events of the two channels to be completely independent.

The fine encoder lattice is the one-dimensional uniform scalar quantizer $\Lambda = \Delta\mathbb{Z}$. The sublattice is $\Lambda' = \xi\mathbb{Z}$, $\xi = K\Delta, K \in \mathbb{N}$. The sublattice thus induces a partition $\Lambda|\Lambda'$ consisting of $\frac{\xi}{\Delta}$ cosets of Λ'. Fig. (1) shows the $\Lambda|\Lambda'$ partition for the case where $\xi = 3\Delta$. For the encoder lattices described, $\mathbf{Q}(x_n) = [x_n]_\Delta$ and $\mathbf{C}(\mathbf{Q}(x_n)) = [x_n]_\Delta \bmod \frac{\xi}{\Delta}$, where $[\cdot]_\Delta$ represents quantization using a uniform scalar quantizer with step-size Δ. The minimum distance $d_{\Lambda'}$ of the encoder sublattice is equal to ξ.

Two descriptions of the coset index $c_n = \mathbf{C}(\mathbf{Q}(x_n))$ are generated by using the multiple description scalar quantizer (MDSQ) proposed in [9]. Fig. (2) shows a MDSQ index assignment matrix with $2k+1 = 3$ diagonals. The matrix entries represent the quantized source index to be communicated, while the row and column labels represent the transmitted Channel 1 and Channel 2 indices respectively. When both channels are received, the decoder can uniquely decode the quantized source index. When only $a_1(\cdot)$ or only $a_2(\cdot)$ is received, the decoder can decode the source index upto an uncertainty specified by the *spread* of the index assignment, $s = 2k^2 + 2k + 1$. In the proposed approach,

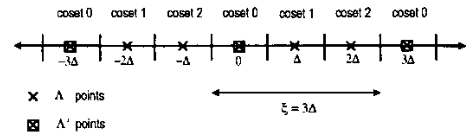

Fig. 1. Λ and Λ' lattices for the one-dimensional case. Λ' partitions the fine lattice into three cosets. $d_{\Lambda'} = \xi \triangleq 3$.

a_1 \ a_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

Fig. 2. MDSQ index assignment with $k = 1, s = 4$. $a_1(\cdot)$ is the Channel 1 index and $a_2(\cdot)$ is the Channel 2 index.

the matrix entry represents the coset index c_n which is to be communicated to the decoder (note that $c_n \in \mathbb{N}$). Thus, the row and column labels corresponding to c_n are transmitted over Channel 1 and Channel 2 respectively.

The decoder uses conventional MDSQ decoding to reconstruct the coset index–in the event of a channel loss, the coset index is reconstructed with corresponding distortion.[2] After reconstructing the coset index $\widehat{c}_n$, the decoder reconstructs x_n as

$$\widehat{x}_n = \widehat{\mathbf{C}}(\widehat{c}_n, \widehat{x}_{n-1}) = \arg \min_{x \in \Delta\mathbb{Z}, \mathbf{C}(x) = \widehat{c}_n} \|x - \rho \widehat{x}_{n-1}\|$$

The value of ξ for encoding should be selected in accordance with eqn. (3), so as to make the probability of decoding failure negligibly small. While encoding x_n, the reconstruction set of decoder predictors $\mathbf{R}_{n-1}$ will, in general, consist of $2^{2(n-1)}$ predictors corresponding to all possible channel behaviours at time instants $1, \ldots n-1$. For the case of one-step predictive encoding $|\mathbf{R}_{n-1}|$ can be reduced to $2(n-1)$ as follows. Let $\mathcal{R}_{i,j}$ denote the event that exactly j channels ($j = 1, 2$) are received at time $n - i$ and no channels are received at times $n - i + 1, \ldots n - 1$. Then, it suffices to consider,

$$\mathbf{R}_{n-1} = \{\widehat{x}_{n-1}(\mathcal{R}_{i,j}) : i \in \{1, \ldots n-1\}, j \in \{1, 2\}\}$$

and $|\mathbf{R}_{n-1}| = 2(n-1)$. To estimate the value of ξ, we need to evaluate the probability of decoding failure in eqn. (3). For this, we need to consider the $4(n-1)$ events given by $\{\mathcal{R}_{i,j} : i \in \{1, \ldots n-1\}, j \in \{1, 2\}\}\{\mathcal{R}_{0,j} : j \in \{1, 2\}\}$. Denoting a decoding failure event as $\mathcal{E}_{df}$, the probability of decoding failure is

$$P_{df} = \sum_{i=1}^{n-1} P(\mathcal{E}_{df} | \{\mathcal{R}_{l,j} : l = i\}) \cdot P(\{\mathcal{R}_{l,j} : l = i\}) \quad (4)$$

$P(\{\mathcal{R}_{l,j} : l = i\})$ represents the probability of the event that $n - i$ was the last time instant at which a non-zero number of channels were received (not including the current time instant). This is easily computed in terms of the channel failure probabilities

$$P(\{\mathcal{R}_{l,j} : l = i\}) = (p^2)^{i-1}(1 - p^2) \qquad (5)$$

Also,

$$\begin{aligned} P(\mathcal{E}_{df} | \{\mathcal{R}_{l,j} : l = i\}) = \; & \tfrac{(1-p)^4}{1-p^2} P(\mathcal{E}_{df} | \mathcal{R}_{i,2} \cap \mathcal{R}_{0,2}) + \\ & \tfrac{2p(1-p)^3}{1-p^2} P(\mathcal{E}_{df} | \mathcal{R}_{i,2} \cap \mathcal{R}_{0,1}) + \\ & \tfrac{2p(1-p)^3}{1-p^2} P(\mathcal{E}_{df} | \mathcal{R}_{i,1} \cap \mathcal{R}_{0,2}) + \\ & \tfrac{4p^2(1-p)^2}{1-p^2} P(\mathcal{E}_{df} | \mathcal{R}_{i,1} \cap \mathcal{R}_{0,1}) (6) \end{aligned}$$

[2] In the event that both channels are lost, the decoder reconstruction is simply the MMSE estimate $\widehat{x}_n = \rho \widehat{x}_{n-1}$.

The individual terms on the right hand side in eqn. (6) can be bounded in terms of the complementary c.d.f. of x_n. Thus, for example, it can be shown that

$$P(\mathcal{E}_{df} | \mathcal{R}_{i,2} \cap \mathcal{R}_{n,2}) \leq 2Q\left(\frac{\xi - (1 + \rho^i)\Delta}{2\sqrt{(1 - \rho^2)\sum_{j=0}^{i-1} \rho^{2j}}}\right)$$

where $Q(\cdot)$ is the Marcum Q-function. The other terms in eqn. (6) can be similarly bounded. The parameter ξ can be selected using eqn. (4), (5) and (6) to ensure that the probability of decoding failure is negligibly small. ξ represents a trade-off between the transmission rate and the probability of decoding failure, with an increase in ξ resulting in an increase in both quantities. We shall see in Section 5 that relatively small values of ξ are adequate to ensure $P_{df} \sim O(10^{-5})$, thereby allowing efficient compression.

3.3. Codebook training

Better compression performance, in a R-D sense, can be achieved by allowing more general scalar quantizers than $\Delta\mathbb{Z}$. We use a training algorithm similar to the algorithm in [9] to generate locally optimal MD scalar quantizers for the proposed approach.

The sublattice minimum distance ξ can be calculated from eqn. (4). The probability distribution $p_\xi(x)$ of the resultant coset indices is given by [10]

$$p_\xi(x) = \sum_{s=-\infty}^{\infty} p(x + s\xi) . \chi_{\{0 \leq x \leq \xi\}} \qquad (7)$$

where $p(x) \sim \mathcal{N}(0, 1)$ and χ is the indicator function.

Since we are MD coding the coset indices, the distribution of data to be coded is given by $p_\xi(x)$ rather than $p(x)$. The reconstruction levels and quantizer partitions of the side and central decoders can then be estimated by iteratively minimizing the appropriate Lagrangian functional for the level-constrained case or the entropy-constrained case.

4. ASYMPTOTIC ANALYSIS

In this section we compute the asymptotic R-D characteristics of the proposed coding technique for the example Gauss-Markov process. In Section 5, we shall compare the asymptotic R-D performance of the proposed approach to that of alternative approaches.

In the following derivation, we make use of the asymptotic analysis for optimum entropy-constrained MDSQ presented in [11]. Assume that a uniform scalar quantizer with step-size Δ is used to encode a scalar source symbol prior to MDSQ, and variable-length codes are used to code the MDSQ Channel 1 and Channel 2 indices. Then, for the case of mean-square distortion, as channel transmission rate $R \to \infty$, [11] shows that the asymptotic central distortion $\overline{d}_0$, and side distortion $\overline{d}_1$ are given as

$$\bar{d}_0 = \frac{\Delta^2}{12}, \; \bar{d}_1 = \frac{2\alpha_k(k+1)^2}{(2k+1)}\Delta^2 \tag{8}$$

and the transmission rate on each channel is given by

$$R = \log_2\left(\frac{1}{(2k+1)\Delta}\right) + h(p) \tag{9}$$

where $2k+1$ is the number of diagonals in the MDSQ index assignment, $\alpha_k = \sum_{u=1}^{k} u^2$ and $h(p)$ is the differential entropy of the source symbol distribution.

For the proposed predictive MD coding approach, the channel rate is, therefore

$$R = \log_2\left(\frac{1}{(2k+1)\Delta}\right) + h(p_\xi) \tag{10}$$

where $h(p_\xi)$ is the differential entropy of the coset index distribution (eqn. (7)). The expected mean-square distortion is

$$D = (1-p)^2\bar{d}_0 + 2p(1-p)\bar{d}_1 + p^2\bar{d}_2 \tag{11}$$

where $\bar{d}_0$ and $\bar{d}_1$ are given by eqn. (8) and $\bar{d}_2$ is the expected distortion when both channels fail at the current time instant. Now,

$$\bar{d}_2 = \sum_{i=1}^{n-1} \bar{d}_{2/\{\mathcal{R}_{l,j}:l=i\}} P(\{\mathcal{R}_{l,j} : l = i\}) \tag{12}$$

$\bar{d}_{2/\{\mathcal{R}_{l,j}:l=i\}}$ is the expected distortion in the event that time $n-i$ was the last time instant when a non-zero number of channels was received. Then,

$$\begin{aligned}
\bar{d}_{2/\{\mathcal{R}_{l,j}:l=i\}} &= E[(x_n - \hat{x}_n)^2] \\
&= E[(x_n - \rho^i\hat{x}_{n-i})^2] \\
&= E[(x_n - \rho^i x_{n-i})^2] + \rho^{2i}E[(\hat{x}_{n-i} - x_{n-i})^2] \\
&\quad -2\rho^i E[(x_n - \rho^i x_{n-i})(\hat{x}_{n-i} - x_{n-i})] \\
&= E[(x_n - \rho^i x_{n-i})^2] + \rho^{2i}E[(\hat{x}_{n-i} - x_{n-i})^2] \\
&= (1-\rho^2)\sum_{j=0}^{i-1}\rho^{2j} + \rho^{2i}\frac{1-p}{1+p}\bar{d}_0 \\
&\quad +\rho^{2i}\frac{2p}{1+p}\bar{d}_1
\end{aligned} \tag{13}$$

where the third step follows since $(x_n - \rho^i x_{n-i})$ is mean zero noise which is independent of $(\hat{x}_{n-i} - x_{n-i})$, and the fourth step assumes $n \to \infty$. Substituting eqn. (13) in eqn. (12) and taking $n \to \infty$ gives

$$\bar{d}_2 = 1 + \frac{\rho^2 p^2}{1-\rho^2 p^2}\frac{1-p^2}{p^2}\left\{\frac{1-p}{1+p}\bar{d}_0 + \frac{2p}{1+p}\bar{d}_1 - 1\right\} \tag{14}$$

which, when substituted in eqn. (11) gives the expected asymptotic distortion of the proposed approach.

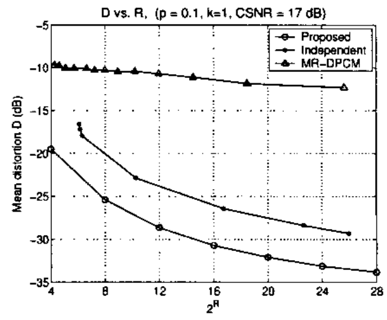

Fig. 3. Comparison of $D - R$ characteristics of the three approaches for channel failure probability $p = 0.1$. $\xi = 15\sigma_w$ for proposed approach.

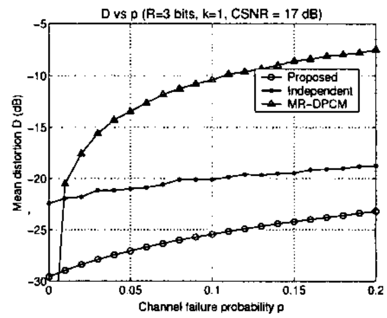

Fig. 4. Comparison of $D - p$ characteristics of the three approaches for channel rate $R = 3$ bits. $\xi = 15\sigma_w$ for proposed approach.

5. RESULTS

Simulations were performed on a first-order Gauss-Markov process with CSNR = 17dB, with two-channel MDSQ used for MD coding. 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.[3] 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 pre-

[3] A high-latency alternative which avoids this assumption is also presented in [1].

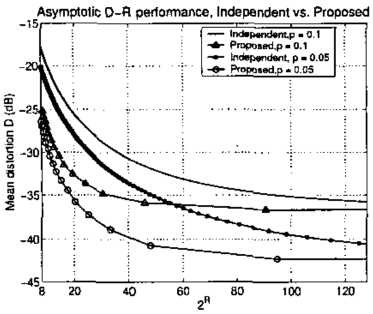

Fig. 5. Comparison of asymptotic D-R characteristics for the proposed scheme and for independent coding, $P_{df} \approx 10^{-5}$. The performance of the proposed scheme is strictly better, and approaches that of independent coding at very high rates.

dictor was estimated as $\widehat{x}_k^j = \rho \widehat{x}_{k-1}^j$. The second alternative approach considered was the independent coding approach, in which each source symbol was coded independently of all other symbols. This approach completely avoids predictive mismatch, at the cost of lower compression efficiency.

Fig. 3 compares the operational D-R characteristics of the three approaches with channel failure probability $p = 0.1$ and the MDSQ index assignment parameter $k = 1$. Fig. 4 compares the performance of the three approaches for fixed channel rate $R = 3$ bits and varying channel failure probability. The value of ξ used for Fig. 3–4 was $15\sigma_w$, for which $P_{df} \sim O(10^{-5})$. For both the proposed approach and the MR-DPCM approach, one independently coded symbol was inserted after every 100 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 independent coding approach by about 5 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 performance of the proposed technique can be further improved by making use of lattice and trellis codes [7]. The use of such codes also allows operation at much lower CSNR values, by reducing the probability of decoding failure, thereby making the proposed approach practical for a wider range of coding scenarios.

Fig. 5 compares the asymptotic D-R characteristics of the proposed approach with those of independent coding, for $P_{df} \approx 10^{-5}$. It is interesting to note that the performance of the proposed approach, while strictly better, approaches that of independent coding at very high rates. This is because, at very high rates, the distortion due to loss of both channels $(\overline{d}_2)$ dominates the MD coding distortion $(\overline{d}_0, \overline{d}_1)$, and is the same for both coding schemes. Note that the performance of the two schemes can also be expected to be similar at very low rates–in the limiting case when $R = 0$, $D_{prop} = D_{ind} = \sigma_x^2$.

[4] The mean distortion plotted, *includes* the effect of decoder failure.

6. CONCLUSIONS

We have proposed a real-time predictive MD coding technique that avoids predictive mismatch without requiring restrictive channel assumptions or high latency. Through the example of a first-order Gauss-Markov process, we have shown the potential performance gains that can be achieved by using the proposed approach. We are currently researching extensions of the proposed approach to higher-dimensional lattice and trellis codes which are expected to significantly improve the performance of the proposed approach.

7. 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," To be presented at ICIP, Sept. 2002.

[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] S.S. Pradhan, *Distributed source coding using syndromes (DISCUS)*, PhD dissertation, University of California at Berkeley, 2001.

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

[10] J. Kusuma, L. Doherty, and K. Ramchandran, "Distributed compression for sensor networks," in *IEEE International Conference on Image Processing*, 2001, vol. 1, pp. 82–85.

[11] V.A. Vaishampayan and J.-C. Batllo, "Asymptotic analysis of multiple description quantizers," *IEEE Transactions on Information Theory*, vol. 44, pp. 278–284, Jan. 1998.