Multiple Description Coding via Polyphase Transform and Selective Quantization

Wenqing Jiang, Antonio Ortega

Integrated Media Systems Center
Department of Electrical Engineering-Systems
University of Southern California
Los Angeles, CA 90089-2564
E-mail: [wqjiang, ortega]@sipi.usc.edu

ABSTRACT

In this paper, we present an efficient Multiple Description Coding (MDC) technique to achieve robust communication over unreliable channels such as a lossy packet network. We first model such unreliable channels as erasure channels and then we present a MDC system using polyphase transform and selective quantization to recover channel erasures. Different from previous MDC work, our system explicitly separates *description generation* and *redundancy addition* which greatly reduces the implementation complexity specially for systems with more than two descriptions. Our system also realizes a Balanced Multiple Description Coding (BMDC) framework which can generate descriptions of statistically equal rate and importance. This property is well matched to communication systems with no priority mechanisms for data delivery, such as today's Internet.

We then study, for a given total coding rate, the problem of optimal bit allocation between source coding and redundancy coding to achieve the minimum average distortion for different channel failure rates. With high resolution quantization assumption, we give optimal redundancy bit rate allocations for both scalar i.i.d sources and vector i.i.d sources for independent channel failures. To evaluate the performance of our system, we provide an image coding application with two descriptions and our simulation results are better than the best MDC image coding results reported to date. We also provide image coding examples with 16 descriptions to illustrate the simplicity and effectiveness of our proposed MDC system.

Keywords Multiple Description Coding (MDC), Balanced Multiple Description Coding (BMDC), Polyphase Transform, Selective Quantization

1. INTRODUCTION

In this paper we present an efficient multiple description coding strategy for robust communication over unreliable channels. An example of such channels are the best-effort packet networks, such as today's Internet, where congestion, routing delay, and network heterogeneity can all contribute to packets missing at the destination. Consequently, large segments of the transmitted data may be lost or useless. This will severely degrade the received signal quality if the missing data is not recovered.

Existing techniques to recover the lost data or mitigate the loss impact include ARQ retransmission,[1] FEC using error-correcting codes[2] and receiver reconstructions using only the received data by exploiting the residual correlation in the encoded data.[3,4] For delay-constrained applications (e.g. real-time audio/video) or multicast applications (e.g. multiparty teleconference), ARQ would obviously not be an appropriate choice. For large bursts of bit erasures (packet missing/lost), error-correcting codes, such as block codes and convolutional codes, can not provide sufficient protection without excessive delay and computation. On the other hand, reconstruction using residual correlation is limited in performance by the amount of residual correlation and it can not be used for an i.i.d. memoryless source.

Recently, Multiple Description Coding (MDC) techniques have been shown to be effective to combat channel failures. Two examples are the Multiple Description Scalar Quantizer (MDSQ)[5–7] and the Multiple Description Transform Coding.[8,9] However, such robustness to channel errors is achieved at the expense of relatively complicated system design. For example, MDSQ requires careful index assignments while MDTC necessitates another correlating transform besides the conventional decorrelating transform. We believe that the increased system complexity can be partly attributed to the fact that both systems generate descriptions with redundancy *implicitly* carried over.

In this paper, we propose an efficient multiple description coding system with reduced complexity both of system design and implementation. In our proposed system, we *explicitly* separate the *description generation* and *redundancy addition*, an idea inspired by a recent technique used within Robust-Audio Tool (RAT),[10,11] and formalized as Signal processing-based FEC (SFEC).[12] Unlike these approaches, in our proposed system, description generation is accomplished using a polyphase transform and each of these polyphase components is coded independently at a source coding rate. Redundancy is then explicitly added to each description by coding other descriptions at a lower redundancy coding rate using selective quantization. In case of channel failures, the redundancy is used to reconstruct the lost descriptions. We will show that our proposed MDC system, although simple, can yield better coding results compared to previously reported systems. The main contributions of our work are *(i) to propose a multiple description coding system based on polyphase transform and selective quantization, and (ii) to study optimal bit allocations for i.i.d scalar sources and i.i.d vector sources under independent channel failure probabilities.*

The remainder of this paper is organized as follows. In the next section, we give an outline of our proposed MDC system together with definitions of some MDC terminologies. Section 3 presents the rate-distortion analysis of our MDC system for i.i.d scalar and vector random sources. MDC image coding results are shown in Section 4 together with system extensions and possible future work. Finally, we conclude our work in Section 5.

2. PRELIMINARIES

In this section, we first characterize the erasure channels we are going to study and give definitions of MDC terminologies. Then we present our proposed MDC system.

2.1. Erasure Channel Characterization

The erasure channel model which we are going to study in this work is a simplified one. The transmitted data is assumed either to be completely lost or received correctly. In essence, we do not consider errors introduced in the physical layer whose protection will be largely dependent on applications of error-correcting codes. Rather we are looking for an alternative data recovery technique for the data link layer. Naturally, to achieve overall better protection capabilities, our scheme can be used jointly with error-correcting codes such as these unequal protection schemes proposed by Davis, Danskin and Sherwood.[13,14,2] Specifically, the erasure channel has two properties: *(1) the receiver knows where erasures occur in the data* and *(2) erasures occur independently.*

To satisfy the first requirement, we can insert synchronization points in the bitstream or tag packets with sequence numbers. Whether the second assumption is appropriate depends on the actual channel itself. For a lossy packet network, if packets are routed through different routes from the source to the destination, then we can safely assume it is. If, however, all the packets are routed through the same route, it seems that a bursty loss model is more appropriate.[11] However, as observed by Bolot et al.,[11] packet losses do not show strong bursty effect unless the network is heavily loaded. The reason is that, although packets in a congested node will experience bursty loss, this is an aggregated loss process. Packets from many different sources are multiplexed together. If the network is reasonably loaded and routers adopt random drops, rather than priority schemes, the loss will not show a strong bursty property for a single source. We briefly mention here that, our proposed MDC system can be easily extended for bursty channel erasure models thus will be a part of our future work.

We study techniques to recover from these erasures. Erasure channel models have been studied for a long time, yet our specific model here has two differences. First, we are considering erasures of large bit segments rather than single bit erasures. Second, we are not considering exact recovery of the erasures. Rather we only want to reconstruct an approximate version of the signal lost due to erasures. We would like to build a system which degrades gracefully in a predictable way in the presence of channel erasures.

2.2. Multiple Description Coding Terminologies

Let us first introduce MDC definitions and then contrast MDC with two other coding techniques: Simulcast Coding (SC) and Layered Coding (LC) or scalable coding.[15,16] Given an information source $\mathcal{X}$ and distortion measure $d(\mathcal{X}, \bar{\mathcal{X}})$.

Two Description Coding[17] involves finding two rate distortion codes $\{C_i, i = 1, 2\}$ such that C_1 achieves the rate distortion pair (R_1, D_1), C_2 achieves the rate distortion pair (R_2, D_2) and (C_1, C_2) achieves rate distortion pair $(R_1 + R_2, D_0)$ with $D_0 < D_1$ and $D_0 < D_2$.

Multiple Description Coding (MDC) involves finding multiple rate distortion codes $\{C_i, i = 1, 2, \cdots, M\}$ such that C_i achieves the rate distortion pair (R_i, D_i), any combinations of more than one codes (total number $\sum_{k=2}^{M} \binom{M}{k}$) achieve smaller distortion (smaller than $min\{D_i\}, i = 1, 2, \cdots, M$), and $(C_1, C_2, \cdots, C_M)$ achieves the global minimum distortion D_0 (rate distortion code $(\sum_{k=1}^{M} R_k, D_0)$).

Whereas the the notation is somewhat cumbersome here, two important observations can be made:

- Each code C_i is independently decodable since it has its own pair of encoding-decoding functions (f_i, g_i).
- Each code C_i carries new information about the original source which indicates that the more codes used for reconstruction, the smaller the overall distortion one can achieve.

These are two features of MDC which make it clear distinct from SC and LC. In a LC system, the source is encoded into multiple bitstream layers $\{L_0, L_1, \cdots, L_M\}$, which correspond to the multiple descriptions $\{C_0, C_1, \cdots, C_M\}$ in a MDC system. However, layers are usually not independent to each other. Higher layers can only be encoded/decoded after lower layers have been encoded/decoded. For example, L_2 is encoded and decoded with the help of L_0 and L_1. Obviously, for perfect channels, LC systems can achieve higher rate-distortion gain compared to MDC systems and are thus more appropriate for communication networks which provide delivery with different priority levels. However, LC systems are more susceptible to channel errors due to this inter-layer dependency on non-priority networks, such as today's Internet.

In a SC system, the source is actually encoded into multiple bitstreams $C_0, C_1, \cdots, C_M$ which can be independently decoded to yield different reconstruction qualities. Normally, each code is specifically designed for a specific class of users. For example, C_0 is for users with the smallest bandwidth while C_M is for users with the largest bandwidth. However, codes do not usually complement each other in the sense that C_i only carries new information compared to codes upstreams $C_0, C_1, \cdots, C_{i-1}$. However for MDC systems each C_i carries new information about the original source. This shows that, for example, (C_0, C_1) will only give reconstruction quality equal to that can be achieved by C_1 in a SC system, yet better reconstruction than that using only C_0 or C_1 in a MDC system. Clearly, this shows that MDC has a better rate-distortion gain compared to that of SC techniques.

In summary, MDC is better than SC in the rate-distortion sense while it outperforms over LC in the channel robustness sense at least in the case when no priorities exist. Of interest in this work is one special type of MDC systems which we define as the BMDC system.

Balanced Multiple Description Coding (BMDC) A multiple description coding which can (1) generate descriptions with equal rates and (2) generate descriptions with equal importance.

We emphasize here that the *equality* is defined only in a statistical sense. For example, if the polyphase components of a random source are two Gaussian sources with equal variances, they will be treated as having same rate-distortion functions. They will have equal average rates if quantized by the same quantizer. The importance of a description (code C_i) is defined as its corresponding distortion D_i achieved at the rate R_i.

The significance of BMDC systems is that their characteristics match well those of a best-effort lossy packet network such as today's Internet. Equal rate means that packets can all have same size and thus make it simple for buffer management during packetization and depacketization. Equal importance indicates that any single packet loss will incur the same average amount of loss of total information to be conveyed. Consequently, random drop when congestion occurs will not lead to drastic changes in the quality of the received data. Similar work on such a balanced system has also been reported in the MDSQ design[5,6] and in the equal energy distribution algorithm for raw image transmission over lossy packet network.[18]

3. THE PROPOSED MDC SYSTEM

In this section, we first outline our proposed MDC system. Then we present performance analysis based on optimal bit allocation for i.i.d random sources.

3.1. System Outline

We propose a multiple description coding system framework shown in Fig.1. The input X is first decomposed into two subsources Y_1, Y_2 via a polyphase transform. Each of these two components is quantized independently by Q_1 and constitutes the primary part of information for its corresponding channel. For reconstruction of the other channel in case of loss, each channel also carries information about the other channel, a coarsely quantized version by Q_2. Then quantized output from Q_1 and Q_2 are multiplexed together for transmission. At the receiver, if data from both channels arrives, fine quantized data of both polyphase components is then used for reconstruction. If one channel data is lost, one fine quantized polyphase and one coarsely quantized polyphase component are used for reconstruction.

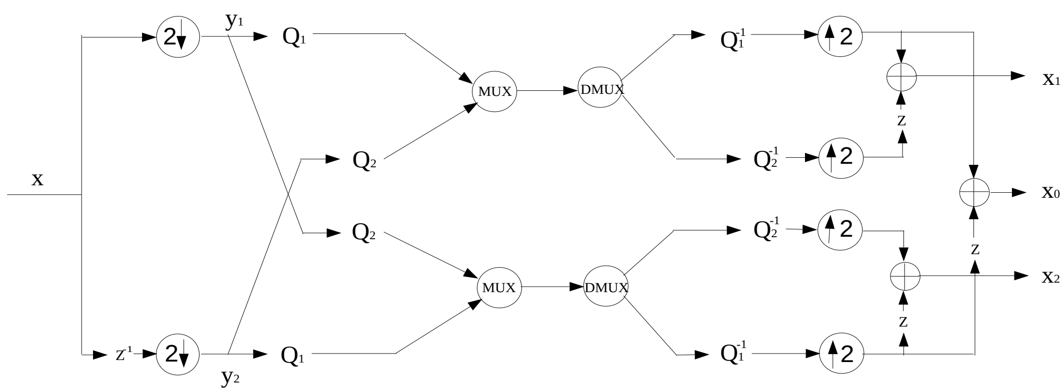

Figure 1. The proposed MDC system.

As one can see, the encoded bit stream explicitly separates the redundant information (coarsely quantized polyphase components) from the primary information (finely quantized polyphase components). It simplifies both the encoding and decoding processes compared to the MDTC system proposed by Wang et. al[8] where a correlating transform has to be used. It reduces the design complexity compared to the MDSQ system proposed by Vaishampayan[5] since both the polyphase transform and selective quantization can be implemented easily. Our system is similar to Jayant's Subsample-Interpolation (SI) approach.[3] However, SI implicitly assumes natural redundancy exists in the original data which does not hold for an i.i.d random source. The difference between our system and the RAT scheme proposed by Hardman et al.[10] is that descriptions in our system corresponds directly to the polyphase components rather than the original signal itself. The polyphase transform is equivalent to interleaving the data for transmission. It turns out that not only can we interleave the data in the spatial/time domain but also we can interleave the data in the frequency domain. This gives our system more flexibility to combat channel errors.

In Fig.2 we show one possible framework of our proposed MDC system operated on correlated input data. In the encoder, the input data, for example an image, is first transformed by a decorrelating transform (e.g. KLT). A polyphase transform is applied to the transform coefficients and each polyphase component forms the primary part of a single description. We call this the *description generation* stage. Next the *redundancy addition* block introduces redundancy among descriptions which identifies for each description what other descriptions it will protect. After this stage, each description is quantized and entropy-coded independently at a source coding rate and the redundant descriptions are coded at a lower redundancy coding bit rate. Quantized descriptions are then independently packetized and sent through the network.

Upon receiving data at the destination, the decoder checks the packet sequence numbers and identifies which descriptions are available and which descriptions are lost. It then decodes the available descriptions and recovers the lost descriptions with the help of the redundancy information. Description merging is used to reorganize the reconstructed descriptions back to the original signal representation. Finally an inverse transform is applied to obtain the reconstructed data.

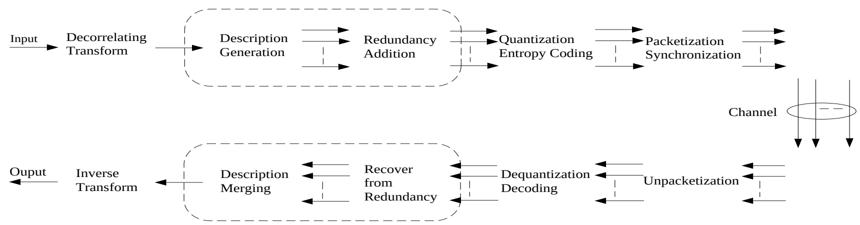

Figure 2. A complete MDC system for correlated input

3.2. Optimal Bit Allocation

We now give performance analysis of our proposed MDC system for i.i.d random sources. We answer the following question: for a given total bit rate, what is the optimal bit allocation between the source coding (primary information) and the channel coding (redundant information) to minimize both the central distortion (D_0) and side distortions ($D_i, i = 1, 2, \cdots, M$)? For the problem of two descriptions coding, this is equivalent to finding the operational optimal regions for the 5-tuple (R_1, R_2, D_0, D_1, D_2). We only study equal rate MDC systems with $R_i = R_0, i = 1, 2, \cdots, M$ in this work.

3.2.1. Scalar Sources

Assume X is an i.i.d zero mean random source with pdf $f(x)$ and variance σ^2. Under the high resolution quantization assumption, its rate-distortion function can be approximated $D_0 = h\sigma^2 2^{-2R_0}$ with h an integral constant defined as $h = \frac{1}{12}\left\{\int_{-\infty}^{\infty}[f(x)]^{1/3}dx\right\}^3$.[19] For Gaussian sources, $h = \sqrt{3}\pi/2$. Obviously, applying the polyphase transform on a memoryless source does not change the distortion-rate function since uniform subsampling basically does not change the characteristics of a random source. Therefore, the polyphase components Y_1, Y_2 will have same distortion-rate function as x with $D_1 = h\sigma^2 2^{-2R_1}$ and $D_2 = h\sigma^2 2^{-2R_2}$.

Let the bit rate for primary information be R_0 and the redundant bit rate be ρ, the side distortion (one description is lost) and effective bit rates are then

$$\begin{aligned} D_{s,i} &= \frac{1}{2}h\sigma^2 2^{-2\rho} + \frac{1}{2}h\sigma^2 2^{-2R_0} \\ R_{s,i} &= \frac{1}{2}(R_0 + \rho) \\ i &= 0, 1 \end{aligned}$$

The corresponding central distortion achieved is

$$D_c = h\sigma^2 2^{-2R_0}$$

at a total bit rate $R = R_0 + \rho$.

With this simple formulation, the optimal bit allocation between R_0 (primary information) and ρ (redundant information) then becomes a constrained optimization problem. Using a Lagrange multiplier, define the cost function J as

$$\begin{aligned} J &= D_c + \lambda D_s \\ &= h\sigma^2 2^{-2(R-\rho)} + \lambda(\frac{1}{2}h\sigma^2 2^{-2\rho} + \frac{1}{2}h\sigma^2 2^{-2(R-\rho)}) \end{aligned}$$

The optimal redundancy bit rate ρ^* can be analytically solved by having $\frac{\partial J}{\partial \rho}|_{\rho=\rho^*} = 0$ which leads to

$$\rho^* = \frac{1}{2}R + \frac{1}{4}\log_2(\frac{\lambda}{2+\lambda})$$

This optimal redundancy ρ^* has very intuitive explanations. Since $\frac{\lambda}{2+\lambda} \leq 1$, ρ^* is in the range of $[0, \frac{1}{2}R]$ for a given total bit rate R. If only side distortion counts (i.e. one description will be lost with very high probability), then the optimal redundancy rate is $\frac{1}{2}R$. This shows that both the primary information and the redundant information part are coded at the same rate and that two channels equally split the total bit rate, each of which is coded at half the total rate available. This is in fact the minimum possible achievable side distortion D_s for our system, however, at the expense of maximum possible central distortion D_c. If only central distortion count (i.e. both descriptions will arrive at the destination with very high probability), the redundancy ρ should always be set to zero. This means that each channel will carry half of the total information. Upon receiving data from both channels, one can get the minimum possible central distortion for the given bit rate R. However, the side distortion achieves its maximum. In-between these two extreme cases, one has the freedom to fine tune the side distortion with respect to the central distortion by choosing different redundancy bit allocations.

As a comparison, we consider MDC for a unit-variance zero mean memoryless Gaussian source. In our simulation, we first generate a sequence of Gaussian i.i.d samples. Then the even samples are quantized by a Lloyd-Max quantizer at a source coding rate R_0 and the odd samples are quantized by the same Lloyd-Max quantizer at a redundancy coding rate ρ. This is our description 1. The description 2 is formed in the same way except that odd samples are quantized at rate R_0 and even samples are quantized at rate ρ. The central distortion D_c is the MSE achieved at rate R_0 of the original source and the side distortion D_s is the average of the MSEs achieved using only description 1 or description 2. We use fixed-length codes for the index coding so the bit allocation is very simple with the only constraint of fixed total coding rate $R = R_0 + \rho = constant$. For example, if the total coding rate is $5bps$, the possible bit allocations are $(R_0, \rho) = (5, 0), (4, 1), (3, 2)$ at which we measure central distortions and side distortions.

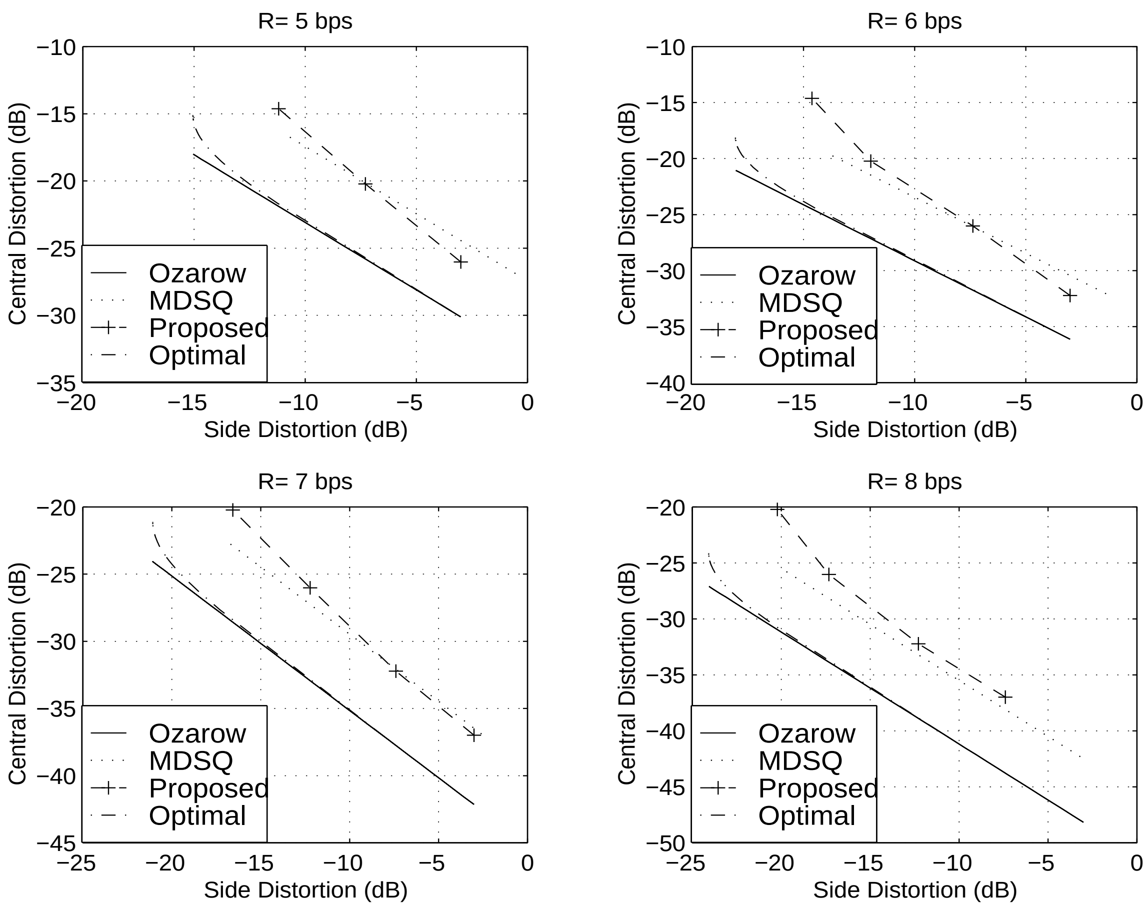

Figure 3. Rate-distortion performances comparison for a Gaussian source $\mathcal{N}(0, 1)$: (1) Ozarow: optimal bound.[20] (2) MDSQ: optimal level-constrained results.[6] (3) Proposed: Lloyd-Max quantizer results with fixed length code. (4) Optimal: results using the rate-distortion function of the Gaussian source.

The optimal lower bound of the achievable set of 5-tuple $(R_1, R_2, D_{s,1}, D_{s,2}, D_c)$ has been given by Ozarow[20] as

$D_{s,1} \geq 2^{-2R_1}$, $D_{s,2} \geq 2^{-2R_2}$, and $D_c \geq \frac{2^{-2(R_2+R_1)}}{1-(\sqrt{\Pi}-\sqrt{\Delta})^2}$ with $\Pi = (1-D_{s,1})(1-D_{s,2})$ and $\Delta = D_{s,1}D_{s,2} - 2^{-2(R_2+R_1)}$. The total bit rate is given by $R = R_1 + R_2$. We also compare with the asymptotic results given by Vaishampayan et al.[6] The optimal level-constrained MDSQ results are $D_c \approx \frac{1}{4}h2^{-R(1+a)}$ and $D_s \approx h2^{-R(1-a)}$ with $0 < a < 1$ and $h = \frac{\sqrt{3}\pi}{2}$. The comparisons are shown in Fig.3 under various bit rates.

As one can see, the results of our proposed system using a Lloyd-Max quantizer on an average is comparable to that of optimal level-constrained MDSQ. It is even better than MDSQ when total bit rates and redundancy rates are both very low. A big gap (about 7dB) still exists from Ozarow's optimal bounds which indicates that improvements can be made if more efficient quantizer can be designed. As a matter of fact, optimal entropy-constrained MDSQ has also been studied by Vaishampayan et al.[6] and significant gain has bees observed. The same is true for our proposed system.

An interesting question is that how well we can do using the proposed MDC system. In other words, for a Gaussian source, if we can design a quantizer which operates exactly on the rate distortion function, can we approach the Ozarow's MDC bounds? Using the optimal bit allocation derived before, the achievable central and side distortions are

$$\begin{aligned} D_c &= \sigma^2 2^{-R+\frac{1}{2}\log_2(\frac{\lambda}{2+\lambda})} \\ D_s &= \frac{\sigma^2}{2}2^{-R-\frac{1}{2}\log_2(\frac{\lambda}{2+\lambda})} + \frac{\sigma^2}{2}2^{-R+\frac{1}{2}\log_2(\frac{\lambda}{2+\lambda})} \end{aligned}$$

Now we plot this optimal results in the same figure (see Fig.3). As one can see, the performance gap narrows drastically and almost approaches the Ozarow's lower bounds at lower redundancy rates. This indicates that, while the proposed system can not achieve the lower bounds within the whole operational range, its performance can be greatly improved if we can design better quantizers. We mention that here the quantizer design is exactly the same as that for single description coding. Therefore we can make use of the state-of-art results from single description coding to reach our multiple description coding goals. As a result, the system design and implementation complexity is expected to be reduced compared to MDSQ systems. Our experimental results on image MDC will further illustrate this point.

Next we consider the channel model in the analysis. Assume that the two descriptions are sent over two different channels with independent channel failure probability p. There are four different situations at the receiver[21]: (a) both descriptions are received. This happens with probability $(1-p)^2$ and distortion $D_c = h\sigma^2 2^{-2R_0}$; (b) one description is lost. This happens with probability $2p(1-p)$ and distortion $D_s = \frac{1}{2}h\sigma^2 2^{-2\rho} + \frac{1}{2}h\sigma^2 2^{-2R_0}$; and (c) both descriptions are lost. This happens with probability p^2 and distortion $D_b = \sigma^2$. The average distortion at the receiver for this channel model is

$$\begin{aligned} D &= (1-p)^2 D_c + 2p(1-p)D_s + p^2 D_b \\ &= (1-p)^2 h\sigma^2 2^{-2R_0} + 2p(1-p)[\frac{1}{2}h\sigma^2 2^{-2\rho} + \frac{1}{2}h\sigma^2 2^{-2R_0}] + p^2\sigma^2 \end{aligned}$$

Since the reconstruction distortion in last case (c) will not be affected by the redundancy allocation, the optimal bit allocation only needs to minimize the first two terms in D. It can be found that, to minimize the average distortion in the presence of channel failures, the optimal bit allocation is

$$\rho^* = \frac{1}{2}R + \frac{1}{4}\log_2 p$$

The basic bit allocation still follows the half-rate splitting strategy. Clearly, the larger the channel failure probability, the more redundancy needs to be allocated to reduce the reconstruction distortion. If p approches 1 then the redundancy approaches $R/2$. This indicates that for very bad channels the maximum protection we can provide is to code the redundancy at half the total rate. If p approaches to 0, ρ may take negative values so we constrain the redundancy rate ρ to be in the range $[0, R/2]$. This indicates that, for nearly perfect channels, there is no need to carry redundancy at all. The threshold failure probability p_t can be calculated by setting ρ to 0 which yields $p_t = 2^{-2R}$. For example, if the total coding rate is $5bps$, then redundancy can be removed if the channel failure probability is smaller than 10^{-3}. On the other hand, it also indicates that no protection can be provided against channels with failure probabilities smaller than this threshold $p_t = 2^{-2R}$.

3.2.2. Vector Sources

We now study MDC techniques for i.i.d random vector sources, specifically random vectors with decorrelated or approximately decorrelated components, such as the output of a KLT/DCT/DWT transform in a transform coder.

Let $X = [x_0\ x_1\ \cdots\ x_{M-1}]$ be an M-dimensional vector with zero mean and with component variances $Ex_i^2 = \sigma_i^2$ for $i = 0, 1, \cdots, M-1$. We assume that $x_i, i = 0, 1, \cdots, M-1$ has the same normalized pdf. With high resolution quantization assumption, for given bit rate R, the minimum overall average distortion achieved is

$$D = h\rho^2 2^{-2\bar{b}}$$

where $\rho^2 = (\prod_{i=1}^{M} \sigma_i^2)^{\frac{1}{M}}$ and $\bar{b} = R/M$ is the average bit rate. As one can see, the optimal bit allocation for vector sources can be derived exactly in the same way as that for i.i.d scalar sources derived in the previous section. The difference is that ρ^2 replaces the σ^2 and bit rates are taken as average bit rates. Due to lack of space, we will skip the derivation details.

However, vector sources provide more flexibility for the packetization process compared to that for a scalar source. Two different approaches are shown in Fig.4. A time frequency plot is shown for an input. In (a), all the frequency components corresponding to one time/spatial location are packed into one packet while in (b) these frequency components are interleaved into two different packets. We will name this type of packetization technique as *Time-Frequency Packetization* which is achieved basically by polyphase transform both in time and frequency domains. Compared to (a), obviously (b) tells more about the time-frequency distribution of the input signal if one packet is lost. For example, in a JPEG-type block image coder, if each block goes into one packet, then one can interleave DCT coefficients from different blocks and pack them into different packets. Compared to a direct mapping between blocks and packets, this time-frequency packetization will avoid block holes in the received images in case of packet losses.

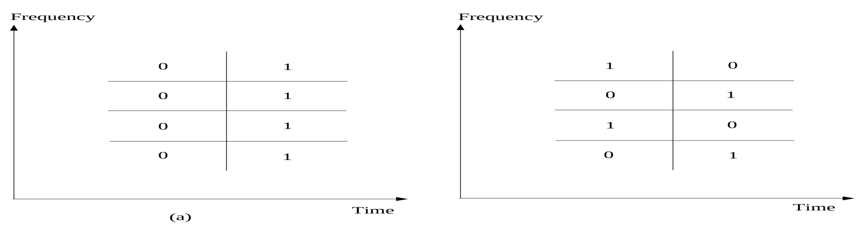

Figure 4. Two different approaches of packetization for vector sources. The blocks with $0s$ are frequency components packed into description 0 while blocks with $1s$ are frequency components packed into description 1. (a)Time packetization. (b)Time and frequency packetization.

4. EXPERIMENTAL RESULTS AND FUTURE WORK

We first present an image MDC example and then we show some possible future works.

4.1. An Image MDC Example

As shown by our analysis, the more efficient the quantization scheme, the better performance of our MDC system. Among these state-of-art wavelet coders,[22–25] we choose the Said-Pearlman wavelet coder due to its simplicity.

In our experiment, the input image is first wavelet transformed and its polyphase components are extracted. Two different types of polyphase transforms are tested on the wavelet coefficients. One is the plain polyphase transform which, for two descriptions coding, is simply group all the even coefficients into one description and all the odd coefficients into the other description. This is done for each row in each subband. The second is a vector form polyphase transform in which we group wavelet coefficients in different subbands corresponding to the same spatial location into a block structure like a block coder such in JPEG. This is simliar to the zerotree structure extensively exploited for coding efficiency in the image coding community.[25,24] All the even blocks then go into one description and all the odd blocks go into the other description.

Let the two polyphase components be y_1, y_2. Then $(y_1(R_0), y_2(\rho))$ constitutes our first description and $(y_2(R_0), y_1(\rho))$ the second description. The Said-Pearlman wavelet coder is used to quantize and entropy code the polyphase components. For example, $y_1(R_0)$ means that y_1 is coded at a bit rate R_0 with the Said-Pearlman coder. If one description is lost, reconstruct from the received data $((y_1(R_0), y_2(\rho))$ or $(y_2(R_0), y_1(\rho)))$ which gives the side distortion. The central distortion is derived using $(y_1(R_0), y_2(R_0))$. The total coding rate is $R_0 + \rho$.

Since the Said-Pearlman coder makes use of the zerotree structure among subbands, the second type of polyphase transform generates slightly better coding results. In Fig.5 we show MDC results for Lena gray-level image (size 512x512) using the Said-Pearlman wavelet coder.[25] The results of two descriptions are plotted and the comparison with a recent MDSQ-based MDC wavelet coder by Servetto et al.[7] is also given. With a fixed total coding rate, our MDC coder achieves better rate-distortion performance in the whole redundancy range.

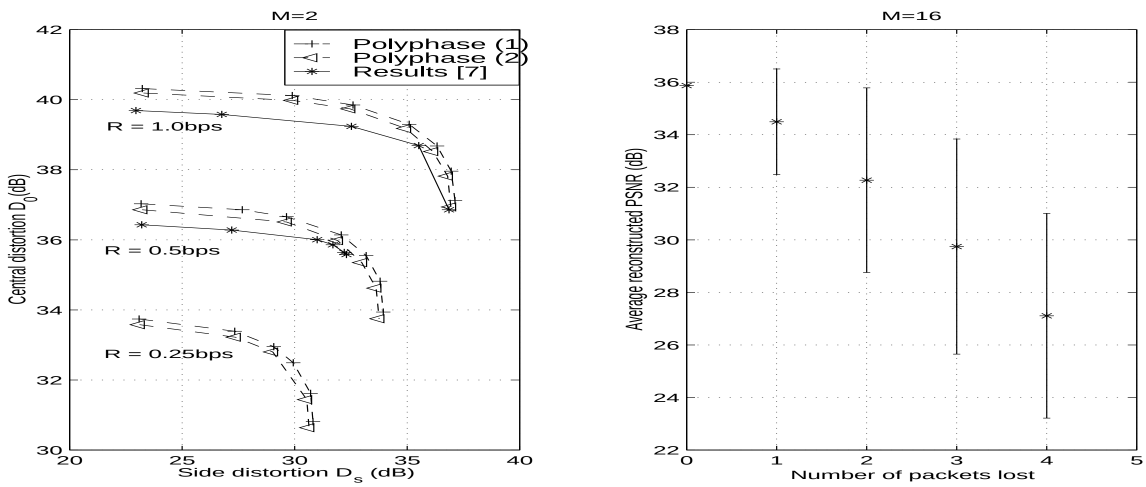

Figure 5. Experimental results with Lena gray-level image. (a) Two descriptions. Polyphase (1): plain polyphase transform. Polyphase (2): zerotree polyphase transform. (b) Performances with independent packet losses.

In the second experiment, we measure the average achieved PSNRs when there are independent packet losses. The input image is first wavelet transformed and then a polyphase transform (the zerotree vector form) is implemented on the wavelet coefficients. The downsampling factor is 16 so we end up with total 16 polyphase components. To protect from channel failures, the redundancy is carried in a sequential way like that used in RAT.[10] That is, packet 1 carries redundancy to protect packet 2 while packet 2 carries redundancy to protect 3 and so on.

Each polyphase component constitutes the primary part in each packet while it also carries redundancy to protect the next polyphase component in sequence. For example, packet 0 carries two parts of information:(1) polyphase component 0 coded at rate R_0 and polyphase component 1 coded at rate ρ. In this experiment, we fix the coding rates with $R_0 = 0.4bps$ and $\rho = 0.1bps$ thus a total coding rate $R = 0.5bps$. Since the total number of wavelet coefficients in each packet is the same, all 16 packets have the same size. We then measure the reconstruction error assuming independent packet losses. For example, assume there are 4 packets lost during the transmission, we first generate the loss pattern independently with 4 erasures. Let one loss pattern be 1 1 1 0 0 1 1 1 0 1 0 1 1 1 1 1 with $1s$ represent received packets while $0s$ lost packets. In this case, packet 4 can be reconstructed by using the redundancy carried by packet 3. The same is true for packet 9 and 11 while packet 5 will be lost without reconstruction. We tested 1000 loss patterns for each case when there are 1/2/3/4 packets lost.

In Fig.5 we show the image MDC results at total rate $0.5bps$ ($0.1bps$ redundancy rate). With different number of independent packets losses, the average PSNRs are plotted using star symbols with the standard deviations in vertical bars. As one can see, reconstructed PSNRs deviates from the mean values with an average standard deviation about 3dB. The quality changes are due to changes in the different loss patterns with consecutive losses lead to the worst reconstructions. However, as we explained before, our polyphase transform based MDC system should be also a BMDC system which means, at least, the reconstruction should be approximately the same for cases with one packet loss. However, the plot shows that, for one packet loss, the average PSNR is 34.5dB with standard deviation about 2dB. This is due to the fact that we assume that the first packet loss is not recovered in the experiment. If the first packet can also be recovered using the redundancy carried by the last packet, then the average PSNR becomes 34.99dB with standard deviation 0.12dB which becomes consistent to the basic concept of BMDC system. Obviously, our system can be easily modified to implement the equal energy packetization scheme as that by Ng et al.[26] for raw image transmission over lossy packet network.

4.2. Future Work

In Fig.6 we show one extended MDC system using context based coding technique. Since Y_1 and Y_2 are polyphase components of the original input, strong correlation or structural similarities existing among polyphase components, which can be used to further improve the quantization efficiency.[22,23] Thus the MDC system performance is expected to be improved.

One interesting problem is to study our system for bursty-error channel models. An example of this will be the mobile wireless channel where fast fading due to the movement of the object will cause rather long time bursty error over the link. We would like to know how well our system can do in this hostile communication environment.

Another interesting question is that whether MDC techniques are applicable to multicast applications. Traditionally MDC techniques have been used toward error-protection applications. However, we believe that MDC techniques, specifically the BMDC system proposed in this work, may also be applicable to multicast applications. The difficulty with multicast applications is that the coding technique has to meet a wide range of bandwidth requirements due to the network heterogeneity.[16] A typical coding technique for a multiparty teleconference is Layered Coding (or Scalable Coding). Although LC has better rate-distortion performance compared to BMDC, BMDC wins over LC in two aspects. First, BMDC is more robust compared to LC since each BMDC code can be independently decoded while LC maintains a strictly ordered sequence of layers. Second, for each coding layer, a separate multicast group has to be set up. This obviously increases the management expenses for both the intermediate routers and end nodes. On the other hand, if BMDC is used for a multicast session, only one multicast group needs to be set up and intermediate router can randomly drop incoming packets if it can not handle the traffic. The end node can also randomly drop the incoming packets when it can not handle. When it has available resources, it can take in as many packets as it can to improve the receiving signal quality.

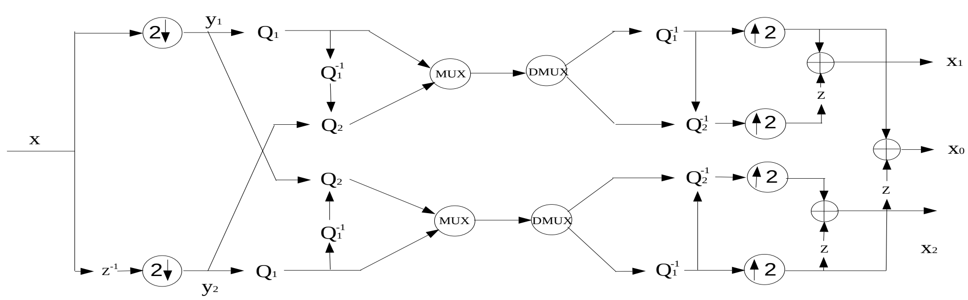

Figure 6. Context-based Quantization System

5. CONCLUSIONS AND ACKNOWLEDGMENTS

In this paper, a MDC system using polyphase transform and selective quantization has been proposed. We give detailed analysis of optimal bit allocation to achieve minimum average central distortion and side distortion for a fixed total coding rate. This is done both for i.i.d scalar random sources and vector random sources under independent channel failure models. Our experimental results have shown that our system implementation, compared to previous proposed systems, is simple yet the achieved MDC results for image coding is better specially at lower redundancy rates. It is also straightforward to generate multiple descriptions using our proposed system. We also give some possible system extensions and future work. The authors would like to thank Vinay Vaishampayan for bringing to attention the reference[6] and Sergio Servetto for providing the image MDC coding results.[7]

REFERENCES

1. C. Y. Hsu, A. Ortega, and M. Khansari, "Rate control for robust video transmission over burst-error wireless channels," *IEEE Journal on Selected Areas in Communications, Special Issue on Multi-Media Network Radios*, to appear, 1998.
2. P. G. Sherwood and K. Zeger, "Joint source and channel coding for internet image transmission," in *Proc. of DCC'97*, 1997.

3. N. S. Jayant and S. W. Christensen, "Effects of packet losses in waveform coded speech and improvements due to an odd-even sample-interpolation procedure," *IEEE Trans. Communications*, vol. COM-29, no. 2, pp. 101–109, Feb. 1981.
4. G. Karlsson and M. Vetterli, "Subband coding of video for packet networks," *Optical Engineering*, vol. 27, no. 574–586, 1988.
5. V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Information Theory*, vol. 39, no. 3, pp. 821–834, 1993.
6. V. A. Vaishampayan and J.-C. Batllo, "Asymptotic analysis of multiple description quantizers," *IEEE Trans. on Information Theory*, vol. 44, no. 1, pp. 278–284, Jan. 1998.
7. S. D. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt, "Multiple description wavelet based image coding," in *ICIP'98*, 1998.
8. Y. Wang, M. Orchard, and A. R. Reibman, "Multiple description image coding for noisy channels by paring transform coefficients," in *Proc. IEEE 1997 First Workshop on Multimedia Signal Processing*, 1997.
9. V. K. Goyal, J. Kovacevic, R. Arean, and M. Vetterli, "Multiple description transform coding of images," in *ICIP'98*, 1998.
10. V. Hardman M. A. Sasse, M. Handley, and A. Watson, "Reliable audio for use over the internet," in *Proc. INET*, 1995.
11. J.-C. Bolot and A. V. Garcia, "The case for FEC-based error control for packet audio in the internet," in *to appear ACM Multimedia Systems*, 1997.
12. M. Podolsky, C. Romer, and S. McCanne, "Simulation of FEC-based error control for packet audio on the internet," in *INFOCOM'98*, San Francisco, CA, Mar. 1998.
13. G. Davis and J. Danskin, "Joint source and channel coding for internet image transmission," in *Proc. of SPIE Conf. on Wavelet Applications of Digital Image Processing XIX*, Aug. 1996.
14. J. M. Danskin, G. M. Davis, and X. Song, "Fast lossy internet image transmission," in *Proc. of 1995 ACM Multimedia Conference*, San Francisco, CA, 1995.
15. D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. on Image Processing*, vol. 3, no. 5, pp. 572–588, Sept. 1994.
16. S. R. McCanne, "Scalable compression and transmission of internet multicast video," Tech. Rep. Ph.D. Thesis, UCB/CSD-96-928, University of California, Berkeley, 1996.
17. A. A. El-Gamal and T. M. Cover, "Achievable rates for multiple descriptions," *IEEE Trans. Information Theory*, vol. IT-28, no. 6, pp. 851–857, Nov. 1982.
18. F. M. L. Ng and J. Kovacevic, "Raw image data transmission over lossy packet networks," *Submitted to IEEE Trans. on Image Processing*, Aug. 1997.
19. A. Gersho and R. M. Gray, *Vector quantization and signal compression*, Prentice Hall Information and System Sciences Series. Prentice Hall, Englewood Cliffs, NJ 07632, 1991.
20. L. Ozarow, "On a source-coding problem with two channels and three receivers," *The Bell System Technical Journal*, vol. 59, no. 10, pp. 1909–1921, Dec. 1980.
21. V. K. Goyal and J. Kovacevic, "Optimal multiple description transform coding of Gaussian vectors," in *Proc. of IEEE Data Compression Conference*, 1998.
22. C. Chrysafis and A. Ortega, "Efficient context-based entropy coding for lossy wavelet image compression," in *Proc. of DCC'97*, Snowbird, UT, Mar. 1997.
23. S. M. LoPresto, K. Ramchandran, and M. T. Orchard, "Wavelet image coding using rate-distortion optimized backward adaptive classification," in *Proc. of VCIP'97*, San Jose, CA, Jan. 1997.
24. X. Wang, G. Wu, and X. Lin, "A zerotree based multirate wavelet image coding method," *Acta Electronica Sinica*, vol. 25, no. 4, pp. 48–51, Apr. 1997.
25. A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 4, pp. 243–250, June 1996.
26. F.M.L. Ng and J. Kovacevic, "Raw image data transmission over lossy packet networks," *IEEE Trans. on Image Processing*, Aug. 1997, Submitted.