On Scalable Distributed Coding of Correlated Sources

Ankur Saxena, *Member, IEEE*, and Kenneth Rose, *Fellow, IEEE*

***Abstract*—This paper considers the problem of scalable distributed coding of correlated sources that are communicated to a central unit, a setting typically encountered in sensor networks. As communication channel conditions may vary with time, it is often desirable to guarantee a base layer of coarse information during channel fades, as well as robustness to channel link (or sensor) failures. The main contribution is twofold. First, we consider the special case of multistage distributed coding, and show that naive combination of distributed coding with multistage coding yields poor rate-distortion performance, due to underlying conflicts between the objectives of these two coding methods. An appropriate system paradigm is developed, which allows such tradeoffs to be explicitly controlled. Next, we consider the unconstrained scalable distributed coding problem. Although a standard "Lloyd-style" distributed coder design algorithm is easily generalized to encompass scalable coding, the algorithm performance is heavily dependent on initialization and will virtually always converge to a poor local minimum. We propose an effective initialization scheme for such a system, which employs a properly designed multistage distributed coder. We present iterative design techniques and derive the necessary conditions for optimality for both multistage and unconstrained scalable distributed coding systems. Simulation results show substantial gains for the proposed multistage distributed coding system over naive extensions which incorporate scalability in a multistage distributed coding system. Moreover, the proposed overall scalable distributed coder design consistently and substantially outperforms the randomly initialized "Lloyd-style" scalable distributed coder design.**

***Index Terms*—Distributed quantization, multistage coding, scalable coding, sensor networks.**

I. INTRODUCTION

DISTRIBUTED source coding (DSC) [1], [2] has witnessed a significant revival of interest since the late nineties, with a growing focus on practical code design. The work of Pradhan and Ramchandran [3] was a notable precursor. The field has eventually seen the emergence of various distributed coding[1] techniques, mostly with an eye towards sensor networks (see reviews in [4] and [5]). The basic setting in DSC involves multiple correlated sources (e.g., spatially distributed sensors) transmitting information to a fusion center, as shown in Fig. 1. The main objective in DSC is to exploit intersource correlations despite the fact that each sensor source is encoded without access to other sources. The only information available to a source encoder about the other sources is via joint statistics (typically extracted from training data).

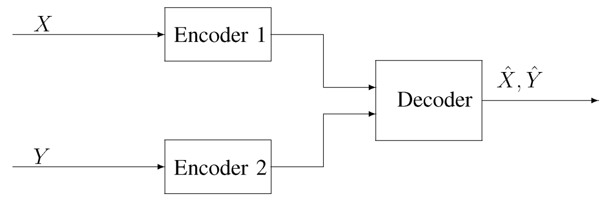


Fig. 1. Distributed coding of two correlated sources.

The communication channels in a sensor field may vary in capacity due to the presence of obstacles or other phenomena such as fading. In such a scenario, it will be beneficial to convey a minimal amount of information even when the channel deteriorates. This motivates the problem of scalable distributed source coding (S-DSC) or successive refinement of distributed correlated sources, which generalizes the traditional problem of scalable coding of a single source [6]–[9].[2] Successive refinement for Wyner–Ziv coding (side information at the decoder) was proposed in [10], and has been studied in [10] and [11] from the information-theoretic perspective of characterizing achievable rate-distortion regions. In this paper, we derive practical iterative algorithms for the design of scalable coding systems within the multiterminal (distributed) setting, i.e., for a S-DSC system. The general S-DSC problem subsumes several important special cases such as multiple-description coding [12], robust distributed coding [13], [14], etc.

Various scalability structures for S-DSC may be implemented, such as tree-structured quantizers or multistage quantizers [16]. In practice, multistage structures are often preferred due to the reduced complexity search (during encoding) and storage (of source codebooks). A leading application is speech coding where multistage vector quantizers are heavily used. In this work we first analyze system design for multistage distributed source coding (MS-DSC) [17], [18]. It may be

Manuscript received June 12, 2009; accepted January 13, 2010. Date of publication February 05, 2010; date of current version April 14, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Huaiyu Dai. The work was supported in part by the NSF by grants IIS-0329267 and CCF-0728986, the University of California MICRO program, Applied Signal Technology Inc., Cisco Systems Inc., Dolby Laboratories Inc., Qualcomm Inc., and by Sony Ericsson, Inc. The material in this paper was presented in part at the IEEE Data Compression Conference, Snowbird, UT, March 2008 and at the IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 2009.

The authors are with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA (e-mail: ankur@ece.ucsb.edu; rose@ece.ucsb.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2042488

[1]Note that the term "distributed coding" in this paper is always employed in the context of source coding and is used interchangeably with "distributed source coding."

[2]We use the term "scalable coding" in the standard source coding sense, namely, rate-scalable, i.e., the source is encoded into a layered bit-stream with a base and enhancement layer, thereby allowing a decoder reconstruction quality to scale with the available communication bit rate.

tempting to assume that simple integration of algorithms for distributed coding [19]–[22] and multistage quantizer design [16], would yield a good MS-DSC coding scheme. However, as we will see, there exists a fundamental tradeoff between exploiting intersource correlation at the base or intermediate layers, and better reconstruction in subsequent enhancement layers of the MS-DSC. Moreover, by allowing for a slight but controlled mismatch between encoder and decoder estimates and reconstructions, intersource correlation can be exploited more effectively.

Our second main contribution in this paper is to extend the proposed MS-DSC system as an efficient initial solution for the general S-DSC problem [23]. While there is no direct conflict between the objectives of distributed quantization and scalable coding in unconstrained S-DSC, and a Lloyd-style DSC design algorithm [19], [21] can be extended to S-DSC, the resulting algorithm depends heavily on initialization. We show how the proposed MS-DSC design can be used as an effective initialization for S-DSC.

It is also desirable that the S-DSC system be *robust* to fades or failures of various communication channels, and utilize all received information to attain maximal efficiency. We incorporate system robustness objectives by leveraging our techniques for robust distributed coder design [19].

As a side note, we should mention that there exists a significant channel coding "school" that adopts ideas from channel coding (see, e.g., [24] and [25]) and may exploit long delays to achieve good performance [26] and [27]. We will, however, restrict our attention to source coding and low delay methodologies in this work.

The rest of the paper is organized as follows. In Section II, we state the problem formally and introduce notation. In Section III, we specify the components of the MS-DSC system and consider a naive design approach that simply combines standard distributed coding and multistage quantizer design algorithms. Section IV describes the proposed MS-DSC algorithm. Section V presents the S-DSC system alongwith efficient initialization that avoids many poor solutions. Simulation results are summarized in Section VI, followed by conclusions in Section VII.

II. Problem Statement

Consider the S-DSC scenario in Fig. 2. For brevity, we will restrict the analysis to the case of two sources and two-layers, but without loss of generality as the model is trivially extendible to an arbitrary number of sources or layers. Here (X, Y) are two continuous amplitude, i.i.d. (scalar or vector) sources which are mutually correlated. The encoder $\mathcal{E}_x$ compresses source X and transmits an index pair (i_1, i_2) where $i_1 \in \{1, \ldots, 2^{R_{1x}}\}$ and $i_2 \in \{1, \ldots, 2^{R_{2x}}\}$. Similarly, the Y encoder $\mathcal{E}_y$ produces index pair (j_1, j_2) where $j_1 \in \{1, \ldots, 2^{R_{1y}}\}$ and $j_2 \in \{1, \ldots, 2^{R_{2y}}\}$. Here R_{1x} and R_{1y} denote the first (base) layer rates while R_{2x} and R_{2y} denote the incremental second (enhancement) layer rates.

We assume that the fusion center reliably receives full information from the base layer, while enhancement layer information from sources X or Y may be lost independently with probability p_x and p_y, respectively. Depending on the subset of information received at the enhancement layer, the fusion center uses decoder $\mathcal{D}_{00}$, $\mathcal{D}_{01}$, $\mathcal{D}_{10}$ or $\mathcal{D}_{11}$ to reconstruct X as $\hat{X}_{00}$, $\hat{X}_{01}$, $\hat{X}_{10}$, or $\hat{X}_{11}$, respectively, and similarly for Y (see Fig. 2). The decoder subscripts indicate whether the enhancement layer information for source X and Y have been received, e.g., decoder $\mathcal{D}_{10}$ is used when only enhancement layer information from X is received. Thus, the decoders $\mathcal{D}_{00}$, $\mathcal{D}_{01}$, $\mathcal{D}_{10}$, and $\mathcal{D}_{11}$ are used with probabilities $p_{00} = p_x p_y$, $p_{01} = p_x(1 - p_y)$, $p_{10} = (1 - p_x)p_y$, and $p_{11} = (1 - p_x)(1 - p_y)$, respectively.

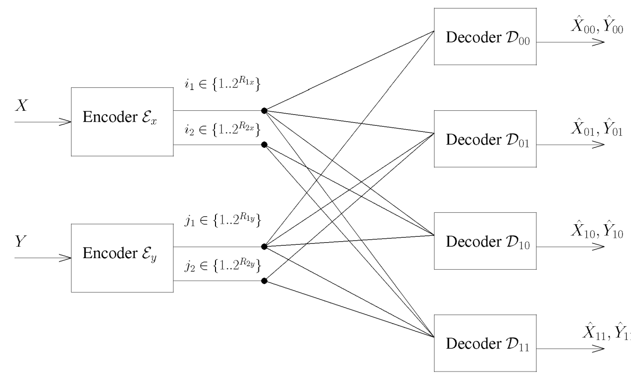


Fig. 2. Scalable distributed source coding.

The distortion incurred when decoder $\mathcal{D}_{00}$ is used is

$$E[\alpha d(X, \hat{X}_{00}) + (1 - \alpha) d(Y, \hat{Y}_{00})] \tag{1}$$

where $d(\cdot, \cdot)$ is an appropriately defined distortion measure and $\alpha \in [0, 1]$ is a weighting factor to govern the relative importance of sources X and Y at the fusion center. Distortion terms for decoders $\mathcal{D}_{01}$, $\mathcal{D}_{10}$ or $\mathcal{D}_{11}$ are similarly defined. Note that we use uppercase letters for random variables and lowercase letters to denote their particular realizations. We use the following streamlined notation to denote the distortion incurred by the different decoders for a data point (x,y)

$$\begin{aligned} D_{00}(x, y) &= \alpha d(x, \hat{x}_{00}(i_1, j_1)) \\ &\quad + (1 - \alpha) d(y, \hat{y}_{00}(i_1, j_1)) \\ D_{01}(x, y) &= \alpha d(x, \hat{x}_{01}(i_1, j_1, j_2)) \\ &\quad + (1 - \alpha) d(y, \hat{y}_{01}(i_1, j_1, j_2)) \\ D_{10}(x, y) &= \alpha d(x, \hat{x}_{10}(i_1, j_1, i_2)) \\ &\quad + (1 - \alpha) d(y, \hat{y}_{10}(i_1, j_1, i_2)), \\ D_{11}(x, y) &= \alpha d(x, \hat{x}_{11}(i_1, j_1, i_2, j_2)) \\ &\quad + (1 - \alpha) d(y, \hat{y}_{11}(i_1, j_1, i_2, j_2)) \end{aligned} \tag{2}$$

where the index pairs are determined by the source values, $\mathcal{E}_x(x) = (i_1, i_2)$ and $\mathcal{E}_y(y) = (j_1, j_2)$. We assume for simplicity that the same weighting factor α is used by all the different decoders. This simplifying assumption can easily be removed if desired.

Next we define the net distortion incurred for a source pair $\{x, y\}$, i.e., the distortion averaged over all the channel loss scenarios, as

$$\begin{aligned} D_{\text{net}}(x, y) = p_{00} D_{00}(x, y) + p_{01} D_{01}(x, y) \\ + p_{10} D_{10}(x, y) + p_{11} D_{11}(x, y). \end{aligned} \tag{3}$$

The S-DSC design objective is to minimize the following distortion cost given rate allocations R_{1x}, R_{2x}, R_{1y}, and R_{2y}; and enhancement layer loss probabilities p_x and p_y:

$$E[D_{\text{net}}(X,Y)] \tag{4}$$

where the expectation is over the source statistics.

Also in this paper, we assume that the base layer information is always received at the decoder. If and when, the decoder receives extra enhancement layer information, it refines the estimate and obtains a better reconstruction. In general, this extra enhancement layer information could be transmitted by the sources only when required if we need to consider the overall energy efficiency of the system. Here we do not attempt to address the question of whether or not to transmit the enhancement layer information and consider the S-DSC problem in similar vein to the classical source coding, where enhancement layer information is used whenever available.

A. *Special Cases*

The S-DSC problem is quite general and subsumes a large number of source coding problems as special cases. First the trivial special cases:

1) *Single Source Vector Quantizer*: If no information is transmitted for source Y (both base and enhancement layer rates for Y are zero) and source X transmits only base layer information, the problem reduces itself to that of a single source vector quantizer.
2) *Distributed Source Coding*: In the case when enhancement layer from both the sources is missing [when $p_x = p_y = 1$ or when there is no enhancement layer transmission ($R_{2x} = R_{2y} = 0$ bits)], the problem reduces to that of typical distributed source coding.
3) *Scalable Coding of a Single Source*: When only a single source is present, the S-DSC problem reduces to scalable coding of a single source.
Next we consider some other special cases where base layer information may also be lost (Note that throughout this paper we will assume for presentation simplicity, that base layer information is always received. The proposed model can easily be generalized to remove this assumption.)
4) *Scalable Multiple Descriptions Coding*: When the two sources are identical $X = Y$, and the base layer may also experience loss, the S-DSC problem reduces to scalable multiple descriptions coding.
5) *Multiple Description Coding*: This is a special case of the above with a single layer (when base layer information can be lost with some probabilities and there is no enhancement layer or equivalently all the enhancement layer information is always lost).
6) *Robust Distributed Source Coding*: In the general case of $X \neq Y$, with a single layer (no enhancement layer), the problem reduces to robust distributed source coding [15], [19].

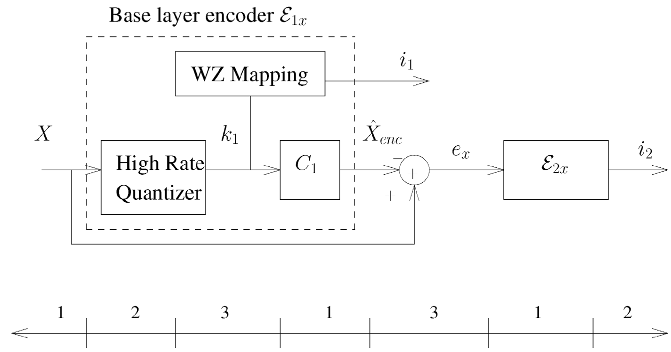


Fig. 3. MS-DSC encoder and an example of Wyner–Ziv mapping from Voronoi regions to (transmitted) indices. The source data points lying in a particular Voronoi region are mapped to an index which is then transmitted. For example, in the above figure, for source points in the second and seventh region, the same index 2 is transmitted, etc.

III. Multistage Distributed Source Coding

We begin this section by describing the various modules of the MS-DSC system and then highlight the major conflict that arises when distributed coding is naively combined with multistage coding. Next we outline an approach to resolve the conflict and propose an iterative design algorithm for the MS-DSC design that efficiently exploits intersource correlations.

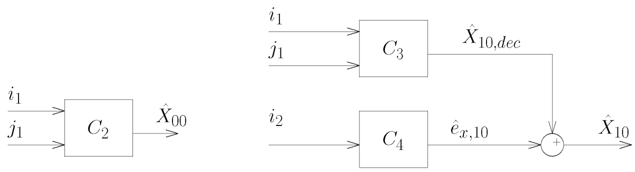
A. *Encoder*

The MS-DSC encoder for source X is shown in Fig. 3. The overall encoder $\mathcal{E}_x$ consists of two stage encoders $\mathcal{E}_{1x}$ and $\mathcal{E}_{2x}$. Input X is fed to the first stage (base-layer) encoder $\mathcal{E}_{1x}$ whose output is an index i_1 and an *encoder* reconstruction value $\hat{X}_{\text{enc}}$. The residual, $e_x = X - \hat{X}_{enc}$ is input to the second stage (enhancement layer) encoder $\mathcal{E}_{2x}$, whose output is an index i_2. Since the sources X and Y are correlated, the encoders $\mathcal{E}_{1x}$ and $\mathcal{E}_{2x}$ will differ from the nearest neighbor quantizers encountered in single-source multistage quantization.

Base layer encoder $\mathcal{E}_{1x}$ consists of a high rate quantizer (used primarily to discretize the source) which maps source X to index k_1 representing Voronoi region $C^x_{k_1}$. Next, a *lossy* (many to one) mapping which we refer to as Wyner–Ziv (WZ) mapping is employed (the name loosely accounts for the fact that the scenario involves lossy coding with side information whose asymptotic performance bound was given in [2]). The WZ mapping block takes in k_1 and outputs index $i_1 = v_1(k_1)$ representing region $R^x_{i_1} = \bigcup_{k_1:v_1(k_1)=i_1} C^x_{k_1}$, to be transmitted over the channel.[3] An example of WZ mapping for a scalar source with $\mathcal{K}_1 = 7$ and $\mathcal{I}_1 = 3$, is also given in Fig. 3.

The *encoder* codebook C_1 is a lookup table where index k_1 is mapped to reconstruction $\hat{X}_{\text{enc}}$, which is used to compute the residual e_x. Base layer encoder $\mathcal{E}_{1y}$ for source Y is similarly defined. The residuals e_x and e_y obtained by the first encoding stage are correlated and motivate distributed coding at enhancement layer to exploit intersource correlations. The second stage

[3]Note that the objective of high rate quantizers is only to discretize the source space. They do not represent the overall quantization. This could be viewed as what is known as a fine-coarse quantizer, where the coarse portion is performed by the WZ module. The WZ module maps many high-resolution cells to each index, and this covers not only the distributed coding aspect but also merging of contiguous cells to produce a much coarser partition.



Decoder $\mathcal{D}_{00}$ for X Decoder $\mathcal{D}_{10}$ for X

Fig. 4. MS-DSC decoders $\mathcal{D}_{00}$ and $\mathcal{D}_{10}$ for source X.

encoders $\mathcal{E}_{2x}$ and $\mathcal{E}_{2y}$ similarly consist of a high rate quantizer followed by WZ mapping. Since the second stage is the final stage in the setting here, no encoder codebook is needed in $\mathcal{E}_{2x}$ or $\mathcal{E}_{2y}$ (in general all except the last MS-DSC stage encoders contain an encoder codebook as in $\mathcal{E}_{1x}$).

B. Decoder

The MS-DSC system consists of four decoders (see Fig. 2), depending on whether enhancement layer information from sources X or Y is received. Decoder $\mathcal{D}_{00}$ is used when only indices i_1 and j_1 are received, and comprises of decoders $\mathcal{D}_{00}^x$ and $\mathcal{D}_{00}^y$ for sources X and Y, respectively. Each of the source decoders $\mathcal{D}_{00}^x$ (see Fig. 4) or $\mathcal{D}_{00}^y$ just consists of a single codebook and outputs base layer reconstruction as $\hat{X}_{00}$ or $\hat{Y}_{00}$, respectively.

Decoder $\mathcal{D}_{10}^x$ (part of $\mathcal{D}_{10}$) comprises two codebooks (Fig. 4). Similar to single-source multistage coding, the decoding is performed in an additive fashion. $\hat{X}_{10,\mathrm{dec}}$ is calculated based on indices i_1 and j_1, using a decoder helper codebook C_3 (which is different from codebook C_2 in (Fig. 4 to provide more flexibility in obtaining a better enhancement layer reconstruction). The reconstructed enhancement layer residual $\hat{e}_{x,10}$ is based on i_2 using the *residual* codebook C_4 as shown. The overall (enhancement layer) source reconstruction, when $\mathcal{D}_{10}^x$ is thus obtained as

$$\hat{X}_{10} = \hat{X}_{10,\mathrm{dec}} + \hat{e}_{x,10}. \qquad (5)$$

Note that the various entities $\hat{X}_{\mathrm{enc}}$, $\hat{X}_{10}$ and $\hat{X}_{10,\mathrm{dec}}$ (corresponding to $\mathcal{D}_{10}$) differ in general. In brief, these entities can be interpreted as follows:

1) $\hat{X}_{10,\mathrm{dec}}$ is the decoder helper codebook output based on i_1 and j_1 and its sole objective is to aid in the reconstruction via (5).
2) $\hat{X}_{\mathrm{enc}}$, based on k_1 is an encoder estimate of X at the base layer in order to derive the residual for the enhancement layer.
3) $\hat{X}_{10}$ is the final source reconstruction values at decoder $\mathcal{D}_{10}$.

The functioning of the Y decoder, and for the other channel loss scenarios is similarly defined in a straightforward manner.

C. Components to Optimize

The MS-DSC design algorithm optimizes the system modules at all layers and for all sources. We will restrict the scope here to the design of all codebooks and WZ mappings. (For simplicity, we will assume that high rate quantizers are independently designed using standard Lloyd's algorithm [28]. Joint optimization with the rest of the system is expected to yield small additional gains.)

D. Naive Design Scheme

We first consider the design scheme that results when distributed coding is directly combined with multistage coding. As it ignores the potential conflict in objectives, we refer to it as "naive" design. In the naive scheme, a base layer distributed coder is designed, while ignoring the enhancement layer, to minimize the base-layer distortion:

$$E[\alpha d(X, \hat{X}_{00}) + (1-\alpha) d(Y, \hat{Y}_{00})]. \qquad (6)$$

The Wyner–Ziv mappings and the decoder codebook (C_2) for both the sources at base layer are designed using a standard distributed coder design algorithm, such as in [19] and [21]. Consequently, encoder estimate $\hat{X}_{\mathrm{enc}}$ and decoder estimates $\hat{X}_{01,\mathrm{dec}}$, $\hat{X}_{10,\mathrm{dec}}$, and $\hat{X}_{11,\mathrm{dec}}$ (for decoders $\mathcal{D}_{01}$, $\mathcal{D}_{10}$ and $\mathcal{D}_{11}$, respectively) are calculated only based on index i_1. Note that there is no encoder-decoder mismatch in this scheme and $\hat{X}_{\mathrm{enc}}(i_1) = \hat{X}_{01,\mathrm{dec}}(i_1) = \hat{X}_{10,\mathrm{dec}}(i_1) = \hat{X}_{11,\mathrm{dec}}(i_1) = E[X|X \in R_{i_1}^x]$, i.e., the estimates are simply calculated as the centroids of the region $R_{i_1}^x$. The encoder codebook C_1 and decoder helper codebook C_3 for $\mathcal{D}_{10}$ (and similarly for $\mathcal{D}_{01}$ and $\mathcal{D}_{11}$) are identical (in order to avoid any potential mismatch between encoder/decoder) and solely based on the common information at both the encoder and decoder, i.e., index i_1.

The residual is calculated as $e_x = X - \hat{X}_{\mathrm{enc}}$ (and similarly for e_y). The resulting training set for $\{e_x, e_y\}$ is used to design a distributed coder for the enhancement layer to minimize the expected distortion corresponding to the last three distortion terms in (4) using a Lloyd-style algorithm for robust distributed coder design [19] in which the various codebooks and enhancement layer Wyner–Ziv mappings are optimized given the *fixed base layer coder*. For more details on robust distributed coder design, we refer the reader to [19].

E. Comments on the Naive Design Scheme

In essence, the naive scheme for MS-DSC design tries to first minimize the base layer distortion term by designing a base layer distributed coder. Given the fixed base layer distributed coder, a robust distributed coder is designed (the term robust distributed coder is used because the enhancement layer channels for the two sources can fail independently and a particular decoder is used depending on the available information) to minimize the remaining distortion terms. The inherent assumption is that ignoring (a) the enhancement layer during base layer distributed coder design and (b) the role of p_x and p_y, may not degrade the performance substantially. For example, when enhancement layer information from both sources is almost always received, p_x and p_y are close to 0 implying that the base layer decoder $\mathcal{D}_{00}$ will be used with very low probability $p_{00} = p_x p_y$. This implies that the base layer distributed coder should not be designed independently while ignoring its effects on the performance of the enhancement layer.

Further, to avoid any potential mismatch, only index i_1 (available at both the encoder and decoder) is used as input for the encoder and decoder helper codebooks, and information from the other source Y (in the form of index j_1) is ignored.

IV. Multi-Stage Distributed Coding Design Algorithm

A. Motivation and Design

The most fundamental deviation of this work from the "natural" approach to MS-DSC is in the use of different codebooks for constructing $\hat{X}_{\text{enc}}$ at the encoder and $\hat{X}_{01,\text{dec}}$, $\hat{X}_{10,\text{dec}}$ and $\hat{X}_{11,\text{dec}}$ at the decoder. In the sequel we will focus on $\hat{X}_{10,\text{dec}}$ (Extension to $\hat{X}_{01,\text{dec}}$ and $\hat{X}_{11,\text{dec}}$ is trivial.) At the decoder, both indices i_1 and j_1 can be utilized to construct $\hat{X}_{10,\text{dec}}$. However, the encoder for source X only has access to index i_1 to construct $\hat{X}_{\text{enc}}$, and does not know j_1. Obviously, there will be a mismatch between $\hat{X}_{\text{enc}}$ and $\hat{X}_{01,\text{dec}}$. A possible way to match $\hat{X}_{01,\text{dec}}$ with $\hat{X}_{\text{enc}}$ will be to make both $\hat{X}_{01,\text{dec}}$ and $\hat{X}_{\text{enc}}$ a function of i_1 alone (as was done in the naive approach). But this may in fact compromise the performance of the enhancement-layer distributed coder. To illustrate this observation, consider the Wyner–Ziv mapping example of Fig. 3: A (scalar) source point in the X space lying in the second region will be mapped to index $i_1 = 2$. The encoder estimate $\hat{X}_{\text{enc}}$ corresponding to this point may actually lie in the middle of the line (the center of mass of all source points X that is mapped to index 2). So, in general, the estimate $\hat{X}_{\text{enc}}$ will be a compromise reconstruction for a union of disparate intervals and, therefore, be coarse. This will obviously lead to large error residuals $e_x = X - \hat{X}_{\text{enc}}$.

The idea is, therefore, to allow for some mismatch between the encoder and decoder's first (or intermediate) layer estimates, and jointly optimize them so that efficient distributed coding at the consequent enhancement layer will more than compensate for any allowed mismatch. This approach is motivated by the realization that the encoder has complete knowledge of the source itself (or after high-rate discretization index k_1), while the decoder obtains additional information from the correlated source Y, in the form of index j_1. This implies that there exists some (elusive) complementary information shared by both ends, which could be exploited, if an appropriate means were devised. Joint design of the distributed coders at both layers must be performed, so that the impact of the enhancement layer and the role of p_x and p_y are not neglected while designing the base layer distributed coder.

We therefore use different codebooks at the decoder and encoder for calculating $\hat{X}_{10,\text{dec}}$ (and similarly $\hat{X}_{01,\text{dec}}$ and $\hat{X}_{11,\text{dec}}$) and $\hat{X}_{\text{enc}}$. The encoder codebook (C_1) now has k_1 as input, and the *decoder helper* codebooks have inputs i_1 and j_1. This flexibility enables optimization of the tradeoff between better exploitation of intersource correlations at the subsequent layer, and the cost of some mismatch in the system. Appropriate design of encoder and decoder codebooks (as well as WZ mappings) will optimize a precise measure of the overall performance, while accounting for the mismatch.

Note that the scheme subsumes single source multistage quantizer design as a special case. Also, when the sources X and Y are uncorrelated, then WZ mappings for the base layer will converge to a union of contiguous cells (the encoder $\mathcal{E}_{1x}$ will act as a fine-coarse quantizer) and both the encoder and decoder helper codebooks will effectively be the same and depend on i_1 only.

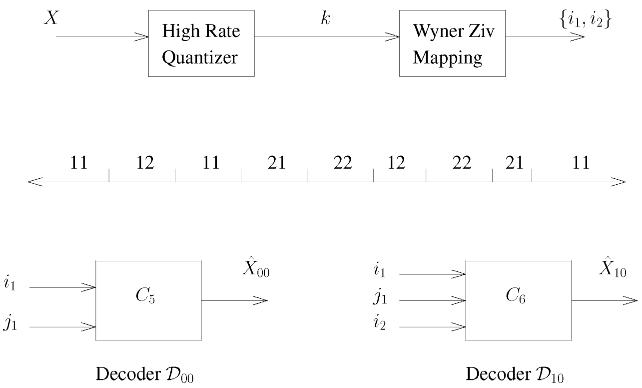


Fig. 5. S-DSC encoder for source X; an example of Wyner–Ziv mapping from Voronoi regions to index pair $\{i_1, i_2\}$. Here the first region is mapped to index pair $\{i_1, i_2\} = \{1, 1\}$ etc.; and decoders $\mathcal{D}_{00}$ and $\mathcal{D}_{10}$ in S-DSC.

B. Update Rules for the Proposed MS-DSC Algorithm

To minimize the cost in (4), the Wyner–Ziv mappings and the various codebooks are optimized iteratively using the necessary update rules outlined below. In the update rules, $C^x_{k_1}$ and $R^x_{i_1}$ represent the regions corresponding to high rate quantizer index k_1 and transmission index i_1 for base layer of source X respectively as described in Section III-A. Similarly $C^x_{k_2}$ and $R^x_{i_2}$ denote the regions corresponding to high rate quantizer index k_2 and transmission index i_1 for enhancement layer of source X. The corresponding symbols for source Y such as $R^y_{j_1}$, $R^y_{j_2}$ are similarly defined. Also, to reduce clutter, we omit superscripts and arguments in the following update rules where obvious from the context, e.g., R_{i_1} instead of $R^x_{i_1}$ or $\hat{e}_{x,11}$ rather than $\hat{e}_{x,11}(i_2, j_2)$.

1) **Base Layer Decoder Codebook** (C_2, at $\mathcal{D}_{00}$)

$$\hat{x}_{00}(i_1, j_1) = \arg\min_{\phi} \sum_{(x,y)\in R_{i_1}\times R_{j_1}} d(x, \phi). \tag{7}$$

2) **Enhancement Layer Decoder Codebooks** (for residuals, at decoders $\mathcal{D}_{01}$, $\mathcal{D}_{10}$ and $\mathcal{D}_{11}$)

$$\begin{aligned}\hat{e}_{x,01}(j_2) &= \arg\min_{\phi} \sum_{e_y\in R_{j_2}} d(x, \hat{x}_{01,\text{dec}} + \phi)\\ \hat{e}_{x,10}(i_2) &= \arg\min_{\phi} \sum_{e_x\in R_{i_2}} d(x, \hat{x}_{10,\text{dec}} + \phi)\\ \hat{e}_{x,11}(i_2, j_2) &= \arg\min_{\phi} \sum_{(e_x,e_y)\in R_{i_2}\times R_{j_2}} d(x, \hat{x}_{11,\text{dec}} + \phi).\end{aligned} \tag{8}$$

3) **Encoder Codebook** (C_1):

$$\begin{aligned}\hat{x}_{\text{enc}}(k_1) = \arg\min_{\phi} \sum_{x\in C_{k_1}} & p_{01} d(x, \hat{x}_{01,\text{dec}} + \hat{e}_{x,01})\\ &+ p_{10} d(x, \hat{x}_{10,\text{dec}} + \hat{e}_{x,10})\\ &+ p_{11} d(x, \hat{x}_{11,\text{dec}} + \hat{e}_{x,11})\end{aligned} \tag{9}$$

where the dependence on ϕ comes from $\hat{e}_{x,01}$, $\hat{e}_{x,10}$ and $\hat{e}_{x,11}$, which are the estimates of e_x at the enhancement layer and $e_x = x - \phi$.

4) **Base Layer Decoder Helper Codebooks** (at decoders $\mathcal{D}_{01}$, $\mathcal{D}_{10}$ and $\mathcal{D}_{11}$)

$$\hat{x}_{01,\mathrm{dec}}(i_1,j_1) = \arg\min_{\psi} \sum_{(x,y)\in R_{i_1}\times R_{j_1}} d(x, \hat{e}_{x,01}+\psi)$$
$$\hat{x}_{10,\mathrm{dec}}(i_1,j_1) = \arg\min_{\psi} \sum_{(x,y)\in R_{i_1}\times R_{j_1}} d(x, \hat{e}_{x,10}+\psi)$$
$$\hat{x}_{11,\mathrm{dec}}(i_1,j_1) = \arg\min_{\psi} \sum_{(x,y)\in R_{i_1}\times R_{j_1}} d(x, \hat{e}_{x,11}+\psi). \tag{10}$$

5) **WZ Mappings (Base Layer)**: For $k_1 = 1 : \mathcal{K}_1$, assign k_1 to index $i_1 = v_1(k_1)$ such that

$$v_1(k_1) = i_1 = \arg\min_{i_1' \in \{1\cdots I_1\}} \sum_{x\in C_{k_1}} D_{\mathrm{net}}(x,y) \tag{11}$$

where the dependence of $D_{\mathrm{net}}(x,y)$ on the index i_1 is specified by (2) and (3).

6) **WZ Mappings (Enhancement Layer)**: For $k_2 = 1 : \mathcal{K}_2$, assign k_2 to index $i_2 = v_2(k_2)$ such that

$$v_2(k_2) = i_2 = \arg\min_{i_2' \in \{1\cdots I_2\}} \sum_{e_x\in C_{k_2}} D_{\mathrm{net}}(x,y) \tag{12}$$

where the sum is over the residuals e_x which lie in the region C_{k_2} and the dependence of $D_{\mathrm{net}}(x,y)$ on the index i_2 is again specified by (2) and (3).

The update rules for the second source Y are straightforward to specify from the above.

V. Scalable-Distributed Source Coding

In the general S-DSC setting (see Fig. 2), encoders for sources X and Y transmit index pairs $\{i_1, i_2\}$ and $\{j_1, j_2\}$, respectively. The source encoder for X directly generates the index pair $\{i_1, i_2\}$. Similar to MS-DSC, decoder $\mathcal{D}_{00}$ consists of a single codebook (per source) and takes indices i_1 and j_1 as input to obtain the reconstruction $\hat{X}_{00}$ (and $\hat{Y}_{00}$). The decoders $\mathcal{D}_{01}$, $\mathcal{D}_{10}$ and $\mathcal{D}_{11}$ also have a single codebook (per source) and decoding is performed directly using the codebook and not decomposed into additive terms as in MS-DSC. A block diagram depicting the S-DSC encoder $\mathcal{E}_x$ and decoders $\mathcal{D}_{00}$ and $\mathcal{D}_{01}$ (for X) is shown in Fig. 5.

Obviously, S-DSC in its general setting, if well optimized, should outperform its constrained variant of MS-DSC. Also, there is no direct conflict between the objectives of distributed and scalable coding in S-DSC, since the design for all the stages can be performed simultaneously and there is no feedback (dependence) for calculating the source error residual for any intermediate stage in S-DSC.

However, the S-DSC encoding complexity grows exponentially with the total rate, i.e., $2^{R_{1x}+R_{2x}}$ for source X (and similarly for Y). Note that the MS-DSC encoding complexity is in the order of $2^{R_{1x}} + 2^{R_{2x}}$ for source X. Moreover, the total storage required for the various decoder codebooks will grow more rapidly in S-DSC. For example, S-DSC's codebook for decoder $\mathcal{D}_{11}$ will have storage proportional to $2^{R_{1x}+R_{2x}+R_{1y}+R_{2y}}$ whereas the corresponding MS-DSC storage for the codebooks is approximately proportional to $(2^{R_{1x}+R_{1y}} + 2^{R_{2x}+R_{2y}})$ due to the additive nature of decoding.

Efficient design of Wyner–Ziv mappings from high resolution quantizer cells to indices is important for good performance of the S-DSC system. The index-assignment problem implicit in WZ mapping is a discrete optimization problem. In the case of S-DSC, we need to map each of the $\mathcal{K}$ cells (formed by the high-rate quantizer) to one of the $\mathcal{I}_1 \times \mathcal{I}_2$ indices for source X (see Fig. 5). One can generalize a distributed quantizer algorithm (such as in [19] and [21]) for the design of S-DSC system. However, random initialization of the WZ mapping in the generalized DSC iterative algorithm (described below in Section V-A) virtually always leads to a poor local minimum. While we do not attempt a global solution to this index assignment problem, we propose to use an optimized MS-DSC system as a highly effective initialization for S-DSC design. Simulation results confirm that the proposed initialization obtains considerable gains over uninformed, randomly initialized solutions.

We next describe the locally optimal Lloyd-style algorithm alongwith its update rules for S-DSC. We then explain how the MS-DSC solution can be used as an efficient initialization for the S-DSC problem.

A. Iterative Design Algorithm

We use similar notation in Section III to denote the various S-DSC modules (Fig. 5). The high-rate quantizer for source X maps the source sample to a prototype associated with the region C_k^x where $k \in \{1\ldots\mathcal{K}\}$ and $\mathcal{K}$ is the total number of prototypes. Next the WZ mapping block maps the prototype to transmission index pair $\{i_1, i_2\}$ where $\{i_1, i_2\} \in \{1,\ldots,\mathcal{I}_1\} \times \{1,\ldots,\mathcal{I}_2\}$. Similar procedure is performed at the second source encoder and indices j_1 and j_2 are transmitted for the base and enhancement layer respectively.

We next assume the squared error distortion measure to obtain simpler and intuitive intepretable results. To minimize the cost in (4), the WZ mappings and the various decoder codebooks are optimized iteratively using the following update rules derived from the necessary conditions for optimality:

1) **Wyner–Ziv Mappings** For $k = 1 : \mathcal{K}$, assign k to index pair $\{i_1, i_2\}$ such that:

$$\{i_1, i_2\} = \arg\min_{i_1', i_2'} \sum_{x\in C_k} D_{\mathrm{net}}(x,y). \tag{13}$$

Here also the dependence of $D_{\mathrm{net}}(x,y)$ on the index pair is specified by (2) and (3).

2) **Decoder Codebook: Reconstruction Values**

$$\hat{x}_{00}(i_1,j_1) = E\left[X | X\in R_{i_1}^x, Y\in R_{j_1}^y\right]$$
$$\hat{x}_{01}(i_1,j_1,j_2) = E\left[X | X\in R_{i_1}^x, Y\in R_{j_1,j_2}^y\right]$$
$$\hat{x}_{10}(i_1,j_1,i_2) = E\left[X | X\in R_{i_1,i_2}^x, Y\in R_{j_1}^y\right]$$

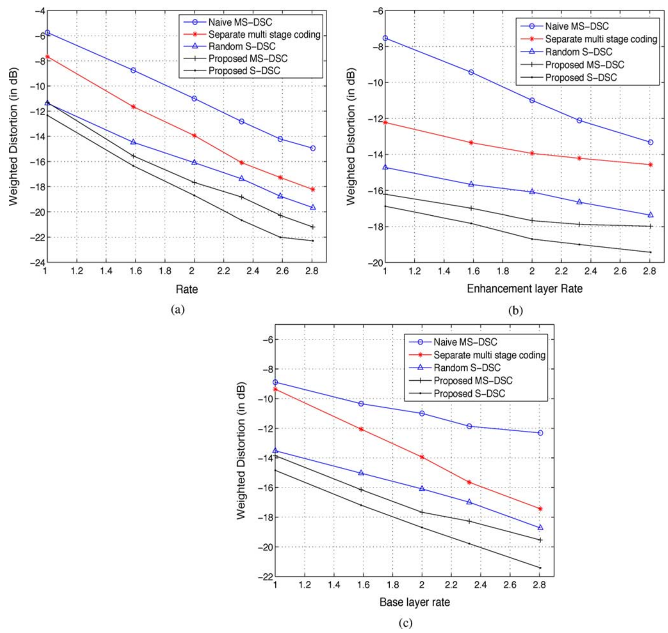


Fig. 6. Performance comparison of naive scheme for MS-DSC, separate (single source) multistage coding, randomly initialized scalable DSC, multistage DSC, and "cleverly" initialized scalable DSC technique. (a) versus transmission rate; (b) versus enhancement layer rate (base layer rate fixed at 2 bits/sample); (c) versus base layer rate (enhancement layer rate fixed at 2 bits/sample).

$$\hat{x}_{11}(i_1, j_1, i_2, j_2) = E\left[X | X \in R^x_{i_1,i_2}, Y \in R^y_{j_1,j_2}\right] \quad (14)$$

Similar rules for Y are easily obtained and will not be reproduced here. In the above update rules, $R^x_{i_1} = \bigcup_{i_2} R^x_{i_1,i_2}$ and $R^y_{j_1} = \bigcup_{j_2} R^y_{j_1,j_2}$.

B. Effective Initialization for S-DSC Design

MS-DSC is a special case of S-DSC under additive encoding/decoding constraints. The proposed scheme for S-DSC will use the optimal MS-DSC system as an effective initialization, and will then remove the structural constraints to apply the iterative algorithm of Section V-A.

In the base layer of the MS-DSC scheme, the source space for X is partitioned into $\mathcal{K}_1$ different cells. These K_1 cells are mapped to $\mathcal{I}_1$ regions via the base-layer WZ mapping module. The residual $e_x = X - \hat{X}_{\text{enc}}$ is then quantized by a high-rate quantizer having $\mathcal{K}_2$ different output cells (regions), which are mapped to the $\mathcal{I}_2$ regions via enhancement layer WZ mapping module. Hence, during design, all the training point samples for source X are associated with corresponding index pairs $\{i_1, i_2\}$.

Now consider a sample X that falls in high-rate quantizer cell $C^x_{k_1}$ and then mapped to index i_1. This sample is then mapped to index i_2 at the enhancement layer (through e_x). We define C^x_k ($k \in \{1 \ldots K\}$ and $K = K_1 \times I_2$) as the set of all source points X that lie in both cells $C^x_{k_1}$ (corresponding to high rate quantizer output index k_1) and $R^x_{i_2}$ (corresponding to index i_2 of the enhancement layer WZ mapping), i.e.,

$$C^x_k = C^x_{k_1} \bigcap R^x_{i_2}. \quad (15)$$

Now, each of the regions C^x_k is associated to an index pair $\{i_1, i_2\}$. So we effectively view the X source space as divided into K different regions, each of which is mapped to an index pair $\{i_1, i_2\}$ via an S-DSC WZ mapping $v(k) = \{i_1, i_2\}$. We can use these $\mathcal{K}$ cells and WZ mappings as an initial solution for the S-DSC algorithm in Section V-A. Note that since the region $R^x_{i_2}$ corresponding to index i_2 (outcome of enhancement layer WZ mapping) is in general a union of possibly, non-contiguous cells, then so is the region C^x_k in (15). A similar construction of the different regions and WZ mapping is performed for source Y.

VI. Simulation Results

We give several examples to demonstrate the gains of (a) the proposed MS-DSC scheme which resolves the conflict between distributed quantization and multistage coding and (b) the proposed MS-DSC initialized S-DSC scheme. In all the simulations, sources X and Y are assumed to be jointly Gaussian with zero means, unit variances and correlation coefficient ρ. The weighting coefficient α of (2) is set to 0.5 to give equal importance to both the sources at the decoder. A training set of 10 000 scalars is generated. The number of prototypes is 60 for the high rate quantizers which are designed using Lloyd's algorithm [28]. We compare four different schemes (a) separate (single-source) multistage coding in which no distributed coding is performed, (b) randomly initialized S-DSC system ('Random S-DSC'), (c) structurally constrained MS-DSC system (Proposed 'MS-DSC'), and (d) proposed S-DSC system which is initialized by MS-DSC ('Proposed S-DSC').

In the first set of experiments (see Fig. 6), we plot the weighted distortion at the decoder (4) versus rate for the various schemes. The probability of enhancement layer loss is $p_x = p_y = 0.2$ and the correlation coefficient is $\rho = 0.97$. In Fig. (a), the same incremental transmission rate is allocated to each layer of each source, i.e., $R_{1x} = R_{2x} = R_{1y} = R_{2y} = R$. As an additional reference, we also plot the performance of the "naive" MS-DSC scheme in this set of experiments. The proposed MS-DSC scheme achieves substantial gains, e.g., 6.8 dB over the naive scheme (at $R = \log_2(3)$ bits) while the proposed S-DSC scheme leads to considerable gains over Random S-DSC approach, e.g., ~3.3 dB at $R = \log_2(5)$ bits/sample. Note that the naive MS-DSC scheme even underperforms separate (single source) multistage coding, as it ignores the potential conflict between the objectives of distributed quantization and multistage coding. In Fig. (b), the base layer rates R_{1x} and R_{1y} are fixed at 2 bits/sample while the enhancement layer rate R_{2x} ($=R_{2y}$) is varied. Again, the proposed MS-DSC and S-DSC schemes outperform their respective counterparts, namely, naive MS-DSC and Random S-DSC, with gains of 9 and 2.6 dB, at rates 1 and 2 bits/sample, respectively. In Fig. (c), the base layer rate R_{1x} ($=R_{1y}$) is varied while the enhancement layer rate $R_{2x} = R_{2y}$ is fixed at 2 bits/sample. Here, again, the proposed MS-DSC and S-DSC schemes outperform naive MS-DSC and Random S-DSC scheme with gains of 7.2 and 2.8 dB, respectively, at rates $\log_2(7)$ and 2 bits/sample, respectively.

In the next set of experiments (see Fig. 7), we fix all the transmission rates $R_{1x} = R_{2x} = R_{1y} = R_{2y} = 2$ bits/sample. The probability of enhancement layer loss for both sources p_x ($=p_y$)

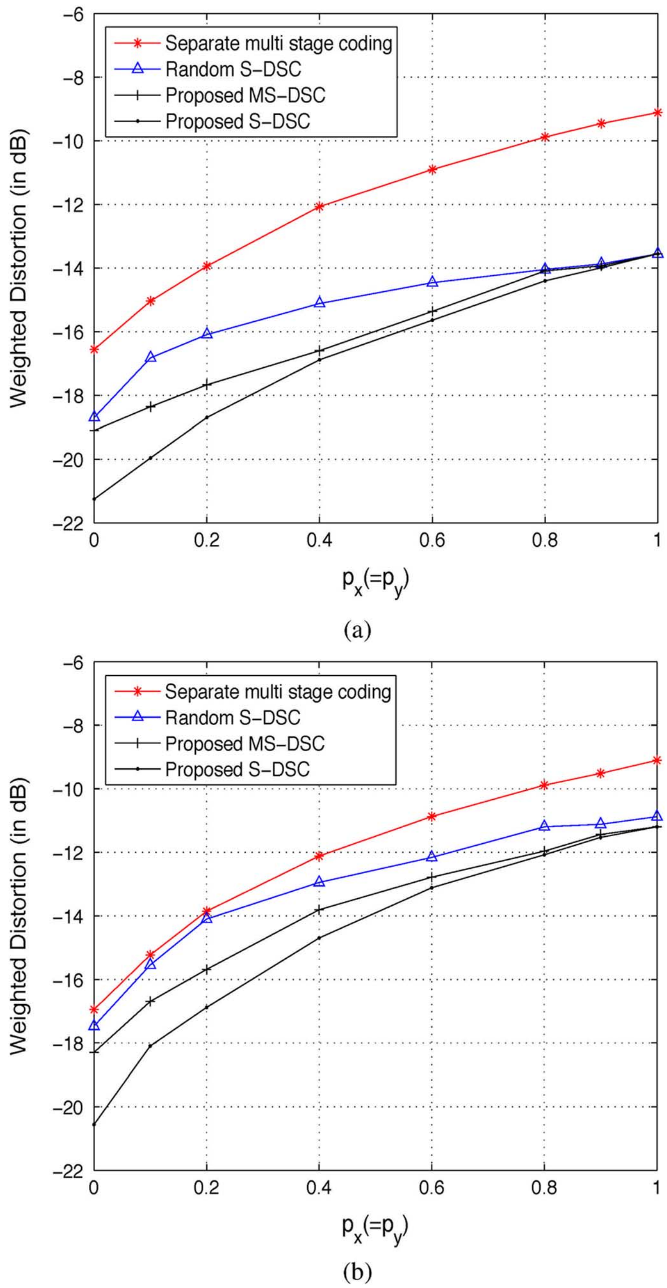


Fig. 7. Performance comparison of separate (single source) multistage coding, randomly initialized scalable DSC, multistage DSC, and "cleverly" initialized scalable DSC technique as the probability of enhancement layer loss p_x $(=p_y)$ is varied. All the transmission rates are 2 bits/sample. In (a) intersource correlation $\rho = 0.97$ while in (b) $\rho = 0.9$.

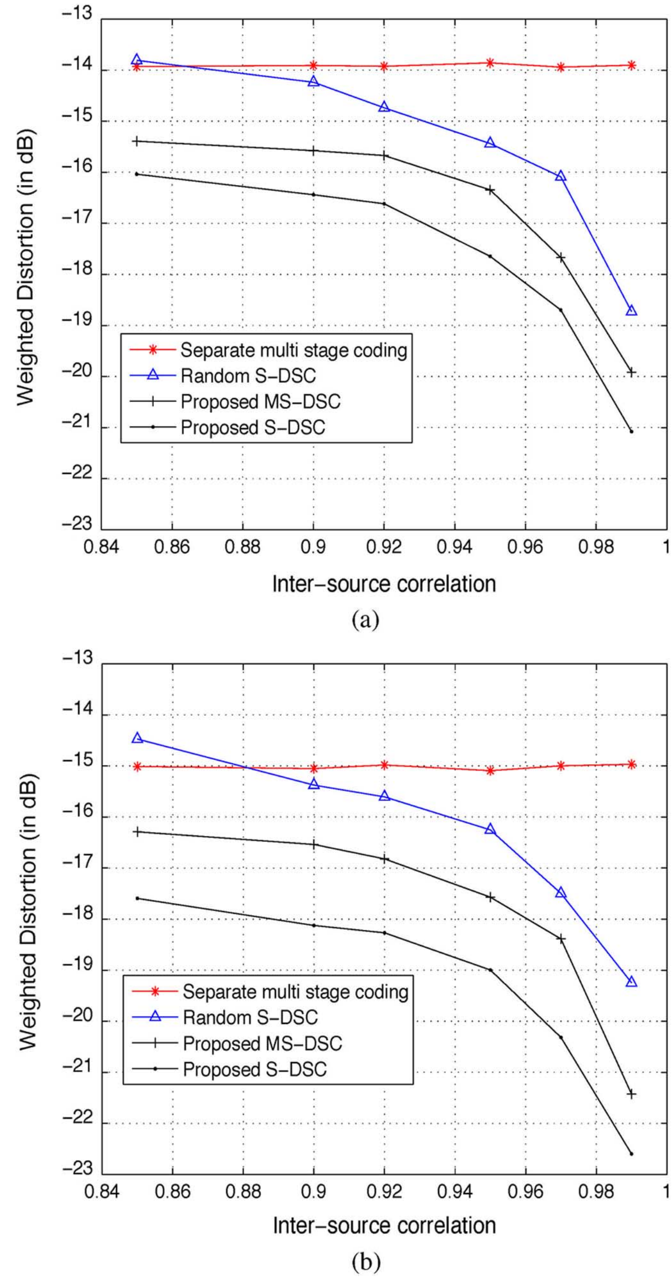


Fig. 8. Performance comparison of separate (single source) multistage coding, randomly initialized scalable DSC, multistage DSC, and "cleverly" initialized scalable DSC technique as the intersource correlation is varied. All the transmission rates are 2 bits/sample. The probability of enhancement layer loss p_x $(=p_y)$ is 0.2 in (a) and 0.1 in (b).

is varied and the weighted distortion is plotted. In (a), the intersource correlation ρ is 0.97 while in (b) ρ is 0.9. Note that when p_x $(=p_y)$ is close to 1, no enhancement layer information is received and the various schemes (Random S-DSC, Proposed MS-DSC and Proposed S-DSC) simply reduce to nonscalable distributed coding, and hence their performance converges. On the other hand, as p_x $(=p_y)$ approaches 0, the proposed MS-DSC and S-DSC schemes consistently outperform the other schemes.

Next we plot the weighted distortion as a function of intersource correlation, for two different probability of enhancement layer loss p_x $(=p_y)$: 0.2 and 0.1, respectively (see Fig. 8). Again, all transmission rates $R_{1x} = R_{2x} = R_{1y} = R_{2y}$ are fixed at 2 bits/sample. Here the separate multistage coding scheme does not utilize correlation between the source, hence its performance is independent of intersource correlation, ρ. We find that the proposed MS-DSC and S-DSC schemes consistently outperform other schemes. Further note that while intersource correlation is utilized by the Random S-DSC approach, the advantage of distributed coding is severely compromised by the random initialization. In fact, in the simulations the performance of the Random S-DSC algorithm is even worse than MS-DSC most of the time, which is strong evidence for the importance of good initialization for S-DSC.

Finally, we note that the proposed methods are extendible to incorporate entropy coding, but such extension is omitted for brevity. Further, for fair comparison and to eliminate atypically poor results, the initialization and design of both the Random S-DSC and MS-DSC algorithms was performed 20 times and the best results are reported. (Optimization using global variants of the technique is beyond the scope of this paper.) In terms of complexity, the proposed S-DSC design takes about 4–5 times longer than the randomly initialized S-DSC design.

VII. Conclusion

We have considered the design of scalable distributed source coders. The proposed S-DSC system is robust to failure of subset of the channels and utilizes all available information for compression efficiency. We first identify the inherent conflict between the objectives of distributed quantization and multi-stage coding and show how to resolve the conflict in the design of an MS-DSC system, a special constrained case of the S-DSC problem. Our scheme allows a "controlled mismatch" between the encoder and decoder reconstructions and jointly optimizes all the components in the MS-DSC system. Next we show that a Lloyd-style iterative S-DSC algorithm is heavily dependent on initialization and will typically underperform the proposed multistage DSC algorithm despite the latter's structural constraint. The proposed MS-DSC algorithm solution is used as an effective initialization for the S-DSC design algorithm. Simulation results show that (a) the proposed MS-DSC scheme consistently outperforms other naive schemes and single source (separate) distributed multi stage coding schemes, and (b) MS-DSC initialized S-DSC consistently and significantly outperforms randomly initialized S-DSC.

References

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

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

[3] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction.," *IEEE Trans. Inf. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.

[4] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.*, vol. 19, no. 2, pp. 51–60, Mar. 2002.

[5] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Process. Mag.*, vol. 21, no. 5, pp. 80–94, Sep. 2004.

[6] B. Rimoldi, "Successive refinement of information: Characterization of the achievable rates," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 253–259, Jan. 1994.

[7] W. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 269–275, Nov. 1991.

[8] V. Koshelev, "An evaluation of the average distortion for discrete schemes of sequential approximation," *Probl. Pered. Inf.*, vol. 17, no. 3, pp. 20–33, Jan. 1981.

[9] E. Tuncel and K. Rose, "Additive successive refinement," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1983–1991, Aug. 2003.

[10] Y. Steinberg and N. Merhav, "On successive refinement for the Wyner–Ziv problem," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1636–1654, Aug. 2004.

[11] C. Tian and S. Diggavi, "Multistage successive refinement for Wyner–Ziv source coding with degraded side informations," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2006, pp. 1594–1598.

[12] L. Ozarow, "Source-coding problem with two channels and three receivers," *Bell Syst. Tech. J.*, vol. 59, no. 10, pp. 1909–1921, 1980.

[13] P. Ishwar, R. Puri, S. S. Pradhan, and K. Ramchandran, "On compression for robust estimation in sensor networks," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun.–Jul. 2003, p. 193.

[14] J. Chen and T. Berger, "Robust coding schemes for distributed sensor networks with unreliable sensors," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun.–Jul. 2004, p. 115.

[15] J. Chen and T. Berger, "Robust distributed source coding," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3385–3398, Aug. 2008.

[16] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer Academic, 1992.

[17] A. Saxena and K. Rose, "Distributed multistage coding of correlated sources," in *Proc. IEEE Data Compress. Conf.*, Mar. 2008, pp. 312–321.

[18] A. Saxena and K. Rose, "On distributed quantization in scalable and predictive coding," in *Sens., Signal Inf. Process. (SENSIP) Workshop*, May 2008.

[19] A. Saxena, J. Nayak, and K. Rose, "On efficient quantizer design for robust distributed source coding," in *Proc. IEEE Data Compress. Conf.*, Mar. 2006, pp. 63–72.

[20] M. Fleming, Q. Zhao, and M. Effros, "Network vector quantization," *IEEE Trans. Inf. Theory*, vol. 50, no. 8, pp. 1584–1604, Aug. 2004.

[21] D. Rebollo-Monedero, R. Zhang, and B. Girod, "Design of optimal quantizers for distributed source coding," in *Proc. IEEE Data Compress. Conf.*, Mar. 2003, pp. 13–22.

[22] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantizers over noisy channels," *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1693–1700, Jun. 2009.

[23] A. Saxena and K. Rose, "Scalable distributed source coding," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 713–716.

[24] X. Wang and M. T. Orchard, "Design of trellis codes for source coding with side information at the decoder," in *Proc. IEEE Data Compress. Conf.*, Mar. 2001, pp. 361–370.

[25] P. Mitran and J. Bajcsy, "Coding for the Wyner–Ziv problem with turbo-like codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2002, p. 91.

[26] J. Garcia-Frias and Y. Zhao, "Compression of correlated binary sources using turbo codes," *IEEE Commun. Lett.*, vol. 5, no. 10, pp. 417–419, Oct. 2001.

[27] A. D. Liveris, Z. Xiong, and C. Eorghiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 440–442, Oct. 2002.

[28] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

Ankur Saxena (S'06–M'09) was born in Kanpur, India, in 1981. He received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, in 2003 and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California, Santa Barbara, in 2004 and 2008, respectively.

He has interned with the Fraunhofer Institute of X-Ray Technology, Erlangen, Germany, and NTT DoCoMo Research Labs, Palo Alto, CA, during the summers of 2002 and 2007, respectively. He is currently a Postdoctoral Researcher with the University of California, Santa Barbara. His research interests span source coding, image, and video compression and signal processing.

Dr. Saxena was a recipient of the President Work study award during his Ph.D. studies.

Kenneth Rose (S'85–M'91–SM'01–F'03) received the Ph.D. degree in 1991 from the California Institute of Technology, Pasadena.

He then joined the Department of Electrical and Computer Engineering, University of California at Santa Barbara, where he is currently a Professor. His main research activities are in the areas of information theory and signal processing, and include rate-distortion theory, source and source-channel coding, audio and video coding and networking, pattern recognition, and nonconvex optimization. He is interested in the relations between information theory, estimation theory, and statistical physics, and their potential impact on fundamental and practical problems in diverse disciplines.

Dr. Rose was a corecipient of the 1990 William R. Bennett Prize Paper Award of the IEEE Communications Society, as well as the 2004 and 2007 IEEE Signal Processing Society Best Paper Awards.