Research Article
Quantization for Robust Distributed Coding

Xiaolin Wu, 1 Abdul Bais, 2 and Nima Sarshar 3

1 McMaster University, Hamilton, ON, Canada
2 Faculty of Engineering and Applied Science, University of Regina, 3737 Wascana Parkway, Regina, SK, Canada S4S 0A2
3 inPowered, San Francisco, CA, USA

Correspondence should be addressed to Abdul Bais; abdul.bais@uregina.ca

Received 29 September 2015; Revised 11 April 2016; Accepted 27 April 2016

A distributed source coding approach is proposed for robust data communications in sensor networks. When sensor measurements are quantized, possible correlations between the measurements can be exploited to reduce the overall rate of communication required to report these measurements. Robust distributed source coding (RDSC) approaches differentiate themselves from other works in that the reconstruction error of all sources will not exceed a given upper bound, even if only a subset of the multiple descriptions of the distributed source code are received. We deal with practical aspects of RDSC in the context of scalar quantization of two correlated sources. As a benchmark to evaluate the performance of the proposed scheme, we derive theoretically achievable distortion-rate performances of an RDSC for two jointly Gaussian sources by applying known results on the classical multiple description source coding.

1. Introduction and Motivation

Sensor networks are usually collections of remote sensors that gather temporal and spatial information about a source distributed in time and space. Each sensor reports its measurements to a central unit which extracts meaningful information about the distributed source being measured. The spatial and temporal correlations between sensor measurements can be exploited to reduce the amount of communication required between the sensors and the central unit to achieve a desired reconstruction quality of the distributed source. We call this problem the distributed source coding (DSC) problem.

For low-cost sensors, remote communication links to the central server are often unreliable. Therefore, at each given point in time, only a subset of all measurements might be available at the decoding unit. Even though the overall communication rate can be minimized by a distributed coding scheme that completely removes the redundancies between the sensor measurements, the reconstruction quality of the sources when a subset of the measurements are available might be unreasonably poor. Thus, robust distributed source coding is a generalization of the well studied multiple description (MD) coding in which all descriptions are describing a single source.

A robust distributed source coding (RDSC) scheme, on the other hand, minimizes the overall communication rate while satisfying a given bound on the maximum reconstruction error when only a subset of the measurements are available. Theoretical aspects of RDSC problem are addressed in [1, 2]. Practical DSC schemes often consist of a quantizer followed by lossless discrete source coders. The correlations between the quantized values of two sensors are used to reduce the overall rate required to communicate them to the server. This reduction is still possible even if each sensor has only access to its own readings, in compliance with Slepian-Wolf’s theorem, as pursued practically in [3]. Moreover, the quantizers can also be designed to optimally exploit this correlation in conjunction with the lossless coders following them. Such schemes are reported in [2, 4–7] mainly for two sources with or without side information. Scalability of distributed coding of correlated sources is considered in [5]. Similarly, [7] demonstrates that for binary symmetric and additive white Gaussian noise channel the correlation between sources can be useful in reducing quantization distortion and protecting data when
transmitted over noisy channels. Discussion on distortion performance gain and quantizer design using intersensor communication is reported in [8].

This work is a practical attempt to address the RDSC problem in the context of optimally designing scalar quantizers. We also introduce a paradigm shift in the problem formulations, which we believe is more appropriate for a distributed measurement context. The goal of conventional DSC and RDSC problems is the reconstruction of a single source X from the encoded versions of a number of noisy measurements of X (one for each sensor). We, on the other hand, assume that each sensor i is measuring a separate random variable X_i, with X_i's being possibly dependent. The objective is to minimize the expected distortion of the reconstructions of all sources.

The rest of this paper is organized as follows. In Section 2, we formulate the problem and present a graph theoretic, iterative algorithm to find a locally minimal solution of the problem in Section 3. To provide a benchmark against which the designs in Section 3 can be compared, a theoretically achievable distortion-rate performance of RDSC of two jointly Gaussian sources is derived in Section 4. Experimental results are reported in Section 5 and the paper is concluded in Section 6.

2. Problem Formulation

Our overall communications model is depicted in Figure 1. In this model, the outputs of two scalar sources, X_1, i = 1, 2, are quantized with two quantizers q_1(·) and q_2(·). The quantizer cell boundaries are denoted by q_{1,0} < q_{1,1} < ··· < q_{1,M} ∈ ℜ, where ℜ represents the real line. Let Q_1 and Q_2 be the discrete random variables as a result of quantizing X_1 and X_2. Q_1 and Q_2 are then fed into their respective discrete lossless encoders. The encoders might or might not take into account the correlations between the two symbols in order to reduce the overall communication rates. Based on this, a number of different coding cases are possible as suggested in [4], which determine the communication rates used in further optimization processes. As an example, for variable length coding of quantized source symbols, if all dependencies between the symbols in the two encoders are ignored, the communication rate would simply be R(Q_1, Q_2) = H(Q_1) + H(Q_2), the sum of the entropy rates of marginal quantized symbols.

On the other hand, if the dependencies of the symbols are fully exploited to achieve the Slepian-Wolf bound, then the communication rate would be R(Q_1, Q_2) = H(Q_1, Q_2) = H(Q_1 | Q_2).

The reconstructions of input vector source X = (X_1, X_2) given only Q_1 or Q_2 are called the marginal reconstructions and are denoted by X̂_1 = (X̂_1,0, X̂_1,1) and X̂_2 = (X̂_2,0, X̂_2,1), respectively. The reconstructed vector source using both Q_1 and Q_2 is denoted by X̂_0 = (X̂_0,0, X̂_0,1). The marginal and joint distortions are therefore

\[
\begin{align*}
  d_i(q_1, q_2) &= E \left( D \left( X, X_i \right) \right) \\
  &= E \left( D \left( X_1, X̂_1 \right) \right) + E \left( D \left( X_2, X̂_2 \right) \right),
\end{align*}
\]

where i = 0, 1, 2 and the expectations are over joint distribution of X_1 and X_2. The quantizer arguments in \( d_i(q_1, q_2) \) and \( R(q_1, q_2) \) are dropped when the dependence on the quantized symbols Q_1 and Q_2 is understood.

The problem can now be formulated as a feasibility problem or the problem of characterizing all the achievable tuples \((d_1, d_2, d_0, R)\). A suitable optimization problem corresponding to this feasibility problem is the following:

\[
\begin{align*}
  \min_{q_1, q_2} & \quad d_0(q_1, q_2) \\
  \text{subject to} & \quad d_1 \leq \delta_1 \\
  & \quad d_2 \leq \delta_2 \\
  & \quad R(q_1, q_2) \leq r
\end{align*}
\]

for a given triplet \((r, \delta_1, \delta_2)\), assuming the optimization problem is feasible. This formulation is identical to multiple description coding (MDC) problem when \( X_1 = X_2 \). The optimal quantizer design for MDC has been considered in [9, 10]. A similar approach to that in [9] is adopted in this paper to deal with the more general RDSC problem. This involves minimization of a Lagrangian similar to those in [4, 9] corresponding to the constrained optimization problem in (2) as given below:

\[
\begin{align*}
  \min_{q_1, q_2} & \quad L(q_1, q_2) = d_0 + \lambda_1 d_1 + \lambda_2 d_2 + \lambda R,
\end{align*}
\]

where \(\lambda_1, \lambda_2, \) and \(\lambda\) are proper Lagrangian coefficients.

3. An Iterative Design Algorithm

The joint optimization of \(q_1\) and \(q_2\) seems to be computationally prohibitive due to the interdependence of the conditional PDFs of each source given the quantized value of the other. We instead adopt an iterative approach in which each quantizer is optimized assuming the other one is fixed and known. Given two initial quantizers \(q_1^0\) and \(q_2^0\), a series of
successively improved quantizers \( q_n^1, q_n^2, n = 0, 1, 2, \ldots \), can be constructed as follows:

\[
q_n^1 = \arg \min_{q_1} L(q_1, q_n^2), \\
q_n^2 = \arg \min_{q_2} L(q_n^1, q_2).
\]  

(4)

The iteration stops after a predetermined number of steps or when the reduction in the value of \( L \) becomes negligible in consecutive iterations.

The actual minimization procedures in (4) are performed with a generalization of the approach in [9]. The optimization problem is cast into the shortest path problem as follows. First, the domain of \( X_1 \) is uniformly prequantized into \( N \) cells of interval length \( \Delta \) each, \( N \) being sufficiently large. Next, we construct an acyclic directed graph \( G \), called quantizer graph, by associating each prequantization cell boundary \( j, 0 \leq j \leq N \), with a node of the graph \( G \). We introduce a directed edge from node \( j \) to node \( j' \), \( 0 \leq j < j' \leq N \), to represent the interval \( x_{ij'} = [x_0 + j\Delta, x_0 + j'\Delta) \), where \( x_0 \) is the smallest value in the support of \( X_1 \) that is taken into account. By including in \( G \) the edges of all pairs of \( j \) and \( j' \), \( G \) becomes a complete directed acyclic graph.

When designing the quantizer \( q_n^1 \) in the \( n \)th iteration with \( q_1 \) fixed as \( q_2^{n-1} \), we assign to the edge from node \( j \) to node \( j' \) the following weight:

\[
W_{ij}(j, j') = (1 - p_1)(1 - p_2)E\left(D\left(X, \hat{X}^0 \mid X_1 \in x_{ij'}\right)\right) + p_2(1 - p_1)E\left(D\left(X, \hat{X}^1 \mid X_1 \in x_{ij'}\right)\right) + \lambda R(X_1 \in x_{ij'}, q_2^{n-1}),
\]

(5)

where \( p_1 \) and \( p_2 \) represent the transmission loss probabilities of each of the two descriptions.

If the interval \( x_{ij'} \) is a cell of the quantizer \( q_1^n \), then the quantity \( W_{ij}(j, j') \) is the contribution of the cell \( x_{ij'} \) to the Lagrangian cost function \( L(q_1, q_2) \). We use the conditional distribution of \( X_2 \) given \( X_1 \in x_{ij'} \) to compute the expected marginal and joint distortions and the rate. Figure 2 illustrates how to calculate the statistics of the reconstruction of \( X_2 \) for fixed \( q_2^{n-1} \) and given \( X_1 \in x_{ij'} \). To be more specific, for the squared error distortion measure, the weights are explicitly calculated in the Appendix.

Now, one can easily see that the problem of designing optimal \( q_1^n \) for fixed \( q_2^{n-1} \) is equivalent to finding the shortest path from node 0 to node \( N \) of graph \( G \) for an appropriate \( \lambda \) value. Given the value of \( \lambda \), this shortest path problem can be solved in \( O(N^2) \) time. The targeted transmission rate \( r \) can be met through a binary search on \( \lambda \).

It should be noted that this procedure will only produce quantizers of connected cells. For a DSC problem where only the performance of the system with both descriptions available is important, limiting the design to those quantizers with connected cells might significantly reduce the performance [4]. On the other hand, when acceptable marginal constructions are required, each individual quantizer will also be close to an optimal quantizer marginally designed for the source, which necessarily requires quantizers of connected cells.

4. Achievable RDSC Performance for Jointly Gaussian Sources

The problem of RDSC for symmetric Gaussian sources is considered in [1, 2]. Given \( n \) sensors that observe noisy versions of a single source \( X \), an achievable distortion-rate performance in reconstruction of \( X \) is derived when any subset of the sensors data are available. We adopt a simple theoretical formulation more suited to our framework in studying achievable distortion-rate performance. In formulations of [1, 2], the goal is the reconstruction of a specific source \( X \) when the sensors observe noisy versions of \( X \). However, in our case, the sensors are to sense different but correlated sources. This necessitates a slight paradigm shift in our formulation. We use known results on classical multiple description source coding of Gaussian sources in [11] to derive achievable distortion-rate performances for a two-dimensional jointly Gaussian RDSC problem. This distortion-rate analysis is useful in evaluating the performance of scalar quantization results in Section 3. We report our results for a symmetric case only, although a general derivation is possible. Also, these theoretical results would only apply to the case where both sensor measurements are known to both encoders.

The region of all achievable distortion-rate pairs for MDC of a Gaussian source is known [11]. For symmetric descriptions of a Gaussian source of variance \( \sigma^2 \), the result of [11] can be written as

\[
d = d_1 = d_2 \geq \sigma^2 2^{-2R} \\
d_0 \geq \left(1 - \frac{\sigma^2 2^{-4R}}{\left[\left|\left|1 - d/\sigma^2 - \sqrt{d^2/\sigma^4 - 2^{-4R}}\right|\right|^2\right]^2}\right).
\]

(6)

where \( |x|^+ = x \) if \( x > 0 \) and is zero otherwise.

Now, consider two jointly Gaussian random variables \( X_1 \) and \( X_2 \), each of which has zero mean and variance \( \sigma^2 \). The joint distribution of these two sources is completely determined by their correlation coefficient \( \rho = E(X_1 X_2) \). These two sources can be written as the sum of
two uncorrelated and hence independent Gaussian sources each of zero mean as \( X_1 = (T + U)/\sqrt{2} \) and \( X_2 = (T - U)/\sqrt{2} \), where \( T \) and \( U \) are the DC and AC components of \( X_1 \) and \( X_2 \). Also, \( \sigma^2_T = 1 + \rho \), \( \sigma^2_U = 1 - \rho \). To describe \( X_1 \) or \( X_2 \) with \( R \) bits, one can equivalently describe their DC and AC components. Applying the results in [11] to \( T \) and \( U \) independently, we arrive at the following achievable distortion-rate performance:

\[
\begin{align*}
    d &= d_T + d_U \\
    R &= R_T + R_U \\
    d_T &\geq (1 + \rho) 2^{-2R_T} \\
    d_U &\geq (1 - \rho) 2^{-2R_U} \\
    d_0 &= d_{x_1} + d_{x_2} \\
    &\geq \frac{(1 + \rho) 2^{-4R_T}}{1 - \left(\frac{d_T}{(1 + \rho) - \sqrt{d_T^2/(1 + \rho)^2 - 2^{-4R_T}}}\right)^2} \\
    &\quad + \frac{(1 - \rho) 2^{-4R_U}}{1 - \left(\frac{d_U}{(1 - \rho) - \sqrt{d_U^2/(1 - \rho)^2 - 2^{-4R_U}}}\right)^2}.
\end{align*}
\]  

Note that \( d \) and \( d_0 \) are the average distortions of reconstructing both sources \( X_1 \) and \( X_2 \). It should once again be emphasized that the coding assumes the values of both sources \( X_1 \) and \( X_2 \) are known at both encoders. A closed form achievable bound on \( d \) can be found at high rates. We derive one such bound by assuming the marginal distortions to be small but still much larger than the smallest marginal distortion promised by their distortion-rate functions. More precisely, we assume \( 2^{-2R} \ll d \ll 1 \). This assumption is usually satisfied for practically meaningful designs considered in the next section. Physically, the assumption promises a small distortion given any of the descriptions alone, while allowing for a large improvement when both descriptions are available. Under this assumption, the minimum achievable joint distortion from (7) is

\[
d_{0,\text{min}} = \frac{(1 - \rho^2) 2^{-2R}}{d}
\]  

for any given marginal rate \( R \) and distortion \( d \). The joint distortion \( d_0 \) in (7) can then be written as

\[
d_0 = \frac{(1 + \rho)^2 2^{-4R_T}}{4d_T} + \frac{(1 - \rho)^2 2^{-4R_U}}{4d_U}.
\]

Minimizing \( d_0 \) under the constraints \( d_T + d_U = d \) and \( R_T + R_U = R \) requires \( d_T = d_U = d/2 \) and \( R_T - R_U = 0.5 \log_2((1 + \rho)/(1 - \rho)) \). Equation (8) follows from inserting these relations back into (9).

5. Experimental Results

In this section, we investigate the performance of our proposed system and report improvements over a naive scheme. When the two descriptions are equally important, that is, \( \lambda_1 = \lambda_2 \), the minimization in (3) can be written in the following equivalent form:

\[
\min_{q_1,q_2} L' (q_1,q_2) = \frac{(1 - \rho) ((1 - \rho) d_0 + \rho d_1 + \rho d_2) + \lambda R.}
\]

In this new form, the parameter \( p \) has a precise meaning as follows. Assume that the two descriptions are transmitted through two independent channels. The probability that each of the two transmissions fails is \( \rho \). Then, \( L' \) in (10) is the expected distortion of the reconstructed sources \( \hat{X}_1 \) and \( \hat{X}_2 \). The relative importance of the marginal and joint distortions is reflected in parameter \( \rho \). For relatively large failure rates \( \rho \) (poor channels), one would expect the two quantizers \( q_1 \) and \( q_2 \), for both sources, to be close to an optimal quantizer for a single Gaussian. The quantizer cells of \( q_1 \) and \( q_2 \) are however somewhat interleaved to reduce the distortion when both quantization outputs are available at the decoder. An example of a pair of optimized quantizers for \( \rho = 0.72 \), \( p = 0.30 \), and \( R = 2.27 \) is shown in Figure 3. While other choices are possible as discussed in Section 3, to calculate the overall communication rate throughout this section, the Slepian-Wolf bound is assumed to be achieved; that is, we assume \( R(Q_1, Q_2) = H(Q_1) + H(Q_2) | Q_1) \).

The joint and marginal distortions for \( p = 0.30 \), \( \rho = 0.72 \), and different marginal rates are depicted in Figure 4. The distortion of the optimal quantizer for a single Gaussian source and the same rate, obtained from the Lloyd-Max algorithm, is also shown in the same figure for comparison. Evidently, the difference between the two is extremely small. In other words, the penalty in the performance of marginal reconstruction is negligible, when jointly designing the quantizers for two sources versus designing the quantizers for each source independently. The distortion-rate performance of the proposed distributed coding scheme, consisting of scalar quantizers and entropy coders, is compared against theoretically achievable performances given by (8). The comparison is carried out in the following manner. For fixed \( p \) and \( \rho \), the quantizers are optimally designed for a target marginal distortion. This design results in a marginal bit rate and a joint distortion, which is then plotted as a function of this marginal rate and is compared to what is theoretically achievable for the same marginal rate and distortion.

Figure 4 also demonstrates the joint distortion achieved with a naive quantization scheme. This quantization scheme

![Figure 3: Quantizers optimally interleaved to reduce the distortion when both quantized values are present, while each of them is close to an optimal quantizer for a single Gaussian source.](image-url)
is represented as "Lloyd-joint" and is based on independently designing an optimal quantizer for $X_1$ and $X_2$ having the same number of cells as $q_1$ and $q_2$ using the Lloyd-Max algorithm. Our proposed method exploits the correlation between the sources to improve the expected distortion. As is evident from the figure, the improvements at low rates are remarkable.

6. Concluding Remarks

This paper considered the practice of RDSC in the context of optimally designed scalar quantizers. A practically convenient coding scheme, consisting of a pair of quantizers and discrete lossless coders to robustly encode two correlated sensor measurements, is proposed in this paper. When the particular form of the lossless coders is chosen, we devised an optimization algorithm for designing the quantizers. It is then shown that a proper choice of the quantizers can significantly increase the overall system performance.

Sensors are assumed to be measuring some random field; therefore, aside from temporal correlations between the measurements of a sensor, a great number of correlations can exist between the measurements of nearby sensors too. Such correlations can be exploited to reduce the amount of communication required to report the sensor measurements to a server.

Appendix

Weights of the Directed Acyclic Graph

One has the following:

$$W_{1n}(j, j) = \left[ \int_{x_{1j} + \Delta}^{x_{1j+1}} \int_{x_{2j} + \Delta}^{x_{2j+1}} d_{x_1} d_{x_2} f_{X_1/X_2}(x_1, x_2) \right]$$

$$+ \sum_{k=0}^{M-1} \sum_{i=1}^{q_{i+1}^j} (1 - p_1)(1 - p_2) \left\| x - \hat{x}_{jjk} \right\|^2 \right]$$

$$- \lambda P_{jj'} \log_2 \frac{P_{jj'}}{P_{jj''}}$$

$$\hat{x}_{jj'k} = \int_{q_{k}^j}^{q_{k+1}^j} \int_{x_{1j} + \Delta}^{x_{1j+1}} \int_{x_{2j} + \Delta}^{x_{2j+1}} d_{x_1} d_{x_2} P_{X_1/X_2}(x_1, x_2) \cdot \frac{P_{jj'}}{P_{jj''}}$$

$$P_{jj'} = \sum_{k=0}^{M-1} P_{jj'k},$$

$$P_{jj'k} = \sum_{i=1}^{q_{i+1}^j} \int_{x_{1j} + \Delta}^{x_{1j+1}} \int_{x_{2j} + \Delta}^{x_{2j+1}} d_{x_1} d_{x_2} P_{X_1/X_2}(x_1, x_2).$$

(A.1)

$P_{X_1,X_2}(x_1, x_2)$ is the joint probability density function of $X_1$ and $X_2$. In the next iteration for finding $W_{2n}$, one simply replaces the role of indices 1 and 2.

Competing Interests

The authors declare that they have no competing interests.

References

[1] J. Chen and T. Berger, "Robust distributed source coding," IEEE Transactions on Information Theory, vol. 54, no. 8, pp. 3385–3398, 2008.
[2] P. Ishwar, R. Puri, S. S. Pradhan, and K. Ramchandran, "On compression for robust estimation in sensor networks," in Proceedings of the IEEE International Symposium on Information Theory, p. 193, June-July 2003.
[3] A. Aaron and B. Girod, "Compression with side information using turbo codes," in Proceedings of the Data Compression Conference (DCC '02), pp. 252–261, Snowbird, Utah, USA, 2002.
[4] D. Rebollo-Monedero, Quantization and transforms for distributed source coding [Ph.D. thesis], Stanford University, 2007.
[5] A. Saxena and K. Rose, "On scalable distributed coding of correlated sources," IEEE Transactions on Signal Processing, vol. 58, no. 5, pp. 2875–2883, 2010.
[6] A. B. Wagner, B. G. Kelly, and Y. Altug, "Distributed rate-distortion with common components," IEEE Transactions on Information Theory, vol. 57, no. 7, pp. 4035–4057, 2011.
[7] N. Wernersson, J. Karlsson, and M. Skoglund, "Distributed quantization over noisy channels," IEEE Transactions on Communications, vol. 57, no. 6, pp. 1693–1700, 2009.
[8] J. Z. Sun and V. K. Goyal, "Intersensor collaboration in distributed quantization networks," IEEE Transactions on Communications, vol. 61, no. 9, pp. 3931–3942, 2013.
[9] S. Dumitrescu and X. Wu, "Lagrangian optimization of two-description scalar quantizers," IEEE Transaction on Information Theory, vol. 53, no. 11, pp. 3990–4012, 2007.
[10] V. A. Vaishampayan, “Design of multiple description scalar quantizers,” *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 821–834, 1993.

[11] L. H. Ozarow, “On a source-coding problem with two channels and three receivers,” *The Bell System Technical Journal*, vol. 59, no. 10, pp. 1909–1921, 1980.
