On Efficient Quantizer Design for Robust Distributed Source Coding *

Ankur Saxena[1], Jayanth Nayak[2] and Kenneth Rose[1]

[1]ECE Department, University of California, Santa Barbara, CA 93106, USA.

[2]IRISA/INRIA, Campus Universitaire de Beaulieu, 35042, Rennes, France.

Email:{ankur, rose}@ece.ucsb.edu, jnayak@irisa.fr

Abstract

This paper considers the design of efficient quantizers for a distributed source coding system. The information is encoded at independent terminals and transmitted across separate channels, any of which may fail. The scenario subsumes a wide range of vector quantization problems. Greedy descent methods rely heavily on initialization, and the presence of numerous 'poor' local optima on the distortion cost surface strongly motivates the use of a global design algorithm. We propose a deterministic annealing approach for the design of all components of a generic distributed source coding system. Our approach avoids many poor local optima, is independent of initialization, and does not assume any prior information on the underlying source distribution. Simulation results show significant gains over an iterative greedy algorithm.

1 Introduction

In a distributed network of sensors, different sensors may be designed to observe various physical quantities, e.g., temperature, humidity, pressure. We may be interested in efficient reconstruction of one or more physical entities at different spatially separated observatories. Typically, the data communicated by networks of sensors exhibits a high degree of correlation. Since the encoders at each sensor location function independently the system will not, in practice, accomplish the maximum possible lossy compression of the sources. Another prominent issue is that of estimation of a source from another correlated source. For example, if a sensor (or a transmission channel) breaks down, then to get a reliable estimate of the data being (or that would be) measured by that sensor, we can only utilize information acquired from the other sensors (channels). To achieve the dual objectives of obtaining the best possible compression efficiency from independent encoders and attaining system robustness, it is necessary that the code *design* at all the terminals be performed jointly for a robust distributed source coding system (see Fig. 1).

*This work was carried out while J. Nayak was with the University of California, Santa Barbara. The work is supported in part by the NSF under grant IIS-0329267, the University of California MICRO program, Applied Signal Technology, Inc., Dolby Laboratories Inc., and Qualcomm Inc.

Proceedings of the Data Compression Conference (DCC'06)
0-7695-2545-8 /06 $20.00 © 2006 IEEE

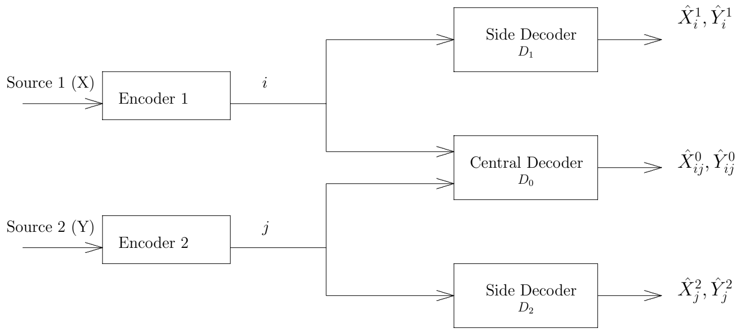

Figure 1: Block Diagram for Robust Distributed Source Coding

The robust distributed source coding model was first proposed and studied in [6] and later in [3] and [4]. As pointed out in [4], the model subsumes a variety of source coding problems ranging from distributed source coding [12, 14], the CEO problem [1] to multiple description coding. A good design for this system should be able to take into account the correlation between the sources as well as the possibility of component failure. An interesting approach to distributed source coding employing source codes (scalar quantization or trellis coded quantization) followed by channel codes (scalar or trellis based coset codes) was presented in [9] and a recent extension utilizing irregular LDPC codes for channel codes has appeared in [15]. These channel coding approaches can conceivably be extended to robust distributed vector quantizer (RDVQ) design. However, current channel coding approaches appear most suitable when the sources can be modeled as noisy versions of each other, where the noise is unimodal in nature. But such approaches are of limited use where the abovementioned simplifying assumptions do not apply. On the other hand, approaches based on the Lloyd-algorithm [8] to design RDVQ will suffer from the presence of numerous 'poor' local minima on the distortion-cost surface and thus will be critically sensitive to initialization. Clever initialization as proposed, for example, in the context of multiple description scalar quantizer design [13], can help mitigate this shortcoming. But such initialization heavily depends on symmetries or simplifying assumptions, and no generalizations are available to vector quantization nor to more complicated scenarios such as RDVQ. Alternatively, a global optimization scheme i.e., a powerful optimization tool that provides the ability to avoid poor local optima, and is applicable to sources exhibiting any type of statistical dependencies, such as deterministic annealing can overcome these issues.

In [7], it has been shown that a deterministic annealing based approach can gain substantially over extensions of the Lloyd-algorithm and various schemes employing heuristic initialization for the case of generic multiple description vector quantizer design. Numerous other applications where deterministic annealing performs better than Lloyd-based iterative algorithms can be found in a tutorial paper [11] and references therein. In this paper, an iterative greedy algorithm for RDVQ design is described which will underline the need of a global optimization approach. We then propose a deterministic annealing approach for optimal RDVQ design.

Deterministic annealing (DA) is motivated by the process of annealing in physics. It is independent of the initialization, does not assume any knowledge about the underlying source distribution and avoids many poor local minima of the distortion-cost surface [11]. In DA, a probabilistic framework is introduced for the encoding rule where each training sample of the input source is assigned to a reproduction value *in probability*. The optimization problem is recast as minimization of the expected distortion subject to a constraint on the level of randomness as measured by the Shannon entropy of the system. The Lagrangian functional can be viewed as the free energy of a corresponding physical system and the Lagrangian parameter as the 'temperature'. The minimization is started at a high temperature (highly random encoder) where, in fact the entropy is maximized and hence all reproduction points are at the centroid of the source distribution. The minimum is then tracked at successively lower temperatures (lower levels of entropy), by re-calculating the optimum locations of the reproduction points and the encoding probabilities at each stage. As the temperature approaches zero, the average distortion term dominates the Lagrangian cost and a hard (non-random) encoder is obtained.

The rest of the paper is organized as follows. In Section 2, we state the problem formally, establish the notation and describe an iterative greedy method based on Lloyd's algorithm for multiple prototype coder design. This will underline the need of a global approach. In Section 3, we derive the DA approach to RDVQ design and provide its update formulae (necessary optimality conditions). Simulation results are given in Section 4, followed by the conclusions in Section 5.

2 Problem Statement and Iterative Greedy Methods

Consider the robust distributed source coding scenario in Fig. 1. For brevity, we will restrict the analysis to the case of two sources, but the model can be extended in a straightforward fashion to an arbitrary number of sources. Here (X, Y) is a pair of continuous-valued, i.i.d., correlated (possibly vector) sources which are independently compressed at rates R_1 and R_2 bits per sample, respectively. The encoded indices i and j are transmitted over two separate channels, which may or may not be in working order, and the channel condition is not known at the encoders. The end-user tries to obtain the best estimate of the sources depending on the descriptions received from the functioning channels. $(\hat{X}^0_{ij}, \hat{Y}^0_{ij})$ denotes the reconstruction values for (X, Y) that are produced by the central decoder D_0, i.e., when information is available from both the channels. If only channel 1 (or 2) is working, then side decoder D_1 (or D_2) is used to reconstruct $(\hat{X}^1_i, \hat{Y}^1_i)$ (or $(\hat{X}^2_j, \hat{Y}^2_j)$). The objective of the robust distributed vector quantizer (RDVQ) is to minimize the following overall distortion function given rate allocations of R_1 and R_2:

$$E[\{\alpha_0 d(X, \hat{X}^0_{ij}) + (1-\alpha_0) d(Y, \hat{Y}^0_{ij})\} + \lambda_1\{\alpha_1 d(X, \hat{X}^1_i) + (1-\alpha_1) d(Y, \hat{Y}^1_i)\} \\ + \lambda_2\{\alpha_2 d(X, \hat{X}^2_j) + (1-\alpha_2) d(Y, \hat{Y}^2_j)\}] \quad (1)$$

Proceedings of the Data Compression Conference (DCC'06)
0-7695-2545-8 /06 $20.00 © 2006 IEEE

where $d(\cdot,\cdot)$ is an appropriately defined distortion measure and $\alpha_n \in [0,1]$ $\{n = 0,1,2\}$ governs the relative importance of the sources X and Y at the n^{th} decoder. In the RDVQ cost of (1), the first two terms contribute to the central distortion when both the channels work. The side distortions are weighted by λ_1 and λ_2 whose specific values depend on the importance we wish to give to the side distortions as compared to the central distortion. In a practical system, λ_1 and λ_2 will often be determined by the channel failure probabilities.

The RDVQ problem comprises the design of mappings from the sources X and Y to indices at the respective encoders and of the corresponding reconstruction values at the three decoders. To minimize the distortion for given transmission rates, the correlation between the sources needs to be exploited. This can be done by sending the same index for different non-contiguous regions of the source alphabet on a channel and then using the information from the other source to distinguish between the different regions. In the case when only one channel is functioning, the RDVQ problem reduces to estimating a signal from another correlated source. On the other hand, if both the channels work and the central decoder is used, the problem reduces to that of correlated source coding. Locally optimal quantizer design techniques for general networks (which encompass the RDVQ model as well) and correlated source coding have been proposed in the literature in [5] and [2, 10] respectively. We next adopt this framework and describe a locally optimum algorithm using multiple-prototypes (MP) for the design of a generic RDVQ system. The MP approach can be viewed as combining histogram or kernel based techniques for estimating source distributions and quantizer design.

Specifically, we have a training set $\mathcal{T} \equiv \{\mathcal{X}, \mathcal{Y}\}$, which consists of m-dimensional i.i.d. vectors. We design a coarse vector quantizer Q_1 for X using a standard VQ design algorithm such as Lloyd's algorithm [8] or DA [11]. Q_1 assigns training set data points to one of the $\mathcal{K}$ regions, C_k^x. The disjoint Voronoi regions C_k^x span the space and a prototype x_k is associated with each of them. Next, each Voronoi region is mapped to one of the $\mathcal{I}$ indices, via a mapping $v(k) = i$, to which we refer as Slepian-Wolf (SW) mapping, since this mapping is the module that in fact exploits inter-source correlation and attempts to approach the theoretical bound promised by [12]. The index i is then transmitted across the channel. An example of SW mapping with $m = 1$, $\mathcal{K} = 8$ and $\mathcal{I} = 3$, is given in Fig. 2. The region associated with index i is denoted $R_i^x = \bigcup_{k;v(k)=i} C_k^x$.

We similarly define the quantizer Q_2, regions C_l^y, R_j^y and prototypes y_l in the Y domain. Here, the $\mathcal{L}$ Voronoi regions are mapped to $\mathcal{J}$ indices via SW mapping $w(l) = j$. At the central decoder, we receive indices in $\mathcal{I} \times \mathcal{J}$, and generate reconstruction values $\hat{x}_{ij}^0$ and $\hat{y}_{ij}^0$. At the side decoder 1 (or 2), the received index is in $\mathcal{I}$ ($\mathcal{J}$), and reconstruction values are $\hat{x}_i^1$ ($\hat{x}_j^2$) and $\hat{y}_i^1$ ($\hat{y}_j^2$).

In the sequel, we will often use the following concise notation

$$D_x(x, \hat{x}_{ij}^0, \hat{x}_i^1, \hat{x}_j^2) = \alpha_0 d(x, \hat{x}_{ij}^0) + \lambda_1 \alpha_1 d(x, \hat{x}_i^1) + \lambda_2 \alpha_2 d(x, \hat{x}_j^2), \quad (2)$$

$$D_y(y, \hat{y}_{ij}^0, \hat{y}_i^1, \hat{y}_j^2) = (1-\alpha_0) d(y, \hat{y}_{ij}^0) + \lambda_1 (1-\alpha_1) d(y, \hat{y}_i^1) + \lambda_2 (1-\alpha_2) d(y, \hat{y}_j^2), \quad (3)$$

$$D_{net} = D_x(x, \hat{x}_{ij}^0, \hat{x}_i^1, \hat{x}_j^2) + D_y(y, \hat{y}_{ij}^0, \hat{y}_i^1, \hat{y}_j^2), \quad (4)$$

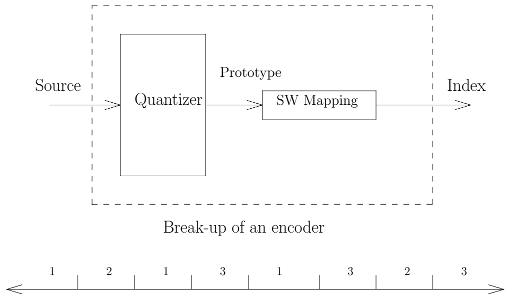

Figure 2: Block diagram of an encoder and an example of SW mapping from prototypes (Voronoi regions) to indices.

to denote the partial distortion terms corresponding to X and Y, and the net expected distortion D_{net} of (1) which we seek to minimize. To minimize D_{net}, the SW mappings v and w, as well as the reconstruction values at various decoders (all of which have been 'randomly' initialized) are optimized iteratively using the following steps:

1. **SW Mapping for X**: For $k = 1, \ldots, \mathcal{K}$, assign region k to index i, such that:

$$v(k) = i = \arg\min_{i'} \sum_{\substack{(x,y)\in\mathcal{T};\\ x\in C_k^x}} [D_x(x, \hat{x}^0_{i'j(y)}, \hat{x}^1_{i'}, \hat{x}^2_{j(y)}) + D_y(y, \hat{y}^0_{i'j(y)}, \hat{y}^1_{i'}, \hat{y}^2_{j(y)})]. \quad (5)$$

2. **SW Mapping for Y**: For $l = 1, \ldots, \mathcal{L}$, assign region l to index j, such that:

$$w(l) = j = \arg\min_{j'} \sum_{\substack{(x,y)\in\mathcal{T};\\ y\in C_l^y}} [D_x(x, \hat{x}^0_{i(x)j'}, \hat{x}^1_{i(x)}, \hat{x}^2_{j'}) + D_y(y, \hat{y}^0_{i(x)j'}, \hat{y}^1_{i(x)}, \hat{y}^2_{j'})]. \quad (6)$$

3. **Reconstruction Values**: For all $i = 1 : \mathcal{I}$ and $j = 1 : \mathcal{J}$, find $\hat{x}^0_{ij}$, $\hat{x}^1_i$ and $\hat{x}^2_j$ such that:

$$\hat{x}^0_{ij} = \arg\min_{a_0} \sum_{\substack{(x,y)\in\mathcal{T};x\in R_i^x,\\ y\in R_j^y}} d(x, a_0), \quad (7)$$

$$\hat{x}^1_i = \arg\min_{a_1} \sum_{(x,y)\in\mathcal{T};x\in R_i^x} d(x, a_1), \quad (8)$$

$$\hat{x}^2_j = \arg\min_{a_2} \sum_{(x,y)\in\mathcal{T};y\in R_j^y} d(x, a_2). \quad (9)$$

The corresponding equations for the reconstruction values of Y have not been reproduced here, but are trivially obtained by symmetry.

We again re-emphasize that it is the SW module that exploits the correlation between the quantized versions of source. The above technique optimizes the SW mappings from prototypes to indices for X and Y, and the final reconstruction values at the various decoders in an iterative manner. We will thus refer to the above design algorithm as the Lloyd Approach (LA). LA inherits from the original Lloyd algorithm the inter-related shortcomings of getting trapped in poor local minima, and dependence on initialization. The sub-optimality of LA will be observed experimentally in the results section. These issues call for the use of a global optimization scheme, such as DA. We next present the DA algorithm and the necessary conditions for optimality in RDVQ design.

3 The Deterministic Annealing Approach

3.1 Derivation

A formal derivation of the DA algorithm is based on facts borrowed from information theory and statistical physics. Here the deterministic encoder is replaced by a random encoder. Detailed derivations and the principles underlying basic DA can be found in [11].

Given the RDVQ setup, we first design quantizers Q_1 and Q_2 for the sources independently using DA [11]. The rationale for this separate design is that as long as the number of prototypes per index is large, then the correlation between the quantized versions of the sources can be fully exploited within the SW mapping modules of the encoders. This means that efficient SW mappings from prototypes to indices is crucial for the overall system performance. The RDVQ DA approach optimizes these mappings and the reconstruction values jointly, is independent of the initialization and will converge to a considerably better minimum.

The coarse quantizer Q_1 for source X assigns each data point in the training set for the source X to a prototype x_k. We define:

$$c_{k|x} = \begin{cases} 1 & \text{if } Q_1(x) = k \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

and $r_{i|k} = \Pr[R_i^x | x_k] = \Pr[x_k \in R_i^x]$ as the probability of a prototype being mapped to an index. Similarly in the Y domain, we define:

$$c_{l|y} = \begin{cases} 1 & \text{if } Q_2(y) = l \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

and $r_{j|l} = \Pr[R_j^y | y_l] = \Pr[y_l \in R_j^y]$.

The probabilistic equivalent of the net distortion function D_{net} in (1) which we seek to minimize is:

$$D = \frac{1}{N} \sum_{(x,y)\in\mathcal{T}} \sum_{k,l,i,j} c_{k|x} c_{l|y} r_{i|k} r_{j|l} [D_x(x, \hat{x}_{ij}^0, \hat{x}_i^1, \hat{x}_j^2) + D_y(y, \hat{y}_{ij}^0, \hat{y}_i^1, \hat{y}_j^2)], \tag{12}$$

subject to a constraint on the *joint entropy* of the system. Here N is the number of data points in the training set.

The joint entropy of the system is $H = H(X, Y, I, J) = H(X, Y) + H(I|X) + H(J|Y)$, since by construction, the source variables X and Y and the transmitted indices I and J form a Markov chain: $J - Y - X - I$. Also $H(X, Y)$ is the source entropy and is unchanged by the encoding decisions for a given training set. The expressions for the other terms in the joint entropy H are:

$$H(I|X) = \frac{-1}{N} \sum_{(x,y)\in\mathcal{T}} \sum_{k,i} c_{k|x} r_{i|k} \log(r_{i|k}), \tag{13}$$

$$\text{and} \quad H(J|Y) = \frac{-1}{N} \sum_{(x,y)\in\mathcal{T}} \sum_{l,j} c_{l|y} r_{j|l} \log(r_{j|l}). \tag{14}$$

So, the optimization of D in (12) subject to the joint entropy H is equivalent to the following Lagrangian minimization:

$$\min_{\{r_{i|k}\},\{r_{j|l}\},\{\hat{x}^0_{ij}\},\{\hat{y}^0_{ij}\},\{\hat{x}^1_i\},\{\hat{y}^1_i\},\{\hat{x}^2_j\},\{\hat{y}^2_j\}} \{L = D - TH\}. \tag{15}$$

where the temperature, T plays the role of Lagrange parameter. In the annealing process, we start from a high temperature and perform annealing at successively lower temperatures. At high temperature, all the reproduction points are at the centroid of the source distribution and a prototype is associated with all the indices with equal probability. As the temperature is lowered, the optimum locations of the reproduction points and encoding probabilities ($r_{i|k}$ and $r_{j|l}$) are calculated to track the minimum. At the limit of zero temperature, the encoding probabilities become *hard* and a deterministic encoding rule is obtained.

3.2 Update Equations for RDVQ Design

At a fixed temperature T, the objective function in (15) is a convex optimization problem in terms of $r_{i|k}$ and $r_{j|l}$. So the expressions for $r_{i|k}$ and $r_{j|l}$ can be analytically computed as:

$$r_{i|k} = \frac{e^{-D_{ki}/T}}{\sum_{i'} e^{-D_{ki'}/T}} \quad \text{and} \quad r_{j|l} = \frac{e^{-D_{lj}/T}}{\sum_{j'} e^{-D_{lj'}/T}} \tag{16}$$

where

$$D_{ki} = \frac{\sum_{(x,y)\in\mathcal{T},l,j} c_{k|x} c_{l|y} r_{j|l} D_{net}}{\sum_{(x,y)\in\mathcal{T},l,j} c_{k|x} c_{l|y} r_{j|l}} \quad \text{and} \quad D_{lj} = \frac{\sum_{(x,y)\in\mathcal{T},k,i} c_{k|x} c_{l|y} r_{i|k} D_{net}}{\sum_{(x,y)\in\mathcal{T},k,i} c_{k|x} c_{l|y} r_{i|k}}. \tag{17}$$

We next give the expressions for the reconstruction values in the case of squared-error distortion measure. The approach is clearly not restricted by this choice of distortion measure.

$$\hat{x}^0_{ij} = \frac{\sum_{(x,y)\in\mathcal{T},k,l} c_{k|x} c_{l|y} r_{i|k} r_{j|l} x}{\sum_{(x,y)\in\mathcal{T},k,l} c_{k|x} c_{l|y} r_{i|k} r_{j|l}}, \quad \hat{x}^1_i = \frac{\sum_{(x,y)\in\mathcal{T},k} c_{k|x} r_{i|k} x}{\sum_{(x,y)\in\mathcal{T},k} c_{k|x} r_{i|k}}, \quad \hat{x}^2_j = \frac{\sum_{(x,y)\in\mathcal{T},l} c_{l|y} r_{j|l} x}{\sum_{(x,y)\in\mathcal{T},l} c_{l|y} r_{j|l}} \tag{18}$$

Observe that these are the standard centroid rules. Also note that the side decoder 2 does not have access to X and the reconstruction of X is done based on the information received from its correlated source Y. By the symmetry in the problem, the decoding rules for Y can be trivially obtained, and will not be reproduced here.

4 Simulation Results

We give two examples to demonstrate the gains of the deterministic annealing approach over the iterative greedy method (LA). To avoid any potential fairness issues, we decided to design the coarse quantizers Q_1 and Q_2 using DA for both the competing approaches. This design could have been done using Lloyd's algorithm for the LA contender. However, Lloyd's algorithm may converge to a poor minimum in the quantizer design. The focus of the paper is on Slepian-Wolf mappings optimization (and reconstruction values) given the initial input quantizers, so we have chosen DA for initial quantizer design to eliminate its impact from the results.

In both the examples, X and Y are assumed to come from a joint gaussian source with respective means 0, variances 1 and correlation coefficient 0.9. The LA algorithm was run 20 times with different initializations while DA was run only once. A scalar RDVQ is designed using a training set of size 4000 in both the examples. The test set consists of 40000 data points. In the first case, the weighing parameters λ_1 and λ_2 for the side decoders are both taken as 0.01 while the rates R_1 and R_2 are 3 and 4 bits respectively. The number of prototypes for X and Y are 64 and 128 respectively. The parameters α_n $\{n = 0, 1, 2\}$ are taken as 0.5, 1 and 0 respectively, i.e., at the side decoder 1 (or 2), only X (or Y) is reconstructed while at the central decoder both the sources are given equal importance. The results are shown in Fig. 3. Here DA outperforms the best solution obtained by LA by $\sim$1.3 dB. The difference between the best and worst expected distortions of LA is itself $\sim$2.9 dB which exhibits the fact that iterative methods are heavily dependent on initialization and are highly likely to get trapped in a local minimum. For the test set, the net distortion obtained by the best LA versus single run DA was -15.18 and -15.95 dB, respectively.

In the second example, the weighing parameters λ_1 and λ_2 are taken to be 0.005 and 0.01 while the rates R_1 and R_2 are 2 and 3 bits respectively. The number of prototypes for both X and Y is 64. The parameters α_n $\{n = 0, 1, 2\}$ are all equal to 0.5 to give equal importance to each source at all the decoders. The results are shown in Fig. 4. The net distortion obtained for the test set for best LA versus single run DA was -12.06 and -13.08 dB, respectively.

5 Conclusions

We have proposed a multiple prototype based deterministic annealing approach for the design of quantizers for a robust distributed source coding system. The approach is general and is applicable to a wide gamut of vector quantization problems such as multiple descriptions, correlated source coding, CEO problem etc.. This

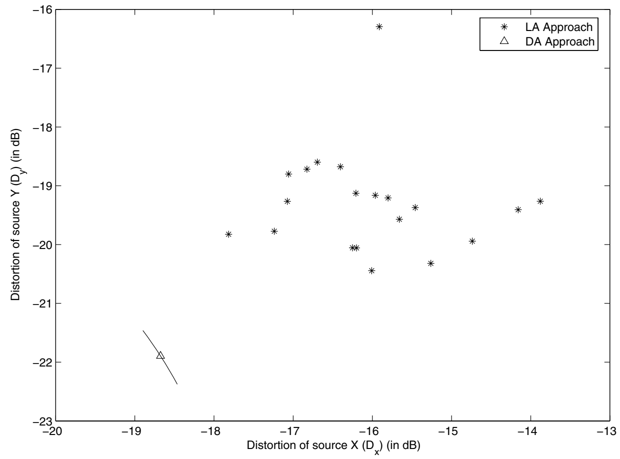

Figure 3: Comparison between LA and DA approaches for $R_1 = 3$, $R_2 = 4$, $\mathcal{K} = 64$, $\mathcal{L} = 128$, $\alpha_0 = 0.5$, $\alpha_1 = 1$, $\alpha_2 = 0$, $\lambda_1 = \lambda_2 = 0.01$. D_{net} from DA is -16.98 dB while LA gives best and worst D_{net} as -15.69 and -12.77 dB, respectively. For ease of comparison, a line along which constant D_{net} = -16.98 dB is drawn.

approach assumes no prior knowledge about the underlying probability distribution of the sources, eliminates the dependence on good initial configurations and avoids many poor local minima of the distortion cost surface. The necessary conditions (and update equations) for quantizer design are derived and presented. Simulation results comparing DA with an iterative Lloyd-like algorithm are shown. Significant improvements confirm the advantage of using a global optimization scheme such as DA for robust distributed vector quantizer design.

References

[1] T. Berger, Z. Zhang and H. Viswanathan, "The CEO problem," *IEEE Trans. on Information Theory*, vol. 42, pp. 887-902, May 1996.

[2] J. Cardinal and G. Van Assche, "Joint entropy-constrained multiterminal quantization," *Proc. IEEE ISIT,* p. 63, June 2002.

[3] J. Chen and T. Berger, "Robust coding schemes for distributed sensor networks with unreliable sensors," *Proc. IEEE ISIT* p. 115, June 27–July 2, 2004.

[4] J. Chen and T. Berger, "Robust distributed source coding," to appear in *IEEE Trans. on Information Theory.*

[5] M. Fleming, Q. Zhao and M. Effros, "Network vector quantization," *IEEE Trans. on Information Theory,* vol. 50, no. 8, pp. 1584-1604, Aug. 2004.

[6] P. Ishwar, R. Puri, S. S. Pradhan and K. Ramchandran, "On compression for robust estimation in sensor networks," *Proc. IEEE ISIT,* p. 193, June 29–July 4, 2003.

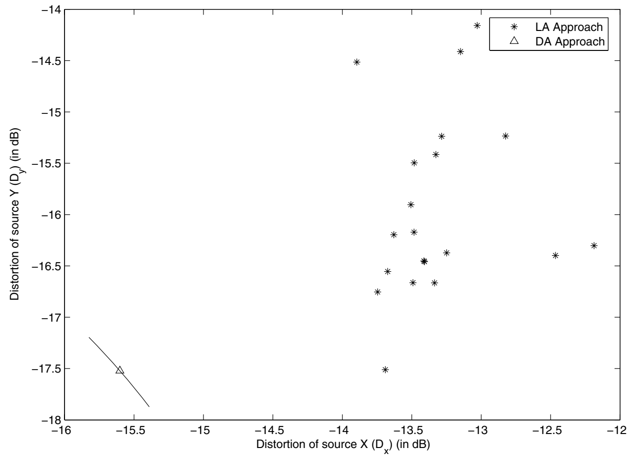

Figure 4: Comparison between LA and DA approaches for $R_1 = 2$, $R_2 = 3$, $\mathcal{K} = \mathcal{L} = 64$, α_n $\{n = 0, 1, 2\} = 0.5$, $\lambda_1 = 0.005$, $\lambda_2 = 0.01$. D_{net} from DA is -13.44 dB while LA gives best and worst D_{net} as -12.18 and -10.54 dB, respectively. For ease of comparison, a line along which constant D_{net} = -13.44 dB is drawn.

[7] P. Koulgi, S. L. Regunathan and K. Rose, "Multiple descriptions quantization by deterministic annealing," *IEEE Trans. on Information Theory,* vol. 49, no. 8, pp. 2067-2075, Aug. 2003.

[8] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. on Information Theory,* vol. 28, pp. 129-137, Mar. 1982.

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

[10] D. Rebollo-Monedero, R. Zhang and B. Girod, "Design of optimal quantizers for distributed source coding," *Proc. IEEE DCC,* pp. 13-22, Mar. 2003.

[11] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *Proc. IEEE,* vol. 86, no. 11, pp. 2210-2239, Nov. 1998.

[12] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. on Information Theory,* vol. 19, no. 4, pp. 471-480, July 1973.

[13] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. on Information Theory,* vol. 39, no. 3, pp. 821-834, May 1993.

[14] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. on Information Theory,* vol. 22, no. 1, pp. 1-10, Jan. 1976.

[15] Z. Xiong, A. Liveris, S. Cheng and Z. Liu, "Nested quantization and Slepian-Wolf coding: a Wyner Ziv coding paradigm for i.i.d. sources," *Proc. IEEE Workshop on Statistical Signal Processing,* pp. 399-402, Sept. 28–Oct. 1, 2003.