Optimization of Correlated Source Coding for Event-Based Monitoring in Sensor Networks *

Jaspreet Singh, Ankur Saxena, Kenneth Rose and Upamanyu Madhow
Dept. of Electrical & Computer Engineering
University of California, Santa Barbara, CA 93106, USA
Email:{jsingh, ankur, rose, madhow}@ece.ucsb.edu

Abstract

Motivated by the paradigm of event-based monitoring, which can potentially alleviate the inherent bandwidth and energy constraints associated with wireless sensor networks, we consider the problem of joint coding of correlated sources under a cost criterion that is appropriately conditioned on event occurrences. The underlying premise is that individual sensors only have access to partial information and, in general, cannot reliably detect events. Hence, sensors optimally compress and transmit the data to a fusion center, so as to minimize the *expected distortion in segments containing events*. In this work, we derive and demonstrate the approach in the setting of entropy constrained distributed vector quantizer design, using a modified distortion criterion that appropriately accounts for the joint statistics of the events and the observation data. Simulation results show significant gains over conventional design as well as existing heuristic based methods, and provide experimental evidence to support the promise of our approach.

1 Introduction

A typical Wireless Sensor Network (WSN) consists of a large number of sensor nodes deployed in order to gather sensory information and communicate it to a fusion center for further processing. Most practical applications involving WSN deployment (e.g., environmental/habitat monitoring) involve continuous sampling of the data field over extended periods of time, thereby generating an enormous amount of data that needs to be transmitted to the fusion center. However, the available resources (bandwidth, power) in typical WSN deployments are often severely constrained. Consequently, there is significant interest in devising data compression strategies that reduce the total volume of transmission without excessive compromise of the overall performance of the network. In particular, there has been much recent effort focussed on exploiting the correlations between the readings of proximate sensors to reduce the total data

*This work is supported in part by the National Science Foundation under grants CCF-0728986, IIS-0329267, CNS-0520335, and CCF-0729222, the University of California MICRO program, Applied Signal Technology Inc., Qualcomm Inc., and Sony Ericsson Inc.

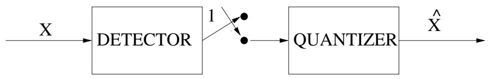


Figure 1: A heuristic encoder: hard detection is used to decide on event occurrences

flow [1, 2, 4, 5, 7, 8, 9]. These distributed coding methods are, in general, inspired by the seminal work of Slepian and Wolf [10] on lossless coding with side information at the decoder, and by the corresponding lossy coding results of Wyner and Ziv [11].

In this work, we propose a new paradigm for compressing sensor data, termed *event-based compression*. The approach is applicable even in the case of a single sensor, while in a distributed scenario, it can be employed to achieve compression gains over and above those achieved by exploiting the correlations between sensor readings. The basic premise underlying our approach is that *not all the data collected by a sensor node is equally important, rather, depending on the application, some portions of data are more important than others*. Hence, it should be possible to heavily compress the less important pieces of data, while still maintaining good quality transmissions for the important parts. For instance, consider a set of acoustic sensors deployed in a forest to monitor the activity of a particular bird species by listening to the calls that it makes. Clearly, the bulk of data collected by the sensor nodes in this scenario would practically be useless for us, since the frequency (duty cycle) of the bird call we are interested in would typically be very small. Given the severe resource constraints, it would therefore be an impractical luxury to blindly compress and transmit all the sensor data to the fusion center (whether or not the correlations across sensors are accounted for). Rather, we must, during compression, allocate the minimal available resources such that every sensor transmits more detailed information during an *event of interest* (in this case the bird call), as compared to the less critical periods where coarser information transmission would suffice.

An event-based compression scheme similar in principle to what we propose here has been studied earlier in [12], wherein a WSN is used to monitor structural vibrations in buildings. In this scheme, a simple threshold rule is applied to the sensor signal. If the signal is more than the threshold (i.e., if an event happens), it is quantized with a good resolution and transmitted, otherwise, a zero data value is sent (Fig. 1). Although such simple decision paradigms can work in certain situations, it is easy to see that they do not generalize to more complex settings, where it is highly non-trivial to correctly detect an event based on the sensor data. For instance, in our previous example of bird call monitoring using an acoustic sensor network, the strong interference caused by the sounds coming from numerous other birds and animals would make it hard to detect the presence of the specific bird we are interested in. It is easy to imagine how, in such circumstances, the simple binary detection rule will result in an unacceptable rate of missed events (leading to no transmission of actual event data), or wasteful false alarms (leading to unnecessary transmission of non-event data), or both. Thus, instead of the hard decision based heuristic framework of [12], a new and more nuanced approach that optimizes sensor data compression, while explicitly accounting for the *probability of event occurrences*, is needed.

The preceding arguments form the basic premise under which we work in this paper. Since events can not be detected reliably at the individual sensors, we instead develop a framework where sensors compress and transmit their data, so as to *minimize the expected distortion during events.* Specifically, we propose a distributed coding technique that is statistically conditioned on event occurrences, to compress the data recorded by different sensors. Thus, the contribution of this paper is twofold. First, the extension of compression to explicitly account for the probability of event occurrence (which is already applicable in the case of a single source) and, second, incorporation of such extensions within a system of correlated sources.

The rest of the paper is organized as follows: Section 2 provides the problem statement. In Section 3, we present the proposed approach, in terms of its underlying principles, and provide a derivation in the simplest case of event-based vector quantizer design for a single sensor. The generalization to multiple sensors is considered in Section 4. Section 5 consists of simulation results, followed by the conclusions in Section 6.

Notation: We denote random vectors/variables in uppercase, and their specific values are shown in lowercase. Bold faced notation denotes a collection of random vectors/variables (across sensors). $\mathbb{E}$ is the expectation operator.

2 Problem Formulation

Consider the setting of Fig. 2. We have M sensors, $\{S_1, \ldots, S_M\}$, recording (possibly correlated) data samples. Let the observation made by sensor S_i be represented by the random vector X_i, and let the collection of sensor observations be denoted by $\mathbf{X} = (X_1, \ldots, X_M)$. (For notational simplicity we dropped the time index). Each sensor compresses its data and transmits it to a fusion center, where a decoder utilizes the summaries received from all the sensors to generate the estimate $\hat{\mathbf{X}} = (\hat{X}_1, \ldots, \hat{X}_M)$. Let $d(\mathbf{X}, \hat{\mathbf{X}}) = \sum_i ||X_i - \hat{X}_i||^2$ be the (squared error) distortion measure. The rate of transmission from sensor S_i is assumed to be R_i bits, and the total rate is $R = \sum_i R_i$. It is assumed that none of the sensors has access to the data recorded by the others.

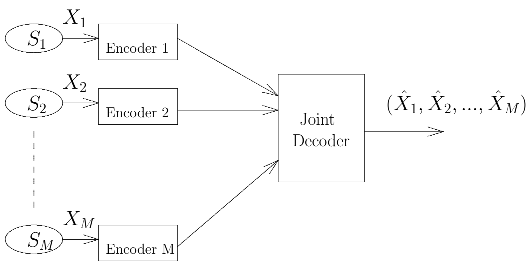


Figure 2: A Sensor Network

The data $\mathbf{X}$ recorded by the sensors at any particular time instant may correspond either to an event occurrence (denoted by E), or no event (denoted by $\bar{E}$). Note that we do not specify what the actual event(s) of interest is(are), since that would be specific to the application. Our aim is to optimize the system design such the *expected distortion during events* is minimized. Specifically, under a constraint on the total transmission rate R, we wish to design the sensor encoders and the joint decoder (Fig. 2), so as to minimize

$$\mathbf{D} = \mathbb{E}[d(\mathbf{X}, \hat{\mathbf{X}})|E] \ . \tag{1}$$

Minimization of the event conditioned distortion (1) would be easy to accomplish if it were possible to reliably detect event occurrences at the sensors, for the sensors could then drop the event-free data, and transmit the event data with the best fidelity achievable under the given rate constraint. However, in a realistic scenario, reliable event detection would not be possible, and hence, it would make sense to develop a *fuzzy* framework, wherein the fidelity of the transmission depends on the probability of event occurrence conditioned on the observed data. This observation is made more precise in the next section.

3 Optimizing the Event-Conditioned Distortion

We begin by considering the case of a single sensor, and illustrate the design ideas crucial to minimizing the expected distortion in event blocks. Multiple sensor scenario is studied in the next section, where correlations between the observations of different sensors are exploited to further enhance the performance.

Consider a sensor S that observes the random vector X. Let the Probability Density Function (pdf) of X, conditioned on the occurrence of an event be $f_{X|E}(x)$, and conditioned on no event occurrence be $f_{X|\bar{E}}(x)$. Further, let the a priori event probability be $P(E)$, so that the unconditional pdf of X is $f_X(x) = P(E)f_{X|E}(x) + (1 - P(E))f_{X|\bar{E}}(x)$. The expected distortion during events, then is

$$D = \mathbb{E}[d(X, \hat{X})|E] = \int d(x, \hat{x}) f_{X|E}(x) dx. \tag{2}$$

Thus, the system must be designed to minimize (2). A simple modification in the above expression will provide additional insight into how such a design will differ from a conventional design that minimizes the unconditional distortion $\mathbb{E}[d(X, \hat{X})] = \int d(x, \hat{x}) f_X(x) dx$. We can rewrite (2) as

$$D = \frac{1}{P(E)} \int [P(E|x) d(x, \hat{x})] f_X(x) dx \ , \tag{3}$$

where $P(E|x)$ is the posterior probability of event given the observation x. The use of brackets in this equation is meant to emphasize that the design can still be done in a conventional manner (i.e., over the unconditional density of X), but instead of using the distortion measure $d(x, \hat{x})$, we must use a *weighted distortion measure*, $P(E|x)d(x, \hat{x})$. Weighing the distortion by the posterior event probability would automatically ensure that the designed system provides better reconstruction of event segments, as compared to non-events. In a sense, we can say that the operation of detection at a particular sensor would be automatically embedded into the system design in a soft manner, thereby eliminating the need for an explicit hard binary detector.

The weighted distortion measure proposed above, establishes, in concrete terms, the *fuzzy* framework we had intuitively thought of in the last section. Since reliable event detection may not be possible, it is best to design the system such that the probability of event occurrence conditioned on the observation is accounted for. Using a distortion measure that is weighted by this probability ensures that it is accounted for in an optimal manner, for it minimizes the ultimate cost criterion of expected distortion during the events.

Having understood the central design ideas for event-based compression, we now demonstrate the approach in the setting of entropy constrained vector quantizer design at each sensor. We first focus on a single sensor scenario.

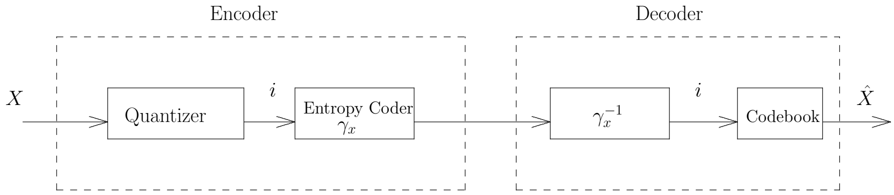


Figure 3: Proposed System in a single sensor case

3.1 Event-Based Vector Quantizer Design for Single Sensor

The system block diagram for an entropy constrained vector quantizer design is depicted in Fig. 3. The source signal X is input to the quantizer Q, whose output index i is fed to an entropy coder γ_x (we will use Huffman code as an entropy code in this work). At the decoder, we have an inverse entropy coder module followed by a codebook to generate the reconstruction $\hat{X}$. Under a given rate constraint, our aim is to jointly optimize the quantizer, the entropy coder, and the reconstruction codebook so as to minimize the expected distortion during events. The problem can alternatively be posed as minimization of the following Lagrangian cost function

$$\mathbb{E}[d(X, \hat{X})|E] + \lambda R \ , \tag{4}$$

where different choices of λ would provide different points in the rate-distortion trade-off. Denoting the length of the entropy coder output by $|\gamma_x(i)|$, we have $R = \mathbb{E}(|\gamma_x(i)|)$. Using the insight from (3), the Lagrangian cost can now equivalently be written as

$$\mathbb{E}[P(E|X)d(X, \hat{X}) + \lambda|\gamma_x(i)|] \ . \tag{5}$$

Note that we have ignored the term $P(E)$ (the prior probability of event) occurring in the denominator of (3), since it is a constant than can be subsumed in the parameter λ.

To design our system, we need to know the source density $f_X(x)$. In a practically deployed sensor network scenario, this density can be estimated by collecting a training data set. Hence, in the sequel, we assume that we have a training set $\mathcal{T}$ which comprises of N independent identically distributed (i.i.d.) observations of the source X. Further, each data point x in $\mathcal{T}$ has been assigned a label $P(E|x)$ to indicate the probability with which an event could have happened. This labeling can be done by an expert who is finally going to analyze the data that has been collected. For instance, in our bird call monitoring example, an Ecologist can assign these labels to the collected training data, depending on what portions of the data he considers to be more important than others.

Given the training set $\mathcal{T}$ and the vector of corresponding labels $P(E|x)$, we now propose an iterative Lloyd-like design strategy to minimize the Lagrangian cost in (5). Our algorithm is a generalization of the algorithm given in [3], in that it appropriately incorporates the conditioning on event occurrences. We denote by R_i^x the region corresponding to index i in the source space for X. The update rules for encoding, codebook update and entropy coder will then be as follows:

1. **Encoder Update**: For all points $x \in \mathcal{T}$, assign x to index i such that:

$$i = \arg\min_{i'} [P(E|x)d(x, \hat{x}_{i'}) + \lambda|\gamma_x(i)|]. \tag{6}$$

2. **Reconstruction Values for X**: For all $i = 1, \ldots, \mathcal{I}$ find $\hat{x}_i$ such that:

$$\hat{x}_i = \sum_{x \in \mathcal{T}; x \in R_i^x} \{x.P(E|x)\} \;/\; \sum_{x \in \mathcal{T}; x \in R_i^x} \{P(E|x)\} . \tag{7}$$

3. **Huffman Coder for X**: Design a Huffman coder based on the current probability distribution of the set of indices, $i = 1, \ldots, \mathcal{I}$.

4 Event-Based VQ Design in a Distributed Setting

In this section, we extend the event-based paradigm to the setting of multiple sensors (Fig. 2), wherein correlations between the sensor readings can be exploited to further enhance the performance. The idea of a weighted distortion measure, developed in section 3 for the case of a single sensor, can be extended directly to the multiple sensor scenario. Specifically, let the joint distribution of sensor observations (modeled by $\mathbf{X} = (X_1, \ldots, X_M)$), conditioned on the event occurrences, be $f_{\mathbf{X}|E}(\mathbf{x})$. Then, the expected distortion during events is $\mathbf{D} = \int d(\mathbf{x}, \hat{\mathbf{x}}) f_{\mathbf{X}|E}(\mathbf{x}) d\mathbf{x}$. This can equivalently be written in the form of (3), with the conditional probability of event occurrence given the observations, i.e. $P(E|\mathbf{x})$, again being used to arrive at a weighted distortion measure. Hence, we will again assume that the design will be based on a training set (of multiple sensor observations), with each entry in the training set labeled to identify their relative importance. In the sequel, we consider only two sources X and Y for simplicity (we use X and Y instead of X_1 and X_2 for notational convenience). The analysis can be straightforwardly extended to an arbitrary number of sources.

Our objective is to minimize the following Lagrangian cost

$$\mathbb{E}[d(X, \hat{X}) + d(Y, \hat{Y})|E] + \lambda(R_X + R_Y) , \tag{8}$$

where R_X and R_Y are the rates for the two sources. Note that in the above formulation, we have given equal importance to both the sources for simplicity. However, the Lagrangian can easily be generalized by introducing different weights for the distortion and the rate of source Y. The unconditional cost, equivalent to (8) is

$$\mathbb{E}[P(E|(X,Y))(d(X, \hat{X}) + d(Y, \hat{Y}))] + \lambda(R_X + R_Y), \tag{9}$$

where $P(E|(X,Y))$ is the posterior event probability.

To minimize the cost in (9), correlation between the sources should be exploited. This may be done by sending the same index for many, possibly non-contiguous regions of a source alphabet on a channel and then using the information from the other source to distinguish between index-sharing regions. We next adopt a multiple prototype framework used in various distributed quantizer design algorithms without accounting for event occurrences (see e.g.,[2, 5, 8, 9]), and seek to explicitly minimize the expected distortion during events.

Our proposed scheme is illustrated in Fig. 4. For the source X, a high rate quantizer Q_x first maps X to one of the $\mathcal{K}$ disjoint Voronoi regions (prototypes), C_k^x. These regions span the source space. Next, each Voronoi region is mapped to one

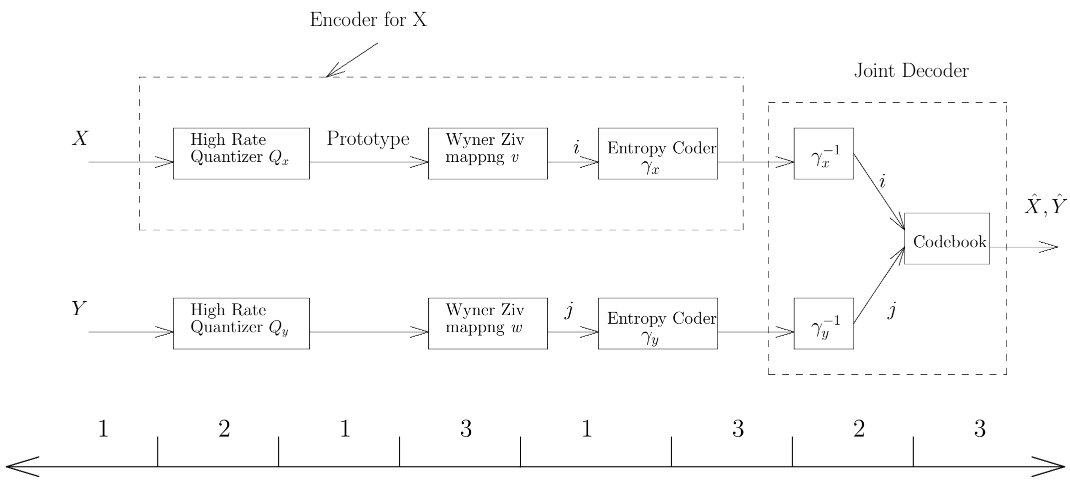


Figure 4: Proposed system and an example of Wyner-Ziv mapping from prototypes (Voronoi regions) to indices.

out of a set of $\mathcal{I}$ indices, via a mapping $v(k) = i$, which we refer to as the Wyner-Ziv (WZ) mapping (the name loosely accounts for the fact that the scenario involves lossy coding with side information). An example of WZ mapping for a scalar source X with $\mathcal{K} = 8$ and $\mathcal{I} = 3$, is given in Fig. 4. 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_y and regions C_l^y, R_j^y. Here, the $\mathcal{L}$ Voronoi regions are mapped to $\mathcal{J}$ indices via WZ mapping $w(l) = j$. The entropy coders γ_x and γ_y operate on the outputs of the WZ mappings for the respective sources, and the coded bits are then transmitted across the channel. (We assume that the WZ mappings efficiently exploit the inter-source correlation and the output indices I and J are approximately uncorrelated, hence the use of separate entropy coders. A Slepian-Wolf coder such as in [1] can alternatively be designed to further improve the performance.) At the decoder, the received bits are first used to obtain the indices i and j, which are then used to get the reconstruction values $\hat{x}_{ij}$ and $\hat{y}_{ij}$.

The transmission rates are given by the expected lengths : $R_X = \mathbb{E}[|\gamma_x(i)|]$ and $R_Y = \mathbb{E}[|\gamma_y(j)|]$. (Again we assume the use of Huffman codes, though any other entropy code can easily be incorporated in our proposed system.)

Given a labeled training set $\mathcal{T}$ which consists of N i.i.d. data pairs (scalar or vector) of (X, Y) and a label $P(E|(X, Y))$ for each data point, we now propose an iterative Lloyd-like algorithm for joint optimization of the WZ mappings v and w, the Huffman coders γ_x and γ_y, and the decoder codebook. We assume that the high rate quantizers Q_x and Q_y are designed independently using a Lloyd's algorithm [6] and kept fixed. (Note that the high-rate quantizers are used primarily to discretize the source and additional performance gains by their optimization will be modest.) We denote the Lagrangian cost incurred for a data pair (x, y) and the corresponding index pair (i, j) by:

$$F(x, y, i, j) = P(E|x, y)(d(x, \hat{x}_{ij}) + d(y, \hat{y}_{ij})) + \lambda(|\gamma_x(i)| + |\gamma_y(j)|). \tag{10}$$

The net Lagrangian cost in (9) which we seek to minimize simply averages the above cost over all the data points. The required update rules can now be specified as:

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

$$v(k) = i = \arg\min_{i'} \sum_{(x,y)\in\mathcal{T}; x\in C_k^x} F(x, y, i', j). \tag{11}$$

2. **Reconstruction Values for X**: For all $i = 1, \ldots, \mathcal{I}$ and $j = 1, \ldots, \mathcal{J}$, find $\hat{x}_{ij}$ such that:

$$\hat{x}_{ij} = \sum_{(x,y)\in\mathcal{T}; x\in R_i^x, y\in R_j^y} \{x.P(E|(x,y))\} \;/\; \sum_{(x,y)\in\mathcal{T}; x\in R_i^x, y\in R_j^y} \{P(E|(x,y))\} \tag{12}$$

3. **Entropy Coder for** X: Design a Huffman coder for the current probability distribution of indices I.

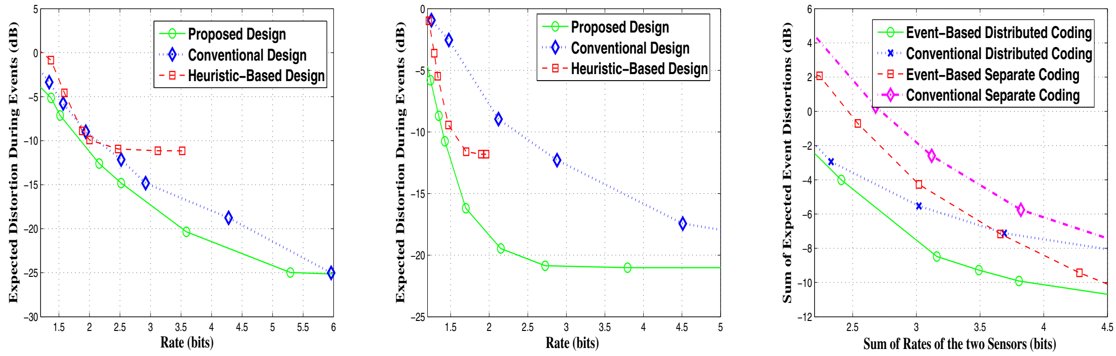
The corresponding update rules for Y can be trivially obtained by symmetry.

5 Simulation Results

Single sensor results: We first consider the case of a single sensor. This will illustrate the gains obtained by designing a quantizer conditioned on the event occurrences. The simulation setting is as follows. We take $f_{X|E}(x) = \frac{1}{2}[\mathcal{N}_{(-2,1)}(x) + \mathcal{N}_{(2,1)}(x)]$, where $\mathcal{N}_{(a,b)}(\cdot)$ is the pdf of a Gaussian random variable with mean a and variance b. In other words, during an event, the data is drawn from an equiprobable mixture of two Gaussians. On the other hand, we take $f_{X|\bar{E}}(x) = \mathcal{N}_{(0,1)}(x)$. The prior probability of event $P(E)$ is assumed to be 0.1. A training set consisting of $3,000$ data points is generated as per the above specifications, and the algorithm derived in Section 3 is implemented for three different schemes :

- **Proposed Design:** Every data point x in the training set is assigned a label (weighing factor) $P(E|x)$, which is the posterior probability of event given x.
- **Conventional Design:** All data points are assumed equally important, hence all have the same label 1.
- **Heuristic-Based Design:** Here a Maximum Likelihood (ML) hard decision rule is employed for event detection. Consequently, all the data points x with $|x| > 1$ are declared as events, and assigned a label 1. On the other hand, training data points with $|x| < 1$ are declared as non-events, assigned a label 0, and are all mapped to the same quantization index (to save rate) with a reconstruction value of $'0'$.

For all 3 schemes, we use an initial codebook of size 128, and run the algorithm with 20 different random initializations. The results obtained are shown in Fig. 5(a), where we plot the expected distortion during events (in dB) against the rate of transmission (in bits). As expected, the proposed scheme outperforms the conventional design which does not account for event occurrences. However, the conventional design is sometimes seen to fair better than the heuristic approach. This occurs because



(a) (b) (c)

Figure 5: Performance comparison: Plots (a) and (b) are for a single sensor scenario. (a) depicts a setting where reliable event detection is hard to achieve, while in (b), reliable event detection is relatively easier (but still not perfect). Plot (c) shows the results for a distributed setting with 2 sensors.

the hard detection of the heuristic approach leads to a significant probability of missed events, thereby causing a large event distortion [1]. Note that the gains provided by the heuristic approach die down as the rate is increased. This happens because, irrespective of the rate, the missed event data points are always reconstructed as $'0'$, and hence the distortion caused by them is a constant that is independent of the rate.

Next, we consider a similar setting as before, except that the density $f_{X|E}(x)$ is changed to be $\frac{1}{2}[\mathcal{N}_{(-4,1)}(x) + \mathcal{N}_{(4,1)}(x)]$, i.e, the event data is now more easily distinguishable from the non-event data, which is still drawn from a $\mathcal{N}_{(0,1)}$ distribution. In this setting, the performance of the conventional design should take a big hit, since the non-event data (which dominates the training set) resembles the event data to a far lesser extent now than before. As can be seen in Fig. 5(b), this is indeed the case. On the contrary, the heuristic approach is expected to perform much better than before, since the hard detector will now become much more reliable. The plot for the heuristic scheme demonstrates this precisely. Note again that the gains in the heuristic approach still saturate, since there is still a non-zero probability of missing an event using the hard detector.

Multiple sensor results: We now present results for a distributed setting. The observations X and Y for the two sensors are again generated from a mixture of Gaussians. In the case of an event, X and Y are drawn from a mixture of 2 equiprobable joint Gaussians, with mean vectors $(2, 2)$ and $(-2, -2)$. When there is no event, X and Y are drawn from a joint gaussian with mean vector $(0, 0)$. In all cases, the variance of X and Y is 1, and their covariance is 0.95. The prior probability of events is 0.1. The number of prototypes and indices used is 64 and 8 respectively, for both the sources. We simulate both scenarios, when the correlations between the sources

[1]On using the Maximum A Posteriori (MAP) rule for event detection, it was observed that the probability of missed events becomes even more significant (due to the small prior event probability $P(E)$), leading to further degradation in the performance of the heuristic scheme. Hence we have reported the results for the ML based heuristic approach only.

are exploited, and when they are not exploited (i.e, separate coding of the sources). Fig. 5(c) shows the results we obtain. The gains achieved by the proposed event based approach over conventional design is evident in both the scenarios.

6 Conclusions

We have provided a simple cost criterion for optimization of compression strategies for event-based monitoring in sensor networks, namely, the minimization of expected distortion during events. In this work, we demonstrated the approach in the setting of distributed vector quantizer design, proposing the use of a labeled training data set that allows us to exploit the joint statistics of events and the observation data. An important problem for future work is the extension of these ideas to unsupervised settings where training data may not be available (e.g., battlefield sensor network to detect and track enemy movements). What would be the best approach for event based compression in such scenarios, where a priori, the sensors have little or no statistical information about the events? In particular, would it be feasible to allow some inter sensor communication as a means to enhance the performance of the overall network?

Acknowledgements

The authors thank the anonymous reviewers for their helpful comments. The first author also thanks Prof. Muriel Médard at MIT for several interesting discussions while he was a visiting student in her research group in 2007.

References

[1] A. Aaron and B. Girod, *Compression with side information using turbo codes*, In Proc. DCC, Mar 2002.

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

[3] P. A. Chou, T. Lookabaugh and R.M. Gray, *Entropy-constrained vector quantization*, IEEE Tran. on Acoustics, Speech and Sig. Proc. , Jan 1989.

[4] T. P. Coleman, A. H. Lee, M. Médard and M. Effros, *Low-complexity approaches to Slepian-Wolf near-lossless distributed data compression*, IEEE Tran. Info. Theory , Aug 2006.

[5] M. Fleming, Q. Zhao and M. Effros, *Network vector quantization*, IEEE Tran. Info. Theory, Aug 2004.

[6] S. P. Lloyd, *Least squares quantization in PCM*, IEEE Tran. Info. Theory , Mar 1982.

[7] S. S. Pradhan and K. Ramchandran, *Distributed source coding using syndromes (DISCUS): Design and construction*, IEEE Tran. Info. Theory , Mar 2003.

[8] D. Rebollo-Monedero, R. Zhang and B. Girod, *Design of optimal quantizers for distributed source coding*, In Proc. DCC , Mar 2003.

[9] A. Saxena, J. Nayak and K. Rose, *On efficient quantizer design for robust distributed source coding*, In Proc. DCC , Mar 2006.

[10] D. Slepian and J. K. Wolf, *Noiseless coding of correlated information sources*, IEEE Tran. Info. Theory , July 1973.

[11] A. D. Wyner and J. Ziv, *The rate-distortion function for source coding with side-information at the decoder*, IEEE Tran. Info. Theory , Jan 1976.

[12] N. Xu et. al., *A wireless sensor network for structural monitoring*, In Proc. Sensys, Nov 2004.