Distributed Encoding of Sensor Data

David L. Neuhoff and Daniel Marco[1]
Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, Michigan, USA
e-mail: {neuhoff,idaniel}@umich.edu

I. Distributed Encoding of Sensor Data

There are many situations for which it would be desirable to measure a spatially varying field in some geographical region of interest, for example, temperature, pressure, light intensity, moisture, CO_2. With the advent of small, low cost sensing devices and radios, it would seem feasible to form a wireless network of closely spaced sensors that samples the field at the sensor locations, and then quantizes, encodes and transmits the measured values to some collector site that, in turn, forms a reproduction of the field. This is one of the reasons why there is much current interest in wireless sensor networks.

Since the field must be sampled, quantized and transmitted, it cannot be known exactly at the collector. However, there will normally be a desired accuracy, for example a target mean squared error (MSE) distortion. The question then becomes: How much communication resources (e.g. power, bandwidth, number of transmitted bits) are needed in order that the collector can produce a field reconstruction with no more than a specified distortion?

A first thought is that the resources consumed by a sensor network increase with the spatial density of its sensors, and therefore, sensors should be deployed in a pattern that minimizes the number needed to attain satisfactory field accuracy. However, fields are often spatially correlated, meaning that neighboring sensors produce correlated measurements. Moreover, such correlations increase as the spatial density of the sensors increases. It seems plausible that such inter-sensor correlations can be exploited so as to permit a dense sensor network to operate essentially as efficiently as one with minimal density, while having additional advantages, such as being resilient to sensor failures, and permitting the attribute field to be measured adaptively or with higher spatial resolution than originally planned. Our goal is to determine if this is correct, and if so to develop suitable coding systems.

In this paper, we ignore transmission issues and focus simply on the total number of bits to transmit to the collector to form a reconstruction of the field with a given MSE. (We assume that all sensors can transmit bits to the collector without error. With this assumption, with total number of bits as the cost measure, and with the style of coding to be described shortly, it can be argued that sensor-to-sensor relaying offers no advantages.) This problem is similar to image coding and transmission, except that the quantization, encoding and transmission are constrained to take place separately at each sensor (pixel location), in contrast to traditional image coding and transmission, wherein the entire image is available for quantization, encoding, and transmission. Due to the need to separately encode values from separate sensors, we pursue a Slepian-Wolf style coding approach [1].

For simplicity and concreteness, let us assume the field is a one-dimensional, continuous-space, zero-mean, stationary Gaussian random process $X(t)$, for example with autocorrelation function $R(\tau) = e^{-\tau}$. The n-th sensor samples $X(t)$, producing $X_n = X(nT)$, where T is the sample spacing. Thus X_n becomes a stationary discrete-space Gaussian random process (If $R(\tau) = e^{-\tau}$, it is Gauss-Markov with correlation coefficient $\rho = e^{-T}$.)

At the n-th sensor site, consider a system that scalar quantizes X_n with a quantization rule Q_n, producing $\widehat{X}_n = Q_n(X_n)$, and then losslessly encodes $\widehat{X}_n$ using a Slepian-Wolf encoder that presumes that the decoder has already decoded $\widehat{X}_1, \ldots, \widehat{X}_{n-1}$. (Actually, this requires encoding a block of $\widehat{X}_n$ values taken at different times, which are assumed to be independent of each other, but of course, correlated with the sensor values at neighboring sites at the same time.)

When this scheme is used to encode the data from N sensors, the average number of bits per unit distance (the *coding rate*) is, approximately, the normalized joint entropy

$$r_T(Q_1, \ldots, Q_N) = \frac{1}{T}\frac{1}{N} H(\widehat{X}_1, \ldots, \widehat{X}_N) \tag{1}$$

and the average distortion at the sample times is

$$d_T(Q_1, \ldots, Q_N) = \frac{1}{N}\sum_{n=1}^{N} E(X_n - Q_n(X_n))^2 \tag{2}$$

To reconstruct the continuous-space process $X(t)$, the $\widehat{X}_n$'s must be interpolated in some way forming $\widehat{X}(t)$. However, we will ignore this issue and simply assume that the mean-squared error between $\widehat{X}(t)$ and $X(t)$ can be approximated by the distortion defined above. (It will actually be a little larger.)

[1]This work was supported by NSF Grant ANI-0112801.

Given some N, typically fairly large, the problem now becomes that of designing $Q_1, \ldots, Q_N$ to minimize $r_T(Q_1, \ldots, Q_N)$ subject to the constraint $d_T(Q_1, \ldots, Q_N) \leq D$, where D is a target distortion. Let $\mathcal{R}_T(D)$ denote the resulting least coding rate at distortion D with sample spacing T. The principal questions are: What quantizers have performance attaining or approaching $\mathcal{R}_T(D)$? And what is the dependence on T of these quantizers and $\mathcal{R}_T(D)$? For example, if for fixed D, $\mathcal{R}_T(D)$ converges to a finite limit as $T \longrightarrow 0$, then we may conclude that, at least in the given scenario, communication resources do not need to increase without bound as sensor density increases. Moreover, if it decreases monotonically with T, then dense sensor networks are at least as efficient as sparse ones.

II. An Open Problem in the Encoding of Continuous-Time Sources

We now point out that quite apart from sensor networks, $\mathcal{R}_T(D)$ is a quantity of basic interest in source coding, because it indicates the dependence on sampling rate of the performance of scalar quantizers with entropy coding when applied to continuous-time sources. Its importance not withstanding, little appears to be known about $\mathcal{R}_T(D)$. There has been work on coding of bandlimited sources and on nonbandlimited sources but with prefiltering [2, 3, 4, 5, 6]. However, in our case, there is no reason to assume the source is bandlimited, and prefiltering is not an option due to the necessity of distributed encoding. Moreover, this work does not make clear the effect of decreasing T. There has also been work on the fundamental tradeoff between sampling rate and quantizer accuracy [7, 8], but not, however, with MSE as the distortion measure. In this remainder of this paper and in the talk accompanying it, we describe our thoughts and preliminary results on the finding and/or bounding of $\mathcal{R}_T(D)$.

For discrete-time sources, much of what is known about the performance of quantizers comes from *high resolution theory* (c.f. [9]), which is applicable when the coding rate in bits per sample is at least of moderate size (at least 2 or 3 bits/sample) or, equivalently, the coding distortion is small. For example, it is known that at high resolution, scalar quantization with high-order entropy coding achieves coding rate within 1/4 bit of the Shannon optimal rate-distortion function.

Unfortunately, high resolution analysis is not applicable to the case of continuous-time sources when sampling at high rates. This is because the coding rate in bits/second equals the coding rate in bits/sample times the $1/T$. Therefore,

$$\mathcal{R}_T(D) = \frac{1}{T} R_T(D) \tag{3}$$

where $R_T(D)$ is the least coding rate in bits per sample needed to encode the samples to distortion D or less. When $R_T(D)$ is moderate to large, it may be approximated using high-resolution methods. However, one can see from the above that no matter how small is D, if $\mathcal{R}_T(D)$ is to decrease, or at least not to blow up, when $T \longrightarrow 0$, then $R_T(D) \longrightarrow 0$. This clearly indicates that the high resolution assumption is not valid when T is small.

Note that Shannon rate-distortion theory for continuous-time Gaussian sources indicates that optimal encoding improves as $T \longrightarrow 0$ (c.f. [10]). The questions we are asking is whether scalar quantization with entropy coding improves as $T \longrightarrow 0$. If so, how close to the Shannon optimal does it come? If not, what then is the best choice of T?

Though bandlimiting is not an allowable option, it might be possible to use results for systems with prefilters to obtain bounds on the performance without prefiltering.

Following the lead of Ziv [11], Zamir and Feder [12] analyzed scalar quantization with entropy coding for discrete-time sources. When applied to our situation, their results show

$$R_T(D) \leq S_T(D) + .754, \text{ for all } D \tag{4}$$

where $S_T(D)$ is the Shannon rate-distortion function applied to samples taken with spacing T. Unfortunately, this result is not too helpful because when multiplied by $1/T$ the .754 blows up. (This shows the futility of bounds of the form $S_T(D)$ plus a constant, in the case of decreasing T.)

It is interesting to note that the system analyzed by Zamir and Feder [12] employed subtractive dithering, which is equivalent to simply translating the cells and levels of the various quantizers, i.e. $Q_n(x) = Q_1(x + \delta_n)$, where δ_n is the *dither* value at time n. In this application, dither was introduced solely to make the analysis tractable. On the other hand, for highly correlated discrete-time sources, e.g. when T is small, the dither can be exploited at the decoder to reduce distortion. In [5] the discrete-time to continuous-time interpolator played this role. Alternatively, one could let the decoded reproduction of X_n be[1]

$$\widetilde{X}_n = \frac{1}{N} \sum_{n=1}^{N} Q_n(X_n) \tag{5}$$

To see the potential benefits of this averaging, consider the case that all quantizers are uniform with many cells of width Δ and assume that the dither shifts adjacent quantizers by Δ/N relative to one another. Finally, consider the extreme case that T is so small the N successive source samples can be assumed to be essentially identical. In this case, the next effect is that N successive quantized values determine the value of X_n to within an interval of width Δ/n. Thus, the net effect is to reduce the quantizer stepsize from Δ to Δ/N and to rduce the distrotion by a factor of N^2. Random dither (uniformly distribution from $-\Delta/2$ to $\Delta/2$) has a similar effect, except the distortion can be shown to decrease by only a factor of N. In either case, as T decreases, we conjecture that for the resulting system $\mathcal{R}_T(D)$ decreases to some finite limit

[1] A centered sliding average would be a little better. The number of terms averaged need not equal N.

as $T \longrightarrow 0$, and that this limit is not far from the Shannon limit $\mathcal{S}(D)$. Without dithering and averaging at the decoder, we believe that scalar quantization with entropy coding is doomed to poor performance when T is small.

References

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Thy.,* vol. 19, pp. 471–480, July 1973.

[2] T.J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inform. Theory*, vol. 11, pp. 558-567, Oct. 1965.

[3] T.J. Goblick and J.L. Holsinger, "Analog source digitization: A comparison of theory and practice," *IEEE Trans. Inform. Theory*, vol. 13, pp. 323–326, April 1967.

[4] W.C. Kellogg, "Information rates in sampling and quantization," *IEEE Trans. Inform. Theory,*, vol. 13, pp. 506–511, July 1967.

[5] R. Zamir and M. Feder, "Rate-distortion performance in coding bandlimited sources by sampling and dithered quantization," *IEEE Trans. Inform. Theory*, vol. 41, pp. 141–4, Jan. 1995.

[6] R. Zamir and M. Feder, "Information rates of pre/post-filtered dithered quantizers," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1340–53, Sept. 1996.

[7] Z. Cvetkovic and M. Vetterli, "Error rate characteristics of oversampled analog-to-digital conversion," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1961–4, Sept. 1998.

[8] Z. Cvetkovic and I. Daubechies, "Single-bit oversampled A/D conversion with exponential accuracy in the bit-rate," *Proc. DCC*, pp. 343-352, 2000.

[9] R.M. Gray and D.L. Neuhoff, "Quantization," *IEEE Trans. Inform. Thy.,* vol. 44, pp. 2325-2383, Oct. 1998.

[10] T. Berger, *Rate Distortion Theory*, Prentice-Hall,Englewood Cliffs, NJ, 1971.

[11] J. Ziv, "Universal Quantization," *IEEE Trans. Inform. Theory*, vol. 31, pp. 344–347, May 1985.

[12] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Trans. Inform. Theory*, vol. 38, pp. 428–436, March 1992.