Communications and Signal Processing Laboratory (CSPL) Report No. 345, University of Michigan at Ann Arbor.

Source coding with feedforward: Gaussian Sources

S. Sandeep Pradhan
Department of Electrical Engineering and Computer Science
University of Michigan at Ann Arbor
email: pradhanv@eecs.umich.edu.

December 1, 2003

Abstract

We explore a mathematical characterization of the functional duality between classical source and channel coding, formulating the precise conditions under which the optimal encoder for one problem is functionally identical to the optimal decoder for the other problem. We then extend this functional duality to the case of coding with side information. We consider several examples corresponding to both discrete-valued and continuous-valued cases to illustrate our formulation.

1 Introduction

With the recent emergence of applications related to sensor networks [1, 2], efficient encoding of information signals in a multiterminal setting has received special attention. One such problem is that of source coding with side information at the decoder [3], where the encoder wishes to represent a source with an index coming from a set to be transmitted to a decoder which has access to some correlated information. In many applications involving estimation of information field (such as seismic, acoustic), the signal to be estimated at the decoder is delay-sensitive, and the signal to be estimated might be available to the decoder either in a delayed and/or noisy form. Thus an efficient encoder must take into account the information available at the decoder at the time of decoding a particular sample.

Motivated by this application, we consider an idealized version of this problem in this paper. Extension of the proposed work to include other practical constraints will be a part of the future work [4]. Consider a stationary discrete memoryless source X with a probability distribution $p(x)$ with some alphabet $\mathcal{X}$, and a reconstruction alphabet $\hat{\mathcal{X}}$. The encoder observes a sequence of independent realizations of the source from the given distribution. Associated with the source, there is a distortion measure $d : \mathcal{X} \times \hat{\mathcal{X}} \to \mathbb{R}^{+}$. The encoder is a mapping from the l-product source alphabet to an index set: $f : \mathcal{X}^l \to \{1, 2, \ldots, M\}$, where l denotes the block-length in encoding and $\frac{1}{l} \log M$ denotes the transmission rate in bits/sample. The distortion measure for a pair of sequences of length l is the average of the distortions of l samples: $d(\mathbf{x}, \hat{\mathbf{x}}) = (1/l)\sum_{i=1}^{l} d(x_i, \hat{x}_i)$, where x_i and $\hat{x}_i$ denote the ith samples of $\mathbf{x}$ and $\hat{\mathbf{x}}$ respectively.

The decoder structure is given by the following: the decoder receives the index transmitted by the encoder, and to reconstruct the ith sample (for $i = 1, 2, \ldots, l$), it has access to all the past samples of the source till $i - 1$. In other words, the decoder is a sequence of mappings $g_i : \{1, 2, \ldots, M\} \times \mathcal{X}^{i-1} \to \hat{\mathcal{X}}$ for $i = 1, 2, \ldots, l$. Let $\mathbf{g}(\mathbf{x})$ denote the l-length vector reconstruction of the l-length source vector $\mathbf{x}$. The goal is to minimize $E\left[d\left(\mathbf{X}, \mathbf{g}(\mathbf{X})\right)\right]$ for a given rate $R = (1/l)\log M$. We refer to this problem as source coding with feedforward. Let $R_{ff}(D)$ denotes the infimum of R over all encoder-decoder pair such that $Ed \leq D$ for some $D > 0$. A schematic of this problem is shown in Fig. 1.

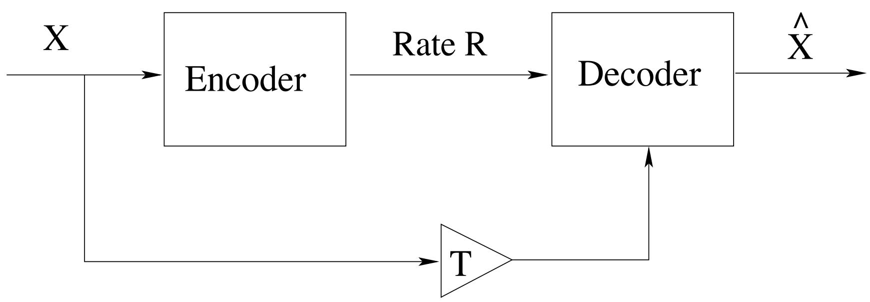

Figure 1: Source coding with feedforward: the decoder, to reconstruct any source sample, has access to all the previous samples in addition to the quantized version of the source.

This problem is formulated and the optimum performance is evaluated in [4]. The main result of [4] is summarized in the following:

1 For stationary memoryless sources, $R_{ff}(D) = R(D)$, where $R(D)$ denotes the optimal Shannon rate-distortion function.

2 For sources with memory, $R_{ff}(D) \leq R(D)$.

3 For a stationary source, the error exponent of source coding with feedforward $E_{ff}(R, D)$ for rate R and distortion D satisfies $E_{ff}(R, D) \geq E(R, D)$ where $E(R, D)$ denotes the standard source coding error exponent for rate R and distortion D.

2 Main Result

In this paper we consider a special case of stationary memoryless Gaussian source with zero-mean and variance σ^2, and with mean squared error as the distortion measure. The main result of this paper is given by the following theorem:

Theorem: For a stationary memoryless Gaussian source with zero-mean and variance σ^2 and distortion measure given by the mean squared error, for any $R > R_{ff}(D) = R(D)$, and for sufficiently large l, $\exists$ a pair of encoder and decoder for source coding with feedforward such that $P_e \le c_1 e^{-c_2 e^{c_3 l}}$ for some $c_1, c_2, c_3 > 0$ where P_e is the probability that a source word is not reconstructed with distortion less than or equal to D.

Note: For stationary memoryless Gaussian source, the source decoding error decays doubly exponentially with block length where for the reconstruction of any source sample all the past samples are available, while for the conventional Shannon source coding, it decays only exponentially [5]. Also note that this problem can be considered as the dual [6] of channel coding with feedback [7, 8] for additive white Gaussian noise channels.

3 Proof of Theorem: Encoding and Decoding strategy

In the following we describe a deterministic scheme involving an encoder-decoder pair with a structure as given above. The inspiration for this comes from the strategy described in [8] for channel coding with feedback.

Let $\beta > 0$ be some constant. Consider the following sequence of recursive functions of the source:

$$Y_{i+1} = Y_i - \frac{\sqrt{\beta^2 - 1}}{\beta^{i+1}} X_i \tag{1}$$

for $i = 2, \ldots, l$, where $Y_1 = 0.5$ and $Y_2 = Y_1 - \frac{1}{\beta} X_1$.

It can be shown that Y_i is Gaussian with mean Y_1 and variance given by

$$\frac{\sigma^2}{\beta^2} \left[1 + \frac{1}{\beta^2} \frac{\beta^{2(i-2)} - 1}{\beta^{2(i-2)}} \right]. \tag{2}$$

Further $\lim_{i \to \infty} Var(Y_i) = (\sigma^2 (1 + \beta^2))/\beta^4$. For the first block of l source samples, Y_{l+1} will be a function of $\{X_1, X_2, \ldots, X_l\}$. Consider a uniform scalar quantizer with M levels and bounded between $-\Delta/2$ and $\Delta/2$, where Δ will be determined later. Thus the step size of this quantizer is Δ/M. The encoder quantizes $(Y_{l+1} - Y_1)$ using the above quantizer and the index of the cell containing it is sent to the decoder. Let $\hat{Y}$ denotes the quantized version of $(Y_{l+1} - Y_1)$.

The decoder reconstruction is given by the following scheme.

$$\hat{X}_i = \beta \hat{X}_{i-1} - \frac{(\beta^2 - 1)}{\beta} X_{i-1} \tag{3}$$

for $i = 3, \ldots, l$ and $\hat{X}_1 = \beta (Y_1 - \hat{Y})$, $\hat{X}_2 = \sqrt{\beta^2 - 1} (\hat{X}_1 - X_1)$.

The encoder and decoder start over for the next block of l samples.

4 Distortion Analysis

Let us model the quantization of $(Y_{l+1}-Y_1)$ as follows. Let $(Y_{l+1}-Y_1) = \hat{Y}+Z$ where Z denotes the quantization noise, and is modeled as independent of $\hat{Y}$ and is uniformly distributed from $-\Delta/2M$ to $\Delta/2M$. Note that this is an approximation that becomes increasingly better as l becomes large. Let D' denote the variance of Z.

Now we calculate the average expected distortion which is given by

$$\frac{1}{l}\left[\sum_{i=1}^{l} E(X_i - \hat{X}_i)^2\right]. \tag{4}$$

Using some algebra it can be shown that

$$E(X_1 - \hat{X}_1)^2 = \frac{\sigma^2}{\beta^2}\left[\frac{\beta^{2(l-1)}-1}{\beta^{2(l-1)}}\right] + \beta^2 D', \tag{5}$$

and for $i = 2, 3, \ldots, l$, we have

$$E(X_i - \hat{X}_i)^2 = \frac{\sigma^2}{\beta^4} + \frac{\sigma^2(\beta^2-1)^2}{\beta^6}\left(\frac{1-\beta^{-2(l-i)}}{1-\beta^{-2}}\right) + D'(\beta^2-1)\beta^{2(i-1)}. \tag{6}$$

Together, the above expressions give

$$\frac{1}{l}\sum_{i=1}^{l} E(X_i - \hat{X}_i)^2 = \frac{D'\beta^{2l}}{l} + \frac{\sigma^2(l\beta^2 - \beta^2 + 1)}{l\beta^4}. \tag{7}$$

Let $\frac{1}{l}\sum_{i=1}^{l} E(X_i - \hat{X}_i)^2 = D$, for some $D > 0$ and let M be chosen such that $M = \beta^{l(1+\epsilon')}$ for some $\epsilon' > 0$ which results in

$$D' = \frac{\Delta^2}{12M^2} = \frac{\Delta^2}{12(\beta^{2l}\beta^{2l\epsilon'})}. \tag{8}$$

Now choose $\Delta = \beta^{l\epsilon'}\sqrt{\frac{\beta^2+1}{\beta^4}}$. Noting that $R_{ff}(D) = R(D) = (1/2)\log(\sigma^2/D)$, we have from the above expressions the rate of transmission to be

$$R = \frac{1}{2}\log\left[\frac{\sigma^2}{D}\right] + \epsilon \tag{9}$$

where

$$\epsilon = \frac{1}{2}\log\left[\frac{D}{D-\delta}\right] + \frac{\epsilon'}{2}\log\left[\frac{\sigma^2}{D-\delta}\right] \quad \text{and} \quad \delta = \frac{1}{12le^{\frac{4R}{(1+\epsilon')}}}\left[12\sigma^2(1-e^{\frac{2R}{(1+\epsilon')}}) + e^{\frac{2R}{(1+\epsilon')}} + 1\right]. \tag{10}$$

Clearly, $\lim_{l\to\infty}\lim_{\epsilon'\to 0}\epsilon(\epsilon', l) = 0$. Hence for sufficiently large l, one can choose ϵ' to make ϵ arbitrarily small.

Now let us analyze the probability P_e that a source word is not reconstructed with distortion less than or equal to D. This probability is upper bounded by the probability that the absolute value of $(Y_{l+1} - Y_1)$ is greater than $\Delta/2$. Thus we have

$$P_e \leq \Pr\left[|Y_{l+1} - Y_1| > \frac{\beta^{l\epsilon'}}{2}\sqrt{\frac{\beta^2+1}{\beta^4}}\right] = 2\mathrm{erfc}\left[\frac{\beta^{l\epsilon'}}{2}\left(\frac{\beta^l\sqrt{\beta^2+1}}{2\sigma\sqrt{\beta^{2(l+1)}+\beta^{2l}-\beta^2}}\right)\right] \tag{11}$$

$$\leq 2\text{erfc}\left[\beta^{l\epsilon'}\frac{1}{4\sigma}\right] \leq c_1 e^{-c_2 e^{l c_3}}, \tag{12}$$

where

$$c_1 = \frac{8\sigma}{a\sqrt{2\pi}e^{\frac{lR\epsilon'}{(1+\epsilon')}}}, \quad c_2 = \frac{1}{8\sigma^2} \text{ and } c_3 = \frac{2R\epsilon'}{(1+\epsilon')}, \tag{13}$$

and we have used the following fact:

$$\text{erfc}(x) \leq \frac{1}{x\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right). \tag{14}$$

5 Conclusions

In this paper we considered the problem of source coding with feedforward where at the time of reconstruction of the ith source sample, the decoder has access to all the past samples. We have shown that for the Gaussian source with mean squared error criterion, if the transmission rate $R > R_{ff}(D) = R(D)$, the probability that a source word is not reconstructed with distortion less than or equal to D decays doubly exponentially with block-length for sufficiently large block-length. Hence, although $R_{ff}(D) = R(D)$ for sources without memory, the source coding error exponent can be increased with feedforward.

Acknowledgments: The author would like to thank Professor Kannan Ramchandran and Dr. Prakash Ishwar of the University of California at Berkeley for inspiring discussions on multiterminal source coding and duality.

References

[1] J. M. Kahn, R. H. Katz, and K. S. J. Pister, "Mobile networking for smart dust," in *ACM/IEEE Intl. Conf. on Mobile Computing and Networking, Seattle, WA*, August 1999.

[2] S. S. Pradhan, J. Chou, and K. Ramchandran, "Distributed compression in a dense sensor network," *IEEE Signal Processing Magazine*, vol. 19, pp. 51–60, March 2002.

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

[4] R. Venkataramanan and S. S. Pradhan, "Source coding with feedforward," *to be submitted to IEEE Trans. Inform. Theory*, December 2003.

[5] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete memoryless sources.* Academic Press, New York, 1981.

[6] S. S. Pradhan, J. Chou, and K. Ramchandran, “Duality between source coding and channel coding and its extension to the side information case,” *IEEE Trans. on Information Theory*, vol. 49, pp. 1181–1203, May 2003.

[7] C. E. Shannon, “The zero-error capacity of a noisy channel,” *IRE Trans. on Information Theory*, vol. IT-2, pp. 8–19, September 1956.

[8] J. P. M. Schalkwijk and T. Kailath, “A coding scheme for additive noise channels with feedback- part II: Band-limited signals,” *IEEE Trans. on Information Theory*, vol. IT-12, pp. 183–189, April 1966.