ON THE TRANSMISSION OF A CLASS OF HIDDEN MARKOV SOURCES OVER GAUSSIAN CHANNELS WITH APPLICATIONS TO IMAGE COMMUNICATION

Igor Kozintsev

Microprocessor Research Labs
Intel Corporation
igor.v.kozintsev@intel.com

Kannan Ramchandran

EECS Department
University of California at Berkeley
kannanr@eecs.berkeley.edu

ABSTRACT

In this work we address the problem of transmission of a useful class of hidden Markov sources with Gaussian observations over a discrete-time power-constrained AWGN channel. Our source model approximates the wavelet representation of natural images, and is motivated by the recent successful modeling approaches in the state-of-the-art image compression and restoration methods. We derive information-theoretic performance bounds, and propose a simple yet efficient joint source-channel coding solution which combines analog and digital modes of transmission. Our results indicate the promise of the proposed approach to the problem of robust image communication.

1. INTRODUCTION AND MOTIVATION

In this paper we propose a novel joint source-channel coding (JSCC) framework for communication over noisy channels. JSCC refers to the joint design of the source and channel coders, and has been applied to the problem of image communication in a number of publications (see, e.g., [1, 2, 3]). Our proposed JSCC solution is motivated by a hybrid analog/digital approach to JSCC [4, 5] and by practical image coding algorithms [6, 7]. *While there is no theoretical advantage of JSCC in the setup of this paper, we are motivated primarily by the simplicity of our (theoretically suboptimal) joint approach that nevertheless performs surprisingly close to the theoretical bounds.*

We consider a class of statistical source models which have recently demonstrated their excellent performance in image compression [6, 7] and denoising [8, 9, 10, 11]. Specifically, we investigate models where observations are independent Gaussian random variables with the variances determined by a hidden Markov random process. This class of models is also very popular in the area of speech processing [12] which may be one of future areas of application of our methods. Similar image models were proposed and analyzed in [13] from a purely source coding (rate-distortion) perspective. Specifically, in one case the source was interestingly modeled as a Gaussian "spike" process, which can be represented as a product of a sequence of binary Bernoulli random variables with a sequence of Gaussian random variables. The source model proposed in [13] is a special case of a 2-state hidden Markov model with Gaussian observations analyzed in this paper. In [14], a problem similar to that proposed here was addressed using an "all-analog" solution wherein the (Gaussian) source samples were linearly encoded and decoded. This is also a special case of our proposed hybrid digital-analog framework for dealing with natural image sources.

This work was supported in part by NSF Grant CCR 99-79381.

In this work we consider the case of communication over a discrete-time power-constrained AWGN channel and we use the mean squared error (MSE) distortion criterion. The rest of this paper is organized as follows. We first describe various hidden Markov-type models for wavelet image coefficients. Then we formulate a joint source-channel coding problem for a simplified 1-D hidden Markov model and propose a hybrid analog/digital JSCC scheme. Finally, we show the surprising performance of this simple scheme with respect to the optimal theoretical bounds for the model applied to real wavelet image data.

2. HIDDEN MARKOV MODELS OF IMAGES

Early classes of subband and wavelet coders were derived from modeling image subband data as being i.i.d. random variables within the image subbands. These approximations have been shown to be inefficient in acknowledging the considerable intraband structure present in the wavelet image subband representation. Fig. 1 shows the scaled wavelet coefficients of the Lena image. Notice the presence of significant dependencies in the transform coefficients. It is these dependencies that are efficiently exploited by current state-of-the-art image compression [6, 7] and restoration [8, 15, 11] algorithms. There can be seen to be two broad classes of stochastic models for wavelet image coefficients that exploit memory, both based on hidden Markov models. The first class we denote as interband (or interscale) modeling because it strives to exploit the redundancy between wavelet coefficients at different scales but at the same spatial location. Hidden Markov trees (HMT) introduced in [16, 8] are excellent representatives of this approach. Wavelet zerotrees (which appeared historically before HMT's, see [6, 17]) represent a popular compression methods that implicitly use the interscale model. Fig. 1 illustrates the interscale tree structure used in zerotrees. The second approach attempts

0-7803-6297-7/00/$10.00

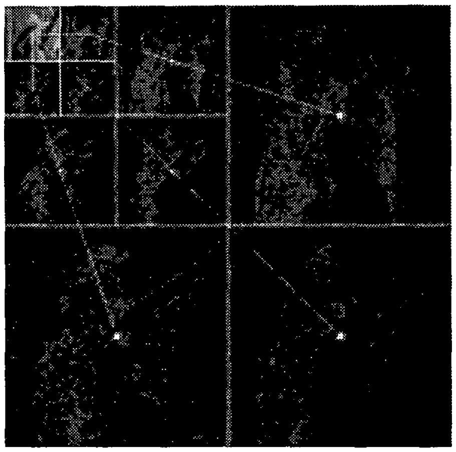

Figure 1: Wavelet coefficients of Lena image (subbands are scaled) and the spatial tree structure.

to model the relations between neighboring wavelet coefficient within a subband and is sometimes denoted as intraband (or intrascale). The intrascale models are typically variations of hidden Markov random fields (HMRF). This approach was introduced in compression in [7] and denoising in [15, 11]. In [11] an HMRF model was proposed to approximate the subband wavelet coefficients of images. Fig. 2 illustrates an HMRF model in one subband using a factor graph conventions (see [11] and references therein for more details about factor graphs). In this plot open circles denote wavelet coefficients X_i that are modeled as zero-mean Gaussian random variables conditionally independent of other coefficients given the hidden states S_i (double circles) that are connected to them. Hidden states correspond to local variance of wavelet coefficients and assumed to form a Markov random field[1].

Both interscale [8] and intrascale [11] models for image wavelet coefficients assume that they are conditionally independent zero-mean (except the low band) Gaussian random variables. The two modeling approaches only differ in modeling the hidden state dependencies: in one case it is a Markov tree, and in another - a Markov random field. Motivated by this, in this paper we consider a simplified 1-D version of the two models outlined above. Though not being the most accurate approximation to the real image data, our modeling approach provides valuable insights into the problem of JSCC for image communication.

3. PROBLEM FORMULATION

The general setup of this work is illustrated in Fig.3 where f and g denote the encoder and decoder functions to be designed for the given source and channel. In the following we describe the source and channel models used in this work and our proposed JSCC scheme. X denotes the input source data, Y is the encoded data and $\hat{X}$ is the decoded estimate of X after transmission over and additive noise channel. The problem is formulated as finding an encoder-decoder pair $\{f, g\}$

$$\{f, g\} = arg \min_{f,g} E[(X - \hat{X})^2],$$

subject to the power constraint on the transmitted signal Y.

[1]Since the states are hidden the overall model for wavelet coefficients within a subband is a HMRF. E.g., state S_i is conditionally independent of all other states given its four neighbors $S_i^N, S_i^E, S_i^S, S_i^W$.

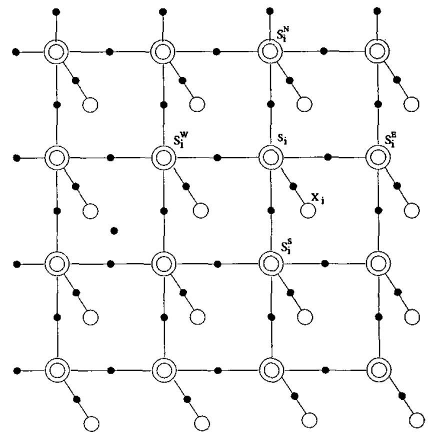

Figure 2: HMRF of wavelet coefficients in a subband.

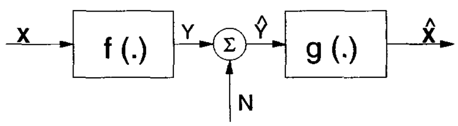

Figure 3: General setup. Given the source and the channel the goal is to design the encoder f and decoder g to optimize the end-to-end performance.

3.1. Source model

The source model which we consider in this work is represented by a doubly stochastic random process. The unobserved random process S_t, $t = 1, 2, \ldots$ is modeled as an ergodic stationary first order Markov chain with the alphabet $A_S = 1, 2, \ldots, n$. This process generates a sequence $s_1, s_2, \ldots$ which is called a state sequence. We assume that the Markov chain process is stationary and has transitional probability matrix with entries $P(i|j)$ - probabilities of going to state i at time $t+1$ from the state j at time t. The observed random variable X_t at time t is normally distributed with zero mean and variance $\sigma(s_t)$ independent of other X_j's and s_j's $j \neq t$, i.e., $X_t \sim \mathcal{N}(0, \sigma(s_t))$. We assume that the source produces a sample every τ_S seconds. This source model is illustrated in Fig. 4 for the special case where the number of states is equal to 2.

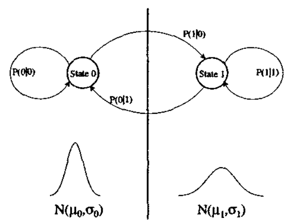

Figure 4: An example of a 2-state HMM with gaussian observations. $P(.|.)$'s represent the transitional probabilities of the Markov chain.

3.2. Channel model

In this work we investigate the case of a power constrained AWGN channel. Let the input to the channel at time t be Y_t. Then the output $\hat{Y}_t = Y_t + N_t$ where N_t is a zero-mean unit-variance Gaussian random variable independent of Y_t. The power constraint has the form $E[Y_t^2] \leq P_Y$. We assume that the channel can be used every τ_C seconds, which is determined by the available bandwidth. Then the overall data rate R is defined as $R = \tau_C/\tau_S$ source symbols per channel use. We also make a standard assumption of a block-N encoder/decoder.

3.3. Theoretical performance

According to information theory, the optimal performance theoretically attainable (OPTA) in our setup (as the block size N allowed to go to infinity) is obtained by evaluating the distortion rate function $D(R_S)$ ($R_S = 1/\tau_S$) of the Gaussian Hidden Markov source at the capacity of an AWGN channel C ($C = \frac{1}{2R} log(1+\frac{P_Y R}{\sigma_N^2})$). The main complication comes from the fact that there is no closed-form expression for the distortion rate function of the Gaussian HMM source. Instead we provide lower and upper bounds on the OPTA.

An obvious upper bound on the OPTA function is obtained by pretending that the source samples are i.i.d. (having a distribution equal to the mixture of Gaussians) and finding the $D(R_S)$ function numerically using the Blahut-Arimoto algorithm [18]. Clearly one can do no better in terms of the resulting distortion by making this memoryless assumption. On the other hand, the bound is also tight if the source is indeed memoryless. A simple closed-form lower bound on the OPTA is obtained by pretending that the states are known at both encoder and decoder. This assumption allows us to use the standard result for $D(R_s)$ function of a parallel Gaussian source. Thus, a lower bound on the Distortion-Rate function is obtained by "reverse waterfilling" [18].

4. THRESHOLD-BASED JSCC SCHEME

Our proposed threshold-based JSCC solution is motivated by the existing state-of-the-art wavelet image coding schemes. The idea behind those schemes is to efficiently represent the locations of significant coefficients and to encode only the values corresponding to those locations. For example, Shapiro's zerotrees [6] do this by exploiting the interband dependencies. Another approach [19, 7] is similar in spirit but uses the intraband dependencies to efficiently encode the locations of significant coefficients. We will use this strategy in our hybrid approach to the JSCC.

The hybrid JSCC system is illustrated in Fig. 5. The input symbols S_t are compared to the threshold T. If the symbol is above the threshold (i.e., it is significant) the symbol's value V_t is send to the receiver using a simple linear scaling (analog transmission). Its location L_t is transmitted using a tandem source/channel coding scheme which achieves arbitrary small probability of error as the block-size N goes to infinity. Analog and digital modes of transmissions share the available channel rate. If the symbol S_t is below the threshold T (it is insignificant) the symbol is not send and reconstructed with its mean (zero) value at the receiver. The scaling coefficient k is chosen to satisfy the power constraint. By varying the threshold T, a desired overall R is attained.

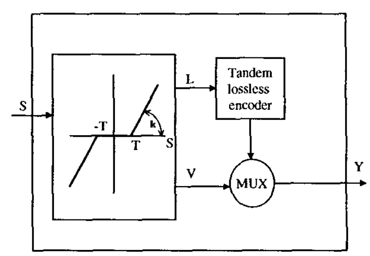

Figure 5: A threshold-based JSCC encoder. Only locations and values of the significant coefficients are send to the receiver.

5. EXPERIMENTAL RESULTS

We performed experiments to assess the performance of our threshold-based hybrid JSCC solution. The parameters of the source are estimated using the Expectation-Maximization (EM) algorithm [12] for a first high-pass subband of Lena image. The source is modeled as a two-state Gaussian HMM with standard deviations $\sigma_0 = 2.53$ and $\sigma_1 = 14.7$ and transitional probabilities $P(0|0) = 0.9835$, $P(1|0) = 0.0165$, $P(0|1) = 0.0648$ and $P(1|1) = 0.9352$ The upper and lower bounds on the D(R) function are shown in Fig. 6 together with the performance of the proposed hybrid system. We have assumed that the digital tandem system achieves the theoretically optimal performance which is an accurate assumption as the block-size N increases. Note that the performance of the proposed hybrid system is below one bound. This is not due to numerical errors. The hybrid system performance is not allowed to be better than the OPTA curve, but it can be better than the bound on the OPTA. In fact, the performance of the hybrid system provides a new, more accurate bound on the OPTA. We would like also to stress

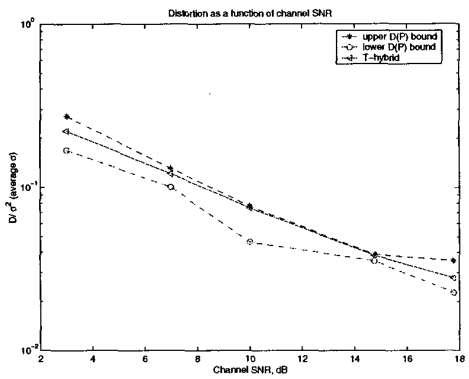

Figure 6: The bounds on the distortion-rate function for a Gaussian HMM model of Lena subband. The parameters of the HMM are estimated using the EM algorithm.

that the performance of the hybrid system is close to the lower bound on the OPTA over a wide range of channel SNRs which suggests it to be a good choice for a JSCC solution.

6. DISCUSSION

We have introduced a novel JSCC solution for transmitting a class of Hidden Markov sources over a power constrained AWGN channel. Our source model approximates the distribution of wavelet image coefficients. We provide the information theoretic bounds on the performance of the proposed JSCC scheme. Our results indicate that the proposed simple hybrid analog/digital coding scheme performs close to the theoretical bounds. We are in a process of extending our work to the case of channel with memory, such as wireless fading channels.

7. REFERENCES

[1] J. Modestino, D.G.Daut, and A. Vickers, "Combined source-channel coding of images using the block cosine transform," *IEEE Trans. on Commun.*, vol. COM-29, pp. 1261–1274, Sep. 1981.

[2] K. Ramchandran and M. Vetterli, *Wireless communications: signal processing perspectives*, ch. Multiresolution joint source-channel coding, pp. 282–329. Upper Saddle River, NJ: Prentice Hall PTR, 1998.

[3] M. Ruf and J. Modestino, "Operational Rate-Distortion performance for joint source and channel coding of images," *IEEE Trans. Image Process.*, vol. 8, pp. 305–320, March 1999.

[4] I. Kozintsev and K. Ramchandran, "A hybrid compressed-uncompressed framework for wireless image transmission," in *Proc. IEEE Int. Conf. Image Proc.*, vol. 2, pp. 77–80, 1997.

[5] N. Phamdo and U. Mittal, "A joint source-channel speech coder using hybrid digital-analog (HDA) modulation," *IEEE Trans. Speech and Audio Proc.*, Submitted 1999.

[6] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Process.*, vol. 41, pp. 3445–3462, December 1993.

[7] S. LoPresto, K. Ramchandran, and M. T. Orchard, "Image coding based on mixture modeling of wavelet coefficients and a fast estimation–quantization framework," *Proc. IEEE Data Compression Conf.*, 1997.

[8] J. Romberg, H. Choi, and R. Baraniuk, "Bayesian tree-structured image modeling using wavelet-domain hidden Markov models," in *SPIE Technical Conference on Mathematical Modeling, Bayesian Estimation, and Inverse Problem*, (Denver, Colorado), July 1999.

[9] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized - gaussian and complexity priors," *IEEE Trans. Info. Theory, Special Issue on Multiscale Analysis*, vol. 45, pp. 909–919, April 1999.

[10] M. Wainwright and E. Simoncelli, "Scale mixtures of gaussians and the statistics of natural images," in *Proc. SPIE Conf. on Vis. Commun. and Image Proc.*, December 1999.

[11] S. Xiao, I. Kozintsev, and K. Ramchandran, "Stochastic wavelet-based image modeling using factor graphs and its application to denoisingg," in *Proc. SPIE Conf. on Vis. Commun. and Image Proc.*, vol. to appear, January 2000.

[12] L. Rabiner and B. Juang, *Fundamentals of speech recognition.* Englewood Cliffs, NJ: Prentice Hall, 1993.

[13] C. Weidmann and M. Vetterli, "Rate-distortion analysis of spike processes," in *Proc. IEEE Data Compression Conf.*, (Snowbird, UT), pp. 82–91, April 1999.

[14] S. Servetto and Y. Bresler, "On the transmission of a class of non-gaussian signals over a power constrained awgn channel," tech. rep., University of Illinois at Urbana-Champaighn, 1998.

[15] M. K. Mıhçak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Processing Letters*, vol. 6, pp. 300–303, December 1999.

[16] M. Crouse, R. Nowak, and R. Baraniuk, "Wavelet-based statistical signal processing using hidden markov models," *IEEE Trans. on Signal Processing*, vol. 42, April 1998.

[17] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-frequency quantization for wavelet image coding," *IEEE Trans. Image Process.*, pp. 677–693, May 1997.

[18] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

[19] R. L. Joshi, T. R. Fischer, and R. H. Bamberger, "Optimum classification in subband coding of images," in *Proc. IEEE Int. Conf. Image Proc.*, (Austin, TX), pp. 883–887, Nov. 1994.