ROBUST SUBBAND IMAGE CODING FOR WAVEFORM CHANNELS WITH OPTIMUM POWER- AND BANDWIDTH ALLOCATION

John M. Lervik and Thomas R. Fischer***

*Department of Telecommunications, Norwegian University of Science and Technology (NTNU)
**School of EECS, Washington State University (WSU)
E-mail: lervik@tele.unit.no, fischer@eecs.wsu.edu

ABSTRACT

Image coding for power- and bandwidth-limited continuous-amplitude channels is considered. We address the problem of power and bandwidth allocation for subband image coding where the goal is to minimize the overall end-to-end distortion. The decomposed image is modeled as a composite source, and an algorithm for allocating power and bandwidth among the subsources of this source is proposed. The algorithm is used to compute estimates of the optimum performance theoretically attainable (OPTA) for subband image communication over a power- and bandwidth-limited AWGN channel. A gracefully degrading subband image coder with dynamic power- and bandwidth allocation is simulated and the performance compared to OPTA and to results of other schemes for robust image communication.

1. INTRODUCTION

During the last decade, subband coding has become a popular method for image compression. A broad range of schemes for coding of the subband signals have been proposed. Often, rate allocation is performed among the subbands, and even within the subbands, to maximize the coding gain and exploit variations in the local image statistics.

There are various methods for rate allocation. Two of the most popular algorithms, producing an overall fixed rate, were proposed by Shoham and Gersho [1] and Westerink et al. [2]. These "equal-slope" algorithms are based on allocating rate to the different sources, such as to minimize the total distortion under an overall rate constraint. The operational rate distortion curve of each source is used in the optimization, and the optimum is found at a point of equal slope of the rate distortion functions, where the total bit budget is exhausted.

On the other hand, Khansari and Vetterli [3] proposed an algorithm for optimum allocation of power to the subsources of a composite source for transmission over a power-constrained channel. The optimum is found at the point of equal slope for the individual distortion power functions.

However, to the authors' knowledge, there exist no algorithms for allocating *both* power and bandwidth. This is an important problem e.g., for mobile image communication and image broadcasting, where the channels typically are both power- and bandwidth-limited. The overall design criteria for a system operating over such channels will be to maximize the quality of the received signal, given the overall channel power and bandwidth constraints.

We propose a method for finding the optimum power and bandwidth allocation among the subsources to minimize the end-to-end distortion. For example, for a subband coder this amounts to finding how much channel power and bandwidth to use for transmitting each subband. A subband image coder optimized using the proposed algorithm is simulated. The performance is compared to the optimum performance theoretically attainable (OPTA) assuming memoryless generalized Gaussian source models, as well as to other published work on image communication.

** The work by the second author was supported, in part, by the National Science Foundation under grants NCR-9303868 and NCR-9627815.

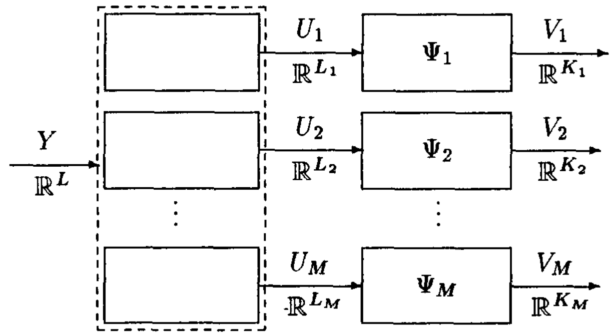

Figure 1. Transmission of a composite source.

2. PROBLEM FORMULATION

Consider the optimum transmission of a *composite* source, Y, over a power- and bandwidth-constrained channel using $r = K/L$ channel symbols per source symbol. This situation is illustrated in Figure 1. The source Y is split into M subsources[1], U_m, $m = 1, 2, \cdots, M$. Each subsource is then individually mapped by a mapping Ψ_m to the modulation signal set, using $r_m = K_m/L_m$ channel symbols per subsource sample and a channel power of S_m per channel symbol.

Assume we know the distortion-power-bandwidth curves, $D_m(S_m, r_m)$, and the probability of occurrence for each subsource, P_m, $m = 1, 2, \cdots, M$, where $\sum_{m=1}^{M} P_m \equiv 1$. Note that $D_m(S_m, r_m)$ is the total distortion, including both quantization distortion and channel distortion, resulting from transmission of subsource m at a rate of r_m channel symbols per source symbol and using channel power S_m.

The optimum power and bandwidth allocation problem can then be formulated as:

$$\min_{\{r_m, S_m\}} D = \sum_{m=1}^{M} P_m D_m(S_m, r_m), \qquad (1)$$

[1] Each subsource might for instance be one subband, or in the case of classification-based coding, one class of samples.

subject to the constraints

$$\sum_{m=1}^{M} P_m r_m \le r, \qquad \sum_{m=1}^{M} P_m r_m S_m \le rS, \qquad (2)$$

and

$$S_m \ge 0, \ m=1,2,\cdots M, \quad r_m \ge 0, \ m=1,2,\cdots M. \qquad (3)$$

The nonnegative constraints on the power and the bandwidth given in Equation 3 are referred to as the *realizability constraints*.

If one has mathematical expressions for the distortion-power-bandwidth functions, the above equations can be solved using the Kuhn-Tucker conditions. However, for practical coding and transmission schemes, an algorithm based on the operational distortion-power-bandwidth performance can be employed.

2.1. A Greedy Algorithm for Power- and Bandwidth Allocation

We propose a modification of the rate allocation algorithm proposed by Westerink et al. [2] to take a channel power constraint into account, in addition to a bandwidth constraint.

Assume we know the operational distortion-power-bandwidth functions, $D_m(S_m, r_m)$, $m = 1,2,\cdots,M$, for all subsources, found by applying either a training set, or synthetic data to the mappings at hand. The algorithm then works as follows: first, we trace the convex hull of the composite distortion-power-bandwidth curve, assuming equally distributed channel power among the subsources, i.e., $S_m = S$, starting with the lowest possible rate for each subsource. It was shown in [2] that this gives the optimum rate allocation for a distortion constraint. Then for the given optimal rate allocation, we trace the convex hull of the composite distortion-power-bandwidth curve for a fixed bandwidth allocation, as found in the previous stage of the algorithm.

The algorithm can be described as follows:

1. Given r and S/N. Allocate the lowest possible rate for each subsource. Let the rate allocated for subsource m be r_m, $m = 1,2,\cdots,M$. Furthermore, let all subsources have the same amount of channel power $S_m = S$, $m = 1,2,\cdots,M$.
2. For subsource m, $m = 1,2,\cdots,M$, calculate
$$s_m = \max_{\{k_m \mid k_m > r_m\}} \left[\frac{(k_m - r_m)}{D(S,k_m) - D(S,r_m)} \right]. \qquad (4)$$
3. Find the subsource for which the maximum of s_m was obtained:
$$n = \operatorname*{argmax}_{\{m \mid m=1,2,\cdots,M\}} s_m, \qquad (5)$$
and update the rate of subsource n: $r_n = k_n$.
4. If the current rate is sufficiently close to the desired rate, i.e., $r - \sum_{m=1}^{M} P_m r_m \le \epsilon$, go to Step 6. Otherwise, continue.
5. Repeat Steps 2, 3, and 4, but do Step 2 only for subsource n.
6. Given rates r_m, $m = 1,2,\cdots,M$. Allocate the lowest possible power for each subsource. Let the power allocated for subsource m be S_m, $m = 1,2,\cdots,M$.
7. For subsource m, $m = 1,2,\cdots,M$, calculate
$$s_m = \max_{\{K_m \mid K_m > S_m\}} \left[r_m \frac{(K_m - S_m)}{D(K_m,r_m) - D(S_m,r_m)} \right]. \qquad (6)$$
8. Find the subsource for which the maximum of s_m was obtained:
$$n = \operatorname*{argmax}_{\{m \mid m=1,2,\cdots,M\}} s_m, \qquad (7)$$
and update the power of subsource n: $S_n = K_n$.
9. If the current power is sufficiently close to the desired power, i.e., $rS - \sum_{m=1}^{M} P_m r_m S_m \le \epsilon$, stop. Otherwise, continue.
10. Repeat Steps 7, 8, and 9, but do Step 7 only for subsource n.

Notice that since the power and rate allocation is found in a two-step process, and not jointly, the algorithm does not guarantee a global optimum.

2.2. OPTA for Image Transmission

Consider finding estimates of OPTA for image transmission over an additive white Gaussian noise (AWGN) channel based on the assumptions: 1) subband signal decomposition, *ii*) separate coding of each subband, and *iii*) each subband is well modeled by a memoryless generalized Gaussian distribution (GGD). ML estimates of the GGD parameters are computed for each subband. Then the greedy algorithm suggested above is used to find the estimates of OPTA.

As an example, the monochrome 512 × 512 images "Lenna," "Barbara," and "goldhill" were decomposed by the 8 × 8 uniform nonunitary parallel FIR filter bank "32I" [4] and the GGD parameters were estimated for each subband. Furthermore, the distortion rate function was computed using Blahut's algorithm [5] for GGDs with shape parameter $\eta \in \{0.4, 0.5, \cdots, 2.5\} = \mathcal{A}_\eta$. Then, for each subband, the ML estimate of η was quantized to the nearest value in $\mathcal{A}_\eta$, and the resulting distortion-power-bandwidth function employed for power and bandwidth allocation using the greedy algorithm.

Note that by using the channel capacity as an estimate for the channel rate, the second iteration of the iterative algorithm is avoided, as the optimal power allocation is $S_m = S$, according to the "water-filling" principle for an AWGN channel [6].

The estimated OPTA for transmission of the subbands can be converted to peak signal-to-noise ratio (PSNR) results for the reconstructed image by compensating for the effect of the nonunitary filter bank "32I" [4], and the difference between the SNR and the PSNR.

Figure 2 shows the estimated OPTA for the 512 × 512 monochrome images "Lenna", "Barbara," and "goldhill," at $r = 0.125$ channel symbols per pixel. For the employed filter bank, and white channel noise, the difference between the SNR for the subband signals and the PSNR for the reconstructed image equals 17.70 dB, 16.44 dB, and 18.06 dB for the images "Lenna," "Barbara," and "goldhill," respectively.

3. SUBBAND IMAGE CODING USING CLASSIFICATION

In the previous section, each of the 64 subbands was treated as a separate source, and power and bandwidth were allocated to the different sources to minimize the overall distortion based on memoryless GGD models for each source.

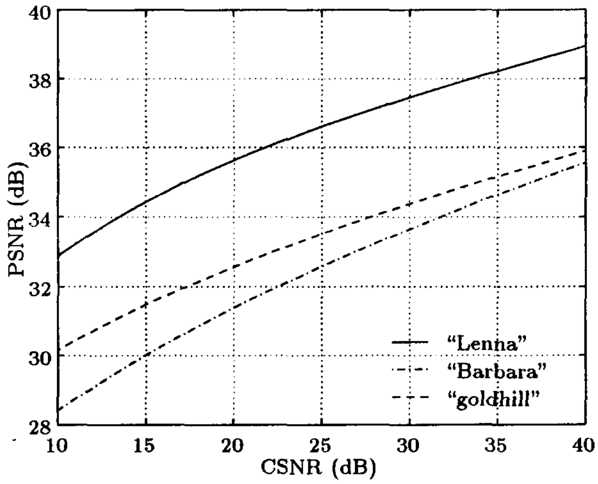

Figure 2. Estimated OPTA for separate coding and transmission of each subband assuming memoryless GGD subbands and an AWGN channel. $r = 0.125$ **channel symbols per pixel**

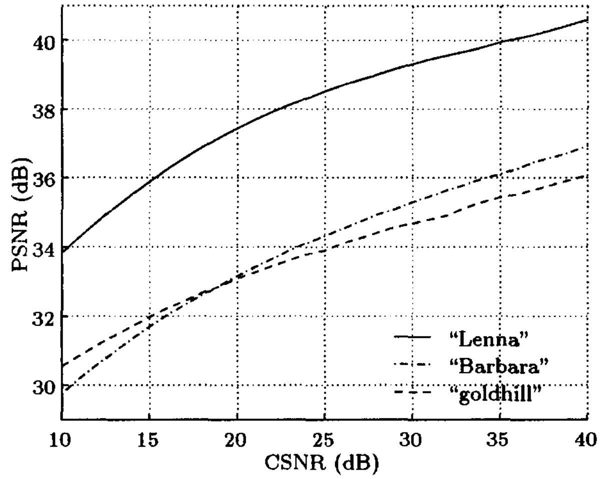

Figure 3. Estimated OPTA for separate coding and transmission of each class assuming memoryless GGD classes and an AWGN channel. $r = 0.125$ **channel symbols per pixel.**

Classification is a powerful alternative to simply applying rate allocation to the different subbands.

3.1. Subband Image Coder

A subband image coder based on the 8×8 uniform nonunitary parallel FIR filter bank "32I" [4] was simulated. All subbands were treated as one composite signal and blocks of 8×8 subbands samples were classified into $M = 5$ classes, $U_1, U_2, \cdots, U_5$. The block energy (l_2-norm) was used for the classification, and each class could contain blocks from any subband.

The subband samples corresponding to the low-energy class 1 were set to zero, i.e., $r_1 = 0$. Then the samples belonging to each of the 4 higher energy classes were mapped to a pulse amplitude modulation (PAM) signal set by the mappings Ψ_m, $m = 2, 3, \cdots, 5$, constructed either as a combination of vector quantization and index assignment or as a direct linear mapping [7]. Samples in classes 2, 3, and 4 were applied to vector quantizers and the corresponding indices assigned to a point in a 81-PAM signal set, using the following number of channel symbols per pixel: $r_2 = 1/4$, $r_3 = 1/2$, and $r_4 = 1/2$. Class 5 samples were left unquantized and mapped linearly one-by-one to continuous amplitude PAM symbols, i.e., $r_5 = 1$.

If there is an error in the block classification table in the receiver, the synchronization in the decoder may be lost, resulting in a break-down of the decoded image quality. Consequently, the block classification table was error protected with a Reed-Solomon (RS) [8] channel code and sent as side information.

At the receiver side, a maximum likelihood (ML) demodulator was used, and the image signal was reconstructed through inverse operations as compared to the encoder side: First, the block classification table was decoded and used to select the correct inverse index assignment function and vector quantizer pair for each block of samples. Then, the reconstructed subband image was applied to the synthesis filter bank, obtaining the reconstructed image.

3.1.1. Block Classification

A slightly modified version of the power and bandwidth allocation algorithm proposed in Section 2.1. was used for block classification.

Consider classification of J blocks into M classes ($J \gg M$), given a mapping, Ψ_m, and a fixed prescribed rate, r_m, for each class m, $m = 1, 2, \cdots, M$. First, we compute the variance-normalized operational distortion-power-bandwidth curves, $D_m(S_m, r_m)$, for all classes, assuming memoryless Laplacian source models. Hence, the distortion for coding block j with the mapping of class m is estimated by $\hat{\sigma}_j^2 \cdot D_m(S_m, r_m)$, where $\hat{\sigma}_j^2$, $j \in \mathcal{J} = \{1, 2, \cdots, J\}$, is the block energy.

Then the following algorithm can be applied for power- and bandwidth allocation [7]: Initially, all blocks belong to class 1. Then the blocks are allocated rate using the class-wise operational distortion-power-bandwidth curves for a fixed channel power, $S_m = S$ $m = 1, 2, \cdots, M$. Finally, power is allocated among the classes, starting with a zero power allocation for all $M = 5$ classes, and iteratively allocating power to the class giving the largest relative distortion decrease.

3.2. OPTA for transmission of $M = 5$ classes

Using $M = 5$ classes and a GGD model for each class, the power- and bandwidth were optimally allocated using the algorithm presented previously, yielding the estimates of OPTA shown in Figure 3. The classification procedure was optimized for $S/N = 40$ dB for the mapping configuration used in the practical image coder, and then used for every other CSNR. Note that the side information is not included in Figure 3. The maximum side information, which must be error-protected in a practical system, is $(\log_2 5)/64 = 0.036$ bits per pixel.

By comparing Figure 2 (using 64 classes as the 64 subbands) to Figure 3 (using 5 classes drawn from the 64 subbands) it is evident that the three images offer a classification gain ranging from less than 0.5 dB for the "goldhill" image, to 1.0 to 1.5 dB for the images "Lenna" and "Barbara."

4. SIMULATION RESULTS

The PSNR performance of the proposed coder for a selected set of channel symbol rates per pixel is displayed in Figure 4. The classification procedure was optimized for $S/N = 40$

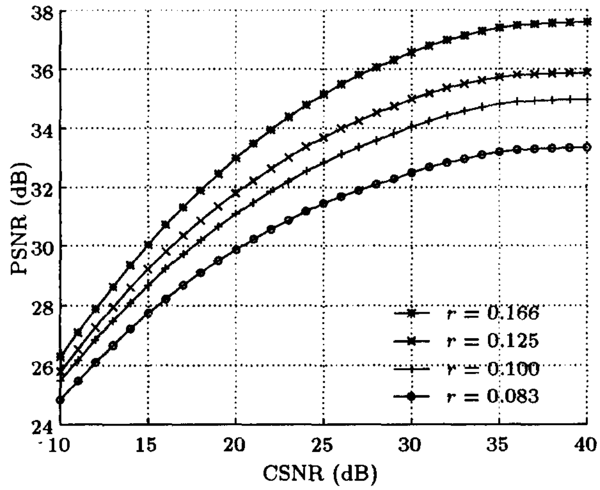

Figure 4. Simulation results for the proposed coder. Image: "Lenna."

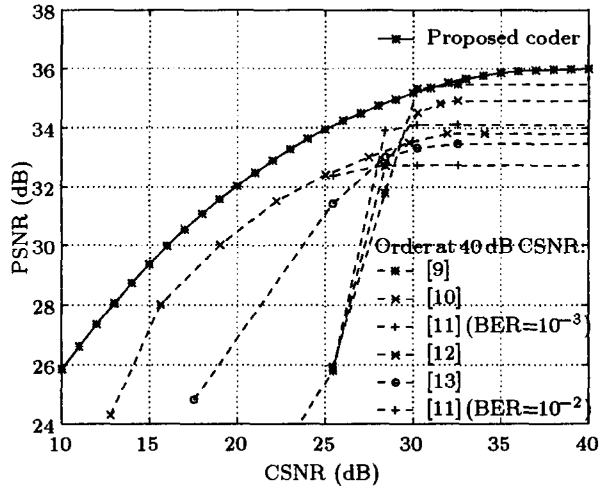

Figure 5. Comparison to other published work. r = 0.125 channel symbols per pixel. Image: "Lenna."

dB. At $r = 0.125$ channel symbols per pixel, the PSNR performance of the proposed coder is approximately 4 dB below OPTA (cf. Figure 3) at high CSNRs, increasing to 8 dB at low CSNRs. The larger deviation at low rates is partly due to the optimization of the coder for high CSNRs.

4.1. Comparisons to other work

Other schemes for robust image coding have been presented by Mohdyusof and Fischer [9], Hung and Meng [10], Tanabe and Farvardin [11], Ruf and Filip [12], and Chen and Fischer [13]. All these schemes report PSNR results as a function of the bit error rate (BER) for transmission over a binary symmetric channel (BSC) at a coding rate of 0.5 bits per pixel.

To compare the PSNR performance of the proposed coder to the other coders for image transmission over an AWGN channel, an appropriate channel symbol rate has to be chosen, for which the CSNR corresponding to a certain BER can be computed.

We assume PAM signaling with equiprobable channel symbols and Gray coding of the indices in the modulation signal set. Thus, for coding at R bits per pixel, a signaling scheme offering R/r bits per channel symbol is required to attain a symbol rate of r channel symbols per pixel. Thus, for PAM signaling the size of the modulation signal set must be $2^{R/r}$.

As an example, the performance of the proposed coder at $r = 0.125$ channel symbols per pixel was compared to the performance of the other coders assuming 16-PAM signaling. Comparisons of the PSNR performance as a function of the CSNR for the image "Lenna" are displayed in Figure 5. Note that the coding results in [12] and [13] are obtained by coders optimized for *each* CSNR, while the proposed coder and the coders in [9], [10], and [11] are optimized for a target CSNR and applied to the CSNRs of interest.

At a CSNR of 24.07 dB the channel capacity of an AWGN channel equals 4 bits per sample. Thus, ideally, subband coding at 0.5 bits per pixel, optimized for error-free transmission conditions, could operate down to this CSNR using $r = 0.125$ channel symbols per pixel.

REFERENCES

[1] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1445–1453, Sept. 1988.

[2] P. H. Westerink, J. Biemond, and F. Müller, "An optimal bit allocation algorithm for subband coding," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, pp. 757–760, 1988.

[3] M. Khansari and M. Vetterli, "Layered transmission of signals over power-constrained wireless channels," in *Proc. Int. Conf. on Image Processing*, vol. III, (Washington, D.C, USA), pp. 380–383, IEEE, Oct. 1995.

[4] S. O. Aase, *Image Subband Coding Artifacts: Analysis and Remedies*. PhD thesis, The Norwegian Institute of Technology, 1993.

[5] R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 460–473, July 1972.

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

[7] J. M. Lervik, *Subband Image Communication over Digital Transparent and Analog Waveform Channels*. PhD thesis, Norwegian University of Science and Technology, 1996.

[8] R. E. Blahut, *Digital Transmission of Information*. Reading, Massachusetts: Addison Wesley Publishing Company, Inc., 1990.

[9] Z. Mohdyusof and T. R. Fischer, "Subband image coding using a fixed-rate lattice vector quantizer," in *Proc. Int. Conf. on Image Processing*, vol. I, (Washington, D.C., USA), pp. 101–104, IEEE, Oct. 1995.

[10] A. C. Hung and T. H.-Y. Meng, "Error resilient pyramid vector quantization for image compression," in *Proc. Int. Conf. on Image Processing*, vol. I, (Austin, TX, USA), pp. 583–587, IEEE, Nov. 1994.

[11] N. Tanabe and N. Farvardin, "Subband image coding using entropy-coded quantization over noisy channels," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 926–943, June 1992.

[12] M. J. Ruf and P. Filip, "Joint source and channel coding applied to the pyramid vector quantizer," in *Proc. Int. Symp. on Information Theory*, p. 391, IEEE, 1993.

[13] Q. Chen and T. R. Fischer, "Image coding using robust quantization for noisy digital transmission," *IEEE Trans. Image Processing*, 1996. (Submitted).