

Rate-distortion-optimal Parameter Choice in a Wavelet Image Communications System

Javier Garcia-Frias, Dan Benyamin and John D. Villasenor

Electrical Engineering Department
University of California, Los Angeles
jgarcia, benyamin, villa@icsl.ucla.edu

Abstract

We describe the methodology and results of parameter optimizations performed with the aid of analytical expressions for the channel-induced distortion in a wavelet-based image communications system. $R(D)$ curves are calculated independently for each subband, with each point on each curve representing (in addition to rate and distortion) an optimal choice of quantizer step size and channel code rate. A global optimization procedure is then performed to identify the best operating point on the $R(D)$ curve for each subband.

1 Introduction

In [1] we presented an analytical framework for describing the distortion in an image communications system that includes wavelet transformation, uniform scalar quantization, run length coding, entropy coding, forward error control, and transmission over a binary symmetric channel. This makes it possible to apply analytical tools to bear on a problem that has previously been approached almost exclusively from an experimental standpoint. In particular, the formulation in [1] offers the possibility to gather and utilize rate-distortion data which would have involved impractical computational costs to acquire in a purely simulation-based approach.

We present here the results of an investigation into parameter choices which give near rate-distortion optimal performance on a per subband basis, and good performance on an image basis. The resource allocation is not fully optimal because we are calculating the total distortion in the wavelet domain as opposed to the image domain, and because the $R(D)$ curves are not strictly convex due to several factors including the granularity of the error correcting codes. The basic idea is to calculate a rate-distortion curve for each subband, and then to apply the algorithm of Westerink *et al* [3] to choose an operating point each subband such that the overall distortion is minimized.

The coder used here includes a dyadic, 5-level wavelet transform using the 9/7 filters, followed by uniform scalar quantization using a quantizer with a step size equal to q and a dead zone of width $2q$ at

the origin. The step size q is constant within any given subband, but is allowed to vary between subbands as part of the optimization. The sequence of integers produced by performing a raster scan of the quantized data in each subband is input to a run-length coder which outputs two bitstreams, one containing runs only and the other containing levels only. Within the level stream, each integer is represented using b bits, where b is the fewest number of bits that will permit representation of all possible output levels. The b -bit symbol for each level includes one bit to identify the sign and $b - 1$ bits representing the magnitude. A similar approach (without a sign bit) is used for representing runs. In addition, the runs are subject to an entropy coding step using exp-Golomb codes [4, 5, 6] is applied. The bitstreams representing the runs and levels are placed into fixed length packets of size N_Z and N_L respectively. The data in each packet are protected using a BCH code, with fixed length headers used at the start of each packet to provide synchronization information [1]. The number of information bits in each packet is a function of the channel code rate, or equivalently, of the number of errors, t_Z and t_L , correctable by the BCH codes used to protect run packets and level packets respectively. As with the quantizer step size, t_Z and t_L will be chosen on a subband specific basis using rate-distortion considerations.

When this coding algorithm is used, it is possible to derive formulas (not presented here for space reasons, but available in [1]) that accurately predict the expected channel-induced distortion as a function of the step size q , the correction capability t_Z and t_L of the BCH codes, and other parameters. The distortion description is purely analytical only when an analytical description of the source is available. For example, the distortion formulas require data such as the expected run length, which can be calculated analytically for a source model such as a i.i.d. generalized Gaussian, but must be found experimentally if the input data is a list of coefficients from a wavelet subband.

2 Rate-distortion Calculations

The goal is to make the parameter choice that minimizes the total distortion subject to a constraint on the overall bit rate. This means choosing, for each subband, a rate in bits/pixel, the quantization step size q , and the protection levels t_Z and t_L to use in the run and level BCH codes. There is some potential confusion because the term “rate” can describe the coding rate in bits per pixel for a given subband, the BCH code rate for the run packets, and the BCH code rate for the levels. These three rates are of course distinct, and are all be subject to optimization. To make the problem tractable we fix the packet sizes to be $N_Z = N_L = 255$, though it would also be possible to make packet size a parameter to be optimized on a subband by subband basis.

We begin the process by calculating a set of rate distortion curves for all of the subbands. To find each point on the $R(D)$ curve for a given subband requires that the distortion equations be applied for each possible combination of t_Z , t_L , and q . Although the step size q appears in the equations for channel-induced distortion, it is also of course the key factor in source coding distortion. To account for the source distortion we introduce an additive term to the channel distortion equations to obtain an expression giving the overall (source plus channel) distortion. In theory there is a small cross term that should also be considered, but in practice the effect of this term is small under the conditions used here, and treating the distortion sources additively is reasonable. Since we are using BCH codes of length 255, there is a relatively small set of possible combinations of t_Z and t_L , and even when we consider approximately 100 different possible choices for q , it is possible using the distortion equations to perform a search to find the combination of (t_Z, t_L, q) that will result in the lowest overall distortion subject to a subband rate constraint. This minimization must be performed at each rate of interest, and gives a set of points approximating the $R(D)$ curve for the subband.

Because of the granularity imposed by the restricted set of available BCH code rates and by the fixed packet size, the achievable $R(D)$ values are not guaranteed to lie along a strictly convex path. This is problematic when applying the algorithm of Westerink because the assumption of a monotonically increasing slope is not met. To handle this we fit the calculated $R(D)$ points using a convex function.

3 Results and Discussion

After fitted $R(D)$ curves were determined for each subband, we applied the algorithm of Westerink *et al.* [3] to obtain the choice of subband rates and corresponding distortions giving the lowest total distortion subject to a total rate constraint. Once the operating point on each fitted $R(D)$ curve is identified, for the actual coding we used the closest experimentally determined (i.e. achievable) $R(D)$ point. Although the distortion minimization was performed in the wavelet

domain, we have found that this corresponds well, though not perfectly, to distortion in the image domain.

The rate-distortion optimizations were performed on 4 image/channel combinations: The 512 x 512 Lena, Goldhill, and Barbara images for a binary symmetric channel (BSC) with bit error rate (BER) 10^{-2} , and the Lena image for a BSC with BER 10^{-3} . For 10^{-3} our results may be overly optimistic – this is because the formulas used to obtain the distortion assume that all errors are detected, which is not always the case for the high-rate codes used at 10^{-3} . Figure 1 shows the wavelet domain PSNR curves for these images as a function of overall (including source coding, channel coding, packetization, header information, etc.) coding rate.

Figure 2 shows, for the same images, the fraction of the total bits that are used for source coding as a function of bit rate. Two characteristics of these curves are particularly noteworthy: First, there is a very high degree of consistency for the source bit percentage for all three of the images coded at BER 10^{-2} . By contrast, in the Lena image at BER 10^{-3} the percentage, as expected, is much higher.

Table 1 gives the optimal parameter settings for each subband for the Lena image at .25 and .5 bpp and BER's of 10^{-2} and 10^{-3} . The subband numbering scheme is shown in Figure 3. The three numbers shown in each table entry represent (from left to right) the fraction of the total number of bits to allocate to the corresponding subband, the fraction of the bits in that particular subband used for source coding, and the quantizer step size. This helps to clarify the extent to which channel code rates vary with subband in an optimized unequal error protection approach.

The graphs and figures presented above, coupled with other analysis using the framework presented here, lead to the following observations:

1. The overall percentage of bits to allocate to source as opposed to channel coding is primarily a function of the channel BER, not of the image. Furthermore, this percentage is approximately constant across the .25 to .5 bits/pixel coding range. The dependence of this percentage on BER is expected. However, the almost complete lack of dependence on the image is more noteworthy.

2. Comparisons between our PSNR numbers and those previous reported in the literature must be made with caution since our results are in the wavelet domain while the other cited results are in the image domain. With this caveat, we note that at a BER of 10^{-2} the PSNR results we obtain are typically 2-3 dB below those provided by the Said and Pearlman algorithm [7] for the noiseless case. The differences are smaller at .25 bpp than at .5 bpp. Our PSNR values are close to, but lower than (by a few tenths of a dB) the source/channel results at BER 10^{-2} reported in [2].

3. Our results appear to be consistent with the theoretical treatment of source/channel coding in by Hochwald and Zeger in [8]. The authors of [8] establish theoretical bounds on the performance achievable in the limiting case of a vector quantizer (VQ) of fixed finite dimension and large resolution (i.e. transmission rate). Since [8] considers VQ followed by a channel coder while we are using scalar quantization, run-length, and entropy coding, the comparison must be treated with caution. However, our result that approximately 70-75% of the bits are used for source coding at BER 10^{-2} , rising to approximately 85% at 10^{-3} corresponds, at both BERs, to a VQ in [8] with vector dimension of between 16 and 64. This is consistent with our use of packets of size 255, each of which contains several tens of source samples.

4. For a BER of 10^{-2} , approximately 96-98% of the total distortion in the optimal resource allocation is due to the lossy portion of the source coding, with the remaining distortion due to channel errors. For the single image we examined at a BER 10^{-3} , the percentage of total distortion due to source coding was over 98% for most bit rates examined. For both 10^{-2} and 10^{-3} the distortion percentages appear to be approximately constant as a function of overall image coding rate in bits/pixel. The dominance of source distortion is partially explained by the relative costs of reducing source and channel distortion. For the BSC, allocating a relatively few number of additional bits to channel coding can reduce the expected residual (after error correction) bit error rate by an order of magnitude. If these bits are instead applied to the source coding, the resulting improvement in distortion will be much smaller. Since obtaining good protection over the BSC is relatively cheap, it is much safer to err on the side of giving the channel coder slightly more bits than is necessary for a balanced distortion contribution.

5. The subband specific parameter choices that resulted from the optimal allocation were quite haphazard. This is somewhat apparent in Table 1, and becomes more so when plots of the allocation as a function of overall bit rate are examined. With the exception of the step sizes, which do respond in intuitive ways to changes in BER and overall bit rate, it is difficult to extract general rules for accurately choosing the parameters shown in Table 1. Subband-specific optimization of the quantization step size seems to be much more important than subband-specific optimization of the rate of the error correcting code. Though we allowed unequal error protection, the marginal additional cost to protect all subbands at the rates needed by the most vulnerable subbands is not large. Equal error protection is vastly simpler, and apparently only slightly less attractive from a PSNR standpoint, than unequal error protection in our coding algorithm.

4 Conclusions

We have presented the results of rate distortion optimization for source and channel coding of wavelet-

transformed image data. Formulas from [1] were used to determine the optimal quantization step size and error protection to use for each subband for each image as a function of channel BER. While joint source/channel coding in image processing has often included the idea of applying error protection to image subbands in accordance with their importance, based on the results here we believe that joint source channel coding approaches that assume fixed rate FEC deserve more investigation, particularly for image source coding algorithms involving run-length coding. Use of equal error protection makes it possible to leverage existing and emerging communications networks which often have error handling built in at network layers that are inaccessible to a source codec. Even in a system in which FEC is not under the control of the image coder there is still ample opportunity to optimize the quantizer, and to employ resynchronization markers and other strategies to match an expected residual bit error rate.

An advantage of our approach in contrast with previous work in joint source/channel image coding is that it supports rate distortion optimizations that would be prohibitively costly to carry out using brute force simulation. The disadvantage that accompanies this, however, is that the framework, and possibly the results presented here are specific to the source and channel coding techniques we are using. More work is needed to determine the extent to which these results apply in other joint source channel coding environments.

References

- [1] J. Garcia-Frias and J.D. Villasenor, "An Analytical Treatment of Channel-induced Distortion in Run Length Coded Subbands," *Proc. DCC '97*, Snowbird, UT, March 1997.
- [2] P. Greg Sherwood, K. Zeger, "Progressive Image Coding on Noisy Channels," *Proc. DCC '97*, Snowbird, UT, March 1997.
- [3] P. H. Westerink, J. Biemond, D. E. Boeke, "An optimal bit allocation algorithm for subband coding," *Proc. ICASSP*, pp. 757-760, 1988.
- [4] J. Wen, P. Meshkat, and J.D. Villasenor, "Structured trees for lossless coding of quantized wavelet coefficients," *Proc. 30th Annual Asilomar Conf. on Sig., Syst., and Comp.*, Pacific Grove, CA, November 1996.
- [5] J. Wen and J.D. Villasenor, "A class of reversible variable length codes for robust image and video coding," *Proc. ICIP 97*.
- [6] J. Teuhola, "A Compression Method for Clustered Bit-Vectors," *Information Processing Letters*, vol. 7, pp. 308-311, October 1978.
- [7] A. Said and W.A. Pearlman "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, June 1996.

[8] B. Hochwald, K. Zeger, "Tradeoff Between Source and Channel Coding," to appear, *IEEE Trans. on Information Theory*.

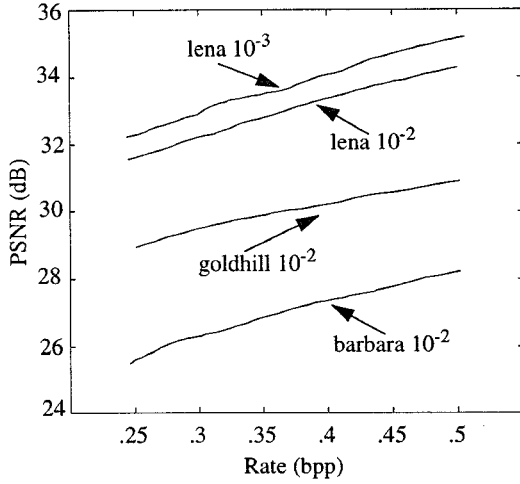


Figure 1: Wavelet domain PSNR for the images Lena ($BER = 10^{-2}$ and 10^{-3}), Goldhill ($BER = 10^{-2}$) and Barbara ($BER = 10^{-2}$).

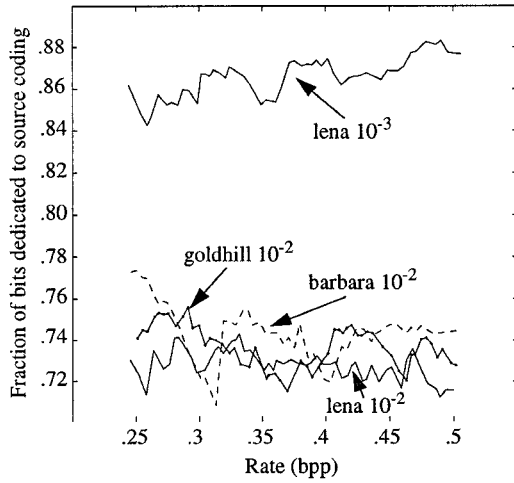


Figure 2: Fraction of the total number of bits used for source coding for the images Lena ($BER = 10^{-2}$ and 10^{-3}), Goldhill ($BER = 10^{-2}$) and Barbara ($BER = 10^{-2}$).

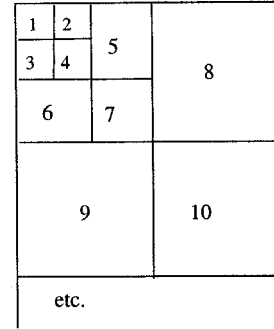


Figure 3: Subband numbering scheme.

Sub No.	Lena BER 0.01 0.5 bpp	Lena BER 0.01 0.25 bpp	Lena BER 0.001 0.5 bpp	Lena BER 0.001 0.25 bpp
1	.023, .67, 21	.043, .64, 41	.019, .80, 21	.031, .88, 41
2	.019, .70, 21	.039, .70, 21	.016, .88, 21	.031, .88, 21
3	.019, .70, 14	.031, .75, 26	.016, .88, 14	.027, .86, 26
4	.019, .70, 16	.031, .75, 32	.016, .88, 16	.031, .88, 16
5	.056, .72, 25	.097, .71, 40	.055, .91, 18	.086, .88, 32
6	.041, .76, 21	.054, .78, 35	.041, .80, 18	.047, .90, 35
7	.041, .71, 21	.051, .76, 39	.041, .92, 15	.054, .84, 37
8	.123, .70, 22	.163, .66, 41	.125, .89, 16	.128, .83, 41
9	.058, .64, 29	.078, .75, 39	.068, .88, 18	.078, .88, 32
10	.060, .66, 27	.082, .74, 39	.064, .92, 17	.078, .91, 32
11	.206, .76, 19	.191, .75, 40	.183, .90, 18	.202, .83, 34
12	.097, .67, 23	.078, .74, 42	.080, .82, 23	.089, .86, 36
13	.095, .73, 20	.054, .71, 41	.107, .85, 15	.066, .89, 34
14	.136, .74, 22	Not sent	.136, .91, 19	.035, .83, 45
15	.004, .35, 40	.008, .35, 40	.029, .84, 21	.008, .44, 39
16	Not sent	Not sent	.004, .38, 25	.008, .38, 25

Table 1: Optimal parameter settings in each subband for the Lena image at .25 and .5 bpp, for BER's of 10^{-2} and 10^{-3} . The three numbers shown in each entry of the table represent (from left to right) the fraction of the total number of bits allocated to the corresponding subband, the fraction of bits used for source coding in that particular subband, and the quantizer stepsize.