PROGRESSIVE WAVELET IMAGE CODING BASED ON A CONDITIONAL PROBABILITY MODEL

Robert W. Buccigrossi
GRASP Laboratory
Computer and Information Science Dept.
University of Pennsylvania
Philadelphia, PA 19104

Eero P. Simoncelli
Center for Neural Science, and
Courant Inst. of Mathematical Sciences
New York University
New York, NY 10012

ABSTRACT

We present a wavelet image coder based on an explicit model of the conditional statistical relationships between coefficients in different subbands. In particular, we construct a parameterized model for the conditional probability of a coefficient given coefficients at a coarser scale. Subband coefficients are encoded one bitplane at a time using a non-adaptive arithmetic encoder. The overall ordering of bitplanes is determined by the ratio of their encoded variance to compressed size. We show rate-distortion comparisons of the coder to first and second-order theoretical entropy bounds and the EZW coder [1]. The coder is inherently embedded, and should prove useful in applications requiring progressive transmission.

Orthonormal wavelet decompositions have proven to be extremely effective for image compression [2, 3, 4, 5, 1]. We believe there are several statistical reasons for this success. Similar to the Fourier transform, wavelets are quite good at decorrelating the second-order statistics of natural signals. The resulting redistribution of variance leads to a reduction in the total first-order entropy of the coefficients relative to the entropy of the original image pixels.

In addition to the redistribution of variance, wavelet subband statistics are significantly non-Gaussian. This property has been exploited in compression, noise removal and texture synthesis (eg., [6, 7, 8, 9, 10]). This observation should be contrasted with frequency-based decompositions, which have approximately Gaussian statistics. Since the Gaussian is the maximal-entropy distribution for a given variance, wavelet-based coders are able to achieve higher degrees of compression than frequency-based coders such as JPEG.

Finally, wavelet decompositions exhibit additional statistical regularity beyond the point statistics of the subbands. In particular, it is clear from casual inspection that the subbands of a wavelet transform are similar in appearance, and are therefore not statistically independent (see figure 2). Such dependencies have recently been utilized in image compression. Shapiro [1] constructed a coder that exploits the fact that a zero coefficient in a bitplane of a subband is likely to indicate a tree of zero coefficients at the same location in all finer scale subbands. Pentland et al. [11, 12] used a vector quantizer to predict fine scale coefficients from coarser scales. Quite recently, several authors [13, 14] have used adaptive arithmetic coding to capture conditional statistics between subbands.

In this paper, we attempt to more directly take advantage of the joint statistics between wavelet subbands at adjacent scales and orientations. We explore the empirical nature of these statistics, develop a simple statistical model, and use it to implement a progressive image compression algorithm.

1. POINT AND JOINT SUBBAND STATISTICS

Figure 1 shows plots of the point statistics (histograms) for a number of subbands of a typical wavelet decompositions. The figure also shows that these statistics are reasonably well fit [6, 10] by distributions of the form:

$$\mathcal{P}_x(x) \propto e^{-|x/s|^p}. \tag{1}$$

The distribution is zero-mean and symmetric, and the parameters $\{s, p\}$ are directly related to the second and fourth moments. Specifically (after consultation with an integral table) one obtains:

$$\sigma^2 = \frac{s^2\,,\left(\frac{3}{p}\right)}{,\left(\frac{1}{p}\right)}, \qquad \kappa = \frac{,\left(\frac{1}{p}\right),\,,\left(\frac{5}{p}\right)}{,^2\left(\frac{3}{p}\right)}, \tag{2}$$

where $,(x) = \int_0^\infty t^{x-1}e^{-t}\,dt$, the well known "gamma" function, and σ^2 and κ are the variance and kurtosis (fourth moment divided by squared variance). Note that by modeling the statistics in this simplistic fashion, we are assuming both independence and stationarity of the subband coefficients, both of which are generally incorrect. Nevertheless, the model provides a reasonable fit to the statistics of many natural images in our database, as can be seen in the examples of figure 1.

Now consider the relationship between wavelet subbands. Figure 2 shows the magnitudes of wavelet coefficients in a four-level pyramid decomposition. It is visually apparent that coefficients with large magnitude tend to occur at the same relative locations in subbands at different scales. This is true when comparing subbands of the same orientation, and also holds (to a lesser extent) across orientations.

To capture this relationship more precisely, consider the joint statistics for two coefficient subbands. Figure 3 shows the conditional statistics for two horizontal subbands at adjacent scales. That is, we plot the conditional histogram $\mathcal{P}(\log_2(|Y|)|\log_2(|X|))$ where X is the "parent" coefficient (i.e., at the coarser scale) and Y is the corresponding "child" coefficient (i.e., at the finer scale, but at the same relative

RWB was supported by NSF Graduate Fellowship GER93-55018. Most of this research was performed while EPS was in the CIS Department at the University of Pennsylvania, and was partially supported by ARO/MURI DAAH04-96-1-0007 and NSF CAREER grant 9624855.

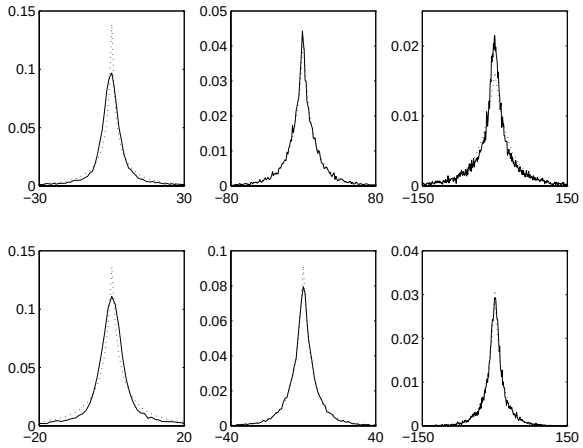

Figure 1. Examples of subband coefficient histograms (solid lines) fitted with the function in equation (1) (dotted lines). The subbands correspond to three different images ("Lena", "Goldhill", "Baboon"), and two different orientations (top row: horizontal, bottom row: diagonal).

location and orientation). When the magnitude of the parent coefficient is large, the expected value of the magnitude of the child appears to be linearly related to the magnitude of parent. When the parent magnitude is small, the magnitude of the child is independent of the parent magnitude.[1] Note that we are considering coefficient *magnitudes*: A simple linear predictor for the raw coefficients will typically fail because the coefficient signs are essentially uncorrelated.

Empirically, the statistics shown in Figure 3 seem to be quite regular across images. We model the mean of the conditional statistics as:

$$E[\log_2(Y)|\log_2(X) = x] = \log_2(2^x\sigma + n) \qquad (3)$$

where σ and n are parameters that can be determined by least squares estimation. Furthermore, the vertical cross-sections of the statistics in figure 3 are highly regular, and after normalization for mean and variance, can be fit by a single lookup table. Figure 4 shows these normalized conditional distributions for a set of natural images. Despite the variety of image content (e.g., we include a face, a landscape, and a CT scan), these cross sections are quite similar.

2. IMPLEMENTATION OF A PROGRESSIVE IMAGE CODER

We have implemented an image coder based on the statistical model described above. We refer to this as the "Embedded Predictive Wavelet Image Coder" (EPWIC). The coder is based on a separable orthonormal wavelet decomposition using 9-tap symmetric (linear-phase) filters designed in [15]. The encoded bit stream begins with a header containing the dimensions of the image and the number of pyramid levels. Coefficients are quantized to 10 bits, which are sent as bitplanes. Subband bitplanes are ordered according to a greedy algorithm, which chooses the bitplane with the highest ratio of encoded variance to encoded size (in bits). The sign bit for each coefficient is sent just after the first "on" bit, as in [14]. For each subband, the quantities (s, p) are

[1] We believe that this occurs because low-amplitude coefficients are dominated by quantization and other noise sources.

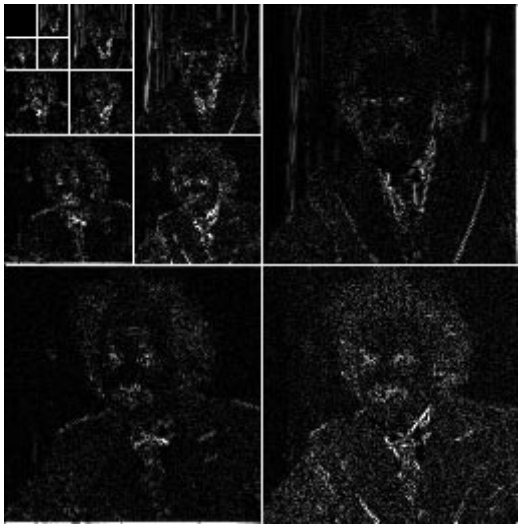

Figure 2. Example coefficient magnitudes of a wavelet decomposition. Shown are absolute values of subband coefficients in a 4-level separable wavelet decomposition of the "Einstein" image. Note that high-magnitude coefficients at adjacent scales tend to be located in the same spatial positions.

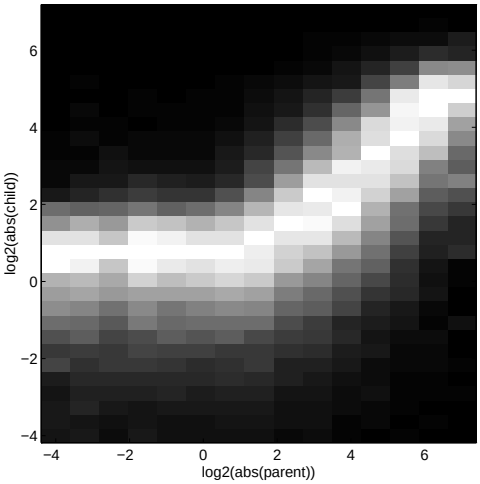

Figure 3. Conditional histogram of coefficient magnitudes in horizontal subbands of two adjacent scales of the "boats" image. Intensity corresponds to probability, except that each column has been independently rescaled to fill the full range of intensities. The right side of the distribution is concentrated about a unit-slope line, indicating that child (fine-scale) coefficients are roughly proportional in magnitude to parent (coarse-scale) coefficients. The left side of the distribution is concentrated about a horizontal line, indicating child coefficients are independent of parent coefficients in this region.

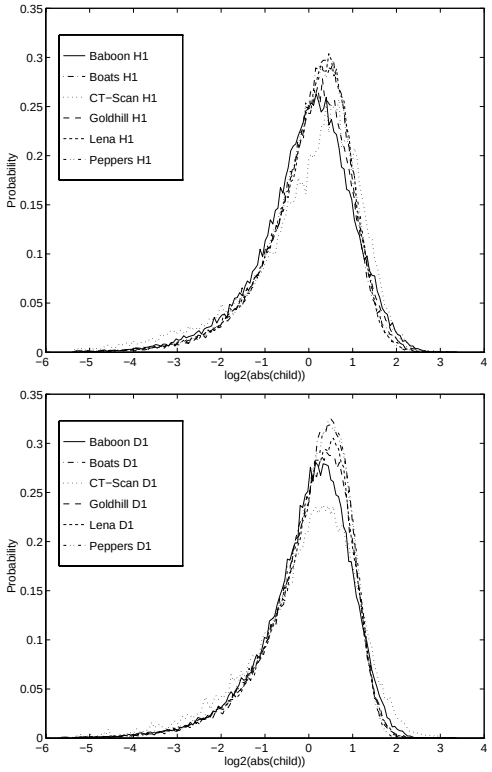

Figure 4. Normalized conditional histograms of child coefficients for a set of six natural images. Top: horizontal subband. Bottom: diagonal subband.

estimated using the relationship in equation (2), and (σ, n) are estimated by least-squares fitting to joint histograms. These model parameters are encoded as 8-bit quantities, and are sent in the encoded bit stream just before the first bitplane of their corresponding subbands.

Using the four encoded parameters, the receiver constructs a numerical table of the joint probability density for each parent-child pair. This is done via Bayes' rule, using a prior density specified by equation (1), and a conditional density specified by a lookup table formed by averaging conditional subband statistics (across all scales, orientations, and images in our database). For each child bit, we calculate the probability that it is one (conditioned on the bits that have already been received for both parent and child) by numerically integrating over the appropriate rectangular region of the joint distribution. A non-adaptive arithmetic encoder [16], which takes advantage of the computed probability values, is then used to encode the actual bit values. Finally, the arithmetic encoding of the bitplane is compared to a run-length encoding of the same bitplane, and the smaller representation is inserted into the encoded bit stream. Note that all bitplanes of the lowpass band are sent using run-length encoding.

3. RESULTS

Figure 6 shows a progressive sequence of EPWIC images. Figure 5 shows a rate-distortion comparison of our coder to the EZW coder and theoretical entropy bounds averaged across six example images. Each curve gives the SNR (in dB) relative to the first-order entropy. The second-order (i.e., conditional across scale) entropy is also shown. Both entropies are computed using the image sample statistics

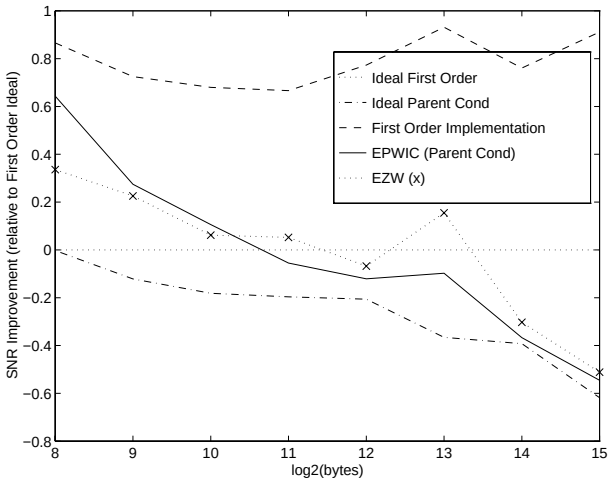

Figure 5. Comparison of rate-distortion curves for a set of actual coders and theoretical entropy bounds. Shown are SNR values (in dB), relative to the first-order entropy (dotted horizontal line), averaged over six images ("Baboon", "Boats", "Goldhill", "Lena", "Peppers", and a CT scan). See text for details.

(histograms). Results from three implemented coders are also shown: a first-order implementation of EPWIC (i.e., subbands encoded independently, using the model in equation (1), the full (conditional) implementation of EPWIC, and the EZW coder [1].[2] The first-order implementation is, on average, about 1dB worse than the first-order entropy. This is due to a combination of incorrectness of the probability model and the inefficiencies of the arithmetic encoder. Both EZW and EPWIC show significant improvement over the first-order coder, and surpass the first-order entropy at higher bitrates. EPWIC outperforms EZW at higher bitrates, but is somewhat inefficient at lower bitrates.

4. DISCUSSION

We have presented an embedded wavelet image coder based on a model of joint statistics. Despite its simplicity, the performance of the coder (in an MSE sense) is excellent. Such a coder is ideal for progressive transmission applications, such as viewing samples of an image database over the internet. The receiver could be constructed to decode the data to an arbitrary dimension and bit depth, depending on the display device and/or user preferences.

Many improvements could be made to EPWIC. The arithmetic coder could be made adaptive, and the representation of the parameters (s, p, σ, n) could be refined. More fundamentally, one would like to improve the underlying statistical characterization. Ideally, one would like to derive a statistical model from the physical constraints of light reflectance and image formation, but this is extremely difficult. Even without such a model based on first principles, alternative point probability models may give improved results (e.g., see [9]). In addition, a model for sign bit probabilities should be incorporated. We are currently developing extended probability models which incorporate subbands other than the immediate parent (e.g., other orientations). Preliminary results show marked improvements.

EPWIC could be extended to handle color images. Given the apparent role of noise in the conditional distribution

[2] We thank Sarnoff Research Center for assistance in the EZW comparisons.

Figure 6. A progressive series of the "boats" image encoded using EPWIC.

(see figure 3), it might also be extended to perform noise removal, as has been suggested in [17]. Finally, we are working to incorporate this and related probability models into other wavelet-based applications, such as image enhancement.

REFERENCES

[1] Jerome Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans Signal Processing*, 41(12):3445–3462, December 1993.

[2] M Vetterli. Multidimensional subband coding: some theory and algorithms. *Signal Processing*, 6(2):97–112, 1984.

[3] J W Woods and S D O'Neil. Subband coding of images. *IEEE Trans. Acoust. Speech Signal Proc.*, ASSP-34(5):1278–1288, October 1986.

[4] H Gharavi and A Tabatabai. Sub-band coding of digital images using two-dimensional quadrature mirror filtering. In *Proc. of SPIE*, volume 707, pages 51–61, 1986.

[5] E H Adelson, E P Simoncelli, and R Hingorani. Orthogonal pyramid transforms for image coding. In *Proceedings of SPIE*, volume 845, pages 50–58, Cambridge, MA, October 1987. Also available as MIT Media Lab Vision Science Technical Report #99.

[6] S G Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Pat. Anal. Mach. Intell.*, 11:674–693, July 1989.

[7] D L Donoho. Nonlinear wavelet methods for recovery of signals, densities, and spectra from indirect and noisy data. In I. Daubechies, editor, *Proc Symp Appl Math*, volume 47, pages 173–205, Providence, RI, 1993.

[8] D. Heeger and J. Bergen. Pyramid-based texture analysis/synthesis. In *Proc. ACM SIGGRAPH*, August 1995.

[9] S Zhu, Y Wu, and D Mumford. Filters, random fields and maximum entropy (FRAME) – towards the unified theory for texture modeling. In *IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition*, June 1996.

[10] E P Simoncelli and E H Adelson. Noise removal via bayesian wavelet coring. In *Third Int'l Conf on Image Processing*, pages 379–383, Lausanne, Switzerland, September 1996.

[11] A P Pentland, E P Simoncelli, and T Stephenson. Fractal-based image compression and interpolation, 1992. U.S. Patent Number 5,148,497, filed 14 Feb 1990, issued 15 Sep 1992.

[12] A Pentland and B Horowitz. A practical approach to fractal-based image compression. In A B Watson, editor, *Digital Images and Human Vision*. MIT Press, 1993.

[13] R Rinaldo and G Calvagno. Image coding by block prediction of multiresolution subimages. *IEEE Trans Image Processing*, July 1995.

[14] E Schwartz, A Zandi, and M Boliek. Implementation of compression with reversible embedded wavelets. In *Proc SPIE*, 1995.

[15] E P Simoncelli and E H Adelson. Subband transforms. In John W Woods, editor, *Subband Image Coding*, chapter 4, pages 143–192. Kluwer Academic Publishers, Norwell, MA, 1990.

[16] T Cover and J Thomas. *Elements of Information Theory*. Wiley, New York, 1991.

[17] B Natarajan. Filtering random noise from deterministic signals via data compression. *IEEE Trans Signal Processing*, 43(11), November 1995.