UC Davis

IDAV Publications

Title

Evaluation and Synthesis of Wavelet Image Coders

Permalink

<https://escholarship.org/uc/item/3232t7rk>

Authors

Lu, Jian
Algazi, Ralph
Estes, Robert R.

Publication Date

1995

Peer reviewed

EVALUATION AND SYNTHESIS OF WAVELET IMAGE CODERS

Jian Lu, V. Ralph Algazi, and Robert R. Estes, Jr.

Center for Image Processing and Integrated Computing (CIPIC)
University of California, Davis, CA 95616, U.S.A.
Phone: (916)752-2387 Fax: (916)752-8894
Email: jian@cipic.ucdavis.edu

ABSTRACT

The design of a wavelet image coder can be divided into three parts: wavelet representation, quantization, and error-free encoding. We evaluate each of these parts individually and synthesize them into complete coders. The evaluation is in the rate-distortion sense; two image quality metrics are used: a perception-based, quantitative picture quality scale (PQS) and the conventional distortion measure, peak signal-to-noise ratio (PSNR). Two representative wavelets, three quantizers, three encoders, and some combinations of these parts are comparatively evaluated. Our results provide an insight into the design issues of optimizing wavelet coders, as well as a good reference for application developers to choose from an increasingly large family of wavelet coders for their applications.

1. INTRODUCTION

Research in wavelet image coding since the late 1980's has resulted in various designs of coders. See, e.g., [2, 3, 10, 12]. As the family of wavelet coders becomes larger and larger, it is important to evaluate different designs carefully in a comparative way so that application developers can make the best of the existing technologies. In addition, the results from a comprehensive evaluation can guide optimum coder design. In this paper, we report our research results in evaluation and synthesis of wavelet image coders.

Our study is confined to still images and the evaluation is based on two distortion measures computed at various bitrates. A common expectation about wavelet image coders is that they produce subjectively better quality images than the standard JPEG coder. However, an objective evaluation must rely on some quantitative distortion measures. A commonly used distortion measure, peak signal-to-noise ratio (PSNR), is based on the mean square error (MSE). The MSE as an image distortion measure has long been recognized as inadequate because of its low correlation with human visual perception. It is particularly inappropriate to use the MSE for evaluating wavelet coders which are largely motivated by the properties of the human visual system (HVS). Therefore, we chose, in addition to the PSNR, another distortion measure called the Picture Quality Scale (PQS) in our study. The PQS is a perception-based, quantitative distortion measure that has been developed in the last few years for evaluating the quality of

compressed images[1, 9].

The design of a wavelet image coder can be divided into three parts: wavelet and related representations, quantization strategies, and error-free encoding techniques. In each part, one has the freedom to choose from a pool of candidates and this choice will ultimately affect the coder performance. Therefore, it is necessary to evaluate each choice independently, i.e., with the other parts of the coder fixed. Then, complete coders can be synthesized and evaluated by combining different wavelet representations, quantizers, and encoders.

The rest of paper is organized as follows: Section 2 gives a brief review of the family of wavelet image coders; Section 3 introduces PQS and PSNR as the two distortion measures used in the evaluation; Section 4 presents results of coder evaluation and synthesis with a discussion; Section 5 concludes the paper.

2. FAMILY OF WAVELET IMAGE CODERS

Generally speaking, a wavelet image coder can be made by selecting a wavelet representation, a set of quantizers, and an error-free encoder. In this section we summarize briefly the options available for each of these parts. A more thorough discussion and complete reference list are given in [8].

2.1. Wavelet Representations

Wavelet representations can be classified into a few general types by their basic "building blocks". Among them orthogonal and biorthogonal wavelets are two popular types that have been used in image coding for some time. Generalizations of wavelet bases include wavelet packets, and multiwavelets which provide greater adaptability and energy compaction at higher computational cost. Additionally, the local extrema or zero-crossings of certain wavelet transforms have been used for representing and encoding images. They belong to a class of primitive-based, non-conventional coding techniques.

2.2. Quantization Techniques

Scalar Quantization (SQ) is a well studied technique. Several types of SQ have been developed for wavelet image coding, including variance-based, HVS-adapted, and entropy-

constrained SQ. Vector Quantization (VQ) techniques have also been proposed and used, ranging from unstructured full-search VQ to lattice VQ. Most SQ and VQ techniques are band-based, which can be generally referred to as “frequency quantization” since each subband corresponds a different frequency range. The recent EZW coder[10] employs “space quantization” on a tree that organizes data across the subbands with the same orientation. More recently, there are studies on joint space-frequency quantization that attempts to fully exploit the space-frequency characteristics of wavelet representations.

2.3. Error-Free Encoding Techniques

Huffman codes are the simplest entropy coding technique. For highly skewed sources, such as quantized wavelet transformed images, Huffman codes are known to be very inefficient. Commonly, run-length encoding the abundance of zeros, when combined with Huffman encoding of the non-zero values, produces good results[3]. Adaptive arithmetic codes work well even with highly skewed sources. One can further improve the efficiency by encoding an *activity mask* (all non-zero values are set to 1) and the non-zero pixels. This is similar to a combined run-length encoding and Huffman coder.

3. DISTORTION MEASURES

3.1. Picture Quality Scale (PQS)

Research into the psychophysics of human visual perception has revealed that the HVS is not equally sensitive to various types of distortion in an image. This directly affects the perceived image quality. The PQS includes five distortion factors of which the first two are derived from random errors and the last three from structural errors. Here we give only a brief description of these distortion factors. Formulas for computing these distortion factors are detailed in [1, 9].

Distortion Factor F_1 is a weighted difference between the original and the compressed images. The weighting function adopted is the CCIR television noise weighting standard. Here the viewing distance is assumed to be four times the picture height.

Distortion Factor F_2 is also a weighted difference between the original and the compressed images. The weighting function is from a model of the HVS. In addition, an indicator function is included to account for the perceptual threshold of visibility.

Distortion Factor F_3 reflects the end-of-block disturbances. The HVS is quite sensitive to linear features in images. In block coders, the error image contains discontinuities at the end of blocks, which explains blocking artifacts in the compressed image.

Distortion Factor F_4 accounts for general correlated errors. Textures with strong correlation are more perceptible than random patterns. Strong correlation in the error image suggests distortions that are more apparent to human observers.

Distortion Factor F_5 is a measure of the large errors that occur for most coders in the vicinity of high contrast transitions (edges). Two psychophysical effects occur in the

vicinity of high contrast edges. On the one hand, the visibility of noise decreases; this is referred to as “visual masking”. On the other hand, the visibility of misalignments increases.

Because the distortion factors $\{F_i\}_{1 \leq i \leq 5}$ are correlated, a principal component analysis is performed to transform them into uncorrelated “sources of errors”, and dominant sources are identified. These errors are then mapped to a PQS value by a model which was obtained from a linear regression analysis with the Mean Opinion Score (MOS)[4]. The final form of PQS is a single value on a scale of 1 to 5.

3.2. Peak Signal-to-Noise Ratio (PSNR)

The traditional distortion measure PSNR is also used in our evaluation. It is defined by

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}} \right)$$

where MSE is the mean square error between the original and reconstructed images.

4. RESULTS AND REMARKS

4.1. About the Experiment

We evaluated a total of 760 encoded images representing a combination of two wavelets, three quantizers, and three encoders, plus the EZW coder for coding two test images (Lenna and Barbara, both 512×512) at 20 bitrates ranging uniformly from 0.1 to 2.0 bpp. The two wavelets used are the orthogonal, 8-tap wavelet of Daubechies(D8)[5] and the biorthogonal, “9-7” wavelet of Barlaud (B97)[2, 5]. All wavelet transforms are computed for 4 dyadic scales, resulting in 13 subbands. All three quantizers are scalar quantizers: the first (Q1) is the non-optimized quantizer like the one used in the EPIC[11]; the second is the HVS-adapted quantizer (Q2) of Lewis and Knowles[6]; the third is an entropy-constrained quantizer (Q3) where a bit budget is optimally allocated to each subband and used as a constraint in the quantizer design. All three encoders are band based, i.e., each band is processed separately. They are: a simple Huffman encoder (E1), run-length encoded zeros plus Huffman encoded non-zero values (E2), and the activity mask based technique, where we QM-encode the mask using a 7-pixel spatial predictive context and the non-zero values using binary tree decomposition (E3). In addition, we tested the EZW coder with the B97 wavelet, tree-structured spatial quantization, and adaptive arithmetic encoding.

The results are computed, organized and presented in several ways. In assessing the choice of wavelets and quantizers, we use the computed entropy H of a quantized wavelet representation as the bitrate, assuming we have an ideal entropy encoder. The two wavelets {B97,D8} are compared for fixed quantizers and the three quantizers {Q1,Q2,Q3} are compared for fixed wavelets. To compare the three encoders {E1,E2,E3}, we fed them with images quantized at the nominal bitrates H , compute the actual output bitrates, and then plot them against H (which is the lower bound on bitrate if the pixels are independent). Finally, we compare the overall performance of a few coders synthesized from

Figure 1: Comparison of two wavelets, {B97,D8}, under quantizer Q2.

different choices of wavelets, quantizers, and encoders. We do this by plotting PQS and PSNR vs. actual bitrates for each assembled coder. Due to the space limit, we present here only a portion of our experimental results from encoding the Lenna image. Additional results can be found in [7, 8].

4.2. Comparison of Two Wavelets

Figure 1 contains two plots comparing B97 with D8 for quantizer Q2. Similar results are found for quantizers Q1 and Q3. In all cases B97 leads D8 in both PQS and PSNR for a large portion of our test bitrate range. For a given bitrate, the lead of B97 over D8 can be as much as 0.39 PQS or 1.2 dB. From another point of view, using B97 one can save as much as approximate 0.2 bpp for a given PQS or PSNR value. Note that filters of B97 and D8 have similar lengths. The advantage of biorthogonal wavelets over orthogonal wavelets is clear in this experiment.

4.3. Comparison of Three Quantizers

Figure 2 compares our three quantizers for wavelet B97. We see little difference between the three quantizers if we look at the PSNR plot. The PQS comparison, however, tells a different story. We find that Q2 is the winner in most cases. For low bitrates, Q2 is sometimes slightly outmatched by one of the other quantizers. At higher rates, Q2's dominance increases. Recall that Q2 is a HVS-adapted quantizer. Its advantage is not obvious at all from PSNR values. The PQS confirms the value of the HVS-adapted quantization. The relationship between Q1 and Q3 in PQS seems image dependent. With its PQS values close to those of Q2, Q3 clearly outperforms Q1 for Lenna in Figure 2, but the competition appears tied for another image we have tested[8].

4.4. Comparison of Three Encoders

Figure 3 shows the output bitrates of three encoders versus computed entropies for Lenna. Similar results were observed for all wavelets and quantizers, therefore, we averaged the results across wavelets and quantizers to produce the composite results shown in Figure 3. We also draw a line of unit slope where the output bitrate equals the entropy. As expected, the simple Huffman encoder (E1) al-

Figure 2: Comparison of three quantizers, {Q1,Q2,Q3}, under wavelet B97.

ways gives a bitrate higher than the entropy, especially at low bitrates where there are a large number of zeros, i.e., when the source is highly skewed. When combined with run-length encoding of the zeros (E2), the results are much better, and only slightly worse than our best, the activity mask based technique (E3). We observe that E3's bitrates are consistently lower than the (independent pixel) entropy, which may appear counterintuitive to some, but is correct since we are exploiting spatial dependencies in the source which are not reflected in the entropy computation. We declare E3 the winner.

4.5. Comparison of Wavelet Coders

We now compare a few complete wavelet image coders synthesized from different wavelets, quantizers, and encoders. A combination of “the best” gives B97-Q2-E3. We also present two other combinations, D8-Q1-E1 and D8-Q2-E3. These coders along with the EZW coder are compared in Figure 4. We see that B97-Q2-E3 is the winner by PQS for most bitrates, with EZW winning at high bitrates. By PSNR, B97-Q2-E3 loses to EZW by a small margin. Also observed from Figure 4, the simple Huffman encoder yields, clearly, the poorest coder by both PQS and PSNR. The performance difference between the best (B97-Q2-E3) and the worst (D8-Q1-E1) coders can be over 3.5 in PQS or 10 dB in PSNR. Of course, an intelligent designer would not choose such a code. Our results only indicate how bad such a brute force design can be. The coder D8-Q2-E3 differs from B97-Q2-E3 only in the choice of wavelets, and its performance is slightly worse.

4.6. Remarks

The purpose of our comparative study is not to simply rank a number of coders. We hope to find out why a coder is good or bad and how to make a good coder. The EZW coder is, in our mind, the state-of-the-art technique in wavelet image coding. The fact that we can make a coder that comes close to or even better than the EZW coder just by assembling available techniques testifies to the value of good synthesis in wavelet coder design. Our results clearly show that all parts (representation, quantization, and error-free encoding) are important in designing wavelet coders. Since wavelets were introduced to image coding, there has been considerable research looking for better wavelets. The

Figure 3: Comparison of three encoders, {E1,E2,E3}. Bitrates are averaged over {B97,D8} and {Q1,Q2,Q3}.

close and good performance of D8-Q2-E3 and B97-Q2-E3 in our study suggests that the effect of different wavelets (of similar filter lengths) may be less significant than that of quantizers and encoders.

The efficiency of error-free encoders is an important issue that has not been much addressed in the context of wavelet image coding. We have shown that a good encoder (e.g., E3) can achieve bitrate lower than the (independent pixel) entropy. On the other hand, an inefficient encoder (E1) produces bitrate much higher than the entropy. The key to a good encoder is to exploit dependency between pixels. Comparing the EZW with B97-Q2-E3, we found that both exploit dependency between quantized coefficients for encoding, which provides the possibility to achieve bitrates below the entropy. The difference is that the EZW exploits both intra- and interband dependencies by encoding the zerotrees while B97-Q2-E3 exploits more intraband dependency by encoding the activity masks. Additionally, the EZW coder is a good example of intelligent organization of data for quantization and encoding.

The PQS quantifies some perceptual characteristics of a coder that can not be revealed by the PSNR; see, e.g., quantizer comparisons in Section 4.3. This testifies to the necessity of perception-based quality metrics such as the PQS for coder evaluation.

5. CONCLUSION

We have evaluated several wavelet image coders comparatively using a perception-based picture quality scale as well as the traditional PSNR. While these results provide a reference for application developers to choose a good wavelet coder for their applications, they also shed some light on issues of optimum design of wavelet coders. Our work shows that an excellent wavelet coder can result from a careful synthesis of existing techniques of wavelet representation, quantization, and error-free encoding. All these parts play a role in making a good coder. Exploiting the dependency of quantized coefficients, including zeros, is an effective way to improve the overall performance of a wavelet coder. Quantizers designed with considerations of the characteristics of HVS are very attractive; their advantages can be quantified when an appropriate distortion measure is used. The effect of variations between asymmetric orthogonal and symmetric biorthogonal wavelets is also noticeable, but seems less significant when compared with the other two factors.

Figure 4: Comparison of four wavelet image coders.

6. REFERENCES

- [1] V.R. Algazi, Y. Kato, M. Miyahara, and K. Kotani, "Comparison of image coding techniques with a Picture Quality Scale," *Proc. of SPIE, Appl. of Digi. Image Proc. XV*, v. 1771, 1992.
- [2] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Proc.*, v. 1, pp. 205-220, Apr. 1992.
- [3] J.N. Bradley, C.M. Brislawn, and T. Hopper, "The FBI wavelet/scalar quantization standard for grayscale fingerprint image compression," *Proc. of SPIE, Visual Info. Proc. II*, Apr. 1992, pp. 293-304.
- [4] CCIR, "Rec. 500-2, Method for the subjective assessment of the quality of the television pictures," v. 11, pp. 165-168, 1982.
- [5] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Phila., PA, 1992.
- [6] A.S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," *IEEE Trans. on Image Proc.*, v. 1, pp. 244-250, Apr. 1992.
- [7] J. Lu, V.R. Algazi and R.R. Estes, "Comparison of wavelet image coders using the Picture Quality Scale (PQS)," *Proc. of SPIE, Wavelet Appl. II*, v. 2491, pp. 1119-1130, Apr. 1995.
- [8] J. Lu, V.R. Algazi and R.R. Estes, "A comparative study of wavelet image coders," submitted to *Optical Engineering* Special Issue on VCIP, 1995.
- [9] M. Miyahara, K. Kotani, and V.R. Algazi, "Objective Picture Quality Scale (PQS) for image coding," *Proc. SID Symp. for Image Display*, 44.3: pp. 859-862, May 1992.
- [10] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Proc.*, Vol. 41, pp. 3445-3462, Dec. 1993.
- [11] E.P. Simoncelli and E.H. Adelson, "Efficient Pyramid Image Coder (EPIC)," Software available from URL: <ftp://whitechapel.media.mit.edu/pub/epic/epic.tar.Z>
- [12] W.R. Zettler, J. Huffman, and D.C.P. Linden, "Application of compactly supported wavelets to image compression," *Proc. of SPIE, Image Proc. Algori. and Appl.*, Vol. 1244, pp. 150-160, 1990.