Title
Comparative Performance of Wavelets and JPEG Coders at High Quality

Authors
Algazi, Ralph
Estes, Robert R.

Publication Date
1997

Peer reviewed

Comparative Performance of Wavelet and JPEG Coders at High Quality

V. R. Algazi R. R. Estes Jr.

Center for Image Processing and Integrated Computing (CIPIC)
University of California, Davis

Qualimage, Inc.
Davis, CA

ABSTRACT

In recent work, we have examined the performance of wavelet coders using a perceptually relevant image quality metric, the Picture Quality Scale (PQS). In that study, we considered some of the design options available with respect to choice of wavelet basis, quantizer, and method for error-free encoding of the quantized coefficients, including the EZW methodology. A specific combination of these design options provides the best trade off between performance and PQS quality. Here, we extend this comparison by evaluating the performance of JPEG and the previously chosen optimal wavelet scheme, focusing principally on the high quality range.

Keywords: Coder performance, wavelets, wavelet transform, JPEG, perceptual distortion measure, image coding.

1 INTRODUCTION

The current situation in image coding is that the existence of standards provides strong disincentives to the introduction of new encoding methods for widespread use. Unless a compelling case can be made that a new algorithm will provide vastly improved performance, the use of a standard has significant commercial advantages. However, when it comes to evaluating the performance of a new image coder, the issue is both complex and difficult to resolve. This is because in the trade off of bit rate for image quality, the evaluation of quality using the mean squared error (MSE) or peak signal-to-noise ratio (PSNR) is inadequate and does not allow for meaningful comparisons. In fact, the development of both wavelet and subband coders has been predicated, in part, on an expected improvement in image quality that has been difficult to substantiate. Of course, within its class of coders, there is a number of options and parameters that will affect the performance. The choice of such intraclass parameters is a substantial problem of its own. This is the problem that we have addressed in previous publications for the case of wavelet encoding of monochrome images.[14–16] In this paper, we expand our study by comparing the best of the wavelet encoders that we have previously considered to a standard JPEG encoder. For that comparison, we will make use of two image quality measures as a function of bit rate. We use the PSNR for reference, because it is widely used, and the Picture Quality Scale (PQS), a perceptually relevant distortion metric. The use of PQS has already been shown to be beneficial in the study of wavelet coders. Here, we focus on high quality image coding.

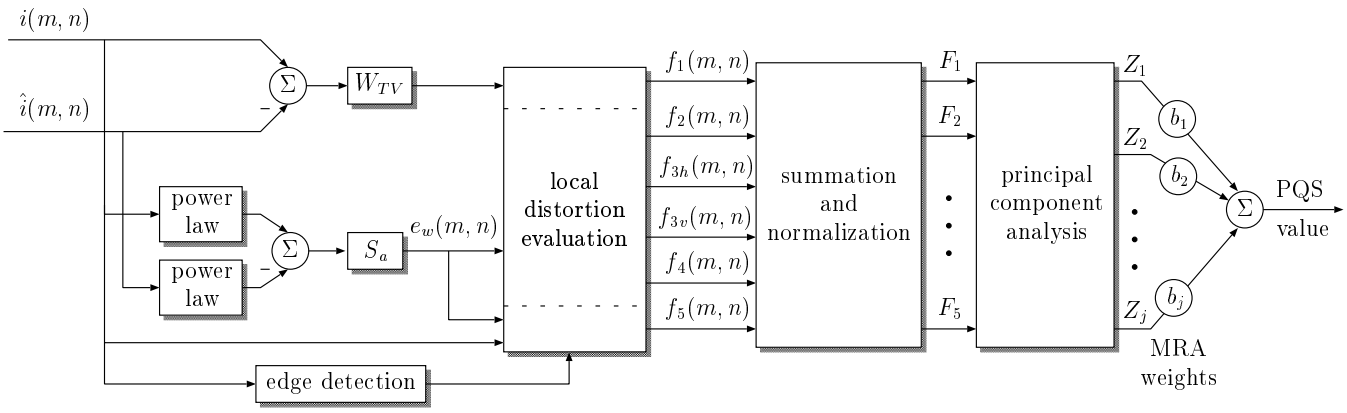

Figure 1: The construction of PQS.

2 TEST IMAGES AND QUALITY MEASURES

The development of a perceptually relevant image quality measure is a complex process that requires specifying carefully the image viewing conditions, among other requirements. A simple distortion metric, such as MSE or PSNR,* provides an indication of relative quality changes for adjustments in the coding parameters. This is useful when encoding a single image using a specific coder, with small changes in coding parameters. The results, however, are coder and image dependent. Further, the correspondence of PSNR to the evaluation of image quality by an observer is weak. Therefore, the use of the PSNR as a quality measure is principally useful when comparing techniques with respect to a specific image. Extending the results obtained for one image to other images is problematic. As for the comparison of coding algorithms, since the PSNR is dependent on image complexity, local behavior of coders in complex portions of the image may determine the PSNR, while visual masking may make such large local errors perceptually unimportant.

A more satisfactory approach is to use a perceptually relevant distortion measure, i.e., one which agrees with evaluations by human observers. One such measure, which we have developed, is the Picture Quality Scale (PQS),[16,17] mentioned earlier and currently illustrated in Figure 1. It is computed from five distortion factors commonly introduced by coders. Local distortion factor images are computed for each factor, then the images are combined using multiple regression and principal component analysis to obtain a single number, the PQS value, representing the quality of a given image. For a set of 75 images, the correlation between PQS and the mean opinion scores (MOS) obtained from subjective testing is 0.93,[17] indicating the potential of PQS as a image quality metric. Its properties and development are discussed next.

2.1 Picture Quality Scale

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, thereby, directly affecting the perceived image quality. The PQS is based on quantitative measures of several distortion factors. Because these distortion factors are correlated, a principal component analysis is done to transform them into uncorrelated "sources of errors." These errors are then mapped to a PQS value using a model obtained from linear regression analysis with the mean opinion score (MOS), a five scale subjective ranking of image quality in terms of perceived distortions that are described in Table 1.[5]

*$PSNR = f(MSE) = 10\log_{10}(255^2/MSE)$, i.e., they are just different representations of the same quantity and, therefore, can be used interchangeably.

Grading Scales	Impairment
5	Imperceptible
4	Perceptible, but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

Table 1: The mean opinion score, a picture quality scale.

2.1.1 Distortion factors

The current version of the PQS includes five distortion factors of which the first two are derived from random errors and the last three from structured errors. Here we give only a description of these distortion factors. Formulas for computing the actual numerical measures are detailed elsewhere.[17] Note that effective perceptual distortion measures are a function of display resolution and viewing distance.

Distortion factor F_1 is a weighted difference between the original and the compressed images. The weighting function adopted is the International Radio Consultative Committee (CCIR) television noise weighting standard.

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.[11] 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 and structured error features in images. In block coders, the error image contains discontinuities at the end of blocks, which explains perceived blocking artifacts in the compressed image.

Distortion factor F_4 accounts for general correlated errors. Errors with strong correlation are more perceptible than random patterns. Strong correlation in the error image suggests more apparent distortion than accounted for by the magnitude of the errors.

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 and errors decreases; this is referred to as *visual masking*. On the other hand, misalignments and blurring of edges is quite objectionable.

2.1.2 Principal component representation of distortion measures

Because the distortion factors $\{F_i\}$ are correlated, a principal component analysis is performed to decorrelate the distortion measures and identify the dominant sources. The MOS data consists of 675 subjective evaluations: 9 viewers, 5 images, 3 coders, and 5 quality levels for each coder. For each of the 75 encoded images, the MOS scores are averaged resulting in 75 MOS evaluations. An eigen analysis of the distortion factors versus the MOS evaluations indicates that the three largest eigenvalues account for 98% of the total error energy. Therefore, the three corresponding eigenvectors transform $\{F_i\}$ into a principal component representation, $\{Z_i\}_{1\le i\le 3}$, and

$$PQS = b_0 + \sum_{i=1}^{3} b_i Z_i$$

where $\{b_i\}_{0\le i\le 3}$ are the partial regression coefficients obtained by multiple linear regression of $\{Z_i\}$ against the MOS.[17] Transforming the results obtained back into the original factor space, we have

$$PQS = 5.797 + 0.035F_1 + 0.044F_2 + 0.01F_3 - 0.132F_4 - 0.135F_5. \quad (1)$$

As defined above, PQS fits the original MOS data with a correlation coefficient of 0.93.

Our extensive experiments indicate that PQS differentiates images encoded at the same PSNR, in accordance with the assessment of image quality by human observers. Note, however, that the MOS scale, and the PQS scale which matches it, is very broad and may not meet the needs of critical applications.

In this paper, we will use several test images and both PSNR and PQS for the comparison of coding algorithms. The use of both measures will allow some statements or conclusions on the relative sensitivity and use of such measures in comparing coders.

Because of our interest in high quality, for which image quality is difficult to evaluate, we will also use a hybrid method that attempts to sharpen the application of PQS across coders. We will compare and match subjectively coders for a single image at the bottom of the quality range of interest, and use the quality measure for tracking the image quality for increasing bit rate. Thus, in such an hybrid scheme, the objective PQS measure is anchored subjectively across coders. This experiment will be described more fully in Section 5.

3 THE WAVELET CODERS

A very common image encoding paradigm consists of three stages: transformation, quantization and encoding. Many wavelet transform encoding techniques, as well as the DCT based JPEG, fit into this framework. More sophisticated techniques have been developed, but we have shown that a proper choice of each of these components results in a code which performs as well as, or better than, some of these newer techniques.[14–16] In particular, a simple such strategy using a biorthogonal wavelet transform, a quantizer designed for the human visual system and a simple "color shrinking" based coder performs better than the widely used EZW code.[23] In the remainder of this section, we summarize and extend our previous results and present the best performing wavelet coder. Later, in Section 5, this coder will be compared with the JPEG results of Section 4.

3.1 Wavelet Representations

For our purposes, a wavelet transform is an octave band, subband decomposition. The original image is split into four subbands using a critically sampled filter bank, and this process is then iterated on the lowest frequency subband to further decorrelate the transform coefficients. Each iteration corresponds to a coarser spatial *scale* in the original image. Here, we consider separable transforms which are computed by independently processing the rows, then the columns, using half band 1-D filters. We also further limit our discussion to 2-D wavelet transforms, i.e., wavelet transforms of images. Finally, we only consider simple, widely used, orthogonal and biorthogonal wavelet transforms, i.e., we leave the study of more sophisticated wavelet image representations, including wavelet packets, multiwavelets, and multiscale edges for future work.

Orthogonal expansions have many advantages for image coding and compactly supported orthogonal wavelets, which correspond to finite impulse response (FIR) filters and can be implemented efficiently, are typically used. The length of the filter used is related to the degree of smoothness and regularity of the wavelet which, in turn, can affect coding performance.[22] With the exception of the 2 tap Haar wavelets, however, linear phase, compact orthogonal wavelet transforms cannot be designed. This leads us to consider biorthogonal wavelet transforms.

In orthogonal wavelet transforms, the QMF filter pairs are derived from a single prototype filter resulting in a set of quadrature filters (QF) which are orthogonal to each other. The biorthogonal case, derived from two prototype filters, results in a set of quadrature filters which are no longer orthogonal to each other, but which are orthogonal to another QF pair used to compute the inverse transform. This generalization allows symmetric, compact wavelet transforms to be designed. Perfect reconstruction is preserved and Mallat's fast algorithm can still be used.

Studies[22,28] have found the performance gain from using filters with more than 8 or 10 taps is not justified, thus, we chose the popular 8 tap orthogonal wavelet of Daubechies (D8)[8] and the "9-7" wavelet of Barlaud (B97),[3,9] which was rated highly in a recent study by Villasenor, Bellzer, and Liao,[24] for our comparative study.

3.2 Quantization Techniques

Quantization techniques generally fall into one of two categories: scalar and vector quantization, respectively, depending on whether the coefficients are quantized individually, or in groups. The design of such quantizers is driven by rate-distortion theory that is somewhat limited. Specifically, the theory assumes an ensemble of images with a known statistical characterization. For such an ensemble, the quantization strategy can be determined and the set of coefficients which need to be transmitted defined. This quantization information, then, can be stored at both the encoder and decoder and does not need to be transmitted. This methodology does not work well in practice, because adequate statistical characterizations are not available. Threshold coders solve this problem in an adaptive fashion, where the set of significant coefficients is allowed to vary from image to image. As such, the location of significant coefficients must also be encoded. Newer, space-frequency quantization techniques result when this location overhead is explicitly considered.

Scalar quantizers are defined by partitioning the real line into a set of intervals and choosing a single value in each interval to represent all values which lie in that interval. Optimally, each coefficient must be quantized based on its probability distribution. If a variable length encoder is used to encode the quantized coefficients, then we can safely restrict our attention to uniform quantizers,[12] in which case only a single parameter, the step size, must be chosen to specify each quantizer.

In the wavelet case, an N-scale transform results in $3N + 1$ subbands, each of which is assumed to contain coefficients with similar distributions, so that quantizer design consists of choosing $3N + 1$ quantizer step sizes. If the wavelet transform is orthonormal, then to minimize the MSE, a single uniform quantizer for all subbands is used. However, our goal is to minimize perceptual error, so that better quantizer designs should be used.

Vector quantization is a generalization of scalar quantization in which vectors, or blocks, of pixels are quantized instead of the pixels themselves. The optimality of VQ over SQ is discussed by Gersho and Gray.[10] We refer to both scalar and vector quantization, as discussed above, as *frequency quantization* methods, since each subband corresponds to a different frequency range. Wavelet representations, however, have both scale (frequency) and space contexts, so that spatial grouping and quantization is possible and desirable.

For reasons mentioned above, the cost of encoding a coefficient is more accurately represented by the cost of encoding its location plus the cost of encoding its magnitude. Techniques which consider both of these costs will outperform those that do not. A relatively simple such approach, the embedded zerotree coder (EZW),[23] was one of the first papers that demonstrated the advantages of wavelet representations over other commonly used representations, such as the DCT. It encodes wavelet transform coefficients in an embedded fashion, i.e., the data stream can be truncated at any point, trading rate for quality, and the image reconstructed. It encodes the quantized coefficients from most significant to least significant bitplane and exploits intra and inter-subband spatial dependencies using a quadtree-like data structure. Newer, space frequency quantizers[27,26] *optimize* coder performance by discarding coefficients which do not contribute proportionately to a reduction in MSE.

In our previous study, in addition to using a single uniform scalar quantizer (Q1), we also considered the HVS frequency weighted quantizer of Lewis and Knowles (Q2)[13] and an entropy-constrained quantizer in which a bit budget is optimally allocated to each subband and used as a constraint for quantizer design (Q3). For the latter technique we used the optimum bit allocation scheme of Chen, Itoh, and Hashimoto[6] with a uniform Laplacian rate-distortion model.[12]

3.3 Error-Free Encoding Techniques

Although not an actual encoding technique, band based Shannon entropy is commonly used in the evaluation of coding performance. A simple encoding technique results if Huffman codes are designed for each band. Care must be exercised, however, to ensure that accurate statistics are used to design these codes. One can design a universal code based on an ensemble of typical images or explicitly transmit the Huffman codes, along with the compressed image data, to the decoder. For highly skewed sources, such as quantized wavelet transformed images, Huffman codes are known to be very inefficient. However, if the most probable symbols (zeros) are removed from

the source and encoded separately, little spatial correlation remains among the nonzero values, which can then be encoded efficiently. Commonly, run-length encoding the abundance of zeros, when combined with Huffman encoding of the nonzero values, produces good results.[4,7]

Adaptive arithmetic codes start with no information about the image and implicitly transmit the model to the decoder in the compressed data stream. Therefore, they are free from the statistical ensemble issues associated with the design of Huffman codes. Binary arithmetic codes, such as the Q-code and QM-code,[18] are more computationally efficient than their multi-alphabet counterparts, but require a mapping from the quantized coefficients to a sequence of binary decisions. A simple technique, which is similar to the run-length encoding discussed above, proves to be very beneficial. The locations of the nonzero pixels are specified by encoding a binary *activity mask* (all nonzero values are set to 1) with standard binary image compression techniques, such as the Joint Bi-level Image Experts Group (JBIG) coder, after which the nonzero pixels are mapped through a balanced binary tree and encoded. Using this *color shrinking*[2] based technique, we often obtain bit rates less than the Shannon entropy (based on independent pixels) due to the significant spatial correlation between the zeros in a wavelet-transformed image.

Three encoding strategies were considered in our previous study: Huffman coding (E1), Huffman coding plus runlength coding of the activity map (E2), and color shrinking technique where we QM-encode the activity mask with a 7-pixel spatial predictive context and the nonzero values using binary tree decomposition (E3).

3.4 Extensions

After reviewing our previous work,[14–16] we have modified the encoding and quantization strategies used there. The conclusions are consistent with earlier results, but we present our modifications here for completeness. In our previous work, a separate, independent bit rate was specified for the lowest frequency LL band and the minimum allowed step size for any band was 1.0. Such quantization, for some images at higher bit rates, leads to cases where the lowest frequency coefficients are quantized as coarsely as the highest frequency coefficients, which is clearly undesirable. Here, we modify our previous quantizers Q1 and Q2 (to obtain Q1′ and Q2′) so that the lowest frequency LL subband is quantized with the same step size as the lowest LH, HL, and HH subbands. When orthonormal transforms are used, Q1′ is now just a uniform quantizer and the step sizes for Q2′ are just scaled versions of those given in Figure 2. This simple quantization matrix scaling mirrors the technique used in JPEG.

<table>
<tr><td>0.80 0.80
0.80 $0.80\sqrt{2}$</td><td>0.64</td><td rowspan="2">0.64</td><td rowspan="3">1.0</td></tr>
<tr><td colspan="2">0.64</td><td>$0.64\sqrt{2}$</td></tr>
<tr><td colspan="2">0.64</td><td>$0.64\sqrt{2}$</td></tr>
<tr><td colspan="3">1.0</td><td>$\sqrt{2}$</td></tr>
</table>

Figure 2: Perceptual frequency weighted, 4 scale, wavelet quantization matrix.

In addition, we modified the encoder. It is still based on the binary QM-code,[18] but uses a different mapping from gray scale to binary and is a complete code in that a separate encoder and decoder are used and the only input to the decoder is the encoded data stream. The entire image is also encoded in one pass. First, we encode whether the pixel is 0 or not using a 7-pixel spatial context. Next, if the value is non-0 then we encode its sign. Finally, we encode the positive integer that remains using a binary tree decomposition or by using a magnitude category based technique similar to that used in JPEG's arithmetic coder.[19] The latter method is simpler and leads to results which are slightly better than the binary tree, so that it is used exclusively and briefly described next. Consider a strictly positive integer i to be encoded. We encode it by encoding $\lfloor \log_2 i \rfloor$ 0s followed by the binary representation of i from most significant bit (MSB) to least significant bit (note that the MSB must be 1).

The magnitude category bits and the most significant 1 bit are encoded in their own states, but a single state is used to encode the remaining less significant bits, since there isn't much correlation between them.[23] We denote this coder E3′.

3.5 Results

In our previous work,[14–16] we evaluated 19 wavelet coding techniques, 18 of which were obtained by considering all combinations of the 2 wavelets (D8 and B97), 3 quantizers (Q1, Q2, and Q3), and 3 encoders (E1, E2, and E3) mentioned in Sections 3.1–3.3. The final technique considered was EZW.[23] In this work we have added quantizers $Q1'$ and $Q2'$ and encoder $E3'$. Four scale, 13 band, wavelet transformations were used. Next, we summarize the previous results and then briefly discuss the modifications.

In our previous work, E3, the "color shrinking" based encoder, was obviously the best encoder for this application. By exploiting the spatial dependencies in an *activity mask*, the transform coefficients can be encoded at a rate that is as much as 0.25 bits/pixel (bpp) below the independent pixel entropy.

When evaluated with respect to PSNR, all quantizers perform similarly, but when the perceptually relevant PQS is used, the advantages of exploiting the HVS become evident, and the quantizer of Lewis and Knowles[13] performed significantly better than the other two considered. At higher rates, the dominance of Q2 was as much as 0.36 PQS, confirming the value of the HVS-adapted quantization.

In all cases, B97 outperformed D8 in both PQS and PSNR for a large portion of our test bit rate range. For a given bit rate, the lead of B97 over D8 was as much as 0.43 PQS or 1.2 dB PSNR. From another point of view, using B97, one can save as much as approximately 0.2 bits/pixel for a given PQS or PSNR value. Note that the filters of B97 and D8 have similar lengths; the advantage of the former over the latter is clear.

Finally, we compared the 18 product coders designed above with the popular EZW code. Our best coder, B97-Q2-E3, out performs the EZW code in most cases.

To test our modifications, we encoded 9 256 images (of varying type and complexity) using the set of codes $\{D8, B97\} \times \{Q1', Q2'\} \times \{E3'\}$ and our implementation of EZW.[1] As before, each code was evaluated using both PSNR and PQS. The results obtained for the "hotel" image are given in Figure 3.

As expected, since uniform quantization (Q1′) minimizes the MSE, the results for Q1′ are better than the corresponding results for quantizer Q2′ in all cases when PSNR is used as a quality measure. Yet, the Q2′ codes perform better than Q1′ codes when evaluated using a perceptually relevant quantizer, emphasizing the shortfalls of PSNR. As before, the B97 results are better than the D8 results, and our codes perform better than EZW. Although not shown, our new coder E3′ performs slightly better than E3.

3.6 Discussion

We have presented some results from a comparative study of different wavelet image coders using a perception-based picture quality scale as well as the traditional PSNR. While our study cannot cover all the aspects of wavelet coder design, we believe that the comparisons are highly representative. 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 three parts play a role in making a good coder: exploiting the spatial dependency between quantized coefficients is an effective way to boost the overall performance of a wavelet coder; quantizers designed with considerations of the characteristics of HVS are very attractive when an appropriate distortion measure is used; and the effect symmetric biorthogonal wavelet perform better than the asymmetric, orthogonal counterparts. Finally, our study testifies to the necessity of perception-based quality metrics such as the PQS for coder evaluation. The approach we take here is certainly not limited to evaluation of wavelet coders. Next, we look at similar comparisons performed for the JPEG coder.

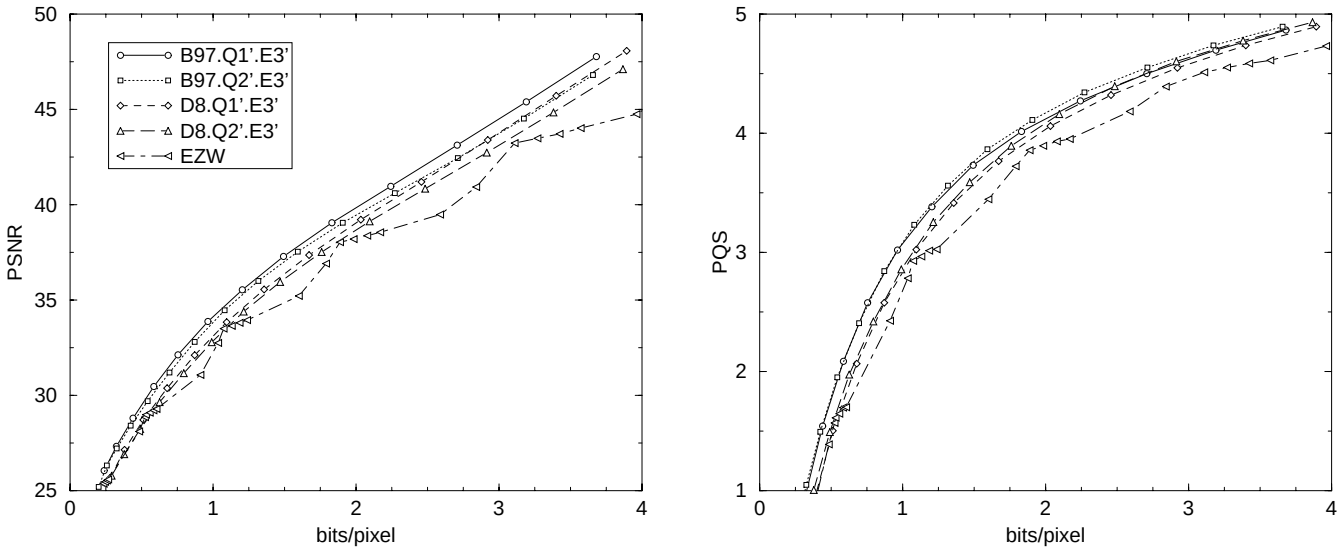

Figure 3: Wavelet encoding results for image "hotel".

4 THE JPEG CODER

The JPEG image coding technique[19] was recently standardized, and since then has become a very popular image encoding technique. In JPEG, images are encoded using a 8×8 discrete cosine transform (DCT). Each image is decomposed into non-overlapping 8×8 blocks which are then transformed with the DCT. The transform coefficients are then quantized independently to exploit the frequency response of the human visual system (HVS). Finally, the quantized coefficients are encoded in a zig zag order, from low to higher frequencies, using a runlength–Huffman technique. In this technique, sequences of zeros are runlength encoded to overcome the inefficiencies of Huffman codes for highly skewed sources.

Of the many extensions proposed in the JPEG standard,[19] a few are important here, since we are comparing the performance of JPEG and wavelet coders using a perceptually relevant quality metric. Of primary importance, JPEG allows for custom quantization matrices to be specified. In this work, we will restrict our attention to the baseline JPEG coder, as implemented by the PVRG-JPEG,[20] and one image dependent rate-distortion quantizer optimization technique (RDOPT[21]). The baseline quantization matrix (shown in Table 2) quantizes high frequencies more coarsely than lower frequencies, exploiting the HVS, but further optimization would to be beneficial. Other techniques exist for designing these quantizers, such as the perceptually based DCTune algorithm,[25] but was not considered due to a lack of its availability.

Any extension that improves the coding efficiency of JPEG would also be useful. Using arithmetic encoding instead of Huffman codes would improve the performance of JPEG considerably. The PVRG-JPEG we use is a two pass algorithm, designing Huffman codes in the first pass that are used to encode the data in the second pass instead of using the default codes suggested in the JPEG standard. As such, its performance is better than baseline JPEG coders which use the built in codes.

Typical coding results obtained with JPEG and the optimized RDOPT are shown in Figure 4, where we see, at high quality, that the large differences indicated by PSNR are much smaller when measured with the perceptually relevant PQS scale. We also note that differences between the two techniques are fairly small. Using PSNR, we find that RDOPT can reduce the bit rate by as much as 1 bpp or improve the PSNR by as much as 5 dB, with

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Table 2: Default luminance quantization matrix for JPEG.

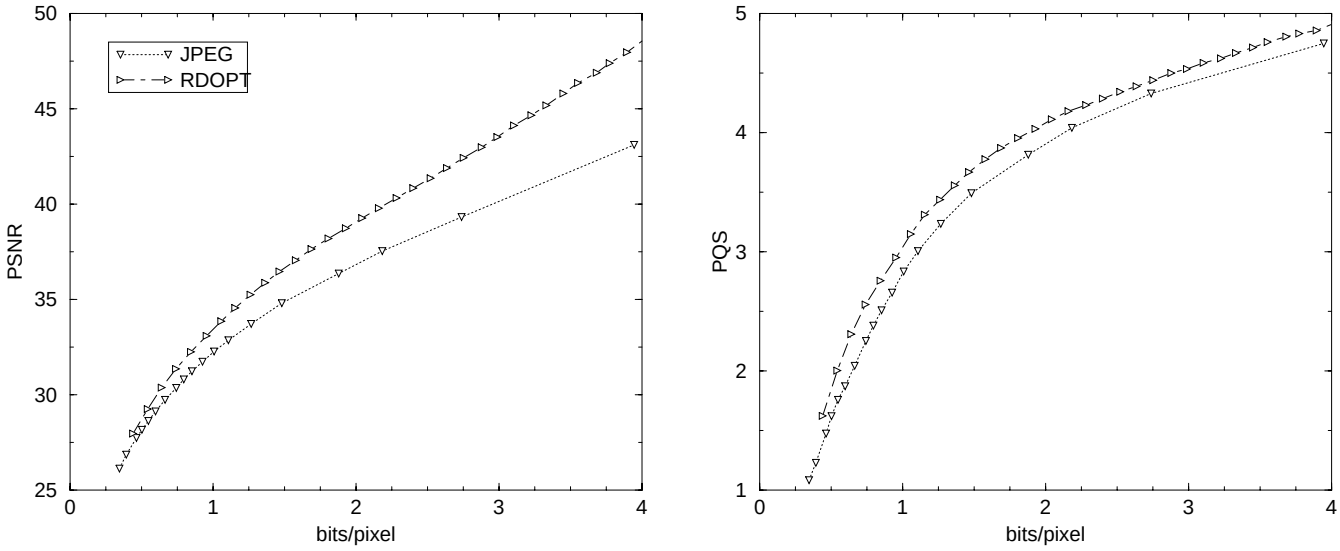

Figure 4: JPEG encoding results for image "hotel".

typical values (for high quality images) of about 0.5 bpp and 3 dB. Using PQS, the bit rate reduction is more modest ($<$ 0.5 bpp) and the PQS improvement is less than 0.4 PQS, with typical values of 0.25 bpp and 0.25 PQS, respectively.

5 COMPARATIVE RESULTS OF WAVELET AND JPEG CODERS

In the comparison of JPEG and wavelet coders, we make use of both PSNR and PQS by similarity to our work on wavelet coders. The use of PQS will also allow us to evaluate the comparative effectiveness of improvements or extensions to standard coders on perceptual grounds. Although we have studied several images, we show graphs and discuss results for a single representative image, since the results are consistent across images and graphs for a representative image are more informative that averages over a set. We use the best wavelet coder (B97-Q2′-E3′) from Section 3 in these comparisons well as the baseline PVRG-JPEG and RDOPT results of the previous section.

In Figure 5 we compare of the performance of the B97-Q2′-E3′ with baseline JPEG and RDOPT, for image

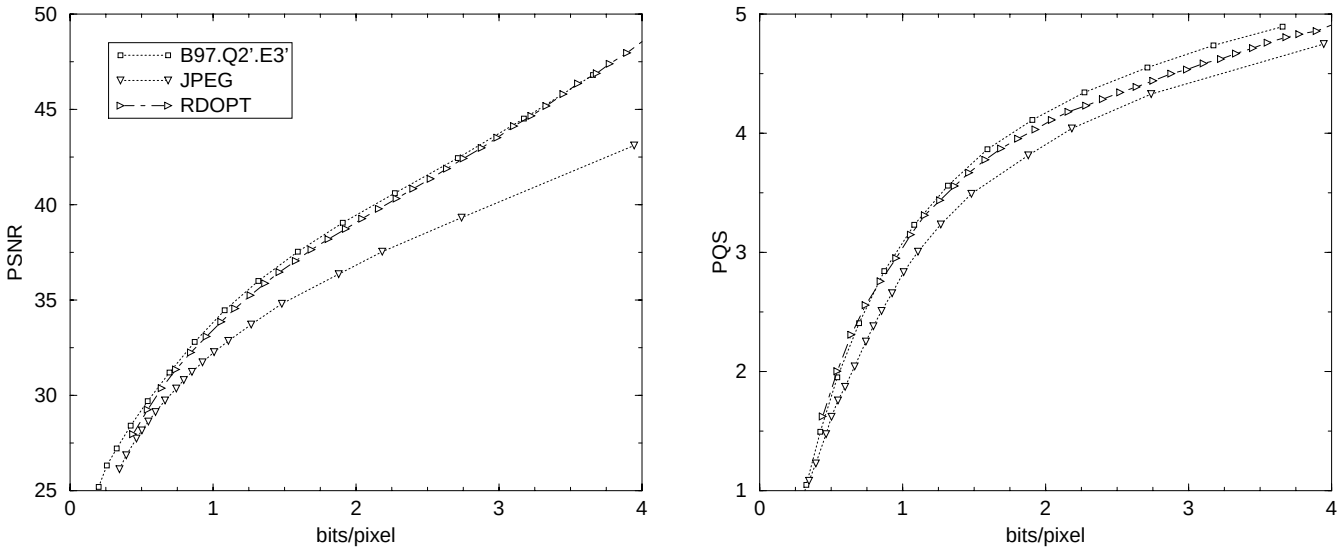

Figure 5: Comparison of JPEG and wavelet codes for image "hotel".

"hotel". We focus principally on the behavior for high quality images (> 3 PQS) and observe that the wavelet coder is consistently better than either JPEG coder. The use of PSNR shows an increasingly large superiority for the wavelet coder as the rate increases, but such a change at high PSNR is not meaningful. The PQS score gives a much more realistic comparison. If we compare the gain in PQS value versus bit rate, we find that the a typical gain of about 0.25 PQS over JPEG and 0.1 PQS over RDOPT. This corresponds to a typical reduction of 0.2 bpp and 0.4 bpp, respectively, but can be as high as 0.3 PQS and 0.7 bpp. Comparing with the PSNR results, we see that the PQS results are moderated at higher bit rates and, furthermore, that the advantages of the perceptually designed wavelet quantizer are not apparent from the PSNR results.

Similar results are found for all test images, with B97-Q2′-E3′ 0.2-0.8 PQS better, or requiring 0.15-0.8 bpp less, than JPEG, with typical values around 0.4 PQS and 0.5 bpp. RDOPT achieves about half of these gains. Using PSNR as a performance measure, the differences are as high as 1.1 bpp and 5.5 dB for JPEG and are misleading.

5.1 Refining the PQS Evaluation

Because there is scattering of values and uncertainty in the PQS scores obtained for a set of images and a set of coders, we considered refinement of the PQS scale to account for any coder dependent bias that might exist. Since we are interested in high quality images, we visually compared images encoded at a quality of about 3 PQS, for a wavelet coder and JPEG. The "hotel" image was encoded at a range of qualities with both coding techniques. We printed these images on a very high quality Fuji Pictography printer in strips of four $2'' \times 2''$ prints. The quality varied gradually from one image to the next across the strip. By comparing the strips of images, we performed a comparison of images across coders. Using a wavelet encoded image at 3.1 PQS, we determined that the JPEG encoded images that bracketed its quality were at 3.06 and 3.23 PQS. From this simple subjective evaluation, we concluded that the PQS scale allows us to reach reasonable conclusion about the relative merit of the coding techniques with respect to their subjective image quality. Similar experiments would also be useful for comparison of PQS across images, but have not been performed at this time.

6 DISCUSSION AND CONCLUSIONS

The very large and growing number of coding schemes, extensions and improvements, makes the comprehensive comparison of their performance a massive effort beyond the scope of this paper. This paper makes two contributions. The first one is with respect to methodology. The comparative use of PQS and PSNR as quality measures points out the value of a perceptual measure, principally at high quality, as the results for PSNR become meaningless. Second, by restricting ourselves two popular coding schemes, and to high quality, we were able to reach some conclusions about their relative performance. The best wavelet coder is consistently superior to the baseline JPEG coder and slightly better than the optimized RDOPT JPEG coder. RDOPT's image dependent optimization bridges the gap between the two classes of coders, but we note that the similar optimization has not been done as for our wavelet coding. In general, JPEG and wavelet codes can be designed with similar quantization and encoding techniques and current research is bridging this gap, but, at the present time, results are sparse, hindering a more systematic comparison. We are currently working on such a comparison.

A similar study also needs to be carried out at lower quality, but the results presented here make it clear that a perceptual quality measure must be used. The PQS scale is applicable to lower quality after a tailoring of the technique to changes in the relative importance of visual artifacts as the image quality decreases. We note, however, that the artifacts are substantially different for wavelet and for JPEG coders, and that a broad equivalence of quality may not be adequate for some applications. Defining and sharpening the coupling between the application and the relevant quality measure is a task that we have set for ourselves.

7 ACKNOWLEDGMENTS

This research was supported in part by the UC MICRO program, Lockheed and Hewlett Packard.

8 REFERENCES

[1] V. R. Algazi and Robert R. Estes, Jr. Analysis-based coding of image transform and subband coefficients. In *Digital Image Processing XVIII*, volume 2564 of *Proc. SPIE*, pages 11–21, 1995.

[2] V. Ralph Algazi, Phillip L. Kelly, and Robert R. Estes. Compression of binary facsimile images by preprocessing and color shrinking. *IEEE Transactions on Communications*, 38(9):1592–8, September 1990.

[3] M. Antonini, M. Barlaud, P. Mathieu, and I Daubechies. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1:205–20, April 1992.

[4] J. N. Bradley, C. M Brislawn, and T. Hopper. The FBI wavelet/scalar quantization standard for grayscale fingerprint image compression. In *Visual Information and Processing II*, volume 1961 of *Proc. SPIE*, pages 293–304, 1992.

[5] CCIR. *Rec. 500-2, Method for the subjective assessment of the quality of the television pictures*, 1982. Vol. 11, pp. 165–168.

[6] J. Chen, S. Itoh, and T. Hashimoto. Scalar quantization noise analysis and optimal bit allocation for wavelet pyramid image coding. *IEICE Trans. Fundamentals*, E76-A:1502–14, September 1993.

[7] C. K. Cheong, K. Aizawa, T. Saito, and M. Hatori. Subband image coding with biorthogonal wavelets. *IEICE Trans. Fundamentals*, E75-A:871–81, July 1992.

[8] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 41:909–96, 1988.

[9] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, PA, 1992.

[10] Allen Gersho and Robert M. Gray. *Vector Quantization and Signal Compression.* Communication and Information Theory. Kluwer Academic Publishers, 1992.

[11] Y. Horita and M. Miyahara. Image coding and quality estimation in uniform perceptual space. Technical report IE87-115, IEICE, January 1987.

[12] A. K. Jain. *Fundamentals of Digital Image Processing.* Prentice Hall, 1989.

[13] A. S. Lewis and G. Knowles. Image compression using the 2-D wavelet transform. *IEEE Transactions on Image Processing*, 1:244–50, April 1992.

[14] Jian Lu, V. Ralph Algazi, and Robert R. Estes, Jr. Comparison of wavelet image coders using the Picture Quality Scale (PQS). In *Wavelet Applications II*, volume 2491 of *Proc. SPIE*, pages 1119–1130, 1995.

[15] Jian Lu, V. Ralph Algazi, and Robert R. Estes, Jr. Evaluation and synthesis of wavelet image coders. In *ICIP*, volume 1, pages 590–3, October 1995.

[16] Jian Lu, V. Ralph Algazi, and Robert R. Estes, Jr. A comparitive study of wavelet image coders. *Optical Engineering*, 35(9), September 1996. To be published.

[17] M. Miyahara, K. Kotani, and V. R. Algazi. Objective picture quality scale (PQS) for image coding. *IEEE Transactions on Communications*, 1996. Accepted for publication. Also available as CIPIC report 92-12, University of California, Davis.

[18] W. B. Pennebaker, J. L. Mitchell, G. G. Landon, Jr., and R. B. Arps. An overview of the basic principles of the Q–coder adaptive binary arithmetic coder. *IBM Journal of Research and Development*, 32(6):717–26, November 1988.

[19] William B. Pennebaker and Joan L. Mitchell. *JPEG Still Image Data Compression Standard.* Van Nostrand Reinhold, New York, 1993.

[20] PVRG: Portable Video Research Group. This JPEG software is available at `ftp:/havefun.stanford.edu/pub/jpeg/JPEGv1.2.tar.Z`.

[21] Viresh Ratnakar and Miron Livny. Extending RD-OPT with global thresholding for JPEG optimization. In James A. Storer and Martin Cohn, editors, *DCC'96:Data Compression Conference*, pages 379–86, Snowbird, Utah, March 1996. IEEE Computer Society Press.

[22] O. Rioul. On the choice of 'wavelet' filters for still image compression. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, volume V, pages 550–3, 1993.

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

[24] John D. Villasenor, Benjamin Belzer, and Judy Liao. Wavelet filter evaluation for image compression. *IEEE Transactions on Image Processing*, 4(8):1053–60, August 1995.

[25] Andrew B. Watson. Visually optimal DCT quantization matrices for individual images. In James A. Storer and Martin Cohn, editors, *DCC'93:Data Compression Conference*, pages 178–87, Snowbird, Utah, March 1993. IEEE Computer Society Press.

[26] Zixiang Xiong, Kanna Ramachandran, and Michael T. Orchard. Space-frequency quantization for wavelet image coding. *IEEE Transactions on Image Processing*, 1997. To appear.

[27] Zixiang Xiong, Kannan Ramchandran, and Michel T. Orchard. Joint optimization of scalar and tree-structured quantization of wavelet image decompositions. In 27^{th} *Annual Asilomar Conference on Signals, Systems, and Computers*, pages 891–5, Pacific Grove, CA, November 1993.

[28] W. R. Zettler, J. Huffman, and D. C. P. Linden. Application of compactly supported wavelets to image compression. In *Image Processing Algorithms and Applications*, volume 1244 of *Proc. SPIE*, pages 150–60, 1990.