Abstract—We implement image compression using various wavelet filterbanks and measure performance with rate distortion characterizations. Various separable filterbanks are chosen and compared. Coefficients in the subbands obtained by wavelet decomposition are quantized using various schemes. The average entropy of the quantized subband coefficients is taken to be a measure of the information rate. The image is then reconstructed from the quantized coefficients, and distortion is measured. Two distortion measures are used: mean square error (MSE) and the mean structural similarity measure (MSSIM), a perceptual distortion measure. The final results are rate distortion operating points for the various quantization and decomposition schemes.

I. INTRODUCTION

 RATE Distortion theory is concerned with representing a source, such as an image, with the minimum number of bits for a given reproduction quality. The necessity for image compression arises from the fact that communication networks and storage devices have limited capacities. The fundamental tradeoff in compression is between rate and distortion.

 Transform coding, a method for image compression, transforms the image into a domain which has a sparse representation, allowing better compression. Natural signals generally have spectral energy distributed inversely proportional to frequency. Therefore, constant-Q filterbank decomposition results in an efficient representation. Multiscale wavelet transforms have traditionally been used for this decomposition [1].

II. DECOMPOSITION AND RECONSTRUCTION

 Discrete wavelet transforms are simple linear transformations that allow signals to be decomposed into relatively uncorrelated coefficients. A separable two dimensional discrete wavelet transform can be described in terms of a standard filterbank as shown in Fig. 1.

 The analysis and synthesis filters in Fig. 1 are chosen to meet the Smith-Barnwell perfect reconstruction condition. Scaling is avoided by normalizing the filters and the delay introduced by the filterbank is compensated at the output to obtain an exact copy of the input image.

 Linear phase filterbanks chosen are reasonably short (less than 36 taps in the synthesis/analysis pair) minimum-order biorthogonal wavelet filterbanks that have been shown to have good impulse and step responses for image compression [2]. We test the performance of five of these filterbanks as well as the Haar wavelet filterbank using our image compression implementation.

 Fig 2. shows an image subband decomposition and reconstruction scheme created by cascading several levels of filterbanks with each successive filterbank operating on the low frequency subband. This decomposition allows the signal energy to be concentrated in the Nth subband. If an equal number of bits are allocated to each subband, the Nth subband receives the most bits per coefficient due to smaller band size. Thus decomposition into a large number of subbands before coefficient quantization results in a more efficient bit allocation.

III. COMPRESSION

 Having split the image into N-subbands, the coefficients must be quantized for compression. Several quantization schemes can be conceived. We considered three:

1) Integer coefficients: coefficients rounded to nearest integer
2) Dropped Subbands: coefficients except Nth band dropped
3) Inverse Variance: 8 bits allocated to Nth band; other bands use quantization step size inversely proportional to the logarithm of band coefficient variance.

Schemes 2 and 3 exploit the capability of the human visual
system to extract information from highly distorted images if their low frequency content is largely intact.

After quantization, the lossy portion of the compression has been completed. Lossless compression on the quantized coefficients is used to further reduce data rate. The output of an ideal lossless coder would have a data rate equal to the entropy of the signal. The entropy of the quantized coefficients is thus a bound on the performance of the compression scheme, a good measure of the data rate required to code the image. The entropy is calculated as the size-weighted average of the entropies of the subbands. If one considers the alphabet of unique symbols in each subband $S$, and determines the relative frequency of each of these unique symbols, $p_{i,S}$, the entropy of subband $S$ is

$$H_S = -\sum_{i \in S} p_{i,S} \log_2 p_{i,S},$$

measured in bits per coefficient. The average entropy of the entire subband coded coefficient set is then

$$H = \sum_{i=1}^{N} \|S_i\| H_{S_i}/\sum_{i=1}^{N} \|S_i\|,$$  \hspace{1cm} (2)

IV. DISTORTION

When an image is compressed using lossy compression, there is inherently a reduction in the quality of the decompressed image. In order to quantify the notion of quality, a distortion measure is introduced. Good distortion measures should possess three qualities: mathematical tractability, computability and subjective meaning, so that small distortions correspond to good perceived quality [3]. We investigated two distortion measures, mean square error (MSE), and the mean structural similarity measure (MSSIM) [4]. MSE is highly tractable mathematically, although its subjective meaningfulness is questionable [5]. Contrarily, MSSIM is not as tractable mathematically, but has good subjective meaning. Both measures are easily computable.

The MSE is calculated as

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [I(i,j) - \hat{I}(i,j)]^2,$$  \hspace{1cm} (3)

where the two images under comparison, $I$ and $\hat{I}$, have size $M \times N$.

The MSSIM is a perceptual distortion measure that operates under the assumption that the human visual system extracts and codes structural information from a scene. The measure considers structural similarity, as well as luminosity and contrast similarities between two images. It operates locally on image blocks and is then averaged over the entire image. MSSIM is bounded between 0 and 1, with equality to 1 when the two images are identical. Since we desire that the distortion between two identical images be 0, the MSSIM we use is actually $-\text{MSSIM} + 1$.

The rate distortion (RD) characterization of a compression scheme succinctly captures the tradeoff between the number of bits used to represent an image and the quality of the reconstruction. Although this RD characterization resembles the RD function from information theory, the two functions are not identical. The information theoretic RD function gives the infimum of the rate for a given distortion, whereas our RD curve simply indicates the minimum rate for a given distortion that we can achieve using our compression scheme. The experimental RD characterizations, obtained by determining the convex hull of the data points is decreasing, implying that with greater compression, there is greater distortion. RD characterizations can be used to directly compare quantization schemes and wavelet filterbanks. The closer to the origin a particular RD curve is, the better the performance.

V. RESULTS

We applied our various image compression schemes on the standard test images, “Barbara,” “Mandrill,” “Fingerprint,” and “Facets.” The experimentally determined RD functions are given in Figs. 3-6. The filterbanks are numbered as in Table I of [2]. Results for integer coefficient quantization are not shown since performance differences for different filterbanks were negligible. Figs. 7-11 correspond to the labels Image 1-6 (except 2) on the RD curves and show examples of test images compressed using those specific decomposition and quantization schemes.

VI. CONCLUSIONS

We have characterized the performance of several biorthogonal filterbanks for wavelet image compression with two different distortion measures and three different quantization schemes. It is clear that the choice of filterbank as well as the quantization scheme affect the resultant compression achieved and distortion observed. Figs. 7 and 11 suggest that the Dropped Subband Quantization scheme gives better compression but much worse perceptual and MS error when compared with the Inverse Variance Quantization scheme for the same filter. Thus, dropping all high frequency band information is not a good idea. Fig. 10 indicates that the computationaly easy to implement Haar basis filters give a high MSE distortion, but perform very well on the perceptual measure. Clearly, selection of appropriate filterbanks for compression depends on the application. Ease of
implementation and performance, based on relevant distortion measures, are often as significant as the level of compression desired.

Fig. 3. Rate-MSE Functions using Dropped Subband Quantization.

Fig. 4. Rate-MSSIM Functions using Dropped Subband Quantization.

Fig. 5. Rate-MSSIM Functions using Inverse Variance Quantization.

Fig. 6. Rate-MSE Functions using Inverse Variance Quantization.

Fig. 7. (a) Original Fingerprint. (b) Compressed Fingerprint using filter 1 with 5 levels of decomposition. Has best MSE (323.6) to bit rate (.0183 bpp) ratio for Dropped Subband Quantization and MSSIM of 0.6715.

ACKNOWLEDGMENT

The authors would like to thank Yegnaswamy Sermadevi, Damon Chandler, and Mark Masry for their guidance in selection of the perceptual distortion measure.

REFERENCES


Fig. 8. (a) Original Facets. (b) Compressed Facets using filter 2 with 2 levels of decomposition and Dropped Subband Quantization. It has an MSE of 161.4, MSSIM of 0.1137 with an entropy of 0.9607 bpp.

Fig. 9. (a) Original Barbara. (b) Compressed Barbara using filter 1 with 4 levels of decomposition and Dropped Subband Quantization. It has an MSE of 331.0, an MSSIM of 0.4886 with an entropy of .0823 bpp.

Fig. 10. Compressed Fingerprint using filter 7 (Haar basis) with 5 levels of decomposition and Inverse Variance Quantization. It has an MSE of 63.7, an MSSIM of 0.1621 and an entropy of 0.6359 bpp.

Fig. 11. Compressed Fingerprint using filter 1 with 5 levels of decomposition and Inverse Variance Quantization. It has an MSE of 88.0, MSSIM of 0.2188 and bit rate of 0.5236 bpp.