01 Jan 2001

Performance Evaluation of Hybrid Coding of Images Using Wavelet Transform and Predictive Coding

S. R. Subramanya

Chaman Sabharwal
Missouri University of Science and Technology, chaman@mst.edu

Recommended Citation

S. R. Subramanya and C. Sabharwal, "Performance Evaluation of Hybrid Coding of Images Using Wavelet Transform and Predictive Coding," *Proceedings of the Fourth International Conference on Computational Intelligence and Multimedia Applications, 2001*, Institute of Electrical and Electronics Engineers (IEEE), Jan 2001.
The definitive version is available at https://doi.org/10.1109/ICCIMA.2001.970507

Performance Evaluation of Hybrid Coding of Images Using Wavelet Transform and Predictive Coding

S.R. Subramanya[‡,§] and Chaman Sabharwal[‡]
[‡]Department of Computer Science
[§] Intelligent Systems Center
University of Missouri–Rolla
1870 Miner Circle
Rolla, MO 65409
{subra,chaman}@umr.edu

Abstract

Image compression techniques are necessary for the storage of huge amounts of digital images using reasonable amounts of space, and for their transmission with limited bandwidth. Several techniques such as predictive coding, transform coding, subband coding, wavelet coding, and vector quantization have been used in image coding. While each technique has some advantages, most practical systems use hybrid techniques which incorporate more than one scheme. They combine the advantages of the individual schemes and enhance the coding effectiveness. This paper proposes and evaluates a hybrid coding scheme for images using wavelet transforms and predictive coding. The performance evaluation is done using a variety of different parameters such as kinds of wavelets, decomposition levels, types of quantizers, predictor coefficients, and quantization levels. The results of evaluation are presented.

1 Introduction

Image compression techniques take advantage of the spatial and spectral redundancies generally present in images to reduce the number of bits required to represent the images. Image compression schemes are broadly classified as either *lossless* or *lossy*, depending respectively on whether the compressed image can be *exactly* recovered or not. Generally, lossy schemes provide a much higher compression ratios than lossless schemes. In addition to the above mentioned redundancies, lossy techniques make use of properties of the human visual system (HVS) such as more sensitivity of the eye for lower frequencies than for higher frequencies, to achieve higher compression. Major characteristics to be considered in any compression scheme are the bit rate (average number of bits per pixel), the peak-signal-to-noise ratio (PSNR) of the reconstructed image with respect to the original, and the speed of encoding and decoding. In this paper, we use the terms *coding* and *compression* synonymously.

There are several lossy compression techniques such as prediction-based coding, transform coding, block truncation coding, vector quantization, and subband coding, etc. [3, 6, 7, 8]. Most of the practical image compression systems and standards are *hybrid schemes*, utilizing a combination of more than one technique, such as: (a) transform coding and predictive coding, (b) subband coding and transform coding, and (c) predictive coding and vector quantization. They also use entropy coding schemes as the final phase to further compress the output of the earlier stages.

In this paper, we propose and evaluate the performance of a hybrid coding scheme for images which uses a combination of wavelet transform and predictive coding. Techniques using wavelets have been popular for the decomposition of images into components of different resolutions (multiresolution). Predictive coding is a simple and effective scheme which could be used in lossless or lossy mode. The performance evaluation parameters are types of wavelets, decomposition levels, types of quantizers, predictor coefficients, and quantization levels. The evaluation metrics are the bit rate and the reconstruction quality measured by PSNR.

The next Section briefly describes the principles of wavelet transforms, DCT coding, and predictive scheme of coding of images. Section 3 describes the proposed hybrid image coding scheme. Experimental results are given in Section 4, followed by conclusions.

2 Background

This section provides an overview of the fundamental techniques of wavelet decomposition, predictive coding, and quantization.

2.1 Wavelet decomposition

Wavelets are mathematical functions which extract different frequency components from a given data, and study each component with a resolution matched to its scale. The Discrete Wavelet Transform (DWT) analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse component represented by *approximation* coefficients and finer components represented by *detail* coefficients. In this sense, the wavelet decomposition of a signal is similar to *subband coding*, where, a signal is passed through several band-pass filters to obtain signal components at different bandwidths, for subsequent analysis and processing. DWT employs two sets of functions called *scaling functions* and *wavelet functions*, which are associated with *low-pass* and *high-pass* filters, respectively. The decomposition of the signal into different frequency bands is obtained by successive low-pass and high-pass filtering of the signal and down-sampling the coefficients after each filtering.

Since a half-band low-pass filter removes half of the frequencies present in the input, the frequency resolution of its output is twice that of its input. After the downsampling, the time resolution is halved since the output contains only half the number of points as the input. Thus the combination of filtering and down sampling increases the frequency resolution but decreases the time resolution. usually, only the low-frequency subbands are successively decomposed into two (finer) subbands, while the high-frequency subbands are untouched. Thus the approximation coefficients (low frequencies) have a high frequency resolution but poor time resolution, while the detail coefficients of level 1 (high frequencies) have a poor frequency resolution but good time resolution. Note that the frequency and time resolutions of the detail coefficients progressively change in successive levels. Thus, in DWT (unlike DFT or DCT), the time localization of the frequencies are not lost. However, the time localization will have a resolution that depends on which level they appear. More details can be found in [2, 1].

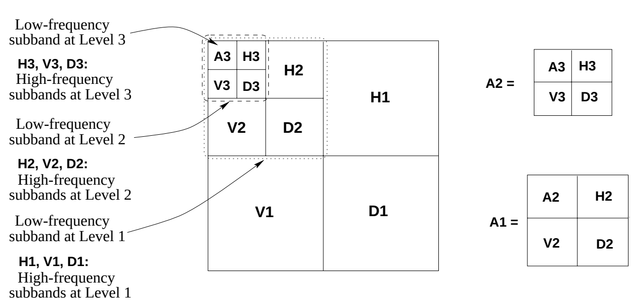

Figure 1. 3-level wavelet decomposition of an image.

The decomposition an input image data up to three levels using wavelets is shown in Figure 1. The data at each level is decomposed into four components, each containing different frequency bands: A, H, V, and D. A (approximation) corresponds to the low-frequency subband, while H (horizontal details), V (vertical details), and D (diagonal details) correspond to high-frequency subbands. The frequencies that are most prominent in the original signal will appear as high amplitudes in that band of the DWT transformed coefficients which includes those particular frequencies, while the frequencies that are not prominent in the original signal will have low amplitudes in those parts of the DWT transformed coefficients which contain those frequencies.

2.2 Predictive Coding

The basic idea behind this scheme is to predict the pixel values of a given image based on the pixels in the neighborhood. Typically, the prediction is a linear combination of the neighborhood pixels. In case an image is to be predicted from a smaller sized image, some kind of interpolation is used. The common schemes are nearest-neighbor, bi-linear, and bi-cubic interpolations. Then the *residual image* is derived, which is the difference between the original and predicted image. The values in the residual image have much less dynamic range and variance compared to the original image. This is beneficial for achieving good compression. In addition, with good prediction, the residual image would have large runs of zero values to which RLE (Run Length Encoding) could be applied.

2.3 Quantization

The basic function of a quantizer is to map a large range of (possibly continuous) values onto a relatively smaller set of (discrete) values. In the context of image coding and decoding, a quantizer has a *decision vector* and a *reconstruction vector*. The decision vector determines the *quantization level* l for any given pixel x: $l = Q(x)$, and the reconstruction vector determines the *reconstructed pixel* x' value for a given quantization level: $x' = Q^{-1}(l)$. There are different types of quantizers which essentially differ in terms of how the decision and reconstruction vectors are determined.

In a uniform quantizer of k levels, the decision vector consists of k equally spaced intervals from the minimum pixel value to the maximum pixel value in the input, where k is the number of quantization levels. The reconstruction vector consists of the mid-points of the decision intervals.

The most commonly used non-uniform quantizer is the *Lloyd-Max quantizer*. In the Lloyd-Max quantizer, the probability distributions of the pixels are used to determine the optimum decision and reconstruction levels, while keeping the quantization error minimum with respect to the mean-square-error metric [4, 5]. This results in non-uniform decision levels. The decision levels are halfway between the neighboring reconstruction levels and the reconstruction levels are the *centroids* of the two adjacent decision levels. These levels are determined by an iterative algorithm.

3 The Proposed Hybrid Coding Scheme

The outline of the proposed hybrid coding of images using Wavelet transform and predictive and differential coding is shown in Fig. 2. The input image is first decomposed up to a certain number of decomposition levels, say n, into low and high frequency components (subbands) using a suitable discrete wavelet transform (DWT). The original image is considered to be at level 1. The low-frequency subband of the highest level A_n is a miniature version of the original image. The elements of A_n are also referred to as *approximation coefficients*. These contain most of the information in the original image. The other subbands, H, V, and D contain details in the image such as edges, lines, and boundaries, typically information corresponding to higher frequencies. The elements in these subbands are also referred to as *detail coefficients*. It is also known that the human visual system (HVS) is sensitive to distortions in the low-frequency components in images (smoother areas) than to distortions in the higher-frequency components (edges, lines, etc.). So

A_n needs to be encoded with minimal loss or no loss of information. A_n is subjected to DCT coding or lossless predictive coding, and the resulting data is entropy coded using Huffman coding. The coefficients in all the H, V, and D could be subjected to higher compression. Subband H at level i, H_i is predicted using the (quantized) coefficients in H_{i+1}, for $i = n-1 \ldots 1$. This is done using any of the interpolation schemes mentioned in Section 2.2. Similarly V_i and D_i are predicted. Then the corresponding residuals which are difference between the original and predicted subbands are determined. For example, the residual of H_i subband, $RH_i = H_i - PH_i$, where PH_i denotes the prediction of the H_i subband. These residuals are the quantized using uniform or non-uniform quantizers. The quantized residuals have much less dynamic range of values compared to the original subbands, and also contain sequences of similar values (runs). These make them amenable for much higher compression.

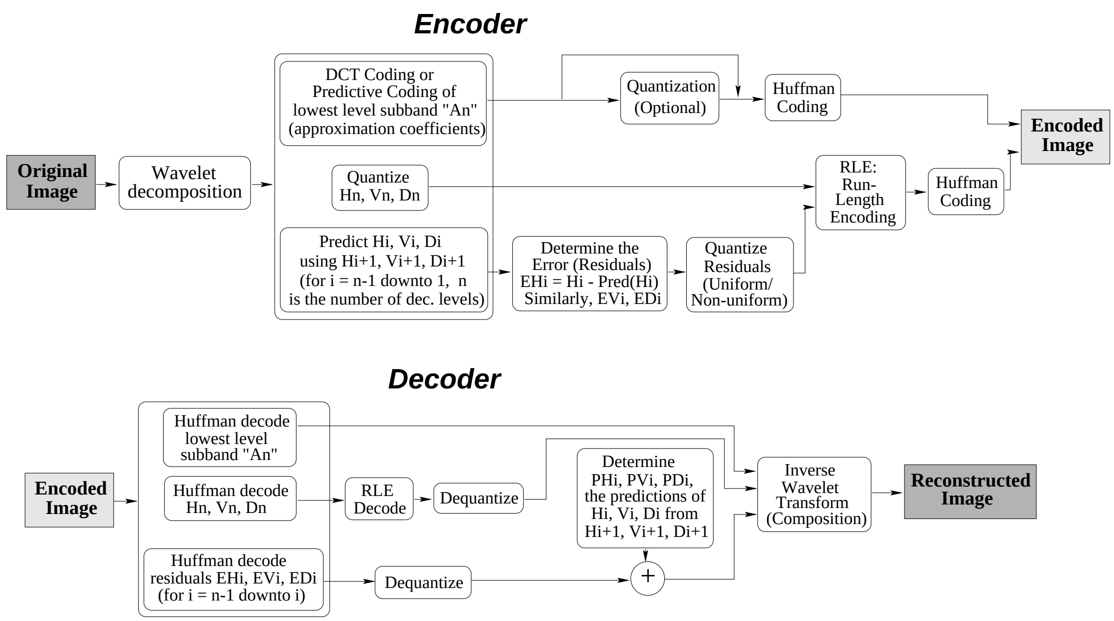

Figure 2. Outline of hybrid coding of images.

The decoding proceeds in the reverse fashion. The encoded low frequency components of the highest level, A_n are decoded. The the other subbands at level n, H_n, V_n, and D_n are decoded and dequantized. Then the subbands at level $i-1$ are successively predicted using the corresponding subbands at level i, which are added to the decoded residual at level i to get the reconstruction of the subbands at level i. These are composed using the inverse wavelet transform to get the low-level (approximation) subband at the next lower level. The procedure is repeated until level 1 is reached. This scheme also offers the advantage of reconstruction at various resolutions which can be used in a multi-use environment consisting of devices with various resolutions (e.g. a high resolution monitor, a low-resolution printer, etc).

4 Experimental Results

The scheme has been implemented using MATLAB on a Sun Ultra 10 workstation. It has been tested on over a hundred different kinds of images.

The parameters used in the evaluation are wavelet types, the number of decomposition levels, quantization levels, types of quantizers, and predictor types. The evaluation metrics are PSNR (Peak Signal to Noise Ratio) and the bit rate. The PSNR roughly corresponds to the quality of the reconstructed image in terms of its closeness to the original. The PSNR for an 8-bit gray scale image is given by: $PSNR = 20\log_{10}(255/\sqrt{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij}-x'_{ij})^2})$ where x_{ij} and x'_{ij} are the pixel values in the original and reconstructed images. The bit rate is the average number of bits per pixel (BPP). The PSNR and BPP for different wavelet decompositions and quantizer levels is tabulated in the table below.

Wavelet Type	Quantizer Levels					
	2 Levels		4 Levels		8 Levels	
	BPP	PSNR	BPP	PSNR	BPP	PSNR
Haar	0.6622	11.34	0.1472	12.11	0.5083	14.82
DB2	0.3698	11.82	0.2203	12.32	0.5046	14.70
DB3	0.8777	9.92	0.2530	11.09	0.3411	13.22

The quantized values using a 2-level quantizer for the H, V, and D subbands at level 2 using DB2 wavelet is shown in Fig. 3 (top). The distribution of the lengths of the runs in the subbands are shown below the corresponding bands.

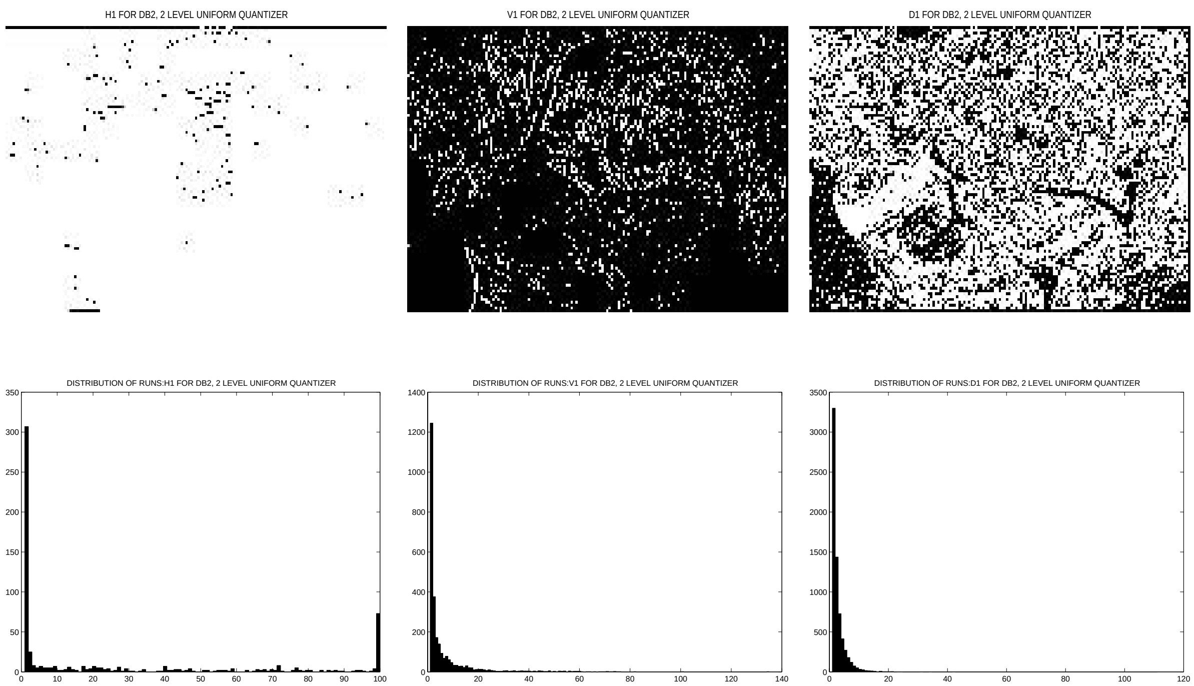

Figure 3. Quantized H, V, and D subbands (DB2, level 2) and the corresponding distributions of runlengths of coefficients.

The original image is shown far left in Fig. 4. Representative reconstruction values using Haar wavelet and 4-level uniform quantizer, and for Haar wavelet and 4-level Lloyd-Max quantizer is shown in the middle and far right in Fig. 4, respectively.

Among different quantizer levels, the 4-levels gave consistently good performance. For this quantizer level, the Haar wavelet performed best in terms of bit rate, followed by DB2 and DB3 wavelets. The results are for two-level decompositions, and for bilinear prediction. Bilinear prediction scheme works better compared to nearest-neighbor and is faster compared to bicubic interpolations. DB3 gave better results with respect to PSNR.

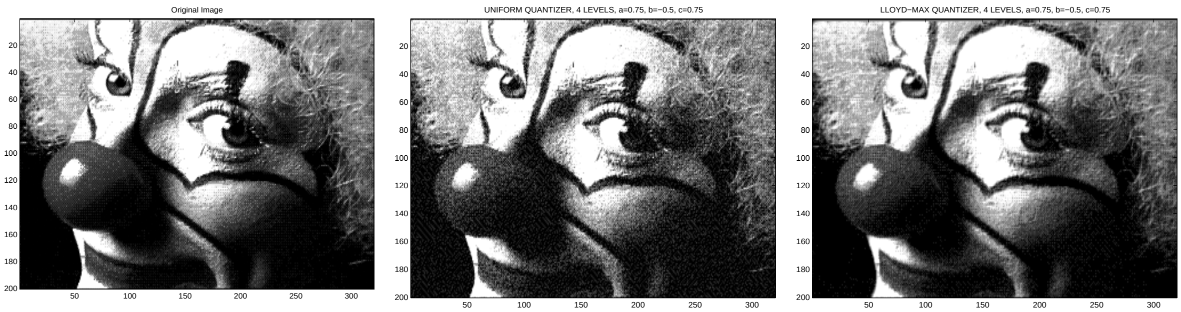

Figure 4. Original and reconstructed images.

5 Conclusions

This paper proposed and evaluated the performance of a hybrid coding of images using a combination of wavelet transform and predictive coding. Wavelet transform was used for successively decomposing the original image and the low-frequency subbands into several levels of low and high-frequency components. Predictive and differential coding was used to encode the low-frequency subband at the highest level. Predictive coding was used to encode the high-frequency subbands H, V, and D at various levels. The performance was evaluated using a variety of parameters related to each of the coding techniques. Using uniform quantizer of four levels gave the best performance for most images for all the types of wavelets that were used. Using Haar wavelet with 4-level quantizer gave the best bit rate, overall. A particular combination of parameters can be selected based on the required image quality and/or compression ratio for an application.

Acknowledgment. We wish to thank Sanjay Shankar for implementing and testing the proposed scheme.

References

[1] C.S.Burrus, R.A.Gopinath, and H.Guo, *Introduction to Wavelets and Wavelet Transforms*, Prentice-Hall, 1998.

[2] A.Graps, ‘An Introduction to Wavelets’, *IEEE Computational Science and Engineering*, Vol.2, No.2, Summer 1995.

[3] A.K.Jain, ‘Image Data Compression: A Review’, *Proc. IEEE*, 69(3), 1981, pp349–389.

[4] S.P.Lloyd, ‘Least Squares Quantization in PCM’, Bell Lab. Memo, July 1957; also in *IEEE Trans. Info. Theory*, IT-28, 129–137, 1982.

[5] J.Max, ‘Quantizing for Minimum Distortion’, *IRE Trans. Info. Theory*, 6(1), 7–12, 1960.

[6] A.N. Netravali and B.G. Haskell, *Digital Pictures: Representation and Compression* (2nd. Ed.), Plenum Press, 1995.

[7] M.Rabbani and P.W.Jones, *Digital Image Compression Techniques*, SPIE, Vol. TT7, 1991.

[8] K.Sayood, *Introduction to Data Compression*, 2nd Ed. Morgan-Kaufmann, 2000.