

Human Visual System Based Wavelet Decomposition for Image Compression

Thomas P. O'Rourke and Robert L. Stevenson

Abstract—

Due to limitations in bandwidth or storage space, many applications require compression of digital images. One of the main hindrances to progress in the area of image communications is the large amount of information contained in an image. In this paper we propose a wavelet-based image compression technique which incorporates some of the properties of the human visual system (HVS). A new wavelet decomposition based on the HVS will be introduced. The wavelet coefficients will be quantized with a vector quantization technique which uses HVS properties to allocate bits to the various subbands. The new HVS based wavelet decomposition will be compared with the existing separable and non-separable wavelet decompositions. Compression techniques based on such considerations yield higher quality results compared to standard techniques. Several experimental results will be shown.

I. INTRODUCTION

In many applications requiring image compression, a human observer is the final receiver of the image information. Human visual system (HVS) properties must be taken into consideration to achieve quality reconstruction of images at rates below 0.8 bits per pixel (bpp) [1]. The HVS spatial frequency response was used as a weighting function for the cosine transform by Nill [2] to compensate for the properties of the HVS. Safranek and Johnston [3] used HVS perceptual considerations in a subband coding context. HVS properties have been used to solve the bit allocation problem for the DCT by Macq[4], for the lapped orthogonal transform by Queiroz and Rao [5], for the wavelet transform with vector quantization (VQ) by Antonini and others [6], and for subband coding by Van Dyck and Rajala [7]. In the proposed compression system, a new HVS based decomposition is presented. The image compression is achieved by VQ of the wavelet transform coefficients. The VQ encoder uses a new bit allocation method based on the HVS.

Subband coding techniques have been found to be useful in the processing of sound and image signals. Subband coding decomposes a signal into several subband signals which have their frequency content concentrated in a particular frequency range. Since more error can be tolerated at frequencies where the observer has lower sensitivity to noise, knowledge of the frequency response of a human observer can be used in the coding of each bandpass signal

[3]. Subband coding is of particular interest since the discrete wavelet transform can be computed using a subband coding structure[8]. The wavelet transform can be considered as a special case of perfect reconstruction subband coding with additional regularity requirements on the filters[9]. General background on subband coding and further information on perfect reconstruction subband filters can be found in Vetterli [10; 11], Smith and Barnwell [12; 13], and Vaidyanathan [14]. See Woods [15] for general coding of images using subbands and [16] for a good overview of image subband coding.

The advantage of the wavelet transform over more familiar transforms such as Fourier or cosine lies in the localization of wavelets in both spatial and frequency domains [17] and the multiresolution nature of wavelets [18]. A discontinuity in a signal, such as an edge in an image, is a very local structure in the spatial domain. Much of the information in an image is contained in the location and magnitude of edges. Contrary to the goal of transform coding, which is to pack the information in as few coefficients as possible, the Fourier transform spreads the discontinuity information across a large range of transform coefficients. The use of the wavelet transform for image compression was suggested in [8] using a separable extension of one-dimensional (1-D) wavelet theory to two dimensions. The use of compactly supported wavelets, introduced in [19], greatly reduces the computational cost of the wavelet transform. A non-separable extension to 2-D is discussed in [20]. Image compression involving the wavelet transform is discussed in [21; 22; 23] and in combination with VQ in [6]. For a good overview of wavelet theory with an extensive bibliography, see [9]. An introduction to the mathematics of wavelets can be found in [24] with more in-depth treatment in [25; 26; 27]. Thorough background on wavelets and signal analysis can be found in [8; 19; 18; 17].

Section II outlines properties of the HVS which are used for image coding. Section III describes two existing wavelet decompositions and concludes with a new wavelet decomposition based on the HVS properties. The compression of the 2-D wavelet decomposition representation is discussed in Section IV. Section IV.A describes how HVS properties are used to allocate bits to the various subbands and VQ of the wavelet transform coefficients is described in Section IV.B. The remainder of Section IV compares the HVS-based wavelet decomposition with other wavelet decompositions and the JPEG compression standard.

To appear in *Journal of Visual Communication and Image Representation*. Revised November 5, 1993.

The authors are with the Laboratory for Image and Signal Analysis, Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA, <http://lisa.ee.nd.edu/>.

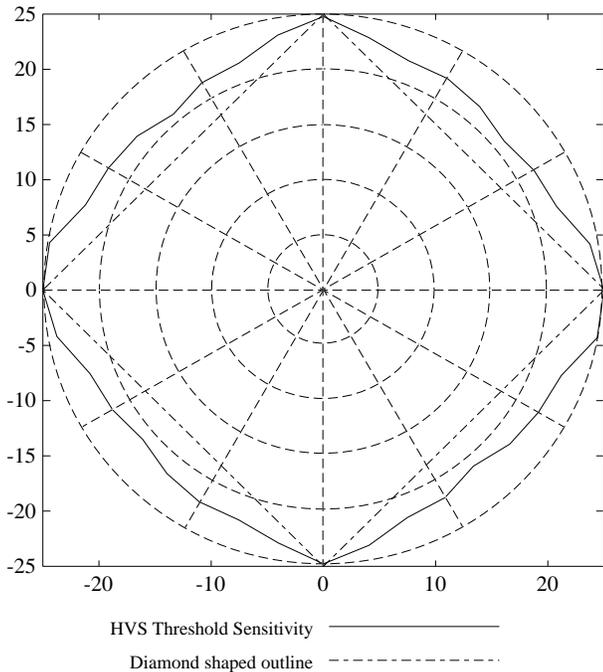


Fig. 1. Resolvability threshold for sinusoidal gratings, cycles per degree.

II. RELEVANT HUMAN VISUAL SYSTEM CHARACTERISTICS

This brief overview of useful information on the human visual system (HVS) is based on research done in the 1960's by Campbell and others [28; 29; 30; 31]. In the proposed compression technique, two important characteristics of the HVS will be incorporated into the compressed data representation. These properties are the orientation sensitivity of the HVS and the contrast sensitivity of the HVS. The proposed technique does not directly exploit the sensitivity masking effect of the HVS. The findings of [28; 29; 30] were based on psychophysical experiments. The subject had to determine the threshold where a sinusoidal grating was just noticeable. In [31], electrodes are used to measure the electric potential of nerves. The subject is exposed to a sinusoidal grating and measurements are made of the changes in potential of the nerves. The threshold was determined by observing at what point the potential changed. These electrophysiological measurements confirm the findings of the previous studies.

In an experiment which examined the sensitivity of the HVS to frequencies at various spatial orientations [28], it was found that the HVS could better resolve frequencies along the horizontal and vertical orientation than along the diagonal orientations. This can be seen in Figure 1 which plots the point for which the spatial frequency at each orientation can no longer be resolved by the HVS. In this experiment, the contrast was held constant and the threshold frequency was determined for various orientations. Notice that this experimental data [28] is roughly diamond shaped as evident by the diamond which is also plotted on the graph. This implies that the passband of the subband coding system should be diamond shaped to

take advantage of this HVS characteristic.

In the same set of experiments [28], the contrast sensitivity of the HVS was also examined. In this experiment, the orientation was fixed and the threshold contrast was determined for various frequencies. For spatial frequencies above 7 cycles per degree, it is found that the logarithm of the contrast sensitivity is a linear function of spatial frequency. The slope of this linear function was found to be different for horizontal and vertical orientations relative to the oblique 45 and 135 degree orientations. The log sensitivity approximately satisfies

$$\log S(f) = \log 500(1 + \frac{1}{-50}f) \quad (1)$$

for the vertical and horizontal orientation and

$$\log S(f) = \log 500(1 + \frac{1}{-40}f) \quad (2)$$

for the oblique orientation, where f is the spatial frequency in cycles per degree [28].

The optical portion of the HVS can be considered as a linear system. Based on the experiments in [30], it was concluded that the neurological portion of the HVS also operates in a linear fashion. The linearity of the HVS is significant since most signals can be expanded in a Fourier series as a linear combination of sinusoids. From the linearity of the HVS, the response to such a signal is a linear combination of the response to sinusoidal gratings.

The HVS has a diamond shaped frequency passband which will be exploited in the design of the proposed image compression scheme. Additionally, the logarithm of the contrast sensitivity is approximated by a linear function which will be used for bit allocation. The perceived quality of the reconstructed images will be improved by consideration of the HVS in the design of the image compression system.

III. SUBBAND CODING WITH WAVELETS

A. Separable Wavelets

The separable wavelet transform [8] is the most common method for extending the one-dimensional wavelet transform to higher dimensions. Filtering is performed at each level first along one dimension and then in the other dimension resulting in the decomposition of the frequency space shown in Figure 2. All 1-D filters used in this paper are Daubechies' compactly supported wavelet filters [19]. The original image can be recovered exactly from its wavelet decomposition provided the transform coefficients are not quantized.

B. Non-separable Wavelets

The rectangular subbands in Figure 2 are a poor match for the diamond shaped HVS frequency response. Ansari and others [32; 33] use a diamond shaped filter which results in the frequency decomposition shown in Figure 3 when the low-pass branch is iterated. The sampling pattern of the subband signals alternates by level between

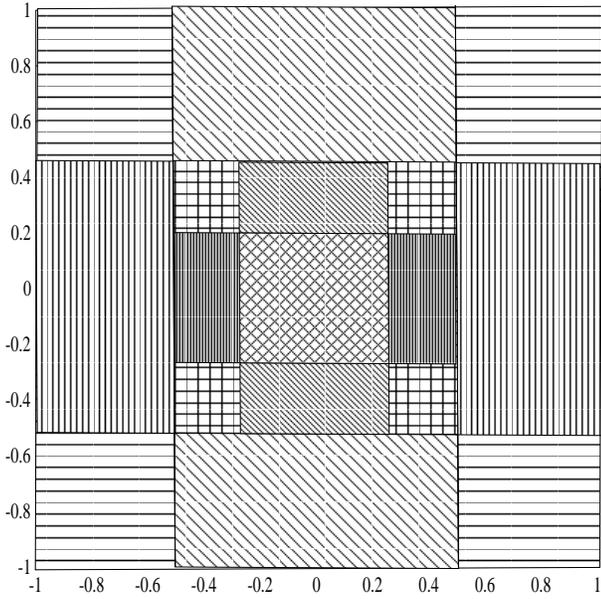


Fig. 2. Separable wavelet decomposition of frequency space, 2 levels.

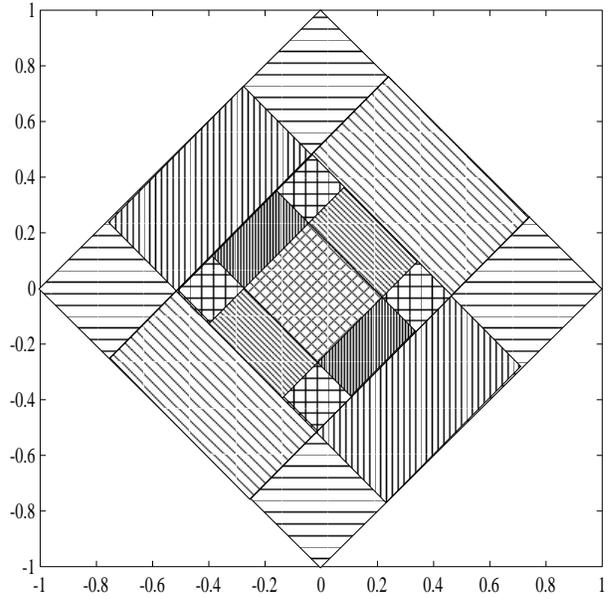


Fig. 4. HVS-based wavelet decomposition of frequency space, 3 levels.

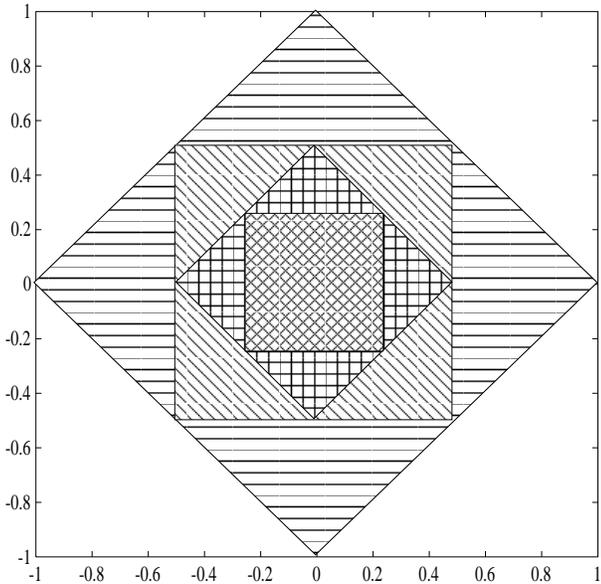


Fig. 3. Non-separable wavelet decomposition of frequency space, 4 levels.

rectangular and quincunx. The 8-tap non-separable orthonormal compactly supported wavelet of Kovačević and Vetterli [20] is the diamond shaped filter used in this paper. Unfortunately, the regularity of this non-separable filter cannot easily be increased. The regularity of the 2-D separable filters could be increased easily by increasing the order of the 1-D filters. The quality of the reconstructed images increases as the regularity of the filters is increased, see Section IV.C.

C. HVS-based Wavelets

The HVS-based decomposition proposed here is intended to avoid the problems associated with the decompositions described in sections III.A and III.B. The first stage of this

decomposition uses a filter with a diamond shaped 2-D frequency response. The input is rectangularly sampled and the output is quincunx sampled. This stage is the same as the first stage in the decomposition in section III.B. Because this stage is not repeated, it does not matter whether the iterated filter would converge to a regular function or not. The next stage and all following stages use a four-band decomposition similar to the one used in section III.A.

The quincunx sampling pattern can be viewed as a rectangular sampling pattern which has been rotated 45 degrees and scaled by $\sqrt{2}$. The new coordinates for the rotated sampling pattern in terms of the rectangular sampling pattern coordinates can be found by [34]

$$\begin{aligned} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}' &= \begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} \cos 45^\circ & -\sin 45^\circ \\ \sin 45^\circ & \cos 45^\circ \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \\ &= \begin{bmatrix} n_1 - n_2 \\ n_1 + n_2 \end{bmatrix} \end{aligned} \quad (3)$$

In the frequency domain, this amounts to a rotation by -45 degrees and a scaling by $\frac{1}{\sqrt{2}}$. The four-band decomposition shown in Figure 2 is rotated and scaled and the new HVS based frequency decomposition is shown in Figure 4. The diamond shaped subbands are meant to fit the diamond shaped sensitivity of the HVS shown in Figure 1. The 1-D filtering is performed along the 45° and 135° diagonal directions.

This HVS decomposition has the advantage over the non-separable diamond shaped wavelet because the regularity of the wavelet can be increased by increasing the order of the 1-D filters from which it is derived. The regularity of the non-separable wavelets cannot easily be increased as is the case for the HVS based wavelets.

The HVS based decomposition is separable along the 45° and 135° diagonals and has all the advantages of the separable decomposition. The HVS based decomposition

has an advantage over standard separable decompositions because it is oriented to fit the diamond shaped HVS sensitivity.

Van Dyck and Rajala use a similar decomposition in [35]. The first stage there is an interpolation of the rectangular sampling pattern to obtain a quincunx sampling pattern. Following stages also use 1-D filters along the 45° and 135° diagonal directions.

IV. WAVELETS FOR IMAGE COMPRESSION

The quantization technique used here is VQ [36]. Good results have been found using scalar quantization, entropy coding, and HVS considerations [21; 22]. VQ was chosen, however, because it has been shown to be potentially superior to any scalar quantization technique [36]. The three wavelet decompositions from Section III are quantized and the resulting reconstructed images are compared. It is shown that the HVS-based wavelet decomposition has better looking artifacts. In comparing the HVS-based wavelet decomposition with VQ to the standard JPEG compression technique, it is found that the HVS-based wavelet VQ outperforms JPEG at high compression rates. HVS-based wavelet VQ performs better in smooth regions but aliasing artifacts are more visually annoying at edges.

A. HVS-based Bit Allocation

The purpose of this section is to describe how knowledge of the HVS is used to determine the allocation of bits to the various subbands of the proposed image coding system. The 1-D case will be examined first and extension to the 2-D case will be described next.

A two-band subband system with a low-pass channel and a high-pass channel can be cascaded to form a system with several subbands of different bandwidth and sampling rate. Due to the downsampling, the sampling rate of the subband signals will be lower than the original sampling rate by a factor of two for every stage the signal passes through. The bandwidth of the subband signal is also decreased by a factor of two for each stage so the sampling rate is proportional to its bandwidth.

The problem of bit allocation is to assign a bit rate in bits per sample to each of the subband signals such that the overall bit rate of the total system, R_T , satisfies an overall bit rate requirement $R_T \leq R_{req}$. The overall bit rate of an M stage subband system is given by

$$R_T = 2^{-M} B_L + \sum_{n=1}^M 2^{-n} B_n \quad (4)$$

where B_n is the bit rate in the n -th subband signal and B_L is the bit rate for the low-pass signal.

The basic idea behind the HVS based bit allocation is that the number of bits allocated to a subband signal should be based on the importance of the information contained in that signal. The importance of the information in a signal is based on the sensitivity of the HVS to that signal from (1) and (2).

The bit allocation for a subband signal w_n will now be explained. Let w_n be band-limited between f_1 and f_2 , see

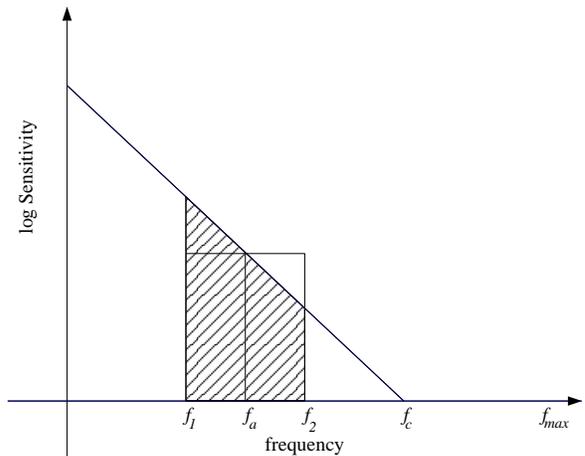


Fig. 5. Log sensitivity approximation

Figure 5. The maximum frequency in the original digital signal is given by

$$f_{max} = \frac{N}{4 \arctan(\frac{1}{2\alpha})} \text{ cycles/degree} \quad (5)$$

where N is the number of samples and α is the ratio of the distance of the viewer from the screen, l , to the screen width, w ; $\alpha = l/w$ [5]. The number of bits allocated to this signal is proportional to the area under the log sensitivity function between f_1 and f_2 .

$$Area = \frac{\log S(f_1) + \log S(f_2)}{2} (f_2 - f_1). \quad (6)$$

Note that the bandwidth of w_n is given by $(f_2 - f_1) = 2^{-n} f_{max}$ and let $f_a = (f_1 + f_2)/2$. From the linearity of the log sensitivity,

$$Area = \log S(f_a) 2^{-n} f_{max}. \quad (7)$$

The number of samples in w_n is proportional to $2^{-n} f_{max}$ so the bit rate B_n for this subband is described by

$$B_n = \kappa \log S(f_a) \text{ bits/sample}. \quad (8)$$

From (4),

$$R_T = \kappa (2^{-M} \log S(f_{aL}) + \sum_{n=1}^M 2^{-n} \log S(f_{an})). \quad (9)$$

The constant κ is fixed to a convenient value depending on the base of the log. Given the overall bit rate requirement R_{req} , the smallest value of f_{max} is determined which will satisfy $R_T \leq R_{req}$ with R_T given by (9). Using this value of f_{max} , the bit rate for each subband is determined by (8) and the overall bit rate will satisfy the given system requirement. Note that the value of f_a depends on f_{max} and increasing f_{max} will decrease the value of $\log S(f_a)$. Increasing the value of f_{max} is equivalent to increasing the ratio of the viewing distance to the image width, α , or

increasing the viewing distance, l , for a particular image. Note that the number of bits allocated to a subband cannot be negative. For this reason, $\log S(f_a)$ is taken to be zero for values of f_a which are above f_c , the sensitivity threshold frequency or cutoff frequency.

Bit allocation for a 2-D system is carried out in a similar manner. The areas in the 1-D case become volumes and the bandwidths are 2-D. The number of samples is proportional to the area of the bandwidth and the bits allocated to the signal is proportional to the volume under the log sensitivity curve for the bandwidth of the subband signal. An equivalent average frequency f_a is determined and either (1) or (2) is used depending on the orientation of f_a . This bit allocation method is used in the coding of wavelet coefficients in section IV.B.

Table 1 shows the target bit allocation for the separable wavelet decomposition and the actual bit rate consumed. Note that the highest frequency bands are dropped altogether.

Unlike the bit allocation method of [7], this bit allocation method does not require knowledge of the rate-distortion curves for the particular vector quantizer being used. The method used in [6] uses bit allocation dependent on the statistics of the image being compressed and a fixed vector quantizer while the proposed bit allocation method is independent of the input signal and the proposed vector quantizer is adaptive.

B. Vector Quantization of Wavelet Coefficients

Each of the subband images in the wavelet decomposition is quantized separately. The image is divided into $n_1 \times n_2$ blocks which are reordered into a sequence of vectors. For each image vector, a codebook is searched and the index of the codevector closest to the image vector is transmitted. Distance is measured as the squared Euclidean distance. The receiver uses the received index to lookup the codevector in an identical codebook.

There are two main stages in the design of a codebook. An initial codebook is obtained in the first stage and this initial codebook is improved iteratively to a locally optimal codebook in the second stage. Several methods for generating an initial codebook are described in [36].

In Antonini and others [6], the codebook is designed in advance using an ensemble of "typical" images called a training set. This approach works quite well if the images being coded are similar to the images in the training set. However, if the image being coded is very different from the images in the training set, the results may be quite poor. The centroid splitting method is used with this training set to generate the initial codebook. The centroid splitting method works well if there are several natural clusters but a typical wavelet coefficient image will have a single dominating cluster near zero. This will lead to several codevectors near zero and a reduced mean square error but increased distortion in the more visually significant vectors with components farther from zero.

In the proposed image coding scheme, the codebook is designed specifically for the image being coded and must

be transmitted along with the indices. Transmitting the codebook can be expensive but it allows for adapting the quantization to the needs of the particular image being coded and avoids the problems of training set adequacy. The pruning algorithm is then used to generate the initial codebook. Like the pairwise nearest neighbor algorithm, the set of codevectors is initially the entire set of training vectors. As a minor modification, the zero vector is added to the codebook since it is the most probable vector. A codevector is selected from the codebook and all vectors which are within a distance less than a certain pruning threshold are deleted from the codebook. The distance metric used is the squared Euclidean distance. Another codevector is selected and the process is repeated. After all codevectors have been selected, there will be no two codevectors which are closer together than the pruning threshold. While the number of codevectors is still larger than the desired codebook size, the pruning threshold is increased and the process is repeated.

After the initial codebook has been designed, it is improved through an iterative process called the generalized Lloyd algorithm (GLA), also called the Linde-Buzo-Gray (LBG) method [36]. This process is known to converge to a locally optimum codebook for any initial codebook. Since this process may converge to different locally optimal codebooks for different initial codebooks, it is important to create a good initial codebook. To avoid excessive computation in the event of slow convergence, the GLA was limited to ten iterations in the implementation. Examination of the VQ of several subimages showed little improvement when the number of allowable iterations was increased but this may not be true of all possible subimages.

The treatment of the zero vector differs from the standard GLA. First, the zero vector is added to the codebook before pruning and is the first codevector selected in the pruning operation. This guarantees that the zero vector is not deleted from the codebook. When the codevectors are recomputed in calculation of the optimum codebook during the GLA, the zero vector is not recomputed. This insures that the zero vector remains as an element of the final codebook. The zero vector does not need to be transmitted with the rest of the codebook since the receiver can assume that the zero vector is present. The use of a zero vector rather than the optimum codevector for that partition cell adds some distortion to the reconstructed image but this distortion will be among the wavelets coefficients with the smallest absolute values. After the inverse wavelet transform, these smallest coefficients correspond to scaled wavelets with small contrast. From Section II, these small contrast wavelets have the least visual impact on image quality.

The codebook size for a particular subband image effects the quality of the reconstruction in that subband. A larger codebook will generally produce a higher quality reproduction. A larger codebook will also require more bits to be transmitted. In the implementation used here, the codebook components are quantized before transmission using an 8 bit uniform quantizer. The total number of bits re-

Table 1. Bit rates by subband for separable decomposition (bpp).

level	orientation	Heavy Compression		Moderate Compression	
		target	actual	target	actual
-3	low-pass	5.899941	6.262672	6.882840	7.275391
-3	horizontal	1.699822	1.673828	4.648521	4.740234
-3	vertical	1.699822	1.660156	4.648521	4.224609
-3	diagonal	0.0	0.0	2.075366	2.070312
-2	horizontal	0.0	0.0	1.297042	1.250488
-2	vertical	0.0	0.0	1.297042	1.186523
-2	diagonal	0.0	0.0	0.0	0.0
-1	horizontal	0.0	0.0	0.0	0.0
-1	vertical	0.0	0.0	0.0	0.0
-1	diagonal	0.0	0.0	0.0	0.0
	overhead		0.00464		0.00586
total bit rate		0.145	0.155	0.447	0.444

Note. Actual bit rates are from separable decomposition of Lenna image.

quired to transmit a codebook of size S is given by

$$B_{codebook} = 8n_1n_2(S - 1) \quad (10)$$

where the block size is $n_1 \times n_2$ so each vector has n_1n_2 components. Note that the zero vector does not need to be transmitted. The 8 bit uniform quantizer is sufficient to reduce codebook distortion to an acceptable level but is certainly not optimal given the distribution of the wavelet coefficients. Improved codebook quantization is an area for future research. Recent work by Zegar and others [37] takes a closer look at transmitted codebook quantization and the tradeoffs involved.

The size of the codebook also effects the number of bits needed to transmit the vector indices. A codebook of size S will require $\lceil \log_2 S \rceil$ bits to represent an index into the codebook. The number of blocks in an image is N_1N_2/n_1n_2 where the image size is $N_1 \times N_2$, the block size is $n_1 \times n_2$, and it is assumed the image size is a multiple of the block size. The total number of bits used to transmit the indices is

$$B_{index} = \frac{N_1N_2}{n_1n_2} \lceil \log_2 S \rceil. \quad (11)$$

Examination of the VQ of several subimages revealed that the most frequently selected codevector is the zero vector. This was to be expected given the distribution of the wavelet coefficients. The subimages were taken from the separable wavelet decomposition of images from the ‘‘Miss America’’ and ‘‘Salesman’’ sequences. The index coding was modified to take advantage of the frequency of zero vectors. A single bit is transmitted to indicate whether the vector is zero or non-zero. If the vector is non-zero, this bit is followed by an index of size $\lceil \log_2 (S - 1) \rceil$ bits into the $S - 1$ non-zero vectors. It was found that the percentage of zero vectors ranged from 80 to 96 percent of the total number of vectors for the images investigated. Caution should be exercised in the use of these numbers since they were obtained from a relatively small number of images. The point is that well over half the vectors are the

zero vector, justifying the use of one bit to indicate a zero or non-zero vector. If 88 percent of the vectors are the zero vector, the number of bits needed to transmit the indices with this improved coding technique is

$$B_{index} = \frac{N_1N_2}{n_1n_2} ((.88)1 + (.12)(1 + \lceil \log_2 (S - 1) \rceil)) \quad (12)$$

$$= \frac{N_1N_2}{n_1n_2} (1 + (.12) \lceil \log_2 (S - 1) \rceil). \quad (13)$$

In the VQ of a particular subimage, it is desirable to have as large a codebook as possible given the allowable bit rate for that subimage. From (10) and (13) the total bit rate for the VQ of a subimage is given by

$$R_{total} = \frac{B_{codebook} + B_{index}}{N_1N_2} \text{ bpp} \quad (14)$$

$$= \frac{8n_1n_2(S - 1)}{N_1N_2} + \frac{(1 + (.12) \lceil \log_2 (S - 1) \rceil)}{n_1n_2} \text{ bpp}. \quad (15)$$

Note that N_1, N_2, n_1, n_2 are taken as given and are not optimized. The codebook size is chosen as the largest value of S for which R_{total} is less than the bit rate allocated to that subband. The largest codebook possible is used given the bit rate constraint. The single bit coding of the zero vectors reduces the bit rate required for transmission of the indices. Taking advantage of this entropy consideration allows more bits for the transmitted codebook which may then contain more codewords. The increased codebook size can reduce distortion in the reconstructed subimage. Also, (15) assumes that 88 percent of the vectors will be zero vectors. The actual bit rate attained will be larger or smaller depending on the actual percentage of zero vectors.

C. Effect of Increased Wavelet Regularity

The concept of regularity distinguishes a wavelet transform decomposition from other subband coding methods.

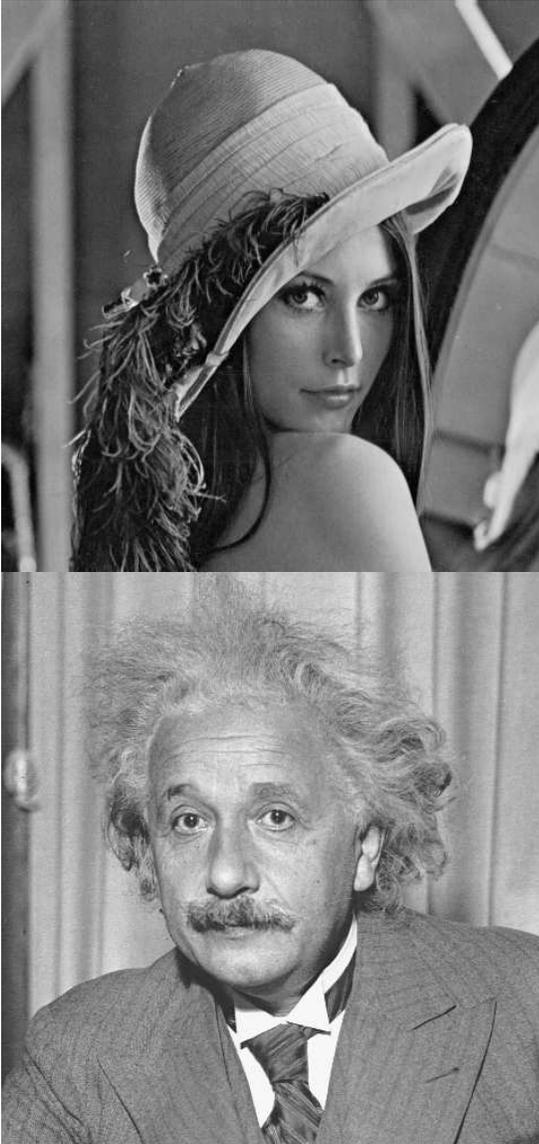


Fig. 6. Original images: (a) Lenna; (b) Einstein.

Regularity refers to the “smoothness” of the basis functions. The trigonometric functions from Fourier analysis are infinitely differentiable but the wavelet scaling functions have only a finite number of derivatives. The regularity of the Daubechies compactly supported scaling functions $\phi(t)$ can be increased by increasing the order of the associated low pass filter $h_0(n)$ [19]. From a signal processing point of view, the regularity of the scaling function is important because $\phi(t)$ is the impulse response of iteratively low-pass filtering and downsampling in the subband structure.

Using scalar quantization and Huffman coding, Rioul [38] takes a close look at the importance of regularity. Although a different quantization method is used, the results in [38] show the regularity of the filters is an important system characteristic. For a given filter order, it was found that more regular filters gave better performance with the Daubechies filters giving the best performance. Addition-

ally for Daubechies filters, the performance improved as the order of the filters was increased but increasing the filter order beyond 12 did not greatly improve the coding performance.

While the peak SNR used to measure performance in [38] may not be a good measure of image quality, it can be seen that the subjective image quality improves as the filter order and regularity increase. Figure 7 shows the reconstructed images from the HVS-based wavelet decomposition after heavy compression for varying filter order.

D. Comparison of the Three Wavelet Decompositions

The three wavelet decompositions presented in Section III are compared in this section. All of these decompositions give perfect reconstruction if the transform coefficients are not quantized. If the transform coefficients are compressed using the VQ method discussed in section IV.B, artifacts of the quantization can be found in the reconstructed image. These artifacts appear differently in each of the three wavelet decompositions.

The separable wavelet decomposition is the most common 2-D wavelet decomposition. As discussed in Section III, the rectangular subbands do not match the diamond shaped response of the HVS. When quantization is introduced, there is an additional problem with the separable wavelet decomposition. Since the filtering is applied in the horizontal and vertical directions, the quantization artifacts are also aligned horizontally and vertically, see Figures 8a, 8b, 9a and 9b. Increasing the filter order will reduce the impact of quantization but the artifacts are still oriented horizontally and vertically, the orientations to which the HVS is most sensitive. The original images are shown in Figure 6.

The non-separable wavelet decomposition addresses the problems found in the separable wavelet decomposition. The diamond shaped subbands are a better fit to the HVS diamond shaped frequency response. The sampling pattern alternates between rectangular and quincunx and the orientation of the filtering alternates as well. As a result, quantization artifacts are not oriented in only two directions. Since the quantization artifacts are spread along more directions, they are less noticeable than the separable wavelet quantization effects which are oriented in only two directions, see Figures 8c, 8d, 9c and 9d. The regularity of the non-separable wavelet decomposition unfortunately cannot be increased as the regularity of the separable wavelet decomposition can be increased.

The HVS-based wavelet decomposition also uses diamond shaped subbands to take advantage of the diamond shaped HVS frequency response. After the first level, all of the filtering is performed along the diagonals. The quantization artifacts are therefore found to be oriented along the diagonal directions where the HVS is least sensitive, see Figures 8e, 8f, 9e and 9f. The orientation of all the artifacts along two directions, even the diagonal directions, is less desirable than spreading the quantization artifacts in four directions as in the non-separable decomposition. For low filter orders, the non-separable reconstruction images



Fig. 7. HVS-based wavelet VQ reconstruction with varying filter order: (a) 4-th order filter, 0.167 bpp; (b) 6-th order filter, 0.167 bpp; (c) 10-th order filter, 0.166 bpp; (d) 12-th order filter, 0.169 bpp. Partial view, zoomed.

look better than the HVS-based decomposition reconstruction images. But the filter order can be increased easily for the HVS-based decomposition. Increasing the filter order reduces the aliasing and the basis functions used in reconstruction are smoother producing quantization artifacts which are less objectionable and improving the quality of the reconstructed image. For higher filter orders, the proposed HVS-based decomposition reconstruction images look better than the reconstruction images from either the non-separable or separable wavelet decompositions.

E. Comparison with JPEG Compression Method

The JPEG still picture compression method [39; 40] has become an accepted standard. For light compression both

JPEG and HVS-based wavelet VQ reproduce the image very well. For moderate compression, blocking artifacts can be seen in Figure 10a and aliasing artifacts are noticeable in Figure 8e. It is difficult to compare the artifacts from the two methods because they are so different. Signal to noise ratios and mean square error are not appropriate measures because they do not take into account the properties of the HVS. It is considered here that the blocking artifacts are more visually significant than the aliasing artifacts. The loss of high frequency information makes Figure 10b look out of focus at close viewing distances. However, this effect is decreased as the viewing distance is increased and is eliminated when the viewing distance is increased to the point where the missing frequencies fall

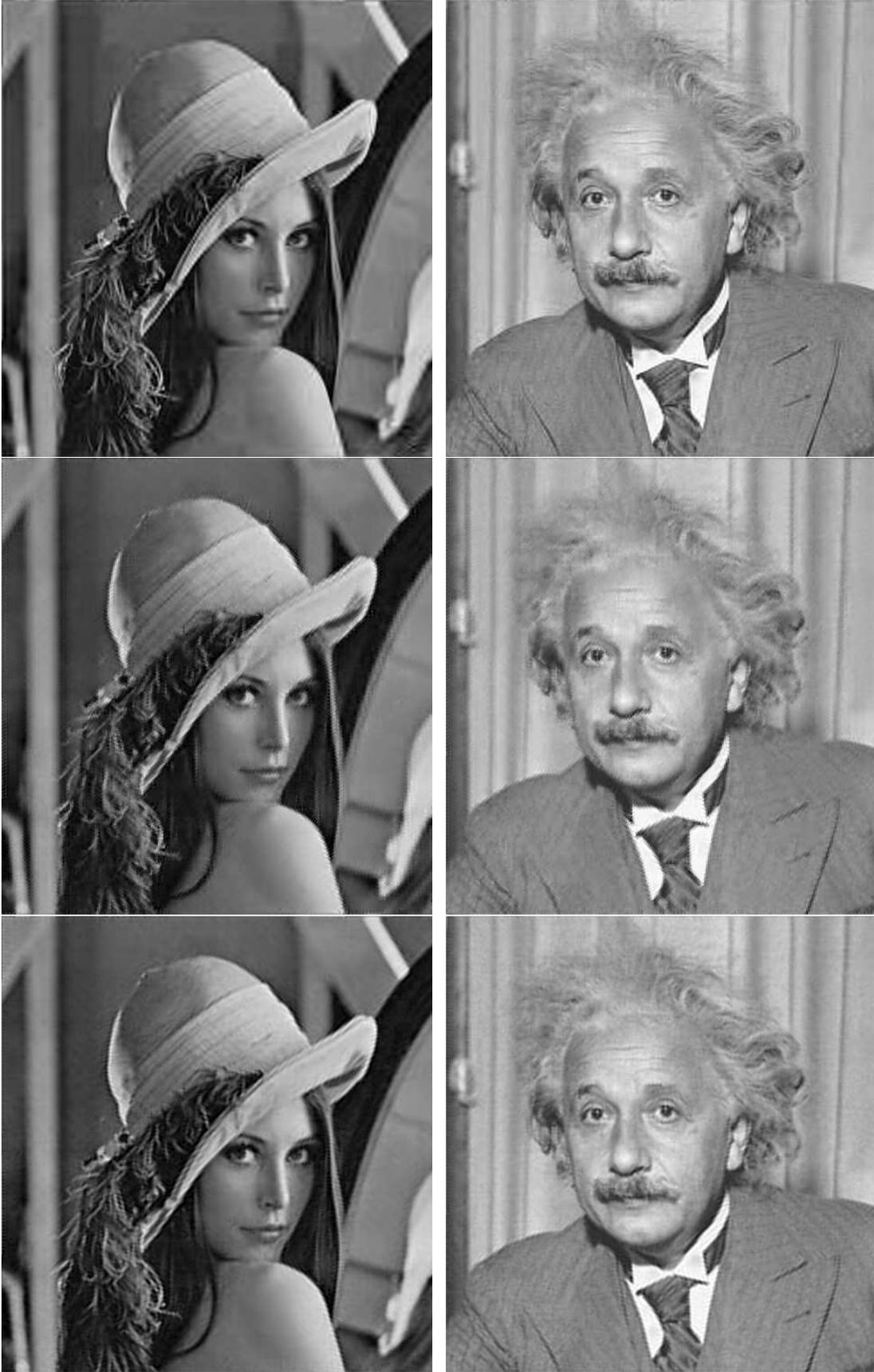


Fig. 8. Reconstruction with moderate compression. Separable wavelet: (a) 0.444 bpp; (b) 0.488 bpp. Non-separable wavelet: (c) 0.430 bpp; (d) 0.403 bpp. HVS-based wavelet: (e) 0.450 bpp; (f) 0.437 bpp.

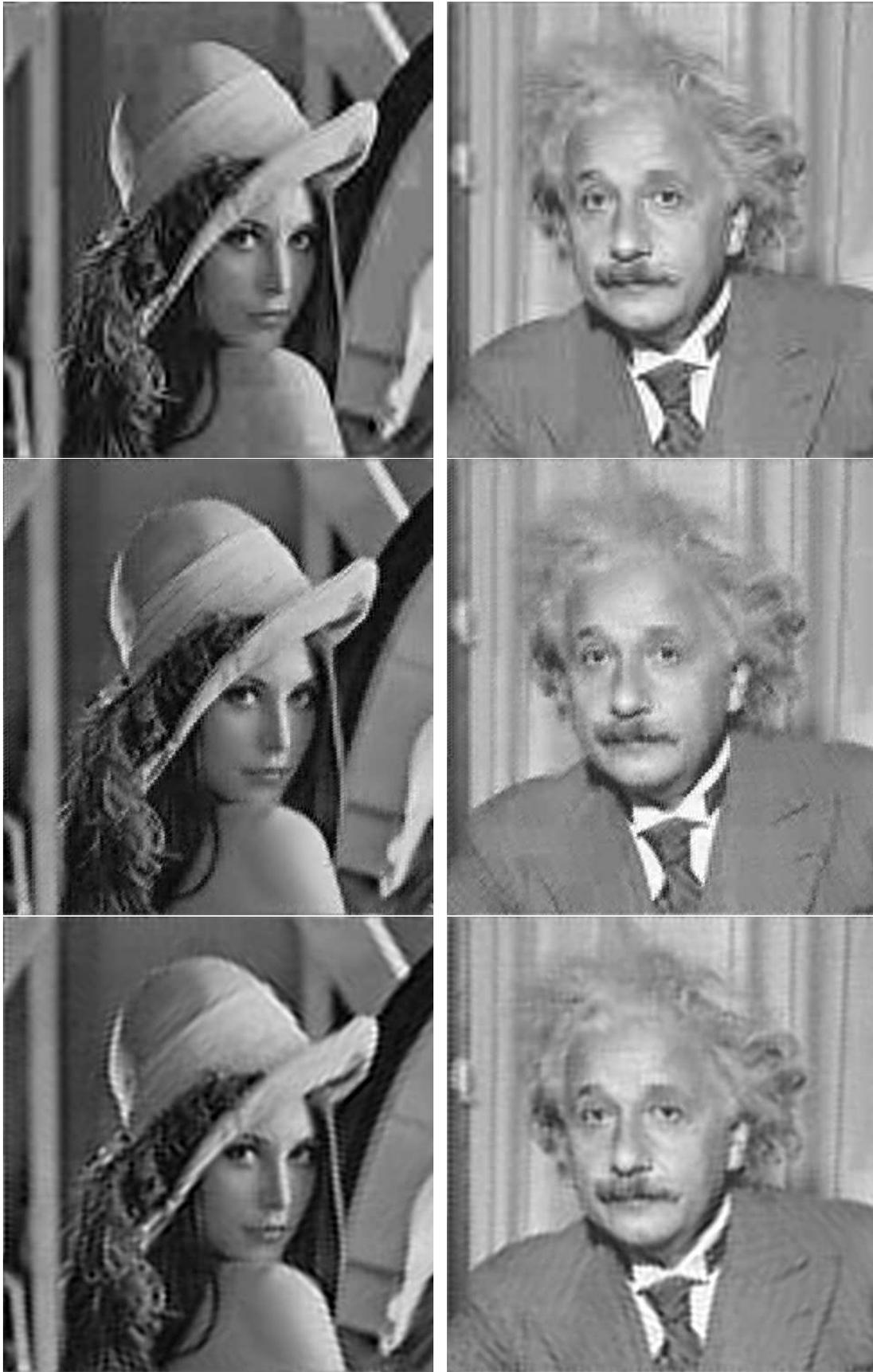


Fig. 9. Reconstruction with heavy compression. Separable wavelet (a) 0.155 bpp; (b) 0.155 bpp. Non-separable wavelet: (c) 0.205 bpp; (d) 0.173 bpp. HVS-based wavelet: (e) 0.169 bpp; (f) 0.165 bpp.



Fig. 10. Comparison of JPEG and HVS-based wavelet VQ reconstructed images: (a) JPEG, 0.458 bpp; (b) HVS-based wavelet VQ, 0.167 bpp.

outside the passband of the HVS. For heavy compression, JPEG loses much of the image information as the number of quantization levels is reduced and the blocking effect takes over, Figure 10b. Most of the high frequency information has been lost in the HVS-based wavelet VQ reconstructed image, shown in Figure 9e, but the low frequency information has been preserved. The image is degraded but not as severely as the JPEG image because the bit allocation method puts more degradation in the frequency regions where the HVS is less sensitive. The main advantage of JPEG, it is relatively cheap compared to the wavelet/VQ method, is due to the existence of a fast DCT algorithm and the use of scalar quantization rather than the more expensive VQ method.

For relatively high bit rates, (≈ 0.9 bpp), both JPEG and HVS-based wavelet VQ produce high quality images with the JPEG picture looking slightly better. As the bit rate is decreased, the JPEG image degrades more quickly

than the wavelet VQ image. The HVS-based wavelet VQ method performs slightly better than the JPEG method at moderate (≈ 0.45 bpp) bit rates and much better at low (≈ 0.15 bpp) bit rates.

V. CONCLUSION

An overview of important properties of the HVS was provided. Existing separable and non-separable wavelet decompositions were discussed and a new HVS-based wavelet decomposition was introduced. A technique for VQ of wavelet transform coefficients with bit allocation based on useful properties of the HVS was described. This quantization was used on the three wavelet decompositions under study and the proposed HVS-based wavelet decomposition had the best reconstruction of a compressed image. The proposed HVS-based wavelet decomposition was also found to be superior to JPEG at low bit rates and comparable to JPEG at higher rates.

REFERENCES

- [1] M. Kunt, A. Ikonopoulou, and M. Kocher, Second-generation image-coding techniques, *Proc. IEEE*, **73**, Apr. 1985, 549–574.
- [2] N. B. Nil, A visual model weighted cosine transform for image compression and quality assessment, *IEEE Trans. Comm.* **COM-33**, June 1985, 551–557.
- [3] R. J. Safranek and J. P. Johnston, A perceptually tuned sub-band image coder with image dependent quantization and post-quantization data compression, in *Proc. IEEE ICASSP 89*, **3**, pp. 1945–1948, Glasgow, Scotland, 1989.
- [4] B. Macq, Weighted optimum bit allocations to orthogonal transforms for picture coding, *IEEE Journal on Selected Areas in Communications*, **10**, June 1992, 875–883.
- [5] R. L. de Queiroz and K. R. Rao, Human visual sensitivity-weighted progressive image transmission using the lapped orthogonal transform, *Journal of Electronic Imaging*, **1**, July 1992, 328–338.
- [6] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, Image coding using wavelet transform, *IEEE Trans. on Image Processing*, **1**, Apr. 1992, 205–220.
- [7] R. E. Van Dyck and S. A. Rajala, Subband/VQ coding in perceptually uniform color spaces, in *Proc. IEEE ICASSP-92*, **3**, pp. 237–240, San Francisco, CA, 1992.
- [8] S. G. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. Machine Intell.* **11**, July 1989, 674–693.
- [9] O. Rioul and M. Vetterli, Wavelets and signal processing, *IEEE Signal Process. Magazine*, **8**, 1991, 14–38.
- [10] M. Vetterli, Multi-dimensional sub-band coding: Some theory and algorithms, *Signal Process.* **6**, 1984, 97–112, Erratum *Signal Process.* **11**, 1986, 191.
- [11] M. Vetterli, Filter banks allowing perfect reconstruction, *Signal Process.* **10**, 1986, 219–244.

- [12] M. J. T. Smith and T. P. Barnwell, III, A procedure for designing exact reconstruction filter banks for tree-structured subband coders, in *Proc. IEEE ICASSP 84*, **2**, pp. 27.1.1–27.1.4, San Diego, CA, 1984.
- [13] M. J. T. Smith and T. P. Barnwell, III, A unifying framework for analysis/synthesis systems based on maximally decimated filter banks, in *Proc. IEEE ICASSP 85*, **2**, pp. 521–524, Tampa, FL, 1985.
- [14] P. P. Vaidyanathan, Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial, *Proc. IEEE*, **78** Jan. 1990 56–93, Correction *Proc. IEEE*, **79**, Feb. 1991, 242.
- [15] J. W. Woods and S. D. O’Neil, Subband coding of images, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-34**, Oct. 1987, 1278–1288.
- [16] J. W. Woods, Ed., *Subband Image Coding*, Kluwer Academic Publishers, Boston, 1991.
- [17] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, *IEEE Trans. Inform. Theory*, **36**, Sept. 1990, 961–1005.
- [18] S. G. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$, *Trans. Amer. Math. Soc.* **315** 1989, 69–87.
- [19] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Comm. on Pure Appl. Math.* **XLI** 1988, 909–996.
- [20] J. Kovačević and M. Vetterli, Nonseparable multidimensional perfect reconstruction filter banks and wavelet bases for \mathcal{R}^n , *IEEE Trans. Inform. Theory*, **38**, Mar. 1992, 533–555.
- [21] R. A. DeVore, B. Jawerth, and B. J. Lucier, Image compression through wavelet transform coding, *IEEE Trans. Inform. Theory*, **38**, Mar. 1992, 719–746.
- [22] P. Delsarte, B. Macq, and D. T. M. Sloock, Signal adapted multiresolution transform for image coding, *IEEE Trans. Inform. Theory*, **38**, Mar. 1992, 897–904.
- [23] A. S. Lewis and G. Knowles, Image compression using the 2-D wavelet transform, *IEEE Trans. on Image Processing*, **1**, Apr. 1992, 244–250.
- [24] G. Strang, Wavelets and dilation equations: A brief introduction, *SIAM Review*, **31**, Dec. 1989, 614–627.
- [25] I. Daubechies, *Ten Lectures on Wavelets*, vol. 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.
- [26] C. K. Chui, *An Introduction to Wavelets*, vol. 1 of *Wavelet Analysis and Its Applications*, Academic Press, Boston, 1992.
- [27] Y. Meyer, *Ondelettes ét opérateurs, I: Ondelettes, II: Opérateurs de Calderón-Zygmund, III: Opérateurs multilinéaires*, Hermann, Paris, 1990.
- [28] F. W. Campbell, J. J. Kulikowski, and J. Levinson, The effect of orientation on the visual resolution of gratings, *Journal of Physiology*, **187**, 1966, 427–436.
- [29] F. W. Campbell and J. J. Kulikowski, Orientational selectivity of the human visual system, *Journal of Physiology*, **187**, 1966, 437–445.
- [30] F. W. Campbell and J. G. Robson, Application of fourier analysis to the visibility of gratings, *Journal of Physiology*, **197**, 1968, 551–566.
- [31] F. W. Campbell and L. Maffei, Electrophysiological evidence for the existence of orientational and size detectors in the human visual system, *Journal of Physiology*, **207**, 1970, 635–652.
- [32] R. Ansari and C.-L. Lau, Two-dimensional IIR filters for exact reconstruction in tree-structured sub-band decomposition, *Electronic Letters*, **23**, 1987, 633–634.
- [33] R. Ansari, H. P. Gaggioni, and D. J. Le Gall, HDTV coding using a nonrectangular subband decomposition, *SPIE Visual Communications and Image Processing ’88*, **1001**, 1988, 821–824.
- [34] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics Principles and Practice, Second Edition*, Addison-Wesley, 1990.
- [35] R. E. Van Dyck and S. A. Rajala, Subband/VQ coding of color images using separable diamond-shaped subbands, in *Proc. IEEE Visual Sig. Proc. Conf.* pp. 235–240, 1992.
- [36] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, 1992.
- [37] K. Zegar, A. Bist, and T. Linder, Universal source coding with codebook transmission, *IEEE Trans. Comm.* 1993, to appear.
- [38] O. Rioul, On the choice of “wavelet” filters for still image compression, in *Proc. IEEE ICASSP 93*, **5**, pp. 550–553, Minneapolis, MN, 1993.
- [39] G. K. Wallace, The JPEG still picture compression standard, *Comm. ACM*, **34**, Apr. 1991, 30–44.
- [40] G. K. Wallace, The JPEG picture compression standard, *IEEE Trans. on Consumer Electronics*, **38**, Feb. 1992, 18–34.