

Image compression using M-adic wavelets

C. A. SUBRAMANIAN* AND K. R. RAMAKRISHNAN

Department of Electrical Engineering, Indian Institute of Science, Bangalore 560040, India.
krr@ee.iisc.ernet.in

Abstract

In this paper, we propose a low bit rate image coding scheme using M-adic wavelet transform. M-adic wavelets arise from solutions to the dilation equation, $\phi(x) = \sum_k \phi(Mx - k)$. The M-adic wavelet transform of an image is computed using tree-structured perfect reconstruction filter banks. The transformed image is coded into a compressed bit stream using a human visual system (HVS) model and a wavelet image model. HVS helps in removing perceptual redundancies, while the wavelet image model provides a framework to exploit the redundancies across different scales of the wavelet-transformed image. Two wavelet image models, namely, zerotree model and web model are generalized to be applicable to M-adic wavelets. Using these models the paper describes algorithms for obtaining an embedded bit stream. Simulation results show that the proposed algorithms can cater to a wide range of applications.

Keywords: M-adic wavelet, image compression, perceptual quantization

1. Introduction

Compression of images has a number of applications: ranging from online product catalogues, browsing on the internet, image database, multimedia to virtual reality. In the recent past, there has been a tremendous development in the field of wavelet transforms. Wavelets are new families of basis functions that yield series expansions of the form $f(x) = \sum_j \sum_k \sum_{r=0}^{M-1} c_{j,k}^r \psi^r(M^j x - k)$ to functions in $L^2(R)$. Unlike the Fourier series, the basis functions are not of infinite duration but are compactly supported in time, with a reasonable frequency localization. Wavelets arise from solutions to the dilation equation, $\phi(x) = \sum_k c_k \phi(Mx - k)$. The case $M = 2$ yields dyadic wavelets. In this paper, we discuss some results for the general case $M \geq 2$ in literature and present image compression algorithms using M-adic wavelet image models.

In order to perform image compression, it is useful to have a theoretical image model that highlights those aspects of an image that benefit compression. These models also help in the development of other image-processing algorithms in the compressed domain. The wavelet image models essentially hypothesize the dependencies across scales of a wavelet-transformed image. The human visual system (HVS) in wavelet domain has also been considered. In this paper, the compression of images is carried out using these models and an entropy coder. The algorithms considered here are embedded, i.e. all encodings corresponding to lower bit rates

*Present address: Silicon Automation Systems Ltd, Bangalore 560 075, India.
subbu@sasi.com

are embedded in the same stream. This has a potential application in progressive transmission.

2. M-adic wavelets

Wavelets decompose the signals into channels that have same bandwidth in a logarithmic scale. Thus, high-frequency channels have larger bandwidth and lower frequency channels narrow bandwidth. These characteristics are well suited for the analysis of low-frequency signals mixed with sharp transitions (spikes). The disadvantage however is that if there are high-frequency signals with relatively narrow bandwidth, the decomposition is not well suited. In order to overcome this problem, M-adic orthonormal wavelet bases have been constructed by Zimmerman¹ as a direct generalization of the dyadic wavelets of Harr bases. M-adic wavelets help zoom in, onto narrow-band high-frequency channels, while simultaneously having logarithmic decomposition of frequency channels (Fig. 1).

Just as a two-channel unitary filter bank is central to the dyadic wavelet transform, a perfect reconstruction (PR) M-band filter bank is central to the M-adic wavelet transform. Consider the M-band filter bank shown in Fig. 2. Equations 1 and 2 mathematically represent the operations of an M-band filter bank.

$$d_i(n) = \sum_k x(k)h_i(Mn - k), \tag{1}$$

$$y(n) = \sum_{i=0}^{M-1} \sum_k d_i(k)g_i(n - Mk). \tag{2}$$

For perfect reconstruction, the system should be lossless. If the filters h_i and g_i constitute a bi-orthogonal set and are complete, then one can get perfect reconstruction. It is easy to see these facts if filters are orthogonal, i.e. the discrete basis generated by the filters is orthogonal.

2.1. M-band Harr system

The Harr system is the simplest of the wavelet representations. Harr functions are piecewise constant functions. Using the indicator function $\chi_{[a,b)}(x)$, the dilation equation takes the form,



FIG. 1. Time-frequency tiling for M-adic MRA, M = 3.

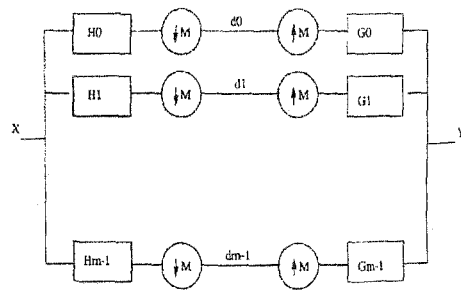


FIG. 2. M-Band filter bank.

$$\phi(x) = \chi_{[0,1)}(x) = \sum_{l=0}^{M-1} \chi_{[\frac{l}{M}, \frac{l+1}{M})}(x) = \sum_{l=0}^{M-1} \frac{1}{\sqrt{M}} \sqrt{M} \phi(Mx - l). \quad (3)$$

The Fourier transform $H_0(\omega)$ of the corresponding scaling filter is given by:

$$H_0(\omega) = \sum_{k=0}^{M-1} \frac{1}{\sqrt{M}} e^{-i2\pi\omega k/M} \cdot \frac{1}{\sqrt{M}} \frac{e^{-i\pi\omega} \sin \pi\omega}{e^{-i\pi\omega/M} \sin \pi\omega/M}. \quad (4)$$

In the dyadic wavelet transform the scaling function and the wavelet are closely related and once the scaling function is defined the wavelet function also gets defined. On the other hand, there is considerable choice in choosing the wavelet functions when $M > 2$. The wavelets defining the space W_j , i.e. $\bigcup_{r=1}^{M-1} W_j^r$ are not unique. Zimmerman¹ generates a set of wavelets corresponding to the scaling function by setting $H_k(\omega) = H_0(\omega - \frac{k}{M})$. The corresponding wavelet filters are nothing but DFT filter banks, i.e. translated versions of the low-pass filters. In this paper, we propose to use discrete cosine transform (DCT) generalization to generate the wavelet bases for some $M \geq 2$. This is mainly to make the filter coefficients real. The $M - 1$ Harr wavelets generated using the DCT matrix are given by the following equations:

$$\psi^r(x) = \sum_{l=0}^{M-1} \sqrt{\frac{2}{M}} \cos\left(\frac{\pi(2l+1)r}{M}\right) \sqrt{M} \phi(Mx - l) = \sum_{l=0}^{M-1} \sqrt{2} \cos\left(\frac{\pi(2l+1)r}{M}\right) \chi_{[\frac{l}{M}, \frac{l+1}{M})} \quad (5)$$

where r takes values $1 \dots M - 1$.

Generalizations of Daubechies wavelets can be found in Heller.² The multiresolution analysis in two dimensions (image) is carried out using separable filtering along rows and columns of the image as it is done for dyadic wavelets.³

3. Wavelet image models

In this section, the various wavelet image models we have used for image coding are described.

- Zerotree model⁴: The zerotree model states that if a wavelet coefficient in a particular scale is insignificant, then coefficients at the same location and orientation at finer scales are likely to be insignificant.^a
- Web model⁵: The web model states that if a wavelet coefficient is significant, then coefficients at finer scales at same location (all orientations) are likely to be significant.
- Human visual system model: This model describes the visual importance of various bands of the wavelet transform. Figure 3 shows the degree of quantization error that can be tolerated on various bands of a 4-band wavelet transformed image without any visual distortion. The graph has been obtained using experimental techniques. A set of

^asee tree representation in Fig. 4.

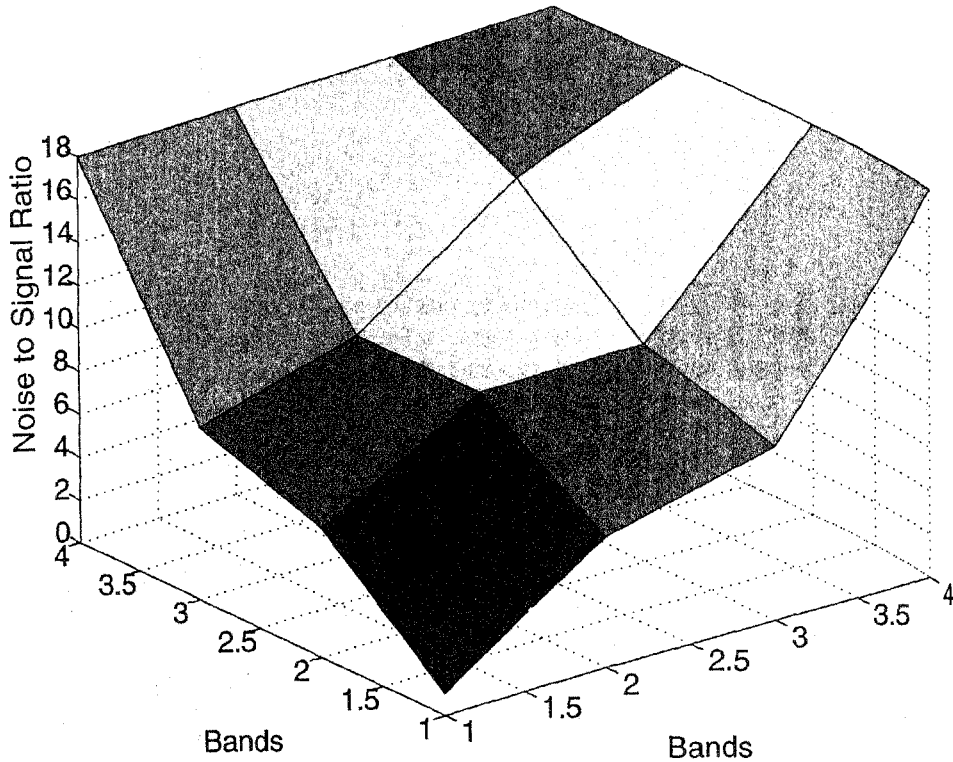


FIG. 3. HVS model for a 4-band wavelet transformed image.

checked patterns of different sizes and orientation were used in the experiment. Madic wavelet transform of these images was computed and the image was reconstructed using the quantized coefficients. The allowable quantization noise in each band with no perceptual distortion in the reconstruction yielded the perceptual model. The wavelet-transformed image is first quantized using the HVS model and then processed into a bit stream using the other models.

4. Zerotree model

4.1. Embedded coding

This section addresses a two-fold problem: (a) obtaining the best quality image for a given bit rate and (b) accomplishing this task in an embedded fashion i.e. in such a way that all encodings of the same image at lower bit rates are embedded in the beginning of the bit stream for the target bit rate.

4.2. Significance map encoding

Significance map is a binary decision, which tells whether a coefficient (in this case the inner product with the basis function) has a zero or a nonzero quantized value. The cost of specifying

ing this map in terms of bitrate is very high.⁴ As the target bit rate decreases, the probability that a coefficient is zero increases (one uses fewer and fewer bits to represent a coefficient) and it has been shown⁴ that no matter how optimal the transformation is in decorrelating the data, the cost of determining the positions of a few significant coefficients takes away a large fraction of the bit budget. Image models help in specifying these significance maps at low bit rates.

4.3. Compression of significance maps

The zerotree model has been used to improve the compression of significance maps of wavelet coefficients. A wavelet coefficient x is said to be insignificant with respect to a threshold T if $|x| < T$. The zerotree is based on the hypothesis that if a wavelet coefficient at a coarse scale is insignificant with respect to a given threshold T , then all wavelet coefficients of the same orientation in the same spatial location at finer scales are likely to be insignificant. Empirical evidence suggests that such an hypothesis is often true. Even though the image is passed through a decorrelating transform, the occurrence of insignificant coefficients is not independent events.^b More specifically, in a hierarchical subband system (such as Mallat tree decomposition), with the exception of the highest frequency subbands, every coefficient at a given scale can be related to a set of coefficients in the next finer scale.

The coefficient in the coarser scale is called the parent and all coefficients in the next finer scale of similar orientation are called children. For any given node at some level that node will have M^2 children except the leaves and the root nodes. Root nodes have $M^2 - 1$ children. For a given parent, the set of all coefficients at all finer scales of similar orientation corresponding to the same location is called descendants. Similarly, for a given child, the set of all coefficients at the same scale and same orientation at coarser scales is called ancestors. For an M-adic-tree structured decomposition the parent-child relationship is shown in Fig. 4.

4.4. Algorithm for image coding

The occurrence of insignificant coefficients is predicted using the zerotree model. The wavelet-transformed image is scanned in such a way (see Fig. 4) that no child is scanned before its parent. Given a threshold T and an element x one can determine whether or not x belongs to the zerotree, if all descendants of x are insignificant with respect to the same threshold T then x belong to a zerotree. If x itself is not a part of any other zerotree rooted above x (*no ancestor of x is in the zerotree*) the x is said to be the zerotree root. However, if x is insignificant but all its descendants are not insignificant, then x is predictably insignificant but does not form a zerotree.

The following algorithm which describes how the zerotree model can be used to get efficient representation of the wavelet-transformed image without sacrificing high bit rate is specifying a significance map.

W_0 to the set of all nodes

C_0 to ϕ

^bcoefficients across scales are related by location and orientation.

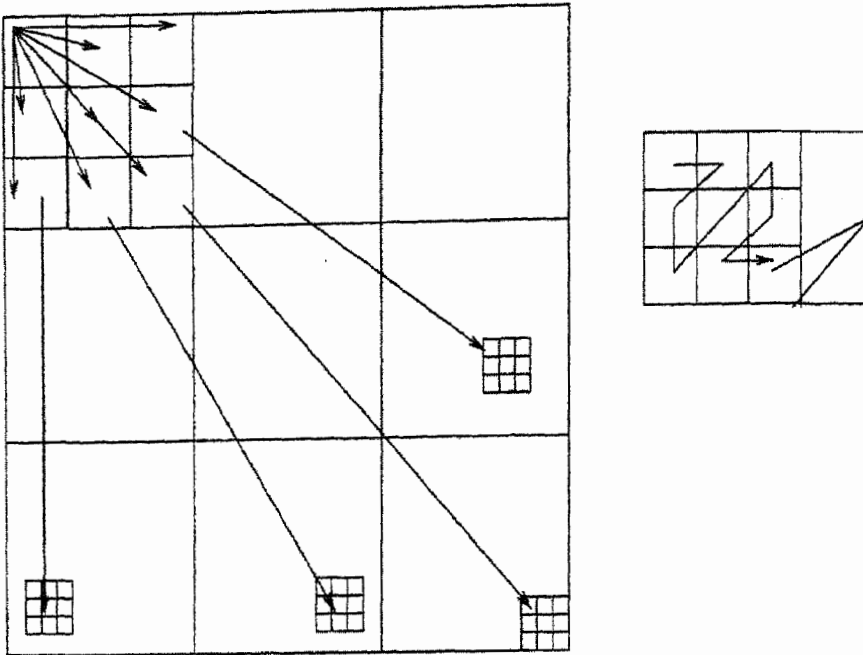


FIG. 4. M -adic wavelet image ($M = 3$) and scan order.

```

 $T_0 = \text{Max}/2$ 
 $i = 0$ 
while bit rate not exceeded.
   $\forall x \in W_i$ 
    if  $|x| < T_i$  and  $x \notin \text{descendants of } C_i$ 
      if  $\exists$  a zerotree rooted at  $x$ .
        Encode End of sub-tree symbol
         $C_i = C_i \cup x$ 
      else if  $x \in \text{descendants of } C_i$ 
        continue.
      else
        Encode magnitude and sign of  $x$ .
         $W_i = W_i \setminus x$ 
      endif
    end  $\forall$ 
   $i = i + 1$ 
   $C_i = \phi$ 
   $T_i = T_{i-1}/2$ 
endwhile

```

The algorithm essentially prunes the M -adic tree and considers only those branches which have significant coefficients. At every refinement step ($T_i = T_{i-1}/2$), only those branches that

were pruned are considered. It is apparent that such a coding is not strictly embedded as some coefficients whose children have a larger magnitude also get encoded even though they themselves are insignificant.

The decoder knows that in a particular iteration the significant coefficient should have had a magnitude between T_i and T_{i+1} and hence can predict the value of the coefficient given its sign. In successive passes, the encoder substitutes the actual values of the coefficients with the prediction error.

5. A wavelet web

This section describes an image-coding algorithm using the web model. Kashyap and Moni⁵ report that the magnitude of the wavelet transform coefficients, when sorted, lie almost exactly on the curve $1/x^\alpha$. Experiments indicate that the hypothesis is true for M-adic wavelets also. A data structure called 'web' (slightly different from Moni and Kashyap⁵) has been defined to order the wavelet coefficients. The coefficients can be chosen from the web in their order of importance. The web essentially predicts both the position and magnitude of the wavelet coefficients.

5.1. Tree of 1D M-adic wavelets

The wavelet basis functions can be represented as $\psi_{j,k}^r = \psi^r(M^j x - k)$ where r varies from 1 to $M-1$, j and k are integers. For each wavelet coefficient(basis function) in the tree, the children correspond to its M dilates given by ψ_{j+1,k_1}^r to ψ_{j+1,k_M}^r .

5.2. Web of two-dimensional wavelets

The web is a two-dimensional data structure. A two-dimensional node $d = (d_1, d_2)$ is an ordered pair of two one-dimensional nodes d_1 and d_2 . As mentioned earlier, the one-dimensional wavelets are indexed on a tree and each node has M children (root level nodes will have $M-1$ children). The children of a two-dimensional node are defined as follows: (x, y) is a child of (d_1, d_2) if x is a child of d_1 and y is a child of d_2 , or if $x = d_1$ and y is a child of d_2 or if $y = d_2$ and x is a child of d_1 . This amounts to saying that the children of (d_1, d_2) are the cross product of the two sets of d_1 , children of d_1 and d_2 , and children of d_2 excluding the node (d_1, d_2) .

The rationale behind defining the children in this manner is the following: To go from an approximation in V_i to V_{i+1} at a particular point in the image, one needs to precisely use all the children defined as above.

With such a representation a node may have up to a maximum of $M^2 + 2M$ children (The maximum number of elements in the two sets is $M+1$ and their cross product can be a maximum of $(M+1)^2$). Also, a node can have up to three parents. Also, note that the root node has no parents at all. We call a set of nodes B a 'web' if for every node $d \in B$ at least one parent of $d \in B$. The only exception to this is the root which has no parents at all. Thus, this structure is similar to a tree with the exception that there can be multiple parents, and there are common children too.

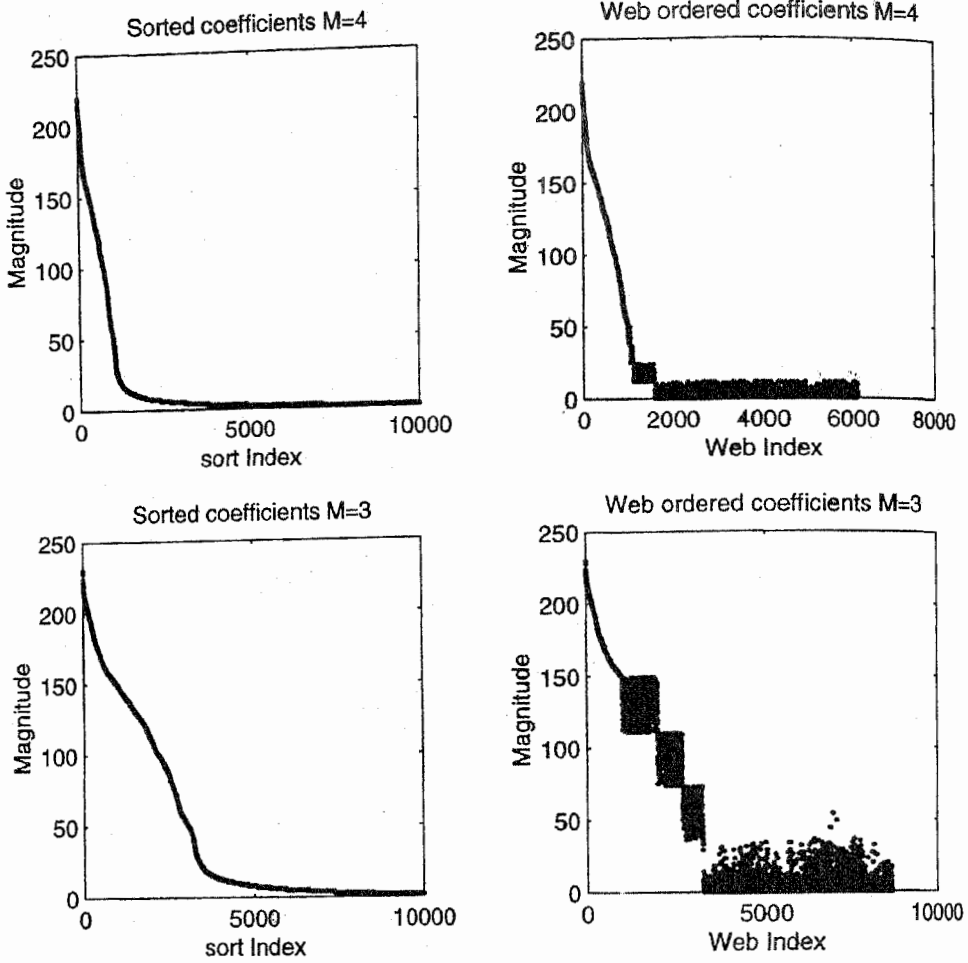


FIG. 5. Sorted and web-ordered data.

5.3. Web-based selection

The problem of encoding the significance map and the value of the significant coefficients is tackled using the web. The children of a web B are defined as follows: $Child(B)$ is the set of nodes that are children of nodes in B but are not themselves in B . The algorithm for choosing nodes adaptively using the web is given here. The algorithms start with a set of root nodes (the nodes corresponding to the inner product with the scaling function $\phi(x)$) and then add nodes to it. At each stage the resulting set of nodes is still a web.

- Set $B =$ root nodes.
- let $\hat{d} \in child(B)$, which has the largest magnitude.
- Set $B = B \cup \hat{d}$.

The last two steps are iterated till the desired number of nodes is chosen. For each node chosen we assign a web index. If the node is chosen at step k it gets a web index of k .

5.4. *Observation on web-based selection*

Figure 5 shows a plot of the coefficient magnitude vs web index (for $M = 4$ and $M = 3$). This shows that most of the wavelet coefficients lie on the curve of the form $y = 1/x^\alpha$. The web index can be equated to sorted index and the error in doing so is shown to be bounded.⁵ We can see that if we transmit the largest 500 or so coefficients with full precision, then we can get a good bound on the error in the estimate of the magnitude of the remaining coefficients. The value 500 simply corresponds to the knee of the web-ordered curve.

5.5. *Compression algorithm*

A compressed representation having about 20,000 nonzero coefficients is usually very close to the original.⁷ The algorithm enables control of both the number of wavelet coefficients encoded and the peak-signal-to-noise ratio (PSNR).

5.5.1. *Brief outline*

As suggested in the previous section, we split the web into two parts for the purpose of encoding. The first part consists of coefficients whose sort index is less than some number L_{initial} and the second part contains the remaining coefficients. In our experiments, we used L_{initial} to be between 500 and 1000 as this captures the main error region. Each of the two parts uses a different method for coding the magnitude and sign of the significant coefficients. In the first part, magnitudes and signs are encoded to full precision (nearest integer) and in the second part only an approximation to the magnitude is coded. Both the encoding schemes utilize a predefined scanning scheme which scans nodes only after its parents have been scanned as shown in Fig. 4.

5.5.2. *Coding with full precision*

Coding the positions

The L_{initial} nodes are obtained by sorting the image and hence the position of these coefficients and their magnitudes must be encoded. The position is encoded by encoding the scan index. This scan index itself is efficiently run-length encoded.

Coding the magnitudes

Since the positions are encoded in decreasing order of magnitude, the magnitudes are differentially encoded. The sign of these coefficients is separately run-length encoded.

5.5.3. *Approximation step*

Before we explain the algorithm, we define the following:

- **C**: the candidates. These are nodes which are yet to be encoded, but are children of some encoded node.

E : the encoded nodes. These are the nodes that have already been encoded. We start by initializing it to L_{initial} .

$W_i, i = 1 \dots Q$: Nodes in the i th quantization level, where Q is the number of quantization levels.

The set W_i corresponds to quantization regions along the web index. In the magnitude versus web-index plot, the magnitude axis is divided into Q quantization levels and the members on the web-index axis in i th level $\in W_i$.

Coding the positions

The algorithm for encoding the positions is given below:

```

Set  $E = \{\text{the } L_{\text{initial}} \text{ nodes}\}$ .
for  $i = 1: Q$  {
  for  $j = 1: N^2$  {
    Let  $d_j$  be the  $j$ th coefficient scanned
    if  $d_j \in (W_i \cap C)$  {
      (i.e. is a candidate to be encoded)
      if Coeff  $> 0$  encode + sign.
      else if Coeff  $< 0$  encode - sign
      update:  $E = E \cup d_j$ .
      update:  $C = C \cup (\text{child}(d_j) \setminus E)$ 
    }
    else if  $d_j = W_i^c$  {
      (is a candidate but Absent or has different N)
      encode Absent
    }
  }
}

```

Notice that this algorithm never considers the possibility of encoding a node not being a candidate (i.e. nodes $\notin C$). The reason is that these nodes can never appear in the web and hence it is unnecessary to scan them. Thus, a large number of scale space positions are eliminated from being scanned. It is clear that the scale space positions are scanned in an adaptive manner, adapting to the nature of the image.

Approximation of magnitudes

In each quantization bin, a straight line optimal in the least square sense is fitted and its parameters are encoded. The decoder can evaluate the value of the straight line for that index. The straight line parameters are transmitted before sending the scale space positions. This will ensure partial reconstruction to the decoder even if the bit rate gets exhausted while coding the scale space positions.

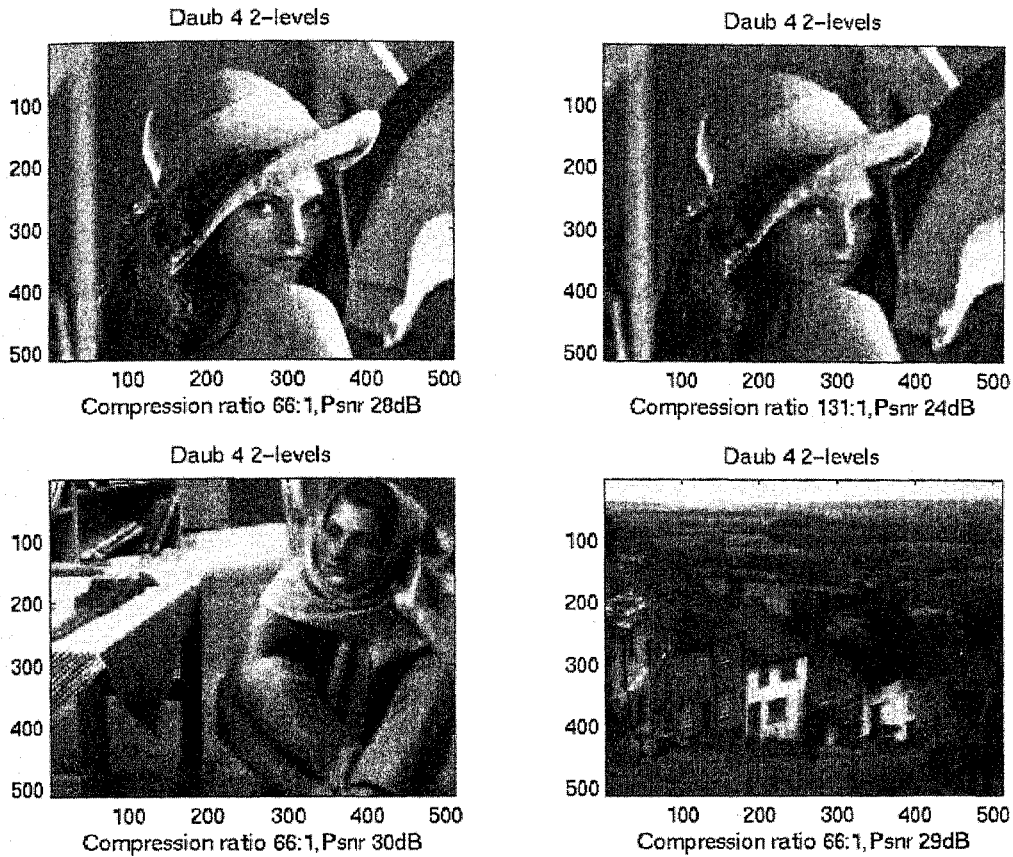


FIG. 6. Reconstructed images, compressed using zerotree model.

6. Conclusion

The above algorithms were tested on a wide range of grey scale images of 8-bit precision and the results are presented in the form of reconstructed images. The distortion metric PSNR, used to evaluate the images, was in the range 25db to 45db for the reconstructed images using the zerotree and the web model.

Needless to say, due to perceptual quantization, the PSNR can be low and the image quality can be good.

The zerotree algorithm achieves a high compression ratio with a linearly increasing distortion. Since the zerotree-based algorithm is a successive approximation algorithm, images at moderate compression ratios are almost lossless. Figure 6 shows reconstructed images of some standard images.

The web-based algorithm achieves very low bit rate coding, but is not asymptotically lossless. The least squares approximation of the wavelet coefficients cannot be refined in a

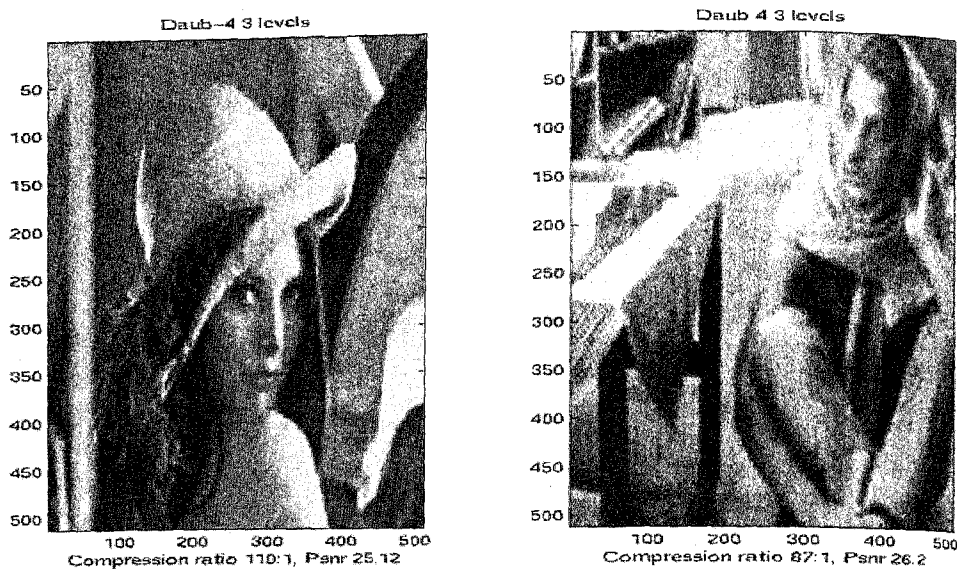


FIG. 7. Reconstructed images, compressed using web model.

systematic manner as in a zerotree. However, the web-based approach has applications where bit rate is very critical.^c In our simulations, the best possible image at very low bit rates was produced by the web model. Figure 7 shows reconstructed images compressed using the web model.

In conclusion, M-adic wavelets with suitable image model can cater to a wide range of applications. However, the mathematics for obtaining a bound on the error using these models, for some given image characteristics, remains open.

References

1. ZIMMERMAN, G. *Projective multiresolution analysis and generalized sampling*, Ph.D. Thesis, University of Maryland, College Park, 1994.
2. HELLER, P. Rank M wavelets with N vanishing moments. *SIAM J. Matrix Analysis Applic.*, 1995, **16**, 502–519.
3. MALLAT, S. G. J. A theory of multiresolution signal decomposition: The wavelet representations, *IEEE Trans.*, 1989, **PAMI-11**, 674–693.
4. SHAPIRO, J. M. Embedded image coding using zerotrees of wavelet coefficients, *IEEE Trans.*, 1993, **SP-40**, 3445–3462.
5. MONI, S. AND KASHYAP, R. L. Image compression using a web of wavelets, *IEEE Trans. Image Processing*, to appear.
6. SUBRAMANIAN, C. A. *Application of M-band wavelets to image and video compression*, M.E. Thesis, Indian Institute of Science, Bangalore, India, 1997.
7. DEVORE, B. J. R. AND LUCIER, B. J. Image compression using wavelet transform coding, *IEEE Trans.*, 1992, **IT-38**, 719–745.

^cstoring image in a smart-card.