

JAN 26 1995

OSTI

# Image Compression Using the W-Transform

William D. Reynolds, Jr.

Mathematics and Computer Science Division  
Argonne National Laboratory  
Argonne, Illinois 60439

## ABSTRACT

We present the W-transform for a multiresolution signal decomposition. One of the differences between the wavelet transform and W-transform is that the W-transform leads to a nonorthogonal signal decomposition. Another difference between the two is the manner in which the W-transform handles the endpoints (boundaries) of the signal. This approach does not restrict the length of the signal to be a power of two. Furthermore, it does not call for the extension of the signal thus, the W-transform is a convenient tool for image compression. We present the basic theory behind the W-transform and include experimental simulations to demonstrate its capabilities.

**Keywords:** W-transform, wavelet transform, multiresolution, image compression, filter banks, subbands

## 1 INTRODUCTION

The concept of multiresolution signal decomposition has received considerable attention in the research community over the past several years. This type of signal decomposition scheme has proved useful in a variety of applications, especially in signal compression and coding. One reason for its widespread use in signal compression is that each resulting frequency band or subband of the decomposition can be quantized and encoded independently from all the other subbands. The corresponding quantization error in each subband is then constrained to that particular band in the reconstruction of the signal.<sup>1</sup> Another reason is that the frequency bands can be matched to some of the properties of the human visual system. Thus, each band can be quantized based on its relative importance to the visual system. Finally, such decompositions can be implemented efficiently by using a pyramidal algorithm.

The traditional approach to multiresolution signal decomposition leads to an orthogonal wavelet representation of the signal.<sup>2</sup> This wavelet representation has been related to an infinitely iterated two-band filter bank, where the low-pass version of the signal at each stage is split into two bands. Combined with additional work by Daubechies<sup>3</sup> and Vetterli,<sup>4</sup> this decomposition technique has evolved into an efficient image compression scheme.<sup>5</sup> However, to ensure proper implementation of this scheme, certain assumptions and constraints must be enforced. First, the input image is assumed to be infinite in length, which requires some method of signal extension (e.g., symmetric or periodic). Second, the filter coefficients are assumed to form an orthonormal basis. Third, the length of the input image is assumed to be a power of two. Finally, an additional regularity or smoothness property is placed on the filter coefficients. This property ensures that in the limit, the filter coefficients lead to a continuous wavelet function. Although techniques have been developed to satisfy these assumptions,<sup>3</sup> such

The submitted manuscript has been authored by a contractor of the U. S. Government under contract No. W-31-109-ENG-38. Accordingly, the U. S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U. S. Government purposes.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED 85  
**MASTER**

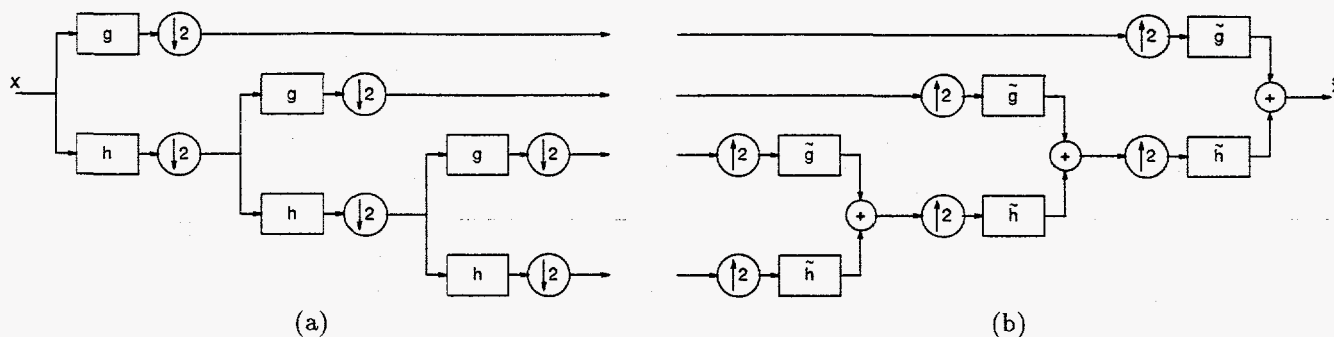


Figure 1: Irregular (octave-band) tree structure filter bank: (a) analysis section (b) synthesis section.

techniques usually impose heavy constraints on the filter design method.

In this paper we describe the  $W$ -transform to obtain a multiresolution signal decomposition.<sup>6</sup> This allows for a broader class of finite impulse response (FIR) filters possessing perfect reconstruction to be used in multiresolution analysis. Unlike the wavelet transform, no assumption is made on the orthogonality of the filter coefficients, and the  $W$ -transform in general leads to a nonorthogonal multiresolution analysis. Also, this approach does not restrict the length of the image to be a power of two, nor does it call for any signal extension. Thus, the  $W$ -transform becomes a convenient tool for image compression. The regularity property is maintained in the  $W$ -transform to allow for smooth approximations to the original image.

## 2 FILTER BANKS AND WAVELETS

This section is not intended to be an exposition on filter banks and wavelets. For our analysis we are concerned with their basic properties and structures. For a more detailed explanation, the works of other authors can be consulted.<sup>7,8</sup> In our analysis, the image is filtered assuming separability, so we will present the results for the 1D case only.

### 2.1 Filter Structure

Consider the diagram in Figure 1a of an irregular tree structure filter bank. In the diagram,  $h$  and  $g$  represent low- and high-pass analysis filters, respectively. Only the low-pass output signal is further split into two bands, which is shown here for a three-level decomposition. In the frequency domain this tree structure leads to unequal-sized frequency bands and is sometimes referred to as an octave-band subband tree structure,<sup>9</sup> where at each level the low-pass signal represents a blurred version of the original signal and the high-pass signal represents the detail (edge) information. We also note that the filter bank is critically sampled, since the decimation factor (2) is equal to the number of subbands at each level.

This octave-band tree structure is also used to perform an orthogonal wavelet decomposition.<sup>10</sup> Observe that as we travel down the tree, the subband bandwidth at each stage decreases, while the corresponding time function width increases. That is, for a large number of decomposition levels we increase the frequency resolution and decrease the time resolution, and vice versa. This fundamental time-frequency trade-off is what the wavelet transform offers in a signal decomposition scheme. The property that sets the wavelet transform aside from the subband coding techniques is the manner in which the filter coefficients are selected. Apart from this difference, the two techniques are essentially equivalent. In our analysis, the  $W$ -transform also uses this type of filter bank





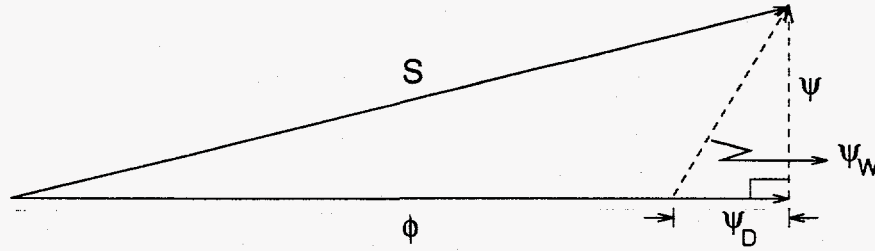


Figure 2: Vector diagram of the W-transform

In the case of the W-transform, we do not make any of the above assumptions. The W-transform treats the signals as finite and does not constrain the length to be a power of two. Note that, although we do not restrict the filter coefficients to form an orthogonal basis, we do not disregard the possibility. Hence, the W-transform leads to a nonorthogonal multiresolution signal decomposition.<sup>6</sup> Feauveau<sup>12</sup> also considers a nonorthogonal multiresolution decomposition. However, we avoid the rigorous mathematical analysis and provide a simple procedure for such a decomposition. This idea can be explained by considering the diagram in Figure 2. In the figure,  $S$  represents the signal vector, and the vector  $\phi$  (scaling function) represents an approximation to the signal vector  $S$ . The orthogonal component to  $\phi$  is the vector  $\psi$ , which represents the traditional wavelet decomposition. The nonorthogonal component to  $\phi$  is the vector  $\psi_W$ , which corresponds to the wavelet used in the W-transform. The component  $\psi_D$  represents the difference or the amount of information lost in using the nonorthogonal component. Although we have not obtained a quantitative measure for the loss in information, this diagram was meant to convey the concept of the W-transform.

Now let us reconsider the filter matrix in (1) and (2). For illustration purposes, let the length of the filters  $h$  and  $g$  be  $N = 4$  and let the signal,  $x$ , be an even finite length signal. Then we have that

$$\mathbf{H} = \begin{matrix} h_3 & \left[ \begin{array}{cccc} h_2 & h_1 & h_0 & \\ & h_3 & h_2 & h_1 & h_0 \\ & & h_3 & h_2 & h_1 & h_0 \\ & & & \ddots & & \\ & & & & h_3 & h_2 & h_1 & h_0 \\ & & & & & h_3 & h_2 & h_1 \end{array} \right] & , & \end{matrix} \quad (10)$$

where the coefficients  $h_0$  and  $h_3$  represent the coefficients that are excluded from the matrix because of the finite extension of the input. To include these coefficients in the matrix, we add them back to the nearest neighborhood that is retained. Thus, we obtain the following matrix

$$\mathbf{H} = \left[ \begin{array}{cccc} h_3 + h_2 & h_1 & h_0 & \\ & h_3 & h_2 & h_1 & h_0 \\ & & h_3 & h_2 & h_1 & h_0 \\ & & & \ddots & & \\ & & & & h_3 & h_2 & h_1 & h_0 \\ & & & & & h_3 & h_2 & h_1 + h_0 \end{array} \right] \quad (11)$$

The  $\mathbf{G}$  matrix is constructed in a similar fashion. Next, we interleave the rows of the  $\mathbf{H}$  and  $\mathbf{G}$  matrices to obtain



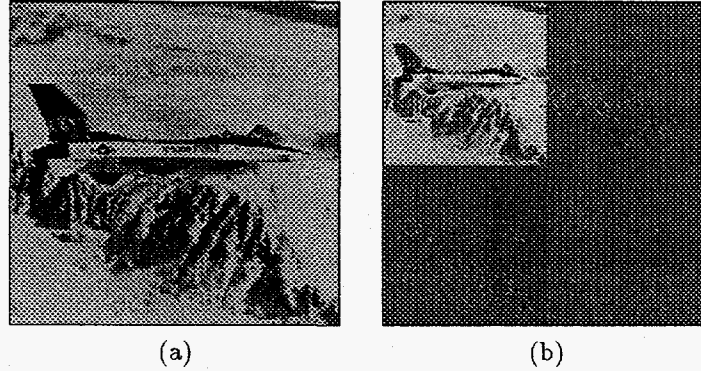


Figure 3: Plane: (a) original; (b) decomposed at one level using the W-transform, which results in 4 subbands. The upper left subband represents a low-pass version of the original, the upper-right represents horizontal frequency components, the lower-left represents vertical frequency components, and the lower-right represents diagonal frequency components.

## 4 EXPERIMENTS

The W-transform was tested on several different monochrome images using different filter lengths and levels of decompositions. The subbands of the transformed image then quantized to yield a desired bit rate. Naturally, this process introduces quantization errors in each subband. However, by properly allocating the number of quantization levels (or bits), these errors can be made almost imperceivable to the human viewer. The number of bits per subband were allocated based on the bit allocation scheme given by Akansu and Liu.<sup>14</sup> Once the bits have been allocated, the subbands are quantized by using differential pulse code modulation (DPCM) with uniform quantization.<sup>16</sup> By using DPCM, the overall bit rate can be achieved while allowing the quality of the reconstructed image to be maintained.

The design of the filter coefficients was based on the spectral factorization method.<sup>7,10</sup> This method involves designing a half-band low-pass prototype filter and then factoring the prototype filter into two spectral factors. That is, let  $T(z)$  represent a half-band low-pass filter and then factor  $T(z)$  as

$$T(z) = H_0(z)H_1(-z), \quad (15)$$

where  $H_0(z)$  and  $H_1(z)$  represent low- and high-pass filters, respectively. If  $T(z)$  has linear phase, the spectral factors  $H_0(z)$  and  $H_1(z)$  also have linear phase. Observe that because of the numerous ways to factor  $T(z)$ ,  $H_0(z)$  and  $H_1(z)$  must be chosen with care to produce desirable frequency responses.<sup>7</sup> We considered two methods for the design of the prototype  $T(z)$ . The first method used the Lagrange interpolation formula, and the second used the Parks-McClellan algorithm.<sup>10</sup> Both design methods yield linear phase filters having some degree of regularity. The spectral factors  $H_0(z)$  and  $H_1(z)$  from the Lagrange method resulted in even-length filters, while the Parks-McClellan method resulted in odd-length filters. Also,  $H_0(z)$  and  $H_1(z)$  satisfy the conditions for perfect reconstruction and in general do not satisfy the orthogonality condition.<sup>10</sup> In order to keep the spectral factors to a reasonable amount, the length of the filters used were  $N = 8, 10$  for the Lagrange method and  $N = 9, 11$  for the Parks-McClellan method.

## 5 RESULTS

The results from the experiments are shown in Figures 4, 5, and 6. The size of each of the original images are  $256 \times 256$  at 8 bits per pixel (bpp). The bit rate, number of decomposition levels, and the peak signal-to-noise

ratio (PSNR) are given along with the reconstructed image. The PSNR is determined by

$$PSNR (dB) = 10 \log_{10} \frac{(255)^2}{\frac{1}{256^2} \sum_{i=1}^{256} \sum_{j=1}^{256} [x(i, j) - \hat{x}(i, j)]^2}, \quad (16)$$

where  $x$  is the original image and  $\hat{x}$  is the reconstructed image. Figure 4 shows the reconstructed image of a two-level multiresolution decomposition. The filters used were determined from the Lagrange method and are of length  $N = 8, 10$ . Figure 5 shows the reconstructed image of a three-level multiresolution decomposition using the filters derived from the Parks-McClellan method. The filter lengths in this case are  $N = 9, 11$ . Figure 6 shows the reconstructed image for a three-level decomposition using the filters of length  $N = 8, N = 10$ , and  $N = 11$ . The bit rate for all three decompositions is 1.062 bpp. All simulations were implemented using MATLAB.<sup>17</sup>

## 6 DISCUSSIONS AND CONCLUSIONS

We have presented the W-transform for a multiresolution signal decomposition. The W-transform resulted in a nonorthogonal decomposition of the input as compared with the orthogonal decomposition of the wavelet transform. The W-transform was also shown to be equivalent to an irregular tree filter bank where the input image is assumed to be finite and the endpoints of the image are handled as described in Section (3.2). Even though only even-sized images were presented in the results, a method for handling odd-sized images was also described in Section (3.2). Furthermore, we demonstrated that FIR filters possessing PR, which in general do not form an orthonormal basis, can be used as filter coefficients in the W-transform. Two methods for determining the filter coefficients were also given.

The reconstructed images for bit rates in the range of 1.5 – 1.3 bpp are visually indistinguishable from the original images. For bit rates in the range of 1.3 – 0.75 bpp some visible distortions are noticeable. Mainly, there is the presence of granular noise, which is inherent in using DPCM. However, this type of visual distortion is much less annoying to the visual system than the blocky effects that result from conventional DCT compression schemes. For bit rates below 0.75 bpp, the edges of the reconstructed image start to become blurred. This is evident from Figure 5d. Overall, the reconstructed images using the W-transform are subjectively acceptable.

## 7 ACKNOWLEDGMENTS

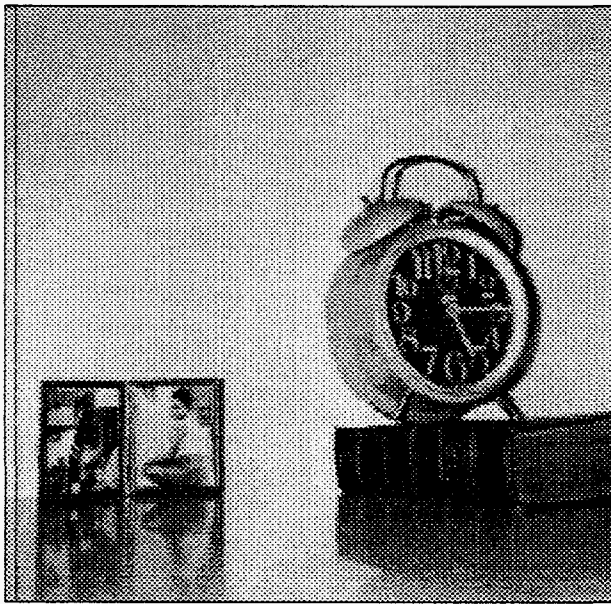
The author thanks Man Kam Kwong for his helpful discussions on wavelet theory, multiresolution analysis and for the diagram in Figure 2. This work was supported by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Computational and Technology Research, U.S. Department of Energy, under Contract W-31-109-Eng-38.

## 8 REFERENCES

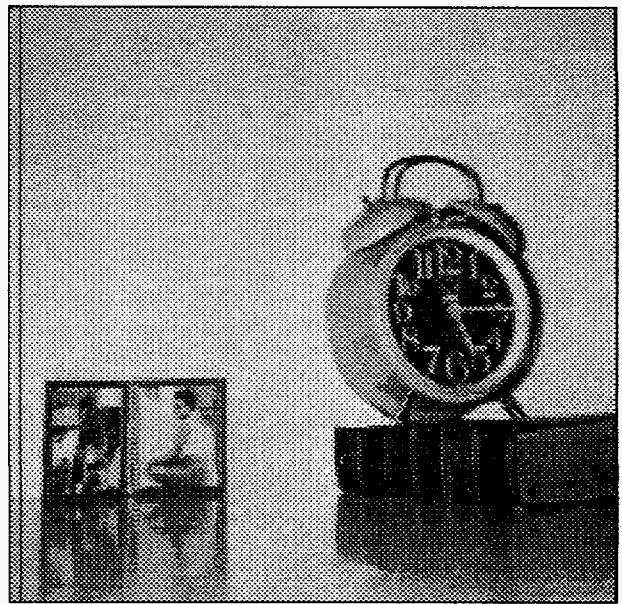
- [1] Hamid Gharavi and Ali Tabatabai, "Sub-Band Coding of Monochrome and Color Images", IEEE Trans. on Circuits and Systems, Vol. 35, No. 2, pp. 207–214, February 1988.
- [2] Stephane G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 11, No. 7, pp. 674–693, July 1989.



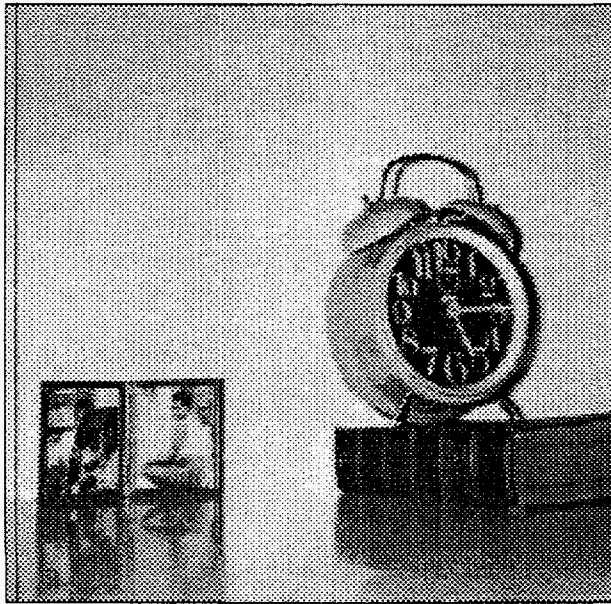
- [3] Ingrid Daubechies, "Orthonormal Bases of Wavelets with Finite Support-Connection with Discrete Filters", *Wavelets: Time-Frequency Methods and Phase Space, Proceedings of the International Conference*, pp. 38-66, Marseille, France, December 14-18, 1987.
- [4] Martin Vetterli and Cormac Herley, "Wavelets and Filter Banks: Theory and Design", *IEEE Trans. on Signal Processing*, Vol. 40, No. 9, pp. 2207-2232, September 1992.
- [5] Marc Antonini, Michel Barlaud, Pierre Mathieu, and Ingrid Daubechies, "Image Coding Using Wavelet Transform", *IEEE Trans. on Image Processing*, Vol. 1, No. 2, pp. 205-220, April 1992.
- [6] Man Kam Kwong and P. T. Peter Tang, "W-Matrices, Nonorthogonal Multiresolution Analysis, and Finite Signals of Arbitrary Length", *Mathematics and Computer Science Division Preprint MCS-P449-0794*, Argonne National Laboratory 1994.
- [7] Martin Vetterli, "A Theory of Multirate Filter Banks", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-35, No. 3, pp.356-372, March 1987.
- [8] Mark J. T. Smith and Thomas P. Barnwell III, "Exact Reconstruction Techniques for Tree-Structured Sub-band Coders", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-34, No. 3, pp. 434-441, June 1986.
- [9] Ali N. Akansu and Richard A. Haddad, *Multiresolution Signal Decomposition*, Academic Press, Inc., San Diego, 1992.
- [10] N. J. Fliege, *Multirate Digital Signal Processing*, John Wiley & Sons, New York, 1994.
- [11] H. J. Barnard, "Image and Video Coding Using a Wavelet Decomposition", Ph.D. thesis, Delft University, The Netherlands, May 1994.
- [12] Jean-Christophe Feauveau, "Nonorthogonal Multiresolution Analysis Using Wavelets", *Wavelets-A Tutorial in Theory and Applications*, ed. C. K. Chui, pp. 153-278, Academic Press, Inc., 1992.
- [13] Edward H. Adelson, Eero Simoncelli, and Rajesh Hingorani, "Orthogonal Pyramid Transforms for Image Coding", *Proc. SPIE: Visual Communications and Image Processing II*, Vol. 845, pp. 50-58, Society of Photo-Optics Instrumentation Engineers, 1987.
- [14] Ali N. Akansu and Yipeng Liu, "On-Signal Decomposition Techniques", *Optical Engineering*, Vol. 30, No. 7, pp. 912-919, July 1991.
- [15] Man Kam Kwong and Biquan Lin, "W-Transform Method for Feature-Oriented Multiresolution Image Retrieval", *Proc. SPIE: Wavelet Applications II*, Vol. 2491, pp. 1086-1095, Orlando, Florida, April 1995.
- [16] Anil K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, N.J., 1989.
- [17] Man Kam Kwong, "MATLAB Implementation of W-Matrix Multiresolution Analysis", *Mathematics and Computer Science Division Preprint MCS-P462-0894*, Argonne National Laboratory 1994.



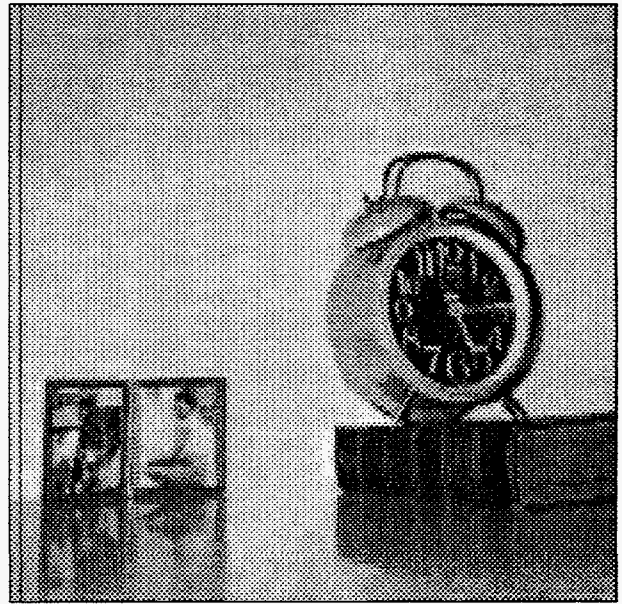
(a)



(b)

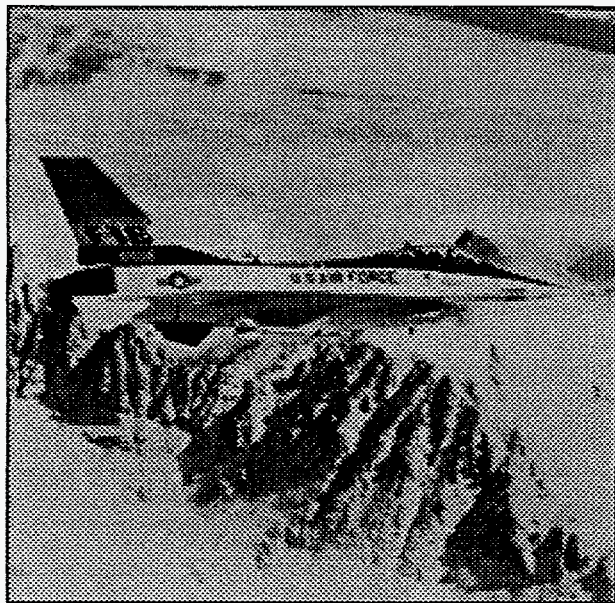


(c)

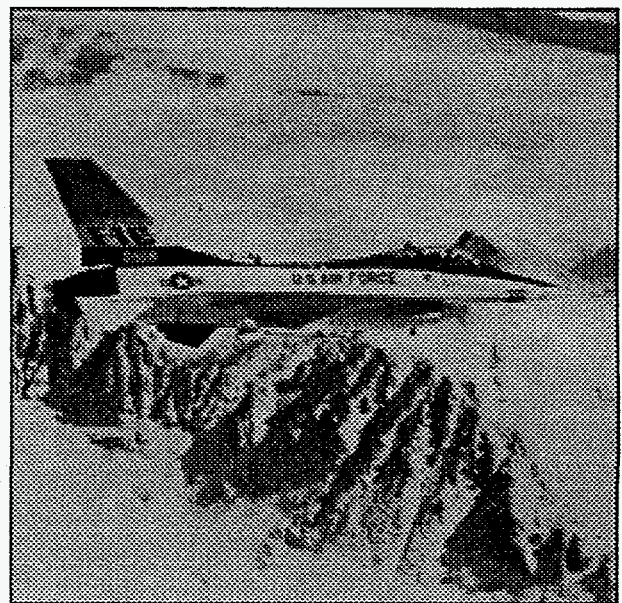


(d)

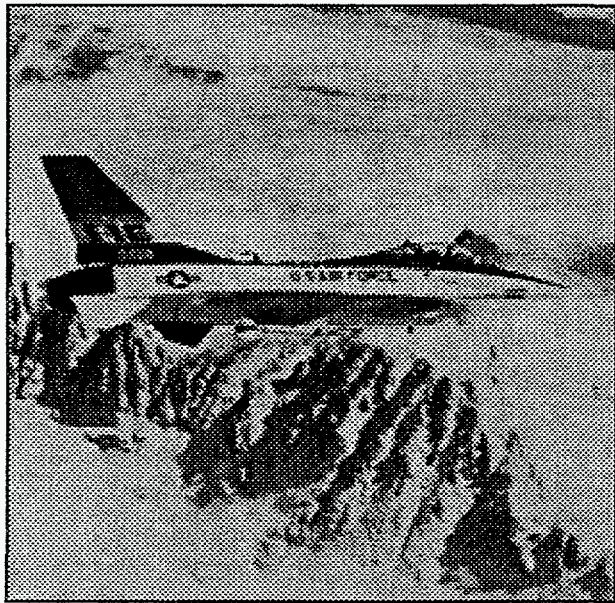
Figure 4: Clock: (a) original 8-bit image; (b) 1.5 bpp,  $PSNR = 38.66$ ; (c) 1.125 bpp,  $PSNR = 37.77$ ; and (d) 0.75 bpp,  $PSNR = 32.76$ . The filters used where designed using the Lagrange formula with (b), (c) of length  $N = 8$ , and (d) of length  $N = 10$ . The number of decomposition levels is  $L = 2$ .



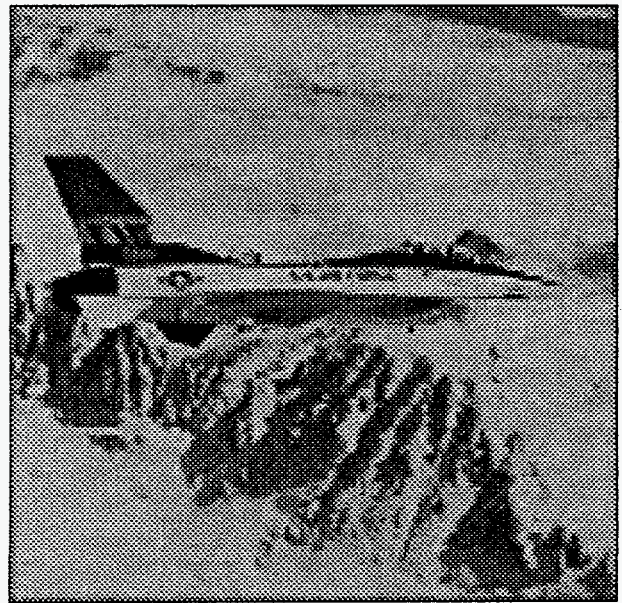
(a)



(b)

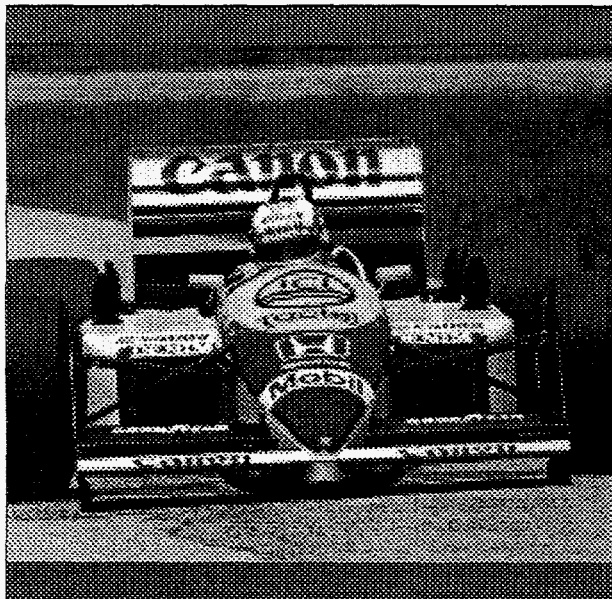


(c)

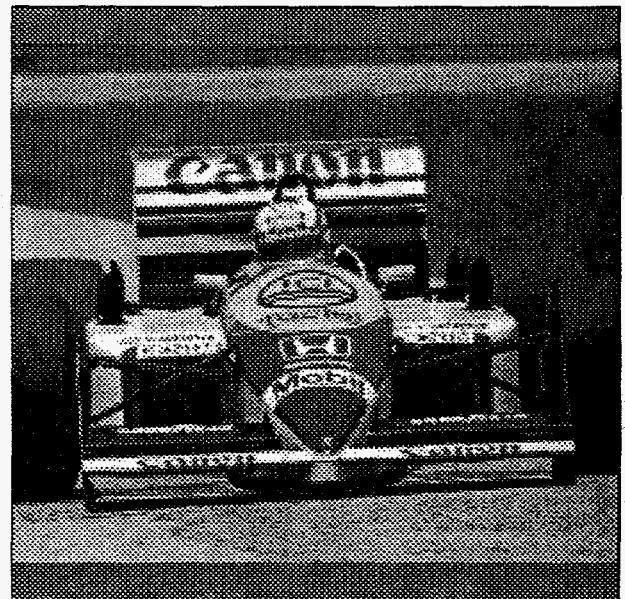


(d)

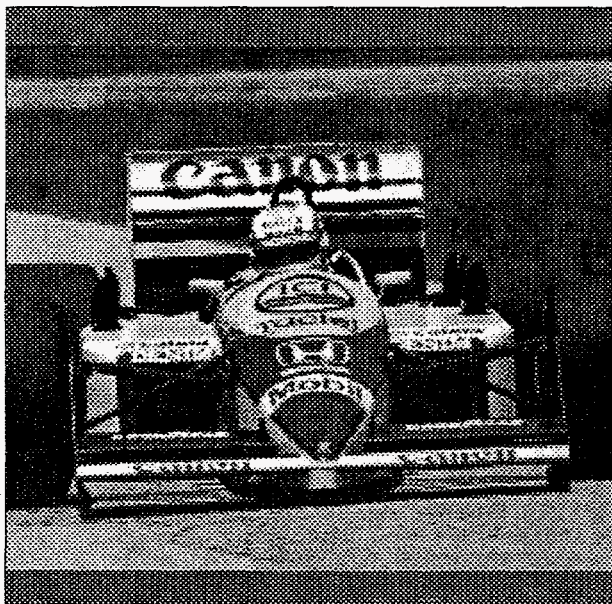
Figure 5: Plane: (a) original 8-bit image; (b) 1.312 bpp,  $PSNR = 36.68$ ; (c) 1.078 bpp,  $PSNR = 35.18$ ; and (d) 0.5625 bpp,  $PSNR = 26.73$ . The filters were designed using the Parks-McClellan method with (b), (c) of length  $N = 9$  and (d) of length  $N = 11$ . The number of decomposition levels is  $L = 3$ .



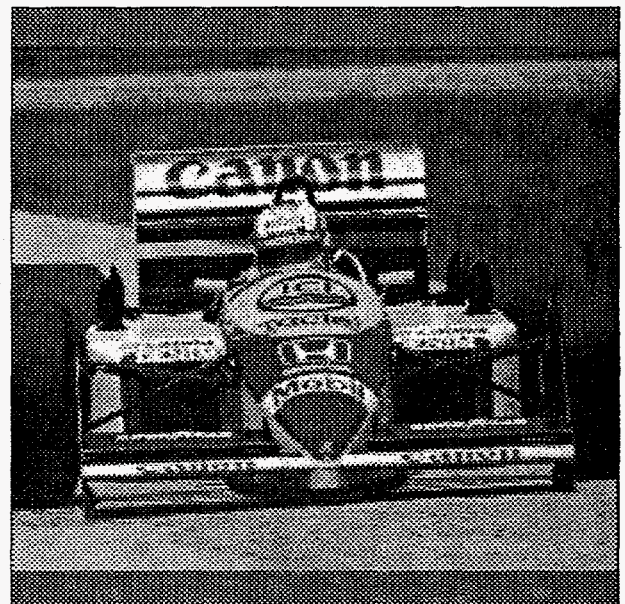
(a)



(b)



(c)



(d)

Figure 6: Car: (a) original 8-bit image; (b)  $N = 8$ ; (c)  $N = 10$ ; (d)  $N = 11$ . The bit rate, PSNR, and number of decomposition levels for (b), (c), and (d) are 1.062 bpp,  $PSNR = 33.03$ , and  $L = 3$ .