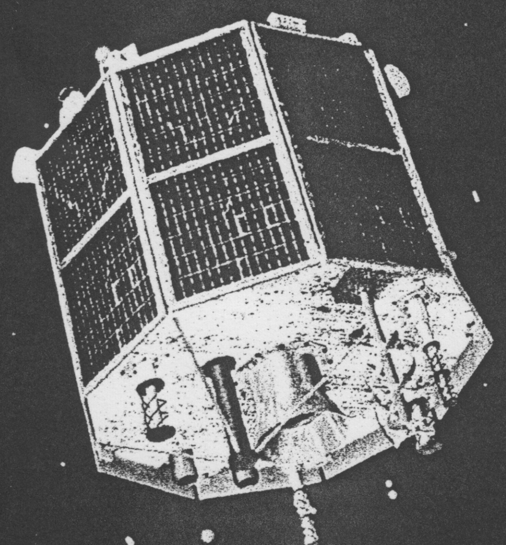
icics



Second International Conference on Information, Communications and Signal Processing ICICS'99

7 - 10 Dec 1999, Singapore

Organised by

Nanyang Technological University

School of Electrical & Electronic Engineering, NTU

Supported By IEEE Signal Processing Chapter, Singapore

(3) ICICS 99 (Singapore) 86

The Need for Adaptive Wavelet Filters in Image Coding – An Analysis

Subhasis Saha, Rao Vemuri

University of California Davis
Lawrence Livermore National Laboratory
7000 East Ave.
Livermore, CA 94550
USA
email : {saha, vemuri1}@llnl.gov

Abstract

Transmission lines are terminated with a load to match the line's characteristic impedance. Matching signal characteristics with filter parameters is the idea behind matched filters. Why not match images to the wavelet filters for best coding? Experiments conducted by compressing images through wavelet filters and integer wavelet transforms suggest that the filter performance indeed is image dependent. It is observed that no wavelet filter outperforms others uniformly while compressing sample images drawn from a large population. In fact, a detailed analysis of the results reveals that certain wavelets perform better on certain classes of images. This experimental observation leads to the hypothesis that for best results, in both lossy and lossless compression, the most "appropriate" wavelet filter should be chosen to match the image being coded.

1. Introduction

In recent years as the world is getting digitally connected and new classes of multimedia services are becoming available in the home and office, the need for storage with greater capacity, lower cost, and greater density is increasing. Despite significant advances in storage technology, uncompressed text, graphics, audio and video data require considerable storage capacity. Similarly for multimedia communications, data transfer of uncompressed images and video over digital networks require very high bandwidth. For example, an uncompressed still image of size 512 by 512 pixels with 24 bit of color requires about 6.29 Mbits of storage. An uncompressed full-motion video (512 by 512 framesize, 30 frames/sec) of 1 minute duration needs 11.32 Gbits of storage and a bandwidth of 188 Mbits/sec. Even if we assume that there is enough storage capacity available, it is impossible to transmit large number of images or play video (sequence of images) in real time due to insufficient data transfer rates as well as limited network bandwidths. So, at the present state of technology, the only solution is to compress multimedia data before its storage and transmission, and decompress it at the receiver for play back.

Modern image and video compression techniques offer a solution to this problem by reducing the storage requirements and transmission bandwidths. Compression ratios of 16 to 32 are quite common and ratios higher than this can be achieved but at the expense of the image quality. Image compression is achieved by exploiting the spatial and spectral redundancy or irrelevancy present in the image data. Video compression techniques exploit in addition to the spatial and spectral redundancy, the temporal redundancy as well. There are many compression techniques that are in part competitive and in part complementary. However, the most important compression technique for still images, the topic of this paper, is the transform coding based on the Discrete Cosine Transform (DCT). This is known as the JPEG standard.

In the past few years, wavelet transform has become a cutting-edge technology in signal processing in general and in image data compression in particular. A wide variety of wavelet-based image compression schemes have been developed. Some of these include the Laplacian Pyramid [1], Shapiro's Embedded Zero Tree (EZW) [2], and Said and Pearlman's SPIHT coding [3]. More complex techniques such as vector quantization, tree encoding and edge-based coding using wavelets have also been developed. Although none of these is part of the standard yet, in the upcoming JPEG-2000 standard, the top contenders are all wavelet-based compression algorithms.

Wavelet-based image coders are typically comprised of three major components: a wavelet filter bank decomposes the image into wavelet coefficients which are then quantized in a quantizer and finally an entropy encoder encodes these quantized coefficients into an output bit stream (compressed image), as shown in Fig. 1. Although one has the freedom to choose each of these components from a pool of candidates, it is often the choice of the wavelet filter bank that is crucial in determining the ultimate performance of the coder. If the performance of the wavelet filter bank is poor in the first place, the schemes for quantization and entropy encoding, however elegant they are, cannot generally provide adequate compensation to maintain significant picture quality. This observation becomes even more relevant in lossless coding because there is no quantization stage and the filter bank's role in any performance loss is more pronounced. In short, the wavelet filter used plays a significant role in both lossless and lossy image coding schemes. In this paper we demonstrate the need for adaptively selecting the most appropriate wavelet filter (for lossy compression) and integer wavelet transform (for lossless compression) in wavelet based image coding schemes while keeping other components unchanged. The methodology for adaptive wavelet filter selection is the topic of another paper [6].

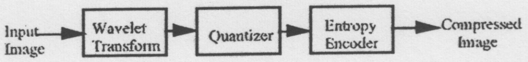


Fig. 1 A typical wavelet-based image coder

1.1 Wavelet Selection

Most of the well known wavelet based lossy image coding algorithms developed so far [2,3,4], use a specific filter bank chosen from a pool of filters designed and developed by researchers over more than a decade. Due to limitations of space, we don't want to go into the details of how these have been designed, but in section two we will look into some of the wavelet features that are important for image compression. Once a filter bank is chosen, the coefficients are hard coded into the algorithm. As a result, the same filters are used for coding and decoding all types of images whether it's a natural image, synthetic image, medical image, aerial image, scanned image, compound image or any other image for that matter. The same is also true in lossless compression context where of course an integer transform derived from the corresponding wavelet filters is used. So far no systematic study has been done to see if using different wavelets for different image classes improves the coding performance. In fact, most of the published results on experiments in lossy image coding used only a few well-known wavelet filters and a few popular test images most of which are natural images such as 'Lena'. As shown in this paper, this static approach of filter (or integer transform) selection may not always give the best quality of service (image quality or compression) from the viewpoint of a specific application.

2. Filter Bank Features

When deciding on a filter bank for image compression, there are many variables to take into account. First of all, there are two well-known wavelet filter families used in wavelet-based image coders viz. orthogonal and biorthogonal wavelets. Orthogonal wavelets are the family of wavelets that generate orthonormal bases of $L_2(R^n)$. Among them the most important ones to image coding are compactly supported orthogonal wavelets. In the discrete wavelet transform (DWT), compactly supported wavelets correspond to finite impulse response (FIR) filters and thus lead to efficient implementations. The popular Daubechies family of compactly supported wavelets is parameterized by an integer that is proportional to the length of the wavelet filter. For compactly supported wavelets, the length of a wavelet filter is related to the degree of smoothness and regularity of the wavelet, which in turn can affect the coding performance. The main attraction of biorthogonal wavelets on the other hand is linear phase of FIR filters (symmetric / anti-symmetric impulse response). One can choose, for example, to build filters with similar or dissimilar lengths for decomposition and reconstruction, or which are nearly orthogonal. Linear phase (symmetric) FIR filters are widely used since such filters can be easily cascaded in pyramidal filter structures without the need for phase compensation.

Subband decomposition comes in several varieties, as these can either have uniform-band splits, octave-band splits, or more generally, non-uniform-band splits. Furthermore, these can be perfect reconstruction (PR), such as many biorthogonal filter sets or conjugate quadrature filter banks, or near-perfect reconstruction like the quadrature mirror filter bank (QMF). Regularity and smoothness of the filters are the other important factors to be decided upon. In summary, the following are the key features that distinguish one wavelet filter from the other, and need be considered while making a choice for image compression.

- Orthogonality
- Linear phase (Symmetric)
- Length of the filters
- Smoothness (Number of zero moments)
- Regularity measure (Holder regularity)
- Order of the filters
- Energy compaction (Coding gain)
- Wavelet coefficient distribution statistics

The following wavelet filters have been used to compress various images in this experiment.

- Orthogonal filters
    - Haar - Orthonormal linear phase filter with two coefficients
    - Daubechies Orthogonal filters of order two, four, and eight with four, eight, and sixteen coefficients respectively.
    - Adelson's symmetric filters with nine coefficients
- Biorthogonal filters
    - Cohen, Daubechies, and Feauveau (CDF) biorthogonal filters e.g. CDF-9/7, CDF-9/11, CDF-13/3 filters
    - Villesenor biorthogonal filters e.g. Vill-18/10, Vill-13/11 and Vill-6/10 filters
    - Odegard biorthogonal filter with 9/7 coefficients
    - Brislawn 10/10 biorthogonal filter

3. Image Features

A spatial domain analysis of various images show that in general, images from different categories tend to have different characteristics. For example, it is a common observation that most of the natural images are continuous in tone compared to the synthetic images most of which are of discrete tone (dynamic range of the pixel bit depth is under utilized). Such images generally have some numerical structures that are not well represented by smooth basis functions. Many medical images like MRI or CT scan contain significant low-intensity (black) regions along image boundaries. Compound images with significant amount of text are a mixture of binary and continuous tone data. Even within a particular category, images vary in many ways with widely varying first and second order Markov statistics. Whereas some are relatively flat, others are very busy having more edges and contours in them. Some are darker and others have more sharpness. So, a spatial domain analysis of these images shows different characteristics like mean, median, standard deviation as shown in Fig. 4. To summarize, the following are the key features that may distinguish between various categories of images.

- Spatial Features -- Mean Median, Mode, Variance, Dispersion, Average Energy, Entropy etc.
- Transform Features -- Mean Deviation, Average Energy, Histogram and Cooccurrence Signatures
- Edge and Boundary Information, Image Activity Code
- Texture
- Higher Order Statistics

Test images in our experiment include natural images (Lena, Barbara, Baboon, and Airplane), synthetic images (Teradata, Ball), binary/compound images (Bengali, Cmpnd1, and Cmpnd2), medical images (mri, nervecell, us, and us1), aerial images (Aerial, Air1, and Air2) and miscellaneous images (camera, couple, seagull, and Finger).

4. Analysis of Experimental Results

We have experimented with a large number of wavelet filters, both orthogonal and biorthogonal with varying lengths, regularity and smoothness, and a large set of images with varying features. Analysis of comparative results using 13 such wavelet filters, on a set of 20 test images of various sizes, as mentioned above, are presented. Test images from different categories viz. natural images, synthetic images, compound images medical images and aerial images have been coded using a variety of wavelet filters. An embedded quantizer and an adaptive arithmetic entropy encoder are used. Up to five levels of decomposition are used. Using the same filters to code an image, the performance of the image coder at four different compression ratios (8:1, 16:1, 32:1 & 64:1) is evaluated. Due to limitations of space, only a small subset (for compression ratio of 16:1) of the results is presented and analyzed here. Very similar results have been observed for other compression ratios viz. 8:1, 32:1, and 64:1. Although similar experiments have been performed in the lossless context on a set of ISO test images using a number of integer transforms implemented using the lifting scheme [5], those are not presented here. However, the observations in lossless coding are very similar to those of the lossy coding presented in this paper.

The results of the lossy image coding experiment are plotted and shown in Fig. 2. The difference between the worst and the best peak signal to noise ratio (PSNR) values for compressing the same image using different wavelets, is anywhere from 1.5 to 6 dB, which is significant. As we can see, although there are a few wavelet filters, mainly biorthogonal that perform generally well for many images there is no single one that outperforms others for all images. Fig. 3(a) through 3(d), where PSNR values using different wavelets have been plotted for various image types separately, give a clearer picture. For natural images (Lena, Barbara, Baboon, and Airplane), biorthogonal filters like the CDF-9/7, and Villasenor-10/18, perform better than Haar and Daubechies' family of orthogonal filters as shown in Fig. 3(a). It is also observed that, for both Barbara & Baboon images, the PSNR values using the same wavelet filter are lower than that of the Lena and Airplane images. This is mainly due to the presence of more sharp textures and edges in those two images than Lena & Airplane. For pure binary text image (Bengali) as well as for images containing both binary data and gray scale images (Cmpnd1, Cmpnd2), Haar filters outperform the rest by more than 3 dB as shown in Fig. 3(b). Fig. 3(c) shows the PSNR values for the aerial images (Aerial, air1 and air2) where the performance of CDF-9/7, Odegard-9/7 as well as Villasenor-10/18 and Villasenor-13/11 biorthogonal filters are very close. However, they all give better PSNR (by about 2 dB) than the orthogonal filters. It can also be seen that the plots of the images in various classes follow a similar pattern. For example, in the case of aerial images in Fig. 3(c), all three images show poor PSNR values using both CDF-9/11 and Brislawn-10/10 filters. PSNR values for four medical images (mri, nervecell, us (ultrasound), and us1) are shown in Fig. 3(d). As expected, all four images don't perform the same, because of the different spatial features for those images. For both ultrasound images (us and us1), Haar filter gives the best PSNR outperforming others by more than 2 dB. However, for mri and nervecell images, the PSNR curves are similar to those of the aerial images in Fig. 3(c) with CDF-9/7 and Villasenor-10/18 giving the best results. Apart from the best filters, Villasenor-6/10 and Odegard-9/7 perform well for the medical images.

Fig. 4 plots the three basic spatial domain features viz. mean, median, and the standard deviation of the various images used in our experiment. It is observed that, images from different classes tend to show similar characteristics. For example, for both natural and aerial images the difference between the median and mean is small whereas for images containing binary data viz. Bengali, Cmpnd1, and Cmpnd2, the difference is much larger. Also for such images both median and mean are higher than those of the natural images. In all these images the median is higher than the mean which means that the image histograms are negatively skewed. However, for medical images, the histograms are positively skewed which implies that the mean is always above the median. The standard deviation for most images is close to or below 50 except for the images containing binary data and the nervecell image where it is much higher than 50. We feel that these observations in the image characteristics are related to the different coding performance using various wavelet filters. In [6] we explore and link the coding performance with these image characteristics.

5. Conclusion and Work in Progress

To conclude, it is observed that for both lossy and lossless compression schemes, no specific wavelet filter or integer transform has performed uniformly better than others on the variety of test images and the performance has been found to be much more image dependent. In fact, detailed analysis of the results as shown in this paper, reveals that certain wavelets perform better on certain classes of images. For example, images containing binary data (compound images) as well as certain types of medical images (ultrasound images) is compressed best using simple Haar wavelets. Natural and aerial images on the other hand are compressed best using different biorthogonal wavelet filters. Also not all natural and aerial images are compressed best by the same biorthogonal filter. Although not presented here, similar observations are made in the case of lossless compression using different integer wavelet transforms. These observations lead us to conclude that for best results, in both lossy and lossless compression, the most "appropriate" wavelet filter should be chosen to match the image class and the characteristics of the individual image being coded.

References

[1] P. J. Burt and E. H. Adelson, 'The Laplacian Pyramid as a Compact Image Code' IEEE Trans. Communications, vol. 31, No. 4, April 1983, pp. 532-540

[2] J. M. Shapiro, 'Embedded Image Coding Using Zerotrees of Wavelet Coefficients', IEEE Trans. Signal Processing, vol. 41, No. 12, Dec. 1993, pp. 3445-3462

[3] A. Said and W. A. Pearlman, 'A New, Fast, and Efficient Image Codec based on Set Partitioning in Hierarchical Trees', IEEE Trans. Circuits and Systems for Video Technology, Vol. 6, No. 3, June 1996, pp. 243-250

[4] M. Vetterli and J. Kovacevic, 'Wavelets and Subband Coding', Englewood Cliffs, NJ, Prentice Hall, 1995

[5] W. Sweldens, "The Lifting Scheme: A New Philosophy in Biorthogonal Wavelet Constructions", Proc. SPIE Vol. 2569, 1995, pp. 68-79.

[6] S. Saha and R. Vemuri, "Wavelet Coding: Matching the Wavelet Filter to the Image Characteristics", In Preparation.

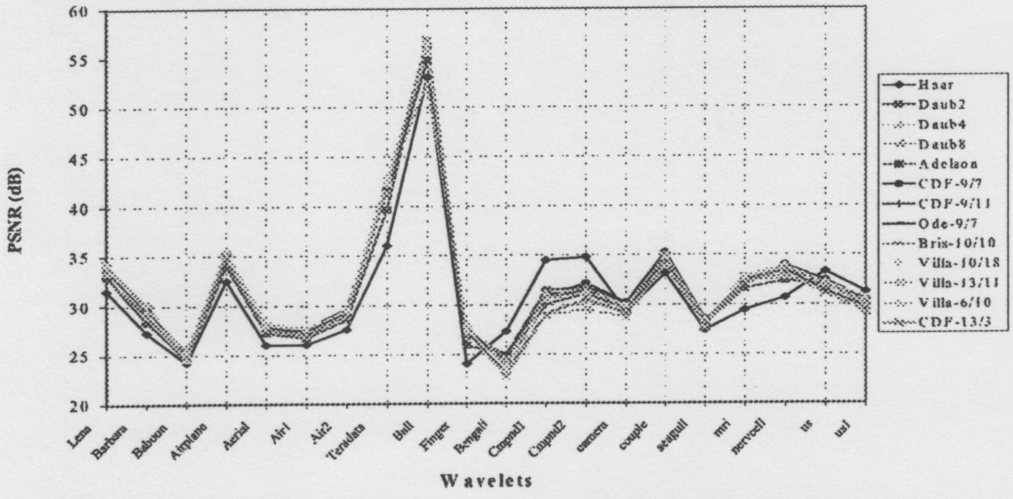


Fig. 2. Lossy Compression Results – PSNR (in dB) for various images using different wavelet filters for a compression ratio of 16:1.

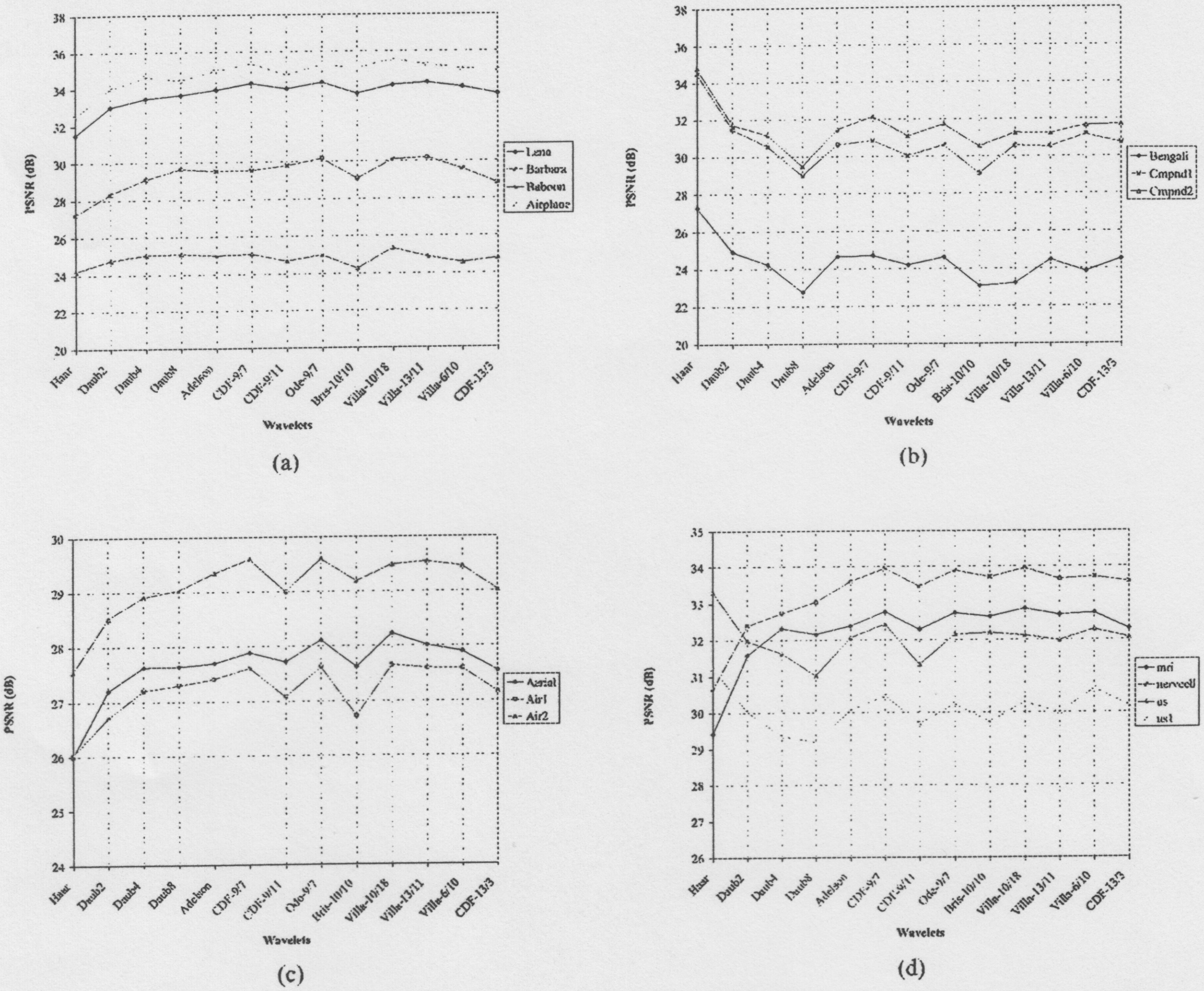


Fig. 3. PSNR results for (a) natural, (b) compound, (c) aerial, and (d) medical images using different wavelets

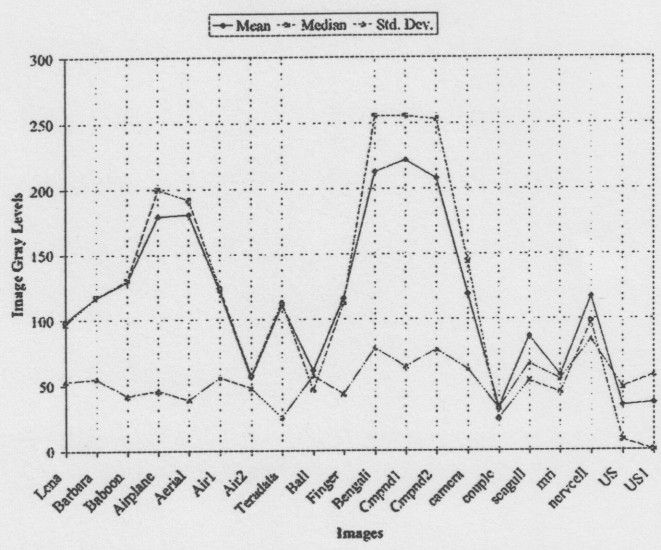


Fig. 4. Image features – mean, median, and standard deviation