EMBEDDED ZEROTREE IMAGE CODING USING LOW COMPLEXITY IIR FILTER BANKS[1]

Charles D. Creusere
Naval Air Warfare Center
Weapons Division
China Lake, CA 93555

ABSTRACT

In this paper, we compare for image coding applications a low-complexity IIR wavelet based on an allpass polyphase decomposition to a pair of linear phase biorthogonal wavelets. To code the wavelet coefficients, we use Shapiro's zerotree algorithm which has the virtues of being both efficient and delivering excellent performance (in a rate-distortion sense). We consider a variety of methods for eliminating filter transients at the image boundaries including circular convolution, symmetric extension (for the biorthogonal wavelets), and direct transmission (for the IIR wavelet). By also coding the filter states in a zerotree form, we find that direct transmission generally performs better than circular convolution. Finally, we show that the use of this IIR wavelet provides equivalent performance to the biorthogonal wavelets at greatly reduced computational complexity.

1. INTRODUCTION

In recent work, Shapiro has developed a simple algorithm for coding images which uses the multiresolutional structure of the wavelet decomposition to achieve very good rate-distortion performance [1]. To further reduce the computational complexity of this coder, we consider here the use of perfect reconstruction infinite impulse response (PR-IIR) filter banks to implement the wavelet decomposition. Such filter banks use noncausal filters in their synthesis sections, requiring that either circular convolution or transmitted filter states be employed to achieve edge transient suppression [2]. For the results presented here, we apply a subclass of the PR-IIR filter bank, the polyphase allpass bank, which has very low computational complexity and has been shown to compare favorably to finite impulse response (FIR) filter banks in actual coding applications [3].

In [3] it has been shown that at higher bit rates it is better (in terms of both objective and subjective performance measures) to code and transmit appropriate initial state vectors to the synthesis filters than it is to perform circular convolution. By explicitly transmitting initial state vectors to the receiver, we guarantee perfect reconstruction (in the absence of coding) without putting any preconditions on the initial states of the analysis filters. Consequently, we are free to symmetrically extend the sequence fed into the analysis bank, eliminating the boundary discontinuity at the leading edge of the image while still achieving perfect reconstruction (despite our use of nonsymmetric filters) in the complete analysis/synthesis system. Thus, we have performance improvements similar to those achieved with symmetric FIR filter banks by using symmetric extension in lieu of circular convolution [4]. Of course, the disadvantage of this method is that we must now transmit side information to the receiver, reducing the number of bits available for coding the wavelet coefficients. Because of the simplicity of its coding scheme, the results presented in [3] are not impressive in terms of rate-distortion performance. In addition, the method used to code the initial state vectors (6 bit PCM) was not very sophisticated, giving circular convolution an advantage at low bit rates.

In this work, we use the same low complexity filters as [3] but with a zerotree coding algorithm based on [1]. In addition, we show how the filter states can be coded and transmitted to the receiver in exactly the same way as the wavelet coefficients themselves, greatly decreasing the bit rate required for this side information at a negligible cost in encoder complexity. The paper is organized as follows. After a brief introduction to 2-band PR-IIR filter banks, we review the zerotree coding algorithm and introduce modifications to handle filter state transmission. Next, we apply a PR-IIR filter bank to actual images and compare its coding results to those achieved using low-complexity biorthogonal wavelet decompositions. Finally, conclusions are presented

2. POLYPHASE FILTER BANKS WITH ALLPASS SUBFILTERS

Shown in Fig. 1 is a 2-band polyphase analysis/synthesis system. Clearly, if

$$A_k(z) \cdot B_k(z) = c \cdot z^{-d}, \quad k = 1,\ 2 \tag{1}$$

for some constant c and an integer d then this system is perfect reconstruction since the DFT (discrete Fourier transform) and IDFT (inverse DFT) blocks cancel out while the remaining delay chains and sampling rate alterations are equivalent to a one sample throughput delay [5]. In the case where the analysis subfilters, $A_k(z)$, are allpass, it is easy to see that the same filters can be used in the synthesis bank but operating in reverse-time, i.e.,

$$B_k(z) = A_k\left(z^{-1}\right), \quad k = 1,\ 2, \tag{2}$$

since

$$A(z) \cdot A\left(z^{-1}\right) = 1 \tag{3}$$

by the definition of an allpass filter, clearly satisfying (1) [6]. In [3] it was found that a very low-complexity allpass filter bank where $A_0(z)$ is 1st order with a pole at 0.375 and $A_1(z)$ is 0th order (i.e., 1) provides the best coding results while minimizing ringing at sharp edges. Consequently, we use these filters in the 2-band analysis/synthesis system shown in Fig. 1 to implement a 2D separable wavelet decomposition.

Because the filters of (2) are entirely maximum phase, they can be implemented stably by simply operating the original allpass filter backwards on a line in the image-- i.e., starting

[1] **Unclassified, Distribution Unlimited**

0-7803-2431-5/95 $4.00 © 1995 IEEE

from the end of a finite-length output block and filtering toward the beginning. Unfortunately, doing this reverse-time processing results in transients at the block edges since the perfect reconstruction property was derived in the z-domain and thus ignores startup effects. There are two straight-forward, stable ways to eliminate these transients: circular convolution [5] or direct transmission of final analysis filter states [2]. The first method eliminates the need to transmit side information to the receiver at the expense of introducing ripples (which require more bits to code) into the wavelet coefficients. The second method, on the other hand, does not introduce these ripples but requires the transmission of side information. As we shall see, the direct transmission method is actually superior to circular convolution at all but the lowest bit rates.

3. ZEROTREE CODING ALGORITHM

As mentioned above, we use the embedded zerotree wavelet (EZW) coding algorithm for all of the results presented here [1]. The fundamental observation around which this coding algorithm is centered is that there is a strong correlation between insignificant coefficients at the same spatial locations in different wavelet scales-- i.e., if a wavelet coefficient at a coarser scale is zero, then it is more likely that the corresponding wavelet coefficients at finer scales will also be zero. Figure 2 shows a 3-level, 2D wavelet decomposition and the links which define a single zerotree structure. If the wavelet coefficient at a given scale is zero along with all of its descendants (as shown in Fig. 2), then a special symbol indicating a zerotree root is transmitted, eliminating the need to transmit the values of the descendants. Thus, the correlation of insignificance across scales results in a net decrease in the number of bits transmitted.

In order to generate an embedded code (where information is transmitted in order of importance), Shapiro's algorithm scans the wavelet coefficients in a bit-plane fashion. Starting with a threshold determined from the magnitude of the largest coefficient, the algorithm sweeps through the coefficients, transmitting the sign (+ or -) if a coefficient's magnitude is greater than the threshold (i.e., it is significant), a ZTR (zerotree root) if it is less than the threshold but the root of a zerotree at the coarsest possible scale, or a 0 otherwise-- this is the dominant pass. Next, for the subordinate pass all coefficients deemed significant in the dominant pass are added to a second subordinate list which is itself scanned, adding one bit of resolution to the decoder's representation of each significant coefficient. The threshold is then halved and the two passes are repeated with those coefficients having been found significant previously being replaced by zeros in the dominant pass (so that they do not inhibit the formation of zerotrees in subsequent iterations). The process continues until the bit budget is exhausted; at this point, the encoder transmits a stop symbol and its operation is terminated. The decoder, on the other hand, simply accepts the bit stream coming from the encoder, progressively building up the significance map and the subordinate list in the exact same way as they were created by the encoder. Because of this precise synchronization, the resolution enhancement bits transmitted during the subordinate pass do not need any location specifiers-- the decoder knows the exact transmission order of these bits because it has reconstructed the same subordinate list as the encoder had at that point in the process.

Modifying the EZW scheme to also code and transmit the final analysis state vectors for transient suppression is straightforward and its rationale is identical to that used to justify zerotree coding of the wavelet coefficients themselves. Basically, since the final states of the analysis subfilters are largely determined by the spatial region nearest the edge of the input image, these states will have a strong correlation with the pixels in this region and will, thus, inherit some of their properties. Specifically, the correlation of insignificance between scales which has been noted in the wavelet coefficients will also be inherited, at least to some extent, by the filter state vectors. Binary zerotrees can thus be constructed for the filter states in the same as the quadtree-based zerotrees are constructed for the coefficients-- this is shown in Fig. 2 (L for the lowpass subfilter state and H for the highpass subfilter state). In general, each box in L or H can be a vector, and the simplest (although certainly not best) way to handle this is to map each vector element into an independent zerotree. The filter states can be added to the dominant list anywhere and will be scanned and transmitted in exactly the same way as the wavelet coefficients themselves. In order to stay consistent with the embedded structure of this coder, we have added the filter state information to the list immediately after the wavelet scale which uses it; the scanning of both the wavelet coefficients and the filter states in the dominant pass basically goes from the upper left-hand corner of Fig. 2 to the lower right-hand corner. As before, the decoder has all of the information it needs to decode the bit stream as long as it knows where the analysis state vectors have been inserted into the dominant list.

4. IMAGE CODING RESULTS

A 5-level wavelet decomposition combined with zerotree coding of the coefficients has been used for all of the results presented in this section. The filters considered here are the allpass polyphase filter bank described in Section 2, the 5/3 biorthogonal wavelet, and the 2/6 biorthogonal wavelet [7] (#A/#S where #A is the length of the analysis lowpass filter and #S is the length of the synthesis lowpass filter). These filters have been chosen for their low computational complexity, and both symmetric extension and circular convolution are considered here to eliminate startup transients. For the allpass filter bank, we are free to select the initial states of its analysis filters as desired if direct transmission is used to eliminate its synthesis filter transients. Consequently, we symmetrically extend the input sequence and process this extended sequence with the analysis bank to initialize its filter states so that no large discontinuities are introduced into the wavelet coefficients. When implementing the allpass filter bank using circular convolution, on the other hand, all of the initial filter states (in both the analysis and synthesis banks) are fixed-- there are no degrees of freedom to manipulate.

Figure 3 shows rate-distortion plots of the coded 'lena' image for the various low-complexity filter bank configurations examined here. The bit rate on this plot is simply the number of bits used to represent one pixel in the coded image while the peak signal to noise ratio (PSNR) is given by

$$\mathrm{PSNR} = 10 \cdot \log\left(\frac{255^2}{\mathrm{MSE}}\right) \qquad (4)$$

where the MSE (mean squared error) between the original image, $x(k,l)$, and its reconstruction, $\hat{x}(k,l)$, is

$$\text{MSE} = \frac{1}{X \cdot Y} \sum_{k=0}^{X-1} \sum_{l=0}^{Y-1} \left(x(k,l) - \hat{x}(k,l) \right)^2 \qquad (5)$$

for an image of size X by Y. The first thing to note about Fig. 3 is that all of the plots are very close together. Also, for all of the filters examined, circular convolution is never better than the other method considered (direct transmission or symmetric extension), and at most bit rates it is worse. For example, the direct transmission method achieves about 0.5 dB higher PSNR at 1.0 bpp (bits per pixel) than circular convolution for the allpass filter bank.

Subjectively, the performance of the allpass filter bank is equivalent to that of the two biorthogonal wavelets at bit rates of greater than 0.25 bpp; in fact, near 1.0 bpp, its reconstructed image looks noticeably sharper than that of the 5/3 wavelet. At lower bit rates, however, ringing artifacts become much more pronounced in the allpass system. When the lena image is coded to 0.125 bpp, the ringing at sharp edges in the allpass-based system is approximately equivalent to that produced using a coder based on a longer 13/11 biorthogonal wavelet [8]. Consequently, for bit rates less than 0.25 bpp, one of these very short biorthogonal wavelets is probably preferable.

One further point which should be considered here is the computational complexity of the various transforms. Both of the biorthogonal wavelets require 6.98 M (million) multiples and 8.38 M additions to decompose and synthesis a 512 by 512 image using a 5-level transform (taking advantage of filter symmetry to reduce the computational burden). The allpass-based system, on the other hand, requires about 698 K (thousand) multiplications and 2.79 M addition using direct transmission and 1.40 M multiplication with 5.59 M additions using circular convolution. Note that these complexity results assume that the allpass subfilters are implemented using the highly efficient structures of [9]. Clearly, then, the allpass filter bank has a significant computational advantage over even the shortest biorthogonal wavelets, especially when direct transmission is used to eliminate edge transients. Their low complexity, combined with their good rate-distortion and perceptual performance at all but the lowest bit rates, provides strong incentive for considering these filter banks in high speed or low cost image coding applications.

The results of our comparison between circular convolution and direct transmission also highlight an interesting question: why is the rate-distortion performance of the direct transmission method which, in general, requires the transmission of side information as good as or better than that of circular convolution which does not require the transmission of side information? The answer is basically that, by eliminating the false edges introduced into the wavelet coefficients by circular convolution, we have made the coefficients easier to code. One indication of this is the entropy of the wavelet coefficients-- the entropy of these coefficients using a 5-level decomposition is 6.08 bits for circular convolution and 5.84 bits for direct transmission (including the transmitted filter states). Even though our coder takes advantage of far more correlation in the coefficients than a simple entropy coder would, this results does indicate that the coefficients created via circular convolution simply have more information in them and thus they must be coded with more bits to achieve a specified quality in the reconstructed image.

5. CONCLUSION

In this paper, we have compared zerotree coders based on low complexity biorthogonal wavelets to one based on a polyphase allpass filter bank and have found that the lower complexity allpass bank does at least as well as the biorthogonal wavelets at all but the lowest bit rates (less than 0.25 bpp). While still comparing favorably in terms of PSNR at very low bit rates, the allpass coder begins to suffer from ringing artifacts near edges similar to those seen with longer biorthogonal wavelets. In addition, modifying the zerotree algorithm to code and transmit the initial filter states (used to eliminate edge transients) improves the overall rate-distortion performance of the system while reducing the filter complexity by eliminating the need to do circular convolution. Making this modification requires only the insertion of a few new nodes into the dominant linked list and does not change the structure or operation of the coding algorithm at all. Thus, for coders which must operate at high speeds (e.g., video) or where low cost is paramount, the polyphase allpass filter banks presented here are an excellent alternative to biorthogonal wavelets.

REFERENCES

[1] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing*, Vol. 41, No. 12, December 1993, pp. 3445-3462.

[2] C.D. Creusere, *Perfect Reconstruction Modulated Polyphase Filter Banks Using Reverse-Time Subfilters*, Ph.D. dissertation, University of California, Santa Barbara, CA, October 1993.

[3] C.D. Creusere and S.K. Mitra, "Perfect reconstruction modulated polyphase filter banks with reverse-time subfilters for image coding applications," *Technical Report CIPR 94-05*, Center for Information Processing Research, University of California, Santa Barbara, CA, March 1994.

[4] M.J.T. Smith and S.L. Eddins, "Analysis/synthesis techniques for subband image coding," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 38, No. 8, August 1991, pp. 1446-1456.

[5] P.P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, Englewood Cliffs, NJ 1993.

[6] T.A. Ramstad, "IIR Filterbank for Subband Coding of Images," *Proc. Int. Symposium on Circuits and Systems*, 1988, pp. 827-830.

[7] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia, PA, 1992, pp. 261.

[8] M. Lightstone and E. Majani, "Low bit-rate design considerations for wavelet-based image coding," Proc. of the SPIE Symposium on Visual Communications and Image Processing, Sept. 1994, Chicago, IL.

[9] S.K. Mitra and K. Hirano, "Digital allpass networks," IEEE Trans. on Circuits & Systems, Vol. CAS-21, No. 5, Sept. 1974, pp. 688-700.

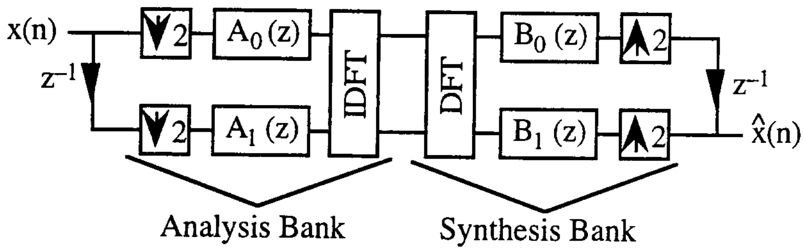

Figure 1: 2-band polyphase analysis/synthesis system.

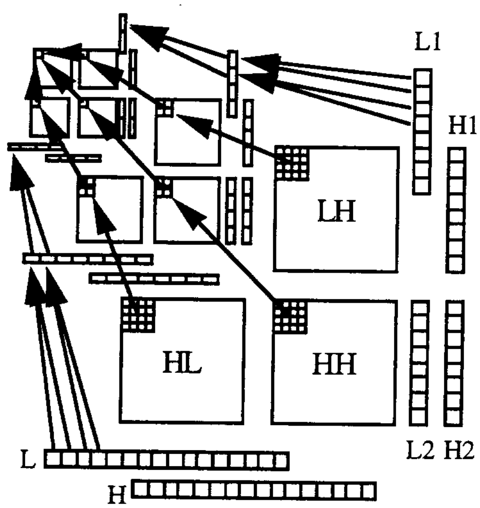

Figure 2: Children to parent links for zerotree computation.

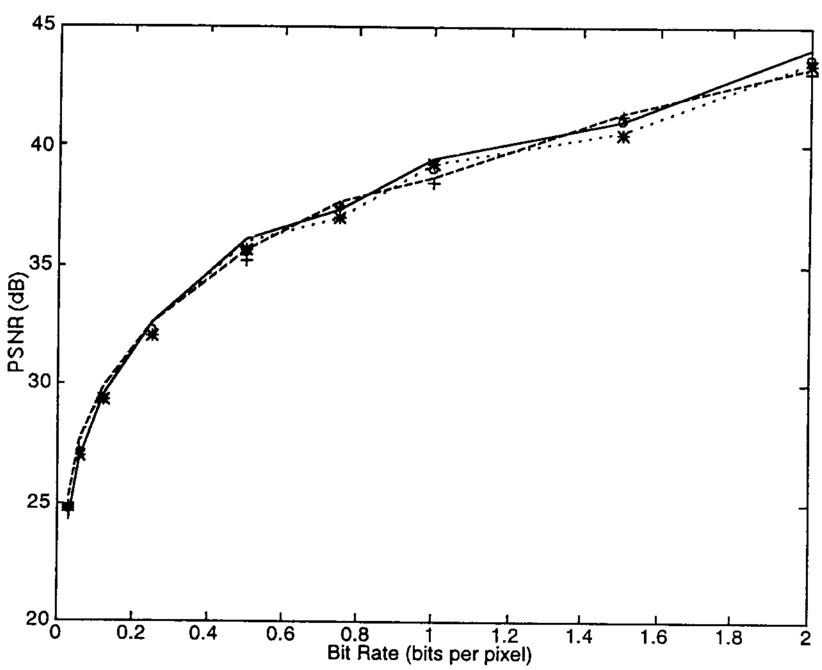

Figure 3: Rate-distortion curves for lena image. Allpass filter bank: solid is direct transmission, circle is circular convolution; 5/3 biorthogonal: dotted is symmetric extension, asterisk is circular convolution; 2/6 biorthogonal: dashed is symmetric extension, plus is circular convolution.