ON THE STUDY OF DIAMOND-SHAPE FILTER BANKS AND APPLICATION IN IMAGE COMPRESSION

Soontorn Oraintara†, Trac D. Tran‡ and Truong Q. Nguyen†

†Electrical and Computer Engineering Department, Boston University
8 St. Mary’s St., Boston, MA 02215
Emails: oraintar@bu.edu, nguyent@bu.edu
‡Electrical and Computer Engineering Department, University of Wisconsin
1415 Engineering Dr., Madison, WI 35706
Email: trac@saigon.ece.wisc.edu

ABSTRACT
This paper discusses non-separable 2-D filter banks with diamond-shape support. The frequency spectrum partitioning is shown to have higher theoretical coding gain than that of the conventional separable approach in image coding application. However, based on the frequency response structure of the subband filters, they can not be implemented with good frequency characteristic although the passband structure is permissible. Some constraints on the passband structure are discussed, and the comparison of the performances between the separable and diamond-shape filter banks is presented. Throughout the simulations, the (9,7) wavelet and its McClellan transformed version are used, respectively, to represent the separable and diamond-shape approaches. Various coding result comparisons are presented.

1. INTRODUCTION

Subband coding is one of the most successful applications of filter banks and wavelets. An input signal is projected into a number of less correlated subspaces using an invertible transform. Each projection is called a subband signal. Figure 1(a) shows a structure of the conventional 2-D separable filter bank with two channels in each direction (horizontal and vertical). \( H_{10}(z) \) and \( H_{11}(z) \) are lowpass and highpass filters for the \( x \) direction and so as \( H_{20}(z) \) and \( H_{21}(z) \) for the \( y \) direction with corresponding decimation matrices \( D_1 = \text{diag}(2, 1) \) and \( D_2 = \text{diag}(1, 2) \) respectively. Since the filters \( H_{2i} \) do not depend on \( x \), \( H_{2i}(D_1(z)) = H_{2i}(z) \)) implies that the same structure is used for the four-channel system.

The separable wavelet transform has been widely applied in image compression. A 2-D input signal (an image) is processed in two directions, \( x \) and \( y \), separately. Figure 2(a) shows the frequency spectrum partitioning of this separable subband decomposition after one iteration. Typically, natural images are modeled as an AR(1) process with correlation factor \( \rho < 1 \). The psd (power spectral density) of the input \( x(n) \) based on this model has highest magnitude at DC (\( \omega = (0, 0) \)) and decays almost monotonically along the radial direction. When the filters are ideal, there is no correlation among the subbands which is one of a requirement for an optimal transform [1]. However, this is not sufficient. From the symmetry between \( x \) and \( y \) directions of the psd, it can be seen that the signal energy in subbands 1 and 2 are comparable. Since the psd of an AR(1) model has, roughly speaking, higher energy at low frequency than at high frequency, non-separable 2-D wavelets are of interest in order to majorize the psd among the subbands.

Non-separable 2-D wavelets, which can be obtained from non-separable filter banks, have been a subject of study for many researchers. They are proposed to better capture the signal energy comparing to the popular separable method. However, no superior coding performance has been reported so far. In this paper, we discuss the possibilities of using non-separable wavelets in image coding. For each iteration, we partition the frequency spectrum into four subbands as illustrated in Figure 2(b) with the corresponding decimation matrix \( D = D_1D_2 = 2I \). There are four shifting vectors \( (k_0, k_1, k_2, k_3) = ((0, 0)^T, (1, 0)^T, (0, 1)^T, (1, 1)^T) \), associating with the decimation matrix \( D \), and thus the corresponding aliasing frequencies are \( \pi k_i, i = 1, 2, 3 \). This frequency parti-
2. DIAMOND-SHAPE FILTER BANKS

2.1. Coding Gain

Transform coding gain (TCG) is a common tool to evaluate the performance of a transform in decorrelating the input signal into subbands. It can be obtained from the ratio of the variances between using direct PCM and the transform, and is equivalent to the ratio between the arithmetic to geometric means of the variances of the subband signals [2]. In this section, we extend the formula of the coding gain given in [3] from the 1-D case to the 2-D case.

Let \( x \) be a Markov process with the inter-correlation coefficient \( \rho (< 1) \), the acf (autocorrelation function) for 1-D case is given by \( R(x) = \rho |a| \). However in 2-D case, the scalar index \( n \) in the acf is substituted by a vector index \( n \), and the absolute value is replaced by a norm

\[
R_{xx}(n) = \rho ||n||
\]  

where the norm can be either \( \ell_1 \) or \( \ell_2 \) norm. If it is \( \ell_1 \) norm, this is called separable acf model which is appropriate to artificial or man-made images, and when it is \( \ell_2 \) norm, it is called isotropic acf model which is appropriate for natural object [4]. Base on the acf, the TCG formula can be modified as follows:

\[
G(\rho) = \prod_{k=0}^{M-1} (A_k B_k)^{-\alpha_k} 
\]

\[
A_k = \sum_m \sum_n h_k(m) h_k(n) \rho ||m-n|| 
\]

\[
B_k = \sum_n f_k(n). 
\]

where the norm in the formula of \( A_k \) is corresponding to the norm in (1). From the top subplot of Figure 3, it can be seen that in the separable acf model, the ideal diamond-shape filter bank has 1.65 dB (at \( \rho = 0.95 \)) higher coding gain than the ideal separable filter bank. However with the (9,7) pair, the separable filter bank has 0.92 dB (at \( \rho = 0.95 \)) higher coding gain than the (9,7) diamond-shape filter bank obtained from the McClellan transformation. Similar results are also drawn from the case of isotropic acf model as shown in the bottom of Figure 3. The reasons for the difference lie at the passband structure of the diamond-shape filters and the iteration scheme in Figure 1(a). Both will be rationalized in the next subsection.

![Figure 3: Comparison of the coding gain for the separable acf model and isotropic acf model.](image)

2.2. The constraints on the passband supports of the filters

For PR systems, in order to successfully cancel the aliasing caused by downsampling [2, 3], aliasing has to appear in pairs. If the supports of the filters satisfy this condition then it is called permissible, which is needed in real implementation with filters having good frequency responses. For more details, the reader is referred to [5]. Figures 4(a) and (b) shows the permisibilities of the conventional separable case and the diamond-shape case where each column is corresponding to the shifted version by the shifting vector \( \pi k_i, \ i = 0, 1, 2 \) and 3. The areas of aliasing are denoted by the highlighted lines. It can be observed that, in both cases, aliasing appears in pairs, and therefore both are permissible. However, in the diamond-shape case, the number of aliasing regions is more than the separable case (the last column of Figure 4(b)), and thus in practice, this extra region is the transition bands of the subbands 1 and 2 and can not have good frequency attenuation around the edges. In other words, the diamond-shape filter bank can not decorrelate signals very well between the subband 1 and 2.

Although the passband structure of the diamond-shape filters is permissible, in practice, the frequency response of the filters cannot be good due to the iterative construction in Figure 1(a). Figure 5(a) and (b) show the 3-D frequency responses of the separable (9,7) filter bank and the diamond-shape (9,7) filter bank. Since each analysis filter is a product filter \( H_j(z) = H_{1a}(z) H_{2b}(D_j^2 z) \) with \( H_{1a}(z) \) and \( H_{2b}(z) \) having diamond-shape supports, it is impossible for \( H_j(z) \) to have zero response at \( \omega = 0 \) or \( \pi \) and
\frac{\pi}{2} \leq \omega \leq \pi$. Therefore there is some energy leakage along the lines $\omega_x = 0$ and $\omega_y = 0$ unlike the separable case. As one can observe from Figure 5, the frequency responses of the filters in the non-separable case are not aligned with the $x$ or $y$ axis. The aforementioned energy leakage partly explains the inferior coding gain of the diamond-shape filter bank in practice.

3. SIMULATIONS

In this section, we present various results in image coding at different compression ratios using both frequency partitioning scheme. Again, the (9,7) wavelet is employed in the comparison between separable and diamond-shape case. All the images are tested at four iteration of the structure in Figure 1(a). Figure 6(a), (b), (c) and (d) show comparisons on Lena, Baboon, Goldhill and Boat respectively. The performance of the diamond-shape wavelet is about 0.5-1 dB worse than of the conventional (9,7), comparable to the inferior coding gain presented in section 2.1. Figure 7(a) and (b) show examples of compressed Baboon images at 32:1 compression ratio using separable and diamond-shape wavelets, respectively. Although, theoretically, the diamond-shape scheme should perform better than the separable one, because of the lack of an efficient filter design method for 2-D filter bank, the frequency responses using McClellan transformation have wide transition bands. Hence, the resulting subband filters can not decorrelate the signal very well. Moreover, since we use two iterations of a diamond-shape two-channel filter bank to produce a four-channel filter bank, as discussed in section 2.2, there are some energy leakage along the lines $\omega_x = 0$ and $\omega_y = 0$ to the highpass subbands. If one can come up with a design method which provide non-separable highpass filter with zero response at the $x$ and $y$ axis, the filter bank’s coding performance should improve significantly.

4. FUTURE STUDY

This paper presents an image coding method using diamond-shape filter banks. Two 2-D models are use to derive the coding gain formulae for the transforms. In theory, the diamond-shape scheme provides better energy compaction due to the symmetry between the horizontal and vertical axis, i.e. the upper bound for the theory coding gain is higher than of the separable scheme. However, in real implementations, the filters can not be ideal and hence the transition bands can not have zero bandwidth. As previously discussed, there is an extra large transition band between subbands 1 and 2, which produces a large amount of aliasing (the last column
Figure 6: PSNR comparisons for (a) Lena, (b) Baboon, (c) Goldhill, and (d) Boat.

Moreover, based on the iteration in Figure 1(a) in order to obtain a four channel filter bank, the frequency response of the highpass filter cannot be zero along the axis $\omega_x = 0$ and $\omega_y = 0$. This results in energy leakage into the highpass subband signals and thus degrades the performance of the coder. The simulation results show that using the diamond-shape filter bank gives about 0.5-1 dB worse PSNR than using the conventional separable method. Note that the filters used in the non-separable case is derived from the McClellan transformation, and thus the energy leakage along the two main axis cannot be avoided. One might be interested in designing 2-D four-channel filter bank with the passband given in Figure 2(b) such that the highpass filter has zero response on the two axis to achieve better performances.

5. REFERENCES


Figure 7: Compressed Baboon’s at 32:1 using (a) separable (9,7) wavelet with 23.98 dB PSNR, and (b) diamond-shape (9,7) wavelet with 23.41 dB PSNR.