Real Time Adaptive PR-QMF Bank Design for Image Coding Using Interior-Point Algorithm

V. Kustov, P. Srinivasan, S. Mitra, and S. Shishkin,
Department of Electrical Engineering,
Texas Tech University, Lubbock, TX, 79409, MS-3102

Abstract- **A fast algorithm for designing two-channel perfect reconstruction quadrature mirror filter (PR-QMF) banks for real time image coding is presented. This interior-point-based optimization algorithm finds a wavelet filter that allows the least amount of energy in the higher frequency subbands. Once such a wavelet filter has been found, only the low-frequency band is used by a chosen advanced lossy image compression algorithm. The higher subbands are completely discarded (approximated by zeros).This technique provides a significant reduction in execution time without an appreciable increase in distortion.**

1. Introduction

Multiscale decomposition of images using PR-QMF filters has been successfully used in many wavelet-based lossy image compression techniques such as the Embedded Zero Tree[1] [EZW] algorithm, the Adaptive Fuzzy Leader Clustering Vector Quantization[2] [AFLC-VQ] algorithm, and others such as in [3]. One disadvantage of these coding algorithms when used for high-fidelity compression of high-resolution images, e.g. 1024x2048 and 2048x4096 pixels Visible Human color medical images, which range in sizes from 6 to 24 Mbytes, is that they require a multiple scanning of the entire wavelet decomposition of an image, thus resulting in an unacceptable coding/decoding time for real time applications as internet transmission or teleconferencing.

An effective means of reducing the computational complexity is to code only the low-frequency subband of the wavelet decomposition. Consequently, only a quarter of the image data set needs to be processed. In addition, we decrease the number of operations to perform the transform, since the higher-frequency bands do not have to be computed (they are approximated by zero). Because we have reduced the data set, a larger part of the data set now fits into the internal memory of a particular processor. This is important since accessing the external memory is the “bottleneck” for most fast processors such as Texas Instruments’ C’6000 DSP’s[4]. An effective use of the direct memory access channels is a challenge for many algorithms for which the order in which the data is processed is determined during the execution time.

We here propose a new algorithm that creates optimal energy compaction filters in real time. The higher frequency bands are discarded (in practice they are never computed) without a significant loss of information. The reduction of the execution time is achieved by reducing the original data set by a factor of four, and by placing a larger portion (if not all the data) of the reduced data set in the internal memory.

Some previous work in this area has been done using different approaches. For example in [5] a similar problem is formulated with respect to the product filter in the frequency domain. In this formulation every locally optimal solution is a global one. However, this optimization problem has a linear objective function with infinitely many linear constraints. The optimization technique proposed in [6] for maximizing theoretical coding gain reduces the problem to an unconstrained optimization at the expense of loss of completeness, i.e. not all the active feasible region is searched, and thus the solution cannot be global[7]. The objective function that expresses the amount of energy in the neglected subbands is a non-convex function of the high-pass filter subject to quadratic and bilinear constraints as well as one linear constraint. Thus, the direct implementation of methods such as the sequential quadratic programming (SQP) will produce only a local minimum, while the direct use of interior-point methods[8] are not applicable due to the non-convex objective function. We here propose a specialized algorithm for this optimization problem for a special filter length.

2. Background

Unless otherwise specified all functions used in this paper are assumed to be real and compactly supported. Functions' arguments and indices are assumed to be integer, if not specified otherwise. All functions are assumed to be bounded. If the upper and lower limits of a summation are omitted, the operation is performed on all values of the argument(s) for which the expression inside the summation is supported. Where the extension to two-dimensions is straightforward one-dimensional notations are used.

The part of the wavelet decomposition created by applying the low-pass filter in both directions is called the low-frequency subband/part of the decomposition. The rest of the decomposition forms the higher frequency subbands.

Let us denote by $h_0(n)$ and $h_1(n)$ the low-pass and high-pass filter, respectively, that form a two-channel filter bank. The filter bank is orthogonal if the following equations hold true[9]:

$$\sum_n h_0(n+2m)h_0(n) = \delta(m), \tag{2.1}$$

$$\sum_n h_1(n+2m)h_1(n) = \delta(m), \tag{2.2}$$

$$\sum_n h_1(n+2m)h_0(n) = 0, \tag{2.3}$$

In the case of quadrature mirror filter (QMF) banks[10] $h_0(n)$ and $h_1(n)$ are related as

$$h_0 = (-1)^{k+1} h_1(N-k), \qquad k = 0..N-1 \tag{2.4}$$

Note that if (2.4) and (2.2) are satisfied then (2.1) and (2.3) are always satisfied. Finally, we require that Fourier transform of $h_1(n)$ have a zero DC component, or in the time/spatial domain:

$$\sum_n h_1(n) = 0 \tag{2.5}$$

3. Description of the technique

The technique can be divided into several steps. First the algorithm amasses statistics about the input image. Based on these statistics the algorithm finds an optimal PR-QMF bank that puts the smallest amount of energy in the higher frequency subbands of the decomposition (alternatively, it compacts the energy in the low-frequency subband, since the transform is orthogonal) for a given input image. The resultant low-frequency subband is used by the coder, such as EZW ALFC-VQ, JPEG, etc. Note that the coding algorithm does not have to be wavelet-based, since it processes the low-frequency band as any other input image. In the decoding part the compressed low-frequency band is decoded first. Then, the inverse wavelet transform is performed using the optimal filter bank with the detail part approximated by zeros. A block diagram of the proposed technique is shown in Fig 3.1

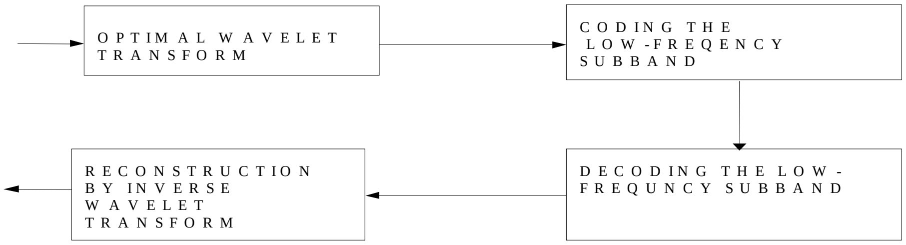

Fig 3.1
A block diagram of the proposed technique
The main compression algorithm acts on a reduced data set.

Note that since the optimal filter bank is image-dependent, the filter coefficients should be stored in the compressed image file. However, since we are dealing with large images, storing either $h_0(n)$ or $h_1(n)$ does not increase the size of the compressed file significantly.

The optimization algorithm finds the optimal $h_1(n)$ of length six (although it can be extended to other filter lengths). Its main idea is to reformulate the problem without loss of generality into a convex optimization problem and then solve the latter by an interior-point method. Lets us describe briefly the main stages of our method: First using (2.2), (2.5), we parameterize the vectors $h_0(n)$ and $h_1(n)$ by the vector $y \in R^5$ to remove the linear equation (2.5) and we rewrite Eq (2.2), and (2.5) using vector notations:

$q^T h_1 = 0, \quad q = [1,1,1,1,1,1]^T$, then:

let $y = Gh_1, \quad G \in R^{6x5} : q^T G = 0$

and write the optimization problem in the form:

$$y^T A_0 y \to \min \tag{3.1}$$

$$y^T A_i y = 0, \quad i = 1,2; \quad y^T A_3 y = 1, \tag{3.2}$$

where matrix A_0 is obtained from the horizontal or vertical autocovariance matrix of the image, and A_1, A_2, A_3 are obtained from (2.2) after the parameterization (to remind: (2.1) and (2.3) follow from (2.2) and (2.4)).

Then we consider the auxiliary optimization problem:

$$x^T (A_0 - \mu A_3) x \to \min \tag{3.3}$$

$$x^T A_i x = 0 \,, \qquad i = 1,2 \qquad x \in R^5 \tag{3.4}$$

and adjust $\mu \in R$ in such a way that the optimal value of (3.3) and (3.4) are equal to zero. When such μ_* is found and the optimal solution x_* is located, the solution of (3.1) and (3.2) can be obtained by

$$y_* = \frac{x_*}{\sqrt{x_*^T A_3 x_*}} \tag{3.5}$$

It remains to describe the solution method for (3.3) and (3.4) when μ is fixed. For any $\lambda \in R^2$, let us denote

$$B = A_0 - \mu A_3 \qquad B = \begin{matrix} \tilde{B} & b \\ b^T & \beta \end{matrix} \qquad A_i = \begin{matrix} \tilde{A}_i & a_i \\ a_i^T & \alpha_i \end{matrix}$$

$$\tilde{A}_\lambda = \tilde{B} + \lambda_1 \tilde{A}_1 + \lambda_2 \tilde{A}_2 \,, \qquad a_\lambda = b + \lambda_1 a_1 + \lambda_2 a_2 \,, \qquad \alpha_\lambda = \beta + \lambda_1 \alpha_1 + \lambda_2 \alpha_2$$

$$f(\lambda) = a_\lambda^T \tilde{A}_\lambda^{-1} a_\lambda - \alpha_\lambda \,,$$

and consider the optimization problem:

$$\text{minimize } (f(\lambda) : \tilde{A}_\lambda \geq 0) \tag{3.6}$$

If λ_* is the solution of (3.6), then $x_\lambda = \begin{matrix} -\tilde{A}_\lambda^{-1} a_\lambda \\ 1 \end{matrix} \in R^5$ is the solution of (3.3) and (3.4). It is easy to verify that the problem (3.6) is convex. We solve this problem using an interior-point method[6]. On all of the test images the interior-point algorithm converges to a global minimum in 15 iterations independent of the input image size.

4. Results and discussion

We tested our algorithm by applying it to different images and then computing the resulting amount of energy in the higher frequency subbands of the wavelet decomposition. This value was compared with the amount of energy in these subbands when same length Daubechies' Maxflat filters[9] were used. These Daubechies' filters were selected because under some frequently met assumptions on the autocovariance sequence of an image, they are near optimal energy compaction filters[11].

To evaluate the overall performance of the algorithm, let us define the figures of merit first (we do not use the coding gain as the figure of merit, because in general optimization for coding gain and energy compaction are not the same[5]). Let us denote by E_d the amount of energy in the higher subbands of a one level decomposition using Daubechies' Maxflat filters[9] and by E_o using optimal filters.

$$T = \frac{E_d - E_o}{(E_d + E_o)/2} x100\%$$

Table 4.1 below shows the algorithm's performance when it was applied to four test images.

Table 4.1 Energy decrease for gray scale images

Image	Improvement (T) in the detail part of the decomposition.
Lena	7.07
Visible Human	3.75
Calgary	18.95
Banff	56.44

Figure 4.1 shows the effect of discarding the detail part on the reconstructed (Calgary) image.

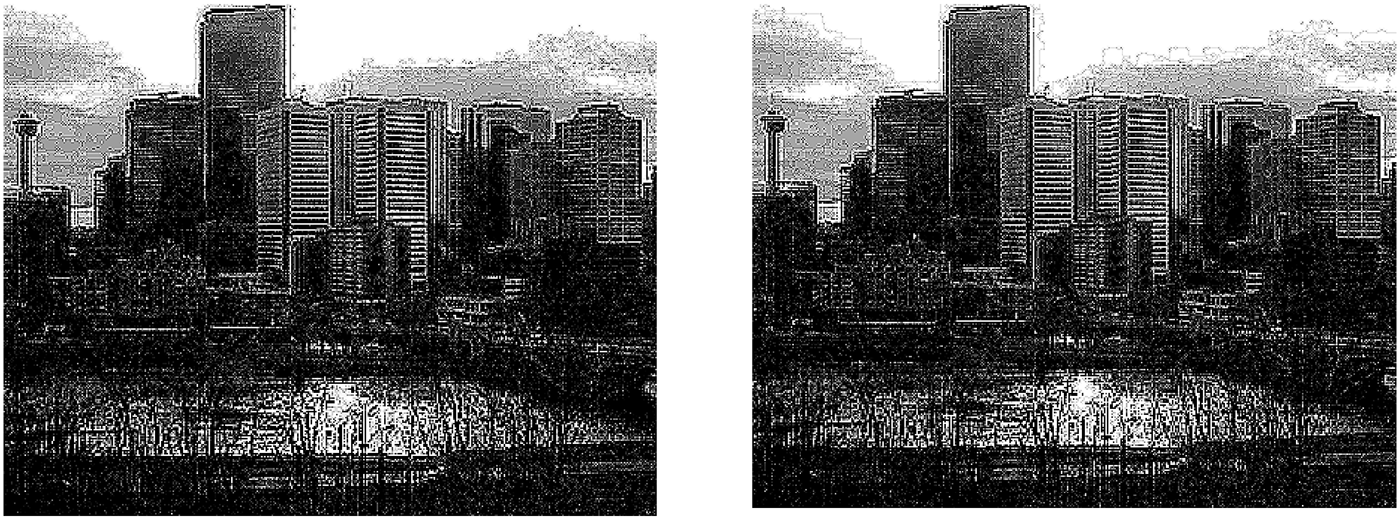

Fig 4.1
Edge enhanced image of the Calgary skyline reconstructed from only the low frequency subband using optimal QMF bank of length 6 (left) and Daub 6 (right)

As it can be seen in Fig. 4.1 the image on the right suffers considerably from the ringing artifact (the right side of the tall building and the clouds). The image on the left does not have such a well-pronounced artifact. Table 4.2 shows the distortion for the Lena when the EZW algorithm used the entire wavelet decomposition, and when it used only the low-frequency subband of the optimal transform. The optimal filters were used only for the first level decomposition. The main (EZW) algorithm used Duabechies' Maxflat filters of length six.

Table 4.2 Distortion for the Lena image at 100:1 compression ratio

Technique Used	MSE	PSNR
EZW	330.93	22.93
EZW with preprocessing[*]	319.82	23.08

*Preprocessing means EZW is applied to the low-frequency subband of the wavelet transform instead of being applied to the original image

As it can be seen the distortion is comparable (there is some increase in PSNR when our technique was used). The overall execution time, however, was reduced 3.7 times compared with the EZW applied to the entire decomposition.

However, at lower compression ratios (30:1) our technique introduces more distortion compared with the coding algorithm applied to the entire image. When lower compression ratios can be tolerated, (a high bandwidth available, or transmission time is not crucial), the DCT-based JPEG is fast and introduces low distortion at low compression rates.

5. Conclusion

We have developed and tested an algorithm that significantly reduces the execution times for lossy image compression algorithms. The reduction in the execution time is achieved by processing only the low-frequency subband of the wavelet decomposition. A novel interior-point -based algorithm for finding globally optimal PR-QMF two-channel filter banks was introduced. To our knowledge this is the first time an interior-point algorithm is used to find an optimal PR-QMF filter. It is important to emphasize that our method does not make any assumptions on the autocovariance sequence of an image. Given any image the algorithm finds a short filter that compacts most of the energy in the low-frequency subband of the image's wavelet decomposition in real time. Only the low-frequency subband is then processed by the coder. This reduction in the amount of data to be processed results in a significantly lower execution time. When the reduction in the data set is sufficient to fit the data into the internal memory of a particular processor, an additional reduction in the execution time is achieved due to a decrease in the access time for each element of the input data set. Our algorithm also allows for fast and more accurate previews of images, when they are transmitted via the internet. Our technique allows image processing application to be implemented in lower-cost devices (primarily DSP's), since it reduces the requirement for a large amount of internal memory, while meeting the required execution time specifications. Our current investigation is also done using biorthogonal filter banks to increase the active feasible region for optimization.

6. References

1. J. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients", *IEEE Transactions on Signal Processing,* VOL. 41, NO. 12, December 1993.
2. S. Mitra, and Shu-Yu Yang, "High Fidelity Adaptive Vector Quantization at Very Low Bit Rates for Progressive Transmission of Radiographic Images", *Journal of Electronic Imaging,* Vol.8, No. 1, pp 23-35, 1999.
3. M. Antonini, M. Barland, P. Mathiu, and I. Daubechies,"Image Coding Using Wavelet Transforms", *IEEE Transactions on Image Processing,* 1(2): 205-220, 1992.
4. V. Kustov, P. Srinivasan, S Mitra, and D. Mehrl " Performance of Wavelet-Based Compression in TI DSP's", Presentation at TI DSPS FEST, Houston, August 1999.
5. P. Moulin and M. Mihcak, "Theory and Design of Signal-Adapted FIR Paraunitary Filter Banks", *IEEE Transactions on Signal Processing,* VOL. 46, NO. 4, April 1998.
6. B. Xuan and R. Bamberger, "FIR Principal Component Filter Banks", *IEEE Transactions on Signal Processing,* VOL. 46, NO. 4, April 1998.
7. T. Farell and G. Prescott, "A Method for Finding Orthogonal Wavelet Filters with Good Energy Tiling Characteristics ", *IEEE Transactions on Signal Processing,* VOL. 47, NO. 1, January 1999.
8. Y. Ye, "*Interior Point Algorithms: Theory and Analysis*", John Wiley & Sons Inc., 1997.
9. G. Strang, T. Nguyen, "*Wavelets and Filter Banks*", Wellesley-Cambridge Press 1997.
10. M Vetterli and J. Kovacechic, "*Wavelets and Subband Coding*", Prentice Hall, 1995.
11. V. Silva and L. De Sa, "Analytical Optimization of QMF Filter Banks ", *IEEE Transactions on Signal Processing,* VOL. 44, NO. 6, June 1996.