SIGNIFICANCE TREE IMAGE CODING USING BALANCED MULTIWAVELETS

Claudio Weidmann, Jérôme Lebrun and Martin Vetterli

Dept. of Electrical Engineering
Swiss Federal Institute of Technology, Lausanne
{weidmann,lebrun,vetterli}@de.epfl.ch

ABSTRACT

Biorthogonal wavelets have been used with great success in most of the recent transform image coders. By using the new *balanced* multiwavelets, one can now easily design fully *orthogonal* linear phase FIR transform schemes. The aim of our work is to assess whether the added orthogonality yields a performance gain compared to traditional biorthogonal transforms. As comparison platform we use the well-known SPIHT codec, which is based on the significance tree quantization (STQ) principle. Without any particular fine-tuning the multiwavelet codec performs within 0.5 dB of SPIHT. A closer inspection shows however that it is hard to improve on this, therefore re-establishing the rule of thumb that strict orthogonality is not a key factor in image transform coding. More details can be obtained on the [WEB] at `http://lcavwww.epfl.ch/~weidmann/mwcoder`

1. INTRODUCTION

In this paper, we want to study the application of the recently presented *balanced* multiwavelets [6, 7, 8] to image transform coding. Wavelet based methods, as pioneered in [1], have become a standard and successful way of implementing transform coding. The underlying filter banks are now well studied, and thus the design procedure is well understood. By the structure of the problem, certain issues are ruled out: e.g. impossibility of constructing orthogonal FIR linear phase filter banks. This is a serious drawback since in many applications, and especially image coding, the following three properties are important: (1) FIR for obvious computational reasons, (2) linear phase to work on finite length signals without redundancy and artifacts, and (3) orthogonality as a necessary condition for the decorrelation of subband coefficients. However, by relaxing the time-invariance constraint, it has been proved recently that new solutions are possible: with multiwavelets, one is finally able to construct [4] an orthogonal linear phase FIR transform system.

Most lossy transform coders can be split into three distinct stages: transform, quantization and entropy coding. While some of these might be combined, their separation not only helps the implementation, but it also enables a clean analysis of the performance impact of different design choices for each stage. In the light of this, we thus decided to modify an existing transform coder (SPIHT [9]) — based on the classic 9/7 biorthogonal wavelet — by replacing the transform stage with one based on *balanced* multiwavelets, that are specially designed for signal compression. Therefore we get a quite fair comparison of the compression performance of the two *wavelets*, since the complexity of the other stages remains the same.

The organization of this paper is as follows: first we give a short introduction to the theory and construction of balanced multiwavelets. Then we briefly describe the SPIHT coder, which we used as our comparison platform. We also overview some implementation details which are peculiar to the multiwavelet (MW) transform. Finally, we discuss the achieved results.

2. MULTIWAVELETS

Generalizing the wavelet case, one can allow a multiresolution analysis $\{V_n\}_{n\in\mathbb{Z}}$ of $L^2(\mathbb{R})$ to be generated by a finite number of scaling functions $\phi_0(t), \phi_1(t), \ldots, \phi_{r-1}(t)$ and their integer translates. Then, the multiscaling function $\boldsymbol{\phi}(t) := [\phi_0(t), \ldots, \phi_{r-1}(t)]^\top$ verifies a 2-scale equation

$$\boldsymbol{\phi}(t) = \sum_k \mathbf{M}[k]\boldsymbol{\phi}(2t-k) \tag{1}$$

where now $\{\mathbf{M}[k]\}_k$ is a sequence of $r \times r$ matrices of real coefficients. The multiresolution analysis structure gives $V_1 = V_0 \oplus W_0$ where W_0 is the orthogonal complement of V_0 in V_1. We can construct an orthonormal basis of W_0 generated by $\psi_0(t), \psi_1(t), \ldots, \psi_{r-1}(t)$ and their integer translates with $\boldsymbol{\psi}(t) := [\psi_0(t), \ldots, \psi_{r-1}(t)]^\top$ derived by

$$\boldsymbol{\psi}(t) := \sum_k \mathbf{N}[k]\boldsymbol{\phi}(2t-k) \tag{2}$$

where $\{\mathbf{N}[k]\}_k$ is a sequence of $r \times r$ matrices of real coefficients obtained by completion of $\{\mathbf{M}[k]\}_k$. Introducing the refinement masks $\mathbf{M}(z) := \frac{1}{2}\sum_n \mathbf{M}[n]z^{-n}$ and $\mathbf{N}(z) := \frac{1}{2}\sum_n \mathbf{N}[n]z^{-n}$, the equations (1) and (2) translate in Fourier domain into

$$\Phi(2\omega) = \mathbf{M}(e^{j\omega})\Phi(\omega) \quad \text{and} \quad \Psi(2\omega) = \mathbf{N}(e^{j\omega})\Phi(\omega). \tag{3}$$

For simplicity and without loss of generality, we restrict ourselves to the case $r = 2$. Furthermore, we will assume that the sequences $\{\mathbf{M}[k]\}_k$ and $\{\mathbf{N}[k]\}_k$ are finite and thus that $\boldsymbol{\phi}(t)$ and $\boldsymbol{\psi}(t)$ have compact support. We assume also that $\mathbf{M}(z)$ verifies a matrix Smith-Barnwell orthogonality condition [6] so that the scaling functions and their integer

0-8186-8821-1/98 $10.00

translates form an orthonormal basis of V_0. Thus, for $s(t) \in V_0$, we have

$$s(t) = \sum_n \boldsymbol{s}_0^\top[n]\boldsymbol{\phi}(t-n) \tag{4}$$

then from $V_0 = V_{-1} \oplus W_{-1}$, we get

$$s(t) = \sum_n \boldsymbol{s}_{-1}^\top[n]\boldsymbol{\phi}(\frac{t}{2}-n) + \boldsymbol{d}_{-1}^\top[n]\boldsymbol{\psi}(\frac{t}{2}-n) \tag{5}$$

and we derive the well known relations between the coefficients at the analysis step

$$\boldsymbol{s}_{-1}[n] = \sum_k \mathbf{M}[k-2n]\boldsymbol{s}_0[k] \tag{6}$$

$$\boldsymbol{d}_{-1}[n] = \sum_k \mathbf{N}[k-2n]\boldsymbol{s}_0[k] \tag{7}$$

and for the synthesis, we get

$$\boldsymbol{s}_0[n] = \sum_k \mathbf{M}^\top[n-2k]\boldsymbol{s}_{-1}[k] + \mathbf{N}^\top[n-2k]\boldsymbol{d}_{-1}[k] \tag{8}$$

These relations enable us to construct a multi-input multi-output filter bank (multifilter bank [7]). In case of a one-dimensional signal, it then requires vectorization of this input signal to produce an input signal which is 2-dimensional. A simple way to do that is to split a one-dimensional signal into its polyphase components. Introducing

$$\begin{bmatrix} m_0(z) \\ m_1(z) \end{bmatrix} := \mathbf{M}(z^2)\begin{bmatrix} 1 \\ z^{-1} \end{bmatrix} \tag{9}$$

and in the same way $n_0(z)$ and $n_1(z)$, the system can then be seen as a 4 channel time-varying filter bank (Fig. 1).

In [6, 7, 8], we showed that if the components $m_0(z)$ and $m_1(z)$ of the lowpass branch have different spectral behavior, e.g. lowpass behavior for one, highpass for the other, it then leads to unbalanced channels that mix the coarse resolution and details coefficients and create strong oscillations. One expect then some class of smooth signals to be preserved by the lowpass branch and cancelled by the highpass. Therefore, we define a multiwavelet system to be *balanced* of order p if the polynomial structure of an input signal is captured up to degree $p-1$ by the lowpass branch coefficients.

Using equivalence results proved in [8], we have designed a *Daubechies like* family of balanced multiwavelets (named BAT). With a Gröbner basis approach [5], we have been able to construct all the multiwavelets of minimal compact support with flipped scaling functions (i.e. $m_1(z) = z^{-2L+1}m_0(z^{-1})$) and symmetric / antisymmetric wavelets for order 1, 2 and 3 of balancing. Fig. 2 and 3 show examples of high order balanced multiwavelets with these properties. Fig. 4 gives an example of another family (inspired by the DGHM multiwavelet [4]) of fully symmetric balanced orthogonal multiwavelets constructed by Selesnick in [10].

3. SPIHT WITH MULTIWAVELETS

This section gives a very short overview of SPIHT and then details some implementation aspects that are peculiar to multiwavelet transforms.

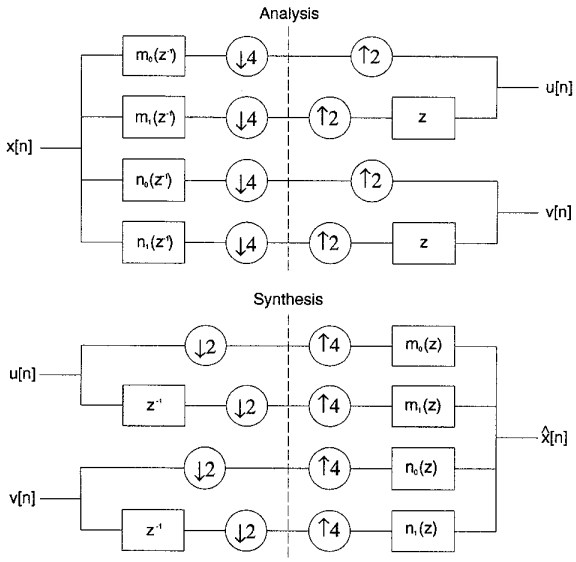

Figure 1: Multifilter bank seen as a time-varying filter bank.

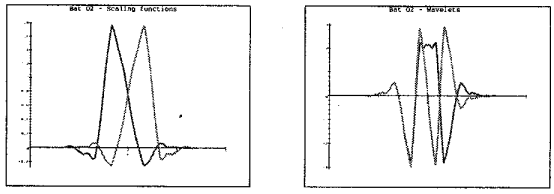

Figure 2: Order 2 balanced orthogonal (BAT) multiwavelet: the scaling functions are flipped around 2, the wavelets are symmetric/antisymmetric, the length of the filters is 8 taps.

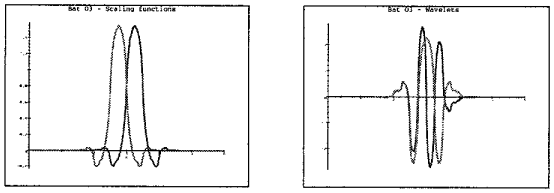

Figure 3: Order 3 balanced orthogonal (BAT) multiwavelet: the scaling functions are flipped around 3, the wavelets are symmetric/antisymmetric, the length of the filters is 12 taps.

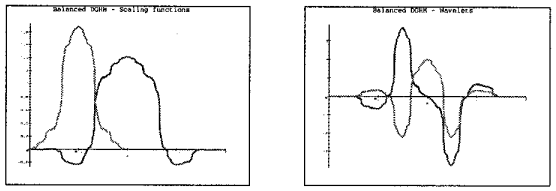

Figure 4: Order 1 balanced DGHM multiwavelet: the scaling functions and wavelets are symmetric/antisymmetric, the length of the filters are resp. 3,7,7,7 taps.

3.1. SPIHT

SPIHT belongs to a class of embedded transform coders which originated with Shapiro's embedded zerotree wavelet scheme (EZW [11]). Here, we describe only its main operating principles. First, the image is transformed into its dyadic (pyramid) wavelet decomposition. The coefficients are then assigned to the nodes of a hierarchical tree, such that the coefficients in the parent nodes are *representative* (in terms of energy) for their offspring nodes. The coder output is generated by successively refining the (quantized) coefficients in order of decreasing magnitude (concept of *significance* [3]). This can be done very efficiently, since the tree structure exploits the self-similarity across scales of the wavelet transform. Finally, an arithmetic coding stage removes the remaining redundancy.

3.2. Multiwavelet Transform

Since multiwavelets are defined for vector-valued signals, one might be tempted to *vectorize* an image signal by grouping pairs of rows or columns together. But besides introducing a fundamental asymmetry, this approach also doesn't fit the notion that the LL-subband represents a coarse approximation of the original image.[1] The latter is a key to the performance of SPIHT, and consequently we decided that no vectorization should be used. This is in fact possible by viewing the MW transform as a time-varying filter bank (see Fig. 1). The coefficients of the two lowpass (highpass) filters are simply interleaved at the output, e.g. in the one-dimensional case we get the following lowpass signal: $[\ldots, L_0[0], L_1[0], L_0[4], L_1[4], \ldots]$. A separable 2-D transform can now be defined in the usual way as the tensor product of two 1-D transforms. Obviously this approach is symmetric in the coordinates, i.e. every combination of horizontal and vertical filters appears at the output. But now we get 16 subbands, instead of the usual 4 with scalar wavelet transforms. The question is how this fits into the pyramid decomposition scheme required by SPIHT.

Let us first observe that the lowpass image, corresponding to the classic LL subband, contains the four lowpass-only bands and will be composed of 2x2 blocks as follows:

$$\begin{array}{ll} L_0L_0[x,y] & L_0L_1[x,y] \\ L_1L_0[x,y] & L_1L_1[x,y] \end{array}$$

Since this lowpass image undergoes further decomposition, we have to verify that it is a coarse approximation of the original. And it is indeed: as we can see in Fig. 3, the two lowpass filters (scaling functions) are flipped version of each other (mutual symmetry) with a group delay of 2. Thus our lowpass image closely resembles the output of a classic biorthogonal wavelet transform. Since the compression performance depends critically on this approximation behavior, we can say that our experimental results clearly confirm these arguments. This no surprise since these arguments were the starting point of the balancing concept (i.e. the preservation/cancelation of discrete-time polynomial signals by the lowpass/highpass branches of the time-varying filter bank in Fig. 1).

[1] The main hurdle is that SPIHT cannot handle vector-valued coefficients without undergoing a major revision of the partitioning and quantization mechanisms.

Another implementation aspect which differentiates multiwavelets from their biorthogonal cousins is the handling of boundary conditions. That is, the extension of a finite support input signal such that the transform coefficients have specific symmetry properties. The goal is to get a non-expansive transform, i.e. input and output dimensionality should be the same. For multirate linear phase FIR filter banks this problem has been treated extensively in [2]. But our lowpass filters are only mutually symmetric and thus we had to develop specific methods of signal extension. In the following we limit ourselves to 1-D transforms, as the extension to 2-D is implicit for separable transforms. Further we assume that the signal has length N, a multiple of 4, and that N is larger than the filter size K (to avoid wrap-around conditions).

With an N-periodic signal extension the coefficients obey the trivial symmetry condition $C[i] = C[i \bmod N]$. But this simple periodic approach may introduce discontinuities and hence add energy to the high frequency bands. Therefore it should be avoided in coding applications, as confirmed by our compression tests. The method of choice is a *symmetric* extension (with period $2N$), whereby the signal $[x_0, \ldots, x_{N-1}]$ is extended to $[\ldots, x_1, x_0, x_0, x_1, \ldots x_{N-2}, x_{N-1}, x_{N-1}, x_{N-2}, \ldots]$ The boundary points have to be doubled due to the even filter support K. To make sure that only N coefficients are needed to reconstruct $[x_0, \ldots, x_{N-1}]$, we have to exploit the following filter symmetries:

$$\begin{aligned} m_0[i] &= m_1[K-1-i] \\ n_0[i] &= n_0[K-1-i] \\ n_1[i] &= -n_1[K-1-i] \end{aligned}$$

where $i = 0, \ldots, K-1$. These equations yield different symmetries for the coefficients $C[n] = \sum_{i=0}^{K-1} h[i]x[n-i-\nu]$, depending on the *filter shift* ν (h stands for any of the four filters). The natural choice $\nu = K/2$ does not work, since then there are no symmetries to compute the coefficients $C[N]$. We have determined the values $\nu = 0$ for order 1 balancing (with $K = 4$), $\nu = 2$ for order 2 (with $K = 8$), and $\nu = 4$ for order 3 ($K = 12$). Then one gets the following coefficient symmetries:

$$\begin{array}{ll} L_0[-i] = L_1[i-4] & L_0[N+i] = L_1[N-i-4] \\ L_1[-i] = L_0[i-4] & L_1[N+i] = L_0[N-i-4] \\ H_0[-i] = H_0[i-4] & H_0[N+i] = H_0[N-i-4] \\ H_1[-i] = -H_1[i-4] & H_1[N+i] = -H_1[N-i-4] \end{array}$$

for $i = 0, 4, 8, \ldots, 4\lfloor N/8 \rfloor$. Hence N coefficients suffice to represent the input signal.

3.3. Results and discussion

Fig. 5 shows that order 2 balanced multiwavelets achieve fairly good results with PSNR values within 0.5 dB of the original SPIHT with biorthogonal 9/7-tap filters. This is remarkable, since the significance tree and arithmetic compression stages have not been fine-tuned to match the time-varying nature of the MW transform. It also means an average of 5 dB improvement in PSNR for "Lena" compared to previous multiwavelet based image coders (DGHM multiwavelet with pre/post filtering [12]). This proves the superiority of the balanced multiwavelet approach over MW

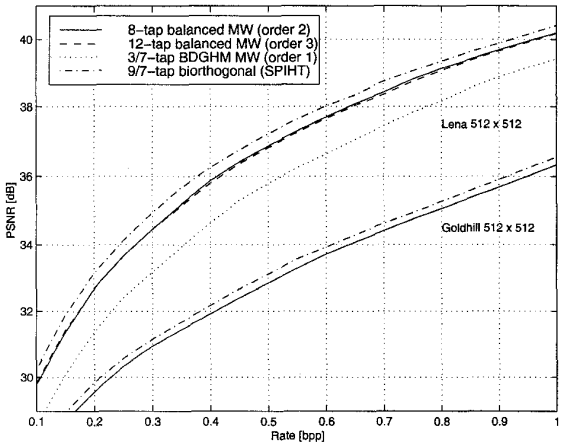

Figure 5: Multiwavelet performance comparison

systems requiring pre/post filtering of the input data because of unbalanced lowpass filters $m_0(z), m_1(z)$. Namely, with balancing one is able to take full advantage of the interesting properties of multiwavelet systems. The balancing order 3 multiwavelets are slightly better than order 2 at rates below 0.2 bpp, above that, they are slightly worse. We also implemented a coder with the new balanced DGHM multiwavelets first introduced in [10]. The bad performance can partly be explained by the fact that there is no easy way to make the transform non-expansive. We had to interpolate a coefficient to get that property. While the formula is exact for real coefficient values, the distortion introduced by the quantization could have been amplified.

On a more subjective level, a visual comparison (Fig. 6 and 7) reveals a disturbing tiling effect for balancing order 2 multiwavelets. If the order 3 MW is used instead, the more familiar ringing artifacts almost "cover up" these tiling effects. We suspect two main reasons for these tiling effects: first, the two lowpass filters introduce a 0.5 pixel phase shift at each iteration, due to their structure and the implementation constraints (border symmetrization). Thus high energy coefficients at an image discontinuity will be less well aligned across scales, hampering the efficiency of the tree prediction as illustrated in Fig. 8. Secondly, the two highpass filters have very different spectral characteristics, and therefore their outputs should be treated separately in the tree/entropy coding stages. Possibly some smoothing filter could be used to lessen the disturbing tiling effect.

Given the relatively small performance gap, the question is whether balanced multiwavelets could outperform biorthogonal wavelets. Fig. 9 shows that the MW transform produces more high amplitude coefficients then the 9/7-tap wavelets (at threshold 8, about 18(20) % of all coefficients are significant). This is a strong indication that the MW transform is not suited for low bitrate compression. Every improvement in successive stages (tree, entropy coding) could as well be applied to biorthogonal wavelets. It is therefore the transform itself, or rather its implementation, which would have to be improved. But this proves difficult, since e.g. the constraints imposed by the down-

Figure 6: Lena at 0.1 bpp, from top to bottom: 9/7-tap biorthogonal (30.23 dB PSNR), 8-tap multiwavelet (balancing order 2, 29.78 dB), 12-tap multiwavelet (balancing order 3, 29.83 dB)

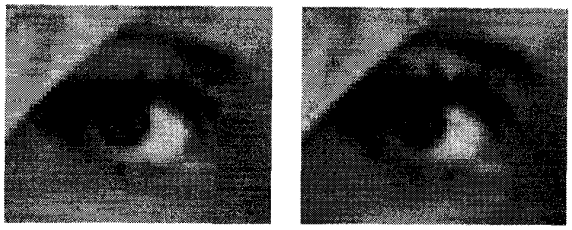

Figure 7: Detail of Lena at 0.25 bpp: 8-tap multiwavelet transform (left), original SPIHT with 9/7-tap filters (right)

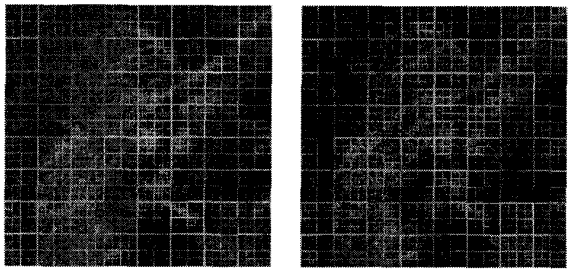

Figure 8: Tree decomposition of the horizontal subband magnitudes: 9/7-tap biorthogonal wavelet (left), 8-tap multiwavelet (right). The latter has less well-localized high-energy coefficients.

sampling factor and the border symmetry conditions imply that the mentioned lowpass phase shift cannot be avoided. On the other hand, it is possible to construct a highpass filter pair with flip-symmetry as in the lowpass branch [WEB]. However the coding results are worse than for the symmetric/antisymmetric highpass construction presented here.

4. CONCLUSION

In this paper, we implemented a *balanced* multiwavelet transform for a significance tree quantization image coder (namely SPIHT). With the introduction of the *balancing* concept, it is possible to design general families of high order balanced multiwavelets with the properties required for practical signal processing (preservation/cancelation of discrete-time polynomial signals in the lowpass/highpass subbands, FIR, linear phase and orthogonality). The results obtained so far show that multiwavelets are finally establishing themselves as a convincing alternative in DSP tools.

On the other hand, our results are also experimental evidence for the well-known fact that strict orthogonality plays a minor role in image transform coding. Design parameters such as filter length, smoothness and regularity have a much heavier impact on performance. The design of non-expansive transforms, which are essential for image coding, is harder in the multiwavelet case. In conclusion, the multiwavelets known so far are still no plug-in replacements for the more traditional scalar wavelets.

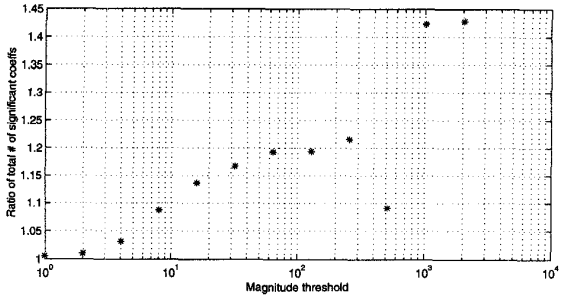

Figure 9: Ratio of the numbers of significant coefficients for 8-tap multiwavelets and 9/7-tap biorthogonal wavelets (Lena image). The multiwavelet transform has over 10% more significant coefficients down to threshold 8, corresponding to about 0.7 bpp.

5. REFERENCES

[1] M.Antonini, M.Barlaud, P.Mathieu and I.Daubechies, "Image coding using wavelet transform", IEEE Trans. on Image Processing, Vol. 1, pp. 205-220, 1992.

[2] C.M.Brislawn, "Preservation of subband symmetry in multirate signal coding", IEEE Trans. on SP, Vol. 43, pp. 3046-3050, 1995.

[3] G.Davis and S.Chawla, "Image coding using optimized significance tree quantization", Proc. DCC'97.

[4] G.Donovan, J.Geronimo, D.Hardin and P.Massopust, "Construction of orthogonal wavelets using fractal interpolation functions", SIAM J. Math. Anal., 1996.

[5] G.-M.Greuel, G.Pfister, and H.Schönemann, "Singular reference manual", University of Kaiserslautern, 1998. `http://www.mathematik.uni-kl.de/~zca/Singular`

[6] J.Lebrun and M.Vetterli, "Balanced multiwavelets", Proc. ICASSP-97.

[7] J.Lebrun and M.Vetterli, "Balanced multiwavelets theory and design", IEEE Trans. on SP, April 1998.

[8] J.Lebrun and M.Vetterli, "High order balanced multiwavelets", Proc. ICASSP-98.

[9] A.Said and W.A.Pearlman, "A new fast and efficient image codec on set partitioning in hierarchical trees", IEEE Trans. on CSVT, Vol. 6, pp. 243-250, June 1996.

[10] I.W.Selesnick, "Multiwavelet bases with extra approximation properties", Proc. 8th IEEE DSP Workshop, Utah, 1998.

[11] J.Shapiro, "Embedded image coding using zerotrees of wavelet coefficients" IEEE Trans. on SP, Vol. 41, pp. 3445-3462, 1993.

[12] V.Strela, P.Heller, G.Strang, P.Topiwala and C.Heil, "The application of multiwavelet filter banks to signal and image processing", IEEE Trans. on IP, to appear, 1998.