Boundary Filters for Segmentation-Based Subband Coding

Alfred Mertins

University of Wollongong
School of Electrical, Computer, and Telecommunications Engineering
Wollongong, NSW 2522, Australia,
Phone +61 2 4221 3410, Fax +61 2 4221 3236, Email: mertins@uow.edu.au

Abstract

This paper presents design methods for boundary filters for the processing of finite-length signals. In particular, it is shown how vanishing moments can be imposed during the design. Applications are found in segmentation-based audio and shape-adaptive image compression.

1 Introduction

Multirate filter banks are widely used in audio and image compression. In segmentation-based audio coding, the input signal is divided into finite-length blocks and each block is encoded separately. This strategy allows one to easily adapt the bit allocation to different segments of the input signal and to directly access parts of the encoded bitstream. Also, by placing the segmentation right in front of attacks one can avoid the problem of pre-echoes [1]. To achieve a high compression ratio, the decomposition scheme should be non-expansive, which means that the total number of subband samples produced from a segment has to be equal to the number of input samples in that segment.

In image compression, multirate decompositions of arbitrarily shaped regions are of significant importance, because they allow object-based rather than frame-based data access. As in the decomposition of one-dimensional (1-D) signals, the problem of achieving non-expansive subband decompositions occurs. Because the objects describing a scene may occur with any shape, the transforms used for compression need to be able to operate on arbitrarily shaped regions of support while maintaining a high compaction gain.

The aim of this paper is to present novel solutions to the problem of optimizing the boundary filters for non-expansive subband transforms. In particular, we address the problem of imposing moment conditions on the boundary filters. The methods presented are not restricted to certain signal lengths or filter characteristics.

2 Construction of Boundary Filters

The use of boundary filters for achieving support-preservative subband decompositions of finite-length signals has been proposed in [2–8]. Unlike circular convolution [9] or symmetric reflection [10], boundary filters are not restricted to particular signal lengths or filter properties such as linear phase.

2.1 General Considerations

The subband decomposition of a length-N signal $\boldsymbol{x}$ can be written as

$$\boldsymbol{y} = \boldsymbol{H}\,\boldsymbol{x} \tag{1}$$

where the matrix $\boldsymbol{H}$ describes the convolution of the input signal with the analysis filters and the down-sampling operation. The vector $\boldsymbol{y}$ contains the subband samples. The synthesis operation can be written as

$$\hat{\boldsymbol{x}} = \boldsymbol{G}\,\boldsymbol{y}. \tag{2}$$

The filter operations can be divided into three parts, namely the processing of the left and right boundaries with boundary filters and the processing of the interior of the signal with the original filters. Correspondingly, the matrices can be partitioned as follows:

$$\begin{aligned} \boldsymbol{H} &= \left[\boldsymbol{H}_1^T | \boldsymbol{H}_2^T | \boldsymbol{H}_3^T\right]^T \\ \boldsymbol{G} &= \left[\boldsymbol{G}_1 | \boldsymbol{G}_2 | \boldsymbol{G}_3\right], \end{aligned} \tag{3}$$

where $\boldsymbol{H}_1$ and $\boldsymbol{G}_1$ contain the boundary filters for the left-hand and $\boldsymbol{H}_3$ and $\boldsymbol{G}_3$ contain the ones for the right-hand side. For the analysis and synthesis operations in partitioned form we get

$$\boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x}, \qquad \hat{\boldsymbol{x}} = \sum_{k=1}^{3} \boldsymbol{G}_k \boldsymbol{y}_k, \tag{4}$$

where vectors $\boldsymbol{y}_k$ are the corresponding partitions of $\boldsymbol{y}$.

If the PR condition $\boldsymbol{G}\boldsymbol{H} = \boldsymbol{H}\boldsymbol{G} = \boldsymbol{I}$ is satisfied, the submatrices satisfy $\boldsymbol{H}_k \boldsymbol{G}_k = \boldsymbol{I}$, where $\boldsymbol{I}$ is an identity matrix of appropriate size. Terms of the form $\boldsymbol{G}_k\boldsymbol{H}_k$ describe projections (not necessarily orthogonal ones) onto the column spaces of $\boldsymbol{G}_k$. If we want to replace one of the matrices $\boldsymbol{G}_k$ by a new (better) matrix $\tilde{\boldsymbol{G}}_k$, we also need to replace the corresponding analysis partition $\boldsymbol{H}_k$ by a new matrix $\tilde{\boldsymbol{H}}_k$. The matrices have to satisfy $\tilde{\boldsymbol{H}}_k\tilde{\boldsymbol{G}}_k = \boldsymbol{I}$ and $\tilde{\boldsymbol{G}}_k\tilde{\boldsymbol{H}}_k = \boldsymbol{G}_k\boldsymbol{H}_k$. Hence, we see that both $\boldsymbol{G}_k$ and $\tilde{\boldsymbol{G}}_k$ must have the same column space and that the columns of $\tilde{\boldsymbol{G}}_k$ can be written as a linear combination of those of $\boldsymbol{G}_k$. With invertible quadratic matrices $\boldsymbol{U}_k$ this can be expressed as $\tilde{\boldsymbol{G}}_k = \boldsymbol{G}_k\boldsymbol{U}_k^{-1}$. For the analysis side, it means that $\tilde{\boldsymbol{H}}_k = \boldsymbol{U}_k\boldsymbol{H}_k$. One would want to manipulate only the boundary filters, so that the modified analysis and synthesis equations may be written as

$$\boldsymbol{v}_k = \tilde{\boldsymbol{H}}_k\boldsymbol{x}, \qquad \hat{\boldsymbol{x}} = \sum_{k=1}^{3}\tilde{\boldsymbol{G}}_k\boldsymbol{v}_k \tag{5}$$

with $\boldsymbol{U}_2 = \boldsymbol{I}$. The optimization of the boundary filters reduces to the optimization of $\boldsymbol{U}_1$ and $\boldsymbol{U}_3$. All invertible matrices $\boldsymbol{U}_1$ and $\boldsymbol{U}_3$ satisfy the PR constraints. However, if paraunitaryness is to be maintained, $\boldsymbol{U}_1$ and $\boldsymbol{U}_3$ have to be orthogonal.

Initial solutions for the matrices $\boldsymbol{H}_1$, $\boldsymbol{H}_3$, $\boldsymbol{G}_1$, and $\boldsymbol{G}_3$ that guarantee PR can be found via the Gram-Schmidt procedure as outlined in [3]. To explain this, we consider a filter bank analysis described as $\boldsymbol{y} = \boldsymbol{F}\,\boldsymbol{x}_E$, where the vector $\boldsymbol{y}$ contains the subband samples that are to be computed. The rows of $\boldsymbol{F}$ contain non-truncated, time-shifted and flipped versions of the analysis filters' impulse responses. $\boldsymbol{F}$ is rectangular and the length of $\boldsymbol{x}_E$ is larger than the length of $\boldsymbol{y}$. The next step is to truncate $\boldsymbol{F}$ to an $N \times N$ matrix. Fig. 1 illustrates the truncation for a two-band decomposition of a length-8 signal. Given the truncated matrix, the method in [3] can be applied to design the required sub-matrices $\boldsymbol{H}_k$ and $\boldsymbol{G}_k$. The matrices $\boldsymbol{H}_2$ and $\boldsymbol{G}_2$ still contain the original impulse responses. The design method can be applied to both paraunitary and biorthogonal filter banks. The result of the Gram-Schmidt procedure is somewhat arbitrary, and one cannot expect to design boundary filters with good properties this way, so that further optimization is needed.

2.2 Vanishing Moments

When applying a filter bank to a finite-length signal by using boundary filters, the problem occurs that the boundary filters will usually not satisfy any moment conditions, even if the original filters do. In the following we will derive a method that enables us to restore vanishing moments for the boundary filters. Our free

$$\begin{bmatrix}
h_0(3) & h_0(2) & h_0(1) & h_0(0) & & & & & \\
h_1(3) & h_1(2) & h_1(1) & h_1(0) & & & & & \\
& & h_0(3) & h_0(2) & h_0(1) & h_0(0) & & & \\
& & h_1(3) & h_1(2) & h_1(1) & h_1(0) & & & \\
& & & & h_0(3) & h_0(2) & h_0(1) & h_0(0) \\
& & & & h_1(3) & h_1(2) & h_1(1) & h_1(0)
\end{bmatrix}$$

Fig. 1. Truncation of $\boldsymbol{F}$.

design parameters are the elements of the matrices $\boldsymbol{U}_1$ and $\boldsymbol{U}_3$, so that we need to find restrictions on these matrices that guarantee the desired moment properties. We formulate the requirements as

$$\underbrace{[\boldsymbol{v}_k^{(0)}, \boldsymbol{v}_k^{(1)}, \ldots, \boldsymbol{v}_k^{(I-1)}]}_{\boldsymbol{V}_k} = \boldsymbol{U}_k \underbrace{[\boldsymbol{y}_k^{(0)}, \boldsymbol{y}_k^{(1)}, \ldots, \boldsymbol{y}_k^{(I-1)}]}_{\boldsymbol{Y}_k} \tag{6}$$

where a vector $\boldsymbol{y}_k^{(i)} = \boldsymbol{H}_k\,\boldsymbol{t}^{(i)}$ is the response to a polynomial input signal $\boldsymbol{t}^{(i)} = \left[1^i, 2^i, \ldots, N^i\right]^T$, and where $\boldsymbol{v}_k^{(i)} = \boldsymbol{U}_k\boldsymbol{y}_k^{(i)}$ is the desired response to this signal. For example, in the case of a two-channel filter bank, a DC input signal, and four boundary filters in $\boldsymbol{H}_1$ it is desirable to have

$$\boldsymbol{v}_1^{(0)} := [\sqrt{2}, 0, \sqrt{2}, 0]^T := \boldsymbol{U}_1\,\boldsymbol{H}_1\,\boldsymbol{t}^{(0)},$$

where it is assumed that $\sum_n h_0(n) = \sqrt{2}$ and $\sum_n h_1(n) = 0$. Similarly, the desired responses $\boldsymbol{v}_k^{(i)}$ for $i > 0$ can be defined according to the properties of $h_0(n)$ and $h_1(n)$.

Given $\boldsymbol{V}_k$ and $\boldsymbol{Y}_k$ the matrix $\boldsymbol{U}_k$ can be described (parameterized) as follows:

$$\boldsymbol{U}_k = \boldsymbol{U}_k^o + \boldsymbol{P}_k\boldsymbol{N}_k^T \quad \text{with } \boldsymbol{U}_k^o = \boldsymbol{V}_k\boldsymbol{Y}_k^+, \tag{7}$$

where $\boldsymbol{N}_k$ contains a basis for the nullspace of $\boldsymbol{Y}_k$ such that $\boldsymbol{N}_k^T\boldsymbol{Y}_k = \mathbf{0}$. The matrix $\boldsymbol{Y}_k^+$ is the pseudo inverse of $\boldsymbol{Y}_k$, and $\boldsymbol{P}_k$ is an arbitrary matrix of appropriate size. If the number of conditions in (6), denoted as I, is small enough to ensure that $\boldsymbol{Y}_k^+\boldsymbol{Y}_k = \boldsymbol{I}$, the requirements (6) are fulfilled exactly. Provided that the nullspace contains more than just the null vector, the elements of $\boldsymbol{P}_k$ may be understood as design parameters, which can be freely chosen to optimize $\boldsymbol{U}_k$ according to other criteria. If I is so large that $\boldsymbol{Y}_k^+\boldsymbol{Y}_k \neq \boldsymbol{I}$, then (6) is approximated in the least squares sense, and there are no further free design parameters for optimization.

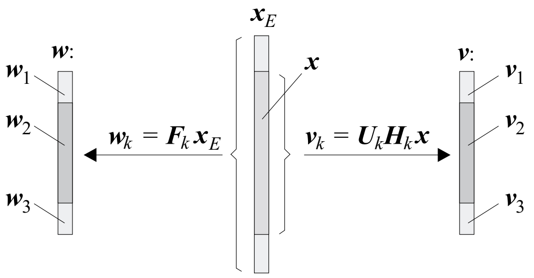

Fig. 2. Computation of $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ from $\boldsymbol{x}_E$.

Don't Care Values

The equation (6) is quite restrictive in the sense that the responses of all filters to the given input signals $\boldsymbol{t}^{(i)}$ have to be specified. To have greater design freedom, we may want to specify only a few (important ones) of these responses and leave the others as don't care values. This can be achieved by deleting certain rows of (6). For the specified responses this yields a parameterization of the form

$$\boldsymbol{U}_k^{\textit{spec.}} = \boldsymbol{U}_k^{o,\textit{spec.}} + \boldsymbol{P}_k^{\textit{spec.}} \boldsymbol{N}_k^T \tag{8}$$

The ones without specification are simply parameterized as

$$\boldsymbol{U}_k^{\textit{unspec.}} = \boldsymbol{P}_k^{\textit{unspec.}} \tag{9}$$

Interleaving $\boldsymbol{U}_k^{\textit{spec.}}$ and $\boldsymbol{U}_k^{\textit{unspec.}}$ then yields $\boldsymbol{U}_k$.

2.3 The Objective Function

We assume that a filter bank has been chosen for a given application (e.g. audio or image coding) because of its good properties. When applying the filter bank to a finite-length input signal, the boundary filters should have similar properties as the filters used in steady state. If this is the case, the same bit allocation can be used at the boundaries and in the interior of the signal, which is very much desirable from a practical point of view. Otherwise, a spatial adjustment of the bit allocation is needed to avoid effects like spatially varying noise when reconstructing the signal from its quantized subband coefficients [8]. In this section, we design boundary filters in such a way that the subband samples produced by these filters resemble (as much as possible) the ones that would be produced by the original filters.

We consider a length-N vector of subband samples computed as $\boldsymbol{w} = \boldsymbol{F}\boldsymbol{x}_E$ with $\boldsymbol{F}$ as in Section 2.1 and $\boldsymbol{x}_E$ being a stationary input process. The vector $\boldsymbol{x}_E$ is longer than $\boldsymbol{w}$ and the matrix $\boldsymbol{F}$ describes the filter bank analysis in steady state. We also consider a length-N part of $\boldsymbol{x}_E$, denoted as $\boldsymbol{x}$, and describe the analysis of $\boldsymbol{x}$ as $\boldsymbol{v} = \boldsymbol{U}\boldsymbol{H}\boldsymbol{x}$ where $\boldsymbol{v}$ has length N. The vector $\boldsymbol{x}$ can be written as $\boldsymbol{x} = \boldsymbol{C}\boldsymbol{x}_E$, so that $\boldsymbol{v} = \boldsymbol{U}\boldsymbol{H}\boldsymbol{C}\boldsymbol{x}_E$. The aim is to minimize the error measure $E\{\|\boldsymbol{v} - \boldsymbol{w}\|^2\}$ for given input statistics through the choice of $\boldsymbol{U}$. When partitioning the vectors into three parts, as explained earlier, we have to minimize

$$\phi(\boldsymbol{U}_k) = E\{\|\boldsymbol{v}_k - \boldsymbol{w}_k\|^2\}, \tag{10}$$

where $\boldsymbol{v}_k = \boldsymbol{U}_k\boldsymbol{H}_k\boldsymbol{C}\boldsymbol{x}_E$ and $\boldsymbol{w}_k = \boldsymbol{F}_k\boldsymbol{x}_E$ with $\boldsymbol{F} = [\boldsymbol{F}_1^T, \boldsymbol{F}_2^T, \boldsymbol{F}_3^T]^T$. Fig. 2 gives an illustration of the concept. The objective function finally becomes

$$\phi(\boldsymbol{U}_k) = \mathrm{tr}\{[\boldsymbol{U}_k\boldsymbol{H}_k\boldsymbol{C} - \boldsymbol{F}_k]\ \boldsymbol{R}_{xx}\ [\boldsymbol{U}_k\boldsymbol{H}_k\boldsymbol{C} - \boldsymbol{F}_k]^T\} \tag{11}$$

where $\boldsymbol{R}_{xx}$ is the autocorrelation matrix of $\boldsymbol{x}_E$.

For the case that no further restrictions are imposed on $\boldsymbol{U}_k$, it can be shown that the optimal matrix $\boldsymbol{U}_k$ is given by

$$\boldsymbol{U}_k = \boldsymbol{F}_k\,\boldsymbol{R}_{xx}\,\boldsymbol{C}^T\boldsymbol{H}_k^T\left[\boldsymbol{H}_k\boldsymbol{C}\,\boldsymbol{R}_{xx}\,\boldsymbol{C}^T\boldsymbol{H}_k^T\right]^{-1}. \tag{12}$$

To include moment conditions, we replace $\boldsymbol{U}_k$ in (11) by $\boldsymbol{U}_k = \boldsymbol{V}_k\boldsymbol{Y}_k^+ + \boldsymbol{P}_k\boldsymbol{Z}_k^T$. One can prove that the optimal matrix $\boldsymbol{P}_k$ then becomes

$$\begin{aligned}\boldsymbol{P}_k \quad = \quad & [\boldsymbol{F}_k - \boldsymbol{V}_k\boldsymbol{Y}_k^+\boldsymbol{H}_k\boldsymbol{C}]\,\boldsymbol{R}_{xx}\,\boldsymbol{C}^T\boldsymbol{H}_k^T\boldsymbol{N}_k \\ & \cdot\left[\boldsymbol{N}_k^T\boldsymbol{H}_k\boldsymbol{C}\,\boldsymbol{R}_{xx}\,\boldsymbol{C}^T\boldsymbol{H}_k^T\boldsymbol{N}_k\right]^{-1}\end{aligned} \tag{13}$$

For paraunitary filter banks where matrices $\boldsymbol{U}_k$ are demanded to be orthogonal, the optimal matrices can be found via a singular value decomposition (svd). This solution to the boundary filter optimization problem has first been presented in [4]. To explain the steps, let us decompose $\boldsymbol{R}_{xx}$ into $\boldsymbol{R}_{xx} = \boldsymbol{L}\boldsymbol{L}^T$. The problem may then be written as a subspace rotation problem [11]:

$$\text{minimize} \quad \|[\boldsymbol{U}_k\boldsymbol{H}_k\boldsymbol{C} - \boldsymbol{F}_k]\,\boldsymbol{L}\|_F^2 \quad \text{s.t.} \quad \boldsymbol{U}_k^T\boldsymbol{U}_k = \boldsymbol{I}. \tag{14}$$

The next step is to compute the svd $\mathbf{U\Sigma V}^T = [\boldsymbol{H}_k\boldsymbol{C}\boldsymbol{L}\boldsymbol{L}^T\boldsymbol{F}_k^T]^T$ and to derive the final solution as $\boldsymbol{U}_k = \mathbf{U}\mathbf{V}^T$.

If both orthonormality and vanishing moments are wanted, no direct solution is available, and the desired parameters have to be found via numerical optimization. This can e.g. be done by using the parameterization (7) to ensure the vanishing moments and adding the orthogonality constraints

$$\boldsymbol{U}_k^T\boldsymbol{U}_k = \boldsymbol{I}$$

to the objective function via the Lagrange multiplier technique.

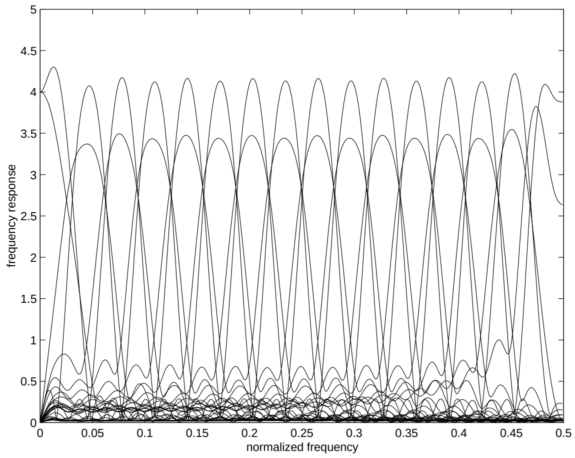

Fig. 3. Frequency responses of optimized boundary filters ($M = 16$ bands, 32 boundary filters).

2.4 Application in 2-D

In addition to simplifying the bit allocation in the boundary regions, the above optimization is also advantageous for 2-D shape adaptive schemes. In 2-D, the alignment of boundary filters operating on adjacent rows/columns that start or stop at different positions has to be taken into account. Filtering in one direction should cause minimal distortion in the second one. Thus, optimizing boundary filters for the 1-D case and using them in a 2-D shape adaptive scheme does not guarantee overall optimality. By properly aligning the filters in the interior of a region and optimizing the boundary filters according to the algorithm above, minimal distortion occurs in the second dimension.

3 Results

We consider a cosine-modulated 16-band filter bank with ELT prototype [12]. In this filter bank, the subband filters have non-linear phase, so that symmetric reflection techniques cannot be applied and boundary filters must be used. The frequency responses for the optimized boundary filters (left boundary) are depicted in Fig. 3. During optimization, a white noise input process has been considered. We see that the optimization yields filters with ideal behavior for DC signals, as demanded via Eq. (6). Also the frequency selectivity is good, so that that good coding properties can be expected.

4 Conclusions

In this paper, closed-form solutions for designing optimal boundary processing have been presented. All methods maintain critical sampling and are applicable to both paraunitary and biorthogonal filter banks. A design example has been presented for a cosine-modulated filter bank, as often used in audio compression, and it turned out that the designed boundary filters have a good frequency selectivity, so that good coding properties can be expected. The method can also be applied to design 2-D region-based wavelet transforms.

References

[1] J. Kliewer and A. Mertins, “Audio subband coding with improved representation of transient signal segments,” in *Proc. European Signal Processing Conference*, Rhodos, Greece, pp. 2345–2348, Sept. 1998.

[2] R. L. de Queiroz, “Subband processing of finite length signals without border distortions,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, San Francisco, USA, vol. IV, pp. 613 – 616, Mar. 1992.

[3] C. Herley, “Boundary filters for finite-length signals and time-varying filter banks,” *IEEE Trans. Circuits and Systems II*, vol. 42, pp. 102–114, Feb. 1995.

[4] A. Mertins, “Time-varying and support preservative filter banks: Design of optimal transition and boundary filters via SVD,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Detroit, pp. 1316–1319, May 1995.

[5] R. L. de Queiroz and K. R. Rao, “Optimal orthogonal boundary filter banks,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Detroit, USA, vol. II, pp. 1296 – 1299, May 1995.

[6] M. Coffey, “Boundary compensated wavelet bases,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, München, pp. 2129–2132, May 1997.

[7] A. Mertins, “Optimized biorthogonal shape adaptive wavelets,” in *Proc. IEEE Int. Conf. Image Processing*, Chicago, vol. 3, pp. 673–677, Oct. 1998.

[8] V. Nuri and R. H. Bamberger, “Size limited filter banks for subband image compression,” *IEEE Trans. Image Processing*, vol. 4, pp. 1317–1323, Sept. 1995.

[9] J. Woods and S. O’Neil, “Subband coding of images,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1278–1288, May 1986.

[10] L. Chen, T. Q. Nguyen, and K. P. Chan, “Symmetric extension methods for M-channel linear-phase perfect reconstruction filter banks,” *IEEE Trans. Signal Processing*, vol. SP-43, pp. 2505–2511, Nov. 1995.

[11] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore: John Hopkins University Press, 3rd ed. 1996.

[12] H. S. Malvar, “Extended lapped transforms: Fast algorithms and applications,” *IEEE Trans. Signal Processing*, vol. 40, Nov. 1992.