UNIFORM FILTER BANKS WITH NONUNIFORM BANDS: POST-PROCESSING DESIGN

Ricardo L. de Queiroz

Xerox Corporation
800 Phillips Rd, M/S 128-27E, Webster, NY 14580
queiroz@wrc.xerox.com

ABSTRACT

In this paper, uniform, critically decimated filter banks are used to approximate nonuniform filter banks wherein different filters have approximately the same magnitude response, but different phase, thus forming a linear periodically time-varying filter whose characteristics are similar to those of a nonuniform bank. This is done by post-processing a number of selected subbands of a uniform bank using a special synthesis filter bank, which combines the selected bands into one. Design methods for the post-processing stage are discussed and design examples are presented.

1. INTRODUCTION

Uniform filter banks are the most common form of subband decomposition systems [1]–[4]. In those, each filter output is critically decimated by the same factor M and the filters have about the same passband width. In a nonuniform filter bank, each filter output is decimated by a particular factor and, yet, it is possible to obtain perfect reconstruction [1]. Also, nonuniform filter banks can be obtained by cascading uniform filter banks as in the case of the discrete wavelet transform and wavelet packets [1]–[4]. Theory and design of nonuniform filter banks can be found in [1],[5]–[7]. Also, nonuniform cosine modulated filter banks were considered in [8]–[10]. The ability to construct nonuniform filter banks facilitates the trade-off of resolution between the two domains (spatial and frequency). We propose a new way to approach the problem, where the filter bank is inherently uniform. However, the filters' passbands can have different width, and different filters can have similar passbands.

We assume a reference uniform paraunitary filter bank having M real FIR filters with length $L = NM$. We also describe a filter bank through its polyphase transfer matrix (PTM), i.e. a multi-input multi-output (MIMO) system relating M polyphase components of the signal to M subbands [1]. The signal is blocked and passed through the analysis PTM $\mathbf{F}(z)$. It is reconstructed from the subbands using the PTM $\mathbf{G}^T(z)$ followed by an unblocking device. See [1]–[4] for details on filter banks, PTM, and paraunitary systems.

This paper contains some theorems whose proofs were omitted due to space limitations. Nevertheless, said proofs appear in a longer version of this paper.

2. MERGING BANDS

We propose to start from a uniform paraunitary filter bank, whose analysis PTM is $\mathbf{F}(z)$, and, by applying a post-processing stage $\mathbf{\Phi}(z)$ to a selected number of filters, to mix subbands together so that a filter passband will actually occupy the passband of a plurality of filters in the uniform design. Let the rows of the analysis PTM $\mathbf{F}(z)$ corresponding to K selected uniform filters be represented in the $K \times M$ PTM $\mathbf{U}(z)$. We want to find a PTM $\mathbf{S}(z)$ of same dimensions such that

$$\mathbf{S}(z) = \mathbf{\Phi}(z)\mathbf{U}(z). \tag{1}$$

Without loss of generality, we can rearrange the order of the filters in $\mathbf{F}(z)$ so that the K selected filters are displaced on the bottom of the matrix. If this is the case, we can devise a PTM $\mathbf{\Phi}'(z)$ such that

$$\mathbf{\Phi}'(z) = \begin{bmatrix} \mathbf{I}_{M-K} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi}(z) \end{bmatrix}. \tag{2}$$

Hence,

$$\mathbf{F}'(z) = \mathbf{\Phi}'(z)\mathbf{F}(z). \tag{3}$$

$\mathbf{F}'(z)$ becomes the actual analysis PTM. We assume $\mathbf{F}(z)$ to be paraunitary, while $\mathbf{F}'(z)$ and $\mathbf{\Phi}(z)$ are not required to be so. In case $\mathbf{\Phi}(z)$ (hence $\mathbf{F}'(z)$) is bi-orthogonal, we would like it to approximate a paraunitary system. We explore 4 methods to design $\mathbf{\Phi}(z)$.

3. APPROXIMATING THE MFR FILTER SET

In a critically decimated system, lower frequency resolution (localization) implies higher spatial resolution [1],[4]. We define the frequency resolution of a filter H as

$$\psi_F(H) = \frac{\int_0^\pi \omega^2 |H(e^{j\omega})|^2 \, d\omega}{\int_0^\pi |H(e^{j\omega})|^2 \, d\omega}, \tag{4}$$

which is basically the second moment ("variance") of the "distribution" $|H(e^{j\omega})|^2$. Let K equal-length real-coefficient filters $H_i(z)$ be constrained by

$$\sum_{i=0}^{K-1} |H_i(e^{j\omega})|^2 = |H(e^{j\omega})|^2, \tag{5}$$

for some real-coefficient $H(z)$ and by

$$\frac{1}{\pi}\int_0^{\pi} |H_i(e^{j\omega})|^2 \, d\omega = c, \tag{6}$$

for some real constant c. The above constraints are characteristics of filters composing a paraunitary filter bank. A set of filters $\{H_i(z)\}$ is defined as having minimum frequency resolution (MFR) if the maximum $\psi_F(H_i)$ is minimized, i.e.

$$\{H_i(z) \mid \min_{H_i} \max_i \psi_F(H_i) \ ; (5); \ (6)\}. \tag{7}$$

Theorem 1 *An MFR set of filters obeys*

$$|H_i(e^{j\omega})|^2 = \frac{1}{K}|H(e^{j\omega})|^2, \tag{8}$$

being composed by spectral factors of $H'(z) = \frac{1}{K}H(z)H(z^{-1})$.

The MFR set has the desirable property of having filters with same frequency response. Hence, one might want to use the MFR set corresponding to the filters contained in $\mathbf{U}(z)$ as $\mathbf{S}(z)$. However, there are inconveniences in this approach. The MFR set may not be internally orthogonal neither orthogonal to the unselected filters. Also, in rare cases, there may not be enough distinct spectral factors. In this case, one might redesign $|H(e^{j\omega})|$ so that the zeros of $H(z)$ are disturbed. In any case, we have to find suitable approximations to the MFR set.

Let $\mathbf{A}$ be a $K \times L$ matrix transforming the signal vector $\mathbf{x}$ (which is obtained by windowing the signal $x(n)$ with a rectangular window of L taps) as $\mathbf{y} = \mathbf{Ax}$. At the next instant the window is shifted by M samples and the process is repeated. Let $\mathbf{B}$ be a given matrix of the same size as $\mathbf{A}$ and let $\mathbf{C}$ be a unitary matrix, while the signal has autocorrelation matrix $\mathbf{R}_{xx}$. Define an error vector as

$$\epsilon = \mathbf{y} - \mathbf{CBx} = (\mathbf{A} - \mathbf{CB})\mathbf{x}. \tag{9}$$

Theorem 2 *The unitary matrix* $\mathbf{C}$ *which makes the product* $\mathbf{CB}$ *to be the closest to* $\mathbf{A}$ *in the sense of minimizing the distance* $J = E\left[\frac{1}{K}\epsilon^H\epsilon\right]$ *(average error variance or error energy) is given by* $\mathbf{C} = \mathbf{Q}_1\mathbf{Q}_2$*, where* $\mathbf{Q}_1$ *and* $\mathbf{Q}_2$ *are unitary matrices derived from the SVD of* $\mathbf{D} = \mathbf{AR}_{xx}\mathbf{B}^H$ *as* $\mathbf{D} = \mathbf{Q}_1\mathbf{\Lambda}\mathbf{Q}_2$.

We can directly apply Theorem 2 for a simplified approximation to MFR sets. Let $\mathbf{U}$ be a matrix whose rows contain the selected filters. (In this case $\mathbf{U}$ has real entries and is an equivalent representation as that of $\mathbf{U}(z)$ [2].) Let us assume we want $\mathbf{\Phi}(z)$ to have order zero, i.e., it is an orthogonal matrix $\mathbf{\Phi}$. The resulting lapped transform matrix $\mathbf{S}$, is given by

$$\mathbf{S} = \mathbf{\Phi}\mathbf{U}. \tag{10}$$

From $\mathbf{S}$, $\mathbf{S}(z)$ can be immediately obtained [1],[2]. If the K MFR filters corresponding to $\mathbf{U}$ are described in the $K \times L$ lapped transform matrix $\mathbf{H}$, and if the SVD of $\mathbf{HR}_{xx}\mathbf{U}$ is given as $\mathbf{HR}_{xx}\mathbf{U} = \mathbf{Q}_1\mathbf{\Lambda}\mathbf{Q}_2$, we can select $\mathbf{\Phi} = \mathbf{Q}_1\mathbf{Q}_2$ so that

$$\mathbf{S} = \mathbf{Q}_1\mathbf{Q}_2\mathbf{U}. \tag{11}$$

This is a simple method to derive a post-processing stage composed only by an orthogonal transform. This method, as expected, yields limited results because of the low order of $\mathbf{\Phi}(z)$. However, it works well in a few cases and provides a powerful method to generate time-varying filter banks, since the post-processing stage can be turned on and off *without transitory states*. Therefore, one might easily implement a filter bank where the filters have time-varying bandwidth (to some extent) without any concern for boundary (transitory) instantaneous filter banks.

Let the signal $x(n)$ be periodic with very large period N_p. Let its Fourier transform be computed over one period as $X(e^{j\omega}) = \sum_{n=0}^{N_p-1} x(n)e^{jn\omega}$. For two signals $x'(n)$ and $x''(n)$ with the same period,

$$E[X'(e^{j\omega})X''(e^{-j\omega})] = N_p, \ {}_{x'x''}(e^{j\omega}) \tag{12}$$

where , ${}_{xy}(e^{j\omega})$ is the Fourier transform of the cross correlation between signals $x(n)$ and $y(n)$. Let the polyphase sequences of an input signal $x(n)$ be $x_i(m) = x(mM + i)$, and let $\mathbf{x}(e^{j\omega}) = [X_0(e^{j\omega}), \ldots, X_{M-1}(e^{j\omega})]^T$. Let $\mathbf{A}(z)$ and $\mathbf{B}(z)$ be given $K \times M$ PTMs and let $\mathbf{C}(z)$ be a $K \times K$ PTM of a paraunitary filter bank. Define the error measure as $\epsilon(e^{j\omega}) = \mathbf{y}(e^{j\omega}) - \mathbf{C}(e^{j\omega})\mathbf{B}(e^{j\omega})\mathbf{x}(e^{j\omega})$.

Theorem 3 *The paraunitary PTM* $\mathbf{C}(z)$ *which makes the product* $\mathbf{C}(z)\mathbf{B}(z)$ *closest to* $\mathbf{A}(z)$ *in the sense of minimizing the distance* $J = E\left[\frac{1}{K}\epsilon^H(e^{j\omega})\epsilon(e^{j\omega})\right]$ *is given by*

$$\mathbf{C}(z) = \mathbf{Q}_1(z)\mathbf{Q}_2(z) \tag{13}$$

where $\mathbf{Q}_1(z)$ *and* $\mathbf{Q}_2(z)$ *are such that* $\mathbf{Q}_1(e^{j\omega})$ *and* $\mathbf{Q}_2(e^{j\omega})$ *are unitary matrices derived from the SVD of*

$$\mathbf{D}(e^{j\omega}) = \mathbf{A}(e^{j\omega})\mathbf{\Gamma}(e^{j\omega})\mathbf{B}^H(e^{j\omega}) = \mathbf{Q}_1(e^{j\omega})\mathbf{\Lambda}(e^{j\omega})\mathbf{Q}_2(e^{j\omega}). \tag{14}$$

$\mathbf{\Gamma}(e^{j\omega})$ *is an* $M \times M$ *matrix with entries* , ${}_{x_i x_j}(e^{j\omega})$.

Theorem 3 can be readily applied to the approximation of an MFR set of filters. If the K MFR filters corresponding to $\mathbf{U}$ are described in the $K \times M$ PTM $\mathbf{H}(z)$, and if the SVD of $\mathbf{H}(e^{j\omega})\mathbf{\Gamma}(e^{j\omega})\mathbf{U}^H(e^{j\omega})$ is given as $\mathbf{H}(e^{j\omega})\mathbf{\Gamma}(e^{j\omega})\mathbf{U}^H(e^{j\omega}) = \mathbf{Q}_1(e^{j\omega})\mathbf{\Lambda}(e^{j\omega})\mathbf{Q}_2(e^{j\omega})$, we can select $\mathbf{\Phi}(z) = \mathbf{Q}_1(z)\mathbf{Q}_2(z)$ so that

$$\mathbf{S}(z) = \mathbf{Q}_1(z)\mathbf{Q}_2(z)\mathbf{U}(z) = \mathbf{\Phi}(z)\mathbf{U}(z). \tag{15}$$

Analytical continuation is only applicable if we know the frequency response for all ω and the transfer function is rational. As the relations to find $\mathbf{Q}_1$ and $\mathbf{Q}_2$ only exist for an individual point in the unit circle, we have an infinite length non-recursive filter solution for $\mathbf{\Phi}(z)$. If the entries of $\mathbf{\Phi}(z)$ are $\Phi_{ij}(z)$, and we only know $\Phi_{ij}(z)$ for every $z = e^{j\omega}$ we are left with a classical FIR filter design problem, where we try to fit a finite-length filter to a known continuous Fourier transform function. However, there is no guarantee that the resulting FIR PTM is paraunitary. The larger N (longer filters) the better chances for a good approximation.

An alternative to those methods is to compute an approximation to the MFR through optimization routines. However, we feel that if we resort to this technique, it will be more productive to optimize $\mathbf{\Phi}(z)$ directly, which we will discuss next.

4. DESIGN THROUGH OPTIMIZATION

An alternative is to directly optimize the post-processing paraunitary filter bank $\mathbf{\Phi}(z)$. In this case we can use any paraunitary filter bank design technique and set a suitable cost function. The cost function may not involve the computation of the MFR set. We know that all MFR filters are spectral factors so that they have the same spectral magnitude. So, we can setup a cost function to minimize the difference in absolute frequency response, while the paraunitariness constraint imposed in the optimization algorithm will do the rest. Let $S_i(z)$ be the i-th filter of $\mathbf{S}(z)$ $(0 \le i \le K-1)$. For example, we can use

$$J = \int_{\omega} \sum_{i} \left| |S_i(e^{j\omega})| - \bar{S}(e^{j\omega}) \right| \tag{16}$$

where $\bar{S}(e^{j\omega}) = \frac{1}{K}\sum_i |S_i(e^{j\omega})|$.

The optimization alternative avoids the filter design and spectral factorization problems found in MFR set approximation. However, optimization techniques are frequently unstable in a sense that no guarantees exist that a global minimum will be found. One method, for example, is to parameterize the filter bank into orthogonal factors and delay stages and to optimize the angles of the orthogonal factors using an unconstrained simplex search algorithm such as the one provided by Matlab^{TM} 4.2. The non-linear relation among angles and cost functions may complicate the process. As in any application involving complex numerical evaluation, the methods discussed here may be effective in some cases but fail in other cases.

5. ALTERNATIVE FILTER MODEL

Let $\mathbf{F}(z)$, the analysis PTM, be decomposed into two PTMs as $\mathbf{F}(z) = \mathbf{F}_1(z) + \mathbf{F}_2(z)$, where $\mathbf{F}_1(z)$ has zero row entries replacing the selected filters, while $\mathbf{F}_2(z)$ retains the selected filters and has zeros elsewhere. Adopt the same notation for the synthesis PTM $\mathbf{G}(z)$. Thus, the overall transfer is

$$\begin{aligned} \mathbf{T}(z) &= \mathbf{G}^T(z)\mathbf{F}(z) = \mathbf{G}_1^T(z)\mathbf{F}_1(z) + \mathbf{G}_2^T(z)\mathbf{F}_2(z) \\ &= \mathbf{H}_1(z) + \mathbf{H}_2(z) \end{aligned} \tag{17}$$

which is basically the sum of complementary LPTV filters, of which we just have interest in $\mathbf{H}_2(z)$. If $\bar{\mathbf{G}}(z)$ is $\mathbf{G}_2(z)$ with the zero rows removed, and since $\mathbf{U}(z)$ is $\mathbf{F}_2(z)$ with the zero rows removed, $\mathbf{H}_2(z) = \mathbf{G}_2^T(z)\mathbf{F}_2(z) = \bar{\mathbf{G}}^T(z)\mathbf{U}(z)$. Note that $\mathbf{H}_2(z)$ is an $M \times M$ PTM with rank K. Each of its row is a filter whose frequency response is hopefully close to be passband on the selected filters' passband and to have large attenuation otherwise. (In effect, this is closer to be true as the filters in the uniform filter bank have higher and higher stopband attenuation.) Therefore, the rows of $\mathbf{H}_2(z)$ may yield a filter close to the desired nonuniform band filter.

Let $\mathbf{D}$ be a $K \times M$ matrix designed to downsample the output of the LPTV filter so that $\mathbf{S}(z) = \mathbf{D}\mathbf{H}_2(z)$. Thus,

$$\mathbf{S}(z) = \mathbf{D}\bar{\mathbf{G}}^T(z)\mathbf{U}(z) = \mathbf{\Phi}(z)\mathbf{U}(z), \tag{18}$$

i.e.

$$\mathbf{\Phi}(z) = \mathbf{R}\bar{\mathbf{G}}^T(z), \tag{19}$$

By precalculating $\mathbf{\Phi}(z)$, we are actually resampling the selected synthesis filters at a lower rate and using the resulting subsampled filters as the post-processing stage to obtain the nonuniform bands.

The resulting filter bank is not necessarily paraunitary, although for the filter banks we have tested it is not far from being so. The filters in $\mathbf{S}(z)$ have linear phase and if the uniform filter bank $\mathbf{F}(z)$ also has linear phase filters, then $\mathbf{F}'(z)$ is very close to being a paraunitary system.

6. COSINE-BASED FILTER BANKS

Some filter banks present a very well organized structure, wherein the filters are samples of sinusoidal functions of different frequencies weighted by a "window". This "window" is a prototype low-pass filter which is modulated to obtain filters uniformly covering the spectrum from 0 to π [1]–[4]. These are called cosine modulated filter banks (CMFB). The discrete cosine transform (DCT) is also a variation on this theme, where the modulating window in an M-tap rectangular box, and so are the other variations of the DCT. The DCT has filters given by

$$f_i(j) = \sqrt{\frac{2}{M}}\alpha_i \cos\left(\frac{(2j+1)i\pi}{2M}\right) \tag{20}$$

where $\alpha_0 = 1/\sqrt{2}$ and $\alpha_{i>0} = 1$, for $0 \le i, j \le M-1$. One instance of the CMFB is the extended lapped transform (ELT) [2] whose filters $(g_k(n) = f_k(L-1-n))$ are given by:

$$g_k(n) = w(n)\sqrt{\frac{2}{M}} \cos\left[\left(k+\frac{1}{2}\right)\frac{\pi}{M}\left(n+\frac{M+1}{2}\right)\right] \tag{21}$$

for $k = 0, 1\ldots, M-1$ and $n = 0, 1, \ldots, L-1$, and where $w(n)$ is a window modulating the cosine terms. This CMFB is used as example and, for the present discussion, any other CMFB is applicable. Let the PTM for an M-channel CMFB or DCT be denoted by $\mathbf{C}_M(z)$. If $\mathbf{U}(z)$ is a set of selected filters from $\mathbf{C}_M(z)$

$$\mathbf{U}(z) = \mathbf{C}_K(z)\mathbf{H}(z) \tag{22}$$

for some LPTV filter $\mathbf{H}(z)$. Because of the modulating structure of CMFB one can check that, for specific selections of filters, $\mathbf{H}(z)$ approximates an MFR set, in the sense that the filters may have similar frequency response and passband coinciding with the passband of the selected filters. The modulating windows for the M- and K-channel CMFB must also be similar for better results [11]. For the most popular selections (i.e. M/K is an integer, the filters passbands occupy contiguous frequency slots, etc.) the approximation is very good. In those cases we can use:

$$\mathbf{S}(z) = z^{-N+1}\mathbf{C}_K^T(1/z)\mathbf{U}(z) \;\rightarrow\; \mathbf{\Phi}(z) = z^{-N+1}\mathbf{C}_K^T(1/z) \tag{23}$$

and $\mathbf{\Phi}(z)$ is the synthesis CMFB of K-channels. Given that a CMFB is easy to design and can possess fast implementation algorithms, it becomes very easy to design and implement a nonuniform filter bank.

A design example is shown in Fig. 1. Note that the band distribution in the second design cannot be approximated by hierarchical transforms.

7. CONCLUSIONS

One application for creating nonuniform bands through post-processing can be found in the field of time-varying filter banks, mainly with approaches that use cascade of post-processing stages [12],[13]. It can also be used for compression of audio and images, where high-pass filters are virtually *shortened* by post-processing to decrease ringing or pre-echo artifacts. These applications will be studied in more detail.

Post-processing stages are not a requirement for the design of the nonuniform band filter banks. The filters can be designed directly. We use the post-processing method because of its analytical simplicity allied with its good results. The increase in computation can be offset by using fast algorithms for each uniform stage, or by discarding marginal coefficients of the resulting filter.

We successfully tested the methods presented here on several filter bank classes. We hope the results presented in this paper may help to bridge the gap between uniform and nonuniform filter banks and to enable the use of uniform filter banks in applications where nonuniform filter banks are required.

8. REFERENCES

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

[2] H. S. Malvar, *Signal Processing with Lapped Transforms.* Norwood, MA: Artech House, 1992.

[3] G. Strang and T. Nguyen, *Wavelets and Filter Banks,* Wellesley, MA: Wellesley-Cambridge, 1996.

[4] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding,* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[5] K. Nayebi, T. P. Barnwell, M J. T. Smith, "The design of perfect reconstruction nonuniform band filter banks," *Proc. ICASSP*, pp. 1781-1784, 1991.

[6] P. Q. Hoang and P. P. Vaidyanathan, "Non-uniform multirate filter banks: theory and design," *Proc. of ISCAS*, pp. 371–374, 1991.

[7] J. Kovacevic and M. Vetterli, "Perfect reconstruction filter banks with rational sampling factor," *IEEE Trans. on Signal Processing,* vol. 41, pp. 2047–2066, June 1993.

[8] J. Princen, "The design of nonuniform modulated filter banks," *IEEE Trans. on Signal Processing,* vol. 43, pp. 2550–2560, Nov. 1995.

[9] J. Lee and B. G. Lee, "A design of nonuniform cosine modulated filter banks," *IEEE Trans. Circuits and Systems II,* Vol. 42, pp. 732–737, Nov. 1995.

[10] J. Li, T. Q. Nguyen, and S. Tantaratana, "A simple design method for near-perfect reconstruction nonuniform filter banks," preprint.

[11] R. L. de Queiroz and R. Eschbach, "Fast downscaled inverses for images compressed with M-channel lapped transforms," *IEEE Trans. on Image Processing,* Vol. 6, pp. 794–807, June, 1997.

[12] R.L. de Queiroz and K. R. Rao, "Time-varying lapped transforms and wavelet packets," *IEEE Trans. on Signal Processing,* vol. 41, pp. 3293–3305, Dec. 1993.

[13] I. Sodagar, K. Nayebi, T. P. Barnwell, "Time-varying analysis-synthesis systems based on filter banks and post-filtering," *IEEE Trans. on Signal Processing,* vol. 43, Oct. 1995.

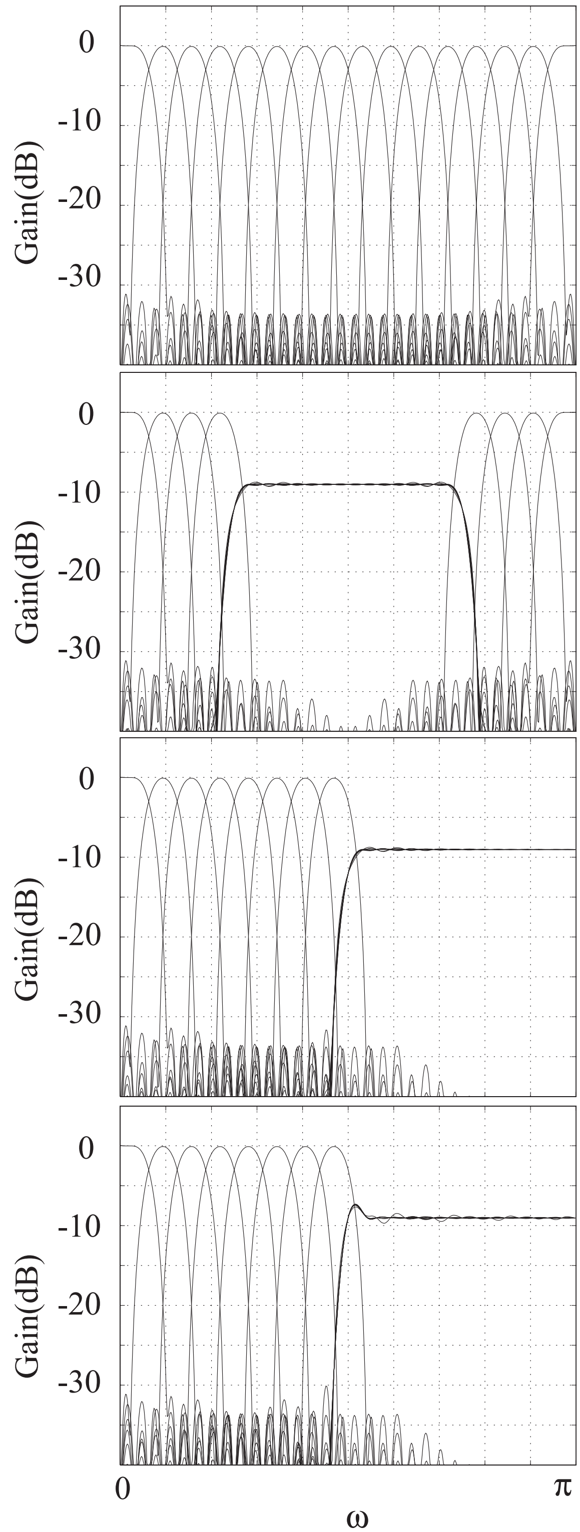

Figure 1: Nonuniform filter banks based on a 16-channel, L=64, CMFB using the inverse CMFB stage. The top graphic corresponds to the uniform bank and the bottom graphic corresponds to the design using the alternative filter model.