Modulated Filter Banks with Minimum Output Distortion in Presence of Subband Quantization

Karine Gosse[1], Tanja Karp[2], Pierre Duhamel[1], and Alfred Mertins[3]
[1]ENST Paris, Dept. Signal, 46 rue Barrault, F-75634 Paris, France, duhamel@sig.enst.fr
[2]Hamburg University of Technology, D-21071 Hamburg, Germany, karp@tu-harburg.d400.de
[3]University of Kiel, Kaiserstr. 2, D-24143 Kiel, Germany, am@techfak.uni-kiel.de

Abstract

This paper presents modulated minimum mean squared error (MMSE) filter banks, designed for subband coding purposes. Uniform scalar quantizers, the synthesis prototype and an additional gain control factor per subband are optimized in a rate-distortion sense. The main advantage of the proposed solution is to keep the efficient modulated implementation of the synthesis filter bank, while reducing the reconstruction distortion introduced by quantization of the subband signals. It is shown that, in many circumstances, the resulting schemes are almost as efficient as plain MMSE filter banks (Wiener filters with optimized quantizers).

1. Introduction

Classically, in filter bank (FB) based coding schemes, the signal to be encoded is split into several decorrelated subband components prior to quantization. The analysis and synthesis FB are usually chosen so as to allow very small reconstruction error (pseudo-QMF), or even perfect reconstruction (PR). These schemes perform lossy compression, and a classical problem is to minimize the reconstruction distortion under a bit-rate constraint. In this case, the "best" quantizers in the rate-distortion sense have to be found, for given analysis and synthesis FB. However, neither PR FB, which are cancelling the reconstruction error in absence of quantization, nor pseudo-QMF filters are optimal any more when subband signals are quantized.

The amount of injected quantization noise was already taken into account in some previous work in the special case of non-uniform quantization [8], [6], mainly in order to recover the PR property. It was further generalized in [3], [11] and [1] to any type of quantization, with an approach similar to Wiener filtering. In [3], we proposed solutions for optimizing jointly the reconstruction filters and uniform subband quantizers, so that the output distortion is minimized under bit-rate constraint, while [11] presents matrix Wiener filters for subband coding. The main difference between both works is that [11] does not address a bit-rate constrained optimization problem, which is done in [3]. In both studies, the synthesis filters obtained do not enable any efficient realization of the filter bank, for example by using fast transforms.

This has been overcome by the approach in [7] for paraunitary filter banks. There a polyphase matrix A is introduced on the decoder side and matrix and scalar quantizers are optimized jointly. Results are compared for different complexities of A (diagonal, tridiagonal or full $M \times M$ matrix) by means of rate-distortion curves.

In this work, the filter bank is not limited to be paraunitary any longer but we constrain the optimized synthesis filters of [3] to be modulated versions of a given prototype, which enables a fast implementation of the filtering process.

At first glance, the only quantity which seems to be tunable is the synthesis prototype filter. However, from our previous study [7], we found useful to introduce a gain control factor (gcf) α_k on each synthesis filter as another set of variables (this corresponds to an additional diagonal matrix A at the decoder). Thus, the problem is to jointly optimize the subband quantizers, the gain factors, and the synthesis prototype filter or a subset of these parameters under bit-rate constraint.

The resulting system is denoted as a modulated Minimum Mean Squared Error (MMSE) filter bank and is shown in **Figure 1**. All MMSE analysis filter banks are the same as in the PR case, to enable the comparison of different coding schemes. The box denoting "other processing" contains the optional gain control factors α_k introduced in this approach.

The paper compares the coding schemes given in **Table 1**, by the means of rate-distortion curves. In the PR scheme, only the quantization steps are tuned. Solutions "mod. MMSE gcf" and "mod. MMSE proto" show the respective roles of both additional parameters to be tuned, while "mod. MMSE gcf proto" uses all degrees of freedom in this modulated context. The solution "MMSE" is

1058-6393/97 $10.00

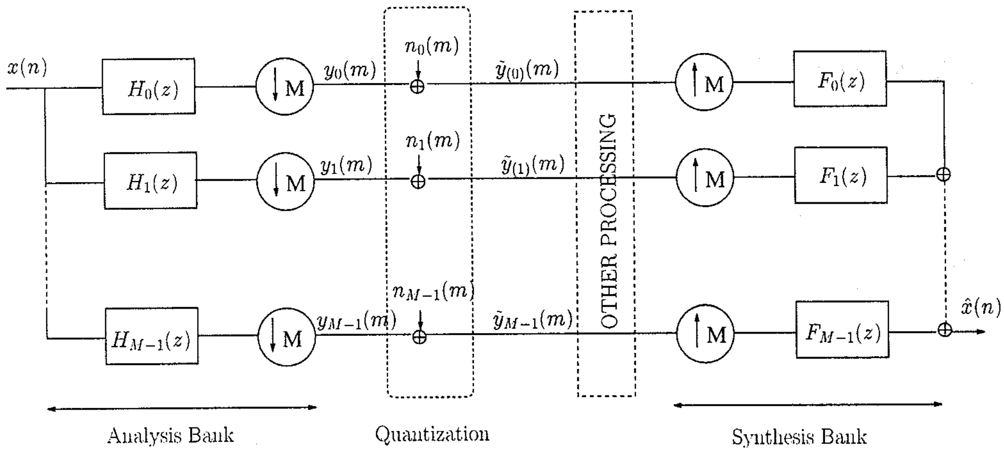

Figure 1. M-band compression system including a filter bank and a quantization stage. In the "mod. MMSE gcf proto" case, the synthesis filters are modulated versions of an optimized prototype and the other treatments include the multiplication by a scalar in order to modify the gain in each subband.

provided in order to show the loss in performance due to constraining the synthesis FB to be modulated.

Note that [2] is also related to this topic: optimal synthesis prototype filters are obtained for a given quantization in the subbands, too, but neither quantizers, nor gain factors are optimized. Moreover, a similar study was carried out in the case of tree-structured filter banks by Gosse & *al.* [4]. A joint optimization of synthesis parameters and subband quantizers is performed so as to preserve the structure, and to limit the number of parameters to be tuned.

2. Coding Scheme Design

Designing a modulated MMSE filter bank requires an expression of the MSE criterion to be minimized, as well as a quantization noise model relating the distortion to the subband bit-rates.

2.1. Distortion Criterion

Since the output signal of the general M-band coding scheme in **Figure 1** is cyclostationary of period M, the distortion of the reconstructed signal can be expressed as:

$$\begin{aligned} D &= \frac{1}{M}\sum_{i=0}^{M-1} \mathcal{E}\left[|\hat{x}(mM+i) - x(mM+i-t_0)|^2\right] \\ &= \sigma_x^2 + \frac{1}{M}\sum_{i=0}^{M-1}\big(\mathbf{F}_i^T\,\mathbf{A}_\alpha^T\,\mathbf{R}_{\tilde{y}\tilde{y}}\,\mathbf{A}_\alpha\,\mathbf{F}_i \\ &\quad -2\,\mathbf{F}_i^T\,\mathbf{A}_\alpha^T\,\mathcal{E}[\tilde{\mathbf{y}}(m)x(mM+i-t_0)]\big) \end{aligned} \tag{1}$$

where t_0 denotes the delay of the filter bank. $N = KM$ is the filter length. It involves the quantized signal vectors $\tilde{\mathbf{y}}^T(m) = [\tilde{y}_0(m),\ \tilde{y}_1(m), \ldots, \tilde{y}_{M-1}(m),\ \tilde{y}_0(m-1), \ldots, \tilde{y}_{M-1}(m-K+1)]$ and their autocorrelation matrix $\mathbf{R}_{\tilde{y}\tilde{y}}$, $f_k(n)$ denotes the n-th coefficient of the k-th synthesis filter, and $\mathbf{F}_i^T$ gathers the i-th polyphase components of the synthesis filters, $\mathbf{F}_i^T = [f_0(i), \ldots, f_{M-1}(i),\ f_0(i+M), \ldots, f_{M-1}(i+M), \ldots, f_{M-1}(i+(K-1)M)]$, σ_x^2 is the variance of the input signal $x(n)$.

The matrix $\mathbf{A}_\alpha$ is a block-diagonal matrix with K diagonal blocks $\mathcal{A}$ containing the gain factors

$$\mathcal{A} = \begin{bmatrix} \alpha_0 & & \mathbf{0} \\ & \alpha_1 & \\ & & \ddots & \\ \mathbf{0} & & & \alpha_{M-1} \end{bmatrix}. \tag{2}$$

Since the subband filters are modulated, the filter coefficients can be written as $f_k(n) = p(n)\cdot w(k,n)$, with $p(n)$ being the prototype filter, and $w(k,n)$ being the modulation function. The modulation matrices $\mathbf{W}_i$ of size $M \times M$ verify $\mathbf{W}_i = (w(k, n+iM))_{k,n=0\cdots M-1}$, and they are gathered in block diagonal matrix $\mathbf{W}_s$ of size $N \times N$ by $\mathbf{W}_s = \mathrm{diag}(\mathbf{W}_0, \ldots, \mathbf{W}_{K-1})$. We also define a matrix $\mathbf{P}$ of MMSE synthesis prototype coefficients:

$$\mathbf{P} = \begin{bmatrix} p(0) & & \mathbf{0} \\ & \searrow & \\ & & p(M-1) \\ p(M) & & \\ & \searrow & \\ & & p(2M-1) \\ & \vdots & \\ \mathbf{0} & & p(N-1) \end{bmatrix} \tag{3}$$

$\mathbf{P}_i$ will denote the i-th column of $\mathbf{P}$. With these notations,

denomination	type of the synthesis FB	gain factors	quantizers	synthesis prototype
PR	modulated	not present	tuned	not tuned (PR)
mod. MMSE gcf	modulated	tuned	tuned	not tuned (PR)
mod. MMSE proto	modulated	not tuned	tuned	tuned
mod. MMSE gcf proto	modulated	tuned	tuned	tuned
MMSE	unconstrained	not present	tuned	all filters tuned

Table 1. Various types of coding schemes enabeling an efficient filter bank implementation

the distortion in (1) becomes:

$$D = \sigma_x^2 + \frac{1}{M}\sum_{i=0}^{M-1}\Big(\mathbf{P}_i^T\mathbf{W}_s^T\mathbf{A}_\alpha^T\,\mathbf{R}_{\tilde{y}\tilde{y}}\,\mathbf{A}_\alpha\mathbf{W}_s\mathbf{P}_i - 2\mathbf{P}_i^T\mathbf{W}_s^T\mathbf{A}_\alpha^T\mathcal{E}\{\tilde{\mathbf{y}}(m)x(mM-t_0+i)\}\Big) \quad (4)$$

or alternatively, equation (5) gives the distortion as a function of the gain factors and the synthesis prototype coefficients, both variables to be tuned in this study.

$$D = \sigma_x^2 + \frac{1}{M}\sum_{i=0}^{M-1}\Big(\mathbf{P}_{i\downarrow}^T\mathbf{W}_{*,i}^T\mathbf{A}_\alpha^T\mathbf{R}_{\tilde{y}\tilde{y}}\mathbf{A}_\alpha\mathbf{W}_{*,i}\mathbf{P}_{i\downarrow} - 2\mathbf{P}_{i\downarrow}^T\mathbf{W}_{*,i}^T\,\mathbf{A}_\alpha^T\,\mathcal{E}\{\tilde{\mathbf{y}}(m)x(mM-t_0+i)\}\Big) \quad (5)$$

with $\mathbf{P}_{i\downarrow}$ equal to $\mathbf{P}_i$ with zero coefficients removed, and $\mathbf{W}_{*,i}$ extracted from $\mathbf{W}_s$, so that the columns corresponding to the i-th polyphase component are kept (not subsampled). Elements of matrix $\mathbf{W}_{*,i}$ are defined as follows:

$$\left[\mathbf{W}_{*,i}\right]_{k,l} = \left[\mathbf{W}_s\right]_{k,lM+i}$$

2.2. Quantization Process

D is also a function of the subband quantizers, and thus of the subband bit-rates. This is highlighted by means of a quantization noise modelization: in subband k, we use the traditional additive model $\tilde{y}_k(m) = y_k(m) - n_k(m)$, with n_k a white noise, of variance $\sigma_{n_k}^2 = q_k^2/12$. In [3], a more accurate noise model did not improve the performances of the MMSE scheme.

The distortion is thus the sum $D = D_f + D_n$ of a term D_n due to quantization noise, plus a filtering term D_f due to the non-PR optimized bank. If the quantization noise tends to zero, the optimal filters tend to the PR solution.

2.3. Bit-Rate Constraint

The fixed bit-rate budget can be expressed as

$$\sum_{k=0}^{M-1} R_k = R_T. \quad (6)$$

In most cases, noise variances $\sigma_{n_k}^2$ and bit-rates R_k can be related by:

$$\sigma_{n_k}^2 = c_k\,\sigma_{y_k}^2\,2^{-2R_k}, \quad (7)$$

The positive parameter c_k depends on the signal statistics in subband k as well as on the type of quantization and/or encoding process used.

For uniform quantizers, R_k and q_k verify $R_k = \log_2(d_k/q_k)$, with d_k, the dynamic range of the signal defined as $d_k = \max_y(y_k) - \min_y(y_k)$; d_k may also be set to $c_k\sigma_{y_k}$.

If the bit-rate is the order 1 entropy of the signals, $c_k = \pi e/6$ under the assumption of Gaussian subband signals, and $c_k = e^2/6$ under the assumption of Laplacian signals.

Note that the distortion expression is common to all bit-rate measures, up to multiplicative constants, thus enabling the use of the same optimization algorithm in all cases.

2.4. MMSE Solutions

Optimizing the free parameters separately avoids the use of constrained optimization methods, since each optimization performed alone (quantization steps, gain control factors, prototype) is simple. The process is then iterated. This procedure, initialized with the classical PR solution, either improves the distortion or stops at the very beginning.

Step 1 : The distortion is a quadratic function in terms of the **prototype coefficients**. Optimizing them consists in solving a set of linear equations given by $dD/d\mathbf{P}_i = 0$ for $i = 0, \ldots, M-1$.

Step 2 : Optimization of **gain control factors**: the criterion appears to be also quadratic in terms of the coefficients α_k.

Step 3 : Tuning the **quantization steps** can be done by minimizing the noise term

$$D_n = \sum_{k=0}^{M-1} c_k\left(\sum_{n=0}^{N-1} f_k(n)^2\right)2^{-2R_k}.$$

(the filtering term does not rely on subband bit-rates). The solutions for the R_k are of classical form (see [10] or [3] for positive solutions).

Each time the iteration of Step 1 and Step 2 has converged, Step 3 is performed, and this until final convergence.

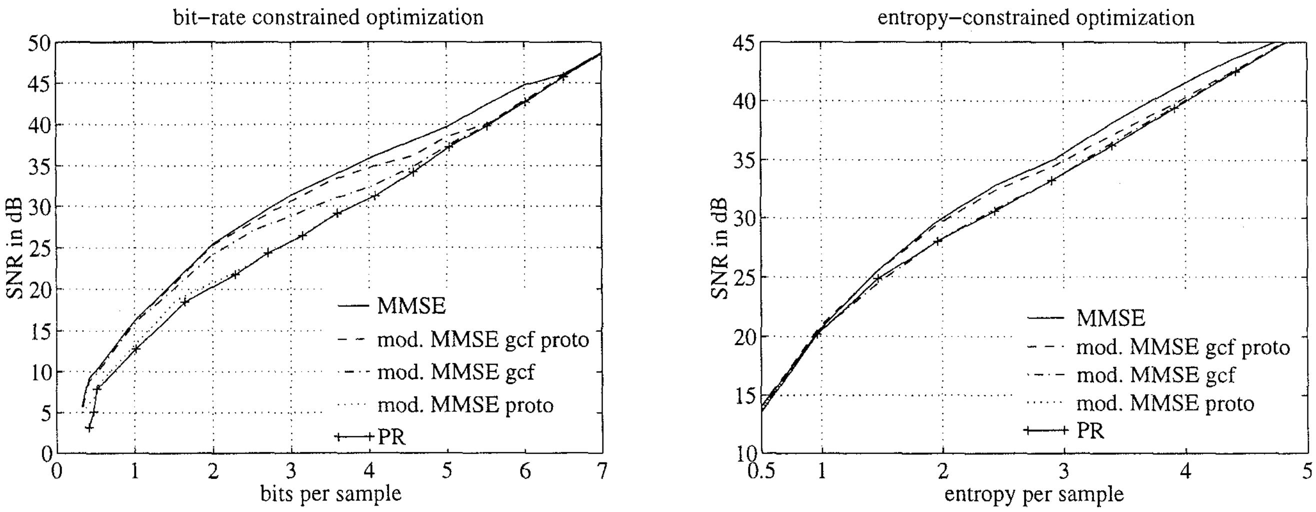

Figure 2. Measured bit-rate and entropy vs SNR (dB) for Vivaldi samples, and all MMSE schemes

The optimization of the whole synthesis filters (MMSE) for fixed quantizers is also straightforward by solving a set of linear equations (given in [3]). Iterate it alternatively with **Step 3**. Tuning the quantization steps in a PR scheme is done by performing **Step 3** once.

3. Simulations in an Audio Context

3.1. Simulation Context

Simulations aim at quantifying the improvement obtained with modulated MMSE FBs with respect to PR FBs, and at highlighting interesting performance/complexity trade-offs.

Analysis FBs. In order to enable the comparison, all schemes in a simulation set have the same analysis bank. Here, we use an 8-band Extended Lapped Transform (ELT, [9]) of length 32.

Input signals. Here, the plain SNR criterion is minimized (see [5] for frequency weighted perceptual criterion). However, in order to show that this approach is not linked to specific properties of the signal, the MMSE filter bank performances were evaluated on various synthetic signals, as well as the audio signal "The four seasons: the spring" by Vivaldi, CD quality. Only the Vivaldi results are shown here.

MMSE FBs in the simulations. Given the analysis bank and N samples of the input signal, the autocorrelation matrix of the subband signals is computed. Then, with the quantization noise model of section 2.3, MMSE filters and quantizers are found with the proposed algorithm. Here, the output SNR, resulting from the uniform quantization and the reconstruction by MMSE filters of the N input samples, is plotted versus bitrate (bits per sample).

3.2. Results

Figure 2a compares the rate-distortion curves obtained by the MMSE scheme, by the various modulated MMSE schemes, and by the PR scheme. The best results are of course obtained by optimizing the whole filters ($\geq$ 3dB improvement in the range of 1.5 - 5 bps), but almost the same performances are reached for bit-rates $\leq$ 4 bps when optimizing both gain factors and prototype, while keeping implementation costs low. The optimization of the gain factors alone, which is less complex, brings less improvement. However, at bit-rates up to 2 bps, this scheme having the lowest optimization complexity since it does not increase the implementation costs, still enables performances comparable to the other ones. Optimizing the prototype filter alone leads to very little improvement.

Figure 2b gives the rate-distortion curves corresponding to modulated MMSE schemes optimized under entropy constraint. The joint optimization of prototype and gain factors still brings an improvement of 2 dB, while increasing implementation costs by one multiplication per subband only. Optimizing the gain factors only does not yield any improvement anymore. An important fact, here, is that improvement is still to be noticed in an asymptotic situation.

Figure 3 shows the frequency responses of the synthesis filters for the different coding schemes. In the cases "mod. MMSE gcf", "mod. MMSE gcf proto", and "MMSE" all available bits are allocated to the subbands zero to three. Indeed, the MMSE optimization "concentrates" the avail-

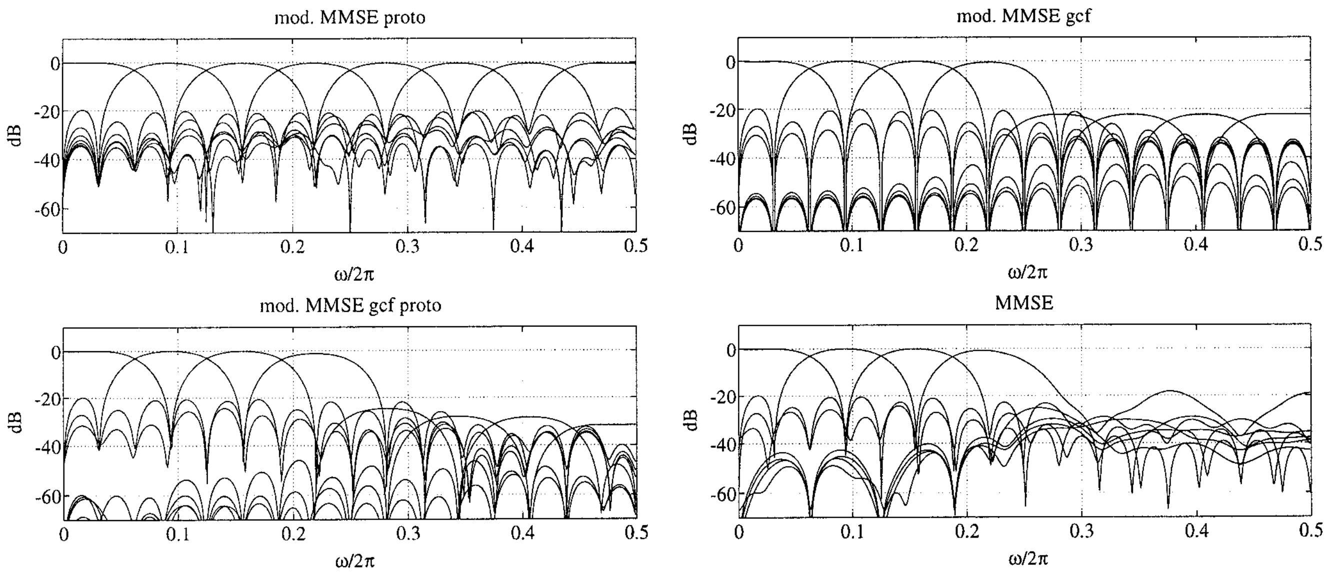

Figure 3. Synthesis filters frequency responses for 3bps

able bit-rate within a number of subbands smaller than in the PR case.

In the transmitted subbands, the first 3 synthesis filters do not differ significantly from the PR synthesis filters (apart from subband three in the "MMSE" case); these subband signals are quantized with a high resolution.

In subband number 3, the MMSE optimization consists also in shaping the quantization noise of schemes "MMSE" and "mod. MMSE gcf proto". The effect is very clear in the "MMSE" system, but it also appears slightly in the "MMSE gcf proto" one, where filter 3 has a band-pass gain smaller than 0 dB.

4. Conclusions

In coding schemes, the usefulness of relaxing the PR property of the synthesis filter bank has already been demonstrated, but this work enables to obtain an noticeable improvement with respect to a SNR criterion, **without increasing the realization complexity** of modulated filter banks. Since the improvement still exists when optimizing the prototype and the gain control factors under entropy constraint, we expect a gain in any coding system including scalar quantization and entropy coding.

References

[1] A. Delopoulos and S. Kollias. Optimal filter banks for signal reconstruction from noisy subband components. *IEEE Trans. on Signal Processing*, 44(2):212–224, February 1996.

[2] A. Dembo and D. Malah. Statistical design of analysis/synthesis systems with quantization. *IEEE Trans. on ASSP*, 36(3):328–341, Mar. 1988.

[3] K. Gosse and P. Duhamel. Perfect reconstruction versus MMSE filter banks in source coding. *Under Review, submitted to IEEE Trans. on Signal Processing*, 1995.

[4] K. Gosse, F. Moreau de Saint-Martin, and P. Duhamel. Filter bank design for minimum distortion in presence of subband qantization. In *Proceedings of the Int. Conf. on Audio Speech and Sig. Proc.*, Atlanta, USA, May 1996.

[5] K. Gosse, O. Pothier, and P. Duhamel. Optimizing the synthesis filter bank in audio coding for minimum distortion using a frequency weighted psychoacoustic criterion. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, USA, Oct. 1995.

[6] R. Haddad and K. Park. Modeling, analysis, and optimum design of quantized M-band filter banks. *IEEE Trans. on Signal Processing*, 43(11):2540–2549, Nov. 1995.

[7] T. Karp, K. Gosse, A. Mertins, and P. Duhamel. MMSE filter banks with reduced complexity. In *Proceedings of EUSIPCO, European Signal Processing Conference*, pages II.1187–II.1190, Trieste, Italy, Sept. 1996.

[8] J. Kovacevic. Subband coding systems incorporating quantizer models. *IEEE Trans. on Image Proc.*, May 1995.

[9] H. S. Malvar. *Signal Processing with Lapped Transforms.* Artech House, Norwood, MA 02062, 1992.

[10] A. Segall. Bit allocation and encoding for vector sources. *IEEE Trans. on Information Theory*, 22(2):162–169, March 1976.

[11] P. P. Vaidyanathan and T. Chen. Statistically optimal synthesis banks for subband coders. In *IEEE Asilomar Conference*, USA, Nov. 1994.