MMSE FILTER BANKS WITH REDUCED COMPLEXITY

T. Karp[1], K. Gosse[2], A. Mertins[3], P. Duhamel[2]
[1]Hamburg University of Technology, Telecommunications Institute, D-21071 Hamburg, Germany
[2]ENST Paris, Dept. Signal, 46, rue Barrault, F-75634 Paris Cedex 13, France
[3]Technical Faculty of the Christian-Albrechts-University, Kaiserstr. 2, D-24143 Kiel, Germany
Email: [1]karp@tu-harburg.d400.de, [2]duhamel@sig.enst.fr, [3]am@techfak.uni-kiel.de

ABSTRACT

This paper focuses on subband coding schemes based on critically decimated paraunitary filter banks. An additional network with matrix $\mathbf{A}$ is introduced on the decoder side. We show here how to optimize its coefficients jointly with the quantization steps in order to reduce the reconstruction mean square error (MSE) on the output signal due to quantization and filtering. This optimization is performed under bit-rate constraint. Of course, the resulting overall MMSE filter bank (including $\mathbf{A} \neq \mathbf{I}_M$) does not allow perfect reconstruction (PR), but the signal-to-noise ratio (SNR) is remarkably better than for the PR solutions.
The main advantage of such an approach is to preserve the original structure of the filter bank while improving the SNR. Here, we apply the method to modulated filters which can be implemented at very low cost. Solutions are given by comparing rate-distortion curves for various forms of $\mathbf{A}$, several paraunitary modulated filter banks, and two bit-rate measures.

1 INTRODUCTION

In a subband coding scheme, the analysis and synthesis filter bank are usually chosen so as to allow PR of the input signal when no quantization is performed in the subbands. Nevertheless, these filters are suboptimal in presence of quantization, and several previous studies [2, 10, 1, 9, 6, 4, 3] already address this problem. Indeed, the design of synthesis filters reconstructing the signal could take into account the amount of quantization noise injected in the subbands, and this should be more relevant than considering other criteria such as frequency selectivity for example.

In [10, 1], optimal synthesis filters are designed according to an MMSE (Minimum Mean Square Error) criterion for given noisy subband signals, yielding matrix Wiener or Kalman filters. This work is generalized in [3, 4] where the quantizers and the subband signals are optimized jointly under bit-rate constraint. However, in neither of the two cases, the resulting synthesis filters allow efficient realizations using fast transforms or lattice implementations. In order to overcome this problem, [2] optimizes the synthesis prototype of a modulated filter bank, enabling low complexity implementations. However, the method is specific to modulated filters and does not tune quantizers. The special case of non-uniform quantization, for which part of the quantization noise is correlated to the signal is treated in [5, 6].

This paper resumes the joint optimization of synthesis filter bank and quantizers of [4], but, here, the synthesis side is constrained to keep its original implementation, what is particularly interesting in the case of computationally efficient modulated filter banks. We focus on paraunitary filter banks and we introduce a network with polyphase matrix $\mathbf{A}$ on the decoder side, as described on Figure 1. Matrix $\mathbf{A}$ and uniform subband quantizers are optimized jointly in order to minimize the output MSE, thus leading to MMSE filter banks. The complexity of the coding scheme can thus be tuned by varying the structure of $\mathbf{A}$. Particular attention is paid to a diagonal and tridiagonal structure of scalar matrix $\mathbf{A}$. Indeed, for the diagonal case, introducing weights in the subbands of a PR scheme enables to reduce the noise contribution of subbands with small SNR in the reconstructed signal. Moreover, by combining adjacent subbands in the tridiagonal case, the fact that they overlap in the frequency domain is taken into account. Note that the case of a polynomial matrix $\mathbf{A}(z)$ can be straightforwardly derived from the results given in this paper. It would give additional degrees of freedom but since we seek low complexity, we do not pay particular attention to this case.

These various forms for matrix $\mathbf{A}$ will be compared by the means of rate-distortion curves. These simulations are performed on modulated filter banks, so as to get the lowest implementation complexity, but all results presented here are valid for any paraunitary filter bank.

2 CODING SCHEME DESIGN

The filter bank based coding scheme is shown in Figure 1, using polyphase matrix representation of the filtering process. When denoting $\mathbf{H}(z)$ the analysis polyphase component matrix, the subband signal vector $\mathbf{Y}(z)$ can be written as $\mathbf{Y}(z) = \mathbf{H}(z)\ \mathbf{X}(z)$, with $\mathbf{X}(z)$,

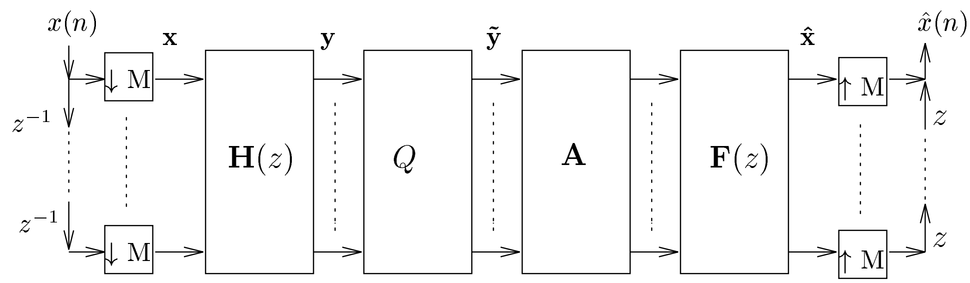

Figure 1: M-band coding system including a filter bank, a quantization stage and a combination matrix

the input signal vector. After uniform scalar quantization represented by Q in Figure 1, we obtain the quantized signal vector $\tilde{\mathbf{y}}(m) = Q(\mathbf{y}(m))$. This operation is modeled by $\tilde{\mathbf{y}}(m) = \mathbf{y}(m) + \mathbf{b}(m)$, where $\mathbf{b}(m)$ is an additive noise vector. The output signal then writes $\hat{\mathbf{X}}(z) = \mathbf{F}(z)\ \mathbf{A}\ \tilde{\mathbf{Y}}(z)$, with $\mathbf{A}$, an $M \times M$ scalar matrix, and $\mathbf{F}(z)$ the synthesis polyphase matrix.

2.1 Distortion Criterion

Let us assume that the overall delay of the coding scheme is zero. To achieve minimum distortion in an MSE sense, we have to solve

$$\mathcal{E}\{||\hat{\mathbf{x}}(n) - \mathbf{x}(n)||^2\} = \frac{1}{2\pi}\mathcal{E}\{||\hat{\mathbf{X}}(e^{j\Omega}) - \mathbf{X}(e^{j\Omega})||^2\} \overset{!}{=} \min \tag{1}$$

Since the filter bank is paraunitary, the following equation holds true:

$$||\hat{\mathbf{X}}(e^{j\Omega}) - \mathbf{X}(e^{j\Omega})|| = ||\mathbf{A}\tilde{\mathbf{Y}}(e^{j\Omega}) - \mathbf{Y}(e^{j\Omega})|| \tag{2}$$

Thus, the distortion can be expressed as

$$\begin{aligned} D &= \mathcal{E}\{||\mathbf{A}\tilde{\mathbf{y}}(m) - \mathbf{y}(m)||^2\} \qquad (3)\\ &= Trace\left[\mathbf{A}\mathbf{R}_{\tilde{y}\tilde{y}}\mathbf{A}^T - \mathbf{A}\mathcal{E}\{\tilde{\mathbf{y}}(m)\mathbf{y}^T(m)\}\right.\\ &\qquad \left.-\mathcal{E}\{\mathbf{y}(m)\tilde{\mathbf{y}}^T(m)\}\mathbf{A}^T + \mathcal{E}\{\mathbf{y}(m)\mathbf{y}^T(m)\}\right]\\ &= \sum_{i=0}^{M-1}\left[\mathbf{A}_i^T\mathbf{R}_{\tilde{y}\tilde{y}}\mathbf{A}_i - 2\mathcal{E}\{y_i(m)\tilde{\mathbf{y}}^T(m)\}\mathbf{A}_i + \sigma_{y_i}^2\right]. \end{aligned}$$

It involves the original and quantized subband signal vectors $\mathbf{y}(m) = (y_0(m), y_1(m), \ldots, y_{M-1}(m))^T$ and $\tilde{\mathbf{y}}(m)$, resp., and the autocorrelation matrix $R_{\tilde{y}\tilde{y}}$ of the latter. $\mathbf{A}_i^T$ denotes the i-th row of $\mathbf{A}$, and $\sigma_{y_i}^2$ the variance of the signal in subband i.

Modeling the quantization error $b_k(m)$ as additive, uncorrelated, white noise, we obtain from Eq. (3)

$$\begin{aligned} D &= \Big[\sum_{i=0}^{M-1}\overbrace{\mathbf{A}_i^T\mathbf{R}_{bb}\mathbf{A}_i}^{D_n} \qquad (4)\\ &\quad + \underbrace{\mathbf{A}_i^T\mathbf{R}_{yy}\mathbf{A}_i - 2\mathcal{E}\{y_i(m)\mathbf{y}^T(m)\}\mathbf{A}_i + \sigma_{y_i}^2}_{D_f}\Big] \end{aligned}$$

These assumptions on quantization noise seem restrictive, but using an elaborated noise model, as in [4], does not improve the performances of the MMSE system.

Because of the additive noise model, the distortion is the sum of 2 terms: one due to quantization noise, D_n, plus a filtering term D_f due to the non-PR property of the bank $\mathbf{F}(z) \cdot \mathbf{A}$. If the quantization noise tends towards zero (high bit-rates), the optimal $\mathbf{A}$ tends towards the identity matrix and the optimal system is close to the paraunitary one.

2.2 Bit-Rate Constraint

We work under the constraint $\sum_{k=0}^{M-1} R_k = R_T$, the fixed bit-rate budget. Classically, given d_k the dynamic range of the signal in subband k, defined as $d_k = \max_y(y_k) - \min_y(y_k)$, the bit-rate R_k and the quantization step q_k are related by : $R_k = \log_2(d_k/q_k)$.

We will also elaborate on the coding scheme performances in presence of an entropy coder, using the order 1 entropy H_k of the subband signals as a bit-rate measure during the optimization (for more details, see the entropy constrained optimization in [4]).

2.3 MMSE solutions

To achieve minimum distortion in an MSE sense, we have to minimize Eq. (4) over the set of coefficients of $\mathbf{A}$ and the subband quantizers. Optimizing jointly all free parameters requires the use of general optimization methods which revealed difficulties in convergence. The problem has been bypassed by optimizing the matrix $\mathbf{A}$ and the quantizers separately, since each optimization step is simple if the other parameters are fixed. The process is then iterated starting from the classical PR system with optimized bit-rates. The procedure either improves the classical solution or stops. However, we could not prove that it reaches a global optimum.

Step 1 : For fixed quantizers, the solution for $\mathbf{A}_i$ yielding minimum distortion is obtained by setting $\frac{\partial D}{\partial A_i^T} = 0$. It writes

$$\mathbf{A}_i^T = \mathcal{E}\{y_i(m)\mathbf{y}^T(m)\}(\mathbf{R}_{yy} + \mathbf{R}_{bb})^{-1}. \tag{5}$$

This equation simplifies if $\mathbf{A}$ is constrained to be diagonal or tridiagonal, as shown below:

$$\mathbf{A} = \begin{bmatrix} \alpha_{0,0} & \alpha_{0,1} & & & \\ \alpha_{1,0} & \alpha_{1,1} & \alpha_{1,2} & & \\ & \alpha_{2,1} & \alpha_{2,2} & \alpha_{2,3} & \\ & & \ddots & \ddots & \ddots \end{bmatrix} \tag{6}$$

For the diagonal case, Eq. (5) reduces to

$$\alpha_{i,i} = (\sigma_{y_i}^2 + \sigma_b^2)^{-1}\sigma_{y_i}^2 \tag{7}$$

In the tridiagonal case, it writes

$$[\alpha_{i,i-1};\ \alpha_{i,i};\ \alpha_{i,i+1}] = [y_{i-1};\ y_i;\ y_{i+1}](\mathcal{R}_{yy}^i + \mathcal{R}_{bb}^i)^{-1} \tag{8}$$

with $\mathcal{R}_{yy}^i = \left[(\mathbf{R}_{yy})_{k,n}\right]_{\substack{i-1 \le k \le i+1 \\ i-1 \le n \le i+1}}$ and $\mathcal{R}_{bb}^i$ defined in the same way.

Step 2 : Tuning the quantization steps, i.e. the bit allocation in each subband, for fixed $\mathbf{A}$ can be done by minimizing the noise term of the distortion:

$$D_n = \sum_{i=0}^{M-1} \left[\mathbf{A}_i^T\ \mathbf{R}_{bb}\ \mathbf{A}_i\right] = \sum_{k=0}^{M-1} \sigma_{b_k}^2 \sum_{i=0}^{M-1} \alpha_{i,k}^2, \tag{9}$$

with $\sigma_{b_k}^2 = q_k^2/12 = d_k^2/12 \cdot 2^{-2R_k}$ being the noise variance in subband k (we assume uniform quantization noise).

The solutions for the R_k are thus obtained by minimizing the Lagrangian functional $\mathcal{L} = D_n + \lambda(\sum_k R_k - R_T)$, in which R_k is constrained to be positive (this is done by setting $R_k = \rho_k^2$, see [4]).

3 SIMULATIONS

In the following paragraph, we compare the performances of various synthesis filter banks: a PR scheme having optimal subband bit-rates, a full MMSE scheme (optimization of the whole synthesis filters, as in [4]), and paraunitary schemes combined with 3 different matrices $\mathbf{A}$ (plain, diagonal and tridiagonal). This comparison is done by the means of rate-distortion curves.

Two simulation sets are performed in order to compare two reference modulated analysis banks: an 8-band ELT [7] with filters of length $N = 32$ and a cosine-modulated bank [8] having $M = 8$ and $N = 128$. Both prototype frequency responses are shown in Figure 2.

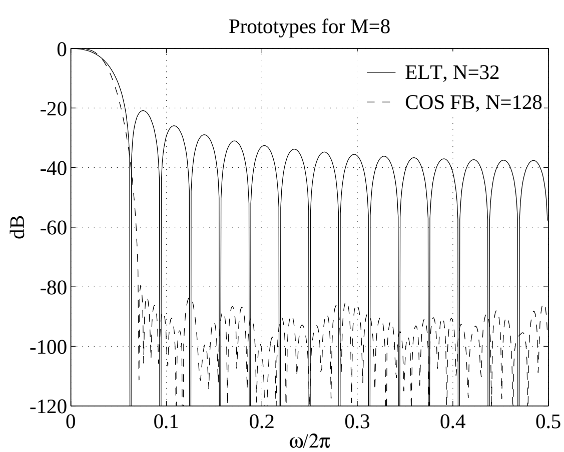

Figure 2: Prototypes for 8-channel paraunitary filter bank

The MMSE filter bank performances were evaluated on various synthetic signals, as well as on the audio signal "The four seasons: the spring" by Vivaldi, CD quality. Only the Vivaldi results are shown here.

3.1 Classical bit-rate measure

Figure 3 shows the rate-distortion curves obtained for the 8-channel ELT analysis filter bank, and for the classical bit-rate measure.

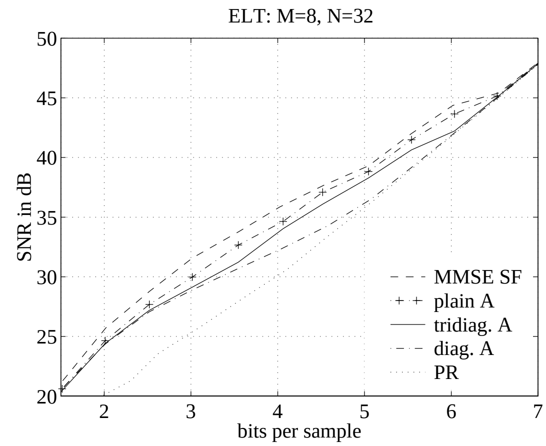

Figure 3: Measured bit-rate vs SNR (dB) for ELT

The best performances are of course obtained by optimizing the whole synthesis filters: the SNR improvement over PR is ≥ 5 dB in the range of 1.5-4.5 bps. However, optimizing the main diagonal of $\mathbf{A}$ only, improves the SNR of ≥ 3 dB for low bit rates up to 3.5 dB. This improvement, obtained at very low cost, reduces quickly for higher bit rates. It can be kept ≥ 3 dB for bit rates up to 4.5 bps by optimizing the tridiagonal structure of $\mathbf{A}$ as described in Eq. (6). Optimizing the whole matrix $\mathbf{A}$, which results in increased implementation costs, yields further improvement. This can be explained by the fact that the stopband attenuation of the prototype is relatively low (≤ 40 dB), so that even non-adjacent subband signals are still correlated.

Figure 4 gives the rate-distortion curves when using the cosine-modulated filter bank of length 128. In this case, all optimization methods yield the same rate-distortion curve. Because of the high stopband attenuation of the prototype, the subband signals are decorrelated and no further improvement can be achieved by combining them with a tridiagonal or plain matrix $\mathbf{A}$. Therefore, for such analysis filters, introducing a diagonal matrix on the synthesis side is the best solution for improving the output SNR of the filter bank. It requires only one additional multiplier per subband, and it enables to use the existing low complexity implementations of modulated filters.

3.2 Entropy-constrained optimization

In this paragraph, the bit-rate measure is the order 1 entropy of the subband signals, thus yielding an entropy-

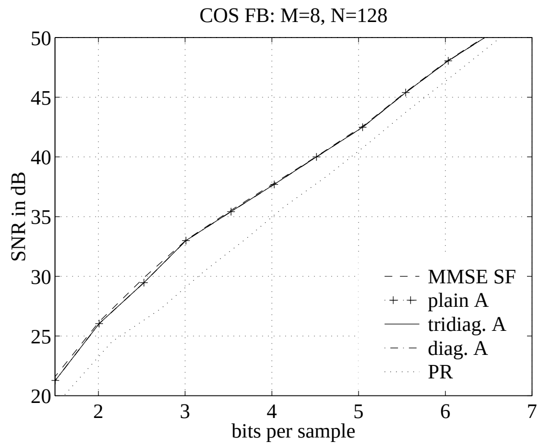

Figure 4: Measured bit-rate vs SNR (dB) for COS FB

constrained optimization of filters and quantizers. Figure 5 shows the curves obtained in this case with the ELT analysis.

Since such simulations evaluate asymptotic performances, the improvement enabled by MMSE SF schemes over PR ones is reduced, here to 0.25 bps in the range 27-45 dB. The same conclusions as in the classical bit-rate case can be given for the plain and tridiagonal matrices **A**. Note that introducing a diagonal matrix **A** is not sufficient any more to improve the performances of a PR scheme.

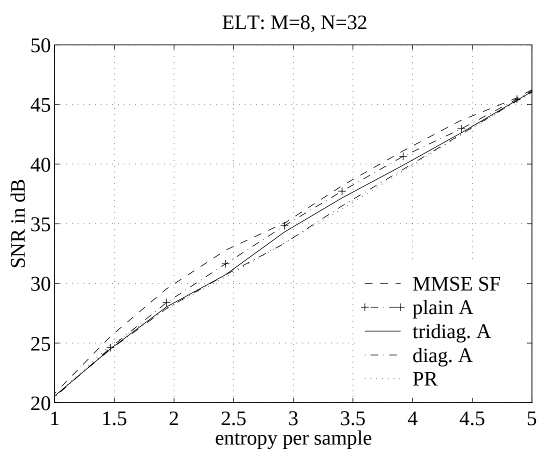

Figure 5: Entropy vs SNR (dB) for ELT

It also appears that this entropy bit-rate reduction obtained thanks to MMSE solutions completely cancels in the case of $N = 128$ modulated filters. The reasons of this fact are certainly manifold. First, we are dealing with asymptotic performances. Then, increasing the selectivity of the analysis filters reduces the potential gain brought by MMSE solutions. This shows that a good frequency selectivity is a crucial criterion for the choice of PR filters. It improves the overall system performances, and brings the scheme closer to optimality (the optimum being the MMSE solution).

4 Conclusion

This paper shows that noticeable output SNR improvements over classical PR schemes can be obtained at very low cost by introducing a diagonal or tridiagonal matrix on the synthesis side of a paraunitary filter bank. For the classical bit-rate measure, a diagonal matrix gives enough degrees of freedom at low bit-rates to get this improvement. However, in presence of an entropy-coding stage, the tridiagonal matrix is necessary. Finally, the solutions proposed here present a good SNR/complexity tradeoff compared to both PR filter banks and fully optimized MMSE schemes.

References

[1] A. Delopoulos and S. Kollias. Optimal Filter Banks for Signal Reconstruction from Noisy Subband Components. *IEEE Trans. on Signal Processing*, SP-44:212–224, February 1996.

[2] A. Dembo and D. Malah. Statistical Design of Analysis/Synthesis Systems with Quantization. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 36(3):328–341, March 1988.

[3] K. Gosse, F. Moreau de Saint Martin, and P. Duhamel. Filter Bank Design for Minimum Distortion in Presence of Subband Qantization. In *Proc. IEEE ICASSP*, Atlanta, USA, May 1996.

[4] K. Gosse and P. Duhamel. Perfect Reconstruction versus MMSE Filter Banks in Source Coding. submitted to *IEEE Trans. on Signal Processing*, 1995.

[5] R. A. Haddad and N. Uzun. Modeling, Analysis and Compensation of Quantization Effects in M-Band Subband Codecs. In *Proc. IEEE ICASSP*, pages III.173–III.176, Minneapolis, USA, April 1993.

[6] J. Kovacevic. Subband Coding Systems incorporating Quantizer Models. *IEEE Trans. on Image Processing*, May 1995.

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

[8] T. Nguyen. Digital Filter Banks Design - Quadratic-Constrained Formulation. *IEEE Trans. on Signal Processing*, SP-43:2103–2108, September 1995.

[9] N. Uzun and R. A. Haddad. Cyclostationary Modeling, Analysis, and Optimal Compensation of Quantization Errors in Subband Codecs. *IEEE Trans. on Signal Processing*, SP-43:2109–2119, September 1995.

[10] P. P. Vaidyanathan and T. Chen. Statistically Optimal Synthesis Banks for Subband Coders. In *Proc. IEEE Asilomar Conference*, USA, November 1994.