

DYNAMIC SUBBAND STRUCTURES FOR ECHO CANCELLATION

Amere Oakman and Patrick Naylor

CSP Group, Dept. of Electrical and Electronic Engineering
London, UK

ABSTRACT

Subband adaptive filters suffer degraded performance when high input energy occurs at the frequencies of subband boundaries. This is seen as increased error in critically sampled systems and as reduced asymptotic convergence speed in oversampled systems. An efficient dynamic frequency decomposition scheme has previously been shown by the authors to be effective in reducing these errors. This paper presents an analytical framework for the evaluation of these non-uniform dynamic subband systems. Simulation results show reductions in MSE of around 5-10dBs for the critical case in addition to increased robustness to coloured inputs.

1. INTRODUCTION

Subband adaptive filters (SAF) are used in system identification applications such as acoustic echo cancellation where the unknown system can be of the order of several thousand taps. They have the main benefits of reduced complexity and possible increased convergence speed due to reduction of eigenvalue spread in the subband signals [1]. Errors in both critically and oversampled SAFs can be shown to be related to signal components around subband boundaries, which manifest themselves in the case of critical sampling as dominating peaks in the final error signal around the subband boundaries [2]. In the case of oversampling, slow asymptotic convergence is observed [3].

This paper examines the use of non-uniform dynamic frequency-subband decomposition (NDS) to substantially reduce these errors. The algorithm chooses the decomposition so as to avoid high-energy signal components around subband boundaries, whilst retaining high decimation factors when possible so as to keep complexity low.

Previous work [4] is extended in this paper by the introduction of an analytical framework appropriate for the study of such non-uniform SAFs. The analysis is primarily focussed on critically sampled schemes, although the framework also applies to the oversampled case.

2. PROPOSED SCHEME

2.1. Filterbank structure

In this section we present an overview of the proposed scheme, which is more fully described in [4]. A non-uniform filterbank (NUFB) is obtained by merging the subbands of a

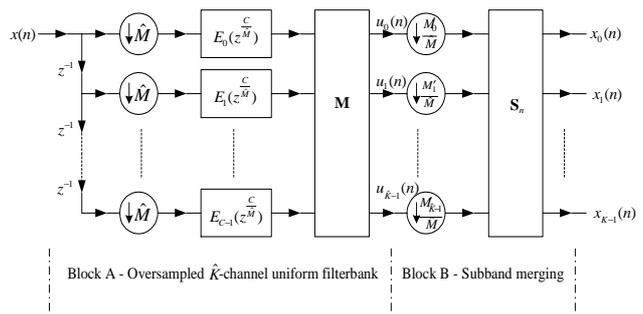


Fig. 1. General dynamic filterbank structure

\widehat{K} -channel uniform filterbank (UFB) and then decimating each of the resulting $K (\leq \widehat{K})$ subbands by an integer factor M_k [5]. This structure is modified to provide an efficient method of NDS. The structure has the general form of Figure 1, where $x(n)$ is the fullband input signal, $E_c(z)$ are polyphase components of a prototype lowpass filter $P(z)$, \mathbf{M} is a modulation matrix, \mathbf{S}_n is a time-varying summation matrix consisting of ones and zeros, and $x_k(n)$ is the k^{th} subband signal. The first stage of decimation is by a constant integer, \widehat{M} , whilst the second stage is by a time-varying integer, $M_{k,n}/\widehat{M}$, for the k^{th} subband. Strictly, k should be denoted k_n since it is time-varying, however the index is dropped for clarity. \widehat{M} represents the lowest decimation factor in the structure and therefore defines the largest subband channel bandwidth, determined by the application and as the greatest common denominator of all possible $M_{k,n}$

$$\widehat{M} = \text{gcd}(M_{k,n}) \forall k, n \quad (1)$$

to provide maximum decimation.

In Figure 1, Block A is an oversampled UFB and Block B is a subband merging section. This structure is preferred since A can be implemented efficiently using fast transforms of the outputs of a decimated polyphase network and B consists solely of adders and decimators, allowing NDS through changes in B only, without the need for intermediate upsampling/downsampling. The synthesis bank is the mirror of Figure 1. The notation used to define a decomposition is a 'split vector', the elements of which indicate the bandwidths of each subband relative to the constituent subband bandwidth, e.g. a 4-band uniform decomposition (with a 4-band

constituent filterbank) is represented as $[1 \ 1 \ 1 \ 1]$ whereas a decomposition with the first two bands merged is $[2 \ 1 \ 1]$.

2.2. Block A implementation

For the case of a critically sampled NUFB, Block A can be implemented as a cosine-modulated filterbank (CMF), by setting $C = 2\hat{K}$ and \mathbf{M} to be a $(\hat{K} \times \hat{K})$ cosine modulation matrix combined with a complex-conjugate summation matrix. This will produce the required, real-valued subbands as long as \hat{M} satisfies (1). Implementing the modulation using a fast DCT gives an overall computational complexity for Block A, in terms of real multiplies per fullband sample period (rmfp),

$$(1/\hat{M})(L_p + ((\hat{K}/2) \log_2 \hat{K} + \hat{K})) \quad (2)$$

where L_p is the length of the prototype filter. Since for this structure $\hat{M} \leq \hat{K}$, the computational cost of the filterbank is slightly greater than for a standard filterbank. However, this can be compensated for by reducing the subband complexity whilst maintaining significant performance advantages over equivalent complexity static structures, as shall be seen. The oversampled implementation of the structure is described in [4].

2.3. Control algorithm

Both critical and oversampled schemes use the same basic approach. K is initialised to $K_0 = \hat{K}$, i.e. all subbands have minimum bandwidth at $n = 0$, giving the greatest resolution for merging decisions. The structural adaptation occurs blockwise, with merging decisions at the end of each block based on the criterion for the two cases. Smaller bandwidth subbands are retained where possible, for efficiency. In the critically sampled case, we attempt to remove large aliasing errors at the subband boundaries, which are a main cause of overall error [2].

3. SUBBAND ERROR ANALYSIS FOR NON-UNIFORM SAFS

3.1. Subband Wiener-Hopf solution

Figure 2 shows the k^{th} and $(k+1)^{\text{th}}$ subband filters of a \hat{K} -subband uniform SAF, including adjacent-band cross-filters only, as in [2]. For the k^{th} analysis filter, the fullband unknown system, and the fullband input and desired signals, respectively, we define

$$\mathbf{h}_k = [h_{k,0} \ h_{k,1} \ \cdots \ h_{k,L_h-1}]^T \quad (3)$$

$$\mathbf{s} = [s_0 \ s_1 \ \cdots \ s_{L_s-1}]^T \quad (4)$$

$$\mathbf{x}(n) = [x(n) \ x(n-1) \ \cdots \ x(n-L_x+1)]^T \quad (5)$$

$$\mathbf{d}(n) = [d(n) \ d(n-1) \ \cdots \ d(n-L_d+1)]^T \quad (6)$$

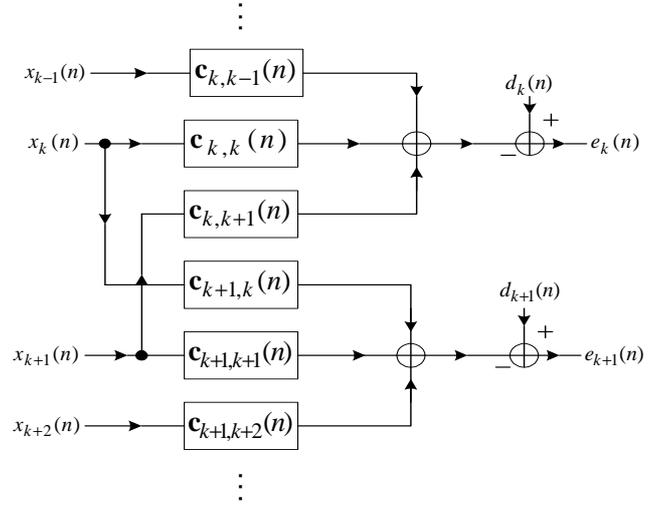


Fig. 2. SAF model, k^{th} and $(k+1)^{\text{th}}$ subbands

where $d(n) = \mathbf{s}^T \mathbf{x}(n)$. The k^{th} subband input and desired signals

$$x_k(n) = \mathbf{h}_k^T \mathbf{x}(\hat{K}n), \quad 0 \leq k \leq \hat{K} - 1 \quad (7)$$

$$d_k(n) = \mathbf{h}_k^T \mathbf{d}(\hat{K}n), \quad 0 \leq k \leq \hat{K} - 1 \quad (8)$$

which yields the subband input and desired vectors $\mathbf{x}_k(n)$ and $\mathbf{d}_k(n)$, respectively, defined analogously to (5) and (6). The three length L_c subband adaptive filters in the k^{th} subband are $\mathbf{c}_{k,k}(n)$ and $\mathbf{c}_{k,k\pm 1}(n)$. This gives the k^{th} subband desired-signal estimate $\hat{d}_k(n)$ and subband error $e_k(n)$ to be

$$\hat{d}_k(n) = \mathbf{c}_{k,k-1}^T(n) \mathbf{x}_{k-1}(n) + \mathbf{c}_{k,k}^T(n) \mathbf{x}_k(n) + \mathbf{c}_{k,k+1}^T(n) \mathbf{x}_{k+1}(n) \quad (9)$$

$$e_k(n) = d_k(n) - \hat{d}_k(n) \quad (10)$$

Dropping the index n to indicate optimal values, the minimum MSE (MMSE) in the k^{th} subband can be written

$$\begin{aligned} J_{k,\min} &= \mathcal{E} \{ e_{k,\min} e_{k,\min}^* \} \\ &= \sigma_{d_k}^2 - 2\mathbf{p}_{k,k}^T \mathbf{c}_{k,k} - 2\mathbf{p}_{k,k-1}^T \mathbf{c}_{k,k-1} - 2\mathbf{p}_{k,k+1}^T \mathbf{c}_{k,k+1} \\ &\quad + \mathbf{c}_{k,k}^T \mathbf{R}_{k,k} \mathbf{c}_{k,k} + \mathbf{c}_{k,k}^T \mathbf{R}_{k,k-1} \mathbf{c}_{k,k-1} + \mathbf{c}_{k,k-1}^T \mathbf{R}_{k-1,k} \mathbf{c}_{k,k} + \\ &\quad + \mathbf{c}_{k,k}^T \mathbf{R}_{k,k+1} \mathbf{c}_{k,k+1} + \mathbf{c}_{k,k+1}^T \mathbf{R}_{k+1,k} \mathbf{c}_{k,k} + \\ &\quad + \mathbf{c}_{k,k-1}^T \mathbf{R}_{k-1,k-1} \mathbf{c}_{k,k-1} + \mathbf{c}_{k,k-1}^T \mathbf{R}_{k-1,k+1} \mathbf{c}_{k,k+1} + \\ &\quad + \mathbf{c}_{k,k+1}^T \mathbf{R}_{k+1,k+1} \mathbf{c}_{k,k+1} \end{aligned} \quad (11)$$

where $\mathbf{R}_{lm} = \mathcal{E} \{ \mathbf{x}_l(n) \mathbf{x}_m^T(n) \}$ and $\mathbf{p}_{lm}^T = \mathcal{E} \{ d_l(n) \mathbf{x}_m^T(n) \}$. Noting that $\mathbf{R}_{lm} = \mathbf{R}_{ml}^T$, that \mathbf{R}_{lm} is in general negligibly small when $|l-m| \geq 2$, and extending the method in [6], we take partial derivatives of $J_{k,\min}$ w.r.t. the individual filter taps in each SAF and set each to zero giving the following set of Wiener-Hopf equations for the whole SAF system

$$\mathbf{R}_{\text{sys}} \mathbf{C}_{\text{sys}} = \mathbf{P}_{\text{sys}} \quad (12)$$

where

$$\mathbf{R}_{\text{sys}} = \begin{bmatrix} \mathbf{R}_{0,0} & \mathbf{R}_{0,1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{R}_{1,0} & \mathbf{R}_{1,1} & \mathbf{R}_{1,2} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{R}_{\hat{K}-1, \hat{K}-2} & \mathbf{R}_{\hat{K}-1, \hat{K}-1} \end{bmatrix}$$

$$\mathbf{C}_{\text{sys}} = \begin{bmatrix} \mathbf{c}_{0,0} & \mathbf{c}_{1,0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{c}_{0,1} & \mathbf{c}_{1,1} & \mathbf{c}_{2,1} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{c}_{M-2, M-1} & \mathbf{c}_{\hat{K}-1, \hat{K}-1} \end{bmatrix} \quad (13)$$

\mathbf{P}_{sys} has the same form as \mathbf{C}_{sys} . Rearranging (12) gives the subband adaptive filter solutions

$$\mathbf{C}_{\text{sys}} = \mathbf{R}_{\text{sys}}^{-1} \mathbf{P}_{\text{sys}} \quad (14)$$

Substituting the expressions for \mathbf{p}_{lm} from (12) into (11) gives the k^{th} subband error

$$J_{k, \min} = \sigma_{d_k}^2 - \mathbf{C}_{\text{sys}, k}^T \mathbf{P}_{\text{sys}, k} \quad (15)$$

where $\mathbf{C}_{\text{sys}, k}$ is the k^{th} column of \mathbf{C}_{sys} and $\mathbf{P}_{\text{sys}, k}$ is the k^{th} column of \mathbf{P}_{sys} . This value will not in general be zero due to the fact that we are only including adjacent band cross-band adaptive filters. The actual value depends upon the level of the non-adjacent aliasing terms, which are usually small.

3.2. Non-uniform subband Wiener-Hopf solution

A non-uniform subband decomposition is constructed through the merging of appropriate adjacent subbands. Therefore the non-uniform filterbank has $\hat{K} - 1$ subband boundaries in it's constituent state, i.e. when no bands are merged. Note that although Figure 2 shows cross-band adaptive filters, this is only for modelling purposes, as they are not used in the NDS implementation. Therefore, the initial state of the system can be represented by

$$\begin{aligned} \mathbf{R}_{\text{sys}, \text{init}} &= \text{diag} \{ [\mathbf{R}_{00} \ \mathbf{R}_{11} \ \cdots \ \mathbf{R}_{\hat{K}-1, \hat{K}-1}] \} \\ \mathbf{C}_{\text{sys}, \text{init}} &= \text{diag} \{ [\mathbf{c}_{00} \ \mathbf{c}_{11} \ \cdots \ \mathbf{c}_{\hat{K}-1, \hat{K}-1}] \} \\ \mathbf{P}_{\text{sys}, \text{init}} &= \text{diag} \{ [\mathbf{p}_{00} \ \mathbf{p}_{11} \ \cdots \ \mathbf{p}_{\hat{K}-1, \hat{K}-1}] \} \\ \mathbf{R}_{\text{sys}, \text{init}} \mathbf{C}_{\text{sys}, \text{init}} &= \mathbf{P}_{\text{sys}, \text{init}} \end{aligned} \quad (16)$$

The set of solutions $\mathbf{C}_{\text{sys}, \text{init}}$ will give rise to an error value that is generally greater than the value in (15) (see [2] for exceptions) due to the fact that no cross-band filters are present. We can continue to use the framework to model the subband errors in the non-uniform case by thinking of the merging of two subbands as the *absence* of the boundary between them and the *presence* of the associated cross-band adaptive filters. In this sense we equate the system in (13) with the non-subband minimum error performance and the system in (16) with the standard critically sampled uniform SAF system. In this work, by considering the MMSE for a merged subband as in (11), we represent general non-uniform SAF error performance by the inclusion or exclusion of the cross-band terms in the system equations. For

example, consider the system equations for a split vector of $[2 \ 1 \ 1]$

$$\mathbf{C}_{\text{sys}4, [2 \ 1 \ 1]} = \begin{bmatrix} \mathbf{c}_{0,0} & \mathbf{c}_{1,0} & \mathbf{0} & \mathbf{0} \\ \mathbf{c}_{0,1} & \mathbf{c}_{1,1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{c}_{2,2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{c}_{3,3} \end{bmatrix} \quad (17)$$

The off-diagonal terms are those included to represent the effect of the removal of the subband boundary due to the merger of the two subbands. The corresponding $\mathbf{R}_{\text{sys}4, [2 \ 1 \ 1]}$ and $\mathbf{P}_{\text{sys}4, [2 \ 1 \ 1]}$ matrices also include the corresponding cross-terms as in (17). We can now use these equations and their solution for $\mathbf{C}_{\text{sys}4, [2 \ 1 \ 1]}$ in order to compute the MMSE for each non-uniform subband as in (15). In the general case this will be the summation of the constituent subband terms plus an adjustment which is the covariance between the constituent desired signals (although in practice this is usually small). The desired signal variance in the l^{th} merged subband composed of I_l constituent bands, $i_{l,0}, i_{l,1}, \dots, i_{l, I_l-1}$, is given by

$$\nu_{d_l}^2 \approx \sum_{i=i_{l,0}}^{i_{l, I_l-1}} \sigma_{d_i}^2 + 2 \sum_{i=i_{l,0}}^{i_{l, I_l-2}} \mathcal{E} \{ d_i(n) d_{i+1}^*(n) \}, I_l > 1 \quad (18)$$

and $\nu_{d_l}^2 = \sigma_{d_{i_{l,0}}}^2$ for $I_l = 1$. This is approximate as we do not include terms of the form $\mathcal{E} \{ d_i(n) d_{i+j}^*(n) \}$ where $j > 2$ which are usually negligible. The subband MMSE for the non-uniform SAF system in the general case is then given by

$$J_{l, \min} = \nu_{d_l}^2 - \sum_{i=i_{l,0}}^{i_{l, I_l-1}} \mathbf{C}_{\text{sys}, i}^T \mathbf{P}_{\text{sys}, i} \quad (19)$$

where \mathbf{C}_{sys} and \mathbf{P}_{sys} are appropriately formed for the particular non-uniform decomposition in question. It can be seen that this is equivalent to the system in (16) when $I_l = 1 \ \forall l$ and equivalent to the system in (13) with $I_0 = \hat{K}$.

4. SIMULATIONS

4.1. 4-band example

A 4-constituent-band SAF system is simulated in order to verify the above analysis. An ensemble of 50 white noise and USASI inputs were applied in a system identification setup for every allowable decomposition (5 in all), with an unknown system of $S(z) = z^{-255}$. In the white input case, as expected, the subband errors and improvements from merges are uniform (except subbands $k = 0$ and $k = \hat{K} - 1$ where there is only 1 boundary and hence the error is lower). This demonstrates that the improvements from merging are in general the same for any split vectors with equal complexity (excluding subbands $k = 0$ and $k = \hat{K} - 1$). In the USASI case, by contrast, the predicted and simulated values of subband error for each split vector, shown in Tables 1 and 2 respectively, illustrate the benefits

Split Vector	Subband MMSE - predicted				Sum
[1 1 1 1]	0.1135	0.1876	0.0553	0.0046	0.3610
[2 1 1]	0.0431		0.0553	0.0046	0.1030
[1 1 2]	0.1135	0.1876	0.0517		0.3528
[2 2]	0.0431		0.0517		0.0948
[4]	1.7×10^{-6}				1.7×10^{-6}

Table 1. subband errors, USASI input, predicted

Split Vector	Subband MMSE - simulated				Sum
[1 1 1 1]	0.1133	0.1879	0.0554	0.0047	0.3613
[2 1 1]	0.0438		0.0555	0.0047	0.1040
[1 1 2]	0.1140	0.1870	0.0515		0.3525
[2 2]	0.0438		0.0509		0.0947
[4]	8.2×10^{-13}				8.2×10^{-13}

Table 2. subband errors, USASI input, simulated

of the non-uniform scheme. The values agree very closely, with differences being due to the stochastic nature of the NLMS algorithm used and the non-adjacent subband cross-terms. In this case, a split vector of [2 1 1] results in an MMSE over 5dBs smaller than the split vector of [1 1 2] which has equal complexity.

4.2. NDS example

We demonstrate the performance of NDS with a coloured input $x(n)$ having a single peak in the spectrum which falls, in certain cases, at a subband boundary. The unknown system is $S(z) = z^{-1023}$ which is realistic in length and white, which allows us to observe the differences in performance due to the input signal characteristics only. Five NDS systems with $\hat{K} = 16$ constituent bands and minimum decimation factor $\hat{M} = 4$ are compared with five uniform SAFs of $K = 8, 7, 6, 5, 4$ subbands. The maximum overall complexity (including the filterbanks) of the NDS systems is set to the complexity of each of the uniform systems. Figure 3(a) shows the performance of the uniform SAFs to be highly variable and not proportional to the complexity of the system. Figure 3(b), shows that the performance of each NDS system is robust and that the error decreases with increasing complexity (K decreasing). It was observed that the dynamic frequency decomposition converged during the first 2000 iterations to solutions that avoid placing a subband boundary in the spectral region with high input energy.

5. CONCLUSIONS

This paper has presented an analytical framework for studying NDS systems. The analysis represents non-uniform subband decompositions through the presence or absence of cross-band adaptive filters. The validity of the analysis is supported by simulations which show that the robustness of NDS to highly coloured inputs yields improvements of

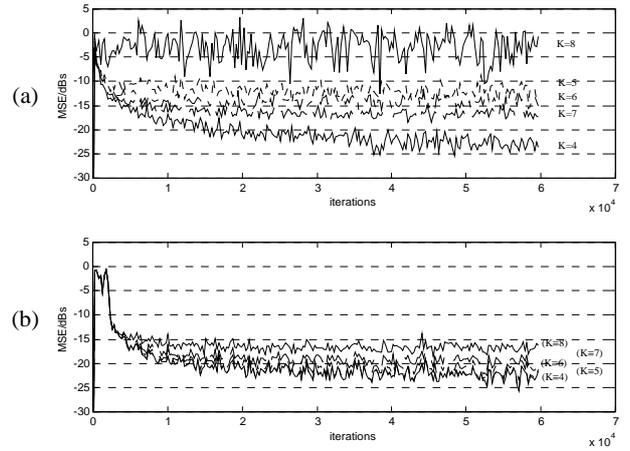


Fig. 3. Uniform and adaptive critical SAF systems - (a) uniform fixed decomposition (b) NDS with fixed maximum complexity

around 5-10dBs in MMSE in the critical case without increasing complexity.

6. REFERENCES

- [1] KELLERMAN, W.: *Analysis and design of multirate systems for the cancellation of acoustical echoes*. in Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, New York, NY, April 1988, pp. 2570-2573.
- [2] GILLOIRE, A. AND VETTERLI, M.: *Adaptive filtering in subbands with critical sampling: Analysis, experiments, and application to acoustic echo cancellation*. IEEE Trans. on Signal Processing, vol. 40, no. 8, 1992, pp. 1862-1875.
- [3] MORGAN, D.R.: *Slow asymptotic convergence of LMS acoustic echo cancellers*. IEEE Trans. on Speech and Audio Proc., vol. 3, no. 2, pp. 126-136, Mar. 1995.
- [4] OAKMAN, A. AND NAYLOR, P.: *Non-uniform dynamic structures for subband adaptive filtering*. in Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, vol. 6, May 2001
- [5] LEE, J. AND LEE, B.G.: *A design of nonuniform cosine modulated filter banks*. IEEE Trans. on Circuits and Systems-II: Analog and Digital Signal Processing, vol. 42, pp. 732-737, 1995
- [6] CHEN, Z. AND YUE, G.: *LMS Algorithm Based on Subband Decomposition*. in Proc. Int. Conf. on Communications Technology, ICCT'98, Oct. 1998, vol. 1, pp. 128-133