

# HIGHLY OVERSAMPLED SUBBAND ADAPTIVE FILTERS FOR NOISE CANCELLATION ON A LOW-RESOURCE DSP SYSTEM

*King Tam, Hamid Sheikhzadeh, Todd Schneider*

Dspfactory Ltd., 611 Kumpf Drive, Unit 200, Waterloo, Ontario, Canada N2V 1K8  
E-mails: {king.tam, hsheikh, todd.schneider}@dspfactory.com

## ABSTRACT

A real-time subband adaptive noise cancellation system on an ultra low-power miniature DSP system is implemented. The system is targeted at personal communication devices where the speaker may be in a noisy environment. The system is implemented on an ultra low-power DSP system that incorporates a DSP core and an oversampled WOLA filterbank. Pre-emphasis filters are used to increase the convergence rate of a leaky LMS algorithm in the oversampled subband implementation. System performance is also improved relative to a fullband implementation due to benefits arising from using subband adaptive filters instead of fullband filters. A 10 dB reduction of noise power is achieved in tests using various noise conditions. The entire DSP system consumes 2.1 mW and can be realized in a package size of 6.5 x 3.5 x 2.5 mm.

## 1. INTRODUCTION

The objective of this research is to implement a subband adaptive noise cancellation system on an ultra low-power, small size, and low-cost platform. The system is targeted for telecommunication (e.g., headsets or mobile phones) or mobile speech recognition applications, where the user is talking in the presence of interfering noise. A robust system should provide significant noise cancellation, fast algorithmic convergence in colored noises, short group delay, and minimal introduction of artifacts into the speech signal. Furthermore, it should have low computational cost and complexity, low memory usage, low power requirements, and small physical size.

It is well known that a noise cancellation system can be implemented with a fullband adaptive filter working on the entire frequency band of interest [1]. The Least Mean-Square (LMS) algorithm and its variants are often used to adapt the fullband filter with relatively low computation complexity and good performance. However, the fullband LMS solution suffers from significantly degraded performance with colored interfering signals due to large eigenvalue spread and slow convergence [2]. Moreover, as the length of the LMS filter is increased, the convergence rate of the LMS algorithm decreases and computational requirements increase. This can be a problem in applications, such as acoustic echo cancellation, that demand long adaptive filters to model the return path response and delay. These issues are especially important in portable applications, where processing power must be conserved.

As a result, subband adaptive filters (SAF) become a viable option for many adaptive systems. The SAF approach uses a filterbank to split the fullband signal input into a number of frequency bands, each serving as input to an adaptive filter. The

subband decomposition greatly reduces the update rate and the length of the adaptive filters resulting in a much lower computational complexity. Further, subband signals are often decimated in SAF systems. This leads to a whitening of the input signals and an improved convergence behavior [3]. If critical sampling is employed, the presence of aliasing distortions requires the use of adaptive cross-filters between adjacent subbands or gap filterbanks [3,4]. However, systems with cross-filters generally converge slower and have higher computational cost, while gap filterbanks produce significant signal distortion. Oversampled SAF systems offer a simplified structure that without employing cross-filters or gap filterbanks, reduce the alias level in subbands. To reduce the computation cost, often a close to one non-integer decimation ratio is used [5].

In this research we propose a SAF system based on generalized DFT (GDFT) filterbanks. The filterbank is a highly oversampled one (oversampling by a factor of 2 or 4). Due to the ease of implementation, low-group delay and other application constraints (explained in Section 3), we chose a higher oversampling ratio than those typically proposed in the literature. The convergence behavior due to the high oversampling rate is analyzed and properly addressed. An LMS-based version of the proposed SAF system is implemented on a DSP system that includes an oversampled filterbank. The DSP system [6,7] has a configurable oversampling rate of 2 or 4. The added computational cost due to sampling the subband signals at a frequency higher than the critical sampling frequency is compensated by the efficiency of the hardware architecture, which has a filterbank coprocessor dedicated to performing subband decomposition of the input signals.

In the following sections, we first present a description of this DSP architecture. We then describe the adaptive noise canceller structure. Finally, a conclusion of the research and the future work is presented.

## 2. THE DSP SYSTEM

Figure 1 shows a block diagram of the DSP system [6,7]. The DSP portion consists of three major components: a weighted overlap-add (WOLA) filterbank coprocessor, a 16-bit block-floating point DSP core, and an input-output processor (IOP). The DSP core, WOLA coprocessor, and IOP run in parallel and communicate through shared memory. The parallel operation of the system allows for the implementation of complex signal processing algorithms in low-resource environments with low system clock rates. The system is especially efficient for

subband processing since the configurable WOLA coprocessor splits the fullband input signals into subbands, leaving the core free to do the adaptive processing on the subband signals.

The core has access to two 4-kword data memory spaces, and another 12-kword memory space used for both program and data. The core provides 1 MIPS/MHz operation and has a maximum clock rate of 4 MHz at 1 volt. At 1.8 volts, 30 MHz operation is also possible. The system operates on 1 volt (i.e., from a single battery). With a system clock rate of 1.28 MHz, it consumes less than 1 mW of power.

The system is implemented on two ASICs. A separate off-the-shelf E<sup>2</sup>PROM provides the non-volatile storage. The chipset can be packaged into a 6.5 x 3.5 x 2.5 mm hybrid circuit.

The system is clocked at a rate of 5.12 MHz for this application. The sampling rate is 16 kHz. Power consumption is 2.1 mW.

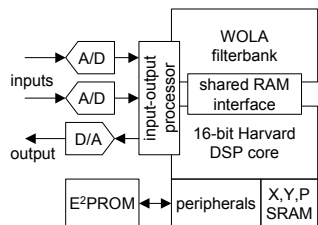


Figure 1: The DSP system block diagram

### 3. SUBBAND ADAPTIVE NOISE CANCELLATION

The SAF system is implemented on DSP system described in Section 2. The adaptive noise cancellation algorithm uses a 16-band stereo configuration of the WOLA filterbank, with an oversampling factor of 2 or 4. For many applications the low group delay requirement does not allow long analysis time-windows. Consequently, high oversampling factors are used to minimize the aliasing distortion found in systems with critical sampling or low oversampling. This results in near-orthogonal subbands, where energy leakage between adjacent bands is small. As a result, prototype filter design constraints become less stringent. As discussed in [6,7], wide gain adjustment of the subband signals leads to considerable distortion in filterbanks with low oversampling ratios. However, it is quite feasible for the WOLA filterbank to apply wide gain adjustment without generating audible distortions.

Figure 2 shows a block diagram of the subband adaptive noise canceller. The system has two inputs: one for the primary signal (voice from speaker with interfering noise), and one for the reference signal (noise only). The signals are received from microphones that are physically placed for good separation of the signals, but not so far apart as to make the transfer function between microphones too complex to be modeled by the adaptive system. For a headset with a boom, the speech microphone is placed close to the speaker's mouth on the inside of the boom and the reference microphone is placed on the opposite side of the boom facing away from the speaker. Each input signal is passed through the analysis filterbank and split into uniform subbands. The analysis filterbank efficiently

decimates the subband signals. The subband processing blocks cancel the noise in the output signal by using a variant of the LMS algorithm that is described in Section 3.2. The subband processing blocks are shown in detail in Figure 3.

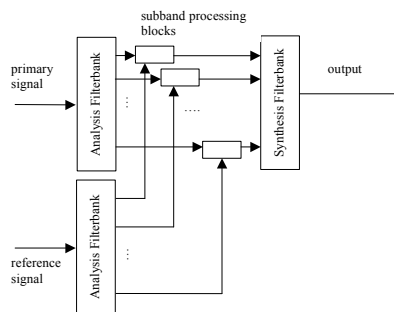


Figure 2: Subband adaptive noise canceller

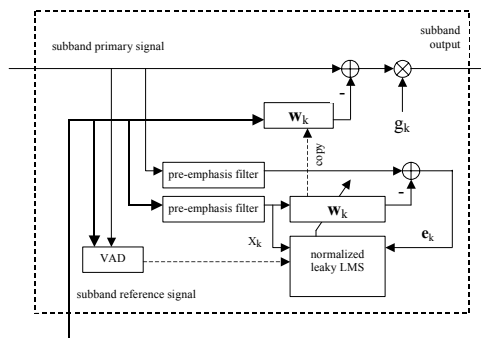


Figure 3: Subband processing block for adaptive noise canceller

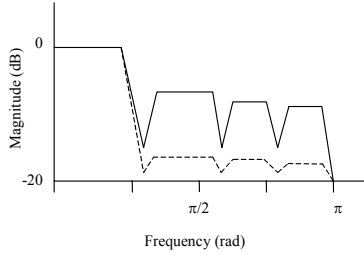
#### 3.1. Pre-emphasis Filters

The oversampled input signals received by the subband processing blocks are no longer white in spectrum. In fact, for oversampling factors of 2 and 4, their bandwidth will be limited to  $\pi/2$  and  $\pi/4$  respectively. As a result, one would expect a slow convergence rate due to eigenvalue spread problem [2]. On the other hand, while the oversampled subband signals are not white, their spectra are colored in a predictable way and can therefore be modified by fixed filters to “whiten” them in order to increase the convergence rate. Thus, the inherent benefit of decreased spectral dynamics resulting from subband decomposition is not lost due to oversampling.

Figure 4 shows a simplified representation of the subband spectra corresponding to white noise input into the filterbank, for a 4-times oversampled configuration. The dashed line shows the spectrum without pre-emphasis. As shown, nearly all the signal power is in the lower quarter of the spectrum. The signal power present in the upper three quarters of the spectrum is decided by the frequency response of the filterbank's prototype low-pass analysis filter.

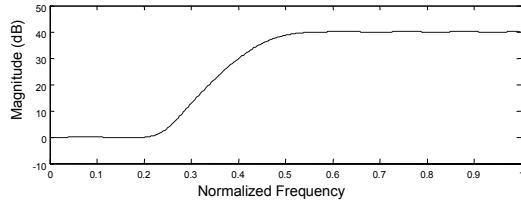
We employ a pre-emphasis filter for each subband to amplify the low-level signal components in the high three quarters of the spectrum to flatten the spectrum, thereby reducing the

signal's autocorrelation matrix eigenvalue spread, and increasing convergence rate. Figure 5 shows the frequency response of a typical pre-emphasis filter employed in the system. The solid line in Figure 4 corresponds to the spectrum of the subband signal after pre-emphasis. The emphasized subband signals are used only for improving the convergence characteristics of the adaptive filters. As shown in Figure 3, in each subband, the adaptive filter coefficients are copied to a mirror filter that processes the non-emphasized version of the signal to obtain the noise-cancelled signal for synthesis.



**Figure 4:** Simplified subband spectrum before pre-emphasis (dashed line) and after pre-emphasis (solid line)

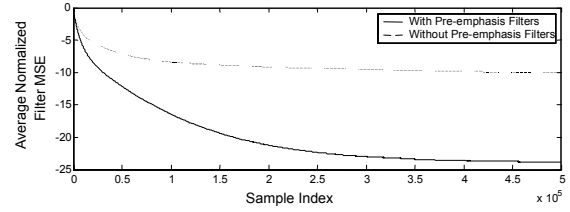
Figure 6 illustrates the change in convergence using a long sequence of white noise input samples into the 16-band WOLA filterbank using an oversampling factor of 4. MATLAB simulations are run with a known finite impulse response system in place to simulate the transfer function between two microphones. The LMS filter mean-squared error (MSE) is the averaged squared difference between the 5 adaptive filter coefficients and the known optimum solution. This value is normalized such that the initial zero values of the adaptive coefficients corresponds to a MSE of 0 dB. The normalized filter MSE is then averaged across the 16 subbands. Note that Figure 6 merely illustrates the difference in average MSE for the finite input sequence; both systems will ultimately converge to the same steady state solution.



**Figure 5:** Pre-emphasis filter response

### 3.2. Subband Adaptation Algorithm

The filter in the  $k$ th subband,  $\mathbf{w}_k$ , is adapted according to equation (1), where  $n$  is the time index,  $\mu_k$  is the LMS step-size parameter,  $e_k$  is the error signal,  $L$  is the adaptive filter length,  $\mathbf{x}_k$  is a vector containing the last  $L$  complex samples of emphasized subband reference signal  $X_k$ ,  $\sigma_k^2$  is an estimate of the power of  $X_k$ , and  $\epsilon$  is a small constant used to avoid division by zero. The normalized and “leaky” variant of the complex LMS algorithm is chosen to ensure stability and convergence to a unique solution [8].



**Figure 6:** Effect of pre-emphasis filter on adaptive filter convergence

$$\mathbf{w}_k(n+1) = (1 - \gamma\mu_k) \cdot \mathbf{w}_k(n) + \frac{\mu_k \cdot \mathbf{x}_k(n) \cdot e_k^*(n)}{L \cdot \sigma_k^2(n) + \epsilon} \quad (1)$$

It is possible to vary the subband LMS parameters such as filter length and LMS step-size parameter  $\mu$ , independent of parameters of adjacent bands since the bands are almost orthogonal. As described below, we have implemented a system with varying values of  $\mu_k$ , constant leakage factor  $\gamma$  across all bands, and 5 complex coefficients for each adaptive filter.

The values for  $\mu_k$  are chosen such that peak noise cancellation in slowly varying noise is achieved within approximately 5 seconds. Faster convergence is possible by increasing  $\mu_k$ , but it comes at the cost of increased artifacts in the enhanced speech. In bands beyond 4 kHz, the filters are more aggressively adapted using increasing values for  $\mu_k$  since the higher bands contain less speech energy and therefore there is less distortion introduced by quickly adapting filters.

The leakage factor  $\gamma$  effectively adds white noise to the input signal and ensures convergence to a unique solution [8]. It also allows the filters to re-initialize themselves by slowly leaking to zero in the absence of input  $X_k$ .  $\gamma$  is chosen such that the factor  $(1 - \gamma\mu_k)$  is very close to 1. This keeps the filter coefficient bias created by using leaky LMS to an acceptable value, while still adding some whitening effect.

The filter length is chosen as a compromise between computational requirements and the system's ability to model the physical system between primary and reference microphones. Filters that are too long will use up all available processing power and will lead to slow convergence. Filters that are too short will result in a truncated model of the system between microphones, and therefore limit the degree of noise cancellation. Since the adaptive filters in our system operate in a decimated domain and are comprised of complex coefficients, they combine to model a fullband system with a comparably more complex response. The 5 complex coefficients per adaptive filter provide adequate modeling capability, while conserving processing resources.

The existence of multiple filters allows the filter updating to be interleaved across successive time slots for efficiency. For example, grouping the subbands into 2 groups of 8, then updating alternate groups at every time slot reduces the computational requirements per time slot by a factor of 2. The

power estimate  $\overline{\sigma_k^2}$  is calculated using a first-order IIR smoothing filter with a time constant of approximately 1 ms.

The constant gain factors  $g_k$  (see Figure 3) are used to scale the noise-cancelled signal before it reaches the subband output and subsequently enters the synthesis stage. We have found that the undesirable leakage of the speech signal into the reference signal in practical systems causes some inadvertent cancellation of speech, particularly in the low frequencies. The static gain factors are set to compensate for this mild low frequency loss. Also, in real-time hardware implementation (reported in Section 4), these gains can be used for microphone equalization.

An optional voice activity detector (VAD) freezes the adaptation of the filters when speech is present. The VAD is particularly useful in physical configurations where microphones are placed such that the speech signal easily leaks into the reference signal. The contamination of the reference signal hinders convergence of the filters. This is avoided by allowing the filters to adapt only when the VAD has detected a pause in speech. The VAD calculates the power in a low band-group and a high band-group. It tracks the changes in the ratio of these powers in order to detect the presence of speech in the primary signal. It is designed to have a bias towards over-detection (false alarms) rather than under-detection (missed speech). A hangover counter is used to prevent the misclassification of trailing portions of speech as noise or silence, thereby improving the reliability of pause detection. Testing shows that activation of the VAD slows down the convergence but does not affect the degree of noise cancellation achieved after convergence.

#### 4. PERFORMANCE EVALUATION

Off-line evaluation tests have been completed for various types of noise (white, pink, car, airplane, babble, and similar noises) in the presence of speech. Table 1 shows the results of a comparison of simulated fullband (128-coefficient FIR) and subband (16 x 8-coefficient FIR) systems using the same input length. The primary input has a 0 dB signal-to-noise ratio (SNR) with no speech leakage to the reference input. The algorithm parameters (filter length,  $\mu_k$  and  $\gamma$ ) are chosen for each system such that SNR improvement in white noise is similar. The results illustrate how the subband implementation performs consistently for various noise conditions, while the fullband implementation does not. As evident from the table, the proposed SAF has a superior performance for both non-stationary (like babble noise) and colored noises (like pink noise) due to the whitening effect of the SAF system and a faster convergence. Informal listening shows very little audible distortion of speech.

A real-time version of the proposed SAF system is implemented on the DSP system described in Section 2. The preliminary results using a variety of double-microphone boom-style headsets show an average improvement (for different types of noise with input SNR in 0-5 dB range) in SNR of 10 dB on a real system. This is promising considering the effects of implementation on a 16-bit block-floating-point system using a real headset that permits leakage of speech into the reference microphone.

**Table 1:** Comparison of simulation results for fullband and subband systems

	SNR improvement (dB)	
	Fullband system	Subband system
<b>White noise</b>	25.5	25.7
<b>Pink noise</b>	18.7	25.3
<b>Airplane noise</b>	17.3	23.0
<b>Babble noise</b>	16.4	25.2
<b>Traffic noise</b>	17.4	25.2
<b>Car noise</b>	20.7	25.6

#### 5. CONCLUSIONS AND FUTURE WORK

An SAF noise cancellation system was developed for a highly oversampled filterbank. The system was implemented on an ultra low-resource platform. To improve the convergence rate, we proposed and implemented pre-emphasis filters to improve the performance of the adaptive subband-LMS algorithm. In real-life environments, the noise cancellation system delivers approximately 10 dB reduction of noise power with little distortion of speech, while requiring modest resources in terms of space and power. It performs well in colored noise and shows faster convergence than a fullband implementation. No other system known to the authors delivers such performance with such small size and low power consumption.

Future work will include a complete evaluation of our real-time system and investigation of optimal design criteria for the pre-emphasis filters, as well as alternate means of subband signal whitening. Also, more research can be done to explore the usage of other adaptation strategies on the DSP system.

#### 6. REFERENCES

- [1] B. Widrow et al., "Adaptive noise cancellation: Principles and applications". *Proc. IEEE*, vol. 63, no. 12, Dec. 1975.
- [2] Haykin, S., *Adaptive Filter Theory*. Prentice Hall, Upper Saddle River, 3<sup>rd</sup> Edition, 1996.
- [3] A. Gilloire and M. Vetterli, "Adaptive Filtering in Subbands with Critical Sampling: Analysis, Experiments and Applications to Acoustic Echo Cancellation". *IEEE Trans. Signal Processing*, vol. SP-40, no. 8, pp. 1862-1875, Aug. 1992.
- [4] J. J. Shynk, "Frequency-Domain and Multirate Adaptive Filtering". *IEEE Signal Processing Magazine*, pp. 14-37, Jan. 1992.
- [5] S. Weiss, "On Adaptive Filtering in Oversampled Subbands", PhD. Thesis, Signal Processing Division, University of Strathclyde, Glasgow, May 1998.
- [6] R. Brennan and T. Schneider, "Filterbank Structure and Method for Filtering and Separating an Information Signal into Different Bands, Particularly for Audio Signal in Hearing Aids". United States Patent 6,236,731. WO 98/47313. April 16, 1997
- [7] R. Brennan and T. Schneider, "A Flexible Filterbank Structure for Extensive Signal Manipulations in Digital Hearing Aids", *Proc. IEEE Int. Symp. Circuits and Systems*, pp.569-572, 1998.
- [8] Hayes, M., *Statistical Digital Signal Processing and Modelling*, John Wiley & Sons, Inc., New York, 1996.