LOW-POWER IMPLEMENTATION OF A SUBBAND FAST AFFINE PROJECTION ALGORITHM FOR ACOUSTIC ECHO CANCELLATION

David Hermann, Etienne Cornu, Tina Soltani and Hamid Sheikhzadeh

AMI Semiconductor
611 Kumpf Dr. Unit 200, Waterloo, Ontario, Canada N2V 1K8
phone: +1 (519) 884 9696, fax: +1 (519) 884 0228, email: david_hermann@amis.com
web: www.amis.com

ABSTRACT

Recent results have demonstrated the performance benefits of using Fast Affine Projection (FAP) algorithms for echo cancellation in oversampled filterbanks. Low-power, low resource implementations of these algorithms are useful for solving acoustic echo problems in size and resource constrained applications such as portable wireless headsets and portable or low-power hands-free phones. This paper describes a novel real-time subband implementation of the Gauss-Seidel Fast Affine Projection (GSFAP) algorithm. The algorithm uses online regularization to eliminate the need for voice activity detectors and greatly simplifies the calibration and deployment of the system. The algorithm is deployed on a DSP system with three processing units, including a filterbank co-processor customized for subband adaptive filtering. The resulting system consumes less than 7 mW while providing up to 30 dB of echo return loss enhancement in a typical hands-free speakerphone application.

1. INTRODUCTION

Acoustic echo is a problem in communications devices with significant acoustic leakage between a speaker and microphone. In small wireless headsets this leakage occurs because of the poor isolation between the speaker and microphone. In hands-free or speakerphone applications, the near-end user speaks into a microphone, but the receiver volume is generally of equal or similar volume and thus is also picked up by the microphone. This leads many speakerphone applications to use only half-duplex communications, degrading the quality of the conversation because one end is muted while the other is talking. In addition to the acoustic echo problem, many audio applications typically require a flexible set of other algorithms for features such as noise reduction and dynamic range compression. This variety of algorithms and the accompanying level of customization require a programmable and flexible audio DSP solution. This DSP solution must also be small and low-power to work within the limited resources of these applications.

Previous work has demonstrated the use of such a DSP system for echo cancellation in Bluetooth based wireless headsets [1]. This solution uses a subband LMS-based algorithm with a highly oversampled filterbank. This approach includes convergence improvements such as spectral whitening by decimation and spectral pre-emphasis filtering. It also uses multiple voice-activity detectors (VADs) to control algorithm adaptation. This implementation works well for short-path echo environments (< 16 ms) like those in headsets, but because of computational and memory constraints it is of limited use in longer echo path environments (> 32 ms) such as those seen when using a hands-free phone inside a car or an office.

In this paper we present a new and novel solution based on the same DSP system as [1] but using an algorithm based on the Gauss-Seidel Fast Affine Projection (GSFAP) [2]. Affine projection algorithms have been shown to be more effective in oversampled subband adaptive filters than LMS-based approaches because of the colouration in the subband signals [3]. Additionally, the use of on-line regularization to control algorithm adaptation eliminates the need for complicated and difficult to calibrate VADs [4]. The resulting algorithm has a simplified structure and reduced memory requirements allowing it to manage longer echo paths on the same DSP system. Careful implementation of the algorithm on the DSP is still required in order to maximize performance in low-power applications.

The DSP system used for this implementation is described in Section 2 including the selection and design of a suitable oversampled filterbank configuration. Section 3 describes the GSFAP-based subband adaptive filter algorithm and its on-line regularization scheme in more detail. Section 4 maps this algorithm to the DSP architecture and describes some of the implementation details. Section 5 presents some experimental results for the system based on a low-power speakerphone application. Finally a discussion of conclusions and future work is presented in Section 6.

2. DSP SYSTEM

The DSP system is composed of a general-purpose 16-bit DSP core, a dedicated input/output processor (IOP), and a weighted-overlap add (WOLA) filterbank co-processor [5]. The entire system is designed for efficient subband signal processing. The three processing units work in parallel to implement complex signal processing tasks in low-resource applications. A block diagram of the system is shown in Figure 1. In this architecture the IOP manages the input and output of time-domain audio samples, while the WOLA filterbank co-processor transforms the time-domain signals to and from the subband domain. The general-purpose DSP core can then be used to implement different subband processing algorithms. The entire system is capable of handling the dual input and output streams required for adaptive signal processing through efficient stereo operations and data management.

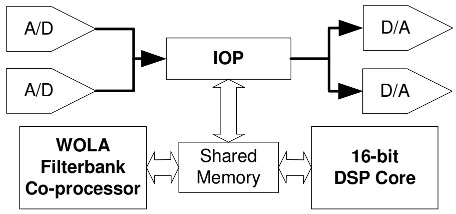

Figure 1 : DSP System Block Diagram

The WOLA filterbank co-processor provides a flexible implementation of a highly oversampled, complex-modulated filterbank. For many audio-processing algorithms, high oversampling factors (> 2) are desirable to allow large gain adjustments in the individual subbands while keeping the analysis and synthesis filter lengths shorter to reduce group delay. However, smaller oversampling factors are advantageous for subband adaptive filtering because they reduce the colouration of the subband spectrum. These requirements may come into conflict when different algorithms must be integrated together.

In order to improve the convergence behaviour and performance of the subband adaptive filters for high oversampling, we employ wider bandwidth analysis filters [6]. A wider bandwidth analysis filter provides additional excitation of the subband signal in the transition band, thereby improving the subband adaptive filter performance. The synthesis filtering then removes these unnecessary components.

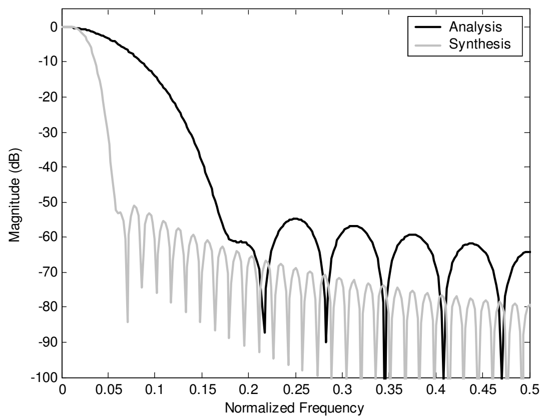

Figure 2 : Filterbank Prototype Filter Frequency Responses

For this application, we selected a filterbank configuration with 16 subbands and a subband oversampling ratio of 4. The synthesis prototype filter is 128 samples long and has the nominal bandwidth for 16 frequency bands. The analysis prototype filter is 32 samples long with a bandwidth twice that of the synthesis window and a wider transition band. The frequency responses of the prototype filters are shown in Figure 2. These filters were designed with windowing using a slightly modified Hamming window. This filterbank configuration provides a low group delay of only 11 ms at a sampling frequency of 8 kHz.

3. ALGORITHM DESCRIPTION

The block diagram for the GSFAP echo cancellation algorithm is shown in Figure 3 and the GSFAP adaptive processor unit is shown in Figure 4. Because the algorithm is a fast version of the affine projection algorithm, the filter coefficients, $h_k(m)$ in Figure 4 are not actually available and the diagram shows only the conceptual flow of the signal processing and adaptation. The GSFAP algorithm that is implemented on the DSP system is the so-called Low-Cost GSFAP algorithm (LC-GSFAP) from [4]. LC-GSFAP uses a sequential partial filter update mechanism to further reduce the complexity of the GSFAP algorithm, particularly for larger filter orders. The algorithm implementation takes advantage of the partial filter updates used in LC-GSFAP to reduce the computational complexity of the adaptive processing.

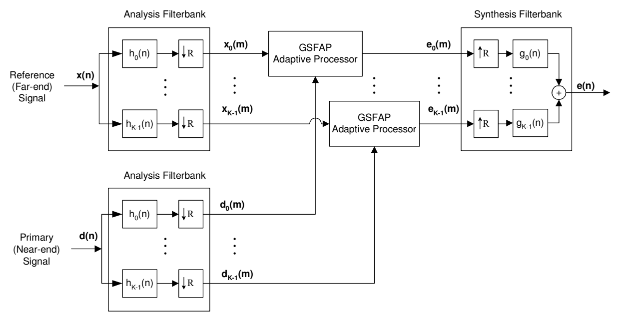

Figure 3 : Block Diagram of Subband Adaptive System

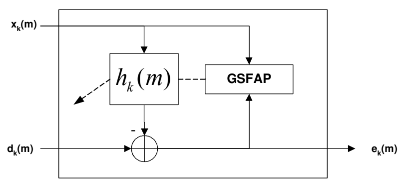

Figure 4 : GSFAP Adaptive Processor

Because of the limited processing power available in low-power applications, a very low order affine projection is required. Even an affine projection of order $N=2$ provides significant performance improvements over LMS-based approaches for oversampled subbands [3]. The algorithm implementation uses an order of $N=2$ and a partial filter update decimation factor of $D= 2$ (i.e., only half of the filter coefficients are updated each block). It has a configurable number of taps (L) within the constraints of the DSP system's available memory and the application requirements. For the filterbank configuration described in Section 2, the maximum length of the echo path that can be modelled by the adaptive system is 64 ms although longer paths can be modelled if the subband oversampling ratio is lowered.

The GSFAP algorithm uses on-line regularization to stabilize the matrix inversion that is part of any affine projection algorithm, and to control adaptation in the presence of near-end disturbances [7]. The regularization parameter for block index m and band k, called $\delta_k(m)$, is estimated using the formula given below.

$$\delta_k(m) = L \cdot \overline{\max\left\{D \cdot (N-1) \cdot \overline{|x_k(m)|^2}, \overline{|d_k(m)|^2}\right\}} \qquad (1)$$

The over-bars in equation (1) indicate first-order attack-release averaging with an instantaneous attack and an appropriately chosen release time constant. The instantaneous attack ensures the stability of the adaptation when a fast increase occurs in the power of either the primary or reference signal. The release time constant of the signal power averaging ($\overline{|x_k(m)|^2}$ and $\overline{|d_k(m)|^2}$) is chosen to be equal to the subband filter length and the release time constant of the overall regularization parameter is chosen to be in the neighbourhood of 0.5-1.0 seconds so that the parameter has some hold-over time between typical short speech pauses [7]. This online regularization parameter efficiently controls the adaptation of the algorithm without the need for an explicit voice activity detector. The computational complexity of calculating this parameter is also an order of magnitude less than that of high-quality VADs.

4. SYSTEM IMPLEMENTATION

This section outlines how the subband adaptive algorithm of Section 3 is mapped to the DSP system from Section 2. Figure 5 shows a diagram of the real-time processing in the system and how it is divided across the three processors.

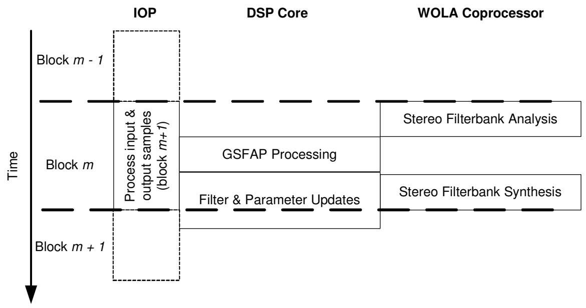

Figure 5 : Algorithm Processing Flow

The IOP runs continuously, collecting input samples for the next block and outputting results from the previous block in an efficient stereo interleaved format. Input and output samples are stored in separate dedicated first-in, first-out (FIFO) blocks of memory. The WOLA filterbank coprocessor operates directly on these two audio streams by computing the analysis and synthesis filterbanks using an efficient stereo filterbank implementation. The output signal is calculated immediately after the analysis filterbank completes so that the synthesis filterbank can start as soon as possible. The DSP core executes the GSFAP algorithm using one channel as the primary signal and the other channel as the reference signal. Additional non-signal path algorithm processing such as internal parameter updates then occurs in parallel with the filterbank processing, thereby maximizing the system's efficiency.

The real-time implementation of the algorithm includes other optimizations to reduce the computational load and improve the efficiency of the system. In particular, the Gauss-Seidel (GS) recursion is done using an optimized mix of 32-bit and 16-bit variables to maintain sufficient precision in the recursion. However this mixed precision requires significant computational resources on a 16-bit platform even for low-order matrix inversions. To reduce the computational load, this portion of the adaptation process is decimated across the bands. We compute the GS recursion only every M blocks, where M is configurable integer value based on the resources available. The decimation is staggered across the bands to maintain an even distribution of the computational load over time. Figure 6 shows how this update scheme is used for 16 bands and an update factor of $M=4$.

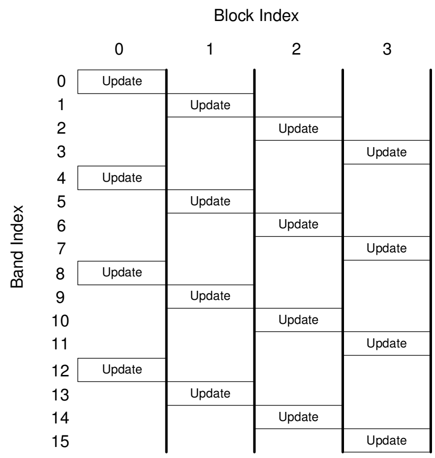

Figure 6 : Staggered Adaptation Update Scheme for $M=4$

Using the efficient GSFAP algorithm and the optimizations of the real-time implementation, it is possible to deploy the algorithm in an extremely resource-constrained environment. The final computational and memory requirements of the real-time GSFAP algorithm are scalable depending on the target system model size, i.e., the echo path to be modelled for echo cancellation. Table 1 shows the minimum system clock and memory requirements of the algorithm for two different echo path lengths. The DSP is powered by a 1.8 V supply and consumes 4.8 mW for a 16 ms echo path or 6.7 mW for a 64 ms echo path when using a fixed system clock frequency of 24 MHz.

Table 1: System Clock and Memory Requirements

Echo Path	Min. System Clock (MHz)	Program Memory (16-bit words)	Data Memory (16-bit words)
16 ms	12.3	1.2 k	1.8 k
64 ms	21.7	2.2 k	4.5 k

5. SYSTEM EVALUATION

The GSFAP subband adaptive system was tested and evaluated in the context of a low-power speakerphone environment. Full-duplex speakerphone and hands-free phone devices require acoustic echo cancellation because of the significant leakage between the receiver and the microphone. The setup for this testing is shown in Figure 7.

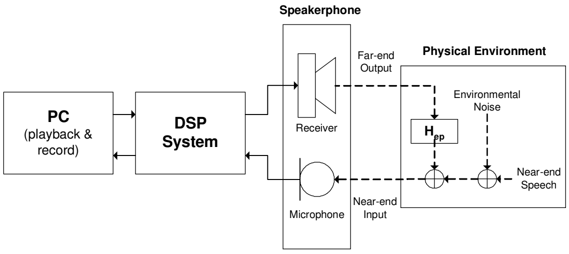

Figure 7 : Speakerphone Test Setup

A reference signal is input into the DSP from a PC and the output of the echo canceller is recorded. The resulting near-end signal is thus only the echo signal filtered through the acoustic echo plant (H_{ep}) plus any noise from the testing environment. Recordings were done with the speakerphone in a room approximately 2.5 m × 2.5 m in size representative of a typical personal office. The SNR of the environment is approximately 30-35 dB (i.e., a quiet room). The algorithm was configured with the filterbank configuration described in Section 2, an echo path model of 64 ms (*L=64*), partial filter update factor of *D=2,* and a GS recursion update factor of *M=2* blocks.

Figure 8 shows the resulting echo return loss enhancement (ERLE) over time based on recordings in this test setup using a white noise input. This shows the real-time performance of the algorithm in typical physical environment, including about 30 dB of ERLE once the initial convergence is achieved. The 30 dB performance limit is caused by the fixed-point nature of the algorithm implementation and the SNR of the environment. Applying the same recordings to a full floating point simulation of the same algorithm produces 35 dB of ERLE over the same timeframe indicating that the real-time, fixed-point implementation only degrades the performance by approximately 5 dB as compared to the simulation.

The speakerphone setup was also used to test the performance of the algorithm with speech inputs. The average ERLE for speech segments was about 20-25 dB for single-talk echo (i.e., no near-end speech disturbance).

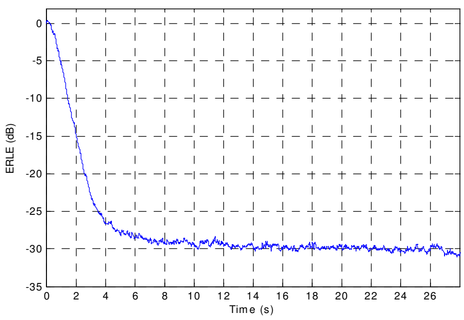

Figure 8 : ERLE for White Noise Input in a Speakerphone Environment

6. CONCLUSIONS

This paper has presented a subband fast affine projection algorithm for echo cancellation, implemented on a low-power DSP system optimized for subband processing. The filterbank was customized to provide improved subband adaptive filter performance through the use of a wider bandwidth analysis filter. The entire system was tested in a real-world speakerphone application for acoustic echo cancellation. It provides as much as 25 dB of ERLE during normal speech and at least 30 dB of ERLE for white noise test conditions. These results can be achieved with less than 7 mW of power consumption.

The convergence speed of the algorithm implementation can be further improved through the implementation of higher order affine projections that compensate better for the oversampling present in the filterbank. There is also ongoing work on a complete echo cancellation solution including residual echo suppression and transmit noise reduction. Deploying the echo cancellation system on a flexible, programmable DSP allows the integration of these additional audio processing algorithms. The GSFAP subband adaptive filter architecture is also being investigated for other adaptive filtering applications such as two-input noise reduction and feedback cancellation.

REFERENCES

[1] J. Johnson, E. Cornu, G. Choy, and J. Wdowiak, "Ultra Low-Power Sub-band Acoustic Echo Cancellation for Wireless Headsets", in *Proc. ICASSP 2004*, Montreal, Canada, May 17-21. 2004, Vol. 5, pp 357-360

[2] F. Albu, J. Kadlec, N. Coleman , and A. Fagan, "The gauss-seidel fast affine projection algorithm", in *Proc. IEEE Workshop on Signal Processing Systems*, October 16-18. 2002, pp 109-114

[3] H. R. Abutalebi, H. Sheikzadeh, R.L. Brennan, and G. H. Freeman, "Affine Projection Algorithm for Oversampled Subband Adaptive Filters," in *Proc. ICASSP 2003*, Hong Kong, China, April 6-10. 2003, Vol. 6, pp 209-212

[4] E. Chau, H. Sheikzadeh, and R.L. Brennan, "Complexity Reduction and Regularization of a Fast Affine Projection Algorithm for Oversampled Subband Adaptive Filters", in *Proc. ICASSP 2004*, Montreal, Canada, May 17-21. 2004, Vol. 5, pp 109-112

[5] T. Schneider, R. Brennan, P. Balsinger, and A. Heubi "An ultra low-power programmable DSP system for hearing aids and other audio applications", in *Proc. ICSPAT 1999*, Orlando, USA, November 1-4. 1999

[6] P. L. De León, II and D. M. Etter, "Experimental Results with Increased Bandwidth Analysis Filters in Oversampled, Subband Acoustic Echo Cancelers," *IEEE Signal Processing Letters*, vol. 2, no. 1, pp. 1-3, Jan. 1995.

[7] H. Sheikzadeh, R. L. Brennan and K. R. L. Whyte, "Near-end distortion in oversampled subband adaptive implementation of affine projection algorithms", in *Proc. EUSIPCO 2004,* Vienna, Austria, September 6-10. 2004