Low-Power And High Speed Robust Frequency-Domain Echo Cancellation On FPGA

Y. Sravani, K. Rambabu, A. Ranjit Kumar

ABSTRACT:- Echo cancellation is one unavoidable module in any voice related communication systems such as telephone, mobile and VOIP. In several applications run time high speed echo cancellation can give better quality of service. Real time echo cancellation is an important feature for hands-free operation of telecommunication equipment like mobile phones. A desirable acoustic echo control should be capable of handling double-talk as well. In this paper, we successfully implement a novel hardware architecture that is based on a robust adaptive algorithm in combination with a two-path model to tackle the double-talk situation. The echo canceller is working in the frequency domain and is improved by bit-width optimization to enhance computational efficiency. In experiments, our implementation of the hardware acceleration of the echo-canceller is fast and outperforms common software implementations running on microprocessors: an implementation with 4 instances of the filter on a Xilinx XC4VFX60 FPGA running at 137MHz can run 40 times faster than software on a 3.2GHz Core 2 Duo Pc. Besides, the hardware acceleration also reduces 90% of the power consumption when compared to a pure soft-core implementation. Our results suggest that the employed hardware architecture is also very energy-efficient.

Key words:- Echo cancellation, Telecommunication, Adaptive algorithm, Bit width, Xilinx.

1 INTRODUCTION

Echo can arise at any points in a voice communication networks, such as hands-free telephony or VoIP. Without proper echo cancellation, it can cause significant degradation in conversation quality. Fundamentally adaptive filters are employed to identify the echo path and cancel the echo. In a normal office room kind of environment the reverberation time will be several 100 ms, which corresponds to several hundred samples at a 8 kHz sampling rate. If the samples difference is large in number then the echo cancellation become more complicated. Another problem is the presence of a strong near-end signal, which is called as double-talk. Under this situation, the adaptive algorithm mistakes the near end signal as an echo and adjusts the filter coefficients in an inappropriate manner. This diverges the algorithm and the echo cancellation fails. In FPGA design for an echo-canceller [2] [3], the problems with double-talk are often not considered. Moreover, the performance of the echo-canceller is greatly degraded by the calculations in time domain, which is often a choice for simplicity. Computations in frequency domain are apparently better in terms of functionality. Power consumption has now been one of the issues when designing applications for embedded systems. Many application domains, including digital signal processing applications, demand longer battery life to deliver better user experience, especially in wireless communication environment. It leads the motivation to invent new techniques to reduce power and energy consumption. Power consumption can be divided into two categories. Static power consumption is due to current leakages of the transistors, while dynamic power consumption is due to switching activity of the circuits. The higher clock frequency, the more dynamic power is dissipated. There are some traditional approaches to reduce the static power consumption such as power gating, where part of the transistors are shut down when they are idle. However, if the circuits are processing data, power gating approach cannot reduce dynamic power consumption. In this paper, we propose a low-power echo-cancellation system based on FPGA device. While traditional echo cancellation system demands high-performance processor which runs at high clock rate (usually more than 3.0GHz) and long processing time, the proposed system exploits FPGA parallelism so it can operate in much lower clock rate while deliver better performance than traditional processor. To summarize, the key contributions of this paper include:

1) A complete and efficient hardware acceleration of the robust two-path sub-band echo-canceller. Our experiments show that the hardware acceleration on a high end Xilinx FPGA can run 30 times faster than a pure software counterparts running on 3.2 GHz Core 2 Duo PC.
2) Power measurement that reveals substantial power usage reduction provided by hardware acceleration implemented on FPGA. In our experiments, the implementation with hardware acceleration reduces 90% of the overall power consumption.

2 BACK GROUND

Our previous work [4] has described the algorithm for echo-cancellation; the following is the summary of the proposed Robust Two Path Adaptive Filter (RTPAF) algorithm. Consider transmitting signals over hands-free telephony systems. Let x(n) be the input calibration signal to the system and y(n) be the return signal. Without echo cancellation, the return signal can be written as

$$y(n) = \sum_{k=0}^{\infty} h(k)x(n-k) + v(n) \quad (1)$$

Where V(n) is the background noise plus the possible speech of the near-end speaker. The echo cancellation is achieved by finding an estimate of the echo and subtracting it from the return signal. Let y(n) be the estimate of the echo, it can be written as

$$\hat{y}(n) = \sum_{k=0}^{\infty} \hat{h}(k)x(n-k) \quad (2)$$

Where H(k) is the estimate of the impulse response of the echo path with a filter length N. The error of the signal is therefore

$$e(n) = y(n)\hat{y}(n) \tag{3}$$

The least-squares method is frequently used to measure the error and a fast convergent NLMS algorithm can be derived. However, it is well-known that large errors in even one data point will seriously degrade the least-squares estimation. This is the case for double-talk where the error function is perturbed badly. In view of this, more resistant methods are often necessary, and for example the use of the least absolute distance criterion or 11 norm can be appropriate in this situation as it tends to give less weights to the effect of gross errors or wild points. There are several ways to combine these properties to give hybrid cost functions. Applying within a stochastic gradient framework using NLMS, the most popular class of functions, known as the Huber function [5] which consists of a parabola prolonged by two tangents, can be written as

$$J(e) = \begin{cases} E[e^2] & \text{if } |e| \le ks \\ E[ks(2|e| - ks)] & \text{if } |e| > ks \end{cases} \tag{4}$$

The first order derivative neglecting the expectation is given by

$$\nabla_h J(e) = \begin{cases} 2ex & \text{if } |e| \le ks \\ 2 \cdot \text{sign}(e)\, x & \text{if } |e| > ks \end{cases} \tag{5}$$

The final NLMS algorithm is given by

$$\hat{h}(n) = \hat{h}(n-1) + \begin{cases} (\mu e(n)/x(n)^T x(n))x(n) & \text{if } |e| \le ks \\ \text{sign}(e(n))(\mu x/x(n)^T x(n))x(n) & \text{if } |e| > ks \end{cases} \tag{6}$$

The delay-less sub-band echo canceller is illustrated in Fig.

1. The echo path is modeled in sub-bands with a set of parallel adaptive filters. The sub-band filters are then collectively transformed to a single full-band filter via a weight transform. In this paper the DFT-FIR weight transform method [6] is used. This full-band filter models the acoustic channel. By separating the paths for adaptation and echo cancellation, the analysis/synthesis system in the signal path, and thus the signal path delay, is avoided whilst the desired features of sub-band processing such as signal decor relation and computational efficiency are retained. The adaptive filter in the m-th sub-band, hm(k), is adapted by the signals in that sub-band, xm(k) and em(k). Depending on how em(k) is constructed, the delay-less sub band adaptive filter can be configured in either a open-loop and closed-loop way. In the open-loop configuration, the error signal em(k) is generated locally in the m-th sub-band.

$$\begin{aligned} e_m(k) &= d_m(k) - \hat{\mathbf{h}}_m^H(k)\mathbf{x}_m(k) \\ d_m(k) &= d(n) \otimes f(n)|_{\downarrow D} \\ x_m(k) &= x(n) \otimes f(n)|_{\downarrow D} \\ \mathbf{x}_m(k) &= \begin{bmatrix} x_m(k) \\ x_m(k-1) \\ \cdots \\ x_m(k-N_s+1) \end{bmatrix} \end{aligned} \tag{7}$$

where Ø denotes the convolution operation, ·|!D denoted D fold down-sampling, and f(n) is the analysis filter. In the closed loop-configuration, em (k) is obtained from the full band error signal e(n) as

$$e_m(k) = e(n) \otimes f(n)|_{\downarrow D} \tag{8}$$

presenting an implementation of the synthesis dependent solution. By utilizing the full-band error signal, it is possible for a closed-loop sub-band adaptive filter to converge to the optimal Wiener solution. Moreover, the closed-loop configuration yields better computational efficiency because no convolution in the sub-bands is necessary. The closed-loop configuration will be employed in this work.

3 IMPLEMENTATION

A. Measurement Procedure

In this paper, we employ the Xilinx ML41 0 [7] development platform to implement the echo-canceller and the power consumption is measured by connecting an ammeter to the power module on the board. The architecture has been implemented on a FPGA platform using VHDL. It is synthesized, placed and routed on the Xilinx XC4VFX60 [8] FPGA device using Xilinx ISE 10.1 FPGA design package. We have setup four implementations and compare the power consumption of each setup. In the first implementation, we run the RTPAF algorithm on Mat lab using a Pentium 4 machine clocked at 3.2 GHz. In one implementation, we solely use a soft-core processor (processor implemented by reconfigurable fabric) to performance the echo cancellation algorithm. The processor is clocked at 200 MHz. In another implementation, we exploit FPGA reconfigurable logic to accelerate the computationally intensive operations such as Fast-Fourier Transforms (FFTs) and convolutions. The soft-core processor is only used for handling control logic and 110 operations. We anticipate that the latter implementation can reduce the power consumption over the former one substantially. The computation core is clocked at 137 MHz while the soft-core processor runs at 180 MHz. In order to maximize the system performance, all the processing cores are implemented using a core generation engine provided by the vendor tools. The core generation engine can generate high speed processing cores by describing the interconnection of the components and the placement of circuits using netlist files. In addition, this approach reduces the development time significantly and the quality of the design can be ensured. The processing cores generated include a 24-bit 128-point FFT core, a 28-bit complex number multiplication core and different configurations of block RAM memory.

TABLE I
IMPLEMENTATION RESULTS OF RTPAF

FPGA CHIP	XC4VFX60
Slices Used	5747 (22%)
DSP48/MULT used	48 (37%)
Block RAM used	17 (7%)
Max. Frequency	136.7 MHz

In the last setup, we implement multiple sub band filters on one single FPGA device and exploit parallelism to improve the performance and reduce the power consumption.

B. Performance Evaluation

In our experiment, the full band filter length is set to N = 1024, the number of sub-bands is M = 128 with a decimation factor of D = 6 4. Clearly, the sub-band algorithm outperforms the full-band algorithm in terms of the tracking efficiency and mis-alignment accuracy. To simulate the double-talk situation, a near-end speech is introduced from sample 80,000 to 100,000. Echo path variation is simulated by switching the desired signal to that from the microphone 24 cm apart from the original one at sample 110,000. This simulates the difficult situation where double-talk is followed immediately by an echo path variation. Simulation results have confirmed that the algorithm performs well during double-talk, and be able to track echo path variation quickly. This is the case even when the near-end speech has a low energy level which resembles the scenario of whispering in front of the microphone. As in [4], the integer size and the fraction size of the fixed-point library are chosen to be 10 and 18 respectively in order to reduce the amount of overflow during computations. In case of any overflow, the coefficients will change dramatically and the results will be invalid. We design and implement the hardware accelerator using VHDL. The accelerator is synthesized, placed and routed using Xilinx ISE 10.1 FPGA design package. Table I represents implementation results of the proposed hardware design on the high-end Xilinx XC4VFX60 FPGA chips In evaluating the performance of the FPGA-based echo canceller, we first assuming one block of data contains 128 samples under a 8kHz sampling rate, the number of clock cycle required for processing the block of data in the frequency domain is measured as 1841. Using our hardware accelerator clocked at 137 MHz, the period of a clock cycle is 7.32 ns. Therefore, the FPGA-based echo-canceller can perform one step of echo cancelling in 13.48 /1S, or equivalently 9.5M samples per second. We also implement an equivalent software version in MATLAB and compiled it to native machine code using the MATLAB supplied compiler (mcc). In evaluating this software version, a test containing 16M sample is provided to the program and the time required to finish all the calculations was measured.

TABLE II
RESOURCE USAGE OF MULTIPLE INSTANCES FOR RTPAF ON A VX4VFX60 FPGA CHIP

No.Of Instances	Speed	Slices used	Speed Up
1	137 MHz	22%	13
2	137 MHz	44%	26
3	137 MHz	66%	29
4	137 MHz	88%	52

TABLE III
POWER CONSUMPTIONS OF RTPAF WITH/WITH OUT HARDWARE ACCELERATION

Set Up	Elapsed Time	Power Consumption
Pure Soft Core	138 s	5900 J
Hardware Acceleration	14 s	598 J

The test was carried out in a Core-2-Duo 3.2GHz machine with 4GB memory, and it took an average of 17.5 seconds to finish the calculations. Therefore, the software performance is 224/17.5 = O.96M samples per second. It shows that the FPGA-based echo-canceller can achieve 10 times speedup when compared with software running on a 3.2GHz PC. Table II summarizes the implementation results when adding more instances of the filter in an XC4VFX60-FFI152- 11 FPGA chip. Ideally, the speedup would scale linearly with the number of two-path adaptive filter instances. In practice, the speedup may become slower when the logic utilization increases because the clock speed of the design deteriorates as the number of instances increases. In our experiment, we found that a XC4VFX60 can pack at most 4 instances of the two-path adaptive filter, so the speedup will be close to 40 times.

C. Power Consumption Evaluation

We measured the power consumptions of the following two setup: (i) only soft-core processor is used; (ii) hardware accelerator handling intensive operations with assistance from soft-core processor. The same test data is provided to the two implementations and we recorded the power consumptions during the core computations in the RTPAF algorithm. The records are shown in Table III. We observed that the power consumption of the setup with hardware acceleration is more energy-efficient and the power consumption is just 10% of that without the hardware acceleration. Despite the lack of direct comparison, we believe that the implementation on FPGA using dedicated hardware accelerators would spend less power consumption than that running on microprocessors. And the proposed architecture is also suitable for devices that are running on limited power supply.

4 CONCLUSION

This paper proposes a novel hardware architecture for two path frequency domain echo cancellation. The proposed echo-canceller is robust against double-talk and is sufficiently fast in tracking echo path variation for real-time applications. Also, bit width optimization reduces the circuit size while maintaining the quality of the result. Multiple instances of our echo-canceller have been packed on to an FPGA to boost the speed of the filter. Our work shows that nearly 90% of the energy consumption is reduced by hardware acceleration experimented on FPGA chips.

REFERENCES

[1] S. L. Gay, "An introduction to acoustic echo and noise control," in Acoustic signal processing for telecommunication, S. L. Gay and J. Benesty, Eds. Kluwer Academic, 2000, ch. 1.

[2] W. C. Chew and B. Farhang-Boroujeny, "FPGA implementation of acoustic echo cancelling network echo cancellation;' in Proc. IEEE TENCON, 1999, pp. 263-266.

[3] S. A. Jang, Y. J. Lee, and D. T. Moon, "Design and implementation of an acoustic echo canceller;' in Proc. IEEE Asia-Pacific Conference on ASIC, 2002, pp. 299-302.

[4] C. Ho., K. Yiu, J. Huo, and W. Luk, "Reconfigurable acceleration of robust frequency-domain echo cancellation," in Proc. of Engineering of Reconfigurable Systems and Algorithms, 2006, pp. 184-190.

[5] P. J. Huber, "Robust estimation of a location parameter;' Annals of Mathematical Statistics, vol. 35, pp. 73-101, 1964.

[6] J. Huo, S. Nordholm, and Z. Zang, "New weight transform schemes for delayless subband adaptive filters," in Proc. IEEE Global Commun. Conj.,2001.

[7] (2008) Ml410 embedded development platform user guide. Xilinx Inc.

[8] (2004) Xilinx Virtex-4 family overview. Xilinx Inc. [Online]. Available: hup://direct.xilinx.comlbvdocs/publications/ds Il2.pdf