Audio De-noising by Spectral Subtraction Technique

Kajol Bayeti[1] Pankaj Rathi[2]
[1,2]Department of Electronics & Communication Engineering
[1,2]SITE, Nathdwara, Rajsamand, India

Abstract— Hearing aids are now used to alleviate hearing impairments. However, more than 60% of impaired persons feel uncomfortable when using their hearing-aids because of the worse intelligibility resulting from bad speech comprehensibility. We believe that hearing impairments can be alleviated by a system with characteristics closer to being body's ones. Research has been deeply involved in developing new algorithms to improve speech intelligibility. Although, lot of researches was held to enhance speech for impaired people, few authors deal with the problem of power consuming. No numerical results are given, making the comparison only with the hearing aids of the market. Most closely related to our approach is the work of, who provide a method for power reducing based on algorithm and hardware optimizations along with the architecture uses the odd/even data lifting. Our work is improved by avoiding the access to the memory for data storage. Instead, we formulate our algorithm under which a given data of the speech signal is segmented at the input and each segment is processed individually. In order to avoid losing of data information, the segments are overlapped and an overlap technique is then used for treatment. This approach provides perfect analysis and efficient computation. Consequently, the framework devised here is made generic and requires simulations and empirical evaluation of the routing scheme in order to be applied. Traditionally, Digital Signal Processing (DSP) algorithms are implemented using General Purpose Processors (GPP) for low rate applications. These devices showed limited capabilities for processing high volume data efficiently in real time. The trends had then been shifted to Special Purpose DSP (SPDSP) and Application Specific Integrated Circuits (ASICs) in order to meet the increased complexity and to gain in performance requirements of these algorithms but with high cost functions. Today, FPGAs are highly preferred for their relatively high capacity, low cost, short design cycle and short time to market. FPGA affords the capability of constant reconfiguration to meet application performances. Dealing with digital speech processing as it pertains to the hearing impaired persons especially for miniaturized system applications; FPGA allows increasing sophisticated features to be built for better sound reproduction while keeping small size and low power consumption of the devices. Fortunately, simulation tools provide us a rapid design and basic information. Similarly, a high-level programming language is an efficient comparison tool for the final output results and system evaluation. In practice, the implementations are often subject to lot of limitations. Using DWT at multi-resolution over disjoint bands remains up to now a practical necessity for perfect design for digital speech processing in particular and herein some references from our work, where the goal is to investigate noise reduction and hardware implementation. This work extends previous research described in. In this we present the implementation of a multi-level one dimension DWT combined to an OLA on FPGA for a bio-inspired medical hearing aid application. The methodology aims to improve in one side better speech quality and in the other side, an efficient flexible reconfiguration and reduced cost functions. The scheme represents architecture for de-noising and frequency shifting. It is realized targeting a DE2 development kit board of Altera (EP2C70F896) and results are compared to that obtained in Matlab. The system provides a generic framework allowing the use of DWT analysis / synthesis with frequency shaping of the speech signal to improved speech intelligibility. We present some simulation results under VHDL and Matlab. Hence, a comparative study is done based on the Mean Square Error (MSE) and the Signal to Noise Ratio (SNR). MOS evaluations are presented for speech intelligibility and the gain obtained by the proposed architecture.

Key words: Audio De-noising, Spectral Subtraction Technique

I. INTRODUCTION

In speech communication, the speech signal is always accompanied by some noise. In most of the cases, background noise of any environment where the source of speech lies, it is the main component of noise signal which adds to the speech communication. |The obvious effect of this noise addition is to make the listening task difficult for a direct listener, there are many negative effects when we process the degraded speech for some other applications. A related problem is processing degraded speech signal in the preparation for coding by the bandwidth compression system of speech. So, speech enhancement not only consists processing speech signals for human listening system but also for further processing prior to listening task. The Main objective of speech enhancement technique is to improve the perceptual aspects of speech signal like overall quality, intelligibility, or degree of listener fatigue. In our work, we study two speech enhancement techniques to enhance the quality of speech in the presence of additive and broadband acoustic noise which is spectral subtraction and FPGA technique. The purpose of the prosody modification is to make one, two or all of these parameters change over a speech segment without affecting the timber. The signal intensity can be easily modified by a multiplication; meaning that by simply amplifying consonant energy will improve their identification. But, the changes of the fundamental frequency or pitch and the duration or speed are not so obvious. Clear speech has better intelligibility that conversational one where significant differences in phonetic, phonological and prosodic features are observed. If the duration-rate decreases, the speech intelligibility increases. However, applying phoneme duration from conversational to clear speech did not improve the intelligibility.

II. REQUIREMENTS & BASIC DESIGN

There are numerous options available for the audio de-noising commercially but there exist a scope of some refinements and cost cutting. This theme defines our

requirements for the proposed scheme. The basic requirements for this is spectral subtraction technique with Quartus representation of VHDL coding. The portability of the system is also a vital feature.

Honouring the pre-request of the requirements, the concept design of audio de-noising device is based on spectral subtraction in which VHDL coding is simulate in Quartus and compare with Matlab coding.

III. SOFTWARE DESIGN IMPLEMENTATION

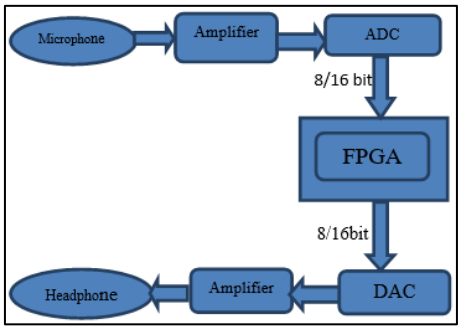

Fig. 1: Block diagram for the audio de-noising

This proposes an efficient reconfigurable hardware design for speech enhancement based on multi band spectral subtraction algorithm and involving both magnitude and phase components. Our proposed design is novel as it estimates environmental noise from speech adaptively utilizing both magnitude and phase components of the speech spectrum. We performed multi-band spectral subtraction by dividing the noisy speech spectrum into different non-uniform frequency bands having varying signal to noise ratio (SNR) and subtracting the estimated noise from each of these frequency bands. This results to the elimination of noise from both high SNR and low SNR signal components for all the frequency bands. We have coined our proposed speech enhancement technique as Multi Band Magnitude Phase Spectral Subtraction (MBMPSS). The magnitude and phase operations are executed concurrently exploiting the parallel logic blocks of Field Programmable Gate Array (FPGA), thus increasing the throughput of the system to a great extent. We have implemented our design on Spartan6 Lx45 FPGA and presented the implementation result in terms of resource utilization and delay information for the different blocks of our design. To the best of our best knowledge, this is a new type of design for speech enhancement application and also a first of its kind implementation on reconfigurable hardware. We have used benchmark audio data for the evaluation of the proposed hardware and the experimental results show that our hardware shows a better SNR value compared to the existing state of the art research works.

A. Amplifier

An amplifier, electronic amplifier or amp is an electronic device that can increase the power of a signal (a time-varying voltage or current). An amplifier uses electric power from a power supply to increase the amplitude of a signal. The amount of amplification provided by an amplifier is measured by its gain: the ratio of output to input. An amplifier is a circuit that can give a power gain greater than one.

B. ADC and DAC

In electronics, an analog-to-digital converter (ADC, A/D, A–D, or A-to-D) is a system that converts an analog signal, such as a sound picked up by a microphone or light entering a digital camera, into a digital signal. An ADC may also provide an isolated measurement such as an electronic device that converts an input analog voltage or current to a digital number proportional to the magnitude of the voltage or current. Typically the digital output is a two's complement binary number that is proportional to the input, but there are other possibilities. There are several ADC architectures. Due to the complexity and the need for precisely matched components, all but the most specialized ADCs are implemented as integrated circuits (ICs).

A digital-to-analog converter (DAC) performs the reverse function; it converts a digital signal into an analog signal. The conversion involves quantization of the input, so it necessarily introduces a small amount of error. Furthermore, instead of continuously performing the conversion, an ADC does the conversion periodically, sampling the input. The result is a sequence of digital values that have been converted from a continuous-time and continuous-amplitude analog signal to a discrete-time and discrete-amplitude digital signal.

C. FPGA

FPGA redirects here. It is not to be confused with Flip-chip pin grid array. A field-programmable gate array is an integrated circuit designed to be configured by a customer or a designer after manufacturing – hence "field-programmable". The FPGA configuration is generally specified using a hardware description language (HDL),

Similar to that used for an application-specific integrated circuit (ASIC). (Circuit diagrams were previously used to specify the configuration, as they were for ASICs, but this is increasingly rare.) Although, lot of researches was held to enhance speech for impaired people few authors deal with the problem of power consuming. No numerical results are given, making the comparison only with the hearing aids of the market.

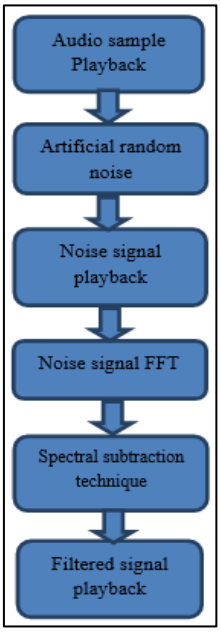

Fig. 2: Flow Chart of Procedure

Once the design is ready then implementation of audio de-noising using spectral subtraction technique is done. Quartus II used for simulation and representation of FPGA design flow and Matlab for FFT plot.

IV. RESULT AND DISCUSSION

Initially we gave 5 seconds speech capture then draw playback of this signal. Audio device is configured for speech signal using microphone in laptop and playback of this signal, after playback artificial noise is added. After completing of noisy data playback plot of FFT and noisy data done and using spectral subtraction technique filtered signal is received in which noise is completely removed. All processed output is shown step by step in figures below.

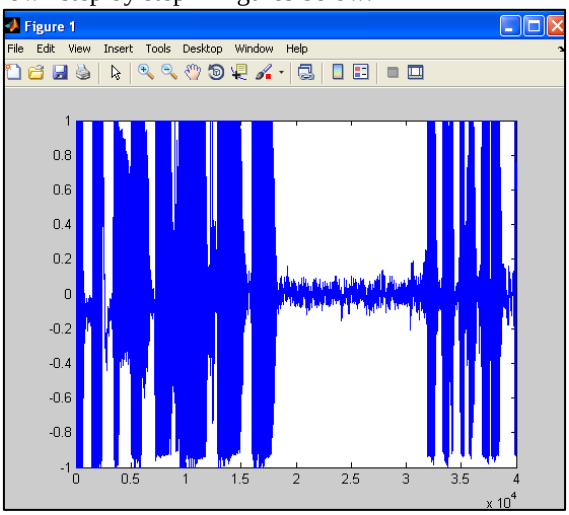

Fig. 3: Playback of audio signal

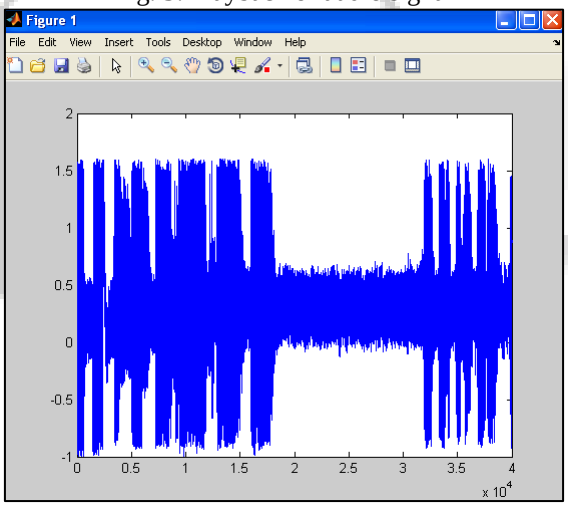

Fig. 4: Playback with noise

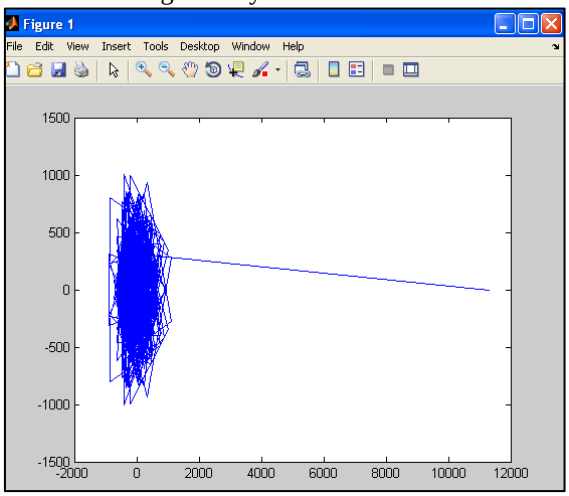

Fig. 5: Noisy data plot

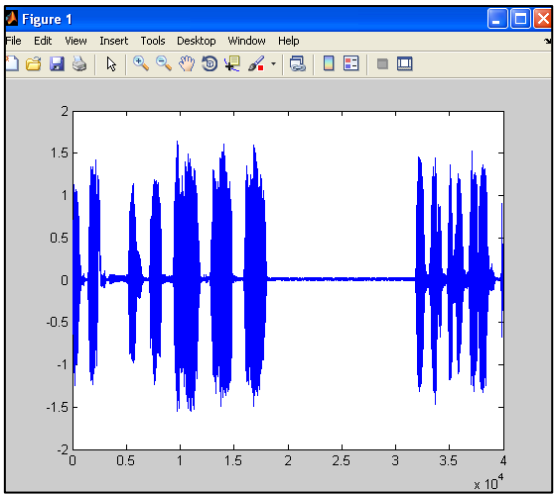

Fig. 6: Filtered Signal

V. SUMMARY

This software is useful for removing noise from audio signals which is given by microphone or any input devices. This investigated the problem of noisy conditions assuming absence of information about the noise. The works reported are carried out using a standard database like street, airport, babble, car, train, station, and exhibition hall and restaurant noise. Spectral subtraction and Wiener filter has been used for noise elimination from the speech signal. From the performance, it has been observed that spectral subtraction is a very efficient method for elimination of all noise in high SNR condition. However, its performance rapidly degrades with increasing of SNR. Overview of currently done work realized that audio in painting problem is still current problem worth to study and resolve. Reconstruction of audio signals of commercial programs is not perfect Spectral subtraction approaches are also suitable for reducing the clicks and broadband distortion from the audio signals. It is still a current research problem.

REFERENCES

[1] Zhang Jie, Xiaoqun Zhao, Jingyun Xu , Zhang Yang, "Suitability of speech quality evaluation measures in speech enhancement", Tongji University Shanghai, China.,2014
[2] University Grant Commission (UGC) RGNF-2012-13-SC-WES-26014, Govt of India.
[3] Dr.Shaila D. Apte "Speech and Audio Processing", Wiley Precise Textbook, 2013
[4] Sanjay P. Patil, John N. Gowdy, "Exploiting the baseband phase structure of the voiced speech for speech enhancement", Clemson University, Clemson USA, 2014.