REAL-TIME DIGITAL SPEECH PROCESSING STRATEGIES FOR THE HEARING IMPAIRED

Neeraj Magotra, Sudheer Sirivara
Department of Electrical and Computer Engineering
University of New Mexico, Albuquerque, New Mexico 87131, USA
Contact: email: magotra@houdini.eece.unm.edu, Ph:505-277-0808

Abstract

This paper deals with digital processing of speech as it pertains to the hearing impaired. The issues described in this paper deal with the development of a true real-time digital hearing aid. The system (based on Texas Instruments TMS320C3X) implements frequency shaping, noise reduction, , interaural time delay, amplitude compression and various timing options. It also provides a testbed for future development. The device is referred to as the DIgital Programmable Hearing Aid (DIPHA). DIPHA uses a wide bandwidth (upto 16 KHz). DIPHA is a fully programmable device that permits us to program various speech processing algorithms and test them *on hearing impaired subjects in the real world* as well as in the laboratory.

1. INTRODUCTION

The ultimate goal of this research is the development of a portable digital hearing aid that conditions the input speech signal based on environmental circumstances and the hearing impaired person's hearing loss. At present we have developed a prototype digital system (DIPHA) consisting of a laboratory-based PC system with a TMS320C30 DSP card and a wearable unit also based on the TMS320C3X DSP chip. The PC-based system is used to fit the patient and the final algorithm is down loaded to a read-only memory (ROM) chip on the wearable unit.

The digital hearing aid is a binaural aid capable of sampling two input microphone signals at sampling rates of up to 32 KHz per channel. The minimum bandwidth of the hearing aid is 9.2 KHz. At present we are researching frequency shaping algorithms, adaptive noise reduction algorithms, amplitude compression and interaural time delay algorithms. The following section describes these algorithms.

2. SPEECH PROCESSING

The block diagram in Figure 1 represents the speech processing techniques currently implemented on DIPHA. The frequency shaping and adaptive noise reduction algorithms have been patented.

Frequency Shaping

This stage of signal processing shapes the speech spectrum (0 upto 16 KHz) to compensate the patient's hearing loss. The speech data is input to a *binaural* equalizer. The equalizer is implemented as two banks of bandpass filters, one for each channel (ear). These filters are designed to have perfectly linear phase and high (>80 dB) band isolation. The therapist can interactively (in real time) choose the number of filters in each bank and select their critical frequency characteristics namely, their cutoff frequencies and isolation between different frequency bands. Because of this design flexibility the same DIPHA unit can be reprogrammed for different languages [1]. Some test results obtained by testing DIPHA on human subjects have been presented in [2]. Generally, DIPHA seems to be able to improve speech discrimination scores for most subjects.

Adaptive Noise Reduction

The noise reduction algorithm used in DIPHA has been designed to work with just one input data channel. Hence DIPHA processes the two (right ear and left ear) single-input single output channels independently. For each channel the input signal is first high-pass filtered to compensate for the low frequency spectral tilt in speech signals [3]. The

HPF is a simple first order infinite impulse response (IIR) filter with tunable cutoff frequency.

The noise reduction algorithm is referred to as the Real-time Adaptive Correlation Enhancer (RACE) algorithm. RACE is essentially an adaptive finite impulse response (FIR) filter [4].

The speech input is used to update the RACE coefficients. These coefficients consist of the estimated autocorrelation coefficients ($\hat{R}_{xx}(m,l)$) of the input channel. The autocorrelation coefficients are updated using a recursive estimator as given by the following equation [4]

$$\hat{R}_{xx}(m,l) = \beta \hat{R}_{xx}(m-1,l) + (1-\beta)x(m)x(m+l)$$

The equation above represents a recursive estimator which corresponds to sliding an exponential window over the data with a time constant (τ in seconds) given by $\tau = 1/((1-\beta)f_s)$ where f_s represents the sampling frequency (sps) and m is the time index, l is the lag index whose magnitude is bounded by L and β is the smoothing constant whose value lies in the range of 0 to 1. After applying suitable gain control [4], the autocorrelation coefficients are used to update the adaptive FIR.

We have shown that for a narrowband signal the amplitude gain and signal-to-noise ratio (SNR) gain are both equal to approximately half the filter length or L. In terms of convergence considerations we have shown that RACE is able to converge rapidly enough so that the short term stationarity of the speech signal [3] does not cause any problems for the algorithm. We have also shown that RACE is able to converge faster than the normalized LMS algorithm used for FIR and lattice adaptive filters.

RACE can be implemented in different configurations, depending on the positioning of a pre-emphasis filter, one of these configurations is shown in Figure 2. RACE is currently being evaluated using computer simulations and human subject testing in various configurations. Anecdotal results, obtained while testing hearing impaired subjects in real-world environments, indicate that the processing strategy is successful. We are currently setting up experiments, with the help of audiologists, to derive quantitative and subjective techniques for evaluating RACE's performance.

Amplitude Compression

Speech amplitude compression is basically the task of controlling the overall gain of a speech amplification system It essentially "maps" the dynamic range of the acoustic environment to the restricted dynamic range of the hearing impaired listener. A crude way of accomplishing this task is peak-clipping. Although computationally inexpensive, this approach introduces harmonics that distort the output speech signal. Our compression algorithms avoid this problem [5]. Compression is achieved by applying a gain of less than one to a signal, whenever its power exceeds a predetermined threshold. A running estimate of the power is obtained by using an estimator similar to the one represented by equation (1). The time constant of estimation is used to modify the attack/release time of the compression algorithm. In [5] two alternative techniques of performing speech amplitude compression have been presented. Our current implementation also allows for look-ahead capability, whereby the algorithm that actually computes the compression factor 'looks ahead' of the incoming data. This option does delay the output by the look ahead time. A block diagram of the algorithm is shown in Figure 3.

Interaural Time-delay

The importance of interaural time difference (ITD) and interaural intensity difference (ITD) in speech perceptiona nd localization has been well documented. Recent research [6] indicates that ITD and ITD might be an even more crucial issue for the hearing impaired. We have implemented an interaural time delay algorithm (see Figure 4 for a block diagram) that involves splitting the input signal into multiple frequency bands, interpolating the data to provide suitable time delay increments, differentially delaying the right and left inputs in each frequency band and then recombining the left and right frequency bands respectively for output to the earphones.

3. RESULTS

The perception of speech is highly subjective in nature and in the final analysis it is the patient's response that determines the success or failure of the signal processing used. To achieve the results described in in the paper we implemented the algorithms described above on custom-built wearable (walkman size) units using the TMS320C31 DSP chip. This unit complements the C30 based unit that operates in the PC environment. Once the audiologist has 'tuned' the patient using the PC based unit the program is burned into a ROM which when inserted into the portable unit customizes it for the patient. Because of the design flexibility built into DIPHA hardware/software it is worth mentioning that the same unit can be reprogrammed for different languages such as English, Navajo, Hindi etc.

Our test results to date indicate that using frequency shaping, noise reduction, amplitude compression and interaural time delay algorithms in various combinations seems to help the hearing impaired.

(1) DIPHA has been able to increase the discrimination scores of severely to profoundly deaf patients by upto 30 %

(2) In real life situations patients have been able to converse normally even in extremely noisy environments, usually they would take off their conventional aids in these situations.

(3) Many of the profoundly deaf patients were able to hear high frequency sounds for the first time and were able to repeat these sounds back to the therapist.

The paper presents the unaided and aided test scores of some of our test subjects as well as analytical test results obtained via computer simulations.

4. CONCLUSIONS AND FUTURE WORK

We have successfully developed and tested a floating-point DSP hearing aid that incorporates wide bandwidth and a great deal of flexibility in adjusting the overall speech processing algorithm to the clients specific needs. We are currently pursuing research to improve our current algorithms, develop new noise reduction algorithms and are considering porting the algorithms to Texas Instrument's newer fixed point (lower power) devices. Also, the algorithms are being tested on human subjects in the real-world environment using the portable devices.

REFERENCES

[1] N. Magotra, J. Stewart, K. Bricker, R. Weaver , "Digital Auditory Device using the TMS320C30," International Conference on Signal Processing Applications and Technology, Cambridge MA, Nov. 1992, pp. 15-19.

[2] N. Magotra, T. Hamill, B. Swartz, "Digital Processing of Speech For the Hearing Impaired," 29*th* Asilomar Conference on Signals, Systems and Computers, Asilomar, CA, Nov 1995.

[3] L. R. Rabiner, R. W. Schafer, "Digital Processing of Speech Signals," Prentice-Hall, 1978.

[4] N. Magotra, F. Livingston, S. Rajagopalan, "Single Channel Speech Enhancement in Real-time," 27*th* Asilomar Conference on Signals, Systems and Computers, Asilomar, CA, Nov 1993.

[5] B. Swartz, N. Magotra, "Speech Amplitude Compression," proceedings of the IEEE ISE conference, Albuquerque, NM, 1994.

[6] H. Simon, I. Aleksandrovsky, "Perceived Lateral Position of Narrow band Noise in Hearing impaired and Normal-hearing Listeners under Conditions of Equal Sensation Level and Sound Pressure Level," Conference on Issues in Advanced Hearing Aid Research, Lake Arrowhead, CA, May 1996.

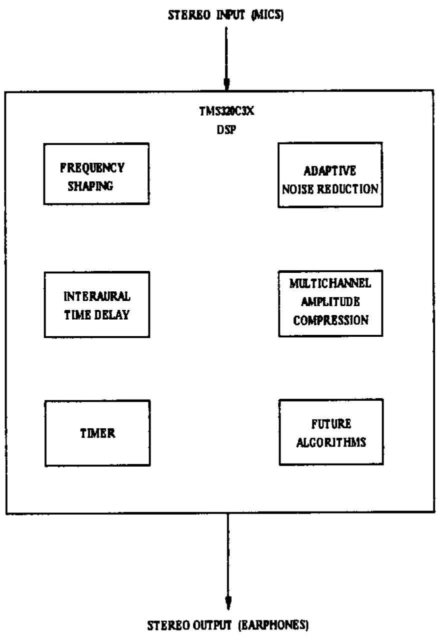

FIGURE - 1: REAL TIME SPEECH PROCESSING FOR THE HEARING IMPAIRED

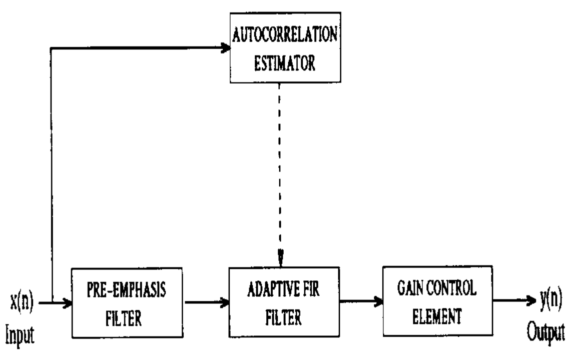

FIGURE - 2: ACE CONFIGURATION

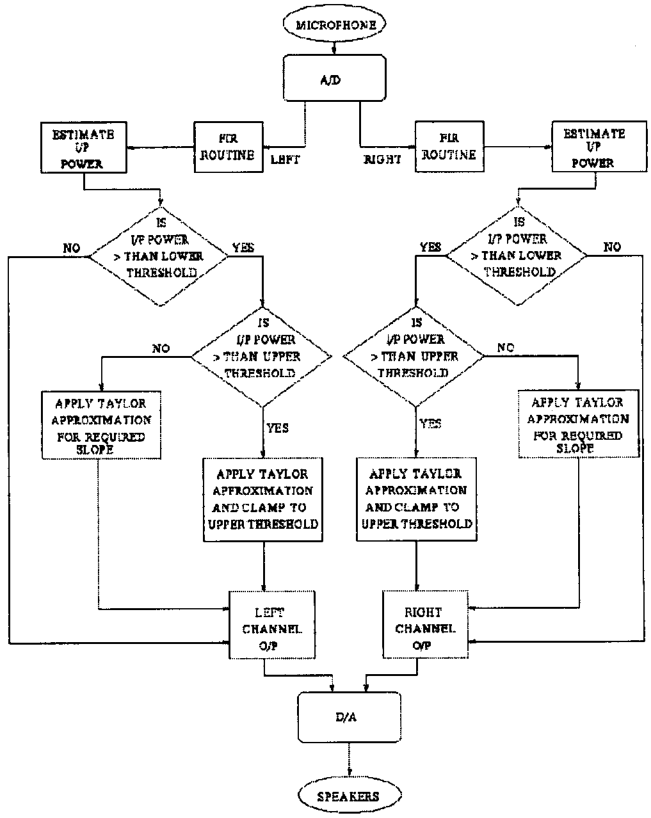

FIGURE - 3 : FLOW DIAGRAM FOR AMPLITUDE COMPRESSION ALGORITHM

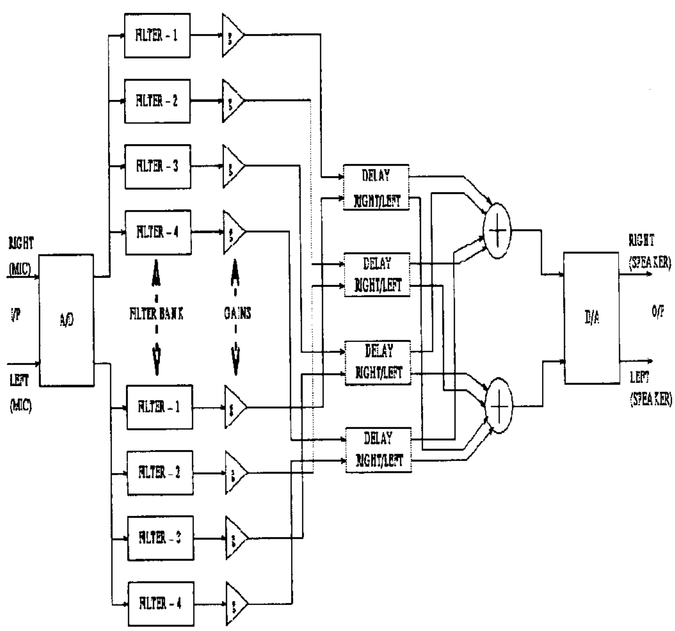

FIGURE - 4 : MULTIBAND FREQUENCY SHAPING & DIFFERENTIAL DELAY SCHEME