

## PROTOTYPE FOR DYSARTHIC SPEECH ENHANCEMENT

PREMA P.<sup>1\*</sup>, VIDHYA S.<sup>2</sup> and LAZAR MATHEW T.<sup>3</sup>

<sup>1</sup>Biomedical Engineering Department, PSG College of Technology, Coimbatore-641004

<sup>2</sup>Biomedical Engineering Division, VIT University, Vellore-632014,

<sup>3</sup>Advisor, Medical Sciences, Engineering and Technology, PSG Institute of advanced studies, Coimbatore-641004,

\*Corresponding Author: Email-ppr@bme.psgtech.ac.in, Fax: +91-422-2573833, Phone: 0422-4344000

<sup>2</sup>svidhyavalentina@vit.ac.in, Fax: +91-422- 2573833, Phone: 0416-2243091

<sup>3</sup>tlm@ias.psgtech.ac.in Fax: 91-416-2243092 Phone: 0422-4344000

Received: November 11, 2011; Accepted: December 01, 2011

**Abstract-** In this paper we present a new prototype which enhances the intelligibility of Dysarthric speech signal of subjects belonging to the category of cerebral palsy. This prototype is designed for Indian English condition. The subject chosen is male in the age group of 30-60years. The speech processing technique is applied to improve the intelligibility by replacement of dysarthric vowel space by best vowel space of the same dysarthric patient so that the voice quality of speaker is retained. The developed algorithm is deployed into the DSP processor TMS320C5416. It is a pc based system. The testing is done on real time with dysarthric subject and intelligibility of about 60 to 70% was achieved.

**Keywords:** Speech Processing, Dysarthria, prototype, codec, JTAG, USART.

### Introduction

Dysarthria is a motor speech disorder caused by paralysis, weakness, or inability to coordinate the muscles of the mouth. Dysarthric speech can be far less intelligible than that of non-dysarthric speakers, causing significant communication difficulties. People with dysarthria may also have problem controlling the pitch, loudness, and rhythm and voice qualities of their speech [1]. Dysarthria may be a sign of neuromuscular disorder such as cerebral palsy or Parkinson disease. Dysarthric speech can vary from mild unintelligibility to severe impairment. In very severe cases, vowels may also be distorted [2]. Intelligibility varies greatly depending on the extent of neurological damage.

The successful implementation of the enhancement and replacement algorithms give way for the construction of a small low powered DSP, which is a voice processor with these algorithms embedded in it. To assist the speech impaired in a usable and acceptable manner, these processor units could be fitted into a wearable device which is portable in nature.

### State Of Art

To date, only very few devices of similar nature exist on the commercial market that attempts to improve the natural dysarthric speech itself. The Speech Enhancer device (Electronic Speech Enhancement Inc) claims to "clarify" dysarthric speech; however, no published controlled research is available. The exact nature of their electronic algorithm is proprietary [3]. In India Very limited work has been carried out on providing assistive devices for speech impaired. No speech enhancers have been developed / marketed in India. There are ongoing researches in speech processing but not in the area of rehabilitation of the

speech impaired and/or developing suitable devices for the speech impaired. So we had made an attempt to develop a prototype that can be used as a wearable device by the speech impaired person.

### Materials and Methods

#### Overview

The main objective of this system is to improve the intelligibility of the dysarthric speech and reduction in ambient noise. The enhancement is done by transforming the vowel space of the dysarthric speech by the best frame of the same dysarthric patient in order to retain the subject voice. The embedded C code was developed and implemented in the DSP processor.

#### Prototype Design

The design of the speech enhancer has various stages like recording of samples creating database, analysis of recorded samples, developing speech enhancement algorithm, testing with the standard board available in the laboratory[5], conversion of code to equivalent c code that can be deployed in the prototype with feature optimization and finally testing and debugging.

#### Speech Data

The speech database for analysis, transformation and synthesis is created by recording the speech samples from male subject of mild case. The database contains CVC words, bi syllabic words, sentences and paragraph levels. The data are recorded in a sound proof environment at a sampling rate of 22 KHz.

#### Analysis of speech samples

The recorded sample is analyzed using PRATT software for its features like formants, pitch, and intensity. The speech processing algorithm for formant modification is

developed for increasing the vowel triangle space of dysarthric input [4]. MATLAB is used for developing the algorithm.

**Algorithm for enhancement:**

Speech waveform from the recorded database or from dysarthric patient in real time is obtained. The sampling rate was fixed to 8 KHz and frame length of 10ms is used. The voiced and unvoiced regions are separated. The voiced region is analyzed for speech feature modification and the unvoiced region is bypassed. The band (0-4KHz) of voiced region is applied to pre-emphasis filter with sample overlap of 50% and windowing to reduce noise and artifacts. Using Levinson-Durbin algorithm the formant is extracted and analyzed. If the formant range is not within the required level for producing intelligible speech, the LPC co-efficient and samples of the best frame from database is used for replacement. The signal is re-synthesized by inverse filtering and de-emphasizing.

The energy for each frame is analyzed and modified to its target level. Finally the two band signals and voiced/unvoiced regions are added and normalized to get reconstructed wave [6]. The reconstructed wave is played back to the audiologists for listeners test. They reported that, they could judge 60-70% of the utterance of the subject and found the signal is intelligible.

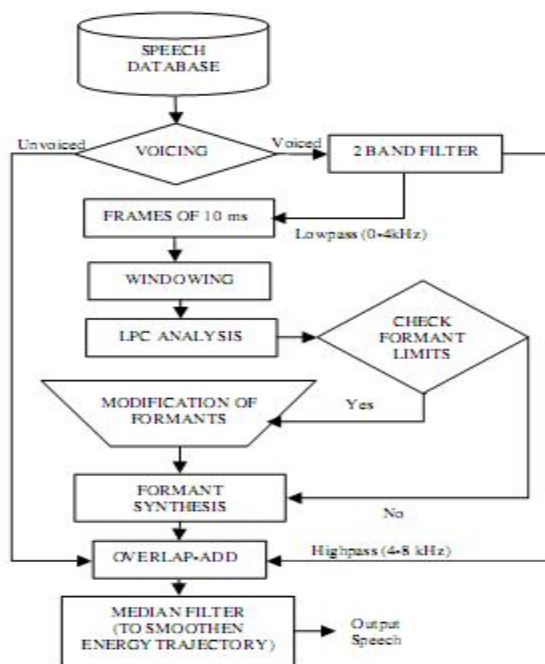


Fig 1 Flowchart of algorithm

**Algorithm testing**

We implemented our algorithm in SIMULINK which provides a graphical user interface for modeling, simulating, analyzing dynamic systems with MATLAB. The simulation results were stored in workspace for visualization and post processing. The code developed in this model is converted into equivalent C code deployed in the prototype and results were compared in each and every stage of implementation with MATLAB results.

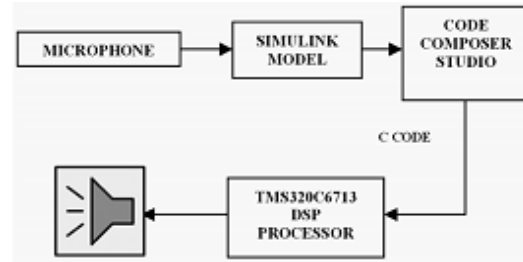


Fig 2 Testing with SIMULINK model

**Implementation in DSP**

The improved speech enhancement algorithm described above is implemented on a DSP system based on TMS320C5416. PC-based speech analysis tools have also been used to analyze all of the necessary pre-adjustment parameters, (i.e. average power, pitch range, V/UV power threshold) in advance. We introduce the DSP to realize all of the analysis-synthesis process in real time. The DSP hardware unit consists of a small board with a 24MHz internal clock, a 16 bit fixed point DSP, 16bit stereo codec TLC320AD77C, variable with sampling frequency of 2KHz to 32KHz, a parallel host interface, 12Kwords RAM ,16Kwords ROM, UART - 16C550 (9600 to 115200 Baud) and other on board features includes two 3.5mm Audio Jack for handset Microphone and Speaker- DB 9F RS232 Serial I/O Port, On Board Standard JTAG Interface connection for optional emulation and Expansion Connectors for Add On accessories. Power Supply - Input 230V, 50Hz (with EMI Filter)Output +/- 5Vdc, 500mA. With the combination of the PC-based speech analysis tools and the DSP hardware unit, a new user can first adjust all of the necessary parameters, and then download the adjusted DSP software into the hardware unit.



Fig 3 Final prototype

**Analysis of output**

The input wave file 'ball.wav' from the recoded database is loaded to the prototype board through the pc. The input sample was processed and the output is obtained in the speaker as well as the reconstructed sample is store in a file. The stored file is analyzed using PRATT software. The result is shown in table I and fig.4, where there is a considerable shift in the formants resulting in enhancement of intelligibility.

Table 1 Result analyzed using PRAAT

Input speech		Enhanced output	
F1(Hz)	388.292	F1(Hz)	401.694
F2(Hz)	1197.77	F2(Hz)	1208.44
F3(Hz)	1916.55	F3(Hz)	2125.49
Intensity	70.2040	Intensity	69.9867
Pitch (f0)	198.878	Pitch (f0)	199.802

**Result**

From the analysis, it is clearly evident that the intelligibility of dysarthric speech has been increased. The result obtained were almost same as that obtained in the MATLAB SIMULINK model tested in laboratory kit

**Conclusions**

After rigorous testing and implementation of the devised algorithms on a hardware unit, initial clinical trials would be carried in centres by audiologist and speech pathologist. After sufficient trials, issues like acceptability, usability, technical issues and ethical issues would be worked upon after which the final prototype would be fitted into a wearable device.

**Acknowledgements**

The Authors take this opportunity to thank All India Institute of Speech and Hearing for providing us with dysarthric samples, SANDS Instrumentation private Limited in developing the prototype board and also Ramya, Nirupa

and Sathivel for their technical support. We also express our special thanks to Society for biomedical technology for funding the project.

**References**

- [1] Medicine Net, "Definition of Dysarthria", <http://medterms.com>.
- [2] The speech-language pathology Website, "Dysarthria", <http://www.home.ica.net>.
- [3] Electronic Speech enhancement Inc., "The Speech enhancer", <http://www.speechenhancer.com>
- [4] Alexander B. Kain et.al (2007) *Improving the Intelligibility of Dysarthric Speech*, Oregon Health & Science University, Portland, USA.
- [5] Texas instruments (2003) *A DSP/BIOS AIC23 Codec Device Driver for the TMS320C6713 DSK*, Texas Instruments, Application Report, SPRA677.
- [6] Nirupa M., Prema P., Vidhya S., and Lazar Mathew T. (2009) *Formant Modification to Improve Intelligibility of Dysarthric Speech*, International conference on Biomedical Information and Signal Processing.
- [7] Ramya D., Prema P., Vidhya S., and Lazar Mathew T. (2010) *International Journal of Computational Intelligence and Healthcare Informatics*, vol.3, No.1, pp.15-19

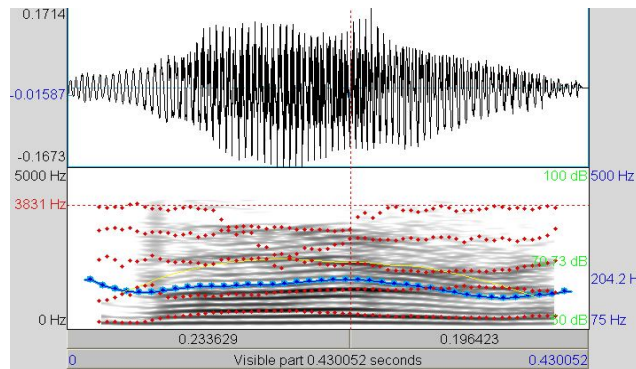


Fig 4 'ball.wav' output waveform before enhancement (spectrogram with formants, pitch and intensity)

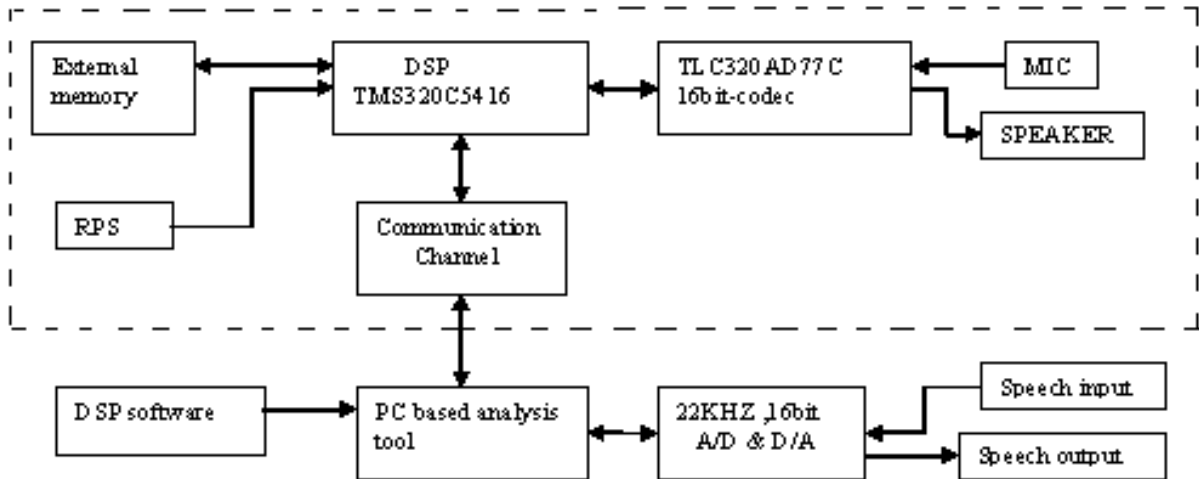


Fig 5 Block diagram of the prototype

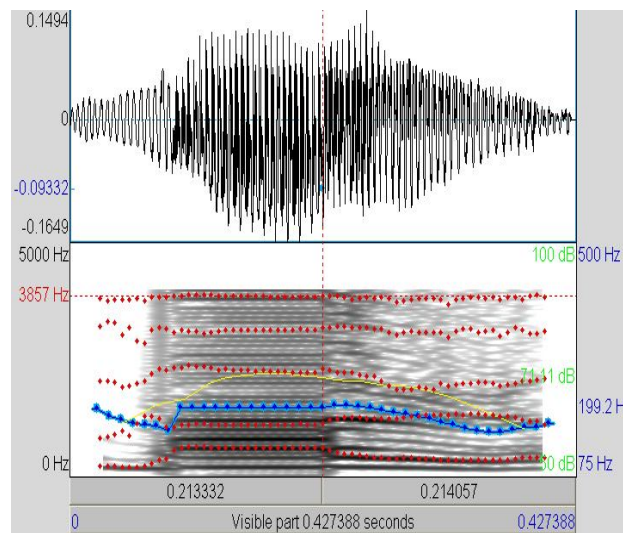


Fig 6 'ball.wav' output waveform after enhancement (spectrogram with formants, pitch and intensity)