REAL TIME FPGA IMPLEMENTATION FOR ATTENTIONAL SELECTION IN MULTI-SPEAKER ENVIRONMENT

Dr. .J. Britto Pari[1] ,Nishanthi. S [2], Swetha Lakshana. M [3], Parameshwari. S [4]
[1]Assistant Professor, Department of Electronics and Communication Engineering,
Sri Sairam Engineering College, Chennai, India [2,3,4]UG Scholar, Department of Electronics and Communication Engineering, Sri Sairam Engineering College, Chennai, India.

***Abstract* — People who are hearing impaired have a difficult time following a conversation in a multi-speaker environment. Current hearing aids can suppress background noise; however there is little that can be done to help a user attend to a single conversation amongst many without knowing which speaker the user is attending to. Cognitively controlled hearing aid receives a single audio channel containing a mixture of speakers that is heard by a listener, automatically separates the individual speakers in the mixture, determines the attended speaker and amplifies the attended speaker's voice to assist the listener. Our novel framework bridges the gap between the most recent advancements in speech processing technologies and speech prosthesis research and moves us closer to the development of cognitively controlled hearable devices for the hearing impaired. The goal of this project is to perform Real Time FPGA implementation of attentional selection in multi-speaker environment.**

I. INTRODUCTION

It is known that the technological advancement in these days are developing as a faster pace. The utilization of the technologies in various sectors are highly employed. Even though we use technology in various sectors, the employment of technology at some places are not used. We know that there are hearing aids available for hearing impaired. But, normal hearing aids can suppress only the background noise.

Cognitively controlled hearing aid separates the individual voices from the mixture of voices and amplifies the desired signal and provides to the user. This enables the hearing impaired people to listen to their required voice from the mixture of voices.

II EXISTING SYSTEM

Normal hearing aids are already in existence. A hearing aid can be worn in or behind the ear. It makes some sounds louder so that a person with hearing loss can listen, communicate, and participate more fully in daily activities. A Hearing aid can help people hear in both noisy and quiet situations. However only about one out of five people would benefit from it.

III PROPOSED SYSTEM

Cognitively controlled hearing aid receives a single audio channel containing a mixture of speakers that is heard by a listener, automatically separates the individual speakers in the mixture, determines the attended speaker and amplifies the attended speaker's voice to assist the listener.

IV AN OVERVIEW OF FPGA

Field Programmable Gate Array (FPGA) is a type of Integrated Circuit (IC) that will facilitate its user to create his own design and program it after it is manufactured. Basically an FPGA is like a blank state that means designer can create their very own configure file or bit file. The bit file that is loaded onto the FPGA will act as a digital circuit that the designer aimed to design. The FPGA does not has a processor to run the software until the designer builds one, it is up to the designer to create a simple file like an AND gate to multi core processor . The configuration of an FPGA needs a Hardware Descriptive Language (HDL) .The utmost important block is primary block that builds up an FPGA is Programmable Logic Block (PLB) which are programmable and are utilized to replicate the functions of logic gates. This logic block can similarly be assigned the function to act as computer memory which permits to store certain values and also between this logic blocks there are reconfigurable wiring circuitry. The FPGA logic block consists of a 4 input LUT (Look up Table), a programmable multiplexer which assists to choose whether the FPGA be programmed in registered or non -registered output and finally a flip-flop for storage. FPGA is a volatile device.

V LITERATURE SURVEY

"Schema-based processing in auditory scene analysis" Bregman A S and McAdams S

In this paper, the contribution of melodic schema-based processes to the perceptual organization of tone sequences is examined. Two unfamiliar six-tone melodies, one of which was interleaved with distractor tones, were presented successively to listeners who were required to decide whether the melodies were identical or different. In one condition, the comparison melody was presented after the mixed sequence: a target melody interleaved with distractor tones. In another condition, it was presented beforehand, so that the listeners had precise knowledge about the melody to be extracted from the mixture

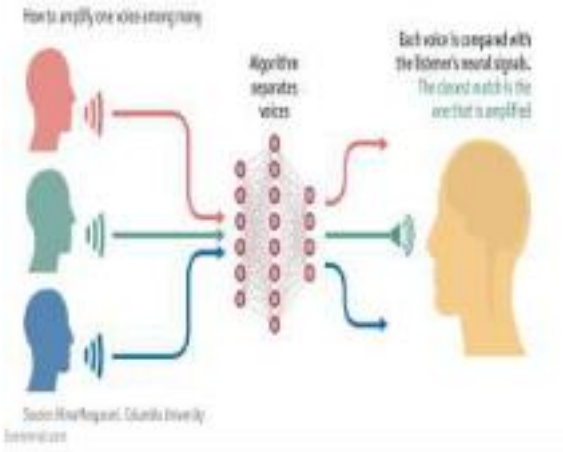
" Talker Seperation and Sequential Stream Seggregation in listeners with Hearing loss" Mackersie C

The purpose of this paper was to examine the relations between the ability to separate simultaneous sentences spoken by talkers of different gender and the ability to separate pitch patterns in a sequential stream segregation task. Simultaneous sentence pairs consisting of 1 sentence spoken by a male talker and 1 sentence spoken by a female talker were presented to 11 listeners with sensor neural hearing loss. Listeners were asked to repeat both sentences and were scored on the number of words repeated correctly. Separate scores were obtained for the male and female sentences. Sequential stream segregation was then measured using series of tones consisting of a fixed frequency (A) and a varying frequency tone (B). Tone series were presented in an ABA_ABA_... pattern starting at a varying frequency either below or above the frequency of the fixed 1000 Hz toneFor both the male and female talkers, better recognition scores were associated with lower fusion thresholds. Results suggest that the importance of streaming in the perceptual separation of talkers may depend on the nature of the information provided by the changing pitch stream.

" A Speech Corpus for multi-talker communications research"Robert. S Bolia

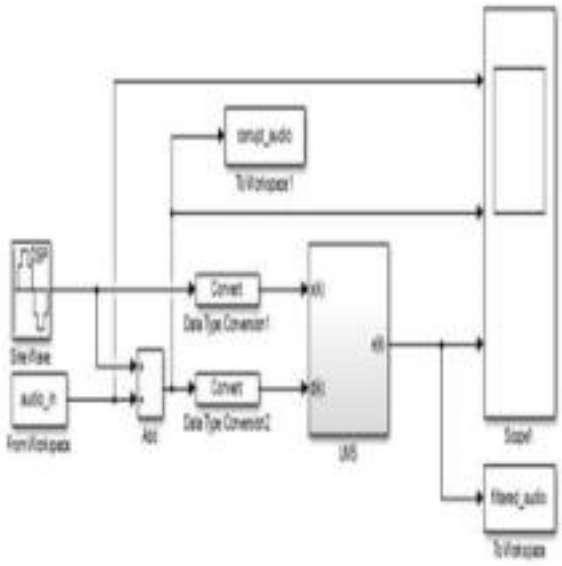
A database of speech samples from eight different talkers has been collected for use in multitalker communications research. Descriptions of the *nature* of the corpus, the data collection methodology, and the means for obtaining copies of the database are presented.

BLOCK DIAGRAM



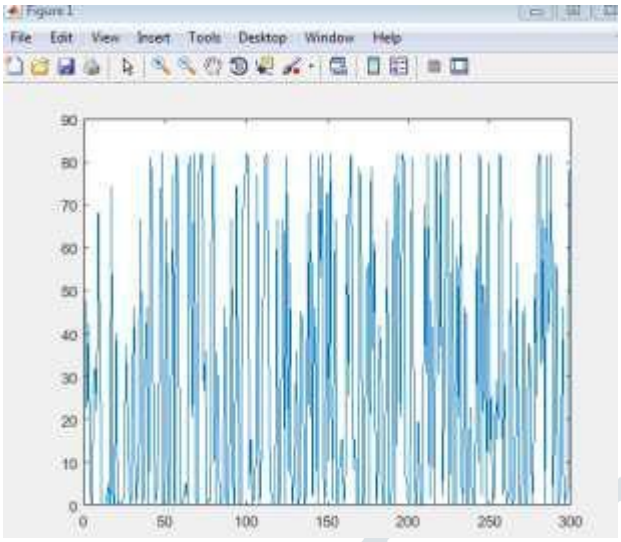
VI WORKING PROCESS

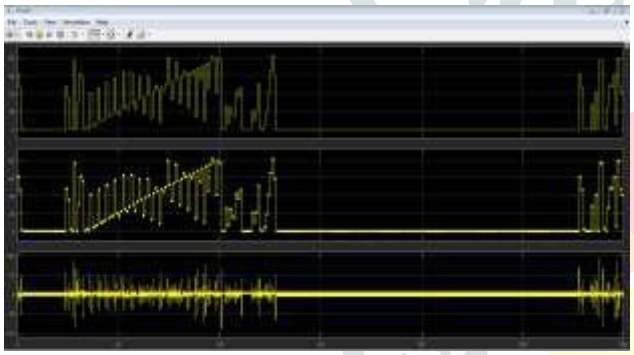
The block diagram is an interface between the user and the listener. The speaker separates the individual voices from the mixture of voices and provides to the listener. The implementation is done in both hardware and software. The software part uses MATLAB and Simulink to verify the results. The MATLAB code is processed and Simulink block is implemented and the implemented block is converted into Verilog code using HDL coder. The converted code is fed to the Field programmable gate array (FPGA) which plays the role of a hardware. The results of both hardware and software is verified and the analysis is matched.



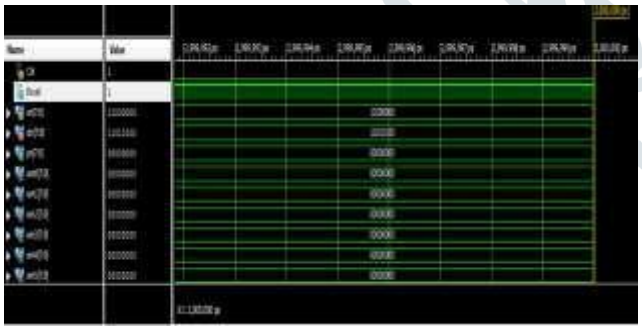
BLOCK DIAGRAM

VII RESULTS AND CONCLUSION





ADAPTIVE OUTPUTS



DESIGN SUMMARY

Device Utilization Summary				
Slice Logic Utilization	Used	Available	Utilization	Note(s)
Number of Slice Registers	0	11,440	0%	
Number of Slice LUTs	10	5,720	1%	
Number used as logic	10	5,720	1%	
Number using O6 output only	5			
Number using O5 output only	0			
Number using O5 and O6	5			
Number used as ROM	0			

TIMING REPORT

```
Timing Summary:
---------------
Speed Grade: -3

   Minimum period: No path found
   Minimum input arrival time before clock: No path found
   Maximum output required time after clock: No path found
   Maximum combinational path delay: 7.827ns

Timing Details:
---------------
All values displayed in nanoseconds (ns)

=========================================================================
Timing constraint: Default path analysis
  Total number of paths / destination ports: 118 / 8
-------------------------------------------------------------------------
Delay:               7.827ns (Levels of Logic = 5)
  Source:            b<2> (PAD)
```

VIII CONCLUSION AND FUTURE WORK

The proposed work on attentional selection in multi-speaker environment is designed and implemented. The desired audio source is separated and implemented in both hardware and software. In software part, it is implemented in MATLAB code and Simulink block. The Simulink block is converted into Verilog code which is being fed to the Field programmable gate array (FPGA). Field programmable gate array plays the role of the hardware. Thus results of both hardware and software are matched and the analysis is verified.

This prototype implementation is still far away from a real product and can be further improved in many aspects. The framework bridges the gap between the most recent advancements in speech prosthesis research and in future it moves closer to the development of cognitively controlled hearing aid for the hearing impaired.

REFERENCES

[1] Bregman A S and McAdams S 1994 Auditory scene analysis: the perceptual organization of sound J. Acoust. Soc. Am. 95 1177–8

[2] Alain C, Dyson B J and Snyder J S 2006 Aging and the perceptual organization of sounds: a change of scene Handbook of Models for Human Aging (New York: Academic) pp 759–69

[3] Mackersie C L 2003 Talker separation and sequential stream segregation in listeners with hearing losspatterns associated with talker gender J. Speech Lang. Hear. Res. 46 912–8

[4] Peelle J E and Wingfield A 2016 Listening effort in age-related hearing loss Hear. J. 69 10–12

[5] Peelle J E and Wingfield A 2016 The neural consequences of age-related hearing loss Trends Neurosci. 39 486–97

[6] Clark J L [and Swanepoel D W 2014 Technology for hearing loss—as we know it, and as we dream it Disabil. Rehabil. Assist. Technol. 9 408–13

[7] Ding N and Simon J Z 2012 Neural coding of continuous speech in auditory cortex during monaural and dichotic listening J. Neurophysiol. 107 78–89

[8] Ding N and Simon J Z 2012 Emergence of neural encoding of auditory objects while listening to competing speakers Proc. Natl Acad. Sci. USA 109 11854–9

[9] Horton C, D'Zmura M and Srinivasan R 2013 Suppression of competing speech through entrainment of cortical oscillations J. Neurophysiol. 109 3082–93

[10] Power A J, Foxe J J, Forde E J, Reilly R B and Lalor E C 2012 At what time is the cocktail party? A late locus of selective attention to natural speech Eur. J. Neurosci. 35 1497–503

[11] Zion Golumbic E M et al 2013 Mechanisms underlying selective neuronal tracking of attended speech at a 'cocktail party' Neuron 77 980–91

[12] Mesgarani N and Chang E F 2012 Selective cortical representation of attended speaker in multi-talker speech perception Nature 485 233–6

[13] O'Sullivan J A et al 2015 Attentional selection in a cocktail party environment can be decoded from single-trial EEG Cereb. Cortex 25 1697–706

[14] Horton C, Srinivasan R and D'Zmura M 2014 Envelope responses in single-trial EEG indicate attended speaker in a 'cocktail party' J. Neural Eng. 11 046015

[15] Dijkstra K et al 2015 Identifying the attended speaker using electrocorticographic (ECoG) signals Brain-Comput.