# Computerised Speech Processing in Hearing Aids using FPGA Architecture

V. Hanuman Kumar
Department of CS & SE
Andhra University,
Visakhapatnam -530003, India

Prof. P. Seetha Ramaiah
Department of CS & SE
Andhra University,
Visakhapatnam -530003, India

***Abstract*— The development of computerized speech processing system is to mimic the natural functionality of human hearing, because of advent of technology that used Very Large Scale Integration (VLSI) devices such as Field Programmable Gate Array (FPGA) to meet the challenging requirement of providing 100% functionality of the damaged human hearing parts. Here a computerized laboratory speech processor based on Xilinx Spartan3 FPGA system was developed for hearing aids research and also presented comparison details of the commercially available Hearing Aids. The hardware design and software implementation of the speech processor are described in detail. The FPGA based speech processor is capable of providing high-rate stimulation with 12 electrodes against conventional 8 electrodes in earlier research. Using short biphasic pulses presented either simultaneously or in an interleaved fashion. Different speech processing algorithms including the Continuous Interleaved Sampling (CIS) strategy were implemented in this processor and tested successfully.**

***Keywords- Speech processing system; VLSI; FPGA; CIS.***

## I. INTRODUCTION

Today digital signal processing applications (DSP), like e.g. speech processing or image processing, need considerable computing performance in order to work in real time. In some cases the performance of DSPs (Digital Signal Processor) is not sufficient to guarantee real time behavior. In these cases the problem is often solved by using dedicated hardware to do some pre-processing of the data or even implement the whole signal processing system. Since hard-wired ASICs (Application-Specific Integrated Circuits) are too expensive and not flexible enough, FPGAs (Field-Programmable Gate-Arrays) have proven to be a viable alternative for implementing DSP algorithms in hardware in the last decade. FPGAs belong to the class of programmable logic devices (PLD) and are reprogrammable or reconfigurable. The reason why hardware implementations of DSP algorithms show a superior performance in many cases compared to a software implementation on a DSP is a different computing paradigm. Compared to a DSP, which implements basically the sequential "von Neumann" computing principle (computing- in-time), FPGAs, can implement algorithms with much stronger parallelism; since the hardware is inherently parallel (computing-in-space). Unfortunately the design of FPGA hardware needs normally more effort than a software implementation on a DSP. Furthermore the development tools like hardware description languages (HDL) e.g. VerilogHDL or VHDL, simulators and synthesis and the way in which the implementation is described differs much from what the system or software engineers are used to, like e.g. describing the problem in tools like Mat lab or in languages like C/C++.

Hearing Aids are devices that provide partial hearing to deaf people through electrical stimulation of the auditory nerve (Wilson *et al*, 1988). A typical device consists of a speech processor, a behind-the-ear unit and an implantable device (Tierney*et al*, 1994). Sound is picked up by a microphone in the behind the- ear unit and sent to the speech processing device. Depending on the speech processing algorithm being implemented, the speech processor extracts various parameters from the input signal and determines the amplitude of the current pulses to be sent to the implant device (Loizou 1998). The amplitude information is transmitted to a receiver/stimulator circuit implanted under the scalp of the patient. This circuit sends current pulses to the electrodes implanted in the cochlea of the patient. As a result, the auditory nerve is excited and transmits nerve pulses to the brainstem and brain, where they are interpreted as sound (Ay *et al*, 1997). Hearing Aids (Cochlear Implants) are not available to most of the world's deaf and severely hearing impaired people due to the high cost (around $30,000 USD).

**Table 1: Comparison of Various Cochlear Implant Systems**

| Components | Parameters | Nucleus Freedom | Clarion HiRes90K | MED-EL MAESTRO | AU-NSTL Hearing Aid |
|---|---|---|---|---|---|
| External unit | Name and Key Features | Freedom:<br>Omni microphone<br>4 sound fields<br>IDR (20-75 dB)<br>Freq range: 100-8000 Hz<br>3 zinc-air batteries<br>(3-5 days) | Harmony:<br>Omni microphone<br>Dual-loop AGC<br>IDR (20-80 dB)<br>Freq range: 150-8000 Hz<br>Li ion batteries<br>(14-24 hours) | OPUS2:<br>Omni microphone<br>Dual-loop AGC<br>IDR (20-75 dB)<br>Freq range: 70-8500 Hz<br>3 zinc-air batteries<br>(3-5 days) | AU-NSTL-CI<br>Omni microphone<br>AGC<br>IDR (20-90 dB)<br>Freq range: 200-5500 Hz<br>Li ion batteries<br>(14-24 hours) |
| | Processing Strategies | CIS<br>SPEAK<br>ACE | CIS<br>MPS<br>HiRes Fidelity | CIS+<br>HDCIS<br>FSP | CIS<br>SPEAK |
| | Number of | 4 | 6 | 4 | 4 |
| RF link | RF carrier | 5 MHz | 49 MHz | 12 MHz | 4MHz |
| | Data rate | 0.5 MB/Sec | 1MB/Sec | 0.6 MB/Sec | .2 MB/Sec |
| Internal unit | Number of Electrodes | 22 | 16 | 12 | 8 out of 12 |
| | Number of Current Sources | 1 | 16 | 24 | 1 |
| | Current Range | 0- 1.75 mA | 0-1.9 mA | 0-1.2 mA | 0-1.2 mA |
| | Total Stimulation Rate (PPS) | 32 K | 83 K | 51 K | 12K |
| | Simultaneous Stimulation | No | Yes | Yes | No |
| Back Telemetry | Impedance Measurement | Yes | Yes | Yes | Yes |
| | Electric Field Imaging | No | Yes | Yes | No |

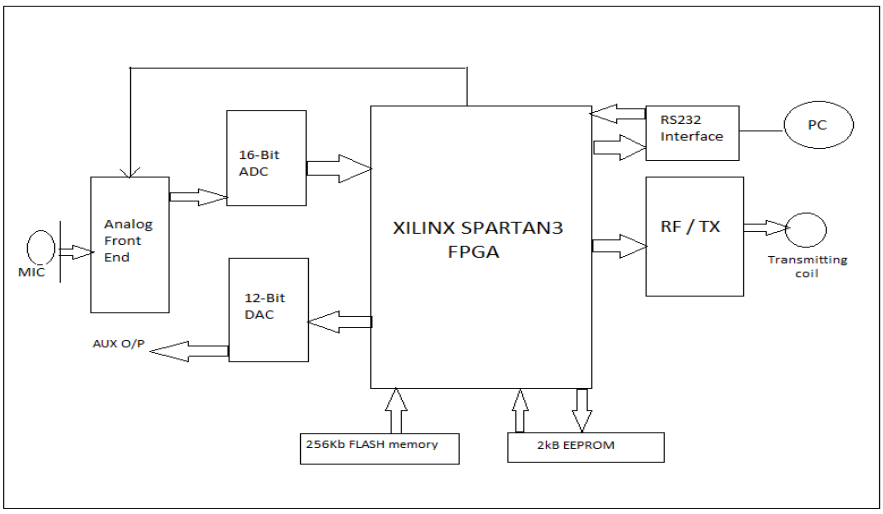


fig:1. Functional Block Diagram for Speech Processing

It is most unfortunate, as approximately 80% of the world's hearing impaired people live in developing countries, with highly limited healthcare budgets and widespread poverty with annual household income around $600 USD, cannot possibly afford the high cost of currently available implant systems. Many researchers are attempting to develop low cost but effective cochlear implant devices [Wilson, 1998], [Kim, 2007]. The three commercially available cochlear implant systems namely, Nucleus Freedom, Clarion HiRes 90K, MEDEL MAESTRO are compared in respect of the parameters: processing strategies, number of maps, RF carrier, data rate, number of electrodes, number of current sources, current range, total stimulation rate, simultaneous stimulation, impedance measurement and electric filed imaging with our developed AU-NSTL Cochlear Implant System [Fan-Gang Zing, 2008]. It is identified that the performance is similar to compared Cochlear implant Systems and parameter comparison information is given in Table 1.

## II. SPEECH PROCESSOR HARDWARE

A laboratory design of computerized speech processor for the Hearing Aids can accept speech or audio signals and transforms them into human understandable processed speech or audio to an implantable cochlear receiver-stimulator of 12 electrodes for making the deaf person to understand the speech or audio tones is designed, developed and evaluated. The working principle of the speech processing system involves capturing sound from the environment, processing sound into digital signals and delivering sound to the hearing nerve via electrode array in cochlea. The core of the system is Speech Processor (Xilinx spartan3 FPGA), 16-bit ADC,12-bit DAC, analog front end circuit includes a fixed gain amplifier, programmable gain amplifier ,an EEPROM and a serial programming interface shown in fig1. The speech processing system can drive a hearing aid receiver stimulator and excite 8-channel electrode array. In a typical application the system works in the following way. Sound waves are converted to electrical signals by the microphone and then fed to analog front-end circuit. An electric condenser microphone can be connected to the front panel auxiliary input socket.

The sound signals are amplified by fixed gain amplifier with a fixed gain of 30dB and based on volume control of speech processing device, programmable gain amplifier amplifies the output of the fixed gain amplifier and then amplified signal is attenuated by the sensitivity potentiometer and then passes to a fixed gain amplifier in the analog front-end circuit with a fixed gain of 30dB. The signal is filtered to eliminate noise before being converted to a 16-bit digital value by the 16-bit ADC, and transmitted to the Xilinx spartan3 FPGA based speech processor in 16-bit zero-padded frames via a serial interface.

A software-controlled programmable gain stage allows independent gain for each channel going into the ADC. The ADC's output can be digitally mixed with the DAC input. The ADC samples the filtered signals at a rate of 64K samples / sec. The 16-bit ADC and 12-bit DAC section is interfaced to Xilinx spartan3 FPGA via Serial Communication Interface (SCI). The FPGA based digital speech processor continually runs a program individually configured for the hearing-impaired user and their hearing prosthesis, is interrupted for every new data sample is received by the processor's SCI. The Xilinx spartan3 FPGA typically stores this sample in memory for future processing and may transfer a modified sample back to the SCI to be transmitted to the DAC, where the sample is converted to an analog signal to drive a hearing aid receiver. For auditory prostheses use, the Programmable Xilinx spartan3 FPGA based processor periodically construct data frames in the special format required for the cochlear implant receiver stimulator which determines the active electrodes and their current levels and then sends the encoded data frames serially with 171 Kbps rate to the RF transmitter.

The RF transmitter is based on ASK modulation and operates at 4MHz carrier frequency. The RF transmitter modulates the incoming encoded serial data and fed to the RF transmitting coil. The RF transmitting coil is seventeen turns, 175 strands Litz wire with High Q value. RF transmitter sends the data frames via the transcutanious inductive coupling to the receiver coil in the laboratory model receiver-stimulator. The receiver stimulator decodes the data and activates the specified electrodes, which stimulate nearby auditory neurons, giving the

user the sensation of hearing. The Xilinx spartan3 FPGA is used as the central processor running at a rate of 326 MHz of core clock frequency and Densities as high as 74,880 logic cells. Up to 1872 Kbits of total block RAM ,up to 520Kbits of distributed RAM, Three separate power supplies for the core (1.2V), IOs (1.2V to 3.3V), and Special function(2.5V) eliminating the need for power- consuming external RAM in cochlear implant applications. The on-chip peripherals consist of SCI-a Serial Communications Interface, a parallel Host Interface, and a Timer Module and relevant control signals are used to access external memories, as well as the encoder in this application. Programs are stored in an external EEPROM and the Xilinx spartan3 FPGA is programmed to boot from EEPROM. The mode pin logic levels are automatically altered when the programming cable is connected to the processor. The speech processing Software CIS for DSPMAP is developed in personal computer (PC) using VerilogHDL and configured Xilinx spartan3 FPGA using JTAG cable using Xilinx ISE Environment.

## III. SPEECH PROCESSOR SOFTWARE

The Hearing Aids are mainly classified into 2 categories Single channel and multi channel Aids.

### *A. Single Channel Implants*

Single channel Hearing Aids consist of a single pair of electrodes inserted in the cochlea. The main advantage of such a device is its simplicity, low cost and small size. The small size could enable it to be developed as a complete Behind the Ear (BTE) system. The two commercially produced single channel Implantable Hearing Aids were the House/3M single channel implants and Vienna/3M single channel implants.

### *B. Multi Channel Implants*

Single electrode cochlear implants are failed mainly due to stimulate only a particular place in the cochlea due to the single electrode used. Thus single electrode cochlear implants can only provide very limited frequency information, since they use only one electrode and perform simple spectral analysis. To better exploit the place/frequency mechanism found in the peripheral auditory system, multi-channel cochlear implants were developed. Multi channel implants provide electrical stimulation at multiple sites in the cochlea using an array of electrodes. An electrode array is used so that different auditory nerve fibers can be stimulated at different places in the cochlea, thereby exploiting the place mechanism for coding frequencies. Different electrodes are stimulated depending on the frequency of the signal. Electrodes near the base of the cochlea are stimulated with high-frequency signals, while electrodes near the apex are stimulated with low-frequency signals.

When researcher's starts designing of multi-channel implants, several issues are raised such as (i) how many electrodes should be used and how many channels are needed to obtain maximum speech understanding? And (ii) as more than one electrode will be stimulated, what kind of information should be transmitted to each electrodes? Researchers experimented with different numbers of electrodes, some are used with more number of electrodes (22) with selected stimulation, while other used a fewer number of electrodes (4-8) and stimulated all of them. Based on how researchers tried to address the second question, different types of speech processing algorithms are evolved. The various speech processing algorithms developed for multi channel implantable Hearing Aids can be divided into three main categories: waveform based speech processing algorithms, feature-extraction based speech processing algorithms and hybrid approach. These algorithms differ in the way the information is extracted from the incoming speech signal and presented to the electrodes for stimulation. The waveform based speech processing algorithms try to present some type of waveform (in analog or pulsatile form) derived from the speech signal by filtering into different frequency bands, while the feature-extraction based speech processing algorithms are try to present some type of spectral features, such as formants, derived using feature extraction algorithms. Hybrid algorithms presents the utilizing both algorithms. A brief coverage of these speech processing algorithms is given in Figure 2.Here the speech processor design using FPGA architecture system software comprises of two important modules, namely Programmable Speech Processing modules and Speech Data Encoding modules as shown in figure3.

The Programmability of the speech processing system design described herein provides the means to develop and test 8 Channel CIS speech processing algorithms. It provides flexibility and programmability according to patient's active electrodes. By using the impedance telemetry and clinical programming software, the audiologist identifies the active electrodes and sends this information to the Speech Processing module via Speech Data Encoding module of Xilinx spartan3 FPGA device. Based on this information Speech Processing Module implements the CIS algorithm for 4 to 8 channel. For example Encoding module sends the information to process 5 channel CIS algorithm, the speech processor changes the coefficients for 5-channel Band Pass filter and process the 5 channel CIS algorithm. The Band pass filters are configured so that the cutoff frequencies are adjusted depending on the number of active electrodes as shown in Table 2 and observe the block diagram of CIS algorithm in the figure4.

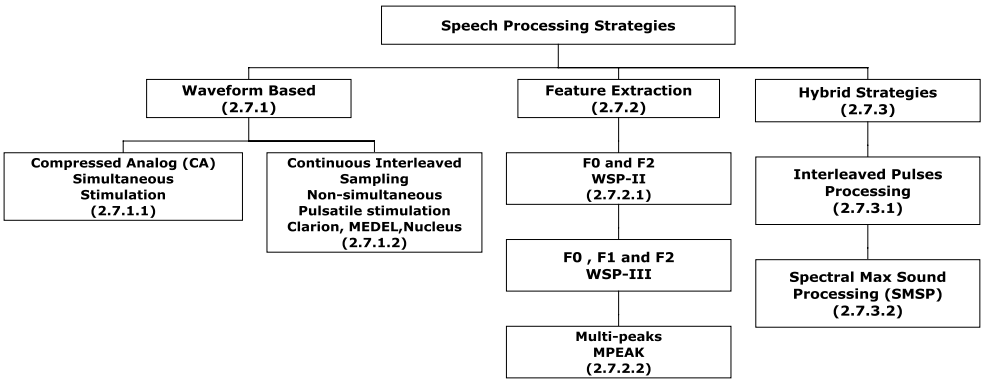


**Figure2:** Various Types of Speech Processing Algorithms

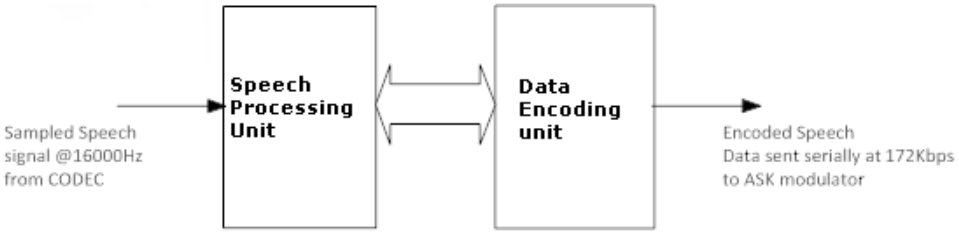


**Figure3:** Functional Modules of the Speech Processing Design

Table 2: Cut-off frequencies of Channels

| Channel number | Lower frequency | Upper frequency | Center frequency |
|---|---|---|---|
| 1 | 200 | 310 | 255 |
| 2 | 310 | 479 | 394 |
| 3 | 479 | 742 | 611 |
| 4 | 742 | 1149 | 946 |
| 5 | 1149 | 1779 | 1464 |
| 6 | 1779 | 2754 | 2266 |
| 7 | 2754 | 4263 | 3508 |
| 8 | 4263 | 6600 | 5432 |

Acoustic inputs from a microphone are directed to analog front-end (AFE) which amplifies the weak acoustic signals. The output from the AFE is Sampled with the sampling frequency of 64 KHz by the AD7683 and operated at 80MHz clock. The input samples which are collected by AD7683 converted to 16-bit digital data and send to Xilinx spartan3 FPGA (1.2V) via programmable interface between ADC and FPGA. The AD7683 interrupts Xilinx spartan3 FPGA whenever the input sample is available. These samples are windowed to produce a frequency response for the spectral analysis comparable with that of the commercially available auditory prostheses.

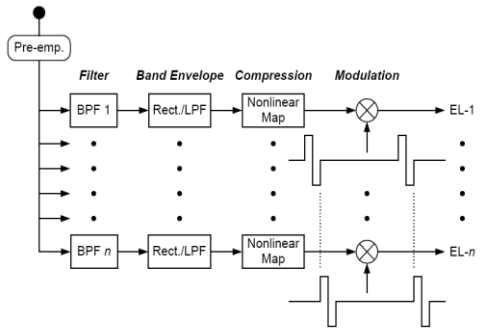


Figure4**:** Block Diagram of CIS Algorithm

The input signals are digitally filtered into 8 band-pass filters using Hamming window finite impulse response (FIR) filtering. Mat lab's Filter Design and Analysis tool (FDA Tool) is used to generate filter coefficients by using Hamming window of length 128bits. The input samples are stored in the circular buffer managed by Xilinx spartan3 FPGA. Each input acoustic sample is fed to a bank of band-pass channels. Each channel includes the stages of band-pass filtering, envelope detection, compression and modulation. The temporal envelope in each channel is extracted with full-wave rectifier followed by 32nd order low pass Hamming window FIR filter. The low-pass filters were designed to smooth the amplitude estimates in time, and were set to a cutoff frequency of 200Hz to allow maximal transmission of temporal envelope cues while avoiding aliasing when a relatively low carrier rates are used.

## IV. PERFORMANCE

A test program was run on FPGA based speech processor with core frequency at 326 MHz as the main processing unit, the ADC sampling rate at 64 kHz and the encoding module of FPGA formatting and transmitting two data frames via the RF

coil. A 128th order FIR program containing 8 band-pass filters runs on Xilinx Spartan3 FPGA processing unit. The input signal "FPGA" to the speech processor is received from the microphone and observed in digital Scope Coder as shown in fig-5 and corresponding responses from 8 channels is observed as shown in fig-6. Since the sampling rate is fixed at 64 KHz, we get 64000 samples for every second (i.e. 64 samples for every 1ms). These samples are fed to the filter bank, containing 8 band pass filters with frequencies ranging from 200Hz to 6600Hz. Rectified signals are generated from each filter, which are then fed to the low pass filter with cutoff frequency of 200Hz for envelope outputs. These 8 envelope signals are compressed using power law compression. These compressed signals of eight band-pass filters are transmitted to the encoding module in a base to apex order. This process is continuous and repeated for every 1.15ms. The FPGA based speech processor using Continuous Interleaved Sampling strategy has been run and tested with a laboratory model implant module.

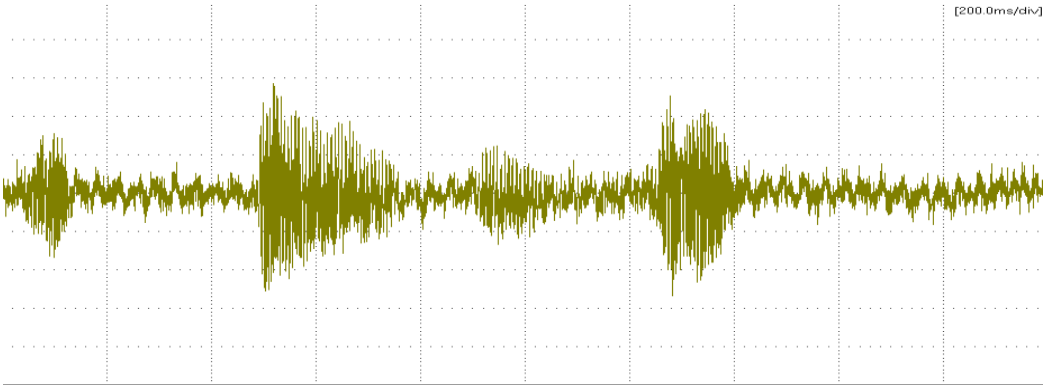


Fig.5: Input signal "FPGA"

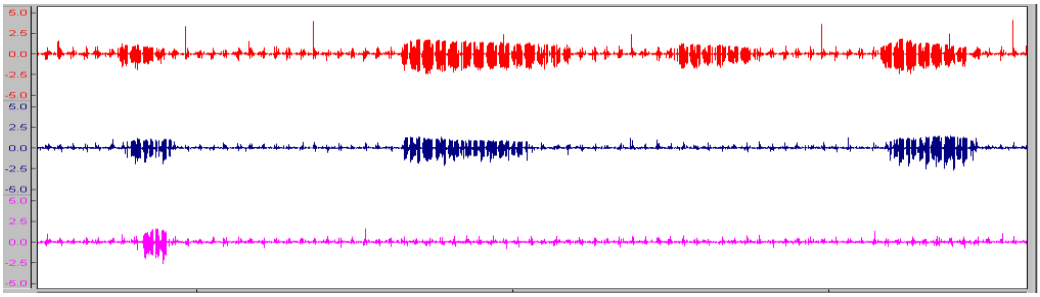

Fig6: Channel responses for signal "FPGA"

## V. CONCLUSIONS

The Computerized Speech Processing in Hearing Aids using FPGA Architecture has been developed particularly for use with a hearing Prostheses. The flexibility and computational power of the developed system allows speech processing using CIS strategy, which is tested and evaluated. Speech processing schemes may be improved by including environmental noise, increase the no of channels and speech intelligibility optimization.

## VI. ACKNOWLEDGEMENTS

The funding and technical support provided by NSTL, Visakhapatnam is acknowledged.

## REFERENCES

[1] Matlab-Homepage "www.mathworks.com"
[2] Ay, S., Zeng, F. -G. and Sheu, B. (**1997**). "Hearing with bionic ears: Speech processing strategies for cochlear implant devices,"IEEE Circuits & Devices Magazine, May 1997, 18-23.
[3] Boëx, C., Pelizzone, M. and Montandon, P. (**1996**). "Speechrecognition with a CIS strategy for the Ineraid multichannelcochlear implant," Am. J. Otol. **17**, 61-68.
[4] François, J., Tinembart, J., Bessat, C., Leone, P., Rossman, F.and Pelizzone, M. (**1994**). "Implants cochle aires: Un processeurportable pour le développement de l'algorithme CIS," in Actesde le conférence DSP 94, Paris
[5] Loizou, P. (**1998**). "Mimicking the human ear: An overview ofsignal-processing strategies for converting sound into electricalsignals in cochlear implants," IEEE Signal Processing Magazine**5**, 101-130.
[6] Loizou, P. and Poroy, O. (**1999**). "A parametric study of the CISstrategy," presented at the 1999 Conference on ImplantableAuditory Prostheses, Asilomar, Monterey, CA.
[7] [Clark, 2006 ] Graeme M. Clark, "The multiple-channel cochlear implant: the interface between sound and the central nervous system for hearing, speech, and language in deaf people—a personal perspective", Phil. Trans. R. Soc. B , 2006, Vol. 361,pp 791–810
[8] [Greenwood, 1990] Greenwood, D. D. (1990). "A cochlear frequency-position function for several species: 29 years later", Journal *of the Acoustical Society of America*, *Vol. 87, 1990, pp* 2592-2605.
[9] [Hirshorn, 1986] Michael S . Hirshorn, Dianne J. Mecklenburg, Judith A. Brimacombe, "Nucleus 22-channel cochlear implant: Preliminary observations", Journal of Rehabilitation Research and Development, April 1986, Vol . 23, No . 2, pp 27-33.
[10] Dua, S. (2010). PC And Speech Recognition Based Electrical Device Control. International Journal of Advanced Computer Science and Applications - IJACSA, 1(July), 46-49.
[11] Anifowose, F. A. (2010). A Comparative Study of Gaussian Mixture Model and Radial Basis Function for Voice Recognition. International Journal of Advanced Computer Science and Applications - IJACSA, 1(3), 1-9.