

Analysis of Speech Processing Strategies in Cochlear Implants

Kouachi Rouiha, Djedou Bachir and Bouchaala Ali
Laboratory of study, research in Instrumentation and Communication of Annaba,
University Badji Mokhtar, BP 12, Annaba, 23000 Algeria

Abstract: Cochlear implants can restore partial hearing to profoundly deaf people; the main function of these prostheses is to electrically stimulate the auditory nerve using an electrode array inserted in the cochlea. The acoustic signal is picked up by a microphone and analyzed. Then the extracted parameters of the signal are coded to generate electrical signals reconstituting the original signal. Currently all commercialized implants are multichannel they allow to stimulate the auditory nerve at different place of the cochlea, exploiting the tonotopic coding of the frequencies. This research will present an overview of various signal processing techniques that have been used for cochlear prosthesis over the years.

Key words: Cochlear prosthesis, speech coding, signal processing

INTRODUCTION

A Cochlear Implant (CI) is a device that provides partial hearing to profoundly deaf people. These people are usually unable to obtain any benefit from conventional hearing aids no matter how loud the sound is. The common reason for this phenomenon is loss of hair cells in the inner ear. In such an event, direct stimulation of the nerve fibers provides a means to restore the patients hearing. Such a device is called a Cochlear Implant device, operates by stimulating electrically the auditory nerve directly thereby bypassing the normal hearing mechanism.

Many factors affect the performance of CI patients, including etiology of hearing loss, duration of deafness, neural survival, speech processing strategy, etc. Because of all these factors, there is a large variability in performance among patients. Performance may vary from a low of 0% correct to a high of 100% correct. To account for the variability in performance manufacturers started introducing several speech processing strategies in their processors. For instance Nucleus 24 processor offers three different strategies, while the Clarion implant processor offers three strategies.

Different speech processing strategies are used in order to obtain optimum patient performance. Selecting the right strategy for each patient is not easy and no formal procedure is currently used to do so.

The purpose of this study therefore was to investigate the performance of the six available speech processing strategies on the Clarion processor in quiet and noise conditions.

Cochlear implants: Research has shown that the most common cause of deafness is the loss of hair cells rather than the loss of auditory neurons. This was very encouraging for cochlear implants because the remaining neurons could be excited directly through electrical stimulation. A cochlear prosthesis is therefore based on the idea of bypassing the normal hearing mechanism (outer, middle and part of the inner ear including the hair cells) and electrically stimulating the remaining auditory neurons directly.

Several cochlear implant devices have been developed over the years. All the implant devices have the following features in common: a microphone that picks up the sound, a signal processor that converts the sound into electrical signals, a transmission system that transmits the electrical signals to the implanted electrodes and an electrode or an electrode array (consisting of multiple electrodes) that is inserted into the cochlea by a surgeon (Fig. 1). In single-channel implants only one electrode is used. In multichannel cochlear implants, an electrode array is inserted in the cochlea 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

Corresponding Author: Kouachi Rouiha, Laboratory of study, research in Instrumentation and Communication of Annaba,
University Badji Mokhtar, BP 12, Annaba, 23000 Algeria

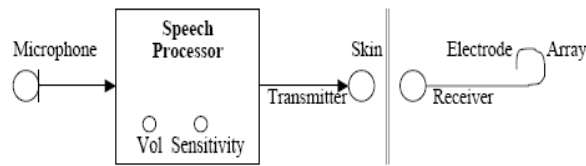


Fig. 1: Components of cochlear implant system

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. The signal processor is responsible for breaking the input signal into different frequency bands or channels and delivering the filtered signals to the appropriate electrodes.

Waveform strategies:

Compressed-Analog (CA) approach: The compressed-analog (CA) approach was originally used in the Ineraid device manufactured by Symbion, Inc., Utah^[2]. The signal is first compressed using an automatic gain control and then filtered into four contiguous frequency bands, with center frequencies at 0.5, 1, 2 and 3.4 kHz. The filtered waveforms go through adjustable gain controls and then sent directly through a percutaneous connection to four intracochlear electrodes. The filtered waveforms are delivered simultaneously to four electrodes in analog form. The CA approach, used in the Ineraid device, was very successful because it enabled many patients to obtain open-set speech understanding. Dorman *et al.*^[3] report, for a sample of 50 Ineraid patients a median score of 45% correct for word identification in sentences.

Continuous Interleaved Sampling (CIS): The CA approach uses analog stimulation that delivers four continuous analog waveforms to four electrodes simultaneously. A major concern associated with simultaneous stimulation is the interaction between channels caused by the summation of electrical fields from individual electrodes. These interactions may distort speech spectrum information and therefore degrade speech understanding.

Researchers at the Research Triangle Institute (RTI) developed the Continuous Interleaved Sampling (CIS) approach^[4] which addressed the channel interaction issue by using non-simultaneous, interleaved pulses.

Trains of biphasic pulses are delivered to the electrodes in a non-overlapping (non-simultaneous) fashion, that is, in a way such that only one electrode is stimulated at a time. The amplitudes of the pulses are derived by extracting the envelopes of bandpassed

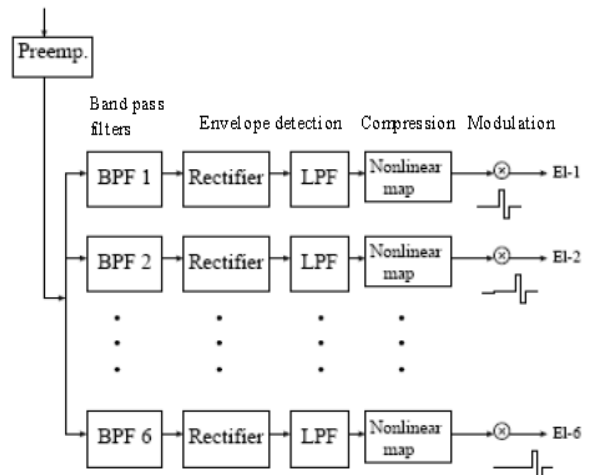


Fig. 2: Block diagram of the CIS strategy

waveforms. The signal is first pre-emphasized and passed through a bank of band pass filters (Fig. 2). The envelopes of the filtered waveforms are then extracted by full-wave rectification and low-pass filtering (typically with 200 or 400 Hz cutoff frequency). The envelope outputs are finally compressed and then used to modulate biphasic pulses. A non-linear compression function (e.g., logarithmic) is used to ensure that the envelope outputs fit the patient's dynamic range of electrically evoked hearing. Trains of balanced biphasic pulses, with amplitudes proportional to the envelopes, are delivered to the six electrodes at a constant rate in a nonoverlapping fashion. The rate at which the pulses are delivered to the electrodes has been found to have a major impact on speech recognition^[1]. High pulse-rate stimulation typically yields better performance than low pulse rate stimulation. Comparison between the CA and CIS approach revealed higher levels of speech recognition with the CIS approach^[4].

Feature- extraction techniques: These are based on extracting the spectral information of the input signal and using this information to generate the stimulus to the electrodes. For proper perception of speech it is important to present the formant frequencies (F1-F3). The frequency of this periodic waveform is called the fundamental frequency (F0). The following three peaks in frequency are called F1, F2 and F3 and are known as formants. The Nucleus implant manufactured by Cochlear Corporation and developed at the University of Melbourne uses these techniques. Some of the techniques used in this device are discussed in the following sections. The Nucleus cochlear implant was a 22-electrode device.

F0/F2: In this scheme^[5,6] F0 is estimated using a zero-crossing detector at the output of a 270Hz lowpass

filter. F2 is estimated using a zero-crossing detector at the output of a 1000 -4000 Hz bandpass filter. The amplitude of the F2 formant is obtained after rectification and lowpass filtering of the output of the bandpass filter. The appropriate electrode (among the 22 electrodes) is stimulated at the rate of F0 pulses sec^{-1} . For unvoiced speech the electrode is stimulated at an average rate of 100 pulses sec^{-1} .

F0/F1/F2: This strategy was an improvement on the previous F0/F2 technique since it also included the first formant F1. A zero-crossing detector was used at the output of a 280-1000 Hz bandpass filter. Two sets of electrodes were now stimulated, one with the F1 formant information and the other with the F2 formant information. The F1 information was used to stimulate the apical electrodes and the F2 information for the basal electrodes. 200 μ sec pulses were used with a separation of 800 μ sec to avoid channel interaction. The pulse amplitudes were proportional to the amplitudes of the F1 and F2 formants and the stimulation rate was still F0 pulses sec^{-1} .

Due to the extra information in the F0/F1/F2 device the performance was improved as compared to the earlier F0/F2 device. The concept of using formant information works well for low-frequency signals, but with higher frequency speech like consonants different strategies had to be used.

MPEAK^[7]: A further improvement over the F0/F1/F2 scheme was the MPEAK (or MULTIPLEAK) that extracted and used high frequency information from the input signal to stimulate the electrodes. A bandpass filter of 800-4000 Hz was used to determine F2. Further three additional bandpass filters (2000-2800 Hz, 2800-4000 Hz, 4000-6000 Hz) were used to extract the high frequency information. Thus four electrodes were stimulated at F0 pulses sec^{-1} for voiced speech and an average of 250 pulses sec^{-1} for unvoiced speech. Due to the availability of the high frequency information the performance of patients improved with this scheme especially for consonants.

The major disadvantage of feature extraction schemes is that they introduce errors in the determination of the formant frequencies. Thus further research sought to look at other techniques to represent speech in cochlear implants.

Spectral maxima sound processor (SMSP): Instead of performing feature extraction the SMSP analyses the speech by a bank of 16 bandpass filters ranging from 250-5400 Hz. The outputs of the bandpass filters are rectified and lowpass filtered (200 Hz) and the six largest outputs are selected from amongst the 16. Only these six electrodes corresponding to the maximum amplitudes are stimulated in each cycle at a rate of 250 pps.

CONCLUSION

Cochlear implant patients have shown widely varying results. This is probably partly due to the history of their deafness and of their implantation

Much of the success of cochlear implants was due to the advancement of signal processing techniques developed over the years. While this success is very encouraging, there is still a great deal to be learned about electrical stimulation of the auditory nerve and many questions to be answered. Future research in cochlear prosthesis should:

Continue investigating the strengths and limitations of present signal-processing strategies including CIS-type and SPEAK-type strategies. The findings of such investigations may lead to the development of signal-processing techniques capable of transmitting more information to the brain.

REFERENCES

1. Wilson, B., 1993. Signal Processing. In: Cochlear Implants: Audiological Foundations, Tyler, R. (Ed.). Singular Publishing Group, Inc., pp: 35-86. ISBN-10: 1879105810
2. Eddington, D., 1980. Speech discrimination in deaf subjects with cochlear implants. J. Acoust. Soc. Am., 68(3): 885-891.
<http://scitation.aip.org/dbt/dbt.jsp?KEY=JASMAN&Volume=68&Issue=3>
3. Dorman, M., M. Hannley, K. Dankowski, L. Smith and G. McCandless, 1989. Word recognition by 50 patients fitted with the symbion multichannel cochlear implant. Ear and Hear, 10: 44-49.
<http://www.ear-hearing.com/pt/re/earhearing/>
4. Wilson, B., C. Finley, D. Lawson, R. Wolford, D. Eddington and W. Rabinowitz, 1991. Better speech recognition with cochlear implants. Nature, 352: 236-238.
<http://cat.inist.fr/?aModele=afficheN&cpsidt=4962682>
5. Clark, G., 1987. The University of Melbourne-Nucleus multi-electrode cochlear implant. Adv. Oto-Rhino-Laryngol., 38: 1-189. Basel: Karger ISBN: 3-8055-4575-4
6. Seligmna, P., J. Patrick, Y. Tong, G. Clark, R. Dowell and P. Crosby, 1984. A signal processor for a multiple-electrode hearing prosthesis. Acta Otolaryngol., 411: 135-139.
<http://journalreview.org/v2/articles/view/6596837.html>
7. Loizou, P., 1999. Introduction to cochlear implants. IEEE Eng. Med. Biol. Mag., 18: 32-42.
<http://www.utdallas.edu/~loizou/cimplants/cpubs.htm>