Precise stimulation in auditory neuroimplants

Søren W. Kristiansen, Hans-Heinrich Bothe

Technical University of Denmark, Centre for Applied Hearing Research, 2800 Lyngby, Email: hhb@elektro.dtu.dk

Introduction

The work concerns auditory neural implants with specific application to cochlear implants. A usual cochlear implant consists of an external and internal part, with the first built by the acoustic receiver, processor and transcutaneous transmitter to the internal part. The internal part consists of receiver electronics and an electrode array inserted into the cochlea. Electrical stimulation of the auditory nerve fiber array is then conducted through the inner hair cells array.

A number of problems with 'electrical hearing' arises, with the most obvious being i) a reduction in available frequency channels to resolve the incoming acoustic stimulation and ii) electrical channel interaction/ crosstalk due to electrode proximity and stimulation methods.

The aim of the project is to reduce the channel interaction by employing a new electrode stimulation method based on phased array channels and to verify the results with a simulator system, which distorts the acoustic signal so that a person with 'normal hearing' has a similar sound perception as a person wearing cochlear implants.

After a brief introduction to auditory neuroimplants, the paper describes three parts of the project, i) the focusing processor, ii) the CI simulator, and iii) a comparison of speech intelligibility with and without focusing processor.

Auditory Neuroimplants

Neural signal processing in the auditory pathway of humans starts in the cochlea in the inner ear and proceeds via the auditory nerve and several processing nuclei towards the primary auditory cortex. On the way, all processing and transmission stages must be functioning in order for a person to perceive and understand the received sound signal. Auditory neuroimplants can be employed if at specific stages of the auditory pathway, the signal flow is significantly distorted or even interrupted. In an intact auditory pathway, sound vibrations mechanically stimulate the primary receptor cells in the cochlea, the inner hair cells. In neuroimplants, sound-related electric current is injected instead at a specific location of the pathway, stimulating the nerve cells electrically. Intact parts towards the cortex can thus, receive a signal that can further be processed in the intact stages towards the brain and understood by the person.

The best known type of neuroimplant is the cochlear implant (CI). CI make use of the tonotopic organization of the cochlea, i.e. the audio energy is mapped onto specific frequency bands equivalent to locations along the spiral nerve fibre array in cochlea. They stimulate the spiral nerve cells with the help of a linear electrode array, fused into the cochlea. Each electrode corresponds to one frequency band and is separately controlled. The principal processing steps are shown in figure 1.

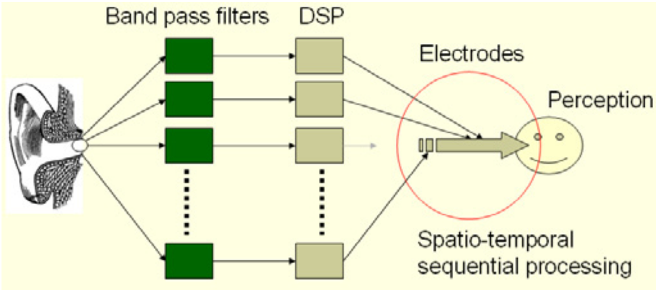

Figure 1: Principle signal processing strategies in a cochlear implant (CI). Acoustic signals are received with an microphone of the external device, bandpass-filtered, further processed within the frequency bands and fed to the electrodes of the implanted device.

The filter parameters are programmable, with the centre frequencies typically located between 500–6000 [Hz]. In most commercial CI, a number of different speech coding strategies are implemented in the DSP blocks. One typical strategy is to activate only a limited number of electrodes at a time, selected by the RMS values of their envelope signals. An activated electrode stimulates the nerve cells in its neighbourhood by AC current injection with a (constant) frequency between 500–1500 [Hz]. The amplitude is given by the present RMS value in the related frequency band. Obviously, the fine structure of the sound signal is removed. Speech perception is still possible with the help of the low-frequency spectral modulations (0–25 [Hz]) of speech-shaped noise signals.

The number of frequency bands and correlated stimulation electrodes vary between 7-22 with spatial distances between neighbouring electrodes of approximately one mm, plus one or more reference electrodes. Stimulation electrodes employ monopolar stimulation using the same reference electrodes. This improves the stimulation efficacy, but yields a relatively wide spatial spread of spiral nerve cell activation, i.e. significant channel interaction. Channel interaction may yield undesired loudness fluctuations caused by a smearing of the spatial mapping or by the creation of artificial stimulation loci (see [1], [2], [3]); it is further, unpredictably influenced by the nonlinear nature of spiral nerve cell excitation. A common method to reduce channel interaction is to time-multiplex electrode activation and to use high-frequency (> 1 [kHz]) AC stimulation techniques.

Our aim is to remove this channel interaction by spatial restriction of the nerve cell excitation through extra current injection in the neighbouring electrodes, as motivated by [4] and briefly illustrated in the following.

Focusing Processor

Let the electrode array use N single stimulation electrodes. Then the principle of the focusing processor is to compensate the influence of neighbour electrodes on the

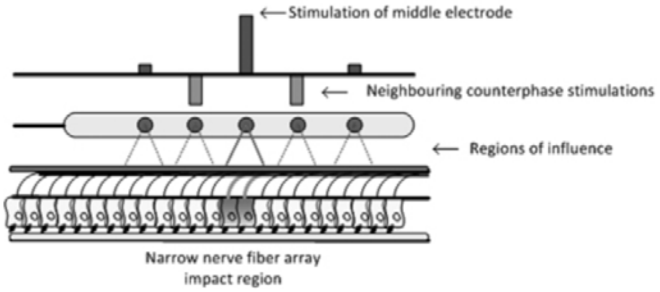

Figure 2: Focused stimulation using spatial counter-phase stimulation. The upper part of the figure shows the counter-phase and counter-counter-phase stimulation of neighbor electrodes for activation of the center frequency band. The spatial region of influence of the electrode on the spiral nerve cells is focused, yielding more precise activation of the cells. For real environmental or speech sounds, several frequency bands carry significant information. The resulting voltages in the electrode array are then calculated as a linear superposition of the voltages related to all active frequency bands.

nerve cells below by a counter-phase current injection/ or voltage on these electrodes.

The case of focused stimulation for one active frequency band is shown figure 2. The upper part illustrates the concerted stimulation of several electrodes with counter-phase voltages of immediate neighbours and respective counter-counter-phase voltages of next neighbours. Counter-phase voltage electrodes have to be adjusted to compensate the influence of the centre electrode on nerve cells below neighbouring electrodes; the voltages of the next neighbours need to compensate the influence from the immediate neighbours, etc. The lower part of the figure shows focused activation of the spiral nerve cell array.

Our approach is now to model the spatial voltage distribution $V_i(x)$ for single frequency band activation by equidistant discrete samples of a cosine function with spatially decreasing amplitude,

$$V_i(x) = \frac{A(i)}{1+|\pi(x-i)|^{2p_i}} \cos(\pi(x-i)), \quad (1)$$

with i being the number of the stimulating electrode or frequency band, x of a neighbour frequency band on which i has influence, $A(i)$ the amplitude of the activation in i, and p_i a parameter that has to be fitted to the efficacy of each electrode contact i in an individual person.

In case of concurrent activation of different frequency bands with electrode voltages $V_i(x)$ (1), the total voltages $V(x)$ are derived from a linear superposition of the $V_i(x)$, yielding

$$V(x) = \sum_{i=1}^{N} V_i(x) = \sum_{i=1}^{N} w_{i,x} A(i) \quad (2)$$

with the linear superposition coefficients

$$w_{i,x} = \frac{\cos(\pi(x-i))}{1+|\pi(x-i)|^{2p_i}}. \quad (3)$$

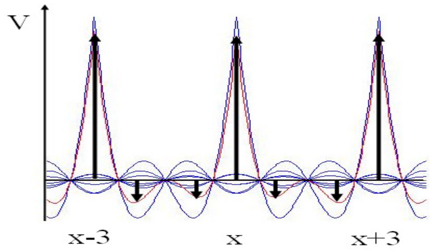

Figure 3: Focused stimulation using linear superposition of weight functions cos(2πx) / [1+|2π x|^2] with electrodes at discrete spatial positions x and stimulation voltage V. The arrows show the total stimulation voltage at electrodes (x) after modification. The total power consumption of the CI is slightly increased due to stimulation in all electrodes.

As an example, simultaneous stimulation of three electrodes at locations x and $x+/-3$ with equal amplitude is illustrated in figure 3. Resulting voltages at the electrode positions are marked by arrows. The total power consumption is in this exemplary case slightly increased by focused stimulation, as intermediate electrodes now also need to be stimulated.

The parameter p_i in (1) and (3) can be determined for individual persons with the help of transadmittance measurements between the implanted electrodes, as described in [4], or estimated a priori by prototype settings.

Simulation System

In order to evaluate the influence of the focusing processor on sound perception, a CI simulation system was designed and implemented. Based on the vocoder principle ([5], [6]), it mimics the processing steps in a CI while assuming that all electrodes have a reasonable good contact with the spiral nerve cells. The principle processing blocks are illustrated in figure 4.

The speech signal is filtered by a gammatone filter bank with the centre frequencies tonotopically ordered according to Greenwood's cochlear frequency-position function [7],

$$f(x) = \alpha(10^{\beta x} - \gamma), \quad (4)$$

with α=165.4, β=2.1 (x expressed in millimetres from the apex), γ=1 for a prototype human cochlea (see figure 5).

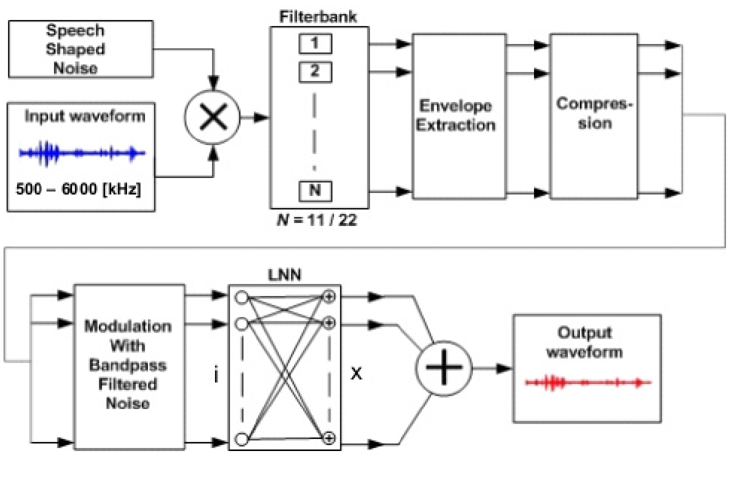

Figure 4: Complete CI simulation system including the linear neural network LNN as focusing processor.

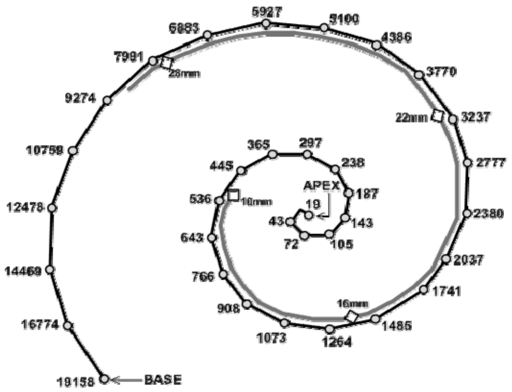

Figure 5: Frequency-position mapping in a prototype human cochlea modeled with Greenwood's function (4). The circles show characteristic frequencies, and the squares the position of four exemplary, equidistantly spaced electrodes at 10, 16, 22, and 28 [mm] from the apex.

Then, the envelope of the signal is extracted for each frequency band with the help of a Hilbert transform; the resulting low-frequency signal is compressed with respect to the level compression mechanisms in the cochlea [6], also considering the difference of dynamic ranges of normal and 'electrical' hearing [5]; each frequency band is modulated with a white-noise carrier signal, bandpass-filtered with the same gammatone filterbank, and fed into the linear neural network (LNN). The LNN incorporates the superposition weights $w_{i,x}$ taken from (3) and some adjustable delays for the compensation of different signal delays in the gammatone filterbank. The reason for the inclusion of these delays is that in our tests, the processed signal is presented to normal-hearing subjects, so that the delays in the frequency bands seem to be applied double; at first in the filterbank of the simulator, and thereafter in the 'natural' filters. Too large unnatural delays must be prohibited with respect to nonlinear excitation mechanisms of the spiral nerve cells through the timing of action potential arrival. We calculated the relative delays across the frequency bands such that the noisy signals at the outputs of the LNN were correlated on a small scale so that the addition of noise does not cancel each other out.

The output signals of the LNN are simply summed-up and sent to the headphones for performance evaluation.

Test with Normal-Hearing Subjects

The resulting simulation system with and without focusing processor was evaluated with four normal-hearing 'naïve' test persons. Two sets of the Dantale test [8] with 25 Danish words and single syllables each were used as text corpus. Speech-shaped noise was added to the input signal in order to simulate a cocktail-party-like effect. Figure 6 shows the mean hit-rates with and without focusing processor for all words and all persons over the signal-to-noise ratio of the input signal for simulators with 11 or 22 virtual electrodes, respectively. The parameters p_i were set to p=1 in both cases.

At first, a clear improvement of the virtual CI with 22 electrodes against 11 electrodes can be reported. The curves of the hit-rates are apart from each other by more than 10dB.

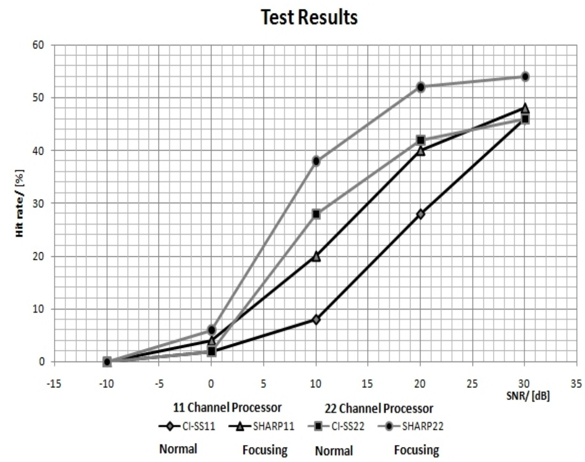

Figure 6: Hit rates for a test with four 'naïve' persons, employing two lists of the Dantale corpus [8] with 25 Danish words of one syllable, each.

This strong difference may be caused by an improvement of the signal for more electrodes, together with the fact that the test persons had no time to develop significant learning effects during the short time of the single test runs. Longer learning periods could well improve the results and make them more similar.

A relatively large improvement of the speech intelligibility can be reported for both virtual arrays, when the focusing processor is switched on. The improvement starts between s/n=0-10 dB and reaches values of 10dB. Maximum hit rates are between 45-55%, i.e. clearly below 100%. The yet low maximum scores may be explained by the similarity of the words/syllables in the text corpus and the fact that 'naïve' test persons are not used to hear or judge signals, which emulate the perception of people wearing CI. We therefore asked the test persons to subjectively evaluate the intelligibility of the single words on a scale between 1 and 10, with 1 being completely non-intelligible and 10 just perfect. The mean values for different s/n-ratios across test persons and words are shown in figure 7. The highest scores were given to the array with 22 electrodes and focusing processor with the exception at s/n=0dB, the second highest to the array with 11 electrodes and focusing processor;

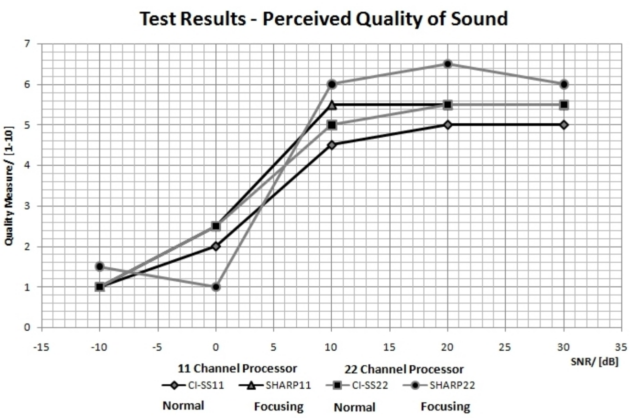

Figure 7: Subjective quality assignments for the words masked by a noise signal for a test with three 'naïve' persons, employing two lists of the Dantale corpus [8] with 25 Danish words of one syllable, each.

below these curves are those of the arrays without focusing processor. It is apparent that even though the array with 11 electrodes and focusing processor has higher subjective intelligibility scores than the array with 22 electrodes and no focusing processor, the actual hit-rates show the opposite behavior. But the differences are relatively small, so that extensive tests with more test persons, a larger text corpus, and several trials per person and corpus could clarify this point.

Conclusions and Future Work

Tests with 'naïve' normal-hearing people show that the employment of a focusing processor for arrays with 11 or 22 electrodes can yield improved speech intelligibility in simple hit-rate tests even though the speech intelligibility for the test material is rated relatively bad by the test persons.

It is now interesting to know how persons wearing cochlear implants evaluate the perception effects of the focusing processor. Such tests are still pending, but they are planned, based on speech intelligibility tests employing words also from different languages.

The simulator system including the paradigms of the focusing processor may further be improved and fitted to test results with persons wearing CI, employing a detailed emulation of the signal processing steps in commercial CI.

References

[1] Shannon, R.V.: Multichannel electrical stimulation of the auditory nerve in man: II. Channel interaction. Hear. Res. **12** (1983), 1-16

[2] Boex, C., C. de Balthasar, M.I. Kos, M. Pellizzone: Electrical field interactions in different cochlear implant systems. J. Acoust. Soc. Am. **114** (2003), 2049-2057

[3] Fu, Q.J., G. Nogaki: Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. J. Assoc. Res. Otolaryngol. **6** (2005);19-27

[4] Honert, C. van den, D.C. Kelsall: Focused intracochlear electric stimulation with phased array channels. J. Acoust. Soc. Am. **121**, 8 (2007), 3703-3716

[5] Fu, Q.-J.: Speech perception in acoustic and electric hearing. PhD dissertation, University of Southern California, CA (1997)

[6] Loizou, P.C.: Speech processing in vocoder-centric cochlear implants. In: A.R. Møller (ed.), Cochlear and brainstem implants. Karger (2006), 109-143

[7] Greenwood, D.D.: A cochlear frequency-position function for several species – 29 years later. J. Acoust. Soc. Am. **87**, 6 (1990), 2592-2605

[8] Elberling, C, C. Ludvigsen: DANTALE, a new Danish speech material. Scandinavian Audiology **18** (1989), 169-175