Lappeenranta University of Technology (LUT)

School of Business and Management

Computer Science Degree Program

Bachelor's Thesis

Esa-Petri Tirkkonen

Using electroencephalography signals to control acoustical processing

Supervisors and Examiners:

Associate Professor Erja Mustonen-Ollilla (LUT)

Dr Tom Campbell (University of Helsinki)

ABSTRACT

Lappeenranta University of Technology

School of Business and Management

Computer Science Degree Program

Esa-Petri Tirkkonen

Using EEG signals to control acoustical processing

Bachelor’s Thesis

2016

29 pages, 1 figures, 0 tables, 0 appendices

Examiners: Associate Professor Erja Mustonen-Ollilla, Dr Tom Campbell

Keywords: EEG, acoustic processing, hearing aid, brain-computer interface (BCI), cocktail party effect, neuroscience, sound processing, hearing, noise, design science.

Current hearing-assistive technology performs poorly in noisy multi-talker conditions. The goal of this thesis was to establish the feasibility of using EEG to guide acoustic processing in such conditions. To attain this goal, this research developed a model via the constructive research method, relying on literature review. Several approaches have revealed improvements in the performance of hearing-assistive devices under multi-talker conditions, namely beamforming spatial filtering, model-based sparse coding shrinkage, and onset enhancement of the speech signal.

Prior research has shown that electroencephalography (EEG) signals contain information that concerns whether the person is actively listening, what the listener is listening to, and where the attended sound source is. This thesis constructed a model for using EEG information to control beamforming, model-based sparse coding shrinkage, and onset enhancement of the speech signal. The purpose of this model is to propose a framework for using EEG signals to control sound processing to select a single talker in a noisy environment containing multiple talkers speaking simultaneously.

On a theoretical level, the model showed that EEG can control acoustical processing. An analysis of the model identified a requirement for real-time processing and that the model inherits the computationally intensive properties of acoustical processing, although the model itself is low complexity placing a relatively small load on computational resources. A research priority is to develop a prototype that controls hearing-assistive devices with EEG. This thesis concludes highlighting challenges for future research.

TIIVISTELMÄ

Lappeenrannan teknillinen yliopisto
School of Business and Management
Tietotekniikan koulutusohjelma

Esa-Petri Tirkkonen

Aivosähkökäyrien käyttö akustisen prosessoinnin ohjaamisessa
Kandidaatintyö

2016
29 sivua, 1 kuvaa, 0 taulukkoa, 0 liitettä

Työn tarkastajat: tekniikan tohtori Erja Mustonen-Ollilla, dosentti Tom Campbell

Hakusanat: EEG, akustinen prosessointi, kuulolaite, aivokäyttöliittymä, cocktailkutsuilmiö, neurotiede, äänen prosessointi, kuulokoje, melu.

Tämänhetkinen kuulolaiteteknologia on suorituskyvyltään heikko meluisessa ympäristössä kuunneltaessa yksittäistä puhujaa useamman puhujan puhuessa samaan aikaan. Työn tavoite on todistaa aivosähkökäyrien käytön mahdollisuus akustisen prosessoinnin ohjaamisessa edellä mainitussa olosuhteissa. Tätä varten raportissa kehitettiin malli aivosähkökäyrien käyttämiseen akustisen prosessoinnin ohjaamisessa. Työssä luotu malli luotiin konstruktiivisen tutkimusmenetelmän mukaan, kirjallisuuskatsaukseen perustuen. Akustisen prosessoinnin metodeina käytettiin: keilanmuodostus tilasuodatusta, mallipohjaista häviöllistä pakkausta ja puheen äänteen alun vahvistusmetodia.

Aiempi tutkimus osoittaa, että aivosähkökäyrät sisältävät tiedon siitä, kuunteleeko puhuja aktiivisesti, mitä kuuntelija kuuntelee ja missä suunnassa kuunneltu äänilähde on puhujaan nähden. Luotu malli käyttää tätä tietoa mahdollistamaan käytettyjen akustisen prosessoinnin metodeiden käytön meluisessa ympäristössä kuunneltaessa yksittäistä puhujaa useamman puhujan puhuessa samaa aikaa. Analyysin perusteella mallin havaittiin tarvitsevan tosiaikaista prosessointia ja perivän käytettyjen akustisen prosessoinnin metodeiden laskennallisen raskauden.

Luotu malli osoittaa teoriatasolla, että aivosähkökäyrien käyttäminen akustisen prosessoinnin ohjaamisessa on mahdollista. Jotta aivosähkökäyriä voitaisiin hyödyntää kuulolaiteissa, on jatkotutkimus tarpeen. Tämän raportin lopussa esitellään esille tulleet mahdolliset tulevaisuuden tutkimus aiheet.

ACKNOWLEDGMENTS

I thank my supervisors Erja Mustonen-Ollilla and Tom Campbell for trusting me and giving me this possibility. Without their support, I could not have finished this task. I thank my roommate Onesfole Kurama for being a good role model for academic studying and for sharing his experiences of academic writing to help me.

TABLE OF CONTENTS

LIST OF FIGURES

GLOSSARY

Acoustics	A science that studies the behaviour of mechanical waves in all forms of matter.
Auditory cortex	A cortical brain region that processes sound.
Beamforming	A spatial filtering technique that operates on sensor arrays to achieve the directional transmission or reception of a signal.
Binaural	Requiring hearing with two ears.
Cocktail party effect	A phenomenon that refers to the remarkable human ability to selectively attend and recognize one source of auditory input in a noisy environment.
Cochlea	An organ that transduces physical pressure waves into electric impulses that the brain perceives as sound.
Constructive research	Problem solving through the construction of models, diagrams, plans, and organizations.
Disabling hearing loss	Hearing loss greater than 40 dB (HL) in the better hearing ear in adults and a hearing loss greater than 30 dB (HL) in the better hearing ear in children.
Electroencephalography (EEG)	A measurement method to record electrical activity that the brain generates.
Envelope enhancement strategy	A sound processing technique that enhances the onsets of intensity increases in the sound envelope.
Hearing-assistive device	A device that processes sound to render that acoustical signal more accessible to the user.
Model-based sparse coding shrinkage (SCS)	A noise-filtering technique based on the neural information theory of sparse coding.
NIHL	Industrial, military, or recreational noise-induced hearing loss.
Noise	Interference with the signal that causes errors upon the reading of that signal.

1 INTRODUCTION

It is estimated that, in the year 2008, 10.7 % of the world's population suffered moderate hearing loss of 35-49 dB(HL) and 2.1% suffered moderately severe hearing loss of 50-64 dB(HL) [1]. Approximately 5 % of the world's hearing loss is industrial, military, or recreational noise-induced hearing loss (NIHL) [2]. Listening to portable music players at high volume causes the vast majority of the youth's NIHL type hearing loss in the United States of America [3]. Estimates are that 12.7 % (30 million people) had bilateral hearing loss (2001-2008) and that 0.12 % of infants in the United States have hearing loss [4]. For persons aged 12-19 years, hearing loss prevalence is 0.31 %, increasing to 6.5 % for the 40-49 year old age group, with the prevalence of hearing loss rising to 79.1 % in 80-year olds [5]. Because more than 350 million people worldwide suffer hearing loss [6], methods to recover hearing or correct for hearing impairment are necessary. Medications and treatments exist to reduce hearing loss caused by acoustic overstimulation within a short period of time – shorter than the time needed for cell death, apoptosis of cochlear hair cells, to develop and complete. These methods require application before permanent apoptosis of cochlea hair cells occurs [2], [7]. The effective treatment time period depends on the method used. For low-level laser therapy [8] and all-trans retinoic acid, the effective treatment time is within three days after exposure [7]. After apoptosis is complete, the damage to hearing is permanent. For this reason, NIHL caused by weeks or years of acoustic overstimulation, is permanent.

Having introduced NIHL and the worldwide scale of the problem of hearing loss, addressing this problem motivates the topic of this thesis. This topic resides at the confluence of auditory cognitive neuroscience [9] with research into brain-computer interfaces (BCIs) [10], [11]. BCIs replace the traditional human-computer interaction channels requiring muscular movement (i.e., mouse, keyboard, or voice recognition) with an alternative pathway between the brain and external devices: This pathway relies on the user's brain signals within the electroencephalogram (EEG) measurable via electrodes placed at several points on the scalp [12]–[14]. Beyond BCI research using desktop computation and ordinary EEG equipment [15], recent BCI research has begun to employ mobile EEG equipment connected to mobile devices [14], [15]. This research relies either on simulations of a smartphone on a mobile tablet device [14] or on implementations within an off-the-shelf smartphone [16]. Electrodes attached to a smartphone can record the influence of neurocognitive decisions about stimulation events on EEG, so as to classify those events and control external devices [16]. The brain's selective attention to auditory stimulation events influences EEG [9], [15], [17]. This thesis thus proposes a new model to guide future research strategy. It is envisaged that this strategy determines how BCI can utilize this attentional influence on EEG to improve the acoustical processing of auditory stimulation events by hearing-assistive devices.

The purpose of that strategy is to drive the development of mobile BCIs that improve the speech-in-noise perception of the hearing-assistive device user. In the cocktail party problem [18], adults younger than 40 years of age, with audiometrically intact hearing, can decipher the conversation of an attended talker amongst a background of multiple talkers – such young adults can perceive speech in noise well [19]. With a damaged auditory periphery, even with a hearing-assistive device [20], or even in elder adults without substantial audiometric hearing loss, this speech perception in noise becomes more difficult [19]: The cocktail party problem becomes more challenging with age, hearing impairment, or both. BCI devices coupled with hearing-assistive devices could not only improve speech-in-noise perception, but also improve the comprehension of speech in that noise [21]. The overarching concept of this thesis is to tackle the cocktail party problem at the source. In controlling the acoustical processing of sound, the BCI removes the unattended acoustical noise input according to EEG information about “what” the user’s brain is attending and “where” in auditory space the user’s brain is listening. The BCI thus bolsters the force of the brain’s pre-existing selective attention to sound by controlling the hearing-assistive device’s acoustical processing prior to neuronal transduction of that sound.

To foreshadow the basis of the approach, what speech sound the user’s brain attends to influences the response of the auditory cortex to speech in multi-talker settings during the cocktail party problem [17]. Also, where the user’s brain attends influences scalp measured EEG, the posterior alpha lateralisation (8-12 Hz) of neuronally generated oscillations, as recorded over the brain’s parietal lobes [17]. The results of auditory cognitive neuroscience experiments thus attest to the feasibility of a new type of BCI. This BCI would harness the EEG input that indexes what and where the user’s brain attends in auditory space. That EEG input is envisaged to influence the output of hearing-assistive devices for neuronal transduction. The output from the mobile BCI via the hearing-assistive device, would tackle the cocktail party problem at the source. That is, this BCI, acquiring and processing EEG on a mobile device, would control the hearing-assistive device’s acoustical processing. This control is envisaged to remove unattended acoustical noise from the hearing-assistive device’s output signal corresponding to unattended talkers. Alternatively, or additionally, this control would enhance the hearing-assistive device’s output signal corresponding to the attended talker. The BCI thus augments auditory selective attention at the input to the human auditory system from the hearing-assistive device.

1.1 Goals and limitations of the study

The goal of this bachelor’s thesis is to construct a model to use information from electroencephalography signals to control acoustical processing to filter out unattended noise information from the amplified sound. The longer-term goal of this method is to increase performance of hearing-assistive devices in noisy and or multi-talker environments. Not just signal processing aids, but also both bone-anchored hearing aids and cochlear implants can potentially benefit from this technology, because they both have sound processing units, in the same manner as conventional hearing-assistive devices do. The aim of model is to improve hearing-assistive devices’ performance in noisy environments and multi-talker situations. Because of this focus of increasing speech intelligibility in multi-talker situations, the model’s performance need not be high under other acoustical conditions. While there may be other uses of the model, for example in audio-processing, these uses are beyond the scope of this thesis. Neither a prototype nor patient testing are reported.

1.2 Concepts and terminology

In this section, the terminology and concepts used in this thesis are introduced, beginning with hearing-assistive device technology, then turning to BCI and electroencephalography technology. This section then focuses on the principles of beamforming methods we propose to use in conjunction with these technologies to solve what we define as the cocktail party problem. Who would be the users of this proposed combination of technologies, who would benefit the most given a disabling loss? Although in 2008, as already noted, 12.8 % of the world’s population suffered moderate-to-severe hearing loss [1], not all had a disabling loss. Rather, to arrive at a definition of a disabling hearing loss, in 2012, an estimate of the world population was 7,003,554,291 people [22], the World Health Organisation estimating that 359.9 million people of those persons had a disabling hearing loss. Disabling hearing loss means 40 dB(HL) hearing loss in the better hearing ear and 30 dB(HL) hearing loss in the better hearing ear in children [23]. According to these 2012 estimates, approximately 5.13% of world population thus suffer disabling hearing losses.

When there is permanent hearing impairment due to such a disabling loss, hearing aids and cochlear implants are the primary approaches to address the impairment. Hearing aids focus on minimizing the effects caused by impairment. A hearing aid is an electrical device that amplifies and processes sound, with the purpose of reducing inconvenience caused by hearing impairment. Normally, this amplification corrects for impaired hearing as measured by audiometry, and processing the sound signal renders speech more intelligible. This sound processing differentiates extant signal processing hearing aids from a normal sound amplifier [24 , Ch. 1–2] and the former generation of analogue hearing aids. Some specialised hearing aids exist that bypass the damaged facets of the peripheral auditory system. Such technologies include bone-anchored hearing aids [25], which transfer sound via vibrations through the skull directly stimulating the cochlea with vibrations such that, in turn, neuronal transduction occurs. Bone-anchored hearing aids are thus a subclass of hearing aids [25]–[27].

Cochlear implants (CI) are electronic devices that function to enable hearing for people with severe or profound hearing impairment in both ears. CIs have two main parts: an internal implant, and an external sound processor and transmitter. CIs bypass most of the peripheral auditory system by placing an electrode array directly in the cochlea. The external part of the CI commonly consists of a microphone, sound processor, and electromagnetic induction transmitter. The internal implant part consists of a receiver, a circuit to convert the signal to electric impulses, and an electrode array within the cochlea [28]. CIs creates better hearing in persons with severe or profound hearing impairment than hearing aids do, especially in noisy environments [29], because sound amplification is of little or no benefit with an inoperative auditory periphery. In other studies, speech comprehension with CIs had a mean of 50.7 % after 3 years of use, for a test of monosyllable perception in a noisy environment that has a signal-to-noise ratio (SNR) of +15 dB [30].

Having defined hearing-assistive devices, brain-computer interfaces (BCI), also known as brain–machine interfaces (BMI), direct neural interfaces (DNI), or mind-machine interfaces (MMI), measure brain activity by invasive or non-invasive methods. This brain activity input to the interface enables the user’s interaction with the surrounding environment [31]. BCI works as any other control system and thus has input, output, and a system that maps brain activity to output (e.g., from brain activity to a physical system command) [31]. Any brain measurement method can be used to create input for BCI, including invasive subdural electrodes on the surface of the cortex or intracranial electrodes inserted into the cortex, as well as non-invasive magnetoencephalography (MEG) or electroencephalography (EEG).

Electroencephalography (EEG) is a monitoring method for recording electrical brain activity [32, Ch. 1]. Building on Richard Caton’s work on dogs and apes, Hans Berger invented electroencephalography, measuring alpha waves in humans for the first time in 1924. Though Berger’s submission was initially rejected as electricity already known to be produced by muscles, a confirmatory experiment measuring EEG directly from the exposed brain of his neighbour’s dog qualified his results for publication in 1929 [33]. EEG measures fluctuations in voltage, which originate from ionic currents inside the neurons in the brain, when specifically pyramidal neurons with aligned axons or dendrites fire synchronously. When a signal is produced electrochemically, there is a difference in the charge across the short distance of each neuron’s membrane. Positive polarity indicates current leaving and negative polarity returning current. Firing of multiple aligned neurons create a primary current, which propagates a secondary or "volume" current that is induced in the brain and liquor thus reaching the skull and scalp. Volume currents cause potential differences on the scalp, which electrodes can measure. Electroencephalography works by placing electrodes on the scalp and then measuring their electrical potential difference. These voltages are amplified and then stored in a recording medium. Normally EEG is non-invasive. That is, electrodes are normally placed on the scalp, but for specific purposes invasive electrodes can be used [31], [34]. Other methods to study brains exist such as: functional magnetic resonance imaging, electrocorticography, nuclear magnetic resonance spectroscopy, single-photon emission computed tomography, positron emission tomography, magnetoencephalography, near-infrared spectroscopy, and the event-related optical signal. Compared to these methods EEG has several advantages for BCI. Mobile low-cost EEG systems are becoming increasingly more common. These systems derive EEG for BCI of a comparable quality to what desktop systems are producing [15]. EEG also offers a higher temporal resolution than the less mobile positron emission tomography and magnetic resonance methods. Recently, it has been shown that multichannel EEG can be acquired using an off-the-shelf smartphone [16]. The time is thus ripe for the widespread valorisation of the concept of mobile EEG-based BCI technology as low-cost enhancements to pre-existing smartphone technology. Such mobile BCI technology could interact with hearing-assistive devices.

Having introduced EEG and the concept of using a mobile BCI, it should be considered that EEG signals contain information of neural activity from multiple brain regions. This EEG data needs to be processed by extracting features. There are different methods for isolating useful features from EEG signal data. Approaches include event-related potentials (ERPs), template-matching, and oscillatory power analysis [17].

Beamforming is a spatial filtering technique that allows the transmission of a signal to, or reception of a signal from, a specific spatial location. Beamforming is used in conjunction of array of receivers or transmitters. This array is needed to enable the use of beamforming, because multiple measurement points are needed for calculation of a signal source. For transmission, an array of transmitters is needed to be able to send the signal to specific spatial location [35, 36]. Transmission uses principles of wave physics to calculate in what spatial locations waves get amplified by one another and in what location those waves negate one another. Using these principles to calculate the timings for transmissions for each element of array, beamforming amplifies the transmitted signal at the receiving spatial location. Beamforming is used when normal temporal filtering such as a finite impulse response filter cannot be used, because signal and interference signal occupy the same frequency bands and are transmitted simultaneously. Signal and interfering noise require transmission from measurably different locations for beamforming to work [35], [36]. Beamforming uses the time difference of the signal's arrival at different nodes of the array to calculate the direction of signal. This time difference can be calculated when the speed of the travelling wave signal is known beforehand, as is the case with sound waves. The distance of the signal source is calculated from distortion and changes of the signal. This distance calculation cannot be applied for cases where distortion or changes in the signal are not measurable or masked by interference signals. Distortion and changes in signals include: deformation of a signal by the transporting medium, amplitude loss, ambient background noise mixing in with the signal, and phase shift. Many fields utilize beamforming including radar, sonar, seismology, communications and astronomy [35], [37, p. 13]. Acoustical beamforming is of particular relevance to the present thesis.

Turning from the acoustical processing technique of beamforming to a germane psychological phenomenon, the cocktail party effect, E. Colin Cherry first described this effect in 1953 [18]. The cocktail party effect is the ability to focus auditory attention to chosen auditory stimuli while filtering out the rest [38]–[40]. The name comes from experience of partygoer in a noisy room: The partygoer can focus on a single conversation in a room filled with music and conversations of other people [18].

This phenomenon could be related to other similar phenomenon, in which a person is able to immediately detect sounds of importance from unattended stimuli. An example of this phenomenon, termed the breakthrough of the unattended, is recognising one's name in a noisy environment when not expecting to hear that name [41], [42]. At first, this phenomenon might be martialled in favour of a concept that the brain processes the entirety of the multi-talker auditory environment. As such, the name only captures attention when that otherwise unattended information is detected as personally relevant by a late selection process. However, it is worth considering that only one third of people do recognise their name [41], [42]. A perspective is that thus two thirds of people indeed do filter out unattended information and arguably do not process that information from a very early stage with an early filter [9]. The remaining third, who do recognise their unexpected name in an otherwise unattended conversation, have poorer top-down control of the concurrent storage and processing of information in working memory [41]: Those persons have poorer cognitive abilities on a working memory task. According to the new early filter model [9], such individuals with low working memory capacity arguably make use of contextual information for top-down control of the auditory system during speech-in-noise perception, in which case the early filter is wide open. With that early filter wide open, one's own name enters the higher auditory cortex and, in turn, the person's brain recognises that name as personally relevant.

The cocktail party effect benefits from being binaural – i.e., stimuli separation works best with stereo sounds, which naturally requires hearing from both ears [38], [43]. Binaural information in audio helps with localization of a sound source, which in turn increases the brain's performance in filtering the unattended sound stimulus [40], [43]. People with more hearing loss in one ear thus perform worse in noisy multi-talker situations [43].

1.3 Structure

This thesis consists of five chapters, including the present introduction chapter. The literature review (chapter 2) then identifies the current limitations of hearing aid technology, followed by the neuroscientific theoretical background for the model constructed in this thesis. There then follows the research problems and methods chapter 3, which, identifying the research questions of the thesis, leads into the constructed model of chapter 4. This acoustical processing model in chapter 4 offers answers to these research questions by constructing and analysing this model, which proposes to use data extracted from electroencephalography signals. The final conclusions and implications chapter points out areas of interest for future development (chapter 5).

1.4 Current limitations of hearing aid technology

Hearing aids and cochlear implants are currently the primary methods to improve hearing for persons with hearing impairment. Current hearing aid technology suffers from poor performance in noisy environments with multiple talkers [38], [44]. This problem exists because of the physical limitations of current technologies [45] and because the hearing of the person who uses a hearing aid is impaired [38]. There are many potential changes in hearing that hearing loss can cause. These changes include: reduced release from noise masking, threshold shift, and reduced temporal resolution [36]. These problems are most commonly caused by regions of dead cochlea. Hearing aids and hearing-assistive devices, in general, exist to alleviate the problems caused by those dead regions. Simple amplification of sound can alleviate a uniform threshold shift, a loss of sensitivity to sound waves uniformly throughout the frequency range. However, if threshold shift is non-uniform, sound processing is needed. Either shifting heard frequencies or non-uniform amplification can alleviate the problems of a non-uniform threshold shift [24, Ch. 1]. Reduced release from noise masking, and reduced temporal resolution, present challenges for hearing aid technology, because no solution exists to alleviate these problems.

The temporal resolution of the auditory system decreases with elevations of gap-detection threshold [46]. This gap-detection threshold is the duration of silence in a sound required for the listener to detect that silence as a gap upon a certain proportion of trials. Gap-detection ability enables the listener to segment individual speech sounds for lexical recognition as well as the subsequent semantic and syntactic processing. Gap-detection is typically poorer in people with hearing loss [47]. Gap-detection is also germane to auditory masking. Auditory masking is a phenomenon whereby sound is rendered inaudible by the presence of another "masking noise" sound. Noisy environments cause such auditory masking. Forward masking is a type of masking whereby a preceding masking noise sound impedes the perception of a subsequent target sound for tens of milliseconds after the offset of the masking noise. Gap-detection abilities are related to the release from forward masking. That is, with higher gap-detection thresholds, the time required to recover from forward masking is longer [44]. Hearing loss itself also increases forward masking [48]. A modulation of masking noise produces stronger elevations of speech reception thresholds – poorer identification of speech in that noise – in persons with hearing impairment than in normal hearing controls [49]. Hearing loss can also affect comprehension in many other ways such as reducing the ability to understand fast speech or speech under reverberatory conditions [50].

Though there a previously demonstrated benefits of the amplifying speech in low level noise through the audiometrically fitted 1980s generation of hearing aids for persons with hearing impairment in terms of lower (better) speech reception thresholds, with levels of noise exceeding 60 dB(A), there was no further benefit of the hearing assistive device [44]. Because reduced temporal resolution hampers the ability to perceive and comprehend sound in a noisy environment, the main approach of signal processing in extant hearing-assistive devices has been implementing noise-reducing algorithms within hearing aids and cochlear implants. Reduced release from noise masking is caused by reduced temporal resolution, reduced sensitivity to binaural cues, and other factors. Improvements in binaural hearing are attainable by means of sound processing via artificially strengthening spatial cues, or with physical design and placement of the hearing-assistive device. Hearing-assistive devices benefit from binaural hearing, but often hearing aids deteriorate binaural hearing by placing microphones far from the eardrums. The head acts as a physical spatial filter, as is termed the head shadow effect [38]. This effect is weaker the further the hearing aid is from the eardrum, because the skull and soft tissues then filter the sounds less. These issues are common because the ear canal is small thus constraining the physical dimensions of hearing aid design. The performance of systems employing directed microphone arrays benefit from noise reduction [51, 52]. Using directed microphones or microphone arrays, the binaural character of hearing can also be recreated [52].

1.5 Beamforming, model-based sparse coding shrinkage, and envelope enhancement strategy

Introducing sound processing to a hearing-assistive device can, at least in part, alleviate problems with noise masking and impairments temporal resolution. There are many algorithms for this purpose. The three methods of choice for this thesis are: model-based sparse coding shrinkage, beamforming, and an envelope enhancement strategy. After an introduction to the notion of filtering, in this section, there is consideration of each of these three methods in turn.

Filtering is a technique whereby processing of a waveform attenuates or eliminates unwanted noise. Filtering can be digital, electronic, or mechanical. A wide array of sciences and practical solutions employ filtering. For this reason, there are multiple filter designs for different purposes and requirements [53, pp. 11–12], [54, Ch. 1]. Noise filtering is a set of filtering techniques specialised for removing noise from a signal. The ideal noise filter would remove all unwanted noise without altering the wanted signal. For most applications, creating an ideal noise filter is difficult if not impossible, because no means exists for distinguishing all types of noise from the wanted signal, necessitating compromise. For this thesis, a model-based sparse coding shrinkage (SCS) noise filtering algorithm was chosen because, considering the effects of noise filtering on improving speech intelligibility for persons with hearing impairment, the SCS algorithm outperforms a Wiener filter algorithm [55]. While the Wiener filter algorithm improves intelligibility of speech in noisy environments for hearing impaired users [56], for persons with normal hearing, none of eight classical algorithms, of which the Wiener filter was one, significantly improved speech intelligibility [57]. Distortion to the speech signal caused by removal of noise could render such filters ineffective. Such findings underscore the view that speech quality and intelligibility do not correlate [57], [58]. Modifying algorithms to avoid distortion to the speech signal, can improve speech intelligibility [58]. The aim of developing SCS alongside other algorithms was to avoid this distortion of speech [55], [59]. The SCS algorithm is a derivation of the sparse coding model of neural coding. Sparse coding represents information in the strong activation of small set of neurons as a compromise between dense and local coding. Sparse coding algorithms are algorithms that produce sparse coding presentation of input to a neural network. Neural network and neural computation studies [60], [61, Ch. 1] have employed neural algorithms including sparse coding algorithms. Such model-based sparse coding shrinkage outperforms a conventional sparse coding shrinkage algorithm [62]. The acoustical processing model of this thesis thus employs model-based sparse coding shrinkage.

Spatial filtering is a technique that selects signals according to spatial origins. Beamforming is both a form of spatial filtering and amplification commonly used for radar, sonar, and other types of transmitter or receiver arrays [35], [36]. Acoustical beamforming based techniques are commonly used in hearing-assistive devices to reduce noise [52]. The acoustical processing model of the thesis uses acoustical beamforming because application of this technique has improved the listeners' understanding of the attended speaker's message in multi-talker situations. The best results occurred when mixing the original signal with the signal output from beamforming [48]. Beamforming algorithms used in hearing-assistive devices include fixed-direction beamformers and adaptive-direction beamformers. Adaptive beamformers adapt to changes of environment. A common adaptive parameter is changing filtering direction according to movements of sound sources. A beamformer is a system that uses a beamforming algorithm. When multiple sound sources are present, one with higher amplitude is commonly selected and this kind of adaptive-direction beamformer removes noise better than a fixed-direction beamformer [63]. The problem with selecting the loudest sound source is that the loudest sound source is not always attended sound source. The loudest sound source can be unwanted in a situation such as cocktail party. Other approaches to adaptive-direction beamformers include algorithmic beamformers. A probability-based adaptive-direction beamformer is also more efficient in terms of the intelligibility and quality of speech compared to a fixed-direction beamformer [64]. Nevertheless, current algorithms make errors on what sound source the user wants to listen to. This thesis focuses on solving this problem by introducing EEG-control of the direction parameter of the beamformer.

Onset enhancement of the speech signal, also known as an envelope enhancement strategy, is an audio processing method whereby the start of each syllable is amplified [65]. This onset enhancement procedure extracts the speech envelope, amplifies the separated signal, and finally combines the amplified signal with the original signal. This envelope enhancement strategy increases the intelligibility of speech in a multi-talker setting [65]. In cochlear implants, an envelope enhancement strategy increases speech intelligibility in reverberating space by improving the localization of sounds [66], [67]. These reasons motivate the inclusion of this envelope enhancement strategy in the acoustical processing model of this thesis.

1.6 EEG signals

Using high-density electroencephalography and a template-matching analysis method, Kerlin and colleagues [17] revealed an attentional gain of dipolar cortical sources, which is time-locked response to speech sounds. The EEG spatial filters that pick up these dipolar sources were derived from the distribution of the auditory N1 [68] to the first word of the sentence [17]. This auditory N1 exhibits distinct spatiotemporal and functional properties to the mismatch negativity component of the auditory ERP [68], though there is also evidence for a theory that attention also increases the generation of that mismatch negativity [69]. The spatial filter for the auditory N1 dipole rather gleans something of the scalp distribution indexing the auditory cortex. The attentional gain was strongest in auditory cortex in the frequency range of 4-8 Hz, which is the speech envelope frequency [17]. In this way, this frequency range reveals “what” is attended. Also, there was a parietally distributed difference in alpha power (8-12 Hz) across hemispheres contralateral to the attended ear. This difference in alpha power at parietal sites indicates the direction of the attended speech source [17] – “where” is attended. Both bone-anchored hearing aids and cochlear implants can potentially benefit from this technology for processing EEG signals in a cocktail party setting, because they both have a sound processing unit, in the same manner as conventional hearing aid does [27]. This thesis addresses these technological limitations by constructing a model to control acoustical processing with electroencephalography signals. Envisaged is the integration of this model into the sound processing of hearing aids or cochlear implants. Principles of this model can also be adapted to control other facets of sound processing, as implemented by the sound processing circuitry of hearing aids and cochlear implants [27].

2 RESEARCH PROBLEMS AND METHODS

2.1 Research problems

There are no known methods or technologies proposed for using electroencephalography signals to control acoustical processing. Acoustical beamforming can produce location data of sound sources [64], but this approach does not produce information of what sound source the user is attending and if the user is attending anything currently. In this context, acoustical processing means a set of audio processing methods, the purpose of which is to increase intelligibility of speech in a noisy environment with multiple talkers. The methods of choice for this purpose are spatial filtering, noise filtering, and an envelope enhancement strategy. All of these three methods need different types of information in the form of parameters. The system can receive envelope enhancement parameters as prior input before usage. The SCS algorithm can either work as a self-learning system or the SCS algorithm's parameters similarly inputted prior to usage. Spatial filtering, however, needs all the time information of the valid location of the target audio source, to prevent filtering that source out, as is a reason for use of electroencephalography signals. Parameters of SCS and envelope enhancement can be changed by information from electroencephalography signals in order to make those parameters dynamic. There are no known previous approaches to using EEG to control acoustical processing. The influence of using EEG to control the chosen approaches on the performance of the system also remains to be determined. Further, combination of the three chosen methods for acoustical processing – spatial filtering, noise filtering, and envelope enhancement – is without known precedent. Thus, I postulate the hypothesis that EEG-control of a combination of these techniques improves speech intelligibility of a noisy audio signal of multiple audio sources. I also postulate that controlling model-based SCS and envelope enhancement parameters dynamically, according to EEG input, can improve performance. The research question of thesis is thus how to use information from electroencephalography signals to control acoustical processing.

2.2 Research method

This thesis uses constructive research and literature review as methods of research. Constructive research, also known as design science [70, p. v], is a research method that creates a construct to solve a problem [71, p. 3]. The goal, the environment, and characteristics of a construct are thus definitive [72, p. 5]. A construct can be both physical and non-physical. Examples of non-physical constructs are mathematical formulas, computer programs, and ideological models. A problem in the context of constructive research refers to the difference between a current situation and the goal situation that is in some way better. This difference is the problem. The construct's function is to enable the removal of this difference. Thus a problem can refer not only to non-optimal conditions, which have a problem that is being solved, but also to conditions that people see as optimal even though further optimisation of these conditions is possible [70, Ch. 1]. An example of this latter case is medical care before and after the invention of X-ray imaging. Before X-ray imaging, doctors thought their ways were optimal in current conditions, but the invention of X-ray imaging enabled doctors to seen inside of the human body in non-invasive ways, thus improving the accuracy of diagnosis and, in turn, doctors' selection of the most appropriate treatments for a condition. In this case, X-ray imaging is a created construct. Real-world problems are often termed wicked problems in the sense that such problems are difficult if not impossible to solve as can be due to: a lack of knowledge, changing and contradicting requirements, or a complicated relationship with related problems. Real-world problems often have no clear end criteria that indicate the resolution of the problem. A lack of resources rather than the solution to the problem is often the reason for putting a stop to the solving of a wicked problem [70, Ch. 1].

This thesis applies constructive research as a method because the problem has attributes suitable for resolution with constructive research. There are the current non-optimal conditions, i.e., hearing aids perform poorly under cocktail party conditions, and the goal condition – hearing aids function ideally under cocktail party conditions. This difference between the current and goal states creates a suitable problem for resolution with a construct. This real-world problem itself shares characteristics with wicked problems. Modern science does not offer an uncontentious account that relates molecular biological systems to all the phenomenal experiences of hearing. As such, it is not completely understood how brain functioning relates to problem. In addition, the problem has contradicting requirements because the end user defines "perfect" functioning of a hearing aid, as is accordingly subjective. For this reason, this thesis addressed a smaller subset of this problem.

3 ACOUSTICAL PROCESSING MODEL

3.1 Constructed model

Using high-density electroencephalography and a template-matching analysis method revealed an attentional gain of a time-locked response to speech sounds under “cocktail party” conditions in the auditory cortex in the frequency range of 4-8 Hz. Kerlin and colleagues also revealed a parietally distributed inter-hemispheric difference in alpha power (8-12 Hz) across hemispheres with an amplitude suppression contralateral to the attended ear [17]. The attentional gain at the frequency range of 4-8 Hz defines “what” is attended and the parietally distributed difference in alpha power (8-12 Hz) indicated the location of the attended sound source. The attentional gain in the auditory cortex in the frequency range of 4-8 Hz is present when the person is attending a particular audio source, as can also be used to determine whether that person is actively listening [17]. Use of this combined “what” and “where” information can control a beamforming method of spatial filtering, SCS noise filtering, and envelope enhancement during sound processing.

There is a division of the model’s function into “idle” and “active” states. The model is in the idle state when the user is not attending any sounds. When the model is in this idle state, spatial filtering and envelope enhancement processing are turned off, allowing all spatial sound information to reach the user. When the user starts attending the sound source, the model switches to the active state. In the active state, the model enables spatial filtering and envelope enhancement processing. When the model switches to an active state, the direction of the attended sound source is calculated. The model then passes the direction of the attended sound source to an acoustical spatial filter. This filtering causes a decrease in amplitude of the signals that come from a different direction from the attended sound source. When the user stops attending to the sound source, the model switches back to an idle state, again turning off spatial filtering and envelope enhancement processing.

It is possible to write this model (Figure 1 below) in pseudo code [73, Ch.10]. The benefit of writing the model into pseudo code is to permit the analysis of the model with the same techniques suitable for analysing code in a normal programming language. This pseudo code implementation (overleaf) adheres to a similar syntax to the programming language Java and follows the object-oriented programming paradigm.

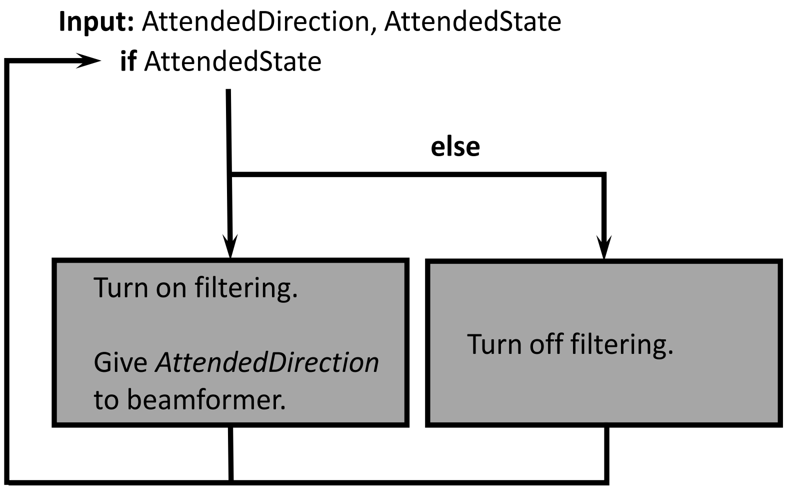

Figure 1: Block diagram of the processing model.

Pseudo code:

```
Algorithm SimpleControlAcousticProcessingWithEEG{
        INPUT: AttendedDirection, AttendedState

        // "//"=means comment, everything after "//" is ignored

        While(device online){
                if(AttendedState == True){
                        activate filtering and update parameters of filters

                }else{ // not attending any sound source
                        deactivate filtering
                }

                OldState= AttendedState
                OldDirection= AttendedDirection

                do{
                        Wait()
                }while(device online
                        AND (NOT(OldState == AttendedState) NOR
                                NOT(AttendedState AND
                                        NOT(OldDirection == AttendedDirection)
                                        )
                                )
                        )
        } // end of while
} //end of algorithm
```

When there is a change in attentional state, or – when sound is attended – a change in the attended direction, this overall simplified program reveals the control in the model. Whilst online, when the user is starting to attend or attend in new direction, the model activates filtering updating filtering parameters. Alternatively, when no particular sound becomes attended by the user, the model deactivates filtering. This program receives the input parameters: *AttendedDirection* and *AttendedState*. The *AttendedDirection* parameter contains the direction of the attended audio source. That direction is the difference in angle between the direction of the sound source and the user of the system. *AttendedState* tells if the person is attending to any sound source or not. The parameters *AttendedDirection* and *AttendedState* originate in EEG signals. A set of algorithms operating on raw EEG input data continuously updates these input parameters for the master program. The purpose of the **while** loop is to keep processing in operation while the device is online. The **if/else** statement determines if the model should be in an idle or an active state. Within the **while** loop, a nested **do-while** loop then serves the purpose of stopping the model from updating sound processors, when there is nothing to update: filters continue to operate, but the model pauses when *AttendedDirection* and *AttendedState* do not change. This pausing occurs by saving the value of *AttendedDirection* and *AttendedState* before the **do-while** loop, then comparing the old to the current value of those parameters. If *AttendedState* changes or if *AttendedState* is "True" and the *AttendedDirection* changes from the old value, then processing continues updating. Otherwise, the algorithm waits, freeing-up computing cycles, thus releasing processing power for other crucial processes including acquiring and processing EEG.

In fuller pseudo code with comments and methods for sound processing in place:

```
Algorithm ControlAcousticProcessingWithEEG{
        INPUT: AttendedDirection, AttendedState

        // "//"=means comment, everything after "//" is ignored

        //AttendedDirection tells the difference in degrees of the current attended sound source from
        //the user. Straight ahead being 0 degree difference, right produces positive degrees and left
        //negative. Thus the sound source at left side of the user is -90 and at right side, 90 straight
        //behind is either -180 or 180. The EEG data processing algorithm modifies AttendedDirection to
        //reflect current data from EEG signals.

        //AttendedState has two possible states, true or false. AttendedState is true when the user
        //attends a sound source and false when user does not attend a sound source. The EEG data
        //processing algorithm modifies AttendedState to reflect current data from EEG signals.

        //New beamforming spatial filter "object"
        //This object controls behaviour of the beamforming spatial filter.
        Object Beamformer = new Beamformer()

        //This object controls behaviour of envelope enhancement processing.
        Object EEProcessor = new EnhancedEnvelopeProcessor()

        //This object controls behaviour of model-based sparse coding shrinkage noise filtering.
        Object SCSFilter = new SCSNoiseFilter()

        //Turn off all filtering, before start-up.
        Beamformer.TurnOffFiltering()
        EEProcessor.TurnOffFiltering()
        SCSFilter.TurnOffFiltering()

        //Compute while the device is online.
        while(device online){

                if(AttendedState == True){

                        //If the beamformer is not filtering already, turn on filtering.
                        if(Beamformer.GetFilteringState() == not filtering){
                                Beamformer.TurnOnFiltering()
                                }
                        //Same for the envelope enhancement filtering.
                        if(EEProcessor.GetFilteringState() == not filtering){
                                EEProcessor.TurnOnFiltering()
                        }
                        //And for SCS noise filtering.
                        if(SCSFilter.GetFilteringState() == not filtering){
                                SCSFilter.TurnOnFiltering()
                        }

                        //Update beamformer's sound source direction.
                        Beamformer.SetSoundSourceDirection(AttendedDirection)
```

```
        }else{ // not attending any sound source

            //if the beamformer is filtering sound, turn beamforming off
            if(Beamformer.GetFilteringState() == filtering){
                Beamformer.TurnOffFiltering()
            }
            //and same for envelope enhancement filtering
            if(EEProcessor.GetFilteringState() == filtering){
                EEProcessor.TurnOffFiltering()
            }
            //and for SCS noise filtering
            if(SCSFilter.GetFilteringState() == filtering){
                SCSFilter.TurnOffFiltering()
            }
        }

        // Save current state of parameters.
        OldState = AttendedState
        OldDirection = AttendedDirection

        // NOT means logical negation, == means logical equal sign
        // NOR and AND mean logical operators of same name
        do{
            Wait() //Save computing cycles by skipping them.
        }while(device online
            AND (NOT(OldState == AttendedState) NOR
                NOT(AttendedState AND
                    NOT(OldDirection== AttendedDirection)
                    )
                )
            )

    } // end of while loop
} //end of algorithm
```

This full algorithm starts by creating the objects of algorithms, which process sound. These objects control the behaviour of those sound processing algorithms. After creating those objects, these objects turn off all filtering before starting the processing of the model. This initial absence of filtering does not stop the person from hearing the sound and thus attending to that sound. Beamforming would misbehave under those conditions, because the filtering direction is, as yet, unknown. In **if/else** statements, the model turns processing on or off depending on whether the person attends to any sound source or not. When the person attends a sound source, the model turns processing on. Simultaneously, the current direction of the attended sound source is an input passed to the beamforming sound processing algorithm. When the person is not attending to sound, the model turns processing off. If no change to *AttendedState* happens, the algorithm waits thus saving computing resources.

3.2 Analysis of the model

Pseudo code shows that the model has low complexity, having only one **while** loop and one **if/else** condition statement. This approach makes the model completely deterministic. For this reason, we know all states in which the model can be. The model can either be in an active or an idle state. Waiting for changes is not a separate state, because that waiting occurs whilst the model is in either of two states. Further, waiting goes unnoticed outside of the algorithm. Crucially, during the active state, the model keeps acoustical spatial filtering updated with the direction of attention from EEG information, such that spatial changes update acoustical spatial filtering. Such spatial changes include changes in locations of the listener or speaker. Failing to do so would cause spatial filtering to filter out the attended sound source.

Computationally, the model needs near real-time processing to be useful for real-time speech comprehension. Not taking into account the computational complexity of beamforming, SCS, and the envelope enhancement algorithms, the model behaves computationally as an O(N) complex algorithm. Each change in *AttendedDirection*, or *AttendedState* parameters, requires 10 or 7 operations. The model has a complexity of 18*N (attending sound) or 17*N (not attending sound), both simplifying as O(N) [74]. **Do-while** causes the algorithm to wait and thus ends processing for that parameter change. Because SCS and beamforming are both computationally more intensive algorithms, the final computational complexity is higher.

Smooth transfer from attending to not attending states could be beneficial for the usability of the model. This switching would make the algorithm feel more pleasant and responsive. Currently the algorithm switches state to an arbitrary value, as is because the system has only two distinct states. Comparing the attentional gain of the current and non-attending state, would yield a scalar value. A dynamic system could use this scalar value to mix non-processed signal and processed signal.

As this chapter has described, the model offers the control of a combination of spatial filtering, noise filtering, and envelope enhancement by using analysed EEG inputs. The model would use EEG input to dynamically control model-based SCS and enhance envelope parameters. It is foreseeable that application of the model would improve the intelligibility of speech in a noisy audio signal with multiple audio sources. This model thus proposes a solution to this thesis's research question as to how to use electroencephalography signals to control acoustical processing.

4 CONCLUSIONS, IMPLICATIONS, AND FUTURE DEVELOPMENTS

4.1 Conclusions and implications

The model constructed in this thesis uses the reflections of user's attention in the EEG input to the BCI to flip between idle and active states of that BCI. Those different states of the BCI, in turn, control beamforming, model-based sparse coding shrinkage, and envelope enhancement of the speech signal in state-specific ways. Such a model serves as a viable framework for using EEG signals to control sound processing. Analysis of the model revealed the model needs near real-time processing to be useful for real-time speech perception and comprehension. Although the model inherits the computationally intensive properties of acoustical processing, the model itself is low complexity placing few additional demands on computational resources, as conducive to real-time computation. Further, analysis of the model revealed that active state acoustical spatial filtering requires regular real-time updating by the direction of auditory attention gleaned from EEG information. Such updating is required for effective operations when the locations of the listener or speaker change.

4.2 Future development

The foregoing arguments highlight the physical make-up of a hearing aid as a priority topic for future investigation using the model developed in this thesis to control that aid's acoustical processing. Ordinary desktop laboratory EEG equipment is suitable for proof-of-concept in a laboratory. Such equipment would not be well suited to the field-testing of a prototype or future valorisation. Limitations include the weight of the EEG equipment, the size and visual appearance of the measuring system, as well as the vaporization of electrolyte solution.

Neuronal activity or other bioelectric or extraneous artefacts, which are unrelated to processing the sound, can obscure the relevant EEG input impairing the system's performance. Such electrical potentials are foreseen as challenges in developing a highly accurate system for selectively amplifying attended sounds. Effective artefact-correction routines may contribute to improved control of the system's performance. Processing information from EEG can also be highly data-intensive. However, the influence of attention in the predominant envelope frequency range of speech, concerning "what" the user is attending, affects the phase-locked representation of speech envelope in EEG (4-8 Hz) [17]. Further, the influence of the direction of attention on the parietal alpha-band (8-12 Hz) of the EEG is also low frequency [17]. The Nyquist limit for recording these frequency ranges is thus 24 Hz, a minimum sampling rate at which to acquire both "what" and "where" information. Processing multi-channel EEG information at such sampling rates, thus need not be particularly data-intensive. EEG data acquisition and processing could thus take place using an extant off-the-shelf smartphone [16].

Such a smartphone is integral to Stefan Debener's group's approach to mobile auditory BCI using small unobtrusive EEG measurement arrays around the ears [16], as addresses the aforementioned concern about the size and visual appearance of the measuring system. Such arrays do not gather EEG from the scalp regions used by Kerlin and colleagues' EEG spatial filters [17] for "what" information about auditory attention. Such arrays also do not acquire data from the parietal sites for "where" information about auditory attention. That said, the mastoids where Debener and colleagues' measurement arrays are located [16] are optimal for picking-up the N1 polarity reversal [68] reflecting well the supratemporal N1 generators. Such polarity reversals are relatively free from other N1 componentry generated outside the auditory cortex and may generate phase-locked EEG input about “what” is attended. Bilateral arrays may also reveal alpha lateralisation features. For the purposes of EEG acquisition through a mobile device to control acoustical processing using the model developed in this thesis, a baseball cap of electrodes might be more appropriate. A calibration session kindred to Kerlin and colleagues’ approach [17] could individualise EEG spatial filters as then might provide the most useful input about “what” is attended for controlling acoustical signal processing. Such a cap would also cover the parietal regions whereby alpha lateralisation reveals provides input about “where” is attended. If such a cap were to contain dry electrodes, the electrolyte vaporization in the field is not a concern. Indeed, recently developed dry electrodes are comfortable and could be worn for extended periods in such a baseball cap.

5 REFERENCES

[1] G. Stevens, S. Flaxman, E. Brunskill, M. Mascarenhas, C. D. Mathers, and M. Finucane, “Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries,” *Eur. J. Public Health*, vol. 23, no. 1, pp. 146–152, 2013.

[2] N. Oishi and J. Schacht, “Emerging treatments for noise-induced hearing loss,” *Expert Opin. Emerg. Drugs*, vol. 16, no. 2, pp. 235–245, 2011.

[3] E. Daniel, “Noise and hearing loss: a review,” *J. Sch. Health*, vol. 77, no. 5, pp. 225–231, 2007.

[4] M. Gaffney, J. Eichwald, S. D. Grosse, and C. A. Mason, “Identifying infants with hearing loss - United States, 1999-2007,” *Morb. Mortal. Wkly. Rep. 2010*, vol. 59, no. 8, pp. 220–223.

[5] F. R. Lin, J. K. Niparko, and L. Ferrucci, “Hearing loss prevalence in the United States,” *Arch. Intern. Med.*, vol. 171, no. 20, pp. 1851–1852, 2011.

[6] C. Mathers, A. Smith, and M. Concha, “Global burden of hearing loss in the year 2000,” *Glob. Burd. Dis.*, vol. 18, pp. 1–30, 2000.

[7] H. J. Shim, H. H. Kang, J. H. Ahn, and J. W. Chung, “Inner ear retinoic acid applied after noise exposure can recover the noise-induced hearing loss in mice,” *Acta Otolaryngol.*, vol. 129, no. 3, pp. 233–238, 2009.

[8] S.-Y. Chang, M.-W. Suh, C. W. Bahk, J. Y. Jung, J.-C. Ahn, P.-S. Chung, *et al.*, “The optimal window time to treat noise-induced hearing loss (NIHL) with low level laser therapy (LLLT),” presented at the Photonic Therapeutics and Diagnostics VIII, San Francisco, California, USA, 2012, vol. 8207.

[9] J. E. Marsh and T. A. Campbell, “Processing complex sounds passing through the rostral brainstem: the new early filter model,” *Front. Neurosci.*, vol. 10:136, pp. 1–33, 2016.

[10] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, *et al.*, “Brain-computer interface technology: a review of the first international meeting,” *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, 2000.

[11] X. Chen, Y. Wang, M. Nakanishi, X. Gao, T.-P. Jung, and S. Gao, “High-speed spelling with a noninvasive brain–computer interface,” *Proc. Natl. Acad. Sci.*, vol. 112, no. 44, pp. 6058–6067, 2015.

[12] Q. Obeidat, T. Campbell, and J. Kong, “The zigzag paradigm: a new P300-based brain computer interface,” presented at the Proceedings of the 15th ACM on International Conference on Multimodal Interaction, Sydney, Australia, 2013.

[13] Q. T. Obeidat, T. A. Campbell, and J. Kong, “Introducing the edges paradigm: a P300 brain–computer interface for spelling written words,” *IEEE Trans. Hum.-Mach. Syst.*, vol. 45, no. 6, pp. 727–738, 2015.

[14] Q. T. Obeidat, T. A. Campbell, and J. Kong, “Spelling with a small mobile brain-computer interface in a moving wheelchair,” unpublished.

[15] M. De Vos, M. Kroesen, R. Emkes, and S. Debener, “P300 speller BCI with a mobile EEG system: comparison to a traditional amplifier,” *J. Neural Eng.*, vol. 11, no. 3:36008, pp. 1–8, 2014.

[16] S. Debener, R. Emkes, M. De Vos, and M. Bleichner, “Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear,” *Sci. Rep.*, vol. 5:16743, pp. 1–11, 2015.

[17] J. R. Kerlin, A. J. Shahin, and L. M. Miller, “Attentional gain control of ongoing cortical speech representations in a ‘cocktail party’,” *J. Neurosci.*, vol. 30, no. 2, pp. 620–628, 2010.

[18] E. C. Cherry, “Some experiments on the recognition of speech with one and two ears,” *J. Acoust. Soc. Am.*, vol. 25, no. 5, pp. 975–979, 1953.

[19] C. Füllgrabe and S. Rosen, “Investigating the role of working memory in speech-in-noise identification for listeners with normal hearing,” in *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, vol. 894, P. van Dijk, D. Baskent, E. Gaudrain, E. de Kleine, A.Wagner, and C. Lanting, Eds. Heidelberg, Germany: Springer International Publishing AG, 2016, pp. 29–36.

[20] S. Kochkin, “MarkeTrakV: ‘Why my hearing aids are in the drawer’: the consumers’ perspective,” *MarkeTrakV*, vol. 53, no. 2, pp. 34–32, 2000.

[21] J. E. Marsh, R. Ljung, A. Nöstl, E. Threadgold, and T. A. Campbell, “Failing to get the gist of what’s being said: background noise impairs higher-order cognitive processing,” *Front. Psychol.*, vol. 6:548, 1–10, 2015.

[22] U.S. Census Bureau, Population Division, “World midyear population by age and sex for 2012,” U.S. Census Bureau, Population Division, U.S. Census Bureau 4600 Silver Hill Road Washington, DC 20233, Population estimate.

[23] B. Duthey, “Background paper 6.21 hearing loss,” *Priority Medicines for Europe and the World Update Report*, 2013.

[24] Harvey Dillon, *Hearing Aids*, 1st ed. Turrramurra, New South Wales, Austrailia: Boomerang Press, 2012.

[25] J. Colquitt, J. Jones, P. Harris, E. Loveman, A. Bird, A. Clegg, *et al.*, *Bone-Anchored Hearing Aids (BAHAs) for People who are Bilaterally Deaf: A Systematic Review and Economic Evaluation*. Southampton, United Kingdom: NIHR Journals Library, 2011.

[26] M. Kompis and M.-D. Caversaccio, “Implantable bone conduction hearing aids,” in *Advances in Oto-Rhino-Laryngology*, vol. 71. Basel, Switzerland: Karger AG, 2011.

[27] P. C. Loizou, “Speech processing in vocoder-centric cochlear implants,” in *Advances in Oto-Rhino-Laryngology*, vol. 64, Møller A. R., Ed. Basel, Switzerland: Karger AG, 2006, pp. 109–143

[28] G. Clark, *Cochlear Implants: Fundamentals and Applications*. New York: Springer, 2003.

[29] J. Hamzavi, P. Franz, W. Baumgartner, and W. Gstöettner, “Hearing performance in noise of cochlear implant patients versus severely-profoundly hearing-impaired patients with hearing aids,” *Audiol.*, vol. 40, no. 1, pp. 26–31, 2001.

[30] R. Gowri, L. Jounghee, M. Chung, J. M., S. Srila, R. Madhumathi, and L. Joseph, “Effectiveness of cochlear implants in adults with sensorineural hearing loss,” Tufts Evidence-based Practice Center, Agency for Healthcare Research and Quality 540 Gaither Road Rockville, Maryland 20850, Technology Assessment Report AUDT0510, 2011.

[31] A. E. Hassanien and A. T. Azar, "Brain-computer interfaces: current trends and applications," in *Intelligent Systems Reference Library*, vol. 74. Cham, Switzerland: Springer International Publishing, 2015.

[32] J. S. Ebersole and T. A. Pedley, *Current Practice of Clinical Electroencephalography*. Philadelphia, Pennsylvania: Lippincott Williams & Wilkins, 2003.

[33] H. Berger, "Über das elektrenkephalogramm des menschen," *Arch. Für Psychiatr. Nervenkrankh.*, vol. 99, no. 1, pp. 555–574, 1933.

[34] B. Fisch and R. Spehlmann, *Fisch and Spehlmann's EEG Primer: Basic Principles of Digital and Analog EEG,* 3rd ed., New Orleans, Louisiana: Elsevier, 2000.

[35] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, 1988.

[36] U. Michel, "History of acoustic beamforming," presented at the Proceedings of the Berlin Beamforming Conference, Berlin, Germany, 2006.

[37] J. Li and P. Stoica, *Robust Adaptive Beamforming*. Hoboken, New Jersey: John Wiley, 2006.

[38] A. W. Bronkhorst, "The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions," *Acta Acust. United Acust.*, vol. 86, no. 1, pp. 117–128, 2000.

[39] B. G. Shinn-Cunningham, "Object-based auditory and visual attention," *Trends Cogn. Sci.*, vol. 12, no. 5, pp. 182–186, 2008.

[40] S. Haykin and Z. Chen, "The cocktail party problem," *Neural Comput.*, vol. 17, no. 9, pp. 1875–1902, 2005.

[41] N. Wood and N. Cowan, "The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel?," *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 21, no. 1, pp. 255–260, 1995.

[42] A. R. Conway, N. Cowan, and M. F. Bunting, "The cocktail party phenomenon revisited: The importance of working memory capacity," *Psychon. Bull. Rev.*, vol. 8, no. 2, pp. 331–335, 2001.

[43] M. L. Hawley, R. Y. Litovsky, and J. F. Culling, "The benefit of binaural hearing in a cocktail party: effect of location and type of interferer," *J. Acoust. Soc. Am.*, vol. 115, no. 2, pp. 833–843, 2004.

[44] A. J. Duquesnoy and R. Plomp, "The effect of a hearing aid on the speech-reception threshold of hearing-impaired listeners in quiet and in noise," *J. Acoust. Soc. Am.*, vol. 73, no. 6, pp. 2166–2173, 1983.

[45] B. Edwards, "The future of hearing aid technology," *Trends Amplif.*, vol. 11, no. 1, pp. 31–46, 2007.

[46] B. R. Glasberg, B. C. J. Moore, and S. P. Bacon, "Gap detection and masking in hearing-impaired and normal-hearing subjects," *J. Acoust. Soc. Am.*, vol. 81, no. 5, pp. 1546–1556, 1987.

[47] R. J. Irwin and S. C. Purdy, "The minimum detectable duration of auditory signals for normal and hearing-impaired listeners," *J. Acoust. Soc. Am.*, vol. 71, no. 4, pp. 967–974, 1982.

[48] G. Kidd Jr, C. R. Mason, V. Best, and J. Swaminathan, “Benefits of acoustic beamforming for solving the cocktail party problem,” *Trends Hear.*, vol. 19: 2331216515593385, pp. 1–15, 2015.

[49] E. L. J. George, J. M. Festen, and T. Houtgast, “Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners,” *J. Acoust. Soc. Am.*, vol. 120, no. 4, pp. 2295–2311, 2006.

[50] S. Gordon-Salant and P. J. Fitzgibbons, “Comparing recognition of distorted speech using an equivalent signal-to-noise ratio index,” *J. Speech Lang. Hear. Res.*, vol. 38, no. 3, pp. 706–713, 1995.

[51] R. A. Bentler, “Effectiveness of directional microphones and noise reduction schemes in hearing aids: A systematic review of the evidence,” *J. Am. Acad. Audiol.*, vol. 16, no. 7, pp. 473–484, 2005.

[52] B. Widrow and F.-L. Luo, “Microphone arrays for hearing aids: An overview,” *Speech Commun.*, vol. 39, no. 1–2, pp. 139–146, 2003.

[53] H. G. Dimopoulos, *Analog Electronic Filters*. Dordrecht: Springer Netherlands, 2012.

[54] A. Saberi, A. Stoorvogel, and P. Sannuti, *Filtering Theory: With Applications to Fault Detection, Isolation, and Estimation*. Berlin, Germany: Birkhäuser, 2007.

[55] J. Sang, H. Hu, C. Zheng, G. Li, M. E. Lutman, and S. Bleeck, “Speech quality evaluation of a sparse coding shrinkage noise reduction algorithm with normal hearing and hearing impaired listeners,” *Hear. Res.*, vol. 327, no. 9, pp. 175–185, 2015.

[56] H. Levitt, M. Bakke, J. Kates, A. Neuman, T. Schwander, and M. Weiss, “Signal processing for hearing impairment,” *Scand. Audiol. Suppl.*, vol. 38, pp. 7–19, 1992.

[57] Y. Hu and P. C. Loizou, “A comparative intelligibility study of single-microphone noise reduction algorithms,” *J. Acoust. Soc. Am.*, vol. 122, no. 3, pp. 1777–1786, 2007.

[58] P. C. Loizou and G. Kim, “Reasons why current speech-enhancement algorithms do not improve speech intelligibility and suggested solutions,” *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 1, pp. 47–56, 2011.

[59] G. Kim, Y. Lu, Y. Hu, and P. C. Loizou, “An algorithm that improves speech intelligibility in noise for normal-hearing listeners,” *J. Acoust. Soc. Am.*, vol. 126, no. 3, pp. 1486–1494, 2009.

[60] P. Földiák and M. P. Young, “Sparse coding in the primate cortex,” in *The Handbook of Brain Theory and Neural Networks*, 2nd ed., Cambridge, MA: The MIT Press, 2002.

[61] H. Cheng, *Sparse Representation, Modeling and Learning in Visual Recognition*. London, United Kingdom: Springer-Verlag, 2015.

[62] P. Jančovič, X. Zou, and M. Köküer, “Speech enhancement based on Sparse Code Shrinkage employing multiple speech models,” *Speech Commun.*, vol. 54, no. 1, pp. 108–118, 2012.

[63] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming forhearing aid applications," in *Handbook on Array Processing and Sensor Networks*, S. Haykin and K. J. R. Liu, Eds. Hoboken, New Jersey: John Wiley & Sons, Inc., 2010, pp. 269–302.

[64] K. Adiloğlu, H. Kayser, R. M. Baumgärtel, S. Rennebeck, M. Dietz, and V. Hohmann, "A binaural steering beamformer system for enhancing a moving speech source," *Trends Hear.*, vol. 19:2331216515618903, pp. 1–13, 2015.

[65] R. Koninga and J. Wouters, "The potential of onset enhancement for increased speech intelligibility in auditory prostheses," *Acoust. Soc. Am.*, vol. 132, no. 4, pp. 2569–2581, Oct. 2012.

[66] B. U. Seeber and J. J. M. Monaghan, "Envelope enhancement for improving hearing in reverberant spaces," in AIA-DAGA 2013, Int. Conf. on Acoustics, Merano, Meran, Italy, 2013, pp. 1071–1072.

[67] C. Freigang, J. W. Browne, and B. U. Seeber, "Does onset enhancement improve availability of envelope ITDs in cochlear implant users?," in *Fortschritte der Akustik - DAGA '14*, 2014, pp. 74–75.

[68] T. Campbell, I. Winkler, and T. Kujala, "N1 and the mismatch negativity are spatiotemporally distinct ERP components: Disruption of immediate memory by auditory distraction can be related to N1," *Psychophysiology*, vol. 44, no. 4, pp. 530–540, 2007.

[69] T. A. Campbell, "A theory of attentional modulations of the supratemporal generation of the auditory mismatch negativity (MMN)," *Front. Hum. Neurosci.*, vol. 8:548, pp. 1–10, 2015.

[70] A. Dresch, D. P. Lacerda, and J. A. V. Antunes Jr, *Design Science Research*. Cham, Switzerland: Springer International Publishing, 2015.

[71] P. Johannesson and E. Perjons, *An Introduction to Design Science*. Cham, Switzerland: Springer International Publishing, 2014.

[72] H. A. Simon, *The Sciences of the Artificial*, 3rd ed., Cambridge, Mass.: MIT Press, 2008.

[73] J. Zobel, *Writing for Computer Science*. London: Springer London, 2014.

[74] S. Arora and B. Barak, *Computational Complexity: A Modern Approach*, 1st ed. Cambridge, United Kingdom: Cambridge University Press, 2009.