A study on the Dik Universe

Development of a bigger dik

Darius Chira,Razvan Bucur,Titas Lasickas, Paulius Riauka
Electronics and Computer Engineering, TEPE4-1005fuk, 2018-08

Bachelor Project

STUDENT REPORT
AALBORG UNIVERSITY

Here you can write something about which tools and software you have used for typesetting the document, running simulations and creating figures. If you do not know what to write, either leave this page blank or have a look at the colophon in some of your books.

AALBORG UNIVERSITY
STUDENT REPORT

Electronics and IT
Aalborg University
http://www.aau.dk

Title:
Project Title

Theme:
Digital Filtering

Project Period:
Fall Semester 2018

Project Group:
ED5-1-E18

Participant(s):
Darius Chira
Razvan Bucur
Titas Lasickas
Paulius Riauka

Supervisor(s):
Daniel Ortiz-Arroyo

Copies: 1

Page Numbers: 23

Date of Completion:
December 5, 2018

Abstract:

This report presents two different filtering methods.
One that involves basic filtering techniques, the other that uses machine learning in order to filter speech. All the samples used in the report have been either recorded using a setup presented or synthetically made. Specific samples have been used to train the neural network.
Different scenarios are described in order to test the filtering method and observe its behavior.

Contents

Todo list

Preface

Here is the preface. You should put your signatures at the end of the preface.

Aalborg University, December 5, 2018

Author 1
<username1@XX.aau.dk>

Author 2
<username2@XX.aau.dk>

Author 3
<username3@XX.aau.dk>

Author 4
<username4@XX.aau.dk>

Chapter 1

Introduction

Humans have the ability to identify the source of a sound around them. In the field of neuroscience, this capability is called sound localization. The brain can determine the location of a sound with very high precision, up to 2 degrees of space. This comes from the brain's capacity to interpret information received from both ears.

Over the years, neuroscientists, have been trying to understand the mechanisms within our brains that are able to determine the location of a sound. They have identified two cues that are essential and sufficient for horizontal sound localization.

In the 1790s, Giovanni Battista Venturi conducted experiments where he played a flute around blindfolded people and asked them to point in his direction. He concluded that the sound amplitude difference between the two ears was the indicator used for determining the direction.

Much later, Malloch proposed that the difference in time between the two ears was the sign used for determining the direction of a sound.

Years later, scientists found neurons in the auditory center of the brain specially adjusted for each indicator: time and intensity differences between the two ears.

Figure 1.1 shows a circle, and a person in the middle. The person is meant to be the listener, while the circle represents a perfectly flat plane around the listeners head.

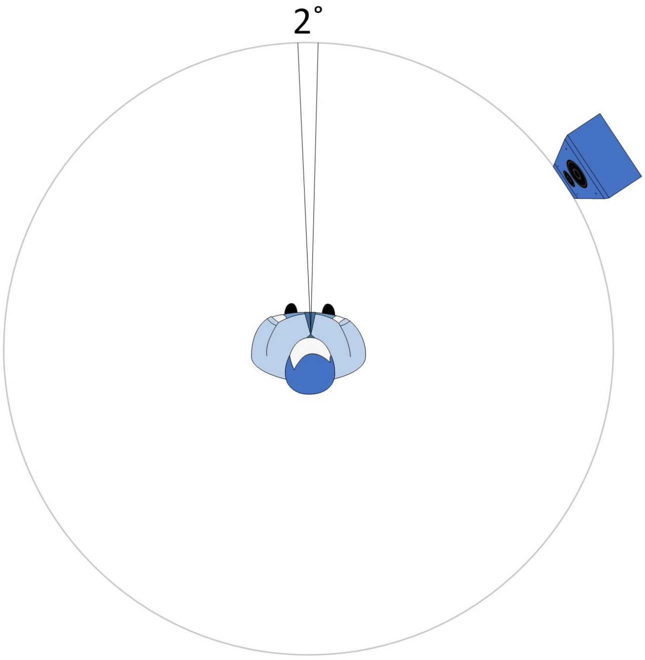

Figure 1.1: Human Hearing Accuracy

Sound coming from the speaker, would reach the right year faster and be louder than the sound that reaches the left year. The brain is able to compare the differences and tell where the sound is coming from.

Reading Guide

Chapter 2 deals with problem analysis. Chapter 3 describes the setup and components used for recording samples. Chapter 4 and 5 focuses on the two filtering methods tested. The results for both methods are presented in Chapter 6. Chapter 7 and 8 present a summary of the work done, suggest possible future uses of the research presented and draw a conclusion of the work carried out.

somebody read this and say what they think
"Chater 3 describes the preparation work for this project"

Chapter 2

Problem Analysis

2.1 Problem Description

The project is about two different approaches towards speech filtering.

The first, involves using two microphones in order to record a two or more persons talking, from different angles. With the sound samples recorded, the next step involves handling them in such a manner that allows to filter one voice out, and keep the other.

Once one voice has been filtered out, the next step is using machine learning. This is done in order to filter out any other noises and return a speech as clean and as clear as possible.

The setup for recording the samples consists of only two microphones placed at approximately the same height as a persons mouth.

2.2 Problem Delimitation

Vertical Sound Filtering

Chapter 3

Development

3.1 Component List

3.1.1 Introduction

To create the basis for development of this project two microphones and a structure to keep a constant distance between them were necessary.

3.1.2 Microphones characteristics

Firstly a decision had to be made whether to use a directional or omnidirectional microphone. Since omnidirectional microphones provide less plosive and wind sounds, does not build up bass it was considered a good option. The fact that it provides equally good audio quality at every angle made it the best choice for the aim of this project.

Voice frequency ranges vary heavily depending on whether it sources from a male of a female. Fundamental voice frequency varies from 100Hz to 900 Hz for men and 350Hz to 3KHz for women. Including peaks to conserve natural sounding voice, a wider frequency range has to be considered. It rises to 8 KHz for males and 17KHz for females [**Seaindia**]. Yet different researches often come up with different results. For example, in phone communications it is accepted to transmit frequency range between 400Hz and 3400Hz. This is the reason some peoples' voices transit poorly over the phone yet for most cases it works fine. This example allows to conclude that smaller frequency ranges could be acceptable but to conserve all of the properties of the human voice microphones do need to be capable of recording at 17kHz.

Our requirements for the microphones were:

- **Affordable price**
  Since the university did not have any planned funding for this semester, it was agreed to aim for budget options. This way it would not be necessary to work towards an agreement with the University to receive funding - some members wanted to have microphones for themselves.

- **Capability to capture the entire range of human hearing frequency**
  Since one of the great benefits of an omnidirectional microphone was it's ability to capture clear sound, it would not have been reasonable to purchase microphones that were not able to capture the full clarity of the input. Since the entire human voice frequency fits within our hearing frequency range, there is no need to have any extra requirements.

- **Appropriate size factor**
  It would be beneficial and logical if the microphones used in the testing would have potential to be a part of the design prototype. Although due to limited budget there is a low chance of affording the microphones with an appropriate size parameters for a hearing aid or earphones, objective to apply a microphone that could be used in the further stages of development remains as one of lower priorities

- **They must to be identical**
  Having identical microphones assures that if there was a delay within recording timings, it would not be caused by hardware differences.

- **Capability to connect multiple microphones to a single computer and record both at the same time**
  For the sake of simplicity regarding sampling and testing, it was agreed that it would be far easier to sample using microphones that can be connected though Universal Serial Bus port instead of Auxiliary one. This way troubleshooting and initiating microphones should be more clear.

3.1.3 Microphones

As requirements for microphones were set, it was first attempted to find them at the University. As it was found that University does not own any microphones that could be applied for the purpose of this project, it was agreed to purchase budget microphones that could come as close to the requirements raised for the project as possible. Bearing in mind that a month-long delivery from Asia is not a viable option, it was decided to order microphones available within an acceptable delivery time .

microphone reference

3.1.4 Distancing

To maintain a constant distance it was decided to attempt modeling a rod with two microphone holders at it's ends. The distance chosen between the microphones was agreed on by referring to a research article in this area: 15.2 centimeters .

reference

Other parameters were found by following the measurements found in the datasheet 3.1. These steps have led us towards the structure development, which after a few iterations has turned into the structure that was used for sampling .

illustration missing (3D model or real one)

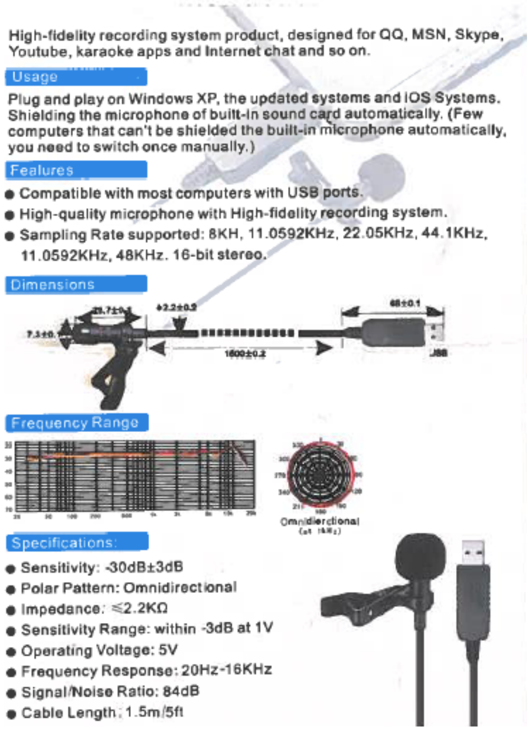

Figure 3.1: Microphones data sheet

https://www.researchgate.net/publication/228749231_Analysis_of_the_Facial_Anthropometric_Data_of_Korean_Pilots_for_Oxygen_Mask_Design

3.2 Setting up the scene

Sampling was done in a lecture/ conference room with side noise: a server, maintaining the virtual conference platform (B107). This area was chosen because it had represented a real life conversation scenario. It is planned to conduct tests with a working system in a different environments to validate the effectiveness of it. Creating a set of samples in every different environment throughout the development phase would take a lot of time and therefore was considered unnecessary for the early stages.

Since the sampling ratio of human ear can only reach up to 20kHz, 48kHz sampling frequency of the microphones the Nyquist frequency and therefore guarantees that there will not be any hearable audio quality losses.

3.2.1 Analog to digital conversion

Since the datasheet of the microphone does not provide any information on the Analog-to-Digital Converter that is used to process the data and there is not default ADC that is used for this process, the only potentially successful way to find the details about the ADC used for the signal processing is to open the microphone and look for the information on the ADC itself. Since the structure seemed to be glued together, we will not attempt to find the details of it and just rely on the specifications sheet provided by the manufacturer.

Chapter 4

Directional Noise Elimination

4.1 Introduction

This chapter deals with directional filtering. The idea is presented and tested. The results and conclusions will be shown afterwards.

4.2 Concept

Figure 4.1 shows a scenario where directional filtering could be used. Source 1 and Source 2 are both talking simultaneously. At the Origin, two microphones left (L) and right (R) are recording both signals.

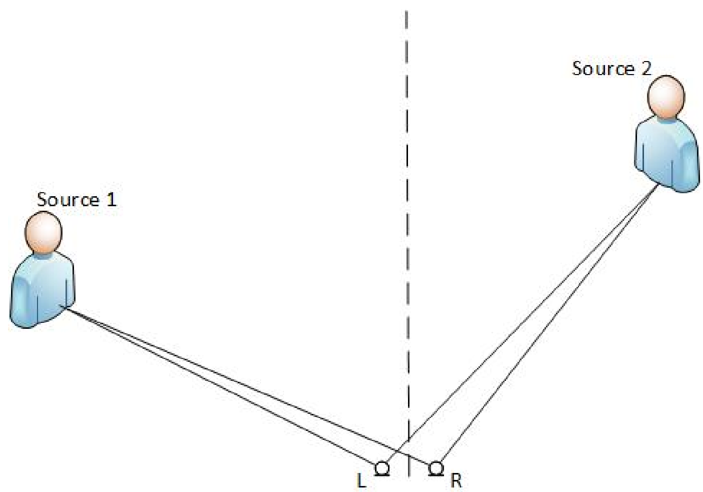

Figure 4.1: Two Persons Talking

Due to the placement of the sources, both microphones will record the sounds, with a certain delay.

4.3 Idea for filtering

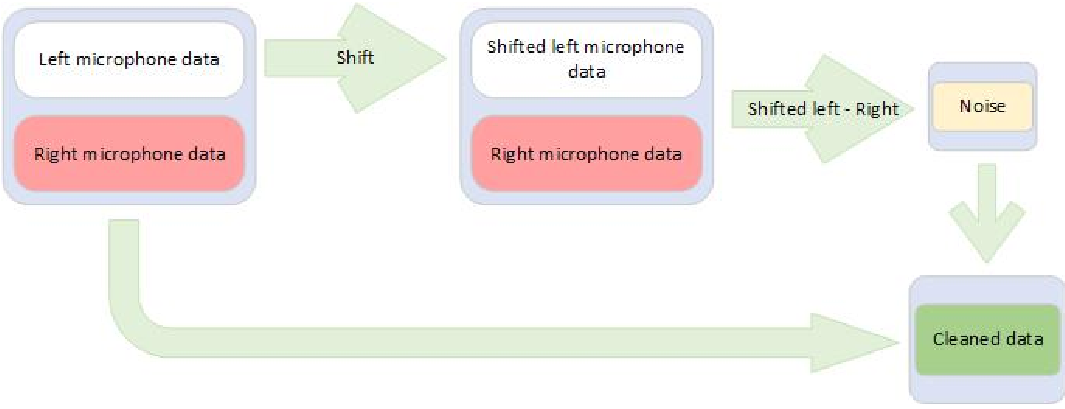

Figure 4.2: Idea diagram

Filtering sound, based on the direction that it's coming from could work by taking data, recorded by one of the microphones, in this case (figure 4.2) we chose left microphone data. Then we shift this data by exact amount of samples, which would make the source we would like to separate recorded by left microphone, exactly aligned with same source, recorded by right microphone.
Then by subtracting right microphone data from shifted left, we should get all the other sounds except the ones from source we would like to separate, this is what we will call noise in directional filtering.
Separating could be done by taking the original data, putting it to one channel. Then subtracting noise from the original data should leave us only with "clean" sound from the direction, we are interested in.

4.4 Development

MENTION WHERE EACH SUBSECTION COMES IN PLAY

4.4.1 Finding delay in samples

By setting the angle of what we want to separate our signal at, we can calculate how big shift in samples that would account for. We are using this delay later, in the directional filtering. Calculations involve using selected angle, gap between microphones, sampling frequency of used microphones and the speed of sound.

DO WE WRITE ABOUT MATH AND EQUATIONS IN MATLAB HERE OR DO WE JUST SKIP IT.

4.4.2 Recording samples

First, in order to record two microphones in separate tracks directly to computer, we need to give access to multiple inputs and outputs, independently to computers' sound card. To achieve that audio driver caller ASIO4ALL was used. Using this driver, together with recording software, which supports multiple input recording, we were able to record different situations with two microphones at the same time.

Recording with two microphones was set in a quiet room with sound panels on the walls and curtains, all of which reduces echo. Using ruler as a guide microphones were set in place and then angles were marked as guides to know from which direction sound is coming from. (Figure 4.3)

process of setting up for first recordings. Last sentence counters delay paragraph in Filtering section (4.3.3)

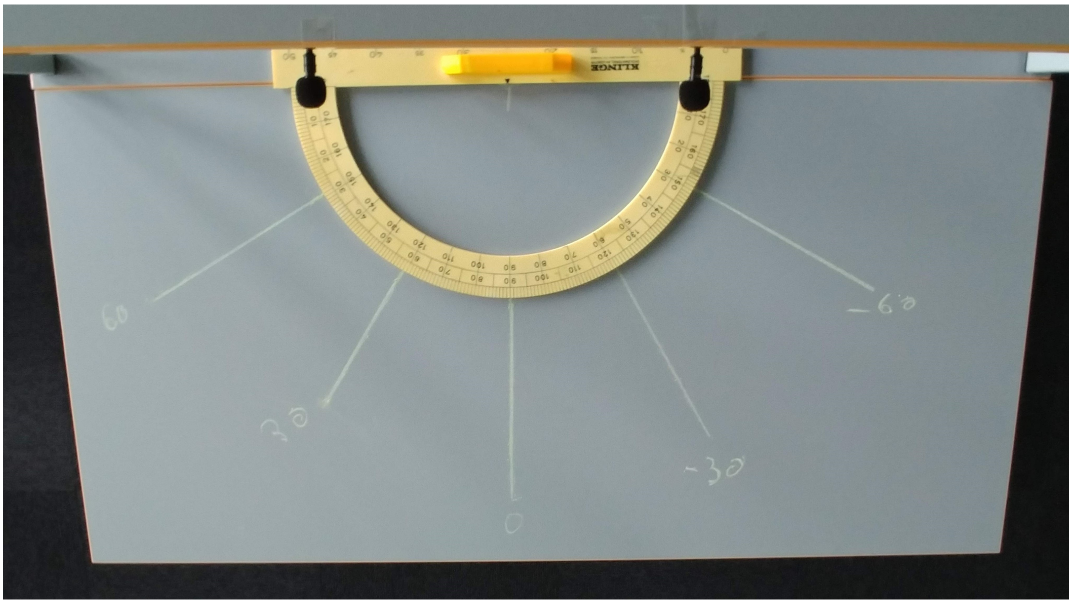

Figure 4.3: Setup for recording

4.4.3 Filtering

All of the directional filtering is done by applying the logic discussed before to the Matlab script.

Prior to shifting signals and filtering, we had to remove any delays, induced by hardware or software. This issue was resolved by starting every recording with loud, sharp sound like a clap or a finger snap right in the middle of the microphones. Using this sound in the beginning we could match loudest peaks, thus eliminating hardware and software induced delay.

To see if we can get the samples, which show that logic behind direction of delay is correct, we set up a recording with just one person first see figure 4.4.

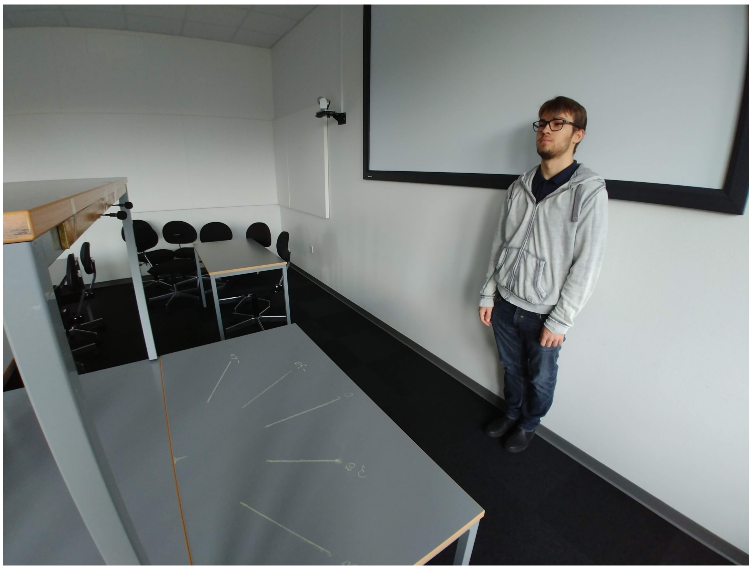

Figure 4.4: Setup for recording with a person

Then we looked at collected data:
When person spoke from the center, (Figure 4.5) shows that recordings match in phase. Green graph is data, captured by left microphone and yellow graph is data, captured by right microphone.

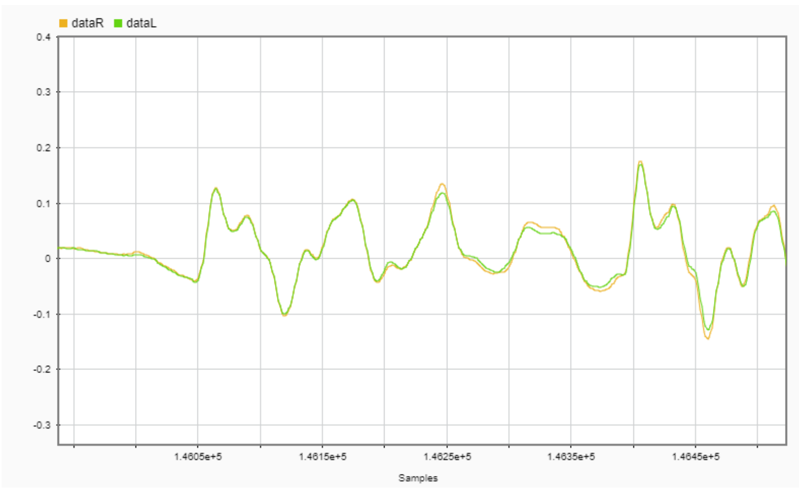

Figure 4.5: Data from speaker in the center

When person spoke from the right, right microphone data led left microphone (Figure 4.6), here blue graph is left microphone data and orange graph represents right microphone data.

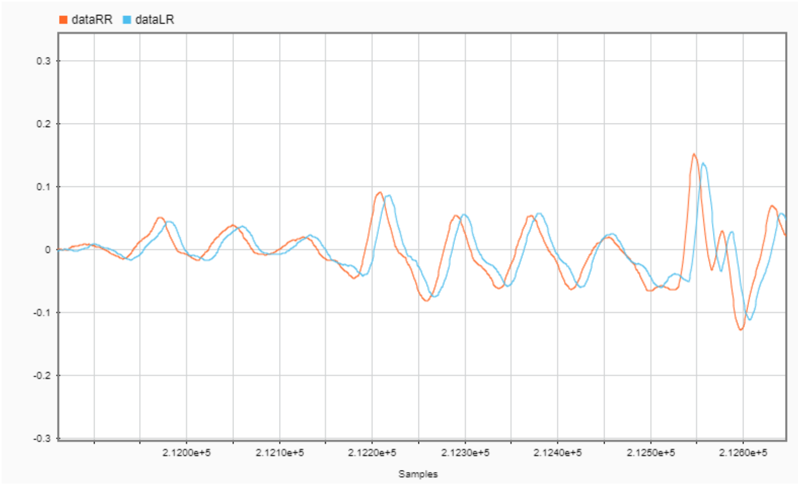

Figure 4.6: Data from speaker in the right side

When he spoke from the left, left microphone was leading the right (Figure 4.7) pink graph is left microphone data and purple graph shows right microphone data.

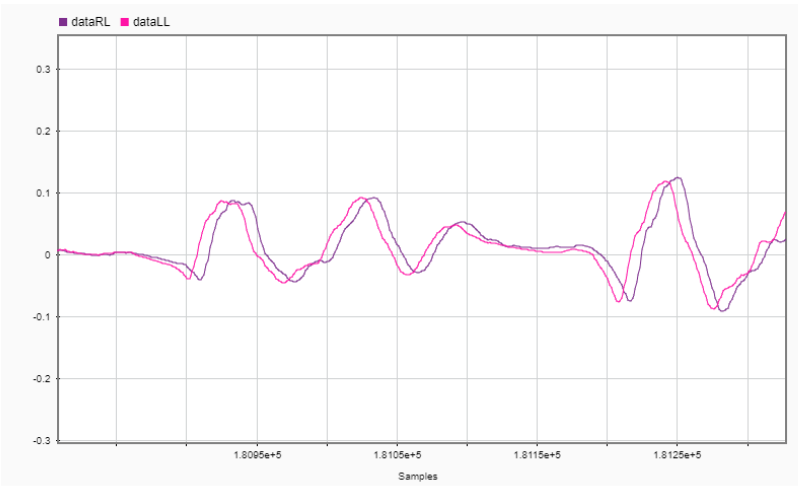

Figure 4.7: Data from speaker in the left side

"For range -1 to 1 20*log10(x) corresponds to dB" Mention?

All of the directional filtering is done by applying the logic discussed in the Idea for filtering section to the Matlab script.

First we took data from the recording with only one person, shifted left microphone to match right microphone data in phase, then subtracted right microphone data from shifted left microphone data - got "Noise". To finish filtering we tried to subtract "Noise" from the Right microphone data.

4.5 Conclusion

Chapter 5

Neural Network Speech Isolation

Chapter 6

Results

Chapter 7

Discussion

7.1 Future Work

Hearing aid technologies have never been improving as fast as they do now. Ever since major smartphone manufacturer companies started investing into wireless wearable technology research, price of PSAP (personal sound amplification products) has decreased while the amount of features has increased. Throughout recent years wireless earphones became predominant in the global market due to new generation of Bluetooth technology. This improvement even further reduces power consumption of wireless technology. The only three differences between hearing aids and PSAPs were:

•The battery life. Due to a far higher number of features and consistent audio stream PSAPs consume a much higher amount of power.

•Hardware design differences. PSAPs are not oriented around sound localization to inform the user about where the sound is coming from regarding natural sources. PSAPs are often not oriented to be invisible to others, they are more often purposely made to stand out and be recognized among its competitors. Fit customization is also often minimal on PSAPs.

•Regulation requirements to produce the hearing aid and license requirements to sell it.

Most other differences lay in software and could be eliminated through a software update.

It is believed that legislative issues could be solved if manufacturers would put effort to reach for an agreement with legislators although it would require a lot of changes since current hearing aid selling process consists of far more than just taking the product off the shelf and swiping it through the register - it is normally performed at hearing clinics, hearing aid is thoroughly adjusted to fit the consumer's ear for long periods of time, warranty for these devices also is taken in a far more serious manner: it comes with included follow-up office visits, checks and cleaning procedures to maintain the highest level of performance. Some companies do express interest to merge the two markets. According to "The State of Hearing Healthcare 2017" by Lindsey Banks, "If Apple Air Pods or Samsung Gear IconX could add in hearing aid functions, that's instant access to over half of the U.S. over night."

https://www.everydayhearing.com/hearing-loss/articles/state-of-hearing-healthcare-2017/
((puthashhere))tech

The value of argument regarding battery life of these two hearables should also heavily decrease in the coming years. At the end of October 2017 Samsung has announced that a considerably new generation of battery has been developed. Currently used lithium-ion batteries seem to have been pushed to it's limit and yet it still takes a fairly long time time to charge in a fast-paced society. This problem has pushed electronics manufacturers to develop energy efficient processors. A new graphene-based battery technology should enable 45 percent more capacity and 5 times faster charging speeds.

These reasons should lead to a breakthrough in battery life factor of next year's electronics. If not at 2019, by 2020 hearing aids should receive this battery update. Combined with improvements in Bluetooth technology, these reasons should encourage both hearing aid and consumer audio manufacturers to increase number of features in coming year's hearing aids as well as PSAPs and might bring the markets closer together.

Chapter 8

Conclusion

In case you have questions, comments, suggestions or have found a bug, please do not hesitate to contact me. You can find my contact details below.

Jesper Kjær Nielsen
jkn@create.aau.dk
http://sqrt-1.dk
Audio Analysis Lab, CREATE
Aalborg University
Denmark

Appendix A

Appendix A name

Here is the first appendix