Requirement Specification

Elias William
Rasmus Olofsson
Malva Eveborn
Oskar Ramsberg
Mohammad Alasmi
Adam Roos

December 12, 2023

Version 1.1

LINKÖPINGS UNIVERSITET

Status

Reviewed	Elias William	2023-10-03
Approved	Martin Skoglund	2023-10-04

Oticon
eliwi204@student.liu.se

Project Identity

Group E-mail: eliwi204@student.liu.se

Homepage: http://www.isy.liu.se/tsrt10/group

Orderer: Martin Skoglund, ISY/Eriksholm Research Centre
Phone: +46 13 28 18 90
E-mail: martin.skoglund@liu.se

Customer: Sergi Rotger Griful, Oticon
Phone: -
E-mail: segr@eriksholm.com

Supervisor: Johanna Wilroth, ISY
Phone: +46 13-28 28 05
E-mail: johanna.wilroth@liu.se

Course Responsible: Daniel Axehill & Gustaf Hendeby
Phone: +46 013-28 40 42 & +46 13-28 58 15
E-mail: daniel.axehill@liu.se & gustaf.hendeby@liu.se

Participants of the group

Name	Responsible	E-mail
Elias William	Project Leader (PL)	eliwi204@student.liu.se
Rasmus Olofsson	Responsible for software (SW)	rosol028@student.liu.se
Malva Eveborn	Responsible for documentation (DOC)	malev556@student.liu.se
Oskar Ramsberg	Responsible for design (DES)	oskra262@student.liu.se
Mohammad Alasmi	Responsible for hardware (HW)	mohal555@student.liu.se
Adam Roos	Responsible for testing (TEST)	adaro336@student.liu.se

CONTENTS

DOCUMENT HISTORY

Version	Date	Changes made	Sign	Reviewer
0.1	2023-09-15	First draft.	All	EW
0.2	2023-09-20	Changes made after feedback from orderer and supervisor.	ME, OR	EW
0.3	2023-09-29	Changes made after feedback from orderer and supervisor.	EW, RO, AR	MA
1.0	2023-10-03	Changes made after feedback from orderer.	EW	AR
1.1	2023-11-23	Changes made after renegotiation of requirements	EW, RA, AR, ME	EW

1 INTRODUCTION

Communicating in noisy environments like restaurants, festivals and exhibitions remain troublesome for people suffering of impaired hearing. Traditionally, hearing aid devices have amplified all surrounding sounds of the user, making them avoid these types of rowdy environments. One way to improve the experience of hearing aid users would be to understand the environment and the intent of the user, enabling focused sound amplification combined with noise reduction.

This project aims to investigate and develop methods for using information about the environment to be able to extract useful sound to the user. This will be done with the assistance of a pair of wearable glasses providing sensors like camera and eye-tracker.

In collaboration with the hearing aid company Oticon and Linköpings University, we aim on this years' CDIO-project to further develop the research field to help people suffering from hearing problems with the focus on real-time implementation. The project will be carried out using the LIPS-model and this document will describe the system and requirements on the product to be developed.

1.1 Partners

The people involved in this project are the customer, orderer, supervisor, and the project group.
Customer: Sergi Rotger Griful (SEGR), Eriksholm Research Centre.
Orderer: Martin Skoglund (MNSK), Division of Automatic Control at LiTH.
Supervisor: Johanna Wilroth, PhD student, Division of Automatic Control at LiTH.
Project group, 6 students at LiTH.

1.2 Definition of terms

Abbreviations used in the document are presented in Table 1.

Table 1: Definition of terms.

Term	Meaning
G3	Tobii Pro Glasses 3.
IMU	Inertial Measurement Unit.
SNR	Signal to Noise Ratio.
CDIO	Conceive Design Implement Operate.
GUI	Graphical User Interface.
LIPS	Project model.
DP	Decision point, a decided time when specific implementation has to be made.

1.3 Aims and goals

The goals and aims in this project are mainly separated into two parts which will be attempted to be combined at a later stage of the project. The first part of the project is to develop software for the hearing aid devices. This includes noise reduction, beamforming and sound tracking. The second part of the project will focus on implementing software algorithms like face detection and orientation estimation for Tobii Pro Glasses 3 (G3) using its available sensors.

1.4 Background information

The main purpose of a hearing aid kit is to amplify the sound that the user want to hear and suppress unwanted background noise. However, hearing aid devices does often have a limited amount of hardware with few sensors. The ability to amplify the wanted sound signals in noisy environments is therefore a difficult task. In order to get a better understanding of the environment more information is needed. One way to analyse the environment further is to use glasses with different sensors such as cameras and inertial sensors in combination with a hearing aid kit. Combining the hardware can create a better view of the environment and clarify what the user want to hear. This will improve the ability to control the hearing aid in a meaningful way by filtering out noise and amplifying what the user want to hear.

1.5 Use

The system is supposed to provide information about the user and the environment around the user by using two subsystem. The G3 system should be able to detect faces, track them and estimate the depth of the detection. The hearing aid devices should be able to amplify sound in a given direction and track the sound objects.

2 SYSTEM OVERVIEW

The system developed in this project revolves around two pieces of hardware. The first piece of hardware is a prototype hearing aid device including amplifiers and a sound-card for communication with a computer. There is one hearing aid device for each ear, in other words two.

The other hardware component is the Tobii Pro Glasses 3 used for collecting data about the surrounding environment of the user. The data can be collected by its different sensors like camera, IMU, magnetometer, microphone and eye-tracker.

2.1 Description of the product

In this section the system is described in more detail.

2.1.1 *Tobii Pro Glasses G3*

The latest generation of Tobii Pro Glasses is used for collecting data of the surrounding enviroment of the user. With sensors such as front facing camera, eye-tracker, microphones, IMU and magnetometer can provide additional information for interpreting the user's surrounding environment. This new generation improves from the previous with all around better and more precise sensors and also extends the sensor suite with a magnetometer. The glasses are powered through a battery pack that also has a slot for a SD-card for recording data. The package can be connected by WIFI or internet cable to a computer [1].

2.1.2 *Prototype hearing aid devices*

The prototype hearing aid devices consists of two microphones each and a in-ear earbud in a so called Receiver in the Ear Canal (RITE) configuration. The sound from the microphones pass through an amplifier before connecting with a computer by a wire. The sound played in the earbud is played from the computer by wire through an external sound-card.

2.2 Product components

This section describes the components of the different hardware products used in the project.

2.2.1 *Tobii Pro Glasses G3*

Figure 1: Tobii Pro Glasses G3. Image reproduced by permission from Tobii AB [2].

Figure 1 depicts the components of the Tobii Pro Glasses G3. In addition to the head unit, the recording unit, which also acts as a battery, is connected to the glasses by a wire.

2.2.2 *Prototype hearing aid devices*

The hearing aid devices can be seen in Figure 2. They consist of two microphones each and one in-ear earbud. The microphones connect to an amplifier with an RCA-connector (red and white). The earbud receives audio from a sound-card (black connector).

2.3 Dependency of other systems

This section describes other components necessary for the system to function properly and also the reference systems used to evaluate estimates.

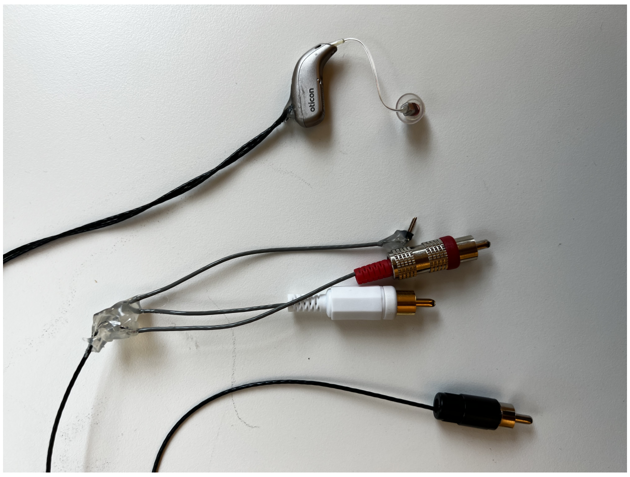

Figure 2: Hearing aid prototype device.

2.3.1 *Amplifier*

A phono pre-amplifier is going to be used to amplify the sound from the hearing aid device microphones. The pre-amplifiers are of model type "NANO-LP1". They consist av 2 input ports and 2 output ports. There are 2 pieces pre-amplifiers used each one for each earpiece.

2.3.2 *Sound-card*

An audio interference is going to be used in the project to record the incoming sound from the microphones. The audio card is of the model "Maya44" which has 4-inputs and 4-outputs. The sound card has a USB connection to connect to a computer device where the input signals can be processed and measured.

2.3.3 *Other Systems*

Previous years project has focused mainly on the Tobii Pro Glasses 2, which is now replaced by G3. The software development of G3 can be inspired by previous year's CDIO-project work, code from Oticon and code for modules from Tobii Pro Glasses 2.

2.3.4 *Visionen*

To evaluate the estimates made the reference system in Visionen Arena is used. Visionen is equiped with a Qualisys motion capture system consisting of a set of cameras that capture IR reflections of a set of reflective markers placed on a target. From the Qualisys system the ground truth can be accessed to compare to visually.

2.4 Included sub-systems

The software that will be developed is going to integrate data from the hearing aid and from the G3 glasses. Data from both devices is going to be inputs or states to modules like image processing, sound processing and noise reduction. The sub-systems themselves will cooperate with one another to produce the output signal to the user.

2.5 Limitations

The project group can only spend a total of 1440 hours on this project, which corresponds to 240 hours per person. Furthermore, the software used in this project can be computationally expensive which demands powerful computers. Another limitation in the project is the new equipment consisting of the hearing aids. Since they never been used before more time may be needed to acquire knowledge of how they behave and to make them work with G3.

2.6 Design philosophy

The project is going to be organised and sectioned modularly to enable effective developments in each section separately. In addition, the design of the product will be developed in order to make it real-time applicable.

3 HEARING AIDS

Following section specifies all the requirements to be developed for using the hearing aid prototypes.

3.1 Hardware

This section describes requirements on the hardware.

Requirement	Version	Description	Priority
1	1.0	One battery holder for both hearing aid devices should be built which powers the hearing aid.	1

3.2 Sound source tracking module

Below are the requirements for the sound source tracking module which only uses measurements from the haring aids.

Requirement	Version	Description	Priority
2	1.0	Be able to recognize an active sound source of interest with automatic speech recognition (ASR) within 3 seconds.	1
3	1.1	Be able to localize the direction of a moving sound source (a human voice), based on measurements from microphones. The sound source will be located at a distance of 1 m to 5 m away from the user in the area in front of the user. The user will be stationary and no additional noise will be present. The maximum mean estimation error will be 10 degrees and with a maximum standard deviation of 10 degrees. This need to be achieved in 5 seconds after the target is in a stand still position.	1

cont. on next page

cont. from previous page			
Requirement	Version	Description	Priority
4	1.0	Be able to localize the 3D position of a stationary sound source (a human voice), located at a distance of 1 m to 5 m away from the user, when the user is stationary and no noise is present with a maximum estimation error of 0.5 meters based on measurements from all microphones.	3
4A	1.0	Be able to estimate the distance to a moving sound source (a human voice) using the amplitude of the measured signal, based on measurements from microphones. The sound source will be located at a distance of 1 m to 5 m away from the user in the area in front of the user. The user will be stationary and no additional noise will be present. The maximum mean estimation error will be 2 meters and with a maximum standard deviation of 2 meters. This need to be achieved in 5 seconds after the target is in a stand still position.	1
4B	1.0	Be able to combine the DOA and range measurements to estimate a 2D position to a sound source (a human voice) live. The sound source will be located at a distance of 1 m to 5 m away from the user in the area in front of the user.	1
4C	1.0	Requirement 3, 4A and 4B should be achieved on recorded data from the hearing aids.	1
5	1.0	Be able to control outlier rejection based on probability of speech being present.	1
5A	1.0	Be able to control outlier rejection of noisy measurements using the Mahalanobis distance for update rejection in the Kalman filtering of both DOA and range estimation.	1
6	1.0	Be able to track the 3D position of sound source (a human voice) moving in front of a stationary user at a distance of 1 m to 5 m from the user with an estimation error of maximum 0.5 m.	2
7	1.0	Be able do all of the requirements from 2 to 6 but with more than one sound source.	3

3.3 Noise reduction module

To be able to analyse the microphone measurements, pre-processing is needed to reduce noise and to help identify sounds/signals of interest. Post-processing is also used to focus on signals of interest. Requirements for the noise reduction module is stated in the table below.

Requirement	Version	Description	Priority
8	1.0	Implement a monaural beamformer using both microphones on a hearing aid to amplify speech in the frontal direction relative to a static user and suppress speech from the opposite direction with a input signal-to-noise ratio (SNR) improvement of 5dB in the hearing aids.	1
9	1.1	Implement binaural beamforming combining data from one microphone from each hearing aid which in a setup of one speech source (signal) in a predetermined direction which the beamformer aims at and a interfer (noise) at least 60 degrees away from the signal direction, the signal should be amplified compared to the interferer.	1
9A	1.0	Computations should be able to be done on recorded data.	2
10	1.1	Implement binaural beamforming combining data from both microphones from each hearing aid which in a setup of one speech source (signal) in a predetermined direction and a interfer (noise) at least 60 degrees away from the signal direction, the SNR improvement should be at least 7 dB.	2
11	1.0	From the input to one microphone two different speech streams, the hearing aids should be able to amplify one of them and suppress the other with a signal-to-interference-plus-noise-ratio (SINR) of 7dB.	3
12	1.0	In a noisy environment, consisting of two speakers placed in front of the user, at least 70 degrees apart, between 1 and 5 meters away, producing white noise, the hearing aids should be able to recognize speech with an SNR of 10dB.	2
12A	1.0	In an environment with a human voice mixed with white noise, with an SNR of 2dB, the hearing aids should be able to reduce the white noise.	1
13	1.0	A filter will be implemented to amplify sound in a given frequency band which is adapted for the user's preference.	2
14	1.0	Be able to amplify different frequency bands depending on the user's PTA (pure tone audiogram).	3
15	1.0	Compress the audio signal to fit into a user's audible range.	3

4 G3

This section will specify the requirements of the G3 system. The system should be able to use its camera to detect, track and estimate the depth of faces in the camera feed as well as estimate the orientation of the glasses.

4.1 Face detection and tracking module

The face detection module should be able to detect faces from the image input stream from the camera of the glasses. The following section lists the requirements for the face detection module.

Requirement	Version	Description	Priority
16	1.1	In a well lit environment, with a face positioned 1 to 5 m from the user in the camera view, from one camera frame a face should be detected with a certainty of at least 90%.	1
17	1.1	The location of a detected face should be estimated with at least an IoU of 0.5 in pixels from the ground truth at a distance of 2 meters from the user.	2
18	1.1	With a stationary camera setup a face moving in front of the camera, at a distance of 1 to 5 meters from the user, with a speed of less than 100 pixels/s should be tracked between frames by detection translation with a maximum disassociation of 5 during a video of 10 seconds.	2
18A	1.0	In a video of a face during 5 seconds, in a well lit environment, with the user stationary and the face between 1 and 5 meters form the user where the face moves with a constant speed from the left to the right side of the image plane. The face should be contained within the tracking box and be tracked with less than 5 disassociations.	1
18B	1.0	Computations should be able to be done on recorded data.	1
19	1.0	With a stationary camera setup a face moving in front of the camera, at a distance of 1 to 5 meters from the user, with a speed of less than 100 pixels/s should be tracked between frames by visual association with a maximum disassociation of 5 during a video of 10 seconds.	2
20	1.0	The face should be tracked between frames taking into account the glasses' orientation change.	2
21	1.0	Requirements 16 to 20 should apply for at least 3 faces in the image.	2

4.2 Depth estimation module

The depth estimation module will use information from the face detection module to estimate depth. The following section lists the requirements for the face detection module.

Requirement	Version	Description	Priority
22	1.1	The estimated distance to a face should be within 30 % of ground truth when the face is further away than 1 m and closer than 5 m from the glasses.	1
22A	1.0	Computations should be able to be done on recorded data.	1

4.3 Eye-tracking module

The eye-tracking module will provide information about where the user is gazing. The following section lists the requirements for the eye-tracking module.

Requirement	Version	Description	Priority
23	1.0	The direction of the gaze should be associated with a point in the image from the glasses' camera.	1
24	1.0	In a stationary setting the module will be able to output a direction that the eyes are pointing towards, relative the glasses, with a maximum estimation error of 10 degrees.	1
25	1.0	The module will be able to output a 3D coordinate position within 0.5 m of ground truth, relative the glasses, from where the eyes gaze intersect when the eyes are looking at something at a maximal distance of 5 m.	1
25A	1.0	Computations should be able to be done on recorded data.	1

4.4 Orientation module

The inertial measurement unit or IMU will be used to help estimate the state and orientation of the glasses. Below is a list of requirements set on the orientation module.

Requirement	Version	Description	Priority
26	1.0	The gyroscope, accelerometer and magnetometer should be used to estimate the roll, pitch and yaw angle when the G3 glasses is static on the ground with a maximum orientation error of 20 degrees each for the roll, pitch and yaw angle.	1
26A	1.0	Computations should be able to be done on recorded data.	1
27	1.0	Gyroscope bias is calculated when the G3 glasses is static on the ground during 10 seconds.	1
28	1.0	The bias for the gyroscope should be compensated for.	1

5 ENTIRE SYSTEM

This section specifies the requirements of the entire system, which include the hearing aid devices and the G3 glasses.

5.1 Relative orientation module

The requirements of the orientation module for the entire system are listed below.

Requirement	Version	Description	Priority
29	1.0	The relative position of the hearing aid devices to the glasses should be calculated and modeled for a predetermined head shape.	1
30	1.0	The sensitivity for mounting errors for the model obtained for requirement 29 should be investigated.	1
31	1.0	The relative position of the hearing aid devices to the glasses is calculated, taking into account the head size of the user.	3

5.2 Sound source tracking and amplification module

The sound source tracking module will be used to track a sound by using the hearing aid devices and the G3 glasses. The requirements are listed in the table below.

Requirement	Version	Description	Priority
32	1.1	Computations should be able to be done on recorded data.	2
33	1.0	Computations should be able to be done on live data.	1
34	1.1	Using the eye-tracking module (see section 4.3) the hearing aids sound source position estimate should be fused with the gaze with the user staring at the sound source.	1
35	1.1	Integrate a beamforming module to be able to amplify sound, in the gaze direction with a point noise source at least 60 degrees from the gaze. The beamformer should amplify the sound source in the gaze direction compared to the noise, in the case that:	1
35A	1.0	The sound source and the user are static.	1
35B	1.1	The sound source is moving and the user is static.	2
35C	1.1	The sound source is static and the user is moving.	1
35D	1.0	The sound source and the user are moving.	2
36	1.0	Be able to track a sound source at a maximum distance of 5 m from the user with an maximum error distance of 0.5 m based on measurements from speech detector (see section 3.2 and 3.3) and align it with visual input from the camera.	2

6 UI

Requirements for the UI are listed below.

Requirement	Version	Description	Priority
37	1.1	Be able to turn on or off the amplification of the sound in the hearing aid.	1
38	1.1	Be able to display the intensity of the amplifying sound in the hearing aid.	1
39	1.1	Be able to change the intensity of the amplifying sound in the hearing aid.	2
40	1.1	Be able to record the sound source for at least 5 seconds from the hearing aid before the sound is amplified.	1
41	1.1	Be able to record the data for at least 10 seconds from the Tobii glasses.	1
42	1.1	Be able to listen to the recorded sound source from the hearing aid before the sound is amplified from the computer.	2
43	1.1	Be able to record the sound source from the hearing aid after the sound is amplified.	1
44	1.1	Be able to listen to the recorded sound source from the hearing aid after the sound is amplified from the computer.	2
45	1.1	Be able to display the user's position and the speaker position relative to the user.	3
46	1.0	Be able to display the ground truth position of the user and the speaker relative to the user.	3
47	1.1	Be able to store input parameters to all algorithms used.	2
48	1.1	Be able to store the output of processing, for example control signal and state estimates.	2

7 PERFORMANCE REQUIREMENTS

This section describes performance requirements on the system.

Requirement	Version	Description	Priority
49	1.1	The delay estimate received by adding delays of the modules from the hearing aid input to the audio output should not exceed 50 ms.	1
50	1.0	The maximal sampling frequency of the hearing aid devices should be higher or equal to 8000 Hz.	2
51	1.0	The processing frequency of the Tobii Pro Glasses 3 data should be at least 5 Hz for online processing.	2

8 POSSIBILITIES TO UPGRADE

Requirements enabling future development are listed in this section.

Requirement	Version	Description	Priority
52	1.0	The project should be well documented with explanations of all modules according to the LIPS-model.	1
53	1.0	A User Manual will be created explaining how to use the system.	1
54	1.0	All code is stored on a GitLab repository.	1
55	1.0	The project should be divided in modules for easy expansion or replacement of modules.	1

9 ECONOMY

The economy requirements are listed below.

Requirement	Version	Description	Priority
56	1.0	Each group member should spend 240 hours on the project.	1
57	1.0	ISY and Oticon should provide up to 40 hours of guidance.	1
58	1.0	One project room provided by ISY should be available for the project.	1

10 SAFETY AND SECURITY REQUIREMENTS

The safety and security requirements are listed below. These requirements are essential in order to ensure the safety of each group member and to minimize the risk of property damage.

Requirement	Version	Description	Priority
59	1.0	The responsible for testing should approve all tests before execution.	1
60	1.0	All members of the project group who will attend the test should read and understand the test plan before the test is performed.	1
61	1.0	The safety rules of Visionen will always be followed.	1
62	1.0	The audio output volume of the hearing aid device should be controllable by the user. Note that the user must control the output audio to a non-damaging volume.	1

11 DELIVERY

The delivery requirements are listed below.

Requirement	Version	Description	Priority
63	1.0	A status report including a time report should be delivered to the orderer weekly.	1
64	1.0	DP2: Delivery of the requirement specification, project plan, time plan and a first draft of the design specification.	1
65	1.0	DP3: Delivery of the design specification and test plan	1
66	1.0	DP4: Mid project check-up	2
67	1.0	DP5: The entire system should be functional. Delivery of test protocols and user guide including a presentation showing that the requirements from the requirement specification have been fulfilled.	1
68	1.0	DP6: Delivery of the technical report, an after study, a poster presentation, the project web page and project movie.	1

12 DOCUMENTATION

Table 17 lists all documents that shall be produced in the project.

Table 17: Documents to be produced.

Document	Language	Aim	Target	Format
Project plan	English	Describes how the project should be conducted and the responsibility of the group members.	Project group, orderer, course responsible	PDF
Time plan	English	Describes how the time on the project should be spent and also how the time should be distributed in the group.	Project group, orderer	Excel spreadsheet
Requirement specification	English	Determination of the requirements imposed on the project and the product.	Project group, orderer	PDF
Design specification	English	Describes how the system should be desgined for to meet the requirements.	Project group, orderer.	PDF
Test plan	English	The test plan describes how the requirements are to be tested.	Project group, orderer	PDF
Technical documentation	English	A collection of all the conducted tests.	Project group	PDF
Project report	English	A report summarizing the whole project and its outcome.	Project group, orderer, course administrators	PDF
After study	English	A reflection of the project as a whole.	Course administrators	PDF
User manual	English	Description of how to use the system.	Course administrators	PDF
Wiki	English	Description of how to use the system regulary updated throughout the project.	Course administrators	Git

13 QUALITY

The following requirements are to be met in order to ensure the quality of the finished product.

Requirement	Version	Description	Priority
69	1.0	Technical documentation following academic standards must be maintained throughout the project.	1
70	1.0	Code must be commented and follow Google code standards [3].	1

REFERENCES

[1] "Tobii official website," https://www.tobii.com/, [Online; accessed September 13, 2023].

[2] T. AB, "Tobii pro glasses 3," https://corporate.tobii.com/sv/media/mediabank, [Online; accessed September 19, 2023].

[3] "Google style guide," https://google.github.io/styleguide/, [Online].