Surround sound tuning based on location

Angeliki Angeletou

Technische Universiteit Eindhoven
University of Technology

Abstract

Surround Sound Systems are designed to faithfully reproduce a movie soundtrack and provide an immersive sound experience to the user, when the loudspeaker placement and the listening position comply with certain standards. Deviation from these standards can cause a distorted or unnatural sound image. In order to address the impact of the loudspeaker setup on the listening experience, two Philips Surround Sound Systems were integrated with Ultra Wideband (UWB) localization technology for the design of two demonstrators. Each demonstrator is composed of two parts; a localization module, providing user's and speakers' coordinates and audio tuning algorithms which use this information for the calculations of the signals that are sent to the speakers. In the present study, the performance evaluation of each component as well as the implementation of the demonstrators are described and discussed.

Executive Summary

In the frame of EUWB project, two Philips Surround Sound Systems, the Philips Ambisound Soundbar and a 5.1 Home Theatre Sound System, were used in the design of two demonstrators, developed to illustrate the use of Ultra Wideband (UWB) Radio Technology for home audio applications. The main goal of the demonstrators is to improve the sound perceived by the listener, when the placement of the loudspeakers in the listening room deviates from the ideal setup.

Although there are available guidelines and standards for an optimal loudspeaker placement, they are generally not followed in real living rooms. As a result, the listener perceives a sound image that is sub-optimal or even might be distorted or unnatural. In order to improve the listening experience, the demonstrators employ localization hardware and software to obtain the loudspeakers' and the listener's positions and apply this information in the calculations of the signals that are sent to the speakers.

The demonstrators are therefore composed of two parts; localization and audio tuning platforms. As the main interest of the project was in the qualities and the limitations of each component, localization and audio tuning were evaluated separately, with tests designed specifically for the specific demonstrator. The results indicate that although there are certain limitations in UWB localization and some acoustic parameters that should be taken into account in audio tuning, the users can measurably benefit from the integration of UWB localization in Surround Sound Systems. The implementation as well as the results of the evaluation tests are further described and discussed in the present document.

Contents

List of Figures

1 Introduction

The surround sound systems are originally designed to deliver an immersive listening experience to the user under certain conditions. One of the most important conditions is the placement of the speakers; the user is required to place the speakers in particular positions defined by standards, such as the ITU standard (ITU-R, 1997) and apply rules of thumb (such as avoiding placement of speakers close to walls or corners). Second, the position of the user plays also a crucial role; the ideal position is usually at the centre of the speakers' setup and it is called the sweetspot. In real households, however, it has been observed that the speaker setup and the listening position of the user are rarely ideal. Walls, furniture, cables and other natural obstacles usually determine the position of the speakers and the user, thus creating a soundscape that does not make use of the product's full capacity.

In order to compensate for this distorted sound image caused by misplacement of speakers and room acoustics, several Surround Sound Systems come with calibration DVDs with test tones and a microphone. The microphone 'listens' the test tone reproduced from all the speakers in the room and sets their levels according to their distance from the microphone (which is regarded as the listening position). Some of them are capable of measuring acoustical parameters of the room and provide equalization for them too. However, the calibration is a time-consuming manual procedure that is usually performed just once, when the speakers are installed for first time. If the speaker setup is slightly changed or the listening position does not correspond to the position of the microphone during the calibration, the complete procedure needs to be repeated.

The present study addresses the surround sound calibration from a different perspective; no manual procedure is needed for the audio tuning, but the system obtains real-time location information of the setup and manipulates accordingly the channels sent to the speakers. For the location tracking of the speakers and the user, Ultra Wideband Radio Technology (UWB-RT) is utilized. UWB-RT is a wireless technology that enables real-time location tracking and short-range wireless communications with high data rates (EUWB, n.d.). Its main advantages are low energy consumption and minimum interference to other electronic equipment compared to existing radio technologies. The EUWB European Project, in the frame of which this study is conducted, focuses on the use of UWB-RT in innovative applications in the domains of home entertainment, automotive, public transport and mobile communications. These applications are studied and developed during EUWB Project by several European industrial partners, including Philips, the carrier of the present study.

In the beginning of EUWB Project, Philips developed use-case scenarios for home entertainment audio applications, combining surround sound audio tuning and loudspeaker and user localization (van Daele & Hafezi, 2008). The scenarios, described in the deliverable D8c-6, were based on the idea that a Surround Sound System could potentially benefit from the addition of UWB localization, since loudspeaker placement is critical for surround sound and UWB is proven to be an efficient, low energy location-tracking technology. Some of the EUWB industrial partners and Universities provided Philips with UWB location-tracking hardware and software in order to be used in combination with surround sound tuning algorithms, developed by Philips. The final outcome of this collaboration is described in this study and includes the integration and the evaluation of two demonstrators; the first demonstrator is based on Philips Ambisound Soundbar and the second utilized a 5.1 Home Theatre Sound System.

Ambisound Soundbar (Figure 1) is a novel Philips sound system which delivers the surround audio experience from a single bar of speakers. The Soundbar has the advantage over normal 5.1 surround sound systems of taking less space, using fewer cables and fitting under any TV set without requiring any complicated setup procedure. The Philips Ambisound audio processing is based on the notch effect; that means that no sound is directly steered towards the assumed listening position, the sweetspot. The sound is steered through the speaker dipoles to the walls and the reflections create the impression that the sound coming from around the user and not from the Soundbar. Ambisound audio processing has been proven in various listening tests to be effective in delivering an immersive surround listening experience to the users.

The Ambisound with location tracking demonstrator, simply mentioned in the rest of the document as Ambisound demonstrator aims to provide the best possible audio to the listener by steering the notch towards her/his position inside the living room. The implementation of this scenario required a localizer that will track the position of the user as well as audio processing that uses this information to adjust the speakers' output accordingly. University of Ilmenau has developed a UWB localizer that provides location coordinates of a person in front of its antenna array. A real time Ambisound audio processing simulator has been developed to use these coordinates to update the gains and the delays for the speaker channels in order to steer the notch towards the user.

The key objective of the 5.1 Home Theatre Sound System Demonstrator is to correct for the unnatural sound image caused by the misplacement of the speakers and the listening position by manipulating the audio channels that are sent to the speakers. For the loudspeaker installation of a 5.1 Home Theatre Sound System setup, the ITU standard is generally recommended in the product instruction sheets. The ITU standard is a symmetric circular setup with the ideal listening position set in the centre (see Figure 2); however it might not suggest the ideal placement for every user or every home; the ideal is defined by the room acoustics and the preference of the user. On the other hand, a completely asymmetrical or a wrong placement of the speakers and the user can cause an audible imbalance in the sound perceived by the user. For example, it is known that proximity to a speaker defines the Sound Pressure Level (SPL) our ear perceives and thus, if a user is very close to a particular speaker, her/his listening experience might be dominated by its channel, which is not as originally intended by the movie sound engineer. Therefore, unlike Ambisound demonstrator where only user's position is required, for the implementation of the 5.1 Home Theatre Sound System demonstrator the positions of each of the speakers and the user are needed.

It should be noted here that the Subwoofer, although it is distributed in the market together with any 5.1 Home Theatre Sound System and with Ambisound Soundbar as part of the product, will not be often referred to in this document, as the loudspeaker's performance is considered more crucial for the sound quality. One reason is that the Subwoofer's sound propagation is considered omnidirectional, namely the sound will be distributed almost equally to all directions. Moreover, the Subwoofer's performance is not influenced or related to the placement of the speakers and the user, but it is rather affected by the room acoustics, such as room dimensions, adjacent surfaces and other obstacles.

The performance of the combined localization and audio tuning demonstrators is a function of both the localization and the audio tuning performance. The performance of localization and audio tuning was evaluated separately for both scenarios, so that a detailed picture of each system could be obtained before the integration. In order to evaluate the effect of the location-based audio tuning on the sound quality and the user's listening experience, listening experiments with sound experts and normal users were conducted. Evaluations of UWB localization for the particular use case scenarios were also performed.

The 1st chapter of the present document provides an introduction to the topic of location-based surround sound tuning. In the 2nd chapter, the principles of the audio tuning algorithms are described. The Ambisound demonstrator evaluation is described in chapter 3, and the 5.1 Home Theatre Sound System Demonstrator follows in chapter 4. For each demonstrator, the integration of the system as well as the evaluation requirements, methods and results are presented.

Figure 1 - Ambisound Soundbar in a living room.

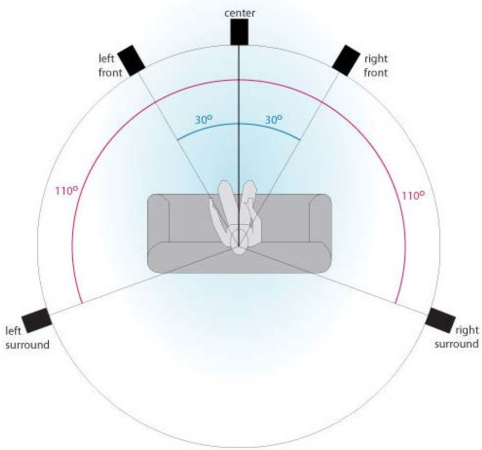

Figure 2 - ITU standard for the placement of speakers.

2 Background

As it was described in the introduction, we used two Philips surround sound systems, 5.1 Home Theatre Sound System and Ambisound Soundbar combined with localization hardware and algorithms in order to correct the reproduced sound image coming from non-optimal loudspeaker setup and/or listening position. For this purpose, two audio tuning algorithms - one for each system - were developed and tested. In the following sections, the basic principles behind the audio tuning algorithms of Ambisound Soundbar and the 5.1 Home Theater Sound System are described.

2.1 Ambisound Soundbar: the notch effect and audio tuning principles

The Ambisound technology is used in Philips Soundbar products, with the aim of creating wide and embracing cinema sound from a single bar, containing 6 loudspeakers in predetermined positions and angles. Out of the 6 speakers, 2 have the centre channel (the 2 inner speakers of both sides) and the other 4 outer speaker drivers have the right and the left channels correspondingly on each side.

In each side, the two outer speaker drivers propagate sound in a lobe shape and in anti-phase, thus creating a dipole. In the lobes' intersection, cancellation of the signals in anti-phase occurs and an area with low SPL is created. Ideally, the SPL would be zero in an anechoic room for particular frequency bands, but in real rooms the SPL is practically some dB s lower in this area than in the rest of the room. This area is called the 'notch' and it is the main effect on which Ambisound Soundbar is based. By delaying one of the two drivers in each dipole, each notch can be steered towards any position within boundaries. In the available in the market versions of the product, the notch is steered towards the position (0, 2.5) in the coordinate system (Figure 3) and therefore the sweetspot, namely the ideal listening position, is located there. This way, most sound energy is distributed around the sweetspot and reflected in the room, thereby creating the impression of embracing sound. The sound pressure level distribution in an anechoic room, at different frequencies is shown in Figure 5, where the lobe-shaped sound propagation of the speakers in the dipole is visible.

The notch width (x), or better, notch aperture (α) (Figure 4) is a function of the distance between the two drivers and frequency. It can be defined as the distance between two points in space where the attenuation at a certain frequency is equal to a defined threshold. The actual distance will also depend on the listening distance from the dipole (D). For a given speaker spacing, the notch becomes smaller with higher frequencies (Figure 5).

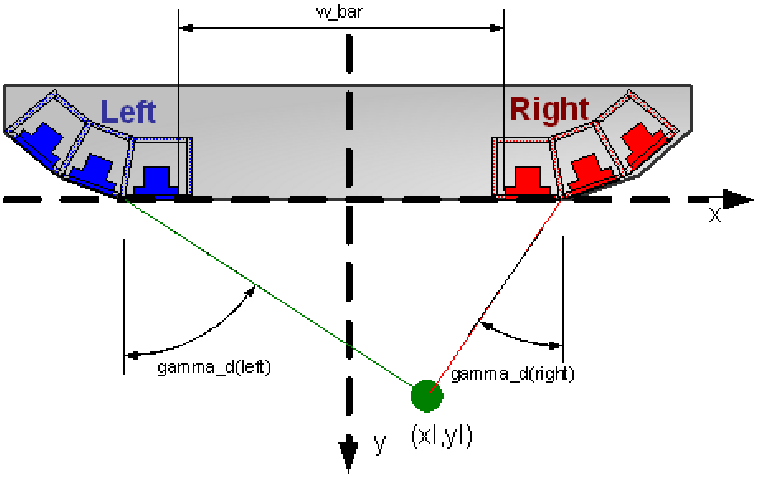

Figure 3 - Top view of speaker driver arrangement in an Ambisound Soundbar. The listening position is represented with the green dot.

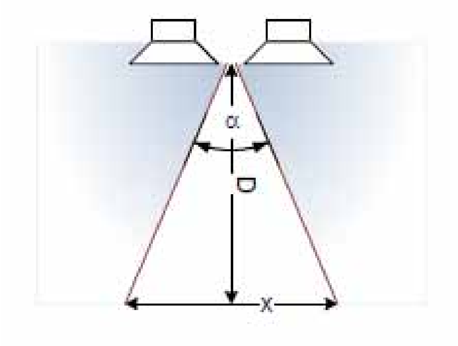

Figure 4 - Notch aperture or width.

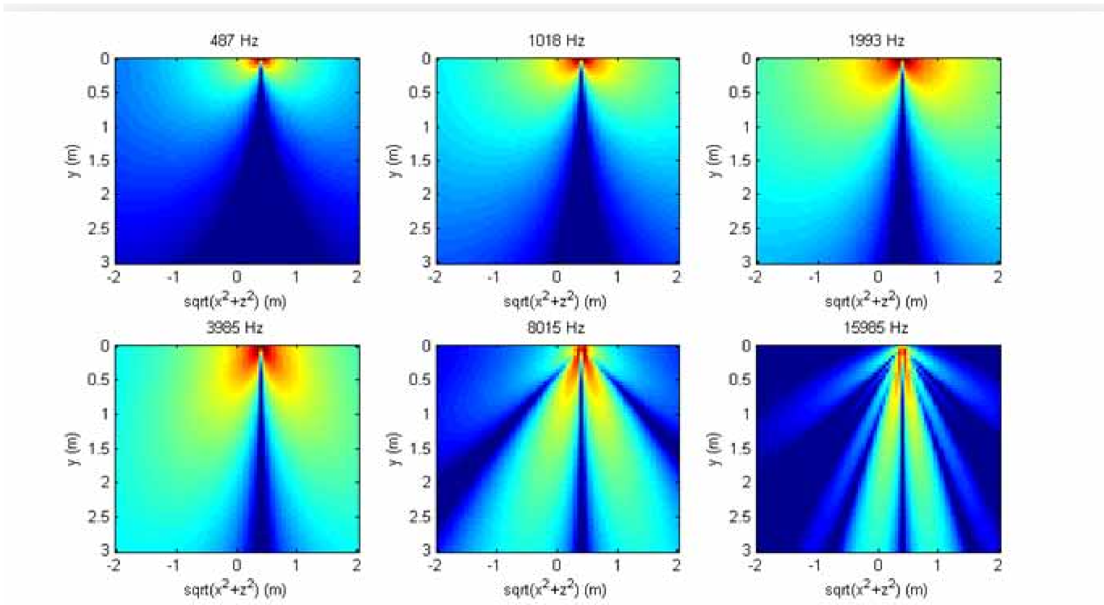

Figure 5 - Sound pressure level distribution in a 3x4m ideal anechoic room, one single dipole playing, at different frequencies. The warm tones (red, yellow) indicate high SPL and the colder tones (blue) low SPL. The notch can be observed as a blue beam between the lobes of the dipole.

2.1.1 Initial angle calculation

The main goal of the Ambisound audio tuning algorithm is to allow the steering of both notches towards the user's position, as it is estimated by the UWB localization module. The rotation of each notch to the desired angle requires delay calculations for each channel, but the gains should remain exactly the same for each outer speaker driver, otherwise the notch that is based on opposite signal cancellation cannot exist.

In Figure 6, the two speakers forming the dipole pair are shown (thick lines) with some indication as to how to calculate the speaker coordinates and initial notch angle γ (gamma).

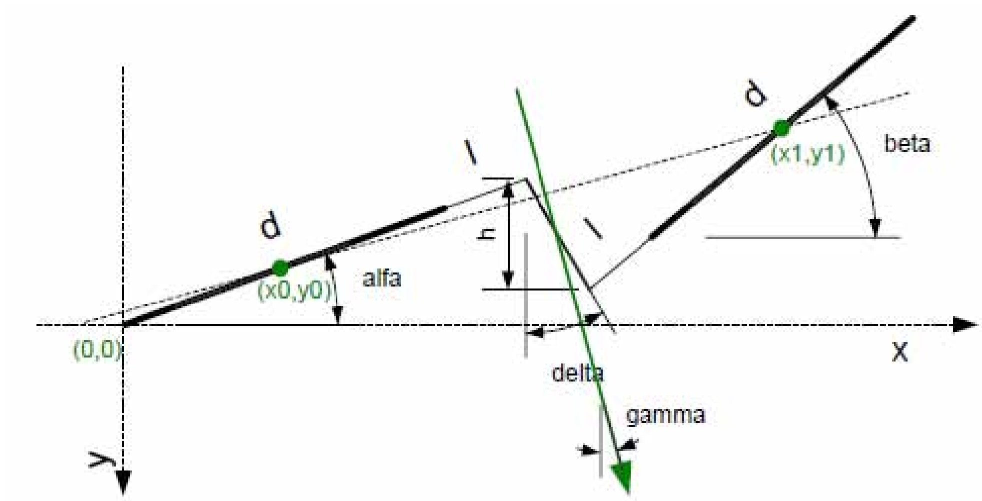

Figure 6 - Representation of the dipole speakers' positions and the initial notch angle (gamma).

The following parameters are defined:

- *d*: driver diameter
- α (alfa): inner driver angle
- β(beta): outer driver angle
- *h*: inline placement ratio (0 = no inline placement, 1 = outer driver is shifted completely towards the front)
- δ (delta): inline placement angle
- *l*: average distance between driver edge and edge between drivers

The driver's coordinates are considered to be the coordinates of the centre of the driver. Calculating the drivers' coordinates and then the distance w between them, the initial notch angle γ can be derived:

$$\gamma = \cos^{-1}\left(\frac{x_1 - x_0}{w}\right) + \varphi_{lPF}$$

In this formula φ_{LPF} is the average in-band phase shift introduced by the low-pass filter on the inner driver.

2.1.2 Desired rotation

The desired rotation of the notch ρ_d of one dipole is the result of subtracting the desired notch angle from the initial (current) notch angle. The desired angles for both right and left dipoles, $\gamma_{d,r}$ and $\gamma_{d,l}$ can be calculated, given the coordinates of the listening position (x_l, y_l), referenced to the centre and the front of the Soundbar, and the distance w_{bar} equal to the distance between the innermost edges of the inner drivers of the dipoles. Assuming that the listening position is always in the far-field of the dipole speakers, the following formula can be used to calculate the desired notch angle for the left and for the right dipole correspondingly. The distance $d_{l,l}$ is the distance between the listening position and the left speaker box, and $d_{l,r}$ the distance between the listening position and the right speaker box.

$$\gamma_{d,l} = \sin^{-1}\left(\frac{\frac{w_{bar}}{2} + x_l}{d_{l,l}}\right)$$

$$\gamma_{d,r} = \sin^{-1}\left(-\frac{\frac{w_{bar}}{2} - x_l}{d_{l,r}}\right)$$

From these angles, the desired rotation ρ_d easily follows.

$$\rho_d = \gamma - \gamma_d$$

The delay with which each channel will be sent to the speaker is then calculated as a function of the sampling frequency f_s, the speed of sound (340m/s) and the required rotation ρ_d.

2.2 5.1 Home Theatre Sound System: audio tuning principles

The audio tuning for 5.1 Home Theater Sound System aims at correcting the sound produced by asymmetric loudspeaker setup by adjusting the signals that arrive to the speakers. The audio tuning algorithm is also mentioned as repanning algorithm, as it re-distributes the channels. The goal of the repanning algorithm is that the sound perceived by the user – and in particular the direction of the sound sources - is as close as possible to the intended sound image, as it would be heard from an ideal symmetric setup. For this purpose, the algorithm creates phantom sound sources that are heard as if they are coming from predetermined positions, called virtual loudspeakers. As it is shown on Figure 7, the repanning algorithm constructs a matrix with delays and gain coefficients. The gain coefficients define the participation of each of the 5 channels (L, R, SL, SR, C) to the signal that arrives to each speaker.

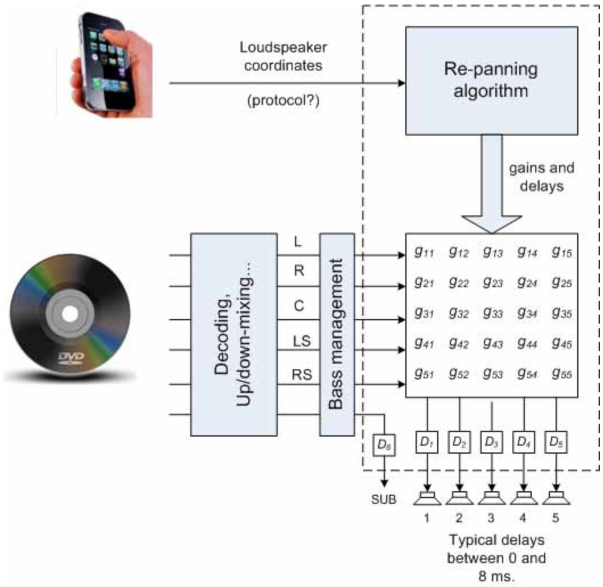

Figure 7 - How the repanning algorithm works.

Gerzon (1992) describes the basic theory of auditory localization, based on the velocity and the energy models. These models allow a phantom sound source direction to be expressed as a weighted sum of the vectors pointing in the direction of the physical loudspeakers. In (van Leest, 2005), the algorithm that calculates the gain coefficients for each speaker utilizing the velocity and energy models are utilized is presented.

The loudspeaker setup consists of S physical loudspeakers (in our case, S=5). The Loudspeaker coordinates are : $L = \begin{vmatrix} x_0 & x_1 & \dots & x_{s-1} \\ y_0 & y_1 & \dots & y_{s-1} \end{vmatrix}$. A virtual loudspeaker can be described by the vector $\underline{p}$ containing its coordinates. Figure 8 shows the virtual and physical loudspeakers. The outcome of the algorithm is the gain vector $\underline{g}$ such that the velocity vector $\underline{v} = L \cdot \underline{g}$ and the energy vector $\underline{e} = L \cdot \underline{g}^{(2)}$ are pointing as much as possible in the same direction as $\underline{p}$. The column vectors $\underline{g}$ are combined in a gain matrix G and adjusted according to each loudspeaker's distance from the listener.

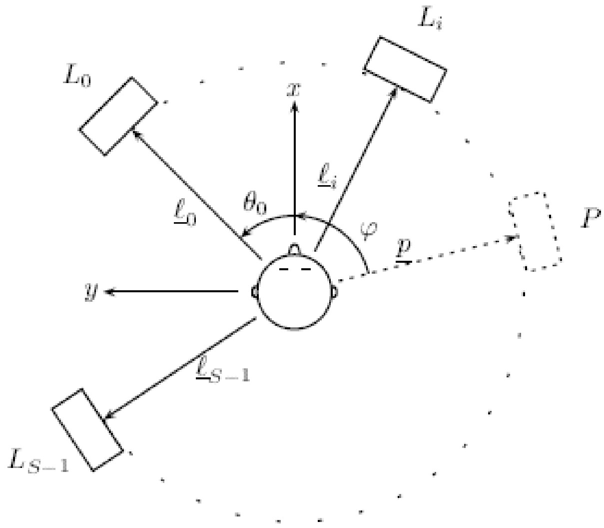

Figure 8 - Diagram of the setup of the speakers and the user. The speakers have the same but different weighted signal and thus the user perceives the sound coming from the phantom sound source direction P.

In order to calculate the delays for each speaker, the maximum loudspeaker distance $d_{\max}$ from the listener is found. Then, for each loudspeaker , the delay D_i (in samples) is given by the following formula.

$$D_i = \left(\frac{d_i - d_{\max}}{343} \right) \cdot f_s$$

where f_s the sampling frequency.

3 Ambisound with location tracking demonstrator

3.1 Integration of the demonstrator

As it was mentioned in the previous section, the Ambisound demonstrator consists mainly of two parts:

- the UWB localizer and
- the Ambisound audio tuning.

The UWB localizer, described in EUWB deliverable (Alvarez et al., 2010) was developed by the University of Ilmenau (UIL). It is comprised of a set of three antennas (two receivers and one transmitter) and a module (dimensions approximately 0.40 x 0.30 x 0.50 in m) that calculates the location coordinates from the data received from the antennas. The module can be used with different sets of antennas – for our use-case, a set of copper thin antennas were utilized as they could be easily mounted on Soundbar chassis.

The UWB transmitter antenna transmits a stimulation signal at a certain measurement rate and this is reflected from static as well as moving objects. In order to separate static from time-variant reflections, a radar signal processing technique is used, known as background subtraction. A simple background subtraction approach requires the averaging of the sequence of all measured reflections. This way, the static background reflections are estimated and can be subtracted. After the background subtraction, Time-of-Arrival (ToA) is estimated. ToA is the time that is required for a transmitted impulse to be reflected on a moving and/or time-variant object and arrive to the receiver antenna. The localization algorithm that follows fuses range estimates from multiple channels receivers. In the arrangement of the UIL UWB localizer with three antennas, the location of the moving target is computed by solving two quadratic equations.

The localizer is capable of detecting one only moving target each time. The user is therefore detected by his movement in space (even the slightest movement, such as breathing) provided that there is no other object moving in the room. Figure 10 shows a photo of the UIL UWB localiser.

In the Background section, the Ambisound audio tuning principles are described. The algorithm, as well as the rest of the audio processing required (volume levellers, equalizers, headroom compensation etc), was implemented in Matlab, using the Simulink platform for the modelling of the processing blocks. The Simulink model is controlled from the Graphical User Interface (GUI). The Advanced Settings panel of the GUI is shown in Figure 11. On the panel, the *Listening Position* was mainly used in order to try out steerings of the notch towards different locations. The two sliders in this field (Distance, Offset) are used to indicate the listening position in steps of 0.5m; distance is the distance between the listening position and the Soundbar and offset is the distance between the central driver of Soundbar and its parallel driver along the listening position. The values in these fields are then used to calculate the desired notch angle and the corresponding delays for the front and surround beamers. Since the audio tuning model receives as input listening position coordinates and not notch angles, the method and the results of the listening tests that follow refer to listening positions and not to notch angles.

Regarding the integration of the two systems, the localizer module is connected through Ethernet with a PC and sends the location coordinates of the user to a predefined serial port. The serial port, the measurement rate and other parameters can be adjusted on the GUI of a software application, written in CVI9.0 National Instruments. The application plots depict in real time the signal received from the antennas and the estimated location coordinates (see Figures 14-19).

The Simulink audio processing model receives audio input from a CD/DVD player and the location coordinates from the serial port and translates them in listening distance and offset. With this information, the model calculates the desired rotation and adjusts the delays of the dipole channels in the Soundbar so that the notch is steered towards the estimated location. The model's output, 6 c hannels with their adjusted delays, are then transferred to the soundcard and then to the amplifier. Finally the amplifier is connected to a 'dummy' (without audio processing, only with speakers) Ambisound model. Figure 9 depicts the complete demonstrator setup.

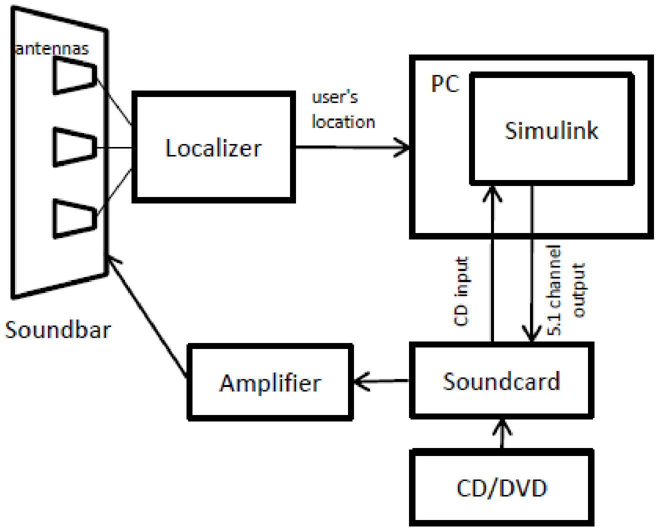

Figure 9 - The block diagram of the demonstrator.

Figure 10 - UWB localizer.

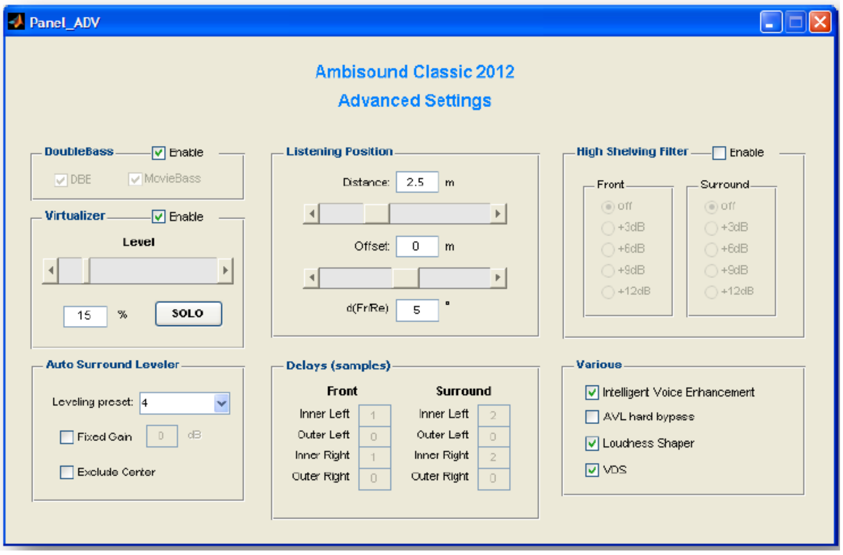

Figure 11 - GUI of audio tuning Simulink model for Ambisound Demonstrator.

3.2 Evaluation requirements

In the following sections, the evaluation methods and results of the UWB location tracking system and the Ambisound audio tuning model will be separately presented and discussed. The goal of the evaluation of the location tracking system is to identify its limitations in regards to the particular use-case and to find out whether and in what ways these limitations affect the performance of the surround sound tuning system. Accordingly, the audio tuning evaluation aims to determine whether the audio tuning provides a clear audible improvement to the sound and define this improvement (or this audible effect) as well as identify the parameters that can affect the audio tuning performance and therefore the listening experience.

In particular, in the evaluation tests conducted on the UWB localizer provided for this demonstrator, certain aspects of the behaviour of this particular system were investigated:

- The way of interaction with the system: passive continuous location tracking without user's intervention, or active location tracking with the user having control over initializing and stopping the location tracking.
- Is the location tracking of the user performed in real-time or are there delays involved.
- The detection area of the localizer: what is the maximum distance in which the localizer can detect.
- The error of the estimated location from the actual location.

All the above aspects and the limitations they impose define the context of use of the demonstrator. For example real-time location tracking would enable continuous location tracking of the user and real-time manipulation of audio. On the other hand, if the location tracking was delayed for some seconds, it would be preferred that the user activates her/himself the location tracking when he decides on he r/his listening position and waits until the audio tuning takes place. Moreover the detection area of the localizer determines the size of the listening space. If the detection area had the size of a tennis court, it could be possible to expand the demonstrator with extra speakers and appropriate audio tuning to send the sound to wherever the user might be inside a house, while a detection area with a size of a living room would fit the use-case of the

Ambisound demonstrator, as described above. Similarly, the localization error should be relatively small so that it can be 'absorbed' by the precision of the notch steering in the Ambisound audio tuning system.

Consequently, the audio tuning evaluation was designed to investigate the behaviour of the Ambisound audio tuning algorithm in the following aspects:

- Precision. What is the minimum notch steering that is required in order to create a clear audible effect? The notch angle is translated in lateral distance (offset) and distance of the user from the Soundbar in the audio tuning interface, therefore the precision is also measured in meters.
- The role of the two most important parameters, audio type and room acoustics in the performance of the algorithm.
- Users' preference. How is the notch steering perceived by the listeners.

At this point, it should be noted that the main goal of the demonstrator is to illustrate the use of UWB location tracking in Surround Sound Systems and therefore its performance and context of use is determined by the current status of the particular location tracking and audio tuning systems. Certainly, some of the limitations imposed by these systems will be addressed in future iterations of the demonstrator. However, the context of use of the final product should be defined mainly by the users' needs and requirements and secondly by technology constraints.

3.3 Location tracking evaluation

3.3.1 Experiment design

Before describing the experiments, it should be noted that the UIL UWB localizer allows for continuous location tracking; the system would update the location of the user automatically every few seconds without user's intervention. The location tracking can only be activated/deactivated by turning the UWB module on/off. While the system allows for passive localization, from some interviews with sound experts, conducted during the audio tuning evaluation study, active localization by means of a button on the remote control was preferred. Thus, the user can choose his\her listening position in the living room and initialize the localization by pressing a button. One reason was that in many situations, steering the notch towards the user would not be desirable, e.g. when the user picks up the phone. Therefore, the advantages of active localization is that the location tracking does not need to follow the user real time and no smart system is required to decide whether localization is meaningful for the user at a particular moment, as the user takes the decision. The selection of the interaction method was also determined by the limitations of the localizer and the Ambisound sound processing; the location tracking was shown to impose a delay of some seconds, as it will be described in the experiments following. Moreover, the update of the localization causes audible interruptions in the audio tuning model in Matlab. Therefore it was decided that active localization would be applied instead: for the purpose of the demonstrator, the user will activate the localization with a button after (s)he has chosen her/his listening position and the audio tuning model will adjust the sound according to the location coordinates retrieved ten seconds after the activation.

The main goal of the experiments that follow was to evaluate the UWB localizer according to the requirements described above and identify the technology constraints, focusing mostly on active localization. During the experiments, the UWB localizer was set in a big room with large windows with almost no furniture. The lack of furniture enabled exploration of the detection area of the localizer without obstacles and ensured that the algorithm would be tested with the minimum background noise, that is a result of the reflections of background objects. The area in front of the localizer was 6m x 6m. The two receiver antennas were separated by 90cm (the length of Soundbar) with the transmitter antenna placed in the middle (see Figure 12).

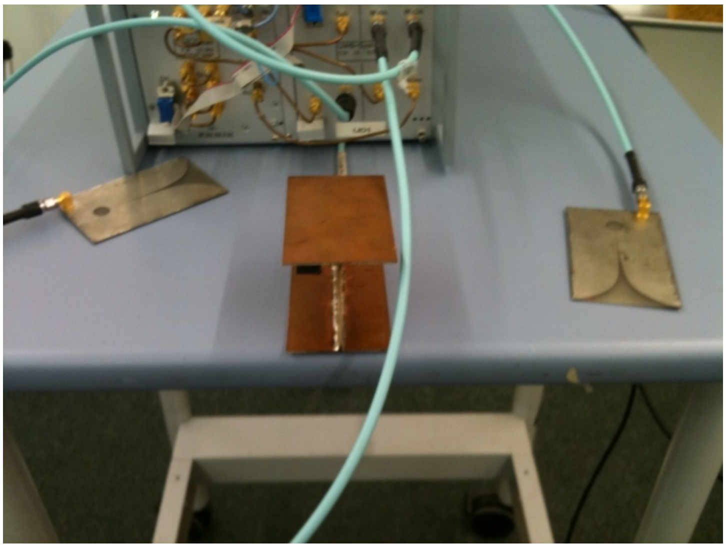

Figure 12 - The copper antennas used for the experiments. During the experiments, the two side antennas were mounted on the sides of a table.

A grid on the floor that consists of crosses placed every half meter in both directions (see Figure 13) was used for identifying the actual position of the experimenter. There was practically no noise, sounds or much movement inside or around the room.

Figure 13 - The experiment room. The white crosses on the floor are the grid on which the experimenter was moving.

A description of the experiments conducted to characterise the behaviour of the localizer, is provided below:

- **1st experiment: Behaviour in an empty space.**

  It is known that the UIL localizer detects movement in space, even when it is subtle (such as air movements). A human is detected by the movement in an empty space and the localizer cannot recognise between a human or an object moving. To investigate the movement detection, the experimenter stayed out of the area of detection (the area in front of the localizer) and records the movement detected by the localizer for 30 seconds. The area of detection is empty. See Figure 14.

- **2nd experiment: Exploring the detection area of the localizer.**

  The experimenter takes different positions on the grid and stays still (hands and head might be shaking but not intense movement) to each position for 20 seconds. The positions taken were (0, 1), (0,2), (0,3), (0,3.5), (0,4), (1, 1), (1,2), (1,3), (1,3.5), (0,4), (2, 1), (2,2), (2,3), (2,3.5), (2,4). See Figure 15.

- **3rd experiment: Investigating the deviation in the limits of the localizer.**

  The average distance from a TV-set is typically 3 - 3.5m.This distance is close to the limits of the localizer, so it was considered important to investigate how the deviation from the actual position increases with distance. The experimenter walks on straight lines parallel to the driver of the central antenna in offset of 0.5m, 1m and 1.5m, stays still at distance 3.5 meters for 15 seconds and returns in the same way too. The experimenter walks at a slow pace. See Figure 16.

- **4th experiment: Velocity of movement and movement tracking.**

  The experimenter performs zig zag on the grid at different pace (from slow to fast). See Figures 17-18.

- **5th experiment : Two subjects at the same time in the room.**

  Two subjects cross walking at a normal walking pace. See Figure 19.

The results in the following sections are observations retrieved by analyzing the screenshots saved from the experiments, as datalogs with position coordinates were not available.

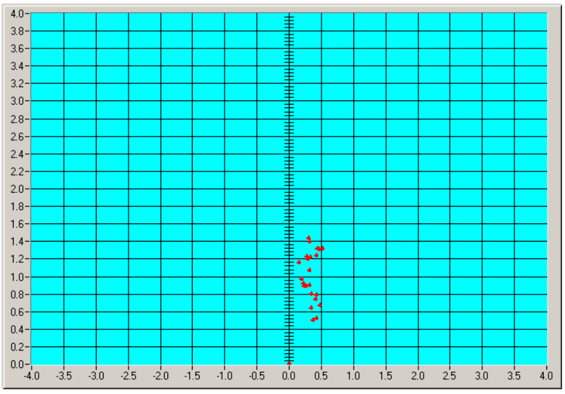

Figure 14 - 1st experiment. Movement detected in an empty space for 30sec.

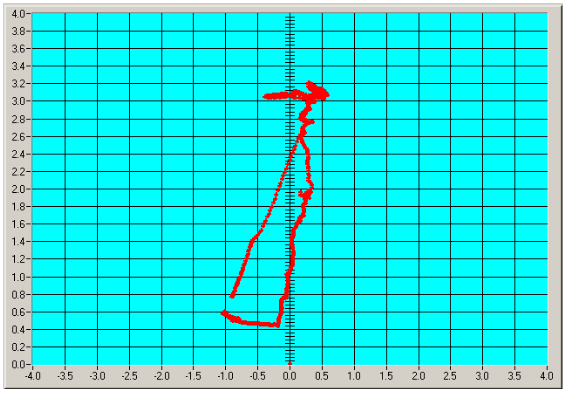

Figure 15 - 2nd experiment. The experimenter takes the position (0,3) and stays for 20sec still.

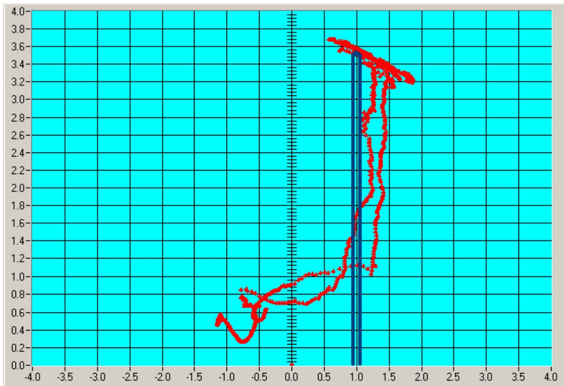

Figure 16 - 3rd experiment. Investigating the deviation in the limits of the localizer. The experimenter walks in a straight line with 1m offset from the antennas. The detected route (red line) converges slowly with the actual route (blue line). The experimenter waits still for 15sec in distance 3.5m.

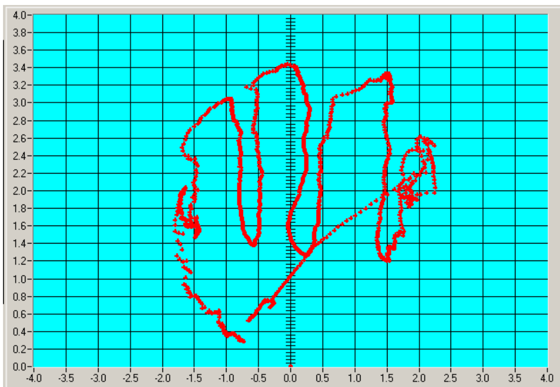

Figure 17 - 4th experiment. Velocity of the movement and location tracking. The user here performs a very slow zig-zag along each line of the grid.

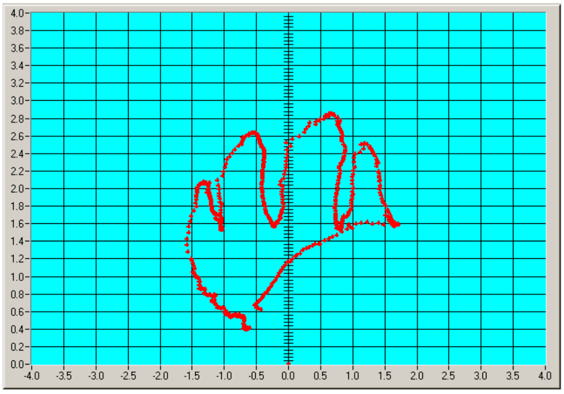

Figure 18 - 4th experiment. Zig zag performed at a normal walking pace.

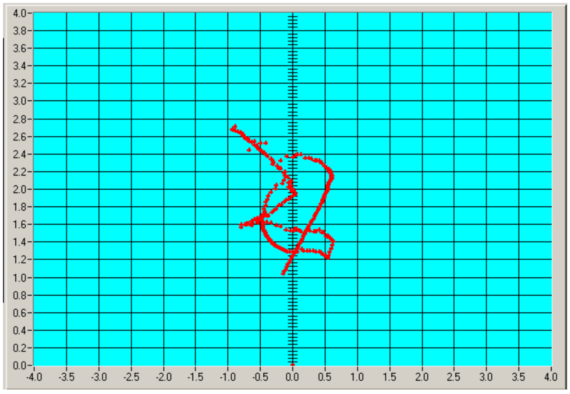

Figure 19 - 5th experiment. Two subjects start from (-1,3) and (1,3) and cross walking at a normal walking pace.

3.3.2 Results

The results are presented per experiment and final conclusions, drawn by the experiments follow.

- **1st experiment: Behaviour in an empty space.**

  In the 1st experiment, the localizer detected movement in the empty space that seems to occur as a result of noise generated by the electronics of the device. Moreover, the samples are averaged over time and space by the localizer, so the random noise appears as a route. For the Ambisound use-case, it would be desired that no movement is detected when no person is in the room. A possible explanation is that the threshold above which a signal is considered significant and therefore is considered as a moving object in the space, is too low.

- **2nd experiment: Exploring the detection area of the localizer.**

  It was observed that the localizer can identify the position of the experimenter with a reasonable error within a square of about 3.5m length and 3m width centered in the central transmitter antenna of the constellation (or in the centre of the Soundbar). The error of the localizer was observed to increase as the distance between the antenna and the experimenter increases.

- **3rd experiment: Investigating the deviation in the limits of the localizer.**

  Deviation between the actual position and the detected position increases significantly above 2.5m - 3m. Moreover the standard deviation of the detected position, which is depicted on the interface as the red mark created when the localizer continues to update the position of the experimenter when the target remains still, increases with distance: for 3.5m distance from the antennas it has dimensions of approximately 1m x 0.40m.

- **4th experiment: Velocity of movement and movement tracking.**

  When a person performs a route, the route that will be drawn depends on the velocity of the move. So, at a normal pace, the route detected from the algorithm is much smaller of the actual route performed. This is happening not only due to constraints of the detection area, but also because the localizer is occupied with averaging which is time-consuming and misses to update the location estimates as often as it is required. Moreover, averaging contributes to the deviation from the actual positions since it is affected by noise.

- **5th experiment : Two subjects at the same time in the room.**

  As expected, the route detected occurred as an average of movements of both subjects, as a result of the averaging algorithm.

In general, the averaging algorithm, although it increases the robustness of the estimated positions, causes delays and deviations from the actual positions and routes. A more sophisticated algorithm combined with a higher threshold for movement could improve the performance of the localizer. Moreover, the averaging of the routes of the two subjects (experiment 5) could be avoided if the algorithm recognized the distance between the detected movements and assumes that they come from different moving objects. Deviation can also be improved by separating the antennas much more apart, e.g. 3-4m, but then the physical integration of the system (antennas mounted on the Soundbar chassis) is not feasible.

3.4 Ambisound audio tuning evaluation

First, a pilot test took place and then the main listening test was designed, followed by a complementary test, where stereo music was used. The purpose of the pilot test was to address at first place some of the evaluation requirements and to collect observations that would be used to optimize the design of the main listening test. In the following section, the listening tests performed are described and the results are presented.

3.4.1 Pilot test

The pilot test took place in the room iLab5 of the Building A in Philips campus in Leuven. The room dimensions are 5.50 x 3.25 m^2 and there is a floor carpet covering the whole floor. Inside the room was a TV, a Soundbar, an amplifier, a soundcard and a desktop PC on a desk with a chair. The Soundbar was placed in the middle of the short wall and about 80cm away from it. The walls of the room are hard and reverbrative and part of the side walls are covered with absorptive foam panels. The opposite wall was also partly covered with foam. Thus, there is a variety of materials with different acoustic properties in the room as in a regular living room. In distance 2.5 meters from the Soundbar, there were two chairs; one placed exactly on the central driver of the Soundbar and the other in 1m offset from the first.

The participants were three sound engineers and a musician/usability expert. They were chosen because of availability and also because they have had training in sound and thus, they would be able to listen and describe the audible effect of the algorithm in much more detail. The experts were given a keyboard, a TV-remote and a paper form. With the TV remote, the users could choose between 5 short audio excerpts (about 30 seconds each) to listen. The excerpts were 5.1 audio recordings taken from movies and were selected so that they each had different audio content. The excerpts were 'Planet Earth', intro from a BBC documentary, with a tribal musical pattern, a scene from 'Twister' with sudden and abrupt sounds of explosions, 2 scenes from 'Casino Royale', one featuring simple dialogue and the other surround city sounds, and a scene from a Beyonce's concert, with music, dialogue and applauses.

Keys on the keyboard were assigned to different settings; each setting was steering the notch towards a particular position. The settings were organized in triads and each setting in the triad was offset by 0.5m (suggested theoretical precision) from the others on y- or x- axis. The settings are represented on the Figure 20 with dots. The users were asked to switch freely between the tracks and note down the differences they observed between a triad of settings. If they could identify a difference, they were asked to indicate in which sound attribute group the difference belongs and to rank the settings for each of the sound attributes. The sound attributes groups that were used for this test -namely timbre, details, spatial, sweetspot, overall (table 1) - are commonly used for the evaluation of the sound quality by experts in other listening tests performed in Philips. Moreover, the experts were also required to fill in which movie track was the most suitable for judging differences between the settings. Upon the completion of the paper form, the participants were interviewed regarding the concept of the sound following the user and the experiment in general and the experimenter was keeping notes.

During the test or in the interview, all participants mentioned that the differences observed between the different settings were subtle. They classified the observed differences between the settings in the attribute categories 'spatial', 'timbre' and 'details'. The tracks that were mostly preferred for tracking the differences were 'Beyonce' and 'Planet Earth'. To the rest of the tracks the differences were judged much more subtle.

The preference to the "Beyonce" and "Planet Earth" tracks can be explained as the audio content in these tracks was distributed in all channels creating a wide soundscape. On the other hand, the "Casino Royale" scene featured mainly dialogue that is played from the centre speaker and not any other significant background sounds, which are usually played by the front and the rear side speakers in a multispeaker sound system (or by the dipoles in Ambisound). The notch steering is based on the dipoles, the two pairs of outer speakers on the Soundbar. If these do not have an important role in the sound image, then it is also expected that the notch steering will not make an important different in the sound image perceived by the listener. The two other tracks – "Casino Royale" track with surround sounds and "Twister" – auditory reference is completely missing; the sounds are sudden and random and no repeatable pattern can be recognised. Therefore, the participants could not listen to the difference between the different steerings of the notch, not only because the difference is subtle, but also because they did not have any expectations regarding the sound in their minds to compare.

The interview questions addressed the concept of the location tracking in combination with sound, possible negative and positive aspects of it and how it could be applied. Three out of four participants mentioned that they like the concept of the sound following the user in the space in general. All four participants said that they want to be aware of the localization and agreed that a button on the remote control enabling location tracking would be a suitable way of giving the control to the user. This was the only time that experiment participants were asked questions regarding the context of use since interaction design was out of the scope of this project.

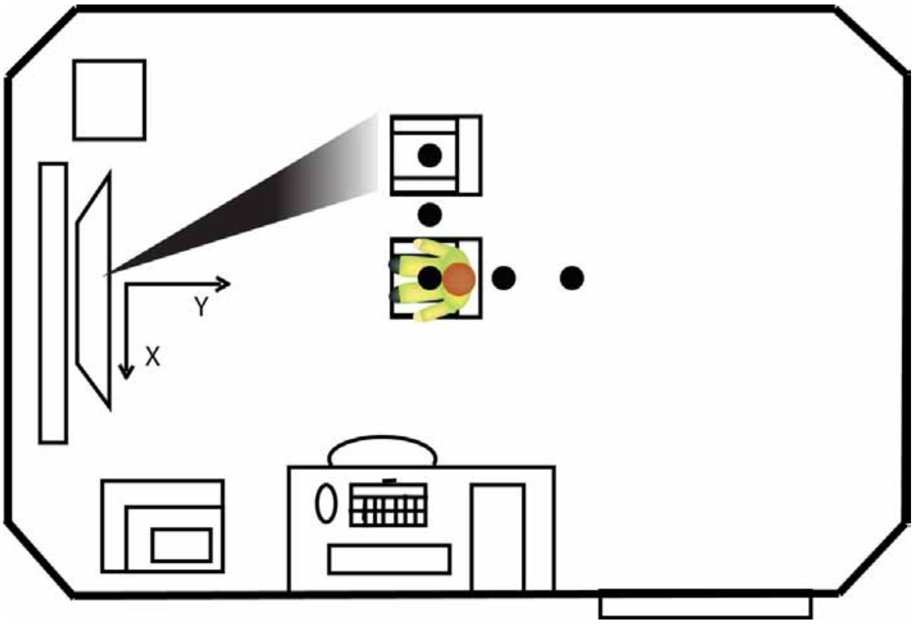

Figure 20 - Pilot test setup.The settings (represented by the black dots) are located 0.5m from each other.

Sound attributes and their sub-dimensions	
Timbre	preference, spectral balance, midrange dominance, boominess, nasalness, muffledness, harshness
Details	preference, voice intelligibility, instrument separation, high frequency extension, distortion and noises
Spatial	preference, width of soundstage, envelopment, embracement, spatial balance, depth, image precision, air
Sweetspot	sweetspot size, colouration, spatial distortion
Overall	preference

Table 1 - These attributes are currently in use for the internal listening tests with experts (also called "golden ears") in Philips.

3.4.2 Main listening test

The conclusions drawn from the pilot tests indicated that the precision of 0.5m is too fine to make an audible difference between the settings and that the experts found it extremely difficult to rank the triads of settings in the various sound attributes provided. Thus, it was decided to configure the settings for the main listening test in steps of 1m and to ask the participants only for preference between two settings each time. Since the audible effect of the notch steering was judged as subtle in most cases, it was decided also this time to use experts as participants.

Movie excerpt 'Beyonce' was preferred by the majority of the participants in the listening test. However, some users mentioned that this track was tiring to be heard over and over again and that the lack of a rhythmic pattern did not help the cognitive process of comparing two different settings. The second preferred movie track was 'Planet Earth'; a tribal simple rhythmic music with clear spatial features. Therefore 'Beyonce' was the main track that was provided to the participants and 'Planet Earth' was decided to be used as a control variable, only in one of the two positions, in order to investigate what is the influence of different types of audio on the results.

As implied by one of the research questions, the effect of the room acoustics on the listening experience was also to be tested. Two acoustically different rooms were chosen for the listening test. The first one, iLab5, is the room where the pilot test took place. The second one, called Pandorix, is a big acoustically certified room. Since the main difference between the rooms was the size, iLab5 will be referred to as the small room and Pandorix as the big room from now on. It was decided to use only two chairs (corresponding to users' positions) with distance 1m from each other, one in the centre and one to the side, in order to keep the setup exactly the same in iLab5 that is a rather narrow room.

Triads in x-axis seemed to have more audible differences than triads in y-axis, but it was also observed later by a sound expert that the soundscape becomes slightly wider or more closed when the configuration is on the y-axis. Therefore, the setup included 4 settings spread on the x-axis and 2 on the y-axis, to test this claim.

3.4.2.1 ***Method***

As it is shown in the figure 21, there are two chairs in distance 2.5m from the Ambisound Soundbar. The chairs are the listening positions from which the listeners are going to make their judgements. The listening position that has 0m lateral distance and 2.5m distance from the centre of Soundbar will be referred to as position 0 and the listening position with 1m lateral distance to the left of Soundbar's central driver will be referred to as position 1. In front of the chairs, there was a computer screen for the registering of the responses (not visible in the figure). Each time the user switched between the settings, the Soundbar steers the notch towards the corresponding position (one of the dots) and the audio file starts over. The settings are in the positions (0, 2.5), (-1,2.5), (1, 2.5), (2, 2.5), (0.5, 3.5), (0.5, 4.5). The setup is symmetric; every chair has two settings to one side, one to the other and two behind and was kept exactly the same in both rooms.

The participant came in, sat on a chair, was given a keyboard and instructions about the experiment. Figure 22 is a photo of the actual setup in the small room. Each time she/he could switch between two audio files of the same movie track (called A and B) configured for different positions. On the screen, the participant had to decide between 3 choices: A , B or "no difference"(see Figure 23). When the user decided, she/he just left the cursor there and presses the button 'Next'. The session was completed when 15 pairs (as many as the different pairs that are formed with 6 settings) of audio files were presented and a decision had been made for each pair by the participant.

The participants were asked to complete 3 sessions: 2 in the position 0 (1 with the track 'Planet Earth' and 1 with the track 'Beyonce') and 1 in position 1 with the track 'Beyonce'. 'Planet Earth' session was in all tests the second in order, so that the users had a break from the repetitive auditions of 'Beyonce' track. 'Planet Earth' was used as a control variable, only in position 0, in order to compare if there were differences between the choices made with 'Beyonce' and those made with 'Planet Earth'. It was decided not to use 'Planet Earth' also for position 1 as the test was already long (on average 40 minutes) and the test required considerable cognitive effort for judging the differences between the settings. In total, 9 sound engineers participated in the test; 5 of them completed the test in the big room, and 4 in the small room.

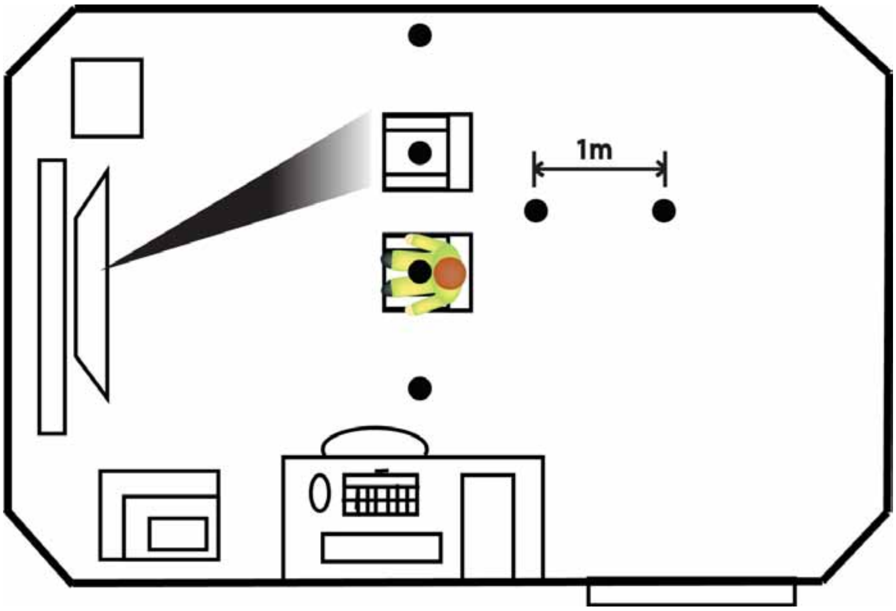

Figure 21- Main listening test setup. The settings are located this time 1m from each other.

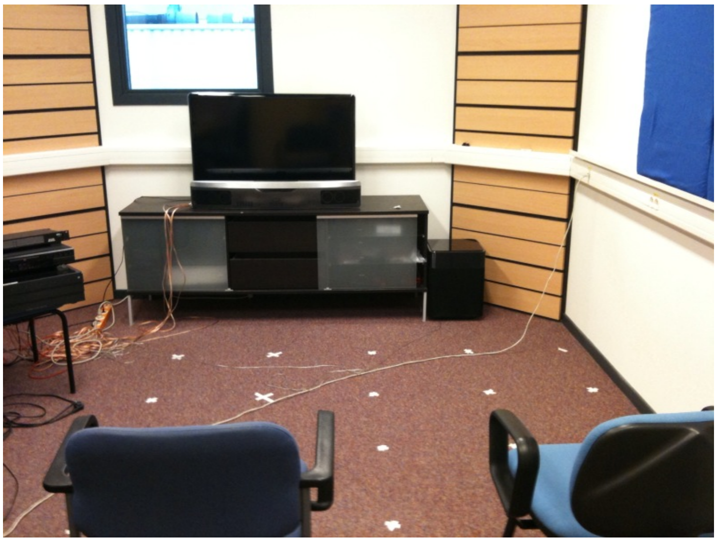

Figure 22 - The small room where part of the main listening test took place.

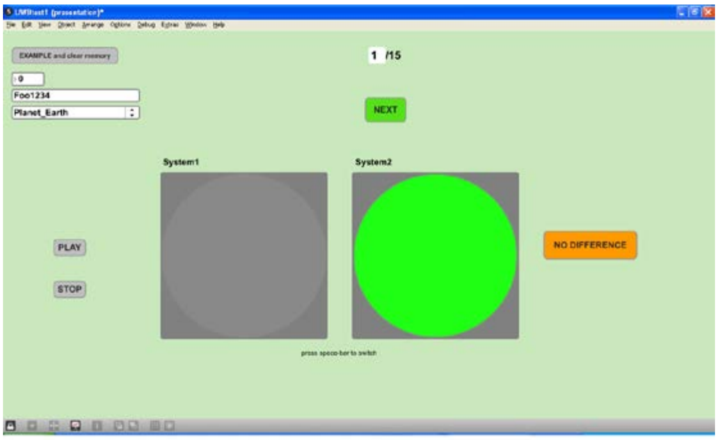

Figure 23: The GUI of the listening test that was presented to the participants.

3.4.2.2 ***Results***

For the results' presentation and analysis, some terms are used that will be explained here, for this section and for future reference.

- Setting refers to the steering of the notch towards a particular position. The participant is able to switch between settings and listen to the audio track configured for the setting of her/his choice. As it can be seen on the diagrams, four settings are in x and two in y axis, separated with 1m distance from each other.
- The intended setting refers to the setting that corresponds to the listening position, in other words, when the notch is steered towards the user.

In position 0, the nine participants who took part, 5 of them in the big and 4 in the small room, listened to two tracks, "Beyonce" concert and "Planet Earth". In Figure 24, the first pie-chart shows the results regarding the intended setting of all the comparisons in total; no clear preference is indicated for the intended setting however 64% of cases there is a preferred setting that indicates a clear perceptive difference made by the steering effect. In the next two charts, showing the results for the intended setting for both tracks in the two different rooms, the room acoustics influence in the responses becomes transparent; in the big room, in 80% of the comparisons, the users indicate preference towards a setting, while in the small room this percentage is 45% and the majority is 'no difference' responses. The next two charts present the results regarding the preference to the intended setting in both rooms for the two different tracks. Although the two songs were completely different in terms of audio content, the results are comparable. Those results indicate that the impact of the room is crucial to the listening experience and to the preference or not towards a particular configuration.

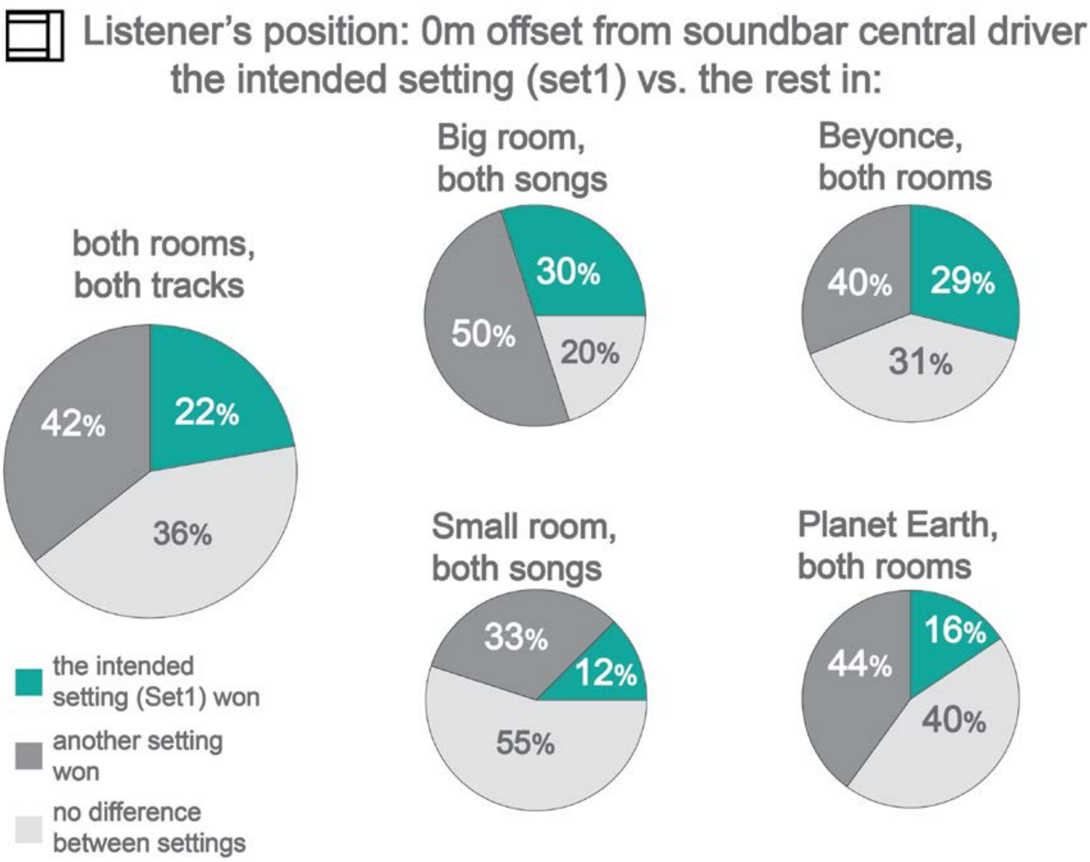

Figure 24: Results regarding the variables room and audio track for position with 0m offset.

Figure 25 presents the results in position 0 for each room on only 'Beyonce concert intro', so as to enable comparison with the corresponding results in listening position 1 (as only Beyonce track was used in position 1). In the first pie chart, it can be observed that the results for the intended setting do not show any clear preference. In particular, the next two diagrams show a great difference between the responses in big and small room (12% and 55% the 'no difference' replies correspondingly). As it can be seen in the room maps, the listeners in the big room indicate more differences between the settings. Moreover, a zone of preferred settings seems to be formed around settings 2, 3 and 4, towards the right side of the room. In the small room, the 'no difference' responses are the majority and no clear preference zone can be identified. The results are similar also for 'Planet Earth', only the participants give more 'no difference' replies in both rooms and especially in the small one.

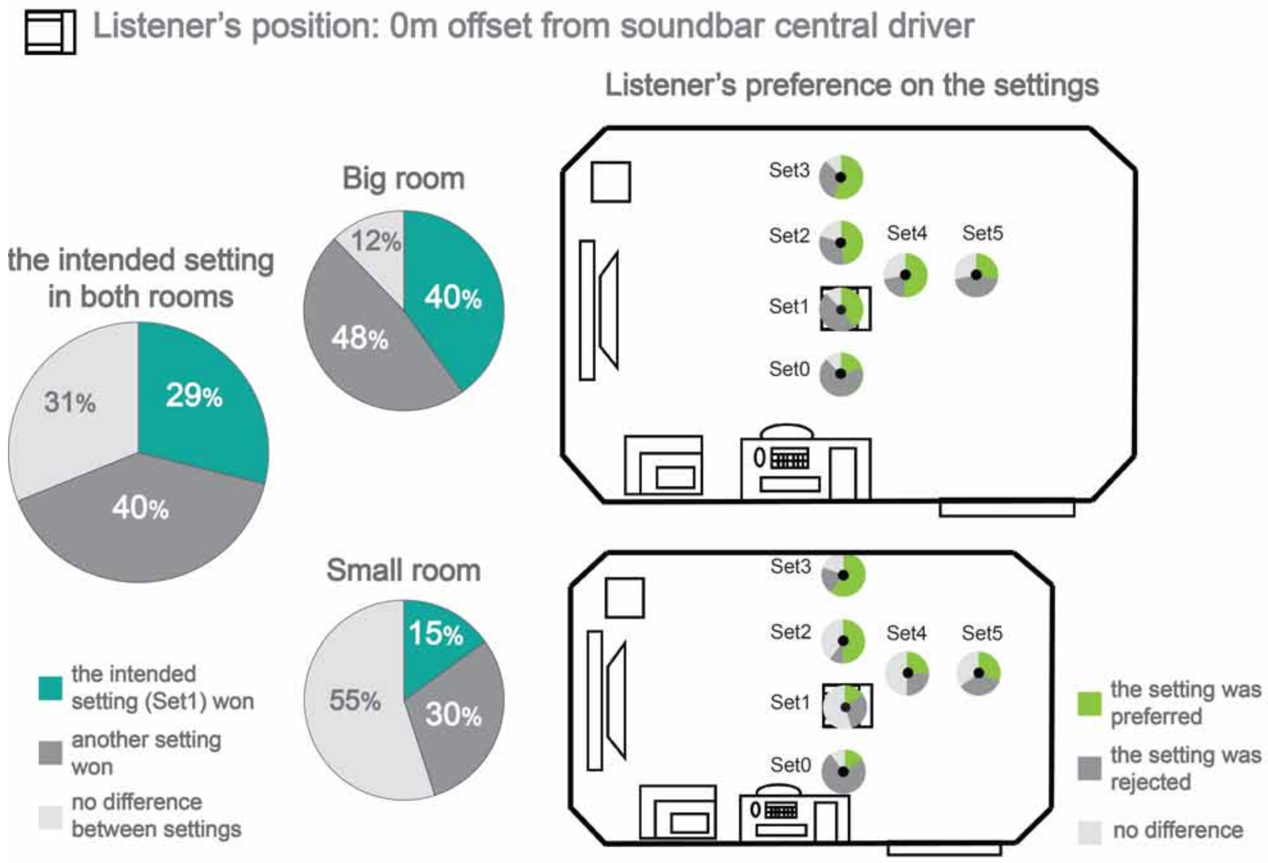

Figure 25: Results for the position in front of Soundbar (0m offset), only 'Beyonce concert'.

Figure 26 presents the results in position 1 for each room on 'Beyonce concert'. The intended setting 'wins' in the comparison by 72% and 30% in the big and small room correspondingly. Comparing the two room maps, the user have more explicit preferences - more differences between the settings are indicated in the big room than in the small. The zone of preference in the big room is formed around settings 2 and 3 and in the small room; setting 3 and setting 0 seem to be the most preferred ones, without though winning above 50% of comparisons. The zones of preference seems to be related to the room acoustics; the sound is steered towards a wall and reflected according to wall's acoustic properties and this reflection is heard together with the sound coming from Soundbar and thus, influences the perceived sound and the listening experience.

In general, stronger preference and fewer 'no difference' answers are given in the big room in both positions. The preference zones appear to be wider than 1m that was the initial precision. For position 0, the preference zones are wider for both rooms than the equivalent ones in position 1, and that is supported also by some sound experts' comments, that all settings sounded better from position 0. In all cases, the zone of preference is located towards a smooth wall and not towards the desk and the equipment. On y axis, as expected, much more differences are heard in the big room and setting 4 is preferred over setting 5, which is reasonable, as the notches converge closer to the position of the listener. Overall, the differences are judged as subtle by the sound experts. However, subtle differences can be potentially very important for the user's long term listening experience from a product.

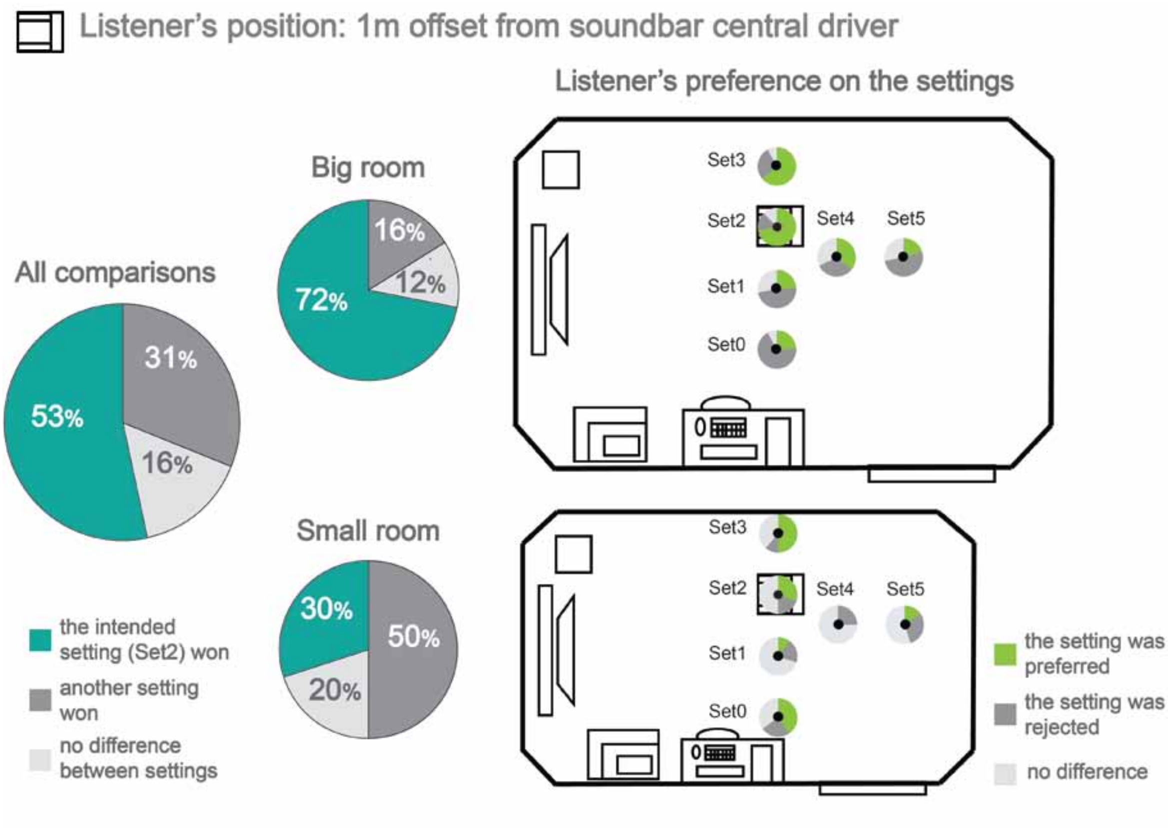

Figure 26: Results for the position with lateral distance 1m from Soundbar.

3.4.3 Stereo listening test

A complementary test was designed with non-expert users as participants. The goal was to test the audio tuning algorithm also for stereo recorded music and to incorporate the observation of the previous test that the precision is wider than 1m.

In the previous tests setup, the size of the available rooms, and in particular the small room limited the setup of the experiment to two user's positions with 1m meter distance between them : the first chair was placed on the central driver of the Soundbar and the second 1m to the side from the first. Thus, for the second experiment a spacious room (10x6 m^2), with reverbrative surfaces (windows), was used and the setup included two chairs placed 2 meters to the side, right and left from the central soundbar driver (see Figure 27). In the previous test, the audio material was two 5.1 recorded movie tracks. In this test, the participants evaluated the sett ings listening to stereo music, a type of audio that is commonly found in TV, radio, music recordings etc. The tracks that were used were extracted from the songs: "Owl city– Hello Seattle" and "Robbie Williams – Feel". It should be noted that Ambisound is designed for 5.1 and not for stereo recordings. However, stereo music in Ambisound is distributed in the dipoles (side speakers) and therefore with this test, the goal was to try out the effect of notch steering on this particular type of audio. The ten participants were all non-experts, without any official ear training experience.

The method that was followed was similar to the main listening test; the participant has to compare each time 2 settings, on a particular audio track and choose whether he prefers the first, the second or he thinks that there is no difference between them. This time, for simplicity reasons, three settings were compared making 3 different comparison pairs: S1 (distance = 2.5m, offset = -2m), S2 (distance = 2.5m, offset = 0m), S3 (distance = 2.5m, offset = +2m). The users had to complete the three comparisons for both tracks.

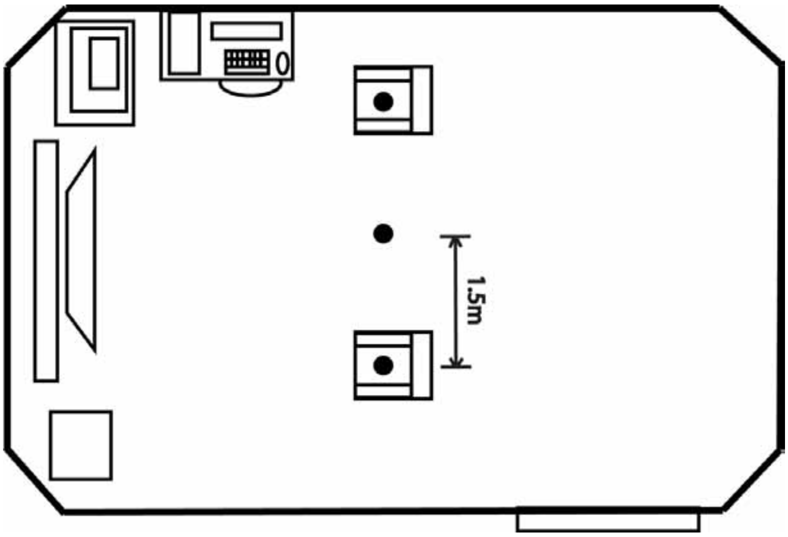

Figure 27: Stereo listening test setup.

3.4.3.1 **Results**

Users were able to identify a difference in the vast majority of comparisons for both tracks. In figure 28, a pie chart appears on each setting that shows the results for the all comparisons in which the particular setting participates, when the participant is sitting on t he chair depicted in the figure. In general, the participants' preference from both listening positions (represented as chairs in Figure 19), has a s light tendency towards the setting that is in the opposite direction of the listening position. When they were asked to choose between S1 and S3, the two opposite settings, half of the users selected the intended setting (setting that steered the notch to his/her chair) and the other half choose the other. Most participants mentioned that the 'intended setting' sounded more "natural" and "surround" and the setting of the opposite direction more "dynamic". This can be explained by the observation that the steering of the notch to a direction shifted the vocals to the opposite direction making the voice louder and thus appearing more "dynamic". The voice intelligibility is widely considered as one of the most important aspects in the human sound perception and it is therefore likely that half the participants were more drawn to choose the setting with the dynamic vocals and the other half the natural more balanced sound setting. Like in the main listening test, also here, the two different tracks did not make a significant difference to the choices of the participants.

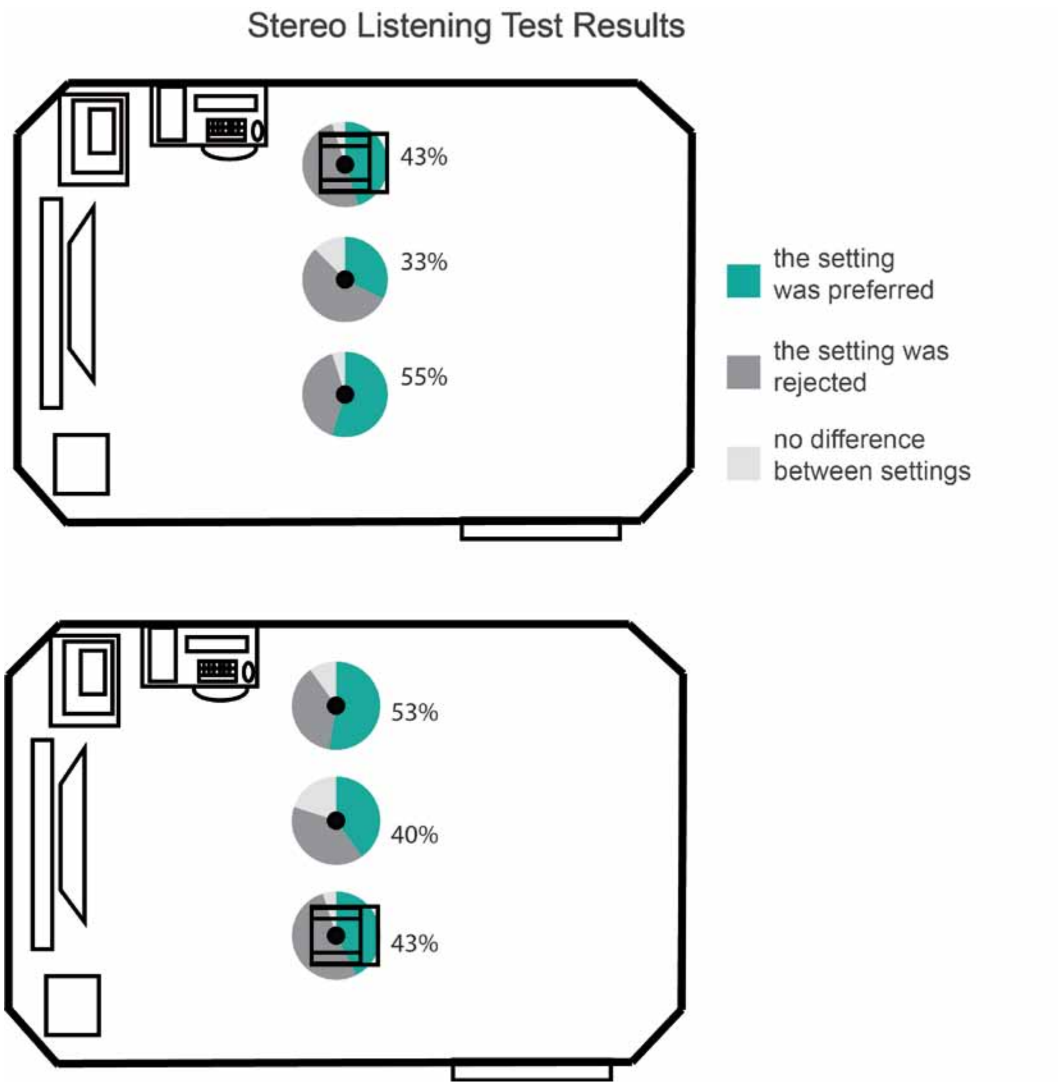

Figure 28: Stereo Listening Test Results

3.5 Discussion

In this section, the key findings from the evaluation sessions are presented. First of all, the localizer due to its size (and cost of about 20.000 euros), but also because it detects based on movement and not on the properties of moving objects, would not fit in a realistic use-case. The module cannot be integrated with the Soundbar chassis and in a real living room there is not only a moving person and a static background; there might be more persons, animals or objects moving (e.g. curtains). The detection area of the localizer (3.5m x 3m) is relatively small and could only cover small living rooms. Moreover, the averaging algorithm, although it adds to the robustness, it cause delays of some seconds and deviation from the actual position. Therefore, for the purpose of the demonstrator, the use of the localizer is only possible if the user initializes the location tracking in his listening position and waits for e.g. 10 sec until the audio tuning takes place.

Regarding the audio tuning, from the pilot test it was observed that the notch steering produces a subtle effect that is audible in some types of audio content. The precision was also found to be equal or more coarse than 1m, therefore satisfying the evaluation requirement that the step of the notch steering should be greater than the error caused by the localizer. A rather unexpected finding was the big difference in how the notch steering effect was perceived in the two rooms; it seems that the big room acoustics accentuated the notch steering effect and allowed the participants to listen more detail and make judgements more easily, while in the small room, in most of the cases, no difference between different notch steerings was heard. Moreover, it was observed that in both rooms, for both positions, the zone of preference, namely the group of settings that concentrated the highest preference in the comparisons, was located towards an empty wall, opposite from where the equipment (sound card, amplifier, pc, desk, cables) was located. Therefore, it can be concluded that the audible effect of the notch steering and how it will be perceived by the listener depends mostly on the room acoustics and the arrangement of obstacles/objects in the listening space, rather than on the position of the listener. Table 2 summarizes the evaluation requirements and the performance of the two systems. In a future iteration of the design and the evaluation of the algorithm, it would be thus valuable to investigate how can the algorithm incorporate room information and adjust the notch steering not only in relation to the user's position but also in accordance with the room arrangement.

Location tracking system	*response time*	- delays of some seconds due to averaging algorithm
	detection area	length ≅3.5m, width ≅3m
	error	- error increases with distance - threshold for movement is low - for 3.5m distance from the antennas, deviation on y axis ≅0.4m and on x axis ≅1m
	way of interaction	- active localization with a button on the remote was preferred and is more suitable due to the response time
audio tuning system	*precision*	- at least 1m step to produce audible notch steering effect (on both axes)
	audio content	-works with : audio content with spatial reference/sound coming from different directions -does not work : dialogue, when the focus is on the centre channel and the front and rear channels carry much less or not at all audio content -stereo sound : the tuning can shift the vocals to the opposite direction of the notch.
	user's preference	- mostly dependent on the room acoustics -in the big room, users hear differences and can make a judgement, while in the small room, much less difference is heard. - the audible effect was mainly identified in the spatial attributes of audio, but also in timbre.

Table 2 - Summary of the evaluation requirements and results for Ambisound demonstrator.

4 5.1 Home Theatre Sound System with user-speaker location tracking demonstrator

4.1 Integration of the demonstrator

The multi-speaker demonstrator required the integration of the following components:

- the UWB localization sensors/nodes that deliver a distance matrix, containing the distances for each pair of them
- an algorithm that calculates the positions of the nodes from the distance matrix
- the audio tuning model, developed by Philips.

The localization part consists of 6 UWB nodes and an I/O board that collects data from the nodes and software in order to inject the data to an online repository available to all project partners. The nodes were developed by CEA-LETI (Pezzin, Buccaille, Alvarez & de Celis, 2010) and the data injection software by UPB. For our use-case, out of the 6 UWB nodes, 5 are mounted on top of the speakers and 1 on the remote control. The node that is mounted on the top of the central speaker is the central node and its coordinates are by default (0,0). Each node estimates its distance from every other node in the network and sends the data to the central node. The data collected from all the nodes forms a 6x6 distance matrix. Each value of the matrix contains the distance between two nodes as seen from the node that corresponds to the row-number, therefore the matrix is asymmetric. Data from magnetometers and gyroscopes is also delivered. Algorithms that apply the data from these inertial sensors in order to find the orientation of the nodes/speakers are developed by Bitgear, but are not used in this demonstrator, because they were delivered late in the project. Together with the distance matrix and the data from the magnetometers and the gyroscopes, a matrix that contains the number of broken connections between nodes is sent through the central node to the I/O board. In order to increase the reliability of the results, the distance matrix delivered to the audio tuning model is the average of 5 matrices, retrieved from the repository.

A positioning algorithm, developed by University of Oulu (Destino et al., 2010) utilizes the retrieved distance matrix and provides the positions of each of the nodes with the centre node as the origin of reference. The algorithm is based on the concept of ranging contraction, which involves the correction of the distance measurements by subtracting a certain value such that they are shorter than the exact ones.

The principles of the audio tuning model are described in the Background section. The algorithm was implemented in MAX/MSP. A simpler version of the algorithm, correcting only for the distance from the listener is implemented in MATLAB and it is also available for the prototype. In the interface of the MAX/MSP model (Figure 30), the user can move the speakers (the five blue dots) and the user (yellow dot) to new positions with regards to the initial position of the centre speaker, in front of the TV set that has coordinates (0, 0) (symbolized with the yellow vertical line). Moreover, the user can see the matrix of gains and delays and reset the loudspeaker positions according to the ITU standard. When the speakers are placed according to ITU standard, the gains matrix is a diagonal matrix with bright green dots. That means that no repanning is needed and each loudspeaker is playing its corresponding channel and only this. With the repanning, the gain matrix is not any more diagonal, because the speakers play a combination of more channels, each with a different gain.

Regarding the integration, one problem that needed to be solved was the broken connections between the nodes. Some of the nodes might temporarily fail to deliver their estimated distances to the central node, because of synchronization problems. This is indicated with the value -1 assigned when the estimated distance was not delivered in the distance matrix. For the purpose of this demonstrator, we developed a simple compensation method that makes use of the asymmetry of the distance matrix. As an example, imagine that out of the five matrices sampled to deliver one distance matrix, between nodes C and FL, C node delivered only 3 times an estimated distance and FL node delivered just 1 estimated distance and the rest, due to broken connection, were not delivered. The compensation algorithm will take the 4 distance values, divide them by 4 and assign the result symmetrically to the elements (C, FL) and (FL, C) in the distance matrix. In case both nodes do not deliver any distance values at all, then the corresponding value from the previous averaged distance matrix will be assigned to the elements of the matrix. The compensation method described here addresses the issue of missing values and increases the reliability of the measurements by averaging them, although it is not optimal. For a future design of a method compensating for broken connections but also for the improvement of the localization performance, advanced geometrical models and/or graph algorithms could be utilized. For example, Dijkstra's algorithm, a g raph search algorithm that calculates the cost of the shortest path in a graph, could be adjusted to calculate the missing distance values.

Unlike the Ambisound demonstrator where the localization concerns only the position of the user, in this demonstrator the positions of both speakers and user are identified with origin the position of a central node. The data is collected in the I/O board, then sent to a Linux virtual machine in the pc and a script injects the data in an online repository, available for all EUWB project partners. The data is finally retrieved in almost real-time (with some seconds delay) with a Matlab script back to the pc.

The audio tuning model receives audio input from a CD/DVD player and the location coordinates in MATLAB format and adjusts the levels and delays of each channel of the 5 speakers accordingly. The model's outputs are then transferred to the soundcard and from there to the amplifier. Finally the amplifier sends audio signal to each of the five speakers. Figure 29 depicts the set up of the final demonstrator.

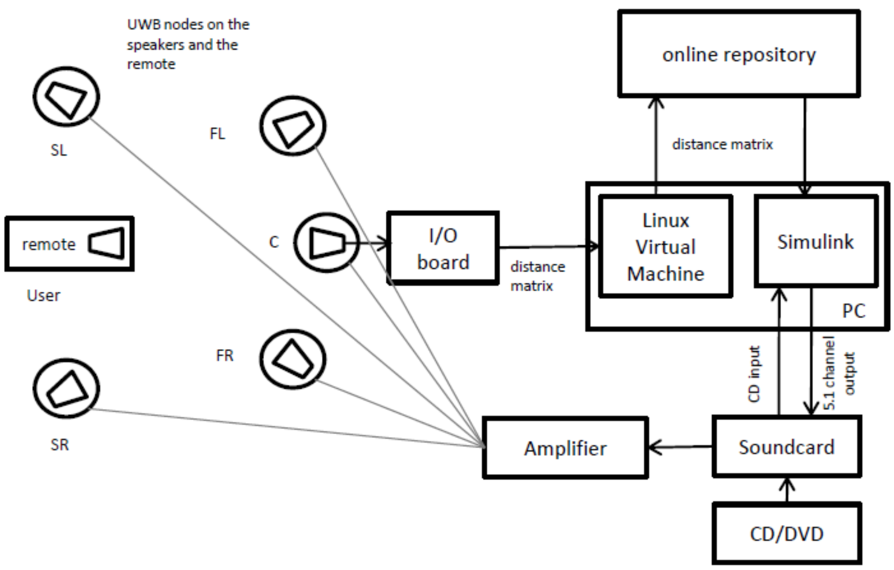

Figure 29: The block diagram of the 5.1 Home Theatre Sound System Demonstrator.

	C	FL	FR	SL	SR	U
C	0	0	0	0	0	0
FL	1.72	0	0	0	0	0
FR	1.64	3.02	0	0	0	0
SL	4.45	3.66	4.61	0	0	0
SR	4.71	4.71	3.7	2.93	0	0
U	3.04	2.91	2.7	2.02	1.87	0

Table 3- The groundtruth distance matrix (in meters) as measured with a measuring tape.

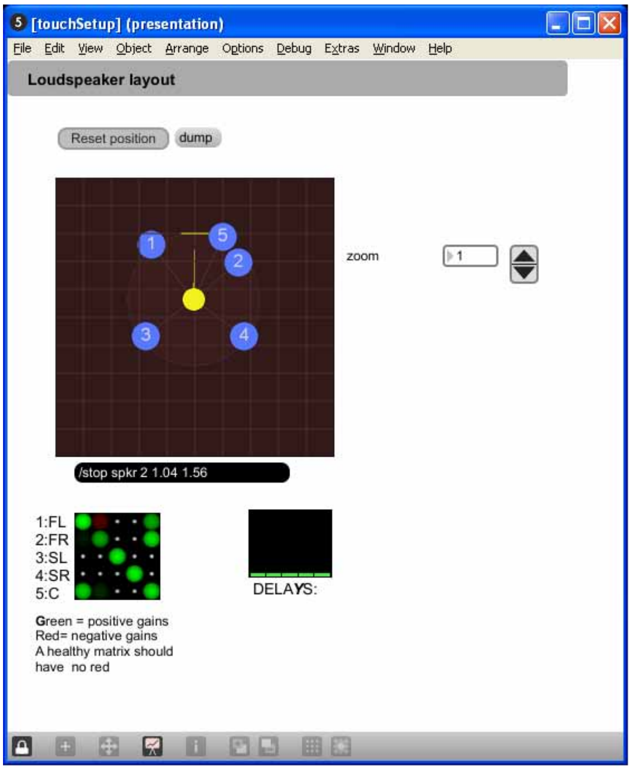

Figure 30 – GUI of the repanning algorithm. By dragging the speakers, the gains and the delays change in the matrices as well as the signals sent to the speakers.

4.2 Evaluation requirements

The evaluation requirements and objectives, described in this section are similar to the requirements described for the Ambisound demonstrator. Regarding localization, we are interested in investigating the error in the estimated distance matrices, the size of the detection area, performance in loss of sight cases (when obstacle interferes in the communication between two nodes) and if there are delays involved. Moreover, the error imposed by the positioning algorithm is also to be tested. In this demonstrator, the method of interaction is active; the user is localized when he activates the UWB sensor on her/his remote control.

The audio tuning evaluation tests were mostly concerned with how the audible effect of the repanning algorithm is perceived and if it improves the listening experience. The role of the room acoustics and the audio type was not investigated and these parameters were constants in our experiment, due to time constraints. Another important requirement for the particular use-case that needs to be addressed, is to define how the misplacement of a speaker without any audio compensation algorithm affects the user's listening experience and consequently the effectiveness of the audio tuning on each of them.

4.3 Location tracking evaluation

4.3.1 Experiment design

The two parts of the multi-speaker demonstrator, localization and audio tuning are tested separately as in the previous demonstrator. The goal of the localization evaluation is to identify the limitations of the UWB LT hardware (UWB nodes) and the software for home audio applications.

For this purpose, some experiments with a known setup were conducted. The nodes were mounted on top of the speakers in a pentagon and the user node is placed in the centre on top of a stand, in order to be in a comparable height with the rest of the UWB nodes. The maximum distance between the nodes (the distance between SR and C) for the test environment was measured 4.7m. In some attempts to place the nodes further apart, it was observed that the distant nodes were not detected or could not deliver estimated distances to the central node. The actual distance matrix (groundtruth) was manually calculated with a measurement tape and this already imposes an in-built error to the measurements presented in this section(table 3).

Four experiments in total were conducted:

- In the first experiment (Figure 31), the UWB nodes are tested without any obstacle intervening in the setup.
- In the second experiment (Figure 32), a person stands behind the user node in the middle of the pentagon.
- In the third experiment (Figure 33), a person stands between the C and the FL nodes.
- In the fourth experiment (Figure 34), a person stands between FR and SR node.

The nodes are mentioned here with the names of the speakers on which they were mounted on. For each experiment, 20 measurements (distance matrices) each sampled every 15 seconds (the distance matrix is an average of 5 matrices and each of them need less than 3 seconds to be retrieved) were retrieved. An additional output of each experiment was 20 broken connection matrices. The broken connection matrices give for each pair of nodes the number of broken connections out of 5 samples used for each distance matrix. As mentioned above, each delivered distance matrix is a result of compensation and averaging of 5 matrices.

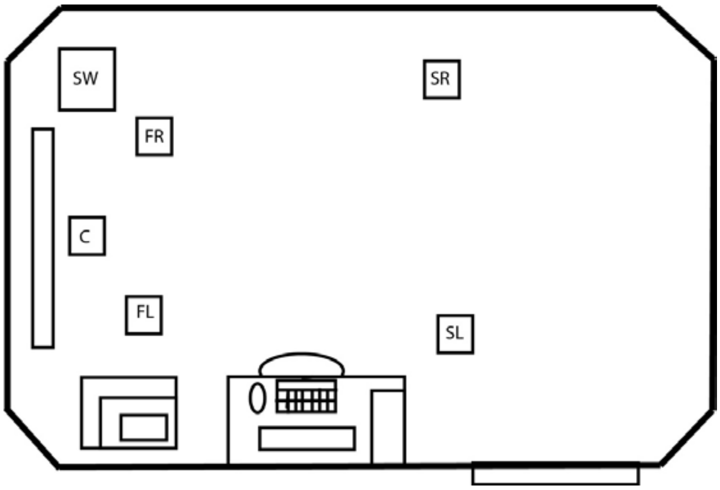

Figure 31 - No obstacle present.

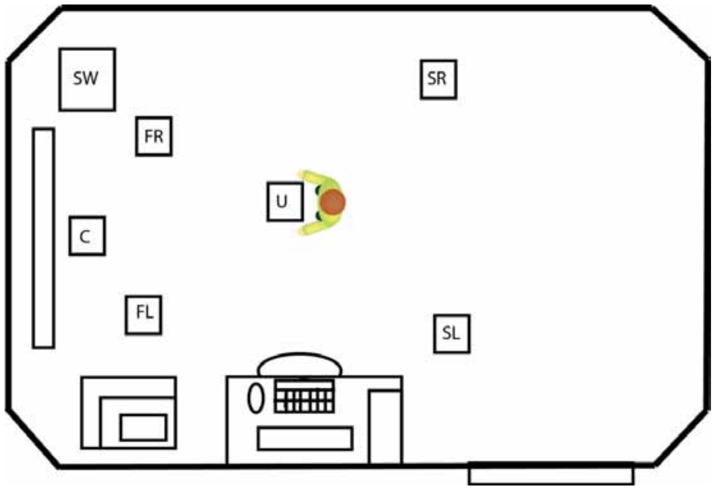

Figure 32 - Experiment with a person behind the remote control.

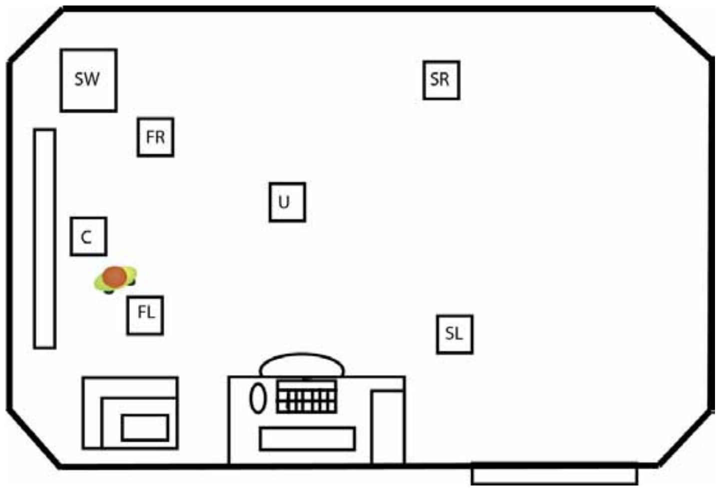

Figure 33 - Experiment with a person between C and FL nodes.

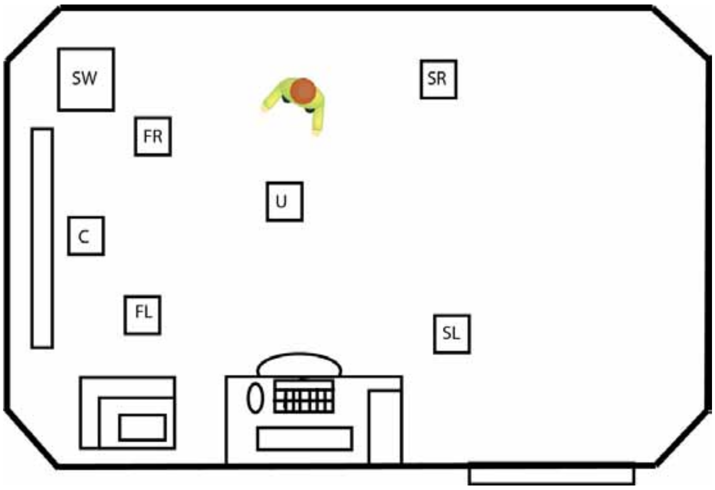

Figure 34 - Experiment with a person between FR and SR node.

4.3.2 Results

4.3.2.1 ***Broken Connections***

Below, the average broken connection matrices for each experiment are presented. The rows represent how many times out of five each node cannot 'see' the others. Tables 4-7 represent the broken connections per each pair of nodes that correspond to the four above-mentioned experiments.

	C	FL	FR	SL	SR	U
C	0	0	0	4.15	0.4	0.55
FL	0	0	0	2.65	0	0
FR	0.05	0.05	0	2.5	0.05	0.05
SL	4.45	3.35	3.4	0	3.35	3.35
SR	0.15	0.1	0.1	2.65	0	0.1
U	0.6	0.05	0.05	2.1	0.05	0

Table 4 - Broken Connections table for the experiment in which no obstacle is present. Light blue is used for 0-1 broken connections, light orange between 1 and 3 and light burgundy above 3.

	C	FL	FR	SL	SR	U
C	0	0.1	0	3.55	0	1.65
FL	0	0	0	2.3	0	0
FR	0	0	0	1.45	0	0
SL	4.75	3.6	2.95	0	2.95	2.95
SR	0	0	0	1.85	0	0
U	0.8	0.45	0.45	2.05	0.45	0

Table 5 - Broken connection table for the experiment where a person is standing behind the remote control.

	C	FL	FR	SL	SR	U
C	0	0.15	0	3.15	0.1	0.3
FL	0	0	0	2.6	0	0
FR	0	0	0	1.85	0	0
SL	4.05	3.1	2.75	0	2.75	2.75
SR	0	0	0	2.2	0	0
U	0.5	0.1	0.1	1.8	0.2	0

Table 6 - Broken connections table for the experiment in which a person is standing between the C and the FL node.

	C	FL	FR	SL	SR	U
C	0	0	0	2.65	0.1	1.3
FL	0	0	0	1.45	0	0
FR	0	0	0	1.55	0	0
SL	4.45	2.25	2.1	0	2.1	2.1
SR	0.05	0	0	1.65	0	0
U	0.8	0.05	0.05	1.5	0.05	0

Table 7 - Broken Connections table for the experiment in which a person is standing between SR and FR nodes.

First of all, a person standing as an obstacle seems not to have an effect on the number of broken connections, therefore the nodes can estimate and deliver distance values. The effect of the obstacle in localization precision will be examined in the next section.

It is obvious observing the matrices that most connections deliver on average distance values more than 4 times out of 5. However, the rate of broken connections is high with the node mounted on SL speaker. In all cases, the pair SL-C scored the highest rate (>4) of broken connections. SL node paired with other nodes also has remarkably high broken connection rates; it seems that this node cannot ‘see’ and ‘be seen’ efficiently from other nodes.

As we can see in the groundtruth matrix, the maximum distance is between SR-C and SR-FL and it would be expected that the maximum broken connections would be assigned to these pairs, but they have almost no broken connections. A possible explanation is that the particular node (the one mounted on SL) is defective and not able to synchronize well with the others.

4.3.2.2 *Relative Localization Error*

The relative localization error is the ratio of the error between an estimated and the actual distance to the actual distance between two nodes:

$$\mathrm{Re}lative\_Error = \frac{|estimated\_dist - actual\_dist|}{actual\_dist}$$

The average relative error per each pair of nodes for each experiment is presented in Table 8.First, it can be observed that the relative error has remarkable variation between different node-pairs, ranging from 0 to 0.40 (with an outlier of 0.66). Moreover, there are identifiable patterns that can be seen on the matrices with the colour-coding. In all experiments the highest relative errors appears to the node pairs with U (around 30% of the actual distance, approximately 1m absolute error). High relative error doesn’t seem to be highly correlated with broken connections rate, because in that case SL node pairs would have high relative error and they don’t (that means that the compensation algorithm worked well for this case).

Comparing the relative errors of the node pairs with obstacle with the corresponding errors without obstacle in experiments 3 and 4, it is observed that the error of the node pairs with an obstacle increases by more than 100%, which is expected as the UWB signal is delayed by the obstacle and this is translated as bigger distance.

Experiment		C	FL	FR	SL	SR	U	all node pairs
1st experiment	C	0	0.07	0.08	0.21	0.03	0.66	mean = 0.18 std = 0.19
	FL	0.07	0	0.13	0.03	0.18	0.32	
	FR	0.08	0.13	0	0.1	0.09	0.27	
	SL	0.21	0.03	0.1	0	0.11	0.3	
	SR	0.03	0.18	0.09	0.11	0	0.05	
	U	0.66	0.32	0.27	0.3	0.05	0	
2nd experiment		C	FL	FR	SL	SR	U	all node pairs
	C	0	0.09	0.07	0.17	0.02	0.34	mean = 0.18 std = 0.17
	FL	0.09	0	0.12	0.11	0.21	0.34	
	FR	0.07	0.12	0	0.1	0.1	0.22	
	SL	0.17	0.11	0.1	0	0.12	0.4	
	SR	0.02	0.21	0.1	0.12	0	0.28	
	U	0.34	0.34	0.22	0.4	0.28	0	
3rd experiment		C	FL	FR	SL	SR	U	all node pairs
	C	0	0.23	0.07	0.21	0.03	0.33	mean = 0.18 std = 0.15
	FL	0.23	0	0.15	0.07	0.2	0.36	
	FR	0.07	0.15	0	0.11	0.1	0.24	
	SL	0.21	0.07	0.11	0	0.13	0.38	
	SR	0.03	0.2	0.1	0.13	0	0.05	
	U	0.33	0.36	0.24	0.38	0.05	0	
4th experiment		C	FL	FR	SL	SR	U	all node pairs
	C	0	0.08	0.07	0.23	0.07	0.32	mean = 0.17 std = 0.14
	FL	0.08	0	0.12	0.05	0.19	0.37	
	FR	0.07	0.12	0	0.1	0.22	0.25	
	SL	0.23	0.05	0.1	0	0.12	0.36	
	SR	0.07	0.19	0.22	0.12	0	0.05	
	U	0.32	0.37	0.25	0.36	0.05	0	

Table 8 - Average relative error for each pair of nodes, per experiment. Light blue is used for error between 0 and 0.15, light orange for error between 0.16 and 0.30 and light burgundy above 0.30 error and in bright orange is the outlier (0.66).

4.3.2.3 ***Positioning Algorithm Evaluation***

The positioning algorithm receives the distance matrix as input and delivers the positions of the nodes in regards to the central node. The algorithm is developed by the University of Oulu and described in deliverable (Destino et. al., 2010).

The evaluation of the algorithm was based on the distance matrices extracted from the above - mentioned experiments. Utilizing these matrices, the algorithm delivered the positions of the speakers with origin the central node (C with coordinates (0,0)) but with unknown coordinate system. From the positions, a distance matrix is constructed and compared to the actual distance matrix. This method was preferred instead of calculating the positioning error by measuring the actual positions of the setup and comparing them with the positioning algorithm results, as the system of reference used in the algorithm was not known. Moreover, the contribution of the positioning algorithm to the error can thus be more clearly represented, since the errors of the constructed distance matrices are comparable with the errors of the direct distance matrices delivered from the nodes.

The results (table 9) show that overall the errors slightly increased for the constructed distance matrices. The U node estimated distances still have the highest errors, but some of them (incl. pair FR-SR with obstacle in experiment 4), are reduced. The pair C-FL presents the highest relative error in most experiments (around 40%) and this is not justified by its values error in the direct distance matrices.

Positioning Algorithm Results								
1st experiment		C	FL	FR	SL	SR	U	all node pairs
	C	0	0.38	0.11	0.23	0.04	0.32	mean = 0.21 std = 0.19
	FL	0.38	0	0.12	0.24	0.19	0.45	
	FR	0.11	0.12	0	0.15	0.07	0.23	
	SL	0.23	0.24	0.15	0	0.27	0.24	
	SR	0.04	0.19	0.07	0.27	0	0.12	
	U	0.32	0.45	0.23	0.24	0.12	0	
2nd experiment		C	FL	FR	SL	SR	U	all node pairs
	C	0	0.36	0.17	0.18	0.03	0.1	mean = 0.18 std = 0.15
	FL	0.36	0	0.12	0.24	0.22	0.24	
	FR	0.17	0.12	0	0.13	0.07	0.07	
	SL	0.18	0.24	0.13	0	0.24	0.3	
	SR	0.03	0.22	0.07	0.24	0	0.19	
	U	0.1	0.24	0.07	0.3	0.19	0	
3rd experiment		C	FL	FR	SL	SR	U	all node pairs
	C	0	0.47	0.14	0.23	0.04	0.16	mean = 0.2 std = 0.2
	FL	0.47	0	0.16	0.2	0.21	0.34	
	FR	0.14	0.16	0	0.15	0.09	0.07	
	SL	0.23	0.2	0.15	0	0.29	0.37	
	SR	0.04	0.21	0.09	0.29	0	0.05	
	U	0.16	0.34	0.07	0.37	0.05	0	
4th experiment		C	FL	FR	SL	SR	U	all node pairs
	C	0	0.38	0.3	0.25	0.08	0.2	mean = 0.2 std = 0.16
	FL	0.38	0	0.09	0.23	0.2	0.34	
	FR	0.3	0.09	0	0.14	0.14	0.09	
	SL	0.25	0.23	0.14	0	0.22	0.36	
	SR	0.08	0.2	0.14	0.22	0	0.05	
	U	0.2	0.34	0.09	0.36	0.05	0	

Table 9 - The same color coding as previously is used here.

4.4 5.1 Home Theatre audio tuning evaluation

4.4.1 Experiment design

The repanning algorithm, used for the audio tuning of 5.1 Home Theatre Sound System was evaluated with a listening test, in which 7 non-sound-experts took part. The participants were chosen because of availability and also because the audible effect of the algorithm appeared to be easily distinguishable. The test was a first approach to the question, whether the repanning algorithm improves the listening experience and in which conditions is it most effective.

The participants, 6 males and 1 female, all working in the IT or human resources domain in Philips, were welcomed by the experimenter in a small room, that resembles a furnished living room. The room was covered with floor carpet, and there were two armchairs, a TV-set, a 5.1 Home Theatre Sound System and decoration items placed in the walls and on the floor. This room is used in some of the Philips listening tests, because it simulates acoustically a small living room as well as it creates a familiar atmosphere to the participants. The room on one side was separated by a curtain from a small electronics lab, where the pc with the repanning algorithm was placed.

The participants were then given a keyboard and were instructed to switch between 2 settings (corresponding to keys) while listening and watching a movie track from the Disney animation movie 'Aladdin'. One setting was the sound as it would be reproduced from a normal 5.1 system and the other would include the repanning algorithm processing, adjusting the audio signals to the speakers according to the actual speaker- and user- coordinates. The participants were required to choose one of these settings or indicate that there is no difference between them in a paper questionnaire. They were also asked to decide whether the difference they heard - if any - belonged to the spatial, timbral or other attributes of sound. The procedure was repeated 6 times, as there were 6 setups with different positions of the user and the speakers. See Figure 35.

When they made their choice, the experimenter would change the position of the speakers and/or instruct the user to change position and make a choice again between the two settings. For the purpose of the experiment, the positions of the user and the speakers were predetermined and measured and were manually entered to the algorithm by the experimenter.

The movie track lasted approximately 3min, but for each of the 6 sessions, the participant was free to listen as long as (s)he needed to make her/his choice. The movie track from 'Aladdin' was selected out of other movie soundtracks because of its rich audio content; the main vocals coming from the centre channel, a choir playing from the front channels and there are background sounds, music and vocals. Like Ambisound, it was observed that the repanning effect was accentuated on soundtracks with rich spatial features. The participants were simultaneously watching the video of the movie track in order to have a visual reference of the direction and the sources of sounds heard.

Unlike the Ambisound audio tuning evaluation, in this listening test, the room and the audio were not used as variables; only one room and one audio track was used. This was because there was insufficient time to develop and implement a test with all possible variables that might influence the audio tuning effect.

Moreover, in all 6 setups the displacements of user and speakers from the actual positions were in steps of 0.5m. The setups were mainly focused on the displacement of the centre speaker and the users, and also on the displacement of both centre and front speaker, as the centre speaker usually concentrates the most important audio content (e.g. dialogue) and receives the most attention from the listeners.

Figure 35- Listening test setups for the 5.1 Home Theater Sound System.

4.4.2 Experiment results

In total 70% of the choices that the participants made, were in favour of the repanning algorithm in the audio processing, 17% indicated preference to the simple 5.1 audio processing without the repanning, and 13% the two settings were found similar (‘no difference’ response).

In the first row’s setups (Figure 36), (a), (c), (e), the user and/or the centre speaker are displaced by 1m. In the second row’s setups, there’s greater deviation from the ideal setup; in one case the user and the centre speaker are both displaced to different directions by 1m and in the other two, displacement of two speakers and/or displacement of the user is attempted.

In the first row, it can be observed, that there’s more variation of the results, however with a clear tendency towards the repanning setting except from (c) setup; there the participants preferred the simple 5.1 audio processing setting. Their comments were that ‘voices were better centralized’ and that ‘the sound was more spatial and voluminous’. A possible explanation is that the (c) setup the user is in-line with the centre speaker and therefore the participant has the main vocals directly coming to her/him with the simple 5.1 audio processing. On the other hand, the repanning creates a phantom centre source in front of the TV, distributing the centre signal to the front speakers, thus 'dissolving' the main vocals and it is possible that the voice intelligibility might be slightly distorted, regardless if the direction is right.

In the second row, the results indicate in all three setups that the participants preferred the repanning setting, in almost all the cases where the difference between the settings could be identified. It is interesting to note that when the participants are sitting in the middle of the setup, in the third column of figure 30, they always prefer the repanning setting. It is possible that the participants, when sitting in the centre chair in front of the TV, have stronger expectations for the direction of the sound and can identify easier where the sound should be coming from. In the particular setups with normal 5.1 processing, the centre channel - which contains the main vocals –is heard as if it is coming from the side, while the participant is expecting it from the front of him, just where the TV set is. The repanning algorithm creates a phantom sound source in front of the TV set –where the centre speaker should be- and therefore the sound seems to come naturally from in front.

The participants also reported that the difference between the settings was spatial, except from the first (a) setup, where the majority found differences in the timbre. In particular, for the (a) setup, they commented that 'the vocals sounded more clear, crisp and real' with the repanning setting. For the other setups, they mentioned that the repanning setting provided 'better balance', 'sounded fuller while the other (setting) sounded mono' and they noticed shift of direction of sound between the two settings.

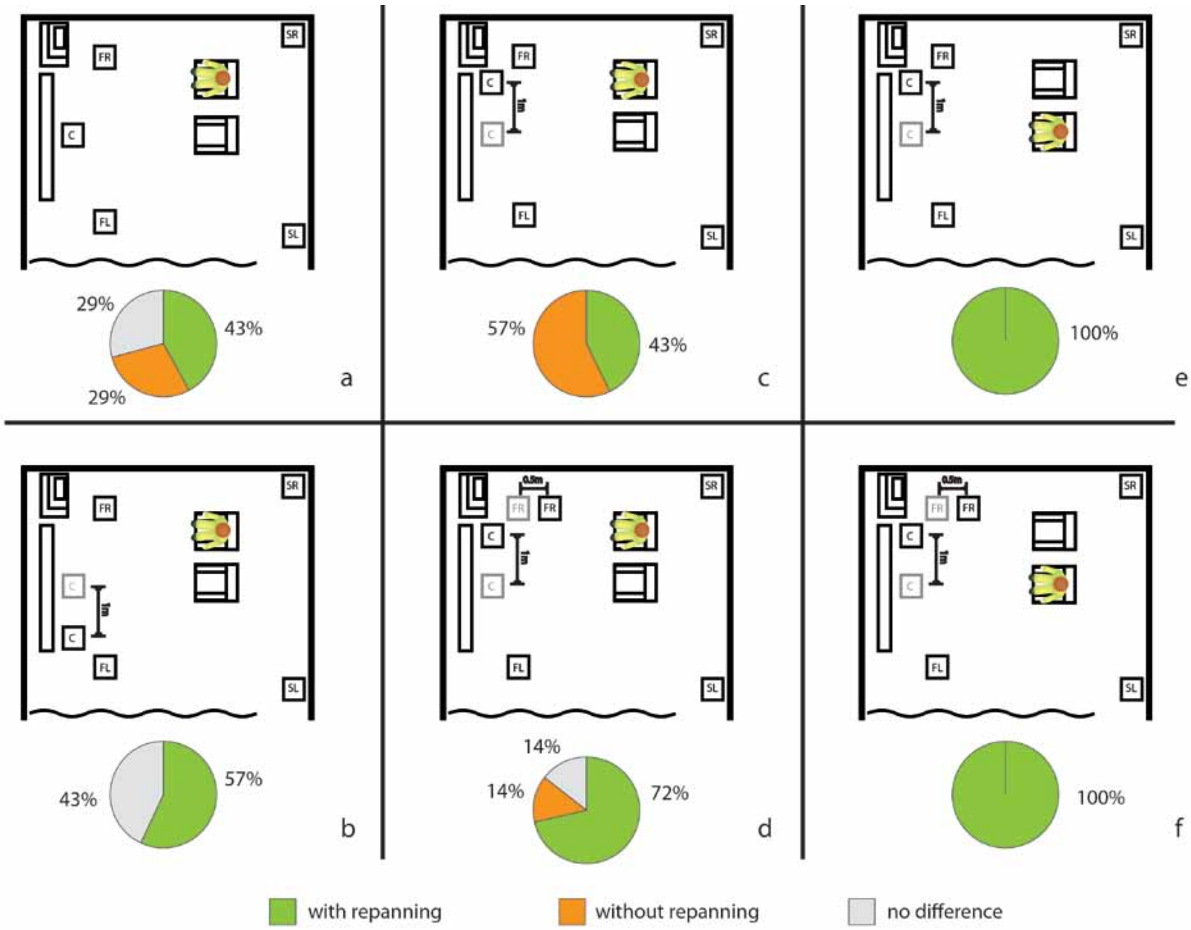

Figure 36 - Listening test results.

4.5 Discussion

Regarding 5.1 H ome Theatre Sound System demonstrator, the results indicated that the UWB localization nodes cover effectively and with reasonable error the typical listening area (width 3.5m, length 5m). However two issues in the localization part need to be tackled in future iterations of the demonstrator. First, the broken connections between nodes that cause missing values in the distance matrix delivered should be addressed with methods of estimating the missing values as well as with techniques to improve the synchronization between the nodes. Second the error cause by an obstacle interfering between two nodes can be measured and corrected; it is yet to be investigated how different obstacles influence the time of flight. Moreover, the relative localization error ranging from close to 0 t o 40% can be translated with the positioning algorithm to a loudspeaker setup very different from the actual (e.g. due to error, two speakers appear swapped) and this could have a negative impact on the audio tuning. Overall, the UWB hardware and software has a robust behavior, serves the purpose of the demonstrator and has the potential to be further developed and integrated in more location-based home applications in the near future.

Regarding the audio tuning, the audio reproduced with the repanning algorithm was shown to concentrate the preference of the vast majority of the users over the simple 5.1 unprocessed sound when the one or more speakers were displaced by 1m. Like with Ambisound audio tuning, the difference was identified mainly in the spatial characteristics of the audio track; the participants reported that the vocals sounded more ‘real’ and the sound was better balanced in comparison to the unprocessed 5.1 audio. The listening tests could be repeated in the future with different types of audio and rooms. More setups, involving one or two speakers per time could also be tested, in order to define the contribution of each speaker in the sound image perceived by the listener. The evaluation requirements and the results are summarized in table 10.

Location tracking system	*response time*	- one matrix needs less than 3 seconds to be retrieved, because of averaging 5matrices, appr. 15 seconds are needed to update information, but this is up to the implementation
	detection area	- was not exhaustively but maximum distance between two nodes that was tried ≅ 4.71m
	error sources	- broken connections between nodes and missing values - obstacles which increase the error - defective nodes -big variation of error between different node pairs - relative error on average ≈ 0.19 (of the actual distance)
	way of interaction	- user decides his listening position and initializes the loc. tracking
Audio tuning system	*precision*	- 1m displacement of just the centre speaker caused a clear audible effect.
	audio content	- works with : audio with spatial references
	user’s preference	- the majority of users (70%) preferred the audio tuning from the simple 5.1 audio processing. - users reported that the difference they heard was spatial

Table 10 - A summary of the evaluation requirements and results for 5.1 Home Theatre demonstrator.

5 Conclusions

The main goal of this project, in the domain of home audio applications, was to provide to the listener the best possible surround sound experience by overcoming the limitations of loudspeaker setup with UWB localization technology. For Ambisound demonstrator, the notch effect was proven to create a perceptible difference that was mainly identified in the spatial attributes of audio, but also in timbre. The users' preference towards particular settings seems to be related with the room acoustic properties and not only with users' position and this is a claim to be further investigated in future listening tests. The concept of active localization, implemented with a button on the remote control, was preferred by the users but also addresses the capabilities and the limitations of the current UWB localization module, available for the demonstrator.

Regarding 5.1 Home Theatre Sound System demonstrator, the results indicated that the UWB localization nodes cover effectively and with reasonable error the typical listening area and can function also when obstacle are present. Certain sources of localization errors were identified and should be tackled in a future design of the location tracking system. The repanning algorithm effect was mainly on the spatial characteristics of the track; the participants (non-experts) reported that the vocals sounded more 'real' and the sound was better balanced in comparison to the unprocessed 5.1 audio. Furthermore, the audio reproduced with the repanning algorithm was shown to concentrate the preference of the vast majority of the users over the simple 5.1 unprocessed sound when the one or more speakers were placed in wrong positions. Although there is still room for improvement, the evaluation experiments until now have provided us with various insights on the topic and with strong indications for the potential of location-based surround sound tuning.

6 Future Work

The concept of a Surround Sound System adjusting to the environment and listener's position and preference has the potential to be further developed and expanded in the near future. In order to design such a system, several parameters have to be taken into account. The participation of these parameters in the final listening experience of the user should be clearly defined and prioritized and new listening tests should be conducted.

For this purpose, two studies would be suggested. The objective of the first study will be to identify the problems with the Surround Sound Systems, as they are used today. Which are the users' requirements regarding surround sound and are they fulfilled by the current Surround Sound Systems? Which sound attributes (e.g. spatial, timbre, bass, voice intelligibility) are the most important for the users' listening experience? How does the 5.1 Home Theatre Sound System performance vary in different living rooms? The second suggested study is concerned with the effectiveness of audio tuning algorithms which correct for these problems. Not only short-term but also longitudinal experiments should be planned, because a subtle difference in sound can lead to a clear identifiable preference after hours of listening over different audio content.

In our study, the role of room acoustics in the effectiveness of location-based audio tuning emerged – particularly in the case of Ambisound demonstrator – and should be tackled in the future work on surround sound audio tuning. Moreover, the fact that the participants reportedly had a strong preference in some particular configurations indicates that the users should be able to customize the sound themselves: not only according to their listening positions, but also according to room acoustics, their taste and even their hearing. Every user has unique ears and auditory organs that might determine together with his taste and habits his preference over particular frequency bands or sound attributes. Thereby the future generation of the Surround Sound Systems should provide the means to the users to optimize their perceived sound. Nevertheless, managing the complexity of giving more control (and therefore new features to the user) in balance with an intuitive and simple user-experience is an ongoing challenge that Surround Sound Systems, along with several other devices, will have to face in the future.

References

Recommendation ITU-R BS.775-1 (1997). Multi-channel Stereophonic Sound System with and without accompanying picture.

EUWB at a glance. (n.d.). Retrieved from www.euwb.eu/project-summary

Van Daele, B., Hafezi, P. (2008). Scenario descriptions for localization application for Home Audio Systems. Deliverable D8c.6, EUWB Integrated Project.

Gerzon, M. A. (1992). General metatheory of auditory localisation. *Audio Engineering Society Convention 92.*

van Leest, A. J. (2005). On amplitude panning and asymmetric loudspeaker set-ups. *Audio Engineering Society Convention 119.*

Alvarez, A., Rakić, D., Destino, G., Waadt, A., Zetik, R. et al.(2010). Implementation of the enhanced LT engine with mobility management (LDR and HDR). Deliverable D4.3.2b, EUWB Integrated Project.

Pezzin, M., Buccaille, I., Alvarez, A., de Celis, L. (2010). LDR-LT platform requirements, feasibility analysis and specification. Deliverable 7.1.1b, EUWB Integrated Project.

Destino, G., Chóliz, J., Zetik, R., Latosa, V. et al (2010). Enhanced LT algorithms with heterogeneous information – initial. Deliverable D4.1.2a, EUWB Integrated Project.

Poilinca, S., Dragomir, D., Wehner, M., Michler, O., Zeisberg, S. (2010). Concepts and possibilities of the positioning device and algorithm evaluation tool POSE for qualifying the suitability of ranging device/positioning algorithm combinations. The 7th Workshop on P ositioning Navigation and Communication (WPNC), p. 270 – 276, 2010, Dresden, Germany,

Legrand, I.C., Newman, H.B., Voicu, R., Cirstoiu, C., Grigoras, C., Toarta, M., Dobre, C. (2009) MonALISA: an agent based, dynamic service system to monitor, control and optimize grid based applications. *Computer Physics Communications*, 180(12), p. 2472-2498.

Acknowledgements

I would like to thank my academic supervisor Jun Hu and my company coach Pejman Hafezi for the pleasant collaboration we had as well as for their support and assistance. Moreover, I would like to thank the sound engineers Gerrit de Poortere, Jeroen Dreessen and Sylvain Choisel as they provided technical assistance and patiently addressed my numerous questions about sound.

Technische Universiteit
Eindhoven
University of Technology