Bournemouth University

Faculty of Science and Technology

BSc (Hons) Music and Audio Technology

May 2020

Binaural In-Ear Monitoring; Humanising Large Stage Performance

by

Michael Conor Dever

DISSERTATION DECLARATION

This Dissertation/Project Report is submitted in partial fulfillment of the requirements for an honours degree at Bournemouth University. I declare that this Dissertation/ Project Report is my own work and that it does not contravene any academic offence as specified in the University's regulations.

Retention

I agree that, should the University wish to retain it for reference purposes, a copy of my Dissertation/Project Report may be held by Bournemouth University normally for a period of 3 academic years. I understand that my Dissertation/Project Report may be destroyed once the retention period has expired. I am also aware that the University does not guarantee to retain this Dissertation/Project Report for any length of time (if at all) and that I have been advised to retain a copy for my future reference.

Confidentiality

I confirm that this Dissertation/Project Report does not contain information of a commercial or confidential nature or include personal information other than that which would normally be in the public domain unless the relevant permissions have been obtained. In particular, any information which identifies a particular individual's religious or political beliefs, information relating to their health, ethnicity, criminal history or personal life has been anonymised unless permission for its publication has been granted from the person to whom it relates.

Copyright

Requests for Information

I agree that this Dissertation/Project Report may be made available as the result of a request for information under the Freedom of Information Act.

Signed: *Michael Conor Dever*

Name: Michael Conor Dever

Date: 20/05/20

Programme: BSc MAT

Acknowledgements

I would like to thank the following people for their support during the development of this project. Your input, knowledge and help was invaluable.

Tom Davis
Darrell Gibson
Dave Haydon
Claire O'Toole
Isabel Orioni
Mandi Abrahams
Miriam Abrahams
Oz Jefferies
Andrew Frengley
Dinis Tavares
Ambrose Seddon
Chris Lyden
Alec Barrie
James Hipperson
James Spinks
Richard Fleming
Darren Stratford
Evan Costello
Joe Jackson

Table of Contents

Abstract

This dissertation outlines a Binaural In-Ear Monitoring system designed to alleviate the issues of isolation expressed by live performers as a result of background attenuation provided by current models of In-Ear Monitors (IEMs). The concept of dynamic monitor mixing is presented, with sensor input affecting parameters within the mix to mimic some of the real world characteristics of acoustic performance. Features were identified by analysing the problems with current models of IEMs, and the needs of musicians, ascertained from interviews and surveys. The system consists of a wearable gyroscopic headband that tracks the head orientation of the user and a distance tag. The tag interacts with Ultra-Wideband (UWB) Radar sensors deployed at the boundaries of the performance space. The hardware integrates with an application deployed within the visual programming language, Max 8 (Cycling74 2020) while using the Spatial Workstation Virtual Studio Technology (VST) plugin as an audio engine to provide 3D binaural audio. The system combines these elements to dynamically pan the mix as the user turns their head, while automatically adjusting the volume levels of sound sources in the mix as the user moves away or towards them, also mimicking elements of an acoustic environment. These dynamic changes allow the performer to experience the space around them, making the performance feel more natural.

0.0 Introduction

Monitoring within today's live industry is fulfilled by floor wedges and In-Ear Monitors (IEMs). The rise of IEMs has introduced the need for more refined monitor mixing. The background isolation offered by current models of IEMs is -20dB ambient attenuation or more, a similar experience to wearing self-isolating headphones. The attenuation can alienate the user from their surroundings. This means the engineer must create a complete listening experience for performers (Burton 2013). The musician relies heavily on the monitor engineer to provide a complete mix as it is the only way for them to hear the entire show (MusicRadar 2020). This development project was undertaken as an exploration into removing the isolation felt by musicians from their audience and fellow band members while their monitoring is fulfilled by IEMs.

1.0 The need for monitoring

In acoustic music, the ability to hear oneself is taken for granted. When a musician strums a chord on their guitar, they receive direct aural foldback from the resonant chamber of the instrument. Similarly once a musician begins to sing, they hear their own voice with perfect clarity. However, if we amplify the instrument, the musician will soon be unable to hear themselves sing. If one cannot hear themselves properly, then their performance will be significantly diminished, no matter their ability (Mellor 2005). This introduces the need for monitoring; a second audio system separate to the primary PA (Public Address) system. These systems range in complexity with some monitor mixes being fed with copies of channels from the front of house (FOH) console. Other systems use a dedicated monitor console capable of providing each performer with their ideal mix (White 2009, p. 1002). These mixes differ greatly from the FOH mix heard by the audience. They are tailored for each individual band member. With mixes primarily consisting of the instrument the musician is playing, if they sing - their vocal too. Backing instruments are usually quieter in the mix, consisting of the lead vocal, kick, snare and bass.

1.1 The History of In-Ear Monitor Development

The claim to the invention of in-ear monitors (IEMs) is disputed. Variations of homemade IEMs began surfacing in the early 1970’s (Verdugo 2015). Stephen Ambrose claims to have created the first iteration of IEMs in 1965 (Hertsens 2011). However, the first significant development in the creation of In-Ear Monitors is credited to Stevie Wonder’s FOH engineer, Chris Lindop (Burton 2013). Wonder was the first artist to experience the freedom of wireless monitoring. Not bound to a single location by a stage monitor. It was an invention that would go on to liberate artists.

During a 1987 Stevie Wonder concert in London, Lindop broadcast Wonder’s monitor mix on an FM radio wave. Wonder tuned in on a specially built FM receiver and received his monitor mix wirelessly. Lindop would go on to found Garwood Communications with Andrew Frengley, David Bowie’s monitor engineer, and electronics engineer Martin Noah. Together they went on to release the Personal Radio System, the first commercially available IEM system. For the purposes of this dissertation, a phone interview with Andrew Frengley was conducted to gain a first hand account of the early developments of the technology. He states:

“We were pushing it [the technology] around. Putting it in people's [ears]. And it was well received, But there were some real problems with it. It was hissy. [It] had some reliability issues. They were glitchy…there was a pressure to try and get synthesised frequencies i.e. the way things work [today], where you can tune content...we also knew that AKG Yeah, Sennheiser and Shure. We're now looking at this as a really viable kind of thing” (Full Extract - Appendix A) (Dever 2020).

These original devices used a single 15mm Sony driver with no isolation (ACS 2019), meaning users had to turn the earbuds up to dangerously high levels to hear themselves over the ambient sound from the stage. However by the time Lindop and Garwood had fixed all the issues with their models, Sennheisser and Shure released the first fully synthetically tuneable in-ear monitor systems in the mid 1990s, effectively killing Garwood with their superior systems.

Van Halen's FOH engineer, Jerry Harvey created the first dual-driver IEM in 1995 by tuning two drivers designed for a hearing aid into bass and treble (Hood 2017). By the early 2000s, the IEM system as we know it today was fully developed: -20dB self isolating earbuds, belt pack receiver and a synthetically tunable MHz (Mega Hertz) transmitter (Milne 2020).

1.2 In-Ear Monitor Systems

In-Ear Monitor (IEM) systems consist of three components: the transmitter, the receiver, and a pair of earbuds. The transmitter is placed off to the side of the stage in an equipment rack. An XLR cable is used to send the monitor mix from the console to the transmitter. The mix is sent wirelessly using RF (Radio Frequency) to a receiver, once they are tuned to the same frequency. The receiver is generally worn as a belt pack. The belt pack also has a volume knob to correct the playback level. Earbuds connect to the beltpack and deliver the monitor mix with clarity (Figure 1).

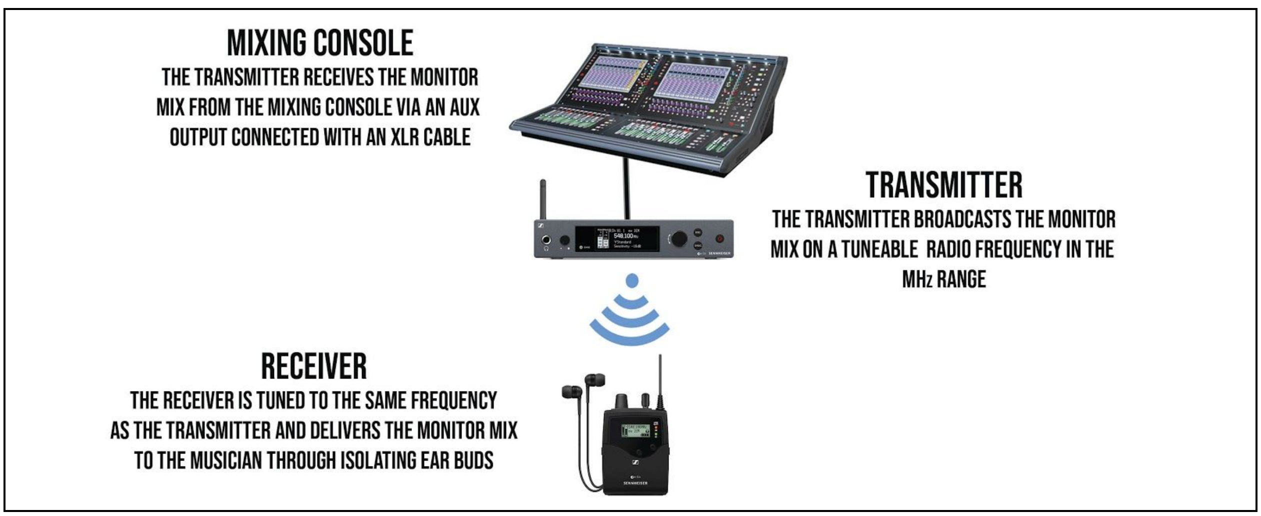


Figure 1: The signal flow of IEMs

1.3 Problems With Current IEM Technology

A number of popular musicians have expressed discomfort felt from isolation experienced due to the standard -20dB background noise attenuation. Some musicians feel that using IEMs deprives them of the feeling they get from the roar of the audience or the shouts of their fellow musicians (Grealy 2017). On large elevated festival stages where the performers are removed from the audience this problem is extenuated. Musicians complain of a loss of feeling, almost unnatural removal from their surroundings.

Research into this topic has primarily relied on interviews with musicians conducted by third parties regarding their monitoring needs. The primary example of this are the issues raised by Jack Steadman in a 2013 interview with Jon Burton in Sound on Sound magazine. He states:

"I find them helpful but don't enjoy them. They can be inconsistent and you need a good engineer to run them. I also feel alienated from the gig. Without IEMs in, I can just walk closer to something on stage to hear it louder; if I'm on in-ears, I feel a loss of control."

Fellow musicians and engineers share Steadman's issues. Drummer Suren de Samam states in the same article: *"you do tend to feel quite isolated from the crowd with in-ears, particularly on bigger stages"*. Ambient microphones are supporting microphones that provide crowd noise. However Sarem goes on to say: *"[Ambient Mics] change your whole mix, and having ambient mics turned up loud with a crowd clapping out of time when you're trying to keep in with a click can be difficult!"* (Burton 2013). Articles from related industry magazines also express these sentiments. Audiologist Keith Gordon lists isolation as one of the primary cons of IEMs in his article *In-Ear Monitors: Tips of the Trade* (2020):

"isolation from external sound that provides hearing protection also removes the connection to the outside world, such as audience noise, and upsets the traditional feel of being onstage...*In a larger room, you'll soon find that your artist may feel isolated. This is very common; in-ears, by design, offer exceptional ambient noise reduction, which in turn can make a player feel cut off from the world around them"* (Gordon 2020).

2.0 Defining a Solution

This dissertation explains the research and development process of BIEMS: the Binaural In-Ear Monitoring System (BIEMS). BIEMS was designed as an additional bundle of hardware and software to integrate with existing In-Ear Monitoring systems. The original proposal raised the prospect of system users wearing a combination of sensors to track both their head orientation and distance relative to other musicians on a stage. The primary object of this was to provide volume and binaural panning attenuation automatically and dynamically to channels in the mix. Changes occur as a result of a change in head orientation; the software pans sound sources automatically following the motion of the head. Volume attenuation is used to simulate changes in volume in a natural environment with volume level changes occurring with movement to or from a sound source. The system was designed as a solution to the problems of isolation incurred from using IEMs. A combination of binaural spatialisation and psycho-acoustics techniques were utilised to mimic the processes of performing in a unified space (Zea 2012).

2.1 Binaural Audio

Binaural techniques simulate the hearing cues created by acoustic interaction between our bodies and our environment. Everyone has an individual pattern of hearing cues that are created by their unique body shape. These cues change as the listener moves (BBC 2012). A key attribution to the way humans hear is the ability to detect where a sound is coming from, referred to as localization (McAnally and Martin 2014).

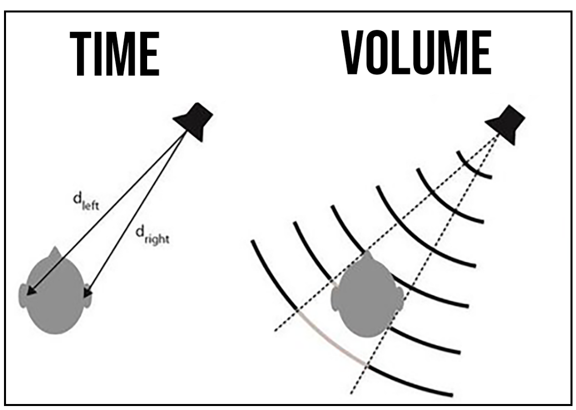


Figure 2: *Localisation of sounds through interaural time difference (ITD) and interaural level differences (ITV)*

Humans localise sources in three ways, time differences (the time taken by the sound to reach the ear), volume differences (Figure 2) and the shape of the head. The architecture of our anatomy dictates how we understand the sounds we hear (Geronazzo et al. 2013). This phenomenon is known as a Head Related Transfer Function (HRTF). HRTFs represent the acoustic filtering that takes place when sound waves travel towards a listener, reflecting off and diffracting around the person's head, face, and external ears (pinnae) before entering the ear canals (Figure 3) (Mokhtari et al. 2008). Binaural microphones (Figure 4) are used to capture HRFTs that can be applied while decoding audio to binaural format (Lalwani 2012).

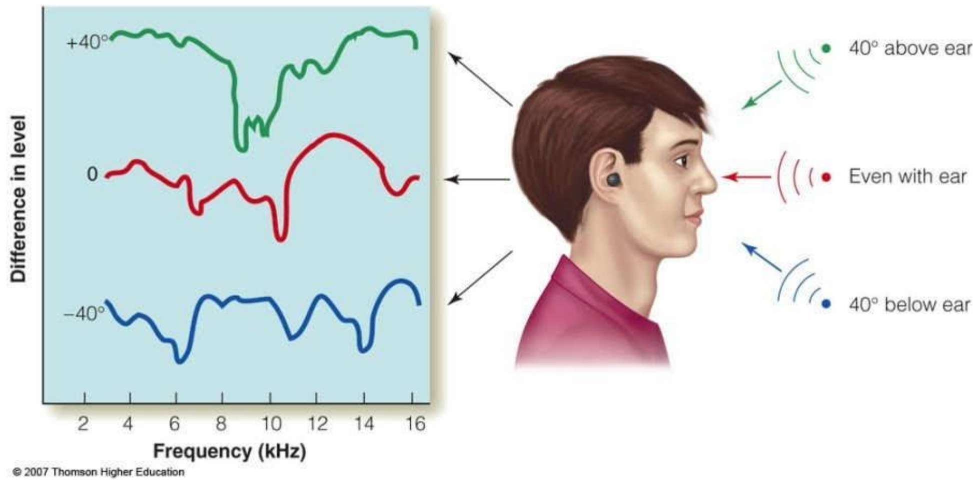


Figure 3: *An example of the acoustic filtering that occurs from the shape of the head and ears at different elevations* (Joe 2020)

Figure 4: *Binaural microphones are utilised to create accurate spatial recordings (Neumann 2020)*

Binaural audio integrates these three factors in the decoding of the audio. It utilises the same two channel L and R format of stereo. The key difference being, stereo does not factor in the natural ear spacing or "head shadow" of the head and ears (Smith 2017). Binaural decoding utilises the three hearing cues allowing for sound localisation 360° around the head in a headphone environment (Figure 5).

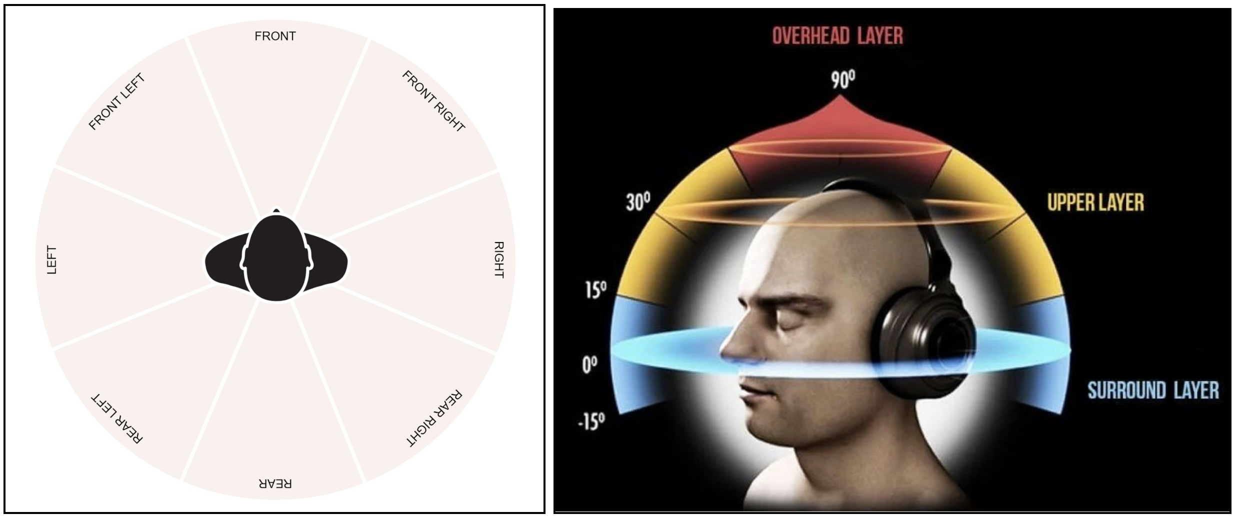


Figure 5: *A visual representation of the 360-degree audible sphere that binaural audio creates around the head* (Braganza 2020)

2.2 Key Research

Binaural technology has seen a sharp uptake in recent years, with many Virtual Reality developers recognising its potential to mimic natural acoustics in a headphone environment (Geronazzo et al 2012). Examining how this technology could be applied to the live industry was the initial primary focus of this dissertation. Research commenced with examining commercial and academic solutions utilising the technology.

KLANG is a German start-up that was recently purchased by industry giant AudioTonix, owners of DiGiCo, Allen & Heath, SSL and Calrec (Hope 2018). KLANG completely changed the landscape of monitoring when they released KLANG:fabrik, the first binaural in-ear monitoring system. KLANG revolutionised the industry by proving binaural technology could offer a real alternative to stereo monitoring (Bennett 2019).

Future examples are available from academic sources. In his 2012 paper B*inaural In-Ear Monitoring of acoustic instruments in live music performance,* Elis Zea describes a method similar to the one described in this dissertation. He presents a binaural monitoring system that utilises a motion camera to track user movement, a sine sweep to measure the impulse response of the performance space. These factors allow the volume to change as users move away from one another and create an accurate representation of the sound of the room. While this system is technically very expansive and highly detailed, one can argue it would not have a commercial application. It is too complicated. In a time-demanding stressful live environment, it would be too time consuming to set up every show. Likewise, KLANG doesn't go far enough. While it is a very good system that already has shown its commercial potential, the only difference between it and standard IEM setup is binaural panning. It does not utilise the full potential of binaural technology and does not at all simulate the space the user is occupying.

2.3 Survey of Key Contacts

Identifying features for BIEMS began by examining the results of a survey of first-party key contacts. The survey asked a variety of questions, around the topic of binaural technology in the context of live performance. It was sent to approximately 30 people. Key contacts were identified as persons employed within the live music industry or by companies that supply the industry. There was only a 20% uptake but these six respondents provided high quality information based on personal experience. Industry roles represented include audio equipment distribution, PA amplification, PA manufacturing to freelance monitor and FOH engineers. The full results of the survey in Appendix B

Question 3's results (Figure 6) show there is some opposition to the technology with 25% of respondents disagreeing to the question *"Do you think musicians would benefit from interactivity in their monitoring?"*. However the majority of respondents expressed they are open to the technology. A R&D Engineer working for Funktion-One Research Ltd. provides a more detailed answer. They responded with: *"To a degree, yes, however in my experience many musicians do not know what to do to their mix to achieve a certain result. Mixing IEMs is far more than volume adjustment"*.

When asked *"What specific features would you like to see in a Binaural IEM system?"*. The reponses proved incredibly beneficial in identifying key features that users would like. A Sales Manager from XTA Electronics who has worked with KLANG, the market leader in the binaural audio space, states: *"The only thing not being done at present is sensing of the listener's head position, however hearing aid technology is way professional audio in-ears".* Head orientation was deemed a necessary feature of BIEM, with audio panning as the user moves their head. A touring sound engineer provides another key feature, stating his desire for a *"complete control of all inputs, including in particular FX returns, with regard to their position and depth in the binaural space".* These answers provided a small insight into the desires of engineers and technologists alike. The responses set the project clear objectives.

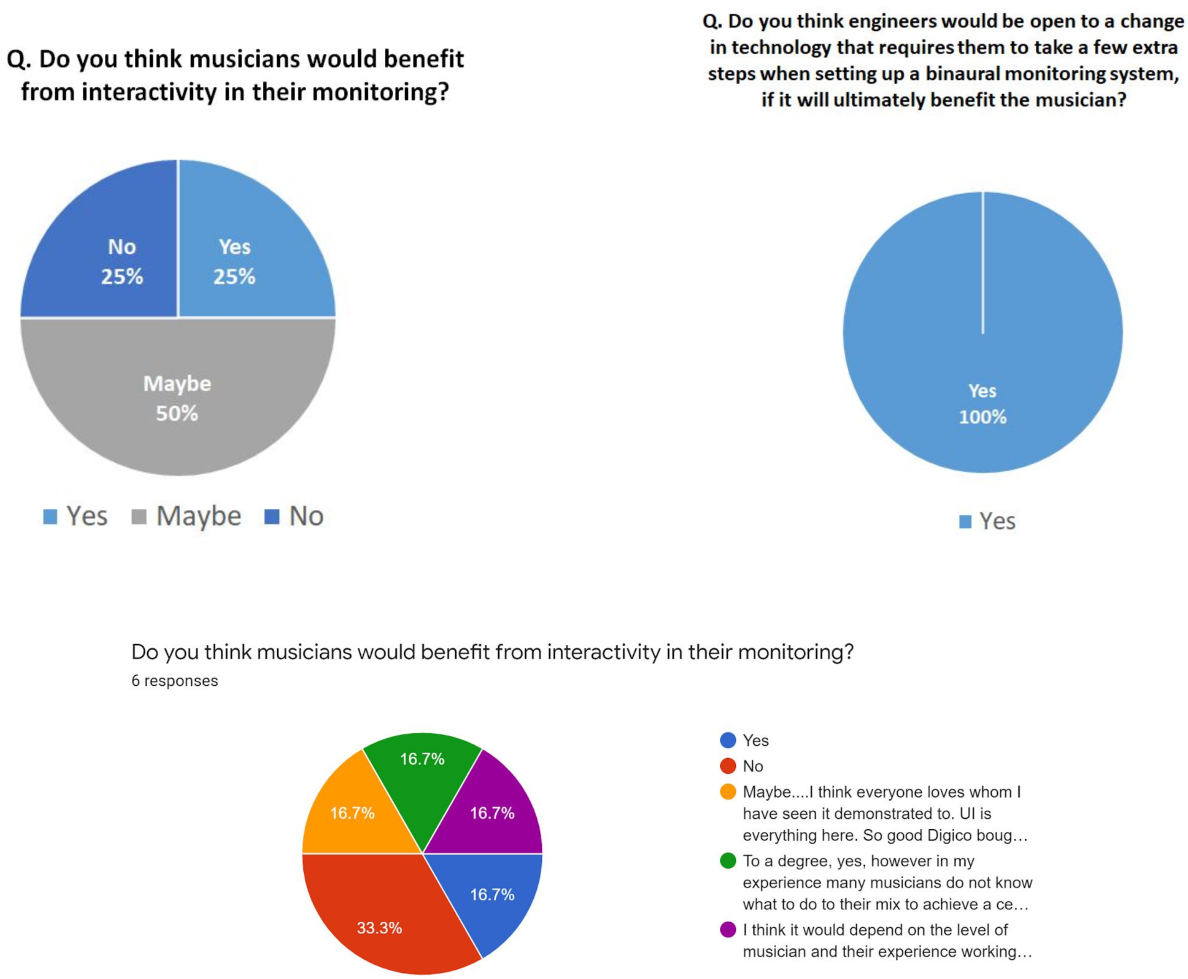


FIgure 6: *Responses from a survey of key industry contacts (Detailed Breakdown - Appendix B)*

2.4 Research Topics

The research was separated into three key topics: Head Orientation, tracking the head orientation of the musician; Distance Tracking, calculating the relative distance between musicians; and Binaural Processing Protocol, selecting which form of decoding and HRTF application to use. Each element required its own research.

2.4.1 Head Orientation

There are a number of technologies that can facilitate head tracking . The primary method is a wearable gyroscope, measuring the yaw, pitch and roll (Jasiewicz et al. 2007) of the head.

Assisted driving technology proved to be the most fruitful area of study with a 2017 journal article on *'An Orientation Sensor-Based Head Tracking System for Driver Behaviour Monitoring'* describing a method for measuring the orientation of a driver's head orientation as they complete driving scenarios. Zhao et al. utilised a BNO-055 high-fidelity 9-axis gyroscope and magnetometer. They it describe as:

"A [logic chip] designed for high-fidelity navigation applications in portable devices. It includes three triaxial sensors for measuring acceleration, rotation speed and magnetic fields, respectively…The logic chip also compensates for the effect of temperature on the sensors and automatically calibrates them".

The on-board smoothing provided by this sensor was a highly desired feature as it didn't require supplemental coding to smooth data within Max 8. The decision to use this sensor relied further on proceedings from the 2017 Academic High Altitude Conference. They measured the orientation of a weather balloon, comparing results from leading market sensors. They noted *"all three IMU's were within 1° of each other, the price of the BNO055 was a fraction of the cost with comparable results*" (Zhaoi et al. 2017). These considerations led to the selection of the BNO055 as the wearable gyroscope sensor.

2.4.2 Methods of Tracking Distance on Stage

Tracking users as they move on-stage proved to be a difficult problem to solve (Appendix C). Ultra-Wideband Radar (UWB) emerged as the most feasible technology to use within this project. It is the most accurate of the three technologies considered (Appendix C) and can penetrate stage lights and decorations. None but the most solid or metallic objects would interfere with the target detection process.

UWB is a radio technology that can use a very low energy level for short-range, high-bandwidth communications over a large portion of the radio spectrum. UWB has traditional applications in non-cooperative radar imaging (SangHyun Chang et al. 2009). UWB radio waves differ from traditional narrowband radio waves, *"[traditional waves have] a narrow frequency range and use harmonic (sinusoidal) or similar quasiharmonic signals"* to transmit information. UWB radio systems also use sinusoidal signals as carrying waves to transmit data, with a key difference. UWB Radar sends information as sectioned, short pulses instead of one transmission. The reduction in signal length allows for improved detected target range measurement accuracy, improved stability observing targets at low elevation angles and increases the probability of target detection and improved stability observing a target (Immoreev and Fedotov 2002). These improvements have allowed for many applications of the technology to in-door range finding and target detection (Chang et al. 2010).

Out Board is a UK-based company which utilises UWB Radar technology in their TiMax TrackerD4 performer (Out Board 2020) stage tracking system. They were contacted for advice on building a budget-conscious UWB Radar system. Company director, Dave Haydon, suggested it would be too difficult to build a system from scratch. Haydon liked the hypothesis of this dissertation and wanted to support the project. A full TiMax Tracker D4 system was loaned to the project by Out Board.

FIgure 7: *TiMax Sensor* ***Figure 8:*** *TiMax Tag*

The TiMax TrackerD4 system was key to solving this complicated and expensive issue. The tracking system has previously been deployed for live performance audio spatialisation, as well as moving-head light control. The TiMax system is impressive, and full credit is deserved to the Out Board team, as they've achieved super-accuracy up to 10cm at 60-80m distance. It utilises between four and eight sensors (Figure 7) distributed around the tracked area for redundancy (only two sensors need to see a tag for tracking to occur); these receive short UWB pulses from tags worn by the tracker subject and use Angle of Arrival (AoA) and Time Difference of Arrival (TDoA) analysis (Sheun et al. 2019) to locate the tag (Figure 8). They are connected together via Cat5e cable to a distributing Timing Hub and POE switch plugged into a laptop running the TiMax Location Engine (Figure 9) and Tracker Translate application (Figure 10). The latter program outputs cartesian *x,y,z* coordinates (Khan 2020) via Open Sound Control (OSC) to Max 8. OSC is protocol that sends messages over Wi-Fi as UDP (User Datagram Protocol) packets (Phillips 2008).

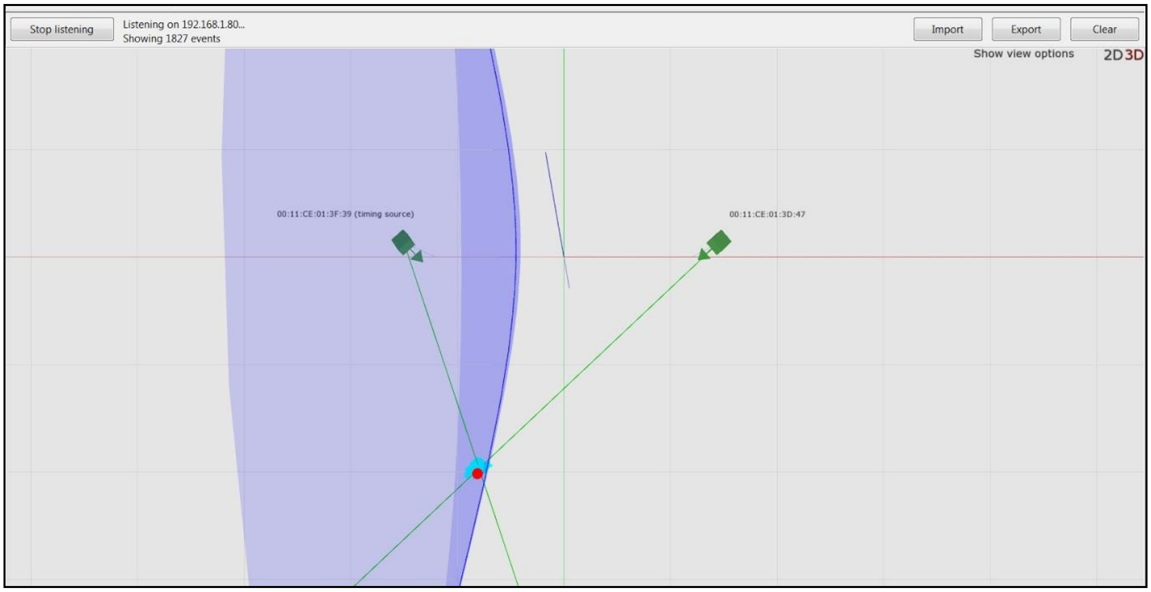


FIgure 9: *TiMax tracking system, the worn tag is represented by the red dot and the sensors are represented by the green icons, the arrow at their front represents their orientation.*

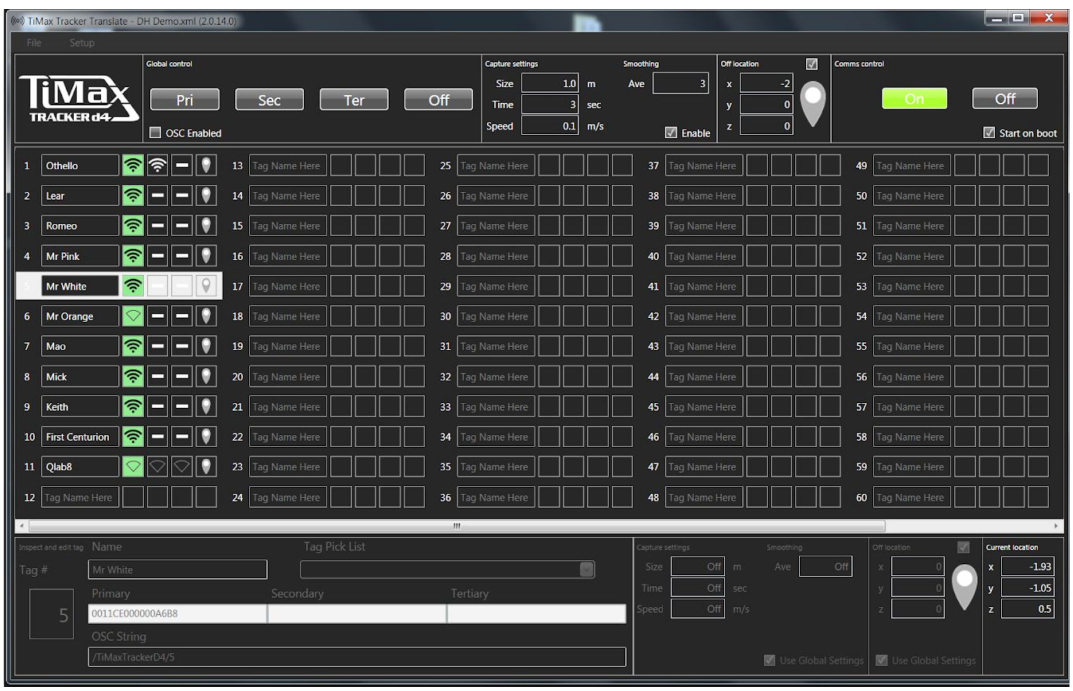


FIgure 10: *TiMax Translate outputs the cartesian xyz coordinates over OSC to a specified IP address*

2.4.3 Binaural Processing

The binaural processing engine was carefully selected, as its level of quality and consistency would affect the primary feature of the system, the audio. The engine would have to be deployed within Max 8, so initially the Institute for Computer Music and Sound Technology (ICST) Ambisonics package for Max 8 was considered. There is a strong case for its use presented in a 2016 paper by D. Rudrich et al. The paper described a method of deploying the ICST Tools to control spatial audio effects for live performance. The latter closely matches the application of BIEMS. Upon initial testing the GUI provided to be accessible, practical and easy to use. However, the audio engine clicked even under low stress testing of two audio sources. The HRTF also proved to be inadequate.

Other considerations were made under the guidance of academic tutorials that have experience with binaural technologies. SPAT (Ircam 2020), IEM Suite (Rudrich 2020) and Spatial Workstation (SW) (Facebook 2020) were all considered and tested. SW had a far superior audio engine to the others and a realistic HRTF. It is developed by Facebook and has received recognition for the quality of the VST plugin (SoundOnSound, 2017). However, the GUI only allowed for one source per channel. Therefore an inventive solution was needed to solve the problem. That came in the form of using the ICST Ambisonics ambimon object (Schacher 2020), scaling the output to a format that was receivable within the SW VST. The ambimon was used as a control mechanism and was not connected to the audio output. This solution delivered high-fidelity audio with a great user experience in an easy to use GUI (Full System Description - Appendix D).

3.0 Development and Implementation

The development and implementation of BIEMS was a stage process that required a lot of testing at each stage of the project. Much like the research stage of the project these were broken down along key lines as described below.

3.1 Input and Output System

The intended application of this system is primarily live sound. Hence the sensible option was to adopt an audio protocol with widespread use within the market. Danté is used by industry leaders Yahamha, Midas, DiGiCo, Allen & Heath and SSL among others (Capps 2020). Due to the Covid-19 pandemic it could not be tested further. Instead, audio sends were simulated using Soundgrid (Figure 11), a program developed by Waves and Digico (Hope 2019). It allowed multitracks to be streamed from the Digital Audio Workstation (DAW) to Max 8. Latency could not be effectively tested with this set up. Future simulations will include a complete monitor mixing console and IEM rig, with audio being sent from the console to the application running on a PC, and back again. This would be a much more effective way to measure true latency for the system's intended purpose. The outputs would then be connected to the rear of the console and connected to the IEM transmitted. The artist would receive a full dynamic binaural monitor mix.

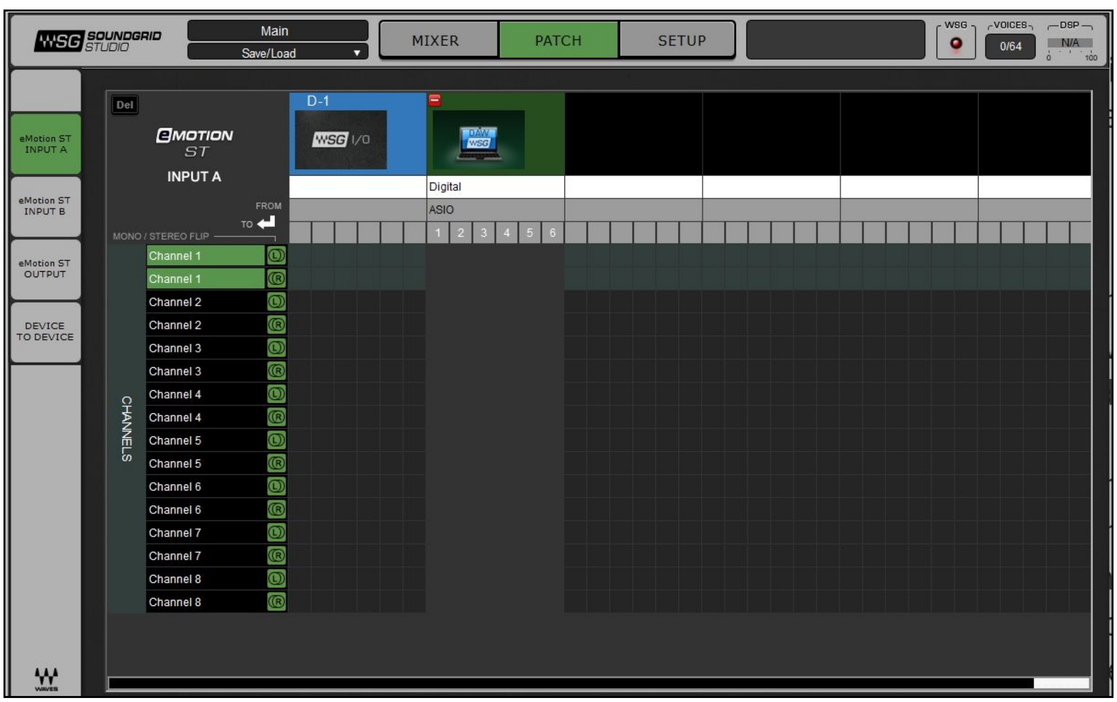


FIgure 11: *SoundGrid Studio software patch window*

3.2 Developed Hardware

The hardware is a low-cost prototype consisting of an Arduino MKR board, a BNO-055 IMU Shield gyroscope, a TiMax D4 Tracker and a 9V battery pack (Datasheet and Schematic - Appendix E), attached to a felt hair band that has been cut to reduce its size (Figure 12).

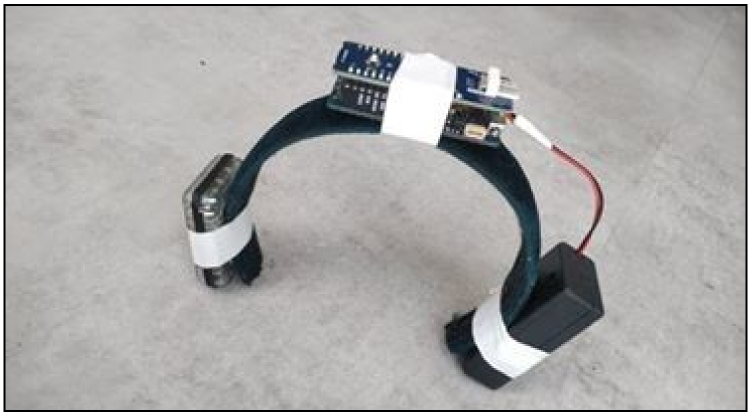


FIgure 12: *the BIEMS weable headband*

The Arduino is programmed to connect to a router that acts as a master server and also receives data from the TiMax sensors. This data is transmitted wirelessly to Max 8. The transmission of OSC messages relies on C code (Figure 13) running on the Arduino's SAMD12 microprocessor (Appendix F). The Arduino connects to the Wi-Fi network on start-up and then sends OSC messages to a specified IP of the computer running the Max 8 program (Full Code - Appendix D).

```
void loop() {
  float heading, roll, pitch;
  if (IMU.eulerAnglesAvailable()) {
    IMU.readEulerAngles(heading, roll, pitch);

    OSCMessage msg("/h/");

    msg.add(heading);
    msg.add("/r/");
    msg.add(roll);
    msg.add("/p/");
    msg.add(pitch);
    Udp.beginPacket(outIp, outPort);
    msg.send(Udp);
    Udp.endPacket();
    msg.empty();

  }
}
```

Figure 13: *The Arduino void sends OSC messages to a specified address (Appendix D)*

3.3 Hardware & Software Setup

Engineers work under intense pressure, with many experiencing burn-out (Vangelova 2008). Generally the reliance on complex technology within the industry creates unforeseen issues and introduces complexity that requires more competence and work from engineers. The TiMaxD4 Distance tracking system (see Section 2.6) requires three to six sensors to be placed at the boundaries of the stage. This is the most intensive requirement of the system. However. it may not be an unreasonable demand as modern stages require networking infrastructure and extra PCs to facilitate Danté networks and multitrack recording.

The BIEMS GUI was designed to simplify the learning curve required to operate it. Many engineers will be familiar with the Ableton channel strips used for level control within BIEMS. Having a Human Computer Interface (HCI) that shares similarities and integrates concepts familiar to the target audience is likely to have more adoption than systems that develop a new approach and discard traditional layouts (Hartson and Hix 1989).

Therefore the software system is as simple as possible. Requiring only four actions from the engineer upon startup of the system.

1. The orientation system features a calibration offset. For accurate results the gyroscope needs to be calibrated by the engineer. This is a simple process. The engineer stands facing the front of the stage. He or she presses the Calibrate Offset button (gray - X) to the right of the GUI and now 0° from the gyroscope output will align with the front of the stage (Figure 14).

2. The second task is to input the room size. This measurement is available in feet or meters. It sets the width of the Monitor (Figure 15) in m^2 to the smallest side of the room. This represents boundaries of the sound field 360° around the user's head. A popup message informs the user of the distance between the circular white lines (smallest side / 10 = line spacing distance).

3. The third condition is selecting the number of inputs required, there are eight available. The matrix located to the left of the monitor controls which channels are affected by the sensor data. If the channel is unselected it will remain stationary, if it is red it will move and attenuate via the sensor input.

4. The final task is to place the channels in relation to the user's current position and orientation. Example: if the user is facing the front of the stage and the drummer was 2M to their rear you would move that channel via a mouse click to 180° on the

ambition on the white line representing 2M (count the white lines from the center, applying the formula to calculate the spacing between lines (smallest side / 10 = line spacing distance) or by viewing the pop-up message.

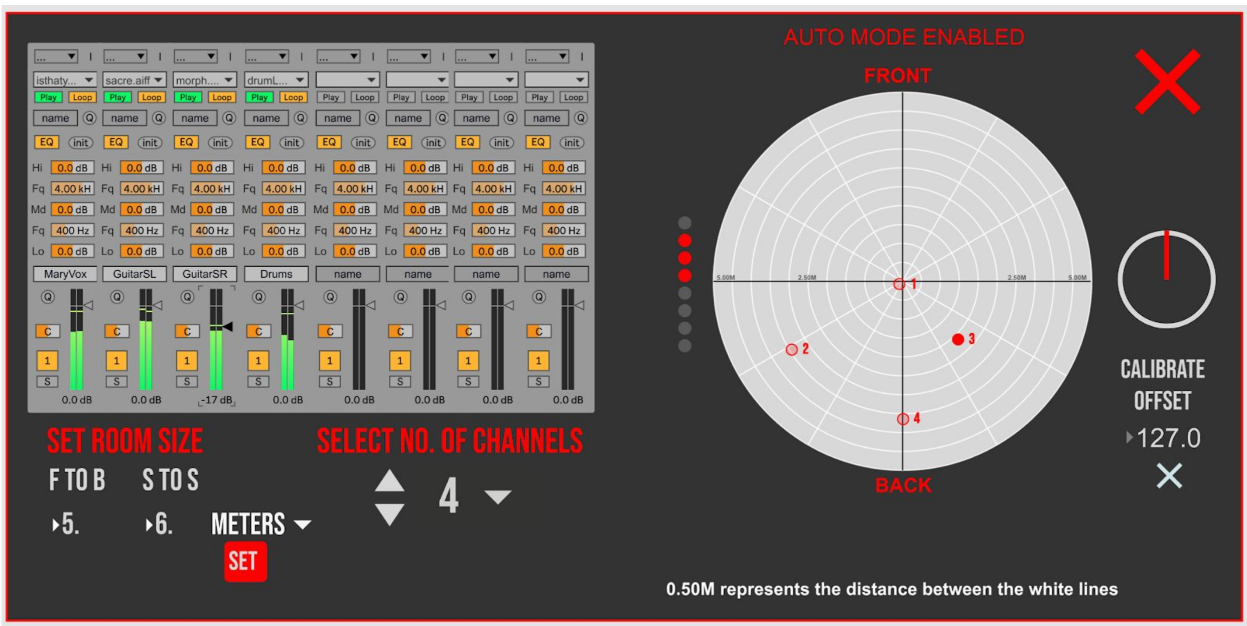


Figure 14: *The BIEMS GUI calibrated with an offset of 127°*

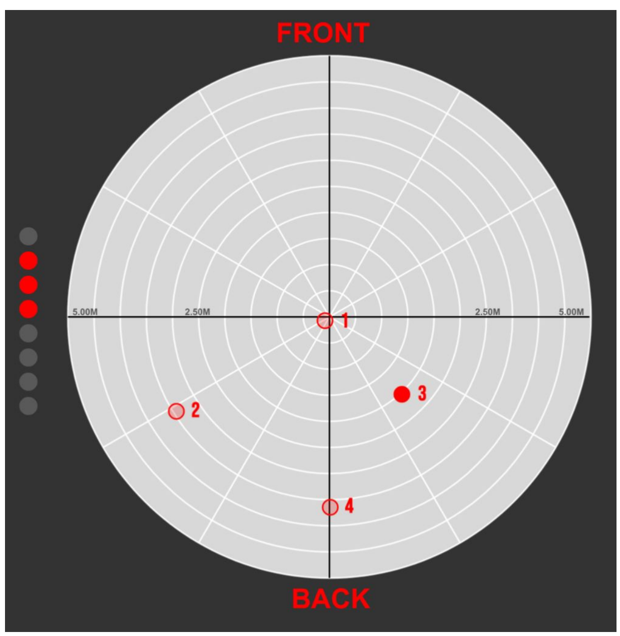


Figure 15: *The Monitor for placing sound sources*

Once the engineer has placed all the sounds, the X in the upper right of the GUI enables Auto Mode. This will commence the auto-panning and dynamic volume changes (Full Description - Appendix D) to the channels highlighted in the matrix.

3.4 Orientation & Distance Calculations

The orientation system functions on simple addition and modulo equations. The raw gyroscope input is added to the calibration offset value. This value is then added to the starting orientation value as specified by the placement of the sound source relative to the user. This value is entered into a modulo equation (Khan 2020) which does not allow it to fall outside the boundaries of 0-360, therefore if the value reaches 370 it will be scaled to be 10 (Figure 16). There is a system of switches and gates that ensure there is no false input to the system and sensor data errors when toggling modes. They capture values before they enter the Monitor and only output the channels selected by the matrix. The smooth of this data via the modulo increases accuracy, ensuring better audial replication.

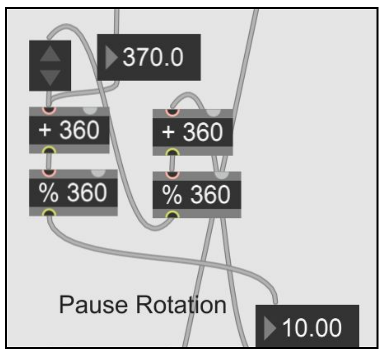


Figure 16: *The cartopol object in Max 8*

The Monitor accepts polar *Azimuth, Elevation and Distance (AED)* values, whereas the TiMax system outputs cartesian *xyz* values. Therefore the TiMax values are converted to polar coordinates via the cartopol object within Max 8 (Figure 17). This object uses the conversion equation $Tan^{-1}(Y/X)$ and outputs the amplitude and angle of the polar equivalent values (Lyon 2012, p.213-214). However, the amplitude is the only value used within the distance system as the azimuth (angle) is calculated by the orientation system (Full Description - Appendix D).

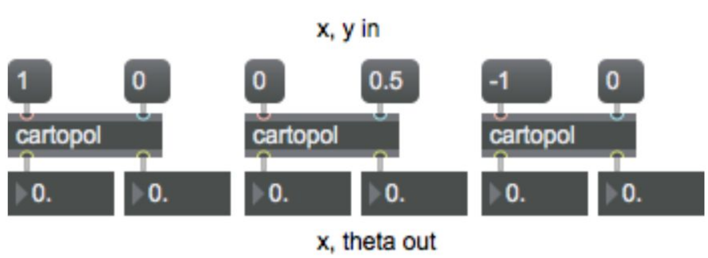


Figure 17: *The cartopol object in Max 8*

3.5 Spatial Workstation Implementation

The data output from the orientation calculation had to be scaled from 0-360 to -1 to 1 to match the Spatial Workstation (SW) input. The distance formula was scaled to set the room boundaries of the SW audio engine. This introduces slight reflections to mimic the acoustic properties of a room. These settings can be tweaked to add or remove extra reflections to increase or decrease the effect of this feature.

The second major feature of the SW implementation is the volume attenuation of sound sources based on their distance to the user. The original intention was to base the volume changes on the inverse square law (Khan 2020) and the SW does use this method, however it can be tweaked to preference.

3.6 Methodology & Coding Validation

The methodology originally selected for the project was the V-Model method. It is an extension of the famed waterfall method (Balaji and Sundararajan Murugaiyan 2012). It is also known as the "Verification and Validation" model. Verification phases are on the left side of the V, while the Validation phases are on the right. Verification and Validation phases are joined by a coding/development phase in a V-shape (Figure 18). The validation section is a testing phase completed after each development stage (Kumar 2019).

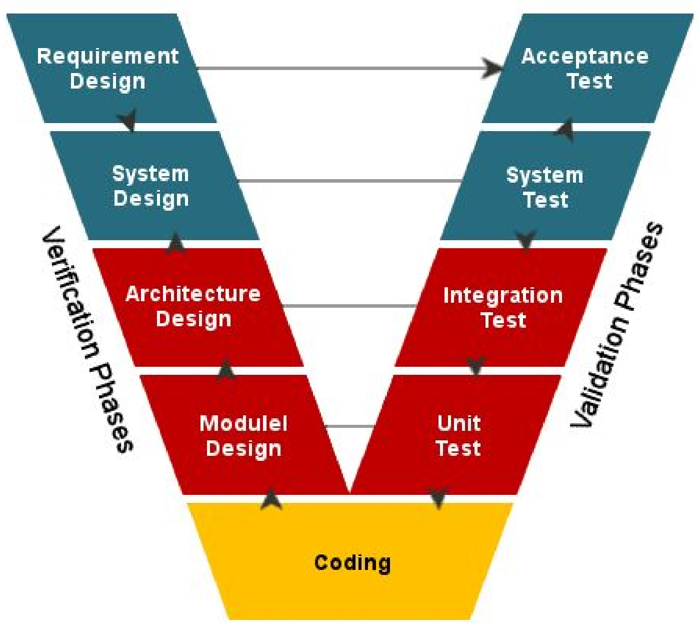


Figure 18: *V-Model Methodology Diagram (TestBytes 2019)*

The complete methodology (Appendix F) outlines five phases in total. The V-Model methodology was well suited to our segmented approach to research and development. Four phases were completed with some delays occurring in the Head Orientation and Distance Tracking calculation validation stages. These delays were caused by bad code, which ultimately led to the system being simplified. Two simple equations were the answer, as most software bugs “arise almost entirely from complexity” (Kanat-Alexander 2012, p.1). The project suffered and experienced teething problems in these early stages, however these delays coding malfunctions lead to more subject knowledge and ultimately a better product.

4.0 Conclusions

The original aim of this project as set out in the proposal was to develop an in-ear monitoring system that removes the isolation felt by musicians when using traditional IEM systems. The developed system was to feature binaural panning techniques (Zea 2012), creating a dynamic mix that changes based on the orientation and proximity (Pulkki 2001) of the performer to their fellow musicians on stage. This aim remained throughout the project

Two out of three of the original objectives were completed in their entirety. However, the final object remains incomplete. This was a public testing phase, where local bands would be invited to perform using the equipment. However, due to Covid-19 this public testing was not able to take place. Instead, multitrack recordings of live sets were put through the system. This testing revealed that this system can run 8 channels of audio with clarity when delivered from a Digital Audio Workstation (DAW) running on the same computer. Measuring latency while using this setup is not an accurate real world demonstration of the system's intended use. In a more developed model, audio would ideally be received from a console, which would perform the bulk of the processing, reducing the load on the PC running the BIEMS software. Therefore, no accurate testing of latency took place. The multitracks also helped in tweaking volume automation when moving between sound sources to mimic further realism. Further testing using this method demonstrated the need for additional gates and logic for data smoothing on the output of the Monitor (Appendix G).

While pleased with the results of this initial study and development, more research is needed in order to achieve a more natural sounding binaural environment. The current system is not robust enough for a live sound performance. This is due to a lack of real world testing capabilities. However, this is something that will be expanded on in the future. Better sensor accuracy is required. One way to achieve this may be using two gyroscopes and averaging the output between them to account for errors. Likewise, increasing the number of boundary TiMax sensors may help distance accuracy.

The field of binaural audio is ripe, in an increasingly technology dependent world, there will be a growing demand for accurate and natural audio replication, not just for recreational Virtual Reality (VR) applications, but also simulations such as pilot training and first responder training. These applications demand high-fidelity accurate binaural audio solutions that are not currently developed to their full potential. Initial results indicate that there would be commercial value in developing this prototype further. Therefore, development of BIEMS will continue.

References

ACS., 2019. The History of In Ear Monitoring with Andy Shiach. in person. Available from: youtube.com/watch?v=0SJhrZoyKQw [Accessed 12 May 2020].

Arduino, 2020. Arduino MKR WiFi 1010 | Arduino Official Store [online]. Store.Arduino .cc. Available from: https://store.Arduino .cc/Arduino -mkr-wifi-1010 [Accessed 11 May 2020].

Ashdown, D., 2018. Biffy Clyro out on Acoustic Tour [online]. Ashdown Engineering. Available from: https://ashdownmusic.com/blogs/news/biffy-clyro-out-on-acoustic-tour [Accessed 26 Apr 2020].

Balaji, S. and Sundararajan Murugaiyan, M., 2012. WATEERFALLVs V-MODEL Vs AGILE: A COMPARATIVE STUDY ON SDLC. International Journal of Information Technology and Business Management [online], 1 (1). Available from: http://jitbm.com/Volume2No1/waterfall.pdf [Accessed 5 Nov 2019].

BBC, 2012. Binaural Sound - BBC R&D [online]. Bbc.co.uk. Available from: https://www.bbc.co.uk/rd/projects/binaural-broadcasting [Accessed 26 Apr 2020].

Bennett, S., 2019. Lifting the lid on DiGiCo/KLANG's immersive IEM integration - Audio Media International [online]. Audio Media International. Available from: https://www.audiomediainternational.com/2019/06/26/lifting-the-lid-on-digico-klangs-immersive-iem-integration/ [Accessed 26 Apr 2020].

Braganza, B., 2020. Understanding Binaural Audio [online]. Headphone Zone. Available from: https://www.headphonezone.in/blogs/audiophile-guide/understanding-binaural-audio [Accessed 26 Apr 2020].

Burton, J., 2013. An Introduction To In-ear Monitoring [online]. Soundonsound. Available from: https://www.soundonsound.com/techniques/introduction-ear-monitoring [Accessed 22 Apr 2020].

Caps, S., 2020. Audinate Resonates Despite Downturn - FNArena [online]. FNArena. Available from: https://www.fnarena.com/index.php/2020/04/09/audinate-resonates-despite-downturn/ [Accessed 4 May 2020].

Cay, E., Mert, Y., Bahcetepe, A., Akyazi, B. and Ogrenci, A., 2017. Beacons for indoor positioning. 2017 International Conference on Engineering and Technology (ICET) [online]. Available from: https://ieeexplore.ieee.org/document/8308143 [Accessed 3 May 2020].

Chang, S., Sharan, R., Wolf, M., Mitsumoto, N. and Burdick, J., 2009. UWB radar-based human target tracking. 2009 IEEE Radar Conference [online]. Available from: https://ieeexplore.ieee.org/abstract/document/4977001 [Accessed 3 May 2020].

Cycling74, 2020. Max 8 Documentation [online]. Docs.cycling74.com. Available from: https://docs.cycling74.com/max8?contentp=Max&contentg=vignettes [Accessed 12 May 2020].

Daggett, W., 2016. Westone introduces new, ambient in-ear monitors at NAMM Show 2016 [online]. video. Available from: https://www.youtube.com/watch?v=9L-H6T5zCyY [Accessed 26 Apr 2020].

Dever, M.C., 2020. A telephone interview with Andrew Frengley of Garwood Communications. Telephone.

Facebook, 2020. *Spatial Workstation* [online]. Facebook360.fb.com. Available from: https://facebook360.fb.com/spatial-workstation/ [Accessed 19 May 2020].

Farinella, D., 2005. Future Sonics. FOH Online [online], 2005, p. 1. Available from: https://web.archive.org/web/20120229154503/http://www.fohonline.com/index2.php?option=com_content&do_pdf=1&id=290 [Accessed 22 Apr 2020].

Frink, M., 2006. Hybrid Monitor Mixing [online]. Mixonline. Available from: https://www.mixonline.com/live-sound/hybrid-monitor-mixing-368975 [Accessed 26 Apr 2020].

Geronazzo, M., Rosenkvist, A., Eriksen, D., Markmann-Hansen, C., Køhlert, J., Valimaa, M., Vittrup, M. and Serafin, S., 2019. Creating an Audio Story with Interactive Binaural Rendering in Virtual Reality. Wireless Communications and Mobile Computing [online], 2019, 1-14. Available from: https://www.hindawi.com/journals/wcmc/2019/1463204/ [Accessed 26 Apr 2020].

Geronazzo, M., Spagnol, S. and Avanzini, F., 2013. Mixed structural modeling of head-related transfer functions for customized binaural audio delivery. 2013 18th International Conference on Digital Signal Processing (DSP) [online]. Available from: https://ieeexplore.ieee.org/abstract/document/6622764 [Accessed 26 Apr 2020].

Grealy, J., 2017. In Ear Monitoring – Getting It Right. Leeds Plaza 2017 [video]. Available from: https://www.youtube.com/watch?v=ucC5G4M_4rU [Accessed 26 Apr 2020].

Hartson, H. and Hix, D., 1989. Human-computer interface development: concepts and systems for its management. ACM Computing Surveys (CSUR) [online], 21 (1), 5-92. Available from: https://doi.org/10.1145/62029.62031 [Accessed 11 May 2020].

Hertsens, T., 2011. Inventor Stephen Ambrose Claims Development of Safer Ear Tip [online]. InnerFidelity. Available from: https://www.innerfidelity.com/content/inventor-stephan-ambrose-claims-development-safer-ear-tip [Accessed 22 Apr 2020].

Hood, P., 2017. You Can Say Thanks To Alex Van Halen [online]. Drummagazine.com. Available from: https://drummagazine.com/you-can-say-thanks-to-alex-van-halen/ [Accessed 22 Apr 2020].

Hope, F., 2018. Audiotonix and Digico acquire KLANG: technologies [online]. PSNEurope. Available from: https://www.psneurope.com/business/audiotonix-digico-klang-technologies [Accessed 26 Apr 2020].

Hope, F., 2019. On tour with Lizzo: The artist's audio engineers take us through their set up - PSNEurope [online]. PSNEurope. Available from: https://www.psneurope.com/live/lizzo-tour [Accessed 4 May 2020].

Immoreev, I. and Fedotov, P., 2002. Ultra wideband radar systems: advantages and disadvantages. 2002 IEEE Conference on Ultra Wideband Systems and Technologies (IEEE Cat. No.02EX580) [online]. Available from: https://www.researchgate.net/publication/3950418_Ultra_wideband_radar_systems_advantages_and_disadvantages [Accessed 10 Jan 2020].

Ircam, 2020. *Spat | Ircam Forum* [online]. Forum.ircam.fr. Available from: https://forum.ircam.fr/projects/detail/spat/ [Accessed 19 May 2020].

Izadi, S., Davison, A., Fitzgibbon, A., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S. and Freeman, D., 2011. KinectFusion. Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST '11.

Jasiewicz, J., Treleaven, J., Condie, P. and Jull, G., 2007. Wireless orientation sensors: Their suitability to measure head movement for neck pain assessment. Manual Therapy [online], 12 (4), 380-385. Available from: https://www.sciencedirect.com/science/article/pii/S1356689X06001160 [Accessed 12 Dec 2019].

Joe, A., 2020. Head-related transfer function | The Headphone List [online]. Theheadphonelist.com. Available from: https://theheadphonelist.com/brain-localize-sounds/hrtf/ [Accessed 26 Apr 2020].

Kanat-Alexander, M., 2012. Code Simplicity. 1st ed. Sebastopol: O'Reilly Media.

Khan, 2020. Intro to the coordinate plane (video) | Khan Academy [online]. Khan Academy. Available from: https://www.khanacademy.org/math/algebra/x2f8bb11595b61c86:foundation-algebra/x2f8bb11595b61c86:algebra-overview-history/v/descartes-and-cartesian-coordinates [Accessed 11 May 2020].

Khan, 2020. Inverse square law (article) | Khan Academy [online]. Khan Academy. Available from: https://www.khanacademy.org/science/electrical-engineering/ee-electrostatics/ee-electric-force-and-electric-field/a/ee-inverse-square-law [Accessed 11 May 2020].

Khan, 2020. What is modular arithmetic? (article) | Khan Academy [online]. Khan Academy. Available from: https://www.khanacademy.org/computing/computer-science/cryptography/modarithmetic/a/what-is-modular-arithmetic [Accessed 11 May 2020].

Kumar, D., 2019. Software Engineering | SDLC V-Model - GeeksforGeeks [online]. GeeksforGeeks.com. Available from: https://www.geeksforgeeks.org/software-engineering-sdlc-v-model/ [Accessed 5 Nov 2019].

Lucid Content Team, 2017. The Pros and Cons of Waterfall Methodology [online]. Lucidchart.com. Available from: https://www.lucidchart.com/blog/pros-and-cons-of-waterfall-methodology [Accessed 5 Nov 2019].

Lyon, E., 2012. Designing Audio Objects for Max/MSP and Pd. Middleton: A-R Editions.

McAnally, K. and Martin, R., 2014. Sound localization with head movement: implications for 3-d audio displays. Frontiers in Neuroscience, 8. Available from: https://www.frontiersin.org/articles/10.3389/fnins.2014.00210/full [Accessed 22 Apr 2020].

McIver, C., Gahl, T., 2017. Positioning and Testing of Inertial Measurement Units. In: Academic High Altitude Conference [online]. Texas. Available from: https://www.iastatedigitalpress.com/ahac/article/id/5559/ [Accessed 14 Dec 2020].

Mellor, D., 2005. Stage Monitoring & Mixing. Sound on Sound [online]. Available from: https://www.soundonsound.com/techniques/stage-monitoring-monitor-mixing [Accessed 26 Apr 2020].

Microchip, 2020. [online]. Content.Arduino .cc. Available from: https://content.Arduino .cc/assets/mkr-microchip_atecc508a_cryptoauthentication_device_summary_datasheet-20005927a.pdf [Accessed 11 May 2020].

Milne, A., 2020. The Ultimate Guide to In-Ear Monitors, Straight from the World's Best Monitor Engineers [online]. Rfvenue.com. Available from: https://www.rfvenue.com/blog/iem-ultimate-guide-redux [Accessed 26 Apr 2020].

Mokhtari, P., Takemoto, H., Nishimura, R. and Kato, H., 2008. Computer Simulation of HRTFs for Personalization of 3D Audio. 2008 Second International Symposium on Universal Communication [online]. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4724498 [Accessed 26 Apr 2020].

MusicRadar, 2020. How to get the most out of in-ear monitors [online]. MusicRadar. Available from: https://www.musicradar.com/how-to/how-to-get-the-most-out-of-in-ear-monitors [Accessed 26 Apr 2020].

Neumann, 2020. KU 100 [online]. En-de.neumann.com. Available from: https://en-de.neumann.com/ku-100 [Accessed 11 May 2020].

Phillips, D., 2008. An Introduction To OSC | Linux Journal [online]. Linuxjournal.com. Available from: https://www.linuxjournal.com/content/introduction-osc [Accessed 3 Jan 2020].

Powell-Morse, A., 2016. *V-Model: What Is It And How Do You Use It?* [online]. Airbrake Blog. Available from: https://airbrake.io/blog/sdlc/v-model [Accessed 6 Nov 2019].

Rudrich, D., 2020. *IEM Plug-in Suite* [online]. Plugins.iem.at. Available from: https://plugins.iem.at/ [Accessed 19 May 2020].

Rudrich, D., Zotter, F. and Frank, M., 2016. Efficient Spatial Ambisonic Effects for Live Audio. In: 29th TONMEISTERTAGUNG - VDT INTERNATIONAL CONVENTION [online]. Graz, Austria: Institute of Electronic Music and Acoustics University of Music and Performing Arts. Available from: http://Efficient Spatial Ambisonic Effects for Live Audio [Accessed 4 May 2020].

SangHyun Chang, Wolf, M. and Burdick, J., 2010. Human detection and tracking via Ultra-Wideband (UWB) radar. 2010 IEEE International Conference on Robotics and Automation.

Schacher, J., 2020. SEVEN YEARS OF ICST AMBISONICS TOOLS FOR MAXMSP. In: 2nd International Symposium on Ambisonics and Spherical Acoustics [online]. Zurich, Switzerland: Institute for Computer Music and Sound Technology, Zurich University of the Arts. Available from: http://ambisonics10.ircam.fr/drupal/files/proceedings/poster/P1_7.pdf [Accessed 4 May 2020].

Shen, C., Wang, C., Zhang, K., Wang, X. and Liu, J., 2019. A time difference of arrival/angle of arrival fusion algorithm with steepest descent algorithm for indoor non-line-of-sight locationing. International Journal of Distributed Sensor Networks, 15 (9), 155014771986035.

Smith, A., 2017. What Is Binaural Audio? [online]. whathifi. Available from: https://www.whathifi.com/advice/binaural-audio-what-it-how-can-you-get-it [Accessed 26 Apr 2020].

SoundOnSound Magazine, 2017. Facebook 360 audio team discuss their tech [online]. Soundonsound.com. Available from: https://www.soundonsound.com/news/facebook-360-audio-team-discuss-their-tech [Accessed 4 May 2020].

TestBytes, 2019. What is V Model in Software Testing | Testbytes [online]. Testbytes. Available from: https://www.testbytes.net/blog/what-is-v-model-in-software-testing/ [Accessed 13 May 2020].

Vangelova, K., 2008. Stress and Fatigue in Sound Engineers: The Effect of Broadcasting in a Life Show and Shift Work. Central European Journal of Public Health, 16 (2), 87-91.

Verdugo, B., 2015. History of In-Ear Monitors [online]. Audiofly. Available from: https://audiofly.com/blog/history-of-in-ear-monitors/ [Accessed 22 Apr 2020].

White, P., 2009. Basic Live Sound. 2nd ed. London: Music Sales.

Zea, E., 2012. Binaural In-Ear Monitoring of acoustic instruments in live music performance. In: 15th International Conference on Digital Audio Effects [online]. Stockholm, Sweden: Sound and Music Computer Group KTH. Available from: http://www.diva-portal.org/smash/get/diva2:602966/FULLTEXT01.pdf [Accessed 22 Apr 2020].

Zhao, Y., Görne, L., Yuen, I., Cao, D., Sullman, M., Auger, D., Lv, C., Wang, H., Matthias, R., Skrypchuk, L. and Mouzakitis, A., 2017. An Orientation Sensor-Based Head Tracking System for Driver Behaviour Monitoring. Sensors [online], 17 (11), 2692. Available from: https://www.researchgate.net/publication/321213176 [Accessed 2 May 2020].

Appendices

Appendix A - Andrew Frengley Interview Extract

This interview with Andrew Frengley was conducted over the phone. The audio was recorded with his permission and can be found in the extras folders.

Interview Extracts

"We were pushing it [the technology] around. Putting it in people's [ears]. And it was well received, But there were some real problems with it. It was hissy. [It] had some reliability issues. They were glitchy.25 units were mothballed very quickly, right. And this production run that was made for Marty Garcia started churning out and that still has customer radio systems on the front of it. And there were probably 50 of those of that order made. Right. Those were the very original, but they were UHF. They were okay. Yeah, they were, but they were fixed frequency devices. Okay. Yeah. So they were basically they had a crystal that had a set frequency you had to work out. Like if you want to attain Your band yet to kind of, you know, get 10 that would work with everything else."

...

"So there were two things going on simultaneously. One was, there was a pressure to try and get synthesised frequencies i.e. the way things work [today], where you can tune content. We might get sued. But we had a bit of a cowboy mentality to that. I think we know, being Engineers ourselves, particularly monitor engineers knew that the idea of giving them [The Artist], a volume pot. They'd Probably always get off. Because it wasn't the technology itself. It was the user who was damaging themselves. We knew we would get off but we also knew that AKG Yeah, Sennheiser and Shure. Were now looking at this as a really viable kind of thing"

Appendix B - Survey of Key Contacts

Each respondent is represented by a number from one to six.

Question 1

Q. What is your role within the music industry?

A:

1. Sales & Application engineer
2. Research & Development Engineer
3. Freelance live sound engineer
4. Sound Engineer
5. Monitor Engineer
6. Applications Engineering Manager

Question 2

Q. Are you associated with a particular music organisation or audio company?

A:

1. Position above is with XTA Electronics and MC2 Audio.
2. Funktion One Research Ltd.
3. No.
4. Freelance.
5. No.
6. POLAR - UK Distributor for PRO, AV & MI Audio Markets.

Question 3

Q. What is your experience with monitoring?
A:

1. Our processors often process ear mixes.
2. I was a freelance FOH and monitor engineer before studying acoustics. I've always enjoyed mixing IEM stages more than wedges for both FOH and mons due to the isolation and quieter stage. Monitor bleed into mics is a real problem and can really screw up the FOH sound, as well as making life difficult for the performers. However, as a performer I prefer a loudspeaker monitor. I don't like headphones of any type due to the physical sensation on/in my ears and the in-head localisation.
3. I tour with a blues rock band, a few band members of which I have moved onto IEMs. I am also the in-house monitor technician at a large South London venue, where the use of IEMs is commonplace.
4. I have worked as a monitor engineer using IEMs and monitors.
5. Engineers point of view.
6. Moderate experience with monitoring, light experience using IEM systems.

Question 4

Q. Have you ever worked with binaural in ear monitoring systems?
A:

1. I work for a company which shares stand space with probably the market leader in this technology,Klang, I have on occasion stepped in while Klang people were not about and gave a brief demo.
2. I've tried binaural systems but not for monitoring. Usually the results are mixed with some localisation in certain directions, but mostly in head or above head. This is likely due to my HRTFs not matching the generic ones.
3. No.
4. No.
5. No.
6. Limited, but typically that artists still struggle to hear when using them due to the loss of loudness summation when NOT wearing both IEM's. Also the lack of efficiency of single dynamic drivers.

Question 5

Q. What specific features would you like to see in a binaural in-ear monitoring system? If any?

A:

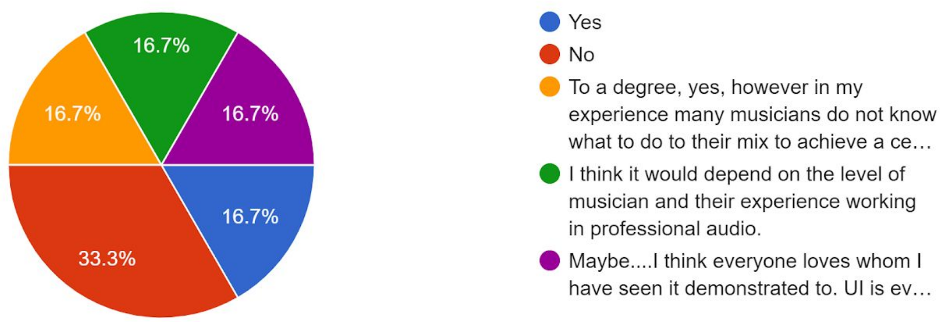
1. The only thing not being done at present is sensing of the listeners head position however hearing aid technology is way professional audio in-ears.
2. Choice of HRTFs or the ability to load a custom set.
3. Complete control of all inputs, including in particular FX returns, with regard to their position and depth in the binaural space.
4. As my experience with binaural in-ear Systems is limited it would be hard to say. I do think it would be important to have a system in which it was easy for festival changeovers etc. It can already be difficult working in stereo with IEMs on a festival changeover, let alone with the challenges I imagine you would have with the more sensitive nature of binaural.
5. No
6. Dual driver systems with internal crossover to increase overall driver efficiency/headroom across the entire frequency spectrum.

Question 6

Q. Do you think musicians would benefit from interactivity in their monitoring?
A:

1. Maybe....I think everyone loves whom I have seen it demonstrated to. UI is everything here. So good Digico bought the company & integrated the UI in their desks.
2. Yes
3. To a degree, yes, however in my experience many musicians do not know what to do to their mix to achieve a certain result. Mixing IEMs is far more than volume adjustment.
4. I think it would depend on the level of musician and their experience working in professional audio.
5. No.
6. No.

Do you think musicians would benefit from interactivity in their monitoring?
6 responses



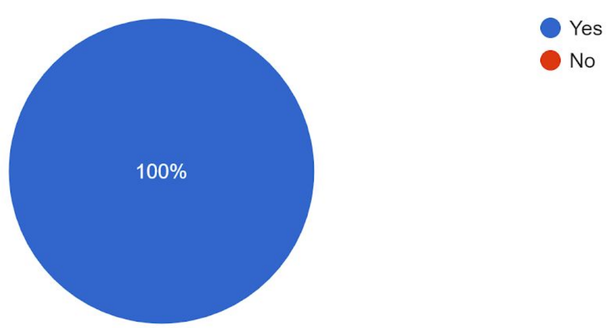
Question 7

Q. Do you think engineers would be open to a change in technology that requires them to take a few extra steps when setting up a binaural monitoring system, if it will ultimately benefit the musician?

A:

1. Yes
2. Yes
3. Yes
4. Yes
5. Yes
6. Limited, but typically that artists still struggle to hear when using them due to the loss of loudness summation when NOT wearing both IEM's. Also the lack of efficiency of single dynamic drivers.

Do you think engineers would be open to a change in technology that requires them to take a few extra steps when setting up a binaural monitoring system, if it will ultimately benefit the musician?
6 responses



Question 8

Q. What is your personal opinion of new technology like this, do you think it a temporary fad or is it here to stay?

A:

1. Better mixes, lower in levels, ability to give musicians limited control. I think it will filter down to lower prices in time.
2. Spatial audio technology and intelligent/auto mixing will absolutely become more and more common in live sound.
3. I think technology should be surplus to requirement. I can see how this could, if carried out correctly, make the musicians monitoring experience better, so I think for some it could become commonplace. Like most things in the business however, it's all down to preference, ease of use, and budget!
4. I think technology like this will struggle to become an industry standard, it would need to be extremely beneficial for the musicians' experience to revolutionise the industry. I haven't researched musicians' experience with binaural in ear monitoring and how they feel it works in a real life environment, so I can't comment. I do however think that it is important to remember that monitoring is a reference to ensure musicians are able to perform accurately, they aren't intended to be FOH mixes. Although I do value the importance of clear sounding monitor mixes and the quality that IEM's provide musician and how this aids in offering them a better all round experience.
5. I think this is a fad, I don't think this technology actually benefits musicians on stage. A well crafted in-ear mix is a fine tuned product of hours of work, and changes based on musicians movements can throw that into chaos. It's an interesting idea but I do not believe it achieves the objective of stage monitoring.
6. I think technologies that are designed to make the setup of 'complex' systems simple to a non-technical user are a marketing ploy from manufacturers and are still heavily under-developed and need further refinement. However, adding complexity to a system for a technically adept engineer should be welcome IF it brings significant benefit to the user/engineer.

Question 9

Q. Anything you would like to add?
A:

1. Looking at porting hearing aid positional tec to pro audio could be interesting
2. N/A
3. It's an interesting concept which I'd be very keen on trying, and I'm sure lots of other people would too. Initial thoughts as always with IEMs is safety, as you can easily damage someone's
4. hearing. Especially if user interactivity is an option, user error could easily result in an accident. Software would have to be rock solid, latency free and completely idiot-proof. Best of luck :).
5. N/A.
6. N/A.
7. No.

Appendix C - Technology Considerations

Head Orientation Sensor & Supporting Hardware

Using a gyroscope to measure the head orientation of the user was the obvious choice for this project as it is a cheap and effective solution. We did however test an additional sensor and microcontroller to those mentioned in the text. These were the MPU-6050 and the ESP2688. These two items were considered due their combined price of £6. However, you really do get what you pay for. The MPU-6050 proved incredibly difficult to work with it as it tended to drift out of sync with its original values after 10-15 minutes of continuous use. The ESP2688 proved difficult to work with and was replaced by the user friendly Arduino MKR1010-WiFi.

MPU-6050

Pros:

- Really cheap, some models start at 0.30p.
- Tutorials and examples are available for a wide range of applications.

Cons:

- Inaccurate over prolonged use.
- Difficult to interface with MKR's SAMD21 chip.

ESP2688

Pros:

- Manufactured in large quantities by generic chinese vendors costs range from £1-10 depending on the quality of manufacturing.

Cons:

- Poor quality of markings on the board.
- No clear instructions.
- No manufacturer examples, library or IDE.

Distance Tracking

There are a number of ways to track distance, but the majority were outside of our financial capabilities. VR technology utilises a relatively new and inexpensive solution, camera motion tracking with depth cameras. High lighting levels are a requirement (Izadi et al. 2011), therefore it is not feasible to track musicians on stage using this technology. The light levels would be constantly altered by the actions of stage lighting.

Bluetooth iBeacons and Eddystone Beacons emerged as another possible solution. They utilise Bluetooth Low Energy, a new bluetooth protocol that expands the range of bluetooth devices, while simultaneously decreasing power usage. The system relies on four beacons in each corner of the space and a bluetooth sensor worn by the user. The towers measure the ping between themselves and the worn sensor and triangulate the user based on the time difference from each beacon. This system would work for our intended purpose, however they are only accurate to 5cm and the cost proved to be too great (Cay et al. 2017).

Motion Camera Distance Tracking

Pros:

- Relatively inexpensive with Microsoft Kinect devices starting at £100.
- Lots of supporting open source C++ libraries developed for Kinect platform.

Cons:

- Inaccurate over large distances. Therefore inadequate for large stage use.
- Needs high lighting levels for accuracy. Stage lighting would disrupt this.

Bluetooth Low Energy Beacon Tracking

Pros:

- Larger area of coverage than motion cameras.
- Bluetooth RF would not be interfered with and can penetrate stage decorations.

Cons:

- Expensive complete systems up to £400
- Already outdated technology that as not seen widespread adoption
- Accuracy drifts and can only be measured accurately to $5cm^2$

Appendix D - Full Max 8

Arduino Code

```
#include <SPI.h>
#include <WiFiNINA.h>
#include <OSCMessage.h>
#include <WiFiUDP.h>
#include <MKRIMU.h>
/* WIFI Setup
 *
 */
#define SECRET_SSID "Binaural Audio Network"
#define SECRET_PASS "00002222"


///////please enter your sensitive data in the Secret tab/Arduino
_secrets.h
char ssid[] = SECRET_SSID ;        // your network SSID (name)
char pass[] = SECRET_PASS;    // your network password (use for WPA, or use
as key for WEP)
int status = WL_IDLE_STATUS;     // the Wifi radio's status

WiFiUDP Udp;                                // A UDP instance to let us
send and receive packets over UDP
const IPAddress outIp(192,168,0,101);        // remote IP of your computer
//const IPAddress outIp(192,168,1,72);
const unsigned int outPort = 10000;          // remote port to receive OSC
const unsigned int localPort = 8888;        // local port to listen for OSC
packets (actually not used for sending)

void setup() {
    //Serial.begin(9600);
    //while (!Serial);

    if (!IMU.begin()) {
      //Serial.println("Failed to initialize IMU!");

     while (1);
    }
    /*
    Serial.print("Euler Angles sample rate = ");
```

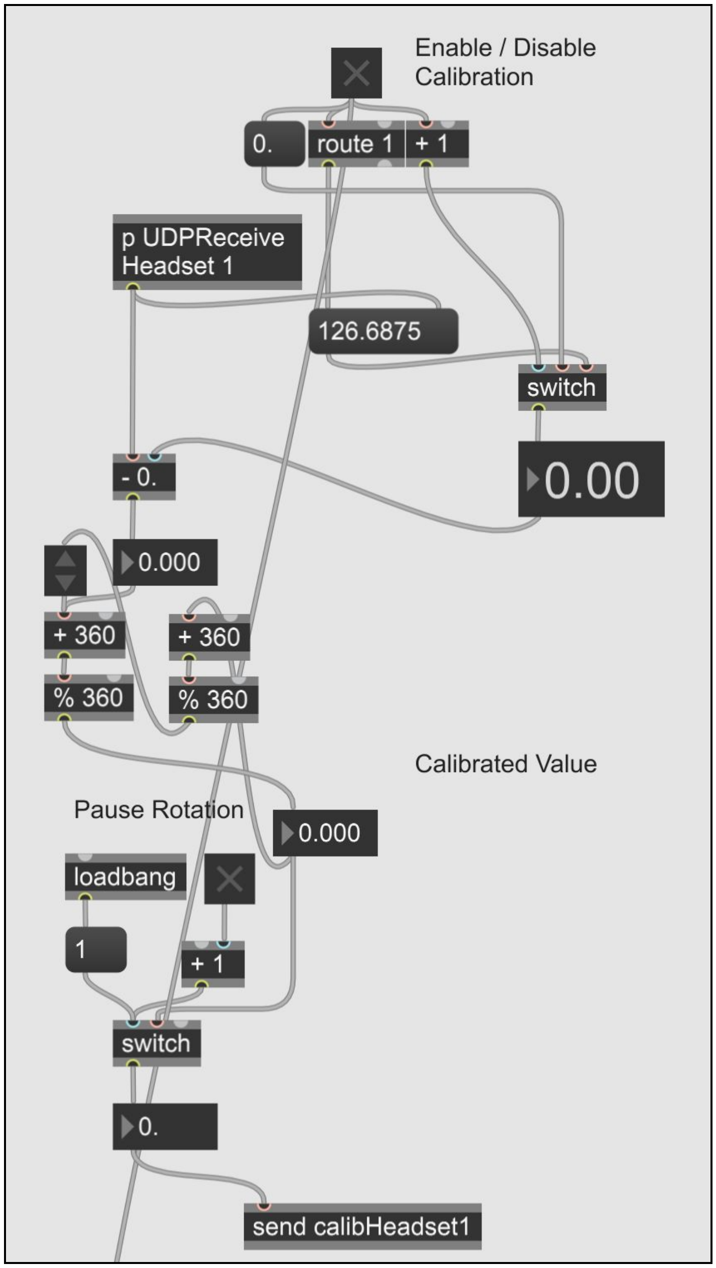
```
    Serial.print(IMU.eulerAnglesSampleRate());
    Serial.println(" Hz");
    Serial.println();
    Serial.println("Euler Angles in degrees");
    Serial.println("Heading\tRoll\tPitch");

    // Connect to WiFi network
    Serial.println();
    Serial.println();
    Serial.print("Connecting to ");
    Serial.println(ssid);
    */
    WiFi.begin(ssid, pass);

    while (WiFi.status() != WL_CONNECTED) {
        delay(500);
        //Serial.print(".");
    }
    /*
    Serial.println("");

    Serial.println("WiFi connected");
    Serial.println("IP address: ");
    Serial.println(WiFi.localIP());

    Serial.println("Starting UDP");
    */
    Udp.begin(localPort);
    //Serial.print("Local port: ");

    //Serial.println(localPort);

}

void loop() {
  float heading, roll, pitch;
  if (IMU.eulerAnglesAvailable()) {
    IMU.readEulerAngles(heading, roll, pitch);
    //Serial.print(heading);
    //Serial.print('\t');
    //Serial.print(roll);
    //Serial.print('\t');
```

```
    //Serial.println(pitch);

    OSCMessage msg("/h/");

    msg.add(heading);
    msg.add("/r/");
    msg.add(roll);
    msg.add("/p/");
    msg.add(pitch);
    Udp.beginPacket(outIp, outPort);
    msg.send(Udp);
    Udp.endPacket();
    msg.empty();

  }
}
```

Head Orientation and Distance Max Inputs

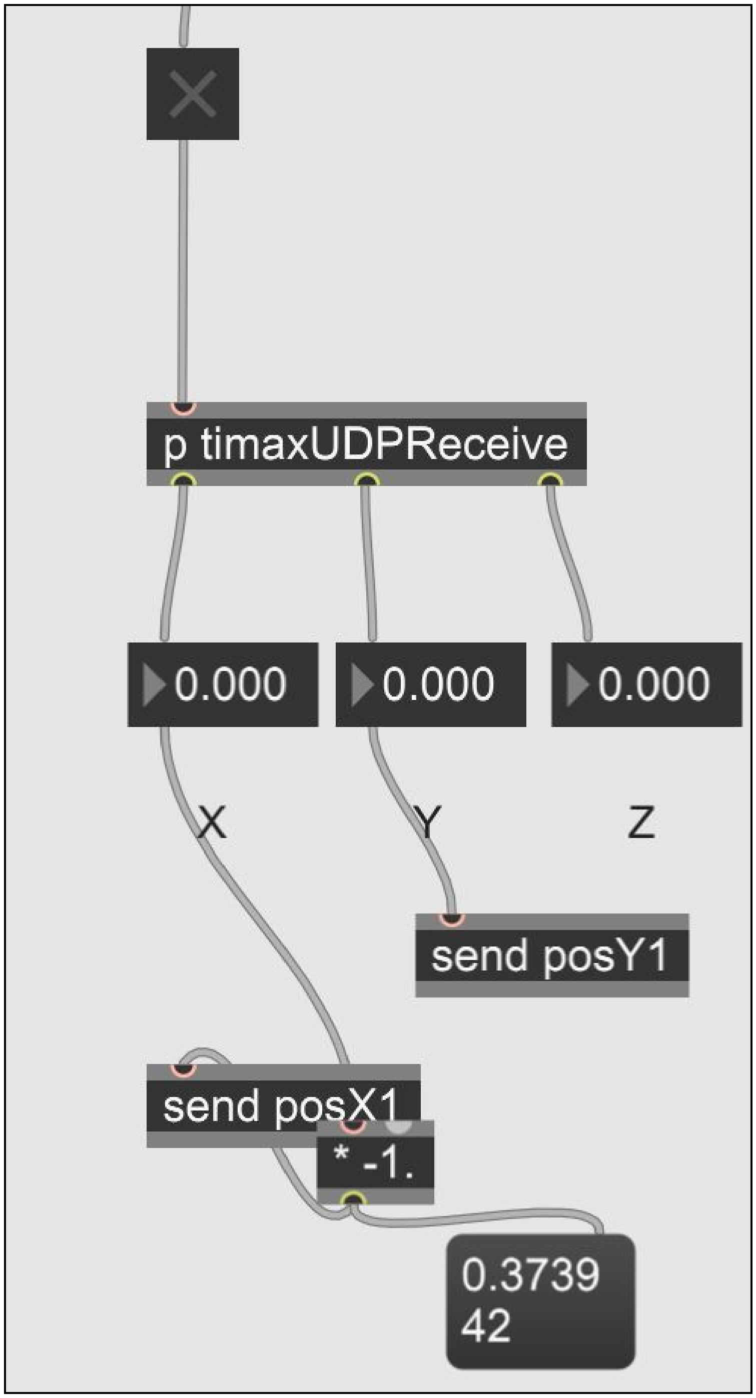
Gyroscope Sensor Input and Calibration System

Previously described in Section 3.3.



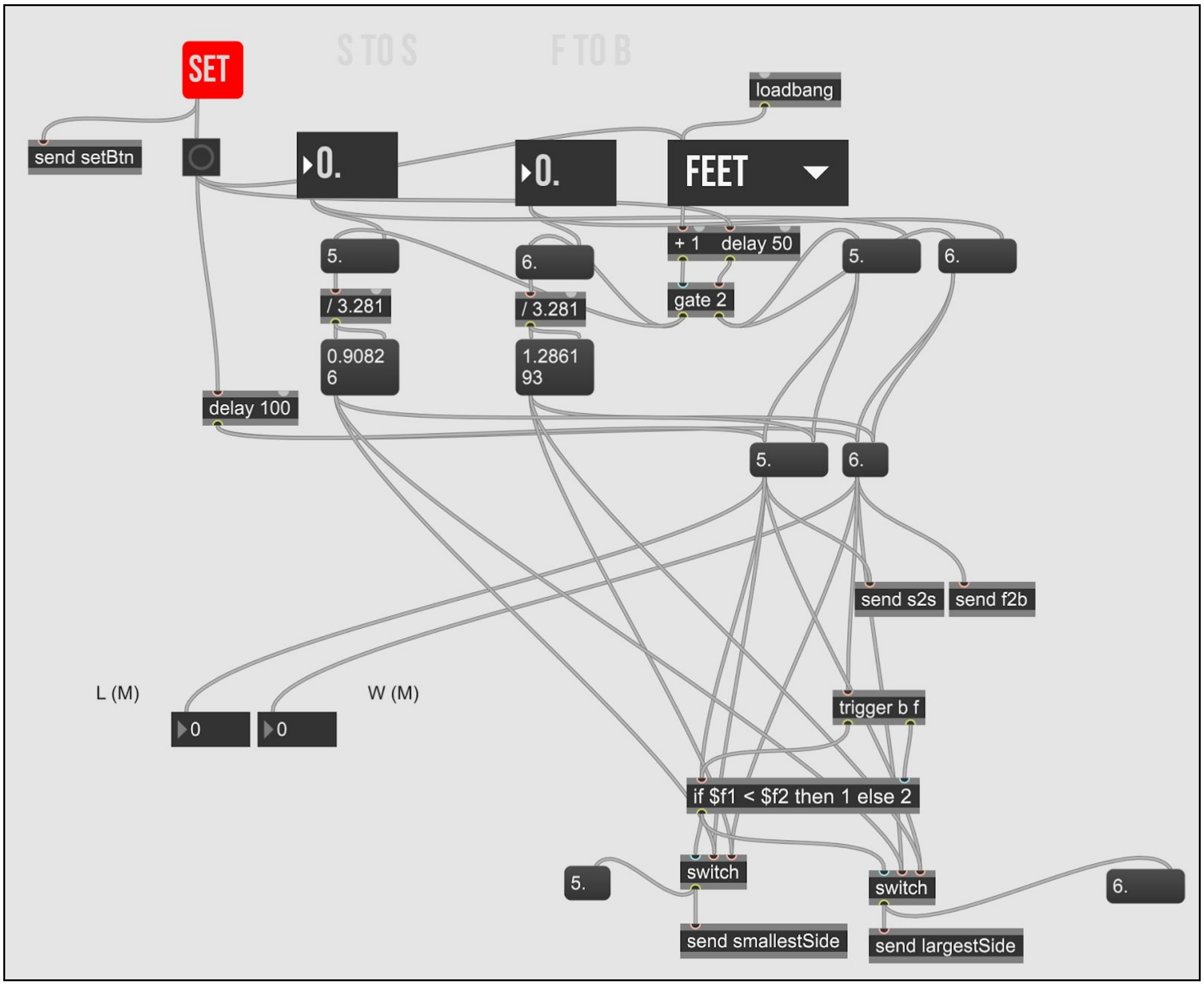
TiMax Distance Data Input

This object receives the OSC data and splits the OSC messages into a format understood by Max 8. It then sends this data to the distance calculation formula described in Section 3.4.



Room Dimensions Input

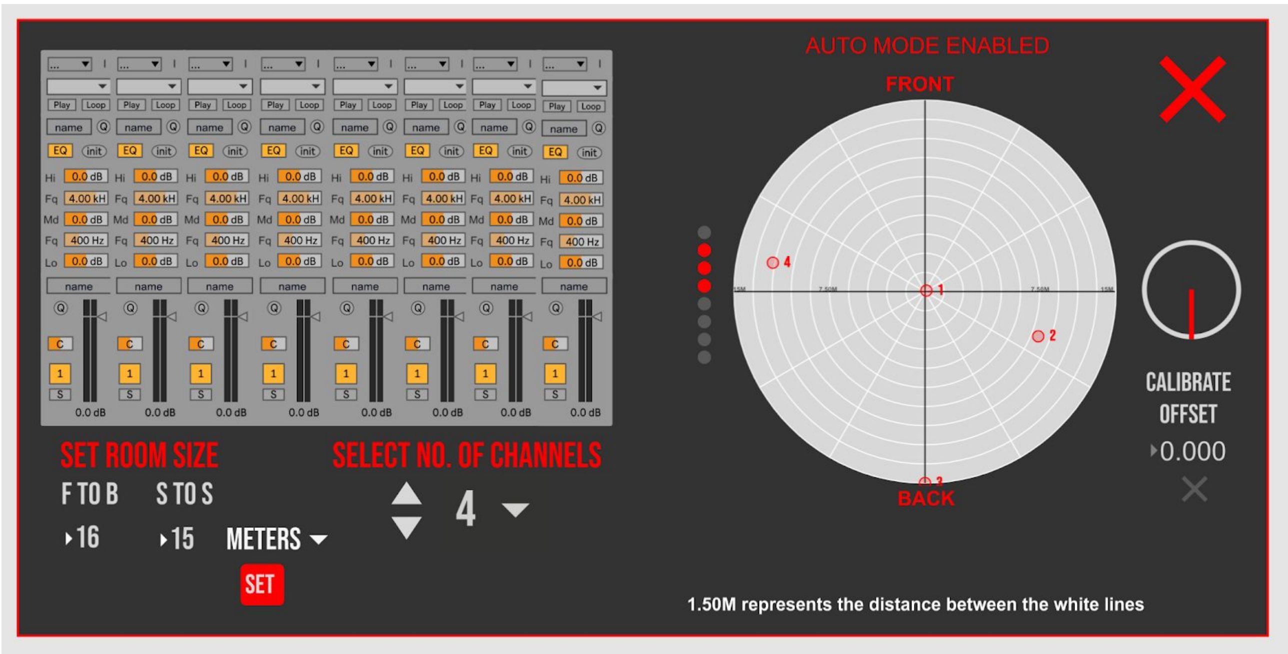
This section of the program requires the engineer to input the room dimensions in meters or feet. The formula automatically converts feet input into meter output, as meters are used by the SW VST. The smallest side of the room is selected as the maximum wide of the sound field. This data is sent to the XY Calculation and SW VST.



GUI

Channel Selector and Monitor

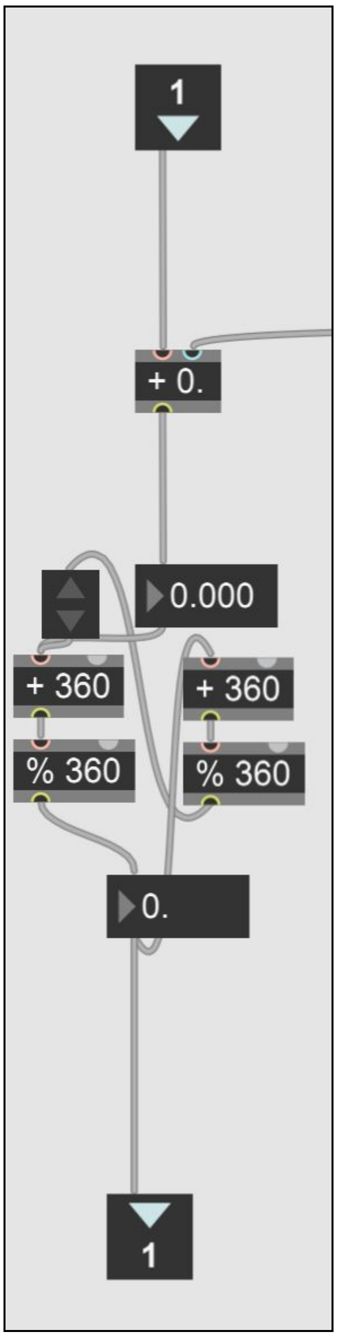
The software setup, GUI and monitor is explained with detail in Section 3.3. The values assigned to the sound sources by the channel selector or by the engineer moving sound sources with a mouse. These values are captured and saved in the program datas as the original values and the gyroscopic change values are added to these initial values.



Calculation Engines

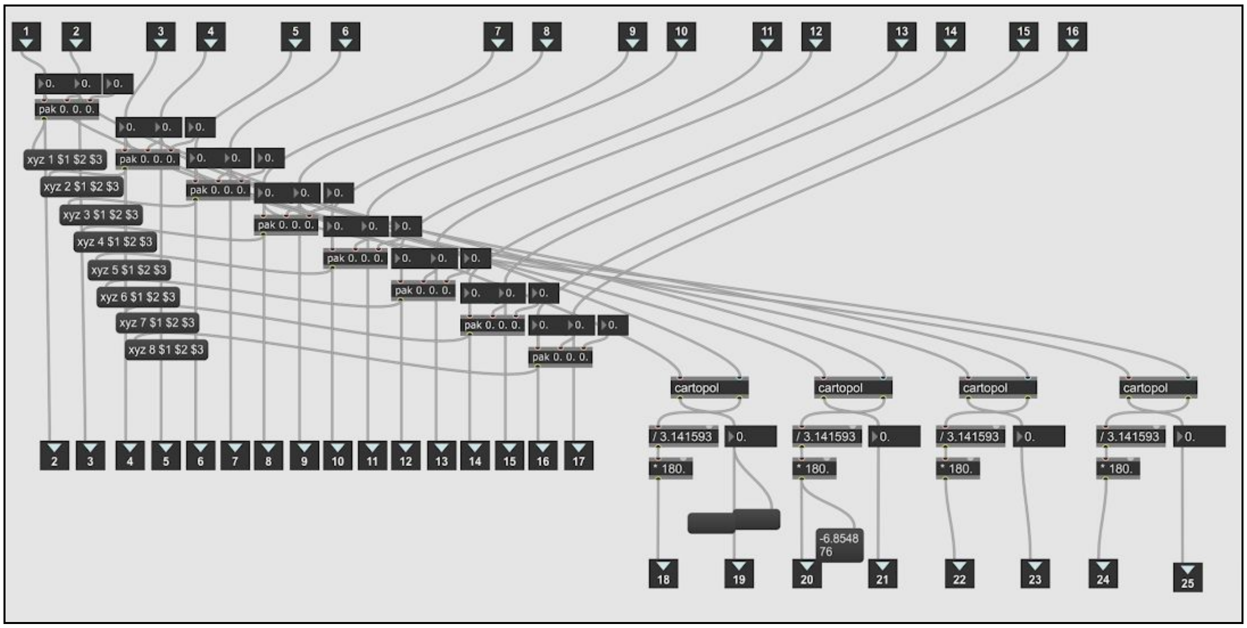
Orientation Addition System

The orientation system uses modulos described in Section 3.4, as well as simple addition and adds the changed value of the orientation to the original values captured upon flicking the automatic mode switch located in the top right corner of the program. The two modes stop sensor data influencing sound sources while trying to located them with the mouse.



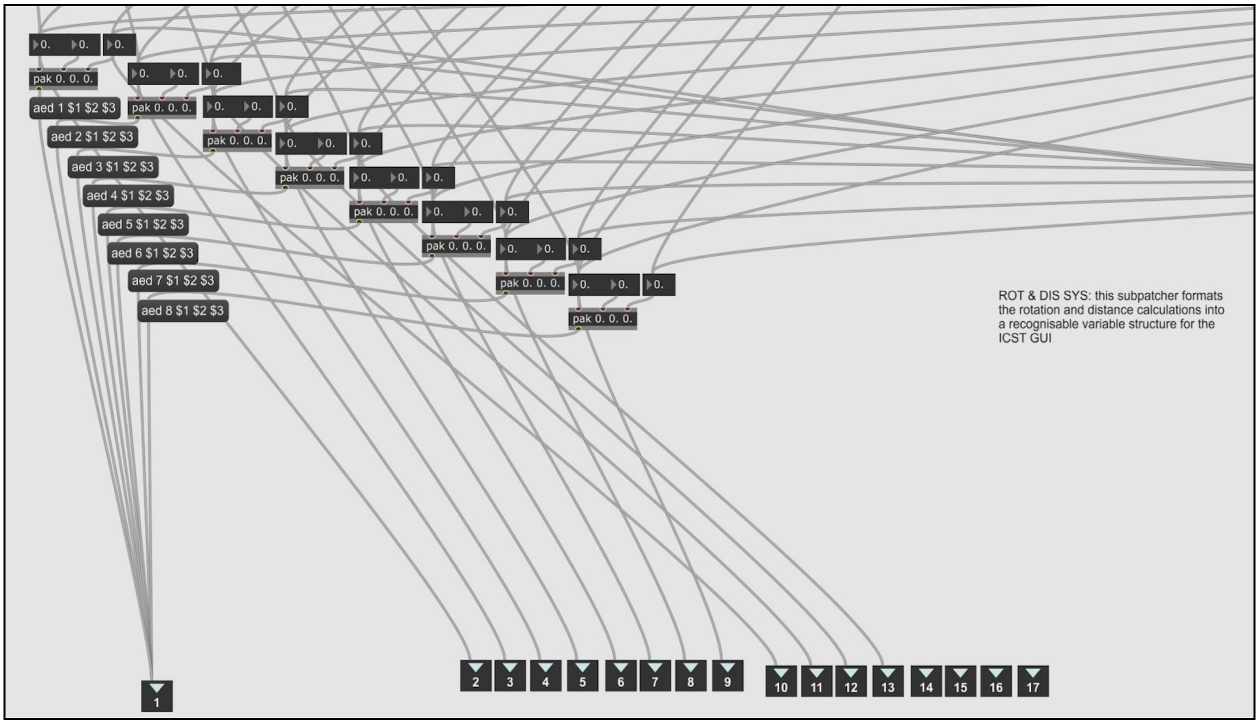
XY Calculations

The TiMax x and y values are scaled to match the AED values understood by the Monitor (Section 3.4). These values are converted using the cartopol object which outputs angle and amplitude. The amplitude value is used in place of the d value assigned by the Monitor or engineers mouse placement.



Ambimon

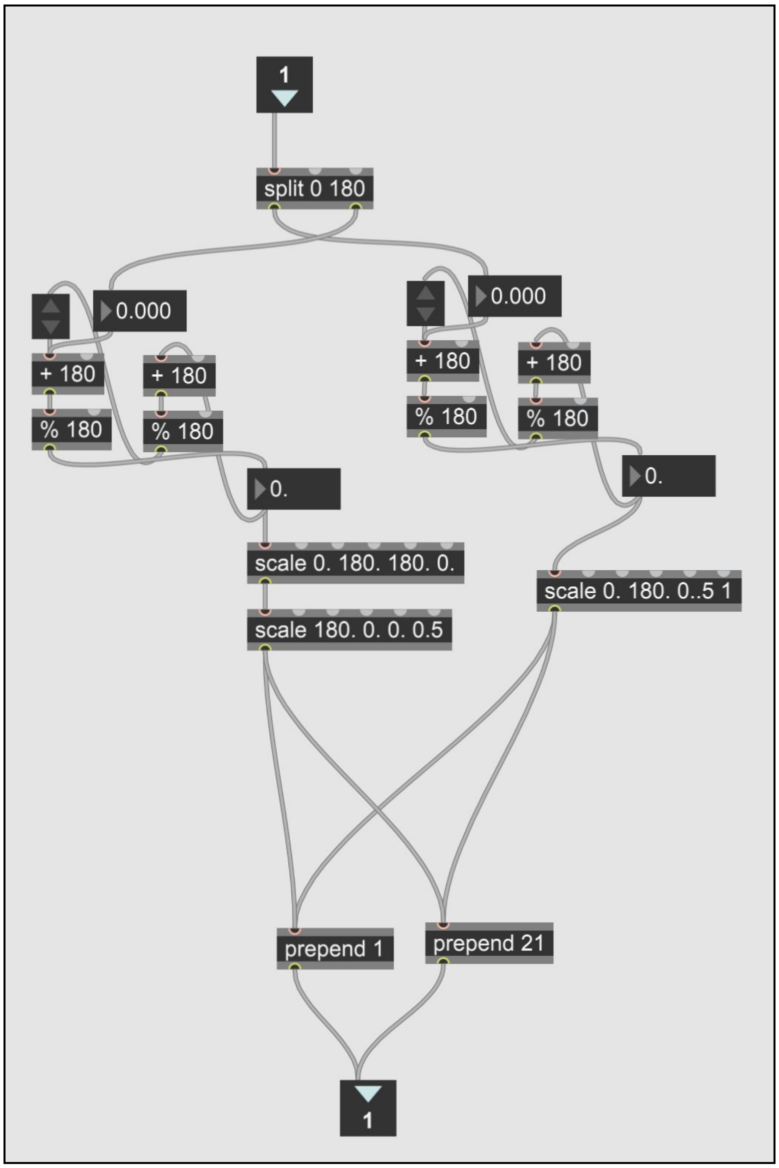
The A values in the AED equation are the head orientation data set and the d values are obtained initially from the Monitor and the TiMax distance system once it is enabled. These values are added together and the prepared in the style of the Monitors formatting system [aed] [channel] [aval] [e=0] [dval] represented in Max 8 as variables ($).



Spatial Workstation (SW)

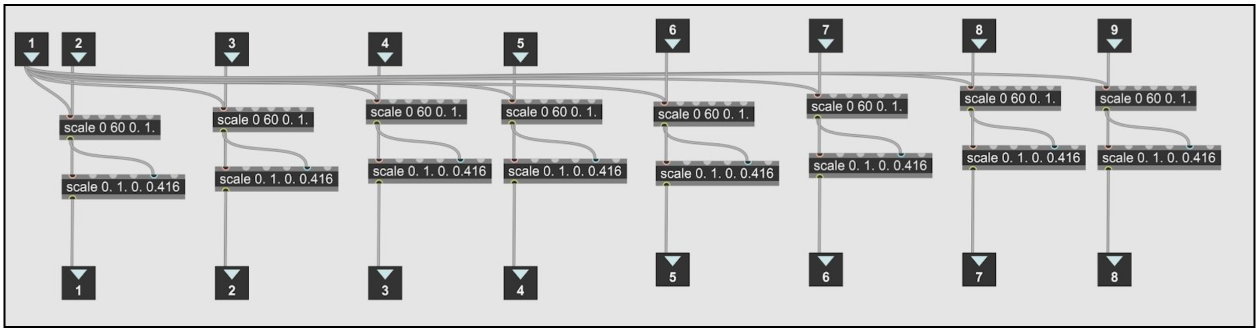
Head Orientation Data Preparation

The orientation data had to be converted from 0 to 360 to a range of 1 to -1. This was achieved using scale objects and a split equation that converts negative Monitor values to positive ones. It also prepends a number that represents the variable that calls that A value in SW.



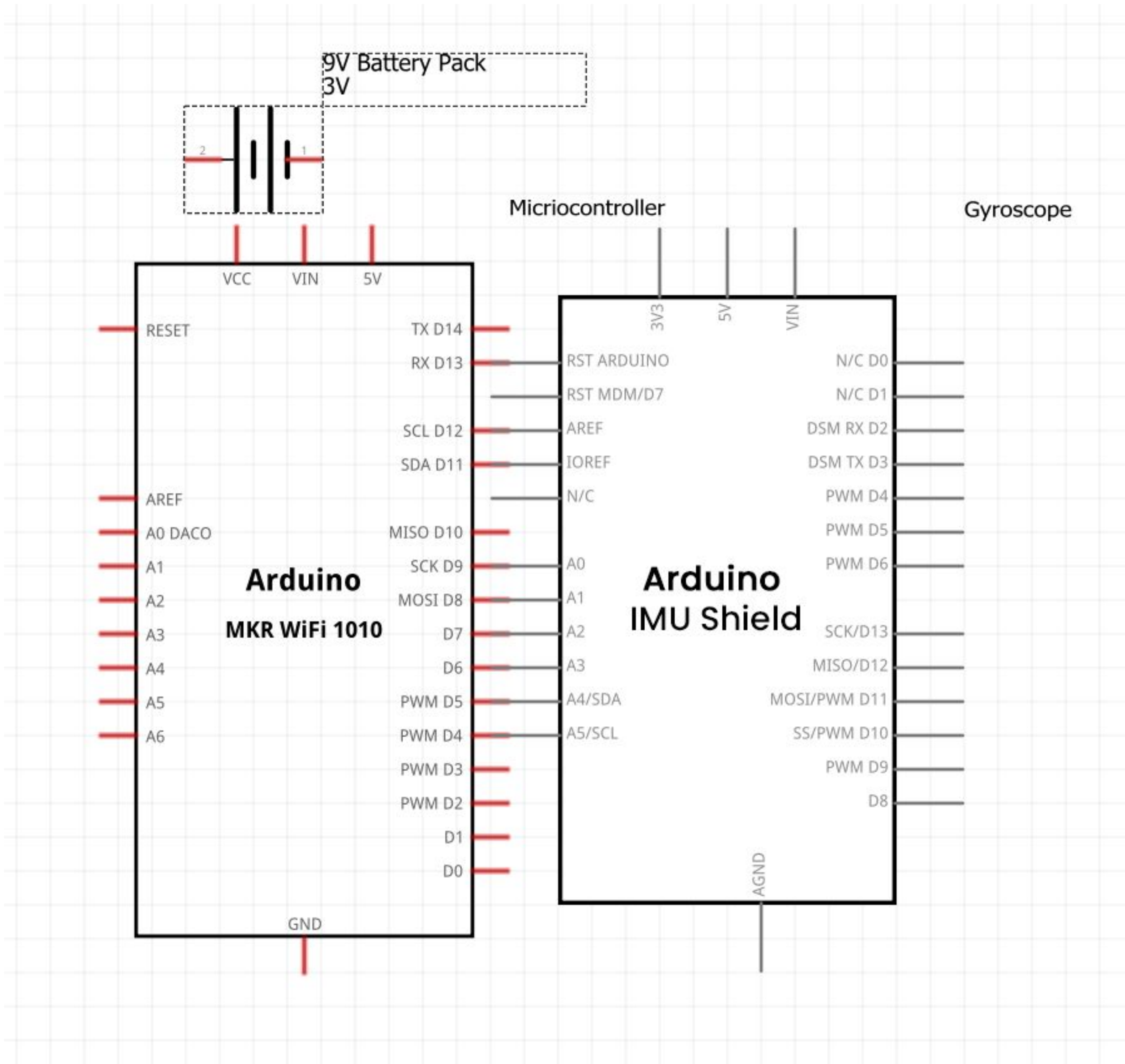
Distance Data Preparation

The distance data preparation differed slightly from the orientation values. The d values did not need to be scaled as they were represented by 0-1 in SW as well as in the Monitor. The room dimension inputs were used to set the room dimension emulation system within SW.

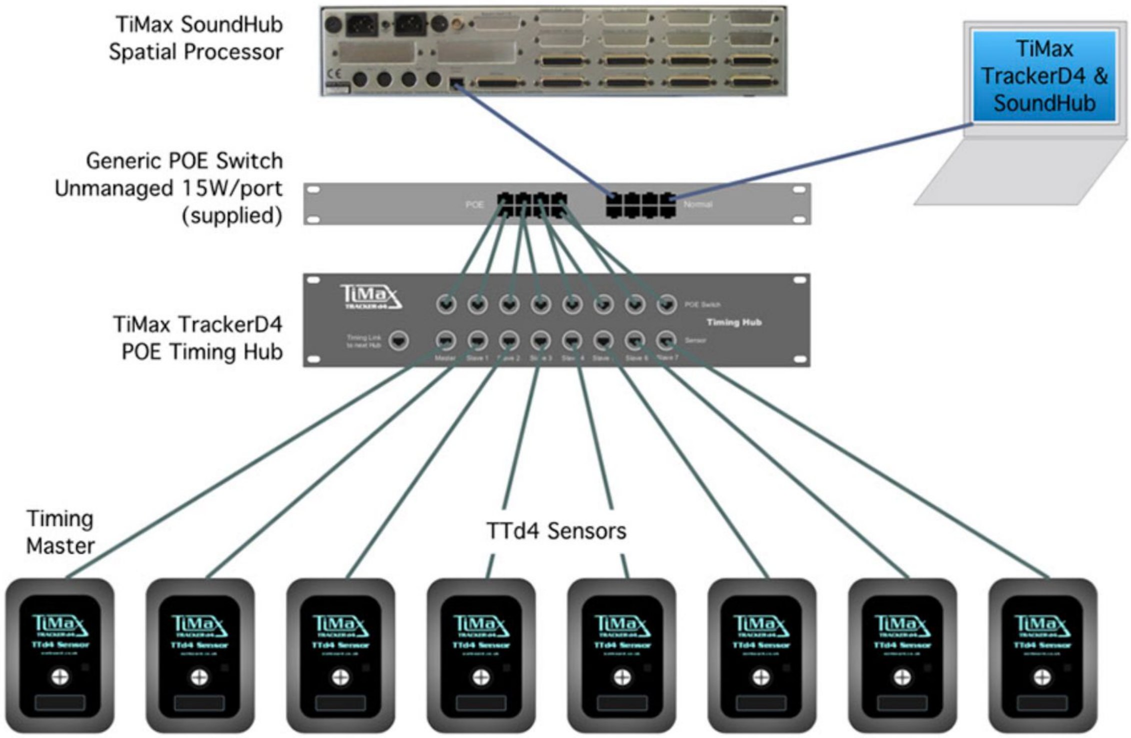


Appendix E - Schematics

Arduino Schematic



TiMax Schematic



Appendix F - Full Methodology

The validation and verification method ensures that the final system will be tested, with bugs identified and eliminated during each development phase, rather than having a big round of bug testing at a late stage in the project, as would happen with the waterfall methodology (Balaji and Sundararajan Murugaiyan 2012). The waterfall approach can be costly, because some bugs come about as a result of poor architecture design, which can be incredibly difficult to fix towards the end of a project (Lucid Content Team, 2017).

4.1 Phase 1: Requirement Analysis & Acceptance Testing

The first step in a project using the V-Model methodology is Requirement Analysis (RA) followed by an acceptance testing stage. The RA stage is also known as a requirement gathering stage. (Powell-Morse 2016) It is completed by identifying the needs and wants of the end user of the system, in our case performing musicians. It is about understanding their requirements and meeting their expectations. Determining the complete feature set at this stage is very important. The Acceptance Testing that follows will involve discussing the requirements identified with members of the touring industry, any additional requirements identified or changes required can then be added before continuing to the next phase.

Validation:

- Proposing head orientation and binaural panning as a way of removing the isolation felt by musicians when using IEMs.

Verification:

- Surveying key contacts within the music industry to validate and add further to the features initially proposed.

4.2 Phase 2: System Design & System Testing

All the phases are of equal importance, without one, or even just a poorly completed phase - the project would have the potential to collapse. However, the System Design (SD) phase is critical in this case. The first step in this process would be completing two of the identified research objectives:

- Identify the pieces of hardware to buy, such as which microcontroller and sensors
- Consider how the device will be placed on the users head, and how it will be powered

These two tasks go hand in hand with identifying the best methods of binaural panning techniques to utilise and the implementation methods are codependent on each other. Using one binaural technique may result in a different hardware requirement, so it is critical to establish if the correct hardware has been selected within the system testing that immediately follows the SD phase.

Validation:
- Researching and then building two different hardware modules for the head orientation tracking.
- Researching methods of distance tracking, with UWB RADAR the most promising

Verification:
- Comparing the MKR1010 and ESP2688, as well as the IMU-Shield (BNO-055). Ultimately deciding on the Arduino MKR1010 WiFi and IMU-Shield (BNO-055).
- No distance tracking methods were affordable, however one was provided by TiMax

4.3 Phase 3: Architecture Design & Integration Testing

Following on from the System Design is a phase that will be utilised in this project primarily to develop the supporting software, and then integrate it with the selected hardware. This was by far the largest stage. Max 8 was identified as the main coding development tool, due to existing knowledge in the programming language. The system can already be split into two main modules, the hardware and the software. Identifying a strong and reliable communication method between the two parts of the system is critical at this stage in the project and will be put under heavy scrutiny during the Integration Testing that will immediately follow this development stage.

Validation:
- Transmitting data from the hardware to Max 8 using OSC.
- Integrating the sensor data into max, building a calibration system to align the gyroscope to the front of the stage.
- Developing a channel selection system that plots sound sources spaced evenly on the Monitor depending on the amount of channels you select, with 8 available.
- Applying the sensor data to current sound source values to move them by the amount of change in the gyroscope value. This required several attempts of verification testing before this section passed.
- Applying the same logic to distance tracking but using the cartopol Max object to translate data from TiMax into a format understood by the Monitor.
- Preparing the data for integration with the SW VST.

Verification:
- Testing the stability of OSC when sending the vast quantities of data from the gyroscope and TiMax tracker. It was found that a home WiFi network was adequate. Instead it is recommended to use a secure separate WiFi router that is not connected to the internet.
- The calibration system was fully tested and required some tweaks to perfect.

- This system was one that came by accident when exploring means of plotting within the Monitor, however the system handled it well and it was decided that it would be a good way to show engineers how to plot channels in the Monitor.

4.4 Phase 4: Module Design & Unit Testing

This is the low-level design section of the project. In this phase items such as the Graphical User Interface (GUI) and design of the wearable headband will be fully completed. This will be the final step of the utilised V-Model methodology. It will be followed by a stage of unit testing, where the system will be lab tested vigorously and refined before the final step in the project.

Validation:

- Developing the GUI followed the principles of simplicity and contrast, with an easy to read interface and familiarity by association with the Ableton styled channel strips.
- Lowstress audio tests were designed using multitracks to test system input and outputs via Waves Soundgrid.

Verification:

- Tutors and colleagues praised the sleek design and further tweaks mean it is fully completed.
- Latency was not accurately using the method. It is unfortunate due to current circumstances that latency was not measurable. The normal console setup was not ascertainable therefore all testing was simulated using multitrack recordings of live performances.

4.5 Phase 5: Finalisation & Release

This is not a phase that's usually included in the V-Model methodology, however one can not underestimate the power of real world testing when developing a product. Real world testing can reveal bugs that couldn't have emerged in lab testing. The user will use (and abuse) the product in ways the developer never thought of. This final step is included so that the system can be tested by a number of real bands in a real scenario. This final stage of course could reveal some practicality issues, and may result in more tweaks before finalisation.

This phase could not be completed as outlined. In future scenarios and testing this would be a key parameter to test and introduce real world problems to the system and the direct feedback from musicians would hugely benefit the project.

Appendix G - Extras Folder

The extras folder contains a video demonstration, Max 8 project files and Arduino code:
https://drive.google.com/drive/folders/1oYow3F_hqXJvxcS0gWgwHNswNi9Unvm9?usp=sharing