UNIVERSITY OF TWENTE.

Faculty of Electrical Engineering,
Mathematics & Computer Science

Creating an embodied music controller that can be used to improvise the arrangement of electronic dance music

Robin van Soelen

M.Sc. Thesis
August 2021

Supervisors:
dr.ir.Dennis Reidsma
MA. Benno Spieker
prof.dr. Mariëlle Stoelinga

Summary

In this thesis, an embodied music controller that can be used to control the arrangement of EDM was developed. This was done with the purpose of making electronic music performances more transparent and engaging.

A prototype was built that consists of three devices. Each device uses an IMU (Inertial Measurement Unit) sensor that is placed on either one of the hands or the body. An LSTM model was used for recognising gestures that are performed with each device. Based on these gestures, corresponding MIDI data will be send out to Ableton Live.

A series of co-design sessions with music producers and dancers was conducted to iteratively improve the prototype and to find preferred mappings between movements and music. Participants were able to use the controller in such a way that they were able to make the adjustments to the music that they wanted to make. Mappings that mimicked familiar actions were preferred.

To gain insight into the perception of an audience on the developed music controller, a survey was filled out by 70 participants. They were shown clips of somebody performing a song with the controller and were asked to comment on their understanding, preferred configurations and the enjoyability of this performance. The results show that while not every action that the performer made was understood, the participants saw a relationship between the movements and the music and they found the performance engaging.

This research illustrated that it is possible to use movement to perform electronic music and that there are people who are willing to watch this. It paves the way for researchers and musicians to apply similar music controllers to physical live concerts.

Contents

Chapter 1

Introduction

The developments in music recording technology have made it possible for individual artists to record an almost infinite amount of layers into their compositions, allowing for a large array of new sonic possibilities but making it difficult for these artists to find enough musicians to play these compositions in a live setting. What makes this even more challenging is the fact that modern musicians are not bound to the use of traditional instruments anymore. The use of sounds that are not reproducible on acoustic instruments, for example, sounds created with wavetable synthesizers, drum computers and the editing of samples, is now very prominent in modern music and even defining for certain music genres [6]. These artificial sounds are sometimes impossible to reproduce in live settings. A common solution for artists to perform this kind of music live is to only perform certain layers of their music live, while playing along with the rest of the song through a pre-recorded backing track. Another approach is to not play any instrument live, but to focus on creatively mixing the different layers of a song together. This method is mostly applied for performances of Electronic Dance Music (EDM).

This type of EDM live concerts has caused some controversy. Mostly built around the argument that performing with a laptop or mixing table is creating confusion about what the artist is doing, causing the audience to feel cheated on [7] [8]. While for some people this might be the case, the number of people that buy tickets for concerts given by DJ's seems to indicate that not everybody feels this way. One explanation for the popularity of these concerts is the fact that closely observing how the artist makes the music is not the main motivator for people to go to concerts. According to Caldwell et al's study on the motivations for going to concerts, the main motivators are the experience of being there, engagement with like-minded people and the novel aspects of a live show [9].

However, the lack of musical context could still be a layer missing in the musical experience. The more familiarity the listener has with the musical context, the more vivid the empathetic experience can become. According to Bahn et al. [9], "this describes a connection of the body to sound production, a kinesthetic empathy with the act of creating sound and the visceral/gestural interaction of the performers in the musical context. The strength of this connection can be seen in the common mimesis of rock guitar performance, or "air guitar"". Additionally, a study investigating the relationship between emotional response and perceived liveness found a correlation between the two [10]. This suggests that perceived causality between gestures and sound indeed plays a role in how people experience music.

One way of increasing this causality between gestures and sound within electronic dance music, is through embodied music controllers. These controllers allow artists to use their movement to control musical parameters that normally require knobs or buttons to control them. In recent years, the use of embodied music controllers has gained some traction in the music industry. For example, the Mi.Mu gloves, co-developed by artist Imogen Heap, have already been used by many artists including Ariane Grande. Embodied music controllers show great promise for the way that electronic music is performed and the expressiveness that is achieved through them. However, only a limited amount of research on applying this technology for the purpose of controlling music the same way as performing artists of EDM would do, has been conducted.

Therefore, it would be interesting to explore the use of an embodied way of performing electronic music that can be used by performing artists to intuitively and expressively perform electronic music.

1.1 Research questions

This thesis will explore this concept through attempting to answer the following research question:

> How do you design an embodied music controller that can be used to improvise the arrangement of EDM in such a way that its expressive power is optimised?

Sub-questions that will be used to answer this main research question are:

- RQ 1: How do EDM performances work?
- RQ 2: What does embodiment contribute to music performance?
- RQ 3: How is a prototype of the proposed system built?
- RQ 4: How can the usability of the proposed system be optimised?
- RQ 5: How can the proposed system be made more expressive?

1.2 Approach

To be able to answer the main research question, it is necessary to answer the sub-questions first. These questions will be answered in the following sections:

RQ 1 and RQ 2 will be answered through literature research, which will be covered in the background section (chapter 2, p. 5). Afterwards, RQ 3 will be answered in the development of the prototype section (chapter 3, p. 17), which will go over the steps that are being taken to achieve building the prototype. RQ 4 will be answered through a series of evaluation sessions with producers and dancers, which will be described in the first evaluation section (chapter 4, p. 33). Finally RQ 5 will be answered in the second evaluation section (chapter 5, p. 49), where a survey will be conducted to get insights into how expressive the prototype is for an audience.

Afterwards, in the general discussion will include a reflection on the performed research (chapter 6, p. 67).

Chapter 2

Background

In this section relevant literature will be discussed. The section will be split into three parts. First, an overview of the methods used by performers of electronic music will be discussed, with the aim of answering RQ 1: How do EDM performances work? Afterwards, the relationship between embodiment and music will be explored, to be able to answer RQ 2: What does embodiment contribute to music performance? Finally, a short overview of the current state of the art will be given.

2.1 Electronic music

Understanding electronic music (music made with modern technologies) is very valuable in the context of designing a music instrument for it. Therefore, in this section an overview will be given on the different types of electronic music, how this music is typically made and the things a performer does to interact with a crowd.

2.1.1 What are common elements in electronic music?

The term electronic music is extremely broad. The term essentially covers all music that is made through analog or digital technologies, while a lot of music that is made this way is not necessarily seen as electronic music. Therefore, this report will narrow this term by focusing on Electronic Dance Music (EDM). This genre in itself is quite broad and has many subcategories like, house, trance, downtempo, drum and bass etc.

Nevertheless the songs within these subcategories all make use of some of the same elements. An example of this is the song structure, which is called the break routine. This routine is a set of three sections, a (1) breakdown, (2) build-up, and (3) drop, that is continuously being looped through. Just like the recognisable structure of a symphony or a pop song helps listeners to know where in the song they are, the break routine does the same for dance music by guiding the listener when to dance. A study on embodied experiences of dancing to EDM music has found a significant increase in movement during the drop section [11]. In this study 16 participants, wearing motion tracking devices, danced in a club setting while dancing to music that made use of this break routine. The amount of motion at each moment throughout the music combined with self reported experiences of pleasure in a survey were analysed. The results indicated that the movement between the sections created synchronous behaviour between participants and the participants found the movement from the build-up to the drop particularly pleasurable.

The switching between these sections is mostly done through creating different levels of energy and tension. Typically the build-up works towards a peak of tension that will be released when the drop comes in. There are a few common methods of achieving tension. One of these is the use of risers, which is most of the time white noise that builds in volume and frequency towards a climax at the end of a build-up. This way the higher end of the spectral field slowly fills up, causing the listener to feel more tension as the riser continues. Another method is the use of snare rolls, where the snare drum pattern will stepwise double in speed until a continuous snare roll can be heard, indicating the peak of the build-up. Finally, using low pass filters on harmonic parts of the song and gradually adding the higher frequencies back is another common method of adding energy and tension [6].

Besides this break routine, Lyubenov [12] describes a few other elements that are common in EDM music. The paper mentions the typical use of certain music equipment within this genre, which are mostly electronic instruments. Examples of such instruments are synthesizers, drum computers and sequencers. The combination of these instruments results in music with a distinct timbre that often is hard to trace back to a specific instrument, since these sounds can not organically be produced in nature. Sometimes acoustic instruments, like drum kits, can be heard through sampling. Sampling is the process of using a short bit of an existing sound and changing the length, pitch or playback speed of the sound to make it fit into another composition. This process can transform an organic sound into a sound that is perceived as more electronic since these adjustments do not occur in nature.

Lyubenov also mentions that EDM songs typically have a high tempo (129-150 beats per minute), which can be explained through the main purpose of the genre which is getting people to dance to it.

In summary, like the name Electronic Dance Music suggests, the genre encompasses music with an electronic sounding timbre, made with the intention of having people dance to it. The electronic timbre is achieved through the use of instruments like synthesizers, drum computers and sequencers. Getting people to dance to the music is achieved through a break routine consisting of a breakdown, build-up and drop combined with a high tempo. Given that the genre is very broad, all the discussed characteristics are generalisations and there are of course exceptions to them.

2.1.2 How is EDM made and performed live?

The emergence of the Electronic Dance Music genre ties closely to the development of drum computers produced by Roland. Specifically, the Roland TR-808 largely shaped the evolution of dance music [13]. Often these drum computers would be used in combination with samples from old disco songs and analogue synthesizers. With the development of technology, the production of EDM shifted from analog devices towards digital ones. Nowadays, most EDM music is made using DAW's (Digital Audio Workstations), which is software made for recording, editing and playing digital audio. Currently they are the heart of both professional and home recording setups. DAW's not only allow the recording of audio, but can also control software instruments using MIDI, which is a widely used protocol for exchanging musical data between devices.

The live performance of EDM can be done using several techniques. The most common ones are:

- *DJ-ing*: Creatively mixing between several fully mixed tracks.
- *Live arrangement*: All the layers of a track are pre-recorded. During the live show the artist improvises an arrangement by choosing which part to play when and applying effects.
- *Playing with a backing track*: Parts of the track are pre-recorded, parts are played live.
- *Looping*: All layers are played live and afterwards looped, allowing one person to play multiple parts in a song.
- *Fully live*: All layers are produced in a live setting.

Often there is confusion about the difference between a DJ and a performing artist of EDM music. The term DJ (Disc Jockey) is mainly used for people who mix songs, generally made by other people, in a pleasing way, while performing artists mainly perform their own music in a way where the performer still has room to improvise [14]. Given the complexity of most EDM tracks, often performing artists choose to do a live rearrangement of the song, some of them are also playing certain layers of the song live, but recreating the entire track live is often impossible.

The most popular DAW for live music performance is Ableton Live, which is optimized and branded for this purpose. What differentiates this DAW from others is its session view that allows the artist to make small clips/loops of music that can be triggered using a MIDI controller. MIDI controllers often consist of a matrix of buttons that can be used to either play MIDI notes, which are often used to trigger drum sounds, or to launch audio clips which allows an artist to use these to create a live arrangement out of pre recorded material. Often additionally to the buttons, the MIDI controllers have knobs which can be mapped to parameters such as filters to make the performance more dynamic, or faders which can be used to adjust the volume of the layers. An example of a MIDI controller which is designed for live performances and has both the button matrix, knobs and faders is the APC mini by Akai Professional (Figure 2.1).

Figure 2.1: APC mini by Akai. An example of a MIDI controller used for live performances [15]

2.1.3 How does a performer of electronic music interact with an audience?

The interaction with an audience is a fundamental part of live concerts. This section will explore how performers of electronic music interact with an audience and how this interaction differs from other types of live concerts.

The main method of interacting with an audience at live concerts is through the music itself, but from a series of interviews with DJ's about their interaction with a crowd, it becomes clear that they do a lot more to establish an interaction [16]. The DJ's mentioned that communicating through body language and facial expressions, through dramatizing their technical movements, seeking eye contact and expressing their enthusiasm they try to enhance the audience's perception of the DJ's presence as a live entertainer. Often they combine these actions by taking the role of VJ, lightning technician or oral entertainer. Besides establishing a presence by outputting all this information, an important part of being a DJ is responding to the signals the audience is sending and adjusting the music based on these. These signals could be communicated through body signals, facial expressions or verbal communication. This acquired information will then be used in the music creation process. Like mentioned earlier there is a difference between DJ's and performing artists of electronic music and this difference affects how this information would be used. A DJ would use the input from the audience to determine which songs to play next, while performing artists have a limited choice of songs they can choose from, but have the freedom to change the arrangement of those songs according to the input they get from the audience.

This interaction with an audience is similar to live concerts of other music genres. The use of body language and facial expressions in order to enhance the stage presence is just as important for any other music group, just like the importance of sensing the mood of an audience and reacting based on this. Essentially the only thing that differs between music groups is how this information can be incorporated into the music, which depends mostly on the freedom to improvise. A study on the perception of liveness within audio recordings found a correlation between improvisation and the perceived liveness [17]. The paper defines liveness "as any decision-making that is made during the performance rather than in advance". Improvisation can occur through improvising a solo or adjusting the length of a section, or it can occur through changing the timbral information by for example, changing the dynamics, the position on an instrument where a note is played or the use of an audio effect. The freedom to improvise is mostly dependent on the members and instruments of a music group. For example, a solo performing artist of electronic music has a lot of freedom to improvise the structure of the song but limited freedom in adjusting the dynamics of the already prerecorded loops. While members of a classical orchestra have little freedom over which notes to play, the differences in dynamics and timbre will be different every time.

2.2 Embodiment and sound

Throughout history, embodiment and music have always been interconnected. The embodied nature of music, the indivisibility of movement and sound, characterizes music across cultures and across time [18]. It has only been the last hundred years or so that the ties between musical sound and human movement have been minimized, which is partly due to the fact that the act of music making has shifted from a public activity towards an expert based activity that has since created a distinction between performer and audience. [19]. This section will explore the relationship between movement and sound and will investigate how to measure and objectify movement.

2.2.1 How does movement relate to music perception and creation?

During the beginning of the twenty-first century, the embodied music cognition theory gained traction. This theory focuses on the idea that the body is the mediator between the external environment and the mind and that therefore the body plays a crucial role in the perception of music [20] [21]. The relationship between music perception and movement also works the other way around, a study investigating the effect of rhythmic music on self paced oscillatory movement, discovered that the perception of music also affects the pacing of movements [22].

Musical activity is often described through "gestures", instead of "movements". According to [23], the main reason for doing this is that the notion of gesture somehow blurs the distinction between movement and meaning. Movement denotes physical displacement of an object in space, whereas meaning denotes the mental activation of an experience. They continue to say that the notion of gesture somehow covers both aspects and therefore bypasses the Cartesian split between matter and mind. Musical gestures can be divided into four main categories [23]:

- *Sound-producing gestures* are the ones that are effective in producing sound. For example, when playing a song on the piano, the sound-producing gestures would be the action of pressing the keys.
- *Communicative gestures* are intended mainly for communication. An example of this could be making a gesture to the audience to sing along, or giving a cue to a band member.
- *Sound-facilitating gestures* support the sound-producing gestures in various ways. In the example of playing the piano, the sound-facilitating gestures would be all the gestures that are made that facilitate the pressing of the key. This could be moving the hands to the position of the next note to be played, or moving the upper arms or body while playing.
- *Sound-accompanying gestures* are not involved in the sound production itself, but follow the music. A clear example of this is dancing. Dancers do not contribute to the music but instead make movements that are synchronous with it.

The type of gestures used varies greatly between dancers and musicians. An example of how the functions of gestures relate to each other within both disciplines is shown in figure 2.2.

2.2.2 How can you measure movement?

Measuring human movement is a complex task and can be done through several different methods.

One of these is through measuring physiological signals. Mechanisms such as muscle activation, for example, can be measured in the form of electrical activity with sensors that are put in contact with the skin (electromyography) [24]. EMG's have been used before to create musical pieces [25]. What is interesting is that these techniques can also be sensitive to 'pre-movements' or muscle tension even if there is no significant visible movement.

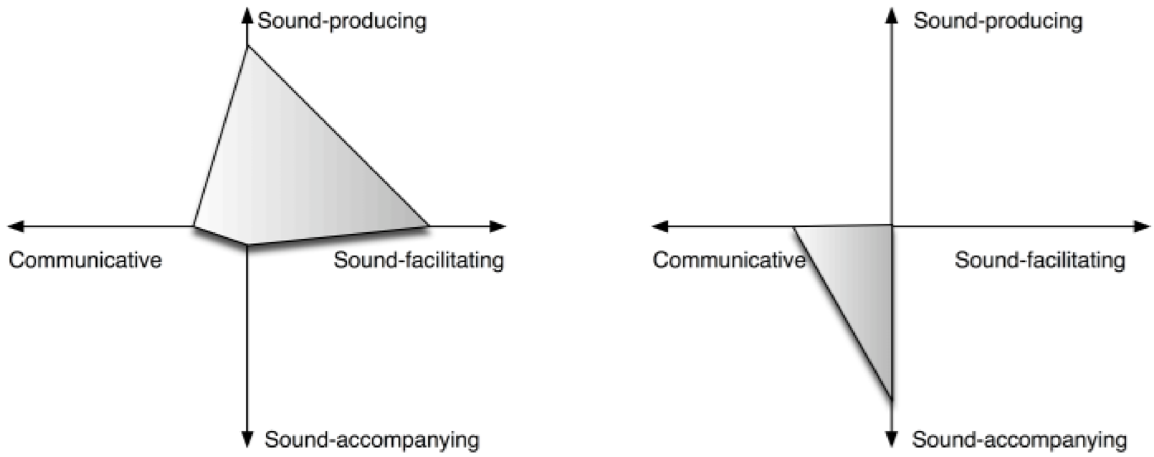

Figure 2.2: "Dimension spaces illustrating how the gestures of a musician (left) and a dancer (right) may be seen as having different functions. Here the musician's movements have a high level of sound-producing and sound-facilitating function, while the dancer's movements have a high level of sound-accompanying and some communicative function." [23]

Another method of capturing movement is through video analysis. One way to do this is through object tracking in video footage. This makes it possible to access continuous information from gestures, after the desired object is manually selected. [26]. Since post-production in most music applications is not desired, often the Microsoft Kinect is used to track movement, which given it has a depth camera, makes it possible to track movement in real time. Only a day after the device had been released as a game controller for the XBOX it had been hacked for the use of other projects. This has ignited some interesting projects, including applications for musical control and expressiveness [27]. A similar device that sparked a lot of interesting musical applications is the Leap Motion. This device makes use of two normal and three infrared camera's that are used to track hand movement. This device has been used to control musical notes, audio effects, parameters on synthesizers and the individual grains of a granular synthesizer [28].

Another common approach of sensing movement is through the use of accelerometers and gyroscopes. An accelerometer can measure the acceleration of an object into a specific direction while a gyroscope uses the earth's gravity to measure the rotation of an object. Often these two are combined into one sensor. The output of this sensor can be used to detect raw movement, but can also be used to learn complicated gestures using machine learning [29] [30]. The sensor has been used for a wide array of musical applications, by attaching the sensor to feet [1], gloves [5] and wrists [31].

Stretch sensors are also a useful tool for detecting movement. They measure deformation and stretching forces such as tension or bending and give an analog output depending on the intensity of the deformation. These sensors have either been used on the body, where for example a sensor can be placed on a finger to measure it's bending after which that data can be transformed into music [32]. Using them on fabric and other deformable objects can also result in some interesting musical interactions [33] [34]. A popular implementation of this sensor is the Seaboard Rise by Roli that has it implemented inside a MIDI keyboard in order to add an extra layer of expressiveness to the the playing [35].

Finally, instead of looking at the movements of the body, it is also interesting to look at the placement of a body relative to a space. For such installations, common paradigms are, for example: body presence/absence, crossing borders, entering/leaving zones. Motion can also be naturally associated with these interactions, for example, by measuring the "quantity of motion" in a particular spatial zone [24]. An example of an installation that tracks the position of a person in a space, while this position is projected on the floor along with game elements is the interactive playground [36].

2.2.3 How can you objectify movement?

Besides looking at the different technologies for sensing movement, it is also interesting to have a look at the different frameworks for objectifying movements.

One of the most common methods of objectifying and evaluating gestures in the field of HCI is through Fitt's law. Fitts' law states that the amount of time required for a person to move an object to a target area is a function of the distance to the target divided by the size of the target. Thus, the longer the distance and the smaller the target's size, the longer it takes [37]. While this method can accurately evaluate the efficiency of a movement, it is rather simplistic and does not cover the intended emotion behind this gesture, which in the case of a musical performance is a valuable part of information.

A method for transcribing human movement that covers the intended emotion is the Laban movement analysis [38]. Laban categorized human movement into four component parts:

- *Direction* is either direct or indirect.
- *Weight* is either heavy or light.
- *Speed* is either quick or sustained.
- *Flow* is either bound or free.

Each unique combination of these components is a different Laban effort, with its own emotional association. This theory has been the basis of some interesting dance

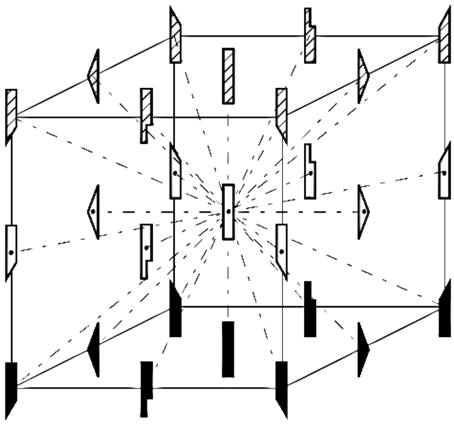

Figure 2.3: The diagram of the positions of the labanotation makes it possible to represent the movement of the part of the body in space (in 3 dimensions) around it. [42]

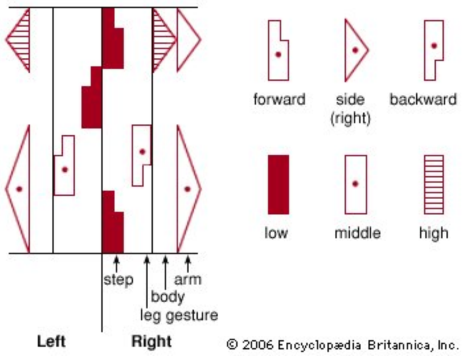

Figure 2.4: Key to labanotation symbols [43]

to music projects [39] [40]. However, using solely the Laban Movement Analysis can be quite limiting for the control of an application since it only covers how a movement is made and not what the movement is.

Consequently, Rudolf Laban also developed a dance notation system called Labanotation, which is one of the most widely used transcription methods for dance [41]. The method is built around vertical bars that have time on the y-axis and the limb that is being moved on the x-axis like shown in Figure 2.4. The direction that the limb will be moved in is represented by a symbol that indicates the direction in a 3D space like illustrated in Figure 2.3. From this notation can be derived that once the 3 dimensional direction of the feet, legs, body and arms can be measured, enough information can be gathered to capture a dance. However, the labanotation method has some additional elements, like symbols that give specification on the limb and rotations. Measuring this as well, makes the capture of a dance already significantly harder, but can be a good framework that can capture a dance completely.

2.3 State of the art

Now that some background information on performance, electronic music and embodiment has been established, a few applications of embodied music controllers will be discussed.

2.3.1 Theremin

Creating music through movement was first done in 1919 when Léon Theremin invented the theremin. This instrument can be played by varying the distance between both hands and two antennas. Here, the distance from one antenna influences the volume and the distance from the other antenna influences the pitch. The instrument is known for being very hard to play and its characteristic sound that has been often used in science fiction movies.

2.3.2 A Motion Recognition Method for a Wearable Dancing Musical Instrument [1]

Another interesting example of an embodied music controller is one by Fujimoto et al, 2009. In this project a system was created that could transform dance steps into music. This was done by attaching a 3-axis accelerometer to both shoes and using them to track the steps of a dancer. They evaluated the impact of delay between the actions of the dancer and the sound that is being being outputted. There appeared to be a huge difference in perceived incongruity with a difference in delay of 100ms. The final system was evaluated by assessing how well the system could track the steps, which is quite successful. However, the idea of creating music using this technique has not been evaluated in this paper.

2.3.3 Vrengt [2]

This paper explores the creation of music as a partnership between the musician and a dancer. This is done by adding two Myo sensors to the arms of the dancer. Using sonification the movements of the dancers will contribute sounds to the music. The paper explores the concept of micro-movements where they look at tiny deviations in movement when the dancer is standing still. Besides movements they also add the sound of the breathing of the dancer to the sound output. The paper did not have a grounded evaluation, but did include the experiences of both the dancers and the musicians who used the system. They seemed to find this form of interaction an enjoyable experience.

2.3.4 Dance Jockey: Performing Electronic Music by Dancing [3]

The Dance Jockey system that allows dancers to compose and perform music through movements. It uses a full body motion capturing suit, consisting of 17 small sensors. They wrote software for this suit using the Jamoma framework, which is a community-driven software built for the MAX/MSP programming environment with plug and play modules. The suit has been programmed to respond to three types of modules: cues, actions and transitions. Here a cue is a set of actions, an action is a certain mapping between a movement and an audio parameter and a transition moves from one cue to the next cue. The suit has been used in front of large audiences who responded positively towards clear mapping between movement and sound, but there was no scientific evaluation included in the paper.

2.3.5 Enhancia: Neova [4]

The Neova (figure 2.5) is MIDI controller in the form of a ring which can be used to add expressive information to sounds through hand movement. Using Neova, artists can control for example the pitch, volume or filters of the instrument they are playing. The movement is recognised using an accelerometer and is being sent wirelessly using an omnidirectional radio link. This ring can help artists to make their performances more expressive. However, it is mostly used in combination with other instruments, which sometimes might not be ideal for performing artists who mainly focus on mixing layers together.

Figure 2.5: Enhancia: Neova [44]

2.3.6 Mi.Mu Glove [5]

The Mi.Mu glove (figure 2.6) is a softly commercialized data glove developed and promoted by artist Imogen Heap. The gloves each have a WiFi connected micro controller and a gyroscope/accelerometer implemented. The gloves are communicating with a software that allows the user to map movements and gestures to MIDI data. The glove is designed to be open ended and to be used in a large variety of applications. This has resulted in the freedom for artists to use the system in a creative way. However, given that every performance using the gloves uses a new connection between gestures and audio, this can cause the audience to get confused of what the artist is doing. Probably due to the system being open-ended, there is little scientific evaluation of the Mi.Mu glove available. However, a large array of artists including Ariane Grande have chosen to incorporate the gloves into their live performances, which could lead to the assumption that these gloves are somewhat successful.

Figure 2.6: Mi.Mu gloves [45]

Chapter 3

Development of prototype

In order to investigate the functionality of an embodied music controller that can be used to control the arrangement of an EDM song, a prototype will be developed that fits these requirements. In this chapter, the creation of this prototype will be discussed. Through this process, RQ 3: "How is a prototype of the proposed system built?", will be answered.

3.1 Goal

In this section, the concept of the prototype will be explained. Thereupon, some requirements that the prototype should fulfill in order for it to be successfully used during evaluations will be discussed.

3.1.1 Shortcomings within the state of the art

Looking at the state of the art, there have been many attempts at creating embodied music controllers. Most of these make use of IMU sensors (Intertial Measurement Unit), with the Mi.Mu glove as the most successful example [5]. Mi.Mu gloves have already successfully been used during large-scale concerts. For example, the concerts given by Ariane Grande and Imogen Heap seem to be well-received by the audience. However, regarding the improvisation of EDM arrangements, there are three areas for improvement.

The first problem is that the Mi.Mu gloves mostly focus on interactions with the voice and on using the gloves to record the material on the spot. When looking at the way that performing artists typically perform their music, they often do not record their material live but instead focus on applying effects and improvising the arrangement of their songs. Therefore, when creating a system that is specifically targeted at performing artists that creatively mix prerecorded audio layers, the controller might function better if it solely focuses on controlling audio effects and triggering audio clips.

Additionally, the Mi.Mu gloves have flex sensors built-in for each finger. These allow the performer to train different hand positions and use these positions to map them onto the respective sounds. However, controlling music using fingers could lack some expressiveness since this is hard to notice from a distance. Therefore, it would be worthwhile to explore a system that only uses IMU data, since it nudges the performer into making larger gestures.

Finally, using only a set of gloves for an embodied controller limits the performer to only using their arms for controlment. This limits the amount of embodiment and expressiveness as well since posture plays a big role in people's perception of a music performance [46]. Therefore, it would be intriguing to add an additional sensor to the body that can track its movement and posture.

3.1.2 Concept

After analyzing the shortcomings in the current state of the art, three main features in need of improvement became apparent:

- The prototype should track movement in three locations. On both hands and on the body.
- The prototype should solely use IMU sensors to track movement.
- The prototype should focus on two actions. Controlling (audio effect) parameters and turning audio clips on and off.

3.1.3 Requirements

In order to build a functional prototype, it is important to set requirements that can be used to guide its development. In his section, the requirements that are essential in order to obtain a prototype of the desired functionality will be listed. These requirements can be grouped into three different themes: the quality, the functionality and the flexibility of a prototype.

Quality of prototype

Through the creation of a prototype the aim is to answer several questions regarding the expressiveness, intuitiveness and user-satisfaction achieved through the use of the music controller. Given that these metrics are often correlated with the performance of a prototype, a high-fidelity prototype that functions similarly to a performance-ready music controller, will be developed. This leads to a set of requirements that the prototype should fulfill in order for it to feel like a real instrument. These requirements are the following:

- **Requirement 1: The prototype should be able to respond in real time.** Timing is a very essential aspect within the performance of a piece of music. It is therefore essential that there is no lag between actions and corresponding sounds.

- **Requirement 2: The prototype should be wireless.** Since the music controller is controlled through movement, it is necessary that the user has freedom to move around. Thus, the user should not be restricted by any wires when performing with the music controller. This creates a requirement for wireless communication between the controller and the laptop and for a battery powered controller.

- **Requirement 3: The prototype should be able to send MIDI information to Ableton Live.** Since the controller should be usable for performing artists, it is crucial that it can be used in their work environment. Ableton Live is the Digital Audio Workstation that is mostly used by most performing artists in live situations. Therefore, it is important that the controller can be used to send MIDI directly into Ableton Live.

Functionality

In order for the prototype to function correctly, there is a set of requirements the prototype should fulfil, which will be outlined more comprehensively below:

- **Requirement 4: The prototype should be able to recognise gestures.** Since recognising gestures is one of the main functionalities of this prototype it is important that it can do this accurately.

- **Requirement 5: The prototype should be able to track combinations of gestures.** To increase the range of actions that can be performed, it is necessary that the prototype can recognise the combinations of gestures and send out corresponding MIDI information based on each combination that is recognised.

- **Requirement 6: The prototype should be able to convert the output of the IMU sensor into MIDI information.** Besides checking for gestures, the prototype should also be able to use the raw output of the IMU sensor to control continuous parameters within Ableton Live.

Flexibility

In order to gain insight into which movements and gestures correspond best to different musical parameters and audio clips, it is necessary that the configuration between the sounds and movements in the prototype is very flexible. Below are the requirements that need to be fulfilled in order for the prototype to be flexible enough.

- **Requirement 7: It should be possible to easily train new gestures.** During the configuration of different mappings between sounds and gestures, it is important that the process of training new gestures can be done easily and quickly.

- **Requirement 8: It should be possible to easily configure new combinations of gestures.** During the evaluation sessions it will be critical to explore different mappings between sounds and gestures. Therefore, it is important that new combinations can be created easily.

- **Requirement 9: It should be possible to easily map the combinations with parameters inside Ableton Live.** When exploring different configurations between sounds and gestures, it is also important that the sounds can easily be changed inside Ableton Live.

3.2 Hardware

This section will cover the decisions that have been made concerning the hardware of the prototype. A picture of the prototype is shown in figure 3.1. This section was split up into several parts, which are also illustrated in figure 3.1. First, the working and choice of IMU sensors will be explained (section 3.2.1). Afterwards, the choice of microcontroller will be justified (section 3.2.2). Later on, the choice of battery and charging structure will be discussed (section 3.2.3). Finally, the wiring of everything will be displayed in section 3.2.4.

Afterwards, some information will be given on the attachment of the sensor. In section 3.2.6, later changes of the prototype will be discussed.

3.2.1 Motion sensor

The motion data will be collected through the usage of IMU sensors. An IMU sensor is an electronic device that measures and reports a body's specific force, angular rate, and sometimes the orientation of the body, using a combination of accelerometers, gyroscopes, and sometimes magnetometers.

The accelerometer measures the angle in which the device is oriented and outputs this for both the x,y and z axis. These outputs are also often referred to as pitch, roll and yaw, like illustrated in figure 3.5.

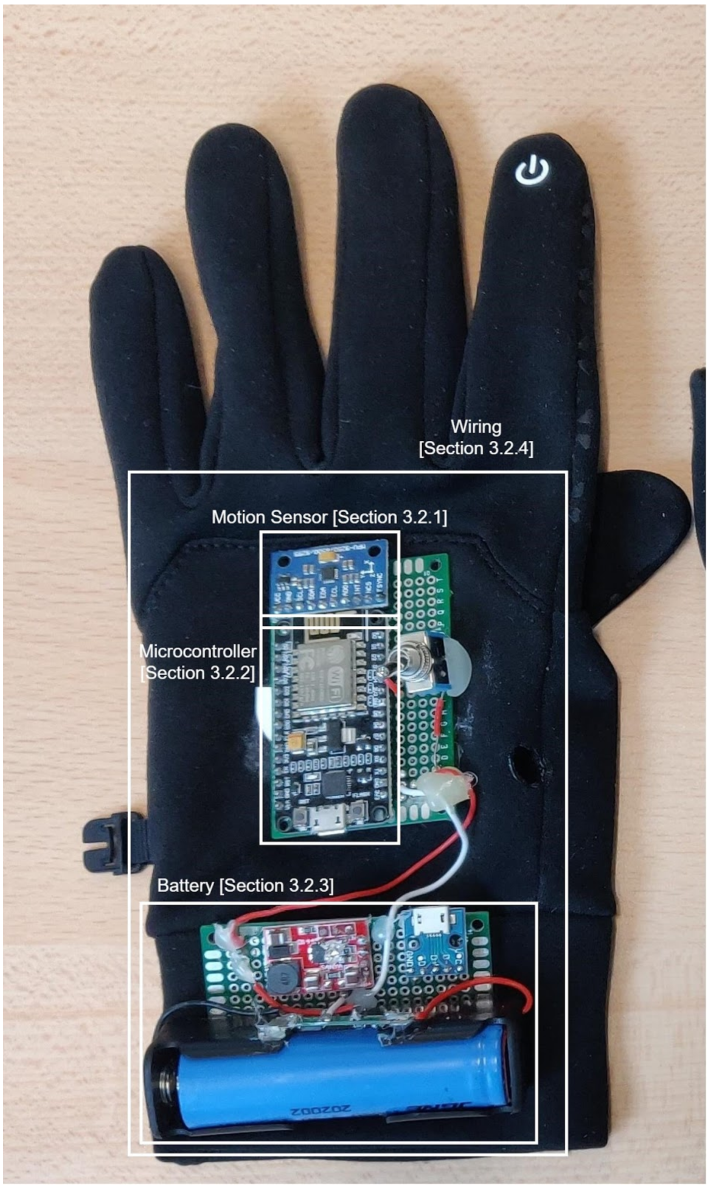

Figure 3.1: First prototype along with the corresponding sections

The gyroscope sensor is a device that can measure the angular velocity of an object and is measured in degrees per second. Angular velocity is the change in the rotational angle of the object per unit of time.

A magnetometer is a sensor that measures the magnetic field of each axis. In most cases, this is the earth's magnetic field. With some calculations, the output of this sensor can be converted into a compass that can measure the orientation.

The most accessible IMU sensor is the MPU6050 (figure 3.3) which combines an accelerometer and a gyroscope. Given that both sensors output in three axes, this sensor can be classified as having 6DOF (degrees of freedom). A common IMU sensor that also has the magnetometer implemented is the MPU9250 (figure 3.4), which has 9DOF. The choice has been made to implement the MPU9250 in the devices on both

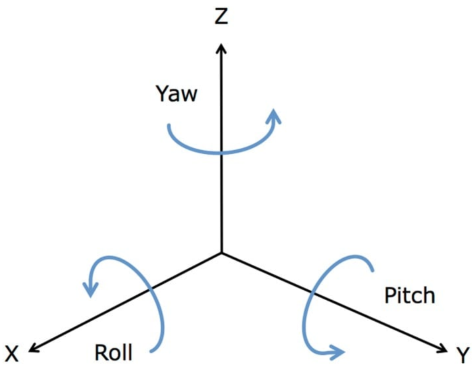

Figure 3.2: Different axis of accelerometer sensor [47]

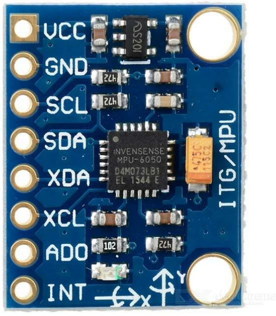

Figure 3.3: MPU6250 [48]

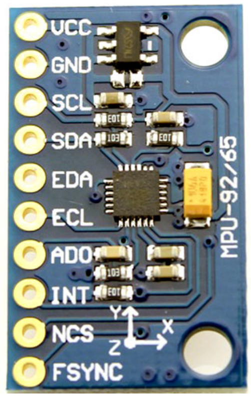

Figure 3.4: MPU9250 [49]

hands, and to implement the MPU6050 on the body.

3.2.2 Microcontroller

To be able to read the data from the IMU sensors and send it to a laptop, it is necessary to use a microcontroller. There are several ways to achieve wireless communication. The three main ones being radio, bluetooth and WiFi. The communication between the microcontroller and the laptop was enabled through the use of Wifi, since both radio and bluetooth have a limited range.. Accordingly, WiFi can be considered as the most advantageous medium since it eliminates constraints in the performer's freedom to move. When looking at how other devices like the Mi.Mu gloves have implemented their wireless communication, it becomes clear that WiFi is most often used for this purpose.

To be able to achieve communication through WiFi, the NodeMCU board will be used (Figure something). The NodeMCU board makes use of an ESP WiFi module and has enough pins for a wide array of prototyping functions. The built-in WiFi module is able to maintain multiple TCP/UDP connections. How these connections are set up will be explained in section 3.3.1.

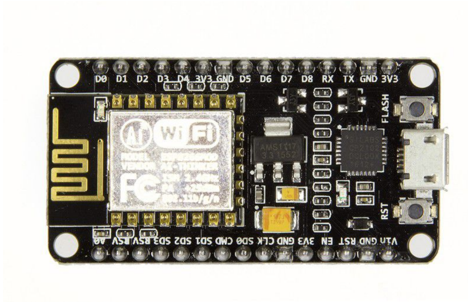

Figure 3.5: NodeMCU [50]

3.2.3 Battery

Using a battery is essential when fulfilling the wireless requirement. However, there is a wide variety of different batteries that could be used. For this project, a rechargeable battery will be used. Since the product will be used frequently, this can be considered as the most sustainable and cost-efficient option.

There are several types of rechargeable batteries that can be used. The most modern and commonplace options are Lithium polymer (Li-Po), lithium-ion (Li-ion) and Lithium Iron Phosphate (LiFePo4) batteries. Among these, the LiFePo4 battery is the most suitable battery for this project for two reasons.

First, lithium batteries are often very unstable and have a tendency to explode if they are overcharged. Given that the batteries will be attached to the body of the researcher as well as to participants during user tests, this should be avoided at all costs. The LiFePo4 batteries, however, are very stable and have a high tolerance for overcharging.

Second, LiFePo4 batteries output a slightly lower voltage (3.2V) as opposed to the Li-Po (4.2V) and Li-Ion batteries (3.6V). This slightly lower voltage makes it possible to attach the battery directly into the 3.3V input of the NodeMCU, thus circumventing the need for a voltage regulator.

While LiFePo4 batteries are less prone to overcharging, it is still necessary to use a suitable charger for this battery. One of the most common chargers for this purpose is the TP5000 (Figure 3.6). This charger supplies a safe amount of power to the battery and can indicate when it is fully charged. To be able to power the charger, a micro USB breakout board is used. Thereupon, a small protection circuit is added to the circuit in order to prevent the battery from undercharging.

3.2.4 Wiring

A schematic of the wiring of the different elements can be found in figure 3.7. To increase the usability of the prototype, a switch has been added to turn the electronics on and off. Also, a LED was added to let the user know that the device is turned on.

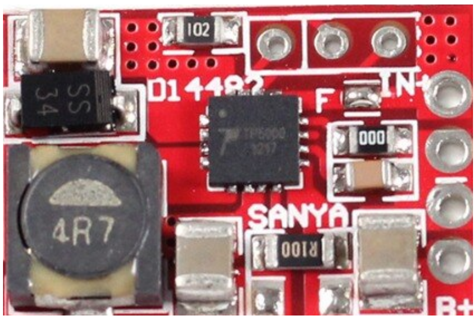

Figure 3.6: TP5000 Battery protector [51]

As visible in figure 3.1, the electronics are glued onto a pair of gloves.

This schematic includes the button that was added after the first set of user tests. This issue will be discussed more comprehensively in section 3.2.6.

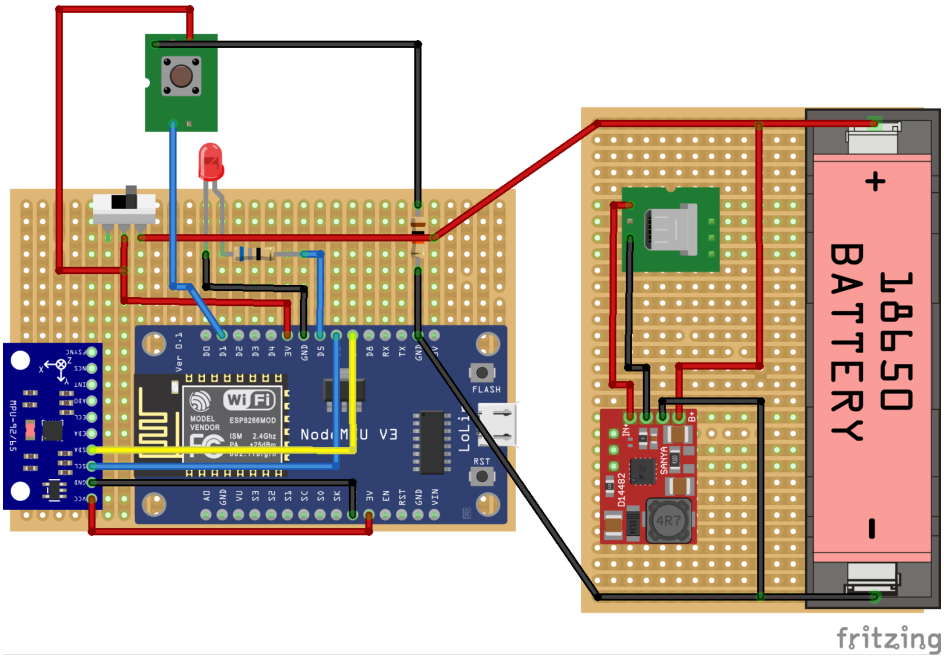

Figure 3.7: Schematic of prototype

3.2.5 Revisions

After the first set of user tests, two changes were made to the prototype in order to ameliorate two points of concern: the vulnerability of the electronics and the indirectness of triggering sounds.

To make the controller more robust, a 3D print for the electronics was designed, which can be seen in figure 3.8. The print was designed in such a way that the electronics that already have been made, fit perfectly inside. Furthermore, the bottom part of the electronics was partially left open, since this enables the usage of the micro

USB ports that are used to charge the battery and uploading new code onto the microcontroller. It also makes it less effortful to replace the battery if necessary. There are two holes in the top of the case. One hole is used to attach the switch, the other is used to add an LED which will indicate when the device is on and if it is connected.

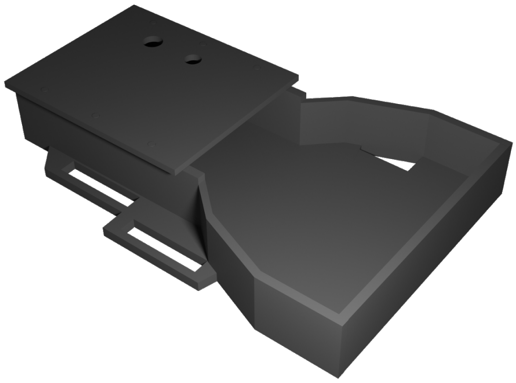

Figure 3.8: Render of designed 3D print

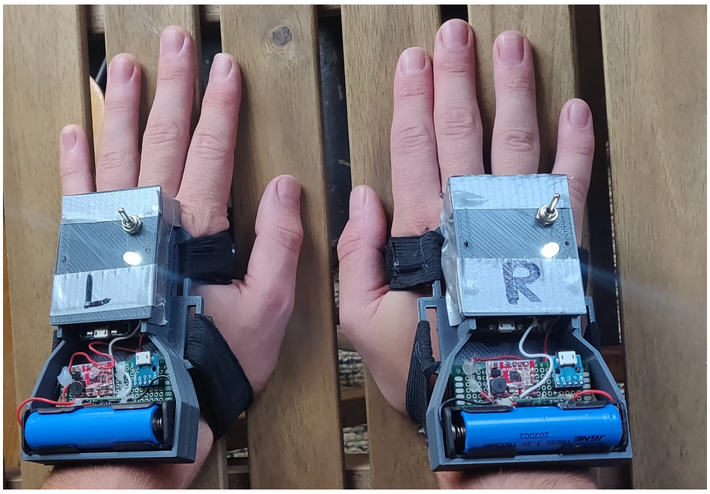

Figure 3.9: Prototype after revisions

The device will be attached with the use of two elastic bands: one around the palm of the hand and one around the wrist, as can be seen in figure 3.9.

The indirectness of the sound is solved by adding a push-button to the device. The implementation of this button in the electronics is already visible in figure 3.7. The button will be attached on the inside of the elastic band that goes through the palm. This enables the user to press the button when he/she closes the hand.

3.3 Software

This section will cover all the information concerning the programming of an interface that can be used to easily create new mappings and perform them. The final interface is shown in figure 3.10. Since this interface is only meant to be used by the researcher, the functionality was prioritised over the aesthetics or user-friendliness of the interface. The interface is created in Python using a library called Tkinter. This chapter will be split up into the following sections: Retrieving data, training gestures, mapping gestures to MIDI, starting a performance, saving projects and revisions.

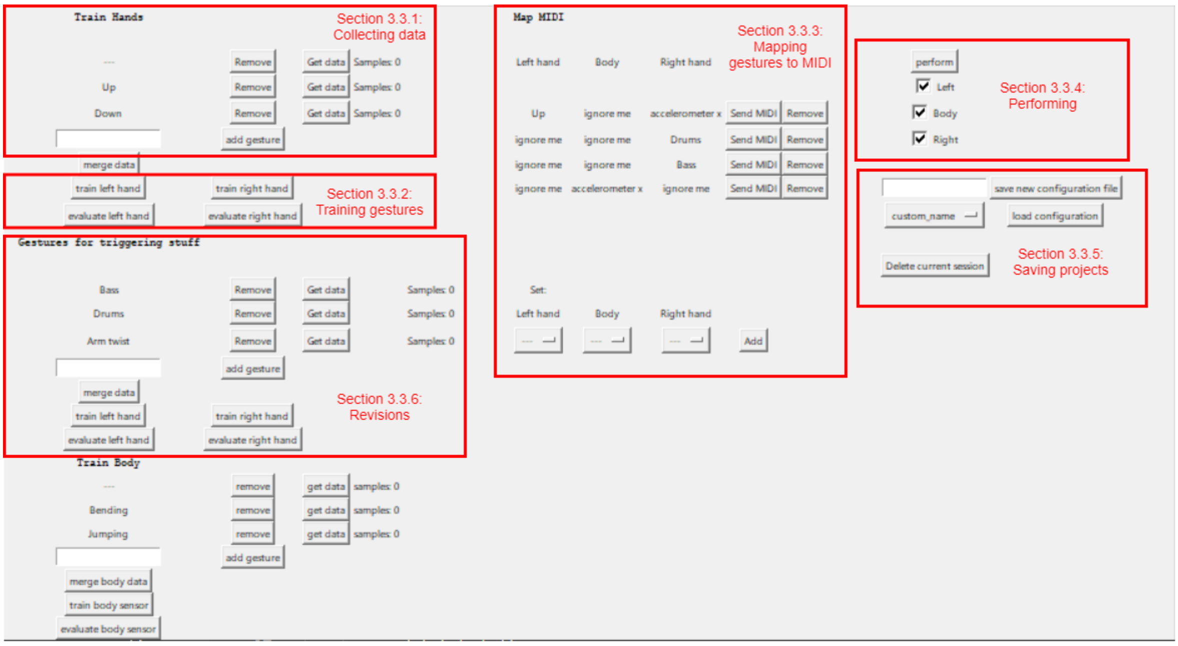

Figure 3.10: Interface for training, mapping and performing gestures

3.3.1 Collecting data

This section will describe how the data is being collected. It will go over how the sensor data is read and sent out from the microcontroller to the laptop. Afterwards, the creation of gestures and how data is assigned to these, will be discussed.

Reading the sensor data

The code that is running on the microcontroller has two main functions: retrieving data from the IMU sensor and sending this data to the laptop. In this section, it will be discussed how the data is retrieved and cleaned up.

The MPU9250 sensor makes use of an I2C serial bus. This bus allows sensors to send a large array of data to a microcontroller using only two pins. These pins are called the SDA (serial data) and SCL (serial clock). The SDA sends the data and the SCL determines when. Writing code that can receive information from an I2C bus can be a lot of work. Luckily, there are libraries that make reading this data simpler. For this project an MPU9250 library that can easily access the nine values that the sensor is outputting.

After the data is read, a running average filter is applied over the data to smoothen out any noise that the sensor is outputting. Afterwards, this data is scaled in such a way that the value ranges between -128 and 128.

Wireless communication

To be able to send the sensor data to the laptop, a socket is created. In this application, a Python script is running a server and the three devices are running a client.

There are two different protocols through which packets could be delivered through a socket: UDP (User Datagram Protocol) and TCP (Transmission Control Protocol). The way TCP works is that it first establishes a connection between server and client and then makes sure that all packets are received. UDP does not need this initial connection and instead sends out all the data without checking if it will be received. This results in a difference where TCP is more reliable and UDP is quicker. Since the prototype needs as little delay as possible and a missing data packet would not make a big difference, the UDP protocol is used.

This UDP socket connection can be established relatively easily by using the socket library for python and the WiFiUDP library for Arduino.

To be able to send data from the devices to the laptop, the device needs to be connected to the WiFi network. To prevent having to update the code on the all three devices whenever the system is used on a different network, they will be connected to a mobile hotspot that is hosted on the laptop. The device also needs an IP address to be able to know where it is sending data to. Fortunately, when using a hotspot the IP address of the host stays the same. Additionally, the socket does not require an outgoing internet connection, which means that the devices can be used anywhere as long as the hotspot on the laptop is turned on.

Creating gestures

Gestures are created through entering the name of a gesture in the input field and pressing on the button “add gesture”. This will add the gesture to the list of gestures. Next to each gesture is the option to remove it or to retrieve data for this specific gesture.

The button that is used to retrieve the data starts a loop that saves all incoming data into a dataframe until a fixed number of data points has been reached. The data is collected based on which device is turned on. It is also possible to collect data for a specific gesture using both devices at the same time.

3.3.2 Training gestures

In the following section, the process of training various gestures using a machine learning model will be elaborated on. To this end, the section will comprise two sub-sections. First, the decision making process of choosing the right model will be elucidated. Second, the model's implementation and evaluation will be explained in more detail.

Choosing a model

To be able to recognise gestures, a machine learning model is used. The goal of this model is to have the data of the IMU sensor as an input and assign a gesture as an output. This can be achieved through supervised learning. Supervised learning is a type of machine learning in which the model is trained on data which has an input and an output. During training the goal for the model is to learn how to arrive at this output. Here the output is a single gesture, therefore this can be called a classification algorithm.

There are several different types of classification algorithms. Some examples are logistic regression, k-nearest neighbor, support vector machines and decision trees. For this specific application, the temporal information within the sequence of data points is of importance. Therefore, a recurrent neural network (RNN) is a suitable option. Specifically, an RNN LSTM model will be used in this project, since they provide a good solution for problems with long-term dependencies that other RNNs suffer from and they have been proven to work well in other other projects with similar tasks [52].

Implementation

The code for the implementation of this bi-directional LSTM model was adapted from a project devised by Barkowiaktomosz on Github. [53] This application was built with the purpose of recognising fitness activity using the built-in IMU sensor in smartphones. Given that the model architecture seemed to perform well for this project, the same architecture is used for recognising gestures for the prototype. The architecture of the machine learning model is quite basic and is shown in figure 3.11. It has a stacked bi-directional LSTM layer, followed by a dense layer.

Figure 3.11: Neural network architecture [53]

Amount of LSTM layers	2
Epochs	10
Learning rate	0.0005
Amount of hidden neurons	50
Batch size	30
Dropout rate	0.5

Table 3.1: Hyperparameters of model

The hyperparameters for the machine learning model were chosen through iterative testing until a combination of parameters was found to work sufficiently well. These parameters are shown in Table 3.1.

Before the data can be fed into the model, it has to be preprocessed. This is done through scaling the numeric variables to a range of -1 to 1. The categorical variables of the gestures are converted into binary, using a method called one hot encoding.

3.3.3 Mapping gestures to MIDI

This section will discuss how the mappings between music and gestures are made. This will be done by first discussing what MIDI is, afterwards explaining how MIDI can be used to control music in Ableton Live and finally how new mappings can be created.

MIDI

The communication of musical data is done through a protocol called MIDI (Musical Instrument Digital Interface). MIDI is a widely used protocol for transferring musical information in real time, by communicating the pitch, velocity and channel of individual notes. It is mostly in external MIDI controllers which are connected by USB. In this project, the MIDI information is supposed to be sent by the software after a gesture has been recognised. To be able to do this, a virtual MIDI port is created. This is done using the software, LoopMIDI, by Tobias Erichsen. To be able to send out MIDI information from a Python script, the Python library Mido is used. Using this library it is possible to connect to a MIDI port and send out MIDI messages through it.

This prototype requires two different types of MIDI messages to be sent. The first type are single short messages that can be used for triggering clips or turning on effects. The second type of messages are controller change values. These will be sent out at a quick and constant rate and are used to change the value of parameters such as volume and audio effects.

Ableton Live

The music will be controlled in the session view in Ableton Live, which is shown in figure 3.12. The session view has the intended use of starting and stopping audio clips that loop. Like seen in figure 3.12, the different instrument tracks are displayed horizontally. It is possible to add any audio effect to any of these tracks. Vertically, the sections of the song are shown, which are visible on the master track on the right. It is possible to combine the audio clips from different sections.

Using MIDI it is possible to control any button or parameter. This can be done through Ableton's MIDI map mode. When entering this mode, the user can select any knob or button in the program and send a MIDI note, which will result in an immediate mapping between that button and that specific note.

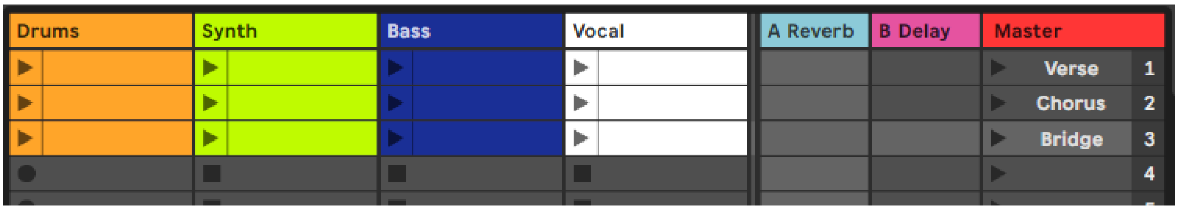

Figure 3.12: Session view in Ableton Live

Creating new mappings

New mappings are created through assigning combinations between gestures or raw sensor outputs. Which can be chosen from a drop-down menu and are added using the "add" button. This will add the combination into a list of gesture combinations and will assign a MIDI note to that combination. During a performance these combinations will continuously be checked for. Once a combination is found, the corresponding MIDI note will be sent out through the virtual MIDI port.

There are two types of combinations that can be added to the mappings. The first type is the combination between one or multiple gestures. Once this gesture occurs, a single note will be sent out that triggers something. The other type occurs when the raw sensor output is added in that mapping. When this combination is recognised, the program will continuously send out the current value of the IMU sensor as MIDI controller change information, until the combination is no longer performed.

Next to each combination is the option to remove it or to send the corresponding MIDI note. Sending MIDI notes manually is useful for mapping these notes to specific parameters or triggers into Ableton Live, as explained in the previous section.

3.3.4 Saving projects

To be able to save different configurations, the program has a file saving system implemented. When a file is saved, a new folder for that configuration will be created. In this folder the gestures, training data, models and mappings will be saved. When a configuration is loaded this folder will be made the main folder. The interface also has a "delete current session" button which will restore the current session to a default template.

3.3.5 Performing

Once the "perform" button is pressed, the program will go into the performance mode. When this mode is enabled, the program will continuously receive data, check for gestures, check for gesture combinations and send the appropriate MIDI information.

To filter out false positives, a running average filter is applied on the recognised gestures.

3.3.6 Revisions

A button was added to the devices on both hands to fix the indirectness of triggering sounds. The goal of this button is to make it possible for the users to press it whenever they want to record a gesture for triggering, and to release it whenever they want the triggering to happen.

To be able to tell the laptop when a button has been pressed, an additional variable is added to the data that the microcontroller is sending out. This variable indicates whether the signal of the button is high or low.

The button will serve as a switch between recording the data for the gestures that will be used to trigger things and the gestures that are recognised continuously. Given that the gestures for triggering things will be a different set of gestures than the continuous ones, a new model for both gloves is created that is trained to recognise the gestures for triggering things. The model and hyperparameters for this model will be the same as the one used for the continuous gestures because this model turns out to work accurately on the triggering data as well.

Chapter 4

Evaluation 1: Co-design session with performers

Now that a functional prototype is developed. A series of tests will be conducted in order to be able to answer RQ 4: "How can the usability of the proposed system be optimised?" This will be done through getting insights into how artists would use a system like this and how this prototype can be improved to better meet the needs of those artists.

4.1 Goals

To be able to answer RQ 4, a set of three goals has to be reached. One of these goals is to get an insight into how well the prototype performs when it is used by different people and how well it performs when new gestures are trained. Additionally, the evaluation serves as a method to find unexpected bugs in the software.

Another goal of the evaluation session is to find a mapping between musical parameters and gestures that are intuitive, logical and expressive. How these gestures are configured will determine how easy the instrument will be to play and how it will be perceived by the audience.

The final goal of this evaluation is to get insights into the usability of the system. The usability of the system covers whether it is easy to use, if it is intuitive and whether the participants would like to use the system during real performances.

4.2 Method

To reach these goals, a series of formative test sessions with dancers and producers will be held. The choice of including dancers is made because of their knowledge in finding movements that fit a certain piece of music. This knowledge might be very valuable when finding a mapping between sounds and gestures. During these different sessions, the participants will be introduced to the system and afterwards encouraged to think about how to successfully apply it in the context of live performances of electronic music and about different mappings between gestures and sounds. The remainder of the session will consist of exploring different ideas the participant might come up with, until at the very end of the session a small semi structured interview will be held. During the course of these sessions adjustments will be made to the prototype based on the results of previous sessions. This way the prototype will be iteratively improved throughout the different sessions until at the end of the evaluation round the prototype will be intuitive and robust.

4.3 Session design

The design of the evaluation session can be split into five phases which are explained in more detail below. These five phases are: Introduction, determining the musical parameters, determining the gestures, iterative mapping and interview.

Phase 1: Introduction

The session will start with an introduction of the system. An example live set with pre-configured gestures will be prepared and will be presented to the participant to give them an initial idea of the workings and possibilities of the prototype. It is important for this live set to explain in a broad sense what's possible and not limit the participant into a certain direction of thinking.

Phase 2: Determining the musical parameters

Prior to the session, the participants will be asked to bring the project file of a song of theirs. After the introduction the participant will be asked to guide the researcher through the file and let them explain what they normally control during live performances. They will find automation curves, clips and audio effects that will be essential to control and can later be mapped onto the prototype.

If the participant did not prepare a song, he/she will be asked to create a very simple song using loops from the internet. This song will then be analysed on the important bits that can be controlled so that these can be mapped onto the prototype.

In the case of testing with dancers, the researcher will have prepared a song and has selected parameters to control. During this scenario the emphasis of the session will mostly be to find gestures that are fitting to these parameters.

Phase 3: Determining the gestures

Afterwards the participant will be encouraged to think of different gestures that can be used to control the previously defined musical parameters. These gestures will be trained by performing the same gestures for a given amount of time with both hands or body.

Phase 4: Iterative mapping

The fourth phase focuses on the mapping between the trained gestures and musical parameters. The participant will be given the task to find a combination between a musical parameter and a gesture. This mapping will be made by the researcher and the participant will be asked to try it out. Afterwards the participant will be asked a series of questions on how the action felt. The action will be evaluated on these criteria.

- Does the action feel right?
- Does the action provide enough control?
- Does the action do what it is supposed to do?
- Does the action make sense?
- Does it add to performance?
- How is this action useful for the performance?

While the participant is trying out new configurations, motives behind different actions will also be discussed. This will be done by occasionally asking one of the following questions:

- Why did you make that movement?
- What did you hope would happen after this?

This phase will be continued until an interaction is found that satisfies these requirements.

The situation might occur when a participant wants to train new gestures or wants to add more content to the song. This can be done by going back to the corresponding phase and continuing from there.

Phase 5: Interview

At the end of the session a small semi structured interview will be conducted. This will be done by first asking some general questions:

- What is your experience with live concerts?
- What equipment do you use?
- What are the things that are essential to control during a live performance of EDM?
- What are gestures that performing artists typically make?
- Do performing artists see the use in controlling their music using movements?

Afterwards some questions about the prototype will be asked:

- Could you see yourself using this in a performance?
- How would you use it?
- Did you think about the perception of an audience when thinking of gestures?
- Do you think the aesthetics of the product matter?
- Do you have any other feedback?

4.3.1 Data processing

The data that will be collected will be in the form of written notes and video recordings that afterwards can be studied. The interview will after the session be transcribed.

Since the data will be mostly qualitative, the retrieved information will be analyzed using reflexive thematic analysis [54]. The data will be coded in order to find recurring themes throughout the different sessions.

4.4 Results

In total ten sessions have been held. The sessions gave some interesting insights into how the instrument was perceived by the participants. After four sessions some major changes have been made to the prototype. Therefore the results will be split up into two sections. The first section will cover the evaluations before the changes to the prototype were made and the second section will cover the evaluations after the changes to the prototype were made.

4.4.1 First evaluations

The first set of evaluations was held with four participants. Each session was about two hours long. The group of participants consisted of three music producers and one dancer. The results from these evaluations are grouped into four categories: the quality of the prototype, the mapping between gestures and sounds and the usability of the prototype

Quality of prototype

During these four sessions the prototype was usable by all participants. The participants managed to use the prototype in such a way that they were able to perform the actions they wanted to make. However, there were three main issues with the current prototype.

One of the issues was the robustness of the prototype. Since the electronics were glued to a pair of gloves, some of the wires kept moving when the users would move their wrists. This movement would cause some of the wires to loosen or break after using them for a longer period of time. During one session, the evaluation had to be paused for approximately 10 minutes because a wire had to be fixed. During another session a wire got slightly loose which caused the connection to temporarily break when sudden movements were made. These breaks in connection cause the devices to restart, resulting in a small period where the device is not sending data to the laptop. This resulted in a bit of frustration in this participant when he tried to perform gestures but the music was not responding correctly.

Another issue arose when the participants were introduced to the system through an example. The gestures in this configuration turned out to be specifically trained to recognise the gestures of the researcher. This resulted in the prototype not being able to pick up some of the gestures that the participants made, which caused frustration in some of the participants. When the participants were able to train the gestures themselves the system responded better to their movements.

Finally, the attachment of the sensors caused an issue as well. For example, the way the body sensor was attached allowed it to slip down after certain movements. This caused the participants to be more aware of sudden movements they had to make. One participant came with the solution to wear the sensor diagonally over the shoulder. This helped with the stability of the sensor but put the sensor in a diagonal position as well, which had an effect on its output. Another participant, added this sensor onto head. This seemed to work better for controlling parameters, however using the head to control quick gestures resulted in a headache. The attachment of the sensor on the gloves created an issue as well, because the gloves were a fixed size and the size of the hands of the participants varied. This caused the sensor to be attached loosely on the hands of participants with smaller hands. This influenced the data from the sensor and created the need for the participants to make sure that the gloves were not falling off.

The imperfections with the prototype caused two of the participants to mention that the prototype should work perfectly, if they would consider using it during a performance. One participant elaborated on this by explaining that when an artist is performing, the artist should be in control of the music and when the prototype does not function properly, the artist can never be fully in control.

Despite the fact that the prototype in its essence was some electronics glued to a pair of gloves, all participants spoke positively about its aesthetics. They argued that the visible electronics and wires would help the audience to realise that the artist is controlling something.

Mapping of gestures

An important aspect of this evaluation was to understand the different mappings between gestures and sounds and to understand which mappings the participants would prefer. After coding the data it became clear that the motivation behind the different mappings could be explained through two different themes.

One recurring theme was the mimicking of familiar gestures. During the sessions, all participants impersonated a gesture or activity that occurs during live concerts in some form. These impersonations can be split up into coming from three different sources.

One type of gestures that was mimicked, was the type of gestures that DJ's make while performing. During the sessions all four participants at least once referred to movements that DJ's make during live shows. Most of the time they would refer to the movement where DJ's are pushing their hands up in the air, often during the drop. During one session, when the participant was exploring gestures related to the movement from the buildup to the drop, he chose to use a hand pumping gesture to activate the drop. Another recurring gesture inspired by DJ's was the turning of a knob. When assigning effects to the raw output of either the x-axis or the y-axis, two participants would motivate their choice by comparing the movement of turning the y-axis to the turning of a knob.

Another type of gesture was the impersonation of playing a physical instrument. When an instrument in the song occurred that could be played on a physical instrument like drums, guitar or keys, two participants would want to trigger that sound by pretending to play that instrument.

Finally, one participant would mimic the gestures that are normally performed by the crowd. As a way to start and stop the lead sound in a song he chose clapping as a gesture. His reasoning was that during for example jazz concerts whenever a solo would start or stop, the audience would applaud.

According to two participants, mimicking gestures of familiar settings helps them with the intuitiveness because it helps to remind them what gestures belong to what sound. They also argue that an audience would understand this coupling and that these gestures enhance the expressiveness as well.

Another motivation behind the chosen mappings was the energy level that posture and the position of the arms can convey. According to multiple participants, low arms and a bent body suggest a low amount of energy, while standing straight with arms pointing upwards indicates a high amount of energy. This reasoning was given when a participant was finding a gesture for the buildup section of a song. He ended up going for the gesture of moving the arms from low to high to increase the energy within this buildup.

This motivation also was used often when different musical parameters were mapped to the x-axis of the accelerometer. Participants would argue that moving the arms up and down is an intuitive way of adding more of something. Where moving the arms up implies that more is added.

Almost all of the mappings that were created throughout the sessions were mapped to movements of the arms instead of movements with the body sensor. The only movement that participants sometimes used for the body was bending downwards, this was mapped to a low pass filter, making the sound more muffled when the participant would bend over.

Mapping of audio

During the sessions, the use of several different audio effects was tested. The participants preferred the mappings which had the most noticeable result in the overall song. When more layers were added to the song, it became harder for them to notice effects that were applied on the individual instruments.

The audio effects that were most applied by the producers were filters. Both high pass filters and low pass filters were applied and very noticeable in the mix. Other successful parameters that were adjusted were reverb, volume and distortion. During one session, a participant also experimented with the panning of different tracks. Unfortunately, this effect was not very noticeable mainly because the speakers were relatively close together during testing.

Usability

All participants managed to use the system in such a way that they could use it to perform the actions they wanted to. The participants especially enjoyed controlling parameters by using the raw output of the sensor. This felt intuitive because there was a direct relationship between the movement and the thing that was being controlled. However there were also two usability issues with this prototype.

One issue was the absence of a direct relationship between action and result when audio clips were triggered. This is partly because of the fact that in order to prevent unwanted triggers of audio, a running average over the gestures was applied. This caused the gestures to be recognised after some delay and created frustration in the participants when the clip was triggered later than they expected.

Another problem with the usability was that performing actions required combinations of actions that did not feel very intuitive. These combinations were introduced as another method to prevent the triggering of unwanted sounds through false positives. During the first two sessions the example song that was given to introduce the system at the start of the session required combinations of moving limbs that were hard to perform without practice. This was partly fixed during the second session when stationary gestures were trained and used. These gestures would be for example holding the arm in a specific way and keeping it there. These gestures allowed for more intuitive combinations because only one hand would do the moving.

4.4.2 Changes to prototype

After the first four testing sessions, it became clear that there were two big problems with the prototype.

The first problem was related to the robustness of the prototype. During several sessions the wires of the prototype would break or get loosened, which would cause interruptions during the sessions since the prototype had to be fixed afterwards. The reason why the wires broke is most likely because of the way the electronics were attached to the hands. By glueing the electronics to a pair of gloves, the wires would have to constantly adapt to the angle of the wrist. Additionally, the gloves were a fixed size which caused them to be too loose on the hands of some of the participants, which would make the electronics even more. Therefore, some adjustments to the way the electronics are being attached to the hands have been made, in a way that the electronics are protected and fixed on the hands. More details on these adjustments can be found in section 3.2.5.

The second problem with the prototype was related to the triggering of audio clips. Triggering clips makes up a big part of the system, however multiple participants have indicated that they are not liking the way it feels to perform this action. This is mainly because of the slow response time of the system recognising the gestures. This caused an unpredictability of when exactly the MIDI is sent to Ableton Live, which can be frustrating in a musical context. The slow response time is caused by the moving average that is applied to the retrieved gestures, which is necessary to filter out the undesirable gestures that are recognised when the performer is switching between different gestures. A solution to this problem is to separate the gestures that are used to select the parameters the user wants to control and the gestures that are used to trigger audio clips. This can be done by adding a button that while it is pressed records the data of the gesture the performer wants to trigger and when it is released immediately recognises the gesture and sends out the corresponding MIDI information.

Another motivation behind adding this button, is the improvement in the accuracy of recognising gestures. This is because of two reasons. The first reason for this improvement is that, when using the button, the user can determine when to start and stop the data collection. This is in contrast with the current prototype, where the data of the sensors is continuously checked against the model. This becomes a problem when a user switches between different gestures. The motion of switching gestures is not added to the model as a gesture but will still be recognised as one, which can cause some unwanted behaviour. The second reason for the improved accuracy is the fact that with this button, the LSTM model will be split up into two models: one for recognising continuous gestures and one for recognising the gestures that trigger things. This halves the amount of gestures that a model has to distinguish between, which in turn improves the accuracy of the model.

More detail on how this button is implemented can be found in section 3.2.5 and section 3.3.6.

4.4.3 Changes to session design

The first view evaluation sessions already provided a lot of useful information on the motivations behind the creation of a mapping between gestures and sounds. Because of this, combined with the fact that the process of training new gestures and mapping these to parameters inside a new Ableton Live set was the most time consuming part of the sessions, while the most valuable information came from the iterative testing afterwards. The decision has been made to skip creating and mapping a new song within the session and work with a template of mapped gestures that was based on the results from the previous sessions. This template will be discussed in section 4.4.4. Eliminating the creation of an original song during the session, causes phase 2 and 3 to drop out of the session design that was presented in section 4.3. The iterative mapping will now be done with the example that is given in the introduction. During the iterative mapping the participants will still have the freedom to change any mapping or to train new gestures.

The interview will also have some changes because some questions resulted in the same answer for every participant.

- What is your experience with performing?

- How much did you feel in control of the music?
- How easy was it to learn how to use the system?
- How much did you think about the perception of an audience while finding mappings?
- Do you see the use of controlling music using movements?
- Could you see yourself using this in performance?
- How would you personally use it?
- Do you have any other feedback?

4.4.4 Template design

For the next set of evaluations, a template will be used. This template includes a set of pretrained gestures, an initial configuration and a song that is recorded in Ableton Live.

The gestures that were trained for this template were based around two findings of the previous test. The first finding was that the participants preferred to use samples that mimicked familiar music instruments or actions. This has been applied by training gestures that do exactly this. Examples of these gestures are mimicking the guitar, bass, keys and drums. The other finding was that the participants thought that the posture of the body and arms correspond to the energy level within a song. Because of this, gestures are trained where the arms are either raised up or are lowered. Some quick gestures which are easily performable are also trained. An example of such a gesture is turning the wrist in a certain direction.

Given the preference for using gestures that mimic familiar musical instruments, a song has been recorded that includes these instruments (guitar, violin, drums, keys and bass). This song has three sections: a verse, a pre chorus/drop and a chorus. The audio effects that can be used are a bandpass filter, reverb, pitch and a beat repeat. The usage of familiar music instruments in this song, resulted in a timbre that does not entirely represent that the timbre of modern EDM music. However, when using risers, filters and other audio effects on the song it starts to adapt this timbre more.

For the mapping between the movements and music the decision has been made to make a distinction between the buttons on both hands. The button on the right device will exclusively be used to trigger new sounds or sections and the button on the left device will be used to turn audio effects on and off. The mapping of these gestures was straightforward: The gestures that mimic the instruments are mapped to triggering the samples of instruments they mimic. The gestures that indicate a change in energy are used to switch between sections of low and high energy. The quick gestures are used to turn audio effects on and off.

4.4.5 Evaluations after changes

After the changes to the prototype were made, six additional evaluation sessions have been held. This group of participants consisted of three dancers, two musicians with producing experience and a DJ. The results of these evaluations are discussed below.

Quality of prototype

The new casing for the devices solved two of the previous issues with the prototype. One of these issues was the robustness of the prototype. Before the new casing was made, the prototype would lose connection or break occasionally throughout the sessions. During this set of evaluations the prototype would stay intact. This caused the sessions to run more smoothly, without interruptions of things breaking. The second issue the casing fixed was the attachment to the hand. The gloves that were used in the first iteration were one fixed size which caused problems, given the varying hand sizes of the participants. Using the elastic bands with velcro, all participants were able to steadily attach the device to their hands, which enabled them to move their hands freely.

Adding a button that collects data when it is pressed and checks this data against a model when it is released, solved a lot of issues with the accuracy of recognizing gestures. However, given that the gestures were still trained by somebody else than the participants themselves, some participants still had issues with performing certain actions. Only this time when the wrong gestures were recognised the participants often blamed it on themselves and felt like they had to practice in order to get better at using the system. During the sessions, four participants stated that they needed more practice in order to perform certain gestures correctly. Two of them stated that at the end of the session they felt like they already had improved their skills of using the system.

Mapping of gestures

Like during the previous set of evaluations, the participants showed a strong preference towards the gestures that mimicked an instrument or action that seemed familiar. The mapping that had been made prior to the sessions included some gestures that mimicked musical instruments like the guitar, violin, drums and bass. These were also the first gestures that all participants remembered and tried first after they were introduced to the system.

Some gestures were made with the purpose of switching audio effects on and off. Since there is no intuitive set of physical representations for these actions, these gestures were created based on being able to perform these movements quickly and easily. Some participants had problems with using these effects. It was not clear to them whether certain effects were either on or off or which motion referred to which audio effect. This caused a bit of frustration in those participants. It was interesting to see that the participants with a background in music making had less trouble with these audio effects when compared to the dancers. Some of the dancers indicated that they had trouble distinguishing the audio effects from each other, which caused them to be confused when trying to turn them off.

The usage of audio effects introduced a new criteria for choosing gestures for a specific action. This criteria is the location where the hand ends up after the effect is turned on. Some mappings forced the user to finish the gesture with their hand in a specific position, which had consequences on the sound when the effect was triggered. Two participants brought this up, but were not particularly bothered by it.

Usability

Adding the button to the prototype seemed to have a positive effect on the feeling of triggering sounds. The problem of not knowing when the action is triggered, stopped occurring all together during these sessions. While there was still some unexpected behaviour at some points, the participants often blamed themselves instead of the prototype.

It was fairly easy for the participants to learn how to trigger something by using the button. Because the music changed exactly at the point that the button was released, it was easy for them to learn how to time the release of the button correctly. However, it took a bit longer for the participants to learn how long to keep the button pressed in order to collect sufficient data for the gesture to be recognised. Initially, the participants sometimes held the button for a very short time which resulted in very little data being collected and therefore in the model not being able to accurately predict the right gesture. After explaining what had happened, the participants were able to adjust to this and were able to more accurately trigger the sounds.

The thing that confused the participants the most was the large amount of gestures and mappings that needed to be remembered in order to use the system correctly. Two participants brought up that they would prefer having a visual representation of the gestures and how they are mapped. Every participant had to ask the researcher, after the initial explanation, about how to perform certain gestures again. However, within the time span of the session, most participants were able to remember almost all gestures and were able to improvise the arrangement of the song in some form.

Given that during this set of evaluation sessions the song and mapping were already created beforehand, the participants had the opportunity to perform an arrangement of the song within the session. This brought up two issues concerning the performability of the system.

The first issue that some participants brought up was the fact that it felt awkward during periods where they did not control anything. At some points after triggering a new sound there was a moment where the participants had to wait until they wanted to perform a new action. When asked about how he would use those moments in a performance, one participant stated that a performer should use those moments to engage with the crowd. However, he also stated that he himself would not feel comfortable doing this. These uneventful moments most likely played a part in the fact that five of the participants brought up that they would prefer to use this system additionally to something else in a performance. They for example wanted to use the system in addition to: vocals, guitar, dance or synthesizers.

The second issue that came to light during these sessions was the fact that it is easy to make mistakes when using the system. As explained earlier, the machine learning model is sensitive to the precision of the movement. This caused troubles for the participants to perform the gestures accurately, given that they were trained by somebody else. This sensitivity caused some frustration within the participants. Several indicated that they would not like to use the system in a performance when the chance for mistakes is this high. One participant stated: "I do see the use, but as of now I would still rather use a launchpad. The system needs to be more advanced to the point where it does what you want it to do at all times, before I would use it in a performance. Of course the occurrence of happy accidents is nice, but I would rather not have this happen on stage".

4.5 Discussion

In total ten evaluation sessions have been held. These evaluation sessions have provided sufficient information to reach the goals that were needed to answer RQ 4. These three goals were to gain insights into the quality of the prototype, into the best mapping between movement and sounds and into the usability of the system. This section discusses the results from each of these.

4.5.1 Quality of prototype

The first goal was to gain insight into the quality of the prototype and to gain information on ways to improve this quality. Throughout the sessions, the quality of the prototype has significantly improved. These improvements occurred in two areas.

One of the areas that improved was the robustness of the prototype. Initially this was an issue because the wires of the prototype kept loosening or breaking. This is because during the first view sessions, the electronics were hot glued on top of a pair of gloves, which worked sufficiently when tested by the researcher. However, given the varying hand sizes of participants, the gloves were very loose on the hands of some of them, which caused the electronics to move around more than expected, causing some of the wires to break. This problem was fixed through creating a 3D printed case that precisely fits the electronics inside. These cases were attached using an elastic band with velcro, which enabled the cases to stay fixed on hands of varying sizes.

Another thing that improved throughout the sessions was the accuracy of performing gestures. The problems with this accuracy were mainly caused by two things. The first problem was the sensitivity of the model. To be able to perform a gesture correctly, the user needs to be very close with the movement to the original training data. This caused issues for the participants when the gestures were trained by someone else. However, given the long amount of time it takes to train new gestures, the decision was made to stick to the pretrained gestures during the second set of evaluations. The other problem that influenced the accuracy of the performance of gestures was the transition between them. This caused a problem because the system was continuously checking which gestures were applied and was reacting to this. The transitions between different gestures were not included in the model and could therefore not be recognised. This caused other gestures to be recognised instead and would sometimes result in unwanted behaviour. This problem was solved by adding a button that can be used to indicate when a gesture is recorded, which allows the users to have control over the exact moment they want to collect data.

Overall the participants were able to use the system well enough to perform the actions they wanted to do, while getting the results they wanted. However, from the interviews it became clear that it is crucial that if a music controller is to be used in live performances that it should not output anything unexpected. A way to improve on this would be to try to make the model less sensitive to specific movements. This might be achieved through gathering more data for each gesture. Having more data for each gesture might increase the range of motions that are classified as a certain gesture.

4.5.2 Mapping of gestures

The second goal of this evaluation session was to gain insights into the preferences for mappings between movements and sounds. Two themes arose from analysing the arguments given for motivating the usage of a specific gesture.

One of these themes was the mimicking of familiar actions. These actions could be to mimic playing a musical instrument or to turn a knob on a DJ booth. One of the reasons why the participants gravitated towards mimicking familiar actions could be that it is easier to remember these. It became apparent that remembering all the gestures was quite hard for some of the participants and that having gestures that mimic familiar actions helped with remembering what they did. Another reason could be that these actions just feel more logical, or enhance the expressivity of the movement.

The second theme that was used to motivate the mappings was the correlation between the position of the arms and the energy level it conveyed. This reason was mostly used for gestures that were used to switch between sections of different intensities, or for affecting parameters used to increase the energy in a buildup. This action could even be seen as mimicking familiar actions, since raising the hands during buildups is a common thing to do at concerts.

While these themes were commonly used between the participants, not every participant preferred using the same gestures. For example, for changing certain audio effects there was some disparity between using the x axis (moving the arm up and down) or the y-axis (twisting the wrist). Some preferred using the x axis because it indicated an increase in energy and some preferred the y axis because it felt like controlling a knob.

4.5.3 Usability

The third goal of the evaluation was to gain some insights into the usability of the system and into ways of improving this. The initial prototype had two big usability problems which were fixed using the addition of a button to both gloves.

The first usability problem was the uncertainty of when certain gestures were recognised. This uncertainty came from a delay between the action of performing a gesture and hearing the sound as a response. This delay was caused by a running average that was applied on the recognised gestures to filter out some of the wrong gestures that occurred during the transitions between different gestures.

The other problem with the usability was that in order to trigger things, combinations between different gestures had to be used. This had to be done in order to prevent unwanted sounds being triggered whenever the user would switch between gestures, since using a combination of three different gestures reduced the chance of accidentally triggering something. Participants indicated that they thought using multiple gestures at the same time was very unintuitive. Fortunately, when the button was added to the gloves this problem disappeared. Since this change made it possible to accurately recognise the gestures, it was not necessary anymore to use unintuitive combinations of gestures. Instead simple and quick gestures could be used to trigger things.

Chapter 5

Evaluation 2: Audience perspective

The previous tests were held with the goal of evaluating intuitivity by testing the prototype with producers and dancers. This section will focus on the perception of an audience when this instrument will be used. With the purpose of answering RQ 5: "How can the proposed system be made more expressive?".

5.1 Goals

The purpose of this evaluation is to get insight into how an audience would perceive a performance given with the developed music controller. This can be assessed through reaching a set of three goals.

The first goal is to get information into how well an audience understands what the performer is doing. This goal is important given that one of the main reasons behind developing this controller was to create more transparency between what the artist is doing and what the audience can observe. The second goal is to test whether the mappings that came out of the first evaluation also seem logical for the audience. The final goal is to assess whether an audience would enjoy watching a performance using the system.

5.2 Method

To be able to gather the necessary information that is needed to reach these goals, a survey is created. This section will go over the design of this survey and how it will be analysed.

5.2.1 Survey design

The survey will be split into three parts, where each part is designed to reach one of the earlier described goals. An additional section is added to gain some general information about the participant.

Part 0: General info

The first page of the survey has two purposes: Introduce the participant to the research and retrieve some general information about the participant.

The participant will be introduced to the survey through a short description of what will happen in the survey and with the contact information of the researcher, supervisor and ethics committee in case they have any questions. Additionally, participants will be asked if they agree with data being used for academic purposes.

Afterwards, a series of general questions about the participants will be asked. This part will consist of questions about the age, musical background, musical preferences and experience with live concerts of the participants. The purpose of these questions is to see if there's a difference in the perception of the prototype between people from these different groups. It would for example be interesting to see whether there is a difference in opinions between people who attend a lot of EDM concerts and people who never go to concerts.

Part 1: Understanding

The first part of the survey will focus on the level of understanding of the participants when they see a performance of someone using the system. This will be done by creating two video's of short performances and asking the participants to describe all the actions they think the artist has performed and how they think these actions affected the music.

The two videos that will be used in the survey will each be about a minute long. While both videos will include elements of both, one video will focus on adjusting the parameters of audio effects and the other video on adding instruments and switching between sections. This is done to see how well each of these types of actions are recognised. An example of a movement in the first video that is used to control the parameters of a buildup action is shown in figure 5.1. An example of different gestures that are used to trigger the loops of instruments is given in figure 5.2.

Afterwards, a question on how much the audience felt that there was a relationship between the movements of the artist and the music will be asked. It will be interesting to see if there is a difference between how much of the video the participants exactly understood and how much they think that there is a relationship between movement and sound. The input of the question will be through a likert scale, with the possibility to motivate the answer afterwards.

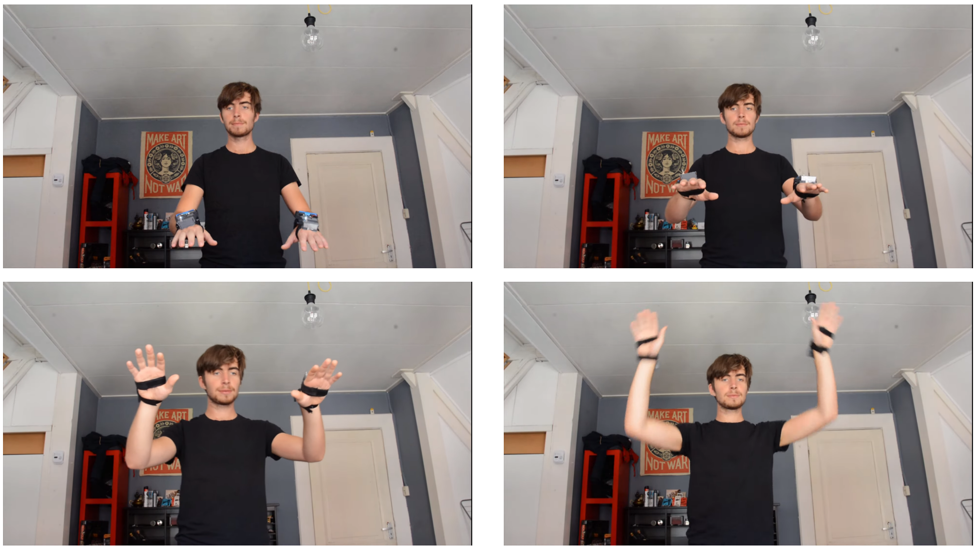

Figure 5.1: Buildup action in video 1

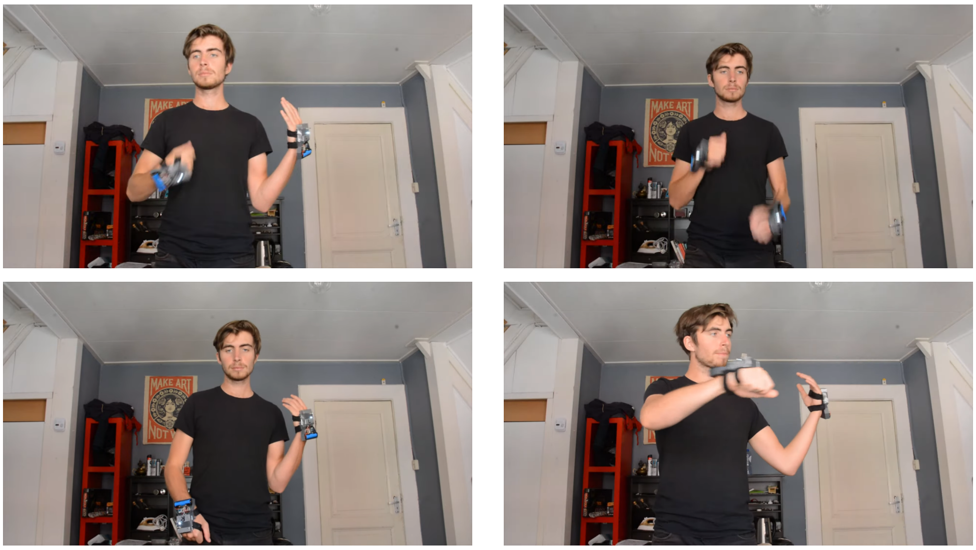

Figure 5.2: Gestures for controlling: Guitar, drums, bass and violin in video 2

Part 2: Mapping

The second part of the survey will focus on the different mappings between gestures and sound. This will be done through two multiple choice questions where several options for gestures are shown. The participants will be asked to choose which gesture they think corresponds best to the affected sound. The goal of these questions is to see if there is a correlation between the preferred mappings that were chosen by the artists during the previous evaluation sessions and the preferred mappings of an audience.

The first question will focus on different ways of applying a low pass filter onto the sound. The four examples include: moving the right arm up and down, twisting the arm left and right, moving the arm up and down but with the palm upwards and bending the body. During the previous evaluation session, multiple reasonings behind different preferences for low pass filter mappings were given. Common reasons were that twisting the arm feels like controlling a knob and that moving the arm up and down makes sense because of the difference in height of the arm. This question serves to see if one of these reasonings is more prevalent within the audience.

The second question focuses on triggering an audio clip of a synthesizer. One of the gestures will be pretending to play the synthesizer, the other gestures are more abstract and quick. Given that during the previous evaluation session the preference of the participants was to imitate playing a familiar musical instrument, the expectation is that this will be the preference of the audience as well.

After each question the participants will be given the opportunity to motivate their answers. The arguments used for choosing a particular parameter can be used for future mappings.

Part 3: Enjoyability

The final part of the survey will evaluate the enjoyability of looking at a performance using the music controller. This will be done by asking a series of questions that cover certain aspects of the music controller. Using multiple questions that are similar but hint at slightly different qualities of the music controller, it becomes possible to find out which qualities are better received than others. Each question can be answered with a likert scale with a possibility to motivate the answer afterwards. The questions that will be asked are the following:

- Do you see the use in controlling music using movements?
- Do you think this music controller would contribute something to a performance?
- Do you think it is enjoyable to look at a performance given with this music controller?

- Would you like to attend a concert given with a music controller like this?
- Do you consider a performance given with the music controller as live music?

5.2.2 Analysing

This section will go over how the data will be analysed.

Part 0: General info

The answers to the general questions can be used to figure out whether there are differences in opinions and understanding between different groups of people. Interesting factors to distinguish the participants would be: age, whether they prefer to listen to electronic (dance) music, musical expertise and live concert attendance. The variables that are interesting to measure these groups against are the sum of assigned points from the questions about the understanding (section 5.2.2.2), whether they saw a relationship between movements and music and the enjoyability questions.

Part 1: Understanding

The questions where the participants are asked to describe what the performer is doing in the video had a text field as input. This requires the researcher to go over each answer and analyse them manually. The answers of the participants will be compared to what is actually happening in the video. This will be done by creating a small list of key actions that the performer is making in a video and assigning a point to the participant for each type of action that is being described. In total three main actions will be used to assess the understanding of the participant: adjusting the parameters of audio effects, triggering an instrument and switching between song sections. The sum of points for each video will be used as a metric for understanding.

The question on whether the participants saw a relationship between the movements can be compared with the sum of points, to be able to see if there is a relationship between the two. This can be done through calculating a Pearson's correlation coefficient. It is also interesting to look at the reasoning given. These open answers can be coded and studied in order to find themes in the answers of the participants.

Part 2: Mapping

Given that the questions concerning the mapping were multiple choice, these questions are fairly easy to analyse. By looking at the occurrence of each answer it becomes possible to see which type of mapping is preferred.

Both questions also provide the option of motivating the answer. These answers can be coded and analysed to be able to find themes within the answers. The themes that are retrieved from these questions can be used in the future to figure out new mappings.

Part 3: Enjoyability

The questions concerning the enjoyability are all likert scale questions. Given that the participants might not want to hurt the feelings of the researcher through giving bad ratings, it is a good idea to mostly look at the relationships between the answers instead of the individual answers.

The open questions where the participants get the possibility to motivate their answers can again be coded and analysed for themes. These themes can be used to figure out what exactly is successful about the system, what is wrong with the system and what can be improved.

5.3 Results

In total 70 responses to the survey have been recorded. In this section the results of this survey will be presented. This will be done by first displaying the descriptive statistics and afterwards the results from a thematic analysis.

5.3.1 Descriptive statistics

The mean and standard deviation of all the likert scale questions are shown in Table 5.1.

The questions where participants were asked to describe all the actions they thought the performer made were coded through assigning points to each type of action they mentioned. For each correct action, one point was assigned. The mean and the standard deviation of the points that were assigned to the descriptions of the first video are shown Table 5.2. The mean and standard deviation for the assigned points of the second video are shown in Table 5.3.

The distribution of preferred gestures for changing the filter is shown in Figure 5.3. The distribution of preferred gestures for triggering the sound of a synthesizer is shown in Figure 5.4.

Question	Mean	Std. deviation
Did you think there was a clear relationship between my movements and the music	4.09	1.069
Do you see the use in controlling music using movements?	4.14	1.044
Do you think this music controller would contribute something to a performance?	4.48	0.728
Do you think it is enjoyable to look at a performance given with this music controller?	4.42	0.860
Would you like to attend a concert given with a music controller like this?	4.24	0.836
Do you consider a performance given with the music controller as live music?	4.19	0.680

Table 5.1: Descriptive statistics of likert scale questions

Action	Mean	Std. Deviation
Triggering instruments	0.42	0.498
Applying audio effects on instruments	0.91	0.294
Switching sections	0.23	0.427
Sum of points	1.54	0.849

Table 5.2: Descriptive statistics of assigned points for understanding of video 1. Here the individual sections have a range from 0 to 1 and the sum of points a range from 0 to 3

Action	Mean	Std. Deviation
Triggering instruments	0.95	0.215
Applying audio effects on instruments	0.76	0.429
Switching sections	0.21	0.408
Sum of points	1.86	0.704

Table 5.3: Descriptive statistics of assigned points for understanding of video 2. Here the individual sections have a range from 0 to 1 and the sum of points a range from 0 to 3

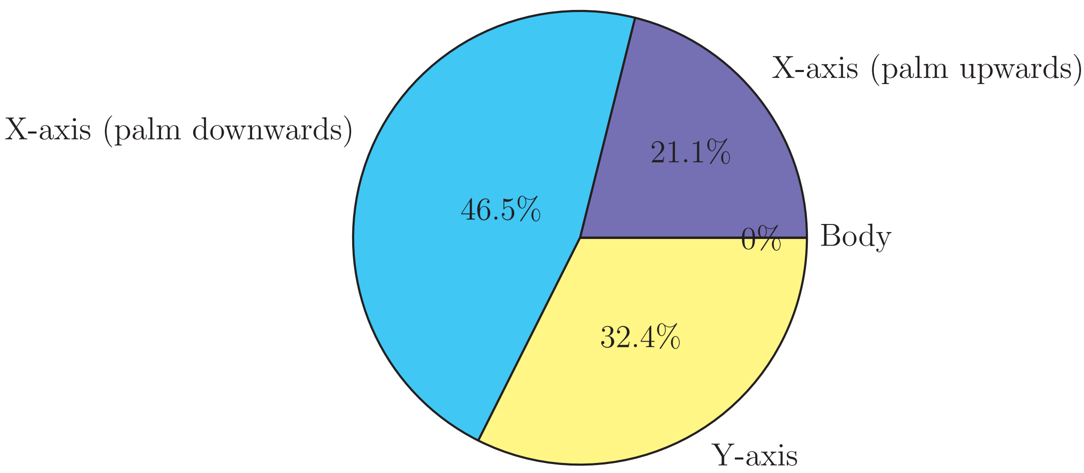

Figure 5.3: Preferred way of adjusting filter

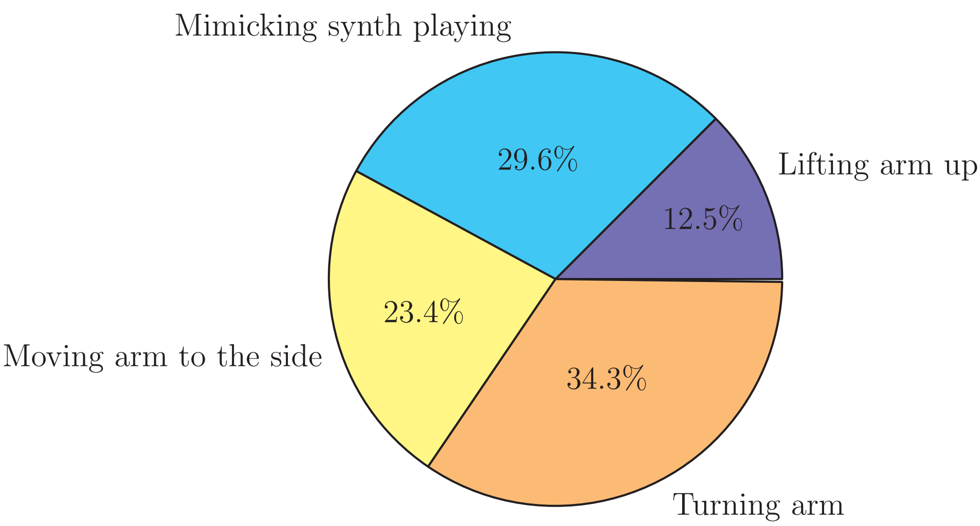

Figure 5.4: Preferred way of triggering synthesizer sound

5.3.2 Thematic analysis

This section will discuss the results of a thematic analysis that was applied on the answers from the open questions of the survey. This will be done by going over each of the open questions that were asked in the survey and by providing the most common type of arguments for each of them.

Do you think there was a relationship between my movements and the music?

The question asking if the participants thought there was a relationship between the movements and the music, got an average rating of a 4.1. After coding the motivations the participants gave for this question, it became clear that there were three types of arguments that occurred often.

It was interesting to see that for the participants there was a difference between being able to fully understand what the performer is doing and thinking that there is a relationship between the actions and what happens in the music. One participant said: “It was not very clear what actions exactly led to certain effects but it definitely felt like your actions influenced the music”. While another participant noted: “I think it’s very clear that the gestures you perform are linked to changes in the music. Although I was also able to recognize what actually happens to some extent, I have the feeling that there’s more intricacies going on than I’ve been able to identify.”

Another argument that was often given to motivate the rating, was that using the actions that mimicked familiar instruments were making it especially clear what was going on. In total this argument was given 12 times. An example of such an argument is: “The second one was very easy to understand what kind of instrument you want to hear with the movement. The first one was harder to understand for me since I never see electric things being produced. but the volume and damping seemed intuitive”

Finally, a few participants were bothered by the lag that sometimes occurred between performing a gesture and hearing the corresponding instrument. While this was not a lag caused by the devices, but by the music software to make sure that the sound would be started in time with the rest of the music, some participants blamed this on a faulty system and were thrown off by this.

Which action seems most logical to control a filter?

The question of which action seemed most logical to control a filter had a divided result. The most preferred option was controlling the x-axis with the palm of the hand pointing downwards (45,7 percent). This option was preferred because of three types of arguments. The first argument was that this gesture looks like controlling the fader of a mixing board. Faders are usually being used to control the volume of a track, however these effects can sound pretty similar. An example of a motivation is: ”You’re controlling volume (or something that feels similar), and although varying the intensity of any filter is (intuitively) done with a fader-like movement (options 1, 2) and with a knob-like movement (option 3), a fader-like movement intuitively describes volume the best to me." Another argument was that "The up-down movement of the hand with the palm down just feels natural for indicating a stronger or weaker effect." Finally, the most common argument for this option was that the movement looks like you are pushing the higher frequencies down and lifting them back up again. An example of such an argument was "Specifically for this filter: it looks like you suppress the sound by pushing down on it". Some participants described this in a more tangible way: "It looks like taking the lid off the pan and letting the music out again. Putting it back on the pan creates that muffled sound."

The second most preferred option was to use the y-axis while turning the wrist (32.9 percent). The main motivation for this option was that the motion looked similar to controlling a knob. This motivation was given by 14 people. Another motivation for this option was that it seemed like it cost the least amount of effort. An example of an argument for this option was: "Looks more like tuning a dial, and has the least effort. But 1 and 2 look cooler".

Finally the third most preferred option was to control the filter with the x axis with the palm upwards (21.4 precent). Most arguments for this option were similar to using the x axis with the palm downwards. The main argument that was made for this option was that the motion looked like a conductor of an orchestra controlling the volume.

Controlling the filter while bending the body was preferred by nobody. The most common reason for this was that it looked uncomfortable and excessive for the artist to perform. This reason was derived from comments that were insight into the arguments for other options in which they mentioned this. Examples of such comments were: "Option 4 is really not that enjoyable for the artist I think haha..." or "My first answer was 4, but that was just to torture musicians. If I'm being real it's 2."

Which action seems most logical to start a synthesizer

The question of which gesture to use for starting a synthesizer loop had mixed answers. It was interesting to see that the gesture that was mimicking the playing of a synthesizer was not the one that was most preferred by most of the participants, but only by 27.9 percent of the participants. While performing this gesture was by most participants seen as most logical, they did not always prefer it. This resulted in reasoning like the following: "I guess 2 seems most logical but 3 and 4 definitely look cooler". In total 4 participants motivated their answer by explaining that it simply looks the coolest.

The most preferred option was the motion of swinging the arm and opening the hand in front of the performer (30.9 percent). While the participants didn't prefer the mimicking of a familiar musical instrument, they ended up finding familiarity in this gesture as well. This particular gesture reminded a lot of participants of dropping something into the music. In total five people mentioned this dropping effect, an example of such a comment is: "It was a short movement that correlates with dropping something in, like this sound."

Another big motivator for participants to prefer a certain gesture was the simplicity of the gesture. This indicates again that one of the motivations for the right gesture has to do with the comfortability of the performer. An exam[ple of an argument for this one was: "Opening gesture makes sense for me to use! All the other ones seem too complicated / unnatural. The first option might also be mixed up with some other commands."

Do you see the use in controlling music using movements?

The average answer to the question of whether the participants saw the use in controlling music using their movements was a 4.2 out of 5. The reasons given for this question varied greatly.

One of the motivations behind giving this score was that using movement to control music is more fun to watch than traditional EDM concerts, because the controller is something new. A few participants describe using movement as "engaging" or as "A way to get a crowd fired up" or "a way of adding more immersion, interactiveness and in case of a live event this makes for a killer show". However some participants saw using movement for the purpose of controlling music more as a gimmick as opposed to something that could stay interesting for a longer period of time. Examples of answers for this were: "It could be fun. And it might inspire some novel music, but it feels like something to use once or twice and then to be set aside." or "It feels a bit like a gimmick, but it's a cool gimmick."

Another theme that occurred multiple times in the motivations the participants gave was that using movement to control music seems like a very simple and intuitive way to control it. This inspired a lot of participants to think of alternative ways of using similar systems examples of this were: "Looks like a fun interaction for people who can't make music but still want to" or "It is a new creative way that perhaps appeals to a different target group. It could be a solution for people with a disability who cannot read notes or are blind, for example. Maybe people who are hard of hearing can read certain beats this way?" or "Sure! It could be used in live performances, get artists more in touch with their music (which I can imagine can be quite challenging when you're creating music electronically), it might be useful in music therapy for people who aren't strong on an instrument or prefer EDM."

Do you think this music controller would contribute something to a performance?

The question asking whether the participants think that the controller would contribute something to a performance, was answered with an average of 4.4 out of 5. The main argument given for this, is that a performance using this music controller would look more engaging. This is because of two reasons.

The first reason is the fact that using the controller looks like a dance performance, which in itself is associated with something that is entertaining to look at. An example of such a reasoning is: "I think it would add dancing to the performance which is fun and artistic." Another participant stated "I think it depends on the movement. If the controller can be combined with dancing moves it can add a lot to a performance. For a performance on stage, the bigger the movements the better, while in a studio smaller movements are more practical."

The second reason for making the performance look more engaging is that it becomes clearer to the audience what the performer is doing. Or at least provide the opportunity for the audience to try to figure out what is going on. One participant said: "It's fun to see how correlating dance moves of the performer would affect their music, or how they would need to have a choreography for playing music." Another participant explained: "We all know the 'DJ pressing play' trope. It can be hard for an electronic music performer to convey their 'art' like a more traditional band can. I think the music controller you've demoed can be a very cool, but also "intuitive" (for the public) way for the public to get an idea of what is actually happening behind the mixpanel."

Also for this question, some participants were inspired to provide new ideas to make the performance even more engaging. Three participants mentioned that they would like to see a performance using multiple people where each person controls a single element of the music. An example of such an idea was: "I could see this work if a group of like 8 artists would each be equipped with these sensors. They would each be controlling just one or two instruments or effects. It could be a nice experience for the audience to slowly figure out whose movements control which sound."

Do you think it is enjoyable to look at a performance given with this music controller?

The question asking whether the participants would enjoy watching a performance given with the music controller, had an average answer of a 4.5 out of 5. This score could be explained through the fact that the main motivation for the previous question was that it would be engaging to look at a concert given with the music controller. Given that a lot of participants already motivated this opinion in the previous question, a lot of answers for this question are along the lines of: "see previous answer".

However, participants added some requirements that would have to be fulfilled to make the concert more enjoyable. One of these requirements is that the performer needs to use the controller in an entertaining way. An example of an answer for this argument is: "I think it goes both ways. If a person plays with it really deadpan and uninspired, then it will obviously not be fun. But if a person not only plays good music, but builds a show around the movements (e.g. dancing), it would be double entertainment." Another requirement for the enjoyability is that the gestures should be chosen wisely: "Yes, but only with the raising/lowering hands motion. Pretending to hold a violin to trigger a violin track feels kinda silly. I think it would be better to embrace its digital nature and not try to imitate real-world equivalents." This answer is particularly interesting since this participant gives an argument for not using gestures that mimic familiar musical instruments.

Another participant stated that since the actions of the performer are more visible, it would raise more questions on what exactly those actions do: "It's more visible, that's nice. However, it raises more questions."

Would you like to attend a concert given with a music controller like this?

The average answer to the question on whether the participant would like to attend a concert given with a music controller like the one shown in this survey, is a 4.2. For this question, it turned out that music taste has an effect on the given score. Examples of answers where this becomes clear are: "I doubt my kind of music would come from this; the music controller lends better to electronic music" or "Yes, no, I don't know. Given that I rarely attend concerts. Music for me is mostly something fun for the background. The most fun I get out of concerts comes from the people I go with."

However, a lot of participants seemed excited about the idea of attending a concert given with the music controller. Big reasons for this are the innovative aspect of the music controller and the fact that it becomes clearer to the audience what the performer is doing: "Would be something new and different! Plus, if I knew exactly what movement would control what effect and everything, you could more easily see how the song is created - would be interesting." Yet, some people were worried that if it were to be a concert only about an innovative instrument, that it would diminish the quality of a show: "Depends on the music, the ambiance, the price, the location. I think they do add a novelty which is interesting. Also, if the concert is about novel instruments, then these are a nice addition. But it should fit and not take away from the show."

Do you consider a performance given with the music controller as live music?

An average rating of 4.2 was given to the question that was asking the participants if they would consider a concert given with the music controller as live music. This question evoked some interesting arguments about the definition of live music.

One participant argued that if the audience can tell that the performer is controlling something, the music is considered live: "If it's clear enough that the actions are controlling the music live, then yes, but if it looks like it's pre-recorded, then no". While some participants believed that the liveness depends on the amount of control that the performer has over the music: "Yes and no. Would depend on the amount of control the music controller provides. Just loop triggering or more than that? But I would still consider it a live performance due to the overall spectacle and energy". This control would enable the performer to adapt the music based on a crowd which for some participants was a criteria: "While the base of the music isn't made live (as with live sets by DJs etc), the manipulation is live and it will always be a little different with each performance. It's possible to react to the crowd live as well." Another participant argued that if the performer is able to make mistakes, the performance can be considered live music: "If you mess up a movement, you mess up the song. So yeah that's live music."

The liveness of a concert given with the music controller was often compared to the liveness of traditional electronic music concerts. Examples of these comparisons are: "I'd consider it the same as a DJ set. The music isn't really created live as it is with an actual band, but it's still mixed live and everything." or "I do feel like it is a different discipline than playing a live instrument, it gives freedom to play with the composition of the song instead of the melody. But it's still very much an artform." Given that these types of concerts are not considered as live music by everyone, this also resulted in people who did not consider this type of performance as live music: "Electronic music for me is different than live music"

5.4 Discussion

This evaluation was held with the purpose of gaining information about the audience's perspective on the music controller and thereby gaining insight into the expressiveness of the system. The retrieval of the information about the audience's perspective was done through a survey. This survey was split into three parts: understanding, mapping of gestures and enjoyability.

5.4.1 Understanding

The first part of the survey was about understanding the actions of a performer that uses the music controller. This was measured by showing the participants a short video of somebody giving a performance using the music controller and by asking the participants to write in detail what actions they think the performer made. The results seem to indicate that not every action is noticed or understood by an audience, but that it is possible to make the audience understand the things that are desired to communicate. During the first video, most of the emphasis was on the adjustment of audio parameters. 91 percent of the participants figured out that audio parameters were adjusted. While only 42 percent of the participants noticed the triggering of new instruments and 23 percent noticed the switching between sections. The second video had more emphasis on the triggering of instruments. In this video 95 percent of the participants noticed that new instruments were added through gestures. 76 percent of the participants noticed audio effects were adjusted and 21 percent noticed the switching between sections. One possible explanation for the disparity between the main action of the video and the other actions is that not all the participants wanted to spend the effort of writing all the actions down, but instead focused on writing down the main action. It was noticeable from the responses that the length of answers varied greatly, where some participants wrote down a whole paragraph and others a single sentence.

The participants gave an average rating of 4.2 out of 5 for the question on whether they saw a relationship between the actions of the performer and the music. It was interesting that quite some participants motivated this answer by stating that they did not understand all the actions but still were very confident that the movement the performer was making influenced the music in some way. This raises the question on how much understanding of the actions of the performer is enough. Should the audience be able to decipher every action or is it enough to know that the performer is just influencing something?

5.4.2 Mapping of gestures

The second part of the survey was about exploring the mappings between movements and sounds. This was done by showing four examples of different movements that could be used to control the same audio effect/trigger and by letting the participants choose which movement looks most logical. The types of mappings were based on two observations from the first evaluation sessions. During these sessions it became clear that most mappings were chosen through analogies of familiar actions. From the responses of the participants it became clear that translating movements into an action that seems familiar is done by an audience as well. For the choice of controlling the filter, the most preferred option was moving the arm up and down with the palm facing downwards. While this action did not mimic the familiar gesture of turning a knob, participants motivated the use of this action through creating a set of other analogies like the resemblance of a fader or putting the lid on the sound and removing it.

Also for the gesture of starting the sample of a synthesizer, the audience did not prefer the obvious mimicking of playing a synthesizer, but instead preferred a slightly more abstract movement, for which the participants were able to think of other analogies like "dropping the instrument into the music". This seems to indicate the most obvious representation of a familiar action is not always preferred, but that the audience either prefers more abstract gestures or that there is an intrinsic "coolness" factor within certain gestures that makes audience members gravitate towards it.

5.4.3 Enjoyability

The third part of the survey was about finding out how enjoyable it is to watch a performance given with the music controller. This was done through five slightly different questions about the perception of the controller which could be answered through a likert scale. The question on whether the participants think it is enjoyable to watch a concert given with the music controller was rated the highest, with a rating of 4.5 out of 5. While the rating for the question on whether the participants would like to attend a concert given with the controller, had a rating of 4.1 out of 5. These ratings are quite high and therefore seem to indicate that there is potential in the live use of the music controller. However, it is important to consider that a number of participants were acquainted with the researcher and therefore might not want to be too harsh about their opinions. Therefore, the reasoning behind the answers was also analysed.

After analysing the reasoning given for each question, three insights into the viewpoint of the participants was found. One thing that became clear was that a lot of participants were excited about the novel aspect of the controller. Some participants described the use of the controller in concerts as more engaging, entertaining and interactive. However, some participants were worried that the controller would quickly turn into a gimmick and that from a certain point on the controller won't be entertaining anymore. The use of gestures might play an important role in avoiding this to happen. One participant indicated that by using gestures that mimic the playing of familiar musical instruments, the digital nature of the music instrument is not embraced enough. This could mean that pretending to play other musical instruments might be perceived as a quirky gimmick, instead of as a music controller that has potential to be used to give concerts with.

It also became clear that the controller often could not change the opinion of participants who already had a negative perception of electronic music. Which was the reason that quite some participants answered negatively to whether they would like to attend a concert given the music controller or whether they considered it live music.

Finally, the responses included a lot of ideas on how to implement this controller in live shows. One idea that occurred often was the concept of using the controllers with multiple people, where each person controls one aspect of the music. Other ideas were to merge the usage into an actual dance performance. The fact that this many people feel the need to contribute their ideas about the application of the controller could either mean that the current way the controllers are applied are not satisfactory enough, or that the participants are just excited about the possibilities the controllers have to offer.

Chapter 6

Final discussion

In this section a discussion of all the research that was done in this thesis will be presented. First changes to the original concept will be discussed. Afterwards, some general requirements that need to be satisfied to make the controller truly intuitive and expressive are established. Later, the limitations of the research that has been carried out will be given. Finally, a list of some of the research that can still be done on the topic of embodied music controllers will be presented.

6.1 Concept changes

At the beginning of the thesis a concept for the music controller was established. Now that the evaluations are performed, it turns out that the controller might work better if two changes to this concept are made.

Firstly, the original concept stated that IMU sensors would be the only input used to control the music. With the purpose of forcing the performer to only make big movements to control the music, as these are easier to pick up by an audience. The addition of the button in the music controllers, added a different input to the devices. However, this button can only be used in combination with the IMU sensor and the opening and closing of the hand is still a gesture that might be visible from far away.

Secondly, the original concept included the use of a third device that is placed on the body. However, after both evaluation sessions it became clear that neither the performers or the audience saw a lot of use in this extra device. The device was rarely used by the performers because of the limited amount of actions were possible to do with them. The main change the sensor was able to pick up from the body, was bending and stretching. However, most performers rather wanted to use other actions that were less bothersome, like moving the arm up and down. The evaluation with audience members illustrated that the use of this body sensor is also not desired by audience members. Therefore, for further use it might be useful to obliterate the use of this device and focus on the two devices in the hands.

These two changes, bring the concept of this music controller closer to controllers like the Mi.Mu gloves. However, the simplicity of only using the device for triggering sound and adjusting parameters, together with the big movements that are needed to do this, provides a unique and possibly more transparent performance as opposed to other embodied music controllers.

6.2 Requirements

The purpose of this thesis was to develop an embodied music controller that could be used to intuitively perform Electronic Dance Music in such a way that the audience is able to see a relationship between the music and the actions of the performer. Two evaluations have been conducted with the purpose of gaining insight into the intuitiveness and expressiveness of the developed music controller. Both evaluations seem to indicate that this type of music controller has the potential to be both intuitive and expressive. There are a two main requirements that need to be fulfilled in order for this to happen.

First, the gestures that are used to control the music play a crucial role in both the intuitiveness and expressiveness of the controller. From the tests with professionals it became clear that using gestures that mimic familiar actions and instruments helps them with remembering and performing the gestures. From the test on the audience's perspective a similar response to these familiar gestures was found. However, it also illuminated that there is a degree to the familiarity of these gestures. If the gestures are too much mimicking obvious gestures and instruments, there is a chance that the performance turns into a gimmick and not into a useful alternative way of performing electronic music.

The second requirement that needs to be satisfied, is that the music controller should work flawlessly before it can be used during any live concert. The participants expressed that they would not like to use the controller in any performance if they are not fully sure that the music controller will behave like they want to. The prototype that was developed for this thesis fulfilled most of the set requirements, however most participants still did not feel confident to use it for their performance. Currently, the main issue with the prototype is the sensitivity of the recognised gestures, which created unwanted results when the user performed a gesture slightly different from the way the gesture was trained. From the second evaluation it also became apparent that participants were very sensitive to lags and faults within the system. Multiple people commented on the lag between the performance of a gesture and the moment the sound was triggered, which was caused through the software making sure that loops are triggered in time with the rest of the music.

6.3 Limitations

This section will discuss some of the limitations of the research that has been performed. In total there were three main limitations that had an impact on the results of this study.

Firstly, one of the limitations was the type of participants that took part in the first evaluation. This evaluation was meant to be held with professional music producers and performers. However, due to the implications of COVID-19, the length of the sessions and the limited number of music producers in the network of the researcher, it was not possible to find professional music producers for this evaluation. Instead, the research has been conducted with a combination of amateur producers and DJ's. While these sessions provided valuable information about the developed music controller, it could be the case that insights from testing with music producers with a ton of performing experience could have been different.

The second limitation of the study was the limitation of time during each co-design session. Each session took about two hours, during which the gestures should have been trained, mapped to music and tested. During the evaluations it turned out that this process takes quite a bit longer than that. This resulted in a limited amount of time that was left for evaluating the trained gestures and mappings. Given that the testing of the controller was the most important part of the evaluation, the decision was made to skip the training and mapping during the last set of tests and instead use a pre-configured set. This decision removed the freedom for the participants to map gestures to their own music, which provided less information about the desired mappings.

The final limitation of this study was the fact that the expressiveness of the music controller was measured through videos in a survey instead of through a live concert. While these videos demonstrate the initial concept, it could be the case that the perception of the music controller would have changed if the participants saw it in a live show. Due to the practicality of gathering information in an online survey and again the implication COVID-19 had on live concerts, the choice was made to do the data gathering in a survey.

6.4 Future work

While this thesis has presented some interesting insights about using embodied music controllers to perform the arrangement of EDM music, there is still a significant amount of work that can be done on the topic.

One of the things that will be interesting to research, is to test how the proposed system will be perceived during an actual live performance. The way an audience of a live concert behaves will be very different from the way they behave when you ask them to fill in an online survey. Besides, a performance with the music controllers will be a very good way to test the robustness of the devices.

Another thing that would be intriguing to research is to have music producers use the music controllers for a longer period of time and give them the assignment to create a performance using them. Allowing the music producers to tinker with the music controllers for a longer period of time, will most likely give some great insights into the gestures and mappings they ended up creating, as opposed to the research presented in this thesis where a mapping had to be created within two hours.

Finally, it could be interesting to test different applications of using the system. As some of the participants of the survey suggested, it could be interesting to see how these controllers could be used in group settings. Alternatively, it could be worthwhile exploring how to incorporate the music controllers into a dance performance. Finally, given that participants of the first evaluation mentioned this, it could be interesting to combine these controllers with playing an instrument or singing.

Chapter 7

Conclusion

This thesis has explored an alternative way of performing EDM music. The aim of this study was to answer the question: “How do you design an embodied music controller that can be used to improvise the arrangement of EDM in such a way that its expressive power is optimised?”. Finding an answer to this question was done through five steps.

Firstly, through a literature research, information was found on how electronic music is made and performed. This research provided information about the needs of performing artists of electronic music. This made it possible to create a more usable and functional prototype. Additionally, through literature, information was gained about the relationship between embodiment and sound as well as on different ways of measuring and objectifying movement. This information was useful for understanding the importance of certain movements in the context of music. It also provided a useful background on how to measure these movements. Afterwards, an embodied music controller was developed that made it possible to perform the arrangement of electronic dance music. This was done through recognising gestures and movements and mapping these to a live set in Ableton live. Later, evaluation sessions with producers and dancers were held to gain insight into the quality of the prototype and the usability of the controller. Through iteratively changing the prototype, the music controller became more robust and usable throughout the evaluations. By the end the participants were able to perform the actions they wanted to make and they enjoyed making them. Finally, the perception of an audience on the developed music controller was measured through a survey. The results of this survey seem to indicate that there is interest in concerts given with music controllers like the one developed in this thesis.

The explored method of performing EDM using movement, has illustrated a way of performing electronic music that provides the audience with visual information about the actions of the performer. Given that the music controller has not yet been used to give a physical live concert, it is still hard to say exactly how successful such a concert will be. However, the evaluations done in this study seem to indicate that the use of gestures and movements show benefits in the way live concerts are experienced by both the performers and the audience members. Whether performing electronic music using movement will become a gimmick or the next new standard for live shows is still hard to say, but it is undeniable that it is a novel and exciting alternative to live shows.

Bibliography

[1] M. Fujimoto, N. Fujita, Y. Takegawa, T. Terada, and M. Tsukamoto, "A motion recognition method for a wearable dancing musical instrument," in *2009 International Symposium on Wearable Computers*. IEEE, 2009, pp. 11–18.

[2] C. Erdem, K. H. Schia, and A. R. Jensenius, "Vrengt: a shared body-machine instrument for music-dance performance," *arXiv preprint arXiv:2010.03779*, 2020.

[3] Y. de Quay, S. Skogstad, and A. Jensenius, "Dance jockey: performing electronic music by dancing," *Leonardo Music Journal*, pp. 11–12, 2011.

[4] Neova enhance your expressiveness. [Online]. Available: https://www.enhancia-music.com/product/neova/

[5] Mi.mu gloves. [Online]. Available: https://mimugloves.com/

[6] S. Matla. The advanced guide to tension and energy in electronic music. [Online]. Available: https://www.edmprod.com/tension/

[7] V. Joaquim and Á. Barbosa, "Are luminous devices helping musicians to produce better aural results, or just helping audiences not to get bored?" 2013.

[8] K. Cascone, "Laptop music-counterfeiting aura in the age of infinite reproduction," *Parachute*, pp. 52–59, 2002.

[9] C. Bahn, T. Hahn, and D. Trueman, "Physicality and feedback: a focus on the body in the performance of electronic music," in *ICMC*, 2001.

[10] M. T. Marshall, P. Bennett, M. Fraser, and S. Subramanian, "Emotional response as a measure of liveness in new musical instrument performance," 2012.

[11] R. Solberg and A. Jensenius, "Pleasurable and intersubjectively embodied experiences of electronic dance music," *Empirical Musicology Review*, vol. 11, p. 301, 04 2017.

[12] B. Lyubenov, "Music analysis of the electronic dance music mus-17," 2017.

[13] P. Kirn, *Keyboard presents the evolution of electronic dance music.* Backbeat Books, 2011.

[14] J. A. Magana, "Performance in edm-a study and analysis of djing and live performance artists," 2018.

[15] Apc mini. [Online]. Available: https://www.akaipro.com/apc-mini

[16] C. Gates, S. Subramanian, and C. Gutwin, "Djs' perspectives on interaction and awareness in nightclubs," in *Proceedings of the 6th conference on Designing Interactive systems*, 2006, pp. 70–79.

[17] O. Bown, R. Bell, and A. Parkinson, "Examining the perception of liveness and activity in laptop music: Listeners' inference about what the performer is doing from the audio alone," 2014.

[18] J. Blacking and B. Nettl, *Music, culture, and experience: Selected papers of John Blacking.* University of Chicago Press, 1995.

[19] D. J. Levitin, *This is your brain on music: The science of a human obsession.* Penguin, 2006.

[20] M. Leman, "Musical gestures and embodied cognition," in *Journ 'e es d'Informatique Musicale (JIM-2012).* University 'e of Mons, 2012, pp. 5–7.

[21] M. Leman *et al.*, *Embodied music cognition and mediation technology.* MIT press, 2008.

[22] M. Peckel, T. Pozzo, and E. Bigand, "The impact of the perception of rhythmic music on self-paced oscillatory movements," *Frontiers in psychology*, vol. 5, p. 1037, 2014.

[23] R. I. Godøy and M. Leman, *Musical gestures: Sound, movement, and meaning.* Routledge, 2010.

[24] F. Bevilacqua, N. Schnell, and S. F. Alaoui, "Gesture capture: Paradigms in interactive music/dance systems," *Emerging Bodies: The Performance of Worldmaking in Dance and Choreography*, vol. 183, pp. 183–193, 2011.

[25] A. Tanaka and R. B. Knapp, "Multimodal interaction in music using the electromyogram and relative position sensing," *NIME 2002*, 2002.

[26] C. Hilliard and S. W. Cook, "A technique for continuous measurement of body movement from video," *Behavior Research Methods*, vol. 49, no. 1, pp. 1–12, 2017.

[27] M.-J. Yoo, J.-W. Beak, and I.-K. Lee, "Creating musical expression using kinect." in *NIME*, 2011, pp. 324–325.

[28] L. Hantrakul and K. Kaczmarek, "Implementations of the leap motion device in sound synthesis and interactive live performance," in *Proceedings of the 2014 International Workshop on Movement and Computing*, 2014, pp. 142–145.

[29] B. Caramiaux and A. Tanaka, "Machine learning of musical gestures: Principles and review," in *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. Graduate School of Culture Technology, KAIST, 2013, pp. 513–518.

[30] J. Wu, G. Pan, D. Zhang, G. Qi, and S. Li, "Gesture recognition with a 3-d accelerometer," in *International Conference on Ubiquitous Intelligence and Computing*. Springer, 2009, pp. 25–38.

[31] J. C. Schacher, "Motion to gesture to sound: Mapping for interactive dance." in *NIME*, 2010, pp. 250–254.

[32] Sound on intuition. [Online]. Available: https://www.stylus.com/kkkzgp

[33] A. Chang and H. Ishii, "Zstretch: a stretchy fabric music controller," in *Proceedings of the 7th international conference on New interfaces for musical expression*, 2007, pp. 46–49.

[34] G. M. Troiano, E. W. Pedersen, and K. Hornbæk, "Deformable interfaces for performing music," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 377–386.

[35] Roli seaboard. [Online]. Available: https://roli.com/products/seaboard

[36] D. Tetteroo, D. Reidsma, B. Van Dijk, and A. Nijholt, "Design of an interactive playground based on traditional children's play," in *International Conference on Intelligent Technologies for Interactive Entertainment*. Springer, 2011, pp. 129–138.

[37] S. K. Card, W. K. English, and B. J. Burr, "Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a crt," *Ergonomics*, vol. 21, no. 8, pp. 601–613, 1978.

[38] R. Von Laban, *Modern educational dance*. Princeton Book Co Pub, 1975.

[39] J. Françoise, S. Fdili Alaoui, T. Schiphorst, and F. Bevilacqua, "Vocalizing dance movement for interactive sonification of laban effort factors," in *Proceedings of the 2014 conference on Designing interactive systems*, 2014, pp. 1079–1082.

[40] L. Souza and S. Freire, “Towards an interactive tool for music and dance: gestures, laban movement analysis and spectromorphology.” *Revista Música Hodie*, vol. 18, no. 1, 2018.

[41] (2013) Labanotation. [Online]. Available: https://www.britannica.com/art/labanotation

[42] Labanotation. [Online]. Available: https://www.dcode.fr/labanotation

[43] Key to labanotation symbols. [Online]. Available: https://www.britannica.com/art/labanotation#/media/1/326646/71523

[44] Enhancia launches neova midi ring controller. [Online]. Available: https://musicplayers.com/2020/11/enhancia-launches-neova-midi-ring-controller/

[45] Mi.mu gloves. [Online]. Available: https://dublin.sciencegallery.com/sound-check-exhibits/mimu-gloves

[46] S. Dahl and A. Friberg, “Visual perception of expressiveness in musicians’ body movements,” *Music Perception*, vol. 24, no. 5, pp. 433–454, 2007.

[47] Linear algebra and rotation. [Online]. Available: https://dev.to/jemaloqiu/learn-julia-12-linear-algebra-and-rotations-4n83

[48] Mpu6250 pcf electronics. [Online]. Available: https://www.pcfelectronics.nl/nl/mpu6050-6-as-gyroscope-accelerometer-sensor.html

[49] Mpu9250 datasheet. [Online]. Available: https://components101.com/sensors/MPU9250-9-dof-mems-sensor-module-datasheet-pinout-features-working

[50] Nodemcu elektor. [Online]. Available: https://www.elektor.nl/nodemcu-microcontroller-board-with-esp8266-and-lua

[51] Tp5000 battery protector. [Online]. Available: https://nl.aliexpress.com/item/32465408605.html

[52] A. Fu and Y. Yu, “Real-time gesture pattern classification with imu data,” 2017.

[53] bartkowiaktomasz. Fitness activity classification with lstms. [Online]. Available: https://github.com/bartkowiaktomasz/fitness-activity-classification-with-lstms

[54] V. Braun and V. Clarke, “Reflecting on reflexive thematic analysis,” *Qualitative Research in Sport, Exercise and Health*, vol. 11, no. 4, pp. 589–597, 2019.