Universität Bremen

Fachbereich 3: Mathematik und Informatik

# Bachelor Thesis

## JazzHandsXR: A Mixed Reality Application for Visualising Hand Structure in Piano Playing

Dmitry Tschernobai

Matriculation No. 612 835 3

February 15, 2025

**First Examiner:** Prof. Dr. Gabriel Zachmann
**Second Examiner:** Dr.-Ing. Robert Porzel
**Supervisor:** Dr. Rene Weller

**Dmitry Tschernobai**

JazzHandsXR: A Mixed Reality Application for Visualising Hand Structure in Piano Playing

Bachelor Thesis, Fachbereich 3: Mathematik und Informatik

Universität Bremen, February 2025

## Statutory Declaration

I hereby declare that I have written this thesis independently, have not submitted it for any other examination purposes, and have not used any resources other than those indicated. All text excerpts, quotes, or content from other authors that have been knowingly used are explicitly referenced as such.

Bremen, February 15, 2025

Dmitry Tschernobai

## Acknowledgements

I would like to extend my deepest gratitude to those who have supported me throughout this journey. To my parents, thank you for your unwavering encouragement and patience as this project stretched over an extra semester. Your support gave me the strength to persevere.

A special thanks goes to my sister, who believed in me even when I doubted my abilities to finish this project. Her faith in me kept me motivated, and I am incredibly grateful for her constant encouragement. I also want to thank my friend Mert, who generously lent me his VR headset, making it possible to bring this project to life.

To my other friends, thank you for always listening when I hit creative or technical blocks and for believing in me throughout the process.

Lastly, a warm thank you to my cat, who lay by my side during countless late nights, offering quiet comfort whenever I felt stuck in the implementation. Each of you played an invaluable role in the completion of this work, and I am deeply thankful.

# Abstract

This thesis introduces JazzHandsXR, a mixed reality application designed to support self-taught piano learners by providing real-time visualization of hand positioning and movement. Traditional piano instruction offers immediate feedback on ergonomic hand posture, which self-learners often miss when using conventional resources like sheet music or apps such as Synthesia. These tools typically focus on note sequences, lacking guidance on hand technique —an essential component for preventing strain and promoting efficient playing.

JazzHandsXR addresses this gap by projecting virtual ”ghost hands” that model correct hand shapes and finger transitions, enabling users to mirror these movements in real time. By providing this interactive feedback, JazzHandsXR gives self-learners valuable insight into effective hand positioning without the immediate need for a teacher’s supervision, encouraging healthier and more productive practice habits.

The application leverages hand-tracking technology to animate virtual hands above a real piano setup, simulating essential techniques and postures needed for fluid playing. Emphasizing hand mechanics like fingerings, transitions, and posture, JazzHandsXR provides a unique focus compared to traditional note-based learning tools. Although a virtual keyboard display aids orientation, the primary focus remains on hand technique, making JazzHandsXR a useful tool for independent piano practice.

In evaluating this system, a user study was designed to assess its usability, effectiveness, and overall learning support. Initial observations reveal that while JazzHandsXR offers substantial benefits, it has certain limitations: the lack of haptic feedback and the simplified animation may affect the precision of technique training. Nonetheless, the flexibility it offers—allowing learners to practice at their own pace and repeat exercises as needed—highlights its value for users without consistent access to in-person guidance.

Ultimately, while JazzHandsXR cannot fully replace in-person instruction, the study’s findings suggest it is an accessible and promising aid for self-learners aiming to develop foundational hand techniques.

# Contents

Chapter 1

# Introduction

”When all is said and done and we’ve built the highest high rises and we built the fastest machines, there’s still going to be room for somebody to tell you a story or write you a piece of music.” – Hans Zimmer [Zim18]

With these words, famous composer Hans Zimmer points out that, no matter how much technology advances, the ability of humans to create music will always be important. But as technology continues to grow, the question is: how much will it be involved in music? Will it help improve the creative process, or will music remain, despite all innovations, a domain where the human touch is indispensable?

This brings me to the topic of my thesis which touches on the field of piano education in combination with multimedia and specifically looks at how Mixed Reality could improve piano playing by helping with hand positioning and potentially preventing injuries while gaining a more adequate playing technique. The question also is, can a mixed reality application make a difference in learning a highly intricate skill like playing the piano?

To answer these questions and explore the world of music making in virtual reality I developed a mixed reality application over the course of writing this thesis, which I am calling JazzHandsXR, for conducting a user study on this topic and to underline the ideas presented with some data.

## 1.1 Motivation

Learning to play the piano is a goal for many people, but it is not always easy to achieve. Traditional piano lessons can be expensive and require a lot of time. On the other hand, learning on your own is often frustrating. Beginners might struggle with using the right fingers, keeping a good posture, or staying motivated without guidance. Meanwhile Mixed Reality offers exciting new possibilities to solve problems. By combining the real world with digital elements, MR can help learners see exactly how they should play, get immediate feedback, and stay motivated through interactive and fun features. This makes learning more accessible and engaging, especially for people who cannot attend regular lessons.

In this section, I aim to explore the reasons why this topic is important and why there is a need for applications like this. Specifically, I will examine the factors that can hinder individual learning progress and the common challenges people face when trying to learn an instrument like the piano starting with piano related muscoskeletal disorders (PRMDs).

When discussing body or hand posture at the piano, we cannot leave out the topic of piano-related injuries. Poor playing technique and improper posture can lead to various physical disorders over time. To better understand this connection, I reviewed existing research to find evidence of injuries and health issues directly linked to inadequate piano technique.

First, I would like to examine the risk factors associated with piano-related injuries. Meaning, I want to answer what happens to the body during these injuries and under what conditions they are most likely to occur. For this I would like to take a closer look at the paper "A systematic review of prevalence and risk factors associated with playing-related musculoskeletal disorders in pianists" by Bragge et al.: The paper systematically reviews the prevalence and risk factors associated with playing-related musculoskeletal disorders (PRMDs) in pianists. PRMDs are common among musicians and can include pain, weakness, numbness, or other symptoms that interfere with playing. The prevalence rates vary widely, ranging from 26% to 93% depending on the study, reflecting inconsistencies in definitions and methodologies. [BBM05]

Key risk factors for PRMDs include poor posture, small hand size, prolonged practice sessions, and inadequate breaks. Some studies also highlight the role of stress and improper technique. However, there is a lack of consensus, as different studies often identify conflicting or unsupported factors. That's an interesting insight for my app, since it emphasizes the importance of addressing these risk factors—such as posture and technique—through technological solutions. There is potential to offer guidance to help pianists avoid PRMDs, improving both performance and physical well-being.

But let us answer the question who it might affect and what the playing technique looked like on a simple level. For this there I want to refer to a similar paper "The prevalence of playing-related musculoskeletal disorders in relation to piano players' playing techniques and practising strategies" by Allsop et al. The study investigates the prevalence of playing-related musculoskeletal disorders (PRMDs) in piano players, examining their connection to playing techniques and practice strategies. A survey of 505 pianists found that 42.4% reported experiencing PRMDs.

Professional pianists were more affected (71.9%) compared to non-professional pianists (38.1%). Teachers and performers also reported more severe symptoms than casual players.[AA10]

Key factors linked to playing-related musculoskeletal disorders include posture, practice habits, and experience. Elevated shoulders and neutral wrist postures were associated with a higher risk of developing PRMDs, while longer practice hours and fewer breaks were found to correlate with more frequent occurrences of these disorders. Additionally, more years of piano playing increased the likelihood of PRMDs, especially for those practicing advanced techniques such as octaves and chords. Interestingly, hand span size was not significantly associated with PRMDs. The study underscores that PRMDs arise from a combination of overuse, misuse, and poor practice conditions, highlighting the need for better education on safe playing techniques and healthier practice habits.

This is interesting to me because, as I understand it, it does not really affect complete beginners. It starts to become a problem when the student has already developed bad habits. This means it plays a bigger role at the amateur level, where players face the challenge of handling a large number of notes for the first time but have not yet developed a proper strategy for hand posture to manage this workload in a single playing session. For my app, this insight suggests that testing should focus on scenarios involving a high number of notes that require certain fingering strategies. The goal is to evaluate whether the playing strategy promoted by the app allows users to maintain proper technique while staying relatively loose and relaxed. This is particularly important because tension and stiffness are common signs of inefficient technique, which can contribute to long-term playing-related injuries.

Now, let's take a step back from PRMDs and consider other, simpler reasons why this app might be relevant. While injuries represent the most serious concern, it's equally important to address how the good playing technique is actually conveyed to the piano student. Clear instruction and communication are crucial for maintaining motivation. Let's assume for a moment that the perfect practice strategy for hand posture problems exists. Even then, there is still the problem that nobody would use this strategy if it were not fun and time-efficient enough to implement. This is where gamification might come in handy. Gamification introduces elements of fun and engagement, such as visualizing the notes being played, highlighting mistakes and introducing challenges while setting achievable practice goals. This app could offer a fun and exciting alternative to traditional training methods, making fingering and hand posture a priority for learners from the very beginning.

Furthermore, gamification can foster a sense of progression and accomplishment, which is crucial in maintaining long-term interest and dedication. Combined with the app's ability to provide real-time feedback, these features ensure that users not only learn effectively but also develop a positive relationship with the instrument. In the end, the success of a learning tool is sometimes less about its raw effectiveness and more about its efficiency. Effectiveness measures whether the tool achieves its intended goal, such as improving hand posture or technique. However, efficiency focuses on how quickly and easily users can reach that goal while maintaining their motivation and engagement.

For an app like this, it's not just important that it works, but that it fits seamlessly into the learner's routine, offering results without requiring excessive time or effort. By prioritizing efficiency, the app can appeal to a wider audience, making it more accessible and practical for users who might not have the time or patience for traditional methods. We will delve deeper in later chapters into the specific features I intend to implement in the app to incorporate elements of gamification, aiming to improve efficiency rather than focusing solely on raw effectiveness.

## 1.2 Hypothesis and Objectives

This section explores the main hypothesis behind the development of my app for this thesis. The hypothesis is formulated in a way that allows for straightforward testing and data collection, ensuring that the research remains focused and feasible. I aim to avoid unnecessary complexity and keep the approach simple and practical. I will also define the goals for this work in a way that allows me to revisit them later and assess whether they have been achieved. The research question guiding this exploration is: How can a Mixed-Reality application, such as JazzHandsXR, support the learning process for hand posture and fingering in piano playing? Alongside this question, I hypothesize the following: Using a Mixed-Reality application as a supportive learning has a positive influence on the user experience and enhances the understanding of hand placement and movement coordination in piano playing towards a given hand posture target. So as you can tell I am not going for the optimal revolution in piano education. And my goal is not to answer the question whether the hand posture actually gets better long term, since I will not provide the theoretically best posture. This would probably be a whole work on its own. I do not aim to completely replace piano lessons, teaching, or in-person instruction, as I believe these are very important, and we will discuss this further in the chapter on Piano Pedagogy. What matters to me is that an app such the one which I am developing serves as an additional positive tool—an alternative that users can turn to without feeling like they are wasting their time. I want to find reasons why students might want to add this training methods on top on their traditional practice. Maybe it will offer some valuable insights, and perhaps it will resonate differently with each individual. Sometimes it's not what the teacher is telling but how he packages this information. This is also the approach taken in this thesis. Such an app will not yet have the power to replace traditional lessons, and we will also discuss in the "State of Mixed Reality" chapter why this is still the case, but it can be a great tool nevertheless for different reasons.

Now Let's formulate some clear objectives that I should reach in order to make the user study and finally test the hypothesis. First, the app aims to improve users' hand coordination through the interactive guidance provided by, what I will from now on refer to as, the Ghost Hand feature. This feature represents a core innovation of the application and will be discussed repeatedly throughout this thesis as it forms the foundation of the proposed improvements. The Ghost Hand works as follows: normally, users would only see a visual representation of the note sequence they need to play, displayed as bars indicating the keys to press. However, this approach does not convey any information about proper hand posture. The Ghost Hand addresses this gap by displaying an animated, semi-transparent hand that overlays the user's own. By aligning their hand with the Ghost Hand, users can not only match the correct notes but also emulate the recommended hand positioning. This dual focus on note accuracy and posture is designed to foster healthier and more efficient playing habits. My app also aims to deepen the user's understanding of proper hand positioning principles by providing informative animations and text elements before exercises are played. By leveraging Mixed Reality technology, the app aspires to enhance the overall learning experience, offering a richer context that allows users to observe and reflect on the key principles of hand posture and technique. This approach helps users absorb essential concepts that are critical for achieving efficient and relaxed playing, setting a solid foundation for future practice. For example, before beginning an exercise, the user could watch an animation showing the correct relaxation of the hand or finger placement in action, accompanied by explanatory text that highlights the importance of it.

Now before we conclude this chapter, I want to specify the target audience, since a given solution might not always fit everybody. The target audience for this app is neither complete beginners that never touched a piano before nor advanced players. Complete beginners face the challenge of learning basic skills, such as pressing the correct notes, and their focus is typically not yet on proper hand posture. I don't mean to suggest that they don't need technique—of course, they need it from the very first note they play. However, testing this might be challenging within the scope of my work. Advanced players, on the other hand, have already developed their hand techniques to a considerable degree and may not require guidance on fundamental aspects of posture. Advanced players require more detailed improvements to their playing technique and, above all, individuality, which is also beyond the scope of my work. Instead, this app is designed for those who are beginning to consider the importance of clean playing with natural movements or for those who may not yet be aware of the concept of adequate hand posture. It addresses players who are at a stage where they are starting to refine their technique and build a solid foundation for healthy and efficient playing habits. To conclude this chapter, I hope that the app will not only offer an innovative tool for improving hand posture and playing technique but also contribute to a deeper understanding of how Mixed Reality technology can be integrated into piano education. Ultimately, the app aims to explore its feasibility, providing valuable insights and advancing the ongoing research in this field.

Chapter 2

# Related Work

With the rise of modern technology, extensive research has been conducted on the topic of piano technology. It is logical to develop a tool to assist with piano playing, as modern technology has significantly contributed to advancements in other fields as well. While searching for similar applications for this thesis, I came across several tools that address specific problem areas and attempt to solve them. Additionally, there are well-conducted user studies evaluating these applications. I will present some of these applications to give an idea of where this bachelor thesis fits within the broader context. The topic appears to be well-researched, as it is highly practical due to the ease of combining head-mounted displays with electric pianos. Since electric pianos produce MIDI input, this data can be forwarded to an application, and with the help of cameras, such as external depth cameras or the cameras on the headsets itself, it is possible to analyze body and limb posture. This way, data can be collected on the accuracy of playing—whether it involves piano pieces, technical exercises, or posture training. There is significant potential for an application to provide accurate feedback, helping to correct mistakes and improve playing. Furthermore, as VR becomes increasingly popular, the number of VR applications specifically for piano education is also growing.

One particular example that closely resembles my idea is VRMonic by Ethan Matzek et al. This is a VR application designed to assist with hand positioning by overlaying the image of an expert's hand performing a specific exercise. The user attempts to match their hand to the expert's, receiving color-coded feedback indicating how closely they align. In this way, the application presents the correct hand form, allowing users to practice their posture and reduce the risk of injury.[Mat+24] First, they collected the hand data with an RGB-D camera and used a pre-trained hand landmark model to extract the joint positions for the hand poses. The target hand posture is calculated as the mean of the joint positions per frame across several trials [Mat+24]. This provides a playback of the ideal hand posture for performing an exercise. The student can then record themselves performing the same exercise and compare the hand positions afterward.

They conclude that this technology has the potential to reduce injuries for both novice and professional musicians, as it provides guided feedback even in times where a real instructor is unavailable to correct their playing. But they also point out the limitations, as it does not provide real-time feedback, instead, they have to compare the playback of recordings. Additionally, it is a VR application, meaning the piano is entirely virtual, which impacts the learning experience. This is where my thesis comes in: my goal is to extend this idea by exploring an approach using the Quest 3 headset, which features color passthrough. This allows me to move beyond virtual reality and utilize a real piano in a mixed reality environment.

One other application is by Johnson et al., who developed an application that focuses on identifying incorrect hand postures and classifying them accurately using machine learning. The system works by recording the student's hands from above with a camera while they play an exercise. The recorded frames are processed to segment the hands from the background, assigning pixels to specific hand regions. Then, a Support Vector Machine classifier analyzes the segmented data to classify the hand posture into one of three categories: (1) correct posture with proper finger curl and wrist height, (2) flat fingers, or (3) collapsed wrist. The system achieved over 90% accuracy in distinguishing these postures. For a given video of a student playing, the system can estimate how often, frame by frame, the student displayed an incorrect posture, such as flat fingers. This provides the student with clear, actionable feedback on areas to improve[JDT20].

This concept is highly relevant to my thesis because it demonstrates that hand posture can be quantified and classified as incorrect. While hand posture is somewhat subjective and players may develop their own techniques, there are still common mistakes that hinder progress, often because the hands are not naturally positioned for optimal piano playing. A key principle is that the hands should remain as relaxed as possible to achieve speed and accuracy. Flat fingers and collapsed wrists are typically stiffer and less relaxed, which is why they can be classified as poor hand posture. But I won't focus too rigidly on the exact wrist position and curl angles, as it can be challenging to define precisely, and it's not the entire picture of piano technique anyway. There is much more depth and nuance to it than just identifying a few incorrect postures. However, I will keep it in mind when developing such a mixed reality experience, as it's important to understand the potential mistakes students might make. The goal should be to guide them in the right direction by raising awareness about the different possibilities and impacts of various hand positions.

Since I plan to conduct a user study on my mixed reality app, I have reviewed similar applications to understand what has been previously achieved with this technology. I came across a case study by Mariano Banquiero et al. focusing on the Passthrough functionality of the Quest 2, which is the predecessor of the headset I intend to use, making it very relevant to my work[Ban+23]. In this study, they evaluated the appearance and usability of a mixed reality learning app for piano. A total of 33 participants played piano pieces using different visualization modes. Their performance was assessed by teachers, and their experiences were captured through questionnaires. The evaluations included aspects such as simulator sickness, user presence, ease of use, and self-reported preferences.

They found that 72.7% of the users like the solid mode where the virtual keyboard overlays the real keyboard with solid colors and the wireframe mode caused more significant symptoms of discomfort but remained within acceptable thresholds. It is pointed out that Passthrough technology currently lacks color rendering and is unstable at short distances, impacting usability in the wireframe mode. Manual calibration was required as the application couldn't automatically detect the piano or its keys. [Ban+23] It appears that I will encounter similar challenges with the next-generation headset. One major limitation is the lack of an API for accessing Passthrough images, which prevents the implementation of image recognition algorithms. As a result, I need to devise a calibration mechanism to ensure the digital clone of the piano remains accurately aligned with the physical keyboard. This case study is a good insight for my own study design.

## 2.1 Theoretical Background

Before going deeper into the subject, I would like to clarify some key technical terms and concepts. To do so, I will provide definitions and explain the necessary details to ensure a clear understanding as we proceed with the thesis. Lets first define the terms for describing the different levels within the virtual reality spectrum by looking at the definition given by Paul Milgrim. He coined the term Reality-Virtuality Continuum, which illustrates the different stages between the real environment and a fully virtual environment. A virtual environment is a world or a region which must be completely modelled by a computer in order to be rendered. So VR creates an entirely synthetic, computer-generated environment so that the real world is completely replaced, immersing the user in a virtual space. Real environments are completely unmoddeled. These make up the two extremes of a spectrum. Wihtin this spectrum there are Augmented Realities (AR) and Augmented Virtualities (AV). AR resides closer to the real environment side of the continuum. It overlays digital information or objects onto the physical world, enhancing but not replacing the user's perception of reality. AR ensures the real world is still dominant, with digital elements augmenting it. Augmented Virtualities add real information to a computer-generated environment. Mixed Reality (MR) is the range that includes the entire spectrum, but excluding the endpoints of a fully real or fully virtual environment. Therefore it is an environment that includes both augmented reality and augmented virtuality and where virtual elements coexist with real elements. In MR, virtual objects are not just overlaid but are anchored to the real world, enabling users to interact with and manipulate both real and virtual objects in a shared space. [MK94]

This definition of mixed reality in particular is very important for this thesis, since the idea of JazzHandsXR is to create a space for interaction between virtual elements such as the ghost hands and the real piano.

## 2.2 Current State of Mixed Reality

To gain a better understanding of the current state of research and to discuss existing applications in later chapters, I want to grasp the current state of Mixed Reality (MR) in this section. This brief chapter examines the possibilities available today and the challenges that characterize the state of MR in 2024. We need to talk about advancements in hardware as well as innovations in sensor technology, eye-tracking, and hand-tracking. Additionally, I want to review prominent development platforms alongside tools and APIs. Finally, let's address technical limitations like latency, tracking accuracy, and rendering requirements and barriers to development. In the paper "Laying the foundation for Augmented Reality in Music Education" Nijs et al. discuss the common devices that can be used for mixed reality in music education that include for example HoloLens, Meta Quest Devices, and Nreal Light, each offering varying degrees of portability and visual fidelity. They argue that AR has found it's way into a lot of fields like education and healthcare and has become more accessible than ever. These devices often integrate with Unity, Unreal Engine, and Mixed Reality Toolkits (MRTK) for development. APIs like OpenXR and ARKit/ARCore are prevalent, enabling cross-platform compatibility. [NB24]

But what are the challenges in the moment right now with those devices? Can they be programmed to do every task, no matter the level of detail? As the paper "Presence and Flow in Virtual and Mixed Realities for Music-Related Educational Settings" by Bruns et al. suggests, there are some challenges that has to be adressed. MR still struggles with technical reliability, especially in terms of tracking accuracy and latency. Breaks in plausibility can disrupt immersion and reduce the perception of presence, which is essential for effective learning. Current HMDs, while more portable and visually advanced, still suffer from artifacts in rendering and environmental integration, making detailed applications unreliable. MR applications are often used for educational gamification and collaborative experiences but remain far from being robust enough for high-stakes applications like surgery or advanced engineering. The cost of development and hardware significantly limits widespread adoption in fields like medicine, where precision and reliability are critical.[Bru+24] I imagine that for music in general every little bit of latency or unreliable features are more important than in other applications because for good training it requires fineness and subtlety. For my app for example that would mean, that i should not focus on the precision of finger placement, since there will be a mistake that will naturally accumulate. I should go for broader movements and general ideas rather than precision, since the devices probably cannot catpure those details reliably. I also found some interesting insights regarding the limitations of the hand tracking technologies in the paper by Amm et al. They argue that although gamified learning elements and remote collaboration tools are highly engaging for beginner and intermediate learners they lack the sophistication to replace traditional teaching methods. For example MR applications for piano education demonstrate limited precision in real-time feedback and fine motor tracking. Hand-tracking technologies, though promising, often lack the resolution needed for detailed and nuanced movements.[Amm+24] This further emphasizes the last point I made for my app. One honorable mention should go to the apple vision Pro, since it is a new headset that has exceptional capabilities in terms of passthrough, display and cameras and that could be of great benefit for those types of mixed reality experiences.

It shows what our future has to offer and how future apps can use technology. It is particularly well-suited for mixed reality applications due to its advanced hardware capabilities, including high-resolution displays and the M2 chip, which ensure smooth and detailed rendering of virtual environments.[Wai+24]

If we continue discussing hardware, this time focusing on piano-related media, we can mention the ROLI Airwave. This is a relatively advanced technology capable of tracking hands above a keyboard and displaying them on a screen as a visual twin. It can provide guidance on proper hand positioning and indicate which notes should be played with specific fingers. However, it is important to note that it uses a synthetic keyboard rather than a real one, and the hands are only displayed on a screen, such as on an iPad. Therefore, it cannot be classified as mixed reality.[Tho24] For my work, it's important to look at technologies like the ROLI Airwave because they show how hand movements can be tracked in a musical context. It helps me understand how these systems work, how they track the hands, and translate that into visual feedback. This is useful for figuring out how such technologies could be applied to piano learning. Plus, it gives me insights into the strengths and weaknesses of these approaches, which helps me come up with new ideas for my app. I believe, for example, that the cameras on top can capture the hands relatively well, but I suspect that there might be positions, like when executing the thumb-under technique, that are not visible from above, which could lead to some lag.

After discussing various technologies, we now set them aside for a moment and shift focus to traditional piano teaching and the pedagogy behind it, which plays a crucial role in training students to learn proper hand positioning and fingering. This is essential for my work, as I need to develop ideas that will function within a mixed reality environment, with traditional piano instruction serving as the benchmark.

## 2.3 Piano Pedagogy

Piano pedagogy has been developed over centuries, shaped by countless teachers, performers, and students. It's the foundation of how people learn to play the piano, covering everything from hand positioning and fingering to overall technique. For my research, understanding these traditional methods is key, because they set the standard that any new tools—like mixed reality applications—will need to match or build upon.

In this chapter, I'll look at the core ideas of traditional piano teaching, focusing on how it trains proper hand posture and fingering. These are not just important for learning to play well, but also for avoiding strain or injuries. I'll also explore the contributions of well-known figures in this field and what their work has brought to piano education.

By looking at these time-tested methods, I hope to find ideas that can be adapted for modern tools. This connection between the old and the new is central to my work, as I aim to create something that respects the traditions of piano pedagogy while also pushing it into the future.

The first point I would like to make regarding classical piano pedagogy is the emphasis on aligning anatomy with music. It should never be forgotten that one plays the piano with their body, making it essential to anatomically consider how movements are executed—for instance, how much tension is present at a given moment. For further arguments, I will refer to Thom's book "What Every Pianist Needs to Know About the Body" to explore this perspective in more depth [TTR03, Chapter 1, p. 4].

An important principle in this context is that pain is always a sign that something is being done incorrectly. Playing the piano is different from sports, where muscles are trained by pushing them to the limit of pain. On the contrary, piano playing should always be pain-free, light, and relaxed. Only when one can perform pieces without exhaustion does it become evident that the necessary tension is applied only where and when it is needed[TTR03, Chapter 1, p. 5].

To provide some context, let's take a simple example. Imagine a student wants to play a passage with repetitive notes, meaning the same note is played multiple times in succession. The student's intuition might be to use the same finger for all the notes since it's already on the key. However, the teacher notices that the student's hand becomes tense while playing this passage. Why does this happen?

Well, when you perform the same movement repeatedly, you activate the same muscles with the same amount of tension each time. This tension adds up and that means the energy is focused on a small area, which leads to pain because that part of the hand becomes overworked and unbalanced.

In this case, the teacher advises the student to use a technique where the same note is played with different fingers by executing a quick finger switch. This introduces varied movements, distributes the workload more evenly, and makes the passage much more comfortable to play.

Another point Thom makes is the concept of the Body Map which is based on the Alexander Technique. A student will naturally focus their concentration on certain areas of the body more than others while playing. This can lead to neglecting certain body parts, causing them to tense up automatically without the student noticing. For example, the teacher might notice that the student's shoulders are raised without them realizing it. The teacher then guides the student to bring their attention to the neck and shoulders, and the student becomes aware that their shoulders are tense and lifted. The student is then encouraged to release them, letting them drop and relax. The goal here is to develop a more accurate Body Map. The Body Map refers to the mental representation of one's body and how the brain perceives its movements and posture. When a pianist has a clear and accurate Body Map, they are better able to control their movements and avoid unnecessary tension. By refining the Body Map, the student learns to become more aware of the different areas of their body and how to keep them relaxed and balanced while playing[TTR03, Chapter 2, p. 17].

I would like to briefly touch on what this means for the teacher. We see that a teacher must have the ability to observe and identify such tensions in a student and then respond effectively. This might involve explaining the identified problem, demonstrating better technique or suggesting specific exercises for the student to practice in order to address and eliminate the issue.

Now that we have understood the general principles of piano technique, let us delve into specific applications of these principles. One widely recognized method is the Taubman Approach, developed by Dorothy Taubman. This method offers a systematic guide to solving technical challenges that often lead to pain or injury, while providing concrete tools to make movements more fluid and ergonomic. The Taubman Approach addresses key aspects such as arm and hand rotation, finger independence, wrist movements, and the coordination of arm and hand, making it an invaluable resource for pianists seeking to refine their technique and prevent physical strain [Urv20].

We should take a look at an example of this method to get a sense of what is being conveyed. At one point in the video, a student struggles with a passage that requires playing powerfully on the black keys. Taubman notices that she is using the fourth finger and argues that no twist or rotation can be achieved when the fourth finger plays a black key followed by the first. She suggests using the fifth finger instead. Although the fifth finger is often associated with weakness, Taubman's point makes sense, as it allows the entire hand to rotate, enabling a powerful attack on the black key. Additionally, this technique introduces movement in the wrist, preventing it from becoming stiff[Urv20, 00:04:30].

Another fundamental principle of the Taubman Approach is co-contraction. Co-contraction refers to the simultaneous activation of opposing muscle groups, such as the flexors and extensors in a joint. While this can be useful for providing stability in certain situations, it can also lead to unnecessary muscle tension and restrict the freedom of movement. In the context of piano technique, co-contraction is undesirable because it creates stiffness and reduces fluidity in the player's movements. When a pianist inadvertently engages co-contraction, the hand or arm becomes less flexible, making it more difficult to execute smooth, controlled motions. This can not only hinder performance but also lead to muscle strain or injury over time, especially during prolonged practice or performance sessions [Urv20, 00:41:10].

Let's illustrate this with an example. The student wants to play fast octaves, which requires the hand to be relatively stretched, and the arm to move up and down.

The student performs the fast octaves by actively moving the arm both up and down to generate the speed. Taubman corrects this by taking the student's arm and instructing her to completely let go of the arm. She then lightly throws the arm and lets it drop, catching it afterward. The student realizes that she only needs to perform the upward movement, and gravity takes care of the rest. This way, she doesn't need to engage both opposing muscle groups, and the co-contraction is eliminated [Urv20, 00:42:50].

I believe the role of the teacher in this section has been effectively conveyed. What stands out most is the individuality required from the teacher to help students with their very different challenges. It's crucial to recognize that piano pedagogy engages many senses. Of course, one can visually observe what the teacher demonstrates and attempt to learn through imitation. However, there comes a point where the fine details of tension are no longer visible—either because the movement is too fast or because it's impossible to see what's happening with the muscles in the hand. At this point, a key aspect comes into play: the feeling of the hands. It seems to be an essential component of this process that the student must feel the teacher's hands in order to develop a sense of relaxation and how it should feel. This tactile experience becomes a vital tool in learning proper technique and understanding the physical sensations of relaxed and efficient playing.

We will now move beyond the classical methods and come back to modern applications, as the question remains of how we can combine both scientific advancements and the piano pedagogy practices into innovative training applications.

## 2.4 Existing Tools and Applications

Since there is extensive research on piano education supported by technologies, I have dedicated a separate section to exploring existing applications. It is crucial to understand what is currently available and how these technologies function to gain a sense of what has already been achieved and identify areas where gaps or limitations still exist. It is important to understand that applications in this field pursue diverse goals and therefore exhibit different features. The selection of features, as it turns out, is crucial, especially for the development of my own application. For instance, if an application is designed to train hand posture, as is the case with mine, it should incorporate 3D animations to effectively visualize hand movements and positions. On the other hand, if the goal is to improve sight-reading skills, 3D hand animations might even be counterproductive. Instead, such applications would focus on visualizing sheet music or providing real-time feedback on note accuracy.

Additionally, other considerations come into play depending on the intended use case. For example, whether the application is meant to be used independently by learners or as a supplement to lessons with a teacher significantly influences its design. An app designed for solo use may require more interactive and explanatory features, while one built to work alongside a teacher might prioritize tools for assessment and progress tracking.

Ultimately, understanding the specific goals and context of these applications helps clarify what has been successfully implemented and highlights opportunities for innovation and improvement in my own work. This contextual awareness ensures that the chosen features align closely with the desired outcomes and user needs.

I would like to explore this idea further by presenting the paper by Amm et al. that has invested significant effort into classifying and categorizing existing applications, some of which I have already introduced. This paper groups the applications into 4 groups based on the context in which they are used, such as whether they are designed to work with or without a teacher, which for example is characteristic of Group 4.[Amm+24]

It also considers whether the goal is to improve sight-reading skills or even to replace traditional sight-reading practices, which are categorized under Groups 1 or 2. Additionally, it examines applications aimed at cultivating a sense of musicality, such as improvisation, rather than focusing solely on the classical approach of playing pre-composed songs.

By introducing this classification, the paper provides a structured framework for understanding the different purposes and functionalities of these applications. This, in turn, can offer valuable insights into the breadth of educational strategies in piano learning and help identify where my own application fits within this spectrum.

The table titled ”Literature Overview of Piano Learning Applications Using Immersive Media Technologies” provides a detailed mapping of 16 applications and their implementation of various features. A closer examination reveals that the majority of these applications include a piano roll feature.

This feature visualizes the notes to be played as blocks rather than traditional sheet music notation. The length of each block corresponds to the duration for which a note should be held. Additionally, predefined songs are a common feature among the applications, meaning that the note sequences are fixed rather than generated randomly.

Another notable observation is that most mixed reality (MR) applications utilize a physical keyboard as the primary playing environment. However, only 2 out of the 16 applications specifically focus on training posture, and just 3 are designed to be used in collaboration with a piano teacher. Furthermore, only 3 applications incorporate 3D virtual hand animations as part of the learning experience.

Now it would be very helpful to compare my own planned features with this table and see what I am implementing and what is not covered. I will go through the list of features and decide which ones I will include in my app and explain why. A piano roll will be included, as it provides an effective visualization tool for beginners. Teaching sight reading won't be a focus, as my app is more about improving posture. Improvisation, expression, and music theory won't be included either, except for a little bit about posture theory, as it is valuable to understand hand positioning and the fingering system.

I will include predefined songs, but more in the form of exercises. There won't be a feature for working with a piano teacher, but the app will give feedback on accuracy, especially focusing on using the correct fingering. Virtual hand animations will be part of the app to help with fingering and posture, and the app will work with a physical keyboard. This means I cover about half of the possible features, focusing on a combination that appears to be less commonly implemented in existing apps.

Since there are 2 more papers that focused their application on hand posture training, I would like to present them. First being the ”Following the master's hands: Capturing piano performances for mixed reality piano learning applications” and the second one ”Hand-by-Hand Mentor: An AR based Training System for Piano Performance”.

Pianoverse captures how expert pianists move their hands and fingers while playing. It uses motion tracking, MIDI data, and music scores to create a synchronized system that shows these movements in 3D over a real keyboard. Learners can watch the pianist's virtual hands, follow their gestures, and see exactly how to position their own hands.[Lab+23]

The app includes tools like movement traces that show upcoming hand movements, and indicators that highlight which keys to press and with which fingers. These features help learners understand the right technique and posture. Tests showed that even beginners could imitate movements better and learn faster by watching these visuals.

The key takeaway I get from this paper is that motion capture is an effective tool for recording hand movements because it captures the essential joint positions, which can then be visualized. It's interesting that the authors chose to use thin purple light indicators instead of a fully detailed hand mesh. I imagine there are pros and cons to this approach, perhaps a fully meshed hand would be too overwhelming, and the light indicators ensure the user's own hands aren't overly obstructed.

However, in my project, I plan to use a fully detailed hand mesh for the visualization. The reason might be that the fingers are easier to see, and it helps to clearly associate the indicators with the correct fingers or avoid mixing them up. I can definitely imagine that being the case with Pianoverse. The Hand-by-Hand Mentor uses augmented reality (AR) to provide real-time 3D hand-motion animations for piano training, based on MIDI input and a pre-trained Hidden Markov Model (HMM). It automatically generates realistic animations of finger transitions, such as crossings and extensions, in sync with a physical piano through a head-mounted display (HMD). The algorithm predicts optimal finger movements and generates natural, realistic 3D hand motion animations, aiming to mimic human-like piano performance. User studies show that it reduces cognitive load compared to video tutorials and improves learning efficiency and quality, especially for beginners who need visual motion guidance[Guo+21].

So in summary, the Hand-by-Hand Mentor demonstrates the potential of AR in piano education by integrating an advanced motion animation generator with real-time feedback, making the learning process more efficient and immersive for beginners.

This is highly relevant to my study and demonstrates the significant advancements in the field, particularly in the generation of hand animations for applications like piano training. It is impressive to see that the state of research has progressed to the point where real-time, dynamic hand-motion animations can be generated based on models such as Hidden Markov Models (HMM). While my approach focuses on using pre-recorded, predefined animations, the research shows how advanced the field has become, and it is encouraging to see that more automated solutions are emerging. I believe that using pre-recorded animations remains a valid and effective method for providing high-quality guidance in piano learning.

Chapter 3

# Methodology

This chapter is all about the app I developed for this study, JazzHandsXR. We'll start by looking at the technologies I used to create a mixed reality piano learning app, including the frameworks and tools I built on and the reasons behind my choice of hardware. After that, we'll take an overview of the core functionalities—how the app is structured, the decisions I made during implementation, and some insights into the development process, including the challenges I encountered along the way.

In the final part, I'll discuss the predefined goals I had for the app's interaction capabilities summarized as affordances so that it is clear what the user can and cannot do in the app. Later, in the discussion, we'll see if these interactions were actually perceived by users as intended. So, let's begin by looking at the technological aspect of development first.

## 3.1 Technology and Software

For my development environment, I chose Unity and am using version 22.3.29f1. In my opinion, Unity offers a very smooth experience for VR development, thanks to the many toolkits that are available for it, and not least because I am more comfortable working with C# than C++. However, that shouldn't be the only reason. An alternative to Unity would have been Unreal Engine, which also offers great VR capabilities. I decided to go with Unity because its toolkits, especially those by Meta, are very well integrated within the platform.

The version I chose is from May 2024 and supports integration with the Meta XR All-in-One SDK, which is essential for VR development on the Meta Quest. During my development, this version has continued to receive updates, and I'm on the currently newest version from October 2024, specifically 69.0.1. This package provides functionalities such as the Camera Rig, basic building blocks like buttons and UI interactions, as well as hand tracking.

To give a clearer idea of the development environment, in the figure 3.1 you can see a screenshot showing Unity in action with an example of the UI from the Meta SDK.

You can see two buttons as part of the interface, which were implemented as part of the app's functionality. This screenshot demonstrates how the development environment looked during the process and gives insight into the tools I was using to create the user interactions in the app.

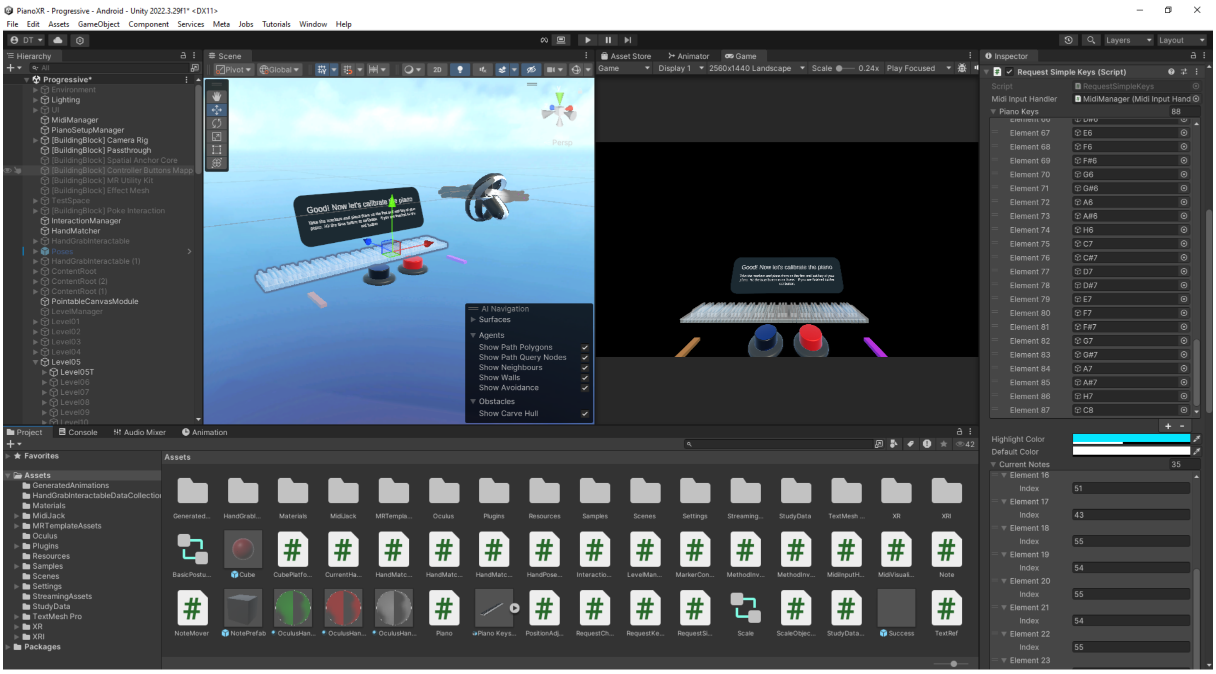

**Figure 3.1** Development Environment

At the core of the setup is the Quest 3, which is connected to the PC via Oculus Link using a USB-C cable. This setup allows me to run the app directly from Unity without the need to generate a build first, making testing a bit easier. Additionally, the PC is connected to a digital piano, which sends MIDI input to the computer via a printer cable, and finally to the Quest.

For programming the C# scripts, I use Visual Studio Community 2022, which allows for direct integration with the Unity engine. This allows for real-time synchronization between the scripts and the Unity environment, making the development process more efficient and streamlined.

When it comes to frameworks and libraries, besides the Oculus Integration library, I only use one other library, which is Midi Jack. All other assets are custom-created and written by myself. Midi Jack allows me to create an interface with the incoming MIDI data, enabling me to write a MIDI input handler. This handler sends events to my other scripts whenever MIDI input is detected—specifically when a key is pressed. In addition to the information about which MIDI channel is being used, I also receive the velocity of the pressed key. The channels can then be mapped to the keys of my modeled digital piano twin. But let us talk more on the implementation later in the chapter Implementation, where the general structure and layout of the code base will be discussed.

I also had to do some modeling in Blender, for example, for the virtual piano, which consists of 88 keys and the different transparent materials, allowing the user to still see the real piano underneath and helping it blend into the real environment. The keys have slightly rounded corners to make them look more realistic, rather than just using simple blocks.

For recording the hand animations I used the Meta Hand Animation Recorder, which is part of the Interaction SDK, and I used Unity's Animator Component to manage and playback these animations.

For version control, I used GitLab to manage the project. This allowed me to track changes and keep a backup of the development process.

## 3.2 Overview

Now, before diving deeper into the code structure I will walk through the app from a user's perspective in this section, highlighting its various features and demonstrating what it would feel like to interact with the app and play through it.

The experience is divided into three parts. The first part is a short introductory sequence, where the basics of interaction are introduced. This allows the user, especially if they are new to mixed reality, to get familiar with the environment and test whether the hand size matches correctly.

The second part is the Piano Setup, where the virtual piano is aligned with the real one. The user can handle this process entirely through various interactions with the UI and buttons. Finally, the third part is the main section, featuring five exercises for the user to play through.

The beginning starts with the user pushing a button with his hand, which triggers a welcome message. After reading it, the user is tasked with matching their hands and grabbing a few blocks to place them on the correct platforms. This serves as a test sequence to ensure that the user is prepared for the slightly more complex movements involved in the Piano Setup. It also verifies whether the hand tracking and the hand-matching algorithm work properly. If the hand size doesn't align perfectly, adjustments have to be made accordingly. In this figure 3.2 you can see the Test Sequence and how it might look like, when someone plays it through.

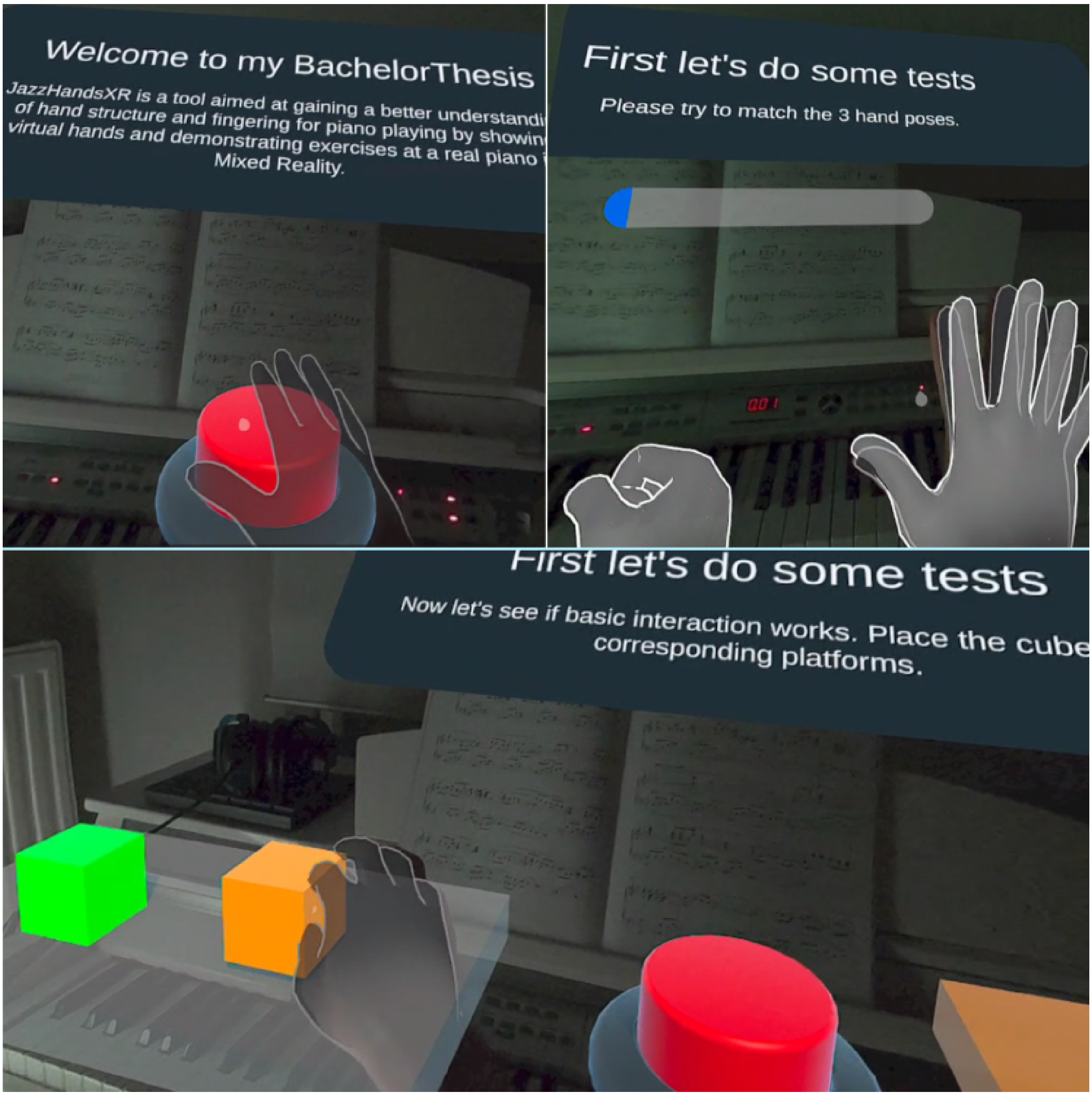


**Figure 3.2** Test Sequence

In the second phase, the digital piano is aligned with the real piano. To achieve this, two markers are placed on the outermost keys on the left and right sides using a grab gesture, and a button is pressed to perform a rough calibration of the piano. After that, users can refine the alignment with additional buttons if the initial calibration wasn't accurate. These controls allow adjustments along the three axes (x, y, z), as well as rotation and the width of the piano. This should be sufficient to create an almost perfect alignment.

Users can test the alignment by playing a few keys to check whether the virtual and real keys match. The following figure 3.3 illustrates what this process might look like.

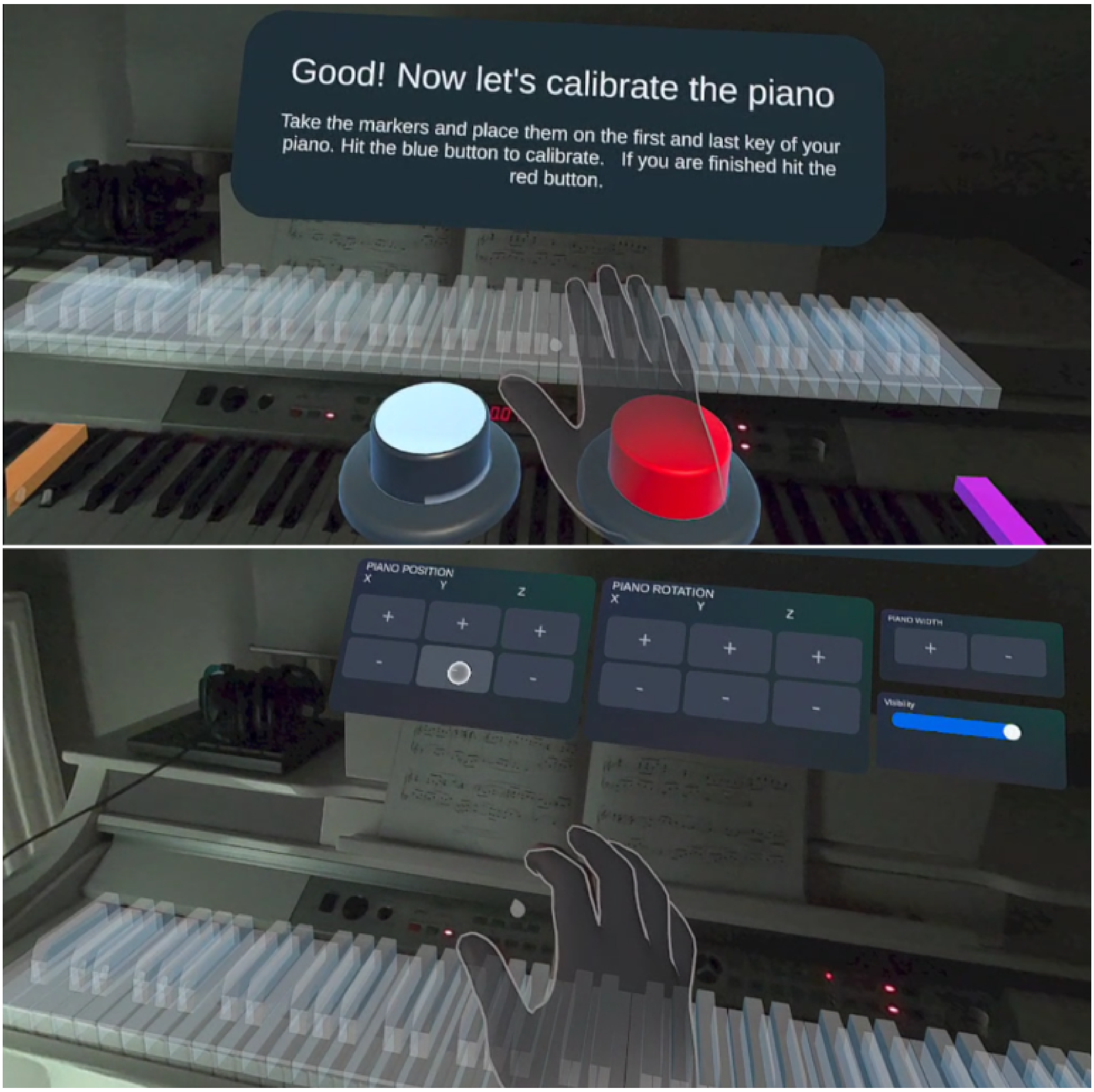


**Figure 3.3** Piano Calibration

Now that the calibration is complete, the third phase involves five exercises designed to teach the user the fundamental principles of proper hand positioning when playing the piano. I'll briefly walk through each of the five exercises, which I designed based on my research for this thesis on the topic as well as my own experience:

- **Recognizing correct hand positioning:** In this exercise, the user learns to visually identify proper hand posture.
- **Adopting the correct hand posture:** Here, the focus is on physically adjusting the hands to achieve the correct posture.
- **Playing a scale with proper hand positioning:** This introduces movement while maintaining the correct posture by performing a basic scale.
- **Practicing a melody:** The user progresses to playing a simple melody, incorporating slightly more complexity through various techniques.
- **Playing chords with both hands:** This final exercise involves playing two-handed chords, designed to push the boundaries of coordination and technique.

On the right side you can see a figure 3.4 that shows parts of the exercises.

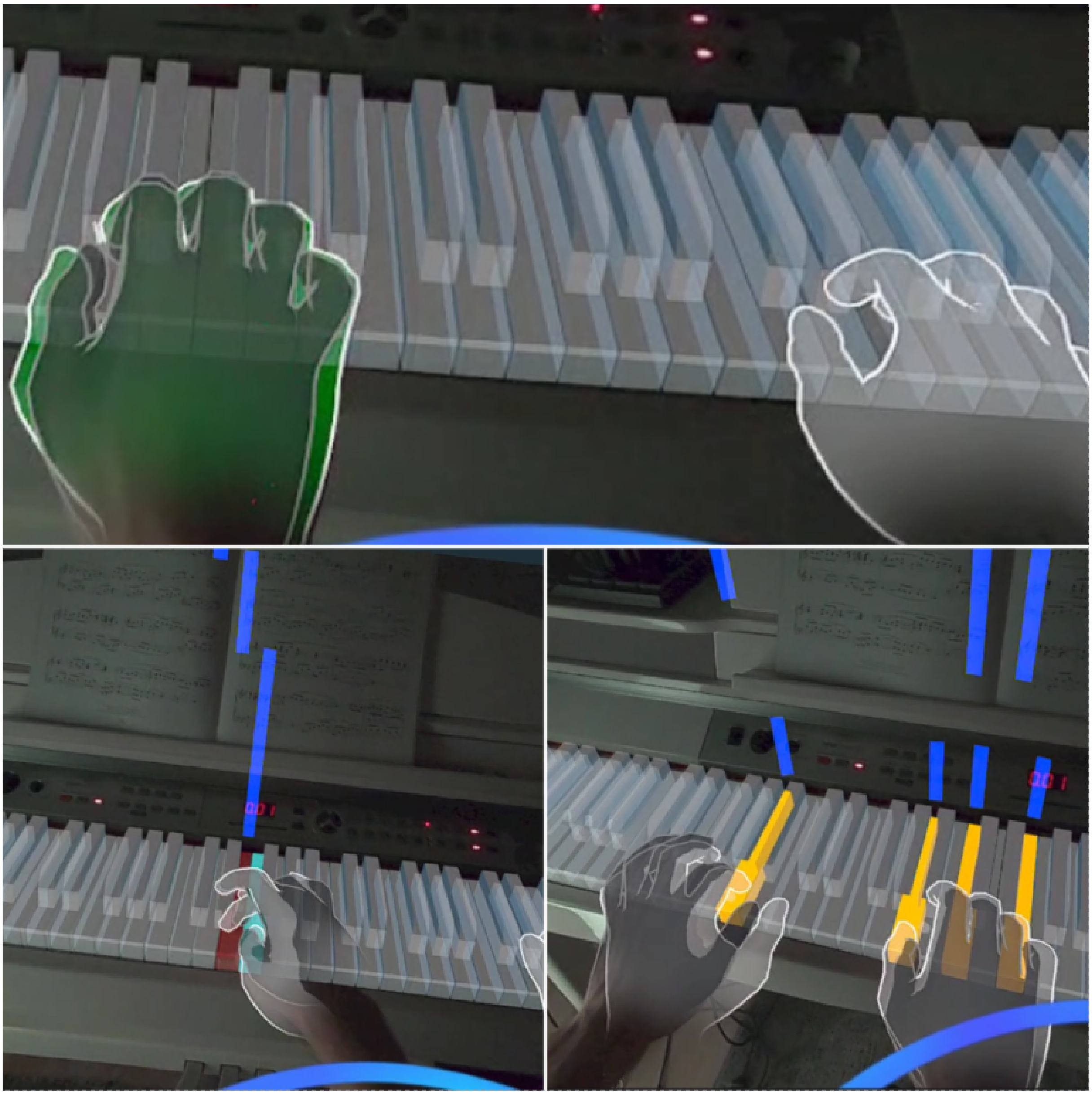

**Figure 3.4** Exercises

## 3.3 Implementation

I would now like to delve deeper into the decisions I made throughout the development process and provide an overview of the general code structure. By doing so, I aim to offer insights into how the application operates behind the scenes. This includes explaining the architectural choices, the logic behind key components, and how various parts of the code interact to create an experience for the user.

The code structure of the application can be broken down into six main parts, each playing a distinct role in ensuring the app's functionality.

- Midi Input Handling
- Midi Visualization
- Piano Setup
- Level Managing
- Ghost Hand Matching
- Study Data Collection

In the figure 3.5 you can follow the core functionality of the app and how the components and scripts interact with each other.

At the heart of the system lies the MIDI Input Handler, which processes incoming signals from the connected digital piano. These signals are then passed to the MIDI Visualizer, responsible for mapping the input data to the corresponding visual feedback on the virtual piano, ensuring real-time representation of user actions. Also it signals the Request Keys script to update the next requested key, if the right key was pressed by the user. For Calibrating the piano, the Piano Setup Manager oversees the alignment process between the physical and virtual pianos. It works closely with the Scale Object and Position Adjuster, which handle scaling, translation, and rotation adjustments to achieve the desired calibration.

User interaction is achieved through the Buttons and UI, which connect to the different levels of the app. These elements guide users through exercises while maintaining an intuitive interface. For the musical exercises, the app utilizes the Request Simple Keys and Chords scripts to generate the tasks like highlighting specific keys, triggering events, or activating virtual hands for guidance. Another critical element is the Hand Matcher and its companion, the Hand Matcher Progress system. Together, they take the user's hand positions, track joint locations, and align them with target configurations to fire an event when they detect the proper technique during exercises.

Finally, the Study Data Collector component consolidates the user performance data during the session. This script gathers the metrics, such as times and accuracy, which we will later analyze to evaluate the app's effectiveness and user interaction patterns.

Now I want to share a few insights on the decisions I took to make the experience as engaging and clear as possible. Once I finished implementing the MIDI visualizer, I had to make a decision on how to request notes from the player. I came up with two possible methods to achieve this. The first method involved showing the notes falling from above, similar to the MIDI visualizer.

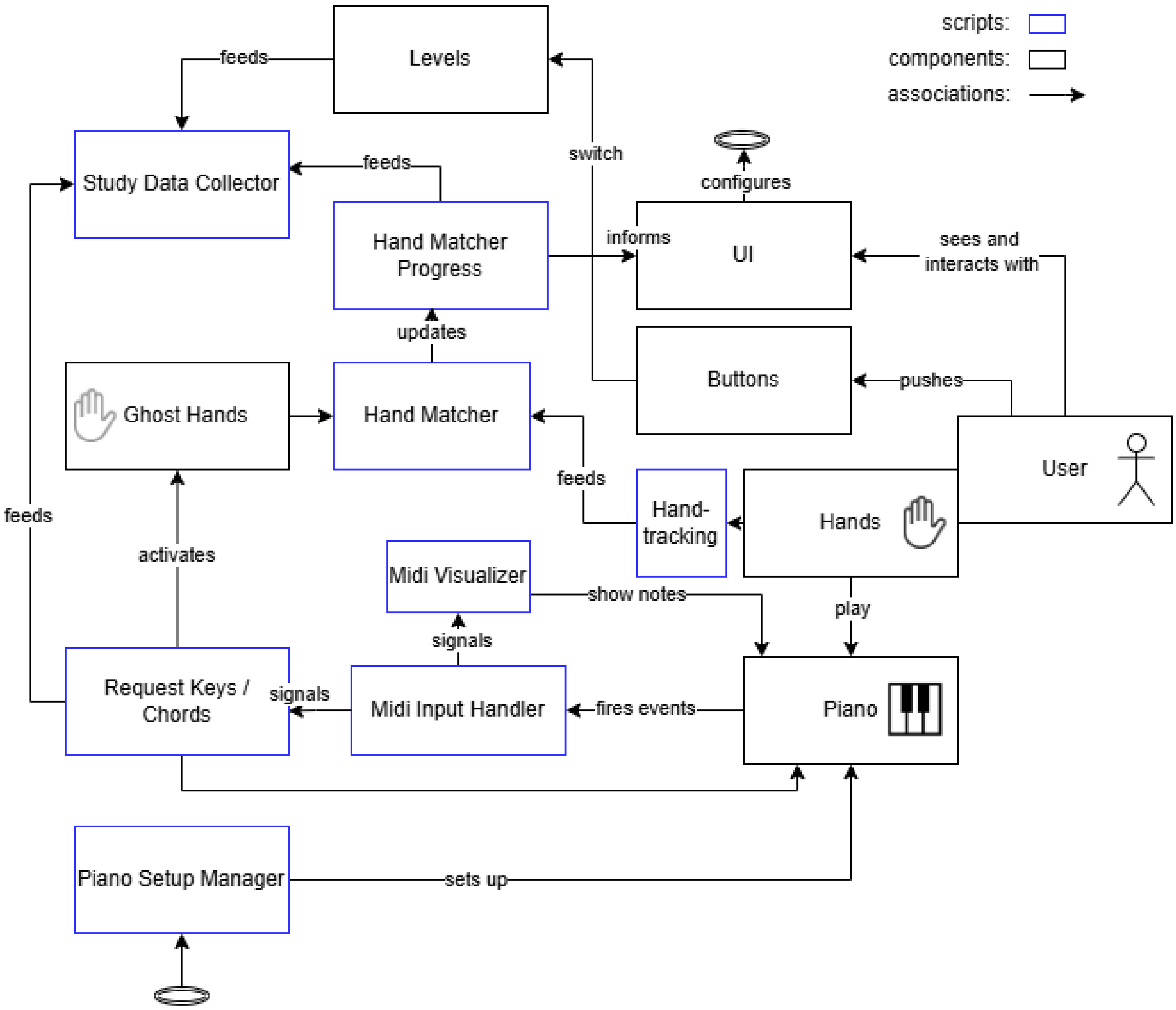


**Figure 3.5** Component Diagram

However, this approach proved to be imprecise, as my MIDI visualizer didn't clearly convey which keys were actually meant to be pressed. I ultimately discarded this idea, as I found during testing that it was difficult to accurately identify the intended keys, much like in software such as Synthesia. Instead, I decided to go with the second method: highlighting the keys directly by applying a highlight material to the key itself. While this method improves clarity by directly marking the required keys, it comes with the drawback of losing the preparation time that falling notes would provide. The time element offered by falling notes, which allows users to anticipate the next key press, is absent when simply coloring the keys. It later turned out that this wasn't an issue, as I had already planned to reduce the difficulty of the exercises. I didn't want the user to feel overwhelmed if they couldn't press the keys at the right time. My goal was to remove the stress, allowing the user to observe the situation without feeling pressured.

The next big decision I had to make was whether to display the ghost hands statically, showing a discrete hand for each note to be played, or to play a continuous animation where the user could track the movement. I decided to use both methods, but with a stronger focus on the static hands. During testing, I quickly realized that the animation wasn't precise enough to indicate exactly which finger should be used for each key. By using targeted snapshots of individual hand positions, I could be much more accurate, positioning them exactly where I wanted to show which finger position to use and how the hand should be placed.

I still kept the animation as an aid, but shifted it slightly to the right so that the user could first observe a smooth movement and then practice using the static hands. This approach also aligned well with my earlier decision to color the keys, providing a clearer, less stressful learning experience.

Additionally, I was faced with the question of how to calibrate the piano, as I didn't have access to the passthrough image for object detection. This feature wasn't supported in the current version, and there was no interface available to allow for this. Therefore, I knew I had to implement the calibration during runtime or start the app in a specific position. Since I didn't want to be location-dependent, I decided to integrate the calibration directly into the app as a playable level with clear guidance on how to achieve the calibration. This might come with a drawback, as performing the calibration manually could lead to inconsistencies, potentially distorting the data. However, the clear advantage is that it allows for usability testing, since calibration is an essential part of such applications. And because I am interested in the big picture—specifically, whether such apps are viable in the real world, given that every user needs to adapt them to their own digital piano — I have integrated calibration as a fundamental component of the application. At that point, I had to choose whether to use a rubber-band principle, where the user could grab and move the piano, or solve it with buttons. I opted for the simpler and more discrete method using buttons, in order to limit the degrees of freedom. With just one free hand movement, the accuracy would have been too low, and I didn't want the piano to wobble or shift accidentally when being grabbed. Therefore, I implemented a 2-marker method, where the user can grab the endpoints of the piano, and once done, press a button to position the piano. Afterward, there is an option to fine-tune the position via the UI. This approach is more controlled and less error-prone, even though it takes more time and is not as intuitive.

Another question is how to visually represent the Ghost Hands in terms of color. After all, a key aspect of this study is to determine whether such hands might actually be a distraction, as they could obscure the user's own hands. I decided on a simple white design, where the contours of the hand are clearly outlined, but the interior remains highly transparent with only a slight white tint, creating the ghost-like appearance. Due to the slightly darker shading of the user's real hand, the distinction remains clear, while the underlying piano keys are still visible enough. We will later evaluate this further through the questionnaire to determine whether this is actually the case. For the levels where users need to match the hands and find the correct hand position, I colored them green and red accordingly to provide feedback on whether they were positioned correctly.

Chapter 4

# Evaluation

After presenting the complete implementation of the application, this chapter focuses on the process of evaluating its effectiveness. While the previous sections introduced the app and its features, the evaluation aims to answer key questions regarding the collected data and the user study conducted.

A crucial aspect of this study is not to assess correct hand positioning in a traditional sense, but rather to analyze whether users can achieve a targeted hand position. Since I do not have the means to validate an objectively correct posture, my focus lies on how well users can match a given reference hand position and whether the visualization methods support this process effectively. And since significant amount of work needs to be done to first determine whether this method makes sense at all before evaluating whether it improves hand posture. After all, the goal is to assess whether this type of app is viable—not to present the ultimate solution, but rather to showcase a possible training method that may not have existed before.

To structure this evaluation, I divided it into two key aspects:

User Study: Explaining the rationale behind the selected metrics, the data collection process, and the specific questions this study aims to answer. Survey and Metrics: The survey includes the questionnaire I used to capture subjective user feedback, while the other metrics focus more on the objective side meaning the recorded performance data, such as timing, note accuracy, and hand similarity with the ghost hand. This chapter will outline how the study was conducted, what data was collected, and how the results contribute to understanding the usability and effectiveness of the application.

## 4.1 User Study

To gain meaningful insights into the functionality of the application, I designed a user study where participants would go through the entire app before answering specific questions. This approach ensures that they experience all relevant features firsthand, allowing for a well-rounded evaluation. A total of 17 participatns took part in the study, with each session lasting approximately 20 to 25 minutes.

Regarding the scientific design of my study, I chose a within-subjects design because I was interested in a direct comparison to address the hypothesis. I strategically decided on a fixed order for the conditions, as learning effects are likely be only contrary to my hypothesis anyway. Additionally, this approach helps prevent participants from being overwhelmed by playing without assistance and ensures a more controlled starting point. Furthermore, the study contains a confirmatory component through the hand similarity data and an exploratory component through the participants' opinions. The focus of the study is primarily qualitative, as only a small number of participants are thoroughly tested. It can be considered more of a laboratory experiment than a field study in order to create a controlled environment, but it still remains realistic.

The study follows a structured process: First, I welcome the participant, briefly explain the purpose of my research, and introduce them to what they can expect during the session. It is important to emphasize that the focus is not on playing the notes correctly—this helps minimize any nervousness or discomfort. Instead, the main goal is to observe the Ghost Hands feature and allow participants to form an opinion on its effectiveness.

After this introduction, the participants play through the app while I primarily observe, only intervening when necessary to answer questions. To better understand what the user is doing I open a stream on my laptop to see exactly what the user sees. This ensures that their interaction with the app remains as natural and unbiased as possible. Once they have completed the session, they fill out a questionnaire designed to capture their experiences and impressions. Finally, they have the opportunity to provide additional verbal feedback, sharing any thoughts or suggestions before concluding the study.

Since I was able to build the app directly onto the Quest and connect the e-piano via USB-C, I was independent of a PC, allowing for a very compact and convenient setup, as shown in the image 4.1.

To minimize distractions, the user study was conducted in a quiet environment with a closed door, ensuring consistent conditions. The study took place in the MZH building at the University of Bremen.

The goal of this user study consists of two fundamental aspects. First, it aims to evaluate whether the app itself can withstand the real-world usage by actual users. This means testing whether the app is intuitive and user-friendly enough to be operated without external assistance. In other words, I want to assess whether my design decisions were effective in creating a seamless user experience. This part of the study is crucial for determining if the app functions as a complete system that users can engage with independently. It also serves as a validation of the overall usability and design, which are critical factors in ensuring the app's success in a potential real-world setting.

**Figure 4.1** Study Setup

Second, and perhaps more importantly, the study seeks to answer the earlier defined scientific question of whether the Ghost Hand feature can truly help users achieve a targeted hand posture while playing the piano. This part of the study focuses on the effectiveness of the Ghost Hand in guiding users toward a specified hand position, which is an essential aspect of good piano technique. The question being tested here is whether this feature can serve as an effective tool for teaching users a suggested hand posture without direct intervention from a teacher. However, it is important to keep in mind that, given the current conditions I have set, it is not possible to truly teach a correct hand posture through this app, as it is still based on my own posture and not the theoretically correct hand position, which would need to be developed with the guidance of many more teachers.

By addressing these two core objectives, the study will contribute both to the practical evaluation of the app's usability and to the scientific understanding of how technology can be utilized to enhance piano learning. The results will inform whether the app can be considered a valuable educational tool in this context, while also shedding light on the broader implications of using interactive feedback systems in music education.

## 4.2 Survey

In this section, I will explain in more detail which questions I asked participants to gather important subjective data about the application. Since the focus of this study is to assess the impact of the Ghost Hand feature, I structured the questions in pairs: the first question evaluates the general experience of a feature—for example, how easy or helpful it was—while the second question specifically examines to what extent the Ghost Hand contributed to that feature.

This approach allows for a clear comparison between the Ghost Hand and the overall feature, reducing potential bias. Otherwise, a feature that is already intuitive might not benefit as much from the Ghost Hand, whereas a challenging feature might seem ineffective even if the Ghost Hand provided support. These edge cases need to be carefully considered. For instance, a participant might report that the notes to be played were generally difficult to recognize but that the Ghost Hand significantly helped in this regard. This would be an ideal case for the Ghost Hand. However, if the notes were already easy to identify, the Ghost Hand would naturally provide less added value. Therefore, simply asking whether the Ghost Hand was helpful would not be sufficient, as it would not reveal how easy the task was in the first place. In the same way, prior experience with the piano or VR applications also plays a crucial role. My hypothesis is that highly experienced players benefit less from the Ghost Hand than beginners. To explore this, I included questions about participants' prior piano experience to compare these aspects directly.

The key factors I wanted to assess include hand positioning—whether the Ghost Hand had a positive influence on it, fingering - whether the Ghost Hand helped in following a specific fingering pattern, and note visibility, examining whether the hand obscured the notes or kept them clearly visible. Additionally, since the app is designed to be used both with and without the Ghost Hand, I aimed to analyze this comparison by asking participants how well they managed without the Ghost Hand. I address this also by asking whether the participant actively paid attention to the Ghost Hand in the first place.

Another important aspect is motivation and enjoyment. My hypothesis is that the app introduces a motivating factor through gamification, distinguishing it from traditional learning methods and presenting a potential innovation in this area.

Since the study also evaluates the application as a whole, usability cannot be overlooked. To assess this aspect, I incorporated a standardized test from the literature: the System Usability Scale (SUS) by Brooke. This scale is designed to provide a general measure of how usable an application is, regardless of its specific purpose. Its strength lies in offering a comparative usability score that applies across a wide range of systems, making it globally relevant [Bro+96]. I included this test because the focus is not solely on the Ghost Hand feature but on this type of piano learning app as a concept. While my implementation serves as one possible example, future applications may take different forms while pursuing the same goal. Therefore, conducting proper usability research is essential to understand the broader potential of such systems.

## Questionnaire

Thank you for participating in this user study! Your feedback is invaluable and will help me improve the app. Please answer the following questions honestly. There are no right or wrong answers.

## General Questions About You

1. Age:________________
2. Gender:
   - ☐ male
   - ☐ female
   - ☐ non-binary
   - ☐ Prefer not to say
3. What is the current date and time? ________________
4. How would you rate your experience with playing the piano?
   - ☐ Never touched a piano before
   - ☐ Beginner
   - ☐ Advanced Beginner
   - ☐ Intermediate
   - ☐ Professional
5. Have you ever used a Mixed Reality or VR application before?
   - ☐ Yes
   - ☐ No

## Usability

Left = **Strongly Disagree**, Right = **Strongly Agree**.

| **Statement** | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| I think that I would like to use this system frequently. | ○ | ○ | ○ | ○ | ○ |
| I found the system unnecessarily complex. | ○ | ○ | ○ | ○ | ○ |
| I thought the system was easy to use. | ○ | ○ | ○ | ○ | ○ |
| I think that I would need the support of a technical person to be able to use this system. | ○ | ○ | ○ | ○ | ○ |
| I found the various functions in this system were well integrated. | ○ | ○ | ○ | ○ | ○ |
| I thought there was too much inconsistency in this system. | ○ | ○ | ○ | ○ | ○ |
| I would imagine that most people would learn to use this system very quickly. | ○ | ○ | ○ | ○ | ○ |
| I found the system very cumbersome to use. | ○ | ○ | ○ | ○ | ○ |
| I felt very confident using the system. | ○ | ○ | ○ | ○ | ○ |
| I needed to learn a lot of things before I could get going with this system. | ○ | ○ | ○ | ○ | ○ |

## Questions About the Ghost Hand Feature of the App

6. How easy was it to match the Ghost Hands?

   ☐ Very Easy
   ☐ Easy
   ☐ difficult
   ☐ Very difficult

7. How helpful were the Ghost Hands in showing you the correct hand posture?

   ☐ Very helpful
   ☐ Helpful
   ☐ Not very helpful
   ☐ Not helpful at all

8. How much attention did you pay to the ghost hands?

- ☐ Very much
- ☐ Some
- ☐ Not very much
- ☐ Not at all

9. How did you manage to play the piano without the ghost hands?

- ☐ I managed pretty well.
- ☐ I kind of managed.
- ☐ I didn't manage well.
- ☐ I got stuck.

10. How easy was it to see which notes to play in general?

- ☐ Very easy
- ☐ Easy
- ☐ Not very easy
- ☐ Not easy at all

11. Was the ghost hand helpful in knowing which notes to play?

- ☐ Very helpful
- ☐ Helpful
- ☐ Not very helpful
- ☐ Not helpful at all

12. How easy was it to see which finger to use for each note in general?

- ☐ Very easy
- ☐ Easy
- ☐ Not very easy
- ☐ Not easy at all

13. Was the ghost hand helpful in knowing which finger to use for each note?

- ☐ Very helpful
- ☐ Helpful
- ☐ Not very helpful
- ☐ Not helpful at all

14. Do you feel that the Ghost Hands had a positive impact on your hand positioning?

- ☐ Yes, significantly
- ☐ Yes, somewhat
- ☐ Barely
- ☐ Not at all

## General Questions about JazzhandsXR

15. How motivating did you find using the app?

- ☐ Very motivating
- ☐ Motivating
- ☐ Not very motivating
- ☐ Not motivating at all

16. Can you imagine this type of app to be used for piano hand posture training in the future?

- ☐ Yes, definitely
- ☐ Yes, kind of
- ☐ No, not really
- ☐ No, not at all.

17. Would you prefer this type of app for piano hand posture training over other methods like books or videos?

- ☐ Yes, definitely
- ☐ Yes, kind of
- ☐ No, not really
- ☐ No, not at all.

18. Do you think such an app can replace piano training in person?

- ☐ Yes, definitely
- ☐ Yes, kind of
- ☐ No, not really
- ☐ No, not at all.

19. Was it fun?

- ☐ Yes, definitely
- ☐ Yes, kind of
- ☐ No, not really
- ☐ No, not at all.

Thanks! That's it. I wish you all the best. Dmitry

## 4.3 Metrics

While subjective feedback provides valuable insights, it is not sufficient on its own to make a well-founded statement about whether users can actually follow the Ghost Hand, thereby influencing their hand posture. To objectively assess this, I developed a metric to measure how closely a user's hand movements align with the Ghost Hand.

For every note played, the current position of all hand joints is recorded and compared to the 17 joints of the Ghost Hand using a matching algorithm. If the distance between corresponding joints falls below a defined threshold (approximately 2 cm), the joint is considered matched. The Hand Similarity metric is then calculated as the ratio of matched joints to the total number of joints 4.2.

At the beginning of the development process, I collected metrics on the number of incorrectly played notes per level and the time spent in the app, which ranged from about 7 to 12 minutes. However, I do not use these metrics for conclusions, as my goal was to remove time pressure and give the participants the chance to focus on quality rather than speed. Therefore, the data on hand similarities is more meaningful. In total, I had the participants play four pieces in the second half of the app: a *C major scale*, *Für Elise* with and without the Ghost Hand, and *Let It Be* as a more challenging piece. The scale consists of 15 notes, which alone gives 255 data points ($17 \cdot 15$). *Für Elise* has 35 notes, totaling 595 data points for each of the two runs. The Beatles piece has 31 notes, as I treat chords as a single note and only record the right hand, which adds up to 527 data points. In total, this gives us 1,972 numbers, which forms the objective data foundation for this study.

**Figure 4.2** Hand Joints

## 4.4 Statistics

Now that we've covered how the study was set up, it's time to take a deeper look at the numbers. In this section, I present the raw statistical results of the user study descriptively. I will summarize participant responses, and analyze the hand similarity data before moving on to the actual discussion. Let's start with the most straightforward data, which I'll simply list, as it is highly uniform and shows no variation. Because of this homogeneity, there is no need for statistical calculations, as every participant shared the same opinion.

The study included 17 participants, 12 of whom were male and 5 female. Among them, 6 had never touched a piano before, 5 were beginners, 3 were advanced beginners, and 3 were at an intermediate level. The vast majority—88.2% had prior experience with mixed or virtual reality. The participants' ages ranged evenly from 22 to 32 years.

The questionnaire included questions with a clear majority response. For example, in Question 6, all but one participant stated that matching the Ghost Hands was at least "easy". Additionally, all 17 participants agreed in Question 7 that the Ghost Hand was at least "helpful" in demonstrating the correct hand posture.

In Question 14, there was once again unanimous agreement: every participant indicated that the Ghost Hand had at least a somewhat positive impact on their hand posture. Furthermore, 23.5% of participants even reported a significant difference. A similar trend can be observed in Questions 15, 16 and 19, where participants indicated that they could at least kind of imagine such an app being used for hand posture training in the future and that they found the experience to be motivating and fun. In fact, 70% of participants even responded with "Yes, definitely".

There were a few questions that showed ambivalence and clearly divided the group, but there were two questions in particular that did this. Question 12, which asked whether it was easy to determine which finger to use for the notes in general, not related to the Ghost Hand, was answered with a majority of 58.8% choosing "not very easy" and 41.2% saying "easy."Additionally, two-thirds of the participants stated that such an app could not replace training with a real person, while one-third of the respondents somewhat agreed with the idea.

Now, let's take a look at the SUS score. The System Usability Scale score for the app is 77.8, with a standard deviation of 9.6. The corresponding 95% confidence interval ranges from 75.5 to 80.1. This suggests that overall usability was rated quite positively, with relatively low variance among participants. A SUS score above 68 is generally considered above average, meaning the app performed well in terms of user experience.

Next, we'll focus on the analysis of the paired questions. For this, I prepared the questions 7, 9, 10, 11, 12, and 13. In each pair, the first question is always a general one, focusing on a feature that can be answered without including the Ghost Hand. For example, question 10 asks how well participants could recognize which notes to play in general. The corresponding paired question 11 then asks how much the Ghost Hand helped in recognizing which notes to play. This way, I can isolate whether the Ghost Hand was perceived as more helpful compared to the same task without it. The first question pair focuses on hand posture—how well participants managed without the Ghost Hand versus how much the Ghost Hand helped them adopt the correct posture. The second pair compares playing the correct notes with and without the Ghost Hand. The third pair examines how well participants could determine the correct finger for each key, both with and without the Ghost Hand.

I should mention that I transformed each response into a numerical scale from 1 to 4 to calculate the mean values for each question. For example, when asked how easy it was to determine the correct fingering without the Ghost Hand, participants gave an average score of 2.35, which corresponds to "not very easy." However, with the Ghost Hand, the average score increased to 3.14, meaning "easy". I have compiled these results into the following table, which also includes the difference between the scores to highlight how much each feature changes when the Ghost Hand is introduced.

| | Without Ghost Hand | With Ghost Hand | Difference |
|---|---|---|---|
| Playing with correct posture (Q9,Q7) | 2.93 | 3.21 | +0.29 |
| Finding the correct notes (Q10,Q11) | 2.73 | 3.00 | +0.27 |
| Choosing the fingering (Q12,Q13) | 2.36 | 3.14 | +0.79 |

**Table 4.1** Comparison of mean ratings with and without the Ghost Hand.

Now, I will analyze the hand similarity data, gather the key statistics for all four pieces, and conduct a one-sided paired t-test for *Für Elise* to determine whether playing with the Ghost Hand leads to a significant difference in adopting a specific hand posture, ruling out randomness as a factor. Additionally, I will use Cohen's d to measure the effect size of this comparison, providing insight into whether the difference is actually noticeable[Coh92].

My approach is as follows: first, I calculate the mean hand similarity per user for every piece. Then, I use these values to compute the overall mean and standard deviation for each piece, allowing for a comprehensive statistical evaluation. I have only included the key statistics in the following table, as the complete dataset for each note is too large to present. Additionally, for *Für Elise*, I have also included the results from the t-test. You can see for example that user 1 played the scale with 0,898 accuracy meaning getting roughly 90% of the joints matched during the playthrough and it has a standard deviation of approximately 20%. The users achieved an accuracy of around 50% for the scale, meaning they matched, on average, half of the joints per played note. For the other pieces, less was matched. For *Für Elise* with the ghost hand, the accuracy was 43%, while without the ghost hand it was only 31%. The p-value is significantly below 0.05, indicating a statistically significant difference between the means with at least a 95% confidence interval. The effect size is 0.43, which suggests a small but noticeable effect, as it falls between 0.2 and 0.5.

**Table 4.2** Hand Similarity Statistics per User and Overall Analysis

| Index | Scale $\mu$ | $\sigma$ | Elise (GH) $\mu$ | $\sigma$ | Elise (No GH) $\mu$ | $\sigma$ | Beatles $\mu$ | $\sigma$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 0,898 | 0,194 | 0,691 | 0,345 | 0,679 | 0,344 | 0,899 | 0,212 |
| 2 | 0,804 | 0,273 | 0,514 | 0,304 | 0,476 | 0,311 | 0,704 | 0,331 |
| 3 | 0,267 | 0,212 | 0,097 | 0,115 | 0,106 | 0,131 | 0,040 | 0,092 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 17 | 0,515 | 0,335 | 0,388 | 0,319 | 0,232 | 0,277 | 0,710 | 0,318 |
| **Overall** | **0,508** | **0,301** | **0,432** | **0,310** | **0,311** | **0,256** | **0,415** | **0,232** |
| **t-Test** | - | | **p-value: 0,000398** | | **Cohen's d: 0,429** | | - | |

Chapter 5

# Discussion

In this chapter, we will move beyond the purely descriptive presentation of the data from the previous chapter and begin to interpret the results. By critically examining the findings, we aim to identify the strengths and limitations of the study, address potential weaknesses, and discuss the broader implications of the results. This discussion will help provide a deeper understanding of the effectiveness of the app, its impact on users, and areas for further improvement and development. Furthermore, we will critically assess whether such technologies have a viable future or whether there are cases where their use may not be beneficial. Where appropriate, I will integrate relevant research to support the discussion.

## 5.1 Results

To begin, we focus on the core question of this study: whether this app can effectively support users in adopting a targeted hand posture. The evaluation results strongly indicate that it can. First, the questionnaire responses suggest that the Ghost Hand was perceived as helpful, improving posture, note recognition, and fingering (see Table 4.1). Second, the objective measurement from *Für Elise* shows an improvement in Hand Similarity from 0.311 to 0.432 accuracy. Additionally, the conducted t-test supports this result, indicating a statistically significant difference in hand posture accuracy when using the Ghost Hand.

However, there are strong points that speak against its effectiveness. While users may perceive the Ghost Hand as helpful, they might not be able to accurately judge whether it truly benefits their playing in the long run. A more in-depth engagement with a piano teacher would be necessary to assess whether what they are practicing is actually correct and aligns with proper technique as a teacher would recommend.

Additionally, while the improvement in hand similarity is statistically significant, it is still far from representing a fundamental change. During the user study, I observed that participants' hands, despite being instructed to stay relaxed, tended to become stiff as they tried to replicate a specific pose.

This suggests that the guidance might be misinterpreted, leading to unnatural and rigid movements instead of fluid and relaxed playing. Ironically, this is the opposite of what the Ghost Hand was intended to achieve. Above all, proper hand posture must be conveyed in an adequate and convincing manner. The key question is whether users can perceive the difference well enough to later reproduce it without the app and integrate it into their playing to self-correct. This is not fully ensured with the app, as users still need to trust their own judgment when determining whether their hand posture is incorrect just because the app indicates otherwise. For example, in Question 8, 30% of users stated that they did not pay much attention to the Ghost Hand, despite the instruction that hand posture—not the notes—was the primary focus of the study. This can occur when users are too occupied with playing the notes, making it difficult to focus on their hand positioning. Additionally, the app may not have provided a strong enough visual cue, such as requiring users to adopt the correct hand posture before proceeding, which could have reinforced the importance of proper positioning more effectively. In addition, the guidance of a teacher would be far more reliable and memorable, as their direct feedback carries more authority and can be tailored to individual needs.

We should also critically examine whether this study was appropriately designed and whether its structure was effective in achieving the intended results. A strong argument in favor of the study is that it was conducted under highly controlled conditions—each participant had nearly identical circumstances, the environment was free from distractions, and the study sessions proceeded smoothly without interruptions. Additionally, data collection was fully automated, requiring no manual intervention, and participants were able to complete the app experience independently. However, there are a few factors that could have influenced the results. First, all participants played through the entire app, meaning that the key segment of data collection was not entirely isolated. Since each user performed their own calibration, a poor calibration could have led to inaccurate results. Nevertheless, I ensured that every participant was satisfied with their setup before proceeding, minimizing this risk.

A second potential issue is the varied experience levels of the participants. While diversity in skill level can be valuable, the study aimed to target a specific user group. Ideally, a user study would be conducted with participants who fit the intended target audience. In this case, the study was conducted with volunteers who were willing to participate, but they may not necessarily be individuals actively seeking to improve their hand posture. Consequently, some participants might not have been fully engaged in refining their technique, which could have affected the outcomes.

Let's discuss whether the self-developed metric is sufficient to answer the study's research questions. One strong argument in favor of the metric is that the Quest's hand tracking is relatively stable and considers the entire hand. This is an advantage compared to evaluating only the fingering, for example, which alone does not provide much insight into the movement of the wrist. The tracking system captures joint positions well, ensuring that different areas of the hand are equally accounted for. However, there are also some limitations to consider. One major issue is that the hand tracking is not robust against occlusions. This is particularly problematic in piano playing, where fingers frequently move under the hand or switch positions quickly.

When occlusions occur, the accuracy of the hand similarity metric may decrease significantly. A potentially better approach would be to use physical tracking markers, such as a specialized glove, ensuring precise joint position tracking while minimizing discomfort for the user.

Another drawback of the current system is its dependency on the viewing angle. The hand tracking relies on the user's hands being visible from the right perspective—if the user tilts their head slightly, tracking accuracy may be affected. In my own tests, which were not part of the study, I achieved an accuracy of around 90%. This suggests that users who closely follow the ghost hand's positioning can achieve high similarity scores. However, individual hand size differences also play a significant role, which the current metric does not account for explicitly.

Another strong argument in favor of the app is its high usability. With a SUS score of 77.8, it can be considered highly usable, suggesting that such a system could work well in real-world scenarios. The clear instructional texts and the simple red button interface likely contributed to this positive usability rating. Additionally, the warm-up tests conducted before the main tasks helped users familiarize themselves with the mixed reality environment, reducing potential usability issues.

However, one drawback of this type of app is that users must first learn how to use it before they can effectively practice their hand posture. The initial setup, including calibrating the piano, requires time and effort—time that could otherwise be spent directly improving hand positioning. This learning curve could be a barrier, particularly for users who are less experienced with technology or who seek immediate practice without additional setup steps.

After thoroughly engaging with the participants, I received valuable feedback on areas that could be improved and potential reasons why the experience might not have been optimal. One notable issue was that users had difficulty seeing their own hands clearly, as they were rendered with a gray shadow. This sometimes led to a cluttered visual experience where the virtual piano, the real piano, the user's own hand, and the ghost hand were all stacked on top of each other, making it harder to focus. Despite this, the concept of the ghost hand was well received. Its transparency level was perceived as just right, helping users understand that it was something they should try to mimic or align their hand with. This is an innovative feature that a real teacher cannot provide. Additionally, some participants mentioned that it was not always clear which finger the ghost hand was using to press a key. A possible improvement could be color-coding the fingers or fading out the inactive ones to make the intended fingering more visible.

The results of the user study indicate that the app concept is highly promising, with the study offering valuable insights into its effectiveness. Based on the data collected, it can be confidently said that the app is capable of facilitating the learning of a targeted hand posture. While the results show positive trends, there are still several areas where adjustments and improvements could be made to further enhance its usefulness and impact. These refinements will be crucial for unlocking the app's full potential in supporting hand posture development. Therefore, in the following chapter, I will delve into the practical benefits of such apps, considering how they can contribute to the skill and whether they can serve as a valuable tool in real scenarios.

## 5.2 Practical Applications

Even if the study provides evidence of a potential benefit, that doesn't necessarily mean such an app can be practically applied. Several other factors need to be considered, and I would like to touch on a few of them here. Let's assume there is indeed an app of this type, perhaps a more advanced version. The next challenge would be figuring out how to integrate it into music lessons effectively.

One cannot ignore the cost of VR headsets and their limited usability. Currently, these headsets tend to be heavy, and it's often difficult to use them for extended periods without experiencing motion sickness or fatigue. Additionally, for such an app to work seamlessly, the piano must also be equipped in a way that allows it to provide data, which adds another layer of complexity. Otherwise the headset has to rely on visuals and haptic data from the hand.

Ideally, the app would be independent of any special equipment and would use some type of image recognition or haptic data caputure to analyze hand positions and provide valuable tips without requiring a specific way of playing or a predefined piece to be played. The app could assess whether the hand posture is correct just by recognizing visual cues regardless of the piece being played. This would be a significant challenge, one that likely lies in the future of such technology.

One area where I see a very strong benefit already is in gamification. Such apps are fun to use and motivating, encouraging users to spend time engaging with them. This is also supported by my user study, as all participants indicated that they enjoyed using the app. The element of fun is crucial for maintaining long-term engagement and encouraging consistent practice, which could make a significant difference in how users approach learning hand posture and improving their skills. Such an app can also be beneficial in covering other aspects of learning in a more indirect way, such as guiding users on how to practice correctly. Often, people are left to practice on their own, and it can easily turn into just playing without following the teacher's recommendations. An app like this could serve as an additional interface to the teacher, helping to remind users of effective practice techniques and encouraging them to focus on the right things. These kinds of apps could also provide a structuring benefit by allowing users to track their progress and document what they have already practiced. The app could retain this information and provide personalized suggestions based on what the user needs or what their teacher has assigned. I see great potential in such apps, as they could cover multiple aspects of practice simultaneously and combine them in a mixed reality experience, making the user's piano progress more tangible and easier to monitor.

If I had to assess the practicality of this type of app for exactly this current time, I would phrase it roughly as follows: realistically, I see this app more as an interesting alternative to try out and expand one's perspective on the piano, but not yet as a robust learning app integrated into one's personal learning process.

## 5.3 Comparison to Presented Tools

In this section, I will compare the app presented in this study to other existing apps and research papers discussed in the Related Work section. The comparison will focus on key aspects such as functionality, user experience, and the effectiveness of hand posture training in music education. By examining these points, I aim to place the presented app within the broader context of existing tools, highlighting both its strengths and areas for improvement.

The application that comes closest to mine is, as I have already presented, VRMonic. One strength that differentiates my app is the successful use of mixed reality. In my app, you don't sit in a completely virtual environment, losing the sense of reality, but instead, you sit at your familiar piano, which could certainly have an impact on user persistence. Nevertheless, VRMonic, for example, has this playback feature that allows users to watch recordings, while I focused on fostering interaction with the app. It is worth mentioning, however, that they created the Oracle hands based on a series of expert data in order to provide a good approximation of the theoretically optimal hand position. However, this results in the hands not appearing detailed enough [Mat+24].

As we discussed earlier, the goal was to determine whether practicing in a musical context is feasible, given the potential inconsistencies in rendering or latency that could arise with such technology. For this app, at least, it turns out that there were no significant inconsistencies, and the latency was essentially negligible. One thing that was noticeable was a slight delay between the passthrough and the rendered hand, but this did not pose a significant obstacle, and the usability was not greatly affected. It's more about the interaction using hand tracking, which still isn't entirely intuitive, as even gestures like gripping can sometimes be misinterpreted. Interactions rely on specific gestures, such as the pinch gesture, so not every command operates flawlessly.

When comparing the app to Pianoverse, I notice that they have made a different choice regarding joint positions [Lab+23]. One strong point, however, is that they also include the arms and shoulders to show how the expert moves with their body. This is a good evolution of the idea, as it now incorporates the entire body. In contrast, I focused solely on the hand. Personally, I find the ghost hands more illustrative than the rendered lines in Pianoverse. I suspect that it's more intuitive to follow a real hand, even if parts of it are covered, than to follow a line. In any case, incorporating the rest of the body is a good approach.

On the other hand, we have Hand-by-Hand Mentor, which attempts to generate highly realistic hands [Guo+21]. These are based on a hidden Markov model, so they aren't recordings, but rather generated solutions for hand movements. I see a significant strength in this approach, as it allows for the analytical calculation of the movements one must perform, which aids in classifying movement patterns and practicing them. However, a generated solution is not as natural as a recording, as it only attempts to approximate real human movements. One of the advantages of a recorded model is that it captures the true, organic nuances of human motion, which might make it feel more authentic and relatable to the user.

The challenge with a generated model is that it could potentially lack those subtleties, making it harder for users to connect with the movement or feel as though they are mirroring a real person.

I believe I have found a middle ground by displaying the hands statically. This way, users are not overwhelmed by fast animations and can take the time they need to focus on the hand position. However, this approach does lose some of the natural fluidity that animations provide, which I use when demonstrating how to assume the correct hand posture. The advantage of static hands is that they allow for a clearer and more deliberate focus on the details of hand positioning, while animations are better for conveying movement and flow. Balancing both approaches could be key in creating an experience that is both intuitive and informative.

In summary, I believe the apps have focused on solving problems related to hand positioning in specific passages of pieces or exercises, always emphasizing the need for feedback during playing. However, what sets my app apart is the inclusion of dedicated isolated exercises specifically for hand posture that don't even require playing. For instance, recognizing incorrect hand positions or finding the correct posture by watching the animation. I feel that this aspect is somewhat underrepresented in other apps, as the focus tends to be on playing scenarios. After all, it is about learning principles and understanding why the hand should be positioned in a certain way so that these principles can be applied correctly in any context, not just in a particular scenario.

Chapter 6

# Conclusion

In this thesis, an application was developed and a user study was conducted to explore the effectiveness of using a mixed reality app with animated hands to teach hand posture for piano playing. The study aimed to evaluate the usability of the app and to investigate whether the 'Ghost Hand' could serve as a tool to guide users towards a more purposeful hand positioning during piano practice. The findings suggest that while the app is not yet a comprehensive tool for teaching theoretically correct hand posture, it has great potential for further development, showing strong statistical evidence that the Ghost Hand can lead to improved accuracy in practicing a targeted hand posture. The foundation for such an app has been established, and with further refinement, it could become a valuable resource for piano learners. The gamified approach offers a compelling alternative to traditional methods and provides an engaging, user-friendly experience that motivates users to continue their practice.

While the method is not perfect, it certainly holds value as an alternative and has a rightful place in the landscape of piano learning tools. Further research and development are needed to improve the accuracy of the hand posture guidance, but the positive reception from users highlights the app's potential as a motivating and accessible tool for aspiring pianists.

## 6.1 Summary of Findings

Playing the piano is a demanding task that requires a high level of precision and a clear understanding of how to position and move one's hands—one cannot simply play by intuition. Incorrect playing techniques can lead to piano-related musculoskeletal disorders, with studies showing that approximately 40% of pianists suffer from injuries [AA10]. Moreover, these issues can often be traced back to poor hand posture. The fundamental flaws in hand positioning can be clearly categorized, such as a collapsed wrist or a "flying five," where the fifth finger is raised unnecessarily. These improper techniques create tension in the hand, which hinders smooth and relaxed playing.

To address this problem, we formulated the hypothesis that a mixed reality app could help users practice a targeted hand posture. However, solving the problem is not just about finding a direct technical solution. Such an app must fit into the broader picture of learning strategies—it needs to be usable so that users can effectively engage with it without being hindered by usability barriers. It must also be motivating to ensure that learning is not negatively impacted by frustration. Lastly, the app must be practically applicable, meaning the setup should require an acceptable level of effort, because no matter how perfect a solution is, it will fail if it is not accessible.

We examined the state of the art and analyzed existing applications. We found that there is a wealth of research on this topic and that certain apps already target specific issues. For example, some use machine learning to detect and classify poor hand posture based on images, while others rely on VR applications to generate optimal hand movements for specific passages. A significant amount of research has been conducted on how pianists move their hands, utilizing tracking technologies to collect data and provide feedback. However, we found that there are still few applications that use overlaid models specifically to train hand posture, drawing the attention away from playing the correct notes. Therefore, I aimed to develop a method to create exercises that address this issue and make them more intuitive through the ghost hand.

To develop such an app, we had to explore the capabilities of modern technologies and their potential limitations. I learned that with the relatively new color passthrough feature of the Quest 3, it is possible to build powerful mixed reality applications that blend the real world with virtual objects, allowing learning to go beyond traditional boundaries. However, there are certain limitations in terms of accuracy. It is not yet very reliable to test detailed fingerings or analyze movements at the finger level, as the hand tracking through the cameras is not capable of detecting the smallest details or handling occlusion optimally. Therefore, the focus was placed more on the hand as a whole rather than on fine-grained interaction with individual notes.

After studying the Taubman Approach and the concepts of body mapping, it became clear that a method was needed to manage tension—ensuring that playing remains as relaxed as possible and that the hand structure adapts to each played note in a way that aligns optimally with anatomy. This allows for the minimal necessary effort while still producing the same sound quality. As a result, the exercises focused on recognizing poor hand posture and shaking out tension, as well as carefully positioning the hands for each note. By removing time and rhythm from the equation, players could fully concentrate on practicing proper hand positioning.

To support these ideas with data, I designed the user study to provide a clear comparison by evaluating hand similarities between playing with and without the Ghost Hand. Additionally, a questionnaire was included to better assess the app's qualities and gather detailed user feedback on how the new feature is perceived and whether it effectively serves its purpose. The results showed statistical evidence that the Ghost Hand positively influences hand posture toward the given target posture and that the app is easy to use, helpful, and enjoyable. However, it is important to note that this evidence does not yet conclusively prove that the hand posture will improve in the long term or that this method provides a robust learning approach. Overall, this provides a solid foundation for further research. Future studies could focus on generating a theoretically optimal hand posture based on expert data and applying this methodology to refine the approach. This could eventually lead to the development of a tool that not only assists in hand posture training but also ensures that players adopt better techniques long term.

## 6.2 Limitations

Here, I would like to discuss the limitations of this work. Although much has been accomplished, it is important to acknowledge the resources available and what might have been left out, but could have been crucial.

Firstly, as mentioned earlier, the Quest 3 was used, which means there was little opportunity to implement image recognition or process the passthrough image in a cheap and easy way. I can imagine that other approaches, such as physical tracking or better hardware with additional cameras from different angles, could have led to improved results.

Additionally, I can imagine a larger scope for the user study. The target group was somewhat off, and it would have been beneficial to include people who are actually focused on improving their hand posture. I can also envision a scenario where the study is conducted with one person over an extended period. This way, we could observe the improvement in hand posture more precisely and have concrete evidence that it truly makes a difference. Ideally, the study would involve two comparison groups: one group learns hand posture without the app, using other methods like books or videos, while the other group learns with the app.

Although I was able to achieve good results with my method, the metrics used are somewhat under-tested. The hand similarities could be somewhat skewed and likely don't fully reflect reality, as they only focus on aspects like joint position. However, the main element we're really concerned with, tension, was overlooked. To address this, other measurement methods would be needed to capture the tension in the hand and convey it to the user, so they can identify weaknesses in their playing.

Another important point is the potential for overfitting to the ghost hands. While it's helpful to show that the hand posture is similar to the ghost hand, whether this truly represents a good hand position in principle is questionable, even when its carefully chosen. What if there are multiple valid ways to play a passage, and the provided hand posture is just one of many possibilities? Therefore, there should be a way to better focus on the principles, for example, by testing the angles between the finger joints, ensuring the relative form is controlled instead of just the absolute position.

## 6.3 Further Development

During the user study, I received a lot of feedback that led me to new ideas for further improving the app. This feedback can also serve as a basis for further research opportunities, allowing this concept to be explored and refined.

One potential improvement would be to incorporate finger positions by highlighting exactly the fingers that need to be used to play a particular note. While this differs from traditional fingering with numbers, it opens up room for investigation into whether a system of glowing fingers or color coding leads to a faster understanding of fingering compared to using numbers.

Additionally, parts of the hand could be emphasized or hidden to focus attention on the key areas. For example, if the majority of the hand positioning is correct but the little finger is consistently raised, it could be highlighted in pulsating red to draw attention to the issue. Another idea could be designing a model that shows the direction the hand should move to be in the optimal position, perhaps using particles that fly in the right direction, attracting the hand like a magnet.

If the idea of hand posture exercises is further developed, it would be beneficial to focus specifically on the wrist, visualizing and analyzing its movement separately. Since wrist rotation plays a crucial role in piano playing, it would be helpful to visualize how the wrist rotates and whether, for example, circular motions are executed correctly. This targeted visualization could allow users to become more aware of their wrist movement and make adjustments to ensure proper technique.

If we take the idea of mixed reality seriously, we could make full use of the virtual world by adapting the playing environment. I could envision a scenario where, for example, the bending of the fingers for proper hand posture is visualized using metaphors. Imagine trying to explain to a child that they shouldn't keep their fingers flat while playing because they are not strong enough to support the weight of the hand. In this case, we could virtually display a model with bridges and weights, using interactive physics to playfully demonstrate why the fingers should be held in a certain way. This approach would not only engage the child but also provide a clear and understandable representation of the concept.

Or let's say we need to practice different techniques that are too many to remember. In that case, we can use strong visual associations to make the learning more memorable. For instance, imagine you're playing a section of a piece that requires a flowing legato technique. At that point, we could display waves of water to signal that this section requires smoother, more connected playing. This visual cue would serve as a reminder to focus on the fluidity of the movement, making the technique easier to internalize by associating it with a natural and recognizable metaphor.

With this openness towards mixed reality technologies, I would like to conclude this work. I see a bright future where these technologies are not used to limit human creativity but to expand it. The potential of mixed reality goes far beyond merely solving existing problems. It offers a unique opportunity to enhance the way we learn, create, and experience the world around us. By allowing us to interact with both the real and virtual worlds seamlessly, we can break down barriers that once constrained us. By visualizing complex concepts in new, immersive ways, we can gain deeper insights and foster innovation. The key is to embrace the endless possibilities and think beyond traditional boundaries, blending technology with the boundless creativity of the human mind. In such a future, we are not simply using technology to replicate what we already know, but to imagine and create new realms of possibilities that were once unimaginable.

Dmitry

Appendix A

# Appendix

## A.1 List of Figures



## A.2 List of Tables



## A.3 Annex

- references
- demo_repo_link.txt
- hand_similarities.xlsx
- survey.xlsx
- sus_score.xlsx
- Thesis.pdf

## A.4 Bibliography

[AA10] Lili Allsop and Tim Ackland. "The prevalence of playing-related musculoskeletal disorders in relation to piano players' playing techniques and practising strategies". In: *Music Performance Research* 3.1 (2010), pp. 61–88.

[Amm+24] Vivienne Amm, Krishnan Chandran, Lars Engeln, and Matthew McGinity. "Mixed reality strategies for piano education". In: *Frontiers in Virtual Reality* 5 (2024). ISSN: 2673-4192. DOI: *10.3389/frvir.2024.1397154*. URL: *https://www.frontiersin.org/journals/virtual-reality/articles/10.3389/frvir.2024.1397154*.

[Ban+23] Mariano Banquiero, Gracia Valdeolivas, Sergio Trincado, Natasha García, and M.-Carmen Juan. "Passthrough Mixed Reality With Oculus Quest 2: A Case Study on Learning Piano". In: *IEEE MultiMedia* 30.2 (2023), pp. 60–69. DOI: *10.1109/MMUL.2022.3232892*.

[BBM05] Peter Bragge, Andrea Bialocerkowski, and Joan McMeeken. "A systematic review of prevalence and risk factors associated with playing-related musculoskeletal disorders in pianists". In: *Occupational Medicine* 56.1 (Nov. 2005), pp. 28–38. ISSN: 0962-7480. DOI: *10.1093/occmed/kqi177*. eprint: *https://academic.oup.com/occmed/article-pdf/56/1/28/4181430/kqi177.pdf*. URL: *https://doi.org/10.1093/occmed/kqi177*.

[Bro+96] John Brooke et al. "SUS-A quick and dirty usability scale". In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7.

[Bru+24] Leonard Bruns, Benedict Saurbier, Tray Minh Voong, and Michael Oehler. "Presence and Flow in Virtual and Mixed Realities for Music-Related Educational Settings". In: *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)*. 2024, pp. 1–7. DOI: *10.1109/IS262782.2024.10704115*.

[Coh92] Jacob Cohen. "Statistical power analysis". In: *Current directions in psychological science* 1.3 (1992), pp. 98–101.

[Guo+21] Ruoxi Guo, Jiahao Cui, Wanru Zhao, Shuai Li, and Aimin Hao. "Hand-by-Hand Mentor: An AR based Training System for Piano Performance". In: *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2021, pp. 436–437. DOI: *10.1109/VRW52623.2021.00100*.

[JDT20] David Johnson, Daniela Damian, and George Tzanetakis. "Detecting Hand Posture in Piano Playing Using Depth Data". In: *Computer Music Journal* 43 (Jan. 2020), pp. 59–78. DOI: *10.1162/comj_a_00500*.

[Lab+23] Katerina Labrou, Cagri Hakan Zaman, Arda Turkyasar, and Randall Davis. "Following the Master's Hands: Capturing Piano Performances for Mixed Reality Piano Learning Applications". In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI EA '23. Hamburg, Germany: Association for Computing Machinery, 2023. ISBN: 9781450394222. DOI: *10.1145/3544549.3585838*. URL: *https://doi.org/10.1145/3544549.3585838*.

[Mat+24] E. Matzek, T. Yankee, O. Kohler, T. Lipke-Perry, N. Kholgade Banerjee, and S. Banerjee. "VRmonic: A VR Piano Playing Form Trainer". In: *2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)*. 2024, pp. 330–334. DOI: *10.1109/AIxVR59861.2024.00056*.

[MK94] Paul Milgram and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays". In: *IEICE Transactions on Information and Systems* E77-D.12 (Dec. 1994). Special Issue on Networked Reality, pp. 1321–1329. ISSN: 0916-8532.

[NB24] Luc Nijs and Bahareh Behzadaval. "Laying the Foundation for Augmented Reality in Music Education". In: *IEEE Access* 12 (2024), pp. 100628–100645. DOI: *10.1109/ACCESS.2024.3427698*.

[Tho24] Felix Thoma. *ROLI Airwave, Controller und Klavier-Lernhilfe*. Oct. 2024. URL: *https://www.amazona.de/roli-airwave-controller-und-klavier-lernhilfe/* (visited on 01/02/2025).

[TTR03] Miles Thom, Mark Thomas, and Gary Roberta. *What Every Pianist Needs to Know About the Body*. GIA Publications, 2003. ISBN: 9781579992064. URL: *https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=3002181&site=ehost-live*.

[Urv20] Ernest Urvator. *Choreography of the Hands: The Work of Dorothy Taubman*. YouTube. Verfügbar unter: *https://www.youtube.com/watch?v=lP1nofzv8j4* [Zugriff am 04. Januar 2025]. 2020.

[Wai+24] Ethan Waisberg, Joshua Ong, Mouayad Masalkhi, Nasif Zaman, Prithul Sarker, Andrew G. Lee, and Alireza Tavakkoli. "Apple Vision Pro: the future of surgery with advances in virtual and augmented reality". In: *Irish Journal of Medical Science (1971 -)* 193.1 (2024), pp. 345–346. ISSN: 1863-4362. DOI: *10.1007/s11845-023-03457-9*. URL: *https://doi.org/10.1007/s11845-023-03457-9*.

[Zim18] Hans Zimmer. *Hans Zimmer Breaks Down His Career, from 'Gladiator' to 'Interstellar' | Vanity Fair*. YouTube. Verfügbar unter: *https://www.youtube.com/watch?v=GGs_NT4iL2c* [Zugriff am 25. Dezember 2024]. 2018.