Pianow - Piano Self-learning Assistant in Mixed Reality

Created by

Tsai Yi Ting*

Supervisor

Dr. Dirk Schnieders

Department of Computer Science

The University of Hong Kong

April 15, 2018

* UID: 3035124520, Email: u3512452@connect.hku.hk

Abstract

Learning to play the piano is a difficult process that requires constant and rigorous practice. While it is believed that 65% percent of human beings are visual learners [1], traditional methods of teaching and learning to play the piano have not been focusing on exploiting the potential of human vision. For example, young children, when learning to play the piano, are very often not allowed to see the names of the keys they press but are instead required to picture the keys in their minds [2]. Meanwhile, in helping new learners picking up the instrument faster, the idea of using falling notes and colored keys as visual hints has proven to be highly promising by many applications. In this project, this idea will be brought into Mixed Reality, a combination that is expected to improve the learning experience, as learners can follow the tutorials without looking away from their own physical piano keyboard. The new application, named Pianow, aims to offer beginners a new way to practice playing the piano by themselves, which is easy, fun, and interactive. Pianow is developed for the HoloLens. It can effectively align virtual assets to the physical piano keyboard and display interactive piano tutorials consisting of falling notes and colored keys.

This interim report will discuss the significance of Pianow in assisting and motivating piano learners and justify why the idea of falling notes and colored keys are chosen. In the methodology section, the solutions to some technical challenges, especially hologram alignment, will be presented. Finally, the result and future direction of the project will be discussed and reported.

Acknowledgement

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background & Problem Statement

The piano keyboard is composed of only black and white color. The pattern of black and white keys on the piano is repeated every 12 notes (i.e. an octave). This includes 7 white keys and 5 black keys (see Figure 1). The reason for this color scheme is that the black keys help distinguish flats and sharps in musical notation and allow the players to quickly determine where they are on the keyboard [3].
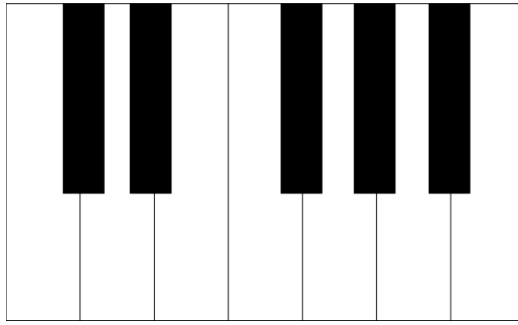
Figure 1: An octave consisted of 7 white keys and 5 black keys

Mastering the piano requires regular and rigorous practice. Professional pianists normally practice 6-8 hours per day and have to take regular breaks. [4] For most piano students, a reasonable amount of practice is around 30 minutes per day [4]. Although, learning to play the piano is time-consuming and repetitive, there are a lot of benefits to playing the piano. Since learners have to quickly differentiate repetitive groups of black and white keys, they acquire the ability to split concentration, which is good for multitasking in real-life situations [5]. Each hand performs entirely different movements, which stimulate multiple parts of the brain. This process is beneficial to hand/eye coordination and also inspires creativity [5].

However, not everyone can attend piano lessons regularly. A 30-minute piano lesson costs USD30-60 on average [6]. Learning to play the piano by oneself is not an easy process, especially for absolute beginners who are not able to read the music sheets. An estimation by CMUSE [7] indicates that 85 percent of piano learners give up on learning at an early stage because they dislike learning the basics. For starters, it is easy to get disheartened if the progress comes slowly.

There have been many attempts to address this problem. The most common form of assistance is self-learning tutorials by means of a video, a book or an application. All these forms, however, have their critical shortcomings. While books are not visually friendly, video tutorials are not interactive enough. From time to time, learners have to alternate between playing the piano and watching the tutorial, which causes great inconvenience. Mobile applications such as Perfect Piano[1], an intelligent piano simulator designed for Android phones and tablets, are more interactive, but they fail to offer such experience on a physical piano keyboard. As a result, the learning experience on one of these applications is drastically different from that on a real piano.

Mixed Reality is an emerging technology that embraces the interactions between the virtual and the physical worlds. Many studies have shown that the application of Mixed Reality in education can significantly boost the effectiveness of teaching as well as the performance of learners. For example, Dickey [8] mentions that the use of virtual content in education facilitates a constructivist learning environment, where learners can proactively participate in practical learning activities rather than just trying to absorb theoretical concepts. Dede [9] also points out that the interactivity brought about by Mixed Reality technology allows knowledge to be obtained through hands-on practices in a virtual environment, which is very difficult to achieve in the real world.

The potentials of Mixed Reality are especially apparent when it comes to learning activities that demand constant practice and high level of interactivity like learning to play the piano. A piano tutorial built in Mixed Reality is expected to offer strong immersive experience and significantly improve the learning outcome.

1.2 Current Solutions

There are several piano self-learning tools that are available on the market. The most prominent solutions include Synthesia, Teomirn and Music Everywhere.

---

[1] Available at: https://play.google.com/store/apps/details?id=com.gamestar.perfectpiano&hl=en
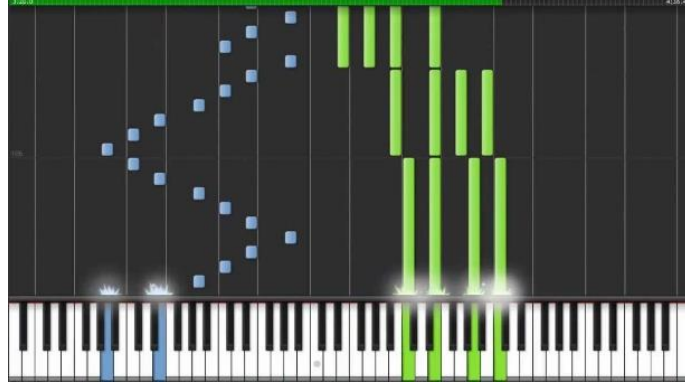
Figure 2: Synthesia's tutorial with falling notes and colored keys[2]

Synthesia [10] attracts special attention due to its successful idea of building a piano tutorial using falling notes and colored keys (see Figure *Figure 2*2). The falling notes match with the keys and let the user know beforehand which notes to be played next. The colored keys simply indicate which keys should be pressed at the moment. The application allows the user to connect to a MIDI piano keyboard, the data from which is then passed back to Synthesia to determine whether a key has been pressed by the user and for how long it has been pressed. With the appropriate piano keyboard, the keys can be illuminated to create the effect of key coloring. Thanks to its intuitive design, Synthesia has made playing the piano easier and more fun to many people. The major downside of Synthesia is that most of the tutorial contents are still displayed on the computer screen, and the application needs to be paired with certain types of digital piano keyboard to enable all functionality.

Figure 3: Teomirn's tutorial with virtual hands and keyboard[3]

---

[2] Youtube Image. Available: https://i.ytimg.com/vi/m5RmYEd1N9I/maxresdefault.jpg
[3] Youtube image. Available: https://www.youtube.com/watch?v=aovJh2SxDYU

Teomirn [11] (see Figure 3). is a Mixed Reality application on HoloLens that teaches the user to play the piano. The design philosophy of this application is to allow the user to imitate piano experts by watching them play. The application supports 2 learning modes: watching a virtual piano expert playing on the physical piano keyboard or playing alongside a virtual hand that appears on a virtual keyboard right above the physical keyboard. Figure 3 illustrates how the second learning mode is designed. The aim of this application is to let the learner observe how a certain piece of music is played by an expert, from which they can learn the correct hand placements, movements, and the best practices when playing the piano. However, with this approach, the user is still forced to divert his attention to the tutorial. Teomirn's solution, although developed for Mixed Reality, can simply be considered a more realistic version of a video piano tutorial.

Figure 4: MusicEverywhere's tutorial with virtual band and keys[4]

Music Everywhere [12] is another application in Mixed Reality that describes itself as a piano improvisation learning system. This system aims at providing the user with piano improvisation skill, which is not often taught but is needed when playing with other artists or in a band. Music Everywhere, hence, addresses a slightly different problem compared to that mentioned in Section 1.1. There are two main components of instruction used in this application: a virtual band along with which the user can play, and a virtual panel placed perpendicular to the piano keyboard to indicate which keys to be pressed or for how long it should be pressed (see Figure 4). Like Teomirn, Music Everywhere does not overlay the tutorial content directly on the physical piano keyboard, which may present a significant obstacle for

---

[4] Twitter Image. Available: https://pbs.twimg.com/media/C230cqvXUAQJG-o.jpg

absolute beginners. The specific focus on piano improvisation skill also renders this application as an inappropriate choice for many piano learners. Additionally, Music Everywhere requires the use of a marker to locate the position of the piano keyboard. In the long term, this is not the best approach to be used on the HoloLens.

1.3 Motivation

The project team is motivated to provide a solution that solves the problems of self-learning piano mentioned in Section 1.1, which is better than or is a complement to the current solutions mentioned in Section 1.2. As Mixed Reality is foreseen to be popular in the near future, a Mixed Reality application is likely to cost less than hiring a human instructor. Using Mixed Reality as the medium of instruction enhances the teaching content with more visual elements, which make the tutorials easier and more intuitive, especially for young children who have little or no knowledge of musical notations. At the same time, a Mixed Reality application should be fun and interactive, which will inspire beginners to keep practicing as they can see their own progress. Unlike Music Everywhere and Teomirn, however, the new application should append all tutorial content directly on the piano keyboard to make full use of the immersitivity of Mixed Reality.

1.4 Objectives

The project team aspires to develop a new Mixed Reality application to assist piano self-learning. The application is named Pianow. Pianow mainly aims to help beginners find playing the piano fun and easy at the early stage so that they have the motivation to keep learning and moving on to later stages. It can also be entertaining even for experienced piano players. In the long term, Pianow is expected to change the way self-taught pianists learn and allow more people to feel the excitement of playing the piano. At this stage of the project, there are 2 objectives to be achieved:

1. Pianow is reasonably accurate and effective in aligning its virtual content with a standard 88-key piano keyboard.
2. Pianow offers at least 2 working tutorials with intuitive user interface and good user experience.

The first objective involves the main technical challenge of the project. It is important that the application has knowledge about where the physical piano keyboard is located in the environment so that virtual contents can be placed appropriately. The second objective is about the practicality of the project, i.e. whether piano tutorials in Mixed Reality are practicable and whether they offer good experience to the user. These two aspects are identified as the two most fundamental issues that the project needs to address at the current stage.

1.5 Design for Effectiveness

Pianow's two methods of instruction are determined to be falling notes and color keys. The choice of these two methods are based on several factors.

The first factor that justifies this approach is the success of Synthesia itself. Synthesia uses the same idea of falling notes and colored keys as the instruction medium, which has proven to be useful for many users, based on multiple ratings and review on App Store and Google Play Store. On Google Play Store, Synthesia has more than a million downloads and holds a rating of 3.8 stars [13]. Meanwhile, on App Store, it is rated at 4.5 [14]. The reviews are generally positive, and a large number of users believe that Synthesia has helped them improve their piano skills in a fun and interactive way.

Nevertheless, from time to time, the most commonly mentioned issue of Synthesia-style tutorial is the fact that it discourages learners from looking at the music sheet, which will form a bad habit in the long term. However, while following the music sheet is one way to learn the piano, learning-by-ear is another method that has proven to be helpful to many people, including professional pianists. This method, proposed by Shinichi Suzuki in the 1930s, is also referred to as the mother-tongue approach, which is advocated by many influential musicians and has become widely accepted in North America [15]. The method involves 2 steps. First, to familiarize the learner with a piece of music by letting them listen to it repeatedly for many times and, second, to let them find out by themselves how to play the piece of music on a keyboard without any printed music sheet [15]. The music sheet will be introduced at a later stage, when the learner's ears have been well-trained [15]. The design of Pianow does not strictly follow this framework because it is hard to guarantee that users have been familiar with a piece of music before they start learning it, and when following the tutorial, users do

not need to find out by themselves which key to press to replicate the memorized melody. Pianow, however, do embrace the idea of intuitive music learning, which allows new learners to learn music in a natural way, without too much emphasis on learning to read the music sheet.

Last but not least, by overlaying the graphical content directly on the piano keyboard, Pianow is likely to boost the productivity of the learning process. The Social Science Research Network states that 65% of human beings are visual learners [1]. Using visual instructions directly on the piano keyboard will probably allow many learners to remember the songs much faster compared to other methods.

Certainly, the effectiveness of this design needs to be formally examined. Case study, expert consultation and user-assisted testing are the intended methods for gathering related information and opinions. In the long term, it is necessary to perform an evaluation of real-life use cases over a long period of time to objectively determine the effectiveness of the application. This requires more resources and is out of the scope of the project at the current stage.

1.6 Scope

Pianow is a Mixed Reality application developed for HoloLens only. The piano keyboard to be used with the application should be a standard 88-key keyboard. The application is expected to be used under stable and slightly dim lighting condition (300-500 lux). At this stage, a tutorial consists of only two elements: falling notes and colored keys. Pianow allows basic tutorial customizations, including choosing song, changing speed and switching the music on or off.

1.7 Work distribution

The following table shows the work distribution for the final year project report.

Task	Bui, Thanh Tung	Tsai, Yi-Ting
Introduction	shared	shared
Methodology - Hologram Alignment	✓	
Methodology - User Interface		✓
Final Result & Evaluation for Hologram Alignment	✓	
Final Result & Evaluation for User Interface		✓
Future Direction	shared	shared
Conclusion	shared	shared

Table 1 Work distribution table

2 Methodology

2.1 System & Equipment Setup

This project requires a range of hardware and software that should be set up to work in coordination with each other during the development process.

2.1.1 Hardware

The major pieces of hardware used in this project are the HoloLens and a standard 88-key piano keyboard. The HoloLens is a powerful headset

designed exclusively for Mixed Reality, and arguably the best of this category on the market. It is also the most widely known Mixed Reality headset with a large supporting community. Hence, it is the most viable option for this project.

Besides, in this project, most of the development tasks are performed in a system with 64-bit dual-core CPU, 8GB of RAM and integrated Intel HD Graphics.

2.1.2 Software

For Mixed Reality development on HoloLens, Microsoft has provided a detailed list of software requirements. According to this list, Unity is required as the officially supported development platform, and Visual Studio has to be used alongside Unity as the debugging and deployment tool for HoloLens [16].

For achieving the first objective of the project, OpenCV and Vuforia are used for keyboard recognition and marker detection respectively. OpenCV library contains implementations of many traditional computer vision algorithms, which are suitable for recognizing objects with standard features like a piano keyboard. Vuforia, on the other hand, is a well-known Augmented Reality SDK with a robust mechanism for marker detection and tracking. It is recommended by Microsoft for application that requires holograms to be placed at certain locations or attached to certain objects [16].

Starting from version 2017.2, Unity has integrated Vuforia as a buit-in module. With appropriate settings, Vuforia can be easily enabled for any Unity projects. However, as OpenCV is not officially supported by Unity and HoloLens, some extra integration steps need to be performed. First, the code for piano keyboard recognition is written in C++ and then compiled to a Dynamically-linked Library (DLL) file for Universal Windows Platform. OpenCV source code also needs to be slightly modified to remove some dependencies on the Windows Desktop Platform before being compiled to DLLs. This modified version is available on the opencv-hololens[5] repository

---

[5] Source code available at: https://github.com/sylvain-prevost/opencv-hololens

by Sylvain Prevost. Finally, all the DLLs are imported to Unity as native plugins, which can be gracefully deployed and run in the HoloLens.

Additionally, to stream data from the locatable camera of the HoloLens to Unity, the HoloLensCameraStrea[6] plugin written by Vulcan Technologies is used. This plugin is necessary since the HoloLens API does not offer a native method for accessing the camera video stream in real-time. The plugin allows the camera stream to be passed to the application at 30 frames per second (fps), which is sufficient for the recognition task.

Finally, some functions that can be implemented using HoloLens API are adapted from the MixedRealityToolkit-Unity [7] repository written by Microsoft. This repository is a collection of pre-made components that can be utilized to accelerate the application development process on the HoloLens. In this project, the features that are adapted from this repository include input handling, spatial mapping and draggable hologram.

2.2 Alignment Methods

Please refers to Ethan's report.

2.3 User Interface & Tutorial Content

Control menu design The User Interface (UI) of the application will mostly be adapted from pre-made assets in HoloToolkit bundle, with appropriate modifications to suit the project. The design of the main control menu UI is kept simple from which the user can perform all actions. The following (see Figure 5) is the draft of the control menu design.

---

[6] Source code available at: https://github.com/VulcanTechnologies/HoloLensCameraStream
[7] Source code available at: https://github.com/Microsoft/MixedRealityToolkit-Unity
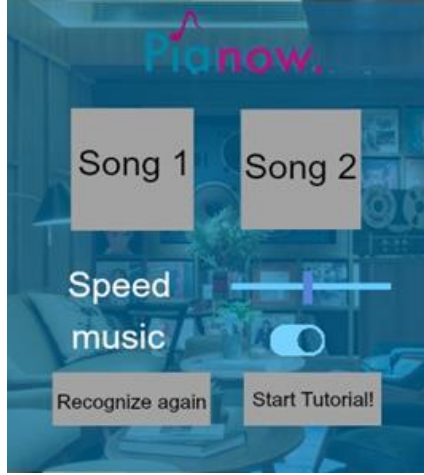
Figure 5: Control menu design (draft)

In the article of principles of user interface design written by Euphemia Wong on the interaction design foundation website, it mentions the importance of interaction between human cognition and the interface design. A good design reduces learning new representation and eliminates the confusion at every touch point [17]. As a result, the UI design consists a clear Pianow logo and no redundant text and functionality. The background and buttons is all set to cyan and blue related colors for consistency. Color is produced in the visual system of human brain and therefore color psychology is a crucial element of human perception as well. Blue is considered to give emotions of trust and relaxation [18] and therefore the color is chosen for the purpose of users' relaxed and trustful piano tutorial experience.

In addition, the button is designed by two important factors. One is the pointing. Pointing means it highlight the button in some visually distinctive manner when the pointer is resting on it. The other one is the activation. Activation means it highlight the button in some visually distinctive manner when the button is activated or pressed. As a result, choosing the pre-made traditional button from HoloToolkit is sufficient. When the user gazes the button, the gaze becomes a circle cursor which satisfies the pointing. When the user taps to press the button, the button shrinks which satisfies the activation [19].

Instruction picture is designed to be plain text with supplementary graph to show each step for OpenCV recognition and marker recognition method. The picture mapped as texture onto the raw image component.

Tutorial content is made in the form of video tutorials, which can be scaled to fit the physical piano keyboard and inserted to Unity scene as media objects. The 3D modeling a L shape model (see Figure 6) is achieved preliminarily on Sketchup. After importing to Unity, the video texture is mapped to the model and it renders the tutorial video.
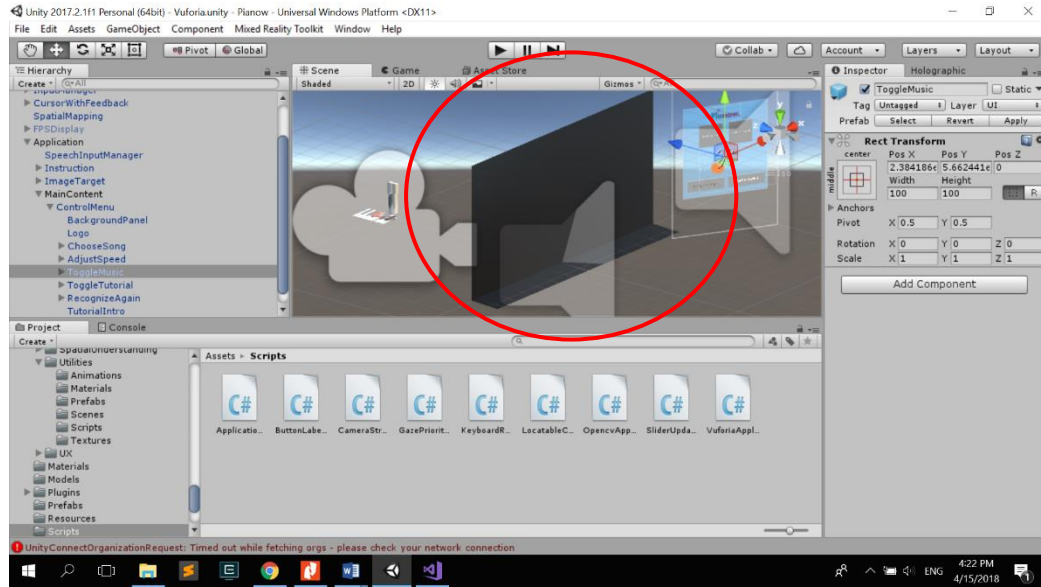
Figure 6: The L shape model to map video texture

2.4 User experience with Stability of HoloLens

Hologram render distances From Microsoft windows development center, there is an article discussing about hologram stability. The HoloLens display should be set to the optimal distance to 2.0m and it could maintain a clear image [20]. The reason to choose an optimal distance is that human visual system integrates multiple distance-dependent signals when it fixates and focuses on an object. The first factor is accommodation. Accommodation is the ability of the eyes to focus on an object when it changes the distance from the person. Second factor is convergence. As an object moves closer to the person, the person's eyes slowly converge, and while an object moves away, the eyes diverge. Accommodation and convergence are related. When the eyes view object at infinity, the eyes become parallel and the eyes accommodate to infinity. The eyes accommodate to a near point when the eyes cross and view a near object. However, when the

accommodation and convergence are in conflict, which means that accommodation and convergence are at different distance, this leads to discomfort. To avoid the discomfort, Microsofts (see Figure 7) suggested to set the distance of the content that the user converges to as close as 2.0m.
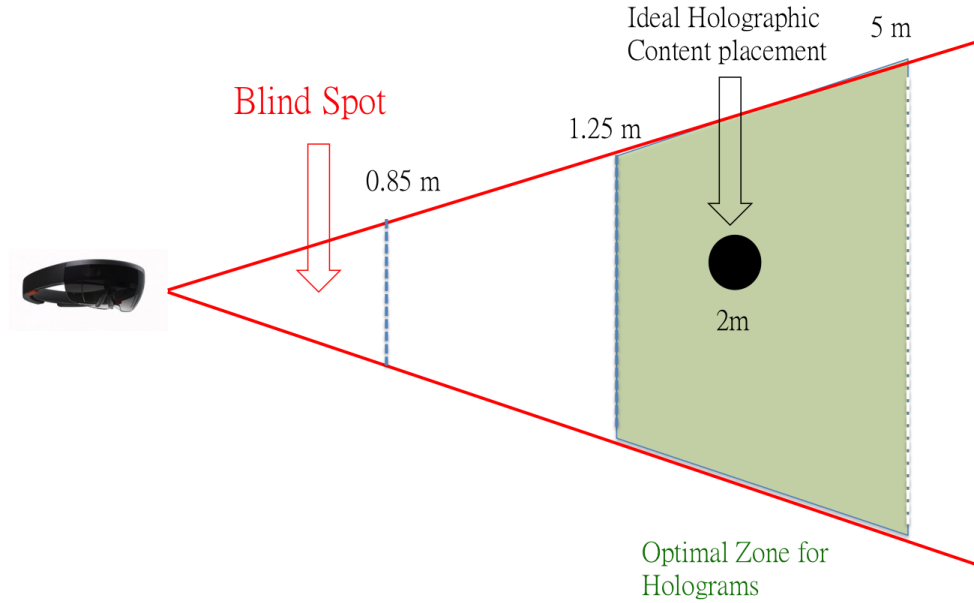
Figure 7: Optimal distance for placing Holograms from the user

However, for the Pianow application, the distance could not be set to 2.0m, since it is unreasonable for user to play the keyboard at a far distance. In this case, a stabilization plane should be set to increase stability. All the holograms that stay on the same stabilization plane would receive the maximum hardware stabilization. The HoloLens would choose this plane automatically. Therefore, for the instruction pictures, it is specially set the stabilization plane to 0.5m. (see Figure 8) The stability worsens when the user gazes back and forth at different distances, but in the application, the user simply needs to see the stationary hologram. And therefore setting a stabilization plane would be the enough to secure the stability even the distance is not 2.0m.
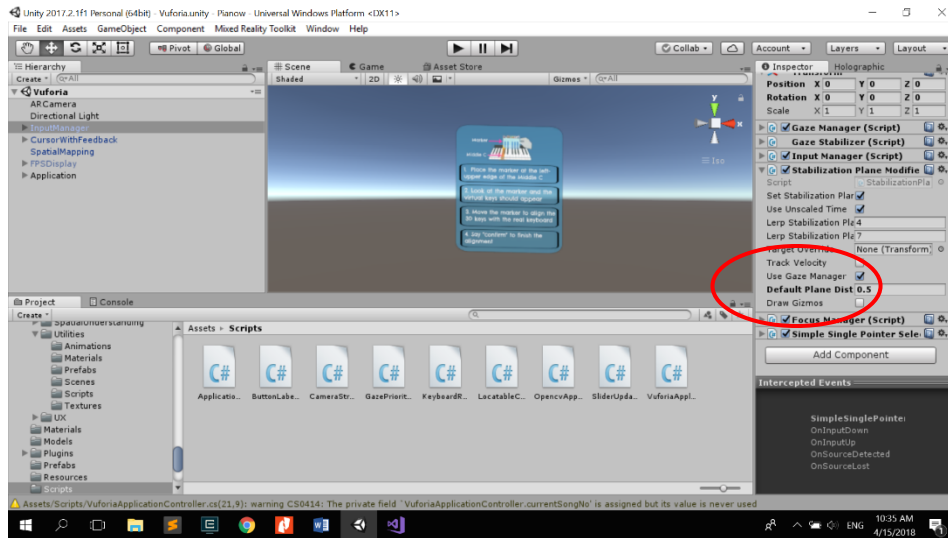
Figure 8: Screenshot of setting stabilization distance to 0.5m on Unity

Box Collider The collider needs to be added to the 3D objects when the user sets the stabilization plane. Different kinds of collider should do the work, however, the mesh collider is more expensive for the performance and as a result the box collider is added.

To distinguish different result of the quality of Holograms, the Microsoft windows development center have given seven different terminologies.

Terminology	Description
Accuracy	The Hologram stays where it is placed and it should be independent to the changes of sparse environment and motion.
Jitter	When the tracking of environment downgrades, high frequency of shaking holograms will appear. This problem could be solved by senor tuning to update sensor calibration information.
Jumpiness	A hologram jumping away from its position could occur when HoloLens is updating the information of the user's environment.

Judder	The holograms with motion sometimes result in double images, and in this case the Frames per second should be set to 60 fps.
Drift	A hologram moving away from its position. To solve this, construct the hologram to be close to the spatial anchor, which is an important point in the world system for hologram to refer to the coordinates.
Swim	A hologram sways corresponding to the user's head movement. This occurs because the hologram is not on the stabilization plane and therefore the stabilization plane should be updated.
Color separation	User normally sees a rainbow effect of the hologram due to rapid head movement. The degree of the color separation is influenced by the speed of the hologram.

Table 2: HoloLens stability terminology

2.5 Testing & evaluation

The testing and evaluation focuses on the physical experiment conducted on 10 subjects of their user experience. Under the review of different research study, testing methods, procedure and evaluation methods are discussed.

User experience experiment
2.5.1 Testing method

To optimize the Pianow application, we conduct user testing with two methods: moderated user testing and session replays.

Moderated user testing is a testing method used for measuring product experience. The researcher directly observes the user in real time and

directs the user to accomplish different tasks. Normally the user gives instant feedback when the researcher asks questions. This is opposite to the idea of unmoderated user testing, which is normally conducted remotely with predetermined tasks for user to follow and complete without the scanning of the researcher. The moderated user testing avoids the downsides of unmoderated user testing which are mentioned in the book named usability testing essentials written by Carol Barnum [21]. Barnum mentioned how it is a significant limitation when the researcher is unable to talk to the users and ask any followed up questions as well after receiving the subjective feedback. Moreover, without moderated user testing, the researcher could learn that the user has dropped off the testing, but they could not understand the reason behind these actions. The moderated user testing on the other hand has more connection and flexibility with the user.

However, despite the advantages of the moderated user testing method, there are still hindering problems. Chris Goward, the founder of WiderFunnel which is a marketing agency based in U.S., describes how the Hawthorne Effect can affect the moderated user testing method. The act of observing one perform an action may affect the outcome . When users know they're being observed, they may be more motivated to complete the action [22]. For example, if the user knows the researcher is looking for a good usability result, the user will be tempted to speak of the good parts more than the bad parts in terms of the usability. After analyzing both pros and cons, although the moderated user testing method may be influenced by bias, the method is most suitable for HoloLens testing, since not every user owns a HoloLens. The HoloLens is a relatively new technology requires technical instructions for the user. It is also easier to realize the problems the user may encounter by observing their gestures, expressions and oral feedbacks.

Session replays is a method to replay the user's testing journey. In our test procedure, the project team records every user's testing process. (figure) The advantage of this method is to have chances to observe the small details that may have been missed out, for example, the hand gestures and the head movements. Moreover, we would record the users' responses when we asked them questions or when they encounter

difficulties. This recording helps us to analyze the different aspects of our application and to distinguish where the problems are immediately.

However, we did not choose to capture what the user saw in the HoloLens, because the video capture requires mixed reality capture [23] and the system will throttle the application to 30Hz rendering in order to match with the video record framerate of 30 fps. This is not the perfect framerate of 60 fps to run the application and the testing result may be biased. Therefore, we chose to record the observation of the user from the researcher perspective only and we further analyze with the oral feedbacks.

2.5.2 Testing preparation

Determining goals It is crucial to have an idea of what the purpose of doing the usability test is. The goal for testing is to understand the perception of user interface and efficiency of usability. In other words, the design of the user interface is visually acceptable and it is intuitive to use the application.

Selecting participants We have set our experiment subjects to 10 people with or without music knowledge background. In this way, we could test learning efficiency difference between absolute beginners and the others. In the article written on UXBooth by Alina Zawadzka [24], author mentions in selecting the subject in user testing and survey it is important to understand the relationship with the company or research team. For example, if the subject is a friend of the project team, the result could be biased and normally has a better result. We have realized this issue, however, since it is more approachable to conduct experiments with friends on campus, we chose to look for the subject we are familiar with. We assume there is no preliminary psychological influence with these subjects.

Venue and sets-up The testing is conducted in the 310 Lab where it is a purpose-built venue for mixed-reality devices and it has good lighting too. The keyboard is placed in the middle of the lab and recording camera is placed at the good angle. The laptop for filling in the form is put and set up. (See Figure 9)
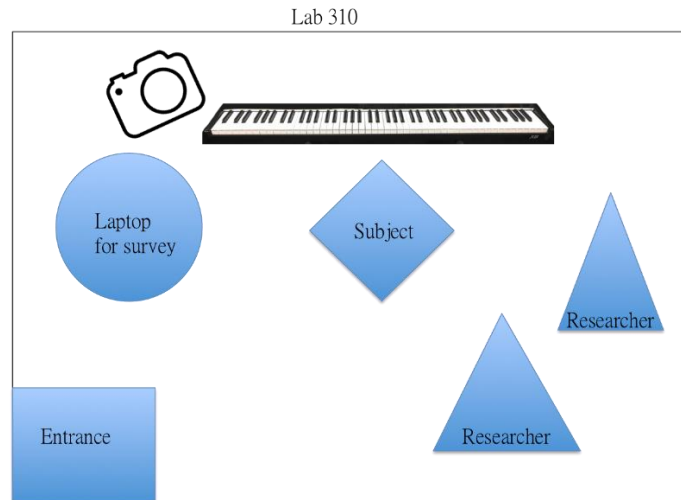
Figure 9: Testing device placement and venue setup

2.5.3 Testing procedure

The testing for each user is set to no more than 15 minutes. Every time only allows one user in the testing lab. The experiment conductor teaches the user the usage and normal tap gesture of the HoloLens. The user is first tested by the application with the Vuforia marker detection first. Followed by testing the tutorial content and menu user interface, the user will follow the tutorial content and press the keys accordingly on the keyboard. After finishing testing the first method, user moves on to the application with the keyboard recognition with OpenCV and follows the tutorial content to see how far the alignment is correct. After trying the application, the user should conduct an online goolge survey and the testing is completed when the user finish the survey. (See Figure 10)
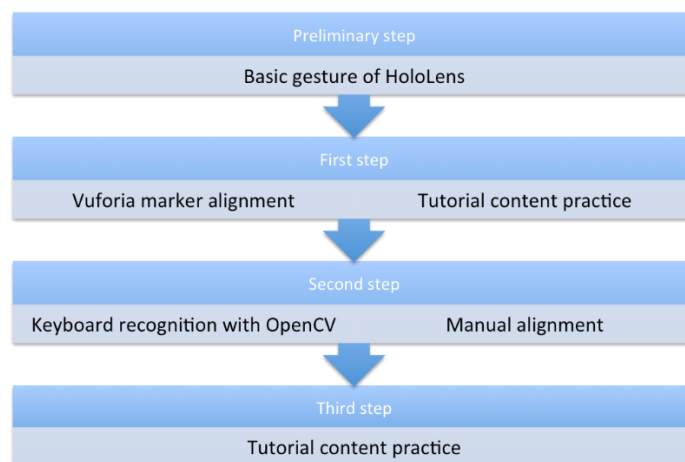
Figure 10: testing procedure with different steps to achieve

2.5.4 Evaluation method

The evaluation of the testing is separated into three parts: oral feedback analysis and gesture, posture analysis, survey analysis.

Survey questions design The survey should contain four parts [25]. A short but clear introduction includes a welcome message, the purpose of this survey, the creator of the survey and the time it takes to finish the survey. Next, screening questions to understand the respondent's background and characteristics for further analysis purpose. Following with the measurement questions, this is the core of the survey which providing the information the researcher is looking for. Open-ended questions should be added for qualitative research. Close-ended questions help the user to narrowing down the possible answers and at the same time help researchers to interpret result easily. Lastly, it should include an ending with message to reassure the user that their input is valuable. Following by the above study suggested, the survey questions are showed below:

Introduction

Pianow experiment survey
Hello! Thank you for your participation! This is a survey for subject who has finished testing the Pianow HoloLens app. The survey is for analyzing the usability and user experience on our app. The following questions will take less then 5 minutes to complete. If you encounter any questions, please seek help from our coordinator.

Screening questions

1. Have you played piano or had any experience with piano before?	Yes/No
2. Do you know how to read music notes?	Yes/No
3. Have you used HoloLens headset before?	Yes/No

Measurement questions

In the experiment we had let the subject test 3 kinds of recognition keyboard method, please answer accordingly.	
4. For OpenCV recognition keyboard do you think the instruction is clear enough for you to understand the process?	Yes/ No/Other
5. For OpenCV recognition keyboard do you think the alignment of the virtual keyboard is accurate enough?	Yes/ No/Other
6. For marker method do you think the alignment of the virtual keyboard is accurate enough?	Yes/ No/Other
7. For drag and drop method do you think the alignment of the virtual keyboard is accurate enough?	Yes/ No/Other
8. Do you have any comment on the recognition method?	Open question
The tutorial has two features falling notes and colored keys. Please answer the question accordingly.	
9. Do you think the falling notes help you to learn piano?	Yes/ No/Other
10. Do you think the colored keys help you to learn piano?	Yes/ No/Other
11. Do you think the control menu need any improvement?	Open question
12. Do you think there should be any improvement for the tutorial?	Open question
This part is to know your overall experience with our app.	
13. Will you consider this app as a tool for you to learn piano?	very likely/likely/not likely/not very likely
14.What improvement do you think is needed for this app?	Open question
15. From 1 to 10 please rate this app. (10 is the highest)	Interval question from 1 to 10

Ending

> Thank you for submitting the survey. We are appreciated for you time. We believe that the efforts from you can really help Piaow team to improve the applications! Have a nice day!

Oral feedback Identify the sentences that the user says they have difficulty with or they are unclear about. Follow the words the user says to understand what they observed first when using the interface.

Gesture and posture observation

Identify if any abnormal gestures or uncomfortable posture are showed. Observe the finger placement on the keyboard when the user plays tutorial.

3 Final Result & Evaluation

3.1 Work Accomplished

All three alignment methods discussed in section 0 has been implemented and tested against real use cases. Meanwhile, the first prototype of the application, including a functional user interface and two working tutorials, is also ready for demonstration.

Below are the results of UI design, tutorial content, stability, and testing experiment. The resolution of the pictures captured through mixed reality capture is 1408x792px, which is low. As a result, besides the picture captured on HoloLens the design pictures are provided as well.

3.1.1 UI design

When running the application, the **Pianow logo** will show first. Below is the project logo design which we decided to keep it simple with our theme color cyan and a music note, and the design photo (see Figure 11) and picture captured which is what the user sees on HoloLens (see Figure 12).

Figure 11: Pianow application logo design

Figure 12: Pianow app starting screen with logo

After showing the logo, the alignment process begins. The **instruction picture** (see Figure 13 and 14) is showed as a Tagalone component where it will move to where the eyes look at following by the HoloLens cameras. However, the alignment process utilizes the locatable camera and the mixed reality camera captured cannot be accessed at the same time. Therefore, the instruction picture result only has the design graph.
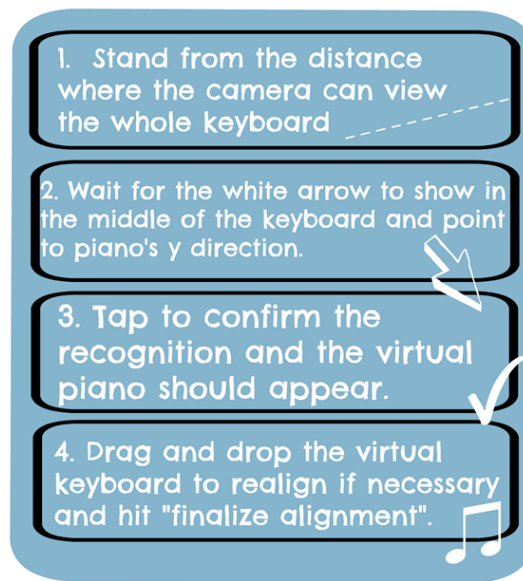
Figure 13: The instruction for keyboard recognition with OpenCV
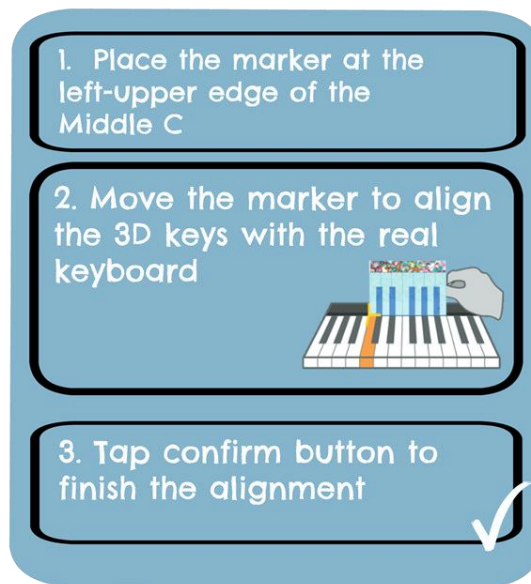
Figure 14: The instruction for Vuforia marker recognition

After following the instructions and finishing the alignment for the keyboard, the **control menu** appears on the right hand side (see figure 15) of the center view. A circle cursor generates after user's gazing. (See Figure 16,17 and 18)
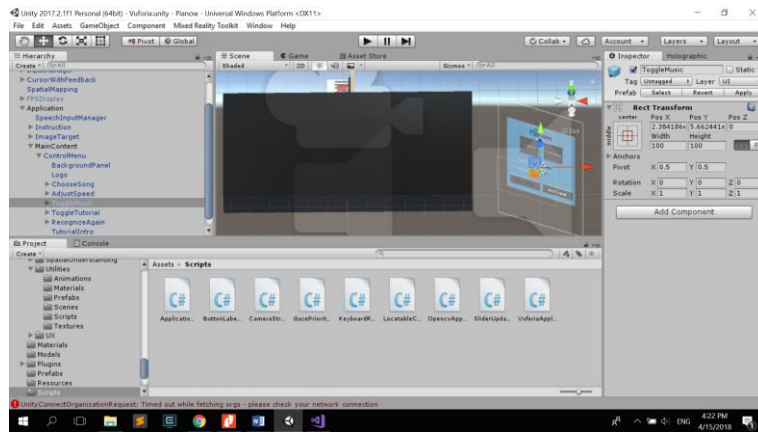
Figure 15 relative position of control menu and tutorial
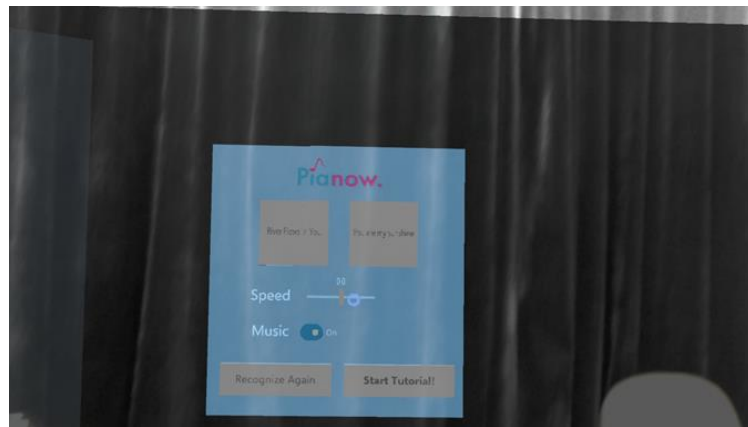
Figure 16: Control menu showed on HoloLens
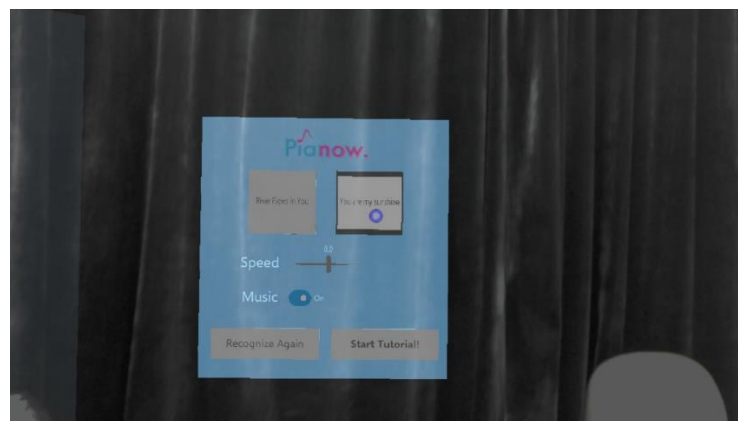
Figure 17: HoloLens cursor on song buttons
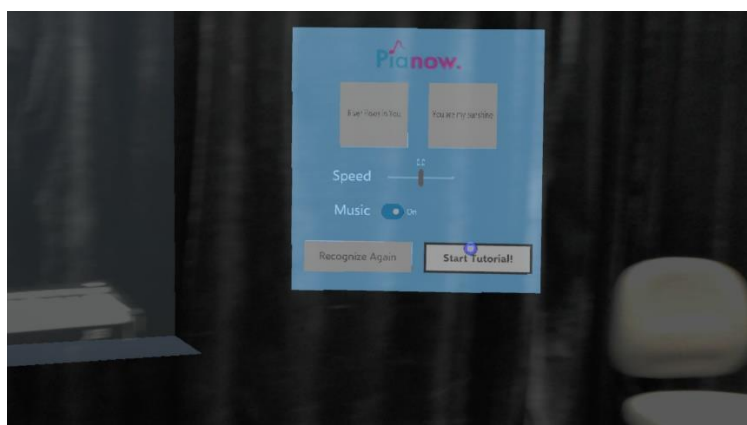
Figure 18: HoloLens cursor on the start tutorial button

3.1.2 Tutorial content

The tutorial in a L shape. The texture mapping of the video maps perfectly on to the piano keyboard and showed in front of the screen. (See Figure 19)
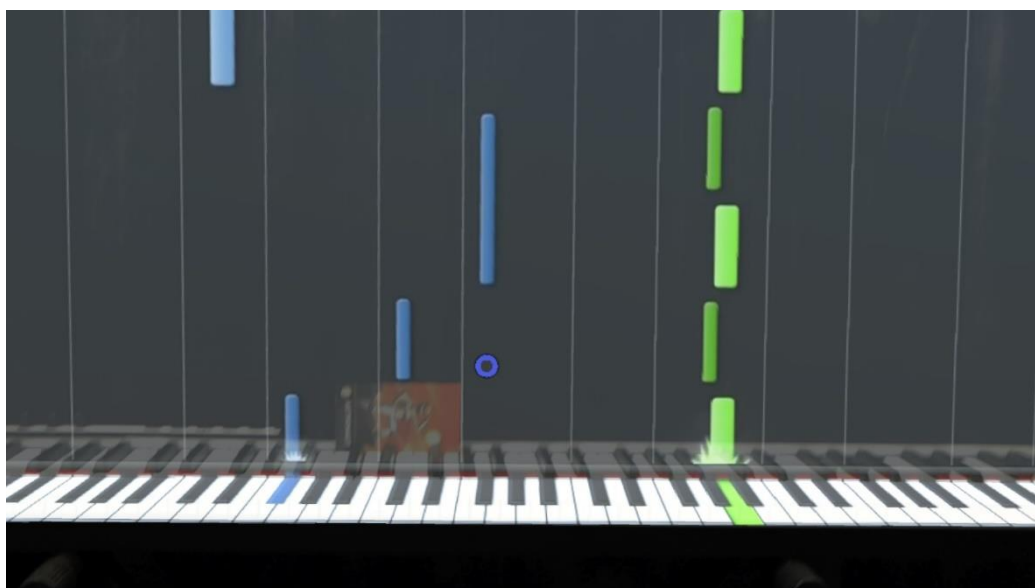
Figure 19:Tutorial content rendered from Synthesia style video

3.1.3 Stability

The stability is applied on the instruction picture to allow the tagalong component to increase stabilization. The method after setting the stabilization plane and box collider did result in a better stability. When the gaze hit on the instruction plane, it did not jitter and shake in a high frequency. However, the result is not able to be captured when the locatable camera is being used. Therefore, there is no figure showing the result.

3.1.4 User experience experiment

Survey result The following tables show the statistics of the answers from the 10 subjects.

1. Have you played piano or had any experience with piano before?

Answer	Number of subjects
Yes	7
No	3

2. Do you know how to read music notes?

Answer	Number of subjects
Yes	7
No	3

3. Have you used HoloLens headset before?

Answer	Number of subjects
Yes	3
No	7

4. For recognition do you think the instruction is clear enough for you to understand the process?

Answer	Number of subjects
Yes	2
No	8

5. For OpenCV recognition keyboard do you think the alignment of the virtual keyboard is accurate enough?

Answer	Number of subjects
Yes	4
No	6

6. For marker method do you think the alignment of the virtual keyboard is accurate enough?

Answer	Number of subjects
Yes	9
No	1

7. For drag and drop method do you think the alignment of the virtual keyboard is accurate enough?

Answer	Number of subjects
Yes	2
No	8

8. Do you have any comment on the recognition method?
Significant answer: Not sure how to put the marker.

9. Do you think the falling notes help you to learn piano?

Answer	Number of subjects
Yes	7
No	3

10. Do you think the colored keys help you to learn piano?

Answer	Number of subjects
Yes	8
No	2

11. Do you think the control menu need any improvement?
    Significant answer: the position of the control menu should be told to
    the user; the speed button could be bigger; it should provide more
    song choices.

12. Do you think there should be any improvement for the tutorial?
    Significant answer: The difference of falling notes between black key
    and white key is not clear. The instruction of the tutorial should be
    given.

13. Will you consider this app as a tool for you to learn piano?

Answer	Number of subjects
Very likely	1
Likely	7
Not Likely	2
Not very likely	0

14. What improvement do you think is needed for this app?

Significant answer: The procedures and instructions need to be clear; The width of the HoloLens is limited and caused the limitation to see both hands of tutorial.

15. From 1 to 10 please rate this app. (10 is the highest)

	points									
Numbers of subjects	1 point	2	3	4	5	6	7	8	9	10
	0	0	0	0	0	1	4	4	1	0

Table 3 User's rate of the application after experiment

Oral feedback From the recordings, the subjects normally speak out the difficulty of understanding the alignment methods in both Vuforia marker method and keyboard recognition with Open CV. In the middle of the experiment, they voiced out the discomfort of wearing the device. However, in the end when playing the tutorial, their happy expression and compliment replies are positive for this application.

Posture and gesture observation There are three things have been observed. The performance of the hand gesture of tapping and dragging the elements on the control menu is quite accurate and like the screenshot captured (see Figure 20) most of the user could control with correct hand gestures. Second, the sitting posture for 2-3 users (see Figure 21) is not comfortable enough and they have a hard time adjusting to the position to play tutorials. Lastly, for the user who has the music background, when following the tutorial, they tend to understand the finger placement (see Figure 22), however, vice versa, the absolute beginners have a relatively awkward hand placement. (see Figure 23)

Figure 20: Good gesture on tapping control menu from one of the subject

Figure 21: Subject with uncomfortable sitting posture when using the app

Figure 22: Subject with musical knowledge background shows a better fingering perfor-
mance

Figure 23:Absolute beginner subject shows a stiff and awkward fingering performance

3.2 Evaluation & Comparison

This section provides an overall evaluation of user experience experiment and stability testing.

Survey evaluation 70 % of our subjects have experience on playing piano and these subjects could read music sheet notes as well. 30% of the rest of the subjects are the absolute beginners our application is targeting in, however, it is also important to observed how the 70% subjects learn on the application.

80% of people think that the instruction for recognition on both OpenCV and marker is not clear enough for them to understand how to conduct alignment. This shows that the instruction needs to be revised. Compared with different key board alignment method, 90% of people agrees with the accuracy of marker method, which shows that the marker method is efficient enough. However, for the OpenCV and drag and drop method only 40% and 20% of subject respectively think the alignment is sufficient. The one person who commented that the marker method is not accurate enough is because he/she didn't understand the instructions.

70 % of people agree that the falling notes and colored keys are efficient for them to learn piano, however there are many user commented on the unclear instruction of where to look at the control menu, since it is placed on the right hand side. One user commented on the UI button design size, however, it varies from different user, and since only one user mentioned about it, the modification is reserved. Another thing to take into consideration is that two users mention the difficulty of distinguishing black keys and white keys from the falling notes, and the tutorial instructions should be added on to the application.

80% of people likely or very likely to consider the Pianow application as a tool for learning piano, which is an excellent result for achieving our objectives. For overall improvements, 4 users mentioned about the small screen view, instruction unclearness, and the need for tutorial instructions.

9 people out of 10 rated the app 7 points out of 10, and 10 is the highest score. The application could be considered to have a good usability with some parts of interface need to be revised.

From oral feedback and posture/gesture observation, we understand that the instruction should be revised since most of them reach out for help at the first stage of experiment when aligning virtual keyboard. As for the awkward sitting gesture, it might be caused by the unfamiliarness of the device, however it needs further proof when we have the access to capture what the user sees. Moreover, the fingering placement could be considered as an option function in the future, to encourage the absolute beginners or young children to learn more efficiently.

Preliminary test result: Marker detection with Vuforia offers the most satisfaction, the least difficulty and the most successful alignment on first try, while manual alignment is on the opposite side of the scale. (see Figure 24)
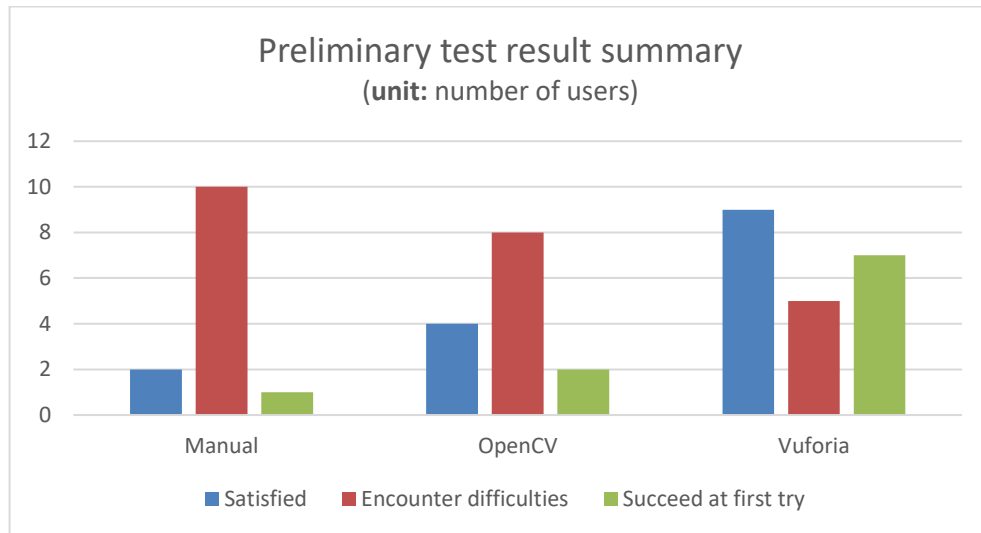
Figure 24: Preliminary test result summary in terms of different methods

3.3 Improvements made

Alignment instruction revised The picture showing the alignment procedure moves up to the top in order to attract user attention. The picture also provides the information of marker and middle c position. After learning from the graphs, it is easier for user to understand the text, since humans are visual creatures and 65% are visual learners according to Social Science Research Network mentioned in Forbes article.[citation] The icon of a tick is also removed to avoid misunderstanding. (see Figure 25 and 26)
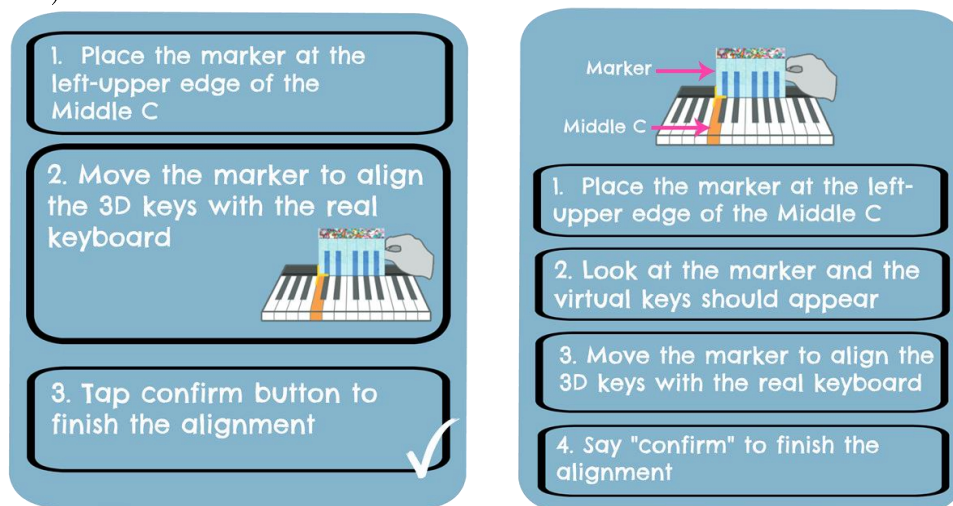

Figure 25: Instruction of Marker method before survey (Left picture)

Figure 26: Revised instruction of Marker method after survey (Right Picture)

Tap to confirm into say "confirm" From the survey, the interface for marker alignment contains too many elements. And a certain amount of users could not find the confirm button to confirm the alignment, and since the screen view already has the instruction picture, the marker and the virtual 7 keys generated by the marker, it is better to remove the confirm button and replace it with voice input. The user could simply gaze at the marker hologram and say "confirm". This is considered to be a natural way to communicate with the application and improve user experience.

Tutorial instruction video Many complain about the unclear information of the tutorial, since not every user has encountered Synthesia style [10] tutorials before. Therefore, a simple instruction video is edited with pictures to show an animation-like introduction. The video mainly shows three aspects of information. First (see Figure 27), there are two elements in the learning tutorial, the falling notes and the colored keys. In this way, we assure that the user understands where to look at when the tutorial starts. Second (see Figure 28), the falling notes have two different sizes. One is the thick one implying the user should press the white key and the other one is the thin one implying the user should press the black key. Lastly (see Figure 29), the green color represents right hand notes and the blue color represents left hand notes. Figure 30 and 31 show that the tutorial instruction video is designed and rendered in front of the control menu.
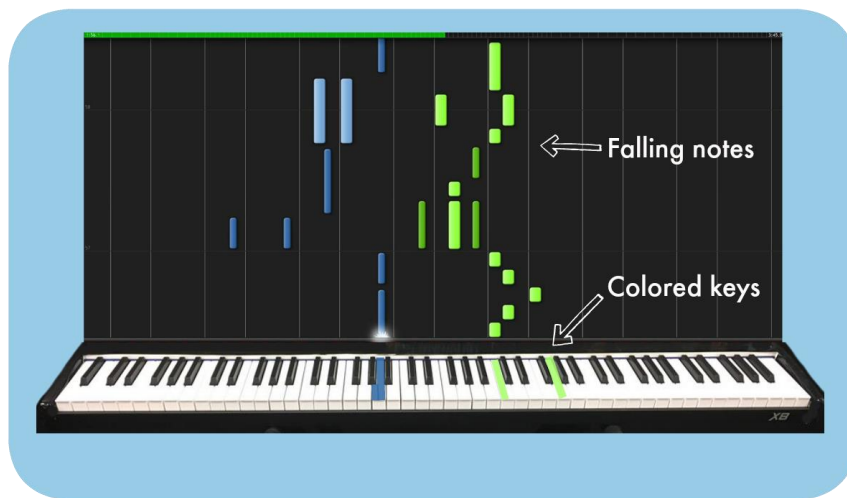

Figure 27: First instruction of tutorial video mentioning two elements of the tutorial
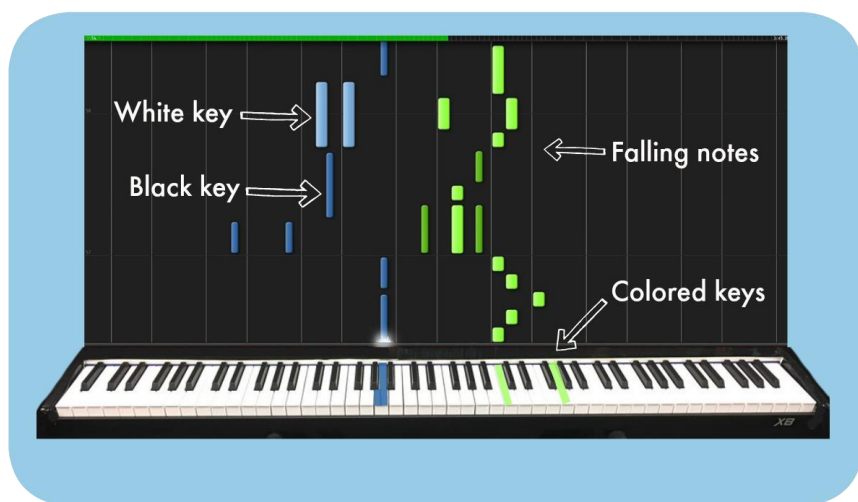
Figure 28: Second instruction of tutorial video mentioning the different sizes of falling notes
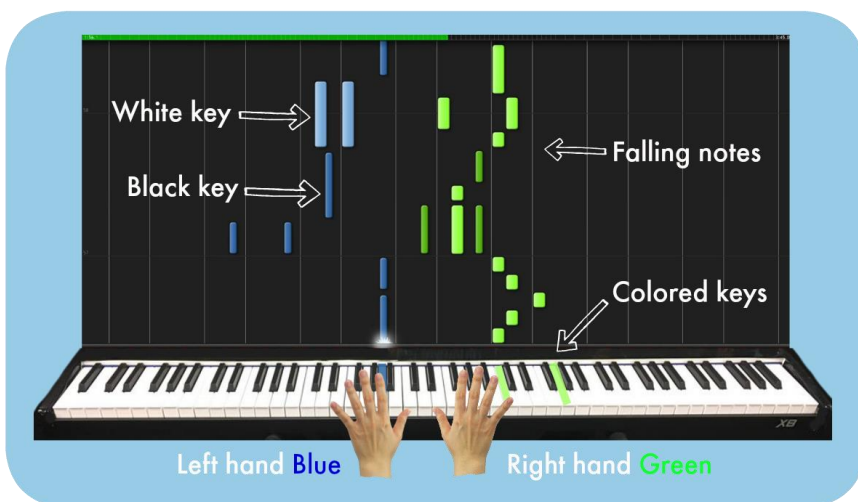
Figure 29:Third instruction of tutorial instruction video mentioning the color refers to left/ right hand

Figure 30: Tutorial instruction video showed on HoloLens
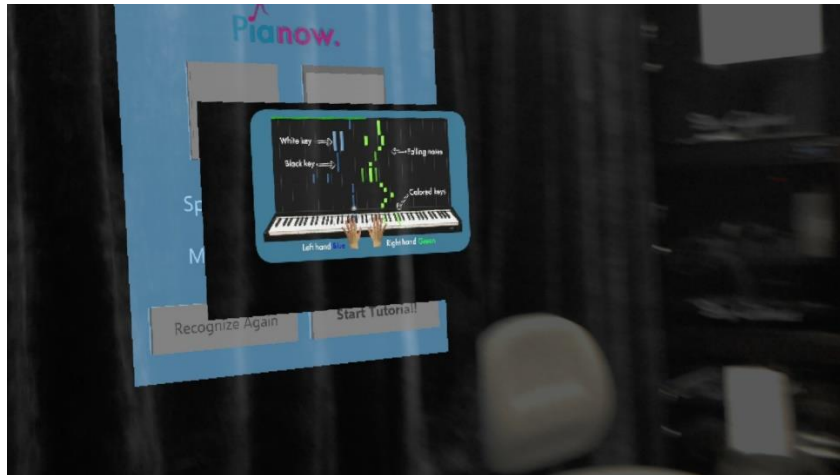
Figure 31: The tutorial insttruction video is rendered in front of the control menu

3.4 Difficulties Encountered

There are 3 major difficulties encountered in this project.

Firstly, the project team lacks experience with most of the tools, such as Unity, .NET development, HoloLens API, OpenCV, etc. This introduces a big learning curve to every development steps and requires extensive effort and patience in researching and debugging.

Secondly, most of the tools related to the HoloLens are new. Therefore, the documentations are not detailed and well-organized enough. The community is also highly immature, so there is not much support if any problem

44

is encountered during the development process. Additionally, since the technologies are still under development, there are many shortcomings. For example, different versions of Unity may cause conflicts with different versions of other software, so it is difficult to decide on a version that can integrate gracefully with every tool. Besides, as the HoloLens' locatable camera cannot be accessed from the Unity Editor, it is impossible to test the application within the editor environment. The consequence is that every testing and debugging process requires deployment to the HoloLens, which significantly slow down the development progress.

Finally, the HoloLens itself has many unexpected limitations that cause the project's approach to change several times. For instance, the video stream from the HoloLens' locatable camera has a much lower resolution than how it is claimed to be. This makes it impossible to achieve a very good recognition result and requires additional alignment method such as manual alignment by the user. Other limitations such as the low frame rate, small field of view and inaccurate spatial mapping mesh also disallowed many solutions that the project team came up with. The disable of mixed reality capture when using the locatable camera also disallows us to understand more of what the user views and compare the stability result.

4 Future Direction

4.1 Further Improvements

In the future, improvements on the robustness of Keyboard Recognition with OpenCV method will be attempted. More specifically, the underlying algorithm will be enhanced to meet the universality criterion. The reason for this is because this method, although cannot offer good results at this stage, is foreseen to be more suitable for future releases of the HoloLens (see Section 4.2 for more details). Also, the implementation of marker detection will be attempted with the Aruco marker featured of OpenCV. This is expected to remove many unused modules from Vuforia and, hence, improve the performance of the application.

In the future, we aim to have more tutorials available, allowing the users to play different music genres. The songs will be separated into different levels of difficulty to let the users choose their own preferences. In this way,

the user could obtain a sense of achievement after completing each level and have the motivation to keep learning. Furthermore, the project team is planning to add the ability to choose only left or right hand to practice in a tutorial, which offers an easier option to beginners. Additionally, the current version of HoloLens has a very limited field of view that the user cannot play both hands at the same time. Almost all participants in our preliminary test encountered the same problem. This new functionality can be a temporary solution to this.

In the further future, the project is aspired to add a personal evaluation system, which will give live feedback to the user on their performance and will also collect relevant data to personalize the user's experience. This functionality, however, may requires Bluetooth connection and requires a Bluetooth-enabled keyboard. Also, it is observed from preliminary survey that some subjects who have not previously learnt to play the piano had the incorrect hand placement. Thus, a functionality that teach users the correct hand placement, such as numbering the keys according to finger indexes, will be helpful.

4.2 Future Prospect

The next generation of HoloLens is expected to have more accurate spatial mapping and locatable camera with higher resolution. These two improvements will greatly enhance the OpenCV keyboard recognition results. Also, if the holograms are more stable or the field of view is wider, the need for an initial offset between the virtual content and the physical piano keyboard can be removed, which will enhance the reliability of the method. For this reason, the implementation of keyboard recognition in OpenCV is still believed to be the best method of recognition in the future, which is easy, fast and accurate.

Additionally, the wider field of view and more stable holograms will also improve the user experience and allow fully immersive tutorials. When the HoloLens is more accessible as a commercial product, Pianow is expected to be one of the applications that led the trend of musical education in Mixed Reality.

To ensure that Pianow has educational practicality, however, it is important to ensure that the application is effective in assisting piano learners. This will necessarily require a formal study which involves a large number of participants over a long period of time to objectively determine the extent to which the application can assist people in self-learning the piano. If the current design of the application is proven to be ineffective, there are still many possible designs that can be attempted, thanks to the flexibility of the current design. For example, most of the formal piano lessons for beginners such as learning the music notes, ear training, finger speed exercises or chord progression can be implemented for Pianow, with the use of the 3D screen and the key overlays. That is to say, Pianow still has the potential to develop evolve, and to succeed side by side with the evolution of Mixed Reality.

5   Conclusion

The objective of the project is to build a piano tutorial application on HoloLens that gives piano learners a better learning experience with its 2 guiding features: falling notes and colored keys. The main difference of this application compared to existing solutions is that all holograms are shown directly on the physical piano keyboard, which makes the tutorials easier to follow. The project uses the HoloLens, OpenCV and Vuforia during the development process. There are two major challenges to be addressed in this project: virtual asset alignment and design for effectiveness. The first objective of the project has been achieved with satisfactory results. Three methods have been implemented, including manual alignment, keyboard recognition with OpenCV and marker detection with Vuforia. Generally speaking, the third method currently offers the best alignment results. However, with view to the future improvements of the HoloLens, the sec-ond method is more prominent as it is expected to offer better user experi-ence. The UI and the tutorial contents of the application have also been implemented, and the application is ready for a demonstration. A prelimi-nary test involving selected participants from outside the development team have also been carried out to collect input on the usefulness and usa-bility of the application. The result is promising, even though there are still room for improvement.

In the future, more functionality such as hand placement lessons and evaluation system will be added. Regarding the effectiveness of the current tutorial design, more in-depth research is needed before any conclusion can be derived. Regardless of the uncertainty, the flexible design of Pianow makes

many future adjustments possible, so its potential is hard to un-dermine. In the long term, the application is anticipated to change the way self-taught pianists learn and be a good foundation for future projects which aspire to exploit the potential of Mixed Reality in musical education.

References

[1]   T. J. McCue, »Why Infographics Rule,« 2013. [Online].
        Available:
        https://www.forbes.com/sites/tjmccue/2013/01/08/what-is-an-
        infographic-and-ways-to-make-it-go-viral/#432c05dd7272.

[2]   S. M. Community, »How Traditional Piano Lessons Cripple Our
        Children,« 2018. [Online]. Available:
        http://www.softmozart.com/our-library/articles/46-how-
        traditional-piano-lessons-cripple-our-children.html.

[3]   L. Pianos, »Why are there Black and White Keys on the
        Piano?,« 2015. [Online]. Available:
        https://livingpianos.com/uncategorized/why-are-there-black-and-
        white-keys-on-the-piano/.

[4]   M. v. p. teacher, »Piano Practice Tips: The 12 Most Common
        Piano Practice Mistakes And How To Correct Them,« 2013.
        [Online]. Available: http://www.musiah.com/blog/2013/august/12-
        common-piano-practice-mistakes.aspx.

[5]   H. F. Revolution, »Top 10 Health Benefits of Playing the
        Piano,« 2015. [Online]. Available:
        http://www.healthfitnessrevolution.com/top-10-health-benefits-
        playing-piano/.

[6]   T. website, »How much do piano lessons cost?,« 2017. [Online].
        Available: https://www.thumbtack.com/p/piano-lessons-cost.

[7]   CMUSE, »How To Get Your Child Excited To Play A Musical
        Instrument,« 2014. [Online]. Available:
        https://www.cmuse.org/how-to-get-your-child-excited-to-play-a-
        musical-instrument/.

[8]   M. D. Dickey, »Teaching in 3D: Pedagogical Affordances and
        Constraints of 3D Virtual Worlds for Synchronous Distance
        Learning,« *Distance Education,* årg. 24, nr. 1, pp. 105-121, 2013.

[9]   C. Dede, »The Evolution of Constructivist Learning
        Environments: Immersion in Distributed, Virtual Worlds,«
        Educational Technology, årg. 35, nr. 5, pp. 46-52, 1995.

[10]   N. Piegdon, »Synthesia,« 2017. [Online]. Available:
http://www.synthesiagame.com/.

[11]   R. company, »Exclusive Interview with Ayato Fujii creator of
Teomirn App,« 2017. [Online]. Available:
http://www.realitydome.com/interview-ayato-fujii-teomirn-app/.

[12]   H. F.Y., S. a. L. B. Das og S. Glickman, »Music Everywhere,«
2016. [Online]. Available:
https://www.etc.cmu.edu/projects/music-everywhere/.

[13]   Google, »Synthesia,« 2018. [Online]. Available:
https://play.google.com/store/apps/details?id=com.synthesia.synt
hesia&hl=en.

[14]   Apple, »Synthesia,« 2018. [Online]. Available:
https://itunes.apple.com/us/app/synthesia/id579518445?mt=8.

[15]   G. Comeau, »Playing by ear in the Suzuki Method: Supporting
evidence and concerns in the context of piano playing,« *The
Canadian Music Teacher,* 2012.

[16]   Microsoft, »Install the tools,« 2018. [Online]. Available:
https://docs.microsoft.com/en-us/windows/mixed-reality/install-
the-tools.

[17]   E. Wong, »Principle of Consistency and Standards in User
Interface Design,« 2017. [Online]. Available:
https://www.interaction-design.org/literature/article/principle-of-
consistency-and-standards-in-user-interface-design.

[18]   m. batagoda, »Color, psychology and design,« 2017. [Online].
Available: https://uxplanet.org/how-color-can-effect-emotion-
ccab0431b1d.

[19]   W. O. Galitz, The Essential Guide to User Interface Design: An
Introduction to GUI Design, 2007.

[20]   M. co, »Hologram stability,« 2018. [Online]. Available:
https://docs.microsoft.com/en-us/windows/mixed-
reality/hologram-stability.

[21]   C. M. Barnum, Usability Testing Essentials: Ready, Set...Test!,
2011.

[22]   CXL, »User Experience Testing: A Conversion-Focused Guide,«
        2017. [Online]. Available: https://conversionxl.com/blog/user-
        experience-testing/.

[23]   M. co, »Mixed reality capture for developers,« 2018. [Online].
        Available: https://docs.microsoft.com/en-us/windows/mixed-
        reality/mixed-reality-capture-for-developers.

[24]   A. Prelicz-Zawadzka, »How to Create an Effective Customer
        Experience Survey,« 2018. [Online]. Available:
        http://www.uxbooth.com/articles/how-to-create-a-customer-
        experience-survey/.

[25]   F. V. Bennekom, »The 7 essentials of successful survey design,«
        2016. [Online]. Available: https://uxmastery.com/successful-
        survey-design/.