PIANOW - PIANO SELF-LEARNING ASSISTANT IN MIXED REALITY
PROJECT PLAN

CREATED BY

BUI, THANH TUNG (UID:3035124491)
TSAI, YI-TING (UID:3035124520)

SUPERVISOR

DR. DIRK SCHNIEDERS

The University of Hong Kong

Abstract

Synthesia is a popular application for learning to play the piano. Its idea of using falling notes and colored keys (see figure 1) has proven to be useful for many learners and has been adapted by many other applications. In this project, we aim to bring the great features of this application to Mixed Reality, a combination that will tremendously improve the learning experience, as the learner can now follow the tutorials without looking away from their own physical piano keyboard. This proposal provides details of the project's motivations, scope and objectives, the methodology to be used as well as how the project will be organized.

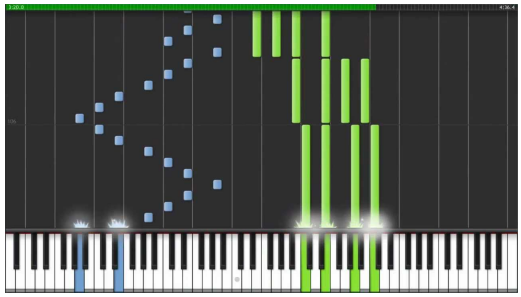

Figure 1: Synthesia-style piano tutorial

1 Introduction

1.1 Background

Mixed Reality (MR) is a recent hit in the realm of technology. As described by Microsoft, "Mixed Reality is the result of blending the physical world with the digital world" [1, para.1]. With the help of Mixed Reality technology, users nowadays, instead of being mere observers, can become part of the digital world themselves. So far, the application of this idea is most frequently encountered in the game and entertainment industry. However, Mixed Reality could be applied in many other aspects of life, such as education. Kerawalla et al. stated that the utilization of Augmented Reality (AR) for training and educational purposes has started since late 1960s, and researchers are still exploring the potential of AR in motivating learners to acquire knowledge and apply them in real life from diverse perspectives [2]. With the interactive nature, Mixed Reality presents great potential in the fields of education that requires lots of interactions such as musical learning.

Piano is actually a relatively easy instrument to pick up compared to other chordophones. Other string instruments like guitar require users to press and pick the string simultaneously [3], while piano does not. However, it's easy for people to feel intimidated to start something new without any instructions or

knowledge of reading the music sheet, which is why some people seek help from personal tutors. However, the cost of hiring a personal piano instructor is relatively high. These reasons may have disinterested many people from learning to play the piano, giving piano the image of a difficult instrument.

From time to time, we often see mobile applications that enable users to play piano on their smartphone like "Pianist HD : Piano +". However, due to the small screen size, the keyboard cannot accommodate all 10 fingers. Moreover, pressing a key stronger or lighter makes a lot of difference to the rhythm, which generally cannot be achieved by tapping on the phone screen. A more effective solution would be to supplement the physical piano keyboard rather than replacing it.

1.2 Current Solutions & Related Works

Synthesia is a piano personal trainer video game and mobile application for Microsoft Windows and Mac OS X, which allows user to follow the falling notes and colored key without the need of understanding the music sheet [4]. The power of the app lies in its intuitive interface with falling notes and colored keys. A beginner will be able to immediately grasp the idea and follow the tutorial without spending much effort. However, much like other piano mobile applications, Synthesia fails to provide the user with the same quality and experience of using a physical piano keyboard, which is precisely the problem that we would like to address in this project.

Music Everywhere is a project done by Carneige Mellon University's entertainment technology center [5]. The application utilizes HoloLens and overlays the holographic instructions in front of the keyboard to assist the learner. It also use a 3D reconstructed hand as another way to guide users through the tutorial. However, even with the animation effects, the holographic guided tips are not shown directly on the piano keyboard, which makes it difficult for users, especially beginners, to follow the instructions.

Teomirn is a HoloLens apps developed by Ayato Fujii, a developer team from Japan [6]. The application allows the users to learn from a virtual piano instructor. It is helpful to present the virtual hands of the instructor to the users so as to give them a better understanding of hand placement and movement. However, similar to Music Everywhere, the instructions are not displayed on the physical keyboard, so the user needs to keep looking at a virtual keyboard to see which note to play next. Such delay is potentially an obstacle for beginner.

1.3 Objectives

In this project, we will develop an application named Pianow. Pianow is a Synthesia-style application in Mixed Reality for piano learners, featuring 2 guiding methods: key coloring and falling notes. There are 2 major objectives to be achieved in this project:

- First, to develop an application that is able to recognize the physical piano keyboard and align virtual assets to it.
- Second, to perform hand detection with as little offset as possible for better visual effect of key coloring.

Additionally, the project also aims to understand the process of learning to play the piano as much as possible and to create at least 2 full Synthesia-style tutorials, plus a control menu for demonstration purpose.

1.4 Flow of Events

The below flow of events illustrates a clearer image of what we are going to build in this project.

1. The user opens the application and scans the piano keyboard until it is recognized by Pianow.
2. Pianow will display the control menu after the piano keyboard has been successfully detected. The control menu will allow the user to:

- Choose a song from the list of tutorials
- Customize the speed at which the tutorial will be played
- Disable / enable the built-in music of the tutorial
- Choose which hand to practice (left or right hand). If left hand is chosen, the notes will be played automatically for the right hand and the user will play the left hand by himself / herself

3. The user finishes his / her customization and choose to start the tutorial.
4. Pianow will start playing the Synthesia-style tutorial consisting of falling notes and colored keys that match the physical piano keyboard.

1.5 Scope

Pianow is an application implemented for HoloLens only. User needs to have the physical 88 keys piano keyboard for this application. Piano will be reconstructed and recognized by this application. Two main methods to guide user would be first having 2 synthesia-style tutorial and second colored key showing on the keyboard. Hand detection would be used to attempt better result for the colored key. The user interface is developed and designed within the objectives only.

2 Methodology

2.1 Prerequisites

2.1.1 Hardware

This project will make use of 2 major pieces of hardware, namely the HoloLens and the Leap Motion. The HoloLens is a powerful headset designed exclusively for Mixed Reality, and arguably the best of this category on the market. Meanwhile, the Leap Motion is generally known for its robust hand tracking mechanism with the accuracy of 0.7 millimeter on average [7]. Even though Leap Motion does not have native support for HoloLens, we still decide to combine these two components as they both excel in their respective field. A more trivial reason is that we have access to both of these two hardware pieces.

Additionally, as recommended by Microsoft for development on the HoloLens [8], the project will also require a PC system with the minimum specifications as follows:

- A 64-bit 4-core CPU
- 8GB of RAM
- A GPU that supports both DirectX 11.0 and WDDM 1.2 driver or their later versions

2.1.2 Software and Technology

For Mixed Reality development on HoloLens, the project will use the following software, as recommended by Microsoft [8].

Software	**Usage**
Visual Studio 2017	For code editing as well as compiling and deploying the application to Windows platform
Unity 2017.1	The main development platform
HoloLens Emulator	For testing the application on a virtual machine, without a physical HoloLens (the use of this program is optional, depending on the availability of the HoloLens)
Vuforia	For recognition of the real-world piano keyboard based on a pre-built 3D model

Apart from the above software, the project may also make use of Blender for modeling the piano keyboard and Synthesia-style tutorial assets. Blender is a free open-source 3D computer graphics software. Windows 10 Professional, Enterprise, or Education edition will also be required if HoloLens Emulator is used [8].

2.2 Implementation

2.2.1 Piano Keyboard Recognition

One of the main objectives to be achieved in this project is to recognize the physical piano keyboard such that virtual assets can be correctly attached to it. Essentially, we need a matching 3D reconstructed piano keyboard that is spatially aligned with the physical piano keyboard as accurately as possible. The virtual assets will then be aligned with the 3D-modelled keyboard and, hence, with the physical keyboard.

To achieve this, there are 2 major tasks we need to accomplished: modeling a piano keyboard and aligning it with the real-world piano keyboards. The former can be done from scratch with Blender. We will construct a standard full-size piano keyboard with 52 white keys and 36 black keys. The size, position and orientation of the keys must match a real-world piano of the same type, which is important for the subsequent step. After having a good model of the piano keyboard, we will fetch this model to Vuforia Model Targets so that it can start recognizing real-world piano keyboards with similar features. Model Targets is the newest feature of Vuforia object recognition that enables the recognition and tracking of objects using pre-built 3D models [9] (see figure 2). This process will automatically align the virtual keyboard with the real-world keyboard.

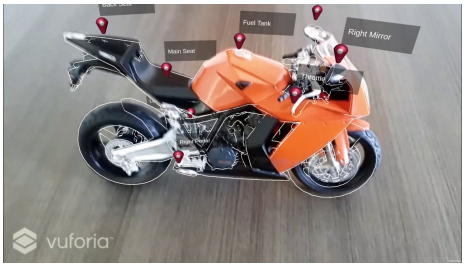

Figure 2: Object recognition and virtual asset alignment with Vuforia Model Targets

2.2.2 Key Coloring & Hand Detection

Adding color to a key after the piano keyboard has been recognized and aligned with a 3D-modelled keyboard in Unity is trivial. However, to make the effect more realistic, we do not want the virtual color to appear on top of the user's hands. This requires the powerful hand detection mechanism provided by Leap Motion. After the user's hands are recognized, they can be placed on top of the virtual color, which will allow better user experience.

2.2.3 Synthesia-style Tutorial

Key and note assets can be modelled in Unity or Blender, then animated in synchronization with the music to create a Synthesia-style tutorial. The tutorial's visuals will then be aligned with the keys in the 3D reconstructed piano keyboard

and, hence, the physical piano keyboard. How good the alignment is depends largely on how well the physical and virtual versions of the piano keyboard are aligned. For this project, we plan to create at least 2 tutorials for 2 different songs for illustration purpose.

2.2.4 Control Menu

The control menu can be modeled in Blender or Unity. A program will be written to handle the logic of this level.

2.3 Risks & Mitigation Strategies

There are 3 major risks that we will attempt to address in this project:

- Vuforia Model Targets is still an experimental feature and has not been officially released, although an early access version can be requested. This entails that the feature might not be available for this project, and even if it is, it might not work as expected. A back-up plan is to make use of the Random Sample Consensus (RANSAC) algorithm for feature-based object detection. RANSAC is able to interpret data that has a large portion of gross error, which makes it suitable for image analysis tasks [10], including object recognition.
- Since Leap Motion uses its own camera for hand tracking, there may be a significant offset between the user's real hand and it's virtual version (see figure 3). To reduce the offset, we come up with 2 approaches: to hardcode a certain level of offset when generating the virtual hand image or to force using the same camera for hand tracking and piano keyboard recognition. We will attempt both and evaluate the results in later phases.

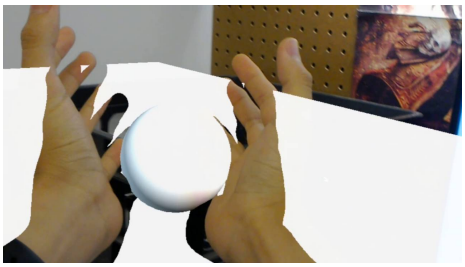

Figure 3: Leap Motion on HoloLens: offset between real and virtual hand images

- Our team do not have sufficient knowledge of the piano. There might be some misunderstanding or uncertainty when developing an application for learning how to play the piano. We will consult people who has more experience with the piano, and may also request help from the university's Music Department if necessary.

If these risks are successfully mitigated, the objectives of the project are very likely to be achievable.

2.4 Testing & Evaluation

For each component of the application, we define a criterion of success and a testing method.

Criterion of Success	**Testing Method**
The application should be able to recognize any standard full-size piano keyboard.	Test on different standard full-size piano keyboard in room lighting condition
The keys should be colored correctly with less than 2mm offset and should not block the user's view of his/her own hands. The offset the user's real hand image and virtual hand image should also be less than 2mm.	Change the perspective by moving only the head[1] to check if the augmented colors as well as the user's hands are positioned and displayed correctly
Synthesia-style tutorial assets should match the real-world keys with less than 2mm offset.	Change the perspective by moving only the head[1] to check if the augmented assets are still aligned correctly with the keys
The control menu should be able to sense the user input after at most 3 tries and should produce the expected feedback within 2 seconds.	Directly interact with the menu

We plan to invite a group of at least 10 people to perform each test and give us their feedback. This sample size is reasonable, considering the scale and scope of the project. We will try to make sure that the samples are as diverse as possible (e.g. people with different skin complexion, different hand size, etc.) to ensure the universal applicability of the application. Subjects need not to have any experience in playing piano.

[1] It is assumed that the user will not walk around the environment, check subsection 1.5 for more details

3 Project Organization

3.1 Division of work

The following table shows how the tasks have been distributed.

Task	Bui, Thanh Tung	Tsai, Yi-Ting
Proposal, webpage and reports	✓	✓
First and final Presentation	✓	✓
Piano keyboard recognition	✓	✓
Key coloring and hand tracking	✓	✓
Synthesia-style tutorials	✓	
Control menu		✓
Exhibition preparation	✓	✓
Testing and evaluation	✓	✓

3.2 Management of Budget & Materials

The team has a total budget of HKD$2000.

Most of the software components we use either are open-source or has a free edition that suffice our needs. A free version of Windows 10 Education can also be obtained from the department's intranet. The Multimedia Laboratory (Room 310, Haking Wong Building) that we have been given access to already has the Leap Motion and a sufficiently powerful PC system. The HoloLens can be booked and borrowed from the Graduate House under the condition that they are returned properly.

Hence, the only additional equipment we may require is a portable 88-key electrical piano for development, testing and demonstration purpose. The piano is priced at around HKD$900 to HKD$1000 on Taobao. We plan to place it in the Multimedia Laboratory for continuous use. The remaining of the budget will be reserved in case we need any add-on packages or resources.

3.3 Schedule

The following schedule presents the deliverables and milestones that have to be delivered by a designated time. Words in **bold** indicate hurdle requirements.

Time	**Deliverables**
1 October	Deliverables of phase 1 - **Project proposal plan** - **Project website**
Mid November	- 3D Reconstruct and recognize piano
End of December	- Testing result with colored key
8-12 January	**First Presentation** - present preliminary implementation
21 January	Deliverables of phase 2 - **Interim report** - Report 3D reconstruct piano result
End of February	- Integrate with Leap Motion hand detection - Develop 2 Synthesia-style tutorials
End of March	- Develop the user interface - Testing and evaluation
15 April	Deliverables of phase 3 - **Final implementation** - **Final report** - Improve user interface design
16-20 April	**Final Presentation**
2 May	**Project Exhibition**
30 May	**Project Competition** (For selected teams only)

4 Conclusion

The objective of the project is to build a Synthesia-stlye piano application for HoloLens that gives piano learners a better learning experience with its 2 guiding features: falling notes and key coloring. 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 will make use of HoloLens, Leap Motion, Unity and Vuforia in the development process. The two major risks that need to be addressed are the unavailability and instability of Vuforia Model Targets feature and the fact that Leap Motion does not have native support for HoloLens. If these two risks can be mitigated, the objectives of the project are likely to be achievable.

References

[1] Microsoft. (2017). Mixed reality, [Online]. Available: https://developer.microsoft.com/en-us/windows/mixed-reality/mixed_reality.

[2] L. Kerawalla, R. Luckin, S. Seljeflot, and A. Woolard, "Making it real: Exploring the potential of augmented reality for teaching primary school," *Virtual Reality*, vol. 10, no. 3, pp. 163–174, 2006. DOI: https://dl.acm.org/citation.cfm?id=1183176.

[3] G. music studio. (2015). Is piano or guitar easier to learn? [Online]. Available: http://gillespiemusic.com/is-piano-or-guitar-easier-to-learn/.

[4] N. Piegdon. (2017). Synthesia, [Online]. Available: http://www.synthesiagame.com/.

[5] S. Glickman, H. F.Y., S. Das, and B. Lee. (2016). Music everywhere, [Online]. Available: https://www.etc.cmu.edu/projects/music-everywhere/.

[6] RealityDome. (2017). Exclusive interview with ayato fujii creator of teomirn app, [Online]. Available: http://www.realitydome.com/interview-ayato-fujii-teomirn-app/.

[7] F. Weicher, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the leap motion controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, 2013. DOI: http://dx.doi.org/10.3390/s130506380.

[8] Microsoft. (2017). Install the tools, [Online]. Available: https://developer.microsoft.com/en-us/windows/mixed-reality/install_the_tools.

[9] Vuforia. (2017). Model targets early access program, [Online]. Available: http://info.vuforia.com/modeltargets.

[10] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981. DOI: http://dx.doi.org/10.1145/358669.358692.