FunPianoAR: A Novel AR Application for Piano Learning Considering Paired Play Based on Multi-Marker Tracking

Hong Zeng¹, Xingxi He¹, * and Honghu Pan¹

¹Department of Mechatronics, College of Mechanical Engineering, Chongqing University, Chongqing, 400044, China

Email: xingxi@cqu.edu.cn

Abstract. This paper describes the design and implementation of a novel augmented reality application, referred as FunPianoAR, that aiming to activate the interests of user and improve the experience of piano learning. The FunPianoAR with a user-friendly interface considers the paired play to further reduce the difficulty of playing the piano for adult novices. For the paired play, one user plays the melody and the other plays the harmony part. The app is developed by using Android Studio 3.0 and artoolkitX 1.0, an open source augmented reality SDK, and installed on the AR smart glasses, Epson Moverio BT300. Due to the textureless features of real piano and registration precision, the application is implemented by using the fiducial marker tracking instead of the markerless recognition. Besides, we divided the piano keyboard into four zones and each zone used a separate marker for tracking to solve the limited field of view (FOV) to some extent. The virtual keys can be accurately superimposed on the piano keys by using multi-marker tracking. An evaluation was also conducted to compare the effects of two types of augmented reality information superimposed, i.e., the instant way without the hints for the next note to be played and the FunPianoAR with the hints for the coming notes to be played. The correctness rate was calculated and the time differences between the notes that two players should play at the same time were collected. The FunPianoAR shows more advantages over the instant way based on the evaluation results.

1. Introduction
It is a prolonged challenge for adult novices to learn the piano because of the difficulty of learning sheet music notation and its mapping to the piano keys [1]. Augmented Reality (AR), a technology where the user sees ‘the real world, with virtual objects superimposed upon or composited with the real world [2], has become a hot research topic in recent years. Therefore, there has been quite a few AR systems or applications proposed by researchers [3-5] to help learn piano. By projecting information onto the piano keys, an AR application or system can ease the difficulty of learning piano for novices because it is no need to learn sheet music notation. Some applications used a virtual piano keyboard to replace the real piano keyboard. for example, Chung-Hsuan at al. [6] presented the Mr. Piano, which is a portable piano tutoring system, and this system can give instructions and feedbacks that help a new learner on fundamental piano skills and the piano's scores with a virtual piano keyboard. Although using the virtual piano keys reduces the cost of learning the piano and makes it easier to learn the piano, there is no real sense of touching the piano keys.

Other applications or systems using a real piano or an electronic piano can be divided into two groups, the first one adopted a fixed projection form, referring to projecting the AR part onto the piano
with a fixed projector, and the second group used a movable equipment such as HMDs.

**Fixed projection form:** Rogers at al. [1] developed P.I.A.N.O to support learning to play the piano with interactive projection, and it turned out that this system could faster learning, require less cognitive load and provide a better user experience. Augmented songbook [7], an augmented reality educational application for young children to raise their music awareness, has a fixed projector to project the AR part onto the paper documents.

**Movable form using HMDs:** Das at al. [4] designed an augmented reality piano learning tool, using a Microsoft HoloLens and an electric piano. The users can watch the virtual hand demonstrations, see and hear example improvisations, and play their own solos. HoloKeys [3] runs on the HoloLens and presents two different ways, called the instant way (no hints for the next note to be played) and the Beatmania way (there are some hints for the coming notes to be played), to superimpose the virtual keys on the real piano. To motivate beginners to learn piano, Honghu at al. [8] developed an augmented reality application that allows for single play and paired play using HMDs.

According to our past work, we found that the user interface will greatly affect the user experience, and a well-designed UI will not only provide a completely better experience, but will also active the fun of learning the piano. The FunPianoAR prompts to draw different virtual keys on the real piano keys to be pressed at the current and the next moment, so that the UI of the app is more concise and helpful for piano learning.

In addition, in order to stimulate the user to learn the piano more, we have added a new and interesting learning mode, which allows the paired play, i.e., one learner plays the right-hand melody and the other partner plays the left-hand harmony.

### 2. The app design

This section mainly discusses the compositions of the APP from the aspects of hardware, software interface and the details of the graphic augmentations. Figure 1 shows a real piano with four fiducial markers for tracking and registration which are pasted by using the double-sided tape and are easy to be teared off. Marker A is on the note f, marker B on the note f1, marker C on the f2, and maker D on the f3. The FunPianoAR is installed on an AR smart glasses, Epson Moverio BT300. The Epson Moverio BT300 weighs 69g with binocular see-through viewing and running Android 5.1. The user sits in front of the piano and wear the smart glasses. Through the smart glasses the user sees a video stream with augmentations including virtual keys, boxes and arrows on the top of the real keyboard.

![Figure 1. A real physical piano keyboard with four markers.](image)

As far as the software interface aspect, a user-friendly interface is designed in the app. The user menus are shown in the Figure 2 in which the left-side main menu offers two options i.e., the “Play Mode” and “Practice Mode”. The practice mode is to get the users familiar with the interface of the app and the wearing of the smart glasses. In this mode, four markers will be used and all the augmentation types will be involved in the practice. The right-side menu is displayed when the “Play Mode” is selected and then the user can select songs, the play speed, and the hand mode, such as the right-hand melody and the left-hand harmony part.
There are three kinds of graphic augmentations provided. The first one is a solid cube, superimposed on the corresponding piano key, indicating the key that needs to be pressed at the current moment. The user is supposed to press and hold the key until the virtual key disappears. A green cube is shown in the Figure 3 (a) and a blue cube in (d) where the colour change between green and blue means the multiple presses on the same key instead of pressing and holding the same key all the time. A red cube means that a black key should be pressed. The second type is a green box or a frame shown in the Figure 3 (a) and (d) beside the cube. The locations of these boxes indicate the keys that needs to be pressed at the next moment. The third type of augmentations is arrows which are shown in the Figure 3 (b) and (c). They are supposed to remind the user of switching the view to different zones, e.g., the three arrows in the Figure 3 (b) indicates that the next virtual key or the green box will show up in zone D. The arrow in the Figure 3 (c) indicates that the user should shift the sight to the left, i.e. to the zone C for the next virtual key will be in zone C and the registration will be located on the marker C instead of marker D. It is worth mentioning that the box and the arrow will not be drawn at the same time.
3. Technology Implementation

This section will delve into the implementation details of the FunPianoAR, including the details of multi-marker tracking based on ARToolkitX [9] and some of the used tools. The application is developed by using Android studio and installed on the BT300, running on Android 5.1.

3.1 ARToolkitX

ARToolkitX is the latest version of artoolkit. We use ARToolkitX to identify and track the marker, calculate the homography matrix to get the camera's gesture, and then help to superimpose the augmented reality information at a specific location. With the help of ARToolkitX, the HMDs handle tracking by recognizing the image markers, computes the accurate position of the augmentations and renders the scene in real time.

3.2 Multi-marker Tracking:

The main shortcoming of BT300, i.e. the limited FOV, poses a thorny problem for the piano learning application. In order to solve this problem, the multi-maker tracking is used to divide the piano keyboard into four zones and each zone of about one octave has a specific marker for tracking and registration, illustrated in Figure 2. The use of this technology also solves the problem of large registration error when the distance to the marker is too far so that the more stable and clearer augmentations are achieved. This app uses four fiducial markers. The registration process includes the following steps: first, the camera's internal parameter matrix is obtained by the checkerboard calibration method [10]. Then, the homography matrix is calculated by the ARToolkitX's marker recognition and tracking technology. Next, the camera poses in combination with the two. Finally, the augmented reality information can be superimposed on the real world according to the camera's real-world posture.

Due to the use of multiple marker tracking, it might be easy to cause misidentification. To eliminate misrecognition, based on the hypothesis that misidentification occurs almost simultaneously in two consecutive frames, the flow chart for a marker recognition method considering two consecutive frames is shown in Figure 4. Based on the known keys to be pressed, the corresponding marker ID is also known. If the recognized marker ID does not agree with the desired ID, to avoid the misidentification, both the IDs from the current and the next frames will be compared.

3.3 Rendering the augmentations

This app uses a two-threaded approach. Tracking and rendering are done in the same thread, and the location of the augmented reality information is changed in the other thread. Finally, the frame rate of FunPianoAR application is about 15FPS. Except the rendering APIs in the ARToolkitX, OpenGL ES is also used to draw the augmentations on the virtual screen for its convenience in handling.

4. Evaluation

In order to compare the influences of different augmentations displayed i.e., the instant way without the hints for the next note to be played and the FunPianoAR with the hints for the coming notes to be played, an evaluation experiment was conducted. The paired play was adopted so that one volunteer played the right-hand melody and the other partner played the left-hand harmony. The two participants are adult novices, not familiar with the five-line staff or piano playing. After playing the practice mode for some time, the participants were asked to play the simplified version of four songs for a testing purpose. The playing process was recorded for result analysis.

Currently the playing accuracy rate only considers the keys pressed correctly and the duration of each note will not be evaluated. Therefore, the accuracy rate is equal to the number of keys pressed correctly divided by the total number of notes to be played in a song.
**Figure 4.** Augmentation drawing process based on multiple markers. “Det” means detection and recognition.

Results and discussion

The calculated accuracy rate is illustrated in Figure 5(a). The x values of 1-4 represent the four songs played by the first participant and 5-8 for the second participant. The mean accuracy of FunPianoAR is 0.93817 while the instant way is 0.90434. Figure 5(a) shows that the accuracy of FunPianoAR is higher than that of the instant way except for a couple of points. It is concluded that in general using the FunPianoAR for augmentation display can yield higher correctness rate than using the instant way. In addition, a couple of points, such as x=7 and 8, where the lower correctness rates obtained for the FunPianoAR are analysed. It is found when those songs are played, sometimes the cubes cannot be drawn because the participant does not turn their sight into the correct zone in time so that the registration on the correct marker cannot be accomplished. As the participants are more familiar with the system, these errors are expected to be reduced.

![Figure 5(a)](image1)

![Figure 5(b)](image2)

**Figure 5.** (a) The accuracy of two types of augmentations. The x values of 1-4 represent the four songs played by the participant 1, and 5-8 by the participant 2.

(b) The time difference curves for two AR interfaces.

When the single play is applied, it is harder to play the different notes for both hands with different tempos but easier to play the notes at the same time. On the contrary, for the paired play, it is easier for both participants to play the different tempos but harder for both of them to play the notes at the same time. Figure 5(b) shows the time difference curves for the two AR interfaces when the paired play is applied. The 6 pair of notes which needs to be played together by the two participants are considered. The average value of time difference for the FunPianoAR is 0.157s (SD=0.092) and for the instant
way is 0.217s (SD=0.1598). The results show that by using the FunPianoAR, it is not guaranteed that the time difference will always be smaller for at least there is no hint for playing together. However, by adding a prompt box, the FunPianoAR yields fewer points with larger deviations. This kind of discrete point happens when the notes appear completely outside the user’s expectation. The FunPianoAR can avoid such “accidents” more and naturally reduces the mean time difference of playing those notes.

5. Conclusion and future work
In conclusion, first the main contributions of this article are summarized:(1) In order to solve limited FOV to some extent, the multi-maker tracking has been used to generate a more stable and clearer augmentations. The piano keyboard is divided into four zones and each zone had a specific marker for tracking and registration. (2) The FunPianoAR presented a novel way to display augmentation on the real piano keyboard, which include the cube, the box and the arrow. (3) An evaluation was done to study the influences of the different augmentation display. It is found that the FunPianoAR is superior to the instant way. (4) This application also considered the paired play mode that one learner plays the right-hand melody and the other partner plays the left-hand harmony.

Our future work will solve the problem of drawing the augmentations when the user does not switch the sight in time. In addition, markerless recognition will be explored in this application.

Reference
[1] Rogers, K., Röhlig, A., Weing, M., Gugenheimer, J., Könings, B., Klepsch, M., ... & Weber, M. (2014, November). Piano: Faster piano learning with interactive projection. In Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces (pp. 149-158). ACM.
[2] Azuma, R. T. (1997). A survey of augmented reality. Presence: Teleoperators & Virtual Environments, 6(4), 355-385.
[3] Hackl, D., & Anthes, C. HoloKeys-An Augmented Reality Application for Learning the Piano.
[4] Das, S., Glickman, S., Hsiao, F. Y., & Lee, B. Music Everywhere—Augmented Reality Piano Improvisation Learning System.
[5] Huang, F., Zhou, Y., Yu, Y., Wang, Z., & Du, S. (2011, August). Piano ar: A markerless augmented reality based piano teaching system. In Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2011 International Conference on (Vol. 2, pp. 47-52). IEEE.
[6] Sun, C. H., & Chiang, P. Y. (2018, August). Mr. Piano: A Portable Piano Tutoring System. In 2018 IEEE XXV International Conference on Electronics, Electrical Engineering and Computing (INTERCON) (pp. 1-4). IEEE.
[7] Rusiñol, M., Chazalon, J., & Diaz-Chito, K. (2018). Augmented songbook: an augmented reality educational application for raising music awareness. Multimedia Tools and Applications, 77(11), 13773-13798.
[8] Pan, H., He, X., Zeng, H., Zhou, J., & Tang, S. (2018, July). Pilot Study of Piano Learning with AR Smart Glasses Considering Both Single and Paired Play. In International Conference on Human Aspects of IT for the Aged Population(pp. 561-570). Springer, Cham.
[9] ARToolkitX, http://www.artoolkitx.org/.
[10] Zhang, Z., Matsushita, Y., & Ma, Y. (2011, June). Camera calibration with lens distortion from low-rank textures. In Computer vision and pattern recognition (CVPR), 2011 IEEE conference on (pp. 2321-2328). IEEE.