Piano Instruction Using AR Hand Presentation and Hand Synchronization by EMS

SHINTARO MORI, University of Tsukuba, Japan
KEIICHI IHARA, University of Tsukuba, Japan
IKKAKU KAWAGUCHI, University of Tsukuba, Japan

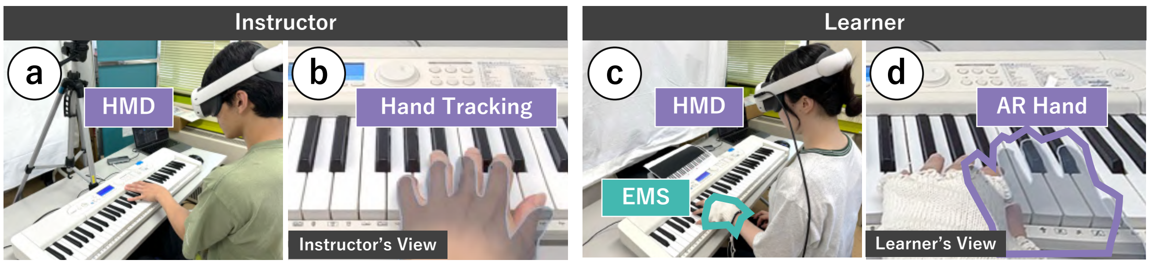

Fig. 1. Overview of the AR and EMS-based piano instruction system. (a) The instructor wears an HMD to capture hand movements during piano performance. (b) H and tracking of the instructor's performance. (c) The learner wears an HMD for AR visualization and EMS electrode pads for tactile feedback. (d) The learner views the instructor's movements as the AR hand and receives synchronized EMS for key presses.

This study addresses the challenge of improving hand positioning, timing, and the uniformity of keystroke strength for beginner piano students by integrating Augmented Reality (AR) for visual guidance with Electrical Muscle Stimulation (EMS) for tactile feedback. Our system is designed to provide immediate, synchronized cues, enhancing the learning experience. During the implementation process, we conducted a preliminary experiment, which led to refinements in both the system and the experimental setup. Following these adjustments, we conducted an experiment. Qualitative results indicate that AR significantly aids in accurate hand positioning, while EMS offers effective timing cues. However, quantitative improvements were not statistically significant, likely due to a small sample size and the limited scope of performance tasks. These findings suggest the need for further refinement of the system, particularly in addressing technical challenges such as AR misalignment and EMS discomfort.

CCS Concepts: • **Human-centered computing** → **Human computer interaction (HCI)**.

Additional Key Words and Phrases: Augmented Reality (AR), Electrical Muscle Stimulation (EMS), Piano Instruction

ACM Reference Format:

Shintaro Mori, Keiichi Ihara, and Ikkaku Kawaguchi. 2018. Piano Instruction Using AR Hand Presentation and Hand Synchronization by EMS. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX).* ACM, New York, NY, USA, 17 pages. https://doi.org/XXXXXXX.XXXXXXX

Authors' Contact Information: Shintaro Mori, University of Tsukuba, Japan; Keiichi Ihara, University of Tsukuba, Japan; Ikkaku Kawaguchi, University of Tsukuba, Japan.

1 INTRODUCTION

In recent years, the number of people who have started playing musical instruments has increased, and online instruction has become popular. However, in the instruction of musical instruments, especially for fine motor skills such as piano playing, online environments present unique challenges. A core aspect of piano performance instruction is the accurate observation and imitation of the instructor's hand movements. Online instruction typically requires multiple cameras and video switching to provide the necessary viewpoints, which imposes technical and operational burdens [8]. Additionally, the fixed viewpoints of these cameras often fail to capture the nuanced finger movements and angles crucial for effective learning.

Previous research has explored methods to overcome these challenges using AR and EMS. In the method using AR [3, 13], the instructor's hand during piano performance is converted into an AR hand, and learners can view the hand shape through a head-mounted display (HMD), allowing instructors to efficiently convey the hand position and movements on the keyboard to learners. Nevertheless, AR hand alone does not effectively communicate the precise timing of key presses, which is critical for piano performance. Additionally, capturing the instructor's sound expression is challenging in AR hand. To mitigate this, a piano roll is used to visually indicate the timing and duration of key presses [1, 3, 27]. However, adding visual elements such as a piano roll along with AR hand to indicate timing may introduce cognitive load [26]. On the other hand, EMS can guide learners in applying the correct pressure on keys by directly stimulating finger muscles [31], but it lacks the spatial guidance needed to inform learners of proper hand positioning. Moreover, the effectiveness of EMS in piano instruction has not been thoroughly evaluated.

This study addresses these gaps by proposing a system that combines AR and EMS to support piano performance learning. The system displays the instructor's hand movements as an AR hand through the learner's HMD while synchronizing EMS signals to guide the learner's finger selection and timing. By integrating these two technologies, the system aims to enhance the learner's ability to accurately mimic the instructor's hand positioning, timing, and keystroke strength. This paper presents the implementation of this system and explores its potential to overcome the limitations of previous methods in online piano instruction.

2 RELATED WORK

Supporting piano performance can be approached in several ways. A learner-based approach to piano instruction focuses on: (1) synchronizing movements and postures, (2) improving sight-reading, (3) enhancing motivation, and (4) encouraging improvisation [6]. This study utilizes AR and EMS to support piano learning. This section reviews research on AR and EMS methods relevant to these aspects.

2.1 AR in Piano Education

AR has been used in piano education in various ways. One approach involves displaying the next notes to play using AR, easing the learning process for beginners [1, 9, 16, 22, 37]. Another involves projecting a piano roll onto the keyboard, which can simplify learning by eliminating traditional sheet music [1, 3, 27].

Expanding on these AR techniques, systems like HoloKeys combine AR with the Internet of Things (IoT) to provide real-time demonstrations and feedback, enabling remote piano lessons [30]. Similarly, the On-call Piano Sensei system offers a portable AR solution that allows piano learning even without access to a physical piano, making piano education more accessible [5].

In addition to these systems, some studies use AR hand to demonstrate hand positioning and technique. For example, systems developed by Labrou et al. and Cai et al. use hand tracking to show instructor hand shapes through HMD[3, 13]. However, relying solely on the AR hand can obscure key press timing and sound subtleties. Combining AR with a piano roll could help, though it may add cognitive load[26]. To mitigate these issues, this study proposes using EMS instead of AR to assist with the timing of key presses and to ensure uniformity in key press strength.

2.2 EMS in Music Learning

EMS has been explored in various contexts to support instrument performance by providing electrical stimulation to specific muscles, aiding in the accurate execution of movements. For example, the Stimulated Percussions system incorporates EMS into percussion learning, assisting learners with mastering complex rhythm patterns [7]. In the context of piano performance, Niijima et al. [23, 25] used EMS to suppress unnecessary muscle activity in novice pianists during tremolo performance, though their study was limited to this specific technique. Furthermore, Niijima et al. proposed a system that uses EMS to help beginners achieve evenness when playing piano scales, a task that requires precise muscle coordination [24]. PossessedHand is another system that uses EMS to control the user's fingers by stimulating the muscles in the forearm, enabling independent control of the hand's 16 joints [33]. This system has been shown to reduce timing errors and string errors in koto (a Japanese harp) performance. However, it can only independently control the proximal interphalangeal (PIP) joints, excluding the thumb, due to the interconnected nature of the finger flexor muscles in the forearm. To overcome this limitation, Takahashi et al. provided electrical stimulation to the lumbrical muscles, which control the metacarpophalangeal (MCP) joints, allowing for independent control of these joints [31]. Building on this work, this study aims to use EMS to mimic the instructor's finger movements on the learner's fingers, providing a novel approach to piano instruction that addresses the limitations of both AR and traditional methods.

3 SYSTEM

We present a piano instruction system that integrates both AR hand and EMS to enhance the learning experience. In this system, an HMD displays an AR hand modeled after the instructor's hand, helping learners accurately replicate hand movements and maintain proper hand positioning during performance. Additionally, EMS provides tactile feedback, improving the precision of keystroke timing and keystroke strength with the fingers. In the following section, we outline the implementation structure of these features within the proposed system.

3.1 System Configuration

Figure 2 shows the configuration of our proposed system. We employed Meta Quest 3 [18] as the HMD and utilized Unity (version 2022.3.09f1) as the development platform for our application. To convert the instructor's hand movements into AR hand, hand tracking was performed using Meta Quest 3. At the same time as acquiring hand tracking information, MIDI (Musical Instrument Digital Interface) information was also recorded. In the Arduino-controlled EMS system, the stimulation timing is determined by MIDI data, and finger selection is manually set based on the instructor's performance. Both the instructor and learner used a Casio LK-520 [4] piano.

3.2 Acquisition of Instructor's Hand Movements in Piano Performance

We acquired the instructor's hand movements in piano performance using hand tracking with Meta Quest 3. However, to obtain precise keystroke information, we also utilized MIDI data from the keyboard.

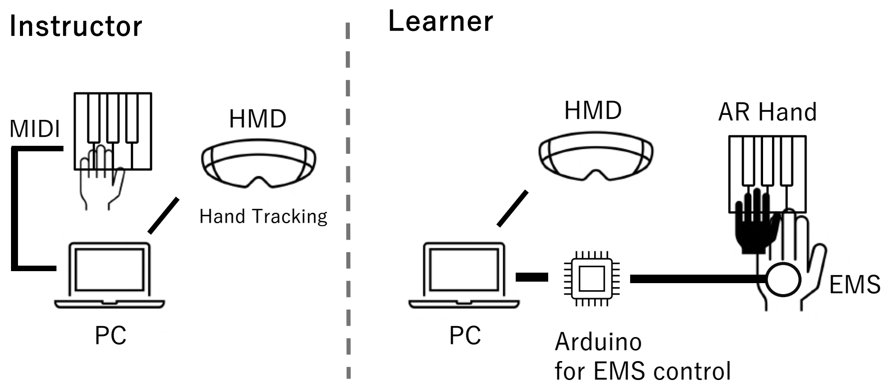

Fig. 2. System architecture for the piano instruction system using AR and EMS. The instructor uses an HMD for hand tracking and a PC for MIDI data capture. The learner uses an HMD for AR visualization, a PC for data processing, and an Arduino-controlled EMS system for tactile feedback.

3.2.1 Hand Tracking and AR Hand Visualization. To capture the instructor's hand movements, we employed the Meta XR SDK [19] for hand tracking within Unity. The tracked hand data are mapped to an AR hand model, which is then recorded as an animation using Unity Recorder [34] at 60 FPS.

The AR hand object used the standard semitransparent texture of Meta XR SDK (Figure 1(d)). To ensure precise spatial alignment between the AR hand and the physical piano, we implemented a calibration procedure using a virtual piano model. This virtual model, identical in size to the physical piano, was placed in the AR space as a reference object. We calibrated the system through a two-step process: first, aligning the virtual piano with the AR hand in the AR environment, then adjusting the virtual piano to overlap with the physical piano. After this position adjustment, the virtual piano was made invisible.

3.2.2 MIDI Information. Hand tracking data acquired from Meta Quest 3 was insufficient for accurate keystroke detection and precise EMS control. Therefore, keystroke or key release positions and keystroke strength (velocity) were obtained from the MIDI information of the piano performance by the instructor. MIDI velocity data provides 128 distinct levels of keystroke intensity. We then processed this detailed MIDI input to determine the optimal timing and finger selection for EMS activation, enhancing the system's accuracy and responsiveness.

3.3 EMS Control

We developed an EMS control system to provide tactile feedback to the learner's fingers during piano practice. The system comprises an Arduino microcontroller, PhotoMOS relays, and a Sanitas SEM 43 [28] as the EMS generator (Figure 3). The Arduino controls the PhotoMOS relays to switch the EMS output ON/OFF, allowing for precise timing of stimulation. In piano performance, when playing a scale, the MCP joint flexes and the PIP joint extends. When playing a chord, the wrist, the MCP joint, and the DIP joint (distal interphalangeal joint) flex [29]. Therefore, following the work of Takahashi et al. [31], we placed the electrode pads on the lumbrical muscles of the hand, which can independently control the MCP joint, to facilitate flexion of the MCP joint during scale and chord playing (Figure 4). However, since the thumb does not have a lumbrical muscle, we placed the electrode on the short abductor muscle of the thumb. The

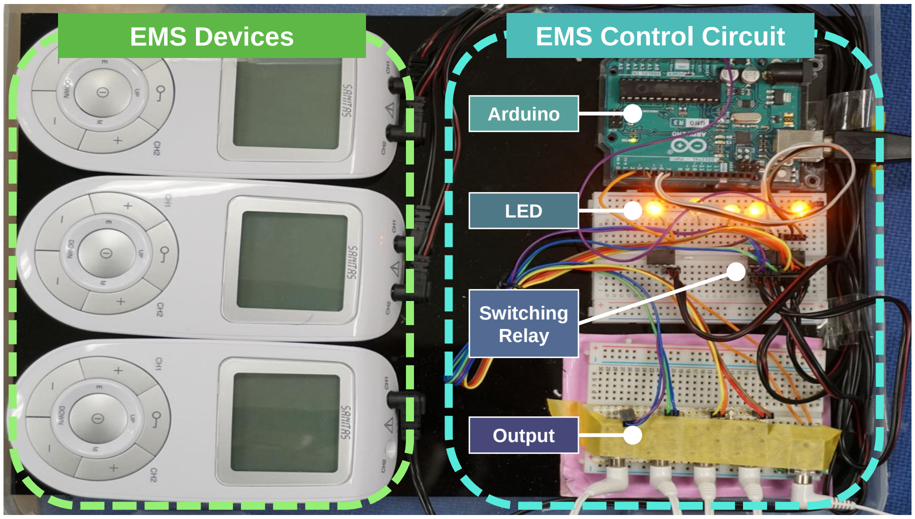

Fig. 3. EMS control module. Left: EMS devices (Sanitas SEM 43) used as stimulation generators. Right: EMS control circuit featuring an Arduino microcontroller for precise timing control, LED indicators for visual feedback, PhotoMOS relays for switching, and output terminals for connecting to electrode pads.

size of the electrode pads was 1×3 cm for the lumbrical muscles and 3×3 cm for the wrist. The strength of the EMS was set to a level where the participants could feel which finger was receiving the EMS stimulation. The use of the EMS generator was limited to a frequency of 150 Hz, a pulse width of 450 μs or less, and a maximum current of 13.6 mA (with 1kΩ resistance), with a total usage time of up to 15 minutes per day. The safety of such EMS stimulation has been confirmed by the guidelines proposed for research purposes [11, 12, 36].

3.4 Playback Interface

To enable learners to play the instructor's piano performance, we implemented a playback interface. This interface plays the instructor's piano performance sound and video while providing AR and EMS feedback according to the playback status. In the experiment, to evaluate the effect of AR, we implemented two different playback interfaces: one for conditions using an HMD and another for conditions using a standard display (Figure 5). The playback interface includes a Start button to replay the instruction screen and a Stop button to stop the replay. A slider indicates the progress of the video and allows the user to check the current position during playback. For conditions using the HMD, the AR buttons are operated by hand gestures. For conditions not using the HMD, the buttons on the display are operated by a mouse.

3.5 Timing Adjustment

Precise synchronization between AR hand animations, EMS feedback, and the instructor's performance is critical for the system's effectiveness. Despite simultaneous capture of video, AR animation, and MIDI data, we observed timing discrepancies during playback. These misalignments likely stem from varying processing delays in different system components. To address this, we implemented a manual calibration process, using the piano performance video as a reference to adjust the timing of AR hand animations and EMS activations. This calibration process fine-tunes the

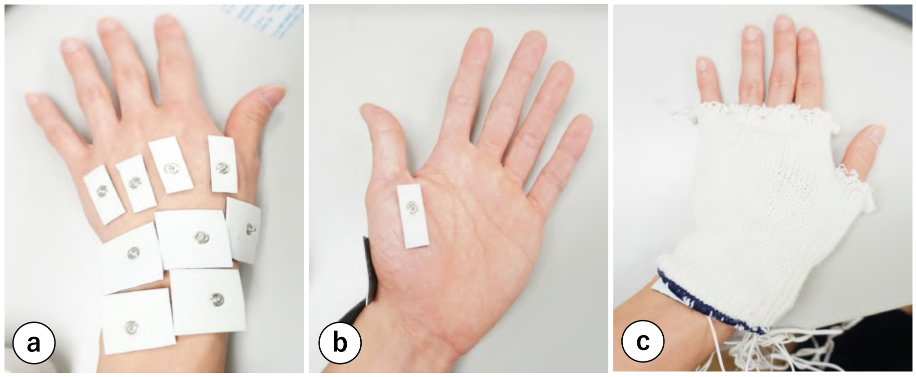

Fig. 4. Electrode pad attachment position. (a) The electrode pads are attached to the lumbrical muscles of the left hand except for the thumb. The wrist electrode pad is used as GND. (b) The electrode pads are attached to the short abductor muscle of the thumb. (c) During the task, the participants wore gloves with the finger parts removed to prevent the electrode pads from coming off.

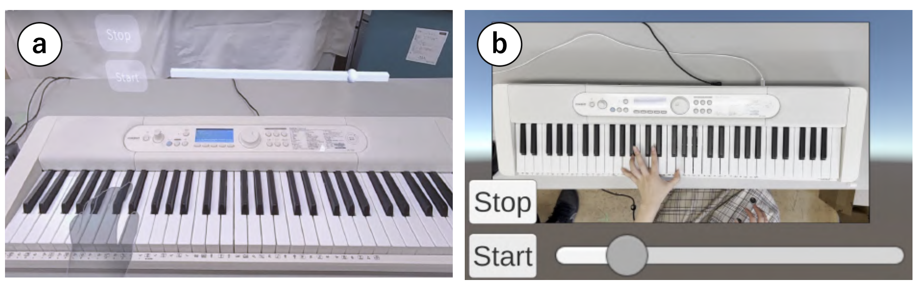

Fig. 5. Playback interface. (a) Interface used for AR and AR × EMS conditions with HMD, showing AR buttons operated by hand gestures. (b) Interface used for EMS-only conditions without HMD, featuring on-screen buttons operated by a mouse.

synchronization based on key moments in the instructor's performance, such as finger presses and releases, to correct for processing delays within the system.

4 PRELIMINARY EXPERIMENT

As part of the system implementation process described in Section 3.1, we conducted a preliminary experiment to explore the application of our proposed system in a simulated remote piano instruction scenario. It is important to note that the system configuration differed from the one described in Section 3.1. The main objectives were to identify areas for improvement, establish design guidelines, and determine an appropriate experimental design for future studies. We compared four conditions combining visual factors and EMS factors in piano instruction.

4.1 System Configuration and Experimental Environment

We set up a real-time remote instruction environment. The experimental environment was divided into two sections: one for the instructor and one for the participant, as shown in Figure 6. The instructor's hand tracking was performed using a webcam and MediaPipe [17], with the data sent to the participant. The instructor used an iPad to monitor rhythm and observed the participant's performance on a display. In the participant's area, hand movements were captured and shared via the control PC. AR hand were displayed using HoloLens 2 [20], which rendered them as a skeletal model. Figures 3 and 4 illustrate the setup for the EMS and AR conditions.

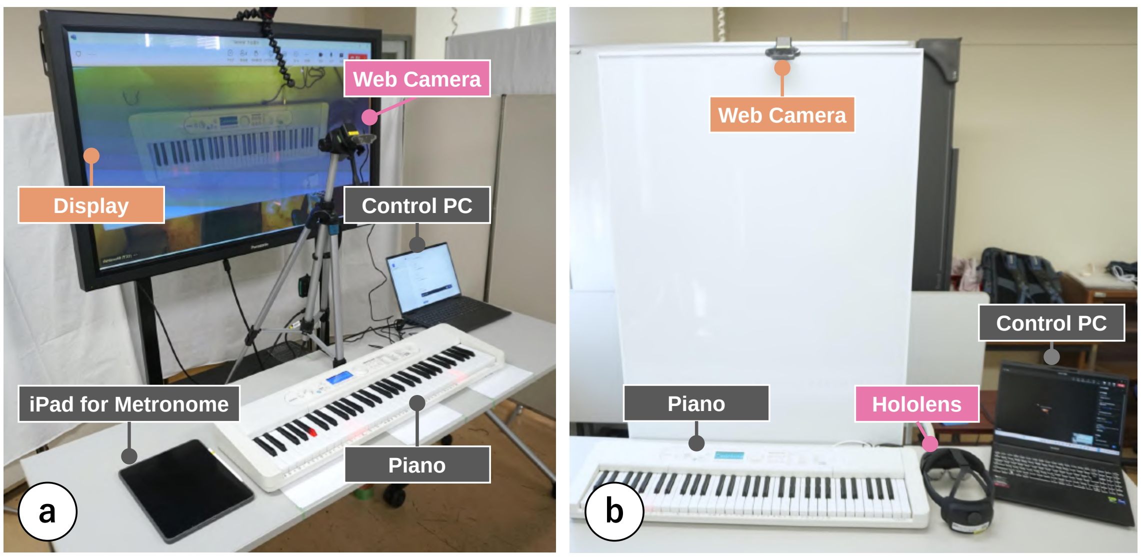

Fig. 6. Experimental environment setup. (a) Instructor's environment featuring a webcam, display, control PC, iPad for metronome, and piano. (b) Participant's environment including a webcam, Hololens, control PC, and piano.

4.2 Instruction Conditions

To verify the effectiveness of AR and EMS, we compared four instruction conditions combining AR visual information and EMS feedback:

- Video: Participants learned without AR and EMS, watching a video of the instructor's hand from an overhead perspective using the Microsoft Teams [21] video conferencing tool.
- AR: Participants learned without EMS, following the instructor's hand using an AR hand model.
- EMS: Participants learned without AR, using EMS on their hand and watching the video using Microsoft Teams.
- EMS × AR: Participants learned using EMS on their hand and following the instructor's hand using an AR hand model.

The instructor monitored the participant's learning progress via a display, as shown in Figure 7. Figure 8 shows the participant's setup for each condition.

Fig. 7. Instructor during the experiment, including display, webcam, control PC, and piano, used to monitor the participant's progress and provide instruction.

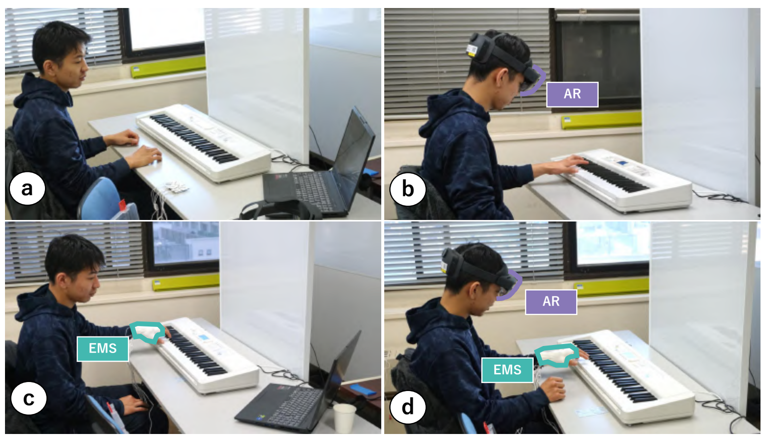

Fig. 8. Participant setup during the experiment for each condition. (a) Video condition: learning by watching a video of the instructor's hand. (b) AR condition: learning by following an AR hand model. (c) EMS condition: learning with EMS feedback while watching the video. (d) AR × EMS condition: learning with both AR hand model and EMS feedback.

4.3 Performance Task

Following Liu et al. [14], scales, arpeggios and chord progressions, were used to assess the effectiveness of the system. The tasks varied in complexity, involving different movements of the fingers and hands at specified tempos.

Scales: Participants played scales starting from C2, moving up two octaves at 100 BPM. The task involved alternating between notes one octave higher and one octave lower, using only the thumb and little finger of one hand.

Arpeggios: Participants performed a C major arpeggio over two octaves (C2 to G4) at 80 BPM, using the thumb, middle finger, and little finger of one hand.

Chords: Participants played the Canon chord progression (C→G→Am→Em→F→C→F→G) at 30 BPM, using one hand for the root note and the other for the three-note chord.

These tasks targeted different aspects of piano performance, such as finger dexterity, hand coordination, and rhythm accuracy.

4.4 Experimental Procedure

Participants completed a pre-experiment questionnaire and practiced the tasks. Each learning phase lasted up to 5 minutes, followed by a post-test for each task. After completing all tasks for each condition, participants completed the System Usability Scale (SUS) [2] and NASA Task Load Index (NASA-TLX) [10] to assess the perceived usability of the system and cognitive workload. These assessment tools were selected based on use in previous piano performance studies by Liu et al. [14, 15]. Additionally, a final questionnaire was conducted at the end of the experiment. This questionnaire asked participants to rank the conditions based on perceived learning effectiveness, provide feedback on the visibility and usability of the AR hand, evaluate the effectiveness of the EMS, and suggest potential improvements for the system.

4.5 Participants

Six computer science students (four males and two females, ages 21-24) participated, all right-handed. The average age was 22.8 years (SD = 0.41). We used a within-subjects design with counterbalanced condition and task order. Participants rated their familiarity with AR/VR, EMS, and piano performance on a 7-point scale, with mean scores of AR/VR: 4.17 (SD = 0.98), EMS: 2.33 (SD = 1.37), and piano performance: 2.33 (SD = 1.86).

4.6 Results

Figure 9 presents the results of the SUS, NASA-TLX and preference. A two-way ANOVA was conducted for visual and EMS factors using these scores. In cases where normality was not observed, an aligned rank transform (ART) [35] was employed to enable ANOVA for nonparametric data. SUS scores showed a significant EMS effect (Visual factor: $p = 0.210$, EMS factor: $p = 0.006 < 0.05$, Visual factor * EMS factor: $p = 0.591$). NASA-TLX showed no significant overall scores, but a significant effect was found on the Temporal scale for EMS (Visual factor: $p = 0.408$, EMS factor: $p = 0.042 < 0.05$, Visual factor * EMS factor: $p = 0.788$). In the final questionnaire, participants showed the highest preference for EMS.

4.7 Discussion

The results indicate that EMS negatively impacted usability but reduced temporal demand, as reflected in NASA-TLX scores. Participants reported visibility challenges with the AR hand due to HoloLens transparency and preferred a more realistic hand model. We recommend a video see-through device less affected by ambient lighting and a more realistic 3D hand model to enhance visibility and learning outcomes.

The final questionnaire revealed that while EMS was helpful for indicating key presses, issues such as incorrect finger stimulation and latency affected focus. Mixed responses suggest potential for EMS, but improvements in control accuracy and timing are needed. Participants also desired more flexible practice time, indicating a need for less rigid learning phases in future studies.

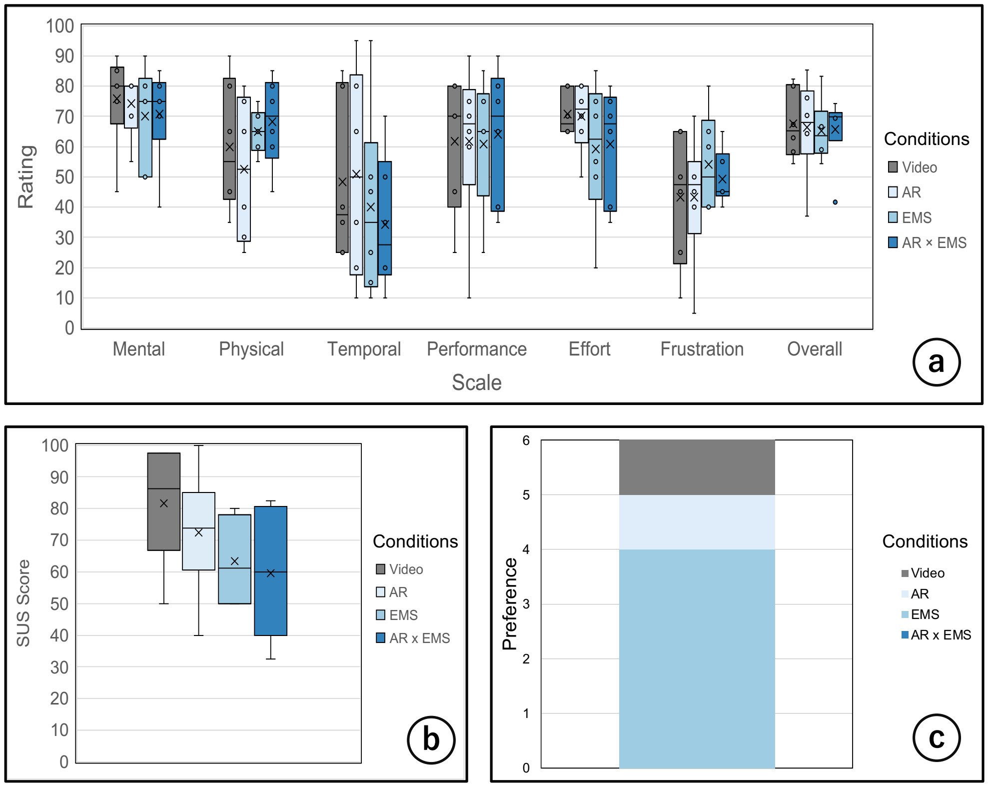

Fig. 9. Questionnaire results. (a) NASA-TLX. (b) SUS Score. (c) Preference.

5 EXPERIMENT

The preliminary experiment highlighted several challenges in real-time remote instruction, leading to system modifications. To improve visibility, we replaced the Hololens 2 with Meta Quest 3, using a video see-through approach to address lighting issues. Additionally, we shifted from real-time to pre-recorded data to enhance EMS stimulation accuracy. We then conducted a laboratory experiment to evaluate the learning experience, comparing three conditions: AR, EMS, and AR combined with EMS(Figure 11).

5.1 Experimental Environment

The experimental environment was set up with the equipment as shown in Figure 10. In addition, depending on the condition, the AR hand display using Meta Quest 3 and the EMS device shown in Figures 3 and 4 were used to share finger movements.

5.2 Performance Task

We designed different scales and arpeggios as performance tasks for each condition, following Liu et al. [14, 15]. All tasks were performed at 120 BPM, with viusal music score provided during the learning phase.

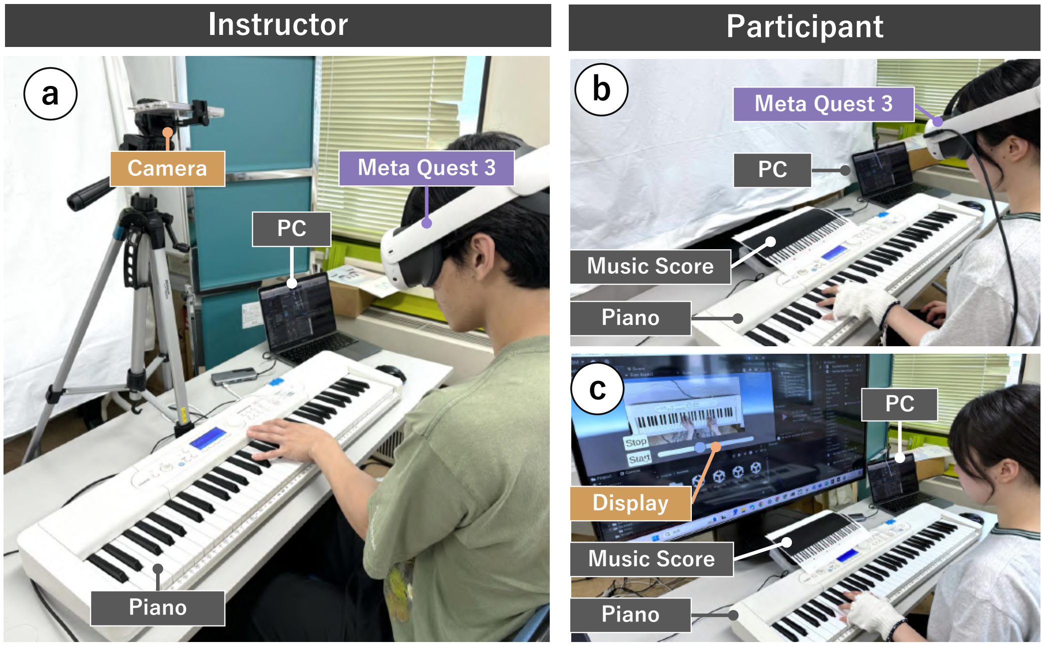

Fig. 10. Experimental environment. (a) Instructor's experimental environment with camera, PC, Meta Quest 3, and piano. (b) Participant's experimental environment with Meta Quest 3, PC, visual music score, and piano. (c) Additional participant view with display, PC, visual music score, and piano.

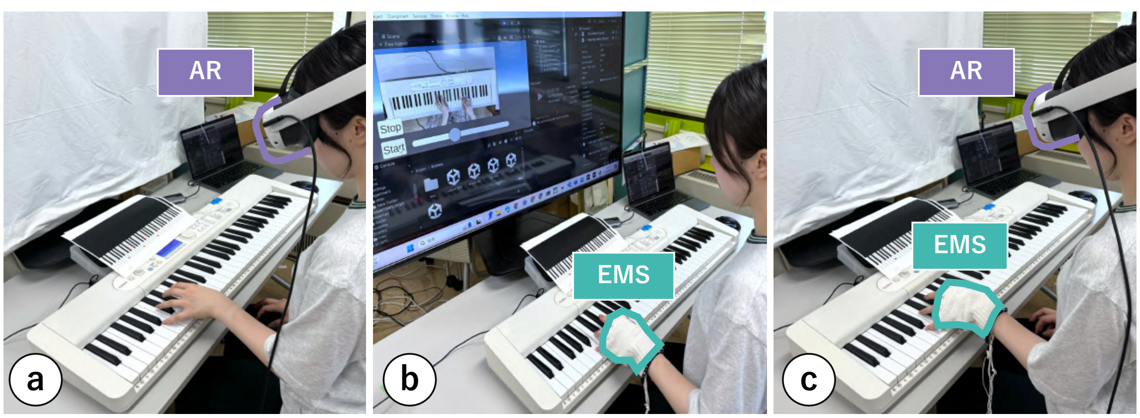

Fig. 11. Participants in each condition. (a) AR condition: learning by following an AR hand model. (b) EMS condition: learning with EMS feedback while watching a video. (c) AR × EMS condition: learning with both AR hand model and EMS feedback.

Scales: Participants played major scales in C#, D, and D#, depending on the condition. The task involved ascending and descending two octaves, alternating between a note one octave higher and one octave lower, using only the thumb and little finger of the left hand.

Arpeggios: Participants performed major arpeggios in A, B, and C, starting from the root note, ascending two octaves, and then descending back. The thumb, middle finger, and little finger of the left hand were used.

5.3 Experimental Procedure

Participants completed a pre-experiment questionnaire and were informed about the protocol. The study included three instruction conditions, each with two performance tasks (scales and arpeggios). The participants practiced each task for 5 minutes, followed by a performance test. After completing all tasks for a condition, they filled out the SUS and NASA-TLX questionnaires. The experiment concluded with a semi-structured interview. Performance tests were recorded using MIDI data obtained through Logic Pro X.

5.4 Participants

Seven computer science students (six males, ages 21-24) participated, all right-handed.The average age was 22.3 years (SD = 1.06). We used a within-subjects design with counterbalanced condition and task order. Participants rated their familiarity with AR/VR, EMS, and piano performance on a 7-point Likert scale (1: beginner, 7: expert). Average scores were: AR/VR: 2.8 (SD = 1.99), EMS: 0.53 (SD = 0.53), and piano performance: 1.5 (SD = 0.71). One participant had prior piano instruction experience (1 year).

5.5 Results

We assessed piano performance using MIDI data analysis and evaluated the learning experience in each instruction condition using SUS and NASA-TLX.

5.5.1 MIDI Data. Interviews with experienced pianists indicated that beginners struggle with consistent keystroke strength (mean years of experience: 11.7 years). We analyzed MIDI data to assess uniformity in key press strength(velocity uniformity) and performance accuracy, applying aligned rank transformation due to non-normal distribution. For velocity uniformity, we calculated the root mean square deviation of velocity following Takahashi et al. [32]. Performance accuracy was assessed using two metrics: (1) the number of keystroke errors and (2) the matching time between instructor and participant MIDI data. Due to non-normal distribution of data (Shapiro-Wilk test, $p < .05$), we applied the aligned rank transformation. We then conducted two-way ANOVAs with condition and performance task as factors for both velocity uniformity and performance accuracy measures. The results did not show significant main effects or interactions for either velocity uniformity or performance accuracy across conditions and tasks (figure 12).

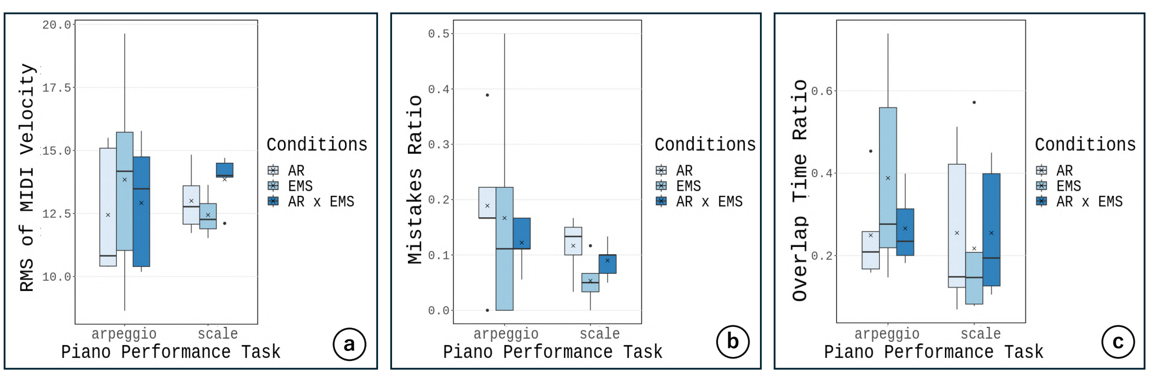

Fig. 12. Performance evaluation across conditions: (a) Velocity uniformity (higher values indicate higher consistency), (b) Keystroke error rate (percentage of errors relative to total performance time), and (c) Timing accuracy (percentage of MIDI event overlap between instructor and participant).

5.5.2 Questionnaire. The results of SUS and NASA-TLX are shown in Figure 13. From the results of SUS and NASA-TLX, normality (Shapiro-Wilk test, $p > .05$) and homoscedasticity (Levene test, $p > .05$) were not rejected. Therefore, one-way repeated measures ANOVA was conducted, and no main effects were observed.

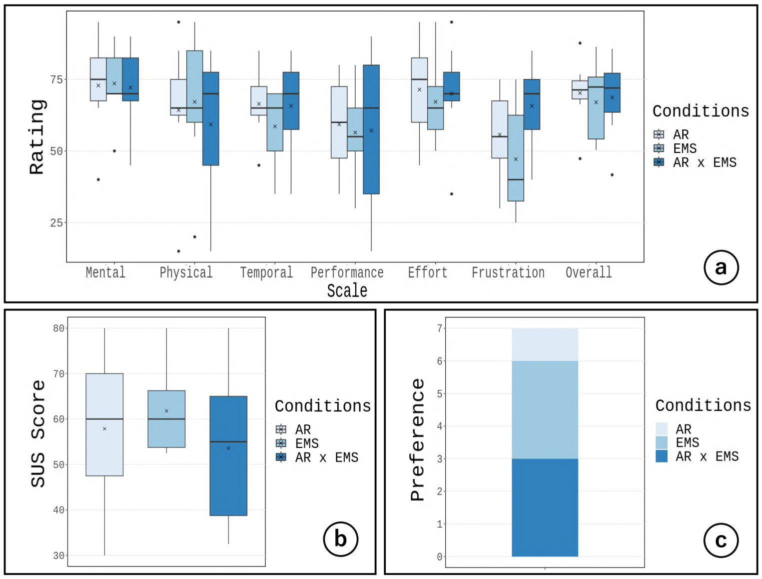

Fig. 13. Questionnaire results. (a) NASA-TLX. (b) SUS Score. (c) Preference.

5.5.3 Interview. We conducted interviews with participants to gather insights about the experimental conditions and performance tasks. In the AR hand condition, participants generally found the hand positioning more intuitive and easier to understand compared to watching video. Many participants (P1-P6) appreciated the visibility of hand position in AR. For example, P7 noted, "It was especially helpful for larger motions like arpeggios." There were also concerns about the AR hand's position accuracy relative to the actual keys. However, some challenges were noted. P1 expressed that "it was difficult to identify which finger was pressing the keys," a sentiment shared by P7. Technical issues such as delays, misalignment between AR and real space, and the weight of the headset were also noted by several participants (P2, P4, P5).

The EMS condition received positive feedback regarding timing and finger control. Participants found it helpful for understanding tempo and timing (P3), identifying which finger to move (P3, P6), and potentially useful for complex finger movements (P4). However, some participants (P2, P3) reported numbness with prolonged EMS use, indicating potential issues for extended sessions.

In the combined AR × EMS condition, several participants observed that the two modalities complemented each other's strengths while offsetting individual weaknesses. P1 remarked that "AR helped in adjusting hand positions, while EMS was useful for monitoring finger movements." P4 similarly appreciated the "Ability to adjust hand positions using AR while simultaneously tracking fingers with EMS." However, P6 pointed out that focusing on the AR visuals "Sometimes detracted from perceiving the EMS stimulation," suggesting a possible cognitive load issue when processing both types of feedback simultaneously.

Regarding the performance tasks, participants had differing opinions on arpeggios and scales. Arpeggios were generally perceived as manageable, with sufficient practice time leading to satisfactory performance. Several participants struggled with the descending parts of scales. P1 remarked that "High notes were harder to press," and P6 noted difficulty in "Maintaining adequate pressing strength, especially with the little finger." Many participants recommended additional practice time, particularly for scales, to address these challenges.

6 DISCUSSION

This study compared three conditions AR, EMS, and AR combined with EMS to evaluate the effectiveness of integrating AR hand and EMS feedback in learning complex finger movements. The discussion highlights key insights and challenges from the experiment.

6.1 Effectiveness of AR x EMS

The combination of AR and EMS showed potential in providing a more comprehensive learning experience by integrating visual and tactile feedback. Some participants (P1, P4) found this approach particularly effective for tasks that required both hand position changes and precise finger control. P1 mentioned, "I was using AR to position my hand and EMS to time my finger movements," highlighting the complementary nature of these modalities. However, the cognitive load of processing both visual and tactile cues simultaneously proved challenging for others. P6 remarked, "I found it hard to focus on the EMS stimulation while also trying to follow the AR hand," indicating that managing dual feedback might be overwhelming for some learners.

Despite these perceptions, the performance metrics did not show significant improvements with the AR and EMS combination. This suggests that while the approach has potential, it may not offer clear advantages in short-term learning scenarios. Further research is needed to optimize the use of AR and EMS for different learning styles.

6.2 Learning Experience and Technical Challenges

The combination of AR and EMS was well-received by some participants, suggesting its potential to enhance the learning experience, particularly for complex movements. However, individual differences in learning preferences were evident, with some participants benefiting more from the combined feedback than others.

Despite improvements with Meta Quest 3, technical challenges persisted. Issues like spatial registration problems and misalignment between the AR hand and physical keys were reported by participants (P2, P4, P5), disrupting the learning process. Additionally, while EMS was helpful for some, it caused discomfort for others (P2, P3) during prolonged use, raising concerns about its suitability for extended practice sessions. These challenges highlight the need for further refinement of both AR and EMS components to ensure a more consistent and comfortable user experience across different learners.

7 LIMITATION AND FUTURE WORK

This study has several limitations. The small sample size and limited participant demographics restrict the generalizability of the findings. The lack of statistically significant results in the quantitative analysis likely stems from the small participant pool and the limited performance tasks, which focused only on scales and arpeggios. Future research should include larger, more diverse samples and a broader range of tasks suited to different skill levels. Additionally, we aim to measure and evaluate changes in participants' playing accuracy—specifically their timing, note accuracy, and finger positioning—before and after system use to better assess learning effectiveness.

Age-specific needs, particularly for young learners who are often piano beginners, also warrant consideration. Younger participants might find the simultaneous use of AR and EMS challenging. The system should be adapted to their cognitive and motor skills, offering more intuitive and age-appropriate guidance. Incorporating gamified practice sessions or step-by-step guidance could help younger users better understand and practice correct hand shapes. However, when using EMS with children, special attention must be given to safety concerns. It is essential to adhere to established guidelines to ensure that the intensity and duration of EMS are appropriately designed for younger users.

Technical issues such as AR misalignment and EMS discomfort need to be addressed. Future refinements should focus on enhancing AR alignment accuracy and making EMS feedback more comfortable and precise. Furthermore, optimizing electrode configurations to reduce the number of ground pads without compromising stimulation efficacy is crucial for improving system usability and participant experience. Additionally, the system's effectiveness may vary by skill level. Beginners might benefit more from both AR and EMS, while experienced users may only need EMS for rhythm. Future studies should explore these variations to optimize the system for different users.

8 CONCLUSION

This study examined the integration of AR and EMS for enhancing piano instruction, particularly for beginners. The experiments compared the effectiveness of the combined AR and EMS approach with AR-only and EMS-only conditions.

The results showed that AR was generally helpful for hand positioning, and EMS provided useful timing cues for finger movements. However, the improvements in performance metrics, such as MIDI data analysis, were not statistically significant. Challenges included mild discomfort from EMS, visibility issues with the AR hand, and occasional misalignment between the AR hand and the physical keyboard.

Future research should involve larger and more diverse samples, with performance tasks tailored to different skill levels. The system's effectiveness may vary between beginners and experienced users, and exploring these differences could help optimize the system. Addressing technical issues such as AR misalignment and EMS discomfort is crucial for improving the system.

In conclusion, the combination of AR and EMS shows promise for piano instruction, but further refinement and addressing the identified limitations are essential for enhancing the learning experience.

ACKNOWLEDGMENTS

The authors acknowledge the use of generative AI in the preparation of this manuscript and the development of the system application. Specifically, Claude and ChatGPT were utilized to assist in the writing process and in certain aspects of the application development. These AI tools contributed to refining the manuscript's language and structure, and aided in generating code snippets for the system application. However, all core research ideas, experimental design, data analysis, and critical interpretations were conducted by the human authors.

REFERENCES

[1] Mariano Banquiero, Gracia Valdeolivas, David Ramón, and M.-Carmen Juan. 2024. A color Passthrough mixed reality application for learning piano. *Virtual Reality* 28, 2 (06 Mar 2024), 67. https://doi.org/10.1007/s10055-024-00953-w

[2] John Brooke. 1996. SUS : A Quick and Dirty Usability Scale. *Usability Evaluation in Industry* (1996), 189–194.

[3] Minya Cai, Muhammad Alfian Amrizal, Toru Abe, and Takuo Suganuma. 2019. Design of an AR-Based System for Group Piano Learning. *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (2019), 20–21. https://api.semanticscholar.org/CorpusID:210694108

[4] Casio. 2021. LK-520 | CASIO. https://www.casio.com/jp/electronic-musical-instruments/product.LK-520/ (Accessed on 06/30/2024).

[5] Pei-Ying Chiang and Chung-Hsuan Sun. 2015. Oncall Piano Sensei: Portable AR Piano Training System. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction* (Los Angeles, California, USA) *(SUI '15)*. Association for Computing Machinery, New York, NY, USA, 134. https://doi.org/10.1145/2788940.2794353

[6] Jordan Aiko Deja, Sven Mayer, Klen Čopič Pucihar, and Matjaž Kljun. 2022. A Survey of Augmented Piano Prototypes: Has Augmentation Improved Learning Experiences? *Proc. ACM Hum.-Comput. Interact.* 6, ISS, Article 566 (nov 2022), 28 pages. https://doi.org/10.1145/3567719

[7] Ayaka Ebisu, Satoshi Hashizume, Kenta Suzuki, Akira Ishii, Mose Sakashita, and Yoichi Ochiai. 2017. Stimulated Percussions: Method to Control Human for Learning Music by Using Electrical Muscle Stimulation. In *Proceedings of the 8th Augmented Human International Conference* (Silicon Valley, California, USA) *(AH '17)*. Association for Computing Machinery, New York, NY, USA, Article 33, 5 pages. https://doi.org/10.1145/3041164.3041202

[8] C. Ailie Fraser, Joy O. Kim, Alison Thornsberry, Scott Klemmer, and Mira Dontcheva. 2019. Sharing the Studio: How Creative Livestreaming Can Inspire, Educate, and Engage. In *Proceedings of the 2019 Conference on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19)*. Association for Computing Machinery, New York, NY, USA, 144–155. https://doi.org/10.1145/3325480.3325485

[9] Dominik Hackl and Christoph Anthes. 2017. HoloKeys - An Augmented Reality Application for Learning the Piano. In *Forum Media Technology*.

[10] S. G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in psychology* 52 (1988), 139–183. https://api.semanticscholar.org/CorpusID:15252590

[11] Michinari Kono, Yoshio Ishiguro, Takashi Miyaki, and Jun Rekimoto. 2018. Design and Study of a Multi-Channel Electrical Muscle Stimulation Toolkit for Human Augmentation. In *Proceedings of the 9th Augmented Human International Conference* (Seoul, Republic of Korea) *(AH '18)*. Association for Computing Machinery, New York, NY, USA, Article 11, 8 pages. https://doi.org/10.1145/3174910.3174913

[12] Michinari Kono, Takumi Takahashi, Hiromi Nakamura, Takashi Miyaki, and Jun Rekimoto. 2018. Design Guideline for Developing Safe Systems that Apply Electricity to the Human Body. *ACM Trans. Comput.-Hum. Interact.* 25, 3, Article 19 (jun 2018), 36 pages. https://doi.org/10.1145/3184743

[13] Katerina Labrou, Cagri Hakan Zaman, Arda Turkyasar, and Randall Davis. 2023. Following the Master's Hands: Capturing Piano Performances for Mixed Reality Piano Learning Applications. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 141, 8 pages. https://doi.org/10.1145/3544549.3585838

[14] Ruofan Liu, Erwin Wu, Chen-Chieh Liao, Hayato Nishioka, Shinichi Furuya, and Hideki Koike. 2023. PianoHandSync: An Alignment-based Hand Pose Discrepancy Visualization System for Piano Learning. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 228, 7 pages. https://doi.org/10.1145/3544549.3585705

[15] Ruofan Liu, Erwin Wu, Chen-Chieh Liao, Hayato Nishioka, Shinichi Furuya, and Hideki Koike. 2023. PianoSyncAR: Enhancing Piano Learning through Visualizing Synchronized Hand Pose Discrepancies in Augmented Reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE Computer Society, Los Alamitos, CA, USA, 859–868. https://doi.org/10.1109/ISMAR59233.2023.00101

[16] Karola Marky, Andreas Weiß, and Thomas Kosch. 2021. Supporting Musical Practice Sessions Through HMD-Based Augmented Reality. *CoRR* abs/2101.00874 (2021). arXiv:2101.00874 https://arxiv.org/abs/2101.00874

[17] Mediapipe. 2024. MediaPipe | ; Google for Developers. https://developers.google.com/mediapipe (Accessed on 6/30/2024).

[18] Meta. 2024. Meta Quest 3: New Mixed Reality VR Headset - Shop Now | Meta Store. https://www.meta.com/quest/quest-3/ (Accessed on 06/29/2024).

[19] Meta. 2024. Meta XR All-in-One SDK - Oculus Developer Center. https://developer.oculus.com/downloads/package/meta-xr-sdk-all-in-one-upm/ (Accessed on 06/29/2024).

[20] Microsoft. 2019. Microsoft HoloLens | Mixed Reality Technology for Business. https://www.microsoft.com/en-us/hololens (Accessed on 06/30/2024).

[21] Microsoft. 2024. Video Conferencing, Meetings, Calling | Microsoft Teams. https://www.microsoft.com/en-us/microsoft-teams/group-chat-software/ (Accessed on 06/30/2024).

[22] D. Molero, Santiago Schez-Sobrino, David Vallejo-Fernandez, Carlos González-Morcillo, and Javier Albusac. 2020. A novel approach to learning music and piano based on mixed reality and gamification. *Multimedia Tools and Applications* 80 (2020), 165–186.

[23] Arinobu Niijima. 2023. Learning Effects and Retention of Electrical Muscle Stimulation in Piano Playing. In *Proceedings of the 2023 ACM International Symposium on Wearable Computers* (Cancun, Quintana Roo, Mexico) *(ISWC '23)*. Association for Computing Machinery, New York, NY, USA, 104–108. https://doi.org/10.1145/3594738.3611373

[24] Arinobu Niijima, Toki Takeda, Ryosuke Aoki, and Shinji Miyahara. 2022. Muscle Synergies Learning with Electrical Muscle Stimulation for Playing the Piano. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 54, 10 pages. https://doi.org/10.1145/3526113.3545666

[25] Arinobu Niijima, Toki Takeda, Kentaro Tanaka, Ryosuke Aoki, and Yukio Koike. 2021. Reducing Muscle Activity When Playing Tremolo by Using Electrical Muscle Stimulation to Learn Efficient Motor Skills. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 123 (sep 2021), 17 pages. https://doi.org/10.1145/3478110

[26] Luc Nijs and Bahareh Behzadaval. 2024. Laying the Foundation for Augmented Reality in Music Education. *IEEE Access* 12 (2024), 100628–100645. https://doi.org/10.1109/ACCESS.2024.3427698

[27] Linsey Raymaekers, Jo Vermeulen, Kris Luyten, and Karin Coninx. 2014. Game of tones: learning to play songs on a piano using projected instructions and games. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI EA '14)*. Association for Computing Machinery, New York, NY, USA, 411–414. https://doi.org/10.1145/2559206.2574799

[28] Sainitas. 2019. For regeneration and pain relief SEM 43 - Digital EMS/TENS | Sanitas Onlineshop. https://sanitas-online.de/en/p/sem-43-digital-ems-tens/ (Accessed on 06/30/2024).

[29] Naotaka Sakai, Michael C. Liu, Fong-Chin Su, Allen T. Bishop, and Kai-Nan An. 1996. Motion Analysis of the Fingers and Wrist of the Pianist. *Medical Problems of Performing Artists* 11, 1 (1996), 24–29. http://www.jstor.org/stable/45440566

[30] Austin Jerald Stanbury, Ines Said, and Hyo Jeong Kang. 2021. HoloKeys: Interactive Piano Education Using Augmented Reality and IoT. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology* (Osaka, Japan) *(VRST '21)*. Association for Computing Machinery, New York, NY, USA, Article 76, 3 pages. https://doi.org/10.1145/3489849.3489921

[31] Akifumi Takahashi, Jas Brooks, Hiroyuki Kajimoto, and Pedro Lopes. 2021. Increasing Electrical Muscle Stimulation's Dexterity by Means of Back of the Hand Actuation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 216, 12 pages. https://doi.org/10.1145/3411764.3445761

[32] Noriyuki Takahashi and Minoru Tsuzaki. 2006. Effects of altered auditory feedback on piano performance. *J. Acoust. Soc. Am. 120* (2006), 3004–3005. https://cir.nii.ac.jp/crid/1010000781869642885

[33] Emi Tamaki, Takashi Miyaki, and Jun Rekimoto. 2011. PossessedHand: Techniques for Controlling Human Hands Using Electrical Muscles Stimuli. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 543–552. https://doi.org/10.1145/1978942.1979018

[34] Unity. 2021. Unity Recorder V1.0 User Manual | Package Manager UI website. https://unitytech.github.io/unity-recorder/manual/index.html (Accessed on 06/30/2024).

[35] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[36] Shunsuke Yoshimoto, Kazuma Aoyama, Hiroyuki Kajimoto, and Atsushi Nishikawa. 2020. Safety Guidelines and Effects of Electricity Applied to the Human Body. *Transactions of Japanese Society for Medical and Biological Engineering* 58, 4-5 (09 2020), 147–159. https://doi.org/10.11239/jsmbe.58.147

[37] Hong Zeng, Xingxi He, and Honghu Pan. 2019. FunPianoAR: A Novel AR Application for Piano Learning Considering Paired Play Based on Multi-Marker Tracking. *Journal of Physics: Conference Series* 1229, 1 (may 2019), 012072. https://doi.org/10.1088/1742-6596/1229/1/012072

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009