PianoSense: Multimodal Visual-Haptic Feedback for Student-Instructor Interaction in Remote Instrument Lessons

Chau Vu
Department of Computer Science
Pomona College
Claremont, California, USA
chvo2022@mymail.pomona.edu

Alexandra Papoutsaki
Department of Computer Science
Pomona College
Claremont, California, USA
alexandra.papoutsaki@pomona.edu

Jingyi Li
Department of Computer Science
Pomona College
Claremont, California, USA
jingyi.li@pomona.edu

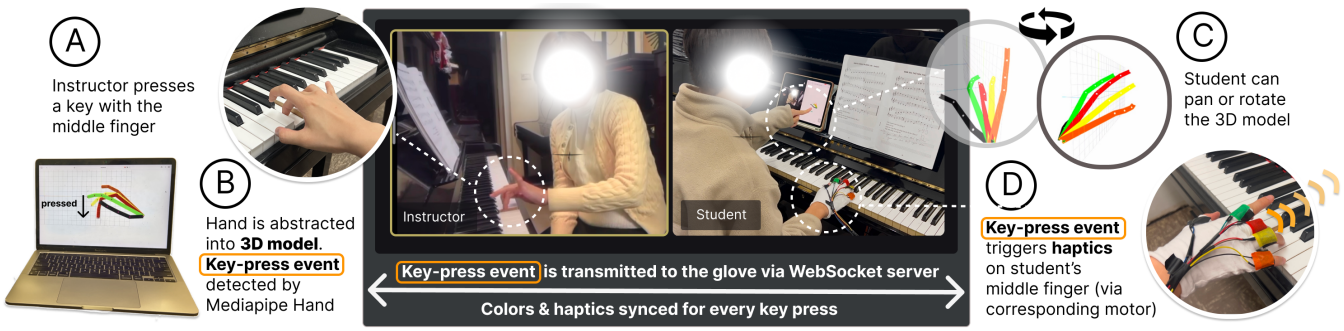


Figure 1: *PianoSense*, a design probe [15] that supports both visual and physical aspects of remote piano learning. (A) The instructor presses a key with the middle finger; (B) the hand is abstracted into a 3D model; key-press event is detected by Mediapipe Hand; (C) the student can pan and rotate the 3D model on the tablet; and (D) the detected middle-finger key press event is synced through a WebSocket server to the student's glove, triggering vibrotactile feedback on the middle finger.

Abstract

Remote piano lessons offer affordability and convenience, yet lack synchronized visual references for hand form and tactile feedback. We present PianoSense, a tool that integrates an interactive 3D visualization of the instructor's hands with a vibrotactile glove conveying key press events. A preliminary design probe with three student–instructor pairs revealed that PianoSense's multimodal cues functioned as a shared instructional language when video alone was insufficient. We also discuss how remote instrument-learning technologies can support reflective, learner-centered practices, and how researchers can best design interaction-aware scaffolding that adapts to evolving instructor and student needs.

CCS Concepts

• **Applied computing** → **Performing arts**.

Keywords

Remote Instrument Learning, Visualization, Haptic Feedback

ACM Reference Format:

Chau Vu, Alexandra Papoutsaki, and Jingyi Li. 2026. PianoSense: Multimodal Visual-Haptic Feedback for Student-Instructor Interaction in Remote Instrument Lessons. In *The 39th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '26), November 02–05, 2026, Detroit, MI, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3830397.3841792

This work is licensed under a Creative Commons Attribution 4.0 International License.
UIST Adjunct '26, Detroit, MI, USA
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2855-6/26/11
https://doi.org/10.1145/3830397.3841792

1 Introduction

Remote instrument learning allows students to learn with experts despite scheduling constraints or geographic and mobility limitations [4, 28]. Learning an instrument like piano relies heavily on visual and physical guidance, which is difficult to replicate in video conferencing platforms. Students observe and coordinate finger movements to produce rhythm [6, 10], while instructors physically guide students' fingers or tap their hands [7]. Both guidance forms are complementary as visual demonstration helps students understand the target hand form, while physical guidance directs them toward the corresponding finger action. However, prior work has typically explored visual and haptic modalities in isolation. Visual approaches have used XR-based visualizations to address only the visual limitations of remote learning [9, 12, 13, 27, 28] while haptic systems support finger movement, motor learning, and force feedback [5, 10, 17, 21, 26]. For example, Passive Haptic Rehearsal explores passive haptic feedback for motor-skill acquisition [5], while other work compares visual and vibrotactile feedback for piano instruction [3] and evaluates vibrotactile gloves, electrical muscle stimulation, and exoskeleton guidance [10]. Existing work also emphasizes skill transfer [5, 11], learning effectiveness [9, 10, 14, 22, 25], or error correction [12, 24, 25], evaluated through note and timing accuracy [19, 24, 29]. While such metrics show technical feasibility and learning gains, they are limiting in instructor-mediated learning. Learning unfolds through continuous exchanges that shape how students interpret instruction and reflect

on their progress, while instructors adapt their teaching to students' needs. Focusing solely on performance outcomes thus risks overlooking these interactions through which learning is experienced.

Motivated by these gaps, we chose piano learning as a case study and designed PianoSense, a multimodal system that complements visual demonstrations of instructors' movements with physical guidance for students. Unlike prior systems using visual or haptic feedback separately, PianoSense integrates both for remote instructor–student interaction. Piano's fine-grained motor demands make it a strong context for exploring multimodal support. We evaluate PianoSense as a design probe [15] to surface opportunities in integrating multimodal feedback into remote piano learning. Through a preliminary study with three instructor–novice pairs, we explore: **(RQ1)** How does a multimodal system, as instantiated by PianoSense, shape interactions and fit in existing teaching and learning practices in remote piano lessons? and **(RQ2)** What new insights emerge for designing for remote instrument learning?

2 PianoSense

PianoSense consists of two integrated components as shown in Figure 1: **(1) a real-time 3D hand visualization** system for observing instructor demonstrations and **(2) a wearable glove** that provided tactile cues to students based on the instructor's finger movements. We use the MediaPipe Hands model [30] to capture the instructor's hand movements (Figure 1A) and reconstruct a real-time 3D hand model in Three.js [2] (Figure 1B). The student can pan and rotate the model to inspect the instructor's hand movements from multiple viewpoints (Figure 1C). The same hand representation is used to detect finger-press events, which are streamed in real time via a Node.js–based WebSocket server [18] to drive the corresponding haptic feedback on the glove. The glove includes coin vibration motors on each finger (Figure 1D) and sends real-time feedback by inferring key presses from fingertip depth displacement (3 cm, empirically determined). End-to-end latency is under 1 second. We make the codebase publicly available at https://github.com/chauvuha/PianoSense-Program.git.

We conducted an IRB-approved preliminary study with three instructor–student pairs, compensating instructors $50 and students $25 for one hour. Instructors had extensive teaching experience (14–30 years) and online teaching experience. Students identified as novices; two (S2 & S3) had prior remote piano learning experience. Instructors and students were placed in separate rooms and connected via a standard video call. The study included two 12-minute conditions: a baseline (video-only) condition followed by a PianoSense condition. Each used different music pieces and scales of comparable difficulty, selected by the instructor. Instructors and students then participated in separate 15-minute semi-structured interviews, analyzed using inductive thematic analysis [1].

3 Preliminary Findings and Discussions

3.1 Learner Observation and Self-Diagnosis

Students used the 3D visual scaffold of the instructor's hand to reflect on their mistakes and attempt self-correction rather than relying solely on the instructor's immediate feedback (**RQ1**). S3 said *"When I made a mistake, I use my own hands to adjust the model and think for myself about where I went wrong. Only after that, I would ask for help [...] Compared to the usual session where I just took in the instructor's feedback."* These behaviors reflect student-driven reflection-in-action [20]. While typically framed as an individual process [8, 20], our findings suggest it can also foster shared understanding with instructors by making students' difficulties more visible and helping them articulate when they need help. Control over the scaffold also supported students' self-paced learning. S2 noted: *"It didn't feel rushed, and I could think more about how to improve my hand posture"*. Future work can explore how student-driven reflection might create space for more meaningful instructor-student exchanges. For example, music learning systems can *support students in articulating their observations of their performance, while giving instructors visibility into those reflections to adapt their teaching better in response.* (**RQ2**). PianoSense's latest version includes a view that reveals when and how students pan or rotate the scaffold, surfacing reflection points for instructors.

3.2 Material- and Interaction-aware Scaffolding

While using PianoSense, instructors and students used visual and haptic cues as a shared instructional language when video alone was insufficient (e.g., I3: *"Can you feel the vibration on this red finger?"*). Their effectiveness depended on (1) the structure of the musical material itself, as pieces usually contain repeating passages of varying difficulty, and (2) the interaction dynamics shaped by students' and instructors' familiarity with each passage after repeated practice (**RQ1**). S2 found the haptic scaffold as most helpful when learning an unfamiliar passage: *"For slower parts where the finger movements are more complex, it is useful at the beginning when breaking into that part."* Instructors similarly noted that visual and haptic cues helped novices map hand forms to fingerings. I2 noted: "*When they see colors, they can clearly distinguish which finger number the is referring to. It would also be helpful for us as instructors when teaching.*" As students became more familiar with repeated passages and instructor expectations for how they should be played, persistent cues (e.g., continuous vibrations) could cause discomfort and reduce focus. S3 noted that repeated vibrations "vibrated a bit too much," motivating requests for more customizable feedback.

Prior learning systems have tailored guidance as learners progress from novice to expert [16] and adapting instruction based on training task difficulty [23]. Our findings extend these approaches to instrument learning, where adaptation is grounded within a single musical piece rather than across discrete tasks. We call for future instrument-learning systems to explore *material- and interaction-aware multimodal scaffolding that adapts within a lesson to both the structural demands of the piece and the evolving instructor-student coordination that emerges from these demands* (**RQ2**). For example, systems could infer familiarity by tracking how often students and instructors repeat specific passages and adjust visual and haptic cue intensity accordingly. PianoSense's latest version points toward this direction by having both manual and automatic haptic intensity controls, enabling instructors and students to co-regulate feedback during the lesson, or allowing the system to automatically adjust intensity by detecting how much of a passage has been practiced.

3.3 Limitations & Future Work

Single-camera tracking was sufficient during typical playing but became less reliable when fingers occluded one another during rapid movements. PianoSense also relies on fingertip vibrotactile feedback, which is effective for discrete key presses but ambiguous during continuous playing. Future work will explore complementary visual and haptic strategies, support the left hand, and evaluate with larger, counterbalanced, and longitudinal study designs.

Acknowledgments

The authors sincerely thank all participants for their time and valuable contributions to the evaluation. We are especially grateful to Anthony Clark for generously sharing his expertise and guidance in developing the haptic glove component.

References

[1] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.
[2] Ricardo Cabello. 2010. Three.js. https://threejs.org/.
[3] Matteo A Coscia and Mazen Al Borno. 2025. Vibrotactile versus visual stimulation in learning the piano. *Interacting with Computers* 37, 6 (2025), 548–552.
[4] Sam Duffy and Patrick Healey. 2017. A new medium for remote music tuition. *Journal of Music, Technology & Education* 10, 1 (2017), 5–29.
[5] Tan Gemicioglu, Elijah Hopper, Brahmi Dwivedi, Richa Kulkarni, Asha Bhandarkar, Priyanka Rajan, Nathan Eng, Adithya Ramanujam, Charles Ramey, Scott M Gilliland, et al. 2024. Passive Haptic Rehearsal for Augmented Piano Learning in the Wild. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 4 (2024), 1–26.
[6] Werner Goebl and Caroline Palmer. 2008. Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research* 186, 3 (2008), 471–479.
[7] Chih-Pin Hsiao, Richard Li, Xinyan Yan, and Ellen Yi-Luen Do. 2015. Tactile teacher: Sensing finger tapping in piano playing. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*. 257–260.
[8] David A Kolb. 2014. *Experiential learning: Experience as the source of learning and development*. FT press.
[9] 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*. 1–8.
[10] Kyungyeon Lee, Daniel S Yang, Kriti Singh, and Jun Nishida. 2025. Hapticus: Exploring the Effects of Haptic Feedback and its Customization on Motor Skill Learning: Tactile, Haptic, and Somatosensory Approaches. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–20.
[11] Ruofan Liu, Yichen Peng, Takanori Oku, Chen-Chieh Liao, Erwin Wu, Shinichi Furuya, and Hideki Koike. 2025. PiaMuscle: Improving Piano Skill Acquisition by Cost-effectively Estimating and Visualizing Activities of Miniature Hand Muscles. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–16.
[12] 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*. 1–7.
[13] 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, 859–868.
[14] Karola Marky, Andreas Weiß, Andrii Matviienko, Florian Brandherm, Sebastian Wolf, Martin Schmitz, Florian Krell, Florian Müller, Max Mühlhäuser, and Thomas Kosch. 2021. Let's frets! Assisting guitar students during practice via capacitive sensing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.
[15] Tuuli Mattelmäki. 2006. *Design probes*. Aalto University.
[16] Tricia J Ngoon, C Ailie Fraser, Ariel S Weingarten, Mira Dontcheva, and Scott Klemmer. 2018. Interactive guidance techniques for improving creative feedback. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–11.
[17] Jun Nishida, Yudai Tanaka, Romain Nith, and Pedro Lopes. 2022. DigituSync: A dual-user passive exoskeleton glove that adaptively shares hand gestures. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology*. 1–12.
[18] Node.js Foundation. 2024. Node.js. https://nodejs.org. Accessed: 2026-01.
[19] Katja Rogers, Amrei Röhlig, Matthias Weing, Jan Gugenheimer, Bastian Könings, Melina Klepsch, Florian Schaub, Enrico Rukzio, Tina Seufert, and Michael Weber. 2014. Piano: Faster piano learning with interactive projection. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. 149–158.
[20] Donald A Schön. 2017. *The reflective practitioner: How professionals think in action*. Routledge.
[21] Nobuhiro Takahashi, Shinichi Furuya, and Hideki Koike. 2020. Soft exoskeleton glove with human anatomical architecture: production of dexterous finger movements and skillful piano performance. *IEEE Transactions on Haptics* 13, 4 (2020), 679–690.
[22] Jingjing Tang, Shinichi Furuya, Hayato Nishioka, Momoko Shioki, Geraint A Wiggins, George Fazekas, and Vincent KM Cheung. 2026. Visualising Pianists' Touch: Transcribing Expressive Piano Performance from Audio to Piano Key Motion. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. 1–11.
[23] Dishita G Turakhia, Yini Qi, Lotta-Gili Blumberg, Andrew Wong, and Stefanie Mueller. 2021. Can physical tools that adapt their shape based on a learner's performance help in motor skill training?. In *Proceedings of the fifteenth international conference on tangible, embedded, and embodied interaction*. 1–12.
[24] Bryan Wang, Meng Yu Yang, and Tovi Grossman. 2021. Soloist: Generating mixed-initiative tutorials from existing guitar instructional videos through audio processing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
[25] Che Wei Wang, Pin Chun Lu, Yun Chen Cheng, and Mike Y Chen. 2025. MR. Drum: Designing Mixed Reality Interfaces to Support Structured Learning Micro-Progression in Drumming. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–17.
[26] Febrina Wijaya, Ying-Chun Tseng, Wan-Lun Tsai, Tse-Yu Pan, and Min-Chun Hu. 2020. VR piano learning platform with leap motion and pressure sensors. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*. IEEE, 584–585.
[27] Xiao Xiao and Hiroshi Ishii. 2010. MirrorFugue: communicating hand gesture in remote piano collaboration. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*. 13–20.
[28] Mayuka Yamashita and Soh Masuko. 2025. AR-Supported Online Piano Lessons with Bidirectional Visualization of Teacher and Student Hand Movements. In *Proceedings of the SIGGRAPH Asia 2025 Posters*. 1–2.
[29] Beste F Yuksel, Kurt B Oleson, Lane Harrison, Evan M Peck, Daniel Afergan, Remco Chang, and Robert JK Jacob. 2016. Learn piano with BACh: An adaptive learning interface that adjusts task difficulty based on brain state. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 5372–5384.
[30] Fan Zhang, Valentin Bazarevsky, Andrey Vakunov, Andrei Tkachenka, George Sung, Chuo-Ling Chang, and Matthias Grundmann. 2020. Mediapipe hands: On-device real-time hand tracking. *arXiv preprint arXiv:2006.10214* (2020).