AR Matchmaking:
The Compatibility of Musical Instruments with an AR Interface

Hyunkyung Shin
School of Music, Georgia Institute of Technology
Atlanta, GA, USA
hshin336@gatech.edu

Henrik von Coler
School of Music, Georgia Institute of Technology
Atlanta, GA, USA
hvc@gatech.edu

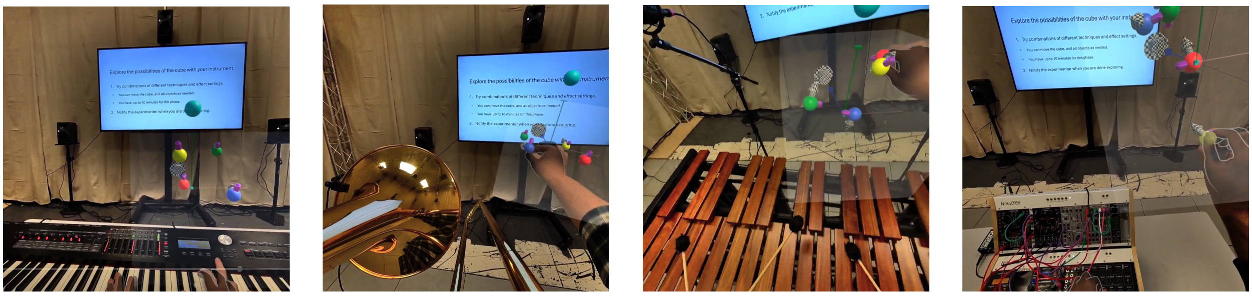

Figure 1: Piano, trombone, marimba and modular synth with the AR interface from musicians' perspectives.

Abstract

Augmented Reality (AR) interfaces offer new possibilities for musical expression by extending the capabilities of acoustic, electronic, and electroacoustic instruments. This study investigates the usability of the ARCube, an AR-based spatial audio controller, with twelve distinct musical instruments played by experienced musicians. We identify usability challenges specific to certain instruments, particularly for two-handed playing, as well as issues related to gesture recognition and cube stability. Our analysis shows that interaction patterns, such as cube placement, sound effect usage, and gesture strategies, vary significantly between instruments. These differences are driven by the physical form of the instruments, the required playing techniques, and user expectations for control and responsiveness. Based on these insights, we suggest directions for developing adaptable AR interfaces that better accommodate diverse instruments and support broader integration of AR technologies into musical practice.

Keywords

Musical Human-Computer Interaction, Augmented Reality Interfaces, Gesture-Based Control, Spatial Audio

1 Introduction

Augmented Reality (AR) technologies have been recognized as a transformative medium that blends the physical and digital worlds, which opens new possibilities for enhancing human experiences [1]. As AR applications have expanded, the design of AR interfaces has been studied to facilitate natural and intuitive interaction between users and augmented environments [2, 5]. Within this context, AR interfaces have been explored to generate, manipulate, and spatialize sound content in embodied ways for musical applications [11, 19, 28].

NIME '25, June 24–27, 2025, Canberra, Australia

While studies on AR in musical contexts have demonstrated its potential for creative musical expression, the enhancement of expressive capabilities in existing physical musical instruments has received limited attention. In particular, there is a need to understand how AR interfaces can meaningfully support and augment traditional musicianship in live performance contexts. Building on this observation, we examine the compatibility of AR interfaces with a diverse set of acoustic, electronic, and electroacoustic instruments. Our focus is not only on the technical integration, but also on how AR interfaces affect playability, control, and creative decision-making during performance.

To structure our exploration, we pose the following research questions:

RQ1 What are the key usability challenges in interacting with an AR interface across different instrument categories?
RQ2 What usability patterns emerge when interacting with an AR interface across individual instruments?

Based on the instrument affordances, we hypothesize that instruments requiring continuous two-handed operation will encounter greater integration challenges, whereas one-handed or hands-free instruments will afford more seamless interaction. In contrast, instruments that can be operated with one hand, such as pianos or synthesizers, are expected to integrate more smoothly. We further expect hands-free instruments to offer the most seamless interaction with the AR interface.

To investigate these aspects of AR-instrument integration, we conducted a user study involving twelve experienced musicians. Each participant engaged with the ARCube [22], an AR spatial audio controller, while performing on their primary instrument. The study included both structured tasks and creative performances to assess usability and expressive potential.

The remainder of this paper is organized as follows. Section 2 reviews relevant background and related work. Section 3 details the technological setup, followed by Section 4, which outlines the experimental design. Section 5 presents the results, and Section 6 discusses the findings.

2 Background and Related Work

2.1 Interfaces for Spatial Sound Control

Spatial sound control has been extensively studied as a means to enhance human-machine interaction and enable efficient task performance through auditory feedback [3]. Traditionally, this research has utilized 2D interfaces for adjusting the position of sound sources [7, 8]. With advancements in technology, interfaces for spatialization have evolved, incorporating haptic feedback [13, 18] and employing actual sensors to create sound spatialization instruments [14].

In the realm of virtual reality, various approaches have been developed for controlling parameters in interactive audio environments [10]. One notable example is the VR music environment LyraVR[1], which features a cube for controlling parameters in three dimensions (3D). This highlights the potential of VR environments to enhance sound spatialization and improve user interaction. Additionally, plugins such as dearVR[2] applicable to virtual reality have emerged for 3D audio production. These plugins are utilized in fields such as mixing, mastering, and sound design.

2.2 AR Interfaces

Augmented Reality overlays virtual objects onto the physical world to prevent collisions [17] and facilitates real-time interaction [6]. In this context, interaction becomes a key concept in AR [15]. To implement such interactions, various AR interfaces are utilized, and depending on the input method, different AR interfaces have been developed, including 3D user interfaces, tangible user interfaces, multimedia interfaces, natural user interfaces, and information browsers [24]. A specific example of these AR interfaces is their use in digital fabrication projects for visualization, where experimental and practice-based studies have emerged in recent years to assist unskilled workers with holographic on-site previews and instructional training in the field of architectural digital fabrication [23]. In Mobile Augmented Reality (MAR), users interact with MAR devices such as smartphones and wearable devices. These devices support a seamless transition from the physical world to a mixed environment with digital entities and enhance accessibility to digital content [6]. For human-robot interaction, AR tools are employed to work with the absolute positions of augmented objects in the physical space [20, 29].

2.3 AR in Musical Performances

Research on integrating AR into musical applications has evolved alongside advancements in Extended Reality (XR) technologies [25]. One of the key advantages of AR in music performance is its ability to provide real-time visual augmentations without fully immersing the user in a virtual space. While Virtual Reality (VR) places musicians entirely within a synthetic environment, AR enables the coexistence of physical and virtual elements. This approach preserves the tactile and acoustic properties of traditional instruments while introducing new forms of digital interaction [27]. Additionally, immersive environments enhance sensory experiences for both musicians and audiences, expanding the expressive potential of musical performance [4].

Studies on AR in musical performance have explored various interaction modalities and hardware configurations. Many AR systems capture musicians' movements and translate them into real-time audio or visual feedback [12]. Gesture-based interaction using spatial tracking has also been a focus and enables musicians to control sound in a more intuitive and immersive manner. [16, 21, 30].

Despite the advantages that digital environments offer for musical performance and instrument utilization, the application of AR in this context remains an under-explored area [9]. This study investigates the integration of AR interfaces with various musical instruments in live performance environments.

[1]https://lyravr.com/
[2]https://www.dear-reality.com/products/dearvr-pro-2

3 System and Setup

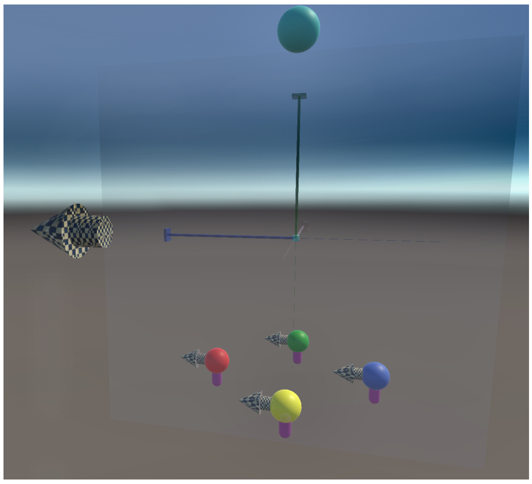

Figure 2: The ARCube interface.

3.1 ARCube Interface

The interface used in this study is the *ARCube*, an augmented reality controller, shown in Figure 2. It has previously been evaluated in a hybrid setup for spatial user interaction [22]. The ARCube is a semi-transparent cube with 30 cm edges, which can be placed anywhere in the physical space and rendered to the user through a head-mounted display (HMD; Meta Quest 3).

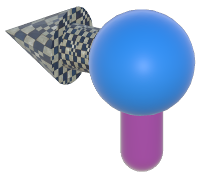

Figure 3: A single object for 6DOF gestural control.

The cube can be freely rotated, with an arrow indicating its forward direction and a sphere marking its top. These visual cues reveal the orientation of the cube to the user. Four color-coded objects (red, green, blue, yellow) are attached to the cube and can be independently repositioned and reoriented both within and beyond the cube. Figure 3 shows such an object in detail. An arrow indicates the front direction, a purple indicator points down. The position and orientation of each object relative to the cube represent a six degrees of freedom (6DoF) pose, expressed using nine parameters: Cartesian coordinates (x, y, z), spherical coordinates (azimuth, elevation, distance), and Euler angles (pitch, roll, yaw). As there are four tracked objects, a total of 36 parameters are transmitted. The flexible placement of the ARCube and associated objects in mid-air or next to physical instruments supports interaction within a hybrid setup that integrates physical and augmented elements.

3.2 Audio Effect

3.2.1 Processing. In the user experiment, participants used the ARCube to control a spatial audio effect. A primary design goal for the audio effects is to ensure their applicability to a wide range of instruments, including acoustic, electronic, and electroacoustic instruments. Since acoustic instruments emit their own sound, an additive audio effect was more effective as it avoided interfering with the original sound. Furthermore, the effects were designed to be sufficiently salient and easy for participants to perceive across all instruments within the experiment.

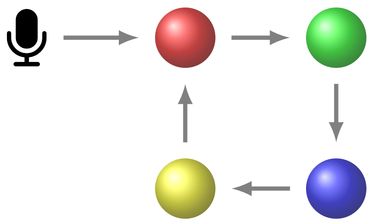

Figure 4: Audio signal flow for the spatial feedback-delay.

Given the guidelines above, a spatial feedback-delay effect was selected. Each of the four objects in the AR interface is connected to a processing unit on the rendering system. The instrument's sound is captured, using a line input or a microphone. For electronic and electroacoustic instruments, the original sound of the instrument is played back on the loudspeaker system without any processing. This part creates the same conditions as for the acoustic instruments. The unprocessed input is sent only to the red object, is then passed on to green, blue, yellow, and back to red, creating a closed loop as visualized in Figure 4. Each object performs the following processing to the incoming sound, before passing it on:

- Delay Time (0 ... 2 seconds)
- Feedback Gain (0% ... 99%)
- Pitch Shift (-2 octaves ... 2 octaves)

The spatial audio effect is rendered on a 28 channel 3D loudspeaker system with 5th order Ambisonics.

3.2.2 Control Mapping. The resulting effect is a feedback-delay with pitch shifting, that is distributed in space. Orientation and position of the objects in the AR interface control the parameters of the processing unit connected to it, as listed in Table 1. Depending on the parameter settings, the resulting effect can range from a spatially distributed rhythmic delay to a dense reverb or flanger effect.

4 Matchmaking Experiment

The matchmaking experiment examined how different musical instruments interact with AR-based spatial sound control, and also identified both instrument-specific challenges and common usability patterns. The experiment included structured tasks, free exploration, and a creative performance to assess the system's usability in live musical expression. While interacting with the setup, a first-person view video was recorded from each participant's HMD to allow additional analyses.

Table 1: Control parameters for each of the audio effects.

AR Object Control Parameter	Effect Parameter
Relative vertical position	Gain of the virtual sound source
Relative horizontal position	Location of virtual sound source
Relative distance to center	Delay time
Rotation around the z-axis	Feedback Gain
Rotation around the x-axis	Pitch shift

Figure 5: Berimbau player during the experiment.

4.1 Instrument Selection

The selection of instruments for the user experiment was based on an extended version of the Hornbostel-Sachs classification [26]. This version includes six categories to encompass the hybrid instrument and human voice. To represent each category, a total of twelve types of instruments were chosen: Marimba, Maracas, Berimbau, Violin, Cello, String Bass, Electric Guitar, Piano, Flute, Trombone, Trumpet, Modular Synthesizer, and Voice.

Table 2: Extended Hornbostel-Sachs Classification for Participated Instruments.

Category	Subcategory	Instruments
Idiophones	-	Marimba
Idio-Chordophones	-	Berimbau
Chordophones	Bowed	Violin, Cello, Double Bass
	Plucked	Electric Guitar
	Struck	Piano
Aerophones	Woodwind	Flute
	Brass	Trombone, Trumpet
Electrophones	-	Modular Synthesizer
Vocalizations	Natural Voice	Voice

4.2 Participants

We recruited 12 experienced musicians (P1–P12) to participate in the study (mean age = 27.2 years), including 8/12 males, 1/12 female, and 3/12 non-binary gender identities. Educational backgrounds ranged from high school completion to doctoral degrees.

In terms of musical experience, 10/12 participants had more than 7 years of experience playing their primary instrument, and 2/12 had 1–3 years of experience with intermediate to advanced levels of musical proficiency. Overall, 10/12 participants identified as advanced level, which included three professional musicians, and the minimum proficiency level was intermediate. This distribution ensured that all participants possessed sufficient musical expertise to evaluate nuanced interactions with an augmented reality interface.

Regarding technological familiarity, 2/12 participants reported a high level of experience with virtual reality (VR) environments, while most others indicated a moderate level of experience. For augmented reality (AR), 3/12 participants reported no prior experience, and the remainder indicated limited or moderate familiarity. In spatial audio, 2/12 participants reported strong experience, whereas 8/12 participants reported a high level of familiarity with electronic musical instruments.

4.3 Tasks

Each experimental session was conducted individually with a single participant and was divided into four parts:

(1) Setup
(2) Tutorial Tasks
(3) Free Exploration
(4) Creative Task

After performing the Tutorial Tasks, participants completed the first section of a survey, which devided into two sections. Once the experiment concluded, participants completed the second section of the survey, and also interview questions.

4.3.1 Setup. Participants were instructed to position the cube within the designated performance area (marked in pink). They were asked to identify a placement that maximized their ability to both play their instrument and manipulate the cube simultaneously, within the physical constraints of their instrument. Furthermore, participants were encouraged to explore and compare multiple placement configurations to assess their impact on interaction efficiency and performance feasibility.

4.3.2 Tutorial Tasks. This tutorial is designed to enable participants to naturally acquire an understanding of interactive sound control through hands-on object placement and parameter adjustment activities. Participants engage with spatial positioning, feedback levels, and pitch effects while playing their instrument. The tasks introduce key concepts in a step-by-step manner.

Table 3: Tutorial Task Groups.

Task Group	Description
Reset (Optional, If Needed)	T1, T2, T3
Delay with All Objects	T4, T5, T6, T7
Short Delay Times	T8, T9, T10
Reset	T11, T12, T13

The following list of tasks was presented to the musicians on a screen:

T1 Mute all effects by moving objects below the cube.
T2 Turn feedback level down (purple indicator down).
T3 Turn pitch effect off (arrow pointing to the front).
T4 Place one object on each of the four upper edges of the cube.
T5 Keep feedback off (purple indicator down) and pitch off (arrow pointing to the front).
T6 Play the instrument and sequentially adjust feedback settings:
- Increase feedback for the red object and play.
- Increase feedback for the green object and play.
- Increase feedback for the blue object and play.
- Increase feedback for the yellow object and play.

T7 Set arbitrary pitch values (arrows up or down) and play.
T8 Gradually move all objects toward the vertical center axis while playing.
T9 Maintain maximum feedback (purple indicator up).
T10 Adjust pitch settings dynamically while playing.
T11 Gradually turn off pitch effects (arrows pointing to the front).
T12 Gradually decrease feedback levels (purple indicator down).
T13 Move all objects below the cube to mute effects.

4.3.3 Free Exploration. Participants were instructed to freely explore the interactive possibilities of the cube in combination with their instrument. They were encouraged to experiment with various techniques and effect settings by adjusting the cube's position and modifying the placement of objects as needed. A maximum duration of 10 minutes was allocated for this phase.

4.3.4 Creative Task. The creative task is consisted of two parts, focusing on planning and executing a short performance with the instrument and the cube. The tasks included:

CT1 Performance Planning
CT2 Performance Execution

Participants were instructed to create a structured plan (CT1) for a short performance integrating their instrument and the cube, with an expected duration of approximately two minutes and the incorporation of the cube's effects. They documented their plan using pen and paper (e.g., notes, scores, sketches) and notified the experimenter upon completion, with a total of 15 minutes allocated for this task. Following CT1, participants proceeded to the performance execution phase (CT2), where they performed a single take of their planned performance. If desired, they were allowed to repeat the task up to two times.

5 Results

The results are based on a qualitative analysis of open-response questions (OQs), designed to gather detailed feedback on participants' experiences at each phase of the task. As summarized in Table 4, key challenges and suggestions are identified by each instrument.

OQ1 Please give detailed feedback on the placement of the cube with your instrument.
OQ2 Did you encounter any problems when placing the cube to play your instrument?
OQ3 What would be necessary to improve the placement options with your instrument?

OQ4 Please give detailed feedback on the use of the cube with your instrument.
OQ5 What are specific problems when playing your instrument with the cube?
OQ6 What are specific opportunities when playing your instrument with the cube?
OQ7 Describe your approach you used to create your performance.
OQ8 What are specific effects you were aiming for?

5.1 Cube Placement and Usability (OQ1, OQ2, OQ3)

5/12 participants (P1, P3, P6, P9, P10) positioned the cube in a way that allowed interaction without interfering with performance (OQ1). 3/12 participants (P3, P9, P10) placed the cube without requiring adjustments, as its initial position aligned with their performance habits. 2/12 participants (P1, P6) needed specific modifications to optimize accessibility. 4/12 participants (P2, P4, P6, P11) encountered stability issues, where the cube unintentionally moved during interaction (OQ2). P2 noted that grabbing the spheres sometimes caused the entire cube to shift, disrupting focus. P6 reported difficulties in aligning the cube perfectly straight, leading to unintended interactions. 2/12 participants (P4, P11) found that even slight movements caused by hand gestures could misalign the cube, requiring frequent repositioning. 3/12 participants (P5, P7, P12) experienced unintended grabs due to hand-tracking sensitivity, which occasionally misinterpreted normal playing gestures as grab motions (OQ2). P7 recommended a Lock/Unlock function to prevent misinterpretation of gestures. Another participant P2 suggested implementing a cube lock button to fix its position once placed, ensuring stability. 2/12 participants (P4, P6) proposed additional constraints that would allow for a stable vertical orientation while still maintaining flexibility in positioning (OQ3). Placement preferences also varied by instrument type. Stationary instruments, such as P1 (Marimba) and P7 (Piano), positioned the cube near their primary playing area to maintain clear accessibility. Bowed string instruments, including P3 (Violin) and P4 (Cello), positioned the cube slightly to the side to avoid obstructing bowing motion. For wind instruments, such as P9 (Trombone) and P10 (Trumpet), placing the cube in direct sight was essential for easy accessibility without disrupting embouchure control.

5.2 Interaction Challenges (OQ4, OQ5)

6/12 participants (P1, P3, P4, P6, P8, P10) found hand-based interaction challenging due to the need for both hands in traditional playing techniques (OQ4). 2/12 participants (P1, P3) adapted by using alternative techniques, such as two mallets in one hand (P1) or left-hand pizzicato (P3). P6 noted that hand-based controls disrupted performance, as guitarists typically use foot pedals. 2/12 participants (P2, P9) suggested alternative input methods, including head tracking and movement-based control using the instrument itself (OQ5).

5.3 Creative Opportunities (OQ6)

8/12 participants (P1, P3, P5, P6, P7, P9, P10, P12) identified new creative possibilities with the cube (OQ6). 2/12 participants (P1, P5) explored rhythmic and harmonic interplay using delay effects. P3 simulated multiple string instruments through pitch shifting. P6 suggested expanding interaction beyond a fixed position for movement-driven sound control. P12 found that percussive vocalizations responded more effectively than sustained singing, leading to rhythmic variations.

5.4 Performance Structuring and Sound Manipulation (OQ7)

7/12 participants (P2, P4, P7, P8, P9, P10, P11) structured their performances based on the cube's capabilities (OQ7). 2/12 participants (P4, P11) used delay-based rhythmic loops. 2/12 participants (P7, P10) explored spatial sound manipulation. P2 transitioned between pitched and percussive textures. P8 found it difficult to distinguish the cube's effects due to their instrument's natural resonance. 2/12 participants (P9, P10) suggested refining the pitch shift effect to improve musical coherence.

5.5 Auditory and Spatial Effects (OQ8)

9/12 participants (P1, P4, P5, P6, P7, P9, P10, P11, P12) integrated delay, pitch shifting, and spatial audio effects into their performances (OQ8). P6 experimented with 3D sound placement to transform familiar effects into new textures. P12 focused on vocal interactions with evolving delays. 2/12 participants (P1, P4) used repeated delays to reinforce rhythmic structures. 2/12 participants (P5, P11) suggested refining pitch control for more precise adjustments and extended sustain.

6 Discussion

6.1 RQ1: Key Usability Challenges Across Instrument Categories

We identified recurring usability challenges across instrument categories, mainly due to physical constraints, gesture recognition limitations, and insufficient feedback. Key usability challenges (UCs) include:

UC1 **Physical Constraints of Two-Handed Instruments:** Participants of two-handed instruments, including bowed chordophones (P3, P4, P5), struck chordophones (P7), and aerophones (P8, P9, P10), found it difficult to interact with the ARCube while maintaining their playing technique. Users often needed to manipulate the cube with one hand, which disrupted natural playing motions and compromised either instrumental execution or effect manipulation.
UC2 **Unintended Gesture Recognition and Misinterpretations:** 5/12 participants (P1, P2, P6, P9, P11) experienced issues with the ARCube misinterpreting gestures, particularly when attempting to adjust individual sound parameters. This caused unintended cube movements and made it difficult to distinguish between grab and pinch gestures. The lack of precise control over spatial placement was especially problematic for electronic musicians (P11), who relied heavily on real-time spatial effect adjustments.
UC3 **Limited Mobility for Performers Requiring Freer Movement:** Instruments that traditionally allow for significant performer mobility, such as plucked chordophones (P6) and idio-chordophones (P2), were constrained by the static placement of the cube. Guitarists (P6) found this particularly restrictive, as their performance style often involves moving between different physical positions. Similarly, the berimbau performer (P2) noted that repositioning the cube during performance was necessary but often difficult.

Table 4: Summary of Open-Response Feedback for individual instruments.

Instrument	Key Challenges	Notable Quotes and Suggestions
Marimba (P1)	Placement stability; Dropping mallets to interact	*"The cube did not always remain vertical when released." "I had to drop a mallet to interact with the spheres."* Suggested **head movements** for control.
Berimbau (P2)	One-handed limitation; Unintended cube movement	*"I would have to move the sphere around, test the feedback with one hand, then proceed to playing."* Suggested a **cube lock button**.
Violin (P3)	Requires both hands; Difficulty adjusting effects	*"I cannot interact with the cube while playing my instrument in any complex way."* Suggested a **way to map adjustments to head movements**.
Cello (P4)	Pitch control was unclear; Difficulty fine-tuning	*"No easy way to tell that pitch was neutral." "The pitch shift seemed extreme even at low settings."* Suggested a **more precise pitch-locking mechanism**.
Double Bass (P5)	Difficulty fine-tuning; Volume balance	*"I overestimated the pitch shift range; harmonies were quieter than expected."* Suggested **a better balance in harmonic adjustments**.
Electric Guitar (P6)	Limited mobility; One-handed interaction	*"The cube limits movement—if it filled the whole room, I could interact on a larger scale."* Suggested a **room-scale cube interaction model**.
Piano (P7)	Hand motion misinterpretation; Effects confusion	*"Initially, the delay and pitch mismatches felt unnatural." "I had to guess how the nodes were affecting sound."* Suggested a **real-time visual feedback system**.
Flute (P8)	Processed sounds masked by instrument; Hard to manipulate objects	*"Distortion and delay made me second-guess my playing."* Suggested **better sound clarity for real-time feedback**.
Trombone (P9)	Two-handed playing; Hard to adjust in real-time	*"If the cube tracked my slide hand, I could play and adjust at the same time."* Suggested **gesture-based tracking for brass players**.
Trumpet (P10)	Improper parameter modification; Limited control precision	*"Pitch and reverb lacked fine control."* Suggested **additional parameter controls for refinement**.
Modular Synth (P11)	Cube moved unexpectedly; Needed better placement stability	*"Misjudged my pinches, causing the cube to move."* Suggested a **fixed forward-facing placement mode**.
Voice (P12)	Feedback balance; Lack of control flexibility	*"I wish I could monitor my input through the speakers." "Pitch and feedback controls should be separate."* Suggested a **snapping function for pitch harmonization**.

UC4 **Auditory Monitoring Limitations with Acoustic Instruments:** Acoustic instruments (P5, P7, P8, P12) struggled with distinguishing the ARCube's processed audio from their natural instrument sound. In particular, aerophone players (P8, P9, P10) reported that delay and distortion effects masked their direct acoustic tone, making real-time interaction more challenging. Vocalists (P12) also expressed difficulty in hearing themselves through the processed signal, which led to an over-reliance on natural vocal monitoring.

UC5 **Unstable Cube Placement and Gesture Locking:** 4/12 participants (P1, P2, P6, P11) noted that unintended cube movement interfered with their performance. A *locking mechanism* to secure the cube in place was suggested to prevent accidental repositioning. This would be particularly beneficial for percussionists (P1) and modular synthesizer performers (P11), who require fine-tuned parameter control.

UC6 **Lack of Visual Feedback on Effect Parameters:** 5/12 participants (P3, P5, P7, P10, P12) expressed difficulty in gauging the cube's current parameter settings. This was particularly problematic for pitch control (P3, P4, P5) and delay feedback (P7, P10).

6.2 RQ2: Instrument Specific Interaction Patterns with AR Interface

AR interaction patterns show both commonalities and differences across instruments (Table 5). Cube placement followed a consistent trend among instruments. Participants positioned the cube based on their default playing posture to ensure ease of hand control while avoiding interference with body movement and existing playing techniques. Participants selected locations where the cube remained visible and accessible without disrupting their natural motion. Regardless of the instrument, cube placement was adjusted to maintain ergonomic efficiency and prevent obstruction of performance gestures. The effectiveness of AR-generated

Table 5: Instrument-Specific AR Interaction Patterns

Instrument	Cube Placement	Effective Sound Effects	Gesture Pattern of Interaction
Marimba (P1)	Centered on marimba, slightly above torso level	Rhythmic variations via delay and pitch shift	One-hand with mallet; Quick one-hand interaction between phrases
Berimbau (P2)	On microphone stand, adjusted for accessibility	Timbre-based rhythmic synchronization using delay	Instrument-integrated movement; One-hand interaction
Violin (P3)	At music stand height, directly in front	Harmonic shifting and simulated ensemble effects	Left-hand pizzicato; Open-string bowing with one-hand interaction
Cello (P4)	Slightly right of microphone for easy access	Layered harmonization via delay and panning	Left-hand pizzicato; Open-string bowing with one-hand interaction
Double Bass (P5)	Near but non-intrusive, easily controllable	Pitch-shifted harmonization and delayed rounds	Left-hand pizzicato; Open-string bowing with one-hand interaction
Electric Guitar (P6)	At head level, centered for equal hand access	Spatial sound expansion with reverb and delay	Foot pedal alternative control; Two-hand with effect control; one-hand interaction
Piano (P7)	Initially right above keyboard, later moved left	Loop feedback and delay for evolving harmonies	One-hand interaction; avoids pressing too many keys at once
Flute (P8)	Centered, no major placement concerns	Spatial diffusion to separate tonal elements	Pre-Setting; Quick one-hand interaction between phrases
Trombone (P9)	Directly in front, avoiding trombone slide	Dynamic pitch shifts and rhythmic delay	Pre-setting only
Trumpet (P10)	Centered, facing a display for visibility	Noise-based modulation and extended techniques	One-hand interaction
Modular Synth (P11)	Fixed position, occasional unintended movement	Freeform spatial movement of synthesized sounds	Both hands used for cube interaction and module control; One-hand interaction
Voice (P12)	Within reach but slightly obstructed by mic stand	Delay-based spatial interaction, echolocation effect	Two-hand interaction; Free hand movement during the performance

sound effects differed across instruments, primarily due to the timbral characteristics. Instruments with rich harmonic content, such as violin (P3) and cello (P4), produced expressive results with pitch shifting and layered harmonization. Percussive instruments, including marimba (P1) and berimbau (P2), effectively integrated rhythmic delay synchronization, which reinforced their attack-based articulation. For wind instruments, such as flute (P8) and trumpet (P10), tonal masking became a challenge. The natural tone of these instruments sometimes obscured the AR-generated effects, making spatial diffusion and dynamic modulation necessary to improve sound clarity.

Participants approached gesture-based interaction differently based on their instrument's physical constraints. One-hand interaction was used by 4/12 participants (P1, P3, P4, P7), who could temporarily disengage one hand during performance. Three participants (P3, P4, P5) adapted by using open-string bowing techniques to free one hand for cube interaction. Two-hand engagement was required for 2/12 participants (P11, P12), whose use of modular synthesizer and voice necessitated real-time control with both hands. Wind and brass players (P8, P9, P10) encountered significant challenges due to the continuous need for two-hand engagement. P9 (trombone) relied entirely on pre-setting effects before performance, while P10 (trumpet) used quick hand gestures to adjust effects in real time.

Two participants (P2, P6) suggested alternative interaction methods. P2 (berimbau) recommended integrating cube control into the natural motion of the instrument. P6 (electric guitar) envisioned a foot pedal system to enable hands-free effect manipulation.

7 Conclusion

This study explored the usability and creative potential of an AR interface across 12 different musical instruments. Key challenges included limited real-time control for two-handed instruments, unintended gesture recognition, unstable cube placement, and difficulty monitoring processed audio. Despite these issues, musicians engaged in spatial sound manipulation, hybrid performance techniques, and interactive improvisation. These findings suggest a need for alternative control methods such as head tracking or foot pedals to support seamless AR integration in musical performance. Furthermore, identified interaction patterns revealed instrument-specific strategies for interface placement, sound effect usage, and gestural control.

8 Ethics Statement

This study was approved by the Institutional Review Board (IRB) at Georgia Institute of Technology. All participants took part voluntarily. All collected data were anonymized and used solely for research purposes.

References

[1] Ronald T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6, 4 (1997), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

[2] Ronald T. Azuma, Yohan Baillot, Reinhold Behringer, Steven Feiner, Simon Julier, and Blair MacIntyre. 2001. Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications* 21, 6 (2001), 34–47. https://doi.org/10.1109/38.963459

[3] Durand Begault, Elizabeth M. Wenzel, Martine Godfroy, Joel D. Miller, and Mark R. Anderson. 2010. Applying Spatial Audio to Human Interfaces: 25 Years of NASA Experience. In *Proceedings of the 40th International Conference: Spatial Audio – Sense the Sound of Space*. Audio Engineering Society.

[4] Florent Berthaut, Myriam Desainte-Catherine, and Martin Hachet. 2011. Interacting with 3D Reactive Widgets for Musical Performance. *Journal of New Music Research* 40, 3 (2011), 253–263. https://doi.org/10.1080/09298215.2011.602693

[5] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. 2001. The MagicBook: Moving Seamlessly Between Reality and Virtuality. *IEEE Computer Graphics and Applications* 21, 3 (2001), 6–8. https://doi.org/10.1109/38.920621

[6] Jacky Cao, Kit-Yung Lam, Lik-Hang Lee, Xiaoli Liu, Pan Hui, and Xiang Su. 2023. Mobile Augmented Reality: User Interfaces, Frameworks, and Intelligence. *Comput. Surveys* 55, 9, Article 189 (2023), 36 pages. https://doi.org/10.1145/3557999

[7] Thibaut Carpentier. 2015. ToscA: An OSC Communication Plugin for Object-Oriented Spatialization Authoring. In *Proceedings of the 41st International Computer Music Conference (ICMC)*. Denton, TX, United States, 368–371.

[8] Thibaut Carpentier. 2018. A New Implementation of Spat in Max. In *Proceedings of the 15th Sound and Music Computing Conference (SMC 2018)*. Limassol, Cyprus, 184–191.

[9] Marko Ciciliani. 2020. Virtual 3D Environments as Composition and Performance Spaces. *Journal of New Music Research* 49, 1 (2020), 104–113. https://doi.org/10.1080/09298215.2019.1703013

[10] Thomas Deacon and Mathieu Barthet. 2023. Spatial Design Considerations for Interactive Audio in Virtual Reality. In *Sonic Interactions in Virtual Environments*, Michele Geronazzo and Stefania Serafin (Eds.). Springer International Publishing, Cham, 181–217. https://doi.org/10.1007/978-3-031-04021-4_6

[11] Bastian Dewitz, Roman Wiche, Chris Geiger, Frank Steinicke, and Jochen Feitsch. 2018. AR Sound Sandbox: A Playful Interface for Musical and Artistic Expression. In *Proceedings of the 9th International Conference on Intelligent Technologies for Interactive Entertainment*. Springer International Publishing, Cham, 59–76. https://doi.org/10.1007/978-3-319-73062-2_5

[12] Vanessa Faschi, Luca Andrea Ludovico, Federico Avanzini, Emanuele Parravicini, and Manuele Maestri. 2024. An Accessible Software Interface for Collaborative Music Performance. In *Proceedings of the 21st Sound and Music Computing Conference*. 150–157. https://doi.org/10.5281/zenodo.14337032

[13] Steven Gelineck and Dan Overholt. 2015. Haptic and Visual Feedback in 3D Audio Mixing Interfaces. In *Proceedings of the Audio Mostly 2015 on Interaction With Sound (AM '15)* (Thessaloniki, Greece). Association for Computing Machinery, New York, NY, USA, 14:1–14:6. https://doi.org/10.1145/2814895.2814918

[14] Florian Goeschke. 2022. The iOSCahedron: Developing a Hybrid Spatialization Instrument. In *Proceedings of the 17th International Audio Mostly Conference (AM '22)* (St. Pölten, Austria). Association for Computing Machinery, New York, NY, USA, 151–154. https://doi.org/10.1145/3561212.3561232

[15] Eg Su Goh, Mohd Shahrizal Sunar, and Ajune Wanis Ismail. 2019. 3D Object Manipulation Techniques in Handheld Mobile Augmented Reality Interface: A Review. *IEEE Access* 7 (2019), 40581–40601. https://doi.org/10.1109/ACCESS.2019.2906394

[16] Rob Hamilton and Chris Platz. 2016. Gesture-Based Collaborative Virtual Reality Performance in Carillon. In *Proceedings of the International Computer Music Conference (ICMC)*. Utrecht, The Netherlands, 336–341.

[17] Max Krichenbauer, Goshiro Yamamoto, Takafumi Taketom, Christian Sandor, and Hirokazu Kato. 2018. Augmented Reality versus Virtual Reality for 3D Object Manipulation. *IEEE Transactions on Visualization and Computer Graphics* 24, 2 (2018), 1038–1048. https://doi.org/10.1109/TVCG.2017.2658570

[18] Frank Melchior, Chris Pike, Matthew Brooks, and Stuart Grace. 2013. On the Use of a Haptic Feedback Device for Sound Source Control in Spatial Audio Systems. *Journal of the Audio Engineering Society* (May 2013).

[19] Ivan Poupyrev. 2000. Augmented Groove: Collaborative Jamming in Augmented Reality. In *Proceedings of the SIGGRAPH 2000 Conference Abstracts and Applications*. Association for Computing Machinery, 77.

[20] Giovanni Santini. 2019. Composing Space in the Space: An Augmented and Virtual Reality Sound Spatialization System. In *Proceedings of the 16th Sound and Music Computing Conference (SMC 2019)*. CERN, Málaga, Spain, 229–233.

[21] Giovanni Santini. 2020. Action Scores and Gesture-Based Notation in Augmented Reality. In *Proceedings of the International Conference on Technologies for Music Notation and Representation (TENOR)*, Vol. 20. 84–90.

[22] Hyunkyung Shin and Henrik von Coler. 2024. ARCube: Hybrid Spatial Interaction for Immersive Audio. In *Proceedings of the 2024 ACM Symposium on Spatial User Interaction (SUI '24)* (Trier, Germany). Association for Computing Machinery, New York, NY, USA, Article 33, 3 pages. https://doi.org/10.1145/3677386.3688883

[23] Yang Song, Richard Koeck, and Shan Luo. 2021. Review and Analysis of Augmented Reality (AR) Literature for Digital Fabrication in Architecture. *Automation in Construction* 128 (2021), 103762. https://doi.org/10.1016/j.autcon.2021.103762

[24] Toqeer Ali Syed, Muhammad Shoaib Siddiqui, Hurria Binte Abdullah, Salman Jan, Abdallah Namoun, Ali Alzahrani, Adnan Nadeem, and Ahmad B. Alkhodre. 2023. In-Depth Review of Augmented Reality: Tracking Technologies, Development Tools, AR Displays, Collaborative AR, and Security Concerns. *Sensors* 23, 1 (2023), 146. https://doi.org/10.3390/s23010146

[25] Luca Turchet, Rob Hamilton, and Anil Çamcı. 2021. Music in Extended Realities. *IEEE Access* 9 (2021), 15810–15832. https://doi.org/10.1109/ACCESS.2021.3052931

[26] Erich M. von Hornbostel and Curt Sachs. 1961. Classification of Musical Instruments: Translated from the Original German by Anthony Baines and Klaus P. Wachsmann. *The Galpin Society Journal* 14 (1961), 3–29.

[27] Rebekah Wilson. 2020. Aesthetic and Technical Strategies for Networked Music Performance. *AI & Society* 38, 5 (2020), 1871–1884. https://doi.org/10.1007/s00146-020-01099-4

[28] Jing Yang, Amit Barde, and Mark Billinghurst. 2022. Audio Augmented Reality: A Systematic Review of Technologies, Applications, and Future Research Directions. *Journal of the Audio Engineering Society* 70, 10 (October 2022), 788–809.

[29] Anıl Çamcı, Aaron Willette, Nachiketa Gargi, Eugene Kim, Julia Xu, and Tanya Lai. 2020. Cross-Platform and Cross-Reality Design of Immersive Sonic Environments. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2020)*, Romain Michon and Franziska Schroeder (Eds.). Birmingham City University, Birmingham, UK, 127–130. https://doi.org/10.5281/zenodo.4813270

[30] İzel Ergül, Tutku Çalış, Esra Yücetürk, Melis Gür, Selen Bulut, and Gökçe Elif Baykal. 2024. Co-Rhythm: Analyzing Children's Performative Gesture-Based Interactions in a Music Composition Tool. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference (IDC '24)* (Delft, Netherlands). Association for Computing Machinery, New York, NY, USA, 686–690. https://doi.org/10.1145/3628516.3659375