Collaborative Mobile Instruments in a Shared AR Space: a Case of ARLooper

Sihwa Park
Media Arts and Technology
University of California, Santa Barbara
Santa Barbara, CA 93106, USA
sihwapark@mat.ucsb.edu

ABSTRACT

This paper presents ARLooper, an augmented reality mobile interface that allows multiple users to record sound and perform together in a shared AR space. ARLooper is an attempt to explore the potential of collaborative mobile AR instruments in supporting non-verbal communication for musical performances. With ARLooper, the user can record, manipulate, and play sounds being visualized as 3D waveforms in an AR space. ARLooper provides a shared AR environment wherein multiple users can observe each other's activities in real time, supporting increasing the understanding of collaborative contexts. This paper provides the background of the research and the design and technical implementation of ARLooper, followed by a user study.

Author Keywords

Shared augmented reality, mobile interface, multi-user, sound recording, collaborative performance

CCS Concepts

•**Human-centered computing → Mixed / augmented reality;** *Collaborative interaction;* Mobile devices;

1. INTRODUCTION

Collaboration is a significant component of music performance wherein a group of people engages in to achieve a shared goal. Group members communicate in verbal or non-verbal ways, such as speech, gaze, and gesture, not to mention sound itself. Musical instruments also affect the way performers collaborate with each other according to how the instruments are designed and what communication capabilities they have. With the advancement of computing technology, designing new types of musical instruments and interfaces to support collective musical creativity has been one of the key topics in computer music and music technology. In this regard, the examples and contexts of collaborative musical interfaces are well documented by Jordà [9], Blaine and Fels [4].

Given by advances in mobile computing, the possibility of the use of mobile devices in NIME has been explored by a body of researchers, artists, and designers suggesting new performance paradigms that blur a boundary between audiences and performers and creating new interfaces for musical expression and experience that utilize the versatility of mobile devices built with various sensors and networking functionality. Furthermore, as Gaye et al. [7] point out, the mobility and collaboration capabilities of mobile devices have enabled the emergence of mobile music making. Since the arrival of smartphones, studies on mobile music interfaces and interactions have increased substantially.

Meanwhile, Augmented Reality (AR) also has been evolved with the increased computing power of mobile devices [1]. In particular, remarkable improvements in simultaneous localizing and mapping (SLAM) techniques have enabled mobile AR to build a map of the user's local environment without prior mapping information and ancillary equipment, such as external tracking devices and image markers, and to create synchronized AR spaces by sharing the map with multiple devices in real time. This markerless, SLAM based mobile AR makes the creation of collaborative AR spaces more feasible and available. Taking advantage of the advanced mobile AR, Google and Apple have been trying to make the shared AR experience more accessible to consumers with their own mobile AR development frameworks, e.g., ARCore[1] and ARKit[2], that allow developers to readily build shared, collaborative AR environments and applications.

The noteworthy features of AR environments for co-located collaboration are seamless interaction and the possibility to support natural communication behaviors that can be seen in unmediated face-to-face collaboration [3]. Most conventional screen-based interfaces for co-located collaboration separate the task space containing objects for tasks from the real world for communication among collaborators, causing a discontinuity. However, collaborative AR interfaces can enable users to see objects and each other simultaneously by fusing the virtual content for tasks with the communication space in the real world. As Billinghurst and Kato [3] insist, this seamless interaction space of which the task space becomes a subspace of the communication space could enhance not only the shared understanding of collaboration contexts but also non-verbal cues for communication.

The mobile instruments for collaborative musical performances inherently exhibit the limitations of the screen-based interfaces for collaboration as mentioned above. Since the small screen of mobile devices becomes the main space for musical tasks, performers need to pay attention to the screen itself, not collaboration. In the case of touchscreens, it becomes apparent because most interactions happen through touch interaction on the screen. Glimpsing others when they are idle, the performers with mobile instruments could rely on each other's sound as the main communication cue, but it would be difficult to have other layers of non-verbal interactions that can be found in face-to-face collaboration.

With collaborative AR, mobile instruments could enhance

[1]https://developers.google.com/ar
[2]https://developer.apple.com/augmented-reality/

NIME'20, July 21-25, 2020, Royal Birmingham Conservatoire, Birmingham City University, Birmingham, United Kingdom.

Figure 1: ARLooper multi-user mode. An iPad user's screen with the multi-touch GUI (left) and an iPhone user's screen showing the activity of the iPad user (right). In the circles of the multi-touch GUI, the names and values of sound filters are shown. A bounding box around a selected waveform appears in both screens.

the capability of communication for musical collaboration. Extending the task space from the 2D screen to the real world, collaborative mobile AR instruments have the potential to make group communication more natural and intuitive by enabling performers to see each other's activities along with virtual content. This enhanced communication could play an important role in supporting the awareness and knowledge of collaborators' activities within the shared workspace, which is one of issues in co-located interaction for musical collaboration [6]. However, the possibilities of collaborative mobile AR instruments have not been adequately investigated by the NIME community.

As a case study of these possibilities, this paper presents ARLooper that is an AR-based mobile interface for collaborative sound recording and performance. ARLooper empowers the musical ensemble to create a shared AR space where multiple users record, visualize, manipulate, and play sounds. The shared AR space is designed to support non-verbal communication by allowing the group members to see each other's sound recording and manipulation activities in real time. Compared to the previous version of ARLooper [13], this paper describes newly designed UI elements for supporting musical collaboration in a shared AR space. It also explains ARLooper's technical structure for collaborative AR in detail, followed by a user study and reflection on the system.

2. RELATED WORK

Before the recent advances in the markerless mobile AR, there has been remarkable research on physical marker-based collaborative AR musical interfaces that are worth recognition. Augmented Groove [14] is a collaborative AR musical interface with head-mounted displays, enabling the user to make music together with 3D virtual user interfaces and content appearing on shared physical 2D markers. Combining a tangible musical user interface with tabletop projection-based AR, Kaltenbrunner and et al. [10] suggested the collaborative models of the reacTable with which performers can play together on the same table for local collaboration or on the distant, networked multiple tables for remote collaboration. However, these systems need special settings and auxiliary equipment to create collaborative AR environments.

As for mobile music instruments, various research on collaborative music making with mobile devices has been conducted, employing the advancement of mobile computing. As an early example of mobile music interfaces, Tanaka [18] presented a PDA-based system that allows a group of users to engage in real-time music making via Wi-Fi networking. CaMus2 [16] used the cameras of multiple mobile phones and marker sheets for collaborative mobile performances. With the emergent of smartphones, the Stanford Mobile Phone Orchestra (MoPho) [12] explored the possibility of smartphones as musical instruments in the various forms for collaborative music performances, suggesting their own paradigm for mobile phone performances. D'Alessandro et al. [5] exploited the networking functionality of mobile devices to build systems for collaborative mobile music making. Keeping the user's focus on the screen, these mobile instruments, however, could have a discontinuity caused by a seam between the task space and the communication space.

Utilizing the advanced mobile AR, ARLooper tries to enable the mobile instruments to benefit from the features of collaborative AR interfaces; By using only mobile devices, the creation of a shared AR environment could become much more convenient; Making the performer be in part of the task space could help with the intuitive understanding of collaboration context and the enhancement of communication among collaborators.

3. DESIGN

This section describes design aspects of ARLooper that aim to support non-verbal communication for collaboration, in addition to elements for musical AR interfaces. The video documentation of ARLooper is available at `https://sihwapark.com/ARLooper`.

3.1 Sound Recording and 3D Visualization

The sound recording interaction of ARLooper is akin to drawing sound-reactive 3D lines. ARLooper begins to record sound from a device's microphone when the user taps a record button. The sound is simultaneously visualized as a 3D tube-like waveform of which the radius and color brightness are determined by amplitude and the center points follow the device's location in space (see Figure 2 left).

3.2 User ID Color

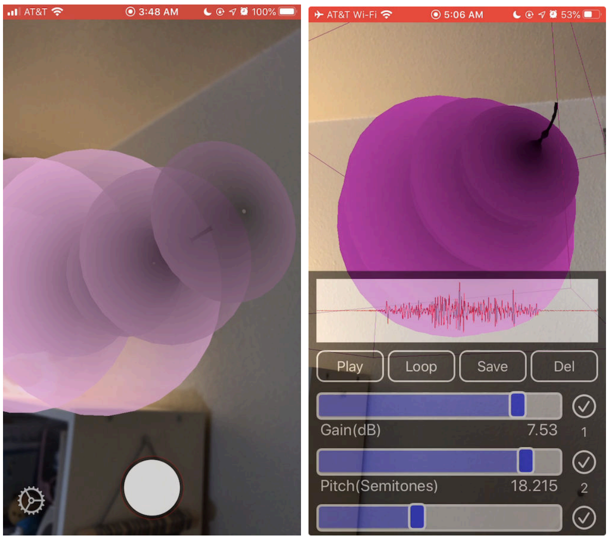

Figure 2: Waveform visualization (left) and sound control GUI (right).

As part of design elements to support collaboration, ARLooper uses color to distinguish the ownership of waveforms. Each user is assigned an ID color automatically that fills the color of waveforms recorded by the user. When a user selects a waveform, a colored bounding box is rendered to represent the user who has selected. When a waveform is playing, the playback position of the waveform is indicated by the ID color of the user who triggered playback. Through the color setting GUI, the user can change their ID color and see the list of joined users and their ID colors (see Figure 3).

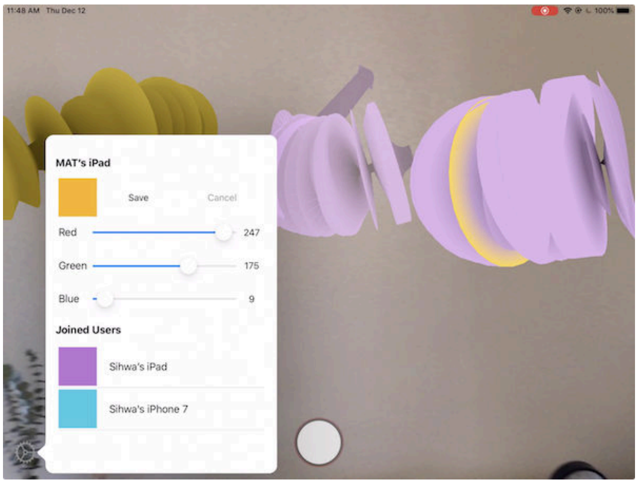

Figure 3: Color setting GUI.

3.3 UIs for Sound Manipulation

ARLooper attaches a set of sound filters and effects to each recorded sound, e.g. gain control, pitch shifter, delay effect, low pass filter, tremolo, and reverb. As shown in Figure 2, the user can change parameters of these filters in a slider-based GUI including a conventional 2D waveform viewer and control buttons. This GUI is presented when the user taps a waveform. Changes to parameters are synced across all devices joined the AR space and part of parameters such as pitch, delay feedback, and delay time are visualized as waveform animation effects.

ARLooper also has a multi-touch GUI for immediate control based on *Catch and Release Interaction* (see Figure 1 left). When the user double taps a waveform, it is fixed at the center of the user's screen with a smaller size and draggable circles appear within a boundary circle around the waveform. The draggable circles are mapped to parameters that are chosen up to five via their associated checkboxes in the slider-based GUI. The distance between the screen center and each circle is proportionally mapped to the parameter value. With this GUI, the user can more promptly and concurrently change multiple parameters. When the user presses a close button or double taps the waveform again, the GUI disappears and the waveform moves back to the original location in an AR space.

3.4 Dominant User and Remote Control

Since the user shares waveforms with others who are in the same AR space, there is a possible conflict in accessing these shared virtual resources. To resolve this matter, ARLooper implements a dominant user protocol by which only one user can control a waveform at a time. If a waveform is selected by a user, others cannot access it until it is deselected. However, if the user is left the waveform after playing it, any other user can remotely change its sound parameters. These rules can make a collaborative situation wherein user A (purple ID) manipulates a waveform recorded by user B (yellow ID) while its sound is coming out from the device of user C (blue ID) (see Figure 1).

4. IMPLEMENTATION

ARLooper is an iOS application built with the Swift programming language, Apple's ARKit 2.0, and an open-source sound synthesis, processing, and analysis framework for Apple devices, AudioKit[3]. Owing to a lack of reference resources for learning how to implement a multi-user AR application with ARKit, I referred to Apple's ARKit sample project SwitShot[4] and adapted portions of the ARLooper system from it. These portions include data encoding and decoding, event synchronization, and the establishment and management of multi-user AR sessions.

4.1 AR World Tracking and Shared AR

ARLooper can be operated in single-user mode or multi-user mode. For both modes, ARLooper follows the world tracking workflow of ARKit which uses a technique called *visual-intertial odometry*. First, it extracts the feature points of physical objects and surroundings from images obtained from the iOS device's camera. The result of the computer vision analysis is combined with the device's motion sensing data to infer the position and motion of the device in the space. This process creates an `ARWorldMap` instance, which includes the data of detected planes in addition to the mapping information of the space. And this instance is shared as a frame of reference to initiate the multi-user AR mode.

4.1.1 Single-user Mode

In single-user mode, the user starts as a host by selecting a host button and begins detecting planes in the given space with the device's camera. During the plane detection, a grid to visualize a detected surface is shown on the screen. The user needs to scan planes enough because the more the detected planes are, the more stable the AR tracking is. The user can start an AR session by pressing a start button.

[3] `https://audiokit.io/`
[4] `https://developer.apple.com/documentation/arkit/swiftshot_creating_a_game_for_augmented_reality`

4.1.2 *Multi-user Mode and Synchronization*

For multi-user mode, ARLooper uses the **MultipeerConnectivity** (MC) framework, which enables peer-to-peer (P2P) networking among Apple devices by supporting the discovery, connectivity, and data communication of nearby devices. To establish the P2P networking, one user needs to be the host by taking steps described for single-user mode. After the host starts the AR session, ARLooper on the host device advertises the session information with the device's name that can be found by nearby devices. Other users then join the session as guests by pressing a join button and selecting the session name from a list of available sessions, thereby establishing the P2P network among the users in the same session. ARLooper on the guest device requests and receives an `ARWorldMap` instance from the host device. In order to relocalize to the shared map, each guest must find a similar view of the surrounding environment that the host detected before starting the session. And the host has to wait until all guests join the same AR session before recording any sound. If the relocation is successful, the GUIs for recording and color settings appear on the guest's screen.

The synchronization of data and user actions is achieved by three types of data communication on the P2P network: First, sending a relatively smaller size of data including the changed parameter values of sound filters or effects and action messages, such as the request of the world map data, the start and end messages of recording, ID color, and the selection or manipulation of a waveform. Second, sending large data, such as recorded sound files or the world map. While the action messages are directly serialized for transmission, the large data is sent to the recipient device after first being stored as temporary files on the sender device for trackable transmission. Last, streaming real-time sound amplitude and device position information for recording. The streaming type is better suited for continuous sensor readings or position updates than the above types of data transmission. The first two types of data communication use the methods of `MCSessoin` objects whereas the streaming transmission is made with the objects of `InputStream` and `OutputStream`. All of the data communication types use the modified version of SwiftShot's `BitStreamCodable` protocol, which extends Swift's `Codable` protocol for data encoding and decoding.

4.2 Sound Recording and Control

ARLooper creates the two graphs of audio nodes: A global graph for recording and an individual graph for waveform manipulation. To record sound, ARLooper creates the global graph in which AudioKit's `AKMicrophone`, `AKMixer`, `AKBooster`, and `AKNodeRecorder` nodes are connected in sequence. `AKNodeRecorder` is also capable of writing recorded sounds to files. The `AKMicorphone` node is also attached to `AKLazyTap` to copy audio data into a buffer that will be used for 3D waveform drawing.

Whenever ARLooper begins recording, it creates an instance of `WaveformNode`, which is a custom subclass of `SCNNode` for rendering 3D content. At each frame of ARKit scene rendering, the 3D position of the device screen center relative to the world coordinates and the amplitude value of the audio buffer that is being recorded are provided to the `WaveformNode` instance that draws a circle. The device position becomes the centroid of the circle and the amplitude determines the radius and brightness. As the device moves in space, the series of circles are connected sequentially, creating a disk-like 3D waveform.

Each `WaveformNode` has its own graph of audio nodes consisting of `AKPlayer`, `AKPitchShifter`, `AKOperationEffect` as a delay effect, `AKLowPassFilter`, `AKTremolo`, `AKCostelloReverb`, `AKBooster`, and several `AKDryWetMixer` nodes. When recording is finished, a recorded sound file is loaded to `AKPlayer`. The parameters of these audio filters and effects are exposed in the slider-based GUI or the multi-touch GUI. To represent the change of waveform status, such as selection, playback, and the changes of delay feedback and time, vertex and fragment shaders written with the Metal Shading Language[5] are applied. The pitch change is visualized as the rotation of a waveform.

5. USER STUDY

Considering that ARLooper is a work-in-progress project, task-based design evaluations would be inappropriate. Instead, the examination of user experience with the interface could provide reflective insights on both the instrument itself and the experience afforded by the interface as Johnston insists [8]. The analysis of reflective and qualitative evaluation could also be helpful in improving the current design as part of an iterative design and development process.

In this regard, I conducted an exploratory user study, including two sessions, questionnaires, a semi-structured group interview, and data analysis. The structure and content of the study are borrowed from other evaluation approaches [15, 11, 17, 2] but adapted in consideration of the design aspects of ARLooper. The main goal of the user study is to evaluate the user experience with ARLooper in a collaborative musical context, attempting to address questions: How can ARLooper support the performers of collaborative musical performances in a shared AR space? How does ARLooper enhance non-verbal cues for communication?

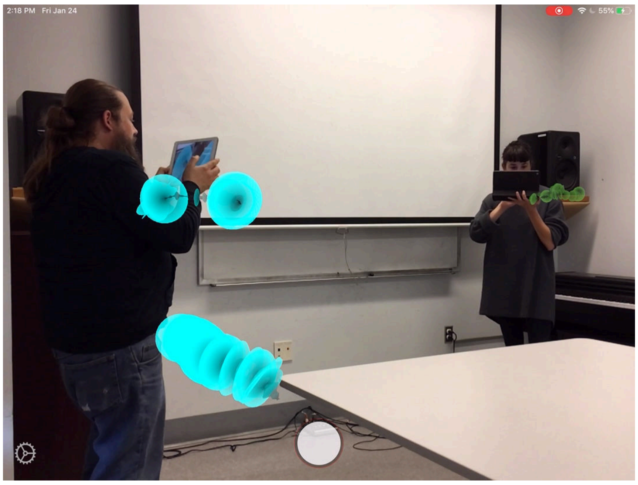

Figure 4: User study with ARLooper.

5.1 Method

One of the UCSB MAT classrooms was chosen as a user study space. One reason for choosing this classroom was its bright lighting that facilitates better AR tracking quality. Three iPads were used. Two devices were given to participants and one device was used as an observer in the shared AR space. Also, the activities of participants during the evaluation were recorded with a DSLR camera.

Participants engaged in two sessions: The first session with an iPad looper music application and the second session with ARLooper. Each session consisted of a solo ex-

[5]https://developer.apple.com/metal/Metal-Shading-Language-Specification.pdf

ploration phase and a group improvisation phase. The sessions were designed to be conducted with a group of two participants. In the first session, participants were given an iPad looper application downloaded from the App Store and asked to explore the interface for about five minutes. They were also suggested to record two or more sounds, and to play and control them. In the group phase, they were requested to improvise together with the same interface for another five minutes. The purpose of the first session was to give a brief understanding of sound looper applications and a baseline experience with mobile music instruments. In the second session, participants followed the same procedures with ARLooper. Before staring the group improvisation, I helped with setting up multi-user mode, explaining how it works. At the beginnings of both solo and group phases, information about how to use the interface, the concepts of the GUIs, and interaction rules in multi-user mode were provided. After the sessions, participants filled out the questionnaire that asked about their musical background and the experience with ARLooper.

With all participants, the semi-structured group interview was administered while watching the recorded videos of the group improvisation in the second session in order to recall their memory. In the interview, participants were given open-ended questions that encouraged them to share their experience with ARLooper during the improvisation.

5.2 Data Analysis and Findings

5.2.1 Participants

The study involved five participants who had experience with collaborative music performance. They were recruited from my graduate program and four of them were also members of the UCSB Center for Research in Electronic Art Technology (CREATE) Ensemble which I also participate in. Two had formal musical training in undergraduate studies and three had private instruction. All had experience with electronic or digital music instruments and had played in a group. Only three had prior experience with mobile music instruments and two among them had experienced mobile AR. The participants divided into three groups of two players and I took part in the third group to match the number of players in a group.

5.2.2 Video Analysis

The videos of the group phases were analyzed to understand user behaviors and activities. Players of both Group 1 (G1) and Group 2 (G2) utilized the room space more with ARLooper compared to their behavior with the non-AR looper application. In the group phase With the non-AR looper, the players stood close to each other to see what their partner was doing on their screen. At the beginning of the group phase with AR Looper, they rather focused on interacting with waveforms, but after playing one or two sounds, they began to see each other through the screen. One player of G2 frequently glanced at their partner during the first session, but during the second session, the player only watched the ARLooper screen. Contrary to the design intention of the multi-touch UI, most players did not drag circles simultaneously but instead adjusted them one by one. Also, it seemed some players had a hard time selecting a waveform that was in distance or overlapped with other waveforms. It was found that the remote control feature was not intuitive to understand without an explanation.

5.2.3 Questionnaire

The questionnaire was designed to evaluate how ARLooper supports the awareness of collaborators in a shared AR space and whether participants find non-verbal communication cues afforded by the design of ARLooper helpful and effective in collaboration. For the below questions, participants were asked to give a score between 1 (strongly disagree) and 5 (strongly agree).

Q1. It was difficult to play.
Q2. I felt connected to the other performers.
Q3. I enjoyed the improvised music making task.
Q4. I felt part of a collaborative process.
Q5. It was difficult to understand what the others are doing.
Q6. I was influenced by the playing of the others.
Q7. I mainly relied on sound to understand what the others are doing.

As shown in Figure 5, the participants generally disagreed that ARLooper was difficult to play (Q1). In terms of the awareness of collaborators (Q2) and a sense of participation (Q4), their experience was shown as quite positive. While the participants were highly satisfied with the collaborative AR experience (Q3), they also perceived the main design aspect for collaborative AR as effective for understanding others' activities (Q5). However, it could not mean that the design of ARLooper facilitates non-verbal cues for collaborative communication because it was clear that the sound was still the main cue for communication, especially in a musical context (Q7). Although most participants were affected by others' playing (Q6), it could be because of the nature of improvising, collaborative musical performances requiring to pay attention to each other's sonic output.

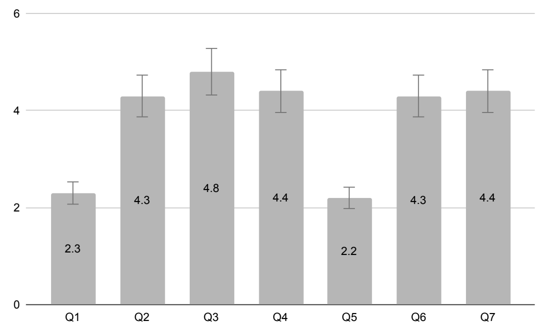

Figure 5: Questionnaire result (n = 5).

The participants were also asked to choose the design elements of ARLooper that they found to be helpful for communication. All participants agreed that real-time shared waveforms and coloring to represent the ownership of waveforms were helpful. While four satisfied with observing their partner with waveforms, only two chose the remote sound control feature. The waveform animation effects and bounding boxes were chosen by three. The participants were also requested to write what other cues apart from sound they tried to find from their partner. Most participants answered that it was body gestures such as head nodding, gaze, and the orientation of the device.

The questionnaire also encouraged the participants to suggest any ideas or comments about the interface. One of the interesting suggestions was that additional information above the waveform such as the name of the creator would be helpful. One participant commented that sound was expected to be spatialized. Some issues were called to attention. The multi-touch UI, for example, needs to loop the playing sound even if no circle is touched. Moreover, holding the tablet during the performance for a long time was ergonomically not ideal. Bounding boxes also need improvements to better represent the occupancy. And a larger space would be better for collaborative AR performances.

5.2.4 Group Interview

During the group discussion, the participants were asked to share their overall experience in collaboration with ARLooper. Most feedback was similar to their written suggestions described above. Overall, sharing waveforms and bounding boxes definitely played a role in understanding the shared context and resources. In terms of working in a shared AR space, one participant liked the fact that the instrument makes the performer move around the space to create waveforms or go from one waveform to another but it is not necessarily associated with collaboration. For this matter, sound spatialization was naturally suggested. Having an additional cue to represent the performer's status waiting for having access to the waveform selected by another performer was also recommended. Finally, the participants suggested that using loudspeakers rather than iPad speakers would result in a higher-fidelity performance environment.

6. DISCUSSION AND FUTURE WORK

ARLooper is an AR-based mobile instrument that supports multi-user sound recording and performance in a shared AR space. In terms of the awareness of collaborators and the understanding of others' activities, its design aspects demonstrate the potential of leveraging collaborative mobile AR in enhancing collaborative mobile music making.

Findings from the user study and the design reflection give a meaningful direction that ARLooper can follow for improvement. First, more stable P2P networking is required to prevent intermittent disconnection caused by intensive data communication in a short period. Second, adding more animation effects to visualize sound parameter changes could provide the user with a more intuitive and immersive user experience and a better understanding of collaborative context. Third, on top of improving the bounding box design, showing who is waiting for access to a waveform could enhance communication in using shared AR resources. Fourth, sound spatialization based on the proximity between the user and 3D waveforms would be a reasonable step to be taken for spatial interaction design in AR. Last, to have a better speaker configuration and a larger space will be an important aspect of actual music performances with ARLooper.

7. REFERENCES

[1] C. Arth, R. Grasset, L. Gruber, T. Langlotz, A. Mulloni, and D. Wagner. The history of mobile augmented reality. *arXiv preprint arXiv:1505.01319*, 2015.

[2] B. Bengler and N. Bryan-Kinns. In the wild: evaluating collaborative interactive musical experiences in public settings. In *Interactive experience in the digital age*, pages 169–186. Springer, 2014.

[3] M. Billinghurst and H. Kato. Collaborative augmented reality. *Communications of the ACM*, 45(7):64–70, 2002.

[4] T. Blaine and S. S. Fels. Contexts of collaborative musical experiences. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Montreal, Canada, 2003.

[5] N. d'Alessandro, A. Pon, J. Wang, D. Eagle, E. Sharlin, and S. Fels. A digital mobile choir: Joining two interfaces towards composing and performing collaborative mobile music. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Ann Arbor, Michigan, 2012.

[6] R. Fencott and N. Bryan-Kinns. Computer musicking: Hci, cscw and collaborative digital musical interaction. In *Music and Human-Computer Interaction*, pages 189–205. Springer, 2013.

[7] L. Gaye, L. E. Holmquist, F. Behrendt, and A. Tanaka. Mobile music technology: Report on an emerging community. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Paris, France, 2006.

[8] A. Johnston. Beyond evaluation : Linking practice and theory in new musical interface design. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Oslo, Norway, 2011.

[9] S. Jordà. Multi-user instruments: Models, examples and promises. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Vancouver, BC, Canada, 2005.

[10] M. Kaltenbrunner, S. Jorda, G. Geiger, and M. Alonso. The reactable*: A collaborative musical instrument. In *15th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'06)*. IEEE, 2006.

[11] R. Laney, C. Dobbyn, A. Xambó, M. Schirosa, D. Miell, K. Littleton, and N. Dalton. Issues and techniques for collaborative music making on multi-touch surfaces. In *Proceedings of 7th Sound and Music Computing Conference*, Barcelona, Spain, 2010.

[12] J. Oh, J. Herrera, N. J. Bryan, L. Dahl, and G. Wang. Evolving the mobile phone orchestra. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Sydney, Australia, 2010.

[13] S. Park. ARLooper: a mobile AR application for collaborative sound recording and performance. In *Proceedings of the International Symposium on Computer Music Multidisciplinary Research*, pages 233–240, Marseille, France, 2019.

[14] I. Poupyrev, R. Berry, J. Kurumisawa, K. Nakao, M. Billinghurst, C. Airola, H. Kato, T. Yonezawa, and L. Baldwin. Augmented groove: Collaborative jamming in augmented reality. In *ACM SIGGRAPH 2000 Conference Abstracts and Applications*, volume 17, page 77, 2000.

[15] R. Pugliese, K. Tahiroglu, C. Goddard, and J. Nesfield. A qualitative evaluation of augmenting human-human interaction in mobile group improvisation. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Ann Arbor, Michigan, 2012.

[16] M. Rohs and G. Essl. Camus2: Collaborative music performance with mobile camera phones. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology*, ACE '07, page 190–195, New York, NY, USA, 2007.

[17] D. Stowell, M. D. Plumbley, and N. Bryan-Kinns. Discourse analysis evaluation method for expressive musical interfaces. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Genoa, Italy, 2008.

[18] A. Tanaka. Mobile music making. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Hamamatsu, Japan, 2004.