

Computer-Supported Musical Improvisation

Torin Hopkins
ATLAS Institute
University of Colorado Boulder
Torin.Hopkins@colorado.edu

ABSTRACT

Music is a connecting force. It blends the nuances of human expression and culture, enabling autotelic, social interactions. Music can move us to dance, or to cry, it is uniquely present in all cultures, and is a celebration of our shared social humanity. Musical improvisation and social musical play have been practiced for thousands of years, among almost all cultures, throughout human history. I aim to leverage newfound knowledge and technologies to imbue new music experiences with social musical play and carry forward values of connection to others through music.

During my PhD I would like to answer three research questions in the following ways: 1) identify the role of technology in music improvisation, 2) identify investigations that may be well positioned currently to support musical improvisation, and 3) create and evaluate systems that push the boundaries of technology and musical improvisation. Throughout this paper I outline the a framework for which I position my work, explain how past projects forward my aim, and seek feedback from the NIME community.

Author Keywords

Music Improvisation, Music Networks, Extended Reality, Neuroscience, Artificial Intelligence

CCS Concepts

•Applied computing → Sound and music computing; Performing arts; •Information systems → Music retrieval;

1. INTRODUCTION

My research focuses on using technology as a means to aid in musical improvisation and the communication of musical ideas between players. In order to create systems that explore technology as a tool to better communicate our shared experiences, I focus on four areas that can enable us to better support the act of musical improvisation: 1) sensing, 2) processing, 3) representing, and 4) networking musical information.

In this paper I discuss related work categorized by the four areas. The various methods for sensing, processing, representing, and networking musical information inspire my work in creating tools for musical improvisation and collaboration. I then discuss my work and how it incorporates these techniques to create various tools for supporting improvisation.

Considering the four areas of focus, I aim to answer three research questions: 1) identify the role of technology in music improvisation, 2) identify investigations that may be well positioned to support musical improvisation, and 3) create and evaluate systems that push the boundaries of technology and musical improvisation.

The paper is organized into four sections. Section 2 describes work in the field of computer-supported music systems that have inspired my work and the work related to developing systems for real-time sensing, processing, representing, and networking musical information. Section 3 describes the work I have undertaken in the field of computer-supported music improvisation. Section 4 then describes what I hope to gain and contribute to the student consortium.

2. RELATED WORKS

Sensing, processing, and representing musical information is quintessential to providing support for musical improvisation. Sensing, processing, representing, and networking information for real-time, in-the-moment, music improvisation is one of the most computationally demanding social activities to support [8]. Though explicit timing requirements for such systems will be discussed in later sections, this chapter focuses on common techniques for sensing, processing, representing, and networking musical information. This chapter is focused on what has been done in the field of music interaction and computing, and will set the stage for how I position my work.

2.1 Sensing Musical Information

Though there are many methods and entire fields dedicated to the aim of sensing musical information (Music Information Retrieval being one), I will discuss the most commonly utilized for collaborative music experiences. Methods for retrieving musical information are commonly executed at a level continuous with information present in sheet music, reducing much of the musical information collected through a lens of pitch and time. This is critical for several overarching goals in music processing and are also a reflection of how we communicate about music to each other and to machines.

Music Instrument Digital Interface (MIDI) is a protocol that does just that—reduces musical information to its basic components of pitch and time. This was originally created



Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

NIME'23, 31 May–2 June, 2023, Mexico City, Mexico.

to connect signals between the various digital instruments emerging at the time. The MIDI protocol has been used ubiquitously ever since, acting as the main protocol for almost all digital musical instruments today, underlies almost all music notation software, and is often used as compressed data for machine learning algorithms and networked musical experiences.

Common sensors such as piezoelectrics, resistive and capacitive sensing are often underlying technologies for MIDI driven controllers and musical information acquisition devices. Often controllers have discrete sensors that are located at a position of the instrument where sound is expected to result from (e.g. keys of a keyboard) and can be associated with a particular pitch. The loudness (also known as dynamics in music) can be determined by how hard the sensor was pressed and timing can be calculated based on internal clocks. This method of converting sensed information to musical information is paramount for not only the creation of new musical instruments, but also the capturing of musical information through sensing technology.

Sound feature analysis is another technique often used to convert audio signals into MIDI information and other musical information. This is often accomplished using a fast-fourier transform (decomposition of complex waveforms into simple sine or cos waves, used for finding a dominant frequency) or other form of audio processing. The techniques used in sound feature analysis vary greatly but often use signal processing techniques to distill musical information from audio. This is used heavily for the use of tuning apps provided on smart phones by directly converting the sound to a note frequency through the fast-fourier transform [28]. This technique has also been used to quickly convert notes to MIDI and convert traditional string instruments into MIDI instruments by simply adding new piezoelectric pickups. Others have used capacitive sensing [1], or simply audio analysis for the aforementioned purpose of learning, composing, or playing a game.

Another common form of ascertaining musical information, first established by Jordá, describes the use of fiducial marker detection for collaborating on musical compositions. Marker detection for music has been used before, but is not generally utilized for note-level musical interaction because of processing delays. The use of fiducial marker detection can be used for higher level musical features such as tone and initiating pre-composed music, which is precisely how it is used in the ReacTable and other collaborative tabletop musical interfaces [14, 15].

2.2 Processing Musical Information

Simple algorithms for improvisation include suggestions for solos, scales, and key information from backing tracks (pre-recorded music to play along to) or songs. This occurs often by containing a database of music that matches a particular key (grouping of notes that match a major scale) [10]. In this way pre-recorded tracks (solo's in MIDI or common chord progressions) can be stored and applied when an algorithm detects that they likely fall into a key. This enables players to begin practice and be told what notes sound harmonious on their instrument. This is very helpful for players who do not have a lot of knowledge in scales or can not identify chords well enough to discern a key.

Deep learning strategies and other machine learning models are used most often for generative purposes. The Magenta team at Google have created a python library that enables musicians and coders to train machine learning models that will produce several types of generative and intel-

ligent responses. One of such models creates a call and response game between the human and the generated output [25]. Following the structure of a general call and response improvisation between two musicians the model generates a reliably interesting musical response. Similarly, Dubnov describes the history of responsive systems using information theory as the guiding mathematics for detection and response, detecting notions of entropy, and developing sophisticated statistical models of musical events [19]. These programs are largely created from MIDI music databases. This technique drastically cuts down on the amount of information that needs to be incorporated for processing. The compressed musical message works with all MIDI capable instruments and can be trained with personal databases.

Physical robots are used to make playing more fun for improvising musicians. In the case of a robotic drummer, Haile, there are considerably less variables in a drumming improvisation than there are in a polyphonic musical jam. Shimon, the marimba playing robot accomplishes a similar task of playing with other musicians but using the entire polyphonic marimba, alongside a piano and guitar player [27]. This task is still accomplished using the MIDI protocol and MAX software, but uses a much larger network to accomplish musical improvisation. Shimon also comes with the additional feature of demonstrating that it is listening by using the motion of its head. Recently, Shimon has been given extra facial features (such as eyebrow and lips), the ability to sing, play, and react to vocalizations from themselves or others [19]. This is accomplished by a combination of machine learning models, particularly Long Short-Term Memory, recurrent neural networks, markov models, and MIDI data. MIDI information informs song related information, markov models handle rhyming, and neural nets and long short-term memory handle the production of speech [19].

2.3 Representing Musical Information

Representing musical information refers to audio, visual, and in some cases haptic information associated with communicating about a musical activity. Representation of musical information includes music notation, information about what other players are doing and trying to communicate, avatars, and a shared virtual jam space.

Traditionally, collaborative musical improvisation is often coupled with an underlying theme or familiar tune. Though jamming is more popular in genres such as jazz and blues music than in many other western music traditions, it is nevertheless a skill available to musicians in all genres of music. A common form for representing music, both for performance and for jam sessions, is the lead sheet. A lead sheet usually consists of the chord root and quality contained within a bar of a given number of beats. A lead sheet may also contain written music, but it is not an essential component of the notation. Additionally, instead of the letter name for a chord, a roman numeral is given instead to represent the chord in reference to the key or tonal center of the song (see Figure 1). Commonly used technology for this type of musical representation is lead sheet software, playable lead sheets, and video-based lead sheets and chord charts found on video sharing services such as YouTube.com.

Some experimental forms of music notation utilize color coordination among the players in order to guide musical action and intention [21]. Chords as well as individual notes have been colored to give greater access to musical knowledge for novice musicians. Experimentally players have reported being able to improvise successfully over genres ranging from pop, blues, rock, advanced jazz [10]. Addition-

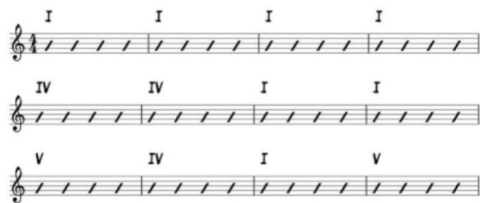


Figure 1: Image of a commonly used form of music notation for musical improvisation.

ally, music notation has been extended beyond musicians to provide other musical actors, such as dancers, with cues for body movement or other sound-movement related affordances [3].

There are similar methods explored in the AR community for learning both guitar and piano [10, 26, 24, 9, 16, 20]. These methods commonly utilize computer-vision to view the neck of a guitar or the keys on a piano to locate the corresponding location in 3D space. These applications can also be supported by colored stickers or fiducial markers to help locate essential objects such as the hand or parts of the instrument. Aside from a few examples where improvisation is supported using an AR-generated band [10], these applications are used to visualize where fingers are to be placed on the instrument, or generally focused on educational applications.

Other examples of scope limiting features include Bryan-Kinns’ daisyphone and LeMo, sequencer based collaboration systems designed to afford novices the ability to remotely jam collaborate on a musical score [4, 18]. The sequencer-based musical artifacts are used to provide players with the ability to plan note placements and rearrange compositions based on preference [14]. In this way novice players can collaborate on a musical score remotely due to non real-time constraints and future planning features of the application.

2.4 Networking Musical Information

Networked musical experiences enable players to play music with each other from anywhere in the world. Networked musical experiences have traditionally been encumbered by slow network speeds and lack of visual content, however, in recent years improvements in network speeds have made the ability to host an improvisational musical activity possible. This has sparked a line of research that explores the different types of networked experiences that can be had for musical improvisation.

Strategies that enable players to capture musical information and send it over a network have to contend with musicians who are affected negatively by processing delays over 11-15 milliseconds expert musicians [22]. This limiting factor precludes many design approaches that would otherwise be possible in real-time. Therefore strategies for combating latency while maintaining a fun jamming experience are critical to their success. Capturing human-human interaction data for these systems is a major concern as well, because the lack of human connection in these systems can have negative effects on the collaboration and expressive capabilities of the collaborators. Numerous techniques have

been used to try to get around the latency issue. These can be categorized into three groups: instantaneous networks, controlled delay systems, and non real-time collaboration strategies (see Figure 2).

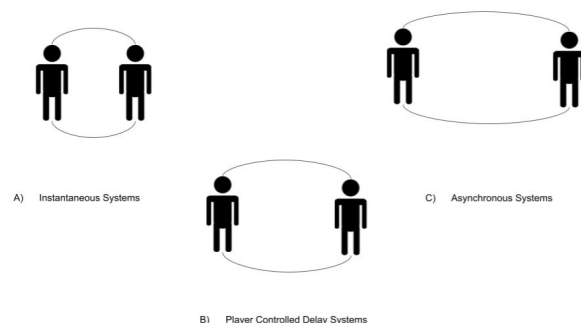


Figure 2: Figure illustrates the three types of networked musical experiences. A) Shows instantaneous systems which largely combat latency issues, B) player-controlled delay systems where music is delayed by a predetermined interval, and C) asynchronous systems which incorporate musical action at different times.

Instantaneous networked musical collaboration systems rely on real-time musical performance and are thus limited by physical distance. Jamkazam [12], JackTrip [5], and Jamulus [13] achieve near instantaneous jamming by requiring players to use an external audio card, direct access to their modem, and an ethernet connection to their personal device. This configuration can support remote collaboration with musicians who are located within a certain distance but does not provide video or any additional real-time graphics [23, 12, 13]. However, these systems are highly sensitive to configuration, server location, network speeds of individual users, latency affected by distance. The AR Drum Circle application can be used in conjunction with one of the mentioned applications by simply using headphones in addition to a head mounted AR display or phone.

The threshold for latency tolerance in rhythmic activities was explored by Schuett, et al. [22], who report that latency over 30msec causes a “tempo dilation”, or the slowing of the tempo as the song is played. In contrast, Chafe, et al. began to observe a tempo decrease at as little as 5 msec of delay (see Figure 3) [6]. This effect is thought to be due to the player attempting to adjust to the latency, or match the heard sound with the sense of internal rhythm held by the player. Up to 70msec can be tolerated by some players, but report that they notice compensating for the delay. Chew et al, investigated these effects on two players playing in an ensemble. Collaborative performance in a setting that is less controlled yielded a higher threshold for the effects of latency. Chew demonstrated that 50ms was a tolerable amount of latency, though noticed, had little impact on the ability for the players to perform together [7]. These latencies are tolerable up to about 75msec, however, players begin to utilize coping strategies and require practice to overcome. This situation is not ideal for music improvisation because listening and responding to what others are playing is paramount to a successful jam session.

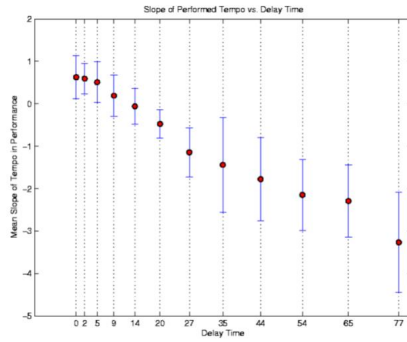


Figure 3: Figure illustrates the perception of delay in a synchronous clapping task. Players demonstrate noticing delay at roughly 5 seconds [6]

3. METHODS AND RESEARCH

In this section I describe my contributions to computer-supported jamming and how my process has evolved over time. I start by describing an example of representing musical information to aid in improvisation (Jam Tabs). I then describe a collaborative musical instrument for novices that senses players and algorithmically promotes collaborative improvisation (Jam Station). Then I describe a networked Augmented Reality application (AR Drum Circle) and an associated perceptual experiment to support its use (How Late is Too Late?).

3.1 Jam Tabs

While there are numerous methods for live scoring improvisation, there are not many resources aimed at assisting novice performance during improvisational activities. Jam Tabs is a notation system that helps coordinate idiomatic musical expression for novice musicians using color-coded instruments and a colored visual display. Jam Tabs utilizes a seven-color notation system, LED cubes and a colored piano keyboard to assist novices in coordinating chord progressions—a common musical element used among improvising musicians in popular western music contexts (see Figures 1, 4, and 5). We have observed that novice piano players could follow chord progressions in tonal music using color coordination while improvising. We found that seven colors in any key provide enough information for players to experience a satisfactorily expressive and creative jam session.

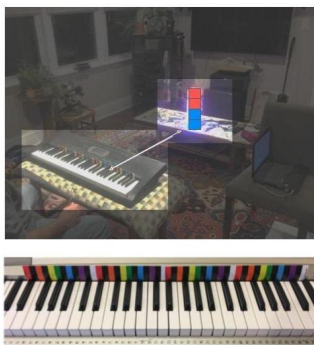


Figure 4: Figure illustrates the Jam Tabs system and the coordination between color indicated on piano and color on a reactive LED cube.

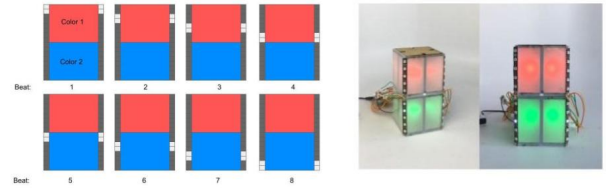


Figure 5: Figure illustrates the Jam Tabs LED cubes notation method. Colors indicate a chord in a key, such as I, ii, iii, IV, V, etc.

Jam Tabs was an investigation into the potential of alternative notations in music. The research question asked in this project was “how does mapping between instrument and notation affect players ability to jam?”. I sought to answer this question by creating an environment where players could jam to backing tracks that had 4 or less chords. The root notes (most prominent note in a chord) were displayed using colored cubes and players were asked to play back the chord progression before improvising—further evidence of the players ability to understand the notation. The players all had limited to no experience in jamming or with musical training, and could see all colors in the spectrum, enabling me to see whether they could utilize the system to have a successful jam.

I varied the music that I showed to the participants and the key (notes associated with the contents of the song) of each song. Some of the keys were more difficult for novices than others (ie. they had notes that required more complex motion to play) and players retained their ability, regardless of key to play notes that they thought sounded “good”. The players then gave feedback about their experience and whether they felt the system helped them attain their goal of jamming. It was found that players benefited from the Jam Tabs system and were able to jam successfully in any key to any song. Though satisfactory jamming was a completely subjective measure, I discovered that players enjoyed having a simple musical notation that mapped onto their instrument (in the case of the piano). For this reason I decided to investigate the use of color, and simpler notation schematics for musical improvisation as a means of notating music while jamming in more of my PhD research.

3.2 The Jam Station

Building on the success I had with color coordination for jamming, the Jam Station continues to utilize color coordination while also testing aspects of communication and collaboration. The Jam Station is a collaborative musical experience, with embedded sensors to detect, evaluate, and reinforce collaboration among players of all skill levels. The system is also gamified, consisting of four instruments that face a centrally located, vertically oriented, display that visualizes various game states. Playing any of the instruments affects the game as musical collaboration is assessed and visualized on the display. The display also includes a “progress bar” which fills up the more players collaborate musically, and to “win” the game, the progress bar must be full, initiating a multi-color light show for the players (see Figure 6 and 7). By creating an environment focused

primarily on collaboration, we aim to support collaborative music-making for musicians of all skill levels.

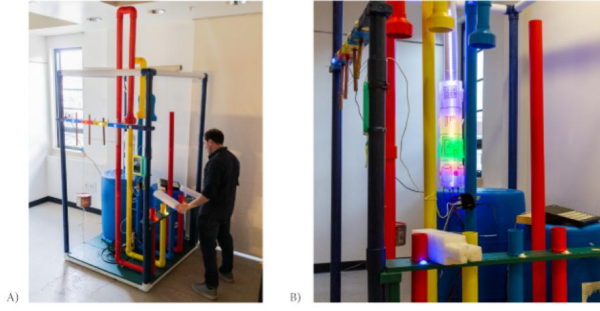


Figure 6: A) Shows the Jam Station from above being played by one person. B) A colored visualization is centrally located between the four players, easily showing the game state.



Figure 7: The Jam Station visualization shows the various game states. The leftmost image shows the instrument being struck. The center image shows the white incrementing lights that demonstrate collaboration is growing and the rightmost image shows the win state of the game.

In effort to explore sensing and processing musical information the Jam Station aims to explore musical evaluation algorithms and visualizations as an aid for musical collaboration among players of all skill levels. We accomplished this by building a collaborative musical experience based on design parameters outlined by Blaine and Fels for collaborative musical experiences, the use of visualization and limited musical range [2]. Through the design process, we hoped to accomplish a sufficiently complex, but not distracting experience for use with novices, as well as expert musicians. In addition to this, we aimed to enhance the inclusion of novices by utilizing collaboration as a feedback mechanism for gamified music interaction.

3.2.1 Visualizing Progress Bar/Collaboration Algorithm Display

A large section of the visualization is dedicated to a series of rising lights that increment towards the top of the progress bar. The rising lights move upwards as the players collaborate (see Figure 7). The collaboration algorithms reward the players for collaborating by moving the lights upward until the top of the progress bar. If collaboration is not detected, the white lights in this section begin to fall back down to the base of the visualizer. This method provided

a simple and detectable way to gauge collaboration in real-time. Participants quickly realize that they are working together to continue moving the lights toward the top of the visualizer, which initiates the “win state” and subsequently produces musical collaboration as a byproduct.

3.2.2 Win State

The top of the display is composed of white plastic and when the white lights that incrementally move upwards based on level of collaboration reach this section a win state is triggered. When this occurs the entire visualization turns into a rainbow of color (see Figure 7). Collaboration is affected positively by the win state because it provides a goal for the players to work towards. Players then can focus on musical collaboration between members of the group as a means to win the game. Winning the game does not end the jam, but promotes further collaboration. If players stop collaborating the rainbow display will go away and the progress bar empties. Because there is no clear ending or limit to musical possibilities beyond note choice this feature promotes open-endedness in the Jam Station design.

3.2.3 Musical Range

Instruments contained within the Jam Station were restricted to three musical notes (C, E, and G). These three notes form what is known as a major triad in music theory and sound harmonious when played simultaneously. This was implemented in order to simplify harmonic expression and reduce the likelihood of dissonance. Restricting the number of musical notes also contributed to a simplification of the collaboration algorithms and subsequently a less cluttered visualization.

3.2.4 Visualizing Notes Played on Each Instrument

The first component of the display show icons of each instrument and visualizes the note being played by the instrument in real-time (See Figure 7). This information is displayed in the form of a color that represents a note that is shared among all of the instruments (red, yellow, and blue). Each instrument icon is dedicated to a specific instrument, so if one player struck the “yellow note”, a yellow color would appear in the segment of the visualization that is designated for that instrument (see Figure 7).

The Jam Station focuses on several critical aspects of music improvisation: collaboration, free-play, and communication. The physical layout, gamification components, and the algorithms associated with measuring collaboration worked together to shift the focus of the activity toward collective action.

3.3 AR Drum Circle

AR Drum Circle, is an augmented reality (AR) platform we developed to facilitate a collaborative remote drumming experience (see Figure 8). We explore the effects of virtual avatars that are rendered in a player’s view to provide the joy and sensation of co-present music creation. AR Drum Circle uses a head mounted AR display, providing players with visual effects to assist drummers in coordinating musical ideas while simultaneously using a latency-optimized remote collaboration service (JackTrip). AR Drum Circle helps overcome barriers in music collaboration introduced when using remote collaboration tools on their own, which typically do not support real-time video, lack spatial information, and volume control between players, all of which

are important for in-person drum circles and music collaboration.

I present the results of two investigations: (1) observations and analysis of 20 videos of live drum circles, and (2) a collaborative drumming case study using a popular networked conferencing application (Zoom). Findings suggest that substantial communicative information including facial expressions, hand-tracking, and eye-contact is lost when using AR alone. Findings also indicate that the AR Drum Circle application can make players less self-conscious and more willing to participate while playing drums improvisationally with others.

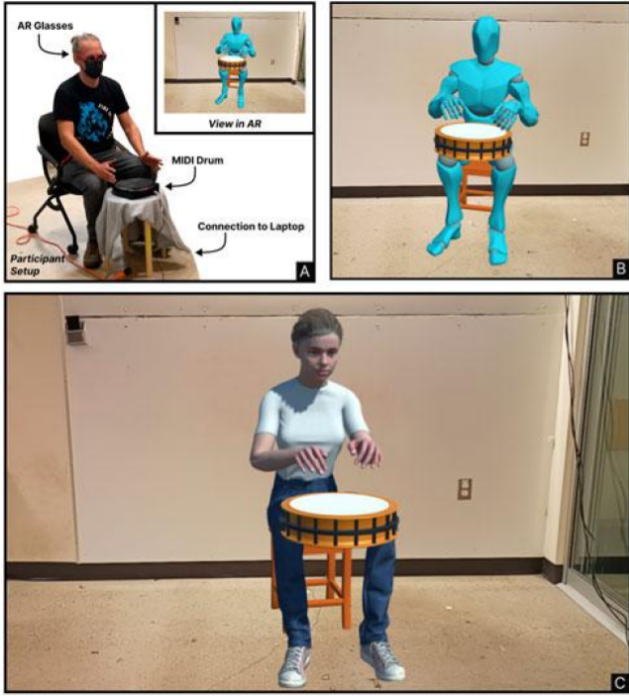


Figure 8: (A) Representation of the Player 1 setup viewing Player 2 in AR, (B) Player 2's avatar as seen from the perspective of Player 1, and (C) Player 1's avatar as seen from the perspective of 2 in AR.

3.3.1 Observations of Drum Circle Videos

To investigate how drummers communicate in a drum circle we captured and analyzed twenty videos of drum circles, ranging from 8–20 min. Participants ranged between two and twenty experienced (regularly participate in drum circles) drummers. Though many types of drums were used in the recorded drum circles, the Djembe (a traditional West African drum) was referenced for the types of drum strikes and rhythmic patterns. Our observations revealed crucial interaction dynamics, such as body language, eye-contact, gestures, and verbal communication that participants use in a drum circle. Using Anvil video annotation software [17] we marked and counted the occurrences of the various communication techniques (see Figure 9). We labeled and counted the instances when players looked at their own hands, or their drum, versus looking at the other player's face, hands, or drum, in a single session. We concluded that verbal communication, head motion, facial expression, and eye-contact play a key role in communication while jamming. These observations were used to inform the behavior of the representations of remote player avatars, by controlling their attention and mannerisms.

Communication technique	Number of occurrences
Eye Contact	31
Head motion for signaling	25
Head motion rhythmic	6
Head/Body Facing	3
Facial expression	13
Head motion rhythmic	6
Looking at own instrument/hands	29
Looking at another's instrument/hands	13

Figure 9: Summary of communication techniques.

3.3.2 Case Study

Building upon our analysis of interaction dynamics in drum circles, we explored how the strengths of AR could overcome the challenges of networked music collaboration tools. Our main design goals were to recreate and expand on some of the dynamics of in-person drum circles (physical, social cues). To address these goals, we developed the AR Drum Circle, an application that enables musicians to participate in a remote drum circle with their friends. Remote participants are rendered as avatars in the local player's field of view using the NReal Light AR glasses. To accommodate the demands of low latency communication in musical collaboration, body language is conveyed through compressed messages. Once received over the network, the compressed messages initiate pre-recorded motion of the avatar to minimize the effects of latency. Audio is transmitted at near-instantaneous speeds using an audio communication service called JackTrip [5]. This audio feed, in conjunction with the visuals rendered by the AR glasses, results in a collaborative experience with near-instantaneous audio transmission as well as a reactive AR avatar.

Gestural Communication

The AR drum circle utilized hours of video as well as motion capture to learn about drumming style and types of interactions between drummers. Studying film made it clear that drummers conveyed information about the music, their current state of enjoyment, and turn-taking with their use of gesture. Data from Goebel and Palmer [11] supports the idea that musicians convey tempo and musical information through collective body motion.

Because human to human visual communication is so important during a jam session we hypothesized that the presence of an avatar that replicates human motion would be useful to have in a networked music system. The AR Drum Circle builds upon this work by creating a mixed methods networked musical collaboration tool for variable latency conditions. The AR Drum Circle has the ability to perform in real time as well as asynchronously through looping, making it a robust and flexible system for remote musical collaboration.

Human Representations in XR

Using AR we aim to enhance co-presence in remote drumming collaborations, traditionally accomplished without visuals. Preliminary observations including video analysis, motion capture, and drummer interviews afforded insight into what features were the most necessary to include. We selected several animations that represented a natural idle drum state (head motion, eye-gaze, and foot tapping). We also used compressed music messages that come directly from the drum pad. Results yielded several key insights: 1) the AR Drum Circle application made players feel more comfortable playing with others because they became less

self-conscious, 2) visual issues are apparent in Zoom, because of limited camera angles, 3) latency appears to be more apparent in the AR Drum Circle app, but when latency is noticed in Zoom it tends to be more distracting.

Through experimentation we explore the differences in 3D spatial configurations in AR using an avatar with the 2-dimensional representations afforded by the commonly used networked conferencing application, Zoom. Results indicated that there are benefits to the 3-dimensional representations, as well as a sense of comfort that can be gained by limiting visual cues. This discovery was made possible through written and interview feedback, and would not have been so easily obtained through traditional quantitative methods. Future versions of the application that include multiple players and video in the AR experience will enable players to select desirable features that are attuned to their comfort level. They will also enable players to communicate amongst themselves spatially by more accurately representing hand gestures and head movements related to musical communication in 3-dimensions.

3.4 How Late is Too Late?

Augmented Reality (AR), with its ability to make people feel like they are in the same space as friends from across the world, is an ideal medium for the purpose of networked musical collaboration. Most conventional systems that enable networked musical collaboration minimize network latency by focusing on the transfer of auditory information at the expense of visual feedback. Studies into human perception have shown that sensory integration of audio and visual stimuli can take place even when there is a slight delay between the two signals. We studied the way changes in network latency affect participants' auditory and visual perception in latency detection, latency tolerance and attention focus; in this paper, we explore the trade-off between the presence of AR visuals and the minimization of latency.

Twenty-four participants were asked to play a hand drum and collaborate with a prerecorded remote musician rendered as an avatar in AR. Multiple trials involving different levels of audio-visual latency were conducted. We then analyzed the subjective responses of the participants together with the recorded musical information from each session. Results indicate a minimum noticeable delay value—defined as the highest amount of delay that can be experienced before two stimulated senses are perceived as separate events, between 160 milliseconds (ms) and 320 ms (see Figure 10). Players also reported that a delay between sound and an accompanying avatar animation became less tolerable at 320 ms of delay, but was never completely intolerable, even up to 1200 ms of delay. We conclude that players begin to notice delay at about 320 ms and most players can tolerate large delays between sound and animation.

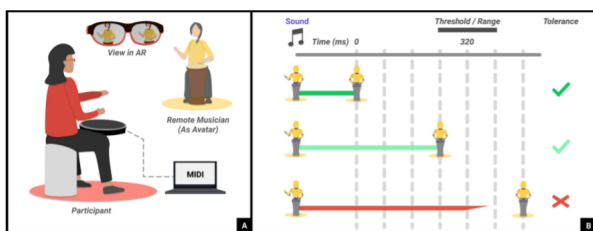


Figure 10: A) A representation of the experimental setup involving a participant collaborating with a remote musician using AR glasses and a MIDI drum, and B) is an illustration of the sound-animation latency threshold being studied.

4. EXPECTED OUTCOMES

At this point in my PhD I am seeking feedback from the NIME community, as well as offering my feedback for others in a similar stage of their development. I aim to achieve a deeper understanding of the interests in the community and explore opportunities for new directions in research. Musical improvisation, being my main focus, is a shared interest among community members. In the spirit of Open Jack, I aim to learn more about utilized technologies and paradigms in the NIME community for musical improvisation. I hope to expand my understanding of the relationship between music and technology, and learn from fellow computer musicians, designers, and practitioners in the field.

I also look forward to strengthening the community by providing any insight to students of all levels. I have been teaching music for 13 years and performing for most of my life. My expertise in music education, musical performance, and facilitating jam sessions may be of use to fellow students who would like to incorporate any insight I can provide.

The ability to contribute to and receive feedback from members of the community is invaluable to my research. I look forward to engaging constructively and improvisationally with fellow researchers if given the opportunity.

5. REFERENCES

- [1] I. Beck. Gtar: The first guitar that anybody can play. 2014.
- [2] T. Blaine and S. Fels. Collaborative musical experiences for novices. *Journal of New Music Research*, 32(4):411–428, 2003.
- [3] Bouchard and Browne. Live structures: Ocular scores.
- [4] N. Bryan-Kinns and P. G. Healey. Decay in collaborative music making. In *NIME*, pages 114–117, 2006.
- [5] J.-P. Cáceres and C. Chafe. Jacktrip: Under the hood of an engine for network audio. *Journal of New Music Research*, 39(3):183–187, 2010.
- [6] C. Chafe, M. Gurevich, G. Leslie, and S. Tyan. Effect of time delay on ensemble accuracy. In *Proceedings of the international symposium on musical acoustics*, volume 31, page 46. ISMA Nara, 2004.
- [7] E. Chew, R. Zimmermann, A. A. Sawchuk, C. Kyriakakis, C. Papadopoulos, A. François, G. Kim, A. Rizzo, and A. Volk. Musical interaction at a distance: Distributed immersive performance. In *Proceedings of the MusicNetwork Fourth Open Workshop on Integration of Music in Multimedia Applications*, pages 15–16. MusicNetwork Barcelona, 2004.
- [8] G. Davies. The effectiveness of lola (low latency) audiovisual streaming technology for distributed music practice. *M. Research diss., Edinburgh Napier University*, 2015.
- [9] M. S. Del Rio-Guerra, J. Martin-Gutierrez, V. A. Lopez-Chao, R. Flores Parra, and M. A. Ramirez Sosa. Ar graphic representation of musical notes for self-learning on guitar. *Applied Sciences*, 9(21):4527, 2019.
- [10] S. Glickman, B. Lee, F. Y. Hsiao, and S. Das. Music everywhere — augmented reality piano improvisation learning system. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 511–512, Copenhagen, Denmark, 2017. Aalborg University Copenhagen.
- [11] W. Goebel and C. Palmer. Synchronization of timing and motion among performing musicians. *Music*

Perception, 26(5):427–438, 2009.

- [12] Jamkazam. Live, in-sync music jamming over the internet. 2022.
- [13] Jamulus. 2022.
- [14] S. Jordà, M. Kaltenbrunner, G. Geiger, and R. Bencina. The reactable. In *ICMC*. Citeseer, 2005.
- [15] N. Klügel, M. R. Frieß, G. Groh, and F. Echtler. An approach to collaborative music composition. In *Nime*, pages 32–35, 2011.
- [16] M. Löchtefeld, S. Gehring, R. Jung, and A. Krüger. guitar: supporting guitar learning through mobile projection. In *CHI’11 Extended Abstracts on Human Factors in Computing Systems*, pages 1447–1452. 2011.
- [17] M. T. Maybury. Multimedia annotation, querying, and analysis in anvil. 2011.
- [18] L. Men and N. Bryan-Kinns. Lemo: exploring virtual space for collaborative creativity. In *Proceedings of the 2019 on Creativity and Cognition*, pages 71–82. 2019.
- [19] E. R. Miranda. *Handbook of artificial intelligence for music*. Springer, 2021.
- [20] Y. Motokawa and H. Saito. Support system for guitar playing using augmented reality display. In *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 243–244. IEEE, 2006.
- [21] L. Raymaekers, J. Vermeulen, K. Luyten, and K. Coninx. Game of tones: Learning to play songs on a piano using projected instructions and games. In *CHI’14 Extended Abstracts on Human Factors in Computing Systems*, pages 411–414. 2014.
- [22] N. Schuett. The effects of latency on ensemble performance. *Bachelor Thesis, CCRMA Department of Music, Stanford University*, 2002.
- [23] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis. 5g-enabled tactile internet. *IEEE Journal on selected areas in communications*, 34(3):460–473, 2016.
- [24] A. J. Stanbury, I. Said, and H. J. Kang. Holokeys: Interactive piano education using augmented reality and iot. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, pages 1–3, 2021.
- [25] N. Statt. Google’s aiduet lets you make music with a virtual pianist. 2017.
- [26] F. Trujano, M. Khan, and P. Maes. Arpiano efficient music learning using augmented reality. In *Innovative Technologies and Learning: First International Conference, ICITL 2018, Portoroz, Slovenia, August 27–30, 2018, Proceedings 1*, pages 3–17. Springer, 2018.
- [27] G. Weinberg and S. Driscoll. The design of a robotic marimba player: introducing pitch into robotic musicianship. In *Proceedings of the 7th international conference on New interfaces for musical expression*, pages 228–233, 2007.
- [28] L. Yousician. Guitartuna: Chords, tuner, songs. 2023.