LINEAR: A MULTI-DEVICE AUGMENTED REALITY ENVIRONMENT FOR INTERACTIVE NOTATION AND MUSIC IMPROVISATION

Giovanni Santini
Hong Kong Baptist University
Hong Kong
giovesan@hkbu.edu.hk

ABSTRACT

LINEAR (Live-generated Interface and Notation Environment in Augmented Reality) is an environment for the generation of real-time 3D interactive graphic notation. The environment is suitable for ensemble improvisative performances featuring acoustic instruments, live-electronics and two Augmented Reality (AR) performers. One AR performer uses an iPhone for drawing virtual trajectories in the space, rendered as a sequence of Virtual Objects (VOs) aligned along the trajectory. VOs trigger samples upon virtual collisions with the iPhone. They are also used as a form of graphic notation for instrumentalists/vocalists: the screen of the iPhone is mirrored to a projector. The second AR performer uses a headset and can use VR controllers to design trajectories used for the spatialization of each audio source in a 3D audio setup. The headset AR performer can use virtual spheres (one per instrument) to control the position of each sound source (one per instrument). The sound of every acoustic instrument is processed live. The mixing of processing effects are controlled by a laptop player. The system has been repeatedly tested during a two-semesters long workshop. The system was also used for two online concerts. Beyond demonstrating the technical and musical viability of LINEAR, the workshop also gave the chance to record student's response to the system. Although the sample size is quite small (four students completed the survey), the answers show encouraging results in terms of engagement and interest. Future work should be conducted to further enhance the user experience and more clearly assess LINEAR's usability and effectiveness as an innovative system for improvisation and musical performance.

1. INTRODUCTION

A new wave of interest for immersive technologies is recently arising, with the hype related to the *Metaverse* and the attempts to actualise it. The Metaverse is a concept about a virtual, distributed, interoperable world that can run in parallel with the real world and includes complex features (from human interaction to trade). It is typically associated with technologies like blockchain and machine learning, and even more strongly with Extended Realities (XR) technologies, i.e. Virtual and Augmented Reality (VR/AR). VR places the user in a completely virtual world, with Virtual Objects (VOs). The only relation to the physical world are the movements of the user detected via motion capture. AR merges virtual objects with the real world and allows interaction with them, possibly enriching the experience with spatial understanding capabilities: the ability to analyze the physical connotation of the surrounding space and have VOs behave accordingly (e.g., by rendering the shadows on the floor). XR applications could become increasingly important for arts making, in term of new ways of expressions and new market segments.

The main interest for the author is to find some creative space opened up by such technological developments. The hybrid reality created by the combination of virtual and physical objects, actors, and environments can reveal uncharted territories for exploration and creation. In designing LINEAR, the aim was to create a complex system, enabling numerous ways of interaction between performers, musical materials and the virtual world. The author wanted to preserve an exploratory attitude and some flexibility of use, without tying the whole design to one single expected outcome. The challenge consisted in doing so while embedding capabilities that are exclusive of AR technology, seen as a new medium for musical expression.

2. BACKGROUND

In the past few years, there has been some research related to musical notation in AR. Mostly, researches exploited the use of the temporal dimension (one of the typical traits of AR) [1, 2]. In some researches, the 3-dimensional spatial nature of AR was also exploited [3, 4, 5]. In most of the cases, researchers adopted graphic notation solutions, as opposed to the traditional descriptive notation. Graphic notation is an umbrella term that refers to numerous different contexts and aesthetics. We could define it as that form of notation that uses graphical solutions that are not part of the Common Western Notation (CWN) lexicon. Graphics can be used either in addition to traditional notation or replace it. An example of a mixed use of CWN and graphic notation could be found in action scores, such as Lachenmann's *Pression* (1969), while an example of a purely graphic scores is Haubenstock-Ramatiì's *Konstellationen* (1971). Graphic notation has also been used in recent technology-based solutions such as real-time scores for animated notation [6], 3D scores [7] and VR scores [8].

Graphic notation has been widely adopted in Augmented

Reality applications for music education, music composition and music performance.

2.1 AR notation in music education

In music education, AR notation seems to be generally conceived as a subsidiary tool to assist traditional music learning. Typically, it replaces (or aids) the descriptive notation of traditional scores with a prescriptive 3D interactive notation that indicates hands or fingers positions requested at a certain time, e.g. [9, 10]. The notational solutions (although not resembling themselves the gestural behavior of the performer) have a clear connection with the physical, spatial displacement of performance actions, rather than with the expected result. An example of this principle is the piano roll. With small variations across different studies, a piano roll is a system that makes use of virtual colored blocks (coming towards the keyboard) to indicate the keys to press at a specific time [11, 12, 13, 14] and Figure 1. Additional indications for dynamics or wrong notes can be delivered through the use of different colors or User Interface (UI) elements.

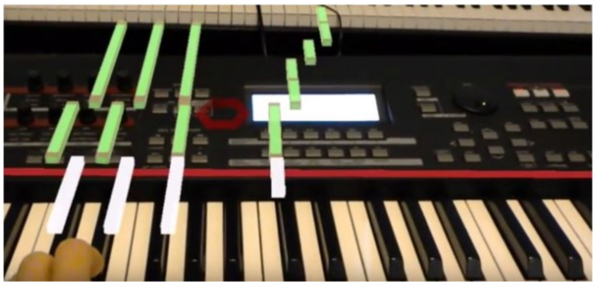

Figure 1. One example of piano roll [15].

2.2 AR for music performance or composition

A different case can be made for AR applications that are not designed for educational purposes but rather as a tool for artistic use (performance or composition or both). Relatively few experiments have been conducted so far, presumably due to the technical challenges and the absence of an already established corpus of background work to help researchers. We can currently identify three directions (not necessarily mutually exclusive):

- AR to solve practical performance/rehearsal issue (performers can visualize a 2D score anywhere without a physical score).
- Immersive scores.
- Interactive notation.

An example of 2D notation in Augmented Reality has been developed in order to facilitate discussion and feedback among musicians. The score (bi-dimensional, written in CWN) is visualized in AR and can be moved at different distances and locked to the point of view of the user(s). This allows for a great flexibility during rehearsal time and favours discussion among musicians, as long as they all have a headset [2].

Smartvox [1] is a browser-based network-distributed environment for synchronized real-time scores. In its recent developments, its use was extended to networked head-mounted displays with Augmented Reality capabilities. This way, a performer can visualize a 2D (animated) score as a semitransparent layer superimposed to the environment. This solution allows for a much freer and comfortable user experience (especially in choir situations) both during rehearsal and performance. This is especially true for animated notation which requires a constant attention from the performer.

In the opinion of the author, the two examples above do not actually represent a case of Augmented Reality notation, but rather a translation of 2D notation in an AR context in order to solve a comfortability issue. Although some interactivity and some use of the time dimension can be identified, these examples do not develop notation using the three spatial dimensions. Essentially, they allow the visualization of a screen with very flexible spatial capabilities. In other terms, the page is in AR but not the score.

In [3], David Kim-Boyle outlines some properties of graphic notation in an AR context. The notation gains an architectural dimension, which becomes essential for the interpretation of the score. Graphic and spatial connotations are then fully exploited in an immersive score.

"In *64x4x4* [...] the physical engagement with the score becomes an essential means of uncovering its various potentialities. The pathways through the score, uniquely instantiated for each performance, may only be discovered when the performers physically navigate the space in which the score is displayed" [3].

Amy Brandon's work has also been widely developed around the use of AR notation. In her works, notation also assumes the function of interface: interactable elements of the graphic notation are also used to trigger samples (*Hidden Motive*, 2018, *Augmented Percussion*, 2019):

"I could add functionalities to the graphic score - make it interactive or animate it. In performance terms, the musician would be able to grab elements of the score, and would be able to trigger audio files in the process" [16]. This combination has been defined notation-interface hybrid [17].

[18] presents a similar concept but more oriented towards sound generation than notation. Pitch values are generated according to the movement in space and then passed to a synthesizer. The sound is panned in 3D according to position in space.

GesturAR [4] is an experimental application that allows to notate performance gestures in the real space. A hand tracking device is used to detect the palm position. When the *record* mode is activated, at each frame, the positions are stored in a trajectory and a line is rendered according to each point of the trajectory. The trajectories can be stored, combined and played back. The coordinates are in real world and associated to an origin point provided by a tracking device which can be positioned on the instrument to be played. The resulting notation has been called embodied interactive notation: the act of notation coincides with the notated act.

The panorama presents a lively, yet relatively small amount of research focusing on AR notation and its connection with VOs, space and physical world. However, the continuous improvement of development frameworks and the increasing generalized interest in XRs seems to slowly foster the growth of the field.

3. DESCRIPTION OF LINEAR

LINEAR is an environment for real-time music improvisation, without a fixed number of performers. The bare minimum number of performers is 4: one iPhone performer, one AR headset performer, one laptop player and one instrumentalist/vocalist. Concerts were performed with a total of 7 performers (4 acoustic performers). The system has been designed in order to be swiftly adapted to a different number of performers and instruments. Each instrument player is required to be miked with at least one microphone for real-time sound processing.

The environment is composed of two different applications, one for iPhone, the other one for an Augmented Reality headset. The two applications are independent from each other but connected to an Ableton Live/M4L project via OSC protocol. The applications have been developed with the framework Unity 3D. The HTC Vive Pro with a ZED Mini VR camera has been used for the headset app.

3.1 The iPhone app

The iPhone app allows the performer to draw trajectories in space by using physical gestures while holding the device. Such trajectories are visualized as a sequence of Virtual Objects (VOs) placed along the the device' movement. By using three different buttons, the iPhone performer can choose between three different types of VOs associated to three different colors and particle effects (graphic effects composed by up to millions of instances of a same fundamental object, a particle). Such effects have a different "energy" and size according to the different speeds of the user's gesture. Various parameters of the particle effects regulate those levels of energy: number, speed, life time and size of each particle. The speed of the device in the moment of the creation of a VO (averaged over 5 frames) is mapped to those parameters in order to deliver different tiers of "excitement".

The visual effects are rendered on the screen of the device, altogether with the real environment. Each VO is connected to a sample stored in a sample library loaded in the Ableton Live project. In the moment of the creation of the VO, the sample is played. It is also played anytime the performer moves the device onto the position where the VO is instantiated.

The screen of the iPhone is mirrored to a projector. Therefore, the instrument/vocal players are able to read the graphic notation. The precise way they are asked to do so is explained later in the article. In order to provide orchestration and behavioral constraints, the screen of the iPhone is divided in as many parts as there are instrumental/vocal performers. In Figure 2, the screen is divided in four parts as there are three instruments and one singer. The names are indicated around the center of the screen, one for each quadrant.

3.2 The AR headset app

The application for AR headset is also based on the creation of virtual trajectories associated to different VOs, each of them corresponding to one of the performers' audio channels (including the iPhone performer). Each object has a different color which is inherited by the corresponding trajectory. The app allows four functions: *select*, *draw*, *play* and *play all*, connected to different inputs on the VR controllers. The *draw* function allows one to draw trajectories in space, visualized as continuous lines. Those trajectories correspond to points of coordinates in space that are communicated to the Ableton Live project via OSC to control a sound spatialization module. The coordinates of those trajectories correspond to the positions of virtual sound sources (one per trajectory). The *select* function enables the choice of different VOs, each of them controlling the coordinate in space of a different sound source: the sound processing (and sometimes amplification) of acoustic instruments and the iPhone player's samples. When a VO is selected, only its trajectory is played (therefore, only the linked sound source is moved in space). The *play* function plays back the trajectory already created for the selected VO. *Play all* plays all trajectories together.

The point of view of the headset performer is mirrored to a projector. The other performers are not asked to "read" the trajectories the same way they do with the iPhone screen. However, when a single VO is selected, the name of the corresponding instrument is shown on the screen and becomes an orchestration indication: the selected object/instrument is a soloist and therefore the rest of the ensemble should adjust their dynamics in order to let the soloist be in the foreground.

Figure 2. A view of the two screens in a concert setting.

3.3 The Ableton Live project

The Abletone live project includes processing modules with an effect chain composed of dynamic EQ, spectral delay, octaver, distortion and a multi-buffer granulator. All the effects are custom-made and created with Max4Live. Each processing chain is conveyed onto a mono signal bus routed to the spatialization module, based on IRCAM's *Spat* (the *spat* track in Figure 3), which positions the different sound sources in a 3D audio panorama. This effect is controlled via OSC by the AR headset app. Although the amount of processing on each mono signal would sometime require more channels for a better result, the author preferred to avoid moving stereo signals in order to better guarantee spatial separation between sound sources. The sound spatialization has been tested in rehearsals and concerts in two different spaces, respectively equipped with a 3D 24.2 setup and with a 7.1 setup. The algorithm used was 3D Vector-based Amplitude Panning (VBAP3D).

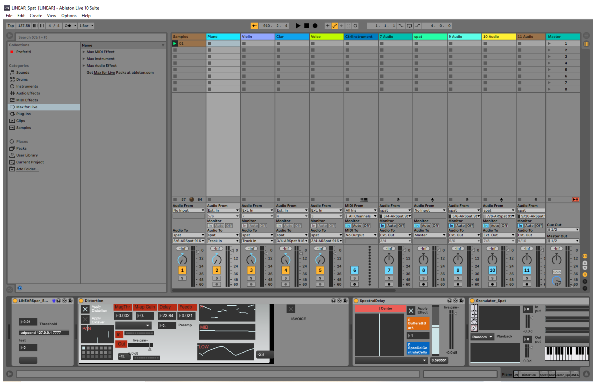

Figure 3. The Ableton Live project for LINEAR.

On each channel, each effect can be separately bypassed and balanced, according to the needs of the different instruments, microphones etc. Presets are available to instantly control numerous combinations of effects and loudness on each channel: e.g. full chain applied to clarinet and violin, only granulator to the voice, only amplification to the piano. Presets and loudness balance are controlled by the laptop player.

4. SCORES AND PERFORMANCE NOTES

The trajectories drawn in space by the iPhone performer form an embodied interactive score [1]: the trajectories represent spatial points of reference, precisely linked to sounds, and are derived from bodily actions. Repeating those actions will reproduce the same sounds. The score is a consequence of the act it is meant to notate.

The mirrored iPhone screen is also read as a graphic score by the non-AR performers. In the example, the screen is divided in four parts, one for each instrumental/vocal performer, who are asked to read only their quadrant and to not produce sound when there is no VO in their slot.

4.1 How to read the graphic score

Each one of the three VO types is connected to a different sample library. The generic sound quality of the samples could be labelled as:

- Long, voice-like low pitches.
- High, articulated, fast attack pitches.
- Noise, breath sound quality, either long or short attack.

The non-AR performers are asked to imitate those sounds by using extended techniques on their own instrument. For example, a violin could imitate as follows:

- 4th string detuned one fifth lower, with continuous movement of the bow between bridge and fingerboard, and alternating the pressure of the right hand while performing microtonal glissandos with the left hand, at most one tone above the open string.
- Fast pizzicatos with very high fingering (close to the bridge) on 1st and 2nd string.
- Muted strings with left-right and up-down bow movements, either slow or fast and either short or long.

The set of techniques required is bigger. Typically, at least 3 techniques for each sound are required to each performer.

Performers need to read left-to-write and imitate the spatial disposition of VOs (e.g., few objects separated = sound alternated with silence, high spatial density = no silence). Ideally, their quadrant should be read as a 5 second loop. Also the “energy” of the effect needs to be taken into account (e.g., steady effect = sustained sound with flat profile, highly cynetic effect = fast articulation on a moving profile).

4.2 Solo instrument information

The point of view of the AR headset is mirrored to a projector. The result is not read as a score by the musicians, but it indicates the name of the instrument corresponding to the virtual source moved in that moment (e.g., voice in Figure 2). The instrument indicated is meant to be treated as a soloist, while the other musicians (if playing) are required to stay in the background. When *play all* is activated, all players should consider themselves as soloists, while when no sound source is selected, all of them should think about staying in the background. These rules might create absurd combinations: sometimes, an instrument could be selected as solo, but no VO is shown on that instrument’s quadrant, and therefore the instrument is not allowed to play. In that case, the other players should stay in the background of an instrument that is not playing. Future iterations of the environment might solve this issue. However, this paradox can also stimulate the seek for creative solutions: how can sounds be in the background of silence?

4.3 How to structure a performance

The iPhone player is the conductor/real-time composer. They create the real-time score that needs to be read by the other performers and also decide the point of view on the VOs, therefore what each perfomer sees in their quadrant. To some extent, the iPhone player decides what the performers will or will not do. The challenge is to create an overall development that delivers some structural interest

[1] “The notation is created as a direct consequence of an embodied act (detected through sensors) and is a 4D representation in space and time of the original gesture, in the form of a trajectory or some other kind of spatial marking” [17].

over time. As a reference for the development of the performance, the author created a performance outline (Figure 4), subdivided in "bars" of approximately 10 seconds. The colored lines indicate the different types of VOs. Rectilinear lines notate a low level of activity, while broken ones do the opposite. This indication translates into the speed of the gesture that creates the VOs. Dynamics are used to indicate the density (FF = screen filled with VOs, PP = one or two Vos in the screen). Indications are also used for audio processing effects (G = granulator, SPD = spectral delay, OKT = octavter, DIST = distortion). The performance outline does not provide other details, such as which instruments should play, which gesture should be used to create VOs and what the point of view should be. All these details, impacting the real-time graphic score, are left to the sensitivity and reactivity of the iPhone player.

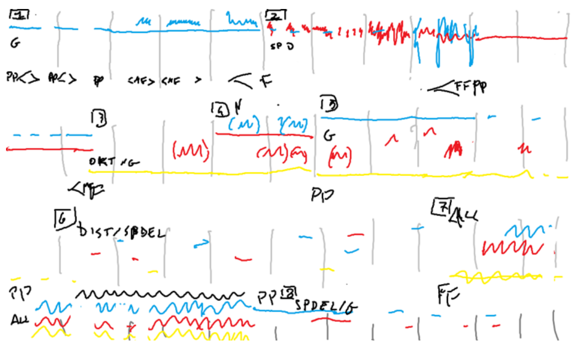

Figure 4. The performance outline for one concert. Different colors indicate different types of VOs, therefore different timbral references for the instruments (and different sets of extended techniques). The shapes of the lines indicate the level of excitement and dynamics describe density.

5. EVALUATION AND DISCUSSION

The current version of LINEAR has been developed through a year of continuous experimentation in a New Music class. Two concerts were made, one mid-term, the other one at the end of the academic year. A small evaluation has been conducted. Out of the 6 students taking part to the final concert, 4 replied to the form. The aim of this informal evaluation was essentially to obtain some feedback and reactions about the notational concept as well as the perceived effectiveness of LINEAR as an innovative tool for musical performance. In order to do so, participants were asked to fill an online form divided in two parts. In the first part, five questions with a 1-5 Likert scale were asked (see Table 1). In the second part, questions with free answers were asked. The forms were submitted through an online anonymous document.

Four free questions were made in the second part of the survey.

- What do you think about the use of notation in LINEAR?
- How would you compare an improvisation using LINEAR with other ways of improvising?
- Use three adjectives to describe your experience with LINEAR.
- Any additional thought/feedback?

The evaluation shows a certain appreciation of LINEAR. In particular, participants seemed to find the visuals engaging and the system quite innovative. Interesting suggestions were made on how to improve the effectiveness and interactivity of the notatio. For example, one response to the first question was: "A set of clearly defined and notated gestures would have been useful; sort of like a taxonomy of possible sounds and their relationship to the virtual objects". Although such relationships were actually presented during the rehearsals, the use of the system would probably benefit from a very structured detailing of playing techniques. While the author approached it quite experimentally (hands-on trial), future iterations should probably include clear indications established in advance.

Strongly disagree (1), Strongly agree (5)	**1**	**2**	**3**	**4**	**5**
The use of spatial movement of the iPhone player is inspiring/effective	-	-	1	1	2
The use of notation helps to find interesting/new solutions	-	-	2	1	1
3D virtual objects contributed to make the experience more engaging	-	-	1	-	3
The sound spatialization is clear/well-working	-	-	1	1	2
You would want to work more with LINEAR	-	-	1	1	2

Table 1. Likert-scale questions and responses.

Another comment stated: "Some special effects/sound can be created when certain movement of phone/VR appeared (e.g. draw a circle and some wind sound is responded)". Including machine learning and visual recognition algorithms to increase the interactivity of the system is a solution that might be experimented in future iterations. This would allow to add pre-composed, fixed gestural material to the performance. The multi-sensory nature of LINEAR was also pointed out, in one of the reponses to the second open quesion of the form: "Improvisation using LINEAR evokes multiple-sense of reactions visually and auditory". The environment is found innovative: in response to the free question number 3, 2 adjectives were repeated twice: *innovative* and *futuristic*. Also, one free comment: "It seems to me that this type of work and technology is very promising, looking forward to future developments".

The result of the performance needs to be considered as part of the evaluation too. The author finds the outcome promising, while challenges and needs of improvement should be pointed out. The overall system has some expressive capabilities, and the reactivity of performers to different graphic changes is sometimes convincing. However, there is some lack of precision in the relationship between score and sound, for what concerns VO-extended technique, enegy-articulation, solo-tutti relationships. The continuous unpredictability of the score is certainly a reason. One of the rehearsals strategies was to practice on steady points of view, rehearsing the techniques until the sound result was adhering to the picture. Part of the reason of the lack of precision could maybe be found in the long time span between rehearsals (typically one month). However, the environment itself can result hard to decipher, especially for the distribution of information across two screens. The ecosystem of mutually listening and reacting performers, with a human-controlled real-time score is fascinating and promising. Future adjustments need to be made in order to render more clearly the different information, thus increasing the efficacy of the notation and the possibility of control.

6. CONCLUSIONS AND FUTURE WORK

LINEAR is designed to be an articulated performance environment that requires continuous practice and exploration to be perfected, similarly to any musical instrument. It is composed of two AR applications for different devices (iPhone and AR headset) and one AbletoneLive/M4L sessions. The application for iPhone is responsible for playing sounds (via Abletone Live) from virtual sources and generating a live AR graphic score. The AR headset app is used to 3D pan virtual sources by controlling an effect based on IRCAM's *Spat*. The live-electronics processing is controlled by a laptop player. For non-AR performers, there are two sources of information to read from. Mainly, the mirrored iPhone screen (divided in four quadrants), containing the AR graphic score. Performance information is contained in a different quadrant for each performer. VO type (color), energy and density are linked to expected performance outcome: extended techniques used, articulation, dynamics and/or density of the sound texture. They also need to read the mirrored AR headset screen to know which instrument is soloist in that moment. For the iPhone player, the performance outline is the primary source of information while creating the live AR score. The AR score is also a form of embodied interactive notation to follow. The headset performer does not need to follow any score or indication. The laptop player follows the lead of the iPhone player in activating the presets.

The environment presents an initial learning curve, as it includes different groups of performers which follow different sets of indication (the non-AR performers, the iPhone performer, the AR headset performer). Additionally, the performance heavily relies on the performers' knowledge of extended techniques. However, LINEAR proved to be viable for rehearsal and performance after a training process.

Future improvements of the system shall include bigger sample libraries with enhanced processing capabilities. Using concatenative synthesis could be a solution to create more lively and differentiated sound results. A more structured definition of performance techniques for each instrument should be considered. Currently, the two AR applications run separately and just talk to Ableton Live. Future enhancements shall allow interoperability between the systems and have VOs on one device impacting rendering on the other device. The clarity of on-screen indications could be enhanced, for example by finding a way to condense all the information needed by the interpreters on one screen. It would also be interesting to find solutions for replacing the laptop player. For example, commands to start a new preset could be designed for the iPhone or for the AR headset app UI.

Recording of one concert

The file is a recording over zoom (the concert was in online format) and mostly the audio is heavily clipped.

Acknowledgments

The developments described in this article have been possible thanks to the collaboration with the *Contemporary Music Workshop* class, held by Dr. Camilo Mendez at Hong Kong Baptist University.

7. REFERENCES

[1] J. Bell and B. Carey, "Animated notation, score distribution and VR-AR environments for spectral mimetic transfer in music composition," in *Proceedings of the International Conference on Technologies for Music Notation and Representation*, Melbourne, Jul 2019, pp. 7 – 14.

[2] Z. Liu, M. Adcock, and H. Gardner, "An evaluation of augmented reality music notation," in *Proceedings - VRCAI 2019: 17th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*, 2019, pp. 1 – 2.

[3] D. Kim-Boyle, "3D notations and the immersive score," *Leonardo Music Journal*, pp. 39 – 41, 2019.

[4] G. Santini, "Composition as an embodied act: A framework for the gesture-based creation of augmented reality action scores," in *Proceedings of the Sound and Music Computing Conferences*, 2020, pp. 357 – 363.

[5] A. Simion, A. Iftene, and D. Gîfu, "An Augmented Reality Piano Learning Tool," in *International Conference on Human-Computer Interaction, RoCHI*, 2021, pp. 16 – 17.

[6] C. Hope and L. Vickery, "The DECIBEL Scoreplayer - A Digital Tool for Reading Graphic Notation," in *Proceedings of the International Conference on Technologies for Music Notation and Representation*, M. Press, Ed., 2015, pp. 58 – 69.

[7] J. Han, "An architectural spatial approach to 3d spatial drum notation," *Tenor 2017 International Conference on Technologies for Music Notation and Representation*, pp. 39 – 49, 2017.

[8] R. Hamilton, "Trois machins de la grâce aimante: A virtual reality string quartet," in *Proceedings of the 2019 International Computer Music Conference,*

ICMC-NYCEMF 2019 - International Computer Music Conference New York City Electroacoustic Music Festival, 2019.

[9] J. Martin-Gutierrez, M. S. Del Rio Guerra, V. Lopez-Chao, R. H. S. Gastelum, and J. F. V. Bojórquez, "Augmented reality to facilitate learning of the acoustic guitar," *Applied Sciences (Switzerland)*, pp. 1 – 7, 2020.

[10] Y. Zhang, S. Liu, L. Tao, C. Yu, Y. Shi, and Y. Xu, "ChinAR: Facilitating Chinese Guqin learning through interactive projected augmentation," in *ACM Interna-tional Conference Proceeding Series*, 2015, pp. 23 –32.

[11] M. Cai, M. A. Amrizal, T. Abe, and T. Suganuma, "Design of an ar-based system for group piano learning," in *Adjunct Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2019*, 2019, pp. 20 – 21.

[12] D. Hackl and C. Anthes, "HoloKeys - An augmented reality application for learning the piano," in *CEUR Workshop Proceedings*, 2017, pp. 140 – 144.

[13] W. Molloy, E. Huang, and B. C. Wunsche, "Mixed reality piano tutor: A gamified piano practice environment," in *ICEIC 2019 - International Conference on Electronics, Information, and Communication*, 2019.

[14] L. Rigby, B. C. Wünsche, and A. Shaw, "PiARno-An Augmented Reality Piano Tutor," in *PervasiveHealth: Pervasive Computing Technologies for Healthcare*, 2020, pp. 481 – 491.

[15] E. Strasnick, A. Chambers, L. Jiang, and T. Xiaonan, "Pianolens: An Augmented Reality Piano Interface." [Online]. Available: https://www.youtube.com/watch?v=5TExa2L1rOM

[16] A. Brandon, "Hidden Motive: on the development of interactive graphic scores for AR and VR," *The Association of Canadian Women Composers (ACWC)*, pp. 19 – 28, 2018.

[17] G. Santini, "Augmented Piano in Augmented Reality," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2020, pp. 411– 415.

[18] K. Yoo and E. Schwelling, "Spatially accurate generative music with AR drawing," in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, 2019, pp. 1 – 2.