Decoupling Physical and Virtual Spaces for New Collaboration Strategies in Co-Located Mixed Reality Instruments

Pierrick Uro
Florent Berthaut
Thomas Pietrzak
pierrick.uro@univ-lille.fr
florent.berthaut@univ-lille.fr
thomas.pietrzak@univ-lille.fr
Université de Lille
France

Marcelo M. Wanderley
marcelo.wanderley@mail.mcgill.ca
McGill University
Canada

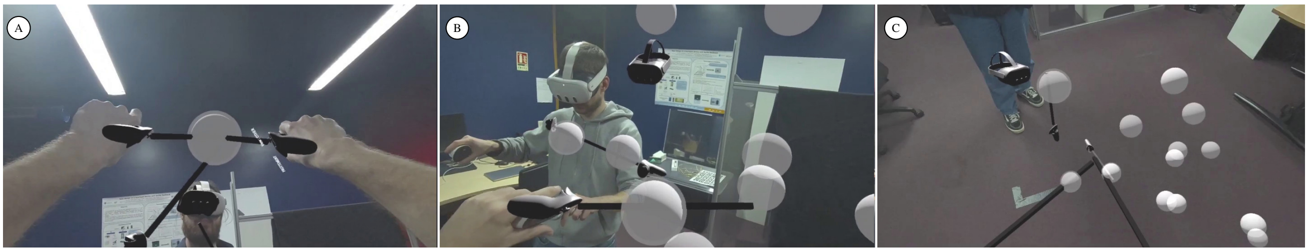

Figure 1: gRAinyCloud enables novel collaboration strategies in Mixed Reality instruments. A) : Two musicians play a virtual sound in a FIXED viewpoint. B) : One musician translated the virtual space to avoid physical conflicts with the other while playing the sounds, the virtual headset shows their actual viewpoint within the virtual space. C) : One musician zoomed out to be able to play fast sequences of sounds, the other musician's activity is therefore displayed as a small avatar within the virtual space.

Abstract

Collaborative co-located Mixed Reality musical instruments combine some of the expressive opportunities of 3D interaction and communication and cooperation of physical multi-user instruments. However in existing instruments, the fixed coupling between the virtual and physical environments constrains the affordances brought by Mixed Reality, such as per-musician free navigation in or multi-scale control of virtual structures.

We designed gRAinyCloud, as a way to reintegrate these lost affordances to a co-located instrument. It allows for the expressive exploration of a set of sounds represented by a virtual structure of shapes placed in the physical space and shared between musicians. Above all, gRAinyCloud enables each musician to freely manipulate their own viewpoint, changing its scale, position and rotation, effectively decoupling the physical and virtual spaces, and to switch between self, other's and absolute viewpoint while playing. We describe the implementation of this decoupling of spaces and analyse its uses and implications for collective musical expression, by relying on a first-person approach.

Keywords

Mixed reality, musical ensembles, decoupling, collaboration, co-located, multi-scale

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

1 Introduction

Collaboration has always been one of the main questions in research on new interfaces for musical expression [4], tackled through the design of novel multi-user instruments [14] or the study of the strategies and experience of musicians in ensembles [27]. Mixed Reality (MR) is defined here as the overlap of physical and virtual environments by Milgram and Kishino's virtuality continuum [19], by either augmenting the physical environment with virtual elements (Augmented Reality) or augmenting the virtual environment with physical elements (Augmented Virtuality). It opens further possibilities for musical collaboration on top of the trove of potential it presents for single-user instruments [2]. In a collaborative context, MR brings new opportunities for musical interaction [24], like the shared access to an instrument such as in Carillon [13] and ARLooper [21], or the dynamic control of privacy like in LeMo [18]. Still, this combination remains largely unexplored [5].

While MR fosters unique opportunities, it also comes with challenges, especially in the case of co-located collaborators. In co-located MR, *i.e.* when the virtual instrument is placed in a shared physical space amongst the musicians equipped with individual headsets or mobile devices, the most common implementation is to harmoniously align the virtual environments of all users with the physical environment [12, 18, 21]. This guarantees a consistency in position for all virtual objects, and a direct correspondence between any potential virtual representation of other users and their physical body. Although this approach provides collision prevention and is highly adequate for closely coupled collaboration, it also puts a tight restriction on what MR can enable for musical interaction. Additionally, since physical and virtual spaces are coupled, they become a resource to share, further restricting potential applications as conflict over space becomes a possibility.

An alternative is the decoupling of the physical and virtual spaces by allowing users to manipulate their viewpoint through navigation metaphors. Various metaphors have been proposed in the context of MR, such as teleportation, flying, or path-planning [2]. A "grab-the-air" metaphor was implemented for a distributed Virtual Reality collaborative instrument [1], but such metaphors have not yet been applied to co-located MR instruments. Still, a "human-joystick" metaphor was used for investigating this decoupling with navigation tasks in co-located MR [8]. In the context of visual MR, where the physical other is visible in addition to the virtual space, decoupling spaces involves only a single avatar in addition to the physical body of the other as a representation of their viewpoint in relation to the virtual space. Although this potential "dual-presence" may be a drawback for task performance [9], it should be less impactful in a creative context such as musical instruments. Overall, this approach enables countless collaborative scenarios that wouldn't be possible with tightly coupled spaces, and has the potential for new collaborative strategies.

In this paper, we present gRAinyCloud, a co-located MR musical instrument that implements this decoupling of spaces, and report on a self-study based on improvisation sessions investigating emerging strategies and guidelines for future applications of this approach for co-located musical instruments in MR.

2 Related work

This sections presents related work in collaborative MR instruments, highlighting the lack of navigation in co-located instruments, and presents strategies in non-musical contexts for addressing concurrent manipulation.

2.1 Collaborative MR Instruments

Collaborative instruments occupy the broad spectrum of MR and further in various forms [5]. They can be characterised based on whether the collaborators are co-located, and if so, how they are embodied from each other's point of view. For instance, Bell implemented a networked instrument that allowed participant, located across Europe and represented as avatars, to explore a sound corpus arranged by timbre similarity [1]. In this context, participants are free to move around in relationship to the environment through a world manipulation or "point-tugging" [10] metaphor where grabbing the air allows to pull or push their viewpoint. Similarly, in a study by Boem *et al.* investigating musical collaboration and social interaction, three Virtual Musical Instruments featured a "flying" metaphor. Its implementation there is not explicit, though it usually consists of allowing participants to shift forward in response to a thumbstick for instance. In each instrument, participants produced music cooperatively through respectively shared control of a sound source, distances between each other and objects mapped to a synthesizer, or contact with each other's avatars triggering sounds. A similar "walking" movement metaphor, where the participant is constrained on the vertical axis, is used in *WAM Jam Party* [7], where networked participants share and interact with virtual audio modules. With co-located collaborators, virtual viewpoints tend to be strictly tied to the physical movement of the user's head, without a way for them to be manipulated freely. This is the case for the performance *Trois Machins de la Grâce Aimante* [12], which features four Coretet instruments independently manipulated by four users who can see each other as avatars. The same applies to LeMo [18], where two participants could operate on the same virtual sequencing interface while also appearing as avatars. This hard coupling of virtual and physical reference frames is also present in co-located instruments without avatar-related representation. It appears in ARLooper [21], with which several participants can record, place and edit audio samples as shapes in the virtual world. These shapes are shared between participants, but can not be edited simultaneously.

The lack of navigation options that appears in these co-located examples can be seen as an avoidance of the conflicts that could arise if the control of the spatial relationship between the physical and virtual environments was accessible. On one hand, this allows a guarantee that at any point of the interaction, the relationship between virtual and physical elements is maintained, which in these examples applies mainly to the participants themselves. If a participant could shift aside the entire virtual environment during a performance with LeMo [18], the avatars of either users would not appear where their physical body is, notably raising collision issues. On the other hand, this approach drastically limits the potential for interactions in co-located instruments. Enabling this decoupling of physical and virtual spaces relies on users accepting a dual representation of the other in relation to both the physical and virtual environments [9], while addressing the challenge of concurrently controlling the spatial relationship between both environments.

2.2 Concurrent Manipulation

When considering the shared control of an element, in our context the spatial relationship between virtual and physical environments, different sharing modalities exist. These have been classified as cooperation levels [17], with the first only including mutual perception. The second level adds individual control of elements within the scene, and the third provides the ability to act on the same element either in independent or codependent ways. The two situations outlined in the third level match strategies for resolving concurrent interaction proposed by Broll [6]. Independent interaction on the same element relates to a constraint based interaction, where users have control over independent parameters of the same object, resulting in a form of degree of freedom separation. In the context of spatial manipulation, this would be akin to a person controlling the position of the virtual origin in relation to the physical world, while an other could control its scale. For dependent interaction, and according to both proposals, a combination of the inputs needs to be computed. However, in the context of musical interaction, collaborators may want to each have complete control over the environment, which can not be accomplished through these strategies.

By introducing an individual mode, collaborators are able to independently manipulate their own viewpoint, effectively decoupling their physical and virtual spaces. This enables the alteration of viewpoints in terms of position, scale, and permits to quickly share someone else's view [8]. When co-located in VR, this mode runs the risk of collision between users [15] and may require complementary information or secondary avatars corresponding to the physical position of the others. Specifically in MR with a view on the physical other, this mode results in a "dual-presence", introducing a perceptual conflict that may impact task performance [9], but may not necessarily hinder more creative activities. Applying this concept to musical interaction may bring forth novel strategies, which is why we decided to implement a musical instrument in MR that can leverage this independent decoupling of spaces.

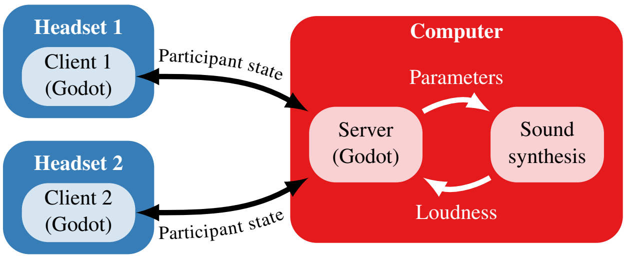

Figure 2: Overview of gRAinyCloud's architecture, showing the communication between its components

3 gRAinyCloud

gRAinyCloud is a MR musical instrument that allows two musicians (although it could be extended to more) to explore a set of sounds associated with virtual shapes placed in the physical space (see Figure 1), triggering them using virtual rods attached to the controllers of a MR headset.

The instrument was implemented using the Godot game engine, with a client on each headset communicating through network messages using OSC with a server running on an independent computer, which exchanges information with a PureData patch handling the sound synthesis (see Fig. 2).

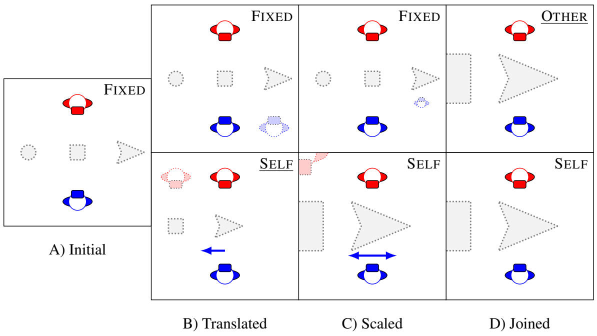

Figure 3: Example sequence of changes in points of view (from Left to Right) : From an initial FIXED and shared view of the virtual space (A), the blue musician translates the virtual space (B), implicitly switching to their SELF view with avatars indicating the decoupled points of view, (C) then they zoom-in on one of the shapes, with corresponding changes in scale of the avatars. (D) Finally, the red musician switches to the OTHER musician's view, restoring a shared view of the virtual objects.

3.1 Viewpoint decoupling

The main purpose of gRAinyCloud is to open novel opportunities of musical collaboration in MR by allowing musicians to manipulate the viewpoint on a shared virtual instrument while preserving the advantages of physical communication and cooperation. To do so, we drew inspiration from previous research in XR, notably on multi-scale collaboration and viewpoint decoupling. Multi-scale collaboration, as presented by Le Chénéchal *et al.* or Nguyen *et al.* [16, 20], enables users to interact within a shared virtual environment from a usual viewpoint (at a 1:1 scale) and from a zoomed-out viewpoint (*e.g.* looking at a miniature of the virtual environment). In the most common setting, one user acts as a conductor, placing guides and cues for other users. Another inspiration is the decoupling of viewpoint which, as demonstrated by Sol Roo and Hachet [22], enables transitioning between various views of the same scene, each affording specific interaction techniques, such as coarse interaction of tangible artefacts in spatial augmented reality and fine-grained interaction of the virtual copies of the artefacts.

In gRAinyCloud, we take advantage of these mechanisms for the design of MR musical instruments. As shown in Figure 3, starting from a FIXED view of the virtual environment shared by both musicians, this viewpoint can be translated, rotated and scaled, therefore decoupling the virtual and physical spaces. This manipulation is performed using a "grabbing-the-air" metaphor [10], by pressing and holding the trigger on one controller for translations and two for rotation/scale. As depicted in Figure 3.B and Figure 1.B, virtual avatars of the musicians (here composed of the headset and controllers) consequently appear in the physical space to visualise the other's viewpoint. Depending on the transformation, the avatars may be translated, rotated and scaled (see Figure 3 B and C).

These manipulations result in three views : FIXED which provides the original alignment of virtual and physical spaces for all, SELF which corresponds to the musicians's own viewpoint transformation, OTHER which allows one to see the virtual space from the other's point of view, thus realigning the two avatars and physical musicians (see Figure 3.C). As depicted in Figure 4, changes between SELF, OTHER and FIXED are performed either explicitly using a joystick (Left/Right to access the FIXED view, Top to access the OTHER view and Bottom to access the SELF view) or implicitly whenever the space is manipulated. Note than when one is in the OTHER mode, changes in viewpoint made by the other are dynamically applied, potentially resulting in changes in virtual objects positions while interacting, but also opening opportunities for collaboration.

Figure 1 shows how these changes are seen by the musicians and how they might enable different ways of interacting and collaborating, such as from moving the viewpoint in order to gain access to virtual shapes while avoiding physical conflicts (see Figure 1.B) or zooming-out in order to more easily access all virtual shapes for more dynamic controls (see Figure 1.C).

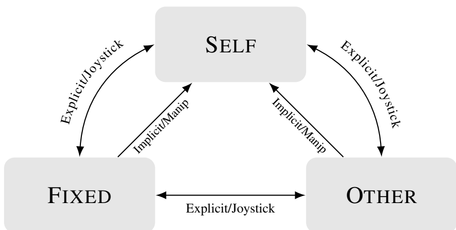

Figure 4: Transitions between FIXED, SELF and OTHER views can be performed explicitly using the joystick to select the views or implicitly, since any manipulation of the virtual space will result in a switch to the SELF view.

3.2 Sound synthesis and Mappings

The decoupling approach presented above could be applied to any virtual space shared using individual MR displays (headsets or mobile screens). In order to test it, we propose a first implementation in an instrument inspired by sound corpus exploration.

gRAinyCloud provides two modes of playing the virtual shapes, both activated when intersecting the shapes : 1) the *percussive* mode corresponds to the sound being triggered at a controllable tempo (and multiple of a shared global clock) and with a controllable low-pass filter cutoff; 2) the *rubbing* mode corresponds to the sound being played as a texture (using granular synthesis) with controllable pitch and Low-Frequency Oscillator amplitude modulation. Users are free to switch between the two modes by pressing a button on the controller.

Parameters of these two modes are controlled through the rotation angles (pitch and yaw) of the controller relative to the performer-shape direction. In addition the distance from the controller to the shape is mapped to the gain, this distance is normalised to the maximum reachable distance, which is the length of the instrument added to the scaled radius of the shape. Each controller of each musician can generate a voice for each shape, allowing for polyphonic interaction with the sounds.

After the synthesis, the loudness is transmitted to the server which displays it on the shapes themselves as a visual feedback for the musical interaction.

4 Evaluation

We proceeded with an evaluation of the impact on strategies for and experience of collective music playing of the viewpoint decoupling, using an approach inspired by a self-use and improvisation method previously used with an Augmented Reality NIME [26].

4.1 Procedure

In addition to playing sessions and discussions during the design phase of gRAinyCloud, two of the authors conducted a more formalised set of three sessions of short (around ten minutes) improvisations interleaved with discussions. Both testers possess a strong background in software engineering, with comparatively limited experience in musical performance. One of them also has some experience in improvisation. The playing sessions were not guided by rules or constraints, but instead built upon the discussions, creating new perspectives and the desire to test various strategies. The playing sessions were recorded from each headset to provide the musicians' point of view, along with a global recording of the physical space, and the generated audio, allowing for detailed examination.

4.2 Observed strategies

We can first look at the employed and discussed collaboration strategies.

4.2.1 No use of the Fixed viewpoint. Throughout the three sessions, the Fixed viewpoint found no real use in contrast to the other views. Our hypothesis is that this is due to the size of the ensemble. Being only two musicians meant that reaching a shared viewpoint only required switching to the Other view. The Fixed viewpoint may have been more relied on with a larger number of collaborators, since the solution to needing to quickly synchronise their point of view would be more complex. This hypothesis still requires validation.

4.2.2 Avoiding physical conflicts. All potential contest over physical space was avoided easily as users could grab the environment and bring it with them while staying physically away from the other. In a way, this strategy results in recreating private spaces for each musician. This can be seen in Figure 5.A, where the user sees the other's controllers and interact with the same virtual objects while avoiding physical conflict.

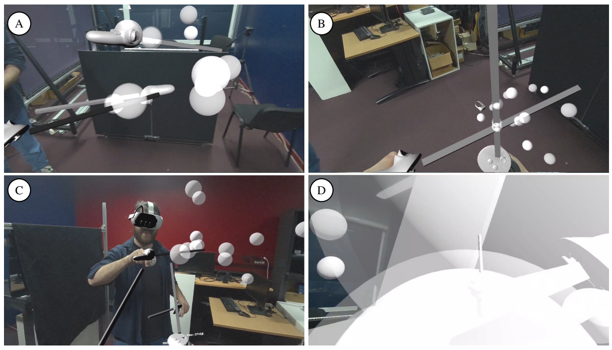

Figure 5: Observed collaboration / playing strategies : A) Moving view in order to avoid physical conflicts while maintaining a visual feedback on the other's activity. B) Zooming-out in order to quickly access all the virtual space. C) Moving the space while the other is in one's view, therefore disrupting the interaction. D) Zooming-in to finely manipulate one specific sound, with the other's controllers and instruments appearing extremely large.

4.2.3 Stealing from / Steering the other. This behaviour also eventually caused someone to "steal" the objects away from the other when they were in Other view, as depicted in Figure 5.C. This is not necessarily negative since, when done on purpose, entangling the interaction of multiple musicians can lead to alternative creative strategies, such as with the Tooka [11] or the Perceptron [23]. This could also be used as a way to conduct a performance, if one of the musicians was given a dedicated role of navigating through the virtual space.

4.2.4 Rejoining / Leaving the other's viewpoint. The option to join or leave the other participant's view has been mainly used for two purposes. The first is to jump to their view and back in order to quickly swap between two setups, which is especially relevant when there is a large scale difference, and was for example employed to switch between viewpoints in Figure 5.B and Figure 5.D. The second is to meet the other physically on an object, usually to engage in a more closely coupled collaboration. This re-alignment has also been performed manually, by progressively shifting the space in order to match the view of the other by placing their avatar on their physical body. This strategy emerged as an answer to the discrete change disrupting an ongoing interaction. This way, it is possible to meet the other physically without interrupting the current interaction with an object.

4.2.5 Zooming in and out. Along the sessions, large changes in scales were used to access other modes of interaction. Zooming-out, *i.e.* reducing the scale of the virtual space, was for instance used for fast selection among a set of small shapes. Zooming-in, *i.e.* increasing the scale of the virtual space, allowed for fine-grained control over a single very large shape. In both cases, the changes resulted in modification of the other's avatar, appearing either as a very small character interacting on a small part of the space (see Figure 5.B, or as a large overlooking figure (see Figure 5.D with the giant controller and rod on the right). Both had an impact on the collective music making experience and behaviour, for example resulting in playful interactions with the small scale avatar.

4.3 Discussion and guidelines

From the formal playing sessions and the overall design process, we can provide a number of guidelines for future investigation of decoupling in co-located MR instruments.

4.3.1 Favour awareness of the others and of the virtual space. The first relate to awareness, *i.e.* the perception and knowledge of others' activity and status within the instrument. The current view should be visible at all times in order to avoid the scenario of being "stolen from" when in the view OTHER. Displaying the view as a small label next to the controller was not sufficient in these sessions, and a more immediately available channel should be used, like for instance changing the aspect of the virtual elements. Complementarily, having the view of the other readily available, especially when they see one's own view, would help keeping the users aware of the potential for a "stealing from" scenario. In addition, the virtual representation of the other is an important consideration for maintaining a sense of copresence, even in co-located MR [25]. In an exploratory instrument such as gRAinyCloud, helping users quickly ascertain the current transformation of the structure becomes especially important when they can jump from a view to a completely different one. In our case, changing the shapes or colours of the elements would help, but it could also be achieved through non-interactable virtual guides.

4.3.2 Increase expressiveness in decoupling. As seen in the strategies, a discrete transition between views is more appropriate without an ongoing interaction, and allowing a smoother or controlled transition between the viewpoints is desirable. Implementing a way to select other musicians either physically or through their avatar can be useful, especially when considering scaled up contexts with more collaborators. Finally, the approach of a personal and shared version can also be applied to the audio itself [18]. In this case, we could imagine controlling whether our own audio output was heard only by ourselves, or if it was broadcast to all, with the potential for listening in on what others might be playing. This would allow a part of private experimenting, which is especially relevant to exploratory instruments such as gRAinyCloud.

4.4 Limitations

The strategies and guidelines reported in this study stem from a self-study approach of two of the authors on a specific instrument. They would gain from being completed through studies investigating the practice of a more diverse group of participants, and with different instruments leveraging the decoupling of physical and virtual spaces. gRAinyCloud remains very simple and lacks the complexity of a more mature instrument, which could see emerge more elaborate strategies. For instance, the present version of gRAinyCloud only fits two collaborating participants for the purpose of this paper and in order to limit the complexity of the implementation and analysis. However, its interaction scheme is scalable to more users, and investigating the strategies from a larger group of collaborators may reveal additional insights. Studying our approach through a variety of instruments may also reveal different strategies that are not linked to the exploratory nature of the instrument, or to other design choices such as the length of the rods, which may have mitigated physical conflicts.

5 Conclusion

In this paper, we presented a novel approach to musical collaboration in co-located MR instruments. It relies on the decoupling of virtual and physical spaces, resulting in different views between which musicians can transition in order to access various modes of playing and collaborating. We presented the instrument gRAinyCloud, which implements this approach and allowed us to conduct an evaluation to observe some of the collaboration strategies that emerge over a short period of practice.

A first perspective of this work is to investigate visualisation and interaction techniques for the guidelines that we propose, in particular how to improve awareness and expressiveness in decoupled collaboration.

Another perspective relates to the effect of the decoupling on the overall collective music making experience. Fixed visualisations of the others activity such as in [3] have shown that users tend to adopt a different state of mind when focusing on the visualisations of others' activity rather than on them directly. In contrast, here musicians can quickly switch between avatars for activity visualisation and the physical musicians. It might therefore be interesting to study the effect that this approach has on the way musicians conceptualise collective musical expression.

Finally the scenography for performances with such a system raise important issues, regarding what the audience should see and how musicians and the audience should see each other [28].

6 Ethical Standards

This research was approved by the author's institution ethics committee under number 2023-743-S122 and it was funded through public grants from the Fonds de Recherche du Québec (FRQ PRISME-ART) and the Hauts de France region.

References

[1] Jonathan Bell. 2023. Networked Music Performance in PatchXR and FluCoMa. In *Proceedings of the International Computer Music Conference*. CUHK-Shenzhen, Shenzhen, China.

[2] Florent Berthaut. 2020. 3D interaction techniques for musical expression. *Journal of New Music Research* 49, 1 (Jan. 2020), 60–72. https://doi.org/10.1080/09298215.2019.1706584

[3] Florent Berthaut and Luke Dahl. 2022. The Effect of Visualisation Level and Situational Visibility in Co-located Digital Musical Ensembles. In *New Interfaces for Musical Expression*. Auckland, New Zealand. https://doi.org/10.21428/92fbeb44.9d974714

[4] Tina Blaine and Sidney Fels. 2003. Contexts of collaborative musical experiences. In *New Interfaces for Musical Expression*. Montreal, Quebec, Canada, 129–134.

[5] Alberto Boem, Matteo Tomasetti, and Luca Turchet. 2024. Harmonizing the Musical Metaverse: unveiling needs, tools, and challenges from experts' point of view. In *New Interfaces for Musical Expression*. Utrecht, The Netherlands.

[6] W. Broll. 1995. Interacting in distributed collaborative virtual environments. In *Proceedings of the Virtual Reality Annual International Symposium*. 148–155. https://doi.org/10.1109/VRAIS.1995.512490

[7] Michel Buffa, Dorian Girard, and Ayoub Hofr. 2024. Using Web Audio Modules for Immersive Audio Collaboration in the Musical Metaverse. In *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)*. IEEE, Erlangen, Germany, 1–10. https://doi.org/10.1109/IS262782.2024.10704108

[8] Weiya Chen. 2015. *Collaboration in Multi-user Immersive Virtual Environment*. phdthesis. Université Paris Saclay.

[9] Weiya Chen, Céline Clavel, Nicolas Férey, and Patrick Bourdot. 2014. Perceptual Conflicts in a Multi-Stereoscopic Immersive Virtual Environment: Case Study on Face-to-Face Interaction through an Avatar. *Presence: Teleoperators and Virtual Environments* 23, 4 (Nov. 2014), 410–429. https://doi.org/10.1162/PRES_a_00209

[10] Noah Coomer, Sadler Bullard, William Clinton, and Betsy Williams-Sanders. 2018. Evaluating the effects of four VR locomotion methods: joystick, arm-cycling, point-tugging, and teleporting. In *Proceedings of the 15th ACM Symposium on Applied Perception (SAP '18)*. New York, NY, USA, 1–8. https://doi.org/10.1145/3225153.3225175

[11] Sidney Fels and Florian Vogt. 2002. Tooka: Explorations of Two Person Instruments. In *New Interfaces for Musical Expression*. Dublin, Ireland, 6.

[12] Rob Hamilton. 2019. Trois Machins de la Grâce Aimante: a Virtual Reality String Quartet. In *Proceedings of the International Computer Music Conference*.

New York, NY, USA.
[13] Rob Hamilton and Chris Platz. 2016. Gesture-based Collaborative Virtual Reality Performance in Carillon. In *Proceedings of the International Computer Music Conference*. Utrecht, The Netherlands.
[14] Sergi Jordà. 2005. Multi-user instruments: models, examples and promises. In *New Interfaces for Musical Expression*. Vancouver, Canada, 23–26.
[15] Jérémy Lacoche, Nico Pallamin, Thomas Boggini, and Jérôme Royan. 2017. Collaborators awareness for user cohabitation in co-located collaborative virtual environments. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*. Gothenburg Sweden, 1–9. https://doi.org/10.1145/3139131.3139142
[16] Morgan Le Chénéchal, Jérémy Lacoche, Jérôme Royan, Thierry Duval, Valérie Gouranton, and Bruno Arnaldi. 2016. When the giant meets the ant an asymmetric approach for collaborative and concurrent object manipulation in a multi-scale environment. In *IEEE third VR international workshop on collaborative virtual environments (3DCVE)*. Greenville, SC, USA, 18–22.
[17] David Margery, Bruno Arnaldi, and Noël Plouzeau. 1999. A General Framework for Cooperative Manipulation in Virtual Environments. In *Virtual Environments '99 (Eurographics)*, Michael Gervautz, Dieter Schmalstieg, and Axel Hildebrand (Eds.). Springer, Vienna, 169–178. https://doi.org/10.1007/978-3-7091-6805-9_17
[18] Liang Men and Nick Bryan-Kinns. 2019. LeMo: Exploring Virtual Space for Collaborative Creativity. In *Proceedings of the Conference on Creativity and Cognition*. ACM, San Diego, CA, USA, 71–82. https://doi.org/10.1145/3325480.3325495
[19] P Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual display. *IEICE Trans. Information Systems, D* 77, 12 (1994), 1321–1329.
[20] Thi Thuong Huyen Nguyen, Thierry Duval, and Cédric Fleury. 2013. Guiding techniques for collaborative exploration in multi-scale shared virtual environments. In *GRAPP International Conference on Computer Graphics Theory and Applications*. Barcelona, Spain, 327–336.
[21] Sihwa Park. 2020. Collaborative Mobile Instruments in a Shared AR Space: a Case of ARLooper. In *New Interfaces for Musical Expression*. Birmingham City University, Birmingham, UK, 190–195. https://doi.org/10.5281/zenodo.4813313
[22] Joan Sol Roo and Martin Hachet. 2017. One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities. In *ACM UIST'17*. Quebec City, Canada. https://doi.org/10.1145/3126594.3126638
[23] Steve Symons. 2024. The Perceptron: A Multi-player Entangled Instrument based on Interpretive Mapping and Intra-action.. In *Proceedings of the 19th International Audio Mostly Conference: Explorations in Sonic Cultures (AM '24)*. ACM, New York, NY, USA, 385–391. https://doi.org/10.1145/3678299.3678338
[24] Luca Turchet, Rob Hamilton, and Anil Çamci. 2021. Music in Extended Realities. *IEEE Access* 9 (2021), 15810–15832. https://doi.org/10.1109/ACCESS.2021.3052931
[25] Pierrick Uro, Florent Berthaut, Thomas Pietrzak, and Marcelo Wanderley. 2025. Behavioral Measures of Copresence in Co-located Mixed Reality. *TVCG* (2025). To appear.
[26] Yichen Wang and Charles Martin. 2022. Cubing Sound: Designing a NIME for Head-mounted Augmented Reality. In *New Interfaces for Musical Expression*. Auckland, New Zealand. https://doi.org/10.21428/92fbeb44.b540aa59
[27] Anna Xambó, Eva Hornecker, Paul Marshall, Sergi Jordà, Chris Dobbyn, and Robin Laney. 2013. Let's jam the reactable: Peer learning during musical improvisation with a tabletop tangible interface. *ACM Transactions on Computer-Human Interaction* 20, 6 (Dec. 2013), 1–34. https://doi.org/10.1145/2530541
[28] Victor Zappi, Florent Berthaut, and Dario Mazzanti. 2022. From the Lab to the Stage: Practical Considerations on Designing Performances with Immersive Virtual Musical Instruments. In *Sonic Interactions In Virtual Environments*. Springer VS. https://hal.science/hal-03790049