Revealing Interactions : Expression and Exploration in Mixed and Virtual Reality

Florent Berthaut

To cite this version:

Florent Berthaut. Revealing Interactions : Expression and Exploration in Mixed and Virtual Reality. Graphics [cs.GR]. Université de Lille, 2023. tel-04425551

HAL Id: tel-04425551
https://hal.science/tel-04425551v1

Submitted on 31 Jan 2024

UNIVERSITÉ DE LILLE

HABILITATION À DIRIGER DES RECHERCHES

Revealing Interactions : Expression and Exploration in Mixed and Virtual Reality

Interactions Révélées : Expression et Exploration en Réalité Mixte et Virtuelle

---

Florent Berthaut

Defended on October 27th 2023

Jury

- **Frédéric Bevilacqua**, Directeur de recherche, IRCAM (Rapporteur)
- **Laurence Duchien**, Professeur des Universités, Université de Lille (Présidente)
- **Thierry Duval**, Professeur des Universités, IMT Atlantique (Rapporteur)
- **Laurent Grisoni**, Professeur des Universités, Université de Lille (Garant)
- **Laurence Nigay**, Professeure des Universités, Université Grenoble Alpes (Rapporteure)
- **Stefania Serafin**, Professeure, Aalborg University, (Examinatrice)

Abstract

Artistic and cultural productions have often followed or even inspired the development of immersive technologies. They have for instance greatly benefited from the generalisation of mixed and virtual reality headsets, mobile augmented reality or projection mapping software that happened in the past two decades. By freely integrating virtual elements within the physical space, these technologies allow for enriching the perception and interactions of artists, ensembles and audiences. My research lies at the intersection of 3D Interaction in Mixed and Virtual Reality and of Artistic, in particular Musical, Expression and Cultural Applications. Part of this research uses mixed-reality displays to reveal the mechanisms of complex or unfamiliar expressive interactions to spectators and collaborators of musical performances, with the goal of understanding and augmenting their experience. The other part investigates opportunities that arise for expression and cultural content exploration when virtual elements are revealed in the physical space, with the development of novel display technologies, interaction devices and techniques. Perspectives of this research include an in-depth study of the perception of digital interactions by the audience, the generalisation of levels-of-detail in the design of expressive interfaces, and the shift towards 3D interfaces that are less centered on technology and more on human expression.

Résumé

La production artistique et culturelle a souvent suivi ou même inspiré le développement de technologies immersives. Elle a ainsi fortement bénéficié de la généralisation des casques de réalité mixte et virtuelle, de la réalité augmentée mobile ou des logiciels de mapping vidéo lors des deux dernières décennies. En intégrant librement des éléments virtuels dans l'espace physique, ces technologies permettent d'enrichir la perception et les interactions des artistes, des collectifs et des spectateurs. Ma recherche se situe à l'intersection de l'Interaction 3D en Réalité Mixte et Virtuelle et des Interfaces pour l'Expression Artistique, en particulier Musicale, et la Médiation Culturelle. Une partie de ces travaux se concentre sur l'utilisation de la réalité mixte pour révéler les mécanismes des interactions expressives complexes ou non familières aux spectateurs et collaborateurs, avec l'objectif de comprendre et d'enrichir leur expérience. L'autre partie montre quelles opportunités émergent pour l'expression et l'exploration de contenus culturels, lorsque des éléments virtuels sont révélés dans l'espace physique, avec le développement de nouvelles technologies d'affichage ainsi que de techniques et dispositifs d'interaction. Les perspectives de ces travaux incluent une étude approfondie de la perception des interactions numériques par les spectateurs, une généralisation des niveaux de détails dans la conception d'interfaces expressives, et le basculement vers des interfaces 3D moins centrées sur la technologie et plus sur l'expression humaine.

Acknowledgements

Contents

1 Introduction

This document covers the research that I have been conducting for the last ten years at the intersection of 3D User Interfaces (*3DUIs*), musical expression and content exploration. In particular, I looked at how *3DUIs*, in mixed and virtual reality, can help understand and enrich the perception of complex digital interactions by diverse roles and in diverse contexts, but also how they may open novel expression and exploration opportunities by closely integrating physical and virtual spaces.

Interaction and expression diversity and complexity

The development of interactive technologies, in particular in the context of artistic expression (music, theater, drawing, dance), has led to an increasing diversity and complexity of interactions with digital content.

This is due to many factors. The first is the diversification of interaction hardware. This hardware can be 1) commercially available, such as control surfaces, 3D displays (including Virtual Reality (*VR*) head-mounted displays), haptic feedback devices. 2) custom built using sensors and hardware development kit 3) a mix of both, with prepared or hacked [129] interfaces.

Similarly, software components for content manipulation, including synthesis and processing of 2D and 3D graphics or sound, have gained diversity, accessibility and customisability. For instance, advanced 3D programming to create and display virtual environments has gone from low level software libraries to accessible game engines with large communities and readily available complex building blocks. The design of interactive systems also benefits from a number of dedicated languages and environments, such as creative programming textual languages (*e.g.* Processing, OpenFrameworks, three.js ...), visual programming / patching environments (*e.g.* PureData), which enable the sharing of modifiable building blocks across communities, making appropriation of software components even easier.

This combination of interactive hardware and software means that users, in particular artists, are free to customize their tool / instrument [146]. In the case of musical expression, they can decide on, or appropriate, the interface, the gestures, the sound processes and parameters, the feedback and the mappings (connections between gestures and changes in the sound) [103]. This has led to the emergence of a large variety of practices and configurations with either subtle or strong differences, which contributes to the richness and diversity of human expression. However, this diversity sometimes constitutes an issue, especially because part or all of these instruments are digital or even virtual. This issue applies to various roles [144] : the main (expert) users, their collaborators and the potential spectators of the interaction.

Physical vs Virtual

Contrary to digital ones, purely physical (*e.g.*, tangible, mechanical, acoustic) interfaces usually provide some amount of familiarity to observers of the interaction, *i.e.*, spectators and collaborators. They may have already performed the same or similar actions, encountered the interface multiple times. Furthermore, physical interfaces rely on laws of physics (known through naive physics [105,140]) and on properties of the human body and behaviour. Observers therefore automatically, and often unconsciously, access cues that enable them to deduce : the causality relation between gestures and responses of the physical environment; the user's intention; the user's skill, errors, level of virtuosity; the potential actions and outcomes. Similarly, physical interfaces provide a number of affordances [91] for the main users which depend on their familiarity with the perception and manipulation physical objects, with the exploration of the physical space, with the perception of their body and so on.

In the case of interaction with digital content, especially as it is increasingly transported to a virtual space [20], the cues used by spectators, collaborators and users are often disrupted. For instance : 1) gestures performed and sensed by the system may be unfamiliar, too subtle or even hidden [145], 2) the manipulated system can be unfamiliar and complex, with hidden mechanisms and states, and a behavior that can not be deduced from its physical appearance, 3) information about the system might not be physically aligned with the interface, leading to a potential lack of feedback and degrading the interaction, 4) the interface might not take advantage of the users' skills in interacting in the physical space with their body, limiting expression opportunities. This may constitute an issue for the audience and collaborators, as we will explain in detail in Section 2, because it may hinder aspects of their experience. Symmetrically, compensating for these issues can enrich interaction and expression opportunities for main users.

Revealing the unfamiliar

One possible solution to the issue of unfamiliar digital and virtual interactions is to apply design constraints to ensure a correct experience for all : relying on transparent interaction metaphors [83]; restraining interaction techniques and devices for example with large familiar gestures and big physical components [38] making sure that all software components are visible at all times and sufficiently simple; using only common hardware and software components. However, I believe that these restrictions limit the creativity of users and designers and over time prevent the emergence of new forms of human expression, which led me to look for another direction.

During the past decade, I have explored the opportunities offered by 3D User Interfaces (*3DUIs*) [44] in Mixed and Virtual Reality (*MR, VR*) [155]. Because they permit the perceptual integration of virtual and physical components, and the combination of virtual and physical spaces, these interfaces can be used to augment and/or enrich interfaces, especially expressive ones, without constraining their design or limiting the user's appropriation. The addition of virtual components becomes in itself a way of appropriating interfaces, for users, collaborators and spectators alike.

Over the course of the research presented in this document, this approach has come to be designated as *"Revealing Interactions"* or *"Revealed Interfaces"*. It echoes and builds upon research in the fields of Human-Computer Interaction, 3D User Interfaces and New Interfaces for Musical Expression, in particular the study of spectator experience [95,144], Collaborative Virtual Environments [16], design of 3D interaction techniques [3,44,106,131], Mixed and Virtual Reality displays [20,77,154,155,160], gesture to sound mappings [63,103] or interface appropriation [98].

Application to musical expression and exploration

This approach naturally applies to artistic / expressive interfaces where spectators and collaborators have an important role to play, and where users strive for novel ways of creating and manipulating digital content.

Mixed and virtual reality displays have a long history of usage in artistic contexts [77], with multiple temporal phases of development that can attributed to changes in technologies and results from research [154,160]. They however often remain limited to the use of generic interaction techniques and devices [22], usually those available in common game engines, and to the virtual transposition of physical artistic interfaces, without embracing novel opportunities created by *3DUIs*.

Here the term *expression* refers to the communication, through the interaction with digital content, of one's ideas, emotions, reactions, intentions, by opposition to functional interactions that align with a precise task that need to be accomplished. In that sense, the quality of expressive interfaces often lies in the interaction itself more than in the result of this interaction. They therefore align with exploratory interfaces where the process of discovering information constitutes the essence of the interaction.

Among artistic practices, this research mostly concentrates on musical expression, for which the issue of unfamiliarity is key. As will be explained in detail in Section 2, musical performance and collective music making with digital instruments are extreme examples of interaction complexity and diversity, of which the experience of spectators and collaborators are essential aspects. Furthermore, musicians have a tradition of "prepared" and hacked instruments, which was amplified with digital instruments. A large part of musical practice is to explore novel playing techniques and gestures, which can benefit from interaction with virtual components added to physical objects and spaces.

Along with the question of expression, this research emphasizes the exploratory aspect of interaction. Indeed, the approach of revealing interactions using 3D interfaces enables the selection of which virtual elements are visible and are interacted with, during both the design and usage phases, allowing users to explore and discover content. Outside artistic expression, I have for instance applied this approach to museum exhibitions, where users can discover information about exhibited physical objects.

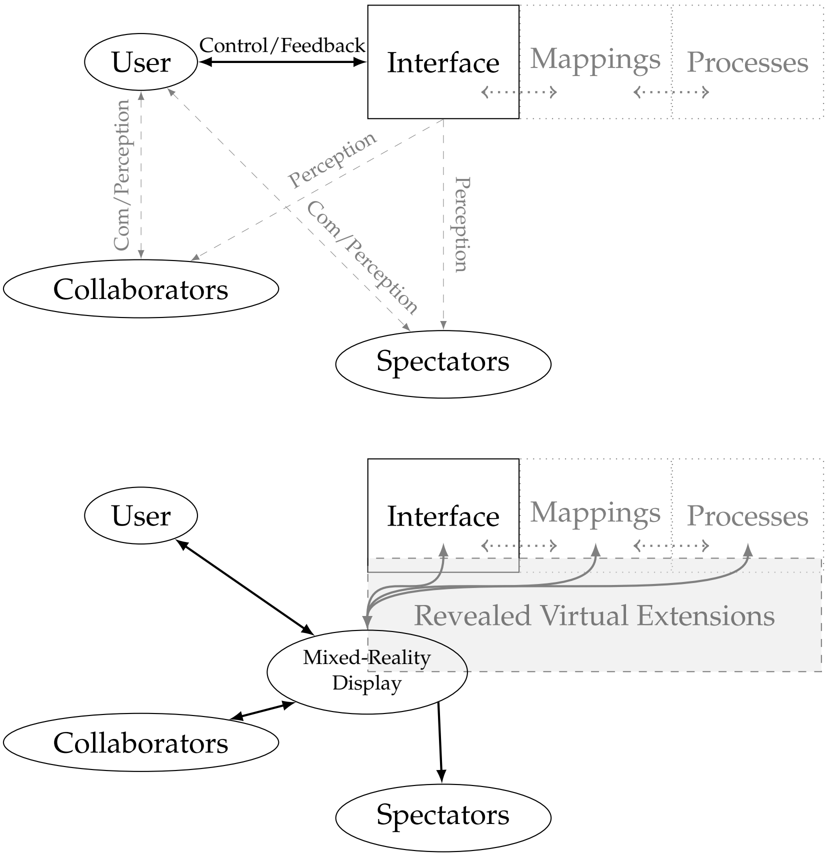

FIGURE 1.1: Top : Contexts of expressive interaction involve various roles of users. Perception and communication are however often disrupted by the diversity and complexity of interactions. Bottom : Mixed-reality displays create shared physical and virtual spaces where revealed virtual extensions of the interfaces may enrich the control and feedback for all roles.

1.1 Research collaborators and projects

During these years of research, I have led and participated in collaborative projects at the local, national and international levels. These collaborations involved other researchers in computer science and human-computer interaction but also in electrical engineering, cognitive sciences, music technology and archaeology. At the local level, I am actively working with researchers from multiple teams of the CRIStAL and L2EP laboratories of the University of Lille, but also with researchers in acoustics and cognitive sciences at the Catholic University of Lille.

Part of the research on spectator experience was conducted during the IXMI project that I led at the University of Bristol (UK), funded by a FP7 Marie Curie Intra-European Fellowship (Grant agreement 330770) in 2013-2015. I had the chance to collaborate with the following colleagues : Sriram Subramanian, Diego Martinez Plasencia, Deepak Sahoo, David Coyle, James Moore, Hannah Limerick, Mark Marshall. The first research paper on the augmentation of audience experience [33] was also written in collaboration with Martin Hachet from INRIA Bordeaux.

The document introduces results from long-term collaborations with colleagues abroad and in France. At the international level, the BOEUF project on digital orchestras is an ongoing project which started in 2014 and is conducted in collaboration with Luke Dahl at the University of Virginia. The study of the scenography of Immersive Virtual Musical Instruments is an ongoing project in collaboration with Victor Zappi at Northeastern University (USA) and Dario Mazzanti (IIT, Italy).

At the local level, the Vibrating Shapes project on spatial augmented reality for actuated instruments is a collaboration with Arthur Paté (ISEN, Junia) with funding from the IRCICA and the Université de Lille. Some of the research in collaboration with colleagues at the Université de Lille was funded through regional projects, including the “Sans Reserve” CPER Mauve project with Patricia Plénacoste, Yvan Peter, Fatma Ben Guefrech and Anthony Beuchey, and the TerRev project funded by a Hauts-de-France region “Stimule” funding with postdoctoral researcher Cagan Arslan.

Important results in the presented research come from the theses of two PhD students that I co-supervised and which originate from my research thematics. In particular, I co-supervised with Laurent Grisoni the thesis of Olivier Capra on the experience of spectators in performances with Digital Musical Instruments [57]. I also co-supervised with Laurent Grisoni the thesis of Vincent Reynaert on the perception of muscular fatigue in virtual reality [147]. Both of which have now successfully defended their PhD theses and were recruited at the Catholic University of Lille, respectively as researcher and lecturer. We are still actively collaborating, respectively on the perception of digital interactions and the design of 3D interaction techniques.

Some results were also obtained in collaboration with PhD student Cagan Arslan, whom I hired after his PhD thesis for a postdoctoral position in a regionally-funded research project on spatial augmented-reality for artistic performances. Research on cross-modal perception of virtual textures was conducted as part of Daetjon Brahimaj’s Phd thesis in electrical engineering at the L2EP.

At the master level, since my recruitment at the University of Lille, I have supervised theses and internships of students in Computer Science, Musicology and Cognitive Sciences, some of which have led to published papers.

Finally, outside the academic domain, I have been strongly implicated in collaborations with artists (musicians, theatre and dance companies, visual artists) and collectives. These collaborations were essential as they 1) allowed me to investigate the experiences and practices of expert users, 2) led to new research questions and methodologies, and 3) facilitated the distribution of research results. I believe they constitute good example of working art-science collaboration, because they led to the production of both artistic works and scientific knowledge.

1.2 Presented contributions

The research presented in this document can be seen as tackling two broad questions :

1. Can mixed and virtual reality be used to understand and enrich the experience of spectators and collaborators of expressive interfaces ?

2. Can 3D user interfaces, by closely integrating virtual and physical spaces, open new opportunities for exploratory and expressive interactions ?

Within these questions, we have looked at various roles of users (as explained in a Computer Graphics and Application paper we wrote with Martin Hachet) [30] : novice and expert users of expressive and exploratory interfaces, collaborators in music ensembles and museum exhibitions, and the audience of performances.

In my work on Mixed-Reality interfaces, I tend to focus on the visual modality for augmentations. This choice is mostly explained by the flexibility in content and data presentation enabled by visual displays. However, some aspects can be transferred to other modalities, as explained for augmentation of the spectator experience in Section 2.3.2, and this will constitute one of the aspects of my future research.

In order to answer these research questions, I combined a variety of methodologies, such as prototyping and design of novel technologies and interaction techniques, quantitative controlled experiment, qualitative semi-controlled experiments with interviews and exploratory sessions, and qualitative "in the wild" studies in museums, artistic residences and performances. Throughout this research, I contributed : Theoretical frameworks and design spaces for design and analysis; Technologies such as novel displays and interaction devices; Interaction techniques; Knowledge on perception and experience.

Because the field of application of my research is principally musical expression, I publish mainly in musical interaction and computing conferences and journals (New Interfaces for Musical Expression, Computer Music Journal, Journal of New Music Research). However, when results can be generalized or extended outside the musical field I contribute to the HCI/VR community, with conferences such as ACM UIST, ACM DIS, ACM ISS, IEEE VR. In addition to scientific results, this work has led to a variety of outcomes such as artistic performances and installations, museum exhibitions, demonstrations for the general public, and the creation of economic activity for artists and collectives.

In the third part of this document, following the description of previous contributions, I propose future research directions which seem essential to continue investigating expressive interactions and 3D user interfaces.

2 Revealing Expressive Interactions

As explained in Section 1, the growing diversity and complexity of interactive systems, especially in artistic contexts, while opening rich opportunities for human expression, can be detrimental to the experience of spectators and collaborators. This is especially true in the context of digital musical performances, where manipulations (musical gestures) and effects (changes in the music) [145] are perceived with different modalities, where manipulations are often hidden, *i.e.*, with small and hidden gestures and sensors, where a large variety of musical parameters can be changed and where effects can be very diverse (from subtle modulations to triggering complete changes). This causality issue [152] does not exist with familiar acoustic instruments, for which, even when they are not visible, one has a clear representation of the link between characteristics of the performers gestures, *e.g.*, its energy, and characteristics of the resulting changes in the sound. Think for example of the understanding one might have of the engagement of a piano player whose hands and keyboard would not be visible, compared to that of an electronic musician in the same conditions.

However, discussing this issue with both musicians and audience members reveals essential differences in their experience. On the performers side, some recognize the importance of making their interactions more transparent to observers, which results in strategies such as amplifying their gestures, placing or adapting their interfaces so that they are clearly visible, choosing clearly identifiable gesture-sound mappings. Even when perceived changes in the music are automated, musicians will tend to synchronise their movements to these changes, essentially choreographing interactions. These practices align with research results that highlight the importance of the visual component of musical performances [142], *e.g.*, with respect to perceived emotions [162] and expressiveness [72,163]. On the other hand, some musicians, even when their performances have a high control / automation ratio, with improvised parts and a complex gestural vocabulary, prefer that the audience focus on the music rather than on the interactions. An extreme example are Autechre's [1] electronic music performances in the dark, with only enough lighting for the musicians to see their controllers. Between these two extremes there is a continuum of practices of *transparency* of interactions, which provide more or less visibility over various components of the performances, *e.g.*, gestures, parameters, interfaces, overall sound process activity, some of them more common, such as screen sharing in live-coding [123], Vjiing or augmented-instruments where the interfaces are familiar [117,136].

On the observer / audience side, the strategies also diverge. For some audience and musical ensemble members, clearly perceiving the interactions is not essential, because they tend to focus on their auditory perception and experience of the music. For others, there is a clear disruption of their experience compared to what more familiar, and often acoustic, instruments might bring, and they are looking for cues to increase their understanding of the performances. Between these two extremes again, a variety of strategies might emerge, as will be shown in Section 2.3.3.

Recognizing these issues has led to the study of the audience experience with digital interactions across disciplines, including :

- how to model the audience experience [88,97,144]
- how to measure it in the lab and "in the wild" [39,118,120,164]
- the influence of various components such as gesture size [38], instrument and repertoire familiarity [40], live visuals [66]
- how to evaluate instruments based on the audience experience [9]

---

[1] https://autechre.warp.net

On the collaborators / ensemble side, issues of interaction perception have historically been linked with concepts such as awareness [16] or embodiment [17,82] in research on Collaborative Virtual Environments (*CVEs*) [137] or Computer Supported Collaborative Work (*CSCW*), which were transposed in the context of Digital Orchestras [51,52,111,130].

The originality of our research compared to previous work lies in the use of 3D User interfaces, particularly in mixed-reality, to enable the integration of virtual content to complement the perception of physical interactions.

In this chapter, I present the research that I have been conducting on the use of mixed and virtual reality to reveal expressive interactions, in particular :

1. Interfaces to support communication and cooperation in digital orchestras (*DO*)
2. Scenography of performances with immersive virtual musical instruments (*IVMI*)
3. Visual augmentations to understand and enrich the experience of spectators with Digital Musical Instruments (*DMI*)

While this research focuses on the case study of musical expression, the same issues have been highlighted in other contexts [144] and the proposed solutions, insights and guidelines can be applied in other fields of HCI.

2.1 Augmenting Communication and Cooperation in Digital Orchestras

Most of my research on *Digital Orchestras* is conducted in collaboration with my colleague Luke Dahl from the University of Virginia. Results are presented on the BOEUF project website [2]. The collaboration started in 2012, when I went as a guest researcher to Stanford University where Luke Dahl was finishing his PhD. Our goal is to provide insights and tools for collaboration in *DOs*, especially in spontaneous ensembles with heterogeneous instruments, such as jam sessions.

In this context, each musician's instrument is not necessarily known in advance by the others. The diversity and complexity of *DMIs*, their lack of visibility [69] and the lack of physical cues or familiarity naturally present in acoustic instruments, might thus hinder communication between musicians and prevent the emergence of group dynamics essential for collective music making [153]. Furthermore, *DOs* often make use of cooperation techniques such as temporally synchronising instruments, or routing sound and controls between them. It has been the case since the first Digital Orchestras in the 1970s such as the League of Automatic composer and The Hub [94]. However, in spontaneous ensembles it is sometimes not possible to access these advanced mechanisms if the instruments rely on diverse software and hardware which have not been configured for interconnection. While standardised protocols such as MIDI or Ableton Link provide some amount of cooperation features (tempo or scale synchronisation), they are very far from what can be done in prepared orchestras with multiple instances of the same instrument [53,139,165] or multi-user instruments [109,110] where a single interface is shared.

Research on *DOs* has led to the formal analysis of dimensions of orchestras [42,168] and to the study of musicians' behavior and practices [171]. Tools have also been designed for the collaborative designing of *DOs* [125], or for interconnecting instruments [104]. However there is still much we do not know about what information is missing in *DOs* compared to acoustic ensembles, and how advanced cooperation can be enabled in the case of spontaneous collective playing sessions. Similarly, while mixed-reality displays have been used for example as a way to enable distant musical collaboration [151], they have not been employed to enrich *DOs*.

In order to tackle these issues, our approach consists in :

1. Creating a model of collaboration in *DOs* and implementing it in software components that can be integrated into instruments to provide access to all modes of collaboration

---

[2] https://bf-collab.net/

2. Designing mixed-reality interfaces to reveal these modes during spontaneous musical sessions with heterogeneous instruments, in order to enrich communication and cooperation
3. Evaluating the impact of these interfaces on musical practice

2.1.1 The BOEUF framework

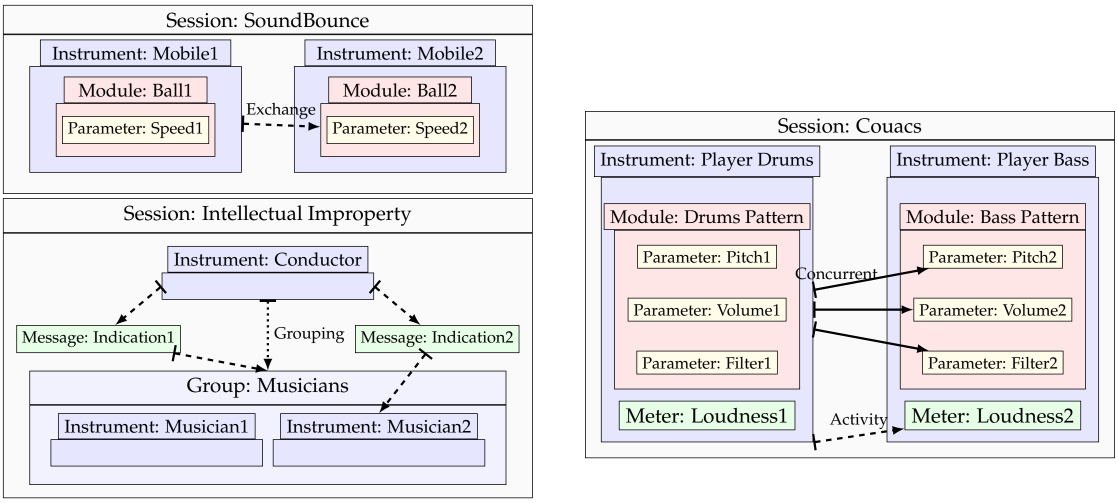

FIGURE 2.1: Three ensembles analysed with BOEUF. Dashed lines show communication modes, dotted lines organisation modes and solid lines cooperation modes.

We started by formalizing *Collaboration Modes* in *DOs* through a conceptual framework, named BOEUF (for BOEUF OrchEstra Unification Framework) [26] The framework was built from literature review and interviews with musicians involved in digital ensembles. It describes all the ways that musicians collaborate, either through the digital instruments (mediated collaboration) or through physical actions (non-mediated collaboration). The framework is organised in three categories, themselves subdivided in subcategories.

Cooperation modes comprise actions that have a direct effect on the produced music. They can be *independent* when they do not affect the musical output the same instrument, *complementary* when they affect the same musical output but not the same input (*e.g.* sensor or sound synthesis parameter) or *concurrent* when they affect the same sensor or parameter and generate potential conflicts. This classification inherits from research on Collaborative Virtual Environments (*CVE*) [126]. *Communication modes* influence the actions of musicians rather than the produced music. Communication modes encompass *Awareness* modes, which are non-intentional and help understand musicians' activity. Awareness is a prominent issue in *CVE* and HCI [16,80] but has also been discussed in the case of *DOs* [84]. Other communication actions are *Indications*, such as text messages, intentional communicative gestures or demonstration of intentions. Finally, *exchanges* comprise transfers of data between musicians. The last category of collaboration modes is *Organisation modes* which do not have any effect on the music but impact the other modes of collaborations. They consist in *Nomination* of various roles in the orchestra, *Grouping* of musicians and *Selection* of musicians to interact with.

Together with these collaboration modes, we defined a set of components that allow for their implementation in spontaneous *DOs*. An example representation of *DOs* with the corresponding components and modes is given in Figure 2.1. Our goal is then to provide implementations of these components for the main programming languages for instrument design and the most common music software.

A first one was developed for the PureData visual programming language : *bf-pd*. It was published in 2017 [70] and has been made available to the public [3] under a free open-source licence. This first implementation provides a set of externals that, when added to an instrument, automatically connect the

---

[3] https://gitlab.cristal.univ-lille.fr/boeuf/bf-pd

instrument to a musical *session* over local network, declare the structure and activity of the instrument and handles access and interconnection with other instruments in the session.

2.1.2 Revealing cooperation opportunities

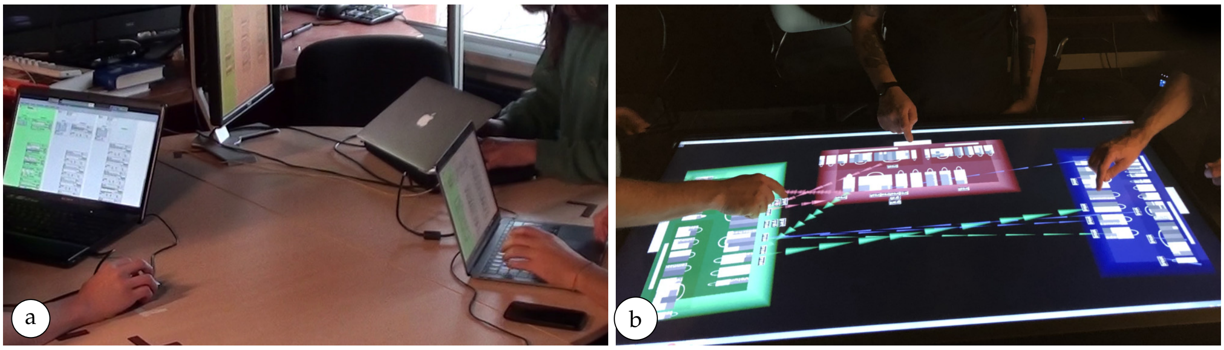

FIGURE 2.2: Interfaces for cooperation in digital orchestras : a) Separate interfaces with a bus paradigm, b) Shared space with a graph paradigm

We then started looking at how to implement modes of collaboration in the context of spontaneous orchestras with heterogeneous instruments. In a first study, published at NIME 2020 [27], we compared two designs of collaboration interfaces that reveal the structure and controls of instruments in the orchestra and enable *awareness* along with *complementary* and *concurrent* modes of cooperation. The two interfaces rely on bf-pd to interconnect three different instruments. This means that musicians can use the output (*e.g.* produced notes or onsets) or parameters variations (*e.g.* changes in effects) of other instrument to modify their instrument, or in the other direction influence others' parameters.

As shown on Figure 2.2, the first design provides for each musician an interface visible only to them which represents all the instruments of the orchestra, *i.e.*, their structure (sound modules and parameters) and their activity. All the collaboration and manipulation of the instruments are performed in PureData patches, therefore with the touchpad/mouse and pointer and graphical buttons and slider. Connections are made following a *bus* paradigm, by assigning parameters to channels to send/receive data, and are not visible to all. The second design provides shared interface, with a *graph* paradigm and manipulated using a large touchscreen. Each musician has a private space to interact with both their and the other instruments, and connections are visible to all. Research on Computer Supported Cooperative Work (CSCW) and on *DOs* has highlighted the importance of private and public spaces of collaboration [52,55,84,130]. In our case, the goal was to evaluate both the paradigm and the shared aspect of the collaboration interface to understand how cooperation modes would be impacted.

The study was performed with two groups of three musicians. Each musician first designed their instrument using PureData, with some constraints on the number of parameters and outputs, in order to obtain different instruments for all. They then performed multiple jam sessions with both designs. We ran a thematic analysis of interviews conducted with each group after they have spent multiple sessions playing with each interface. From these, we deduced insights and guidelines for the design of collaboration interfaces for heterogeneous orchestras. For instance, collaboration interfaces should :

- provide default outputs for each instruments to facilitate interconnection
- encourage complementary and concurrent cooperation with controls shared by default,
- include expressive collaboration controls, not simply graphical faders or buttons to interact with shared parameters
- visualise all connections to reveal the cooperation between musicians

These results now need to be extended by studying instruments with diverse interfaces (here all three instruments had the same interface) to understand how collaboration controls can be mixed with individual controls.

2.1.3 Augmenting awareness

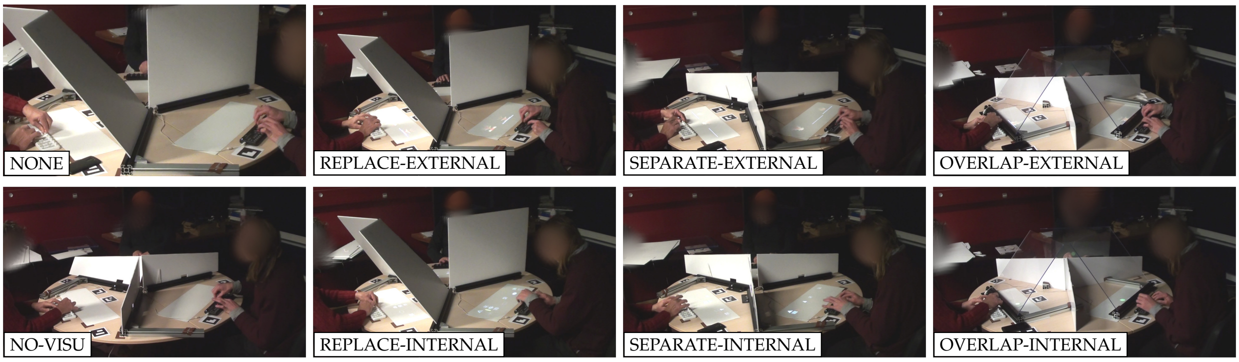

FIGURE 2.3: Studying the *Awareness* communication modes in digital orchestras with various situational visibility and visualisation levels conditions.

For the second study, published at NIME 2022 [28], we focused on the *awareness* modes of communication. We wanted to investigate which representation of instrument activity and which method of display would best help musicians understand what the contribution of others. Previous studies have shown the importance of visualising the contribution of other musicians in an orchestra [132]. Here we tested two *visualisation levels* : *external*, where only the spectrum of the audio output of each instrument was displayed, and *internal*, where the activity and manipulation of the various components (*e.g.* sounds/loops) are represented. These levels were displayed according to 5 situational visibility conditions, *i.e.*, how they were visible with respect to the other musicians : either not shown, replacing the physical musicians, shown separately (projected near the instruments), shown overlapped with the others using an optical combiner (*e.g.*, semi-transparent mirror). All resulting conditions are shown in Figure 2.3.

5 groups of 3 musicians took part in the experiment. Paradoxically, contrary to the context of heterogeneous instruments, we chose to give the same instrument (composed of melodic, rhythm and noise tracks) to all musicians, in order to increase the confusion around the contribution of each instrument, and used improvisation tasks in order to encourage collaboration. Through questionnaires and interviews, we extracted insights on what should be revealed and how.

In particular, our results suggest that, in order to be effective in increasing awareness and facilitating collaboration, mixed-reality visualisations of instruments should display the internal activity of other instruments and that they should be shown grouped and close to each user's instrument. These results contradict our original hypothesis that musicians would prefer seeing the other musicians and their musical activity overlapped. Interestingly, we also saw a shift in focus of musical practice due to the visualisations, from active listening to others to a more *cerebral* or internalised process. This brings us back to the idea that the choice of revealing expressive interactions should be left to users, here to collaborators. The study now needs to be extended to address heterogeneous instruments, *i.e.*, with different internal structure and activity, which will require the design of unified visual representations.

2.2 Scenography of Immersive Instruments

The second research direction on *Revealing Expressive Interactions* was conducted with my colleagues Victor Zappi (Northeastern University, USA) and Dario Mazzanti (IIT, Italy), following research in our respective PhD theses. We published in 2014 a paper at the Sonic Interactions in Virtual Environments workshop at IEEE Virtual Reality [36], recently extended as a book chapter [172]. In this project, we focus on what we called *Immersive Virtual Musical Instruments (IVMIs)*, *i.e.*, instruments that use virtual components visualised with virtual or mixed-reality displays.

FIGURE 2.4: Scenography of Immersive Instruments : a) Dimension Space, b) Scenography with separate stereoscopic screens for the audience and performers, c) Scenography with a semi-transparent screen that reduces transportation but increases visibility

IVMIs have a long history in scientific research, instrument design and performance, with early work such as Lanier's The Sound of One Hand [119] and many instruments since [154]. With the democratisation of commercial VR equipment, there has been an increase in the number of immersive performances, with a variety of scenographies, *i.e.*, arrangements of the virtual and physical spaces. Indeed, some of them require all spectators and performers to be immersed with VR headsets, others rely on a large stereoscopic screen shared by audience and performers, others provide mixed-reality video see-through tablets for the audience to perceive the virtual content placed on a physical stage. They therefore affect aspects such as how much spectators and performers are transported [17] to the virtual space, how much they can see each other, what range of interaction and collaboration possibilities they open. However, there has been little research on what impact these scenographical choices have : How do they influence the performer and audience experience ? What virtual and physical components of the instruments the should be revealed to the audience ? How much of the performers should the audience be able to see and vice-versa ?

Our goal was to understand the different dimensions of immersive scenographies and how the affect the users experience. We designed a dimension space, inspired by [41], for scenographies of IVMIs. It can be seen in Figure 2.4.a. It was built by analysing immersive performances in and outside the scientific literature, including from our own experience as designers of *IVMIs*, and by looking at important design guidelines for both digital musical instruments and virtual environments. It describes scenographies with 6 dimensions, which address the audience experience (*spectators transportation*, *spectators awareness*), the performers' experience (*performers transportation*, *spectators visibility*) and dimensions of performance design (*interaction spectrum*, *ensemble potential*). For example, *spectators transportation*, relying on the transportation dimension of shared spaces by Benford et al. [20], describes how much the audience is surrounded by the virtual components of the performance, from a few virtual objects visible in mixed-reality on the physical stage, to being inside a virtual environment, completely cutoff from the physical space.

The dimension space is not designed for providing absolute measurements of these dimensions, but rather for discussing scenographies (even for a same instrument) and for encouraging novel designs. In fact, we used it to analyse a number of performances, such as the one in Figure 2.4.b, discussing their relative advantages and limitations. We also showed that the dimension space could indicate new design directions for existing scenographies, such as the use of optical combiners in addition to stereoscopic screens, as exemplified with the design shown in Figure 2.4.c.

Research on this topic now continues with the development of a framework for prototyping immersive instruments and performances [4] and the study of the impact of the dimension on aspects of the users

---

[4] The IVMI-builder framework combines the Godot game engine and PureData https://gitlab.univ-lille.fr/ivmi/ivmi-builder

experience such as presence or co-presence [111].

2.3 Augmenting the Audience Experience

This section of my research started in 2013 in the Potioc team at INRIA, it was then the focus of the IXMI project that I led at the University of Bristol, and finally became the center of Olivier Capra's PhD thesis [57].

It began with the analysis (from personal experience, discussions with electronic and acoustic musicians, and literature review) that the spectator experience with *Digital Musical Instruments (DMIs)* suffers from :

- small/hidden gestures and sensors
- complex mappings between sensors and musical parameters
- complex and partly autonomous musical processes

Compared with acoustic instruments, for which gestures transmit many information [71] and where there is a physical link between the gestures and sound, it is therefore very difficult for the audience to understand the relation between the musician's actions and the resulting music and to perceive the actual engagement of the musician in the performance. This is a common analysis in the field of New Interfaces for Musical Expression [97,152], to which solutions have been proposed which imply constraining the instrument design, for example with transparent (or familiar) interaction metaphors [83], more visible interfaces, or amplified gestures [38].

Instead, our first idea was to employ augmented reality displays in order to visually enrich performances with information about the mechanisms of *DMIs* without placing constraints on their design, thus preserving the musicians freedom and expression.

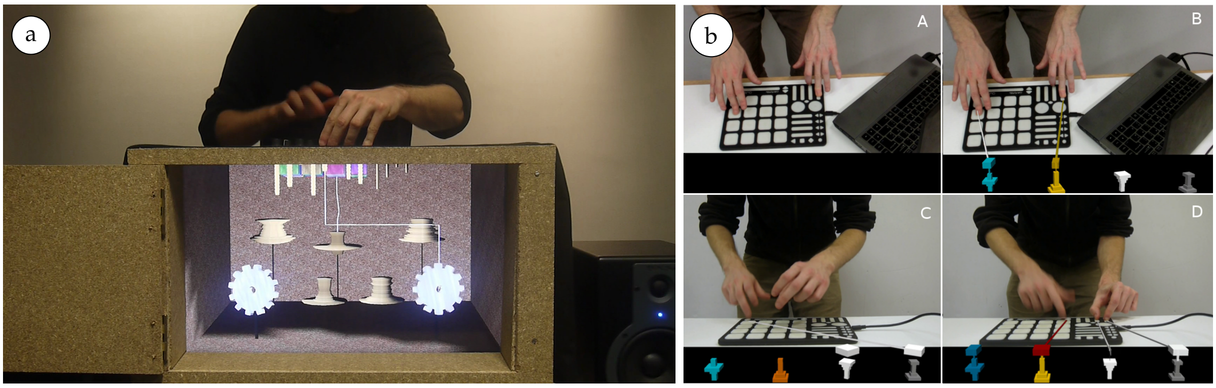

FIGURE 2.5: a) First implementation of visual augmentations for the audience following the Rouages approach using a monoscopic display with head-tracking. b) Some of the experimental conditions to study attributed agency with augmentations overlaid on videos of short performances

We first proposed the *Rouages* approach, in a paper published at NIME 2013 [33], shown in Figure 2.5.a. It consists in augmenting physical instruments with virtual elements which inform the three common levels in *DMIs* structure :

- Virtual sensors that amplify small gestures and sensors
- Simplified representations of the sound processes, which highlight their activity and level of autonomy
- Virtual links between sensors and sound processes that show the gesture to sound mappings

After a first implementation, this approach led us to explore three main questions: What are the components of the spectator experience ? How are they altered in performances with *DMIs* ? Can we enrich this experience by providing information on the mechanisms, and if so what quantity and type of information should be given ?

2.3.1 Attributed Agency, Subjective and Objective Comprehension

During the IXMI project, we started envisioning the experience of spectators through the lens of causality perception. We drew inspiration from the work on self agency, in particular that of Wegner and Wheatley [167] on *Apparent mental causation*, which propose a model to explain the perception by one individual of their own agency, *i.e.*, “I was responsible for this particular action or event”. According to them, this self-agency happens a posteriori by unconsciously comparing the intention (cause) and action (result) on three criteria :

1. *Exclusivity* : there should only be one cause for the result
2. *Consistency* : the nature of the cause should match the nature of the result
3. *Priority* : the cause should closely precede the result

When looking at the perception of interaction with *DMIs*, it becomes clear that the criteria are all broken :

1. *Exclusivity* : perceived changes in the music can come either from the musician’s direct actions or from automated and sequenced events, sometimes from a mix of both, making it difficult for the audience to perceive what was their cause
2. *Consistency* : gestures with a certain type and scale are not necessarily mapped to changes with the same type and scale, *e.g.*, a soft discrete press of a button can lead to a strong continuous change in the sound
3. *Priority* : changes in the music can happen with a delay, *e.g.*, at the next beat or bar, after the gesture was performed

We therefore decided to transpose Wegner and Wheatley’s model to what can be named *attributed agency*. It relies on the same criteria but applied to the perception of the results of someone else’s actions.

In a first paper published at NIME 2015 [25], we designed instruments that would each amplify the disruption of one the criteria. We then used visual augmentations, following the *Rouages* approach in order to compensate for the disruptions. We recorded short performances with the instruments, shown with and without the augmentations (see Figure 2.5.b) to participants of a controlled in-the-lab experiment. Our results suggest that compensating for the exclusivity and consistency criteria led participants to assign a higher agency to the musician in the case of exclusivity and gave them a higher confidence in rating agency in the case of consistency. These first results confirmed our intuition that *attributed agency* was a useful theoretical framework for assessing the experience of the audience with digital interactions, but also that visual augmentations were an essential tool for studying this experience and also enriching it.

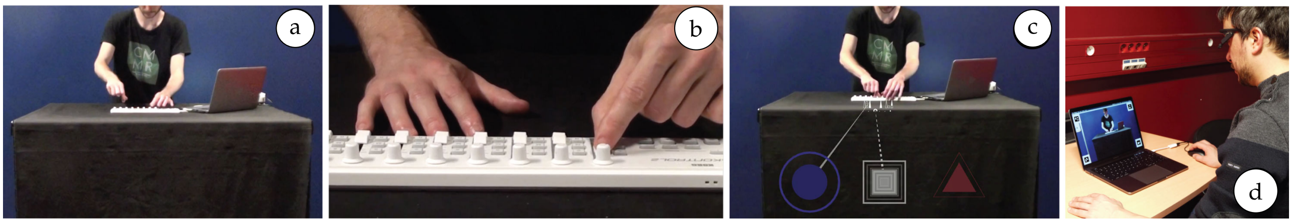

FIGURE 2.6: Conditions (a) *Control*, b) *Explain*, c) *V-AUG*) and d) experimental setup used to evaluate subjective and objective comprehension of DMIs

The notion of attributed agency was then refined during Olivier Capra’s thesis. In a paper published at ACM DIS 2020 [59], we proposed to study the audience experience by comparing subjective comprehension, *i.e.*, how much the audience believe they understand the musician’s actions, and objective comprehension, *i.e.*, how much they actually understand it. In order to obtain their subjective comprehension, we relied on a set of design challenges first proposed by Bellotti *et al.* [14], and adapted to the audience experience by Fyans *et al.* [89]. For the objective comprehension, we asked participants to identify in short sequences who from the musician or automated processes was responsible for a particular change in the sound, giving us an accurate measurement of what they really understood. In the study, we compared conditions with augmentations that follow the *Rouages* approach (*V-AUG*), with augmentations that on the contrary show uncorrelated activity (*V-DIS*), with preliminary explanations of the instrument (*Explain*) and without additional information (*Control*). The stimuli and experimental setup are shown in Figure 2.6.

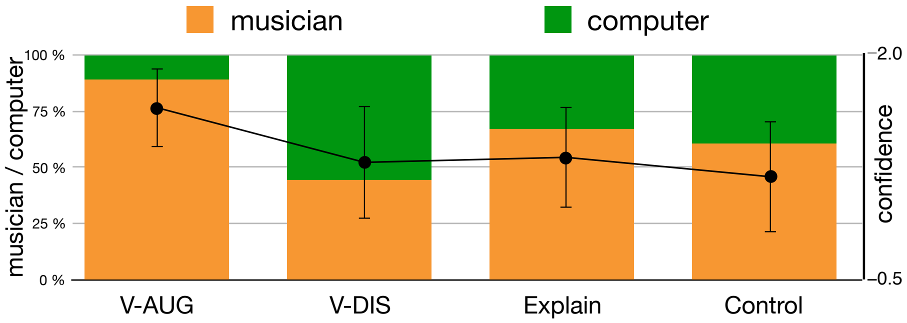

FIGURE 2.7: Results from the AVAA study: the estimated musician contribution is higher with visual augmentations (V-AUG), than when explaining the instrument (Explain), or providing no information (Control). It is lower when disruptive visual cues are given (V-DIS).

Our results suggest that visual augmentations, although they do not necessarily increase objective comprehension, have an effect on subjective comprehension. Furthermore, as shown in Figure 2.7, the attribution of agency to the musician instead of the computer is higher when adding augmentations (*V-AUG*) than without (*Control*), and so is the participant's confidence. Visual augmentations therefore increase the trust in musician's engagement in the performance, even when their contribution is limited.

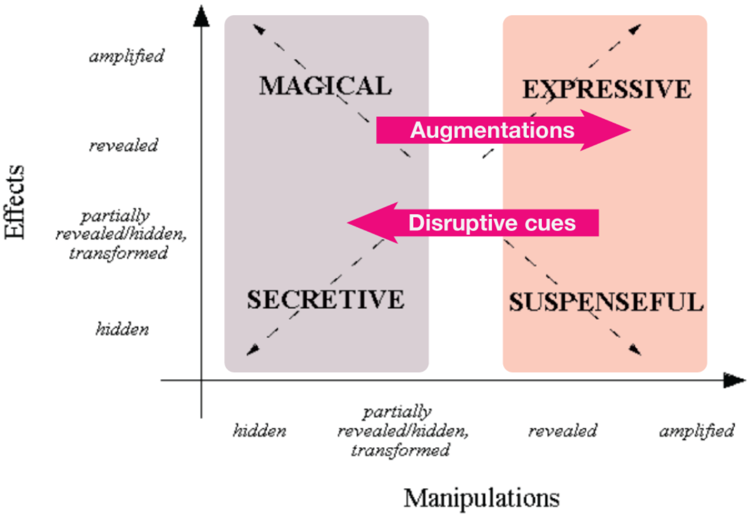

FIGURE 2.8: Visual augmentations can either reveal or disrupt the link between gestures (manipulations) and resulting changes in the sound (effects), leading to a novel opportunity for designing the spectator experience.

Interestingly, visual augmentations, when they are not correlated with the instrument activity (*V-DIS*), *i.e.*, when they show random events, seem to lead spectators to attribute less agency to the musician. They can therefore be used to induce more magical experience, as shown in Figure 2.8 (adapted from [144]), for the audience, providing a novel tool for artists to play with the audience experience.

2.3.2 Spectator Experience Augmentation Techniques

In parallel, we investigated the different ways that can be used to enrich the audience experience with additional information without altering the design of performing tools. We named them *Spectator Experience Augmentation Techniques (SEATs)*. They can be described as techniques that enrich the audience experience without requiring changes in design of the expressive interface itself, or influencing the content of the performance (contrary to participatory performances).

In a short paper published at NIME 2020 [61], we proposed a taxonomy of *SEATs*, enabling their analysis and design, based on 11 dimensions : *Spatial and Temporal Alignment* describe how the augmentations overlap with the performance; *Temporal and Spatial density* describe the quantity of available additional information; *Spatial and temporal control* define over if the presentation of the information can be modified by the audience; *Nature and Modality of presentation* describe how the information is displayed; *Content nature and reactivity* define what information is provided and if it is updated in real-time or not; *Agents* define if the augmentations take into account multiple performers.

In the study presented above [59], we had in fact been comparing *SEATs*, in this case preliminary auditory explanations and real-time visual augmentations. While the results suggest an advantage of using visual augmentations, it is still not clear which aspect of these augmentations is the most efficient in increasing subjective comprehension. This highlights the need for more study of the impact of each of the taxonomy dimensions.

During Olivier Capra's thesis, we also proposed a conceptual pipeline which could be implemented in order to integrate *SEATs* in an efficient manner in performances [58]. This pipeline introduces a new role within the performance ecosystem, that of the *augmenter*, which would consist in acting as an intermediary between performers and the audience, playing with the quantity and nature of the provided information, and therefore influencing the audience experience as a novel dimension of musical performances.

2.3.3 Revealing Levels of detail

In order to implement the *Spatial Control* dimension of *SEATs*, and therefore allow spectators to adjust the quantity of displayed information, we need 1) to understand what impact the quantity of information has, in particular on members of the audience with diverse levels of expertise and 2) to provide interfaces for the audience to adjust the augmentations.

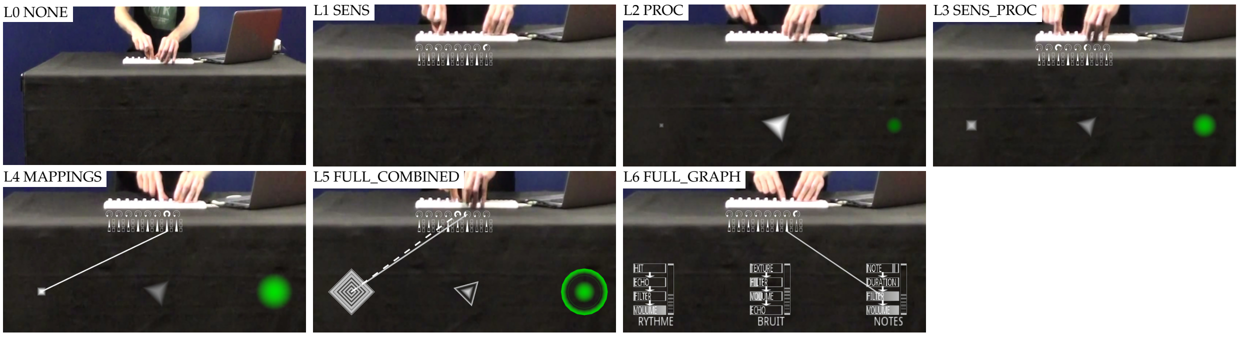

FIGURE 2.9: Investigated Level of details for Visual Augmentations here on a DMI with a control surface and three sound processes (rhythm, bass and noise sequences with various audio effects)

This led us to study the visual augmentations with a controllable level-of-detail. We published a paper at NIME 2020 [60] which provides first insights on this matter. It was then selected for publication as an extended version in the Computer Music Journal [62]. Starting from the *Rouages* approach, we proposed that each component of the *DMIs* can be visualised in more or less detail. For instance, augmentations of the sensors and gestures may : not be shown at all, display only if a sensor is active or not, display the value of the sensor, display the value and type of sensor … Similarly, each component can be displayed or not : only the sensors but not the sound processes, only the sound processes …

As shown on Figure 2.9, out of all possible combinations, we selected 7 global levels of detail (*LODs*) in visual augmentations, which range from no added information at all, to a full representation of the type and value of each sensor and of the structure, type and value of each parameter of each sound process.

In a controlled experiment, we then evaluated during a first block both objective and subjective comprehension of filmed short performances with a group of novices and a group of experts (electronic musicians). To that extent we relied on the same methods (design challenges and contribution assessment) described above. In a second block, we let participants from the same two groups select in real-time during the short performances the *LOD* that they preferred and then the *LOD* that helped them understand the best. We also interviewed them and analysed their LOD selection strategy.

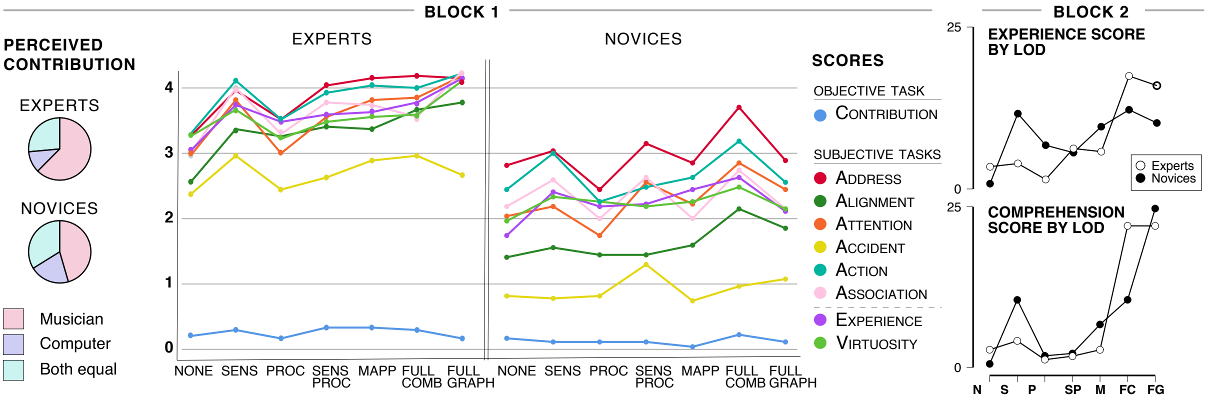

FIGURE 2.10: Block 1 : Evaluation of perceived contributions by expertise and effect on objective and subjective comprehension of diverse LODs of visual augmentations. Block 2 : Chosen LODs when asked for best experience and best comprehension

Our results, depicted in Figure 2.10 suggest that there is a clear difference in perception between experts and novices, the former generally perceiving a higher contribution of the musician (and higher than it actually is), and reporting generally a higher subjective comprehension even though their objective comprehension was not much higher than novices. Although *LODs* had a statistically significant effect on subjective comprehension, they did not seem to impact novices and experts equally. Some *LODs* provided lower subjective comprehension, such as displaying only the activity of sound processes (a configuration usually found in visuals accompanying musical performances). The *LOD* that provided all information but in an abstract manner (*FULL COMB*) was the most efficient in particular for novices, while the *LOD* that provided too much detail led them to a lower subjective comprehension. Finally, when participants were allowed to choose their *LOD* and asked how they would use that opportunity during performances, a few clear strategies emerged, such as 1) displaying all details at first to generate a clear mental model of the instrument before turning almost everything off to enjoy the performance, 2) keeping only the sensors visualisation to focus on the musician's gestures, 3) adapting the LOD depending on musical complexity. However, interpersonal variations within these strategies call again for a way for participants to personalise the information given by the visual augmentations.

In order to enable these very diverse strategies of LOD selection by spectators, we proposed implementations based on mixed-reality displays. As shown in Figure 2.11 we designed three configurations for displaying visual augmentations with controllable *LODs*. Individual displays using mixed-reality headsets or mobile devices enable each spectator to individually set their preferred *LOD*. Shared displays with augmentations overlaid on top of the filmed performance can be placed on or around the stage, visible to all. In that case, various *LODs* can be displayed on the screens, set either by votes from the audience, by the performers, or by an *augmenter*. Finally, grouped AR displays, for example with optical combiners at various angles as shown on Figure 2.11.c can be set to provide

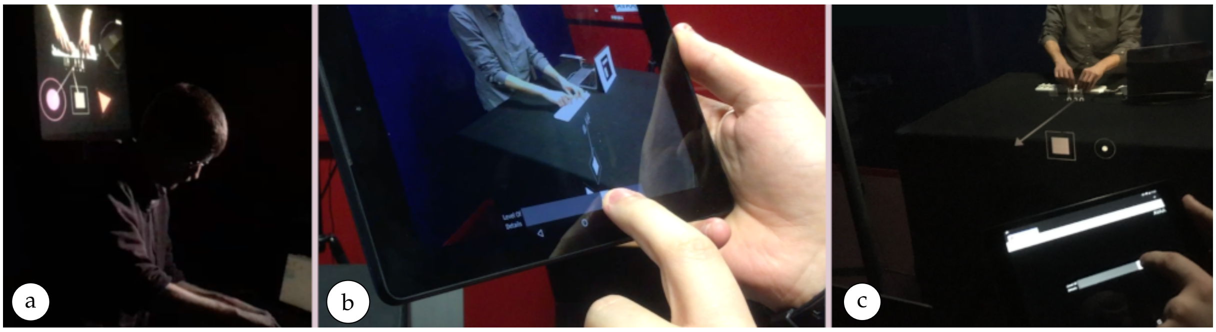

FIGURE 2.11: Implementation of visual augmentations with controlled Level-Of-Detail : a) Overlaid on a projection of the performance b) Individual mobile mixed-reality display with LOD selection c) Shared optical mixed-reality display with LOD selection

different *LODs* depending on the spectators position within the audience, and again changed dynamically.

2.4 Conclusion and Future directions for revealing interactions

Throughout this chapter, I have shown how visual augmentations in mixed-reality, when carefully designed, can enrich the experience of collaborators and spectators of performances with Digital Musical Instruments, improving their subjective comprehension, their trust in performers' engagement, and facilitating cooperation. Our results also demonstrate the importance of scenography in mixed and virtual reality contexts, in particular what information is revealed and how. This research finally highlights the diversity of the strategies used by both spectators and artists in their appreciation of performances, which should be taken into account by providing them control over the augmentations. These are important steps in understanding the perception of expressive interactions, which can be used to inform the design process.

I believe that there are two essential questions, shared between contexts of musical collaboration and audience experience that require further investigation. The first is the design of unified models and representations of *DMIs*. In order to implement visual augmentations that display the activity and structure of instruments, that facilitate communication and cooperation in heterogeneous *DOs* and understanding by the audience, it is essential to be able to extract sufficient information from the software, to organise this information in a structured manner and to display it in an homogeneous manner. This requires the investigation of : 1) models to describe instruments and other expressive tools and software components to extract relevant information from common music software and pass it to visualisation software; 2) Unified representations of the activity and structure of heterogeneous instruments, in particular using a level-of-detail approach that will facilitate the representation of complex instruments, but also adapt to various types of instruments (gestural controllers, control surfaces, virtual instruments . . .), as explained in Section 4.2.

The second is the refinement of our understanding of the perception of expressive digital interactions. We are following up on preliminary work from Olivier Capra's thesis on the use of physiological sensors such as the grip force and on the study of motor simulation [107], in order to retrieve dynamic and objective measurements. This constitutes one of my main future research directions, as described in Section 4.1. We also envision studying the influence of other dimensions of performances on the audience experience, such as the aesthetic of the augmentations (the *content nature* of our taxonomy) and the transportation of spectators in immersive performances.

I strongly believe that these results also apply to other digital tools in public contexts. Their transparency could then be increase in order to facilitate cooperation or support trajectories [18]. But they could also be made more "magical" or opaque for security or anonymity reasons.

3 Revealed Interfaces

The research described in the previous chapter focused on the roles of collaborators and spectators of digital interactions. In this chapter, I present the research that I have been conducting on the main users in contexts of content exploration, *e.g.*, selecting content in design processes or discovering information in museum exhibition, and of musical expression.

Expression is an aspect often left out during the design of 3D interaction techniques and devices and of 3D displays. Generic 3D interactions and visual augmentations often focus on efficiency in time and error with design guidelines that favor reducing complexity/dimensionality to reduce errors [44]. Similarly the research on 3D displays tend to aim for technological perfection [1], often leading to technical solutions that isolate users (*e.g.*, VR headsets). They tend to leave out or reduce human engagement in discovering / exploring virtual content, by providing all information at once in as best as possible quality. On the contrary, here my focus is on expression and exploration capabilities of 3D interfaces, *i.e.*, on designing interaction / visualisation techniques that are not always more efficient than the state-of-the-art but which might trigger changes in perception, practice and leave room for user appropriation.

Following the *Revealing Interactions* paradigm, these *3DUIs* take advantage of interactions with the physical space, revealing virtual controls or feedback within it. They span three main directions :

1. Augmented visual feedback to enrich the perception of gestural, tactile and tangible interaction
2. 3D Interaction techniques that reveal expressive opportunities in physical and virtual spaces
3. Novel mixed-reality displays that emphasize exploration and expression

3.1 Augmented Visual Feedback

Travelling through the real-virtual continuum of Milgram and Kishino [133], physical interfaces can be complemented with virtual content, from a few virtual elements aligned with the physical space to a full environment in which users are transported. In this section, I present research in which we investigated the use of virtual content, at various degrees of immersion, as a way to enrich or alter the perception of users during interaction with gestural, tactile and tangible interfaces.

In particular, we looked at how visual augmentations can benefit users of tangible control surfaces, *e.g.*, with physical buttons, sliders and so on, in the appropriation of their tool. We also showed that adding 3D visual feedback, in our case stereoscopic rendering, can influence tactile perception on touchscreens. We finally studied the impact of manipulating visual feedback, through changes in Control-Display Ratio, on the user experience in gestural musical instruments.

3.1.1 Augmented Visual Feedback on Control Surfaces

As pointed out in Section 1, expert users in creative contexts such as music production and performance, graphics design, video editing and visual performance, are given the opportunity to customize their tools, leading to a strong diversity in setups and *improving core task performance* [85] For example, they can choose the interface (usually commercial or custom control surfaces) [93], the software

---

[1] in a way this is similar to photorealistic rendering which has long been the goal of research on virtual reality

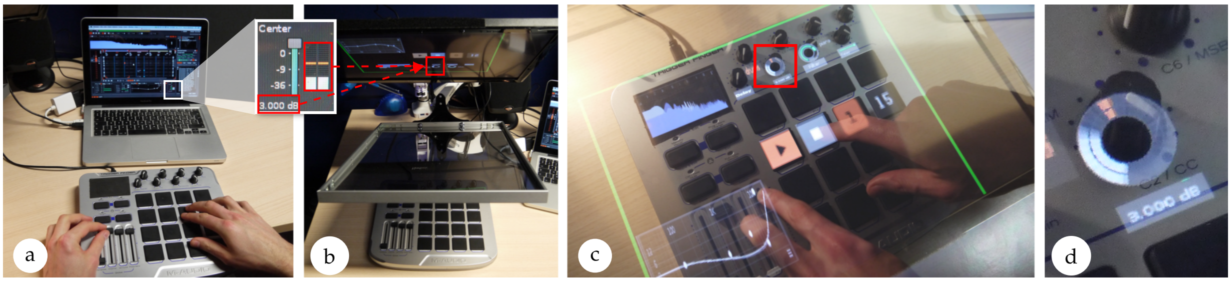

FIGURE 3.1: ControllAR principle : a) Graphical elements from creation / performance software. b) They are rearranged on a separate screen placed above an optical combiner. c) They appear overlapped with the control surface. d) They are used to reveal information about sensors and parameters

they use, the plugins for synthesis and processing of content, and the mappings between sensors and controls over the software.

One aspect that is usually not easily modifiable is the visual feedback on the structure or status of the content and controls that they are manipulating. This feedback is usually either too complex and mostly fixed on the software graphical user interface (*GUI*) or displayed on the control surface with LEDs or small screens but very limited, fixed and/or very different from the main software representations. It is usually not possible for the users to select and place visual feedback where and when needed during interactions. Consequently, they tend to either avoid looking at their screen and rely on limited cues for interacting, therefore missing useful information.

In a paper published at ACM ISS 2016 [32], we studied the combination of graphical interface remixing [49,159] and mixed-reality displays. As shown on Figure 3.1, the *ControllAR* software (available under a free open-source licence[2]) allows users to grab visual elements from their main software *GUI*, modify their shape, colors and visibility, and place them freely on their control surface with the help of an optical combiner. It relies on real-time capture of parts of a desktop window, which are then transformed, and rendered on a separate screen. The screen is placed above a semi-transparent mirror, in a fish-tank setup [99,166], below which the user places their control surface. To them, remixed graphical elements appear overlapped with the various sensors and provide selected information. The control surface behaves as a mixed object [68] combining physical and digital properties.

We ran a qualitative study with 10 expert users (electronic musicians) in order to understand what visual feedback they would appropriate if given the choice. From the sessions, we were able to understand that the information that they missed (and were able to retrieve using *ControllAR*) originates from three parts of their interactive system, as depicted in Figure 3.2 :

1. the *mappings*, for which they extracted labels and colors identifying to which parameters the sensors were assigned (see Figure 3.2.a,b,c)
2. the *parameters*, for which they displayed the exact values with units, and the context, *e.g.*, the modified curve or position within a sound (see Figure 3.2.d,e,f)
3. the *processes*, meaning the status or activity of the system *e.g.*, current tempo, position within timeline, vu-meter . . .

Comments from the users were also very positive, with participants stating that they felt that the system brought back information that they had lost with their usual setup, and pointing out that the tool could also be used to provide information to other members in an orchestra, which brings us back to the importance of the *BOEUF* framework. Finally, the study led us to suggest three guidelines for interactive systems designers to facilitate the appropriation by users : 1) Facilitate remixing graphical

---

[2] https://gitlab.univ-lille.fr/mint/controllar

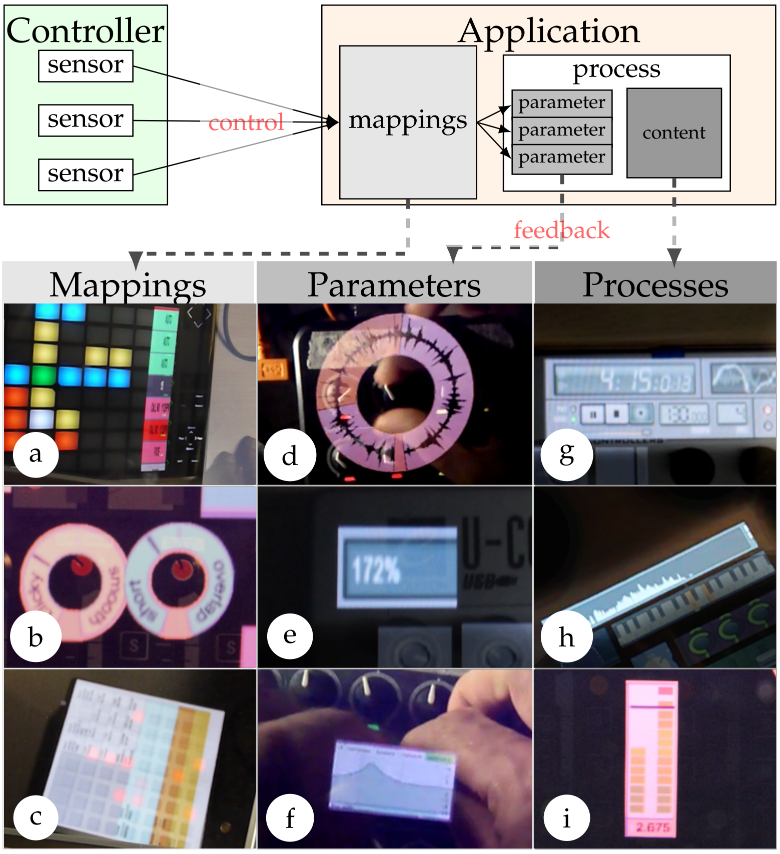

FIGURE 3.2: ControllAR design space extracted from the qualitative study with expert musicians. It shows from which parts of the DMIs the information chosen by participants originated, and the diversity of augmentations designed.

user interfaces (detaching / reorganising visual elements); 2) Correctly expose the mappings, parameters, activity and structure of the application; 3) Facilitate augmentation of hardware controllers by leaving room for integrating visual elements.

3.1.2 Augmented Visio-Tactile Texture Perception

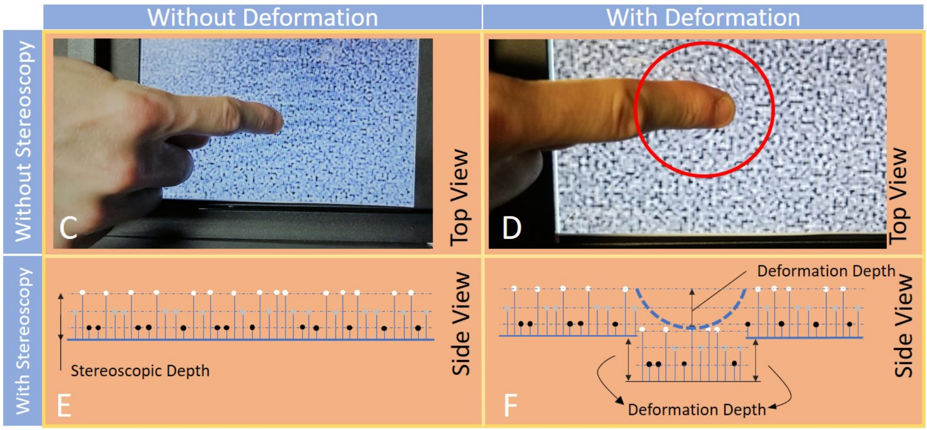

FIGURE 3.3: Experimental conditions for the study of cross-modal visual/haptic perception of virtual texture roughness

Beyond enabling custom visual feedback for expert interaction, 3D displays, by simulating virtual elements in the physical space, can be used to alter and enrich users' perception. For instance, the cross-modal interaction between haptic and visual perception [21,113,122,170] has been the subject of much research and has led to techniques such as pseudo-haptic feedback [121,161] or their combination for 3D manipulation [48] and texture perception [96] in virtual reality.

In the context of Detjon Brahimaj's PhD thesis, we investigated adding stereoscopic rendering in order to influence the tactile perception of roughness of a virtual texture. The tactile feedback that was used relies on ultrasonic vibrations to generate a squeeze-film effect [11], reducing the friction between the user's finger and a surface. By modulating the amplitude of these vibrations, and therefore the friction, according to the finger movement, one can create synthetic textures with controllable waveform and frequency. The study, presented at IHM 2023 [45], relies on previous categorization of tactile textures produced with ultrasonic friction reduction according to common adjectives [73], which allowed us to define sets of parameters that generate smooth and rough textures.

Our goal was to see if adding visual feedback that create a sense of visual depth to the tactile device could alter the perception roughness. The independent variables were the tactile texture type (Rough/Smooth), the presence of stereoscopy or not and the presence of a visual deformation below the finger. We chose to use gradient noise texture to generate stimuli with equivalent roughness but variable appearance. Noise period was chosen small enough so as to not give an impression of geometry, *i.e.*, which would have led participants to expect tactile ridges. Textures from the resulting conditions, shown in Figure 3.3, are composed of random dots that appear either all on the surface (without stereoscopy) or ranging from the surface to 5mm below it (with stereoscopy). When touching the surface, it either does not change or appears pressed. Across these conditions, participants were asked to independently evaluate tactile and visual roughness, along arbitrary scales to which they were accustomed.

Our results show a statistically significant effect of both deformation (confirming results on pseudo-haptic feedback) and stereoscopy on perceived tactile roughness, but only in the case of smooth tactile textures, which could show a preponderance of tactile feedback. In the other direction, increase in tactile roughness led to a higher perceived visual roughness. These results confirm that virtual visual feedback, when co-localised, can help enrich interaction with touch surfaces, amplifying the perceived roughness. They could for example be applied to the exploration of virtual reconstructions of physical objects or of rich volumetric data.

3.1.3 Augmenting Gestural Interaction in Virtual Reality

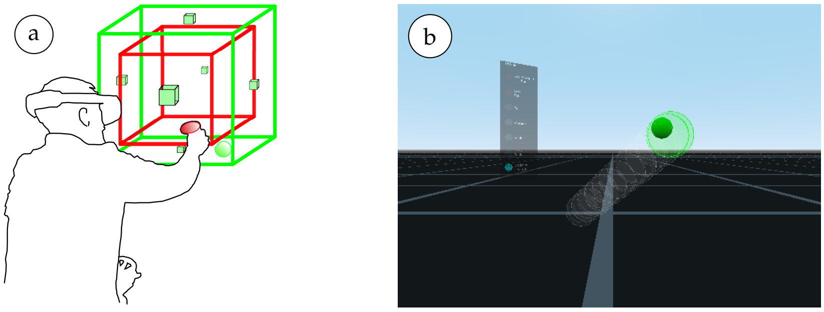

FIGURE 3.4: Studying how to augment gestural interaction in VR a) Influence of Control-Display ratio on user experience with IVMIs. b) Influence of auditory guide, gesture speed and regularity

When they are fully immersed in a virtual environment, *i.e.*, when wearing a *VR* head-mounted display, the perception and actions of users can be altered by playing with perceptual cues, in particular visual rendering. In the context of Vincent Reynaert's PhD thesis, we looked at how gestures performed in *VR* can be influenced by visual and auditory modalities, and what impact this has on the user experience, including the perceived muscular fatigue.

In the first study, published at Interact 2021 [149], we looked at the influence of gesture rhythm : speed, regularity and auditory guide. As shown in Figure 3.4.b, participants were asked to follow a target on a two dimensional trajectory, with varying speed, temporal regularity of the trajectory and the presence of an auditory guide that highlighted the speed of each segment of the trajectory. Our results indicated in particular that augmenting gestures with sound guides led to a stronger sense of presence and an increased level of perceived fatigue. This effect may be due to the fact that, contrary to studies on the effect of sound on endurance which employ complex music and tend to show a positive effect [158], our auditory guide highlights gestures and the user's effort, and act as a feed-forward mechanism [75].

The second study was published at NIME 2021 [148]. It focuses on the alteration of the perceived user's gestural amplitude by changing the Control-Display Ratio [[78];casiez2008impact], *i.e.*, the ratio between the movements that users see in the virtual space (their hand position was represented by a sphere) and the gestures they are actually performing in the physical space. Research has shown that some amount of discrepancy is possible without being noticed by the user, leading to the design of techniques for redirecting gestures and navigation [6,138]. In our case, as shown in Figure 3.4.a, the case study was an *IVMI* consisting in a control cube that allowed the user to mix 6 audio loops by moving a virtual pointer. The gain of each loop was mapped to the inverse of the distance to the pointer, and the overall gain mapped to the gesture speed. Figure 3.4.a shows in green the virtual control cube within which the user perceives their are moving their hand and in red the physical cube within which they are actually moving.

Our independent variables were the gestural sizes (boxes of 20cm, 40cm and 60cm) and visual sizes (boxes of 20cm and 60cm), leading to 6 experimental conditions with large and small visual boxes and physical gestures either not modified, amplified or reduced by the visual feedback. Our results suggest that a CDR that leads to gestures perceived larger than they actually are results in a stronger sense of presence, perhaps due to the increased perceived engagement. Similarly, changes in CDR can be used to increase the precision in virtual controls (with large gestures resulting regularity in small movements of the pointer), providing that the modification remains small, otherwise they lead to an increase in perceived control difficulty. Based on these results, we suggest the development of novel techniques for musical expression in virtual environments that would allow users to adjust their CDR freely and dynamically, to choose between more engagement/presence and more control accuracy.

Overall, the two studies demonstrate that visual and auditory augmentations can influence gestural interaction and that they can be used to enrich the user experience.

3.2 Interaction Techniques for Expression in Mixed-Reality

We have shown that added or altered visual feedback, in mixed and virtual reality, can enrich existing tangible, tactile and gestural interactions. But in the design of interaction techniques itself, in virtual and physical spaces, expressive opportunities can emerge when looking at the components of interaction differently.

In particular, I studied how generic 3D interaction techniques can be used and adapted for musical interaction, how the physical space is itself a rich source of material for expression and how combining virtual and physical spaces through a mixed-reality display and interaction technique can open opportunities for artistic expression in music, dance and theater.

3.2.1 Expressive 3D Interaction techniques

3D interaction techniques are often designed with efficiency in mind, *i.e.*, they should permit to accomplish tasks in the minimum amount of time and with as few errors as possible. In an expressive context however, efficiency is not necessarily the most important design factor and errors might be interesting and even desired. In an article published in the Journal of New Music Research in 2020 [22],

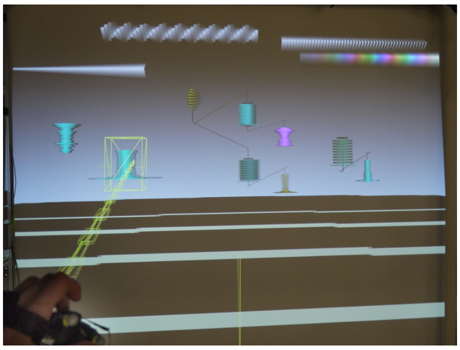

FIGURE 3.5: The Piivert 3D interaction technique, that combines graphical selection and multiple finger sound excitation with pressure sensors. Common 3D selection and manipulation techniques were modified to support expression.

I conducted a survey of how 3D interaction techniques have been used and could be used in a context of musical expression. Starting from commonly accepted categories of interaction, *i.e.*, selection / manipulation, navigation and application control, I showed how each can be integrated and extended with respect to fundamental notions of musical interactions such as instrumental gestures [54], control intimacy [169], input complexity and player freedom [108]. An example is given in Figure 3.5, with the Piivert interaction device and technique [31] that relies on 3D selection and manipulation for slower selection and modulation instrumental gestures [54] and on pressure sensors for fast excitation instrumental gestures. It therefore takes advantage of the interaction diversity and efficiency offered by common techniques while compensating their limitations (in this case for gestures that require very low latency and haptic feedback).

Another example can be found in the use of 3D navigation techniques. In the Versum instrument [10], the musician navigates freely in a virtual space composed of orbiting sound sources represented as 3D shapes, which constitutes an interactive composition. It therefore uses a *general movement* type of navigation. Extending it with other types of navigation techniques, such as targeted movements (*e.g.*, teleportation to a point of interest) or specified trajectory movements (*e.g.*, planning a path across sound sources) would then enrich expression capabilities by allowing for discrete or semi-automated changes in the sound.

I finally provided a number of perspectives for the design of expressive 3D interaction, such as selection and manipulation techniques for chords of virtual objects, dedicated input devices and (re-)introducing ambiguity in interaction. I plan on exploring this last aspect further as detailed below in Section 4.3.

3.2.2 Expressive exploration of the physical space

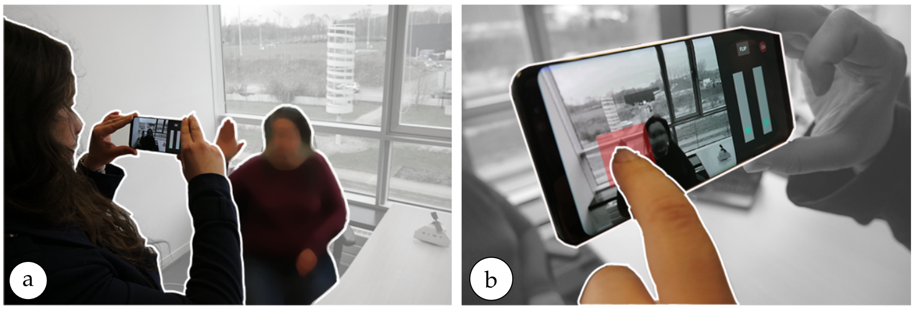

FIGURE 3.6: Phone with the Flow : a) Users select a scene in the physical space with interesting movements using the integrated camera b) Touching the screen isolate part of the scene and sonifies the optical flow within it

In the physical space also, interactions can be turned into expressive opportunities. In the context of Cagan Arlan's PhD thesis, we studied how movements in the physical space can be captured in real-time and used as controls for sound synthesis. In a paper published at NIME 2018 [5], we used optical flow analysis on mobile devices to capture rich information on movements within physical scenes. As seen in Figure 3.6, the user can select a target space by pointing the mobile camera at it and touching one or multiple zones on the resulting image. The optical flow is then computed within these zones, which results in a vector of amplitudes for each motion direction. Components of this vector can then be mapped to sound synthesis parameters.

We proposed a design space based on dimensions such as the source of movement, camera movement, use of the touchscreen, and mappings, which demonstrates the extent of opportunities for expression, and we evaluated musicians' feedback on the system, which reveals their strategies of expressive movement sonification.

3.2.3 Revealed virtual controls in spatial augmented reality

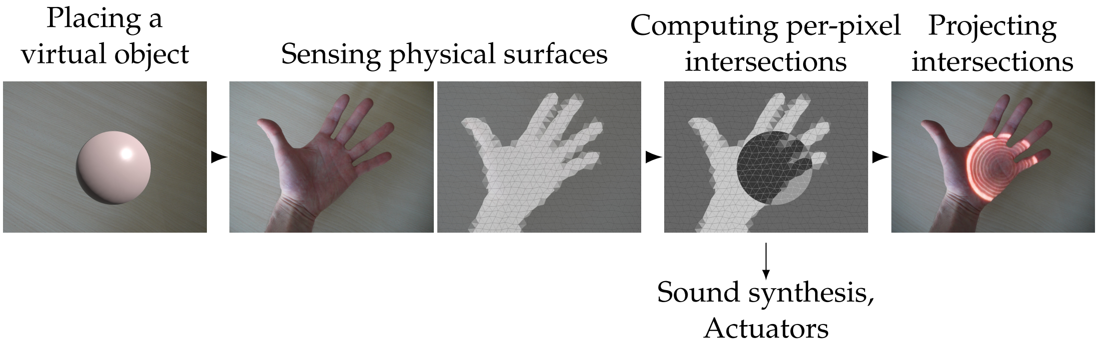

FIGURE 3.7: Pipeline for revealing virtual content with Rivill : Intersections between virtual objects and the physical space (captured with a depth camera) are computed on the GPU for each fragment and then re-projected in the physical space with volumetric content. Intersection properties (size, position) can be sent to other software.

In the context of Cagan Arslan's PhD thesis, we also started looking at integrating expressive virtual interactions within the physical space. This research was first presented at NIME 2017 [24] for which we investigated how to augment gestural musical instruments with additional expressive capabilities by placing virtual controls in the physical space, either at a fixed position, attached to the instrument or to the user's body. We were interested in studying how gestural interaction could be enriched by reintroducing intangible objects, which, like the tangible components of instruments, facilitate the emergence of gestures and playing techniques [98], but which would not constrain existing gestures thanks to their intangibility.

The proposed technique falls into the broad category of *Spatial Augmented Reality (SAR)* [37]. It relies on depth cameras or 3D tracking of physical objects and on one or several projectors. It consists in detecting intersections between captured surfaces of the physical space and virtual objects, and re-projecting these intersections on the physical surfaces, hence revealing slices of virtual objects with diverse internal content (text, images, videos, volumetric textures).

This approach, depicted in Figure 3.7, is implemented in a software named Rivill [3]. The intersection computing and rendering pipeline runs entirely on the GPU and consists in 2-3 passes : 1) meshes of the captured physical space are rendered to a *depth texture*; 2) all virtual objects are rendered from the projector and for each of their fragments, the aligned pixel in the *depth texture* is tested for intersection with the object (using signed distance functions for primitives and with an additional pass for more complex meshes) 3) if there is an intersection the color of the fragment is computed according to the object content, projected, and this intersection is output by aggregating intersection detail in a texture

---

[3] Released under a free open-source licence at https://gitlab.univ-lille.fr/mint/rivill

passed back to the CPU. Intersection information can then serve to control external processes such as sound synthesis or actuators, closely coupling them with pixel-level interaction with the virtual shapes.

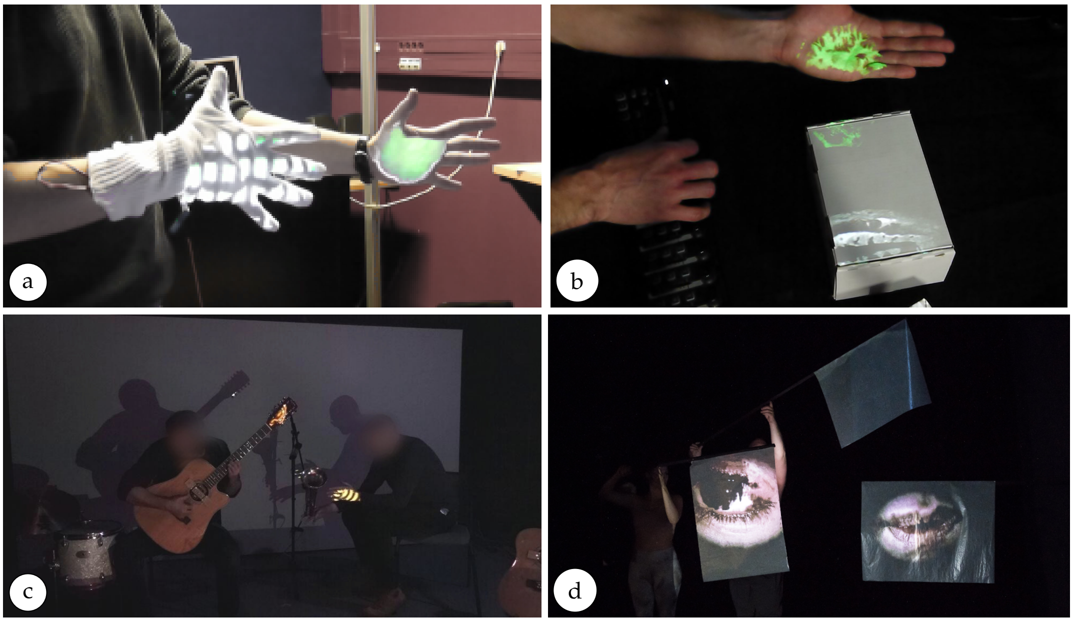

FIGURE 3.8: Projects using revealed virtual objects for expressive interaction. a) Revgest, b) Volume sequencer, c) Vibrating Shapes, d) Terres Rares

Our NIME 2017 paper describes a design space for the use of these virtual objects to augment gestural instruments : how they can be placed in the physical space, what type of control and feedback they enable, their visibility for audience and performers. We evaluate the system by reimplementing existing gestural instruments from the literature and demonstrating how they can be augmented with the revealed virtual objects. For example, Figure 3.8.a depicts a remake of the XthSense instrument [79]. Our remake relies on flexion and pressure sensors to retrieve finger movements (sensed with EMGs in the original version). Our system allowed us to extend these fine-grained finger movements with live-looping controls attached around the users' hand, providing another level of musical complexity, and allowing the musician to combine layers of musical control.

This first project however remained limited to discrete controls or continuous ones but with only a few dimensions. In order to increase the level of player freedom in the mappings [108] I studied how volumetric textures, *e.g.*, synthesised or captured with tomography, can be used as a parameter space. The results, published at NIME 2021 [23], are comprised of a design space that describes how the textures can be revealed and mapped to musical parameters, and an example instrument, the Volume Sequencer, depicted in Figure 3.8.b, that allows one to sequence the movements of volumes and reveal them with gestures and tangibles. A short study in VR demonstrated that dimensions such as the flexibility of intersection surface and amount of visual feedback on the intersection have an effect on perceived player freedom and agency.

Revealing virtual objects in the physical opens many opportunities for artistic expression, which led us to a number of collaborations with artists : the Embodied Sculptures dance performance [4] with choreographer Renaud Wiser and plastic artist Marie Lelouche, electronic music and dance workshops with the Otium collective in Lille in high schools and for the NAME festival, interactive installation at the Bordeaux Grand-Théâtre with Charles Petillon and Otium. A larger project is the collaboration with the Éolie Songe theatre company and the director Thierry Poquet for the Terres Rares cyber-opera [5]. The collaboration was funded with a Stimule project from the Région Hauts-de-France. During two

---

[4] http://embodiedsculpture.marielelouche.com/

[5] https://terrev.univ-lille.fr

years, with Cagan Arslan hired as a postdoctoral researcher on the project, we helped the company design the 3D interaction with virtual objects used in the third act of the opera. One sequence from this act can be seen in Figure 3.8.d, with two actors revealing animations by passing flags through virtual spheres. We followed the development of the opera during residences, three public performances in Lille and Grenoble, and a derived artistic installation presented at the Experimenta Digital Art Exhibition. This led us to start investigating the appropriation of our technology in these diverse artistic practices, in particular looking at design choices and the experience of expert users, using qualitative methods such as reflexive thematic analysis [46]. While we are still analysing data from the cyber-opera, we were able to study another use of our technology.

In a paper published at NIME 2022 [4], we report on the two years design of actuated acoustic instruments (guitars, drums, trumpet) controlled with revealed virtual shapes using *Rivill*, in collaboration with professional improvisers from the Muzzix collective, and with colleagues from JUNIA/Isen (Arthur Paté, Paul Cambourian). The *Vibrating Shapes* project [6] involves creating actuators that either hit strings, vibrate drum heads or send air to tubes connected to a trumpet. Their parameters (amplitude, frequency, regularity) are mapped to the features of intersections with virtual shapes : shape intersected or not, position inside the shape, intersection size. In turn, the vibration patterns are displayed on the shapes by mapping their amplitude to the scale on concentric spheres, providing feedback for musicians and increasing the transparency for the audience. Figure 3.8.c shows the two musicians interacting virtual shapes which are connected to actuators on an acoustic guitar and a trumpet. During the residences that led to the first performance with the instruments, we analysed design choices and feedback from the musicians. Our results for example suggest that adding the virtual shapes enabled an extended gestural and sonic vocabulary, with musicians being able to develop specific new playing techniques not possible otherwise. Interestingly, they reported having internalised the virtual shapes as part of their instrumental space, but also perceiving them as musical entities with which they were able to dialog.

Overall, *revealed virtual controls* provide a novel way to integrate virtual and physical interactions that open many opportunities for interactive environments [92].

3.3 Novel Mixed-Reality Displays

Technological choices for mixed and virtual reality displays place strong constraints on the interaction with virtual content. Notably, they influence if and how this interaction can be collaborative, they define its level of transparency for observers and they push for certain gestures and techniques. In this section, I describe research that I conducted at the University of Bristol and at the University of Lille on novel mixed-reality displays that take advantage of the exploration of physical and virtual spaces, specifically in contexts of musical expression and cultural mediation. They include :

1. Displays based on optical combiners, which overlap virtual and physical spaces so that both can be manipulated
2. Revealable volume displays, which allow users to reveal virtual content around or inside physical objects and visible by all
3. A novel mixed-reality display for musical performances that creates a shared virtual-physical interaction space

3.3.1 Extended Optical Combiners

In a paper published at ACM UIST 2014 [127], we explored a class of mixed-reality displays that mix physical and virtual volumes by using planar optical combiners, *i.e.*, two-way or semi-transparent mirrors. Because of their planarity, they create stigmatic pairs [115] which means that the reflections preserve the geometry of the physical space. Hence all users perceive reflections and the physical space behind mirrors identically aligned. A number of display and interaction opportunities arise

---

[6] https://vibrating-shapes.univ-lille.fr/

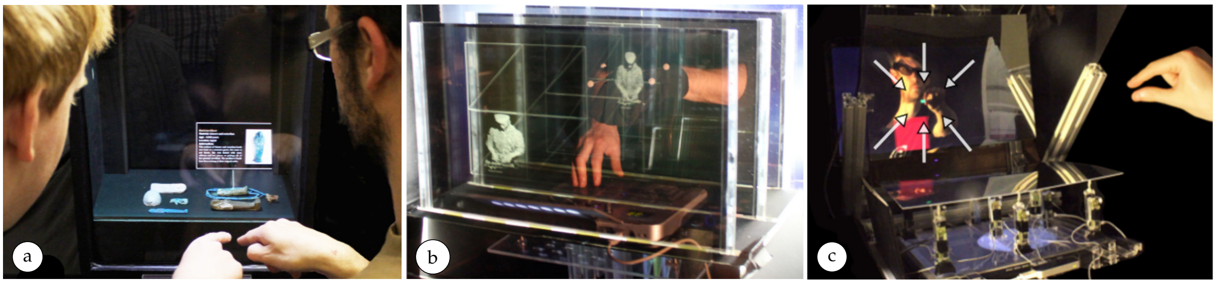

FIGURE 3.9: Combining virtual and physical spaces : a) finger reflection as a 3D pointer inside a physical space b) hand reflection to interact with a volumetric display c) a shape-changing + semi-transparent mirror for optical and virtual manipulations

from this simple property. For example, as shown in Figure 3.9.a, the physical space reflected in a semi-transparent mirror can be used to interact with physical objects placed behind it. Here the user's index finger is tracked and acts as a 3D pointer to select non-reachable physical objects for non-reachable physical volume. The same principle can be used : 1) to interact with a volumetric display, *e.g.*, a depth cube [156] as seen in Figure 3.9.b, 2) to combine virtual content either with the reflected physical space, *e.g.*, in a magic mirror [74], or 3) in the other direction, to place virtual content freely inside a physical space as I will explain in the next section.

We then pushed the idea further with non-planar, shape-changing mirrors in a paper published at ACM ISS 2015 [35]. As seen in Figure 3.9.c, we used dedicated actuators [100], controlled with an underlying screen to change the shape of a flexible mirror. The flexible mirror is reflected on a 45° semi-transparent mirror behind which is placed a stereoscopic screen. The user therefore sees their potentially transformed reflection overlapped with virtual objects. The system then relies on the idea of mirror brushes, meaning that local deformations are applied to the mirror. These are view dependent, *i.e.*, computed according to the main user's point of view. By carefully controlling the deformations, the user can manipulate the reflection of the physical space around them (*e.g.*, moving objects around interactively) or their own appearance. Another example application, shown in Figure 3.9.c, is to enable 3D manipulation with the physical hand of the user acting as a 3D pointer. The reflected hand can be scaled up or down with the shape changing mirror in response to hand poses, to select virtual objects of the corresponding size (here a small green sphere or a large red cube).

3.3.2 Revealable Volume Displays

One subclass of *AR* displays from the class of *Combined AR* presented above is especially interesting in contexts such as museum exhibitions where users are involved in exploratory activities of finding information regarding physical objects. In a paper published at IEEE VR 2021 [15] we present them as *Revealable Volume Displays (RVDs)*, a reference to *Swept-Volume Displays (SVDs)* [157]. *SVDs* rely on automated motion of a physical surface, usually rotated around the vertical axis or translated vertically, on which the corresponding slice of virtual content is projected. With sufficient motion speed and framerate and due to persistence of vision, users perceive a full virtual volume. However, due to mechanical constraints the volumes usually remain small and are difficult to align with physical objects.

RVDs instead rely on users to display the virtual content. They combine the technology described for *Revealed Virtual Controls* in Section 3.2.3 with an optical combiner, *e.g.*, a glass panel, behind which are placed physical objects. When they move their hand or any held object in front of the panel, users intersect the physical space and objects placed behind the panel through reflections. But they also intersect virtual objects in front of the panel, and the re-projected slices of this virtual content are in turn reflected in the panel and appear overlapping the space behind the panel. The virtual objects can be placed around or inside physical objects. For example, Figure 3.10.a depicts an exhibition with annotations overlapping stuffed birds. A visitor places the reflection of a small foam board inside a

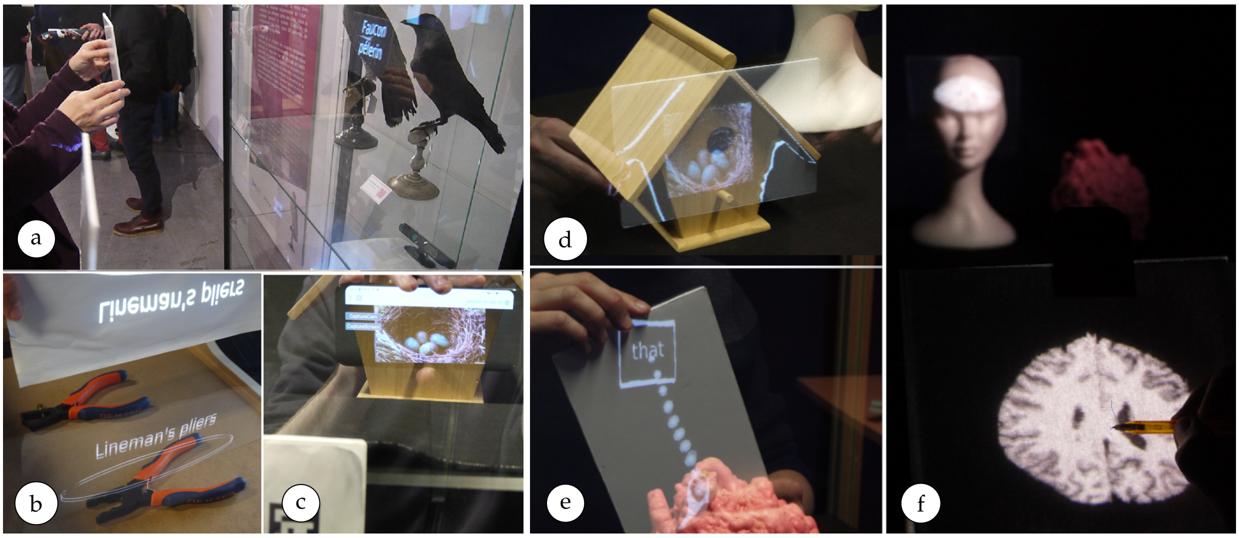

FIGURE 3.10: Revealable Volume Displays : Diverse implementations a) with a vertical panel used in a public exhibition, b) with a horizontal glass panel and c) using a mobile display instead of a projector. Guiding and advanced techniques : d) highlighting physical contours e) using dynamic guides for exploration f) combining RVDs with physical interactions

bird. The surface of the board is captured using the depth camera seen on the bottom right and the intersection with a text box is re-projected on the board. The reflection of the text therefore appears at the correct depth inside the bird, visible by all.

Contrary to mobile AR devices, *RVDs* provide a shared view of the virtual content, because the augmentations are visible by all users facing the glass panel. Contrary to mixed-reality headsets or other see-through stereoscopic displays [101] , *RVDs* do not suffer from convergence-accommodation conflicts [102] (which leads to visual fatigue) because the content is displayed in the physical space at the correct depth. *RVDs* however have clear limitations : 1) only slices of the content are displayed, contrary to full volumes with *SVDs* or stereoscopic displays 2) the content appears only if the users interact with it (which is however interesting in a museum exhibition context) 3) the inverted interaction caused by the mirror can be unfamiliar; 4) perceiving the position of reflected surfaces with respect to exhibited objects can be difficult due to the absence of some visual depth cues such as occlusions and cast shadows.

In the paper, we provide a design space for *RVDs*, showing the diversity of implementations (see Figure 3.10.b and 3.10.c), content and interaction techniques, such as the combination of reflected and physical interactions presented in Figure 3.10.f. We also evaluate three techniques for compensating for issues 3 and 4. In particular, we show that highlighting the contours of physical objects, as depicted in Figure 3.10.d, increases users efficiency in finding virtual content. This efficiency can also be increased using dynamic 3D guides [75], shown in Figure 3.10.e. Finally, the exploration of virtual content can be facilitated by amplifying changes in depth, for example by placing content inside spheres instead of boxes and showing the contours of the shape.

Our *RVDs* have been used in multiple exhibitions, such as with the Natural History Museum of Lille, the Espace Culture of the Université de Lille, and the Maison de l'Archéologie (Pas-de-Calais).

3.3.3 Reflets : a mixed-reality display for artistic performances

Going back to the field of musical expression, the Reflets project [34] takes *RVDs* one step further by placing users on both sides of the optical combiner. As shown on Figure 3.11.a, a glass panel creates a shared space between performers and spectators, and a depth camera / projector pair allows for revealing content on both sides, appearing in this shared space. Consequently, spectators can here

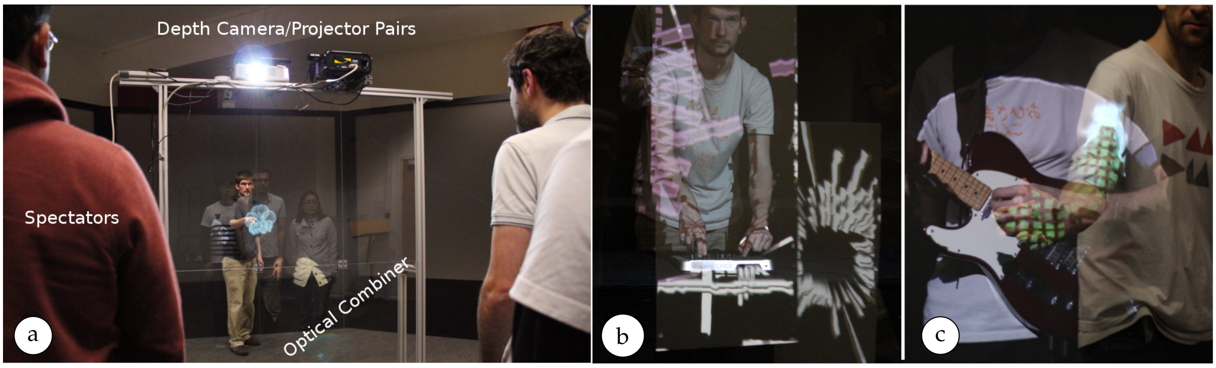

FIGURE 3.11: Reflets : a) the two spaces are combined by the glass panel and camera/projector pairs allow for revealing virtual objects, b) it can be used to reveal mechanisms of DMIs c) and to enable collaborative performances

see the reflection of a blue virtual shape revealed from their side but overlapped with the performer's hand, while the performer sees the reflection of their hand overlapping the virtual shape revealed by one spectator's body. Both perceive the same content, either through or reflected by the panel. Thus the panel, instead of dividing the performance space in two, unifies performers and spectators spaces and enables a variety of augmented performance scenarios.

From workshops with diverse artists, we designed four scenarios, two of which are shown in Figure 3.11. In Figure 3.11.b, spectators reveal components of a *DMI*, exposed following the *Rouages* approach, but here requiring spectator interaction to be visible. Once revealed, the visual augmentations, here showing sensors extensions and the representations of two musical loops, are displayed with a correct perceptual alignment for all. In Figure 3.11.c, performers are on each side of the panel. A guitar player is on one side, inside a white sphere. On the other side, a gestural performer grabs musical phrases from the guitar by intersecting the sphere with their reflection. They then move inside a large virtual box (containing a green grid) to control effects applied to the captured musical phrases.

While this display provide correct visual depth cues for all and is therefore particularly suited for the audience, it can be difficult for the performer to select virtual objects that are not being revealed. *Reflets* could then easily be combined with the use of Mixed-Reality headsets for the performers. Finally, although it has been designed with a focus on musical performances, *Reflets* could be extended to other artistic practices, such as theatre or dance.

3.4 Conclusion and Future Directions for Revealed Interfaces

In this chapter, I presented contributions on 3D interaction and mixed-reality displays which rely on closely integrating the physical and virtual spaces, notably through the use of optical combiners to preserve physical depth cues.

Our results first show how perception, appropriation and experience can be enriched by augmenting tactile, tangible and gestural interfaces with carefully designed visual feedback / virtual elements. They have implications for the design of interfaces, which should either directly integrate these visual augmentations or leave room for their integration, preferably in way that can be customized by expert users. We extended interaction techniques in the virtual and physical spaces, providing design spaces for new musical instruments, and we showed how existing instruments and playing techniques can benefit from these augmentations. Finally, we proposed novel mixed-reality display technologies that offer shared experiences of combined virtual and physical spaces [20], especially adapted to collective experiences such as museum exhibitions and artistic performances. We provided insights to facilitate their design and use. Although these results are obtained in expressive (artistic performance) and exploratory (texture discovery and museum exhibition) contexts, they could be exploited in other fields of application which involve mixing physical and virtual content in collective contexts such as industrial design or medical training.

Future research directions on these interfaces involve :

1. Taking advantage of rich revealed content : More than simple controls such as buttons/sliders, the exploration of 3D textures with the user's hands and physical object should enable complex multi-dimensional mappings
2. Combining revealed interfaces with tangible devices providing passive or active haptic feedback in order to enable interactions at various granularities, as explained in more detail in Section 4.2
3. Taking advantage of expression to bypass or compensate for technological limitations, *e.g.*, pushing forward the principle of *RVDs*, as will be explained in more detail in Section 4.3

4 Future Research Directions

In the previous chapters, I have described the research that I have been conducting for the last ten years. It has led me to the study of mixed-reality displays to enrich the experience of audience and collaborators, and to the creation of interaction and visualisation devices and techniques for artistic expression and content exploration. In addition to the questions that I have mentioned in the conclusions of each chapter which require further investigation, in this chapter I want to insist on research directions that have been little explored yet, and which I believe are essential for the future of expressive interfaces.

The study of visual augmentations for the audience has led to the question of how the experience of observers of digital interactions can be modelled and measured. In particular, we have proposed the notion of *attributed agency*, which describes the amount of control that an observer believes the user has over the system, *i.e.*, the perceived causality between observed actions and effects [144]. However our understanding of attributed agency is constrained to a coarse level due to the use of questionnaires and interviews as measurement tools, which prevents us from precisely analysing the effect of individual components of digital interactions (gestures, interfaces, mappings, system behavior, feedback, ...) and addressing the underlying perceptual and cognitive aspects of observed interactions. The first research direction, detailed in Section 4.1, therefore aims at diving deeper into the perception of digital interactions, through the use of physiological sensors.

Another aspect that appears to pervade much of my research is the notion of *levels of detail*. Whether they allow spectators to select the information they require to improve their experience of digital musical performances, or members of ensembles to understand the activity of other instruments, or even musicians to access multiple levels of control of the sound (*e.g.*, from musical processes to synthesis parameters), the ability to organise and access multiple levels of complexity of *DMIs* seems an essential aspect that needs to be studied further. As explained in Section 4.2, I plan on investigating levels of detail further, building on research from my PhD thesis and applying them to multiple aspects of the design of expressive interfaces.

Finally, I believe that future research on *3DUIs* and musical expression can not forgo a global reflection on the impact of its technological orientation, which can not continue blindly in the current environmental and social context. The goal for the next phase of my research career is therefore to fit contributions on 3D interactions in with the perspective of sustainability and even degrowth. In particular, this means systematically investigating solutions that favor human expression in the physical space over technological components, and shared devices over a multiplication of individual ones. This approach, which could be named *Expressed Virtuality*, is described in more detail in Section 4.3.

Both first directions can serve as subjects for PhD theses and even nationally funded projects. The third one, as it encompasses many aspects, could trigger international collaboration and constitute the basis for a large-scale European project in collaboration with research teams, artists and luthiers.

4.1 Revealing the Perception of Digital Interactions

We have shown that the audience experience, in particular their subjective comprehension, can be enriched using visual augmentations that reveal the interaction mechanisms, coping for disruptions of attributed agency criteria. However, in part because we remained at a short performance level, it is not clear yet how this comprehension emerges from the perception of a sequence of actions and effects

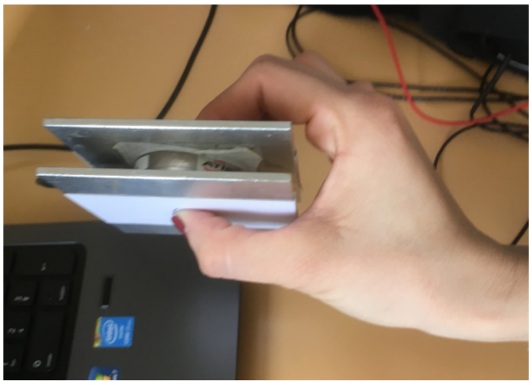

FIGURE 4.1: Held between fingers, the grip-force sensor (strain gage) measures small variations of pressure which can arise from motor simulation when watching an action or hearing action verbs

in the interaction with a digital system, *i.e.*, what is the influence of the various temporal and spatial components of the interaction. For instance, we do not know if the gesture amplitude has more weight than the gesture to sound consistency criteria, or when the judgement of attributed agency actually happens during an observed interaction. In the research we conducted so far, we mostly relied on post-stimuli questionnaires and interviews. Although they provide valuable information, they fail to capture fine temporal cues on the spectator's perception. For example, it is not possible to obtain the reaction to individual actions and effects, nor to understand the effect of sequences of interactions. Previous work has relied on physiological sensors [120] and continuous evaluation [12,39] in order to retrieve on-the-fly data for a single or many spectators of musical performance. Bin et al. [39] combine on-the-fly data, retrieved with mobile interfaces given to spectators, and post-hoc interviews, allowing for a richer analysis on audience experience.

This first research direction will therefore explore novel methodologies to refine our understanding of the audience experience with digital, especially musical, interactions. In particular, we will build on the preliminary results obtained at the end of Olivier Capra's thesis [57] using a *grip force* sensor, *i.e.*, a strain gage sensor that measures small variations of pressure applied when holding it between index finger and thumb, as depicted in Figure 4.1. This sensor has the advantage of being less intrusive than other physiological sensors (an important aspect in contexts of public performances) while providing information on cognitive processes through the measurement of motor simulation during the observation of actions. Motor simulation results from the activation of mirror neurons [90], meaning that the same neural structures activate during one's actions than during one's observation of similar actions, and that this activation results in efferent signals to the body, either congruent with the action or not. This activity can be measured through brain activity but also directly from muscular activity, for example through grip-force. Components of observed actions are reflected in the physiological signals in the observers. For instance, grip force has been shown to vary according to observed gesture intensity [43], Because digital interaction, in particular interfaces musical expression, introduce levels of complexity and semantic to simple physical gestures, notably through mappings, it is important to be able to sense higher-level processes or components of cognition. In fact, grip force highlights changes in motor activation when listening to action and non-action verbs and between affirmative and negative sentence contexts [2,86], indicating the possibility of discriminating higher-level of cognition, such as attributed agency. In fact, outside grip force sensors, auditory-motor interactions have been thoroughly documented in the literature [173], with effect of higher level components of the interaction on motor and brain activity. For example, familiarity (through practice) [116] with a stimulus and congruency [8] between gestures and resulting sounds have been shown to influence cerebral activity when observing musical actions. Similarly, fMRI studies have shown [56,141] the influence of expertise in cognitive processing of observed actions.

Our hypothesis here is that the grip force sensor could be used to analyse the components of digital musical expression and their effect on audience experience. For example, preliminary results from Olivier Capra's thesis suggest that there could be an effect of instrument familiarity on motor simulation [107] detected with a grip force sensor. However, they were obtained from only a few subjects and aggregate multiple non-transparent instruments and mappings, preventing from a finer analysis. We will therefore systematically explore variations of the components of digital interactions (gestures, sensors, mappings, processes, sequences of actions) and study their effect on motor simulation and attributed agency.

Physiological sensing such as grip force, will give us a temporally accurate response to perceived musical interactions. However it also strongly hides many aspects of the rich experience lived by spectators, especially in a context of public performance. In order to integrate these aspects, we will combine physiological measurements and qualitative methods for studying the experience. For example, the NIME community is more and more taking into account reflexive methods such as microphenomenology [143], which enable the rich recounting of moments of the user experience.

With this research direction, we plan to provide the HCI and NIME communities with novel knowledge on the spectator experience, including the effect of practice, mapping complexity, gesture choices on aspects such as subjective comprehension, perception of errors and virtuosity. These results will help inform the design of public and collaborative interfaces. On the long term, this methodology should allow us to move from single gestures, to sequences of gestures in short performances, to the study of full performances in-the-wild.

4.2 Designing with Revealed Levels of Detail

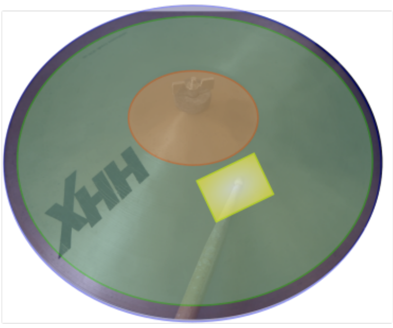

FIGURE 4.2: Example LODs on a cymbal : starting from selecting broad zones in blue, green and orange, users as they gain expertise will be able to focus on fine-grained changes in position within small areas such as the one in yellow, however with some amount of randomness on the exact hit position

The use of levels of detail (*LOD*) is a recurring approach in HCI, music computing and computer graphics, sometimes under appellations such as *multi-scale*, *hierarchical*, *tree-like models*. In computer graphics, providing multiple levels-of-detail is essential to reduce computing load. As the distance to virtual objects increases, they are rendered with fewer and fewer pixels, therefore their shape can be simplified to reduce the number of processed vertices and faces. In *HCI*, Zoomable User Interfaces [13], hierarchical menus [7] and more recently AR interfaces [76], have provided ways of selecting levels of information with zooming or mid-air selection techniques. Multiscale selection and navigation has also been used in 3D user interfaces in order to combine coarse and fine grained interactions with virtual environments [65,114]. Hierarchical representations are also common in information visualisation with techniques such as Treemaps or sunburst diagrams, that enable the exploration of complex information grouped by semantic levels. In music computing, LODs appear in models of musical structures but also in the design of mappings [87]. For instance, in my PhD thesis, I used an hierarchical / LOD approach to extend the live-looping musical expression technique, with the Drile instrument that allowed to manipulate trees of musical loops [29]. As shown in our research, the *LOD* approach can also be used to select relevant information for the audience, allowing them to focus on aspects that are important to enrich their experience of musical performances.

However, I believe that *LOD* has not been enough investigated, in particular in the design of musical interaction, and taking advantage of mixed-reality interfaces. When looking at acoustic instruments, even as simple as a cymbal played with a drumstick, musicians travel through *LOD* as they gain expertise. As one can imagine from looking at Figure 4.2, starting with hitting the cymbal as a whole, they then move to the selection of broad zones (rim, body, bell), then to specific points in each zone, accessing more detailed control as they progress, with however always some amount of variability from *LODs* that they do not control. As we have seen in audience strategies for choosing visual augmentations, expertise also leads to choices of more detailed information about *DMIs*, and collaboration in heterogeneous ensembles might require simplified representations of processes for communication but detailed representation for cooperation. In all cases, *LODs* play an important role which has yet to be fully comprehended.

This research direction will examine how the design of interactive systems and more specifically digital musical instruments may benefit from a *LOD* approach, including :

- Design methodologies that emphasize the combination of mappings between diverse granularities of gestures (*e.g.*, hand pose / small finger movements), and of sound parameters (*e.g.*, synthesizer preset / individual synthesis parameter)
- Mixed-reality interfaces that allow performers to access various levels of input/output complexity [108] either one at a time depending on their expertise, or combining the *LODs* as novel playing techniques
- Visual augmentations of *DMIs* with controllable level of detail for collaborators in digital ensembles, so that they can switch between seeing the simplified overall activity or the detailed variations of all parameters

A first step could be to reimplement the Drile instrument [29] and push *LODs* in its design by 1) creating representations for complexity levels of the musical tree and 2) adding a temporal *LOD* (recording loops with various levels of temporal quantization).

4.3 Long-term research : Expressed Virtuality

On a longer term, my goal is to push research on Mixed and Virtual Reality towards less technology-centered and more *"human expression"-centered* systems. This approach, that could be named *Expressed Virtuality* in reference to the well-known real-virtual continuum [133] has influenced my research in the past years, and I believe it should be formalized and systematically explored. An example are *Revealable Volume Displays*, presented in Section 3.3.2, which transform the automated movements of surfaces traditionally used for volumetric displays into expressive human movements captured with a depth camera, leaving room for a vocabulary of gestures and exploration techniques. Following on that, many components of interactive systems can be transferred from the digital/virtual space, where they need to be processed and rendered, to the physical space, where they will be supported by expression and expertise of users.

This direction, with its focus on expression rather than technology, also provides an opportunity for tackling the issue of reducing the technological footprint of *3DUIs*, which is essential given their energy and resources consumption. In that, it meets the efforts of many research communities (such as Eco NIME [128] or Sustainable HCI [47]) on taking environmental issues into account, in particular aiming at more frugality and sustainability. While this may be easier in the case of tangible or partly acoustic interfaces, it is surely more difficult with 3D interfaces which rely on an important stack of technologies for rich sensing and feedback. 3D Interfaces, by limiting the need for travelling or for building physical objects through virtual simulations, are also too often seen as a way to advance or inform on environmental issues [67] rather than constituting themselves an issue. However, one can take advantage of mixed-reality technologies to selectively move components of *3DUIs* across the real-virtual continuum in order to minimize processing and rendering resources. An example could be a virtual instrument such as Drile [29] which would consist in musical tree structures with virtual nodes and actuated acoustic leaves, leaving all sound production (and potentially some of the interaction) to the physical space. Because the purpose of expressive applications is often to produce physical results for an audience, transferring manipulations or effects to the physical space might in fact have a positive impact on the spectator experience by increasing their familiarity with parts of the system.

In addition to introducing environmental constraints in the design of *3DUIs*, this approach will take advantage of existing practices and gestures, *e.g.*, techniques in acoustic lutherie and instrumental gestures, and allow for extending these physical techniques and gestures, as exemplified by the *Vibrating Shapes* project described above. In order to do so, we will put the emphasis on technologies that bring back control and feedback on virtual content to the physical space such as *passive haptic feedback* [1] or *spatial augmented-reality* [37]. We will apply our methodology to design interaction devices and techniques but also visual, auditory and haptic displays.

Another essential aspect of *Expressed Virtuality* is the transfer of knowledge and tools from 3DUIs to physical interfaces such as acoustic musical instruments.
Ideas that have been explored for the interaction with virtual environments, such as visual guides, virtual tools, could be transposed to physical counterparts, therefore enriching practices such as musical playing techniques.

Finally, inserting expression in interaction means that our results might not be optimal when compared with existing techniques and technologies on common efficiency criteria. This might lead to the creation of specific evaluation methodologies that combine quantitative and qualitative tests, with a focus on the user experience, following third wave HCI practice.

Here are some potential directions to explore *Expressed Virtuality*.

4.3.1 Processing and rendering in the physical space

The first involves taking digital processes from various stages of *3DUIs*, such as sensors, interaction techniques, mappings, synthesis, and replacing them with physical components. On the musical side, this idea very much relates to the notion of intimacy [169] in *DMIs*, *i.e.*, reducing the gap between physical gestures and digital processes. Intimacy can be achieved by refining sensing through increased sampling rate and resolution or enriching haptic/visual/auditory feedback. A good example is the Caress instrument [134] which replaces traditional FSRs which values are mapped to synthesis parameters, with piezoceramic transducers that directly reflect the fine-grained vibrations originating from users' interactions with a surface as audio signals.

However, intimacy can also be achieved by completely leaving the digital/virtual space. This is the method that we chose to apply for the *Vibrating Shapes* project, where all sound synthesis happens on the acoustic guitars and can be influenced directly by the musicians' gestures without further sensing. In this case, only the end of the control chain is transferred.

But other components could be converted as well. In many *IVMIs* and more generic virtual environments, touching virtual objects involves 1) sensing gestures such button presses or finger contact with a surface, 2) generating a feedback signal based on sensors values, 3) rendering the haptic feedback with various technologies. One could imagine bypassing these steps by placing all finger to surface interaction in the physical space. For example, let's imagine a loudspeaker in which the central magnet, held by the user in their dominant hand, would be separated from the coil, held in their non-dominant hand and attached to a wooden box. As shown in Figure 4.3, when the user approaches their non-dominant hand from a virtual object, the system outputs some audio signal to the coil, creating a oscillating magnetic field. To touch this virtual object, the user then simply brings their dominant hand and the magnet closer to their non-dominant hand, which will make both the magnet and the coil vibrate. This will in turn generate both haptic and auditory feedback from only very fine interaction performed in the physical space without the need for a sensing/actuating loop.

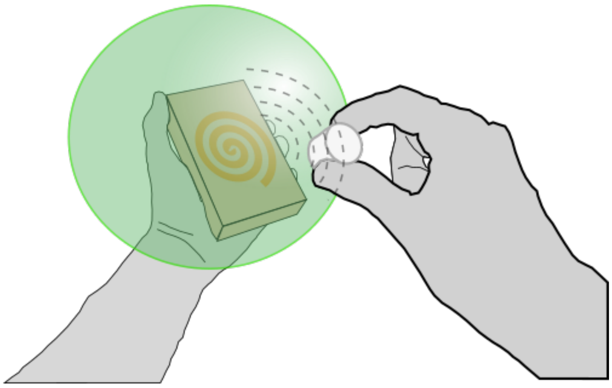

FIGURE 4.3: Mockup of a 3D interaction device to explore virtual objects. An audio signal is sent to the coil when the left hand enters the sphere. The magnet held in the right hand will then oscillate and generate vibrations when in proximity with the sphere, which results in auditory and haptic feedback.

In extreme cases, components of the 3D interaction can be removed. For instance, in the *Vibrating Shapes* project, musicians reported having internalised the position of virtual shapes in the physical space, perceiving them as part of their instrument. In this case, the virtual controls could remain hidden (although with consequences on the audience experience), removing the need for displaying them with a projector, at least after practice sessions.

4.3.2 Finding solutions in human expression

Another direction is to look at common issues within *3DUIs* through the lens of expression. For instance, in my research on 3D interaction techniques for musical expression [22], I advocated for turning disambiguation techniques for 3D selection, *e.g.*, when multiple objects are intersected by a virtual ray when only one should be selected, into *"ambiguation"* techniques. Instead of making sure that only one object is selected using mechanisms such as temporal threshold, depth cursor and so on [3], this idea could for example result into a continuous weighted selection of groups or parts of objects. This fuzzy selection mechanism could then be the starting point for the expressive exploration or manipulation of sound sources, of elements for 3D sculpting. A mockup of this idea is shown in Figure 4.4.

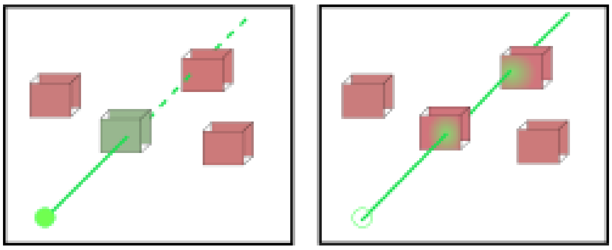

FIGURE 4.4: Mockup of an ambiguous 3D selection technique. Left : the virtual ray stops on the first object, the one behind would be accessible through disambiguation Right : with an ambiguation technique, the selection is weighted over both objects and within their surface

Another example of issue is the latency induced by 3D tracking and projection. In *RVDs* for instance, fast hand movements are not captured and processed fast enough for the intersections to remain properly projected onto physical surfaces. Technical solutions have been proposed in the literature, such as fast moving mirrors [135] or the combination of depth and fast IR cameras [124]. However, another possibility is to rely on expert interaction. In fact, by learning spatial trajectories, users could align their movements with automated projected intersections. One could also imagine relying on expressive techniques such as live-looping, *i.e.*, recording and directly playing back control data as a loop. Users could record a gesture at a slow speed, with depth data being captured, before repeating it faster, with the playback of depth data synchronised thanks to the detection of gesture start. This would allow users to learn how to quickly explore large parts of volumetric content while preserving the projection alignment and without resorting to complex technological solutions.

4.3.3 Evaluation

On the evaluation side, methodologies will need to retain benefits of controlled experiments while avoiding a focus on efficiency. It should be possible to compare our interaction techniques with previous ones, and to evaluate variants of novel technologies, including through hypothesis testing, without necessarily ranking them. I have already started integrating more closely these practices in my past research, but putting the focus on expression requires that all potentially useful methods should be explored, including :

- Bayesian statistics [112] that acknowledge subjectivity and uncertainty in evaluation and leave more room for discussion (similarly to the switch from p-value centered analysis to confidence intervals [81])
- Reflexive Thematic analysis [46] and other qualitative frameworks that insist on a reflexive approach to analysis, involving users (and especially experts) as much as possible
- Performance-led research [19] and in-the-wild trials [50] [64] that involve spending more time with users in an ecological context rather than in the lab

For instance, an interesting lead that originates from our work on *Vibrating Shapes* with expert musicians is the potential of free improvisation [150] for the discovery of gestures, playing techniques, technological constraints. Combining these sessions of free improvisation with diverse evaluation methods could lead to novel insights on expressive *3DUIs*.

5 Conclusion

In this document, I presented the research that I have been conducting on combining 3D User Interfaces, in Mixed and Virtual Reality, and expressive / exploratory interfaces.

The results highlight the interest of *Revealing Interactions* with *3DUIs* in order to :

1. Understand the experience of spectators and collaborators of expressive interfaces, looking at awareness, attributed agency, subjective comprehension
2. Enrich their experience with visual augmentations that reveal the components of expressive interactions : subtle/hidden gestures and sensors, controllable parameters and their values, structure and activity of the underlying synthesis and effects processes
3. Extend existing physical and virtual interfaces by adding virtual controls and altering users' perception
4. Open novel opportunities for expression and exploration of content by closely integrating physical and virtual spaces and relying on users' expertise with physical interfaces

This document also underlines the importance of artistic, and in particular musical, expression as a research axis for human-computer interaction. Many results, such as the design of *RVDs* or knowledge on attributed agency, were indeed obtained while looking at music related issues. We also demonstrated the rich insights on interaction that can be gathered from expert musicians. The other way around, results from *HCI* research, such as redirected interaction or GUI remixing, inspired our research on musical expression, which therefore constitutes a rich application domain.

Finally, I strongly believe that the environmental question brought up in Section 4.3 constitutes more of an opportunity than a challenge for the design of expressive 3D interfaces and for the fields of NIME and HCI. In addition to a reduced and shared use of immersive technologies, promoting free and open-source hardware and software components is essential to ensure long term re-usability. Overall, having to reduce our fields' overall technological footprint will trigger the search for creative solutions, which can take advantage of the capability of navigating the real-virtual continuum.

6 References

1. Merwan Achibet, Maud Marchal, Ferran Argelaguet, and Anatole Lécuyer. 2014. The virtual mitten: A novel interaction paradigm for visuo-haptic manipulation of objects using grip force. In *2014 IEEE symposium on 3D user interfaces (3DUI)*, 59–66.
2. Pia Aravena, Mélody Courson, Victor Frak, Anne Cheylus, Yves Paulignan, Viviane Deprez, and Tatjana A Nazir. 2014. Action relevance in linguistic context drives word-induced motor activity. *Frontiers in human neuroscience* 8: 163.
3. Ferran Argelaguet and Carlos Andujar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3: 121–136.
4. Cagan Arslan, Florent Berthaut, Anthony Beuchey, Paul Cambourian, and Arthur Paté. 2022. Vibrating shapes : Design and evolution of a spatial augmented reality interface for actuated instruments. In *New Interfaces for Musical Expression*. `https://doi.org/10.21428/92fbeb44.c28dd323`
5. Cagan Arslan, Florent Berthaut, Jean Martinet, Ioan Marius Bilasco, and Laurent Grisoni. 2018. The Phone with the Flow: Combining Touch + Optical Flow in Mobile Instruments. In *NIME 2018 - 18th International Conference on New Interfaces for Musical Expression*. Retrieved from `https://hal.science/hal-01809210`
6. Mahdi Azmandian, Mark Hancock, Hrvoje Benko, Eyal Ofek, and Andrew D Wilson. 2016. Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences. In *Proceedings of the 2016 chi conference on human factors in computing systems*, 1968–1979.
7. Gilles Bailly, Eric Lecolinet, and Laurence Nigay. 2007. Wave menus: Improving the novice mode of hierarchical marking menus. In *Human-computer interaction – INTERACT 2007*, 475–488.
8. Marc Bangert and Eckart O Altenmüller. 2003. Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC neuroscience* 4, 1: 1–14.
9. Jerônimo Barbosa, Filipe Calegario, Veronica Teichrieb, Geber Ramalho, and Patrick McGlynn. 2012. Considering audience's view towards an evaluation methodology for digital musical instruments. In *NIME*.
10. Tarik Barri. 2009. Versum: Audiovisual composing in 3d.
11. Cagatay Basdogan, Frederic Giraud, Vincent Levesque, and Seungmoon Choi. 2020. A review of surface haptics: Enabling tactile effects on touch surfaces. *IEEE Transactions on Haptics* 13, 3: 450–470. `https://doi.org/10.1109/TOH.2020.2990712`
12. Mehmet Aydin Baytas, Tilbe Göksun, and Oğuzhan Özcan. 2016. The perception of live-sequenced electronic music via hearing and sight. In *Proceedings of NIME*.
13. Benjamin B Bederson and James D Hollan. 1994. Pad++: A zooming graphical interface for exploring alternate interface physics. In *Proceedings ACM UIST*.
14. Victoria Bellotti, Maribeth Back, W Keith Edwards, Rebecca E Grinter, Austin Henderson, and Cristina Lopes. 2002. Making sense of sensing systems: Five questions for designers and researchers. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 415–422.

15. Fatma Ben Guefrech, Florent Berthaut, Patricia Plénacoste, Yvan Peter, and Laurent Grisoni. 2021. Revealable Volume Displays: 3D Exploration of Mixed-Reality Public Exhibitions. In *IEEE Conference on Virtual Reality and 3D User Interfaces*. `https://doi.org/10.1109/VR50410.2021.00023`

16. Steve Benford, John Bowers, Lennart E. Fahlén, and Chris Greenhalgh. 1994. Managing mutual awareness in collaborative virtual environments. In *Proceedings of VRST*, 223–236.

17. Steve Benford, John Bowers, Lennart E. Fahlén, Chris Greenhalgh, and Dave Snowdon. 1995. User embodiment in collaborative virtual environments. In *Proceedings of CHI 95*, 242–249.

18. Steve Benford, Gabriella Giannachi, Boriana Koleva, and Tom Rodden. 2009. From interaction to trajectories: Designing coherent journeys through user experiences. In *Proceedings of ACMI CHI*.

19. Steve Benford, Chris Greenhalgh, Andy Crabtree, Martin Flintham, Brendan Walker, Joe Marshall, Boriana Koleva, Stefan Rennick Egglestone, Gabriella Giannachi, Matt Adams, and others. 2013. Performance-led research in the wild. *ACM Transactions on Computer-Human Interaction (TOCHI)*.

20. Steve Benford, Chris Greenhalgh, Gail Reynard, Chris Brown, and Boriana Koleva. 1998. Understanding and constructing shared spaces with mixed-reality boundaries. *ACM Transactions on computer-human interaction (TOCHI)* 5, 3: 185–223.

21. Wouter M. Bergmann Tiest and Astrid M. L. Kappers. 2007. Haptic and visual perception of roughness. *Acta Psychologica* 124, 2: 177–189. https://doi.org/`https://doi.org/10.1016/j.actpsy.2006.03.002`

22. Florent Berthaut. 2020. 3D interaction techniques for musical expression. *Journal of New Music Research*. `https://doi.org/10.1080/09298215.2019.1706584`

23. Florent Berthaut. 2021. Musical Exploration of Volumetric Textures in Mixed and Virtual Reality. In *International Conference on New Interfaces for Musical Expression (NIME 2021)* (2021 NIME proceedings). `https://doi.org/10.21428/92fbeb44.6607d04f`

24. Florent Berthaut, Cagan Arslan, and Laurent Grisoni. 2017. Revgest: Augmenting Gestural Musical Instruments with Revealed Virtual Objects. In *International Conference on New Interfaces for Musical Expression*. Retrieved from `https://hal.archives-ouvertes.fr/hal-01518579`

25. Florent Berthaut, David Coyle, James W Moore, and Hannah Limerick. 2015. Liveness through the lens of agency and causality. In *Proceedings of NIME*.

26. Florent Berthaut and Luke Dahl. 2016. BOEUF: A Unified Framework for Modeling and Designing Digital Orchestras. *Lecture notes in computer science* 9617: 153–166. `https://doi.org/10.1007/978-3-319-46282-0/_10`

27. Florent Berthaut and Luke Dahl. 2020. Adapting & Openness: Dynamics of Collaboration Interfaces for Heterogeneous Digital Orchestras. In *New Interfaces for Musical Expression*. Retrieved from `https://hal.science/hal-02561064`

28. 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*. `https://doi.org/10.21428/92fbeb44.9d974714`

29. Florent Berthaut, Myriam Desainte-Catherine, and Martin Hachet. 2010. Drile: An Immersive Environment for hierarchical live-looping. In *New Interface for Musical Expression*, page 192. Retrieved from `https://hal.archives-ouvertes.fr/hal-00530071`

30. Florent Berthaut and Martin Hachet. 2016. Spatial Interfaces and Interactive 3D Environments for Immersive Musical Performances. *IEEE Computer Graphics and Applications* 36, 5: 82–87. `https://doi.org/10.1109/MCG.2016.96`

31. Florent Berthaut, Martin Hachet, and Myriam Desainte-Catherine. 2010. Piivert: Percussion-based Interaction for Immersive Virtual EnviRonmenTs. In *Symposium on 3D User Interfaces*, p 15–18. Retrieved from `https://hal.science/hal-00522637`

32. Florent Berthaut and Alex Jones. 2016. ControllAR: Appropriation of visual feedback on control surfaces. In *Proceedings of the 2016 ACM on interactive surfaces and spaces* (ISS '16), 271–277. `https://doi.org/10.1145/2992154.2992170`

33. Florent Berthaut, Mark T. Marshall, Sriram Subramanian, and Martin Hachet. 2013. Rouages: Revealing the Mechanisms of Digital Musical Instruments to the Audience. In *New Interfaces for Musical Expression*, 6 pages. Retrieved from `https://hal.science/hal-00807049`

34. Florent Berthaut, Diego Martinez Plasencia, Martin Hachet, and Sriram Subramanian. 2015. Reflets: Combining and Revealing Spaces for Musical Performances. In *Proceedings of NIME*. Retrieved from `https://hal.inria.fr/hal-01136857`

35. Florent Berthaut, Deepak Ranjan Sahoo, Jess Mcintosh, Diptesh Das, and Sriram Subramanian. 2015. Dynamir : Optical Manipulations using Dynamic Mirror Brushes. In *ACM International Conference on Interactive Tabletops and Surfaces*, (Proceedings of ACM international conference on interactive tabletops and surfaces). Retrieved from `https://hal.science/hal-01222512`

36. Florent Berthaut, Victor Zappi, and Dario Mazzanti. 2014. Scenography of immersive virtual musical instruments. In *VR Workshop: Sonic Interaction in Virtual Environments (SIVE), 2014 IEEE*. `https://doi.org/10.1109/SIVE.2014.7006285`

37. Oliver Bimber and Ramesh Raskar. 2019. *Spatial augmented reality: Merging real and virtual worlds*. AK Peters/CRC Press.

38. SM Astrid Bin, Nick Bryan-Kinns, A McPherson, and others. 2017. Hands where we can see them! Investigating the impact of gesture size on audience perception.

39. SM Astrid Bin, Fabio Morreale, Nick Bryan-Kinns, and Andrew P McPherson. 2017. In-the-moment and beyond: Combining post-hoc and real-time data for the study of audience perception of electronic music performance. In *IFIP conference on human-computer interaction*.

40. SMA Bin, AP McPherson, N Bryan-Kinns, and others. 2016. Skip the pre-concert demo: How technical familiarity and musical style affect audience response. In *Proceedings of NIME*.

41. David Birnbaum, Rebecca Fiebrink, Joseph Malloch, and Marcelo M. Wanderley. 2017. 2005: Towards a dimension space for musical devices. In *A NIME reader: Fifteen years of new interfaces for musical expression*, Alexander Refsum Jensenius and Michael J. Lyons (eds.). Springer International Publishing, Cham, 211–222. `https://doi.org/10.1007/978-3-319-47214-0_14`

42. Tina Blaine and Sidney Fels. 2003. Contexts of collaborative musical experiences. In *Proceedings of the 2003 conference on new interfaces for musical expression* (NIME '03), 129–134. Retrieved from `http://dl.acm.org/citation.cfm?id=1085714.1085745`

43. J Blampain, L Ott, and YN Delevoye-Turrell. 2018. Seeing action simulation as it unfolds: The implicit effects of action scenes on muscle contraction evidenced through the use of a grip-force sensor. *Neuropsychologia* 114: 231–242.

44. Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola, and Ivan Poupyrev. 2004. *3D user interfaces: Theory and practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA.

45. Detjon Brahimaj, Florent Berthaut, Frédéric Giraud, and Betty Semail. 2023. Cross-modal interaction of stereoscopy, surface deformation and tactile feedback on the perception of texture roughness in an active touch condition. In *IHM'23 - 34e Conférence Internationale Francophone sur l'Interaction Humain-Machine AFIHM*. Retrieved from `https://hal.science/hal-04014962`

46. Virginia Braun, Victoria Clarke, Nikki Hayfield, and Gareth Terry. 2019. Thematic analysis. In *Handbook of research methods in health social sciences*, Pranee Liamputtong (ed.). Springer Singapore, 843–860.

47. Christina Bremer, Bran Knowles, and Adrian Friday. 2022. Have we taken on too much?: A critical review of the sustainable HCI landscape. In *Proceedings of the 2022 CHI conference on human factors in computing systems*, 1–11.

48. David Brickler, Matias Volonte, Jeffrey W. Bertrand, Andrew T. Duchowski, and Sabarish V. Babu. 2019. Effects of stereoscopic viewing and haptic feedback, sensory-motor congruence and calibration on near-field fine motor perception-action coordination in virtual reality. *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 - Proceedings*, March: 28–37. https://doi.org/10.1109/VR.2019.8797744

49. John Brosz, Miguel A. Nacenta, Richard Pusch, Sheelagh Carpendale, and Christophe Hurter. 2013. Transmogrification: Causal manipulation of visualizations. In *Proceedings of the 26th annual ACM symposium on user interface software and technology* (UIST '13), 97–106. https://doi.org/10.1145/2501988.2502046

50. Barry Brown, Stuart Reeves, and Scott Sherwood. 2011. Into the wild: Challenges and opportunities for field trial methods. In *Proceedings of ACM CHI*.

51. Nick Bryan-Kinns. 2013. Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies* 71, 1: 76–90.

52. Nick Bryan-Kinns and Patrick GT Healey. 2004. Daisyphone: Support for remote music collaboration. In *Proceedings of the 2004 conference on new interfaces for musical expression*, 27–30.

53. Ivica Bukvic. 2022. Latency-, sync-, and bandwidth-agnostic tightly-timed telematic and crowdsourced musicking made possible using L2Ork tweeter. In *Proceedings of the international conference on new interfaces for musical expression*. https://doi.org/10.21428/92fbeb44.a0a8d914

54. Claude Cadoz and Marcelo Mortensen Wanderley. 2000. Gesture-music. *Trends in gestural control of music*.

55. Cem Cakmak, Anil Camci, and Angus Forbes. 2016. Networked virtual environments as collaborative music spaces. In *Proceedings of the international conference on new interfaces for musical expression* (2220-4806), 106–111. Retrieved from http://www.nime.org/proceedings/2016/nime2016_paper0022.pdf

56. Beatriz Calvo-Merino, Daniel E Glaser, Julie Grèzes, Richard E Passingham, and Patrick Haggard. 2004. Action observation and acquired motor skills: An FMRI study with expert dancers. *Cerebral cortex* 15, 8: 1243–1249.

57. Olivier Capra. 2020. The experience of spectators of digital interactions : Benefits of visual augmentations and the role of attributed agency in electronic music performances. Université de Lille. Retrieved from https://theses.hal.science/tel-03213044

58. Olivier Capra, Florent Berthaut, and Laurent Grisoni. 2017. Toward augmented familiarity of the audience with digital musical instruments. In *Proceedings of CMMR*.

59. Olivier Capra, Florent Berthaut, and Laurent Grisoni. 2020. Have a SEAT on Stage : Restoring Trust with Spectator Experience Augmentation Techniques. In *Designing Interactive Systems (DIS)*. https://doi.org/10.1145/3357236.3395492

60. Olivier Capra, Florent Berthaut, and Laurent Grisoni. 2020. All you need is LOD : Levels of detail in visual augmentations for the audience. In *Proceedings of the international conference on new interfaces for musical expression*, 67–72. https://doi.org/10.5281/zenodo.4813236

61. Olivier Capra, Florent Berthaut, and Laurent Grisoni. 2020. A Taxonomy of Spectator Experience Augmentation Techniques. In *NIME 2020 - The 20th International Conference of New Interfaces for Musical Expression*. Retrieved from https://hal.science/hal-02560931

62. Olivier Capra, Florent Berthaut, and Laurent Grisoni. 2021. Levels of Detail in Visual Augmentations for the Novice and Expert Audiences. *Computer Music Journal*. Retrieved from https://hal.science/hal-03106626

63. Baptiste Caramiaux, Jules Françoise, Norbert Schnell, and Frédéric Bevilacqua. 2014. Mapping through listening. *Computer Music Journal* 38, 3: 34–48.

64. Alan Chamberlain, Andy Crabtree, Tom Rodden, Matt Jones, and Yvonne Rogers. 2012. Research in the wild: Understanding "in the wild" approaches to design and development. In *Proceedings of the designing interactive systems conference*, 795–796.

65. Morgan Le Chenechal, Jeremy Lacoche, Jerome Royan, Thierry Duval, Valerie 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 *2016 IEEE third VR international workshop on collaborative virtual environments (3DCVE)*, 18–22. `https://doi.org/10.1109/3DCVE.2016.7563562`

66. Nuno N Correia, Deborah Castro, and Atau Tanaka. 2017. The role of live visuals in audience understanding of electronic music performances. In *Proceedings of the 12th international audio mostly conference on augmented and participatory sound and music experiences*, 29.

67. Laura D Cosio, Oğuz'Oz'Buruk, Daniel Fernández Galeote, Isak De Villiers Bosman, and Juho Hamari. 2023. Virtual and augmented reality for environmental sustainability: A systematic review. In *Proceedings of the 2023 CHI conference on human factors in computing systems*, 1–23.

68. Céline Coutrix and Laurence Nigay. 2006. Mixed reality: A model of mixed interaction. In *Proceedings of the working conference on advanced visual interfaces* (AVI '06), 43–50. `https://doi.org/10.1145/1133265.1133274`

69. Luke Dahl. 2012. Wicked problems and design considerations in composing for laptop orchestra. In *NIME*.

70. Luke Dahl, Florent Berthaut, Antoine Nau, and Patricia Plenacoste. 2017. bf-pd: Enabling Mediated Communication and Cooperation in Improvised Digital Orchestras. In *International Symposium on Computer Music Multidisciplinary Research*. Retrieved from `https://hal.science/hal-01577942`

71. Sofia Dahl, Frédéric Bevilacqua, and Roberto Bresin. 2010. Gestures in performance. In *Musical gestures*. Routledge, 48–80.

72. Sofia Dahl and Anders Friberg. 2007. Visual perception of expressiveness in musicians' body movements. *Music Perception: An Interdisciplinary Journal* 24, 5: 433–454.

73. Maxime Dariosecq, Patricia Plénacoste, Florent Berthaut, Anis Kaci, and Frédéric Giraud. 2020. Investigating the semantic perceptual space of synthetic textures on an ultrasonic based haptic tablet. *VISIGRAPP 2020 - Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* 2: 45–52. `https://doi.org/10.5220/0008979800450052`

74. Jean Baptiste De la Riviere, Nicolas Dittlo, Emmanuel Orvain, Cédric Kervégant, and Mathieu Courtois. 2010. Holocubtile: 3D multitouch brings the virtual world into the user's hands. In *ACM international conference on interactive tabletops and surfaces*, 311–311.

75. William Delamare, Thomas Janssoone, Céline Coutrix, and Laurence Nigay. 2016. Designing 3D gesture guidance: Visual feedback and feedforward design options. In *Proceedings of the international working conference on advanced visual interfaces* (AVI '16), 152–159. `https://doi.org/10.1145/2909132.2909260`

76. S. DiVerdi, T. Hollerer, and R. Schreyer. 2004. Level of detail interfaces. In *Third IEEE and ACM international symposium on mixed and augmented reality*, 300–301. `https://doi.org/10.1109/ISMAR.2004.38`

77. Steve Dixon. 2006. A history of virtual reality in performance. *International Journal of Performance Arts & Digital Media* 2, 1.

78. Lionel Dominjon, Anatole Lécuyer, J-M Burkhardt, Paul Richard, and Simon Richir. 2005. Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. In *IEEE proceedings. VR 2005. Virtual reality, 2005.*, 19–25.

79. Marco Donnarumma, Baptiste Caramiaux, Atau Tanaka, and others. 2013. Muscular interactions combining EMG and MMG sensing for musical practice. In *Proceedings of NIME*.

80. Paul Dourish and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on computer-supported cooperative work*, 107–114.

81. Pierre Dragicevic. 2016. Fair statistical communication in HCI. *Modern statistical methods for HCI*: 291–330.

82. Thierry Duval, Thi Thuong Huyen Nguyen, Cédric Fleury, Alain Chauffaut, Georges Dumont, and Valérie Gouranton. 2014. Improving awareness for 3D virtual collaboration by embedding the features of users' physical environments and by augmenting interaction tools with cognitive feedback cues. *Journal on Multimodal User Interfaces* 8: 187–197.

83. Sidney Fels, Ashley Gadd, and Axel Mulder. 2002. Mapping transparency through metaphor: Towards more expressive musical instruments. *Organised Sound* 7, 2: 109–126.

84. Robin Fencott and Nick Bryan-Kinns. 2010. Hey man, you're invading my personal space! Privacy and awareness in collaborative music. In *NIME*, 198–203.

85. Leah Findlater and Joanna McGrenere. 2010. Beyond performance: Feature awareness in personalized interfaces. *Int. J. Hum.-Comput. Stud.* 68, 3: 121–137. `https://doi.org/10.1016/j.ijhcs.2009.10.002`

86. Victor Frak, Tatjana Nazir, Michel Goyette, Henri Cohen, and Marc Jeannerod. 2010. Grip force is part of the semantic representation of manual action verbs. *PloS one* 5, 3: e9728.

87. Jules Françoise, Baptiste Caramiaux, and Frédéric Bevilacqua. 2012. A hierarchical approach for the design of gesture-to-sound mappings. In *9th sound and music computing conference*, 233–240.

88. A Cavan Fyans, Michael Gurevich, and Paul Stapleton. 2009. Spectator understanding of error in performance. In *CHI'09 extended abstracts on human factors in computing systems*, 3955–3960.

89. A Cavan Fyans, Michael Gurevich, and Paul Stapleton. 2010. Examining the spectator experience. In *Proceedings of new interfaces for musical expression*, 1–4.

90. Vittorio Gallese. 2003. The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 358, 1431: 517–528.

91. James J Gibson. 1977. The theory of affordances. *Hilldale, USA* 1, 2: 67–82.

92. T. C. Nicholas Graham, Leon A. Watts, Gaëlle Calvary, Joëlle Coutaz, Emmanuel Dubois, and Laurence Nigay. 2000. A dimension space for the design of interactive systems within their physical environments. In *Proceedings of the 3rd conference on designing interactive systems: Processes, practices, methods, and techniques* (DIS '00), 406–416. `https://doi.org/10.1145/347642.347799`

93. Saul Greenberg and Michael Boyle. 2002. Customizable physical interfaces for interacting with conventional applications. In *Proceedings of the 15th annual ACM symposium on user interface software and technology*, 31–40.

94. Scot Gresham-Lancaster. 1998. The aesthetics and history of the hub: The effects of changing technology on network computer music. *Leonardo Music Journal*: 39–44.

95. Stefan Greuter, Florian Mueller, and Thuong Hoang. 2022. Designing public VR installations: In *Designing interactive systems conference* (DIS '22), 792–806. `https://doi.org/10.1145/3532106.3533454`

96. Sebastian Günther, Julian Rasch, Dominik Schön, Florian Müller, Martin Schmitz, Jan Riemann, Andrii Matviienko, and Max Mühlhäuser. 2022. Smooth as steel wool: Effects of visual stimuli on the haptic perception of roughness in virtual reality. In *CHI conference on human factors in computing systems*, 1–17.

97. Michael Gurevich and A Cavan Fyans. 2011. Digital musical interactions: Performer–system relationships and their perception by spectators. *Organised Sound* 16, 2: 166–175.

98. Michael Gurevich, Paul Stapleton, and Adnan Marquez-Borbon. 2010. Style and constraint in electronic musical instruments. In *International Conference on New Interfaces for Musical Expression*, 106–111.

99. Martin Hachet, Benoit Bossavit, Aurélie Cohé, and Jean-Baptiste de La Rivière. 2011. Toucheo: Multitouch and stereo combined in a seamless workspace. In *Proceedings of the 24th annual ACM symposium on user interface software and technology*, 587–592.

100. John Hardy, Christian Weichel, Faisal Taher, John Vidler, and Jason Alexander. 2015. ShapeClip: Towards rapid prototyping with shape- changing displays for designers. In *Proceedings of the SIGCHI conference on human factors in computing systems* (CHI '15), xxx–xxx. `https://doi.org/10.1145/2702123.2702599`

101. Otmar Hilliges, David Kim, Shahram Izadi, Malte Weiss, and Andrew Wilson. 2012. HoloDesk: Direct 3D interactions with a situated see-through display. In *Proceedings of the SIGCHI conference on human factors in computing systems* (CHI '12), 2421–2430. `https://doi.org/10.1145/2207676.2208405`

102. David M Hoffman, Ahna R Girshick, Kurt Akeley, and Martin S Banks. 2008. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision* 8, 3: 33–33.

103. Andy Hunt, Marcelo M Wanderley, and Matthew Paradis. 2003. The importance of parameter mapping in electronic instrument design. *Journal of New Music Research* 32, 4: 429–440.

104. Nick Hwang, Anthony T Marasco, and Eric Sheffield. 2021. Collab-hub: A system for creating collaborative telematic music performances with web-based instruments and creative development platforms. In *Proceedings of the web audio conference*.

105. Robert J. K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. Reality-based interaction: A framework for post-WIMP interfaces. In *Proceedings of ACM CHI*, 201–210. `https://doi.org/10.1145/1357054.1357089`

106. Jacek Jankowski and Martin Hachet. 2013. A survey of interaction techniques for interactive 3D environments. In *Eurographics 2013-STAR*.

107. Marc Jeannerod. 2001. Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage* 14, 1: S103–S109. https://doi.org/`https://doi.org/10.1006/nimg.2001.0832`

108. Sergi Jordà. 2004. Digital instruments and players: Part i—efficiency and apprenticeship. In *Proceedings of the 2004 conference on new interfaces for musical expression*, 59–63.

109. Sergi Jordà. 2005. Multi-user instruments: Models, examples and promises. In *Proceedings of NIME 05*, 23–26. Retrieved from `http://dl.acm.org/citation.cfm?id=1085939.1085948`

110. Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. 2007. The reacTable: Exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st international conference on tangible and embedded interaction*, 139–146.

111. Byungdae Jung, Jaein Hwang, Sangyoon Lee, Gerard Jounghyun Kim, and Hyunbin Kim. 2000. Incorporating co-presence in distributed virtual music environment. In *Proceedings of the ACM symposium on virtual reality software and technology* (VRST '00), 206–211. https://doi.org/`http://doi.acm.org/10.1145/502390.502429`

112. Matthew Kay, Gregory L Nelson, and Eric B Hekler. 2016. Researcher-centered design of statistics: Why bayesian statistics better fit the culture and incentives of HCI. In *Proceedings of ACM CHI*.

113. Roberta L Klatzky and Susan J Lederman. 2010. Multisensory texture perception. In *Multisensory object perception in the primate brain*. Springer, 211–230.

114. Regis Kopper, Tao Ni, Doug A Bowman, and Marcio Pinho. 2006. Design and evaluation of navigation techniques for multiscale virtual environments. In *IEEE virtual reality conference (VR 2006)*, 175–182.

115. Dietrich Korsch. 2012. *Reflective optics*. Academic Press.

116. Amir Lahav, Elliot Saltzman, and Gottfried Schlaug. 2007. Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience* 27, 2: 308–314.

117. Otso Lähdeoja, Marcelo M Wanderley, and Joseph Malloch. 2009. Instrument augmentation using ancillary gestures for subtle sonic effects. *Proceedings of Sound and Music Computing*: 327–330.

118. Chi-Hsia Lai and Till Bovermann. 2013. Audience experience in sound performance. In *Proceedings of new interfaces for musical expression*, 170–173.

119. Jaron Lanier. 1993. The sound of one hand. *Whole Earth Review*.

120. Celine Latulipe, Erin A Carroll, and Danielle Lottridge. 2011. Love, hate, arousal and engagement: Exploring audience responses to performing arts. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 1845–1854.

121. Anatole Lécuyer. 2009. Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. *Presence: Teleoperators and Virtual Environments* 18, 1: 39–53.

122. Susan J Lederman and Susan G Abbott. 1981. Texture perception: Studies of intersensory organization using a discrepancy paradigm, and visual versus tactual psychophysics. *Journal of Experimental Psychology: Human perception and performance* 7, 4: 902.

123. Sang Won Lee. 2019. Show them my screen: Mirroring a laptop screen as an expressive and communicative means in computer music. In *NIME*, 443–448.

124. Jiajun Lu, Hrvoje Benko, and Andrew D. Wilson. 2017. Hybrid HFR depth: Fusing commodity depth and color cameras to achieve high frame rate, low latency depth camera interactions. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (CHI '17), 5966–5975. https://doi.org/10.1145/3025453.3025478

125. Joseph Malloch, Stephen Sinclair, and MarceloM. Wanderley. 2008. A network-based framework for collaborative development and performance of digital musical instruments. In *Computer music modeling and retrieval. Sense of sounds*. 401–425. https://doi.org/10.1007/978-3-540-85035-9_28

126. D Margery, B Arnaldi, and N Plouzeau. 1999. A general framework for cooperative manipulation in virtual environments. In *In proc. Of EGVE 99: Eurographics workshop on virtual environments*, 169–178.

127. Diego Martinez, Florent Berthaut, Abhijit Karnik, and Sriram Subramanian. 2014. Through the Combining Glass. In *ACM Symposium on User Interface Software and Technology*. Retrieved from https://hal.science/hal-01222458

128. Raul Masu, Adam Pultz Melbye, John Sullivan, and Alexander Refsum Jensenius. 2021. NIME and the Environment: Toward a More Sustainable NIME Practice. In *NIME 2021*.

129. Andrew P McPherson and Victor Zappi. 2015. Exposing the scaffolding of digital instruments with hardware-software feedback loops. In *International Conference on New Interfaces for Musical Expression*.

130. Liang Men and Nick Bryan-Kinns. 2019. LeMo: Exploring virtual space for collaborative creativity. In *Proceedings of the 2019 on creativity and cognition*. 71–82.

131. Daniel Mendes, Fabio Marco Caputo, Andrea Giachetti, Alfredo Ferreira, and J Jorge. 2019. A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In *Computer graphics forum*, 21–45.

132. Tim Merritt, Weiman Kow, Christopher Ng, Kevin McGee, and Lonce Wyse. 2010. Who makes what sound?: Supporting real-time musical improvisations of electroacoustic ensembles. In *Proceedings of the 22nd SIGCHI australia*.

133. Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12: 1321–1329.

134. Ali Momeni. 2015. Caress: An enactive electro-acoustic percussive instrument for caressing sound. In *Proceedings of the international conference on new interfaces for musical expression*, 245–250.

135. Gaku Narita, Yoshihiro Watanabe, and Masatoshi Ishikawa. 2015. Dynamic projection mapping onto a deformable object with occlusion based on high-speed tracking of dot marker array. In *Proceedings of ACM VRST*. `https://doi.org/10.1145/2821592.2821618`

136. Dan Newton and Mark T Marshall. 2011. Examining how musicians create augmented musical instruments. In *Proceedings of NIME*, 155–160.

137. Thi Thuong Huyen Nguyen and Thierry Duval. 2014. A survey of communication and awareness in collaborative virtual environments. In *2014 international workshop on collaborative virtual environments (3DCVE)*, 1–8.

138. Niels Christian Nilsson, Tabitha Peck, Gerd Bruder, Eri Hodgson, Stefania Serafin, Mary Whitton, Frank Steinicke, and Evan Suma Rosenberg. 2018. 15 years of research on redirected walking in immersive virtual environments. *IEEE computer graphics and applications* 38, 2: 44–56.

139. Jieun Oh, Jorge Herrera, Nicholas J Bryan, Luke Dahl, and Ge Wang. 2010. Evolving the mobile phone orchestra. In *NIME*, 82–87.

140. Stephen J Payne. 2007. Mental models in human-computer interaction. *The Human-Computer Interaction Handbook*: 89–102.

141. Karin Petrini, Frank E Pollick, Sofia Dahl, Phil McAleer, Lawrie McKay, Davide Rocchesso, Carl Haakon Waadeland, Scott Love, Federico Avanzini, and Aina Puce. 2011. Action expertise reduces brain activity for audiovisual matching actions: An fMRI study with expert drummers. *Neuroimage* 56, 3: 1480–1492.

142. Friedrich Platz and Reinhard Kopiez. 2012. When the eye listens: A meta-analysis of how audio-visual presentation enhances the appreciation of music performance. *Music Perception: An Interdisciplinary Journal* 30, 1: 71–83.

143. Courtney N. Reed, Charlotte Nordmoen, Andrea Martelloni, Giacomo Lepri, Nicole Robson, Eevee Zayas-Garin, Kelsey Cotton, Lia Mice, and Andrew McPherson. 2022. Exploring experiences with new musical instruments through micro-phenomenology. In *Proceedings of the international conference on new interfaces for musical expression*. `https://doi.org/10.21428/92fbeb44.b304e4b1`

144. Stuart Reeves. 2011. *Designing interfaces in public settings: Understanding the role of the spectator in human-computer interaction*. Springer Science & Business Media.

145. Stuart Reeves, Steve Benford, Claire O'Malley, and Mike Fraser. 2005. Designing the spectator experience. In *Proceedings of the SIGCHI conference on human factors in computing systems*.

146. Nathan Renney, Benedict Gaster, Tom Mitchell, and Harri Renney. 2022. Studying how digital luthiers choose their tools. In *Proceedings of the 2022 CHI conference on human factors in computing systems*, 1–18.

147. Vincent Reynaert. 2021. Étude de la fatigabilité des bras lors d'interactions gestuelles en mid-air au travers de la réalité virtuelle. Université de Lille. Retrieved from `https://theses.hal.science/tel-03580555`

148. Vincent Reynaert, Florent Berthaut, Yosra Rekik, and Laurent Grisoni. 2021. The Effect of Control-Display Ratio on User Experience in Immersive Virtual Musical Instruments. In *International Conference on New Interfaces for Musical Expression*. Retrieved from `https://hal.science/hal-03265169`

149. Vincent Reynaert, Florent Berthaut, Yosra Rekik, and Laurent Grisoni. 2021. The Effect of Rhythm in Mid-air Gestures on the User Experience in Virtual Reality. In *18th IFIP TC 13 International Conference on Human-Computer Interaction, INTERACT 2021*, 182–191. `https://doi.org/10.1007/978-3-030-85613-7/_13`

150. Alain Savouret. 2010. *Introduction à un solfège de l'audible: L'improvisation libre comme outil pratique*. Symétrie.

151. Ruben Schlagowski, Dariia Nazarenko, Yekta Can, Kunal Gupta, Silvan Mertes, Mark Billinghurst, and Elisabeth André. 2023. Wish you were here: Mental and physiological effects of remote music collaboration in mixed reality. In *Proceedings of the 2023 CHI conference on human factors in computing systems*, 1–16.

152. W Andrew Schloss. 2003. Using contemporary technology in live performance: The dilemma of the performer. *Journal of New Music Research* 32, 3.

153. Frederick Seddon and Michele Biasutti. 2009. A comparison of modes of communication between members of a string quartet and a jazz sextet. *Psychology of Music* 37, 4: 395–415. `https://doi.org/10.1177/0305735608100375`

154. Stefania Serafin, Cumhur Erkut, Juraj Kojs, Niels Chr. Nilsson, and Rolf Nordahl. 2016. Virtual reality musical instruments: State of the art, design principles, and future directions. *Computer Music Journal* 40, 3: 22–40. `https://doi.org/10.1162/COMJ_a_00372`

155. Maximilian Speicher, Brian D Hall, and Michael Nebeling. 2019. What is mixed reality? In *Proceedings of the 2019 CHI conference on human factors in computing systems*, 1–15.

156. Alan Sullivan. 2004. DepthCube solid-state 3D volumetric display. In *Stereoscopic displays and virtual reality systems XI*, 279–284.

157. Chuanwei Sun, Xiaowei Chang, Mingming Cai, and Jingao Liu. 2014. An Improved Design of 3D Swept-Volume Volumetric Display. *Journal of Computers* 9, 1: 235–242. `https://doi.org/10.4304/jcp.9.1.235-242`

158. Leon Szmedra and DW Bacharach. 1998. Effect of music on perceived exertion, plasma lactate, norepinephrine and cardiovascular hemodynamics during treadmill running. *International journal of sports medicine* 19, 1: 32–37.

159. Desney S. Tan, Brian Meyers, and Mary Czerwinski. 2004. WinCuts: Manipulating arbitrary window regions for more effective use of screen space. In *CHI '04 extended abstracts on human factors in computing systems* (CHI EA '04), 1525–1528. `https://doi.org/10.1145/985921.986106`

160. Luca Turchet, Rob Hamilton, and Anil Çamci. 2021. Music in extended realities. *IEEE Access* 9: 15810–15832.

161. Yusuke Ujitoko and Yuki Ban. 2021. Survey of Pseudo-Haptics: Haptic Feedback Design and Application Proposals. *IEEE Transactions on Haptics* 14, 4: 699–711. `https://doi.org/10.1109/TOH.2021.3077619`

162. Bradley W Vines, Carol L Krumhansl, Marcelo M Wanderley, Ioana M Dalca, and Daniel J Levitin. 2011. Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. *Cognition* 118, 2: 157–170.

163. Jonna K Vuoskoski, Marc R Thompson, Eric F Clarke, and Charles Spence. 2014. Crossmodal interactions in the perception of expressivity in musical performance. *Attention, Perception, & Psychophysics* 76, 2: 591–604.

164. Chen Wang and Pablo Cesar. 2017. The play is a hit: But how can you tell? In *Proceedings of the 2017 ACM SIGCHI conference on creativity and cognition*, 336–347.

165. Ge Wang, Nicholas Bryan, Jieun Oh, and Robert Hamilton. 2009. Stanford laptop orchestra(SLOrk). In *Proceedings of the international computer music conference*, 505–508.

166. Colin Ware, Kevin Arthur, and Kellogg S Booth. 1993. Fish tank virtual reality. In *Proceedings of the INTERACT'93 and CHI'93 conference on human factors in computing systems*, 37–42.

167. Daniel M Wegner and Thalia Wheatley. 1999. Apparent mental causation: Sources of the experience of will. *American psychologist* 54, 7: 480.

168. Gil Weinberg. 2005. Interconnected musical networks: Toward a theoretical framework. *Comput. Music J.* 29, 2: 23–39. `https://doi.org/10.1162/0148926054094350`

169. David Wessel and Matthew Wright. 2002. Problems and prospects for intimate musical control of computers. *Computer music journal* 26, 3: 11–22.

170. T. Aisling Whitaker, Cristina Simões-Franklin, and Fiona N. Newell. 2008. Vision and touch: Independent or integrated systems for the perception of texture? *Brain Research* 1242: 59–72. https://doi.org/`https://doi.org/10.1016/j.brainres.2008.05.037`

171. 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 Trans. Comput.-Hum. Interact.* 20, 6. `https://doi.org/10.1145/2530541`

172. 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. Retrieved from `https://hal.science/hal-03790049`

173. Robert J Zatorre, Joyce L Chen, and Virginia B Penhune. 2007. When the brain plays music: Auditory–motor interactions in music perception and production. *Nature reviews neuroscience* 8, 7: 547–558.