Spatial Interfaces and Interactive 3D Environments for Immersive Musical Performances

Florent Berthaut, Martin Hachet

▶ **To cite this version:**

HAL Id: hal-01374911
https://hal.inria.fr/hal-01374911

Submitted on 8 Oct 2016

Spatial Interfaces and Interactive 3D Environments for Immersive Musical Performances
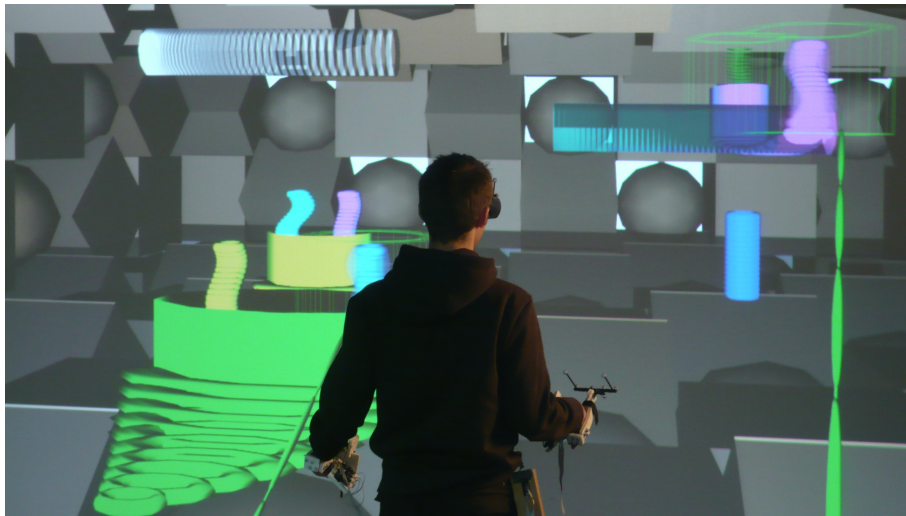
Florent Berthaut and Martin Hachet

Figure 1: A musician plays the Drile instrument while being immersed in front of a large stereoscopic projection. Using the *Piivert* device and techniques and the *Tunnels* modulation tools, he manipulates musical structures represented by *3D reactive widgets*.

1   Introduction

Interactive 3D graphics, immersive displays and spatial interfaces have shown undeniable benefits in a number of fields. In particular, theses technologies have been extensively used in industry, where real-time modifications of virtual mock-ups as well as immersive visualization allows for the optimization of conception cycles. On the other hand, the power of such technologies and interfaces is still under-explored in domains where the main target is to enhance creativity and

emotional experiences. This article presents a set of work we conducted with the final goal of extending the frontiers of music creation, as well as the experience of audiences attending to digital performances.

Our general approach is to connect sounds to interactive 3D graphics with which musicians can interact, and that can be observed by the audience. For example, imagine that a musician faces a huge stereoscopic screen where large composite 3D objects represent musical structures composed of multiple sounds. She can navigate around and interact with these structures. By looking at the appearance of the visual objects, she can easily infer which sound is associated to which visual object. She can select an object and move it to specific places to modify the related sound. For example, by bringing the object closer to her, she increases the amplitude of the sound. By sliding it through a dedicated tool, she modifies its pitch. For any modification, the visual appearance of the object is modified accordingly. Now, she wants to play with other musicians, co-located in the same space, or at a distance. She follows what the others are doing just by looking at their audio-visual 3D objects and the virtual tools that they use. She can prepare musical structure for them to play, too. During all this process, the audience benefit from a rich visual feedback that allows them to follow what musicians are doing.

We are convinced that such interactive 3D environments and spatial interfaces may serve the purpose of expressiveness, creativity and rich user experiences. Compared to more traditional digital music performances, they opens new opportunities for music creation. On the other hand, the use of these technology also opens new research questions that need to be tackle with a great care. Indeed, the final user experience depends on the efficiency of each level of the musical application, from very technological considerations such as the latency of the controllers, to human factor aspects such as the choice of audio-visual mappings, passing through the design of the best suited interaction techniques.
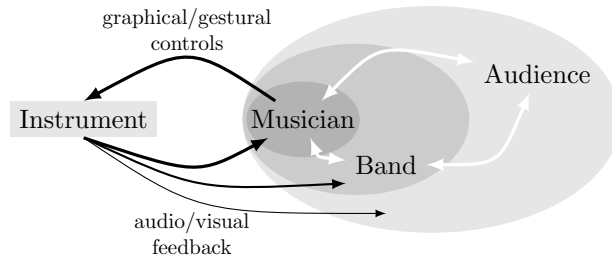
Figure 2: Interactions occur between each circles, and between the circles and the instrument.

To address these questions, we defined three circles for interactive 3D musical performances. The first circle is dedicated to the musician where he or she needs to interact with the sounds in a very precise and expressive way. In the second circle, we take into account the band. The challenge is thus to favor relevant collaborations between the musicians. Finally, the audience is part of a third

circle. Its members should experience rich immersion in the performance. We target advanced interactions between each of the circles on the one side, and the audio-visual content on the other side, both at an input and output level. Of course, it also exists strong interactions between the circles themselves. For example, depending on the actions of a musician, the band reacts, which in turn has an impact on the audience.
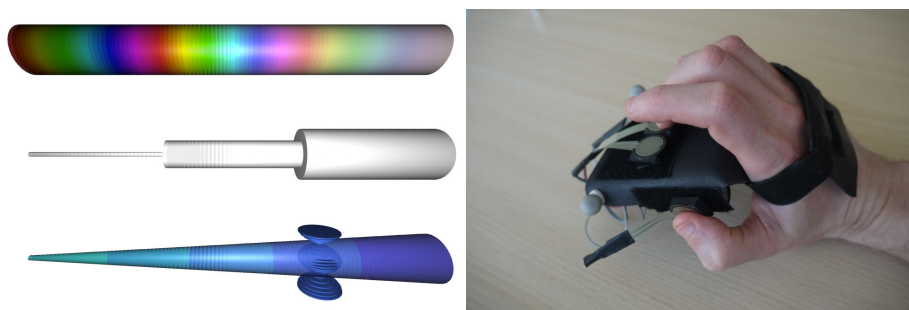
2 First circle: The musician

Figure 3: Left: Audio/visual *Tunnels* dedicated to sound modulations: Color/pitch (top), size/volume (middle) and combined pitch/volume (bottom), with a *reactive widget* being manipulated. Right: The *Piivert* input device with markers for 3D interaction and pressure sensors for musical gestures.

The first circle connects the musician with his or her instrument.

With acoustic instruments, musicians perform gestures which generate (excite) or modulate sounds. The energy of gestures is mechanically converted into vibrations and modulations. With Digital Musical Instruments (DMIs), this physical link is broken, and needs to be rebuilt, in order to restore energy transfer to and from the instrument. This also means that the *mappings* between gestures and changes in the sound can be defined freely. The amount of change in the music is not bound to the energy of the gesture, thus a single gesture may have any impact on any aspect of the sound, and even on multiple sounds at once. To create and modulate sounds with DMIs, the standard approach is to control each of the sound parameters by way of dedicated devices (e.g. mixers and MIDI controllers). Virtual approaches have also been proposed where sounds are controlled by way of virtual sliders and knobs operated by a standard pointing or touch device. Contrary to hardware sensors, these virtual components can be dynamically created and adapted in order to facilitate the control of multiple sound processes.

To extend the possibility of standard approaches, we have explored the control of sounds by way of 3D graphical representations we called *3D reactive widgets*. These audio-visual objects provide some of the feedback on the sound

that is lost with digital instruments and allow sounds to be manipulated through adapted 3D interaction techniques.

They provide visual feedback on the value of sound parameters. To determine the relevant mappings between the sounds and their visual representations, we conducted a set of psychometric experiments [1]. For example, these experiments showed that users preferences followed physical principles, such as having the volume mapped to the size of the widgets, and semantic ones, such as mapping the pitch to the color brightness of the widget. They also showed that multiple audiovisual mappings can be combined on a single *3D reactive widget* without degrading perception, for example to allow users to visually perceive the values of both volume and pitch parameters of an audio loop simultaneously.

Manipulating a sound then amounts to manipulating the appearance of the associated widget, making the musical interaction more transparent to the user. We designed graphical modulation tools called *Tunnels* [3] for changing the appearance of the *3D reactive widgets*, and consequently the sound they embed. Each *Tunnel* displays a scale of values, discrete or continuous, for one or several graphical parameters. Figure 3 left for example shows three *Tunnels*. The one on top changes the color along a continuous scale, which is mapped to the pitch of the sound. The one in the middle changes the size along a discrete scale, which is mapped to the volume. The one on the bottom changes both the pitch along a discrete scale and the volume along a continuous scale. *Tunnels* behave like virtual sliders. When a musician drags a *3D reactive widget* inside a *Tunnel* with a color scale ranging from dark red to light blue, the color of the widget is set accordingly, consequently modifying the pitch of the associated sound. Research on musical have demonstrated the importance of one-to-many mappings [7], where one gesture controls several sound parameters, which lead to more expressive instruments. The *Tunnels* allow for controlling several sound parameters at once with complex scales, but at the same time provide visual feedback of the value of each parameter separately.

More than 2D graphical interfaces, 3D User Interfaces are well suited to represent, navigate and interact in scenes with complex 3D structures. We take advantage of this to extend the musical technique of live-looping with the "Drile" instrument [2]. The hierarchical live-looping technique used in "Drile" allows musicians to build and manipulate complex musical tree structures of looped musical manipulations. Figure 1 shows a musician interacting with two musical trees.

However, 3D selection, manipulation and navigation techniques and devices need to be adapted for expressive musical interaction. To that extent, we developed *Piivert* [?], a 3D interaction device and set of techniques that take into account these specific requirements. *Piivert* is depicted on Figure 3. First it divides interaction according to well-known instrumental gestures categories [6] and their temporal and physical constraints. Excitation gestures, which generate sound, require low latency and are therefore performed on *Piivert* using pressure sensors located below each finger. These also provide the passive haptic feedback needed for precise instantaneous gestures (taps, rolls ...). On the contrary, modulation (changing the sound) and selection (choosing part of the

4

instrument) gestures are done through the 3D interface, respectively with the *Tunnels* and the Virtual Ray technique. To provide additional feedback to the musician, we extended the standard virtual ray metaphor by modifying the appearance of the ray according to the amount of energy sent to excite a *reactive widget* when hitting or pressing the pressure sensors.

The results of a study we conducted suggest that *Piivert*, by separating excitation and selection gestures, increases the temporal accuracy and reduces the error rate in sound playing tasks in an immersive environment.

In order to perform high-level commands on the instrument, whilst keeping with the temporal constraints of musical gestures, *Piivert* provides a vocabulary of percussion gestures. For example, flams (two taps in a fast sequence) with different fingers and in different orders can be used to start/stop the recording of a loop, delete it, or activate different playing modes for the other fingers. Using both hands and gestures such as flams (two fingers) and rolls (three or more fingers), a large set of commands can be triggered while preserving normal playing of notes and chords with individual or simultaneous fingers hits.

3 Second circle: The band

Figure 4: Left: Drile played by two musicians on each side of a semi-transparent 3D screen. The musician on the other side of the screen has selected a *reactive widget* using virtual rays from his both hands. Right: Two musicians interact through the Reflets system. One musician uses a 3D interface to grab and process the sound of the guitar played by the musician reflected in the mirror.

The second circle connects the instrument and musician to the other musicians of the band.

In acoustic orchestras, non-verbal communication between musicians allows them to synchronize their actions, follow a musical structure, exchange and improvise together. With Digital Orchestras (DOs), there is a loss in the awareness of what other musicians are playing making, synchronization and exchanges more difficult.

For instance, Figure 4 depicts the Drile instrument being played by two musicians, using a two-sided semi-transparent screen. An optical combiner is

placed at a 45 degrees angle, with projector screens below and above it, forming a Z. Each side of the combiner only reflects one projection, therefore displaying the instrument only from the corresponding musician's point of view. With this setup, musicians can both perceive the virtual components of the instruments and each other directly.

3D user interfaces also open new possibilities for musical collaboration in DOs. For example, musicians may cooperate on the same musical processes and parameters, by interacting with the same 3D *reactive widget* placed in a shared virtual space. For example, in Drile, an expert musician may prepare loops that they then pass on to other musicians. Furthermore, different interaction techniques and/or access to the musical structure can be given to the musicians depending on their expertise. In Drile again, expert musicians may access the whole trees of loops, while beginners may only access the higher-level nodes, which require less complex gestures in order to produce musically satisfying results.

With the Reflets project [5], we push the collaboration further. As depicted in Figure 4, a large vertical optical combiner is placed between the musicians of the band, combining spaces on each side of it. Musicians therefore perceive their reflections next to or overlapping the musicians on the other side. Reflets enables collaboration with both physical and virtual instruments. Figure 4 shows a scenario with a guitarist and another musician playing a gestural controller. Short loops from the guitar can be grabbed by the other musician simply by reaching through the reflection of the guitar, and manipulated through gestures within a control box. With Reflets, the 3D interface provides both visual feedback and novel collaboration opportunities, while preserving non-verbal communication between the musicians. Various other scenarios for collaboration were explored during workshops with musicians, dancers and circus artists from the Bristol artistic scene, leading to public performances.

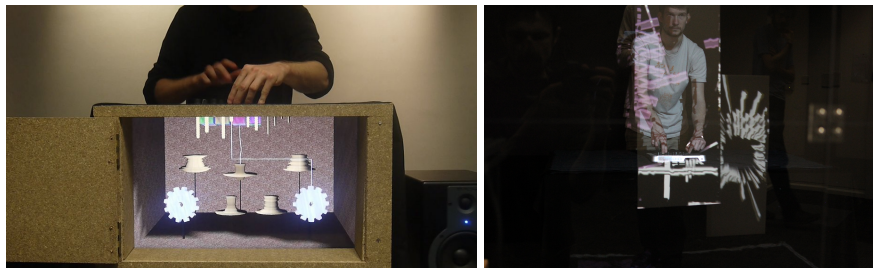# 4   Third circle: The audience

Figure 5: Left: A Digital Musical Instrument is augmented by the Rouages system, revealing its mechanisms. Right: Reflets allows spectators to explore these mechanisms through a large scale semi-transparent mirror.

The third circle adds the spectators to the digital performance equation.

In performances with acoustic instruments, the visual feedback, such as musicians' instrumental gestures but also general body movements, has been shown to have a strong impact on the emotion perceived and felt by the audience. With DMIs, this visual component is greatly impaired. Due to the variety of physical interfaces and sound synthesis/processing techniques, it is very hard for the audience to perceive the relation between musicians' gestures and the musical result. Many DMIs also feature automated sound processes, so that the musical result is not only dependent on the gestures performed. Finally, the mappings between sensor values and sound parameters values can be very complex, with changes in scale and cardinality. The *familiarity* that spectators have with acoustic instruments that they have played or seen played before, and with the physical principles that they experience in everyday life, does not exist anymore with digital ones. This leads to the well-known issue of *liveness*. Not perceiving the engagement of musicians with DMIs, i.e. how much they are in control of the music being played, may degrade the experience spectators have during performances.

Our approach is to augment the instrument from the audience point of view, while preserving the interface designed for the musician's expression. With the *Rouages* project [4] depicted on Figure 5, we propose to reveal the mechanisms of DMIs using an augmented reality display by : i) amplifying gestures with virtual extensions of the sensors, ii) representing the sound processes and the amount of control they require and iii) revealing the links between the sensors and sound processes.

The feedback from audience members of demonstrations and public performances was generally positive, with spectators commenting on the fact that they could more easily understand what was happening in the instrument, and what was the actual impact of musicians' gestures. In addition, the results of a study suggest a positive effect on the audience perception. We specifically designed DMIs that represented commonly found issues. We showed videos of performances with these DMIs, with and without visual augmentations, to participants. We then asked them to rate the perceived control and their confidence in their rating. Augmentations had a significant positive impact on the rating of perceived control when they represented changes in scale or in nature between gestures and resulting changes in the sound, for example when a hit gestures triggers a continuous change of pitch. Also audience members were more confident in their ratings when the changes were in part automated and in part done by the musicians, meaning that they perceived better what was the exact impact of the musicians.

In order to be used in actual performances, these 3D augmentations need to be perceived consistently by all members of the audience, no matter their position in front of the stage. With *Reflets*, we propose a novel mixed-reality display that uses the specific configuration of performances. It relies on spectators revealing the augmentations on the stage side of the optical combiner by intersecting them with their bodies or props from the other side. During a performance, they may therefore explore the inner mechanisms of the instruments

being played. Because of the flat optical combiner, 3D content revealed by one spectator is visible and appear consistently for all spectators. Figure 5 shows a spectator revealing augmentations, extensions of the sensors and representations of a loop and a sound sample, of a DMI using large white panels.

On the contrary to many DMIs, 3D virtual instruments such as Drile already provide visual feedback useful for the audience on the links between gestures and the sound parameters, for example through graphical tools such as *Piivert* and the *Tunnels*. However, the scenography of performances with these instruments need to fill a number of requirements, such as musicians' immersion, audience immersion, musician visibility, audience visibility and continuity between physical gestures and virtual tools. For example, the same screen cannot be used for both the musician and spectators, as the rendered perspective is adjusted for the musician as he moves, and therefore does not match the average audience viewing position. To cope with this issue, we can setup two separate screens which mark out the virtual space. The musician screen renders the 3D environment with stereoscopy and head-tracking, providing correct depth perception of the instrument. The audience screen renders the scene from the side and a point of view in the center of the spectators. They perceive both the physical musician, his gestures and the virtual rays that he manipulates to interact with the 3D musical structures.

5   Conclusion

Interactive 3D environments and spatial interfaces open up very interesting opportunities for musical performances. They offer novel playgrounds to musicians who can explore new dimensions in music creation. They also favor the emergence of interactive installations where audiences can experience new forms of performances. Exploring new directions in immersive musical performances is also fruitful for research in spatial interfaces: it feeds challenging research questions that can find interesting applications outside the scope of music.

References

[1] Florent Berthaut, Myriam Desainte-Catherine, and Martin Hachet. Combining audiovisual mappings for 3d musical interaction. In *Proceedings of the International Computer Music Conference (ICMC10)*, pages 357–364, New York, USA, 2010.

[2] Florent Berthaut, Myriam Desainte-Catherine, and Martin Hachet. Drile : an immersive environment for hierarchical live-looping. In *Proceedings of NIME*, pages 192–197, Sydney, Australia, 2010.

[3] Florent Berthaut, Myriam Desainte-Catherine, and Martin Hachet. Interacting with 3D Reactive Widgets for Musical Performance. *Journal of New Music Research*, 40(3):253–263, 2011.

[4] Florent Berthaut, Mark Marshall, Sriram Subramanian, and Martin Hachet. Rouages: Revealing the Mechanisms of Digital Musical Instruments to the Audience. In *Proceedings of NIME*, Daejeon, South Korea, 2013.

[5] Florent Berthaut, Diego Martinez, Martin Hachet, and Sriram Subramanian. Reflets: Combining and revealing spaces for musical performances. In *International Conference on New Interfaces for Musical Expression*, 2015.

[6] Claude Cadoz. *Les nouveaux gestes de la musique*, chapter Musique, geste, technologie, pages 47–92. Éditions Parenthèses, 1999.

[7] Andy Hunt and Ross Kirk. Mapping strategies for musical performance. *Trends in Gestural Control of Music*, pages 231–258, 2000.