



3D interaction techniques for musical expression

Florent Berthaut

► To cite this version:

Florent Berthaut. 3D interaction techniques for musical expression. Journal of New Music Research, inPress, 10.1080/09298215.2019.1706584 . hal-02430634

HAL Id: hal-02430634

<https://hal.science/hal-02430634v1>

Submitted on 7 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3D Interaction techniques for Musical Expression

Florent Berthaut

Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France

ARTICLE HISTORY

Compiled September 14, 2019

ABSTRACT

As Virtual Reality headsets become accessible, more and more artistic applications are developed, including immersive musical instruments. 3D interaction techniques designed in the 3D User Interfaces research community, such as navigation, selection and manipulation techniques, open numerous opportunities for musical control. For example, navigation techniques such as teleportation, free walking/flying and path-planning enable different ways of accessing musical scores, scenes of spatialized sounds sources or even parameter spaces. Manipulation techniques provide novel gestures and metaphors, e.g. for drawing or sculpting sound entities. Finally, 3D selection techniques facilitate the interaction with complex visual structures which can represent hierarchical temporal structures, audio graphs, scores or parameter spaces. However, existing devices and techniques were developed mainly with a focus on efficiency, i.e. minimising error rate and task completion times. They were therefore not designed with the specifics of musical interaction in mind. In this paper, we review existing 3D interaction techniques and examine how they can be used for musical control, including the possibilities they open for instrument designers. We then propose a number of research directions to adapt and extend 3DUIs for musical expression

KEYWORDS

3D user interfaces, Immersive virtual musical instruments, Virtual Reality, New interface for musical expression

1. Introduction

Historically created with applications such as industrial design in mind, Virtual Reality was rapidly appropriated for artistic purposes. In 1993 Jaron Lanier presents *The sound of one hand* (Lanier (1993)), a musical instrument which allows one to navigate in a Virtual Environment (VE) and manipulate different instruments using a virtual hand technique. However, until the recent development of affordable 3D display and tracking technologies, the design and use of musical instruments based on VEs and 3D graphical interaction techniques has remained mostly limited to researchers of the Virtual Reality and New Interfaces for Musical Expression fields. Such instruments have been designated as Virtual Musical Instruments in Mulder (1998) and Pressing (1997), Immersive Virtual Musical Instruments (IVMI) in Berthaut, Zappi, and Mazzanti (2014) or Virtual Reality Musical Instruments in Serafin, Erkut, Kojas, Nilsson, and Nordahl (2016). As expensive and sizeable Cave Automatic Virtual Environments

and large 3D projections are replaced by more compact commercially available Head Mounted Displays with integrated tracking, IVMIs are being appropriated by musicians and instrument designers.

Commercial musical applications have therefore been proposed. Some can be designated as musical games, where expression remains limited, such as Beat Sabers Beat Games (n.d.). Other applications target music production in VR such as Sound-Stage Hard Light Labs LLC (n.d.), Modulia Studio Thérémix Tech (n.d.), AliveInVR WeARVR (n.d.) or The Music Room The Music Room (n.d.).

Others are complete performance spaces in which both spectators and artists are immersed such as Electronauts Survios (n.d.) or WaveXR WaveVR (n.d.).

In parallel, research on 3D user interfaces (3DUI) has led to a great number of technique for the selection, manipulation and navigation of visual objects in VEs. However, still very little research has been conducted on how to use these techniques for expressive musical interaction. Moreover, existing IVMIs and especially publicly released ones such as those described above mostly focus on mimicking physical interfaces, relying on 3D sliders, buttons and keys.

In this paper, we discuss opportunities for musical expression in VEs brought by existing 3D interaction techniques and possible extensions of these techniques with musical interaction in mind, in the hope that these will inform the design of future IVMIs.

In the fields of Virtual Reality and 3DUIs, D. A. Bowman, Kruijff, LaViola, and Poupyrev (2004) provides a global review of interaction techniques. More recently, Jankowski and Hachet (2013) focus on 3D navigation techniques and the use of 2D input devices, and Argelaguet and Andujar (2013) focus on 3D selection techniques. Mendes, Caputo, Giachetti, Ferreira, and Jorge (2019) reviews 3D manipulation techniques for desktop, touchscreen and immersive environments. Here we will review interaction techniques that can be used for musical purposes and investigate how to expand their expressive capabilities. On the musical interaction side, Pressing (1997) gives a first analysis of the opportunities and drawbacks of IVMIs. Serafin et al. (2016) provide an extensive review of IVMIs. The authors focus on general guidelines for design and include instruments without visual feedback. Berthaut et al. (2014) classify IVMIs based on scenographic choices when performing them, including the degree of immersion of both musicians and spectators. Instead we focus only on the interaction aspect and instruments which involve graphically selecting, manipulating and exploring musical processes and structures which are visually rendered in the form of 3D elements of a virtual environment. We therefore also consider both virtual, augmented and mixed reality displays, but also non-immersive setups, for example with a monoscopic monitor with keyboard and mouse input or a touchscreen. In this paper, we will refer to these instruments as *3D Musical Instruments* (3DMIs).

Our contribution is two-fold.

- First, we review 3DUI techniques developed in research or industrial contexts, highlighting their potential for musical expression and providing examples from existing instruments.
- Second, we propose a number of research directions to further adapt and extend 3DUIs for musical expression.

2. 3D Interaction techniques and opportunities for musical interaction

3D Interaction techniques are generally divided into navigation, selection, manipulation and application control, as stated by D. A. Bowman et al. (2004). We focus only on the first three, leaving out 3D menus which do not open specific opportunities for expression compared to their 2D counterpart.

This classification finds an interesting parallel in the selection / modification / excitation musical gestures categories by Cadoz (1999).

In this section, we first analyse how sound can be represented within a VE so that it can be interacted with. We then describe each category of interaction techniques, not going into the details of optimizing for performance but rather looking at the musical expression opportunities that they afford. We illustrate each section with pictures and screenshot from the Drile instrument. Finally, in order to visualise how existing 3DMIs have taken advantage of 3DUIs and to reveal what opportunities remain for future explorations, we propose in Table 1 a (non-exhaustive) list of 3DMIs which highlights for each of them the interaction techniques used.

2.1. Graphical representations of sound processes and parameters in VEs

So they can be manipulated, musical processes and parameters need to be represented as 3D graphical elements of the VE. From a computer graphics perspective, these elements can be hollow 3D meshes, whose surfaces are defined by vertices with 3D coordinates, or full volumes defined by sets of voxels, such as the ones obtained from tomographic imaging. We can then separate the representations of musical processes and parameters into three categories.

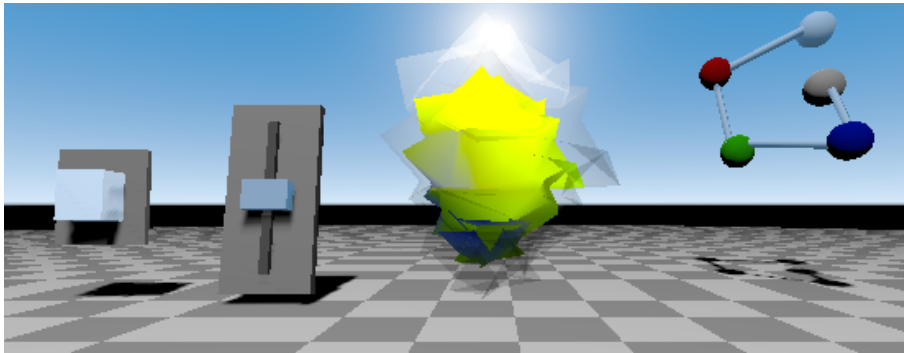


Figure 1. Representations of sound processes and parameters : Virtual controls with button and slider (left), Dynamic/Reactive widget (center), Structure (right)

2.1.1. Virtual sensors

Representations of the first type mimic sensors of physical control surfaces and traditional 2D graphical user interfaces. These include 3D buttons, linear and angular potentiometers, XY (2-dimensional) pads and extensions of these such as 3-dimensional control volumes where a sound parameter can be assigned to each axis. 3DMIs such as WAVE by Valbom and Marcos (2005) make for example use of 3D buttons for the control of audio files and midi loops and notes. Although they provide a lower entry fee (Wessel and Wright (2002)) for musicians due to their familiar aspect and behaviour,

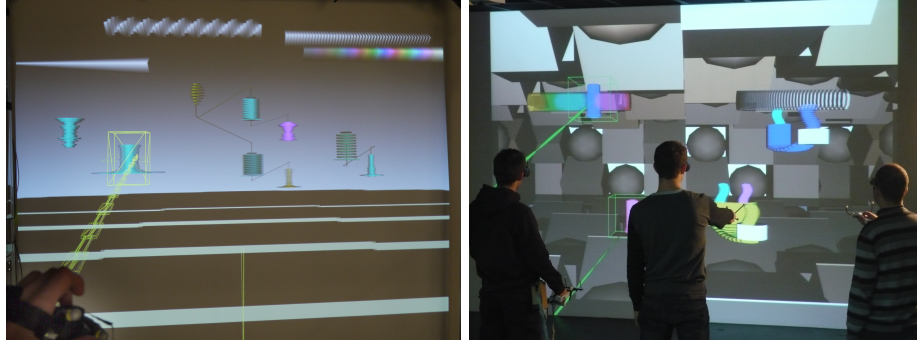


Figure 2. Left : In Drile, hierarchical live-looping structures with sound processes are displayed as trees of 3D reactive widgets. The graphical parameters (e.g. color, transparency, size, shape, orientation) of these widgets are connected to sound parameters. Right: Examples of selection techniques used in Drile : Virtual rays for the leftmost musician (in green, one for each hand) with added visual feedback on excitation gestures, 3D cursors (in white) for the two other musicians which limit them to process level controls

these representations may limit the control and feedback possibilities offered by 3DUIs.

2.1.2. *Dynamic/Reactive widgets*

Representations of the second type take advantage of visualisation possibilities offered by VEs. These visual objects can change shape and material in order to provide feedback on musical processes, e.g. reflecting loudness, spectrum, phase and so on. They also offer more advanced control possibilities by relying on physics such as the projectiles used in Q3osc by Hamilton (2008). Finally they may combine feedback and control, as in the case of the 3D reactive widgets proposed by Berthaut, Desainte-Catherine, and Hachet (2011) and depicted on Figure 2 left.

2.1.3. *Structures and scenes*

The third type of representation are groups of 3D visual elements. These representations are fairly common in more traditional applications of virtual reality, such as the visualisation of chemical and biological structures or architectural and automobile design. They may be connected, allowing for the visualisation of complex musical structures, such as the live-looping trees by Berthaut, Desainte-Catherine, and Hachet (2010), which are depicted on Figure 2 left. They may also be freely arranged within the VE. For example they are placed on axes of concatenative synthesis in VrGrains by Zappi, Mazzanti, Brogni, and Caldwell (2012) or compose a score of spatialized sound sources in Versum by Barri (2009). In addition, these structures and scenes can be built with sets of reactive / dynamic 3D widgets, augmenting their feedback and control capabilities.

2.2. *3D Selection*

The first category of 3D interaction techniques is the 3D selection. Its objective is to allow users to designate one or several elements of a VE with which they are going to interact.

Several taxonomies have been proposed for selection techniques. D. A. Bowman and Hodges (1997) analyse them by decomposing the underlying tasks, namely object indication, confirmation of selection and feedback. Poupyrev, Weghorst, Billingham, and

and Ichikawa (1998) organize them by selection metaphor. A first level separates ego-centric and exocentric metaphors, the former then including virtual hand and pointer metaphors while the latter is composed of world-in-miniature and automatic scaling. Argelaguet and Andujar (2013) propose a classification based on the components and mechanisms of selection techniques : selection tool (e.g. rays, cones, cubes and spheres), selection tool control (e.g. hand, hand+viewpoint), selection tool degrees of freedom, control display ratio (i.e. is the transformation of the selection tool amplified or reduced compared to physical movements of the user), motor and visual spaces relationship, disambiguation mechanism (i.e. how to avoid selecting multiple objects in dense VEs), selection trigger and feedback. Argelaguet and Andujar (2013) defines four usability goals for selection techniques, namely that they should be rapid, accurate and error proof, easy to understand and control and yield low levels of fatigue.

From a musical point of view, 3D selection techniques can be used in two ways. They can serve as the equivalent of musical selection gestures defined by Cadoz (1999). In that case, they do not directly produce changes in the sound, but rather allow one to choose components of an instrument on which excitation (i.e. triggering sound) and modification (i.e. modifying properties of the sound) are going to be performed. However, 3D selection gestures can also serve as excitation gestures, either instantaneous and triggered when entering a 3D audiovisual object or continuous when moving within a selected object.

In the following subsections, we describe selection techniques organized by object indication metaphor, provide examples of existing instruments making use of them and demonstrate the expressive possibilities they afford.

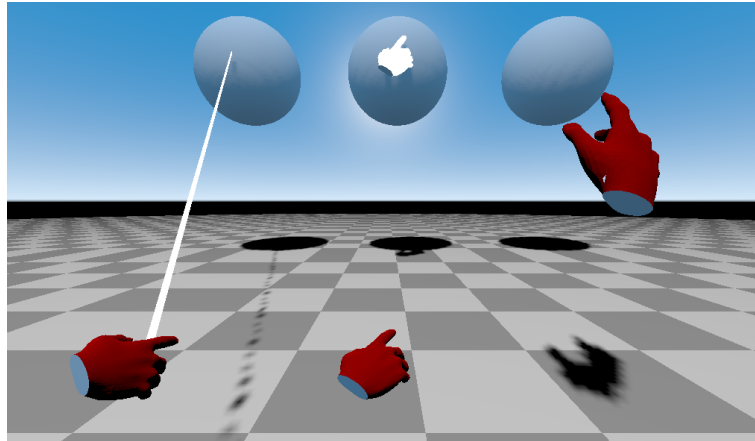


Figure 3. Illustration of the main categories of selection techniques (the user’s hand is displayed in red): virtual ray (left), virtual cursor (center), occlusion (right)

2.2.1. Virtual hand / 3D cursor

The first category of selection techniques relies on a 3D cursor which the user transforms (i.e. translate, rotate, scale) so that it overlaps with the chosen object. This cursor can have various shapes, from a 3D hand to a simple box, which can in turn be static or dynamic.

Because they rely on collision with virtual objects for selection, techniques from this category require that the cursor can be moved anywhere these objects are. While this is not an issue for virtual objects within arm reach or when navigating to distant objects

before selection is possible, it is often the case that users want to select distant objects. In that case, it is necessary to decouple the control, i.e. physical gestures performed to move the cursor, and the display, i.e. the movement of the cursor in the VE, for example by modifying the control-display ratio in order to allow one a larger working area than the one physically reachable. For instance, the Go-Go technique proposed by Poupyrev, Billinghurst, Weghorst, and Ichikawa (1996) increases the control-display ratio when the user’s arm extension reaches a threshold, allowing them to select out of reach objects, but at the same time reducing the selection accuracy.

In existing 3DMIs, Mäki-Patola, Laitinen, Kanerva, and Takala (2005) uses a virtual hand for selecting virtual xylophone bars which are then freely placed in the VEs. The same technique is used in Naef and Collicot (2006) to select sound sources. In Valbom and Marcos (2005), the authors make use of virtual cursors to select 3D widgets which act as controls of midi and audio processes. Berthaut et al. (2010) proposes to use 3D cursors for novices users of the Drile instrument as depicted on Figure 2 right. In that case, the control-display ratio is fixed on purpose, so that novices can only select and manipulate the closest elements of tree structure, which give access to high-level controls over the musical structure, but are not able to select more distant elements, therefore preventing them from modifying the structure itself.

2.2.2. Virtual pointing / ray

The second type of object indication possibility is the virtual ray or virtual pointing. In its simplest version, the ray origin and direction are respectively controlled by the user’s hand position and orientation. There are several issues which arise when using virtual ray techniques. The first is the lack of precision at a distance. Being mostly controlled through orientation, for a given rotation the displacement of the ray is larger at a distance than it is at close range, therefore reducing the accuracy for the selection of small distant targets or in cluttered scenes. The second is the occlusion of a target by another object along the ray, preventing the user from directly selecting it. Several techniques have been proposed to remedy these issues, some of which we describe in section 3.4. The ray itself can be a simple straight line, a dynamically bended line (de Haan, Koutek, and Post (2005)), or a cone as proposed by Steed (2006).

Berthaut et al. (2011) propose to use virtual rays as part of the Piivert interaction device and technique. Musicians use them to select a reactive widget which is then triggered using pressure sensors. This device is depicted on Figure 7. Additional visual feedback shows the activity of pressure sensors on the rays, in the form of propagating oscillation in radius. These are depicted on Figure 2 right, used by the leftmost musician. Mäki-Patola et al. (2005) use virtual rays as sticks to hit 3D xylophone bars.

2.2.3. Occlusion / Image-plane

The last category of object indication metaphors rely on a ray projected from the user’s viewpoint through a cursor either aligned with the user’s hand or moved by it, as described by Pierce et al. (1997). Interestingly, the shape of the cursor can be used to disambiguate selection, for example limiting it to objects which fit in the space between the thumb and index as in the HeadCrusher technique.

To our knowledge there has not been any use of occlusion / image-plane selection techniques in the context of musical interaction in immersive environments, although they provide the possibility of using subtle musical hand and finger gestures to select distant 3D objects, therefore combining intimate and large scale interactions. However,

occlusion metaphors also include shooting metaphors in which projectiles (dynamic widgets) are thrown along a direction within the VE, and trigger sound events depending on the targets that they reach such as q3osc by Hamilton (2008).

2.3. 3D Manipulation

In this section, we present interaction techniques which relate to the manipulation of elements of the VE. They generally follow a selection phase, during which one or several objects were chosen to be manipulated. These can be subdivided into three categories: spatial transformations, including rotation, scaling and translation, shape and structure manipulation, and material manipulation.

Musical uses of object manipulations are very diverse. By associating objects shape, position, orientation, scale and the various parameters defining their appearance (e.g. colour, texture, transparency) with musical parameters, they allow for complex mappings between gesture and sound, as recommended by Hunt and Kirk (2000). In addition, these parameters can be uni (e.g. movement on a single axis) or multi-dimensional (shape) and therefore

Groups of virtual objects can also be used to manipulate musical structures such as scores, audio graphs or spatialized scenes. The same properties can be used to provide feedback on the status of musical parameters and processes.

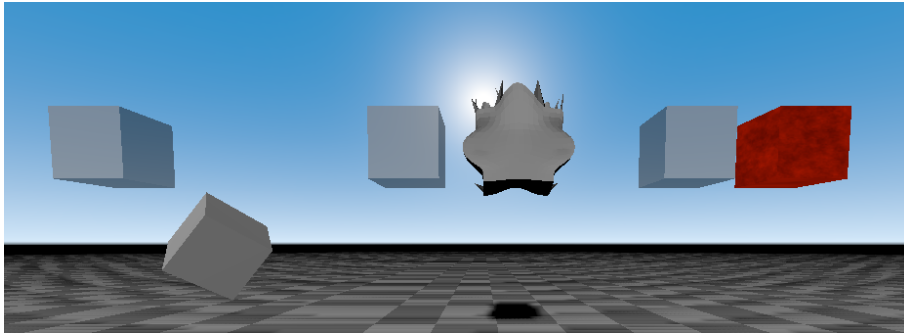


Figure 4. Illustration of the main categories of manipulation techniques on a box : spatial transformation (left), shape manipulation (center), material manipulation, here texture (right)

2.3.1. Spatial transformation

The most important category of manipulation techniques addresses the issue of placing virtual objects in the environment. For an extended review of them, refer to Mendes et al. (2019). These spatial transformations can be performed through mid-air manipulation, such as simply attaching the object to the virtual ray or cursor used for selection and moving it accordingly. If it is attached to a virtual hand as in the case of the HOMER technique by D. A. Bowman and Hodges (1997), the object can be both translated and rotated at the same time. Bimanual interaction can finally be used to perform all three spatial transformations at once. In the case of 2D inputs, similar interactions can be done using multitouch gestures, as demonstrated with the Toucheo system by Hachet, Bossavit, Cohé, and de la Rivière (2011) which allows for the manipulation of a 3D object above a multitouch surface. Another possibility is the use of 3D transformation widgets such as the tBox proposed by Cohé, Dècle, and Hachet (2011) for touchscreen interaction.

In a musical context, Zappi, Mazzanti, Brogni, and Caldwell (2011) describe a virtual reality performance where 3D objects can be moved through the environment to change the spatial position of the associated sound sources. A similar interaction is also proposed in Naef and Collicot (2006). Mäki-Patola et al. (2005) use spatial transformations to arrange virtual xylophone bars in the VE.

2.3.2. Shape/structure manipulation

The second category of manipulation techniques involves controlling the shape of deformable object and the structure of composite objects. Many tools exist for direct sculpting of 3D surfaces, from a 2D or 3D input, such as those available in modeling software such as Blender. Other tools rely on Constructive Solid Geometry, in which a shape is built from a sequence of boolean operators (union, difference, intersection) and spatial transforms on 3D primitives (spheres, boxes, cylinders). Much research has been done on modeling from sketching, as described Olsen, Samavati, Sousa, and Jorge (2009). Hybrid techniques also exist, such as the Mockup Builder by De Araùjo, Casiez, and Jorge (2012), which combines 2D sketching on a multitouch surface and 3D extrusion above the surface.

In musical context Mulder (1998) investigates the use of sculpting metaphors for multidimensional control of musical parameters with a 3D glove metaphor. Oliver and Pickles (2007) describe a 3D instrument in which complex 3D shapes are associated with various sound synthesis techniques, controlled by deforming the shapes using a gamepad. The Drile instrument by Berthaut et al. (2010) allows musicians to build 3D tree structures associated with a hierarchical live-looping technique in which each node represents a loop of events sent to its children.

2.3.3. Material manipulation

The third category of object manipulation techniques pertains to the control of the material of objects, i.e. their colour and texture. Research on the subject has led to tools such as direct painting on surfaces described by Agrawala, Beers, and Levoy (1995) and the placement and alignment of textures on 3D meshes Gingold, Davidson, Han, and Zorin (2006).

An example use of material manipulation for musical interaction can be found in the Tunnels by Berthaut (2012). These widgets, 5 of which are shown arranged on a 3D path on Figure 5 left, act as graphical sliders which modify the graphical parameters of objects passed through them. For instance, a color Tunnel with a continuous color scale allows for a continuous change in pitch of a sound object. Multiple graphical parameters can be modified at once by a single Tunnel and scales can be changed dynamically through a preset system.

2.4. 3D Navigation

The third category of interaction technique pertains to moving through a VE. Various classifications have been proposed for navigation techniques, including organising them according to their goal or the type of performed movement. For example, Tan, Robertson, and Czerwinski (2001) defines three goals, namely Exploration, Search and Inspection. Mackinlay, Card, and Robertson (1990) describe three types of movement, namely General movement, Targeted movement and Specified coordinates movement. In the following sections, we rely on that second classification.

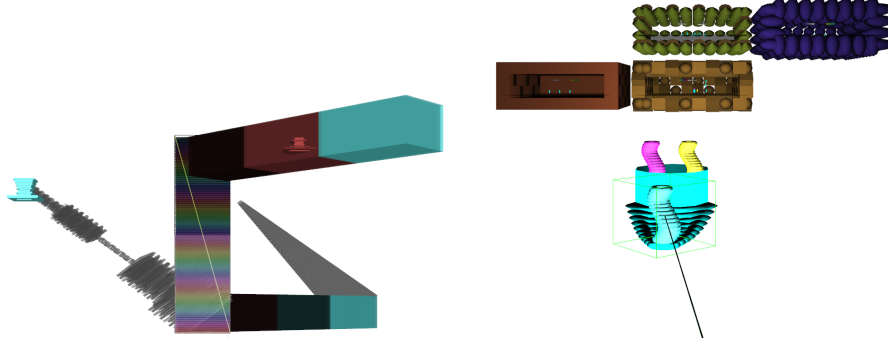


Figure 5. Left : Graphical manipulation tools used in Drile. Tunnels are 3D sliders which modify the appearance of the visual representations of sound processes that they intersect. Right : Targeted movement navigation between musical scenes (sets of sounds and effects) in Drile.

3D navigation open numerous opportunities for musical expression. Musical parameters can be mapped to the visibility, distance, position and orientation of virtual objects relative to the user’s position, or to the spatial transformation of the user relative to the VE. Each movement type then affords a particular set of musical controls that we describe in the following section. More advanced techniques such as multiscale navigation by Kopper, Ni, Bowman, and Pinho (2006) offer additional possibilities such as different levels of control on the same musical process. This might allow one to switch between the process, note and timbre levels of the dimension space for digital musical instruments proposed by Birnbaum, Fiebrink, Malloch, and Wanderley (2005) by navigating at different levels of zooms.

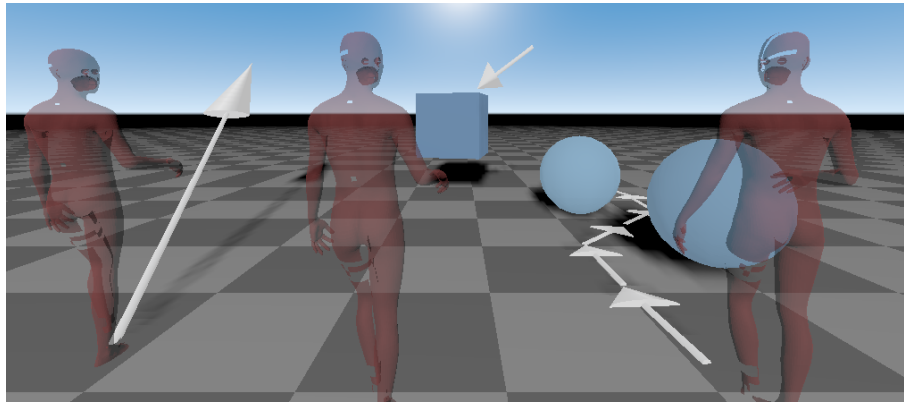


Figure 6. Illustration of the main categories of navigation techniques: General movement (left), Targeted movement (at one corner of the cube, center), Specified trajectory (right)

2.4.1. General movement

The first category of 3D navigation technique pertains to free and continuous movements through the VE. These techniques include the panning/rotation/zooming functions common in 3D modeling software, First Person Shooter-like walking and flying. Much research has been conducted on general movement techniques, comparing their efficiency in different environments. D. Bowman, Koller, and Hodges (1997) reviews travel techniques in viewpoint motion control. Usuh et al. (1999) compares real walk-

ing, walking in place and flying in VEs. Cirio, Marchal, Regia-Corte, and Lécuyer (2009) proposes a redirected walking method that allows for travelling in an infinite VE while physically walking in a restricted physical space.

Most 3DMIs which rely on navigation use general movement techniques. Barri (2009) describes a musical instrument in which the musician flies among various 3D shapes representing spatialized sound sources. Hamilton (2008) and Hamilton (2011) use First Person Shooter movement as basis for the development of 3D musical instruments. Frameworks 3D presented in Polfreman (2009) involves moving along a 3D timeline where patterns are placed and connected with temporal relations. Plumage described in Jacquemin, Ajaj, Cahen, Ollivier, and Schwarz (2007) is a 3D interface for flying in sets of granularized sound sample. In Wozniowski, Settel, and Cooperstock (2006), the authors propose to use movements in a VE and head-movements to select and activate sound sources.

2.4.2. Targeted movement

The second category of navigation techniques pertains to moving to a specific target inside the VE. This movement can be immediate, e.g. teleportation, or follow a path to the target either automatically or at a user defined speed, as presented in Moerman, Marchal, and Grisoni (2012). In addition, some techniques allow users to preview and refine their final position relative to the target, such as Navidget described in Hachet, Decle, Knodel, and Guitton (2008).

Not many 3DMIs have made use of target movements, although it provides an opportunity for instantaneous controls, such as modification of multiple parameters at once when moving through a parameter space or discrete activation of sound sources. In Berthaut et al. (2010), the authors use targeted movement to move between scenes which contain different effects and sound sources.

2.4.3. Specified trajectory movement

The third category of navigation techniques pertains to predefined navigation routes along which the user can move. These provide interesting possibilities for musical applications, as they can be used as a metaphor for moving along a timeline. In Rodet et al. (2005) the user follows a virtual groove with a force feedback device, their position controlling the playback of written music. A series of 3D rhythm games also make use of these techniques, from Rez (United Game Artists (2001)) to Thumper (Drool LLC (2016)), in which the player triggers musical events along a 3D path.

	Selection			Manipulation			Navigation				
	3D Cursor	Virtual ray	Occlusion	Spatial transform	Shape - structure	Material	General	Targeted	Specified trajectory	Sound repres.	Input Device
Sound of one Hand	X			X				X		Reactive	Glove
WAVE	X									Sensors	Tracker Mouse
Drile		X		X	X	X		X		Reactive Structure	Piivert
Virtual Xylophone	X	X								Sensors	Wand Glove
Versum							X			Reactive Structure	Mouse Keyboard
Phase									X	Reactive Structure	Phantom
Fijuu				X	X					Reactive	Gamepad
q3osc			X				X			Reactive	Mouse Keyboard
Plumage							X			Reactive	Mouse Keyboard
Virtual grains	X									Reactive	Wand
Virtual sculpting				X	X					Reactive	Glove
Shoggoth			X				X			Structure	Mouse Keyboard
Hybrid Reality	X			X	X	X				Reactive	Wand
Frameworks 3D			X						X	Sensors Structure	Mouse Keyboard

Table 1. List of existing instruments and the corresponding 3D interaction techniques

3. Adapting 3D interaction techniques for musical interaction

In the previous section, we highlighted the potential of 3D interaction techniques for musical expression. However, there are also limitations in existing techniques with respect to common guidelines for the design of digital musical instruments, such as increasing musician’s input complexity and freedom (Jordà (2003)), providing sufficient feedback (Marshall and Wanderley (2006)), taking the audience experience into account Barbosa, Calegario, Teichrieb, Ramalho, and McGlynn (2012), or allowing musicians to rely on existing musical practice or expert gestures.

In this section, we propose a number of directions for extending 3D interaction techniques in the context of musical interaction. We believe they constitute a strong basis for future research in 3DUIs and NIMEs.

3.1. Multiple objects selection/manipulation

Most research on 3D User Interfaces focus on the selection and manipulation of one object at a time and the issue of disambiguating these operations in the case of dense environments or moving targets. However, in most musical instruments a number of playing techniques rely on independent gestures at the hand or even finger level, performed in parallel on multiple components of the instruments.

Little research has been done on multiple object selection, with the notable exception of Lucas (2005) and Stenholt (2012). In addition to serial multiple selection which consist in a sequence of single object selection, parallel techniques were proposed. They include moving a resizable selection volume (e.g. in the PORT technique), drawing a selection lasso then projected or extruded into the VE and selecting by similarity / proximity. However, they all require a preselection phase, such as iterating modifications of the selection volume or drawing the lasso or selecting a seed. They also only provide undifferentiated selection, i.e. all objects are assigned to the same hand / finger.

For musical purposes, necessary extensions of these techniques are:

- the ability to dynamically and continuously change the selection without going through a preselection phase each time. This will allow for changing sets of selected instrument components such as in chords
- the ordering and differentiation of selected elements, so that they can be separately assigned to e.g. fingers for subsequent activation and manipulation

3.2. Hybrid/Articulated techniques

With the aim of improving the efficiency and accuracy of interaction in VE, research on 3DUIs tend to combine existing techniques to take advantage of their respective advantages. Combining techniques can be done by articulating them temporally or by hybridising them. For example, D. A. Bowman and Hodges (1997) propose the HOMER technique which combines the efficiency of the virtual ray for selection with the accuracy of the virtual cursor for manipulation. Tan et al. (2001) combine a general movement (flying) technique with targeted movement (orbiting). Argelaguet and Andujar (2013) proposes a hybrid selection technique that disambiguates virtual selection by combining it with occlusion selection.

In a musical context, hybridisation and articulation will increase the musician’s input freedom by allowing multiple ways of interacting with the same virtual element.

For instance, volume control of spatialized sound sources placed in the VE could be both discrete, by jumping to a position using a go-to technique and continuous with a free flying technique, allowing for two controls of the same musical parameters. In the case of articulated techniques, it will also be necessary to investigate how to switch in a fluid manner between them.

3.3. Musical disambiguation / guiding

3DUIs often require mechanisms for guiding and disambiguating the interaction within complex VEs. For example, because of the lack of precision of the virtual ray technique at a large distance and with occluded objects, techniques were developed that attempt to ensure only the desired object is selected. An example is *IntenSelect* by de Haan et al. (2005) which bends the virtual ray according to the user's behaviour. These disambiguation mechanisms can also depend on heuristics, i.e. rules which define how to adapt the interaction depending on the objects characteristics. For example, for the spatial transformation of objects Gosele, Stuerzlinger, et al. (1999) proposes to rely on affordances of objects such as tables and floors to automatically place the manipulated objects on their surface.

Finally, in specified trajectory movement, users are left with only a few degrees of freedom, the rest being specified by the designer, leading to a guiding movement through the VE.

An interesting extension of these techniques is to implement heuristics and behavioral disambiguation or guides which depend on musical constraints. For example, when selecting objects associated to different pitches, selection can be guided relative to a pre-defined scale instead of by spatial constraints. Temporal constraints can also be added for example when manipulating objects associated with musical parameters, so as to quantize the changes and synchronize them to a particular tempo. 3D navigation can also be guided according to general musical principles. For example, in order to prevent large changes in volume or other musical parameters, the speed of movement can be limited in parts of the scene or depending on the proximity with predefined sound sources.

3.4. Non optimal techniques

Research on 3DUIs tend to focus on optimizing the techniques to ensure the spatial and temporal performances of selection, manipulation and navigation.

For example, based on Fitt's law (Fitts (1954)) and the Optimized Initial Impulse Model (Meyer, Abrams, Kornblum, Wright, and Keith Smith (1988)), selection techniques have been proposed which optimize selection by modifying the distance to the target (through a change in control-display ratio), increasing the target size, or both. For example, the *Sticky Ray* by Steinicke, Ropinski, and Hinrichs (2006) bends the virtual ray so that it always selects the closest object, therefore potentially removing the need for the slow corrective movements which follow fast ballistic movements during the selection. Another issue in selection is the *Heisenberg effect* described by D. Bowman, Wingrave, Campbell, and Ly (2001), which is the fact that the position of the virtual cursor or ray changes when the selection is confirmed, e.g. the user's hand moves slightly when they press the confirmation button.

While accuracy remains essential in the context of musical interaction, one of the main differences with other application fields is the room that musical instruments

must leave for errors, approximations and more continuous and ambiguous results. This often leads to more input freedom for the musician and creates opportunities for the appropriation of the techniques and the development of musical new gestures.

It would therefore be beneficial for musical interaction in VEs to investigate techniques that integrate this ambiguity, or even *re-ambiguate* existing techniques. For example, selection techniques could allow for partial selection of objects, providing for each of them the selection ratio. Navigation techniques could be designed which would prevent accurate movement through the VE, for example by adding simulated natural events such as gusts of wind or earthquakes. Finally, the Heisenberg effect could be amplified, leading to artefacts in the selection process which would in turn be sonified.

3.5. *Dedicated input devices*

Research in 3DUIs has led to the design of a large number of input devices. These are usually very dependent on the proposed interaction techniques and application field.

Musical interaction in VEs can be enriched by designing specific input devices which complement the graphical interaction techniques.

An important aspect of these devices is the presence of haptic feedback. This presence is one of the main guidelines proposed by Serafin et al. (2016) for the design of IVMI. It is in fact well known that musical interaction benefits from haptic feedback O'Modhrain (2001) whether it is force feedback, vibrotactile feedback or even passive feedback with the musician interacting with a tangible prop. Although common devices for 3DUIs usually integrate some amount of haptic feedback, there is still research to be done on music specific feedback in that context.

Another interesting aspect is the choice of sensors on input devices used for 3D interaction. Usual and commercial input devices rely on a combination of buttons, joysticks and more recently 2D touchpads. However, these provide only limited expressiveness compared to what has been proposed in the field of NIME. We therefore believe that other sensor types used in multimodal interaction techniques such as the microphone (which could be leveraged to integrate expressive control rather than voice commands) or eye/gaze tracking should be utilised when designing for 3D musical interaction.

For instance, audio input as proposed by Momeni (2015) , with an input device equipped with piezos capturing subtle tapping and scratching would enable the combination of intimate control and complex manipulation techniques on a large scale.

Similarly, per finger sensing would allow for the design of expressive parallel selection and manipulation techniques with chords of objects. An example of that is Piivert, designed by Berthaut et al. (2011) and depicted in Figure 7, which integrates a pressure sensor for each finger and enables the use of expert percussion gestures such as taps, flams and rolls to complement graphical interaction techniques.

3.6. *Visual feedback for the audience and orchestra*

The last direction for adapting 3D interaction techniques to Although collaborators have soon been taken into account in the context of Collaborative Virtual Environments for example in Benford, Bowers, Fahlén, and Greenhalgh (1994), the information required for musical interaction differs from other domains. In particular, as pointed out by Blaine and Fels (2003) it is essential that musicians are aware of the contribution of each other to the overall produced music. They should also be able to give each other indications relative to upcoming musical events or actions, e.g. changing tempo



Figure 7. The Piivert device feature pressure sensors for each finger, which enable percussion gestures and expressive excitation gestures to be performed in combination with graphical 3D interaction techniques.

or starting/ending a part.

An even stronger difference lies in the role of the spectator, an important stakeholder of musical interaction as pointed out by Barbosa et al. (2012), rarely taken into account in research on VEs. 3D interaction techniques should therefore be developed with transparency (Fels, Gadd, and Mulder (2002)) and gesture continuity (Berthaut et al. (2014)) in mind, that is spectators should be able to perceive and understand how the musician’s gestures are linked to the interaction techniques and how these interaction techniques impact the music. An example solution would be to provide an alternate view of the interaction technique for the audience which would highlight the activity and engagement while removing the visual complexity required by the musician.

4. Conclusion

In this paper, we first reviewed 3D interactions techniques and what opportunities they open for musical interaction. We then propose a set of directions for extending 3DUIs with musical interaction in mind. In particular, we advocate for the investigation of multiple objects selection/manipulation techniques, hybrid/articulated techniques, musical-centric guiding and disambiguation, re-ambiguation of existing techniques, dedicated input devices and the design of visual feedback targeted at the audience and orchestra members.

We believe that the exploration of these directions could be facilitated by integrating existing 3D interaction techniques in frameworks for the development of 3DMIS, such as UDKOSC (Hamilton (2011)) , OSC-XR (Johnson, Damian, and Tzanetakis (2019)), and Chunity (Atherton and Wang (2018)).

Finally, we hope that this review and proposed guidelines will encourage instrument designers and researchers from the New Interfaces for Musical Expression and 3D User Interfaces fields to continue exploring 3D Musical Instruments.

References

- Agrawala, M., Beers, A. C., & Levoy, M. (1995). 3d painting on scanned surfaces. In *Proceedings of the 1995 symposium on interactive 3d graphics* (pp. 145–ff).
- Argelaguet, F., & Andujar, C. (2013). A survey of 3d object selection techniques for virtual environments. *Computers & Graphics*, 37(3), 121–136.

- Atherton, J., & Wang, G. (2018). Chunity: Integrated audiovisual programming in unity. In *Proceedings of the international conference on new interfaces for musical expression* (pp. 102–107).
- Barbosa, J., Calegario, F., Teichrieb, V., Ramalho, G., & McGlynn, P. (2012). Considering audience’s view towards an evaluation methodology for digital musical instruments. In *Proceedings of the international conference on new interfaces for musical expression*. Ann Arbor, Michigan: University of Michigan. Retrieved from <http://www.nime.org/proceedings/2012/nime2012.174.pdf>
- Barri, T. (2009). Versum: audiovisual composing in 3d..
- Beat Games. (n.d.). *Beat saber*. Retrieved 2019-09-10, from <https://beatsaber.com/>
- Benford, S., Bowers, J., Fahlén, L. E., & Greenhalgh, C. (1994). Managing mutual awareness in collaborative virtual environments. In *Proceedings of vrst* (pp. 223–236).
- Berthaut, F. (2012, March). Beyond the Tunnels: Advanced 3D graphical modulation. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)* (p. 139-140). United States. Retrieved from <https://hal.archives-ouvertes.fr/hal-00793580>
- Berthaut, F., Desainte-Catherine, M., & Hachet, M. (2010, June). Drile: An Immersive Environment for hierarchical live-looping. In *New Interface for Musical Expression* (p. page 192). Sydney, Australia. Retrieved from <https://hal.archives-ouvertes.fr/hal-00530071>
- Berthaut, F., Desainte-Catherine, M., & Hachet, M. (2011, October). Interacting with 3D Reactive Widgets for Musical Performance. *Journal of New Music Research*, 40(3), 253-263. Retrieved from <https://hal.archives-ouvertes.fr/hal-00633750>
- Berthaut, F., Zappi, V., & Mazzanti, D. (2014, March). Scenography of immersive virtual musical instruments. In *VR Workshop: Sonic Interaction in Virtual Environments (SIVE), 2014 IEEE*. Minneapolis, United States. Retrieved from <https://hal.archives-ouvertes.fr/hal-01222454>
- Birnbaum, D., Fiebrink, R., Malloch, J., & Wanderley, M. M. (2005). Towards a dimension space for musical devices. In *Proceedings of the 2005 conference on new interfaces for musical expression* (pp. 192–195).
- Blaine, T., & Fels, S. (2003). Contexts of collaborative musical experiences. In *Proceedings of nime 03* (pp. 129–134). Singapore, Singapore.
- Bowman, D., Koller, D., & Hodges, L. F. (1997). Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Ieee proceedings of vrais'97*.
- Bowman, D., Wingrave, C., Campbell, J., & Ly, V. (2001). Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments.
- Bowman, D. A., & Hodges, L. F. (1997). An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Si3d '97: Proceedings of the 1997 symposium on interactive 3d graphics* (pp. 35–ff.). New York, NY, USA: ACM.
- Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2004). *3d user interfaces: Theory and practice*. Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc.
- Cadoz, C. (1999). Musique, geste, technologie. In *Les nouveaux gestes de la musique*. Éditions Parenthèses.
- Cirio, G., Marchal, M., Regia-Corte, T., & Lécuyer, A. (2009). The magic barrier tape: a novel metaphor for infinite navigation in virtual worlds with a restricted walking workspace. In *Vrst '09: Proceedings of the 16th acm symposium on virtual reality software and technology* (pp. 155–162). New York, NY, USA: ACM.
- Cohé, A., Dècle, F., & Hachet, M. (2011). tbox: a 3d transformation widget designed for touch-screens. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 3005–3008).
- De Araùjo, B. R., Casiez, G., & Jorge, J. A. (2012). Mockup builder: direct 3d modeling on and above the surface in a continuous interaction space. In *Proceedings of graphics interface 2012* (pp. 173–180).
- de Haan, G., Koutek, M., & Post, F. H. (2005, October). Intenselect: Using dynamic object rating for assisting 3d object selection. In *Ipt / egve 2005* (pp. 201–209).
- Drool LLC. (2016). *Thumper*. Retrieved 2019-06-03, from <https://thumpergame.com/>

- Fels, S., Gadd, A., & Mulder, A. (2002). Mapping transparency through metaphor: towards more expressive musical instruments. *Organised Sound*, 7(2), 109–126.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6), 381.
- Gingold, Y. I., Davidson, P. L., Han, J. Y., & Zorin, D. (2006). A direct texture placement and editing interface. In *Proceedings of the 19th annual acm symposium on user interface software and technology* (pp. 23–32).
- Gosele, M., Stuerzlinger, W., et al. (1999). Semantic constraints for scene manipulation. In *in proceedings spring conference in computer graphics' 99 (budmerice, slovak republic)*.
- Hachet, M., Bossavit, B., Cohé, A., & de la Rivière, J.-B. (2011). Toucheo: multitouch and stereo combined in a seamless workspace. In *Proceedings of the 24th annual acm symposium on user interface software and technology* (pp. 587–592).
- Hachet, M., Decle, F., Knodel, S., & Guitton, P. (2008). Navidget for easy 3d camera positioning from 2d inputs. In *3dvi '08: Proceedings of the 2008 ieee symposium on 3d user interfaces* (pp. 83–89). Washington, DC, USA: IEEE Computer Society.
- Hamilton, R. (2008). q3osc: or how i learned to stop worrying and love the game. In *Proceedings of the international computer music association conference*.
- Hamilton, R. (2011). Udkosc: An immersive musical environment. In *Icmc*.
- Hard Ligh Labs LLC. (n.d.). *Soundstage*. Retrieved 2019-09-10, from <https://www.viveport.com/apps/2a0ab90e-d6f3-4e48-8347-2cfaaf52d5e4?gl=us>
- Hunt, A., & Kirk, R. (2000). Mapping strategies for musical performance. *Trends in Gestural Control of Music*, 231–258.
- Jacquemin, C., Ajaj, R., Cahen, R., Ollivier, Y., & Schwarz, D. (2007). Plumage: Design d'une interface 3d pour le parcours d'échantillons sonores granularisés. In *Proceedings of the conférence francophone sur l'interaction homme-machine(ihm'07)*.
- Jankowski, J., & Hachet, M. (2013). A survey of interaction techniques for interactive 3d environments. In *Eurographics 2013-star*.
- Johnson, D., Damian, D., & Tzanetakis, G. (2019). Osc-xr: A toolkit for extended reality immersive music interfaces.
- Jordà, S. (2003). Interactive music systems for everyone: exploring visual feedback as a way for creating more intuitive, efficient and learnable instruments. In *Proceedings of the stockholm music acoustics conference (smac03)*.
- Kopper, R., Ni, T., Bowman, D. A., & Pinho, M. (2006). Design and evaluation of navigation techniques for multiscale virtual environments. In *Ieee virtual reality conference (vr 2006)* (pp. 175–182).
- Lanier, J. (1993). The sound of one hand. *Whole Earth Review*.
- Lucas, J. F. (2005). *Design and evaluation of 3d multiple object selection techniques* (Unpublished doctoral dissertation). Virginia Tech.
- Mackinlay, J. D., Card, S. K., & Robertson, G. G. (1990). Rapid controlled movement through a virtual 3d workspace. In *Acm siggraph computer graphics* (Vol. 24, pp. 171–176).
- Mäki-Patola, T., Laitinen, J., Kanerva, A., & Takala, T. (2005). Experiments with virtual reality instruments. In *Proceedings of the 2005 international conference on new interfaces for musical expression (nime05), vancouver, bc, canada*.
- Marshall, M. T., & Wanderley, M. M. (2006). Vibrotactile feedback in digital musical instruments. In *Nime '06: Proceedings of the 2006 conference on new interfaces for musical expression* (pp. 226–229). Paris, France, France: IRCAM — Centre Pompidou.
- Mendes, D., Caputo, F. M., Giachetti, A., Ferreira, A., & Jorge, J. (2019). A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In *Computer graphics forum* (Vol. 38, pp. 21–45).
- Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Keith Smith, J. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological review*, 95(3), 340.
- Moerman, C., Marchal, D., & Grisoni, L. (2012). Drag'n go: Simple and fast navigation in virtual environment. In *2012 ieee symposium on 3d user interfaces (3dvi)* (pp. 15–18).

- Momeni, A. (2015). Caress: An enactive electro-acoustic percussive instrument for caressing sound. In *Proceedings of the international conference on new interfaces for musical expression (nime-15)*, baton rouge, la, usa (Vol. 31).
- Mulder, A. G. (1998). *Design of virtual three-dimensional instruments for sound control* (Unpublished doctoral dissertation). Simon Fraser University, Canada.
- Naef, M., & Collicot, D. (2006). A vr interface for collaborative 3d audio performance. In *Proceedings of the 2006 international conference on new interfaces for musical expression (nime06)*, paris, france.
- Oliver, J., & Pickles, S. (2007). Fijuu2: a game-based audio-visual performance and composition engine. In *Nime '07: Proceedings of the 7th international conference on new interfaces for musical expression* (pp. 430–430). New York, NY, USA: ACM.
- Olsen, L., Samavati, F. F., Sousa, M. C., & Jorge, J. A. (2009). Sketch-based modeling: A survey. *Computers & Graphics*, 33(1), 85–103.
- O'Modhain, M. S. (2001). *Playing by feel: incorporating haptic feedback into computer-based musical instruments* (Unpublished doctoral dissertation). Stanford, CA, USA. (Adviser-Chafe, Chris)
- Pierce, J. S., Forsberg, A. S., Conway, M. J., Hong, S., Zeleznik, R. C., & Mine, M. R. (1997). Image plane interaction techniques in 3d immersive environments. In *Proceedings of the 1997 symposium on interactive 3d graphics* (pp. 39–ff).
- Polfreman, R. (2009). Frameworks 3d: composition in the third dimension.
- Poupyrev, I., Billinghamurst, M., Weghorst, S., & Ichikawa, T. (1996). The go-go interaction technique: non-linear mapping for direct manipulation in vr. In *Acm symposium on user interface software and technology* (pp. 79–80).
- Poupyrev, I., Weghorst, S., Billinghamurst, M., & Ichikawa, T. (1998). Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques. In *Computer graphics forum*.
- Pressing, J. (1997). Some perspectives on performed sound and music in virtual environments. *Presence*, 6(4), 482–503.
- Rodet, X., Gosselin, F., Mobuchon, P., Lambert, J.-P., Cahen, R., Gaudy, T., & Guedy, F. (2005). Study of haptic and visual interaction for sound and music control in the phase project. In *Proceedings of the 2005 international conference on new interfaces for musical expression (nime05)*, vancouver, bc, canada.
- Serafin, S., Erkut, C., Kojs, J., Nilsson, N., & Nordahl, R. (2016). Virtual reality musical instruments: State of the art, design principles, and future directions. *Computer Music Journal*, 40(3), 22–40.
- Steed, A. (2006). Towards a general model for selection in virtual environments. In *3d user interfaces (3dui'06)* (pp. 103–110).
- Steinicke, F., Ropinski, T., & Hinrichs, K. (2006). Object selection in virtual environments using an improved virtual pointer metaphor. In *Computer vision and graphics* (pp. 320–326). Springer.
- Stenholt, R. (2012). Efficient selection of multiple objects on a large scale. In *Proceedings of the 18th acm symposium on virtual reality software and technology* (pp. 105–112).
- Survios. (n.d.). *Electronauts*. Retrieved 2019-09-10, from <http://www.electronautsvr.com>
- Tan, D. S., Robertson, G. G., & Czerwinski, M. (2001). Exploring 3d navigation: combining speed-coupled flying with orbiting. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 418–425).
- The Music Room. (n.d.). *Chroma code*. Retrieved 2019-09-10, from <http://www.musicroomvr.com>
- Thérémix Tech. (n.d.). *Modulia studio*. Retrieved 2019-09-10, from <https://www.modulia-studio.com>
- United Game Artists. (2001). *Rez*. Retrieved 2019-06-03, from <https://en.wikipedia.org/wiki/Rez>
- Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., & Brooks, F. P., Jr. (1999). Walking , walking-in-place , flying, in virtual environments. In *Siggraph '99*:

- Proceedings of the 26th annual conference on computer graphics and interactive techniques* (pp. 359–364). New York, NY, USA: ACM Press/Addison-Wesley Publishing Co.
- Valbom, L., & Marcos, A. (2005). Wave: Sound and music in an immersive environment. *Computer and Graphics*.
- WaveVR. (n.d.). *Wavexr*. Retrieved 2019-09-10, from <http://wavexr.com>
- WeARVR. (n.d.). *Aliveinvr*. Retrieved 2019-09-10, from <https://www.wearvr.com/apps/aliveinvr>
- Wessel, D., & Wright, M. (2002). Problems and prospects for intimate musical control of computers. *Computer music journal*, 26(3), 11–22.
- Wozniowski, M., Settel, Z., & Cooperstock, J. (2006). A spatial interface for audio and music production. In *Proceedings of the international conference on digital audio effects (dafx)*, 2006.
- Zappi, V., Mazzanti, D., Brogni, A., & Caldwell, D. (2012). Concatenative synthesis unit navigation and dynamic rearrangement in vrgrains.
- Zappi, V., Mazzanti, D., Brogni, A., & Caldwell, D. G. (2011). Design and evaluation of a hybrid reality performance. In *Nime* (Vol. 11, pp. 355–360).