Chapter 6
Spatial Design Considerations for Interactive Audio in Virtual Reality

Thomas Deacon and Mathieu Barthet

Abstract Space is a fundamental feature of virtual reality (VR) systems, and more generally, human experience. Space is a place where we can produce and transform ideas and act to create meaning. It is also an information container. When working with sound and space interactions, making VR systems becomes a fundamentally interdisciplinary endeavour. To support the design of future systems, designers need an understanding of spatial design decisions that impact audio practitioners' processes and communication. This chapter proposes a typology of VR interactive audio systems, focusing on their function and the role of space in their design. Spatial categories are proposed to be able to analyse the role of space within existing interactive audio VR products. Based on the spatial design considerations explored in this chapter, a series of implications for design are offered that future research can exploit.

6.1 Introduction

Technologies like virtual reality (VR) offer many ways of using space that could benefit creative audio production and immersive experience applications. Using VRs affordances for embodied interaction and spatial user interfaces, new forms of spatial expression can be explored. Running parallel to VR research efforts in sonic interaction in virtual environments(SIVE), much of sonic practice exists as applied design, either as music making tools [110], experiential products [106], or games [102]. Commercial work is influenced by academia, but it is also based on broader professional constituencies and practices not related to sound and music interaction design.

T. Deacon (✉)
Media and Arts Technology CDT, Queen Mary University of London, London, United Kingdom

M. Barthet
Centre for Digital Music, Queen Mary University of London, London, United Kingdom
e-mail: m.barthet@qmul.ac.uk

Much of VR design practice is communicated as professional dialogues, such as platform or technology best practice guides [120, 121], or reviews of "lessons-learned" in industrial settings [105, 122]. Within these professional dialogues, previous research, new technological capabilities, and commercial user research are collected together to inform communities on how to best support users and task domains. For the field of SIVE, and sound and music Computing(SMC) more broadly, there is still work to be done to bridge commercial practice and academic endeavours. Despite recent works [6, 77], there is a paucity of design recommendations and analysis regarding how to build spaces, interfaces, and spatial interactions with sound. For the potential of VR to be unlocked as a creative medium, multi and interdisciplinary work must be undertaken to bring together the disciplines that touch on space, interaction, and sound.

Studying how people make immersive tools, in commercial and academic settings, requires a means of framing how spatial design decisions impact users. This brings up two problems, what role do commercial artefacts have in broadening research understanding, and how is relevant knowledge generated from such products? Objects, prototypes, and artefacts create a context for forming new understanding [46]. By analysing an artefact design, research can discover (recover and invent) requirements to create technological propositions related to domain-specific concerns [82]. This is because an artefact collects designers judgements about specific design spaces [33], for instance how to solve interaction problems, and what aspects are of priority to users at different points in an activity. However, this means we cannot recover the needs of design by direct questioning the users alone. A broader research picture is needed, one that integrates action with tools, users, and reflection on devices. So, to develop an understanding for future design interventions, research should gather diverse data to understand the existing practice and perceived professional constituencies.[1]

Section 6.2 sets out the problem of space in more detail, highlighting important contributions to the design of VR sound and music interaction systems. Section 6.2 also describes the suitability of typologies to spatial analysis for this research. Following on from this, Sect. 6.3.1 outlines the approach taken to the design review and typology, indicating how relevant work was identified, selected, and coded. Section 6.3 sets out a typology of interactive audio systems in VR, and presents case studies of spatial design in the field. Section 6.5 looks across analyses and offers ways to understand the design space of VR for SMC. Based on findings and reflections, Sect. 6.6 proposes actionable design outcomes for further research, then Sect. 6.7 draws the work to a close.

---

[1] Prototypes are any representation of a design idea, regardless of medium, and an artefact is a product or interactive system created for a design intervention/experiment [46].

6.2 Background

6.2.1 Terminology

This chapter analyses the spatial design of interactive audio systems (IAS) in VR. IAS refers to any sound and music computing system that involves human interaction that can modify the state of the sound or music system, however, we do not review information-only auditory displays or audio-rendering technologies. While both auditory displays and rendering technologies do include interactivity in their operation, this chapter is interested in the use of interactive sound as the primary function in the VR application, rather than when sound is used as an information medium or renderer of spatial sounds without interactive feedback beyond head rotation. No doubt there are significant overlaps in theory and application, that would be valuable to explore, but trying to address all aspects in one chapter requires a different focus.

The following research areas pertain to spatial interaction with user interfaces (UI)s:

- Spatial user interface (SUI): Human-computer interaction (HCI) with 3D or 2D UI that is operated through spatial interaction, graphically or otherwise [59].
- Three-dimensional user interface (3DUI): A UI that involves 3D interaction [16].
- Distributed user interface (DUI): UIs that are distributed across devices, users, or spatial access points [89].

There are also many terms to describe virtual spaces used for sound and music; in particular, this research is concerned with immersive VR technology, following the definition provided in [6]:

- Virtual—to be a virtual reality, the reality must be simulated (e.g. computer-generated).
- Immersive—to be a virtual reality, the reality must give its users the sensation of being surrounded by a world.
- Interactive—to be a virtual reality, the reality must allow its users to affect the reality in some meaningful way.

The term VR can refer to the hardware systems for delivering immersive experiences and to refer to the immersive experiences themselves. Hardware systems can include commercial head-mounted display (HMD) technology, such as Oculus or HTC Vive, through to complex stereographic projection-based Cave Automatic Virtual Environment (CAVEs) [12]. The key thing is that in these immersive environments the visual system and interaction capacities are mediated through technological means. In the case of social virtual reality (SVR), described in Chap. 8 of this volume, communication layers (speech, posture, and gesture) may or may not be mediated through technological means, for instance co-located users may share a virtual world via

HMD but speech communication is unmediated. Or remote SVR users' communication could be completely mediated by avatar representations and voice over internet protocol (VoIP) technology.

6.2.2 Standing on the Shoulders of Giants, but Which Ones?!

SMC and SIVE are linked to the larger research field of HCI, so it is common practice to adopt HCI research findings on how best to design systems. Below, Sect. 6.2.2.1 describes two examples of how interaction methods are used in the design of VR for IAS. But as research in VR for SMC has developed, researchers have needed to define and collect design principles specific to sound and music in VR, this work is reviewed in Sect. 6.2.2.2.

6.2.2.1 Adapting Existing VR HCI Frameworks to Audio System Design

To establish a dialogue around spatial considerations, there is a need to adopt findings from other VR HCI disciplines. But as with the adoption of HCI evaluation frameworks within new interfaces for musical expression (NIME) [78, 91, 98], critical understanding of the target domain (SMC) needs to be established [70, 81]. For instance, making expressive systems for musical creation or sonic experiences has different design requirements than usability engineering [42], or demonstrations of interaction techniques [8]. This is not to say that usability engineering is not important, but rather the goal of design and evaluation needs to expand to include sonic aesthetic qualities for audio-first spatial scenarios.

Selection and Manipulation Techniques

Object selection and manipulation is fundamental to VR environments where users perform spatial tasks [52]. At a basic level, there are two main categories that describe 3D interaction for VR: Direct and indirect interaction techniques [5]. Object manip-
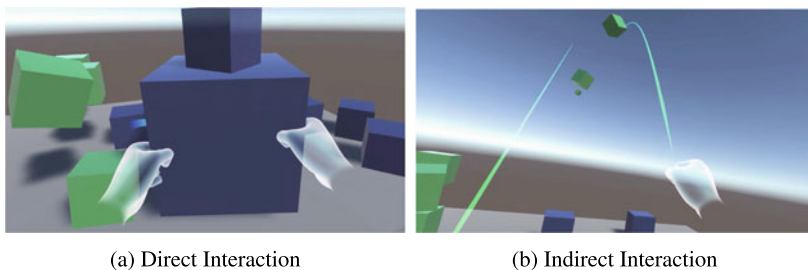

(a) Direct Interaction   (b) Indirect Interaction

Fig. 6.1 Selection and manipulation mechanics in VR

ulation examples of direct and indirect techniques can be seen in Fig. 6.1. Direct interaction refers to having 'virtual hands'; similar to touching and grabbing objects in the real world. A benefit of direct interaction is that control maps virtual tasks identically with real tasks, resulting in more natural interaction [5]. Indirect interaction refers to virtual pointing; like using a laser pointer (ray-casting) that can pickup and drop objects in space. Indirect interaction lets users select objects beyond their area of reach and require relatively less physical movement. Overcoming the physical constraints of the real world provides substantial benefits for the design of virtual spaces, as the arrangement of elements can expand beyond body-scaled interaction. Across both direct and indirect mechanics, interaction should be rapid, accurate, error proof, easy to understand and control, and aim for low levels of fatigue [5]. Depending on how they are designed, both direct and indirect interactions enable spatial transformations of objects, including rotation, scaling, and translation.

In adapting this research to sound and music interfaces, we must ask how techniques impact musical processes and practices. For example, [13] describes the trade-offs designers make when picking different control systems for virtual reality music instrument (VRMIs). Work that has received less attention in SMC includes how to design for some of the unique properties of VR media. The affordances of VR expand into non-real interaction, so there is a fuzzy middle ground between direct and indirect interaction. For instance, the Go-Go technique enlarges a user's limbs to be able to 'touch' distal objects [74]. In broader VR research, techniques like the Go-Go are described under the term homuncular flexibility [93]; the ability to augment proprioceptive perception of action capacity in VR, adapting interaction to include novel bodies that have extra appendages or appendages capable of atypical movements. An example of this type of research into IAS can be found in [27], where magical indirect interaction was implemented to have audio control objects float towards the user based on pinch actions (via Leap Motion sensor attached to the HMD).

User Interface Elements

Reviewing 3DUI for immersive music production interfaces, [11] proposes three categories of representation for sound processes and parameters: Virtual sensors like buttons and sliders, dynamic/reactive widgets, spatial structures; Fig. 6.2 provides examples. These different representation categories provide a set of design templates for audio production SUIs. For instance, fine-grained individual parameter control may be better suited to sensor devices with precise control relationships. Whereas, if spatio-visual feedback is required about an audio process being applied, a dynamic widget is a suitable device to explore. Spatial structures can be used to represent sequencers and relationships between parameters; as Sect. 6.4 indicates later, several VR audio systems use these to represent either modular synthesis units or whole musical sequencers.

6.2.2.2 Audio-Specific Design Frameworks

Design for IASs in VR is a developing field, surfacing the potential for new forms of sound and music experience [20]. But the opportunities and constraints of VR require critical analysis. For instance, embodied interfaces may offer benefits in productivity and creative expression [62], but we still do not know if the same effects are gathered by embodied interfaces in VR. Alongside this gap, there are gaps in design understanding, with only a few design frameworks addressing how to create VR interfaces and interactions for sound and music [6, 11, 77]. Across these works, a deep level of design analysis around the fundamentals of perception, technology, and action is prevalent. But, in terms of design knowledge to aid designers conceptualising space, and the construction of audio interactions and experiences in it, information is limited. Below is a review of the spatial aspects implicated in the design guidelines of existing VR music system research.

Reviewing VRMI case studies, Serafin *et al.* outline nine principles to guide design, focusing on immersive visualisation from performers' viewpoint [77]. Design principles support design focus on levels of abstraction, immersion, and imagination. Their review of works features many examples of hybrid virtual-physical systems and also highlights that VRMI are well suited to multi-process instruments given SUI affordances. Regarding system design their principles offer robust advice for musical performance but there is a lack of detail on how to go about designing different types of spaces and interactions. For instance, within the principles, an emphasis is put on making experiences social, but no guidance is provided on the design or evaluation of social experience in VR. However, aspects of the case studies do draw attention to spatial factors such as menu design can 'cloud' the performance space; in large interfaces, the mixture of control device and interface design means arm movements
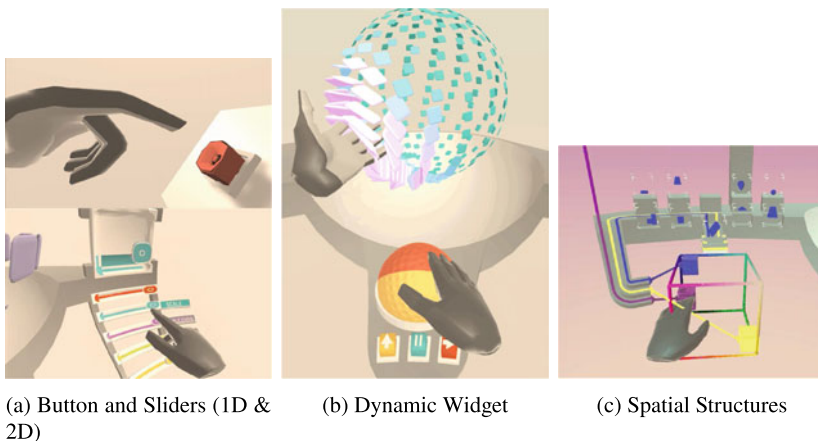

(a) Button and Sliders (1D & 2D)   (b) Dynamic Widget   (c) Spatial Structures

Fig. 6.2 Types of spatial UI for sound processes. Images from Leap motion VR UI design sprint, reproduced with permission from owner, Ultraleap limited

and travel distances can be tiring; and the inclusion of physical control systems supports natural, body-based interaction.

Addressing *Artful Design* for VR sound interaction, Atherton and Wang describe a series of design lenses with subordinate principles using case study analysis [6]. Their work focuses on the idea of creating totally immersive sonic VR. A central concept of their work is the difference between designing for *doing* as distinct from *being* in VR: *"doing is taking action with a purpose; intentionally acting to achieve an intended outcome. In contrast, we define being as the manner in which we inhabit the world around us"* [6]. Expanding on [77]'s suggestion to exploit the 'magical' opportunities of VR, Atherton and Wang highlight that designers should experiment with *virtual physics*, *scale and user perspective*, and *time*, however, these seem to be general principles for VR interaction rather than sound-specific opportunities. Within their discussions spatial concepts emerge, for instance, designers can phase levels interactivity to create different spaces for action in a scene. An actionable design idea relating to this is to guide gaze attention throughout a space related to narrative elements; want people to stop doing and slow down, just put something in the sky above them, as it is not an ideal place to work or interact. Atherton and Wang highlight that designers need to determine different languages of interaction. Design concepts should move beyond functional language towards things that map well to sonic expressions, e.g. instead of physical descriptors like speed of movement and gravity on an object, an interaction language would be *intensity* and *weight and weightlessness*. For Atherton and Wang, play, and particularly social play, is a synthesis of *doing* and *being*, as it is both an activity and a state. Designers can support play by:

1. the lowering users' inhibitions and encouraging them to play;
2. engaging users in diverse movement;
3. allowing users to be silly;
4. making opportunities for discovery in virtual space.

Related to play and interaction, on the social level, designers should provide subspaces within larger worlds and engineer collective interaction scenarios.

6.2.3 Typologies and Spatial Analysis

A typology is a classification of individual units within a set of categories that are useful for a particular purpose. Typologies support the evaluation of a number of different indicators in an integrated manner, based on the identification of relevant links or themes. Within architecture, design typologies are a common method of spatio-visual analysis [24, 72]. The teaching of architectural systems uses an ordered set of types to define areas of interlocking design [22], for instance, in Fig. 6.3 the concept of form is described using a series of types and representative examples.

But typologies can also represent 'spatial qualities' regarding interaction, see Fig. 6.4 where different creative spaces (meeting rooms, maker spaces) can possess

positive and negative attributes for certain activities (socially inviting or separating, playful or serious) [84]. It is this interpretive layer within a set of similar objects that makes typologies a valuable analysis method. We can step out from just the formal representation of space and shape and ask, how does this form or behaviour impact human needs and experience.

Compared to a systematic literature review, a design typology includes references to artefacts regardless of whether it has received formal user evaluation or received previous research analysis. The reasoning is that much of the work happening in the VR music field is happening outside academia, so rather than reflecting design parameters only within previous academic dialogues, design understanding should also be based on practice.

Compared to a taxonomy, typology is preferred for this work, as the separation of types is non-hierarchical and potentially multi-faceted. Classification is done according to structural features, common characteristics, or other forms of patterns across instances. Within a typology, there is no implicit or explicit hierarchy connecting different research artefacts and products in VR. Also depending on the granularity of the type suggested, a single artefact may exist within two types simultaneously. Using typologies, themes of significance can be traced across systems, these patterns may describe best practices, observe patterns in interaction, explain good designs, or capture experience or insight so that other people can reuse these solutions.

6.3 Design Analysis

6.3.1 Methodology

As a formal process the typology was built upon identification, selection, and coding of audio-visual virtual spaces.

Identification: Literature gathering was achieved by parsing VR examples from the Musical XR literature dataset. Practice and product examples were gathered across the first author's thesis research period using search engines, internet forums, interviews, and social media [25].[2]
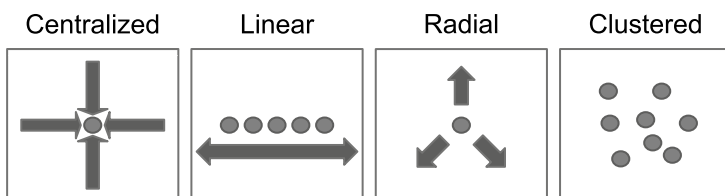

Fig. 6.3 Example of a spatial typology of form within architecture, adapted from [22]

---

[2] https://github.com/lucaturchet/Musical_XR_publication_database.

SPACE TYPE

1. PERSONAL SPACE
allows for concentrated 'heads-down' work (thinking, reading, writing), deep work, and reflection; requires reduced stimulation to avoid distraction.

2. COLLABORATION SPACE
is used for group work, work-shops, face-to-face discussions, client meetings, or student–teacher consultations.

3. PRESENTATION SPACE
is used to share, present, and consume knowledge, ideas, and work results in a one-directional way (presentations or exhibitions)

4. MAKING SPACE
is used for model making and building; allows experimentation, play, noise, and dirt.

5. INTERMISSION SPACE
connects other space types; is used for breaks, recreation, and transfers; includes hallways, stairs, cafeterias, and outdoor areas.

SPATIAL QUALITY

A: KNOWLEDGE PROCESSOR
space can store, display, and foster the transfer of information and knowledge (tacit, explicit, and embedded knowledge).

B: INDICATOR OF CULTURE
space suggests a specific be-havior, either through common sense, written or unwritten rules, rituals, labels, and signs.

C: PROCESS ENABLER
space can provide specific spatial structures or technical infrastructure that might guide or hinder the work process.

D: SOCIAL DIMENSION
space influences social interac-tions and facilitates meetings and personal exchanges.

E: SOURCE OF STIMULATION
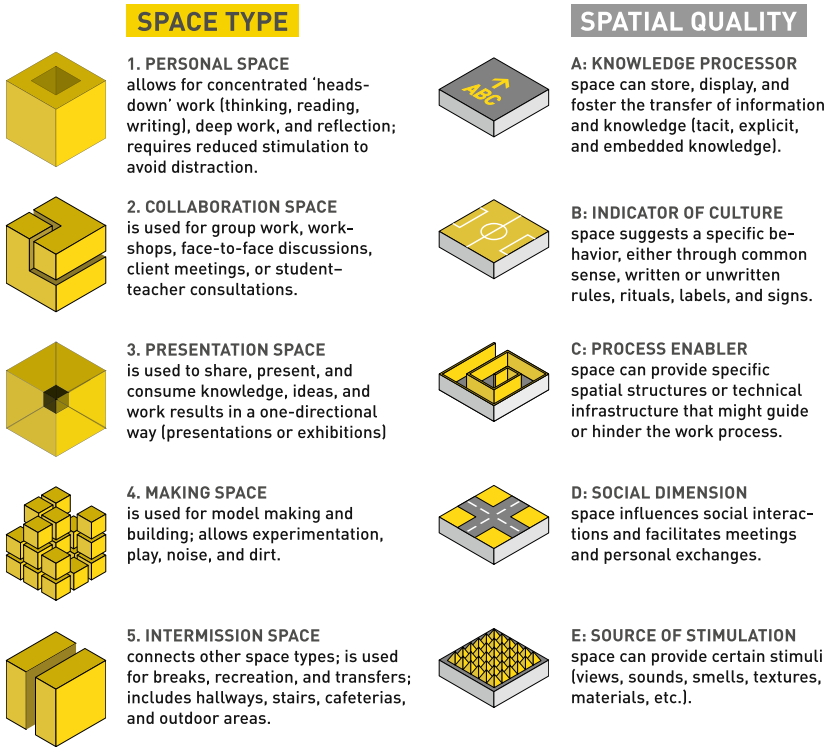space can provide certain stimuli (views, sounds, smells, textures, materials, etc.).

Fig. 6.4 Example of a spatial typology within design, taken from [84]. Reprinted from design studies, 56, Thoring *et al.*, creative environments for design education and practice: A typology of creative spaces, 54–83, Copyright (2018), with permission from Elsevier and Katja Thoring

Selection: Findings were assessed for relevance to the analysis. Cases were included on the basis of the following criteria; (1) Is the system based on immersive VR technology via an HMD? (2) Is the primary function or design intention of the artefact related to sound or music?

Coding: A form of deductive and inductive thematic coding was undertaken, based upon thematic analysis [17]. An inductive approach involves allowing the data to determine your themes, whereas a deductive approach involves coming to the data with some preconceived themes you expect to find reflected there, based on theory or existing knowledge. For this research, the deductive element was the setting of top-level coding categories (UI, Space Use, Social Engagement, Skill Level, Interactions) that probe how a VR IAS was constructed, the questions used are available in Table 6.1. The inductive coding reflects themes within the deductive categories based on the interface designs. Coding sources would involve: Use of the VR system where possible; review online video sources; analysis of images; and review of documentation and published literature. In each activity, notes and open-coding were undertaken on system design using qualitative data analysis software.

Table 6.1 Coding system developed for typology. Bold codes indicate deductive code categories, italics are inductive themes

Code	Description
UI	**What are the types of UI exploited in the VR interface?**
Screen-like	3D or 2D UI is used in VR that behaves like a standard screen menu or workspace
3D Objects	3D UI is used for information and action
None	No conventional UI or SUI is provided to users, such as an open world terrain or an external musical controller
Physical	No functional UI or SUI offered inside is VE, but external hardware musical controller used
Space Use	**How is space used in this device?**
Sonic	The positions of people or objects in space has an impact on sound processing, space as a functional element of the sound design process
Visual	Interactive visual feedback provided based on positions or orientation in space of people or objects
Social Engagement	**What number of users was the application designed to support?**
Solo	Single user spaces with no intelligent-agent interaction
Collaborative	Multi-user or single/multi user with intelligent-agent interactions
Collective	Massive multiplayer environments, both human and agent-based
Skill Level	**Was the system designed for novices, experts or both?**
Novice	
Expert	
NA	No formal user study conducted
Interactions	**What is the flow of action and the related system response?**
Sonic-Visual	Coupling between sound and visual features, where sound changes visual features
Visual-sonic	Coupling between visual and auditory information, where the visual information changessound properties
Sonic-sonic	Audio input used to control system features that relate to sound

After this, the deductive sweep was undertaken where the sources, open-codings and notes were reviewed in the context of each deductive category, and this resulted in the inductive themes that can be found in Table 6.1.

6.3.2 Typology of Virtual Reality Interactive Audio Systems

Here a typology of VR IASs is proposed, delineating how different systems overall function and the use of space in their design. The referencing of work in this section differentiates between commercial products and academic publications, using two

different reference sections for clarity. The typology is split into two broad categories within which VR products and research are discussed:

1. Type of Experience/Application—here we collate instances of products and research by their function as a sound and music system in VR.
2. Role of Space—in this phase we look across the different types of systems to suggest how the design of space can be categorised.

6.3.2.1 Type of Experience

Most implementations of interactive VR sound and music systems fall into one or several of the categories in the subsequent list. Many cited products have no formal user testing results available.

- **Audio-Visual Performance Environment**: Audience-oriented systems for playback or live performance of compositions with audio-visual interactions [14, 51, 101, 109]. For audience-oriented systems, interactivity is related to being part of a social group of spectators, rather than being able to interact sonically.
- **Augmented Virtuality (AV)**: A VR HMD acts as a visual output modality alongside physical controllers or smart objects, creating a AV system [34, 43, 100]. This descriptor excludes augmented reality (AR) technologies, such as HoloLens, as the visual overlay effect is considered different to the total re-representation of visual stimuli that occur in VR [99].
- **Collaborative**: Some form of collaborative interaction occurs in the VR audio system (human or agent-based). The interaction must be to directly make sound/music together [12, 25, 51, 63, 103, 110, 119], rather than more presentational systems like an audience cohabiting with performers in a virtual shared space; denoted by the *Audio-Visual Performance Environments* category. Examples and design considerations are described in Sect. 6.4.
- **Conductor**: Controlling audio-visual playback characteristics of pre-existing composition [51, 117].
- **Control Surface**: VR as a visual and interactive element to manipulate an existing digital audio workstation (DAWs) functionality, e.g., Reaper [104].
- **Generative Music System**: Partial or total algorithmic music composition, where the sound is experienced in VR space, and/or controlled by spatial interaction in VR [57, 116].
- **Learning Interface**: VR systems to support the learning of music, either as performance tutoring, theory, or general concepts in music such as genre [48].
- **Music Game**: Systems where gameplay is oriented around the player's interactions with a musical score or individual songs. A good example is Beat Sabre [102], the highest selling VR game of all time at the time of publication.
- **Narrative and Soundscape**: Pieces that integrate interactive audio in virtual reality [85, 116].
- **Physics Interaction**: Physics-based sonic interaction systems [27, 106].

- **Sandbox**[3]: Designed like visual programming languages for digital sound synthesis—such as Pure Data, Max/MSP, and VCVRack—these VR sandboxes use patching together of modules to create sound. [112–114]
- **Sequencer**: Drum and music sequencers in VR. As sequencing is a common thing in many musical applications, this category refers to interfaces that are either just a sequencer or use sequencing somewhere within their interaction design [27, 63, 103, 110, 112, 119].
- **Spatial Audio Controller**: Mixer style control of spatial audio characteristics of sources and effects [9, 25, 27, 43, 69, 90, 104].
- **Sounding Object**: Virtual object manipulation with parametric sound output [67, 68].
- **Scientific Instrument**: VR systems designed to test an audio or interaction tool/feature, a good example is a VR-based binaural spatialisation evaluation system [35, 73].
- **VR DAW**: Virtual audio environment, multi-process 3D interfaces for creation and manipulation of audio. Important feature is the recording of either audio or performance data from real-time interaction. Interface abstraction and control metaphors may differ significantly to conventional desktop DAWs [12, 27, 88, 103, 110, 119].
- **VRMI**: Virtual modelling and representations of existing acoustic instruments or synthesis methods [9, 12, 19, 31, 34, 51, 56, 61, 66, 68, 71, 80, 110, 114, 118].

Overlaps and Contrasts

Due to the broad design scopes of some systems, an artefact can appear in multiple categories, or exist in a space between two categories. For instance, [51] is in *Audio-Visual Performance Environments*, *Collaborative*, *Conductor*, and *VRMI*. While [12] is a technically a *VR DAW*, the audio and interaction design concept is highly idiosyncratic, so it becomes closer to a *VRMI*. The following statements intend to clarify any issues regarding overlaps in terminology.

- *Sounding Objects vs. Physics Interaction*: Both types refer to physics-based interactions, sounding objects are when the mesh structures of objects are the source of sound generation/control (e.g. scanned synthesis of an elastic mesh), whereas physics interactions include collision-based interactions for sound generation or use of physics systems to control single or multiple audio features (e.g. parameters or spatialisation). The interested reader might refer to Chap. 2 for more details on these topics.
- *VRMI vs. Sandbox*: While both can refer to synthesis methods, sandboxes are specifically modular construction environments, whereas synthesis methods in VRMIs would be a closed form of synthesiser e.g. playing a DX7 emulator in virtual reality.

---

[3] Category name and description sourced from [4].

6.3.2.2 Role of Space

Many of the systems outlined above offer novel interaction methods coupled with 3D visualisation. Looking at how space is used in VR music and audio systems provides a different way to group research and design contributions. For simplicity, the following categories are presented as discrete areas, but dimensions would also be suitable (i.e. systems could belong to several categories, see [15, 39] for examples of dimension-based classification for digital musical instrument (DMI)).

Space as a holder of elements for musical input/sonic control The most dominant form of spatial design is to use space as a container for interactive elements that either produce sound or control sound in some way. Within this category, key differences are whether menu-based SUI is used [103], or more object-based 3DUI is exploited [12]; this is discussed further in the next section. Other works include: [19, 27, 31, 56, 61, 63, 66–69, 71, 88, 100]. [104, 109, 110, 112–114, 118]

Space as a medium of sonic experience In these sorts of systems, space is woven into every aspect of user experience or system design. For instance, in [9], the sonic operation of the VR system makes no sense if users do not engage in collaborative spatial behaviours [9]. In this category, the relationship of spatial interaction to system feedback can be predominantly passive, like a recorded soundscape [85], or fully interactive, like an audio-visual arts piece that maps spatial input to output modalities [90]. In some cases, visual space may only be a supporting medium for a spatial sonic experience [85]. It is worth noting that **spatial audio controllers** are not instantly considered as part of this category. As **spatial audio controllers** deal with controlling and manipulating elements, they are considered to be part of the *Space as a holder of elements for musical input/sonic control* category. Rather, this category holds experiences where spatiality is more intrinsically involved in the interaction between elements and user experience, whereas in a controller system it is a functional relationship. Other works include: [43, 57, 80, 106].

Space as a visual resource to enhance musical performance In this category, space is primarily used for its visual and spatial representation opportunities rather than as a direct control system or as an intrinsic part of the sonic experience derived from the system. Designers use space as an extra layer to a music performance or system, for example, this can be to:

1. Present performers' with enhanced visual feedback related to their Playing of a musical instrument [34];
2. Provide a space for an audience to contribute to a collective experience of musical performance [14]; or
3. Use space as a place for an audience to convene for a music performance in VR [101, 109].

6.4 Spatial Design Analysis Case Studies

The state of the art in VR audio production and immersive musical experiences include single-user and collaborative approaches. In the following case studies, the spatial and social design decisions are discussed; noting that each of the systems serves different purposes as musical experiences. Our motivation is to further detail design typology categories, by understanding and comparing the decisions VR designers make. Reviews are broken into four areas: *single user systems*, *collaborative systems*, *collective systems*, and *spatial audio production systems*. The reason we focus on these previous areas, only within immersive music and interactive sound production, is so that design comparisons and implications can have some level of shared context. We chose the field of immersive music as a point of shared interest between academia and industry. But it would be valuable to probe design decisions comparatively between broader fields of SIVE design, for instance, auditory display and sound production systems; however, this would be a different contribution.

6.4.1 Single-User Systems

Figure 6.5 shows the *music room* [118], an *instrument space* containing multiple VRMIs that are designed to be played with the VR controllers, following a DAW-like workflow of perform and record, then arrange and edit. Instruments include a drum-kit, laser harp, pedal steel guitar and a chord harp. The spatial setup mimics a
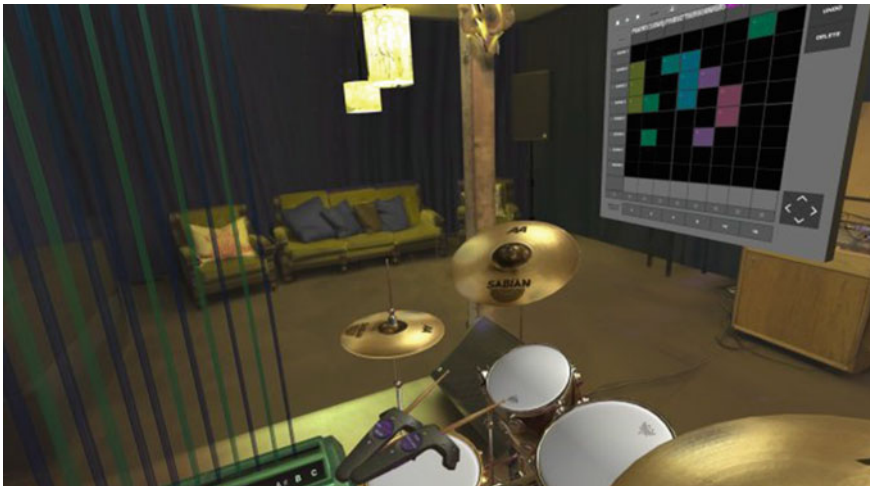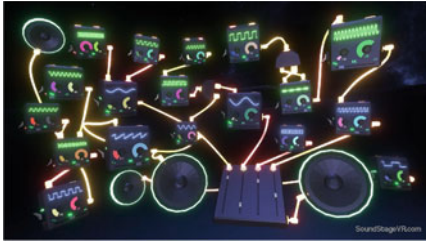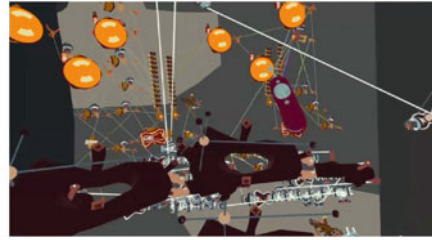

Fig. 6.5 Single-user VR spatial design considerations A—*music room* instrument space, with drum kit instrument being used and the recording panel UI visible, displaying previously recorded data

(a) SoundStage VR for HTC Vive. Sound control devices float in space and are connected by wires, emulating a modular synthesis workflow.

(b) A Mux composition, highlighting the complexity of the workspace.
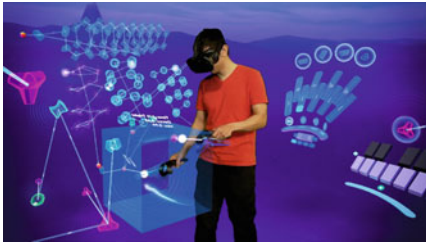
(c) LyraVR user editing a 3D parameter cube widget surrounded by other interface elements, highlighting the spatial complexity of node-edge musical structures.

(d) LyraVR: view of 3DUI elements that float in fixed space.

Fig. 6.6 Single-user VR spatial design considerations B—sandboxes, node-edge structures and modular systems

conventional studio. In Fig. 6.5 we can see spatial 2D graphical user interface (GUIs) presenting recorded information and menu function, while 3DUI objects are used to represent instruments, and a 360 photograph of a real studio provides the visual backdrop. A design decision of the space was to situate all instruments in a circle around the user, presumably to be able to play all the VRMIs in a small physical space. Two areas are utilised for the UI, *action space* and *display space*. The action space is for the VRMIs, and the display space, further away from the user, provides a conventional GUI. To be able to interact with the distant GUI, laser pointers are used.

Sound Stage [114] (Fig. 6.6a) and *Mux* [113] are modular instrument building *Sandboxes* in VR. Users can define their own systems to perform music through those systems. Both are multi-process VRMIs designed for room-scale interaction. In these systems, a user surrounds themselves with modules and reactive widgets, and 'patches' them up using VR controllers. While stimulating and highly interactive, the resulting virtual spaces can be complex and messy spatial arrangements (author's opinion); Fig. 6.6a shows an example of a sound system made with Mux, highlighting the spatial-visual complexity. One possible reason arrangements become complex is

because spatial organisation is arbitrary and user-defined. A novel spatial feature is that speaker scale controls source loudness, and this turns a slider or number UI into a 3DUI interaction process.

The *LyraVR* [112] and *Drops* [106] are two examples of *Sandbox* systems that build the temporal behaviour of the composition using spatial relationships. Figure 6.6b and c show *LyraVR* a musical 'playground' where users build music sequences in space to create audio-visual compositions. The node-based sequencer allows the creation of units in free space. Although aimed at single users, such interaction and playback method would be scalable to collaborative systems. *Drops* is a VR 'rhythm garden', where a user creates musical patterns using the interaction of objects and simulated gravity. The system requires setting up of object nodes ('eggs') that releases 'marbles' that make a sound when they strike other surfaces—the size and release frequency of marbles can be manipulated by the user. By adding more surfaces and modifying planes of movement for marbles, the musical composition is built using a 'physical' design process. In *LyraVR*, *Mux*, and *Sound Stage*, users interact with sound elements via spatial node-edge structures, and this gains a level of immediacy for musical changes at the cost of vision-spatial complexity. But the embodied control of temporal musical behaviour via the arbitrary positioning 3DUI does create an experimental creative process driven by interaction in space.

### 6.4.2 Collaborative Systems

Block Rocking Beats [103], *LeMo* [63], and *Polyadic* [25] are collaborative music making (CMM) *Sequencers* . However, the systems have different approaches to spatial design for collaborative interaction. Both *LeMo* and *Polyadic* are the only collaborative systems in this review that have undergone formal user studies [25, 63, 64].

Block Rocking Beats, Fig. 6.7a and b, enables avatar-based (head and hands only) remote collaborative music production in a virtual sound studio for up to three people. The space is modelled like a futuristic studio, adapting a conventional layout of production equipment areas and multiple screens. The environment provides a sequencer interface for each user while project information is displayed on a single large screen within the environment, and this provides some level of shared visual information. Additionally, reactive systems alter environment appearance in sync with music created. As a spatial layout, users' positions are fixed in the space, a few meters from each other in a semi-circle facing the front screen. The layout limits the capacity to view each other's workspaces and may inhibit forms of mutual monitoring. Regarding avatar design, the character's design is highly stylised, and the 'hand' representation is designed like a tapered wand. The taper is designed to enlarge the usable sequencer area, as when buttons are designed at a normal scale the size of the controller would hit multiple buttons.

The *LeMo* allows two co-located users avatar-based CMM in VR, using a variety of sequencer instruments [63–65]. Depending on experimental condition, different

(a) *Block Rocking Beats* collaborative composition environment, two users pictured working in their own areas.

(b) *Block Rocking Beats*, view of single user display area.
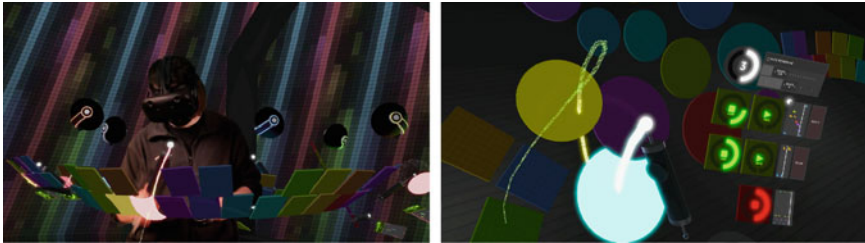
(c) *EXA* collaborative composition environment. Freely configurable space for multiple people to musically perform and sequence sounds.

(d) *EXA* collaborative composition environment. View of a user inputting information with an instrument context window open to the side.
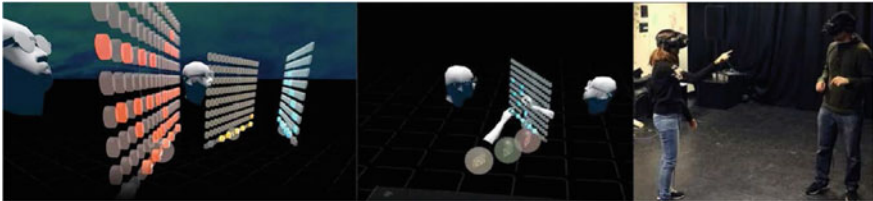
(e) *LeMo* collaborative composition environment [75–77]. Freely configurable space for multiple people to sequence sounds. Transparent SUI used to sequence the sounds, sequencers can be minimised into 'bubbles' to rearrange space.

(f) *Polyadic* collaborative drum sequencer [28]

Fig. 6.7 Collaborative VR music production interfaces

spatial features would be activated, such as private workspace areas and spatially reactive loudness. Studies of *LeMo* evaluated visual and sonic workspace design, based on the concept of public and private territory, developing design implications for SVR; for detailed findings please consult Chap. 8 of this volume. Barring the experimental findings, as a spatial design, compared to *Block Rocking Beats* and *Polyadic*, *LeMo* allows users to move and rotate their workspaces to accommodate social interaction around the task of music making, commonly using face-to-face or side-by-side arrangements (see Fig. 6.7e). A novel design feature of note is that SUI sequencers can be minimised into 'bubbles' to rearrange space. As these sounds are spatially located, the bubble acts as both a UI and an audio object. Additionally, the inclusion of 3D drawing as a communication medium enables a variety of annotation behaviours. Like *Block Rocking Beats* and *Polyadic*, avatar design was rudimentary offering a head with gaze direction, however, the use of Leap Motion as the input device enables more detailed hand representations. These were used for functional input and social communication, e.g. waving and pointing.

Polyadic enables collaborative composition of drum loops to accompany backing tracks for two co-located participants [25]. The system is designed to be instantiated in two user interface media, VR and Desktop. The design motivation of *Polyadic* was to compare VR and Desktop media concerning usability, creativity support, and collaboration. In order to create a fair comparison of media, constraints were imposed on the design of both media types. This limited the design of features to only use control methods that could work equally across both conditions, namely a standard step sequencer with per step volume and timing control. In the VR condition, the environment uses fixed placement of 3DUI sequencers made up of virtual sensor buttons and sliders, see Fig. 6.7f. Low fidelity avatars were utilised to allow rudimentary social cues. Avatars used a sphere head with 'sunglasses' to indicate gaze direction and two smaller spheres to indicate hands, enabling simple spatial referencing. Additionally, each user's workspace and interface actions were replicated within the other users' environment, enabling referencing and looking at what the other is doing.

EXA [110], Fig. 6.7d, is a collaborative *Instrument Space* where multiple users can compose, record, and perform music using instruments of their own design. *EXA* differs from the previous examples as users input musical sequence information in real time using drum-like instruments, rather than pressing sequencer buttons. Once sequences are made they can be edited using menus and button presses. Similar to *LeMo*, EXA allows users to freely organise their workspace in line with collaborative needs. Also, the custom design of VRMIs introduces idiosyncratic uses of space in order to perform each VRMIs. Like others, *EXA* utilises simple head and hands avatars.

6.4.3   Collective Systems

The following reviews are special cases, social VR platforms designed for musical experiences, pictured in Fig. 6.8. As predominantly music visualisations in VR, there is limited sonic interactivity for users. So the focus is on how these spaces act as collective social experiences in VR. For broader discussion of music visualisation in XR, see [92]. While not sound production platforms in themselves, the experience of a collective engagement in VR, related to audio-visual performance, is an area of immersive entertainment where new production tools and design experience are required.

The WaveVR [101], Fig. 6.8a, is a cross-platform social VR experience, like going to a 'gig' in VR. Artists can use it to make audio-visual experiences for audiences across the world. As a virtual space, the shared focus of a stage is used for most performances, but the virtual space is reconfigured for each 'gig'; similar to different theatre performances all taking place on the same stage. In one instance, music toy spaces were designed for the audience to interact with musical compositions, these took the form of objects that change the level of audio effects based on spatial position or touch interaction. As the objects cannot all be controlled by one person, this creates a collective 'remix' of the content [111]. For further reviews of some individual 'gigs' in *The WaveVR* see [6].

Volta is an immersive experience creation and broadcasting system [108]. Performances are rendered in VR using artists' existing tools and workflows, such as parameter mapping a DAW to drive visual feedback systems. In addition to the VR performance, a mixed reality (MR) experience is also broadcast to streaming platforms like Youtube and Twitch. Volta differs to *The WaveVR* in its production method for the artist. In *The WaveVR* developing a performance environment can take a development team months to build, and a significant cost. *Volta* cuts down production time by integrating existing tools with spatial experience design templates (e.g. particle systems), into a streamlined production process for real-time virtual performance environments. [4]

### *6.4.4   Spatial Audio Production Systems*

In the following review of spatial audio production systems in VR, all systems use binaural spatial sound presented over headphones (Chaps. 3 and 4 provide an effective introduction to such audio technology). It is possible for some of the systems (*DearVR Spatial Connect*, *ObjectsVR*) to be used with speaker arrays but the design implications of this are not considered in this review.

Addressing spatial audio production, both the *Invoke* [25] & *DearVR Spatial Connect* [104] systems allow users to record motion in VR to control sound objects.

---

[4] The first author supported the design of early prototypes of Volta XR, interested readers can review the design development at https://thefuturehappened.org/Volta.

(a) The WaveVR's space for an audience to spectate the audio-visual experience.
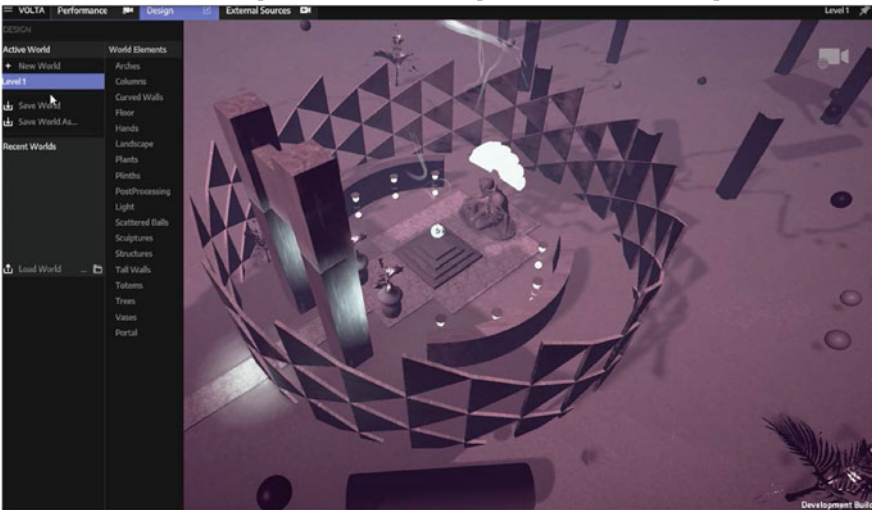

(b) Volta XR's spatial creation workflow, picturing an audience space in development.

Fig. 6.8 Collective music experience spaces in VR

The main functional difference between the systems is that *DearVR Spatial Connect* uses a DAW to host the audio session with the VR system acting as a control layer for spatial and FX automation, while *Invoke* is a self-contained collaborative spatial audio mixing system. The systems also differ in their design approach to space and sonic interaction.

Figure 6.9a shows *Invoke*, a collaborative system that focuses on expressive spatial audio production using voice as an input method. The system utilises a mixture of direct and indirect spatial interaction to record spatial-sonic relationships. A *Voice Drawing* feature allows for the specification of spatio-temporal sonic behaviour in a
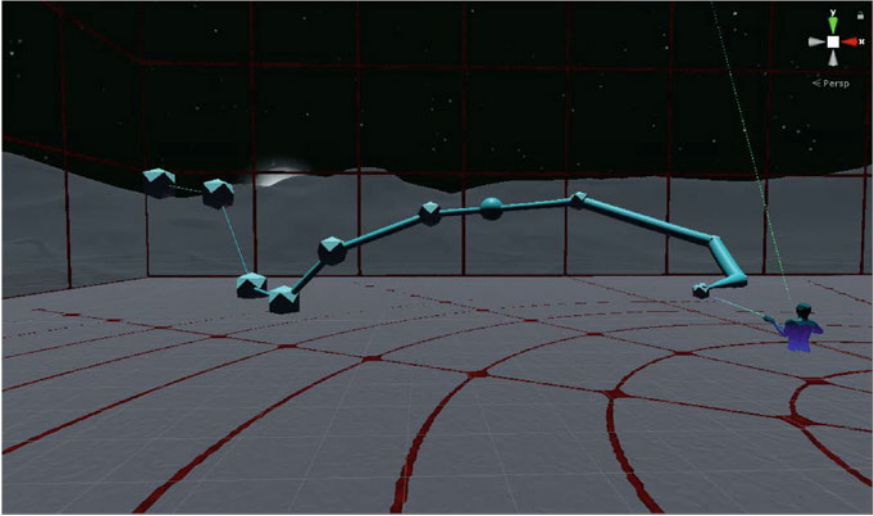
(a) Invoke spatial audio controller. Image is of a user manipulating audio trajectory points.
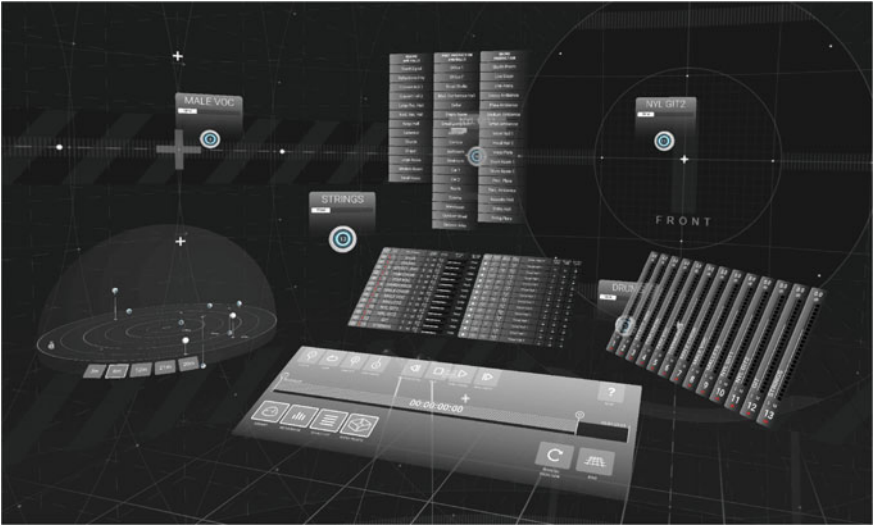
(b) DearVR Spatial Connect system with all interface modules active and five sources positioned in space.

**Fig. 6.9** VR spatial audio production systems

continuous multimodal interaction. Voice input is recorded as loudness automation, while a drawn trajectory controls the location of the spatialised audio over time. Using an automated process the trajectory is segmented in a bézier curve with multiple control points for further collaborative manipulation. The UI design uses a mixture of 3DUI (audio objects, trajectories) and semi-transparent 'screens-in-space' (hand menus, world-space menus). Spatially, users can navigate the virtual space using teleport functionality, all menus travel with the user when they teleport. *Invoke* is the only system in this review to implement more detailed avatar design, each user is represented by a body, head and arms, utilising additional sensors on each user to provide accurate body-to-avatar positioning. This enabled detailed forms of social interaction and spatial awareness [25].

Figure 6.9b shows *DearVR Spatial Connect*, a professional spatial audio production application. The system uses indirect interaction method to control objects in space; a laser pointer controls position while the VR controller thumb-stick controls distance from the centre. The design of the surrounding space adds no features beyond the interface panels and 3DUI (e.g. sound sources), as users commonly project a 360 video into the production space. Also, the user is 'pinned' to the centre of the space, again in line with the rendering perspective of spatial audio for 360 video. One issue of the central design is a lack of perspective on multiple objects that may be distant from the centre. Also, fatigue and motion noise (distant object 'wobble' more spatially) impact control of objects at a distance (dependent on input device design and user-based ergonomic factors like strength and motor control) [5]. Comparing this to *Invoke*, which does not constrain users to the central listening position when mixing audio objects, users can freely teleport around to gain different sonic and visual/interaction perspectives. This is important as the spatio-temporal mixing of sound creates a complex field of trajectories and sound objects [25].[5]

ObjectsVR is a system for expressive interaction with spatial sound objects. The system provides spatio-temporal interaction with electronic music using 3D geometric shapes and a series of novel interaction mappings, examples can be seen in Fig. 6.10. User hand control is provided via a Leap Motion, and the experience is rendered using a HMD. As a spatial audio control system, object positions were a mixture of direct manipulation and 'magical' physics-based interaction. Users could pick up and throw sounds around the space, but an orbiting mechanic meant that sound objects would always move back within grabbing distance. A novel spatial feature of this environment was the use of contextual UI when users grabbed certain objects. When a user grabbed objects that had 3D mappings, a 3D grid of points would appear to provide relative positioning guidance. When released the grid fades away. System design and evaluation investigate users' natural exploration and probes the formation of understanding needed to interact creatively in VR, full details of the evaluation can be found in [27]. [6]

---

[5] The first author participated in formal beta testing of the DearVR Spatial Connect product.

[6] *ObjectsVR* was a single-user system designed and tested by the first author during a research internship at a VR experience design firm.
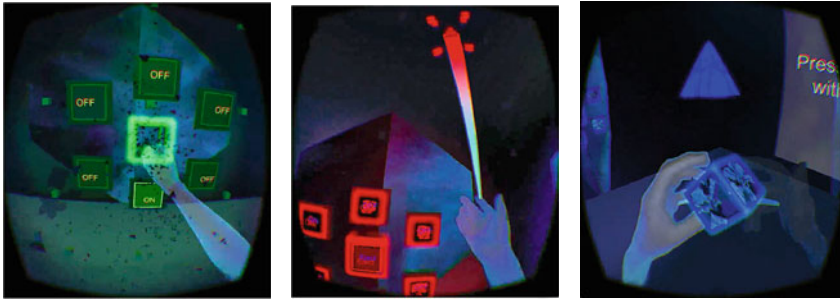
(a) User experimenting with Audio-Visual Feedback used to indicate an error of inserting the wrong type of item into a holder

(b) User utilising magnetic grabbing technique to grab object at a distance. Object controls drum track sound effects

(c) User exploring a bass cube transferring it between both hands

Fig. 6.10 ObjectsVR interface user interaction examples

6.5   Discussion and Implications

6.5.1   Spatial Design Considerations

Consolidating the reviews of products and research, a series of design parameters emerge.

Complexity of spatial representation

Based on the analysis of *Sandbox* systems (Mux and Sound Stage), it is suggested that an unrestricted patching metaphor may be too visually complex for applications like collaborative audio production in VR. Also, systems that build the timing of compositions in space, *LyraVR* & *Drops*, suffer from spatial-visual complexity issues. Similar to visual programming languages [36], when all points of state-change are presented in one space (a low level of abstraction), the information becomes diffuse, and errors may become more frequent. Also, when space is used for functional relationships, like musical time, visual design cannot bracket the visual complexity without the design of abstractions. Related to these issues, the impacts of these design features is unknown for collaborative systems. Future research could design systems to observe spatial organisation patterns undertaken by users to make sense of, and work with arrangements.

Screens-in-space and workspace zones

For certain information (selection menus, settings, note sequences), systems use either conventional 2D information presentation in a floating screen (*Music Room, Block Rocking Beats, EXA, DearVR Spatial Connect*), or attempt to redesign information using forms of 3DUI (*Lyra, Mux*). Also, as described in the *Music Room* analysis, space can be delineated into different action or information presentation spaces. The decision to locate functionality in screens or more novel 3DUI is an

important one for collaborative systems, as each different method offers different access points and levels of shared visual information for collaboration. For instance, in *LeMo*, each SUI could be minimised into a bubble for easy arrangement and organisation. A temptation of VR design could be to embody all interaction in 'physical' 3DUI, such as novel interaction widgets or spatially multiplexed 3DUI (see Fig. 6.2). But this could result in added spatio-visual complexity like in *Sandbox* systems, to deal with this there would be a need for contextual interaction layers (e.g. when I put a cube here its different from when I put it there), or function navigation using button combinations on controllers (VR 3D modelling software do this [107]). Another impact of using entirely 3DUI is that it could limit the amount of shared visual information, as arrangements of 'physical' objects naturally obscure each other. However, 3DUI may provide more access points to embodied collaboration.

Level of acoustic spatial freedom
Related to spatial audio the ability to move from the centre position is a key design decision that needs to be made, especially for collaborative audio production software. For single-user apps, being able to manipulate arrangements, away from the sweet spot is of value. For collaborative apps, multiple users located at the sweet spot would severely impact normal social interaction.

Workspace organisation
For workspace organisation, it should be considered whether fixed or movable UI is preferred for certain audio production tasks. For instance *LeMo*, *EXA*, and *Invoke* each utilised methods for users to reorganise the SUI, while artefacts like *Block Rocking Beats* and *Polyadic* did not.

Control, Play and Expression
Designers should consider how playful they make spatial audio experiences, or whether specific control and sound automation is the design target. For instance, in the *ObjectsVR* system spatial audio objects had 'magical' interaction, contrasting this, *DearVR Spatial Connect* emulates DAW automation. What is missing here is more examples of user experience in mixed systems, and environments to playfully explore spatial sound interactions with levels of direct control and serendipity. Related to making experience of control more expressive, integrating different modalities provides opportunities to expand on the DAW control paradigm, such as in *Invoke*.

Egocentric spatial design
Related to the previous two features, some systems (e.g. *Mux, Music Room*) tend towards egocentric spatial patterns, with devices and elements situated around the user, oriented to one spatial viewpoint. While making sense for an individual application, these forms of design decisions need to be carefully considered in collaborative systems.

Avatar Design
An issue of importance to collaborative systems is avatar design and the spatial behaviours that they enable. For instance, inside *LeMo*, the use of the Leap Motion

compared to standard VR controllers enabled more detailed forms of hand gesturing. Within HCI work has already begun to evaluate avatars based on the constraints of commercial VR [53]. What this area should focus on is moving beyond the so-called *Minimalist Immersion* in VR using only simplistic avatar design. Within *Invoke*, the avatar design utilised a more detailed body representation, offering beneficial characteristics for social space awareness, as users can interpret gaze and body orientations along with hand gestures. This highlights an important area of further research for collaborative and collective systems, where there should be detailed evaluations of the avatar designs' impact music production activities.[7]

6.5.2 Role of Space and Interaction

Comparing the separation of the *Role of Space* with previous research on the *space of interaction* [75], similarities emerge. River and MacTavish analyse space, time and information concepts within HCI across a series of paradigms [75]:

- Media Spaces [86]—media types
- Windows, icons, menus, pointer (WIMP) [47]—user space management
- Tangible user interface (TUI) [44]—space-body-thing interaction
- Reality-based interaction (RBI) [49]—emerging embodied interaction styles
- Information spaces [10]—interaction trajectories and navigation of information
- Proxemic interactions [37]—social spatial relationships

The key spatial dimensions that emerge are:

Dimension 1   *Media and Space Management ↔ Meaning through interaction*
Dimension 2   *Personal and physical ↔ Social and behavioural*

Dimension 1 describes the difference between conventional GUI design (e.g. WIMP) versus approaches using space and the embodiment of technology (e.g. RBI). Dimension 1 relates to the previous analysis on the *Role of Space* (Sect. 6.3.2.2):

- Space as a holder of elements for musical input/sonic control
- Space as a medium of sonic experience
- Space as a visual resource to enhance musical performance

Dimension 2 highlights how space influences personal and social interactions. This is because information is distributed across technologies and is also embedded into contextual spaces, from immediate personal space through to social groups and larger collective social interaction spaces. Looking at these ideas together, a framework of research emerges for VR IAS spatial design. The functional uses of space in VR IAS relates to traditional understanding in the design of media types, user space management, and TUI. While *space as a medium of sonic experience* can benefit from

---

[7] Preprint available at https://hal.archives-ouvertes.fr/hal-03099274.
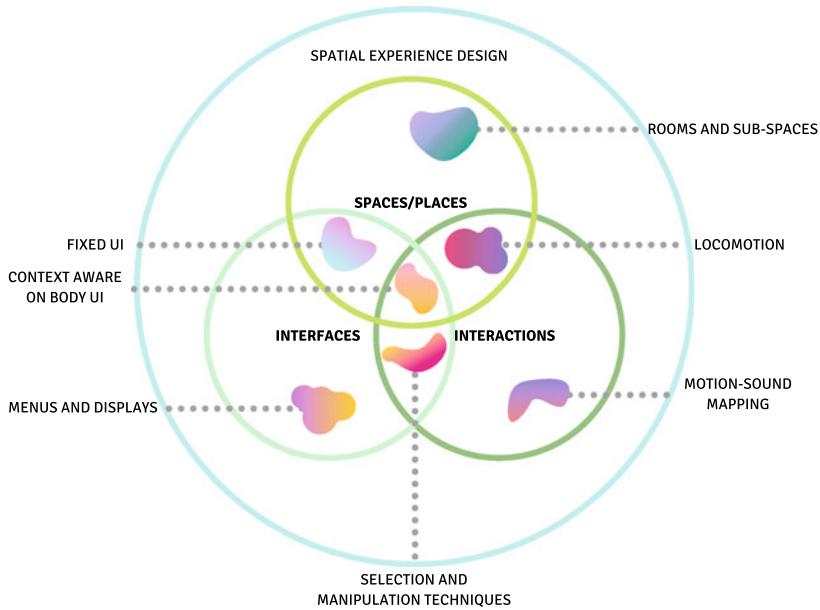
Fig. 6.11 Spatial experience design in VR IAS Venn diagram

research in the areas of RBI, and information spaces. Finally, proxemic interaction can inform things like social spaces for musical enhancement. But this doesn't go far enough. What needs to be included in space for interactive audio is an understanding of architectural space. This is because VR designers must make important decisions regarding space as an element of user experience. Regarding social aspects, as highlighted earlier in Fig. 6.4 [84], we can design space for functions, activities and for their spatial quality. We must design spaces for intimate individual action, shareable group interaction, and visibility and safety in large collective action spaces. Acoustically the sorts of choices we make here matter too. For example, using simple voice chat algorithms could make voice intelligibility poor and yield something similar to 'zoom fatigue' [7]. Instead, we can utilise spatially aware audio communications to deliver intelligible audio for each user in an area of space [60], a commercial approach to this already exists that can handle hundreds of listener-sources across a space [115].

We suggest that spaces need elevated priority in our VR design and evaluation practices. To support this process, we suggest three top-level spatial categories that need to be addressed through interdisciplinary design work: *spaces/places, interfaces, interactions*. Visualised in Fig. 6.11, some of the elements discussed in this chapter are positioned within the different design spaces; for instance, VR selection and manipulation techniques sit between *interfaces* and *interactions*. For brevity, only the category of *spaces/places* is discussed in detail below, as previous research within *interfaces* and *interactions* is already well documented in this chapter and other

research [6, 12, 77]. The categories scaffold future design by drawing together topics, theories, and previous art. Addressing elements that overlap with *spaces/places* in Fig. 6.11, we can use the Venn structure to ask new questions about the interaction of spaces in feature design. For instance, context-aware on body UI refers to the idea that if we have more specific spaces for interaction we can also tune the needs of UI to be relevant to that moment in space and time. The notion of putting it on our body, like a virtual smart watch, means that this design element is part of both *interfaces*, *interactions*, and *spaces/places*. Implicit in such simple categories is the equalising of spaces as a design concern alongside more thoroughly investigated work like spatial interfaces and spatial interaction. Fully describing such a framework is out-with the capacity of this chapter; instead, it is offered as a proposition for the research field to further explore together.

6.5.2.1 Spaces/Places

Spaces are the architectural layouts and areas that form features of a virtual environment used for sound and music activities in VR. An example of a space can be seen in Fig. 6.12. In that figure a central production area is enclosed in a grid/cage structure, bounding it off from the wider spatial setting of floating 'sand-dunes' and night sky. But what does it mean to design for experience within space, and how does this related to an IAS? Borrowing from human geography and architecture [22, 87], some spatial concepts to consider are:

1. Boundaries;
2. Form and space;
3. Organisations and arrangements;
4. Circulation (i.e. movement through space);
5. Proportion and scale;
6. Principles and metaphors (e.g. Symmetry, Hierarchy, Rhythm).

Places are spaces with fixed or emergent social meaning [32]. We can aim to design the spatial qualities of spaces, for instance, the typology of [84] in Fig. 6.4, gives designers ways to conceptualise creative spaces. We can ask, what is the space type (e.g. personal or collaborative), and what is the intended spatial quality (e.g. knowledge processor or process enabler)? Then we can ask, within those boundaries, what are other spatial characteristics i.e. comfort, sound, sight, spaciousness, movement, aliveness/animus?

As architecture, human geography, and interior design are such deep disciplines, interdisciplinary work needs to be done here to produce a dialogue around the design of space for sonic and musical expression. One area of mutual influence to consider is the design of immersive installations that involve technology to alter user experience. VR can learn from techniques and theories in this area [3], as well as be used to prototype systems for physical installation.
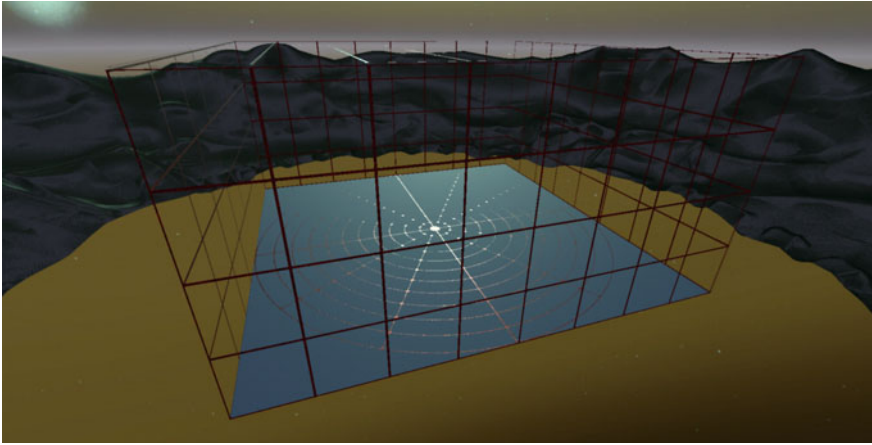
Fig. 6.12 Example of a VR IAS space, *invoke* artefact's spatial audio composition area

6.6 Research Directions and Opportunities

6.6.1 Embodied Motion Design

Echoing the design principles within Atherton and Wang's work [6], motion, embodiment and play are important design spaces to explore. However, human motion and spatial analysis is not a new discipline for computing and technology, with special research groups such as the International Conference on Movement and Computing (MOCO) and the ACM SIGGRAPH Conference on Motion, Interaction and Games (MIG). Within these existing dialogues, the role of embodiment is a central topic of design [83] (see Chap. 7 for further details). What would differ in virtual spaces is a form of synthesis, or symbiosis, between visual and proprioceptive embodiments. The plural is intentional, as virtual environments may introduce the idea that embodiment is not a fixed state, with avatars and motion feedback being augmented by the virtual setting. A research problem in this area is determining appropriate vocabularies for low-level and high-level motion so that systems of motion analysis and mapping can be utilised in an informed way. But the difficulty in VR IAS is systems will often need to utilise data from only the headset and controllers, where many previous approaches have been developed using high-resolution motion capture data [29]. Also, motion design is not just a single person experience. Take for instance dancing in a crowd. Research into virtual *togetherness* through joint embodied action is a rich direction for collaborative and collective systems to explore [40].

6.6.2 Designing for Collaborative Sound and Music in Virtual Reality

There is a paucity of design and evaluation frameworks addressing social experiences in sound and music VR. While work is ongoing in this area. For instance, Men and Bryan-Kinns' chapter in this volume (Chap. 8), to address the gap in design knowledge for VR, design perspectives from other embodied CMM and HCI research provide valid considerations for the design of SVR. The following integration of research from other fields intends to offer SMC actionable research directions to support collaboration in VR.

Adapting Tangible User Interface Research

An area of potential influence on spatial design for social VR is to look at how TUIs are designed to support spatial collaboration. For example, [65]'s research on CMM in VR shows similar results to co-located CMM using TUIs [96], regarding the design of public and private workspaces. When designing TUIs for co-located CMM, spatial orientation and configuration are important design areas. The *Hitmachine* is a tangible music-making tool for children, focused on creating and understanding collective interaction experiences [38]. To understand interactions with devices like the *Hitmachine*, there is a need to design social interactions and technology together. For designers, this means specifying and evaluating how people *distribute attention*, *share attention*, *dialogue*, and engage in *collective action*. To analyse designs in context, spatial formations of peoples' positions and orientations can be analysed to understand different constructions of social play in CMM [38]. Observations of social engagement around *Hitmachine* found that the configuration of space (people, furniture, and music interfaces) altered the level of social interaction. Also, regarding the design of space in VR, research findings from VR CMM resemble the results from the *Hitmachine* analysis [64]: How spatial encounters are set up for music interaction impact social interaction. So, to design collective interaction spaces, how basic spatial partitions are implemented matters.

Another TUI design principle of relevance is to provide multiple access points to a collaborative task [45, 76]. This means devising multiple spatial ways for different users to act on the same object, creating a form of DUI. Research suggests that the more access points participants have to a collaborative task improves how equitable participation is [76]. Increasing the tangibility is also said to improve participation. This is because users can complement what each other are doing in spatial tasks, using space as an organiser of the shared activity [76]. Adapting tangibility to VR means designing the affordances of objects appropriately to allow collective spatial interaction, while keeping in mind that we can move beyond some of the constraints embedded in physical reality. A good example of this is in VR *Sandboxes*. In physical reality, physics governs layout patterns of blocks whereas in VR elements can be placed in any part of 3D space. This in turn impacts the design of modules how users connect them [6]. But as mentioned previously, idiosyncratic design patterns within *Sandboxes* may need additional support for collaboration, and this is where previous TUI work could be integrated [97].

Collectively, these similarities suggest that as a form of spatial collaboration, VR CMM can benefit from other non-VR research findings regarding spatial interaction to design systems. But, directly importing collaborative design concepts from other media should be done carefully, and thoroughly evaluated for any differences in results across media (see [25] for a media comparison study focusing on this).

Designing for Embodiment in Collaboration
Embodied spatial input and avatar representation are key features of VR for supporting intimacy [54], awareness and coordination [41], and control [1]. Spatial media, such as VR, has the capacity for visual and spatial abstraction of UI, something needed for the complex requirements of expert music production [28]. The following examples highlight some specific opportunities to support spatial collaboration.

Augmented Object Interaction   The affordances of embodied interaction in SUI offer possibilities to transform how joint action on complex digital objects can occur [1, 2, 8, 21, 55].

Awareness Support   Embodied control and spatial representation in VR can ameliorate mutual understanding issues in shared workspaces compared to other media [79]; support informal awareness to co-ordinate actions given shared visual information [30]; provide pointer mechanisms that support referencing of content and environmental objects [23, 94, 95]; allow for the recording of embodied motion, as a form of embodied memory within an environment [58, 63]; provide novel mechanisms for the division of labour and workspace organisation [64].

Spatial Problems   Space is a powerful organiser of human memory and can change how we solve problems [18, 50], and VR, compared to WIMP systems, is suggested to alter problem-solving strategy in spatial tasks [50].

These considerations have in common an influence on the interaction space in collaboration. This suggests that the collaborative process in sound and music production could be improved by designing support for augmented interaction and awareness. For example, in a common studio environment, usually, a shared screen (or set of screens), a keyboard and mouse, mixing desk with dedicated audio outboards, are the tools in the hands of audio producers. In contrast, in an embodied VR interface, the possible interaction space can centre around collaborative spaces where functionality is engineered to support mutual access and modification, adapting levels of visibility and position based on collaborative needs.

6.6.3 Spatial Audio Production for Immersive Entertainment

VR provides an ostensibly promising environment for spatial audio production, it is an example of professional workflow that could benefit from further research into interaction methods in VR. The spatial nature of the technology, and action in it, could support problem issues encountered when making audio compositions in space (e.g. transformation of spatial reference frames between self and audience)

[25, 26]. Regarding the previous analysis, a highly significant research area would be the management of complexity in the information design of spatial representation. The impact of these improvements would be felt within fields such as immersive entertainment, where spatial audio technologies allow the engineering of soundscapes that represent real or imagined sonic worlds, using the location of sounds in space as a critical component of audience experience. In particular, there is an under-explored research opportunity in VR to enable more collaborative practice for spatial audio production. This addresses a need in professional audio production communities that look to make content for immersive entertainment.[8]

6.7 Conclusion

Much of how we design VR is based on borrowed design principles. We import ideas from other disciplines and hope they 'fit'. But to capitalise on any opportunities for enhanced expression and new forms of sonic entertainment presented by VR, we must set out how we design, what that involves, and what that excludes. Given such a broad focus embedded in the concept of space, the first goal of any schematic representation of design types and guidelines is to find suitable descriptors to collect the features relevant to domains of research. For researchers, this means setting out the design rationale behind systems clearly, so that over time we can understand the emerging practice and propose novel directions. This research offered the beginning of this process for the design of IAS for VR, setting out the different functional types both research and commercial interests pursue while reflecting on the way space is implicated in their design. This provides a framework for spatial design, highlighting a set of actionable areas for future design research. From our perspective, a key missing piece is guidance about how to design spatial social experiences in VR for engagement with sound and music. We need to define the transitions between individual, collaborative and collective interaction when it comes to audio interaction. A stepping stone in this gap is more research into avatar design for SIVE, as to start assessing spatial transitions in social activity we need to understand virtual embodiment as the vessel that affords basic social communication beyond speech. Looking forward, we should begin to think about what it means to be an immersive application designer that is audio-first. Realising that practice will need to integrate concepts from acoustics, architecture, phenomenology, HCI and SMC, this calls us to think about transdisciplinary pedagogical models to support development in the field.

---

[8] Narrative and physical experiences that engage an audience member in a fictional world, for instance immersive VR theatre production.

References

1. Aguerreche, L., Duval, T., Lécuyer, A.: Comparison of Three Interactive Techniques for Collaborative Manipulation of Objects in Virtual Reality in CGI (Computer Graphics International) (Singapore, 2010).
2. Aguerreche, L., Duval, T., Lécuyer, A.: Reconfigurable Tangible Devices for 3D Virtual Object Manipulation by Single or Multiple Users. Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology, 227–230 (2010).
3. Akpan, I., Marshall, P., Bird, J., Harrison, D.: Exploring the effects of space and place on engagement with an interactive installation in ACM Conference on Human Factors in Computing Systems (CHI) (ACM Press, New York, New York, USA, Apr. 2013), 2213.
4. Andersson, N., Erkut, C., Serafin, S.: Immersive Audio Programming in a Virtual Reality Sandbox English. in Audio Engineering Society Conference:2019 AES International Conference on Immersive and Interactive Audio (Audio Engineering Society, Mar. 2019).
5. Argelaguet, F., Andujar, C.: A survey of 3D object selection techniques for virtual environments. Computers and Graphics (Pergamon) **37**, 121–136 (2013).
6. Atherton, J., Wang, G.: Doing vs. Being: A philosophy of design for artful VR. Journal of New Music Research **49**, 35–59 (2020).
7. Bailenson, J.N.: Nonverbal Overload:ATheoretical Argument for the Causes of Zoom Fatigue. Technology, Mind, and Behavior 2 (Feb. 23, 2021).
8. Baron, N.: CollaborativeConstraint : UI for Collaborative 3D Manipulation Operations in IEEE Symposium on 3D User Interfaces (2016), 269–270.
9. Barrass, S., Barrass, T.: Musical creativity in collaborative virtual environments. Virtual Reality **10**, 149–157 (2006).
10. Benyon, D., Höök, K., Nigay, L.: Spaces of Interaction in Proceedings of the 2010 ACM-BCS Visions of Computer Science Conference (BCS Learning & Development Ltd., Swindon, GBR, Apr. 2010), 1–7.
11. Berthaut, F.: 3D interaction techniques for musical expression. Journal of New Music Research **49**, 60–72 (2020).
12. Berthaut, F., Desainte-Catherine, M., Hachet, M.: DRILE: an immersive environment for hierarchical live-looping in Proceedings of the International Conference on New Interfaces for Musical Expression (NIME) (2010), 192–197.
13. Berthaut, F., Hachety, M., Desainte-Catherine, M.: Piivert: Percussion-based interaction for immersive virtual environments. IEEE Symposium on 3D User Interfaces (3DUI), 15–18 (2010).
14. Berthaut, F., Martinez, D., Hachet, M.: Reflets : Combining and Revealing Spaces for Musical Performances. Proceedings of the International Conference on New Interfaces for Musical Expression, 116–120 (2015).
15. Birnbaum, D., Fiebrink, R., Malloch, J.,Wanderley, M. M.: Towards a dimension space for musical devices in Proceedings of the International Conference on New Interfaces for Musical Expression (NIME) (2005), 192–195.
16. Bowman, D. a. et al.: New Directions in 3D User Interfaces. International Journal **5**, 3–14 (2006).
17. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qualitative research in psychology **3**, 77–101 (2006).
18. Burgess,N.: Spatial memory: how egocentric and allocentric combine. Trends in Cognitive Sciences **10**, 551–557 (2006).
19. Cabral, M. et al.: Crosscale: A 3D virtual musical instrument interface in 2015 IEEE Symposium on 3D User Interfaces (3DUI) (Mar. 2015), 199–200.
20. Çamcı, A., Hamilton, R.: Audio-first VR: New perspectives on musical experiences in virtual environments. Journal of New Music Research **49**, 1–7 (2020).
21. Chénéchal, M. L., Lacoche, J.: When the Giant meets the Ant An Asymmetric Approach for Collaborative and Concurrent Object Manipulation in a Multi-Scale Environment. IEEE Symposium on 3D User Interfaces, 277–278 (2016).

22. Ching, F. D. K.: Architecture: Form, Space, & Order Fourth edition (Wiley, Hoboken, New Jersey, 2015).
23. Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., Looser, J.: Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. International Journal of Human-Computer Studies **69**, 401–414 (2011).
24. Colquhoun, A.: Typology and Design Method. Perspecta, 71–74 (1969).
25. Deacon, T.: Shaping Sounds in Space: Exploring the Design of Collaborative Virtual Reality Audio Production Tools PhD thesis (Queen Mary University of London, 2020).
26. Deacon, T., Bryan-Kinns, N., Healey, P. G., Barthet, M.: Shaping sounds: The role of gesture in collaborative spatial music composition in Creativity and Cognition (ACM, San Diego, 2019), 121–132.
27. Deacon, T., Stockman, T., Barthet, M. in Bridging People and Sound: 12th International Symposium, CMMR 2016, Sáo Paulo, Brazil, July 5–8, 2016, Revised Selected Papers (eds Aramaki, M., Kronland-Martinet, R., Ystad, S.) vol 10525, 192–216 (Springer International Publishing, Cham, 2017).
28. Duignan, M., Noble, J., Biddle, R.: Abstraction and Activity in Computer- Mediated Music Production. Computer Music Journal **34**, 22–33 (2010).
29. Durupinar, F., Kapadia, M., Deutsch, S., Neff, M., Badler, N. I.: PERFORM: Perceptual Approach for Adding OCEAN Personality to Human Motion Using Laban Movement Analysis. ACM Transactions on Graphics **36**, 6:1–6:16 (Oct. 2016).
30. Ens, B. et al.: Revisiting collaboration through mixed reality: The evolution of groupware. Computer Supported Cooperative Work **131**, 81–98 (2019).
31. Fillwalk, J.: ChromaChord : A Virtual Musical Instrument in 2015 IEEE Symposium on 3D User Interfaces, 3DUI 2015 - Proceedings (2015), 201–202.
32. Gardair, C., Healey, P. G. T., Welton, M.: Performing places. Proceedings of the 8th ACM conference on Creativity and cognition - C&C '11, 51 (2011).
33. Gaver, W. W.: What Should We Expect From Research Through Design? In ACM Conference on Human Factors in Computing Systems (CHI) (2012), 937–946.
34. Gelineck, S., Böttcher, N., Martinussen, L., Serafin, S.: Virtual Reality Instruments capable of changing Dimensions in Real-time in Enactive (2005).
35. Geronazzo, M. et al.: The Impact of an Accurate Vertical Localization with HRTFs on Short Explorations of ImmersiveVirtual Reality Scenarios in 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (Oct. 2018), 90–97.
36. Green, T. R. G., Petre, M.: Usability Analysis of Visual Programming Environments: A 'Cognitive Dimensions' Framework. Journal of Visual Languages and Computing **7**, 131–174 (1996).
37. Greenberg, S., Marquardt, N., Ballendat, T., Diaz-Marino, R., Wang, M.: Proxemic Interactions: The New Ubicomp? Interactions **18**, 42–50 (Jan. 2011).
38. Grønbæk, J. E. et al.: Designing for Children's Collective Music Making: How Spatial Orientation and Configuration Matter in Nordic Conference on Human-Computer Interaction (NordiCHI) (2016), 23–27.
39. Hattwick, I., Wanderley, M. M.: A Dimension Space for Evaluating Collaborative Musical Performance Systems in Proceedings of the International Conference on New Interfaces for Musical Expression (NIME) (2012), 429–432.
40. Himberg, T., Laroche, J., Bigé, R., Buchkowski, M., Bachrach, A.: Coordinated Interpersonal Behaviour in Collective Dance Improvisation: The Aesthetics of Kinaesthetic Togetherness. en. Behavioral Sciences 8, 23 (Feb. 2018).
41. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., Greenhalgh, C.: Object focused interaction in collaborative virtual environments. ACM Transactions on Computer-Human Interaction **7**, 477–509 (2000).
42. Hix, D., Gabbard, J. L. in Handbook of Virtual Environments chap. 28 (2014).
43. Honigman, C.: The Third Room : A 3D Virtual Music Paradigm in Proceedings of the International Conference onNewInterfaces for Musical Expression (NIME) (2011).

44. Hornecker, E., Buur, J.: Getting a grip on tangible interaction in ACM Conference on Human Factors in Computing Systems (CHI) (2006), 437.

45. Hornecker, E., Marshall, P., Rogers, Y.: From Entry and Access - How Shareability Comes About in Designing pleasurable products and interfaces (2007).

46. Houde, S., Hill, C.: What do prototypes prototype? Handbook of Human Computer Interaction, 1–16 (1997).

47. Hutchings, D. R., Stasko, J.: Revisiting Display Space Management: Understanding Current Practice to Inform next-Generation Design in Proceedings of Graphics Interface 2004 (Canadian Human-Computer Communications Society, Waterloo, CAN, May 2004), 127–134.

48. Innocenti, E. D. et al.: Mobile Virtual Reality for Musical Genre Learning in Primary Education. en. Computers & Education **139**, 102–117 (Oct. 2019).

49. Jacob, R. et al.: Reality-based interaction: a framework for post-WIMP interfaces in ACM Conference on Human Factors in Computing Systems (CHI) (2008), 201–210.

50. Jin, Y., Lee, S.: Designing in virtual reality: a comparison of problem-solving styles between desktop and VR environments. Digital Creativity 6268 (2019).

51. Jung, B., Hwang, J., Lee, S., Kim, G. J., Kim, H.: Incorporating Co-Presence in Distributed-Virtual Music Environment in Proceedings of theACMSymposium onVirtualReality Software and Technology (Association for Computing Machinery, New York, NY, USA, Oct. 2000), 206–211.

52. Jung, J. et al.: A Review on Interaction Techniques in Virtual Environments. Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management, 1582–1590 (2014).

53. Kolesnichenko, A., McVeigh-Schultz, J., Isbister, K.: Understanding Emerging Design Practices for Avatar Systems in the Commercial Social VR Ecology in Proceedings of the 2019 on Designing Interactive Systems Conference (Association for Computing Machinery, New York, NY, USA, June 2019), 241–252.

54. Kolkmeier, J., Vroon, J., Heylen, D.: Interacting with virtual agents in shared space: Single and joint effects of gaze and proxemics. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **10011 LNAI**, 1–14 (2016).

55. Lages, W.: Ray, Camera, Action ! A Technique for Collaborative 3D Manipulation. IEEE Symposium on 3D User Interfaces, 277–278 (2016).

56. Lages, W., Nabiyouni, M., Tibau, J., Bowman, D. A.: Interval Player: Designing a virtual musical instrument using in-air gestures in 2015 IEEE Symposium on 3D User Interfaces, 3DUI 2015 - Proceedings (2015), 203–204.

57. Le Groux, S., Manzolli, J., Verschure, P. F. J.: VR-RoBoser: Real-Time Adaptive Sonification of Virtual Environments Based on Avatar Behavior in Proceedings of the International Conference on New Interfaces for Musical Expression (NIME) (2007).

58. Lilija, K., Pohl, H., Hornbæk, K.: Manipulation Who Put That There? Temporal Navigation of Spatial Recordings by Direct Manipulation in CHI Conference on Human Factors in Computing Systems (Association for Computing Machinery, 2020).

59. Lubos, P., Bruder, G., Ariza, O., Steinicke, F.: Touching the Sphere: Leveraging Joint-Centered Kinespheres for Spatial User Interaction. Proceedings of the ACM Symposium on Spatial User Interaction (SUI'16), 13–22 (2016).

60. Lugasi, M., Rafaely, B.: Speech Enhancement Using Masking for Binaural Reproduction of Ambisonics Signals. IEEE/ACM Transactions on Audio, Speech, and Language Processing **28**, 1767–1777 (2020).

61. Mäki-patola, T., Laitinen, J., Kanerva, A., Takala, T.: Experiments with virtual reality instruments in Proceedings of the International Conference on New Interfaces for Musical Expression (NIME) (2005), 11–16.

62. Melchior, F., Pike, C., Brooks, M., Grace, S.: Sound Source Control in Spatial Audio Systems in Audio Engineering Society Convention (Rome, Italy, 2013).

63. Men, L., Bryan-Kinns, N.: LeMo: Supporting Collaborative Music Making in Virtual Reality. IEEE 4TH VR Workshop SIVE (2018).

64. Men, L., Bryan-Kinns, N.: LeMo: Exploring virtual space for collaborative creativity in Creativity and Cognition (ACM, San Diego, USA, June 2019), 71–82.

65. Men, L., Bryan-Kinns, N., Bryce, L.: Designing spaces to support collaborative creativity in shared virtual environments. PeerJ Computer Science **5**, e229 (2019).

66. Moore, A. G., Howell, M. J., Stiles, A. W., Herrera, N. S., McMahan, R. P.: Wedge: A musical interface for building and playing composition-appropriate immersive environments in 2015 IEEE Symposium on 3D User Interfaces (3DUI) (Mar. 2015), 205–206.

67. Mulder, A., Fels, S. S., Mase, K.: Mapping virtual object manipulation to sound variation. IPSJ Sig Notes **97**, 63–68 (1997).

68. Mulder, A., Fels, S. S., Mase, K.: Design of Virtual 3D Instruments for Musical Interaction. Graphics Interface, 76–83 (1999).

69. Naef, M., Collicott, D.: A VR Interface for Collaborative 3D Audio Performance in Proceedings of the International Conference on New Interfaces for Musical Expression (NIME) (2006).

70. O'Modhrain, S.: A Framework for the Evaluation of Digital Musical Instruments. en. Computer Music Journal **35**, 28–42 (Mar. 2011).

71. Palumbo, M., Zonta, A.,Wakefield, G.: Modular reality: Analogues of patching in immersive space. Journal of New Music Research **49**, 8–23 (2020).

72. Plowright, P. D.: Revealing Architectural Design: Methods, Frameworks and Tools (Routledge, 2014).

73. Poirier-Quinot, D., Katz, B.: Assessing the Impact of Head-Related Transfer Function Individualization on Task Performance: Case of a Virtual Reality Shooter Game. en. Journal of the Audio Engineering Society **68**, 248–260 (May 2020).

74. Poupyrev, I., Billinghurst, M., Weghorst, S., Ichikawa, T.: The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR. Proc. UIST '96 (ACM Symposium on User Interface Software and Technology), 79–80 (1996).

75. River, J., MacTavish, T.: Research through Provocation: A Structured Prototyping Tool Using Interaction Attributes of Time, Space and Information. The Design Journal **20**, S3996–S4008 (July 2017).

76. Rogers, Y., Lim, Y.-k., Hazlewood,W. R., Marshall, P.: Equal Opportunities: Do Shareable Interfaces Promote More Group Participation Than Single User Displays? Human-Computer Interaction **24**, 79–116 (2009).

77. Serafin, S., Erkut, C.,Kojs, J., Nilsson,N. C.,Nordahl, R.: VirtualReality Musical Instruments: State of the Art, Design Principles, and Future Directions. Computer Music Journal **40**, 22–40 (2016).

78. El-shimy, D., Cooperstock, J. R.: User-Driven Techniques for the Design and Evaluation of New Musical Interfaces. Computer Music journal **39**, 28–46 (2015).

79. Smith, H. J., Neff, M.: Communication Behavior in Embodied Virtual Reality in ACM Conference on Human Factors in Computing Systems (CHI) (2018).

80. Snook, K. et al.: Concordia: A musical XR instrument for playing the solar system. Journal of New Music Research **49**, 88–103 (2020).

81. Stowell, D., Robertson, A., Bryan-Kinns, N., Plumbley, M. D.: Evaluation of live human-computer music-making: Quantitative and qualitative approaches. International Journal of Human-Computer Studies **67**, 960–975 (2009).

82. Suchman, L., Trigg, R., Blomberg, J.: Working artefacts: ethnomethods of the prototype. The British Journal of Sociology **53**, 163–179 (2002).

83. Svanæs, D.: Interaction Design for and with the Lived Body : Some Implications of Merleau-Ponty ' s Phenomenology. ACM Transactions on Computer-Human Interaction **20**, 1–30 (2013).

84. Thoring, K., Desmet, P., Badke-Schaub, P.: Creative Environments for Design Education and Practice: A Typology of Creative Spaces. en. Design Studies **56**, 54–83 (May 2018).

85. Trommer, M.: Points Further North: An acoustemological cartography of non-place. Journal of New Music Research **49**, 73–87 (2020).

86. Trumbo, J.: The Spatial Environment in Multimedia Design: Physical, Conceptual, Perceptual, and Behavioral Aspects of Design Space. Design Issues **13**, 19–28 (1997).
87. Tuan, Y.-F.: Space and Place: The Perspective of Experience. **4**, 513 (1978).
88. Valbom, L., Marcos, A.: Wave: Sound and music in an immersive environment. Computers & Graphics **29**, 871–881 (2005).
89. Vanderdonckt, J.: Distributed user interfaces: how to distribute user interface elements across users, platforms, and environments. Proc. of XI Interacción, 20–32 (2010).
90. Wakefield, G., Smith, W.: Cosm : a Toolkit for Composing Immersive Audio- Visual Worlds of Agency and Autonomy in Proceedings of the International Computer Music Conference (ICMC) (2011).
91. Wanderley, M. M., Orio, N.: Evaluation of Input Devices for Musical Expression : Borrowing Tools from HCI. Computer Music Journal **26**, 62–76 (2002).
92. Weinel, J. in Technology, Design and the Arts-Opportunities and Challenges 209–227 (Springer, Cham, 2020).
93. Won, A. S., Bailenson, J. N., Lanier, J. in Emerging Trends in the Social and Behavioral Sciences 1–16 (2015).
94. Wong, N., Gutwin, C.: Where are you pointing? Proceedings of the 28th international conference on Human factors in computing systems - CHI '10, 1029 (2010).
95. Wong, N., Gutwin, C.: Support for Deictic Pointing in CVEs : Still Fragmented after All These years? in Computer Supported Cooperative Work (2014), 1377–1387.
96. Xambó, A., Laney, R., Dobbyn, C., Jordá, S. P.: Multi-touch interaction principles for collaborative real-time music activities: towards a pattern language. Proc. of ICMC'11, 403–406 (2011).
97. Xambó, A. et al.: Exploring Social Interaction With a Tangible Music Interface. Interacting with Computers 28 (2016).
98. Young, G., Murphy, D.: HCI Models for Digital Musical Instruments: Methodologies for Rigorous Testing of Digital Musical Instruments. International Symposium on Computer Music Multidisciplinary Research (CMMR) (2015).
99. Zhou, F., Dun, H. B. L., Billinghurst, M.: Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. Proceedings - 7th IEEE International Symposium on Mixed and Augmented Reality 2008, ISMAR 2008, 193–202 (2008).
100. Zielasko, D. et al.: Cirque des Bouteilles : The Art of Blowing on Bottles in 2015 IEEE Symposium on 3D User Interfaces, 3DUI 2015 - Proceedings (2015), 209–210.

Products and Grey Literature

101. Arrigo, A., Lemke, A.: Wave http://wavexr.com/. Austin, TX, USA, 2016.
102. Beat Saber - VR Rhythm Game Beat Games. 2019.
103. Block Rocking Beats http://blockrockingbeats.com/. 2016.
104. DearVR Spatial Connect https://www.dearvr.com/products/dearvrspatial-connect. Düsseldorf, Germany, 2018.
105. Designing For Virtual Reality en. https://www.ustwo.com/blog/designing-for-virtual-reality/. 2015.
106. Drops https://drops.garden/. 2018.
107. Gravity Sketch https://www.gravitysketch.com/. Aug. 2017.
108. Kane, A.: Volta https://volta-xr.com/. 2021.
109. Kane, A.: Volta https://www.voltaaudio.com. London, UK, 2019.
110. Kinstner, Z.: EXA: The Infinite Instrument https://store.steampowered.com/app/606920/EXA_The_Infinite_Instrument/. Grand Rapids, Michigan, USA, 2017.
111. Lee, J., Strangeloop: The Lune Rouge Experience The WaveVR. 2017.

112. LyraVR http://lyravr.com/. 2018.
113. Mux https://store.steampowered.com/app/673970/MuX/ http://playmux.com/. 2017.
114. Olson, L., Havok, R., Ozil, G., Fish, R.: Soundstage VR https://github.com/googlearchive/soundstagevr. 2017.
115. Spatial Audio API https://www.highfidelity.com/. 2021.
116. The Garden https://www.biomecollective.com/the-garden. Dundee, UK, 2019.
117. The Last Maestro https://www.maestrogames.com/. 2021.
118. The Music Room http://www.musicroomvr.com/. 2016.
119. Tranzient https://www.aliveintech.com. 2019.
120. Virtual Reality Best Practices en-US. https://docs.unrealengine.com/en-US/SharingAndReleasing/XRDevelopment/VR/DevelopVR/ContentSetup/index.html.
121. VR Best Practice https://learn.unity.com/tutorial/vr-bestpractice. 2017.
122. VR Design : Best Practices en-US. http://blog.dsky.co/2015/07/30/vr-design-best-practices/. July 2015.