IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 17, 2020, accepted January 5, 2021, date of publication January 19, 2021, date of current version January 28, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3052931

Music in Extended Realities

LUCA TURCHET[1], (Member, IEEE), ROB HAMILTON[2], AND ANIL ÇAMCI[3]
[1]Department of Information Engineering and Computer Science, University of Trento, 38122 Trento, Italy
[2]Department of the Arts, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[3]School of Music, Theatre and Dance, University of Michigan, Ann Arbor, MI 48109, USA

Corresponding author: Luca Turchet (luca.turchet@unitn.it)

ABSTRACT The intersection between music and Extended Reality (XR) has grown significantly over the past twenty years, amounting to an established area of research today. The use of XR technologies represents a fundamental paradigm shift for various musical contexts as they disrupt traditional notions of musical interaction by enabling performers and audiences to interact musically with virtual objects, agents, and environments. This article both surveys and expands upon the knowledge accumulated in existing research in this area to build a foundation for future works that bring together Music and XR. To this end, we created a freely available dataset of 260 publications in this space and conducted an in-depth analysis covering 199 works in the last decade. We conducted this analysis using a list of conceptual dimensions belonging to technical, artistic, perceptual and methodological domains. This review of the literature is complemented with a set of interviews with domain experts with the goal of establishing a definition for the emergent field of Musical XR, i.e., the field of music in Extended Realities. Based on the results of the conducted review, a research agenda for the field is proposed.

INDEX TERMS Music, extended reality, virtual reality, augmented reality, augmented virtuality, mixed reality, digital musical instruments, Internet of Musical Things.

I. INTRODUCTION

Throughout history, instrument makers and musicians have explored technologies specific to their eras to create musical instruments and perform for audiences. The introduction of digital technologies has provided musicians, designers, software developers and engineers with access to previously unimaginable ways to create, control and disseminate music. Today's Extended Reality (XR) technologies represent a paradigm shift in musical interaction possibilities, allowing artists as well as audiences to take an active role in shaping musical processes presented in virtual and immersive contexts. Virtual objects, agents and worlds, created, controlled and defined through musical interaction schema, act as the bridge between traditional analog and physical musical realities and those that exist purely as rendered constructs within a digital music landscape.

Over the past few decades, artists and researchers have increasingly devoted their attention to the use of XR for musical applications and experiences. The rapidly developing field defined here as *Music in Extended Realities (Musical XR)*, currently constitutes an established area of research that brings together practitioners from a wide range of fields including music composition, game studies, computer science, engineering, visual arts and cognitive science.

This paper seeks to identify the field of Musical XR and its unique creative and expressive affordances. To document the artistic practices and technological research that have shaped the development of this field, we present a comprehensive and systematic review of published research focusing on Musical XR as documented in over 260 publications. This review focuses on a broad range of Musical XR research, encompassing technical, artistic, perceptual and methodological domains. To better understand the broad spectrum of academic studies and artistic projects that are often labeled as Virtual Reality (VR), Augmented Reality (AR), Augmented Virtuality (AV), Mixed Reality (MR) and Extended Reality (XR), this paper investigates current trends and conceptualizations of how the term "Musical XR" has been and could be defined. To this end, a set of interviews featuring academic researchers, industry practitioners, artists and engineers, each deeply immersed in musical interaction and XR research, are presented.

In Section II, we provide a broad overview of Musical XR as it relates to modern practices and considerations in this domain. In Section III, we discuss the evolution of XR technology since the 1960s, leading up to modern XR platforms. In Section IV, we survey existing perspectives on

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaogang Jin.

how to distinguish between different types of extended realities. This conversation facilitates our definition of Musical XR in Section V, where we offer both audio-centric and multi-sensory perspectives on what constitutes Musical XR. This section also summarizes the feedback we gathered from domain experts on what in their opinions comprises Musical XR. In Section VI, we provide examples from prior surveys in this field and specify the contribution of the current article. In Section VII, we detail the methodology we employed to compile and analyse a database of publications that deal with Musical XR. Then in Section VIII, we offer and discuss the results of this analysis. Finally in Section IX, we propose a research agenda for the field of Musical XR based on our analysis of the publication database.

II. MUSICAL XR: AN OVERVIEW

Significant academic and industrial research has been carried out in the design and development of immersive and augmented visual systems, with the goal of rendering novel ''virtual'' visual environments or ''augmenting'' our visual fields by placing rendered virtual constructs within views of real space. At the same time, a significant body of work has investigated the creation of rendered auditory scenes that are capable of articulating space and motion for independent sound sources with high fidelity. While non-visual and non-aural XR technologies, such as those that leverage virtual touch, smell and taste [147], [151], have also been studied, the primary trends in hardware and software engineering for XR systems, as well as the primary commercialized tools available to the public for experiencing XR content, have focused on the pairing of rendered visual and auditory stimuli. As such, our definition of extended reality and the concentration of the research presented in this paper, largely focus on audio-visual systems and the content created for such systems.

The use of XR in artistic applications bridges sensory modalities and carries with it any number of active and passive affordances that shape the experience of a given work. In a musical definition of XR, auditory considerations – both fundamentally creative or musical, and those that bear spatial and semantic information – are blended with extra-aural sensory data. Among multimodal XR experiences, there are those that privilege music and sound, either conceptually or functionally. We have previously referred to these as ''Audio-First'' experiences [34]. These kinds of audio-centric or audio-first XR works closely align with the definitions of Musical XR that this paper explores. At the same time, a continuum of theoretical and practical considerations in XR research needs to be taken into account in order to fully unpack what makes a VR, AR, AV, or MR experience both ''Musical'' and ''XR''.

For many early-21st century consumers, researchers and practitioners, any determination as to whether an experience is an example of VR, AR, AV, or MR seems to be based on a loose aggregate of factors, with a strong bias towards visual modes of presentation. Even if we accept the abbreviation ''XR'' itself as an amalgamation of VR, AR, AV, and MR – a viewpoint most definitely not shared by all – the role of music, sound and interactivity within such systems still is cause for discussion and spirited debate.

What this term ''Musical XR'' might represent to different communities of technical researchers, scholars and artists is by no means a foregone conclusion. The criteria that separate our descriptions of ''virtual'' reality environments from those considered ''mixed'' and ''augmented'' are indefinite and ill-defined at best. The roles of sound and music within an XR experience can range widely from passive background elements to fully interactive and controllable phenomena, where sound acts as a fundamental driver of, is driven by, or exists completely independently of visual stimuli. Rendered components within XR can be presented using stereoscopic head-mounted displays (HMDs), within room-scale CAVE-style projection systems [46], on mobile devices or, especially when sound is a fundamental driver of the experience, with no visual presentation at all. Interaction layers that tie together rendered and real-world visual and aural content currently employ a number of controllers and tracking methods including but not limited to n-degree of freedom head, hand and skeleton tracking, depth and plane detection, multi-input hand-held controllers, and GPS-based mobile location tracking.

III. THE EVOLUTION OF XR

Although it is only within the past decade that affordable and powerful VR systems have been commercially available to non-specialist users, research into VR technology has taken many forms since Morton Heilig's patent describing a head-mounted display in 1960 [72] and Ivan Sutherland's first prototype of this kind of technology in 1968 [140]. Myrion Krueger's 1977 project Videoplace made use of computer vision to bring users into a virtual environment [85]. Throughout the 1980s several research labs, including VPL Research and NASA Ames Research Center, have created successful implementations of head-mounted displays [58]. In 1992, the Electronic Visualization Lab at the University of Chicago introduced the first CAVE system, which is a room-scale stereoscopic VR environment based on surround full-wall projections and motion tracking [46]. Between then and the early 2000s, game companies such as Sega (with their unreleased Sega VR headset) and Nintendo (with the Virtual Boy and the Power Glove) have attempted to commercialize VR technology albeit unsuccessfully [23], [155]. In 2010, Microsoft announced its Kinect depth-tracking camera system, a relatively accessible technology that opened up many advanced applications wherein users can be mapped into a virtual environment with point-cloud and skeletal-tracking information [168]. The following year, Oculus announced its work on a modern HMD, with the first development kit made available to the public in 2012, and new hardware and software releases continuing through the publication of this paper [51]. Since then, numerous affordable and robust VR hardware systems have been introduced, ushering in the most commercially successful era of VR technology to date [65], [78], [105].

Augmented Reality systems have also found commercial success and ubiquity within the context of mobile computing, whereby on-device cameras, accelerometers and depth sensing technologies have made the insertion of rendered visual elements into dynamic views of real-world environments. At the same time, recent advances in commercially available AR headsets, such as the Microsoft Hololens and Magic Leap 1, have enabled new enterprise, consumer and academic applications, expanding the user base of wearable AR technology.

Another frontier of XR involves web technologies, which aim to bring AR and VR capabilities to the web by enabling users to interact with browsers using XR devices [75], [90], [119]. One of the foremost representations of the state of the art in this space is the WebXR Device API,[1] an initiative of the World Wide Web Consortium to standardize access to VR and AR hardware.

IV. VIRTUAL, AUGMENTED AND MIXED REALITIES

Previous efforts have been made to define those elements that are fundamental to our ability to perceive rendered environments and constructs as respectively "immersive" and "real". Rosenblum and Cross put forth "Immersion", "Interaction" and "Visual Realism" as necessary elements for successful and convincing virtual reality systems [1]. In 1994, in a discussion of taxonomies for visual display technologies for AR and VR, Milgram and Kishino proposed their Virtuality Continuum stretching between purely "Virtual" environments and those that were entirely "Real". As shown in Figure 1, the liminal space between these polar ends encompasses Augmented Reality and Augmented Virtuality, as well as a blending of elements from both real and virtual environments, referred to as "Mixed Reality" [99]. Benford *et al.* attributed successful immersive environments and users' perceived sense of belonging to such environments as functions of a two-dimensional plane plotting levels of Artificiality (ranging from the Physical to the Synthetic) and Transportation (from the Local to the Remote), as seen in Figure 2 [10].

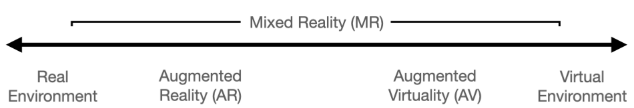


FIGURE 1. The Milgram-Kishino Virtuality Continuum (adapted from [99]).

As presented in this paper, XR can be considered an umbrella term that encompasses systems and experiences commonly referred to as Virtual Reality [131], Augmented Reality [6], [7], Augmented Virtuality (wherein elements of reality are themselves inserted into virtual environments) [97], and Mixed Reality [108]. VR technologies completely

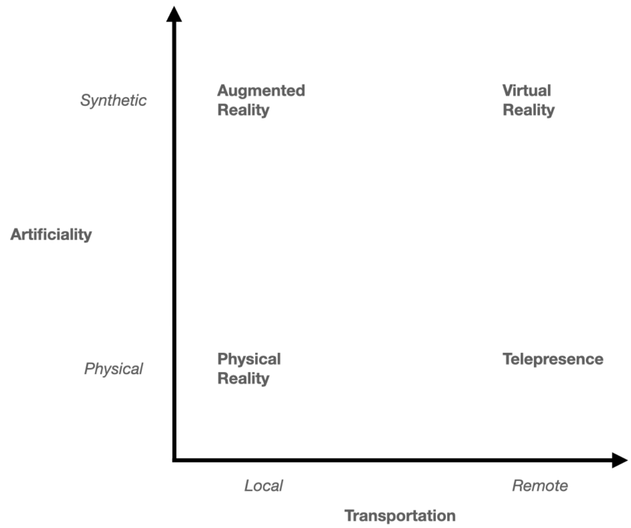


FIGURE 2. The Benford *et al.* Spectrum (adapted from [10]).

immerse users within a computer-generated virtual environment (VE), which can be navigated and interacted with in ways that mirror our own human interactions with the real world. VR systems often rely on the use of an HMD, occluding any visual stimuli from the outside world and instead presenting a stereoscopically rendered visual space. In CAVE systems, although the user's vision is not directly occluded, they are encapsulated in a room-scale environment with stereoscopic visuals displayed or projected on the surrounding walls [46]. By contrast, AR systems allow users to see the real world, but with virtual objects superimposed upon and within it. This is often achieved by means of hand-held devices (e.g., phone, tablets) or wearables (e.g., smart glasses) that overlay digital information onto the user's field of view. Therefore, AR tends to supplement rather than supplant our interaction with reality.

Less generally discussed than AR or VR is Milgram and Kishino's region of Augmented Virtuality. Both AR and AV environments blend elements from real and virtually-rendered environments, with the primary distinctions revolving around the function of individual elements and the type of environment within which these are placed. AV environments insert "real" actors or objects within a virtual scene, inverting the relationship between real and virtual objects observed in AR. Following the taxonomy proposed by Milgram and Kishino, both AR and AV environments can be considered Mixed Reality. According to the authors, MR can be conceived as a display paradigm in which real and virtual objects are presented together at the same time, and is therefore denoted on their Virtuality Continuum as any position between "Real" and "Virtual" poles [99]. This way, MR combines elements from reality, AR, AV, and VR, blending aspects of both the real and virtual world. Speicher *et al.* offer an overview of modern interpretations of MR, highlighting views that align with or diverge from Milgram and Kishino's conceptualization [138]. As such, the defini-

[1] https://www.w3.org/TR/webxr/

tion of MR appears to be in an ongoing evolution, much like the modern technologies that underlie this concept.

As helpful as these continua and schema are in articulating the similarities and differences between virtual, augmented and mixed reality environments, each suffer from an exclusive focus on visual stimuli as criteria for establishing immersive rendered realities. As the modern field of XR continues to grow, it will be necessary to construct models wherein non-visual stimuli are not only accounted for but also regarded as fundamental aspects of the user's sense of presence and immersion [47], [133], [136], [161]. To accommodate not only sound, but touch, taste and smell – essentially every human sense for which a virtual analogue does or might exist – such models would need to treat both visual and non-visual stimuli as being capable of supporting immersive and interactive user engagement. [2]

V. DEFINING MUSIC IN EXTENDED REALITIES

By extending current conceptualizations of VR, AR, AV, and MR systems and experiences to be more inclusive of non-visual sensory modalities, and by focusing on sound and music as primary drivers of the rendered realities themselves, we can first arrive at a rather broad definition of Music in Extended Realities, or *Musical XR*, wherein virtual and real elements of multimodal nature can be intermixed.

A. AN AUDIO-FIRST APPROACH

The roles of sound and music within Musical XR systems can vary widely, ranging from diegetic to non-diegetic uses of sound and music, from fixed audio playback to interactively triggered or procedurally generated sound events, and from head-locked stereo sound presentations to dynamic spatial audio implementations based on head or location tracking. As such, when considering an Audio-First approach, we differentiate XR experiences where music and sound are fundamental drivers of a user's experience from those in which music and sound play relatively simple, auxiliary or passive roles. Musical XR projects can take a diverse range of forms, including virtual musical instruments, immersive concert experiences, generative musical systems and gamified musical environments. This is why defining what constitutes an Audio-First and, moreover, a Musical XR experience is more easily achieved by identifying a continuum across several criteria, wherein sound and music are situated as essential modalities of information in multisensory creative work.

In Musical XR projects, the manner in which sound and music elements are controlled through manipulations of their properties (e.g., pitch, amplitude, spatial, and timbral characteristics) must be in someway connected to the user and their behavior within the environment. Static and non-responsive audio playback simply placed in an XR environment does not generally qualify a project to be considered Musical XR.

[2]While Azuma states this explicitly in their definition of AR systems, the presented survey focuses almost entirely on visual AR systems [7].

The role of sound and music in such projects should be fundamental to the functionality or aesthetics of the XR experience. Simply put, in a Musical XR system, if the audio stream was to be disabled, valuable information pertaining to this system would be lost, resulting in a fundamentally degraded user experience. For instance, games with background music or simple sonic indicators of user interactions generally would not be considered Musical XR experiences, whereas games that are governed by musical or sonic structures, or those that are focused on the performance or control of such structures, generally would.

To establish the illusion of immersion and self-presence in a virtual space, the sonic elements of a Musical XR system should convey spatial information that either augments the physical space with spatial consistency or suggests an alternative space through dynamic spatialization cues such as reverb, delay and filtering. Simple mono or stereo audio presentations fail to convey spatial detail and as such would not strongly suggest a work to be classified as Musical XR.

One of the most defining qualities of Musical XR is the extent to which the sound material is presented and perceived to be within what we consider a "Musical" or organizational framework. Composer Edgard Varèse famously described music itself as "organized sound", differentiating content that simply exists as audible waves from that which was purposefully curated to exist within a specific framework [149]. The mere existence of sound within an XR environment is not sufficient to consider it an example of a Musical XR. Instead, sonic elements must reveal structural and organizational properties of the system.

B. SENSORY MODALITIES

As we previously described, central to the nature of XR systems and experiences is the integration of stimuli being presented by more than one sensory modality at the same time. Not only does the combination of diverse sensory inputs – each representative of some aspect of the non-real environment or experience itself – contribute to the senses of immersion and presence in XR, but each modality presents a unique set of affordances with which designers and developers can craft meaningful and engaging user experiences. We briefly describe relatively common sensory modalities explored in Musical XR projects as follows:

1) VISUAL

Due in no small part to the physiological prominence human perception places on visual stimuli, research into XR has historically and overwhelmingly been focused on visual rendering and presentation technologies. Advances in computer graphics and systems that support stereoscopic rendering of virtual scenes have fueled academic, corporate and consumer interest in VR and AR technologies, often at the risk of ignoring other sensory modalities. The majority of XR projects, even those that do privilege audio and musical content, have significant visual components using head-tracked stereoscopic HMD or CAVE-like projection systems, or rendered objects placed into AR scenes using head-mounted displays, retinal projection or non-stereoscopic computer or mobile-device screens.

2) AUDITORY

As previously described, audio content within Musical XR projects should ideally focus the user engagement significantly on the role of perceptually organized music and sound, as befitting the description of an Audio-First experience. Musical XR systems often act as tools for creating music, either performative or compositional in nature, or use musical content as a conceptual driver for the experience as a whole. Of primary importance within a Musical XR experience is not only the incorporation of sound and music, but also their spatialization, correspondence to elements of other sensory modalities, and their interactive functions.

3) HAPTIC

Although the virtuality of the elements in an XR application is at odds with the sense of touch, haptics are nevertheless fundamental to how we establish a sense of presence and agency in any given environment. The hand-held controllers found in modern XR systems incorporate haptics into user interactions, albeit in a limited fashion. These interactions involve the manipulation of virtual objects or avatars via touch or force-based input mechanisms and the vibro-tactile feedback that the virtual environment provides through such mechanisms as actuators and motors. Touch being the foremost modality that links a musician to their instrument, haptic interactions are a fundamental yet under-explored aspect of how we engage with Musical XR systems.

4) PROPRIOCEPTIVE

Musical expression is tightly intertwined with movement. We perform and react to music by moving our bodies. Senses of position and movement are therefore intrinsic to our relationship with Musical XR systems not just in room-scale experiences that leverage physical locomotion, but also in stationary experiences where the user might orient themselves physically to perceive the virtual environment surrounding them. In XR systems, motion and positional trackers included in HMDs and hand-held controllers are employed to integrate these senses into how the user explores and engages with the virtual environment.

5) SMELL AND TASTE

Although prior research has been carried out on the use of virtualized smell and taste in XR systems [151], these sensory modalities are not currently in widespread use in XR contexts outside of specific research environments. As such, they have yet to become significant components of existing Musical XR systems.

C. EXPERT INTERVIEWS

In an effort to achieve a comprehensive definition of Musical XR that could be generally agreed upon by the Musical XR community, we interviewed 8 experts who have made extensive contributions to the field in both academia and industry. Interviews were conducted via email by asking the following two questions: *How would you define XR as it relates to musical applications?* and *What do you think are the technical, conceptual or modal requirements for a system or experience to be called Musical XR?*

Most of the experts approached these questions from the viewpoint of sensory modalities and interactivity. While none of the participants posited a particular sensory modality as an absolute requirement, several participants highlighted a need for Musical XR to address as many modalities as possible. According to Berthaut, an XR musical instrument displays its virtual components through one or more sensory channels. Zappi mentions that visual output, even though not strictly necessary, can be extremely powerful in establishing a sense of presence in Musical XR. Furthermore, he argues that an extended space that completely lacks any materiality places greater emphasis on haptics to bring out the nuances of musical expression in XR. Indeed, Kuchera-Morin identifies the inability to use all of our senses as a shortcoming of computational systems in general. In that respect, she sees Musical XR as a means to extend computational systems into a multi-sensory platform through the integration of sounds, visuals, haptics and gestural control. According to Kuchera-Morin, this would bring us closer to modeling natural human interactions with holistic computational systems that can process the flow of complex information involved in these interactions.

The central role of interactivity in musical XR was highlighted by other experts as well. According to McLeran, while user interactions are not mandatory for a VR system to be focused on music, such a system should at least be interactive to the extent that the user can control their viewpoint. In these systems, the user input coming from handheld controllers and head-trackers constitute the only "real" aspect of the user's experience in virtual space.

A common sentiment among the experts was that Musical XR should exploit the affordances of XR technology to offer experiences that go beyond what is possible in the real world. According to Serafin, such possibilities are bound only by our current understanding of multi-sensory perception. Wang argues that a Musical XR system should justify the use of technology by offering an experience that is uniquely suited to the underlying medium. Similarly, Berthaut posits that Musical XR systems should leverage the ability to have components that are virtual or intangible. According to McLeran, while musical AR can be as simple as the simulation of space via reverberation, musical AR and VR systems should strive to approach, and then go beyond, what we experience in real life.

The experts have generally refrained from identifying specific technologies. More explicitly, Ciciliani argued that Musical XR cannot be reduced to a specific set of tools. According to him, a Musical XR is a virtual environment experienced as a performance space wherein the user can suspend their disbelief and substitute the parts or whole of the physical world with an artificial version. Serafin gave an inclusive definition of Musical XR as any experience where real-world musical artifacts – and the perception of such artifacts – are reproduced or extended with the help of technology. According to Cook, such experiences would require a level of computer mediation by either augmenting or replacing the user's sensory space, and tracking their gestures for control purposes.

The role of meaning, defined as the combination of ''doing'' and ''being'', serves as a framing mechanism for Wang's understanding of Musical XR, in so far as any experience within a Musical XR should allow for meaningful musical interactions and should similarly foster a ''meaningful sense of belonging and existing''. Going even further, Cook postulates that music itself – through the transformative ability of organized sound to extend one's reality – is in fact an extended reality.

D. SO, WHAT IS MUSICAL XR?

Based on our review of the literature, interactions with experts and our own practices in this domain, we define Musical XR through prerequisites in four main areas:

1) EXISTENCE OF VIRTUAL ELEMENTS

These elements can be presented in one or more sensory modalities. They can function as an intangible augmentation of a physical space or object, or they can be presented fully virtually. These elements can include agents and objects situated in an environment, or the environment itself.

2) SPATIAL PERSISTENCE

The virtual elements that make up a Musical XR system and the user share a persistent 3-dimensional space. Although this environment can be virtually navigated, the system maintains a constant spatial relationship between the virtual elements and the user.

3) INTERACTIVITY

A Musical XR system should, at the very least, be interactive to the extent that it will respond to the user's position and head orientation, and render the virtual space accordingly. We refer to this kind of interactivity as passive participation, where the user observes the virtual space but does not exert control over it. An active participation system not only takes into account the user's spatial position and orientation, but also presents elements that can be manipulated by the user.

4) SONIC ORGANIZATION

A Musical XR system should situate sonic organization as a fundamental aspect of its conceptual and/or technical implementation. For instance, such organization may govern the structure of the virtual environment or the unfolding of the events in it. The ways in which the user interacts with this environment can be informed by or designed to facilitate musical expression. The role of auditory elements in conveying meaning, intent and focus in a Musical XR system is leveraged to the extent that the organization of these elements becomes an inextricable component of the user experience.

Additionally, there are preferable, if not mandatory, qualities of Musical XR systems: such systems should address as many senses as possible, nearing the experience to full immersion. When possible, these systems should implement room-scale experiences to address the proprioceptive sense. If a virtual element in a Musical XR system addresses multiple senses, its representations in different modalities should be mapped to each other in space as closely as possible. To better establish a sense of presence and agency, these systems should offer as many affordances for interaction as possible.

VI. PRIOR SURVEYS IN MUSICAL XR

To the best of our knowledge, a comprehensive and systematic review of the Musical XR field has yet to be undertaken. Such review of peer-reviewed studies in this complex domain could reveal developing trends in XR research and, therefore, support the development of design guidelines and best practices for musical applications that leverage XR. Moreover, providing a clear and complete overview of the field is useful for identifying not only existing trends and established methods, but also future opportunities for research and artistic projects.

Previous studies have surveyed selected aspects of this broad field. Serafin *et al.* conducted a review of virtual reality musical instruments (VRMIs) up to 2016 [130]. Loveridge offered a survey of the emerging field of networked music performances in VR through 2019 [89]. Berthaut reviewed existing 3D interaction techniques and examined how they can be used for musical control [14]. Atherton and Wang devised an overview of recent musical works that employ the VR medium [4], while Çamcı and Hamilton identified research trends in this field through a set of workshops that focus on Audio-first VR [34].

This paper can be distinguished from such previous surveys for the following reasons: i) it aims to capture a broad picture of the Musical XR field, accounting for technological, artistic, perceptual and methodological perspectives; ii) it covers diverse music-related application domains (e.g., performance, education, composition); iii) it offers an in depth-analysis of Musical XR studies, which are classified into several categories based on their main features; iv) from this analysis, it identifies open research questions and proposes a research agenda.

This endeavour may help novices gain a basic understanding of Musical XR research. Moreover, practitioners in this field may find this paper useful for conceiving new research directions. Ultimately, this paper seeks to bridge existing research areas and communities in Musical XR and foster interdisciplinary collaborations that engage with the challenges and opportunities in this field.

VII. METHODOLOGY

The review was guided by the following research questions:

- What are the common technological, artistic, and perceptual considerations for practitioners in the field of Musical XR?
- What is the distribution of works across various domains of application in this field (e.g., performance, education) and who are the stakeholders (e.g., performers, students)?
- Which tools and methods do the authors employ to build Musical XR systems?
- How do the authors conceptualize, design and evaluate their systems?
- Which artistic forms or practices have been favored by the Musical XR community?
- Are there existing blind spots in the field and, if so, can these inform future directions?

To address such questions, we created a publication database that encompasses works related to Musical XR going back to the 1990s. We then performed a quantitative analysis of the papers that were published between 2011 and 2020, in recognition of the technological advances and the widespread availability of tools in the last decade that helped establish the modern field of XR.

A. CREATION OF THE PUBLICATION DATABASE

We started our study by collecting an extensive pool of scientific and artistic research publications. Specifically, this phase focused on identifying and selecting relevant papers that could address our research questions. We used IEEE Xplore Digital Library, ACM Digital Library, Elsevier Scopus, Google Scholar, and Web of Science in order to search for relevant contributions. From all the retrieved contributions, we adopted the following inclusion and exclusion criteria in order to select the studies of interest for our work: i) we removed studies that were unrelated to the specified research questions (i.e., for each search result, the title or the abstract of the article had to foreshadow theoretical or practical aspects related to our investigation); ii) we only considered the contributions that were published in peer-reviewed journals, conferences, and book chapters, thus eliminating Bachelor's, Master's, or PhD theses, patents, and technical reports); iii) we removed non-English papers and duplicate papers (e.g. those that appeared in different databases or in duplicate languages).

Such criteria led to a total of 260 works that were included in the publication database, which was finalized in January 2021 and, therefore, encompasses articles published in 2020 and before. The database is freely accessible online.[3]

The next sections detail our clustering approach and the results of the analysis applied to 199 entries in the database spanning from 2011 through 2020. In the following discussions, we also refer to some non-peer reviewed publications that nevertheless represent formative works in the field, such as influential artistic projects.

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

B. CLUSTERING

We clustered the retrieved publications under the following categories and subcategories:

- **Year:** the year of publication;
- **XR medium:** the XR medium which the described study is most strongly based on (i.e., AR, AV, or VR);
- **Study type:** the nature of the study (i.e., theory-based or application-based);
- **Primary function:** the main purpose of the Musical XR system (i.e., performance, education, composition, sound engineering, entertainment, perception study, development);
- **Target user:** the intended user of the Musical XR system (i.e, performer, student, composer, audience member, studio engineer, developer);
- **Social experience:** whether a Musical XR system offers an individual or multi-user experience;
- **Connectivity:** whether the Musical XR system functions in a standalone mode or within a network of systems.

Notably, we acknowledge that some studies have multiple functions and users, but for the purpose of clustering across the primary function and target user dimensions, we considered only the most prominent application of a given project. For instance, some systems such as VRMIs may be devised to play music, but they can also be utilized for compositional purposes. Along the same lines, such systems may primarily target the performer, whereas certain performance contexts may allow these systems to be used by the audience.

VIII. RESULTS AND DISCUSSION

This section reports the analysis of the publications in the database and their distribution across the aforementioned clusters with a focus on publications from 2011 to 2020.

A. HISTORICAL DISTRIBUTION

In Figure 3, we show a distribution of Musical XR projects over the years. The earliest entry in the database is Krueger's 1977 paper, which describes PSYCHIC SPACE; this is a musical system that places the user in a virtual space and involves interactivity, visuals, computer vision, and computationally generated sounds [85]. However, this work can be considered an outlier, as it is only from the 1990s that studies at the intersection of music and XR started to gradually appear in the literature, likely as a result of the growing interest in computer music among researchers who had institutional access to specialized equipment.

The distribution of database entries is relatively sparse until the 2000s, then shows a slow growth in the number of publications until the beginning of the 2010s, followed by a rapid increase since then. Indeed, the last decade – and particularly the last few years – has seen a dramatic increase

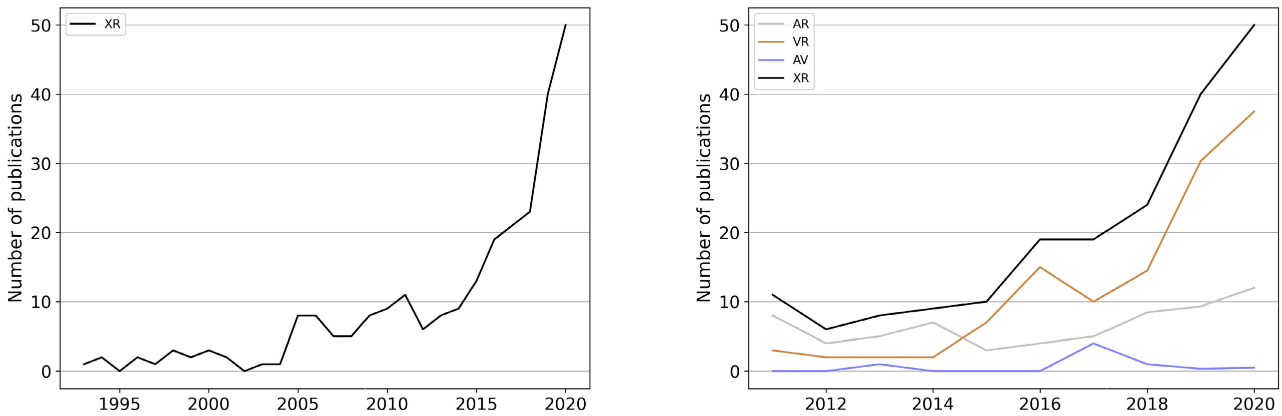


FIGURE 3. Historical distribution of Musical XR publications according to the year of publication. Left: years 1993–2020; right: years 2011–2020 with trend lines for individual platforms. An aggregate of all studies are shown with the XR trend line.

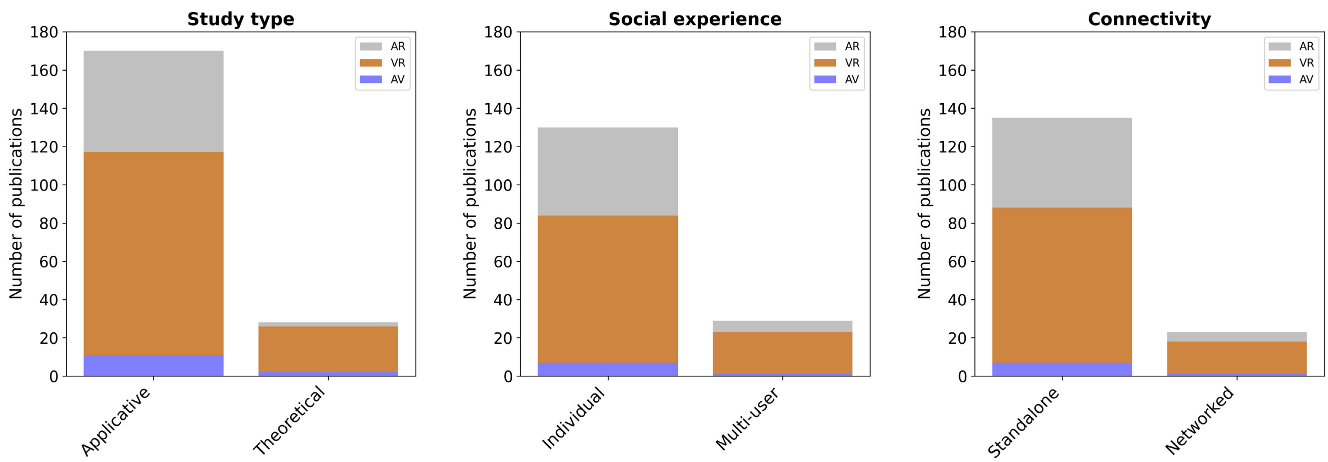


FIGURE 4. Distribution of Musical XR publications according to the study type (left), social experience (middle), and connectivity (right), by XR medium (years 2011–2020).

in Musical XR research, likely due to the greater availability and affordability of XR hardware and software tools.

As illustrated in Figures 3 (right) and 4 (left), our analysis identified that the last decade has seen a significant increase in research that focuses on VR than AR, while only a handful of studies have investigated AV (see e.g., [52], [82]). It can also be seen that, from 2015 onward, there has been a dramatic increase in the number of studies dealing with VR.

B. STUDY TYPE

The vast majority of studies on Musical XR have an applicative nature, as seen in Figure 4 (left). Nevertheless, various theoretical studies have also been carried out (e.g., [3], [24], [126]). These range from design principles and evaluation frameworks, to position papers, perspectives, and philosophical considerations.

1) APPLICATIVE STUDIES

Applicative studies can be grouped into three primary categories: those that describe tools for developing Musical XR systems, those that describe Musical XR systems themselves across various domains of application (which may involve a user evaluation), and documentations of artistic projects.

Taken as a whole, Musical XR systems represented in the database are multifaceted and address a range of primary functions and target users as identified in Section VII-B. Examples for performance are reported in [36], [71], [143], for education in [8], [49], [56], [74], for composition in [37], [45], [64], [102], and for sound engineering in [11], [83]. Other XR systems have focused on entertainment [79], [128], [166], including music visualization [55], [156] and music-related cultural heritage applications [60], [114], [142]. Such systems include VRMIs [26], [57], [91], [137], and interfaces for audience members [96]. Notably, research on VRMIs has focused on both reimagining conventional instruments in VR [67], [91], [159], [169] or imagining entirely new instruments [26], [57], [102], [137].

Works that document specific Musical XR compositions and performances as well as artistic applications of the developed systems are numerous. Examples include [4], [43], [69],

[76], [86], [109]. However, only a minority of such works report an evaluation that focuses on the experience of the performers or audiences.

2) THEORETICAL STUDIES

In recent years, a number of researchers have formulated guidelines for Musical XR practitioners and perspectives on the field as a whole (see e.g., the seminal work reported in [117]). These works can be grouped under the following subcategories:

a: DESIGN FRAMEWORKS

In 2014, Berthaut *et al.* proposed a framework to classify and design VRMI stage configurations within the context of performance scenography [20]. The framework is based on 6 dimensions:

1) musician immersion: the immersion of the musician(s) in the VE;
2) audience visibility: the level of perception of the audience by the musician(s);
3) audience immersion: how well the audience perceives the VE and the instrument;
4) musician visibility: the perception of the musician(s) by the audience;
5) gestural continuity: how the musical gestures performed by musicians in physical space are connected to the graphic feedback of the instrument's metaphor, as perceived by the audience;
6) from virtual to physical: how the virtual and physical spaces are merged.

In 2016, Serafin *et al.* proposed a set of nine principles for designing VRMIs [130]:

1) design for feedback and mapping: design should focus on sound, visual, touch and proprioception in tandem, as well as consider the mappings between these sensory modalities;
2) reduce latency: limit any delay between the player's gestures and the generated multisensory feedback;
3) prevent cybersickness: avoid "wrong" mappings (e.g., between vision and proprioception) that can cause the user to experience motion sickness;
4) make use of existing skills: do not copy but rather leverage expert playing techniques derived from interactions existing in the real world;
5) consider both natural and "magical" interaction: reflect on the use of interactions that conform to real-world constraints as well as interactions not limited by them (e.g., the laws of physics or human anatomy);
6) consider display ergonomics: reflect on the potential strain and discomfort introduced by wearing VR devices;
7) create a sense of presence: devise methods that allow the player to experience the sensation of "being there" in a computer-generated environment;
8) represent the player's body: consider the representation of a player's body in the virtual world by tracking and mapping the real body to a virtual representation;
9) make the experience social: create social shared experiences while the instrument is being played.

In 2017, Johnson and Tzanetakis proposed a series of design guidelines for VR-based musical pedagogy systems [82].

Guidelines for performance feedback:

1) visual feedback shouldn't prevent a user from focusing on aural feedback;
2) reduce cognitive load by limiting the amount of concurrent feedback;
3) provide terminal performance feedback upon completion of the practice task.

Guidelines for VE design:

1) visibility of system status;
2) choose metaphor(s) that naturally match the application task space;
3) match between system and the real world;
4) create a sense of presence;
5) consider display ergonomics;
6) consider controller ergonomics;
7) represent the player's body;
8) ensure that users' avatars provide a familiar, accurate, and relevant frame of reference;
9) allow users to alter point of view, or viewpoint.

Guidelines for interaction in VR pedagogy:

1) limit nonessential interaction during practice;
2) recognition rather than recall;
3) make use of existing skill.

In 2019, Turchet proposed a class of smart musical instruments and a set of nine principles for their design [145], envisioning the new subcategory of "Smart VRMIs" that are characterized by embedded intelligence (i.e., featuring context-awareness and proactivity) and connectivity within the Internet and local networks:

1) embed as much as possible;
2) design for interconnection;
3) make it upgradeable and updatable;
4) make it ready to use;
5) make it personalizable;
6) make it smart, but maintain the musician's sense of control;
7) add signature features and make it beautiful;
8) bear in mind the cognitive load and the learning curve;
9) address the security and privacy issues.

In 2020, Atherton and Wang described a practical philosophy for creating artful designs for musical experiences in VR based on a duality between doing (i.e., interacting) vs. being (i.e., observing) [4]. The proposed philosophy consists of a set of lenses and eighteen principles to achieve good design in VR, covering both the craft and the humanistic dimensions of the medium:

1) audio should be dynamically generated;
2) audio should be immersive;

3) audio should be interactive;
4) don't port: make things that would be impossible in the physical world;
5) design to balance doing (action) and being (reflection);
6) look up! use gaze to modulate between doing and being;
7) drive interaction design with aesthetics;
8) multimodality is a virtue;
9) make space for being alongside doing in interaction;
10) create worlds that enhance doing and being through animus;
11) balance stylization and realism;
12) design for virtual embodiment;
13) the body is an implicit medium where being supports doing;
14) movement matters;
15) play is both an activity and a state: a synthesis of doing and being;
16) replicate baseline social interactions; redesign the rest;
17) support many kinds of social engagement;
18) design for social doing and social being.

b: EVALUATION FRAMEWORKS

These design frameworks can be utilized to evaluate Musical XR works. However, some authors have explicitly devised evaluation frameworks for this purpose as well.

Serafin *et al.* proposed a three-layered evaluation framework for VRMIs based on the set of nine design principles previously described [130]. The first layer concerns interaction modalities, including considerations on the alignment between input and output modalities, as well as the perceptual integration and mapping based on perceptual, sensorimotor, and cognitive abilities and capacities of users. The second layer is a VR-specific layer that accounts for cybersickness, virtual body representation and ownership, and presence. The third layer aims to evaluate the goals, practices, and experiences of VRMI players.

Mazzanti *et al.* offered a framework for the description and evaluation of participatory live music systems, including those for Musical XR. The framework, which they apply to their Augmented Stage system [96], involves the following six metrics:

1) control design freedom: how freely audience interaction can be designed with the platform;
2) system versatility: overall performance setting up simplicity and performer's comfort on stage;
3) audience interaction transparency: clearness of the relation between audience manipulation and its effects;
4) audience interaction distribution: to what extent interaction can be located towards the participants (strongly centralized interface vs. every participant holds one);
5) focus: how easily the audience can freely focus on different performance aspects (the stage, their interaction, visuals, music, etc.);
6) active/passive audience affinity: how much the non-interacting and interacting audience experience can be similar.

For the evaluation of audience experiences in Musical XR, Pirchner proposed the Interactive Real-time Measurement of Attention, a tool for gauging audience engagement during performances of audiovisual computer music in situated contexts [115]. This is a mobile application that allows audience members to report which aspects of a performance they are mostly engaged with in real time on an "Attention Triangle" which spans across Music/Sound, Visuals, and Performer.

c: POSITION STATEMENTS

Various authors have offered critical perspectives on Musical XR practices. For instance, Hamilton highlights the mediating role of rule systems in virtual spaces as having a key influence on interactive musical expressions, casting these mediation layers as intrinsic components and unique attributes of any and all Musical XR works [68].

Berthaut raises an important point in [14], stating that existing 3D interaction techniques are not designed with musical interaction in mind, but are rather geared towards minimizing task completions times. Based on his review of existing techniques, the author proposes a number of directions to improve the usability of such techniques in musical applications. These include the investigation of multiple object selection/manipulation schemes, hybrid/articulated approaches, music-centric guidance and disambiguation, re-ambiguation of existing techniques, dedicated input devices and the design of visual feedback targeted at the audience and performers.

Reflecting on his own artistic involvement with VR, Ciciliani highlights areas of significance in the design of musical systems for this medium. These include the representation of performers in the virtual space while they are simultaneously present in the real space, the conscious composition of space and movement, the use of polyspatiality, the use of the 3D environment as both instrument and score, and the mapping of performers' interactions in the virtual environment [43].

In [31], Çamcı presents a number of considerations that can inform the design and implementation of new creativity support tools for spatial audio in VR applications. The author argues that such tools should enable both novice and expert users to make creative use of spatial sound and that such an endeavor would necessitate the formulation of novel user interaction schemes. Such schemes should combine common user interface elements found in existing spatial audio design tools with new ones that are informed by the inherent properties of spatial sound and the affordances of new VR platforms.

C. PRIMARY FUNCTION

Figure 5 (left) illustrates the distribution of the reviewed applicative works across the primary function dimension with indicators of the specific XR medium. As shown here, performance is the most common primary function, followed by education. Less attention in this space seems to have been devoted to composition, entertainment and sound

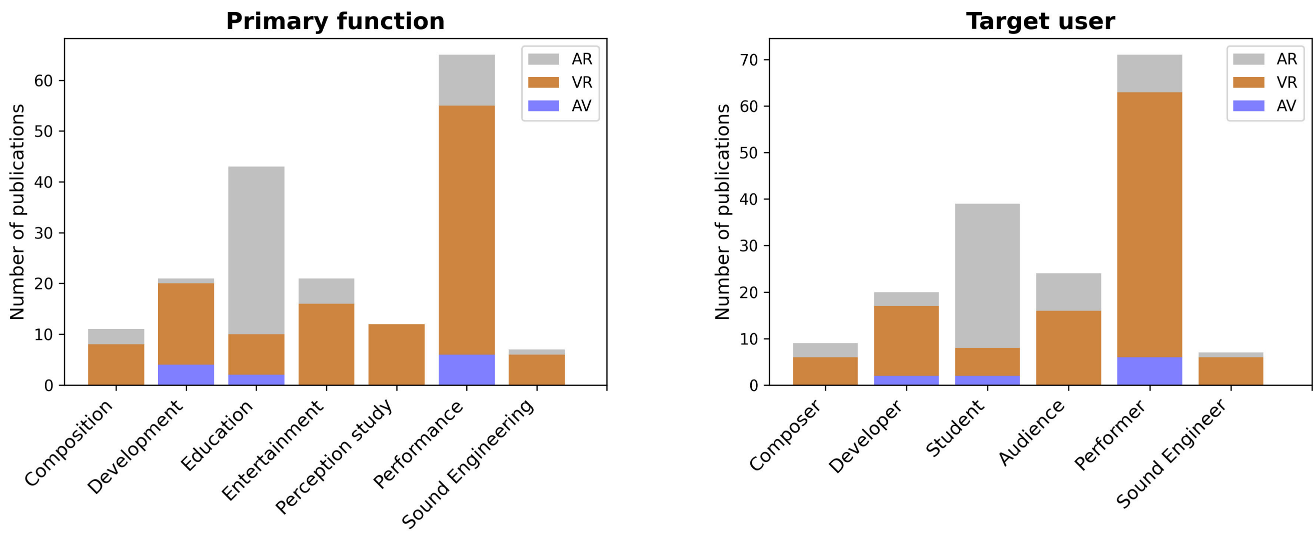


FIGURE 5. Distribution of Musical XR publications according to the primary function (left) and user (right), by XR medium (years 2011–2020).

engineering applications, whereas a considerable percentage of applications deal with perception and development. Overall, the majority of research in this area has been focused on VR, followed by AR, while only a handful of works have focused on AV. Whereas VR seems to have a prominent role in music performance and perception-related studies, AR finds a primary role in educational applications. The rest of this section surveys the most representative works in the various functional dimensions identified.

1) COMPOSITION

Although a limited number of Musical XR systems directly target composers, the analysis suggests that XR-based tools can be effective in supporting compositional processes as evidenced in [24], [43]. The authors of [64] provide discussions based on projects exploring the embellishment, augmentation, and extension of environmental cues, spatial mapping, and immersive potential of scalable multichannel audio systems for VR and AR. The authors also present a system enabling users to treat virtual soundfields as a fundamental building block for spatial music composition and sound design.

In a different vein, Masu *et al.* propose the concept of VR Open Scores, which are score-based virtual scenarios where an aleatoric score is embedded in a VE [94]. The aims underlying this study are to create a VE where users can immerse themselves in the visual elements of a score while listening to the corresponding music, and to help users develop a personal relationship with both the system and the score. Along a similar line, the study reported in [9] describes a system where musical notation is presented in AR to musicians wearing an HMD. In another project, Santini describes a system for creating AR gesture-based scores in the context of experimental instrumental composition [126].

Composers looking to present musical scores to performers in XR paradigms must often create new tools as well. In Coretet [67], pre-composed MIDI scores are dynamically rendered as finger-placement markers along the neck of virtual stringed instruments, providing real-time visual guidance for performers.

Throughout history, musical instruments have evolved from one stage to another not only thanks to technological progress, but also thanks to the requests for novel features by performers and composers. In addition to their uses in performance situations, VRMIs also find use in music composition. These instruments often lack a tradition of practice and a historical repertoire, and as such have not yet become firmly established as accessible instruments for a majority of performers and composers. At the same time, performance and composition practices involving VRMIs are fundamental to the longevity of these tools [150]. Furthermore, the role of improvisational and collaborative composition is highlighted in a number of Musical XR projects. For instance, [37] explores the idea of real-time music composition in VR through a network shared among audience members and performers. Such examples indicate the participatory potential of XR in facilitating novel composition practices.

2) DEVELOPMENT

Researchers have devised both hardware and software tools to support the development of Musical XR systems. A notable example of one such software tool is Chunity [2], which embeds the ChucK programming language [154] into the Unity development environment. Another example is OSC-XR [81], a toolkit that aims to simplify the process of designing immersive music environments by providing pre-built controllers for the Open Sound Control protocol as well as Unity scripts for designing custom ones. On the other hand, Palumbo *et al.* describe an environment for visual programming within VR that supports live, fine-grained, multi-user collaboration, through a framework for operational transformations based on graph structures [110].

Various hardware tools have also been conceived for Musical XR applications [14]. One such tool called Piivert, is an input device for manipulating virtual elements; it combines 6 degrees-of-freedom tracking and pressure detection, allowing several new techniques for 3D musical interaction [17]. [11] reports a comparison between physical controllers and hand/gesture tracking for audio mixing in VR. This study reveals that physical controllers are preferred and generally rated as more preferable than gesture-tracking controls in terms of perceived accuracy, efficiency, and satisfaction.

A growing number of projects employ emerging XR technologies for the web in Musical XR contexts (see e.g., [30], [113]). It is plausible to expect that more attention to be devoted in the upcoming years to Musical XR systems based on web technologies amounting to a sub-field of practices in Musical XR. This will entail effective integration of visual and audio technologies for the web, such as the WebXR API and the Web Audio API.

3) EDUCATION

Education is viewed among the foremost areas to be influenced by VR technology [152]. This is true for music education as well. A number of projects have already demonstrated the use of XR in musical training. The authors of [22] investigated the use of VR exposure training to reduce music performance anxiety. The results of the study suggest that music institutions can benefit from installing virtual systems, which musicians can utilize to train themselves within a variety of concert scenarios in order to improve their anxiety management skills. Along the same lines, [128] reports a prototype and a preliminary evaluation of a VR experience that allows children to sing in a virtual choir with the aim of preventing or helping to cope with social anxiety. These projects demonstrate the effectiveness of XR technology in providing performers with a surrogate for being physically on stage, supporting the findings of other studies conducted in non-musical domains (see e.g., [135]). Other VR systems have been proposed as pedagogical tools to train primary and secondary education students on music-related concepts such as rhythm and acoustics [129] and genre identification [49].

In the context of AR, different paradigms and technologies have been devised. One of the aims motivating such systems is that of enabling students to visually monitor their practice and have fun while doing so. One of the first examples of AR systems for musical training is reported in [104], which was devised to support playing the guitar. The system uses a camera to track the movement of a player's fingers in order to detect chords. Chord names are displayed on a computer screen superimposed on the camera image to allow the performer to check which chord they are playing. Similar approaches are reported in [50] and [48].

Several studies have explored the projection of virtual elements onto physical musical instruments to support education. The authors of [163] report a system where music is visualized as animated characters walking on a physical keyboard. The system takes advantage of walking as one of the most fundamental human rhythms to promote musical understanding. A similar system was then adapted for children [162]. The piano has been the instrument mostly investigated using such a paradigm [21], [41], [120], [122], [124], [158], [162], [163], [165], while similar systems exist for guitar [88] and drums [164].

Other studies have focused on the use of HMDs to deliver virtual elements of visualizations supporting music learning superimposed onto real instruments. Examples can be found for piano [27], [42], [63], [66], guitar [144], and drums [106]. In a different vein, the VRmin is an instrument resulting from the augmentation of a real theremin with visualization displayed in VR [82].

Notably, whereas most of the studies conceived for education target adult learners (typically for self-training), a growing number of works are being devised for children [5], [49], [87], [128], [162]. For instance, Farinazzo *et al.* proposed MUSIC-AR, a set of software that uses AR for teaching children sound properties, such as timbre, pitch and sound intensity [56].

The analysis of the publication database reveals several challenges in the use of XR tools in music education. As described in [129], the issues pertaining to accurate and expressive gesture-based control of musical sound in VR make current musical interfaces in this medium more useful for entertainment rather than effective pedagogical applications. As a consequence, for skill learning with conventional musical instruments, the use of AR in conjunction with an actual instrument appears to be a more valuable option than using VR at the moment. This approach allows for the preservation of the ergonomics of the original instrument while at the same time allowing additional visual feedback to assist the performer (see e.g., [88], [163]). Nevertheless, as pointed to by the authors of [129], VR may be helpful in various types of musical training applications, including remote rehearsals, addressing stage fear through exposure training, and developing rhythmical skills.

As highlighted by Çamcı and Hamilton, there is a need for more accessible tools capable of facilitating novel learning experiences for inexperienced users who wish to utilize immersive audio as a medium for artistic expression [34]. Current VR authoring models, in which the design and review stages are often separated, can hinder the learning process [32]. As a consequence, there is a need for of new creativity support tools, including those for VR audio design, that enable content creation within VR, effectively integrating the learning, design and exploration processes.

Notably, while the review identified a number of XR systems conceived for students and dedicated to self-training, there were no systems designed specifically for educators except for one study that addressed teacher-student interactions [27]. Moreover, existing studies have focused on novices, whereas limited effort has been made towards developing systems for expert and intermediate-level musicians.

4) ENTERTAINMENT

While entertainment systems, and more specifically VR games, constitute a major technological and conceptual contributor to the commercial growth of the VR industry, it should be noted that peer-reviewed academic publications from the creators of such systems are not common. Successful game titles such as Rock Band VR[4] and Beat Saber,[5] both of which feature rhythm-based score-driven gameplay, are popular examples of Musical XR systems for entertainment that are not directly represented in academic literature, and therefore under-represented in this database. Research detailing immersive systems for the presentation and enjoyment of musical performances, as well as the exploration of the therapeutic benefits of ensemble singing include [128] and [121]. Abstract, dynamic and interactive immersive visual XR presentations that accompany music can be seen in [153], [156] and [35], while a faithful AR representation of members of a string quartet performing can be found in [143]. An additional survey of a number of music-first VR entertainment experiences is detailed in [4], including both works from industry as well as academia.

5) PERCEPTION

One major challenge for practitioners of Musical XR is the capability of a system to allow users to experience a strong sense of presence, a trait also common to non-musical XR systems [47], [133], [134]. A similar consideration applies to the sense of social presence [25], [107], which is a key component of live musical performance.

Presence is an important aspect of networked musical interaction, which appears to be a growing consideration in the Musical XR literature. However, in the context of VR-based networked music performances, only a limited number of studies have attempted to evaluate the sense of presence experienced by musicians [89]. Results of a study reported in [127] suggested that the use of VE for collaborative music making among networked musicians could enhance the feeling of being together. It is worth noting that the proliferation of Musical XR-based networked interactions partly relies on the development of computer network technologies that can address issues of bandwidth and latency [39], [80], [148]. And particularly in AV and VR systems, the sense of presence can be severely hampered by simulation sickness, a sensation theorized to be resulting from conflicting information from the vestibular and visual senses, leading to nausea, disorientation, headaches, sweating and eye strain [40].

6) PERFORMANCE

Many XR systems have been developed for live music contexts, targeting audiences, individual performers and ensembles. One of the earliest and most referenced musical performances in VR is *The Sound of One Hand* by Jaron Lanier in 1993 [86]. The composer performed live on a plain stage, wearing a single instrumented glove and an HMD. While immersed in a VE, Lanier could play different types of VRMIs while his viewpoint was projected onto screens for the audience. Other examples of immersive systems designed for individual performers are reported in [15] and [16], where 3D graphical elements enable efficient and simultaneous control and visualization of musical processes. Systems devised for multiple performers have been devised for both AR and VR. For details see Section VIII-E2.

Other XR systems have targeted both the performers and the audience. Berthaut *et al.* developed Reflets, an AR environment that allows both performers and audience members to display virtual content on stage, such as 3D virtual musical interfaces or visual augmentations of instruments and musicians [19]. Santini developed a system that enables audiences to see the view of a pianist wearing a AR-based HMD thanks to projectors [125]. Other examples include live music performance ecosystems such as [96], where audience members use their smartphones to superimpose virtual contents on the real stage.

The VR performance work Carillon takes place in an game-like environment designed as a functioning futuristic carillon tower wherein networked participants collaboratively interact with a singular virtual instrument [71]. Kaneko *et al.* report a preliminary investigation on an XR-based system devised for remote audiences of live music concerts [84]. The system attempts to enable remote audiences to exchange nonverbal communication with body actions between them in a VR environment.

Ensemble performance within XR spaces has been explored with a number of projects focused on smaller sized ensembles. In 1992, Dolby presented a virtual string quartet featuring pre-recorded audio performances and simple performer avatars which could be viewed by single audience members wearing an HMD [54]. Audio from the four-channel performance was spatialized depending on audience member position and head-rotation within the virtual space. Designed in 2018, Coretet is a VRMI and networked performance space that presents a fully-interactive family of bowed and plucked string instruments such as the violin, viola, cello and doublebass for real-time solo or ensemble performance [67]. Up to four co-located performers can share a single virtual space while seeing each other's instrument and avatar. Audio is generated procedurally using physically-modelled strings from the Synthesis Toolkit [44]. Performances of fully composed ensemble works [69] are presented to audiences using projected views into the virtual space via an array of virtual cameras. An AR string quartet is also presented in [12] in which individual performers playing physical instruments were filmed and presented in an augmented mobile environment for research and entertainment purposes.

Musical performance situations within Musical XR systems encounter their own set of unique complications, fundamentally based around the disparity of experience between that as presented to performers of XR instruments or systems and that as presented to the viewing audience [20].

[4]https://www.rockbandvr.com/

[5]https://beatsaber.com/

For graphically rich and immersive XR performance environments, performers are often equipped with stereoscopic HMDs and controllers capable of displaying depth and interacting with objects with multiple degrees of freedom. In sharp contrast, due in part to the lack of ubiquity in XR technologies as well as the cost and complication required to outfit a typical audience with the necessary equipment, most audiences in such performance situations are presented with a two-dimensional projection of a rendered space, via projector and screen or large-format display. As such any semblance of depth and scale as experienced within the rendered XR space is lost and the experience itself fundamentally differs between performer and audience.

Sonic expressions of space and depth have long been a staple of electroacoustic and electronic musical performance, using multi-channel speaker arrays driven by spatial algorithms such as ambisonics [92], vector-based amplitude panning [118] or wave field synthesis [13] to immerse audiences and performers alike in a rich and nuanced auditory space. Motion and action in such spaces can be heard and understood, though hyper-specific and subtle motions might not be immediately understandable. The sonification of motion and gesture is itself a mature field within music and audio engineering [73], requiring a combination of artistic sensibility and technological rigor. As such the performative aspects of sonic actions and actors within extended environments can and has been explored within musical concert settings.

7) SOUND ENGINEERING

XR offers many opportunities for the mixing and spatialization of sound sources in innovative ways that are not possible with conventional hardware and software tools. In particular, VR offers a unique method of sound source visualization, allowing sound engineers to directly control sound sources in an immersive context. While there are commercial products, such as DearVR[6] that exploit some of these opportunities, only a small number explore sound engineering in AR (e.g., [38]) and VR (a review on 3D audio production tools is available at [95]).

The study presented in [83] reports a system for sound engineering in VR, where predefined sound sources could be attached to objects in the VE and manipulated with the Oculus Touch controllers. Along the same lines, the study reported in [11] investigated the design and evaluation of alternative audio mixing interfaces based on a VE and associated controllers. The study showed the value and potential of using VR to visualize and control sound sources in an articulated and convincing digital environment suitable for audio mixing tasks. On the other hand, the study reported in [59] investigated the use of VR within the context of the 3D Stage Paradigm interfaces, showing that VR can be as efficient as conventional desktop applications, but a preference for VR was observed. Other studies have explored the use of 3D visualizations for mixing purposes [61], [62].

[6]https://www.dearvr.com

D. TARGET USERS

The results of the target user analysis largely mirror that of the primary function analysis: a majority of studies were targeted to performers and students, followed by audiences and developers (see Figure 5 (right)). Only a handful of studies have been devised specifically for composers and sound engineers. This highlights an area, where the potential of XR can be more extensively utilized. Although there is existing work on participatory [96] and non-participatory [146] audience engagement in Musical XR, there is still room for new Musical XR systems that specifically target audiences.

Only a limited number of Musical XR systems target musicians with specific levels of expertise (e.g., beginners, intermediates, experts). Whereas some systems have been devised for users with no or little musical proficiency [45], most systems are aimed at intermediate levels of musical proficiency rather than virtuosity. It was also found that the vast majority of systems target adults. Nevertheless, a growing number of works cater to children, although mostly within the context of musical education [129], [162].

E. SOCIAL EXPERIENCE

Figure 4 (middle) illustrates the distribution of studies across the social experience dimension. This shows that existing research has primarily focused on the development of Musical XR systems targeting individual experiences (e.g., VR systems for individual musical activities where a single user is immersed in one virtual world). This, however, contrasts with one of the fundamental aspects of music, that is its ability to create shared social experiences. Nevertheless, a recent body of literature shows a growing interest among researchers and practitioners in creating XR technologies that support musical collaborations.

1) INDIVIDUAL EXPERIENCES

Systems for individual use have been created for AR, AV and VR as well as for all types of target users and primary functions identified in Section VII-B. A number of interfaces have been developed for individual composition (for VR see e.g., [102], for AR see e.g., [111]), education (for VR see e.g., [49], [103], for AR see e.g., [101], [124], for AV see e.g., [82]), performance (for VR see e.g., [26], [57], [169], for AV see e.g., [52]), and entertainment (for VR see e.g., [55], [157], for AR see e.g., [53]).

2) MULTI-USER EXPERIENCES

Multi-user experiences in XR can be achieved through either Shared Virtual Environments (SVEs) or independent VEs for each user in settings where the users can be co-located, geographically displaced, or both. Musical XR multi-user systems have been primarily designed for collaborative experiences, but a few systems allow users (e.g., audience members) to experience the XR content independently from each other (e.g., [37], [146]).

Whereas SVEs have been extensively researched for entertainment, education, work, and training purposes, there has been limited research on the creative aspects of collaboration in such environments. That being said, the last few years have seen a growing interest in SVEs for musical collaboration in both AR and VR [29]. A recent system for collaborative music making in VR is LeMo, where two musicians are provided with a shared musical interface based on a step sequencer, with which they can co-create 16-beat music loops [98]. Specifically, LeMo includes two HMDs equipped with Leap Motion sensors for hand tracking, and two PCs synchronized on a local area network. Using a similar approach, the performance project Carillon allows up to three performers to play a shared virtual musical instrument within an SVE [71].

Other works have leveraged musical collaboration for ensemble performances of different virtual musical instruments within an SVE while performers inhabit the same physical space. One prominent example in this category is Coretet [67], previously discussed in VIII-C6. Other systems have focused on collaborative music making among geographically displaced performers [28]. We will discuss such networked XR performances more extensively in Section VIII-F2.

Collaborative musical experiences have also been investigated in the context of AR (see, for instance, [139]). Poupyrev *et al.* proposed the Augmented Groove, an interface for collaborative music making where users who are manipulating physical cards on a table generate musical sounds in conjunction with virtual objects visualized through an HMD [116]. A more recent example is ARLooper [111], where multiple users can interact with each other through an AR mobile interface that supports real-time synchronization of activities such as sound recording, playback, and manipulation.

Another strand of research in this area has focused on the creation and evaluation of XR systems that involve audience participation in the music creation process. In an example involving AR, Zappi *et al.* proposed a multimodal platform for Hybrid Reality live performances, where artists and projected interactive virtual objects were presented on the same physical stage [167]. Audience members wearing 3D glasses could directly modify the scene and its audio-visual features thanks to an RGB/infrared camera system that tracks their gestures. Similarly, Mazzanti *et al.* developed Augmented Stage, an AR environment where virtual objects are superimposed on the performance stage [96]. Using smartphones and tablets, audience members, could interact with virtual objects and control audio-visual aspects of the performance, while observing the performance on stage. The system relied on large-scale posters placed on the stage that acted as AR markers.

However, as far as VR applications are concerned, we could not identify many systems that facilitate the active participation of audience members in the music creation process. In one example, SpectraScore VR [37] offers a set of modular tools developed for use with VR peripherals to allow for music composition, performance and viewing in real-time across the network. Audience members' movements are tracked via Google Cardboard headsets powered by smartphones. Using OSC messages, their movements are transmitted as scores to performers, who in turn improvise music from this notation using VR theremins controlled by Leap Motion controllers. Other recent endeavours focus on the integration of remote audiences within a live music performance ecosystem in VR (see e.g., the preliminary investigation reported in [84]).

As noted in [20], when creating an immersive VR experience for the audience, it is important to acknowledge the challenges that the medium can pose for the spectators. These challenges are mostly related to the issue of visibility of the musicians, as audience members wearing head-mounted displays are unable to share a performer's experience. A preliminary attempt to address these issues using XR technology is reported in [18], where the authors developed a system that exposed to the audience the internal mechanisms of digital musical instruments using 3D visualization.

F. CONNECTIVITY

Figure 4 (right) illustrates the distribution of works along the connectivity dimension, showing that the vast majority of Musical XR systems have been conceived as standalone systems typically targeting a single user. Nevertheless, recent literature indicates a growing research interest in the creation of connected systems that are capable of communicating with a plethora of devices and enabling multi-user interactions.

1) STANDALONE SYSTEMS

The systems designed for individual use listed in Section VIII-E1 are standalone systems. Examples of this category include [163] for AR, [52] for AV, and [91] for VR. No system for individual use was identified which had a distributed nature, although these kinds of networked architectures are envisioned in the endeavours of the Internet of Musical Things community [148].

2) NETWORKED SYSTEMS

Networked musical interactions in XR may happen both in co-located settings (i.e., where users share the same physical environment) and remote ones (i.e., where users are geographically displaced and connected over a communication network) [123], [148].

As far as co-located interactions are concerned, different systems have been proposed (see e.g., [67]). One example is the system reported in [146] where a smart guitar was wirelessly connected to a laptop controlling a VR headset. The system was conceived to enable the audience member equipped with the headset to view virtual objects manipulated by the performer as they interact with the smart guitar sensors.

Regarding remote interactions, we could not identify any study involving AR, while only a handful of studies leveraging VR exist (see e.g., [37] for playing digital instruments together, and [28] for vocal ensembles). The latter are placed in the context of the networked music performance field [100], [123]. As reported in [89], current video-conference systems involved in networked music performance typically are characterized by a higher degree of latency for the video stream compared to systems dedicated to the audio communications. As a consequence, studies have revealed that musicians often ignore the video image of each other in order to maintain a steady rhythm. In this context, Loveridge suggests that VR has the potential to be a suitable visual platform for collaboration in networked music performances as well as a valid alternative to existing video-conference systems used in these contexts.

An example of a Musical XR performance involving both telematic audio as well as networked SVE remote interactions in virtual worlds is reported in [70], where an ensemble of local and networked performers using both traditional and virtual musical instruments played live for an audience surrounded by an ambisonic array of eight speakers and four projection screens, arranged to provide an immersive audio and video experience. Another example is presented in [93], where authors describe a performance system created in the virtual-world platform Second Life by the Avatar Orchestra Metaverse. This system made use of pre-scripted avatar movements to trigger sound samples within the virtual world. The New Atlantis project features a multi-user networked musical design and development environment – built within the Unity game engine and featuring a highly customized set of auxiliary audio and control scripts and classes – optimized for musical performance across wide area networks [132].

Remote interactions in VR have also been investigated in the context of networked music therapy. The study reported in [141] describes a system enabling patients with spinal cord injuries to participate in online therapeutic group singing. The system did not leverage the Internet, but involved a local area network where connected musicians were placed in different rooms of the same building. Notably, authors emphasized the need for an integrated solution for low-latency audio and VR in networked contexts.

IX. RESEARCH AGENDA

XR leverages the simulation of visual, auditory and other types of sensory stimuli to activate our sense of presence in wholly virtual or augmented spaces. As such, XR shows great potential for facilitating both new and reimagined experiences. The application of XR technology to the composition, consumption and analysis of music has been explored both as a means to reproduce or extend existing models of musical expression and creativity, as well as a means to construct novel musical instruments and experiences that could not exist within a purely ''real'' context. This creates unique opportunities for musical interaction design, wherein musical interfaces can be liberated from physical constraints and user input can be mapped to sound via arbitrarily defined mechanics underlying a Musical XR system [34]. This way, XR can evoke levels of abstraction, immersion, and imagination not possible with conventional musical interfaces, ushering in entirely new ways of engaging with music for both the performer and the listener.

The progress of Musical XR research is inexorably bound to that of XR hardware and software technologies. However, the development of new tools and techniques for XR itself needs to be accompanied by artistic and scientific inquiry aimed at identifying theories and standards that can be applied to XR as a whole, as well as those that are idiomatic to the application domains. This kind of research necessitates a high degree of interdisciplinarity to seamlessly and simultaneously address technological, artistic, cognitive and social dimensions in XR. Our analysis of the body of work in Musical XR has revealed a set of past and current research trends as well as areas that require further attention within each of these dimensions. As a result of this analysis, we propose a research agenda that is designed to support the forward momentum of the Musical XR field. The following points lay out a road map for this with the aim of expanding the context of this research field from the perspectives of technological advances, user perception and behaviour, social and cultural aspects, and creativity.

A. HARDWARE

To progress the design of XR hardware tools specific to musical applications (e.g., wearable devices, tracking systems) in order to allow finer gestural control and more responsive feedback, with a particular focus on the multisensory nature of musical interactions.

Of primary importance in the design of VRMIs are the methods utilized for gestural control of musical sound. Whereas the interfaces of traditional musical instruments often feature finely tuned and sophisticated designs that heavily rely on tactility, such refined constructions cannot generally be found in VRMIs. Control of VRMIs is often based on non-specialized input devices provided with a given VR or visual display system, the majority of which are not capable of supporting the subtle control schemes necessary to play traditional musical instruments. Furthermore, the vast majority of VRMIs are not equipped with haptic feedback mechanisms (for notable exceptions see e.g., [33], [77], [112], [159]). Moving forward, the Musical XR community will need to explore not only how to best leverage ready-made hardware solutions in Musical XR projects, but also how to implement new hardware platforms that are idiomatic to musical applications. To that end, there are numerous opportunities in this area to explore the intersections between Musical XR studies and the vast body of existing research on new musical interfaces for musical expression (NIME).

B. SOFTWARE

To develop software tools capable of supporting creativity, expression, education and performance in Musical XR; this entails the design of XR software with domain-specific features as well as user interfaces that cater to a wide range of levels of musical expertise.

Much like the broader XR community, Musical XR practitioners rely on common XR software tools, such as game engines, to develop interactive applications. Although a few domain-specific Musical XR software platforms already exist (e.g., [2], [81], [160]), there is a need for development tools specifically conceived for Musical XR, which are capable of addressing the idiosyncrasies of designing for auditory and musical applications. This would involve taking into consideration the immersive qualities of sound on one hand, and the domain-specific needs of music practitioners on the other. Such tools should also natively support sophisticated dynamic audio synthesis and processing, moving away from the predominately sample-based audio workflows that are standard in most platforms that support XR development.

C. BEST PRACTICES

To adopt existing best practices for Musical XR design, development, and evaluation, as well as to define new best practices as new creative techniques and additional hardware and software tools emerge.

Various design principles for Musical XR systems have already been formulated (e.g, [4], [130]). Nevertheless, with each new technology, our understanding of, and therefore involvement in, the immersive medium evolves, necessitating new sets of guidelines to offer comfortable and engaging experiences to creators, performers and audiences alike. The wide adoption of both existing and future frameworks for the design, development, and evaluation of Musical XR practices will contribute to the sustainability of such practices. These guidelines should also address topics of accessibility and inclusiveness to ensure that the Musical XR field as a whole can cater to a wider portion of the population.

D. SOCIAL EXPERIENCE

To create multi-user systems that can capture the collaborative nature of musical engagement; in particular, to advance our understanding of how to design shared virtual and augmented realities, as well as systems that support XR-based interactions across local and remote networks.

A majority of Musical XR projects so far have been based on single-user experiences, which can contradict the nature of musical collaboration as a social experience. The ability of music and sound to propagate through space, creating a shared experience for groups of performers and listeners should be increasingly leveraged in Musical XR applications. Therefore, fundamental research is needed to devise better methods and design interfaces that can capture this essence of musical collaboration. The lack of ubiquity of current XR hardware systems and the cost and complexity involved in their use have been a deterrent for the immersion of audiences in the same spaces as Musical XR performers. As a result, only a limited number of projects have addressed the topic of XR-based musical interactions over co-located and remote networks [67], [89], [148]. Recent advances in browser-based XR technologies are likely to contribute to research in this area, making the web a promising delivery pipeline for Musical XR content. The connection of Smart Musical Instruments to XR technology reported in [146] represents another promising avenue for both artistic and scientific research. Accordingly, the implementation of XR-based forms of musical interaction among performers and audience members in both co-located and remote settings is gaining major focus among researchers [148].

E. PERCEPTION

To progress our understanding of human perception in Musical XR environments and in XR-mediated musical interactions across a variety of applications (e.g., in education, sound engineering, performance) and users (e.g., audience members, performers), with a particular focus on multisensory aspects.

Research that focuses specifically on human perception within Musical XR contexts has been limited so far. The design of future Musical XR systems would benefit from fundamental research on how different stakeholders interact with virtual elements and spaces in musical contexts. Of particular importance is the manner in which we perceive sound within multimodal XR environments, the effects of latency and the supposition of virtual bodies on our perceived sense of self and our ability to exert musical control, and the manner in which we communicate subtle cues and affectations within shared virtual environments. This requires further research and practice aimed at unravelling the mechanisms underlying multisensory perception, to understand which aspects best contribute to the sense of individual and social presence, and how to avoid artefacts that are detrimental to the musical experience, such as simulation sickness and loss of orientation.

F. COMPOSITION AND PERFORMANCE

To foster the use of XR media in music composition and performance, so as to stimulate the growth of the community of practitioners in Musical XR and, as a consequence, the amount of artistic output (including the creation of a repertoire for VRMIs).

The design of conventional musical instruments, such as the violin or the electric guitar, have been refined over decades or even centuries. Furthermore, extensive musical repertoires have accompanied the development of these instruments. By contrast, XR-based musical interfaces such as VRMIs have appeared only in recent years and often do not make their way into large communities of users. Such instruments therefore, lack a history of design evolution and repertoire, as well as a solid user base that would normally support such evolution both culturally and financially. For Musical XR to gain a stronger foothold in the artistic landscape, composers and performers will need to envision long-standing practices in this domain. At the same time, the pursuit of virtuosity, and the ability of Musical XR music and musicians to display mastery over the form and function of their chosen instrument or system is necessary to instigate this progress. Needless to say, the sustainability of such artistic and technological efforts is intrinsically tied to the continued availability and accessibility of the hardware and software platforms that will facilitate this kind of work.

G. PEDAGOGY

To establish new pedagogical practices that can support both teaching and learning experiences in ways that are specific to Musical XR.

A number of XR-based systems have been devised to support musical education, especially for self-training. However, major gaps exist in our understanding of the effectiveness of such practices. There is a need for not only novel learning experiences that can effectively exploit the affordances of XR for music education, but also extensive evaluation of the benefits of adopting these experiences. As distanced learning becomes more accessible due to technological advances across the world, and more necessary due to health-related and social restrictions placed on a wide range of communities, XR technology has gained a unique position to offer immersive and engaging social and pedagogical engagement in a range of domains including music. An increased focus on the pedagogical uses of Musical XR can leverage traditional oral forms of instruction and pedagogy across the increasingly distanced and digital divides of 21st century societies.

H. DELIVERY PLATFORMS

To create platforms capable of bringing Musical XR experiences to wider audiences through research into how technology can be leveraged to make Musical XR experiences more accessible, and how it can transform the ways in which musical artists engage with audiences.

The audience experience of Musical XR performances is limited due to practical issues related to current technology. For instance, with performances that involve VR, it is often not feasible to bring an entire concert audience into the virtual space due to cost, maintenance and sanitary considerations. Parallel to the continual growth in the accessibility of XR technologies, we need to envision delivery platforms. To this end, software systems that can mitigate the current limitations in concert experiences involving XR performances will need to be devised to render such performances more engaging and immersive for the audiences. Furthermore, the aforementioned ongoing work on browser-based XR experiences can help Musical XR practitioners to engage with broader audiences.

I. INTELLIGENT INTERFACES

To enhance Musical XR interfaces with context-awareness and proactivity features, and provide ad-hoc services to their users to support musical engagement.

Opportunities to progress the state of the art of Musical XR interface design include the enhancement of such instruments with artificial intelligence capable of conferring them with the ability of being aware of their context of use (including the user, activity, location, temporal aspects, and social dimensions), proactively providing services, and wirelessly communicating with a plethora of connected devices and the cloud. Such opportunities envisioned in [145] with respect to Smart VRMIs have yet to be explored for the most part.

J. STANDARDIZATION

To define standards for Musical XR that will allow practitioners to avoid fragmentation and facilitate interoperability among different XR platforms.

What emerges from our survey of the literature is a picture of Musical XR as a field rather fragmented at the technological level, where artists and researchers have often focused on the development of bespoke systems and individual technologies. A project created using one engine or framework is generally not compatible with or easily portable to not only other engines or frameworks, but also with the previous versions of the tools used in the design of the project. Audiovisual output and hand-held control systems are for the most part proprietary and not interoperable, much to the frustration of content and system creators. Such fragmentation significantly hinders the development and successful adoption of Musical XR technologies. Standardization initiatives across the industry, such as WebXR and OpenXR, can counteract this trend, by providing interoperability, compatibility, and effective cross-platform operations. Similarly, the Musical XR community can strive to formulate or agree on standards, such as communication protocols, input mapping schemes and immersive audio processing and spatialization APIs to establish a level of platform independence for future work in this domain.

X. CONCLUSION

In this article we have attempted to define the field of Musical XR, using a hybrid approach encompassing interviews with experts and a systematic review of relevant peer-reviewed literature. The first step was to build a database of publications related to both theoretical aspects and practical applications of Musical XR, currently comprising 260 entries. This database was evaluated in various ways: we presented the historical distribution of the entries in the database, portraying the field of Musical XR since 1977; we analyzed 199 entries from the last decade based on the theoretical or applicative nature of the study, target user, primary function, the social aspect of the XR experience and its level of connectivity; we discussed experimental, artistic and perceptual studies, surveyed evaluation metrics, and offered a comparative overview of what constitutes the state of the art in Musical XR.

Based on our findings we discussed trends and themes that emerged in the last decade, and how projects from this period have been pushing boundaries of the field. We also identified new directions that could benefit the practitioners and users of Musical XR. The results of our investigation shed light on many areas of the field that still require further technological, artistic and theoretical inquiry. Whereas certain areas of the Musical XR field are more mature (e.g., XR performances, VRMIs design), others are relatively under-developed (e.g., networked XR performances, XR systems that support composition and sound engineering). Clearly, the evolution of these fields are directly influenced, by developments in XR technology.

We believe that XR shows its greatest potential for musical applications when it facilitates experiences that cannot be encountered in the real world. We also believe that XR technologies enable novel musical experiences that extend beyond those offered by conventional musical interfaces, such as analog or digital musical instruments. Despite the depth and breadth of the studies discussed in this paper, Musical XR as a field is still in its infancy. We hope that this paper will stimulate discussions across a broad range of communities and encourage a growing body of researchers to take on new technical and artistic challenges in Musical XR.

ACKNOWLEDGMENT

The authors would like to thank the field experts who agreed to participate in the interviews: F. Berthaut (Assistant Professor, University of Lille), M. Ciciliani (Professor, University of Music and Performing Arts Graz), P. Cook (Emeritus Professor, Princeton University), A. McLeran (Lead Audio Programmer, Epic Games), J. K.-Morin (Professor, University of California, Santa Barbara), S. Serafin (Professor, Aalborg University Copenhagen), G. Wang (Associate Professor, Stanford University), and V. Zappi (Assistant Professor, Northeastern University).

REFERENCES

[1] L. J. Rosemblum and R. A. Cross, "The challenge of virtual reality," in *Visualization & Modeling*, W. R. Earnshaw and H. J. Vince, Eds. New York, NY, USA: Academic, 1997, pp. 325–399.

[2] J. Atherton and G. Wang, "Chunity: Integrated audiovisual programming in unity," in *Proc. Int. Conf. New Interfaces Musical Expression*, 2018, pp. 102–107.

[3] J. Atherton and G. Wang, "Curating perspectives: Incorporating virtual reality into laptop orchestra performance," in *Proc. Int. Conf. New Interfaces Musical Expression*, 2020, pp. 154–159.

[4] J. Atherton and G. Wang, "Doing vs. being: A philosophy of design for artful VR," *J. New Music Res.*, vol. 49, no. 1, pp. 35–59, Jan. 2020.

[5] F. Avanzini, A. Baratè, M. Cottini, L. A. Ludovico, and M. Mandanici, "Developing music harmony awareness in young students through an augmented reality approach," in *Proc. 4th Int. Conf. Comput.-Hum. Interact. Res. Appl.*, 2020, pp. 56–63.

[6] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," *IEEE Comput. Graph. Appl.*, vol. 21, no. 6, pp. 34–47, Nov. 2001.

[7] R. T. Azuma, "A survey of augmented reality," *Presence*, vol. 6, no. 4, pp. 355–385, 1997.

[8] A. Barmpoutis, R. Faris, L. Garcia, L. Gruber, J. Li, F. Peralta, and M. Zhang, "Assessing the role of virtual reality with passive haptics in music conductor education: A pilot study," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2020, pp. 275–285.

[9] J. Bell and B. Carey, "Animated notation, score distribution and AR-VR environments for spectral mimetic transfer in music composition," in *Proc. Int. Conf. Technol. Music Notation Represent.*, 2019, pp. 1–9.

[10] S. Benford, C. Greenhalgh, G. Reynard, C. Brown, and B. Koleva, "Understanding and constructing shared spaces with mixed-reality boundaries," *ACM Trans. Comput.-Hum. Interact.*, vol. 5, no. 3, pp. 185–223, Sep. 1998.

[11] J. Bennington and D. Ko, "Physical controllers vs. hand-and-gesture tracking: Control scheme evaluation for VR audio mixing," in *Audio Engineering Society Convention 147*. New York, NY, USA: Audio Engineering Society, 2019.

[12] I. Bergstrom, S. Azevedo, P. Papiotis, N. Saldanha, and M. Slater, "The plausibility of a string quartet performance in virtual reality," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 4, pp. 1352–1359, Apr. 2017.

[13] A. J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Amer.*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[14] F. Berthaut, "3D interaction techniques for musical expression," *J. New Music Res.*, vol. 49, no. 1, pp. 60–72, Jan. 2020.

[15] F. Berthaut, M. Desainte-Catherine, and M. Hachet, "Interacting with 3D reactive widgets for musical performance," *J. New Music Res.*, vol. 40, no. 3, pp. 253–263, Sep. 2011.

[16] F. Berthaut and M. Hachet, "Spatial interfaces and interactive 3D environments for immersive musical performances," *IEEE Comput. Graph. Appl.*, vol. 36, no. 5, pp. 82–87, Sep. 2016.

[17] F. Berthaut, M. Hachet, and M. Desainte-Catherine, "Piivert: Percussion-based interaction for immersive virtual environments," in *Proc. IEEE Symp. 3D User Interfaces (DUI)*, Mar. 2010, pp. 15–18.

[18] F. Berthaut, M. Marshall, S. Subramanian, and M. Hachet, "Rouages: Revealing the mechanisms of digital musical instruments to the audience," in *Proc. Conf. New Interfaces Musical Expression*, 2013, pp. 1–6.

[19] F. Berthaut, D. M. Plasencia, M. Hachet, and S. Subramanian, "Reflets: Combining and revealing spaces for musical performances," in *Proc. Conf. New Interfaces Musical Expression*, 2015, pp. 1–6.

[20] F. Berthaut, V. Zappi, and D. Mazzanti, "Scenography of immersive virtual musical instruments," in *Proc. IEEE VR Workshop, Sonic Interact. Virtual Environ. (SIVE)*, Mar. 2014, pp. 19–24.

[21] A. Birhanu and S. Rank, "KeynVision: Exploring piano pedagogy in mixed reality," in *Proc. Extended Abstr. Publication Annu. Symp. Comput.-Hum. Interact. Play*, Oct. 2017, pp. 299–304.

[22] J. Bissonnette, F. Dubé, M. D. Provencher, and M. T. M. Sala, "Virtual reality exposure training for musicians: Its effect on performance anxiety and quality," *Med. Problems Performing Artists*, vol. 30, no. 3, pp. 169–177, Sep. 2015.

[23] S. Boyer, "A virtual failure: Evaluating the success of Nintendo's virtual boy," *Velvet Light Trap*, vol. 64, pp. 23–33, Sep. 2009.

[24] Z. Buckley and K. Carlson, "Towards a framework for composition design for music-led virtual reality experiences," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2019, pp. 1497–1499.

[25] S. T. Bulu, "Place presence, social presence, co-presence, and satisfaction in virtual worlds," *Comput. Edu.*, vol. 58, no. 1, pp. 154–161, Jan. 2012.

[26] M. Cabral, M. Zuffo, A. Montes, G. Roque, O. Belloc, M. Nagamura, R. R. A. Faria, F. Teubl, C. Kurashima, and R. Lopes, "Crosscale: A 3D virtual musical instrument interface," in *Proc. IEEE Symp. 3D User Interfaces (DUI)*, Mar. 2015, pp. 199–200.

[27] M. Cai, M. A. Amrizal, T. Abe, and T. Suganuma, "Design and implementation of AR-supported system for piano learning," in *Proc. IEEE 8th Global Conf. Consum. Electron. (GCCE)*, Oct. 2019, pp. 49–50.

[28] P. Cairns, H. Daffern, and G. Kearney, "Immersive network music performance: Design and practical deployment of a system for immersive vocal performance," in *Audio Engineering Society Convention 149*. New York, NY, USA: Audio Engineering Society, 2020.

[29] C. Çakmak, A. Çamcı, and A. G. Forbes, "Networked virtual environments as collaborative music spaces," in *Proc. Int. Conf. New Interfaces Musical Expression*, 2016, pp. 106–111.

[30] C. Çakmak and R. Hamilton, "Composing spatial music with Web audio and WebVR," in *Proc. Web Audio Conf.*, 2019, pp. 1–5.

[31] A. Çamcı, "Some considerations on creativity support for VR audio," in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2019, pp. 1500–1502.

[32] A. Çamcı, "Teaching immersive media at the 'dawn of the new everything,'" in *Proc. Int. Audio Mostly Conf.*, 2020, pp. 229–232.

[33] A. Çamcı and J. Granzow, "Hyperreal instruments: Bridging VR and digital fabrication to facilitate new forms of musical expression," *Leonardo Music J.*, vol. 29, pp. 14–18, Dec. 2019.

[34] A. Çamcı and R. Hamilton, "Audio-first VR: New perspectives on musical experiences in virtual environments," *J. New Music Res.*, vol. 49, no. 1, pp. 1–7, Jan. 2020.

[35] A. Çamcı, Z. Özcan, and D. Pehlevan, "Interactive virtual soundscapes: A research report," in *Proc. Int. Comput. Music Conf.*, 2015, pp. 118–125.

[36] A. Çamcı, M. Vilaplana, and L. Wang, "Exploring the affordances of VR for musical interaction design with VIMEs," in *Proc. Int. Conf. New Interfaces Musical Expression*, 2020, pp. 1–6.

[37] B. E. Carey, "Spectrascore VR: Networkable virtual reality software tools for real-time composition and performance," in *Proc. Int. Conf. New Interfaces Musical Expression*, 2016, pp. 1–2.

[38] A. M. Cassorla, G. Kearney, A. Hunt, H. Riaz, M. Stiles, and D. T. Murphy, "Augmented reality for DAW-based spatial audio creation using smartphones," in *Audio Engineering Society Convention 148*. New York, NY, USA: Audio Engineering Society, 2020.

[39] M. Centenaro, P. Casari, and L. Turchet, "Towards a 5G communication architecture for the Internet of musical things," in *Proc. 27th Conf. Open Innov. Assoc. (FRUCT)*, Sep. 2020, pp. 38–45.

[40] E. Chang, H. T. Kim, and B. Yoo, "Virtual reality sickness: A review of causes and measurements," *Int. J. Hum.–Comput. Interact.*, vol. 36, no. 17, pp. 1658–1682, Oct. 2020.

[41] P.-Y. Chiang and C.-H. Sun, "Oncall piano sensei: Portable AR piano training system," in *Proc. 3rd ACM Symp. Spatial User Interact.*, Aug. 2015, p. 134.

[42] J. Chow, H. Feng, R. Amor, and B. C. Wünsche, "Music education using augmented reality with a head mounted display," in *Proc. 14th Australas. User Interface Conf.*, 2013, pp. 73–79.

[43] M. Ciciliani, "Virtual 3D environments as composition and performance spaces," *J. New Music Res.*, vol. 49, no. 1, pp. 104–113, Jan. 2020.

[44] P. R. Cook and G. P. Scavone, "The synthesis toolkit (STK)," in *Proc. Comput. Music Assoc.*, 1999, pp. 164–166.

[45] W. Costa, L. Ananias, I. Barbosa, B. Barbosa, A. De' Carli, R. R. Barioni, L. Figueiredo, V. Teichrieb, and D. Filgueira, "Songverse: A music-loop authoring tool based on virtual reality," in *Proc. 21st Symp. Virtual Augmented Reality (SVR)*, Oct. 2019, pp. 216–222.

[46] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart, "The CAVE: Audio visual experience automatic virtual environment," *Commun. ACM*, vol. 35, no. 6, pp. 64–72, Jun. 1992.

[47] J. J. Cummings and J. N. Bailenson, "How immersive is enough? A meta-analysis of the effect of immersive technology on user presence," *Media Psychol.*, vol. 19, no. 2, pp. 272–309, Apr. 2016.

[48] F. De Sorbier, H. Shiino, and H. Saito, "Violin pedagogy for finger and bow placement using augmented reality," in *Proc. Asia Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, 2012, pp. 1–5.

[49] E. D. Innocenti, M. Geronazzo, D. Vescovi, R. Nordahl, S. Serafin, L. A. Ludovico, and F. Avanzini, "Mobile virtual reality for musical genre learning in primary education," *Comput. Edu.*, vol. 139, pp. 102–117, Oct. 2019.

[50] M. S. Del Rio-Guerra, J. Martin-Gutierrez, V. A. Lopez-Chao, R. F. Parra, and M. A. R. Sosa, "AR graphic representation of musical notes for self-learning on guitar," *Appl. Sci.*, vol. 9, no. 21, p. 4527, Oct. 2019.

[51] P. R. Desai, M. D. Parle, P. N. Desai, K. D. Ajmera, and K. Mehta, "A review paper on oculus Rift—A virtual reality headset," *Int. J. Eng. Trends Technol.*, vol. 13, no. 4, pp. 175–179, Jul. 2014.

[52] J. Desnoyers-Stewart, D. Gerhard, and M. Smith, "Augmenting a MIDI keyboard using virtual interfaces," *J. Audio Eng. Soc.*, vol. 66, no. 6, pp. 439–447, Jun. 2018.

[53] B. Dewitz, R. Wiche, C. Geiger, F. Steinicke, and J. Feitsch, "AR sound sandbox: A playful interface for musical and artistic expression," in *Proc. Int. Conf. Intell. Technol. Interact. Entertainment*. Cham, Switzerland: Springer, 2017, pp. 59–76.

[54] T. Dolby, *The Speed of Sound: Breaking the Barriers Between Music and Technology: A Memoir*. New York, NY, USA: Flatiron Books, 2016.

[55] K. Enriquez, M. Palacios, D. Pallo, and G. Guerrero, "SENSE: Sensory component VR application for hearing impaired people to enhance the music experience," in *Proc. 15th Iberian Conf. Inf. Syst. Technol. (CISTI)*, Jun. 2020, pp. 1–6.

[56] V. F. Martins, L. Gomez, and A. G. D. Corrêa, "Teaching children musical perception with MUSIC-AR," *EAI Endorsed Trans. e-Learning*, vol. 2, no. 5, p. e3, Mar. 2015.

[57] J. Fillwalk, "ChromaChord: A virtual musical instrument," in *Proc. IEEE Symp. 3D User Interfaces (DUI)*, Mar. 2015, pp. 201–202.

[58] S. S. Fisher, E. M. Wenzel, C. Coler, and M. W. McGreevy, "Virtual interface environment workstations," in *Proc. Hum. Factors Soc. Annu. Meeting*, vol. 32. Los Angeles, CA, USA: Sage, 1988, pp. 91–95.

[59] W. Gale and J. Wakefield, "Investigating the use of virtual reality to solve the underlying problems with the 3D stage paradigm," in *Proc. 4th Workshop Intell. Music Prod.*, 2018, pp. 1–4.

[60] R. Gaugne, F. Nouviale, O. Rioual, A. Chirat, K. Gohon, V. Goupil, M. Toutirais, B. Bossis, and V. Gouranton, "EvoluSon: Walking through an interactive history of music," *Presence, Teleoperators Virtual Environ.*, vol. 26, no. 3, pp. 281–296, May 2018.

[61] S. Gelineck and D. Korsgaard, "An exploratory evaluation of user interfaces for 3D audio mixing," in *Audio Engineering Society Convention 138*. New York, NY, USA: Audio Engineering Society, 2015.

[62] S. Gelineck and D. Overholt, "Haptic and visual feedback in 3D audio mixing interfaces," in *Proc. Audio Mostly Interact. With Sound (AM)*, 2015, pp. 1–6.

[63] L. Gerry, S. Dahl, and S. Serafin, "Adept: Exploring the design, pedagogy, and analysis of a mixed reality application for piano training," in *Proc. Sound Music Comput. Conf.*, 2019, pp. 2891–2892.

[64] R. Graham and S. Cluett, "The soundfield as sound object: Virtual reality environments as a three-dimensional canvas for music composition," in *Proc. Audio Eng. Soc. Conf., AES Int. Conf. Audio Virtual Augmented Reality*. New York, NY, USA: Audio Engineering Society, 2016.

[65] M. P. J. Habgood, D. Wilson, D. Moore, and S. Alapont, "HCI lessons from PlayStation VR," in *Proc. Extended Abstr. Publication Annu. Symp. Comput.-Hum. Interact. Play*, Oct. 2017, pp. 125–135.

[66] D. Hackl and C. Anthes, "HoloKeys—An augmented reality application for learning the piano," in *Forum Media Technology*. St. Pölten, Austria: St. Pölten Univ. of Applied Sciences, 2017, pp. 140–144.

[67] R. Hamilton, "Coretet: A 21st century virtual interface for musical expression," in *Proc. Int. Symp. Comput. Music Multidisciplinary Res.*, 2019, pp. 1010–1021.

[68] R. Hamilton, "Mediated musical interactions in virtual environments," in *New Directions in Music and Human-Computer Interaction*. Cham, Switzerland: Springer, 2019, pp. 243–257.

[69] R. Hamilton, "Trois machins de la Grâce Aimante: A virtual reality string quartet," in *Proc. Int. Comput. Music Assoc. Conf.*, 2019, pp. 1–5.

[70] R. Hamilton, J.-P. Caceres, C. Nanou, and C. Platz, "Multi-modal musical environments for mixed-reality performance," *J. Multimodal User Interfaces*, vol. 4, nos. 3–4, pp. 147–156, Dec. 2011.

[71] R. Hamilton and C. Platz, "Gesture-based collaborative virtual reality performance in carillon," in *Proc. Int. Comput. Music Conf.*, 2016, pp. 337–340.

[72] M. Heilig, "Stereoscopic-television apparatus for individual use," U.S. Patent 2 955 156, Oct. 4, 1960.

[73] T. Hermann and A. Hunt, "Guest Editors' introduction: An introduction to interactive sonification," *IEEE Multimedia Mag.*, vol. 12, no. 2, pp. 20–24, Apr. 2005.

[74] F. Huang, Y. Zhou, Y. Yu, Z. Wang, and S. Du, "Piano AR: A markerless augmented reality based piano teaching system," in *Proc. 3rd Int. Conf. Intell. Hum.-Mach. Syst. Cybern.*, Aug. 2011, pp. 47–52.

[75] M. Hudak, S. Korecko, and B. Sobota, "Advanced user interaction for Web-based collaborative virtual reality," in *Proc. 11th IEEE Int. Conf. Cognit. Infocommunications (CogInfoCom)*, Sep. 2020, pp. 000343–000348.

[76] M. Hughes and A. Johnston, "URack: Modular audio-visual composition with unity and VCV rack," in *Proc. Int. Conf. New Interfaces Musical Expression*, 2020, pp. 321–322.

[77] I. Hwang, H. Son, and J. R. Kim, "AirPiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback," in *Proc. IEEE World Haptics Conf. (WHC)*, Jun. 2017, pp. 213–218.

[78] G. Jakus, J. Guna, S. Tomažič, and J. Sodnik, "Evaluation of leap motion controller with a high precision optical tracking system," in *Proc. Int. Conf. Hum.-Comput. Interact.* Cham, Switzerland: Springer, 2014, pp. 254–263.

[79] J. Janer, E. Gomez, A. Martorell, M. Miron, and B. de Wit, "Immersive orchestras: Audio processing for orchestral music VR content," in *Proc. 8th Int. Conf. Games Virtual Worlds Serious Appl. (VS-GAMES)*, Sep. 2016, pp. 1–2.

[80] X. Jiang, H. Shokri-Ghadikolaei, G. Fodor, E. Modiano, Z. Pang, M. Zorzi, and C. Fischione, "Low-latency networking: Where latency lurks and how to tame it," *Proc. IEEE*, vol. 107, no. 2, pp. 280–306, Feb. 2019.

[81] D. Johnson, D. Damian, and G. Tzanetakis, "OSC-XR: A toolkit for extended reality immersive music interfaces," in *Proc. Sound Music Comput. Conf.*, 2019, pp. 202–209.

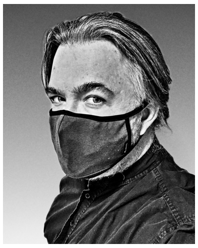
[82] D. Johnson and G. Tzanetakis, "Vrmin: Using mixed reality to augment the theremin for musical tutoring," in *Proc. Int. Conf. New Interfaces Musical Expression*, 2017, pp. 151–156.

[83] D. Jordan, F. Müller, C. Drude, S. Reinhold, V. Schomakers, and M. Teistler, "Spatial audio engineering in a virtual reality environment," in *Proc. Mensch Und Comput. Tagungsband*, 2016, pp. 1–4.

[84] T. Kaneko, H. Tarumi, K. Kataoka, Y. Kubochi, D. Yamashita, T. Nakai, and R. Yamaguchi, ''Supporting the sense of unity between remote audiences in VR-based remote live music support system KSA2,'' in *Proc. IEEE Int. Conf. Artif. Intell. Virtual Reality (AIVR)*, Dec. 2018, pp. 124–127.
[85] M. W. Krueger, ''Responsive environments,'' in *Proc. Nat. Comput. Conf.*, 1977, pp. 423–433.
[86] J. Lanier, ''The sound of one hand,'' *Whole Earth Rev.*, vol. 79, pp. 4–30, 1993.
[87] L. Bruno, G. D. C. Ana, N. Marilena, and D. L. Roseli, ''Augmented reality musical app to support Children's musical education,'' *J. Comput. Sci. Inf. Technol.*, vol. 5, no. 4, pp. 121–127, Sep. 2017.
[88] M. Löchtefeld, S. Gehring, R. Jung, and A. Krüger, ''Using mobile projection to support guitar learning,'' in *Proc. Int. Symp. Smart Graph.* Berlin, Germany: Springer, 2011, pp. 103–114.
[89] B. Loveridge, ''Networked music performance in virtual reality: Current perspectives,'' *J. Netw. Music Arts*, vol. 2, no. 1, p. 2, 2020.
[90] B. MacIntyre and T. F. Smith, ''Thoughts on the future of WebXR and the immersive Web,'' in *Proc. IEEE Int. Symp. Mixed Augmented Reality Adjunct (ISMAR-Adjunct)*, Oct. 2018, pp. 338–342.
[91] T. Mäki-Patola, J. Laitinen, A. Kanerva, and T. Takala, ''Experiments with virtual reality instruments,'' in *Proc. Int. Conf. New Interfaces Musical Expression*, 2005, pp. 11–16.
[92] D. G. Malham and A. Myatt, ''3-D sound spatialization using ambisonic techniques,'' *Comput. Music J.*, vol. 19, no. 4, pp. 58–70, 1995.
[93] G. F. B. Martín, ''Social and psychological impact of musical collective creative processes in virtual environments; the avatar orchestra metaverse in second life,'' *Musica/Tecnologia*, vol. 30, pp. 73–85, Aug. 2018.
[94] R. Masu, P. Bala, M. A. Ahmad, N. D. N. Correia, V. Nisi, N. Nunes, and T. Rom ao, ''VR open scores: Scores as inspiration for VR scenarios,'' in *Proc. Int. Conf. New Interfaces Musical Expression*, 2020, pp. 109–114.
[95] J. D. Mathew, S. Huot, and B. F. G. Katz, ''Survey and implications for the design of new 3D audio production and authoring tools,'' *J. Multimodal User Interface*, vol. 11, no. 3, pp. 277–287, Sep. 2017.
[96] D. Mazzanti, V. Zappi, D. G. Caldwell, and A. Brogni, ''Augmented stage for participatory performances,'' in *Proc. Conf. New Interfaces Musical Expression*, 2014, pp. 29–34.
[97] D. McFadden, A. Tavakkoli, J. Regenbrecht, and B. Wilson, ''Augmented virtuality: A real-time process for presenting real-world visual sensory information in an immersive virtual environment for planetary exploration,'' in *Proc. AGU Fall Meeting Abstracts*, Dec. 2017, Paper IN32B-07.
[98] L. Men and N. Bryan-Kinns, ''LeMo: Exploring virtual space for collaborative creativity,'' in *Proc. Creativity Cognition*, Jun. 2019, pp. 71–82.
[99] P. Milgram and F. Kishino, ''A taxonomy of mixed reality visual displays,'' *IEICE Trans. Inf. Syst.*, vol. 77, no. 12, pp. 1321–1329, Dec. 1994.
[100] R. Mills, *Tele-Improvisation: Intercultural Interact. Online Global Music Jam Session*. Cham, Switzerland: Springer, 2019.
[101] D. Molero, S. Schez-Sobrino, D. Vallejo, C. Glez-Morcillo, and J. Albusac, ''A novel approach to learning music and piano based on mixed reality and gamification,'' *Multimedia Tools Appl.*, vol. 80, pp. 1–22, Sep. 2020.
[102] A. G. Moore, M. J. Howell, A. W. Stiles, N. S. Herrera, and R. P. Mcmahan, ''Wedge: A musical interface for building and playing composition-appropriate immersive environments,'' in *Proc. IEEE Symp. 3D User Interfaces (DUI)*, Mar. 2015, pp. 205–206.
[103] M. Moth-Poulsen, T. Bednarz, V. Kuchelmeister, S. Design, and S. Serafin, ''Teach me drums: Learning rhythms through the embodiment of a drumming teacher in virtual reality,'' in *Proc. Sound Music Comput. Conf. (SMC)*, 2019, pp. 269–273.
[104] Y. Motokawa and H. Saito, ''Support system for guitar playing using augmented reality display,'' in *Proc. IEEE/ACM Int. Symp. Mixed Augmented Reality*, Oct. 2006, pp. 243–244.
[105] D. C. Niehorster, L. Li, and M. Lappe, ''The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research,'' *i-Perception*, vol. 8, no. 3, 2017, Art. no. 2041669517708205.
[106] S. Nilsson, V. Vechev, A. Yeh, and C. Hedler, ''Holo beats: Design and development of an ar system to teach drums,'' in *Proc. 12th Student Interact. Design Res. Conf.*, 2016, pp. 241–253.
[107] K. L. Nowak and F. Biocca, ''The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments,'' *Presence, Teleoperators Virtual Environ.*, vol. 12, no. 5, pp. 481–494, Oct. 2003.
[108] Y. Ohta and H. Tamura, *Mixed Reality: Merging Real and Virtual Worlds*. Berlin, Germany: Springer-Verlag, 2014.
[109] Z. Özcan and A. Çamcı, ''An augmented reality music composition based on the sonification of animal behavior,'' in *Proc. Audio Eng. Soc. Conf., AES Int. Conf. Audio Virtual Augmented Reality*. New York, NY, USA: Audio Engineering Society, 2018, pp. 1–8.
[110] M. Palumbo, A. Zonta, and G. Wakefield, ''Modular reality: Analogues of patching in immersive space,'' *J. New Music Res.*, vol. 49, no. 1, pp. 8–23, Jan. 2020.
[111] S. Park, ''ARLooper: Collaborative audiovisual experience with mobile devices in a shared augmented reality space,'' in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2020, pp. 1–4.
[112] A. Passalenti, R. Paisa, N. C. Nilsson, N. S. Andersson, F. Fontana, R. Nordahl, and S. Serafin, ''No strings attached: Force and vibrotactile feedback in a virtual guitar simulation,'' in *Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR)*, Mar. 2019, pp. 28–31.
[113] N. Petrovic, ''Augmented and virtual reality Web applications for music stage performance,'' in *Proc. 55th Int. Sci. Conf. Inf., Commun. Energy Syst. Technol. (ICEST)*, Sep. 2020, pp. 33–36.
[114] S. Phunsa, ''Applying augmented reality technology to promote traditional thai folk musical instruments on postcards,'' in *Proc. Int. Conf. Comput. Graph., Multimedia Image Process.*, 2014, pp. 1–5.
[115] A. Pirchner, ''IRMA (interactive real-time measurement of attention). A method for the investigation of audiovisual computer music performances,'' in *Proc. Int. Comput. Music Conf.*, 2019, pp. 1–7.
[116] I. Poupyrev, R. Berry, J. Kurumisawa, K. Nakao, M. Billinghurst, C. Airola, H. Kato, T. Yonezawa, and L. Baldwin, ''Augmented groove: Collaborative jamming in augmented reality,'' in *Proc. ACM SIGGRAPH Conf. Abstr. Appl.*, 2000, p. 77.
[117] J. Pressing, ''Some perspectives on performed sound and music in virtual environments,'' *Presence, Teleoperators Virtual Environ.*, vol. 6, no. 4, pp. 482–503, Aug. 1997.
[118] V. Pulkki, ''Virtual sound source positioning using vector base amplitude panning,'' *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466, 1997.
[119] X. Qiao, P. Ren, S. Dustdar, L. Liu, H. Ma, and J. Chen, ''Web AR: A promising future for mobile augmented reality—State of the art, challenges, and insights,'' *Proc. IEEE*, vol. 107, no. 4, pp. 651–666, Apr. 2019.
[120] L. Raymaekers, J. Vermeulen, K. Luyten, and K. Coninx, ''Game of tones: Learning to play songs on a piano using projected instructions and games,'' in *Proc. CHI Extended Abstr. Hum. Factors Comput. Syst.*, Apr. 2014, pp. 411–414.
[121] J. Rees-Jones and H. Daffern, ''The hills are alive: Capturing and presenting an outdoor choral performance for virtual reality,'' in *Proc. Audio Eng. Soc. Conf., AES Int. Conf. Immersive Interact. Audio*, 2019.
[122] K. Rogers, M. Weber, A. Röhlig, M. Weing, J. Gugenheimer, B. Könings, M. Klepsch, F. Schaub, E. Rukzio, and T. Seufert, ''Piano: Faster piano learning with interactive projection,'' in *Proc. 9th ACM Int. Conf. Interact. Tabletops Surf. (ITS)*, 2014, pp. 149–158.
[123] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, ''An overview on networked music performance technologies,'' *IEEE Access*, vol. 4, pp. 8823–8843, 2016.
[124] F. E. Sandnes and E. Eika, ''Enhanced learning of jazz chords with a projector based piano keyboard augmentation,'' in *Proc. Int. Conf. Innov. Technol. Learn.* Cham, Switzerland: Springer, 2019, pp. 194–203.
[125] G. Santini, ''Augmented piano in augmented reality,'' in *Proc. Conf. New Interfaces Musical Expression*, 2020, pp. 411–415.
[126] G. Santini, ''Composition as an embodied act: A framework for the gesture-based creation of augmented reality action scores,'' in *Proc. Sound Music Comput. Conf.*, 2020, pp. 357–363.
[127] M. F. Schober, ''Virtual environments for creative work in collaborative music-making,'' *Virtual Reality*, vol. 10, no. 2, pp. 85–94, Sep. 2006.
[128] S. Serafin, A. Adjorlu, L. Andersen, and N. Andersen, ''Singing in virtual reality with the Danish national children's choir,'' in *Proc. Int. Symp. Comput. Music Multidisciplinary Res.*, 2019, pp. 241–253.
[129] S. Serafin, A. Adjorlu, N. Nilsson, L. Thomsen, and R. Nordahl, ''Considerations on the use of virtual and augmented reality technologies in music education,'' in *Proc. IEEE Virtual Reality Workshop K-12 Embodied Learn. Through Virtual Augmented Reality (KELVAR)*, Mar. 2017, pp. 1–4.
[130] S. Serafin, C. Erkut, J. Kojs, N. C. Nilsson, and R. Nordahl, ''Virtual reality musical instruments: State of the art, design principles, and future directions,'' *Comput. Music J.*, vol. 40, no. 3, pp. 22–40, Sep. 2016.

[131] W. R. Sherman and A. B. Craig, *Understanding Virtual Reality: Interface, Application, and Design*. San Mateo, CA, USA: Morgan Kaufmann, 2018.
[132] P. Sinclair, R. Cahen, J. Tanant, and P. Gena, ''New atlantis: Audio experimentation in a shared online world,'' in *Proc. Int. Symp. Comput. Music Multidisciplinary Res.* Cham, Switzerland: Springer, 2016, pp. 229–246.
[133] M. Slater, ''Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments,'' *Phil. Trans. Roy. Soc. B, Biol. Sci.*, vol. 364, no. 1535, pp. 3549–3557, Dec. 2009.
[134] M. Slater, ''Grand challenges in virtual environments,'' *Frontiers Robot. AI*, vol. 1, p. 3, May 2014.
[135] M. Slater, D.-P. Pertaub, and A. Steed, ''Public speaking in virtual reality: Facing an audience of avatars,'' *IEEE Comput. Graph. Appl.*, vol. 19, no. 2, pp. 6–9, Mar./Apr. 1999.
[136] M. Slater and A. Steed, ''A virtual presence counter,'' *Presence, Teleoperators Virtual Environ.*, vol. 9, no. 5, pp. 413–434, Oct. 2000.
[137] K. Snook, T. Barri, M. Bolles, P. Ericson, C. Fravel, J. Goßmann, S. E. Green-Mateu, A. Luck, M. Schedel, and R. Thomas, ''Concordia: A musical XR instrument for playing the solar system,'' *J. New Music Res.*, vol. 49, no. 1, pp. 88–103, Jan. 2020.
[138] M. Speicher, B. D. Hall, and M. Nebeling, ''What is mixed reality?'' in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, May 2019, pp. 1–15.
[139] M. Sukan, O. Oda, X. Shi, M. Entrena, S. Sadalgi, J. Qi, and S. Feiner, ''ARmonica: A collaborative sonic environment,'' in *Proc. Adjunct 23nd Annu. ACM Symp. User Interface Softw. Technol. (UIST)*, 2010, pp. 401–402.
[140] I. E. Sutherland, ''A head-mounted three dimensional display,'' in *Proc. Fall Joint Comput. Conf. I*, 1968, pp. 757–764.
[141] J. Tamplin, B. Loveridge, K. Clarke, Y. Li, and D. J. Berlowitz, ''Development and feasibility testing of an online virtual reality platform for delivering therapeutic group singing interventions for people living with spinal cord injury,'' *J. Telemedicine Telecare*, vol. 26, no. 6, pp. 365–375, Jul. 2020.
[142] K. L. Tan and C. K. Lim, ''Development of traditional musical instruments using augmented reality (ar) through mobile learning,'' in *Proc. AIP Conf.* Melville, NY, USA: AIP Publishing LLC, 2018, Art. no. 020140.
[143] E. Thielen, J. Letellier, J. Sieck, and A. Thoma, ''Bringing a virtual string quartet to life,'' in *Proc. 2nd Afr. Conf. Hum. Comput. Interact., Thriving Communities*, Dec. 2018, pp. 1–4.
[144] C. Torres and P. Figueroa, ''Learning how to play a guitar with the HoloLens: A case study,'' in *Proc. 44th Latin Amer. Comput. Conf. (CLEI)*, Oct. 2018, pp. 606–611.
[145] L. Turchet, ''Smart musical instruments: Vision, design principles, and future directions,'' *IEEE Access*, vol. 7, pp. 8944–8963, 2019.
[146] L. Turchet, M. Benincaso, and C. Fischione, ''Examples of use cases with smart instruments,'' in *Proc. 12th Int. Audio Mostly Conf. Augmented Participatory Sound Music Experiences*, Aug. 2017, pp. 47:1–47:5.
[147] L. Turchet, P. Burelli, and S. Serafin, ''Haptic feedback for enhancing realism of walking simulations,'' *IEEE Trans. Haptics*, vol. 6, no. 1, pp. 35–45, 1st Quart., 2013.
[148] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, ''Internet of musical things: Vision and challenges,'' *IEEE Access*, vol. 6, pp. 61994–62017, 2018.
[149] E. Varèse and C. Wen-Chung, ''The liberation of sound,'' *Perspect. New Music*, vol. 5, no. 1, pp. 11–19, 1966.
[150] J. C. Vasquez, K. Tahiroglu, and J. Kildal, ''Idiomatic composition practices for new musical instruments: Context, background and current applications,'' in *Proc. Conf. New Interfaces Musical Expression*, 2017, pp. 174–179.
[151] C. Velasco and M. Obrist, *Multisensory Experiences: Where the Senses Meet Technology*. Oxford, U.K.: Oxford Univ. Press, 2020.
[152] D. Velev and P. Zlateva, ''Virtual reality challenges in education and training,'' *Int. J. Learn. Teaching*, vol. 3, no. 1, pp. 33–37, 2017.
[153] J. Wagner, F. Lingenfelser, T. Baur, I. Damian, F. Kistler, and E. André, ''The social signal interpretation (SSI) framework: Multimodal signal processing and recognition in real-time,'' in *Proc. 21st ACM Int. Conf. Multimedia (MM)*, 2013, pp. 831–834.
[154] G. Wang, P. R. Cook, and S. Salazar, ''ChucK: A strongly timed computer music language,'' *Comput. Music J.*, vol. 39, no. 4, pp. 10–29, Dec. 2015.
[155] R. Webster and A. Clark, ''Turn-key solutions: Virtual reality,'' in *Proc. Int. Design Eng. Tech. Conf. Comput. Inf. Eng. Conf.*, vol. 57052. New York, NY, USA: American Society of Mechanical Engineers, 2015, Art. no. V01BT02A052.
[156] J. Weinel, ''Cyberdreams: Visualizing music in extended reality,'' in *Technology, Design and the Arts-Opportunities and Challenges*. Cham, Switzerland: Springer, 2020, pp. 209–227.
[157] J. Weinel, ''Visualising rave music in virtual reality: Symbolic and interactive approaches,'' in *Proc. EVA London*, Jul. 2020, pp. 78–84.
[158] M. Weing, A. Röhlig, K. Rogers, J. Gugenheimer, F. Schaub, B. Könings, E. Rukzio, and M. Weber, ''PIANO: Enhancing instrument learning via interactive projected augmentation,'' in *Proc. ACM Conf. Pervas. Ubiquitous Comput. Adjunct Publication*, Sep. 2013, pp. 75–78.
[159] S. Willemsen, R. Paisa, and S. Serafin, ''Resurrecting the tromba marina: A bowed virtual reality instrument using haptic feedback and accurate physical modelling,'' in *Proc. Sound Music Comput. Conf.*, 2020, pp. 300–307.
[160] A. Willette, N. Gargi, E. Kim, J. Xu, T. Lai, and A. Çamcı, ''Cross-platform and cross-reality design of immersive sonic environments,'' in *Proc. Int. Conf. New Interfaces Musical Expression*, 2020, pp. 127–130.
[161] B. G. Witmer and M. J. Singer, ''Measuring presence in virtual environments: A presence questionnaire,'' *Presence, Teleoperators Virtual Environ.*, vol. 7, no. 3, pp. 225–240, Jun. 1998.
[162] X. Xiao, P. Puentes, E. Ackermann, and H. Ishii, ''Andantino: Teaching children piano with projected animated characters,'' in *Proc. 15th Int. Conf. Interact. Design Children*, Jun. 2016, pp. 37–45.
[163] X. Xiao, B. Tome, and H. Ishii, ''Andante: Walking figures on the piano keyboard to visualize musical motion,'' in *Proc. Conf. New Interfaces Musical Expression*, 2014, pp. 629–632.
[164] T. Yamabe, H. Asuma, S. Kiyono, and T. Nakajima, ''Feedback design in augmented musical instruments: A case study with an AR drum kit,'' in *Proc. IEEE 17th Int. Conf. Embedded Real-Time Comput. Syst. Appl.*, vol. 2, Aug. 2011, pp. 126–129.
[165] Q. Yang and G. Essl, ''Visual associations in augmented keyboard performance,'' in *Proc. Conf. New Interfaces Musical Expression*, 2013, pp. 252–255.
[166] V. Zappi, D. Mazzanti, A. Brogni, and D. Caldwell, ''Concatenative synthesis unit navigation and dynamic rearrangement in vrgrains,'' in *Proc. Sound Music Comput. Conf.*, 2012, pp. 403–408.
[167] V. Zappi, D. Mazzanti, A. Brogni, and D. G. Caldwell, ''Design and evaluation of a hybrid reality performance,'' in *Proc. Conf. New Interfaces Musical Expression*, vol. 11, 2011, pp. 355–360.
[168] Z. Zhang, ''Microsoft kinect sensor and its effect,'' *IEEE Multimedia Mag.*, vol. 19, no. 2, pp. 4–10, Feb. 2012.
[169] D. Zielasko, D. Rausch, Y. C. Law, T. C. Knott, S. Pick, S. Porsche, J. Herber, J. Hummel, and T. W. Kuhlen, ''Cirque Des.bouteilles: The art of blowing on bottles,'' in *Proc. IEEE Symp. 3D User Interfaces (DUI)*, Mar. 2015, pp. 209–210.

LUCA TURCHET (Member, IEEE) received the master's degrees in computer science from the University of Verona, in classical guitar and composition from the Music Conservatory of Verona, and in electronic music from the Royal College of Music of Stockholm, and the Ph.D. degree in media technology from Aalborg University Copenhagen. He is currently an Assistant Professor with the Department of Information Engineering and Computer Science of University of Trento. His scientific, artistic, and entrepreneurial research has been supported by numerous grants from different funding agencies including the European Commission, the European Institute of Innovation and Technology, the European Space Agency, the Italian Minister of Foreign Affairs, and the Danish Research Council. He is a Co-Founder and the Head of Sound and Interaction Design at Elk. His main research interests are in music technology, the Internet of Things, extended reality, human-computer interaction, and multimodal perception. He is an Associate Editor of IEEE Access.



ROB HAMILTON received the Ph.D. degree from the Stanford University's Center for Computer Research in Music and Acoustics. He explores the converging spaces between sound, music and interaction. His creative practice includes mixed and virtual-reality performance works built within fully rendered networked game environments, procedural music engines and mobile musical ecosystems. His research focuses on the cognitive implications of sonified musical gesture and motion and the role of perceived space in the creation and enjoyment of sound and music. He currently serves as an Associate Professor of music and media at the Department of Arts at Rensselaer Polytechnic Institute, Troy, NY.

ANIL ÇAMCI received the Ph.D. degree from the Academy of Creative and Performing Arts at Leiden University in affiliation with the Institute of Sonology in The Hague, and the Industrial Design Department at Delft University of Technology. He is currently an Assistant Professor of performing arts technology with the University of Michigan. His work investigates new tools and theories for multimodal world making at the intersection of extended reality, human-computer interaction, spatial audio and electronic music. Previously, he worked at the University of Illinois at Chicago, where he led research projects on interaction design and immersive audio in virtual reality contexts, and Istanbul Technical University, where he founded the Sonic Arts Program. His research and artistic work has been featured in leading journals and conferences. He has been granted several awards, including the Audio Engineering Society Fellowship, ACM CHI Artist Grant, and NIME Best Installation Prize.