Sonic Scope: New Approaches to Audiovisual Culture

The area~ System: Exploring Real and Virtual Environments Through Gestural Ambisonics and Audio Augmented Reality

Sam Bilbow

Published on: Feb 15, 2021

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

ABSTRACT

The area~ System: Exploring Real and Virtual Environments Through Gestural Ambisonics and Audio Augmented Reality

[Sam Bilbow](#), University of Sussex

In this paper, I outline the development and evaluation of the area~ system. area~ enables users to record, manipulate, and spatialise virtual audio samples or nodes around their immediate environment. Through a combination of ambisonics audio rendering and hand gesture tracking, this system calls attention to the ability of non-visual augmented reality (AR), here, audio augmented reality (AAR), to provide new aesthetic experiences of real and virtual environments. The system is contextualised within the move in computational art, and indeed, broader human computer interaction research, towards multisensory interaction. In particular, area~ is situated in the creative practice of works using multisensory AR as a medium to create expressive computational art.

*Through an autobiographical design study, these experiences are discussed in relation to the research question: “**How can we better understand relationships between virtual and real environments through gesture and virtually placed audio augmented reality objects?**” This hypothesis study proposes that new aesthetic experiences can result from the system and are waiting to be tested through user studies. The adoption of the Project North Star open-source AR head-mounted-display (HMD) could expand the possibilities of the area~ system by reducing the need to be tethered to a laptop and table for hand gesture input.*

In discussing the future development of the system and my research, I propose a devising practice-led method for creating and evaluating new multisensory AR (MSAR) Experiences; as well as the tantalising prospect of adding interaction between other sensory modalities to the area~ system, such as vision or smell, which would be made possible by the use of this open-source HMD.

1. Introduction

An augmented reality (AR) system is widely accepted as a system that utilises the combination of ‘real’ and ‘virtual,’ is interactive in real-time, and registered in three dimensions.¹ Although this definition entertains the prospect of multisensory applications, in the last twenty years, the vast majority of AR applications have dealt with solely visual displays of information.² Indeed, a systematic review from

2018 found that 96% of AR usability studies conducted between 2005-2014 augmented only the visual sense.³ In cross-modality psychology research, this ocular-centrism (the perceptual and epistemological bias ranking vision over other senses) is normally explained by a 'textbook' explanation: "the idea that vision is the most important modality is supported by numerous studies demonstrating visual dominance."⁴ Fabian Huttmacher argues that ocular-centrism can be critiqued through the lenses of:

- A methodological-structural explanation: "Research on vision is often easier than research on other modalities and that this is the result of an initial bias toward vision that reinforces itself."⁵
- A cultural explanation: "The dominance of the visual is not a historical constant, but rather a result of the way (Western) societies are designed."⁶

Within computational arts and specifically within the composition of works using AR as their medium, ocular-centrism has been met in recent years with a push towards multisensory interaction methods.⁷ Meanwhile, within human-computer interaction (HCI) research we have seen the same general movement through the development of applications and studies with multisensory interaction methods.

Within AR and virtual reality (VR) technologies, the trend of ocular-centrism is compounded by the dominance in development of visual head-mounted-displays (HMDs). There has been an increase in the popularity and development of virtual reality HMDs, such as HTC's Cosmos / Vive, the Oculus Quest 2, and Valve's Index.⁸ Augmented reality has seen its own parallel surge in development of HMDs; these include the Magic Leap 1, Nreal Light, and Microsoft's HoloLens 2.⁹ As well as being primarily visual devices these AR HMDs are generally expensive, in the range of £1000 - £2,500. Both VR and AR HMDs often require development licences, being locked into certain software frameworks and provision of personal data to parent companies such as Facebook, HTC, and Microsoft.

Alternatives to these closed-source AR offerings are possible in the form of DIY microcontroller-based solutions like the one described in this paper. Another option is Leap Motion's open-source Project North Star AR HMD designs, released in 2018.¹⁰ These are 3D printable and require approximately £300 - £400 of electronics to get started. This move has helped democratise the use of AR as a medium, tool, and collaborative technology, allowing the maker community to develop tools that aid in critiquing the tensions and various relationships between real and virtual environments.¹¹ Furthermore, it has presented a unique opportunity for computational artists and digital media researchers that want to develop for an AR HMD but either cannot afford to spend £1000-£2500 on one, or cannot justify buying closed-source devices due to clashes in research ethics due the provision of personal data required by parent companies.

1.1 Computational 'ARt' with *area~*

This paper contributes towards the holistic development of AR applications by exploring one of the main attributes of an AR system: the combinational relationship of real and virtual environments, specifically within audio augmented reality (AAR). In doing so, I attempt to form a deeper understanding of AR's aesthetic and material capabilities as an artistic medium.

The *area~* system is a gestural sound sampler that uses hand and head tracking to place and manipulate virtual audio sources in the user's environment, heard through bone conduction headphones which transmit sound directly to the cochlear without occluding the user's hearing. This allows the user to experience virtual audio environments overlaid seamlessly onto the real audio environment. Through gesture, the user can interact with and shape the combined real and virtual audio environment surrounding them.

This paper will address the research question: **"How can we better understand relationships between virtual and real environments through gesture and virtually placed AAR objects?"** Understanding these relationships allows for the facilitation of engaging experiences within computational arts through richer multisensory human-computer interaction.

1.2 Technologies

The three technologies used in *area~* are gestural hand tracking, rotational head tracking and ambisonics. The gestural hand tracker used in the system is a Leap Motion LM-010 Controller, a USB infrared camera device that provides location and orientation data output of individual finger joints (and therefore hands) when they are presented above the device. The Leap Motion Controller (LMC) has been adopted in a multitude of settings such as being mounted on VR headsets and converting hand gestures to MIDI.¹² UltraLeap are now investigating the use of this same technology with gesture-based public information screens to help combat the "hygiene risks of touch screens."¹³

Rotational head tracking is achieved via an inertial measurement unit (IMU). This small and inexpensive component provides orientational data output at 20 times a second. When affixed to the head via a headset or headphones, it is a relatively easy and cheap way of implementing head tracking into the system.

Ambisonics is an audio format that allows for full-spherical audio capture and playback, meaning that it includes sound sources above and below the listener as well as the conventional horizontal plane.¹⁴ There are four recorded channels (referred to as *A-Format*) that, unlike regular mono, stereo or surround sound formats, contain no information about the speakers that the signal should be delivered to. Rather, these channels can be encoded in such a way as to describe a three-dimensional field of sound referred to as *B-Format*, allowing the producer or artist to think in terms of sound sources in three dimensions rather than conventional speaker placement. *B-Format* can be decoded through 'virtual microphones,' any number of which can be placed within this three-dimensional sound field to provide standard channel outputs.

For example, in *area~*, I have used a RØDE Soundfield NTSF-1 microphone array comprised of 4 microphones. The *A-Format* output is encoded to *B-Format* by an audio plugin. A software library decodes the *B-Format* to two responsive, binaural, virtual audio output channels. This all occurs in real-time, so that the microphones inside the three-dimensional sound field rotate proportionally as the user moves their head, providing realistic changes to what is heard.

2. Literature Review

2.1 Augmented Reality

Augmented reality was first defined by Caudell and Mizell in the field of aeronautical research as technology used to “augment the *visual field* of the user with information necessary in the performance of the current task [emphasis added],” reflecting the ongoing tendency for AR to focus on augmenting the visual rather than taking a multi-sensory approach.¹⁵ This paper focused on increasing the efficiency of manufacturing workers using a head-mounted display (HMD) with registration systems such as head-position tracking. Further conceptualisation came a year later when Rosenberg concluded that overlaying audiovisual information in the form of “virtual fixtures” could increase teleoperator performance by reducing demands on “taxed sensory modalities” and aiding in a perceptual simplification of the workplace.¹⁶ In their seminal 1994 paper, Milgram and Kishino conceptualised the “virtuality continuum,” a spectrum of ‘mixed reality’ (MR) scenarios into which most, if not all, of today’s AR/VR applications and devices fit.¹⁷

From the late 1990s, AR has been defined as a technology that fits three specifications outlined by Azuma:¹⁸

Figure 1. Ronald T. Azuma, "Optical see-through HMD conceptual diagram," in "A Survey of Augmented Reality," *Presence: Teleoperators and Virtual Environments* 6, no. 4 (1997): 355-85.

1. **Combines real and virtual**
2. **Interactive in real-time**
3. **Registered in 3-D**

The paper also contains an extensive survey of existing AR applications and devices of the time, from industries as disparate as military aircraft, entertainment, and robot path-planning. Most of these applications use either HMD 'optical see-through' (Figure 1), or 'video see-through' - "combining a closed-view HMD with one or two head-mounted video cameras" - again reflecting the understanding of AR as a primarily visual technology.¹⁹

2.2 Multisensory AR

AR has seen a recent surge in development, not only through the release of new HMDs (optical see-through), but also through the introduction of (video see-through) mobile AR frameworks.²⁰ This has resulted in the increasing use of AR in disciplines from neuroscience to the arts.²¹ Most of these developments continue to approach VR via a conventional ocular-centric approach. However, following Lindeman, Noma and Goncalves de Barros' taxonomy of multi-sensory AR (the use of more than one sensory overlay), the concept of multi-sensory AR has become an increasingly popular topic of

discussion.²² Chevalier and Kiefer argue that AR should be seen as “inseparable from a multisensory ecosystem” rather than consisting only of a visual overlay, and that it is inhabited by “modes of sensing, modes of perceptual mediation, computational relationships between sensing and mediation, human participants, and their environment.”²³ Viewing AR from a fundamentally “experience-focused and conceptual perspective,” Schraffenberger identifies ocular-centrism as an issue in AR research and practice and suggests that new perspectives on the medium must be considered in order to “inspire and facilitate new and different forms of both AR and AR research,” one of which is the investigation into “multimodal and interactive environments.”²⁴

2.3 Examples of MSAR

Examples of experience-driven multisensory applications of AR include “Augmented Reality Flavours,” which details the creation of a pseudo-gustatory AR interface by using a head-mounted olfactory pump to inject smells into the users nose whilst they eat a plain cookie to create the illusion of a flavoured cookie.²⁵ This illusion is further enhanced via the visual alteration of the cookie through the HMD. An example of auditory focused AR is the sound installation study “Listening Mirrors,” in which bone conduction headphones are used as a sensory overlay in conjunction with wearable and sculpted feedback devices for audio augmented reality:

“These headphones transmit sound directly to the [cochlear], bypassing the outer ear and ear drum, and so do not intervene in natural hearing. This allows our system to mediate the sonic environment by creating a mix of the real sound environment and digital reprocessing of the same environment, collected through the phone microphone.”²⁶

Bone conduction headphones here serve as an aural equivalent to visual ‘optical see-through’ devices, i.e. “hear-through.”²⁷ Likewise, Tikander outlines the necessary development of the aural equivalent of visual ‘video see-through’ devices, i.e. “mic-through,” in his paper positing an augmented reality audio headset.²⁸ Somatosensory, or haptic AR interfaces have also been explored, with electrical muscle stimulation being used to add physical forces to mixed reality environments with results showing increased realism.²⁹

In a broader HCI context, multisensory interfacing is beginning to gain more traction, with increasing amounts of novel auditory,³⁰ haptic,³¹ olfactory, and gustatory applications offering rich multisensory experiences.³²

2.4 AR as a Medium in Computational Art

Computational art - the use of computational programming or software as a tool, performance aid, medium or collaborator to create art - has seen the use of AR as a medium since the early 2000s. In 2008, Grasset et al. presented case studies of AR art exhibitions, with the conclusion that for effective design of artworks, “the bond between the virtual and the real, established by the MR [mixed reality] concept should hold in all conditions the project could be used in.”³³

More recently, Papagiannis has drawn attention to a broader multisensory view of AR in attempting to understand an aesthetic for AR applications, suggesting “AR is beginning to expand in new ways, beyond visual frames and into the full human sensorium.”³⁴ In 2015, the Tate Sensorium multisensory art exhibition used novel mid-air haptic devices to augment visual art.³⁵ The results of the included study found that multisensory experiences lead to audiences finding artwork more emotionally engaging compared to solely visual experiences. Chevalier and Kiefer highlight the nascent use of newer AR technologies by artists. They argue that AR has significant potential for creative exploration, and that it is a medium for creating “new nuanced and fine-grained emergent aesthetic experiences” and define AR as “real-time computationally mediated perception.”³⁶

The use of AR as a medium for the composition of computational art raises the question of what the materiality, both for user and developer, of such a digital tool or piece of software would look, sound, feel, taste or smell like. As Papagiannis highlights, “understanding the capacities of the technology and its constraints to exploit the technology to artistic use by envisioning novel applications and approaches, and developing new aesthetics and conventions beyond previous traditional forms.”³⁷

3. The area~ System

The *area~* system, which stands loosely for ‘augmented reality environmental audio,’ aims to afford users the ability to spectromorphologically (defined by Smalley to concern spatial, temporal and textural qualities of sound) manipulate sounds from their environment into a *virtual audio environment*.³⁸ Through bone conduction headphones and head tracking, this sound field is heard in synchronicity with their actual environment. The system was created in order to explore and reveal the relationship between real and virtual environments.

3.1 Hardware Implementation

Figure 2. *area~* hardware (left) and *area~* PCB (right)

The on-desk hardware for the *area~* system shown in Figure 2 on the left includes (a) a laptop running the *area~* Max MSP patch, (b) a 4 channel input audio interface, (c) an Ambisonic microphone, and (d) a Leap Motion Controller.³⁹

The wearable hardware used for the *area~* system comprises 2 sections: (e) a belt pouch containing a PCB-mounted ESP32 microcontroller and 18650 Li-Ion cell (both shown in Figure 2 on the right),⁴⁰ and (f) a pair of bone conduction headphones, with (g) a mounted inertial measurement unit (IMU) for tracking head orientation.⁴¹

The IMU and ESP32 are connected via a detachable 1.5m heat shrunk cable that runs from the back of the bone conduction headphones, down the length of the user's back and into the belt-mounted transmitter pouch. The IMU data is transmitted to the laptop via Bluetooth from the ESP32 with the integration of an Arduino library.⁴² Audio is transmitted to the headphones via Bluetooth from the Max MSP patch running on the laptop.

The only hardware that needs to be accessible for the user is the Leap Motion Controller and the wearable hardware system. The laptop and audio interface are ideally hidden from the user. The microphone should be placed in a location that will provide the user with access to sounds that they wish to manipulate.

3.2 Software Implementation

Figure 3: *area~* Max MSP patch

The patch (Figure 3) uses the RØDE SoundField audio plugin shown in Figure 4 to encode the A-Format ambisonics microphone input into B-Format (a three-dimensional sound field), or what I will refer to as the *ambisonic palette*. This *ambisonic palette* is not heard by the user; instead, they can sculpt from it, forming their own audible (B-Format) *virtual audio environment* through hand gestures. I have defined these gestures in Max MSP with help from the IRCAM Leap Motion library; they occur over three stages of user interaction: **record**, **manipulate**, **spatialise**.⁴³

Figure 4. RØDE, “SoundField by RØDE Plugin,” accessed February 10, 2021, <https://www.ode.com/soundfieldplugin>.

1. **The recording or ‘sampling’ stage** is initiated by making a left-hand grab above the LMC. The longer lasting the grab, the longer the portion of audio from the *ambisonic palette* is sampled. The three-dimensional coordinates of the hand above the LMC correlates with the location of audio recorded (this is achieved by mapping the hand coordinates to a virtual microphone inside the *ambisonic palette*), essentially allowing the user to record sounds around their person in three dimensions. Upon letting go of the grab gesture, the sample plays on repeat (using the karma~ Library) through the bone conduction headphones, thus setting up the session’s *virtual audio environment*. [44](#)
2. **The manipulation stage** is automatically initiated after the ending of the previous grab gesture and uses translational (x, y, z) and rotational (roll, pitch) values from both hands when above the LMC. There are two audio effects being manipulated, with parameters from these effects mapped in different ways to the translation and rotation of the user’s hands.
 - The first effect is a band-pass filter which accentuates certain audio frequencies of the sample. The frequency, strength, and gain of the filter is determined by the parameter mappings detailed in Figure 5.

The second effect is a semi-random granular synthesiser. This selects and copies a section of the sample and deconstructs it into several hundred grains. The section of the sample granulated, and the individual grain duration is determined by the parameter mappings detailed in Figure 6.

Parameter	Mapping	Explanation
Granulised sample section starting position	<ul style="list-style-type: none"> • Right hand roll • Linear map from LMC right hand roll sensor range (scaled 0 – 1) to 0 – 100% position on sample 	Where in this sample should I start the granulation?
Granulised sample section ending position	<ul style="list-style-type: none"> • Right hand pitch • Quadratic map from LMC right hand pitch sensor range (scaled 1 – 0 – 1) to 0 – 100% position on sample 	Where in this sample should I end the granulation?
Randomised grain duration lower bound	<ul style="list-style-type: none"> • Left hand roll • Linear map from LMC left hand roll sensor range (scaled 0 – 1) to grain duration lower bound (0ms – entire sample section length) 	Lower bound for randomised grain duration.
Randomised grain duration upper bound	<ul style="list-style-type: none"> • Left hand pitch • Quadratic map from LMC left hand pitch sensor range (scaled 1 – 0 – 1) to grain duration upper bound (0ms – entire sample section length) 	Upper bound for the randomised grain duration.

Figure 6. Granular synthesiser parameter mappings.

When the

user decides to end manipulating the sample, they can do so by performing a grab with both hands. Once this happens, the band-pass filter and granular synthesis parameters are frozen for that sample.

3. **The spatialise stage** begins once the manipulation stage is ended by the user. The three-dimensional space above the LMC is mapped to the *virtual audio environment*, in which the user is currently listening to the sample that they have recorded. The user can use their right hand to move the sample around the *virtual audio environment*. For an example of the effect this has, moving the hand between the two extremes of the x-axis (left to right) results in hearing the sample move from ear to ear. The spatialise stage is ended by grabbing with the right hand.

Once the spatialise stage has ended, the user has the option to repeat the process 7 more times, allowing for the creation of a *virtual audio environment* comprised of up to 8 spatialised audio samples, or what I refer to as *nodes*.

As a side note, the audio signal arriving in the two conduction pads in the headphones are the signals of two equidistantly spaced virtual microphones inside the *B-Format virtual audio environment* that decode it into two channels, left and right. Further interaction comes from user head movement which, at all times, is mapped to the revolution of these two virtual microphones around the central point of the *virtual audio environment*. This means if there is a *node* playing to the left of the user, rotating the head 90° anticlockwise results in the *node* now sounding as if it is in front of the user's face. This is achieved via the ICST Ambisonics Library and is elaborated on in the sections 4.3.2 Experiences and 5.1 Audio Augmented Reality, but for now it is worth mentioning that it allows for immersion into a combined *real and virtual audio environment*. ⁴⁵

To summarise, the patch can be categorised into having two inputs: audio from the user's environment and hand gesture, and one output: the *virtual audio environment*. In the background, this audio input is decoded into the *ambisonic palette* (inaudible), which is acted on by the user's hands to form one audible output: the *virtual audio environment*, which is comprised of up to 8 *nodes*. Through the choice of sensory overlay (bone conduction) and integration of head tracking, this *virtual audio environment* is experienced synchronously with the user's real, multisensory environment.

4. Study

4.1 Autobiographical Design Method

Originally, the study was planned for late March and would involve several user interaction studies. However, due to the UK lockdown in response to the unfolding COVID-19 pandemic, this was postponed until later in the year. Instead, I investigated the system using an autobiographical design method, framing it as a hypothesis study to better understand relationships between virtual and real environments, in the hope of developing a practice-led method for creating and researching multisensory augmented reality experiences.

Autobiographical design is defined as “design research drawing on extensive, genuine usage by those creating or building the system.”⁴⁶ Neustaedter and Sengers define “genuine usage” here to mean that changes are “based on the true needs of the researchers, rather than them pretending to have needs expected of target users.”⁴⁷

Due to the lockdown, and therefore inability to conduct in-person user-tests, this research method was beneficial as I was spending large amounts of time with the system. Moreover, there are several suggested requirements of employing this research method that Neustaedter and Sengers highlight that are true of the *area~* system:

- The existence of a genuine usage of the system
- The system being already developed
- The ability for fast tinkering
- Record-keeping of the design process
- Long-term usage of the system⁴⁸

Furthermore, as AR technology moves towards being a component of future general personal computing, there is a need for first-person research methods that take into consideration the effects of prolonged system usage as well as the arising relationship between user and system.⁴⁹ Moreover, these methods have been found to be specifically relevant to wearable systems.⁵⁰ A disadvantage of using autobiographical design as a research method is its inability to establish generalisability (also the case with ethnography, case studies and participatory workshops), which is why I still intend to conduct wider usability studies in the future.

4.2 Design

The study was designed as a cycle, in order to promote fast tinkering, record-keeping, and long-term usage in line with autobiographical design guidelines.

1. Over three sessions, ideally during the same week, the system is used with a logbook at hand in order to facilitate record-keeping of hardware setup, *node* manipulation and completed *real and virtual audio environment* listening experience remarks.
2. After each session, the notes are formalised into a database and categorised as ‘user experience remarks’ or ‘improvement remarks’ (subject to increase in categories).
3. At the end of the week, the three sessions’ notes are summarised into a ‘check-in’ document, where user experience remarks are collated, and improvement remarks are further categorised into lists pertaining to the area of the system that needs improvement or change.
4. Those changes are then made to the system and the cycle restarts.

4.3 Results

4.3.1 Hardware Location

I have observed that an environment with a lower noise floor is desirable and have implemented a normaliser to deal with re-recording loud background / ambient noises during successive *node* recording. I found myself basing the choice of microphone placement on what I wanted the *virtual audio environment* to sound like. Since the system uses an ambisonic microphone, consideration of the spherical 360° field of the microphone would lead to a richer *ambisonic palette* and subsequent *virtual audio environment*.

Two sessions were based outdoors and involved natural sounds such as birds, trees and wind, as well as passing cars. One of the sessions was inside and took place at the same time my partner was on a Zoom call, and therefore the *ambisonic palette* was invariably based on her speech and my movement and action inside the room. I want to look into placing the microphone inside bushes, trees, etc. rather than in open spaces to explore aesthetically the *virtual audio environment* that arises from such placement. Furthermore, the relationship between *real and virtual audio environment* would be quite different. Through their ears, the user would be rooted in their position in the environment, but because of the inherent blend between hearing and bone conduction, they would simultaneously experience the sonic environment of the bush mediated by the Max MSP patch.

4.3.2 Experiences

The system takes considerable time to set up, especially when documenting with video, audio, and notes. This process could certainly be streamlined further. Overall, I was pleased with the sound quality; the microphone picks up the environment very clearly. However, I remarked that the manipulation stage could feature more interesting real-time auditory effects on the *nodes*. The blending of *real and virtual audio environment* is achieved well via the bone conduction headphones and there was a subconscious registration via the head-tracking that gave me the very real impression that there was a 3D environment of *nodes* around my body.

The IMU on the bone conduction headphones sometimes provides erratic and erroneous data, leading to accidental revolutions of the virtual audio environment around the head. Despite technical difficulties such as this, in one autobiographical design session, when I took the headphones off, I wrote that “I felt like I’d been disconnected from something” and that “my senses felt heightened before I took them off, not only to the *virtual audio environment*, but now more sensitive to the audio content of my real environment.”⁵¹

If I managed to capture an infrequent environment sound, such as a particular bird call, or a sentence of spoken word, the fact that the patch is set up to loop the samples gave a certain permanence to that otherwise impermanent sound. On multiple occasions I couldn't tell if the sounds of birds I was hearing originated from a *node* or from within a tree.

4.3.3 Arising Interactions

The maximum sample length is currently 28 seconds, but I have found myself mainly sticking to shorter loops, creating quite repetitive and rhythmic sequences. I remarked that grains from the synthesiser sounded like a permanent record of my gestures' effect on the environment. I liked the playfulness of being able to record a sound from a certain location in my *real audio environment* and place it in different location in my *virtual audio environment*.

Despite the wearable hardware being wireless, hand interaction with the *area~* system is inherently limited to being in range of the LMC (often placed on a table). I found myself wanting to be able to move around my environment whilst being able to record new *nodes* and hear existing *nodes* move relative to my body position.

Overall, the results from my autobiographical design method have shown that *area~* is an effective tool for examining the combinatorial relationship between real and virtual environments. Despite the system's hardware setup requiring some further work to allow for quicker start up and more accurate head tracking, it has provided me with novel aesthetic experience through:

- The blending of real and virtual auditory environments to create a third, augmented environment that was greater in experiential nature than the sum of its parts (not simply a combinatorial layering)
- The ability to spectromorphologically manipulate sounds in real-time in this third environment with the body
- The potential for creating believable illusions of real-world sound sources from these manipulated and spatialised virtual sounds.

5. Discussion

5.1 Audio Augmented Reality

In 2.1 *Augmented Reality*, I detailed the importance of the relationship between real and virtual environments in the definition of AR systems. As a medium for creating computational art, the three specifications outlined by Azuma still hold true and should be discussed in the creation of such art or the development of their interactive systems of composition e.g. *area~*.

1. Combines real and virtual

- Using bone conduction headphones means that the user's unmediated perception of the environment is allowed to continue without interference, allowing the effective combination of real and virtual environments.

2. Interactive in real-time

- The system is made interactive in real-time through the gestural mappings (both head and hand) to sound effect parameters, and the perception of those effects through the mediating sensory display (bone conduction headphones).

3. Registered in 3-D

- The *area~* system is registered in 3-D via the IMU data allowing for the mapping of head movement to *virtual audio environment*.
- The mapping of hand movement (in the recording stage) to the spatial location recorded in the *ambisonic palette* (detailed in 3.2 *Software Implementation*), also contributes to the notion that the *virtual audio environment* and *real audio environment* are aligned.

Due to the fact that our environment is perceived by our senses (visual, auditory, somatosensory, olfactory, gustatory, to name a few), there is a potential bandwidth issue if we constrain virtual environments to solely visual displays of information for the user — highlighted by Rosenberg when he mentions “taxed sensory modalities” in the workplace.⁵² Whilst Rosenberg's bandwidth issue is still pertinent as I am engaging the user in primarily their auditory sense, the experiences emerging from this initial autobiographical design study have confirmed that audio augmented reality is able to provide new aesthetic experiences of coexistent real and virtual environments through the use of sensory overlay technology, creatively mutable environments, and expressive input parameters (e.g., hand and head gesture).

This suggests that to fully realise the materiality and affordances of AR as a medium for computational art, including creative and functional uses of multiple senses in AR must be further

explored (see 6.1 *Devising a practice-led method*).

5.2 Interaction

The tables of parameter mappings shown in 3.2 *Software Implementation* may seem like a chaotic mess, however, time has been spent making these mappings intuitive. For example, the representation of volume to a vertical scale corresponds with findings that non-musicians are relatively competent at attributing the size of air-gestures to heightened musical dynamics.⁵³ As for the band-pass filter's horizontal pitch mappings, this is mainly based on the visual representation of frequencies on a horizontal scale often found in the user interface of such filters. Nevertheless, it is pertinent to mention that psychomusicological research finds a correlation between pitch representation and horizontal space i.e. lower pitch to the left, higher pitch to the right.⁵⁴ Although this is often attributed to the internalised representation of horizontal pitch on pianos by keyboard-playing musicians,⁵⁵ it has been found that this effect also propagates in non-musicians.⁵⁶

In contrast, interaction with the granular synthesiser is not intended to be intuitive; instead, I have opted to hide or black-box the interaction through a mix of linear and quadratic mappings on each hand. This is in order to stir curiosity in the user and induce play, as found in 4.3.3 *Arising Interactions*, but also to be tested in wider user studies, as mentioned earlier “I remarked that grains from the synthesiser sounded like a record of my gestures’ effect on the environment.”

Indeed, whilst outlining the material epistemologies of digital music instruments (DMIs), Magnusson describes black-boxed DMIs as containing “knowledge of its inventors, which means that the users of the instrument do not need to have a deep knowledge of its internal functions,” furthermore clarifying that there is a “continuous oscillation between a mode of conceptual (system design) engagement with the instrument and [an] embodied (performative) relationship with it.”⁵⁷ This ‘oscillation’ displays, in turn, an underlying synergy between DMI development and the autobiographical design process, perhaps due to the similarities in requirements of the processes outlined in 4.1 *Methodology*. This synergy has led to the use of ABD in the development of many DMIs and interactive music systems.⁵⁸

5.3 Computational Art

Visit the web version of this article to view interactive content.

Figure 7. Sam Bilbow, "Area~ 360° Video / Ambisonics Documentation," filmed 2020 in Brighton, United Kingdom, <https://youtu.be/SPd-f2EXulQ>. This is an interactive 360° YouTube video: click and drag on the screen to view the demonstration from different angles.

As a system for creating computational art in the form of in-situ AAR experiences, *area~*'s artistic output is firstly a real-time experience. In order to document these experiences of the system however, I have included the automated recording and saving of both the *ambisonic palette* (the ambisonic recording of the *real audio environment*), and the users *virtual audio environment* as separate *B-Format* .wav files in the project directory. These separate *B-Format* .wav files could be merged and decoded from *B-Format* to any number of speakers for a multi-channel installation of field recordings/manipulations of environments made with *area~*. This also leaves open the potential for *area~* to be used as a compositional tool, with possible applications in sound art and soundscape composition.

I also chose to merge and decode a set of these recordings to binaural ambisonic format, and have time-synced this recording with a 360° video and a screen recording of the patch that was taken during the system's use. This allows my performance of the *virtual audio environment* to be experienced second-hand with headphones. Dragging the screen in different directions with mouse/touchscreen emulates hearing the difference in environment if I moved my head in those directions. Optionally, an inexpensive smartphone VR headset (£5 - £15) can be used to heighten the interactivity of the experience. A screenshot of this is shown in Figure 8 and a link can be found in the references.⁵⁹ The potential for users' experiments with *area~* to be captured and re-experienced interactively could have interesting applications, effectively allowing users to explore each others' lived aural experience of the system.

6. Future Research

6.1 Devising a Practice-Led Method

The novel experiences that emerged from the results of the study into the uses of AR as a medium for computational art have encouraged me to devise a practice-led method for researching multisensory AR (MSAR). It begins with the ideation of an MSAR Experience (which describes a possible human-to-sense interaction). These are classed as Snippets, Scenes, and Spaces depending on the number of senses engaged,

sensory intensity, and interaction size. The hardware and software that enable the experience are classed as MSAR Instruments and MSAR Environments respectively. Instruments are categorised into Wearable, Tangible and Situated Instruments. This taxonomy will eventually aid not only in the creation, but in the evaluation of future MSAR Experiences.

6.2 Project North Star

Figure 8. Participant using a Project North Star demo to resize virtual cheese.

As mentioned in 4.3.2 *Arising Interactions*, the ability to move around whilst recording *nodes* could address the rigidity of the hand gestures being tethered to the Leap Motion Controller, which was often placed on a table, but is inherently connected to the laptop. One way of achieving this would be changing the wearable portion of the hardware. Instead of the IMU paired with bone conduction headphones, I could move towards the use of an HMD paired with the bone conduction headphones. This would most likely be the open-source AR HMD design by Leap Motion: Project North Star, for which there is a large community of makers.⁶⁰ The HMD integrates the same LMC (hand tracking), as well as an Intel RealSense T261 for inertial measurement data (body position tracking, not just head).

The headset being used with the Project Esky software toolkit to resize virtual objects in a demo is shown in Figure 9.⁶¹ The headset requires a computer to power it, although a portable compute pack for the HMD is in development by CombineReality.⁶² Regarding *area~*, the use of body position tracking rather than just head orientation tracking, and the move from a desk-placed LMC to head-mounted, opens avenues for different and potentially more interactive gestures i.e., allowing for the re-manipulation of, or interacting with existing *nodes* in a larger environment. Rather than feeling like they were rooted to an instrument, the system would move towards being more of an extension of the performer, and with this, new aesthetic experiences of real and virtual space would be explored, for

example enabling the recording of previously inaccessible environments: in the city centre, or by the beach, where the user wouldn't be able to set up a table.

6.3 Multisensory *area(s)~*

In broader terms of my practice-led method, Project North Star provides an open-source Wearable Instrument, and through preliminary engagements, has shown that it is an extremely malleable platform for creating multisensory AR Experiences, as well as mounting further Wearable MSAR Instruments.

I envisage the future of *area~* to not only include visualisations of nodes, but haptic interactions with *nodes*. I see no reason why olfactory displays of *nodes* couldn't be integrated too, either via North Star mounted injection, or in zones of interaction in a multisensory *area~* installation / performance space. This would allow for a more "modalities-encompassing" approach to understanding nuances in the relationship between the real and virtual environments of AR when it is used as a medium for expressive computational art.⁶³

Acknowledgements

This work was funded by the University of Sussex Leverhulme Trust Doctoral Scholarship Programme.

Bibliography

Aftershokz. "Aftershockz Aeropex." Accessed February 10, 2021.

<https://aftershokz.co.uk/products/aeropex>.

Apple. "ARKit." Accessed February 10, 2021. <https://developer.apple.com/augmented-reality/arkit/>.

Azuma, Ronald T. "A Survey of Augmented Reality." *Presence: Teleoperators and Virtual Environments* 6, no. 4 (1997): 355–85.

Barde, Amit. "Design Considerations for a Wearable, Bi-Modal Interface." PhD diss., University of Canterbury, 2018.

Billinghurst, Mark, Adrian Clark, and Gun Lee. "A Survey of Augmented Reality." *Foundations and Trends in Human-Computer Interaction*, 8, no. 2-3 (March 2015): 73-272.

Brooks, Jas, Steven Nagels, and Pedro Lopes. "Trigeminal-Based Temperature Illusions." In *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-12. New York: Association for Computing Machinery, 2020.

Caramiaux, Baptiste, Frédéric Bevilacqua, and Norbert Schnell. "Towards a Gesture-Sound Cross-Modal Analysis." In *Gesture in Embodied Communication and Human-Computer Interaction: 8th International Gesture Workshop*, 158-70. Berlin: Springer, 2010.

Caudell, Thomas, and David Mizell. "Augmented Reality: An Application of Heads-up Display Technology to Manual Manufacturing Processes." In *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, 659-69. Kauai, HI, USA: Institute of Electrical and Electronics Engineers, 1992.

Cecchinato, Marta E., Anna L. Cox, and Jon Bird. "Always On(Line)?: User Experience of Smartwatches and Their Role within Multi-Device Ecologies." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3557-68. Denver: Association for Computing Machinery, 2017.

Chevalier, Cécile, and Chris Kiefer. "Towards New Modes of Collective Musical Expression through Audio Augmented Reality." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 25-28. Virginia, USA: Association for Computing Machinery, 2018.

Chevalier, Cécile, and Chris Kiefer. "What Does Augmented Reality Mean as a Medium of Expression for Computational Artists?" *Leonardo* 53, no. 3 (February 2019): 1-9.

CombineReality (@CombineReality). "Check Out This Homebrew Portable #ProjectNorthstarRig." Twitter post, April 21, 2020. <https://twitter.com/CombineReality/status/1252638433870143488>.

Constanzo, Rodrigo. "Tool: Karma~ (Sampler/Looper External) | Cycling '74." Last modified 14 May 2015. <https://cycling74.com/tools/karma-samplerlooper-external>.

Cycling '74. "Max MSP." Accessed February 10, 2021. <https://cycling74.com/products/max>.

Desjardins, Audrey, and Aubree Ball. "Revealing Tensions in Autobiographical Design in HCI." In *Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18*, 753-64. Hong Kong, China: ACM Press, 2018.

Dey, Arindam, Mark Billinghurst, Robert W. Lindeman, and J. Edward Swan. "A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014." *Frontiers in Robotics and AI* 5 (April

2018): 37.

Eliasson, Olafur. "Wunderkammer." Accessed February 11, 2021.

<https://olafureliasson.net/press/wunderkammer>.

Espressif. "ESP32." Accessed February 10, 2021.

<https://www.espressif.com/en/products/socs/esp32/overview>.

Gerzon, Michael. "Periphony: With-Height Sound Reproduction." *Journal of The Audio Engineering Society* 21, no. 1 (1973): 2–10.

Godøy, Rolf Inge, Egil Haga, and Alexander Refsum Jensenius. "Playing 'Air Instruments': Mimicry of Sound-Producing Gestures by Novices and Experts." In *Gesture in Human-Computer Interaction and Simulation: 6th International Gesture Workshop*, edited by Sylvie Gibet, Nicolas Courty, and Jean-François Kamp, 256–267. Berlin: Springer-Verlag Berlin Heidelberg, 2006.

Google. "ARCore." Accessed February 10, 2021. <https://developers.google.com/ar>.

Grasset, Raphaël, Eric Woods, and Mark Billinghurst. 2008. "Art and Mixed Reality: New Technology for Seamless Merging Between Virtual and Real." *IMedia-Space* 1: 10.

Hutmacher, Fabian. "Why Is There So Much More Research on Vision Than on Any Other Sensory Modality?" *Frontiers in Psychology* 10 (October 2019): 2246.

Kiefer, Chris, Dan Overholt, and Alice Eldridge. "Shaping the Behaviour of Feedback Instruments with Complexity-Controlled Gain Dynamics." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 343–348. Birmingham, UK: NIME, 2020.

Leap Motion. "Geco MIDI – Leap Motion Gallery." Accessed February 10, 2021. <https://gallery.leapmotion.com/geco-midi/>.

Leap Motion. "HTC Vive Setup." Accessed February 10, 2021. <https://developer.leapmotion.com/vr-setup/vive>.

Leap Motion. "Project North Star." Accessed February 10, 2021. <https://developer.leapmotion.com/northstar>.

Lidji, Pascale, Régine Kolinsky, Aliette Lochy, and José Morais. "Spatial Associations for Musical Stimuli: A Piano in the Head?" *Journal of Experimental Psychology: Human Perception and Performance* 33, no. 5 (2007): 1189–1207.

Lindeman, Robert W., and Haruo Noma. 2007. "A Classification Scheme for Multi-Sensory Augmented Reality." In *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology - VRST '07*, 175. Newport Beach, CA: Association for Computing Machinery.

Lindeman, Robert W., Haruo Noma, and Paulo Goncalves de Barros. "An Empirical Study of Hear-Through Augmented Reality: Using Bone Conduction to Deliver Spatialized Audio." In *2008 IEEE Virtual Reality Conference*, 35-42. Reno, USA: Institute of Electrical and Electronics Engineers, 2008.

Lopes, Pedro, Sijing You, Alexandra Ion, and Patrick Baudisch. "Adding Force Feedback to Mixed Reality Experiences and Games Using Electrical Muscle Stimulation." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1-13. Montreal: Association for Computing Machinery, 2018.

Lopes, Pedro, Sijing You, Lung-Pan Cheng, Sebastian Marwecki, and Patrick Baudisch. "Providing Haptics to Walls & Heavy Objects in Virtual Reality by Means of Electrical Muscle Stimulation." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1471-82. Denver: Association for Computing Machinery, 2017.

Maggioni, Emanuela, Erika Agostinelli, and Marianna Obrist. "Measuring the Added Value of Haptic Feedback." In *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, 1-6. Erfurt: Institute of Electrical and Electronics Engineers, 2017.

Magic Leap. "Magic Leap 1." Accessed February 10, 2021. <https://www.magicleap.com/magic-leap-1>.

Magnusson, Thor. "Of Epistemic Tools: Musical Instruments as Cognitive Extensions." *Organised Sound* 14, no. 2 (2009): 168-76.

Martin, Charles P. "Percussionist-Centred Design for Touchscreen Digital Musical Instruments." *Contemporary Music Review* 36, no. 1-2 (2017): 64-85.

McDuff, Daniel, Christophe Hurter, and Mar Gonzalez-Franco. "Pulse and Vital Sign Measurement in Mixed Reality Using a HoloLens." In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, 1-9. Gothenburg Sweden: Association for Computing Machinery, 2017.

Microsoft. "Microsoft HoloLens." Accessed February 10, 2021. <https://www.microsoft.com/en-us/hololens>.

Milgram, Paul, and Fumio Kishino. "A Taxonomy of Mixed Reality Visual Displays." *IEICE Transactions on Information and Systems* E77-D, no. 12 (1994): 8.

Narumi, Takuji, Shinya Nishizaka, Takashi Kajinami, Tomohiro Tanikawa, and Michitaka Hirose. "Augmented Reality Flavors: Gustatory Display Based on Edible Marker and Cross-Modal Interaction." In *CHI '11: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*, 93. Vancouver: Association for Computing Machinery, 2011.

Neustaedter, Carman, and Phoebe Sengers. "Autobiographical Design in HCI Research: Designing and Learning through Use-It-Yourself." In *DIS '12: Proceedings of the Designing Interactive Systems Conference*, 514–523. Newcastle Upon Tyne: Association for Computing Machinery, 2012.

Nreal. "Nreal Light." Accessed February 10, 2020. <https://nreal.ai/product/>.

Oculus. "Oculus Quest 2." Accessed February 11, 2020. <https://www.oculus.com/quest-2/>.

Oxford Reference. 2020. "Ocularcentrism." Accessed February 9, 2021.

Papagiannis, Helen. "The Critical Role of Artists in Advancing Augmented Reality." In *The Next Step: Exponential Life*, 124–39. BBVA-Open Mind, 2017.

Papagiannis, Helen. "Working towards Defining an Aesthetics of Augmented Reality: A Medium in Transition." *Convergence: The International Journal of Research into New Media Technologies* 20, no. 1 (2014): 33–40.

RØDE. "SoundField by RØDE Plugin." Accessed February 10, 2021. <https://www.ode.com/soundfieldplugin>.

Rompapas, Damien Constantine, Daniel Flores Quiros, Charlton Rodda, Bryan Christopher Brown, Noah Benjamin Zerk, and Alvaro Cassinelli. "Project Esky: Enabling High Fidelity Augmented Reality on an Open Source Platform." In *ISS '20: Proceedings of the 2020 ACM International Conference on Interactive Surfaces and Spaces*, 61–63. Lisbon, Portugal: Association for Computing Machinery, 2020.

Rosenberg, Louis. "Virtual Fixtures: Perceptual Tools for Telerobotic Manipulation." In *Proceedings of IEEE Virtual Reality Annual International Symposium*, 76–82. Seattle: Institute of Electrical and Electronics Engineers, 1993.

Rusconi, Elena, Bonnie Kwan, Bruno Giordano, Carlo Umiltà, and Brian Butterworth. "Spatial Representation of Pitch Height: The SMARC Effect." *Cognition* 99, no. 2 (2006): 113–29.

Schacher, Jan, and Philippe Kocher. "Ambisonics Spatialization Tools for Max/MSP." In *Proceedings of the 2006 International Computer Music Conference*, 274–77. New Orleans: Michigan Publishing, 2006. http://decoy.iki.fi/dsound/ambisonic/motherlode/source/ICST_Ambisonics_ICMC2006.pdf.

Schraffenberger, Hanna. "Arguably Augmented Reality: Relationships between the Virtual and the Real." PhD diss., University of Leiden, 2018.

Schraffenberger, Hanna, and Edwin van der Heide. "Multimodal Augmented Reality: The Norm Rather than the Exception." In *Proceedings of the 2016 Workshop on Multimodal Virtual and Augmented Reality - MVAR '16*, 1-6. Tokyo, Japan: Association for Computing Machinery, 2016.

Seah, Sue Ann, Marianna Obrist, Anne Roudaut, and Sriram Subramanian. "Need for Touch in Human Space Exploration: Towards the Design of a Morphing Haptic Glove - ExoSkin." In *Human-Computer Interaction - INTERACT 2015*, edited by Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler, 18-36. Cham: Springer International Publishing, 2015.

Sheffield, Eric, Edgar Berdahl, and Andrew Pfalz. "The Haptic Capstans: Rotational Force Feedback for Music Using a FireFader Derivative Device." In *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2. Brisbane: NIME, 2016.

Smalley, Denis. "Spectromorphology: Explaining Sound-Shapes." *Organised Sound* 2, no. 2 (1997): 107-26.

Sound Music Movement Interaction - ISMM. "Leap Motion Skeletal Tracking in Max." . Last modified November 7, 2014. <http://ismm.ircam.fr/leapmotion/>.

Spence, Charles, and Jozef Youssef. "Olfactory Dining: Designing for the Dominant Sense." *Flavour* 4, no. 1 (2015): 32.

Tikander, Miikka, Matti Karjalainen, and Ville Riikonen. "An Augmented Reality Audio Headset." In *Proceedings of the 11th International Conference on Digital Audio Effects*, 4. Espoo, Finland: DAFX, 2008.

Timmers, Renee, and Shen Li. "Representation of Pitch in Horizontal Space and Its Dependence on Musical and Instrumental Experience." *Psychomusicology: Music, Mind, and Brain* 26, no. 2 (2016): 139-48.

Turchet, Luca. "Smart Mandolin: Autobiographical Design, Implementation, Use Cases, and Lessons Learned." In *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion - AM'18*, 1-7. Wrexham, UK: Association for Computing Machinery, 2018.

UltraLeap. "Leap Motion Controller." Accessed February 10, 2021. <https://www.ultraleap.com/product/leap-motion-controller/>.

UltraLeap. "Touchscreens: Business as Usual Isn't an Option." Accessed February 9, 2021. <https://www.ultraleap.com/company/news/blog/contactless-technology/>.

Unander-Scharin, Carl, Åsa Unander-Scharin, and Kristina Höök. "The Vocal Chorder: Empowering Opera Singers with a Large Interactive Instrument." In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*, 1001–10. Toronto: Association for Computing Machinery, 2014.

Valve. "Valve Index." Accessed February 9, 2021. <https://store.steampowered.com/valveindex>.

Vi, Chi Thanh, Damien Ablart, Elia Gatti, Carlos Velasco, and Marianna Obrist. "Not Just Seeing, but Also Feeling Art: Mid-Air Haptic Experiences Integrated in a Multisensory Art Exhibition." *International Journal of Human-Computer Studies* 108 (December 2017): 1–14.

Vi, Chi Thanh, Asier Marzo, Damien Ablart, Gianluca Memoli, Sriram Subramanian, Bruce Drinkwater, and Marianna Obrist. "TastyFloats: A Contactless Food Delivery System." In *Proceedings of the Interactive Surfaces and Spaces - ISS '17*, 161–70. Brighton: ACM Press, 2017.

Vuforia. "Getting Started." Accessed February 14, 2021. <https://library.vuforia.com/getting-started/overview.html>.

Weis, Tina, Barbara Estner, Cees van Leeuwen, and Thomas Lachmann. "SNARC (Spatial–Numerical Association of Response Codes) Meets SPARC (Spatial–Pitch Association of Response Codes): Automaticity and Interdependency in Compatibility Effects." *Quarterly Journal of Experimental Psychology* 69, no. 7 (2016): 1366–83.

Wikipedia. "Ambisonics." Accessed February 10, 2021. <https://en.wikipedia.org/w/index.php?title=Ambisonics&oldid=953930554>.

Winer, Kris. "ESP32 Arduino Library." Accessed February 10, 2021. <https://github.com/kriswiner/ESP32>.

Media Cited

Bilbow, Sam. "Area~ 360° Video / Ambisonics Documentation." YouTube video, 2:37. July 17, 2020. Accessed February 10, 2021. [\[2020\] area~ 360° video / ambisonics documentation](#).

Biography

Sam Bilbow is a Brighton-based creative coder and 2nd year PhD student, exploring computational art and multisensory augmented / mixed reality technologies. He completed his undergraduate in Music Technology in 2015 with First Class Honours, specialising in hardware hacking and instrument design. He then went on to complete his MA in Music and Sonic Media in 2019 with Distinction, further specialising in augmented reality and spatial audio composition / design. Sam is a recipient of The Leverhulme Trust Doctoral Scholarship Programme at Sussex University titled “From Sensation to Perception & Awareness” for his PhD project: “Impact on human perception and expression, using augmented-reality technology as a medium for installation art.”

Footnotes

1. Ronald T. Azuma, “A Survey of Augmented Reality,” *Presence: Teleoperators and Virtual Environments* 6, no. 4 (1997): 356. [↵](#)

2.

See Mark Billinghurst, Adrian Clark, and Gun Lee, “A Survey of Augmented Reality,” *Foundations and Trends in Human-Computer Interaction* 8, no. 2-3 (March 2015): 73-272;

Hanna Schraffenberger and Edwin van der Heide, “Multimodal Augmented Reality: The Norm Rather than the Exception,” in *Proceedings of the 2016 Workshop on Multimodal Virtual and Augmented Reality - MVAR '16* (Tokyo, Japan: Association for Computing Machinery, 2016), 1-6. [↵](#)

3. Mark Billinghurst et al., “A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014,” *Frontiers in Robotics and AI* 5 (April 2018): 37. [↵](#)

4.

“Ocularcentrism,” Oxford Reference, accessed February 8, 2021,

[https://www.oxfordreference.com/view/10.1093/oi/authority.20110803100245338#:~:text=Quick%20Reference,be%20the%20opposite%20of%20phonocentrism\);](https://www.oxfordreference.com/view/10.1093/oi/authority.20110803100245338#:~:text=Quick%20Reference,be%20the%20opposite%20of%20phonocentrism);)

Fabian Huttmacher, “Why Is There So Much More Research on Vision Than on Any Other Sensory Modality?” *Frontiers in Psychology* 10 (October 2019): 2246. [↵](#)

5. *Ibid.* [↵](#)

6. *Ibid.* [↵](#)

7.

See Billinghamurst, Clark, and Lee, “A Survey of Augmented Reality”;

Schraffenberger and van der Heide 2016, “Multimodal Augmented Reality”;

Helen Papagiannis, “Working towards Defining an Aesthetics of Augmented Reality: A Medium in Transition,” *Convergence: The International Journal of Research into New Media Technologies* 20, no. 1 (2014): 33-40;

Cécile Chevalier and Kris Kiefer, “Towards New Modes of Collective Musical Expression through Audio Augmented Reality,” in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Virginia, USA: Association for Computing Machinery, 2018), 25–28. [↵](#)

8.

Leap Motion, “HTC Vive Setup,” accessed February 10, 2021, <https://developer.leapmotion.com/vr-setup/vive/>;

Oculus, “Oculus Quest 2,” accessed February 11, 2020, <https://www.oculus.com/quest-2/>;

Valve, “Valve Index,” accessed February 9, 2021, <https://store.steampowered.com/valveindex>. [↵](#)

9.

Magic Leap, “Magic Leap 1,” accessed February 10, 2021, <https://www.magicleap.com/magic-leap-1>;

Nreal, “Nreal Light,” accessed February 10, 2020, <https://nreal.ai/product/>;

Microsoft, “Microsoft HoloLens,” accessed February 10, 2021. <https://www.microsoft.com/en-us/hololens>. [↵](#)

10.

Leap Motion, “Project North Star,” accessed February 10, 2021, <https://developer.leapmotion.com/northstar/>;

Leap Motion is a hand tracking device company that recently merged with Ultrahaptics. [↵](#)

11. Damien Constantine Rompapas et al., “Project Esky: Enabling High Fidelity Augmented Reality on an Open Source Platform,” in *ISS ’20: Proceedings of the 2020 ACM International Conference on Interactive Surfaces and Spaces* (Lisbon: Association for Computing Machinery, 2020), 61-63. [↵](#)

12.

For VR headsets see Leap Motion, “HTC Vive Setup,” accessed February 10, 2021, <https://developer.leapmotion.com/vr-setup/vive>.

For converting hand gestures to MIDI see Leap Motion, “Geco MIDI – Leap Motion Gallery,” accessed February 10, 2021, <https://gallery.leapmotion.com/geco-midi/>. ↵

13. “Touchscreens: Business as Usual Isn’t an Option,” Ultraleap, accessed February 8, 2021, <https://www.ultraleap.com/company/news/blog/contactless-technology/>. ↵

14. Michael Gerzon, “Periphony: With-Height Sound Reproduction,” *Journal of The Audio Engineering Society* 21, no. 1 (1973): 2. ↵

15. Thomas Caudell and David Mizell, “Augmented Reality: An Application of Heads-up Display Technology to Manual Manufacturing Processes,” in *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences* (Kauai, USA: Institute of Electrical and Electronics Engineers, 1992), 660. ↵

16. Louis Rosenberg, “Virtual Fixtures: Perceptual Tools for Telerobotic Manipulation,” in *Proceedings of IEEE Virtual Reality Annual International Symposium* (Seattle: Institute of Electrical and Electronics Engineers, 1993), 81. ↵

17. Paul Milgram and Fumio Kishino, “A Taxonomy of Mixed Reality Visual Displays,” *IEICE Transactions on Information and Systems* E77-D, no. 12 (1994): 8. ↵

18. Ronald T. Azuma, “A Survey of Augmented Reality,” *Presence: Teleoperators and Virtual Environments* 6, no. 4 (1997): 356. ↵

19. *Ibid.*, 362. ↵

20.

For recent examples see:

Apple, “ARKit,” accessed February 10, 2021, <https://developer.apple.com/augmented-reality/arkit/>;

Google, “ARCore,” accessed February 10, 2021, <https://developers.google.com/ar>. ↵

21.

For an example of the use of AR in neuroscience, see:

Daniel McDuff, Christophe Hurter, and Mar Gonzalez-Franco, “Pulse and Vital Sign Measurement in Mixed Reality Using a HoloLens,” in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology* (Gothenburg: Association for Computing Machinery), 1–9.

For an example of the use of AR in the arts, see:

Eliasson, Olafur. “Wunderkammer,” accessed February 11, 2021.

<https://olafureliasson.net/press/wunderkammer.>

22.

See Robert W. Lindeman, Haruo Noma, and Paulo Goncalves de Barros, “An Empirical Study of Hear-Through Augmented Reality: Using Bone Conduction to Deliver Spatialized Audio,” in *2008 IEEE Virtual Reality Conference* (Reno, USA: Institute of Electrical and Electronics Engineers, 2008), 35–42;

Billinghurst, Clark, and Lee, “A Survey”;

Schraffenberger and van der Heide, “Multimodal Augmented Reality”;

Papagiannis, “Working towards Defining an Aesthetics”. [↵](#)

23. Cécile Chevalier and Chris Kiefer, “What Does Augmented Reality Mean as a Medium of Expression for Computational Artists?” *Leonardo* 53, no. 3 (February 2019): 7. [↵](#)

24. Hanna Schraffenberger, “Arguably Augmented Reality: Relationships between the Virtual and the Real” (PhD diss., University of Leiden, 2018), 7. [↵](#)

25. Takuji Narumi et al., “Augmented Reality Flavors: Gustatory Display Based on Edible Marker and Cross-Modal Interaction,” in *CHI '11: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems* (Vancouver: ACM Press, 2011). [↵](#)

26. Chris Kiefer and Cécile Chevalier, “Towards New Modes of Collective Musical Expression through Audio Augmented Reality,” in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Virginia, USA: Association for Computing Machinery, 2018), 26. [↵](#)

27. Lindeman and Goncalves de Barros, “An Empirical Study of Hear-Through AR,” 35. [↵](#)

28.

Tikander, Miikka, Matti Karjalainen, and Ville Riikonen, "An Augmented Reality Audio Headset," in *Proceedings of the 11th International Conference on Digital Audio Effects* (Espoo, Finland: DAFX, 2008), 4.

The phrase "mic-through" is from Lindeman, Noma, and de Barros, 35.

└

29. Pedro Lopes et al., "Providing Haptics to Walls & Heavy Objects in Virtual Reality by Means of Electrical Muscle Stimulation," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver: Association for Computing Machinery, 2017), 1471–82. └

30. Amit Barde, "Design Considerations for a Wearable, Bi-Modal Interface," (PhD diss., University of Canterbury, 2018), 77. └

31.

See Lopes et al., 1471–82; Eric Sheffield, Edgar Berdahl, and Andrew Pfalz, "The Haptic Capstans: Rotational Force Feedback for Music Using a FireFader Derivative Device," in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Brisbane: NIME, 2016), 2;

Emanuela Maggioni, Erika Agostinelli, and Marianna Obrist. "Measuring the Added Value of Haptic Feedback." In *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)* (Erfurt: Institute of Electrical and Electronics Engineers, 2017), 1–6;

Sue Ann Seah et al., "Need for Touch in Human Space Exploration: Towards the Design of a Morphing Haptic Glove – ExoSkin," in *Human-Computer Interaction – INTERACT 2015* (Cham: Springer International Publishing, 2015), 18–36;

Chi Thanh Vi et al., "Not Just Seeing, but Also Feeling Art: Mid-Air Haptic Experiences Integrated in a Multisensory Art Exhibition," *International Journal of Human-Computer Studies* 108 (December 2017): 1–14. └

32.

Jas Brooks, Steven Nagels, and Pedro Lopes, “Trigeminal-Based Temperature Illusions,” in *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York: Association for Computing Machinery, 2020), 1-12;

See Charles Spence and Jozef Youssef, “Olfactory Dining: Designing for the Dominant Sense,” *Flavour* 4, no. 1 (2015): 32;

Chi Thanh Vi et al., “TastyFloats: A Contactless Food Delivery System,” in *Proceedings of the Interactive Surfaces and Spaces - ISS '17* (Brighton: ACM Press, 2017), 161-70.

↵

33. Raphaël Grasset, Eric Woods, and Mark Billinghurst, “Art and Mixed Reality: New Technology for Seamless Merging Between Virtual and Real,” *IMedia-Space* 1 (2008): 8. ↵

34. Papagiannis, “Working Towards Defining an Aesthetics,” 39. ↵

35. Chi Thanh Vi et al., “Not Just Seeing, but Also Feeling Art: Mid-Air Haptic Experiences Integrated in a Multisensory Art Exhibition,” *International Journal of Human-Computer Studies* 108 (December 2017): 1-14. ↵

36. Chevalier and Kiefer, “Towards New Modes of Collective Musical Expression,” 7. ↵

37. Helen Papagiannis, “The Critical Role of Artists in Advancing Augmented Reality,” in *The Next Step: Exponential Life* (BBVA-OpenMind, 2017), 127. ↵

38. Denis Smalley, “Spectromorphology: Explaining Sound-Shapes,” *Organised Sound* 2, no. 2 (1997): 107-26. ↵

39.

For more information see Cycling '74, “Max MSP,” accessed February 10, 2021, <https://cycling74.com/products/max>;

Wikipedia, “Ambisonics,” accessed February 10, 2021, <https://en.wikipedia.org/w/index.php?title=Ambisonics&oldid=953930554>;

UltraLeap, “Leap Motion Controller,” accessed February 10, 2021, <https://www.ultraleap.com/product/leap-motion-controller/>. ↵

40. Espressif, “ESP32,” accessed February 10, 2021, <https://www.espressif.com/en/products/socs/esp32/overview>. ↵

41. Aftershokz, “Aftershockz Aeropex,” accessed February 10, 2021, <https://aftershokz.co.uk/products/aeropex>. ↵
42. Kris Winer, “ESP32 Arduino Library,” accessed February 10, 2021, <https://github.com/kriswiner/ESP32>. ↵
43. Sound Music Movement Interaction – ISMM, “Leap Motion Skeletal Tracking in Max,” last modified November 7, 2014, <http://ismm.ircam.fr/leapmotion/>. ↵
44. Rodrigo Constanzo, “Tool: Karma~ (Sampler/Looper External) | Cycling ’74,” last modified May 14, 2015, <https://cycling74.com/tools/karma-samplerlooper-external>. ↵
45. Jan Schacher and Philippe Kocher, “Ambisonics Spatialization Tools for Max/MSP,” in *Proceedings of the 2006 International Computer Music Conference* (New Orleans: Michigan Publishing, 2006), 274–77. ↵
46. Carman Neustaedter and Phoebe Sengers, “Autobiographical Design in HCI Research: Designing and Learning through Use-It-Yourself,” in *DIS ’12: Proceedings of the Designing Interactive Systems Conference* (Newcastle Upon Tyne: Association for Computing Machinery, 2012), 514. ↵
47. *Ibid.* ↵
48. *Ibid.*, 514–522. ↵
49. Audrey Desjardins and Aubree Ball, “Revealing Tensions in Autobiographical Design in HCI,” in *Proceedings of the 2018 Designing Interactive Systems Conference 2018 - DIS ’18* (Hong Kong: ACM Press, 2018), 753–764. ↵
50. Marta E. Cecchinato, Anna L. Cox, and Jon Bird, “Always On(Line)?: User Experience of Smartwatches and Their Role within Multi-Device Ecologies,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver: Association for Computing Machinery, 2017), 3557–3568. ↵
51. From author’s personal notes (unpublished). ↵
52. Rosenberg, “Virtual Fixtures,” 81. ↵
- 53.

Rolf Inge Godøy, Egil Haga, and Alexander Refsum Jensenius, "Playing 'Air Instruments': Mimicry of Sound-Producing Gestures by Novices and Experts," in *Gesture in Human-Computer Interaction and Simulation: 6th International Gesture Workshop*, ed. Sylvie Gibet, Nicolas Courty, and Jean-François Kamp (Berlin: Springer-Verlag Berlin Heidelberg, 2006), 256–67;

Baptiste Caramiaux, Frédéric Bevilacqua, and Norbert Schnell, "Towards a Gesture-Sound Cross-Modal Analysis," in *Gesture in Embodied Communication and Human-Computer Interaction: 8th International Gesture Workshop* (Berlin: Springer-Verlag Berlin Heidelberg, 2010), 158–70. [↵](#)

54. Renee Timmers, and Shen Li, "Representation of Pitch in Horizontal Space and Its Dependence on Musical and Instrumental Experience," *Psychomusicology: Music, Mind, and Brain* 26, no. 2 (2016): 139–48. [↵](#)

55.

See Pascale Lidji et al., "Spatial Associations for Musical Stimuli: A Piano in the Head?" *Journal of Experimental Psychology: Human Perception and Performance* 33, no. 5 (2007): 1189–1207;

Elena Rusconi et al., "Spatial Representation of Pitch Height: The SMARC Effect," *Cognition* 99, no. 2 (2006): 113–29.

[↵](#)

56.

Tina Weis et al., "SNARC (Spatial-Numerical Association of Response Codes) Meets SPARC (Spatial-Pitch Association of Response Codes): Automaticity and Interdependency in Compatibility Effects," *Quarterly Journal of Experimental Psychology* 69, vol. 7 (2016): 1366–83;

Godøy, Haga, and Jensenius, "Playing 'Air Instruments'," 256–267. [↵](#)

57. Thor Magnusson, "Of Epistemic Tools: Musical Instruments as Cognitive Extensions," *Organised Sound* 14, no. 2 (2009): 171 [↵](#)

58.

See Chris Kiefer, Dan Overholt, and Alice Eldridge, “Shaping the Behaviour of Feedback Instruments with Complexity-Controlled Gain Dynamics,” in *Proceedings of the International Conference on New Interfaces for Musical Expression* (Birmingham: NIME, 2020), 343-348;

Charles P. Martin, “Percussionist-Centred Design for Touchscreen Digital Musical Instruments,” *Contemporary Music Review* 36, no. 1-2 (2017): 64-85;

Luca Turchet, “Smart Mandolin: Autobiographical Design, Implementation, Use Cases, and Lessons Learned,” in *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion - AM'18* (Wrexham, UK: Association for Computing Machinery, 2018), 1-7;

Carl Unander-Scharin, Åsa Unander-Scharin, and Kristina Höök, “The Vocal Chorder: Empowering Opera Singers with a Large Interactive Instrument,” in *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14* (Toronto: Association for Computing Machinery, 2014), 1001-10. [↵](#)

59. Sam Bilbow, “Area~ 360° Video / Ambisonics Documentation,” July 17, 2020, accessed February 10, 2021, [\[2020\] area~ 360° video / ambisonics documentation](#). [↵](#)

60. See “Project North Star,” accessed February 10, 2021, <https://developer.leapmotion.com/northstar>. [↵](#)

61. Rompapas et al., “Project Esky,” 61-63. [↵](#)

62. CombineReality (@CombineReality), “Check Out This Homebrew Portable #ProjectNorthstarRig,” Twitter post, April 21, 2020, <https://twitter.com/CombineReality/status/1252638433870143488>. [↵](#)

63. Schraffenberger and van der Heide, “Multimodal Augmented Reality: The Norm Rather than the Exception,” 1. [↵](#)