Towards a Taxonomy for Immersive Music Performance

Dr Kyriacos Michael

Senior Lecturer in Songwriting and Music Composition at the School of Creative Arts, University of Hertfordshire, UK

Keywords

Introduction

In the context of music performance, the term 'immersive' is largely undefined, posing challenges for creators seeking to develop works that incorporate immersive characteristics. By synthesizing various theories and identifying effective practices, this paper seeks to provide a taxonomy of immersive characteristics for the music performance medium. Thus, the taxonomy aims to inform potential practitioners, and how its application may lead to more positively perceived immersive performances. The following characteristics have been identified as key developmental components of immersivity in music performance: *proximity*, *envelopment*, *sound* and *visual processing*, as well as *audience engagement.* This list is not exhaustive, but it offers a set of key components that can be employed to produce distinctive and meaningful immersive experiences for audiences.

Contextualization

Immersive music practice as a performance medium has a long history, one that has been continually developed and shaped by new ideas and technologies. As this paper is primarily focused on the practice of immersivity within the live music environment, this brief history explores two key components that contribute to how sound is perceived by listeners in that medium - which in turn, will highlight some of the key characteristics further explored in the taxonomy. The first is the manner in which composers have used space, such as the

arrangement of musicians and audiences to take advantage of our spatial hearing capacity. The second relates to the development of sound-reproduction systems together with our improved understanding of human hearing, with the aim to develop sound source localization and spatialization techniques.

Some composers have recognized that the music is intrinsically connected to the space in which it is delivered, influencing the material and the way it is perceived by audiences. Thus, innovative means have been employed that deviate from the traditional front–back format, such as unusual arrangements of the performers and audience, and spatialization techniques which include the movement of musicians.

An example of this practice are antiphonal works (call-and-response performances), which used separated choirs in the space going as far back as the chanting of psalms in biblical times[1]. Composers such as Berlioz, in his *Requiem* (1837), recognized that the use of the space could be advantageous by placing four large brass sections at the four cardinal points, thus composing with the space in mind; something he referred to as 'architectural music'.[2] In *The Unanswered Question* (1908), Charles Ives placed three distinct instrumental layers (strings, woodwind and brass) with distinct compositional roles in different sections of the hall to enhance dynamism. Henry Brant, a spatial music pioneer, regularly separated performers and instrumental groups with the aim to create timbral and textural differentiation for compositional clarity, or in his words 'to make complexity intelligible'.[3]

In *Millennium II* (1954), Brant arranged the conductor, brass and percussion on stage with a single voice placed up on a balcony, with brass along the walls moving in a programmed manner – this is an early example of spatialization techniques.

---

[1] Richard Zvonar, 'A History of Spatial Music' (2000), eContact!, https://econtact.ca/7_4/zvonar_spatialmusic.html, accessed 12 September 2024.

[2] Enda Bates, 'The Composition and Performance of Spatial Music' (PhD diss., Trinity College Dublin, 2009), 117. Available at
www.tara.tcd.i.e./bitstream/handle/2262/77011/Bates,%20Enda%20The%20Composition.pdf?sequence=1, accessed 12 September 2024.

[3] Maria Anna Harley, 'An American in Space: Henry Brant's "Spatial Music"'. American Music, 15.1 (1997), 70–92 (75).
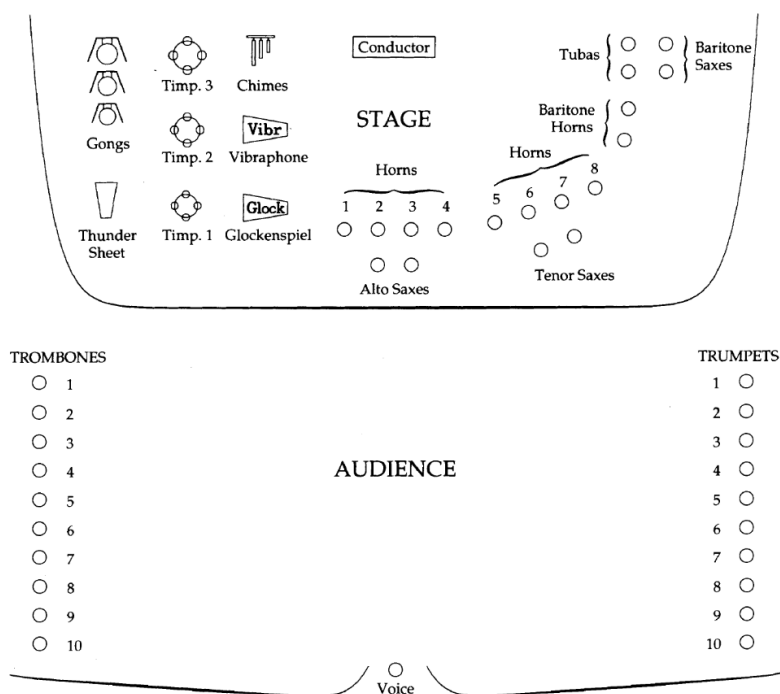
Figure 1 Spatial arrangement of musicians in *Millennium II* (1954), from Harley, 'An American in Space', 76

By 1951 developments in *musique concrete* by the composer Pierre Schaeffer allowed for the projection of recorded material. Loudspeakers in a cross formation (front, left, right and back) with a supplementary fifth speaker placed overhead were spatialized using a diffuser – this is an early example of constructing a three-dimensional acoustic space. Stockhausen advanced immersive music practices such as in *Gesang der Jünglinge* (1956), which incorporated synthesized and natural elements projected through five sets of loudspeakers, four sets surrounding the audience and one on stage. Such innovative arrangement techniques were further experimented within his work *Gruppen* (1955–57), in which three orchestras and three conductors are placed around the audience, and *Carré* (1959–60) which placed four orchestras in a square around the audience. *Sirius* (1975–77) goes a step further, using a square auditorium with the audience facing the centre, performed through an eight-speaker array using diffusion of electronic elements to exaggerate the spatial movement of the projected sounds, in addition to four live soloists (trumpet, soprano, bass clarinet, and bass) placed high on opposite sides of the auditorium.
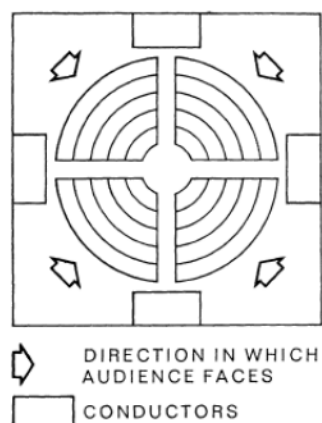
Figure 2 Stockhausen, *Carré* (1959–60), from Bates, 'Composition and Performance of Spatial Music', 136

The continuing development of this field can be evidenced in many notable works from the 1960s onwards. *Terretektorh* (1965–66) by Xenakis featured 88 musicians in a circular arrangement with the audience distributed among them, while *Genesis* (1962 - 63) by Górecki arranges the orchestra in different geometric shapes, and then repositions the audience between each movement.[4]
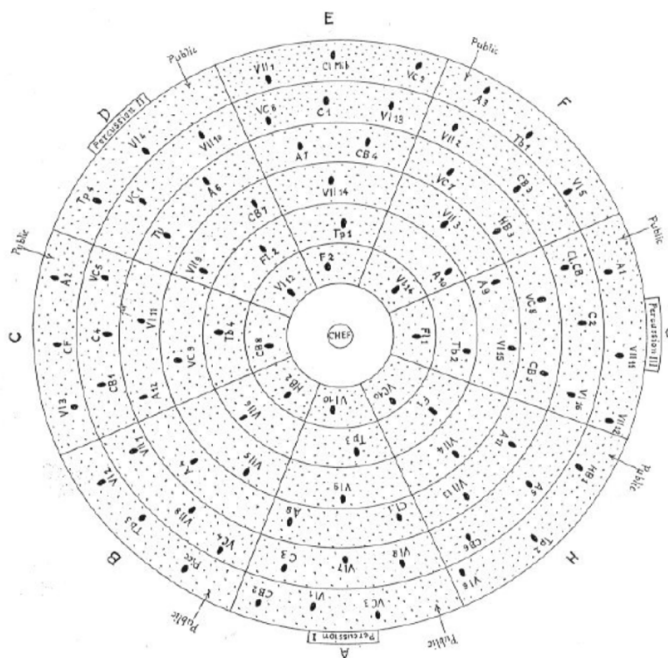

Figure 3 Xenakis, *Terretektorh* (1965–66), from Bates, 'Composition and Performance of Spatial Music', 142

More methods by various composers employed four, six and eight channel speaker arrays along

---

[4] Danuta Mirka, 'Górecki's "Musica Geometrica"', The Musical Quarterly, 87.2 (2004) 305–32.

with a mixture of projected recorded material and live performers. A culmination of many of the techniques discussed were integrated by Pierre Boulez in his grandiose piece *Répons* (1986). This highly ambitious work of electroacoustic music places a 24-piece orchestra in the centre of the auditorium with the audience in the round, surrounding them. Six satellite soloists and six loudspeakers are placed above the round, closer to the auditorium walls surrounding the audience and orchestra.[5] Boulez attempted to create real-time spatialization with the six soloists captured electronically and projected in specific trajectories by the loudspeakers. The effect meant that the audience would find it difficult to localize the paths of the individual sounds, whilst creating separation between the central orchestra and the soloists.
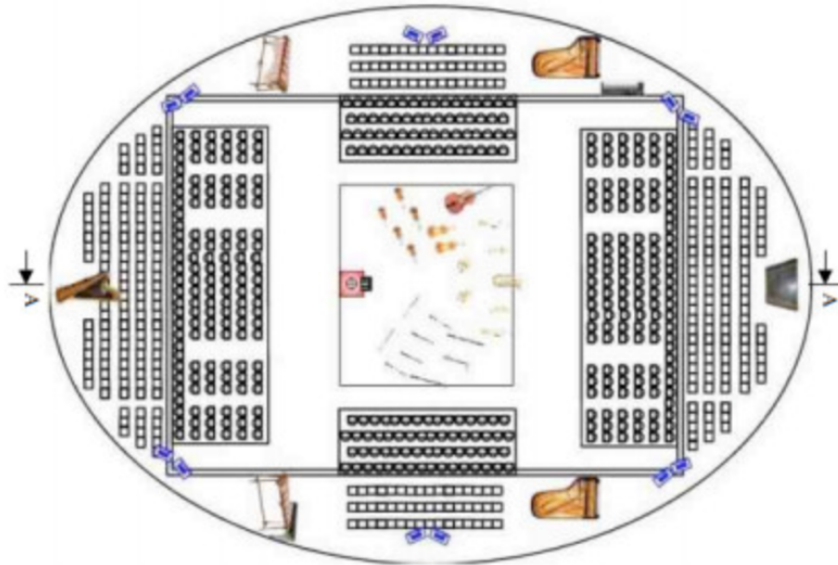

Figure 4 Arrangement of *Répons* by Boulez (1985), from Vagne, 'Pierre Boulez – Répons', vagnethierry.fr

Technological developments allowed composers to create even more immersive soundscapes using larger speaker-arrays and techniques such as ambisonics for better control of sound-sources within the 3D soundfield. Natasha Barrett's *Microclimates III–VI* (2007) is performed over a 3D 16-speaker array, with eight loudspeakers in a central ring around the audience, and four additional speakers above and below.[6]

---

[5] Thierry Vagne, 'Pierre Boulez – Répons', Musique Classique & Co., 10 March 2013, https://vagnethierry.fr/pierre-boulez-repons/, accessed 12 July 2021.

[6] Natasha Barrett, 'Microclimates III–VI' (n.d.), www.natashabarrett.org/mc3-7cp.html, accessed 8 August 2018.
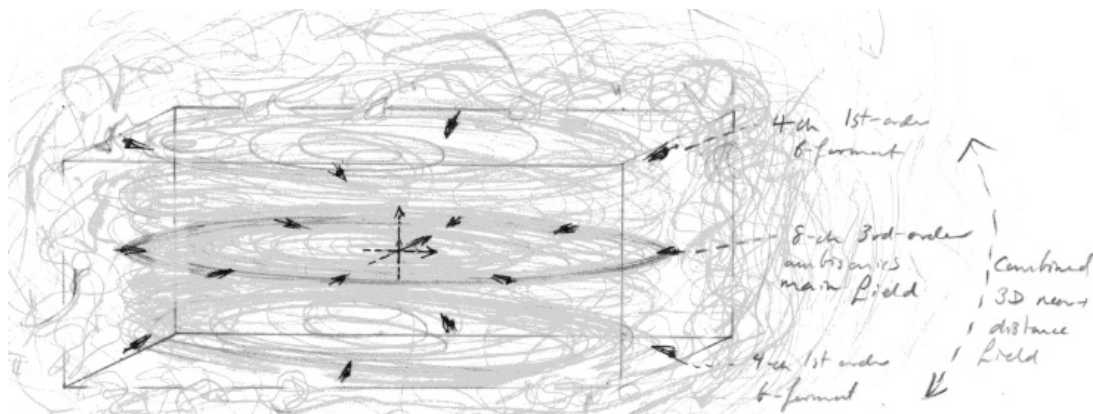
Figure 5 Loudspeaker arrangement for *Microclimates III–VI* (Barrett, n.d.)

Bjork's concept tour *Cornucopia* (2019) presented a 50-person choir, flutes, harp, various percussion instruments, electronics and a reverberation chamber; all realized on a layered stage with a 360-degree sound system, with the aim to create a unique multimedia event that combined music performance, a 3D sound environment, theatre and visuals. The choir would leave the stage and move along the audience to create an evolving soundscape as their voices reverberated throughout the performance space.[7] Beyond the physical space and the arrangement of musicians and performers within it, the single most influential aspect in the development of immersive music is technology. Many of the innovative approaches explored by pioneering composers and artists discussed have to some degree been influenced by new technologies – Bjork's *Cornucopia* is just one such example.

Stereo remains the main commercial system of sound reproduction since Alan Blumlein patented two-channel stereophonic sound in 1931. From HiFi systems to headphones and mobile technology, it remains the standard playback system since its wider use in the 1950s. Stereophony and its expansion into surround-sound systems (including those for cinema) are key developmental stages that need discussing, because they play a key role in our understanding of localization and spatialization, which become central to the application of immersivity in music performance. Stereophony's popularity is in its ability to recreate the illusion of spatial characteristics in just two or more loudspeakers using phantom imaging, where the apparent location of the sound source depends on its panning, amplitude and depth across the available loudspeaker system. This is because the signal emanating from each speaker is heard by each ear with some delay (Figure 7).[8] Composers and producers are able to control this image, building on the human ear's ability to distinguish localization and create

---

[7] ProSound Web, 'Realising a Vision: Deploying Immersive Sound for Bjork & "Cornucopia" in NYC', ProSoundWeb, 12 August 2019, www.prosoundweb.com/realizing-a-vision-deploying-immersive-sound-for-bjork-cornucopia-in-nyc/2/, accessed 12 September 2024.

[8] Francis Rumsey, Spatial Audio (Routledge, 2001), chap. 3: 'Two Channel Stereo and Binaural Audio'.

spatial understanding.[9] Spatial characteristics can be further enhanced by utilizing bigger loudspeaker systems such as quadrophonic, 5.1, octophonic, and much larger loudspeaker systems.
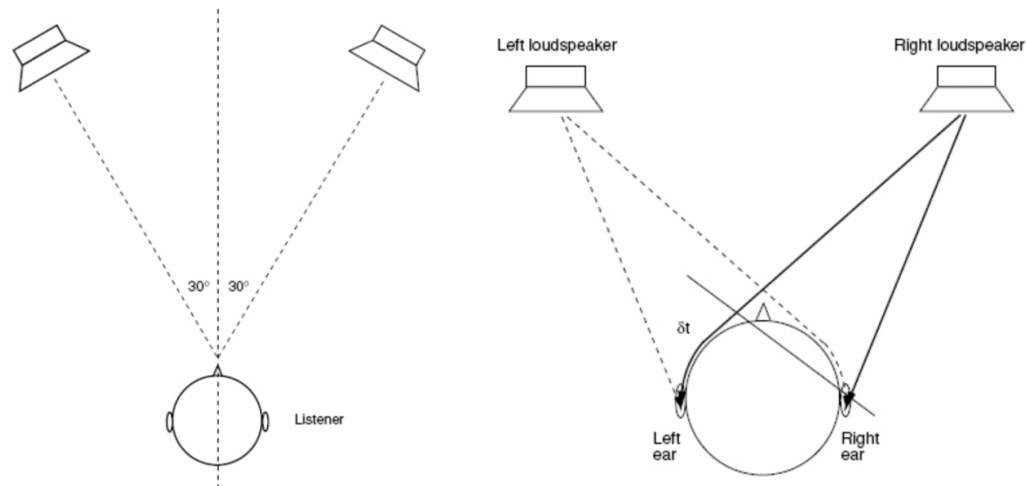

Figure 6 sweet spot   and   **Figure 7** phantom imaging.
From Rumsey, *Spatial Audio*, chap. 3

Stereophony is reliant on listening position, and works best in the 'sweet-spot'. In two-channel loudspeaker systems this is an equilateral triangle with the listener at a 30° angle from each speaker (Figure 6). Due to stereophony's reliance on listening position, the use of larger loudspeaker systems can affect the quality of the sound image emanated. Quadraphonic has been largely unsuccessful because the 90° angle creates instability in the perceived image.[10] 5.1-channel surround had established itself as a serious contender, but the sound suffers on the sides and particularly the back, where the 140° angle at the rear creates a less stable image (Figure 8). Thus, it has not fared well in the field of spatial music, where a more robust spatial image is required. This issue can be slightly mitigated by using a 7.1 surround-sound system, which adds two more channels at the centre-left (CL) and centre-right (CR) positions.

---

[9] Mike Wozniewski, Zack Settel and Jeremy.R Cooperstock, 'A Framework for Immersive Spatial Audio Performance', Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME06), Paris, France, 144–49, https://doi.org/10.5281/zenodo.1177021.

[10] David Malham, 'Spatial Hearing Mechanisms and Sound Reproduction', Semantic Scholar, (2002), www.semanticscholar.org/paper/Spatial-Hearing-Mechanisms-and-Sound-Reproduction-Malham/273df87e05595cd9cd5d512f1cbbe73a29878839, accessed 12 September 2024.
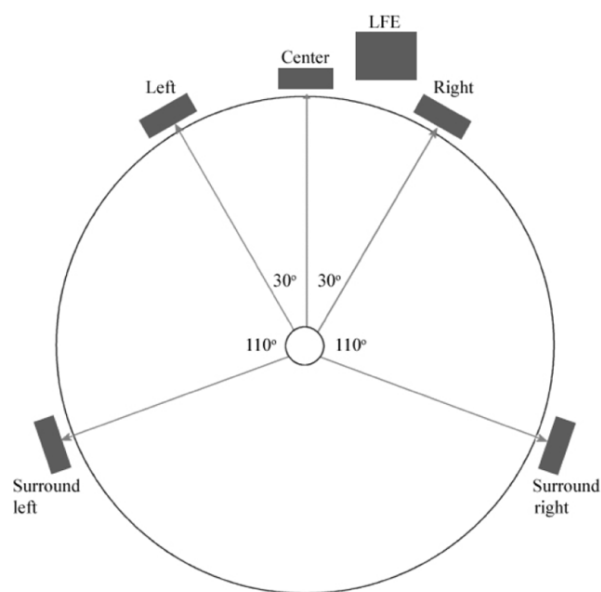
Figure 8 5.1 surround sound (Bates, 'Composition and Performance of Spatial Music', 37)

A 6-speaker hexagonal array with the minimum 60˚ separation angle is required for a much more reliable image. However, the octophonic loudspeaker array, expressed by Jonty Harrison as the 'main eight',[11] is the most common configuration, where evenly spaced speakers at 45°, placed uniformly around the audience creates a higher quality spatial image.[12] Spatial works with even richer content, such as those in diffusion works, require even larger loudspeaker systems. These enhance the listener's ability to detect the localization of audio signals because the signals can be spatially separated, and therefore help listeners clarify complex content, especially pitch and timbre.[13]

Although the octophonic array is reliable and fairly accessible, it still does not provide sound cues on the vertical plane. The 10.2 system designed for cinema, attempted to incorporate vertical sound imaging by adding two height channels at the Front Left and Front Right. An additional Centre Rear channel was installed to reduce the hole at the back from the original

[11] Adam Stansbie (Stanovic), 'The Acousmatic Musical Performance: An Ontological Investigation' (PhD diss., City University London, 2013), 48.

[12] Robert Dow, 'Multi-Channel Sound in Spatially Rich Acousmatic Composition', 4th Conference 'Understanding and Creating Music', Caserta, November 2004, 23–26, http://decoy.iki.fi/dsound/ambisonic/motherlode/source/rdow-multichannelsound.pdf, accessed 12 September 2024.

[13] Maria Anna Harley, 'Spatiality of Sound and Stream Segregation in Twentieth Century Instrumental Music', Organised Sound, 3.2 (1998), 147–66.

standard 5.1 model, and a 2nd LFE channel, one on each side for greater lateral separation.[14] A 22.2 system goes further in attempting to create a fully enveloping sound space; 10 middle layer channels, 5 at the front for a strong frontal image, 9 upper layer channels with a centre channel facing downwards, and 5 lower layer channels, that includes the two sub-channels (Figure 9).[15]
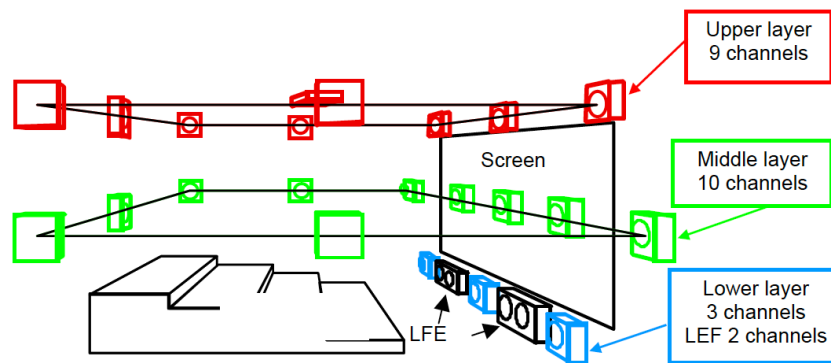

Figure 9 22.2 surround sound (Hamasaki, et al., '5.1 and 22.2 Multichannel Sound Productions', p. 383).

To create a full 360° sound-field and fully envelop the listener, a higher number of loudspeakers need to be deployed. These formats provide the composer with more scope to 'realize spatial detail and differentiation'.[16] Hence, there is a trend in these fields for more Permanent High Density Speaker Array's, such as the IEM-Cube at the Institute of Electronic Music and Acoustics that features a 24-speaker array; the *Motion Lab* at Oslo University featuring a 47-speaker array; the highly influential *Espace de projection* by IRCAM that features 75-speakers; the Acousmonium in Paris; and the BEAST system at Birmingham University, founded by Jonty Harrison in 1982, which can boast up to 100 speakers. These complex loudspeaker systems have afforded composers the tools to create complex immersive environments delivered effectively to all audience members, regardless of listening position. Ambisonics, a technique first introduced in the 1970s by Michael Gerzon, allows spatial audio information to be recorded and stored so that it can be accurately reproduced as a 3D soundfield. Ambisonics technique produces a much wider optimum listening position, meaning the audience is enveloped in a sonic experience that is not as compromised by their position

---

[14] Agnieszka Roginska and Paul Geluso, Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio (New York: Routledge, 2017), chap. 6.

[15] Kimio Hamasaki, et al., '5.1 and 22.2 Multichannel Sound Productions Using an Integrated Surround Sound Panning System', Audio Engineering Society Convention 117, Tokyo, October 2004, 383–387, at 383. https://aes2.org/publications/elibrary-page/?id=12883, accessed 12 September 2024.

[16] Dow, 'Multi-Channel Sound in Spatially Rich Acousmatic Composition', 1.

and movement.[17] Ambisonics allows the composer to choose the number of speakers based on their intentions; therefore, proving a flexible format which does 'not specify a particular loudspeaker array, neither in terms of quantity nor placement'.[18]

Many of the sound systems discussed are primarily based on channel-based audio which delivers specified signals to a desired number of loudspeakers. A disadvantage of channel-based audio is that it requires the content to be reproduced on an identical system as the one on which it was generated, unless time-consuming conversation is applied.[19] The means in which we now consume sound and music are so varied, ranging from mobile phones, laptops, home hi-fi's, soundbars, headphones, to complex cinema systems, means that the channel-based audio platform is problematic when trying to reproduce immersive audio. The *object-based audio* approach has demonstrated many advantages, and in its versatility seeks to be not only a solution to this spatial audio problem, but potentially the future of consumer listening.

Object-Based Audio (OBA) is a broad term that refers to the production and delivery of sound based on audio objects'.[20] Object-based audio is 'agnostic of actual reproduction loudspeakers setup', as the coded information renders the sound-object's position for custom playback, and therefore breaks the bond between the production method and the reproduction setup.[21] Audio objects in the production stage are coded with metadata that describe their position in the 3D soundfield and any time-specific position trajectories.[22] The key aspect of this technology is that the renderer translates the audio scene for the best possible listening experience on the available monitoring system. Simply put, object-based audio technologies allow content creators to place a collection of sounds within a soundfield that can accurately reproduce immersive audio for a variety of consumer-friendly products. Many OBA technologies are already in circulation, such as Dolby Atmos, DTS-X, Auro 3D and Sony 360 Reality Audio,

---

[17] Thomas Mckenzie, Damian Murphy and Gavin Kearny, 'Diffuse-Field Equalisation of First-Order Ambisonics', Proceedings of the 20th International Conference on Digital Audio Effects (DAFx -17), Edinburgh, September 2017, www.dafx17.eca.ed.ac.uk/papers/DAFx17_paper_31.pdf, accessed 12 September 2024.

[18] F. E. Henriksen, 'Space in Electroacoustic Music: Composition, Performance and Perception of Musical Space' (PhD diss., City University London, 2002), 88. Available at https://openaccess.city.ac.uk/id/eprint/7653/, accessed 12 September 2024.

[19] Schuyler Quackenbush and Jürgen Herre, 'Mpeg Standards for Compressed Representation of Immersive Audio', in Proceedings of the IEEE, 109.9 (2021), 1578–89.

[20] Christian Simon, Mateo Torcoli and Jouni Paulus, 'MPEG-H Audio for Improving Accessibility in Broadcasting and Streaming', Audio and Speech Processing (2019), 1–11 (3), doi.org/10.48550/arXiv.1909.11549.

[21] Jurgen Herre, et al., 'MPEG-H 3D audio – The new standard for coding of immersive spatial audio', IEEE Journal of Selected Topics in Signal Processing, 9.5 (2015), 770–779, at 771, https://doi.org/10.1109/JSTSP.2015.2411578.

[22] Philip Coleman, et al., 'An Audio-Visual System for Object-Based Audio: From Recording to Listening', IEEE Transactions on Multimedia, 20.8 (2018), 1919–31, doi:10.1109/TMM.2018.2794780.

to name a few. Although these technologies have primarily featured in film and game audio, they will be fundamental to the development of immersive music performance, particularly when they are creatively used to complement live performers. However, development will be sluggish until more permanent enveloping sound-systems become available in performance spaces.

From this very brief history of immersive music performance and the technological advances that have influenced its development, we can infer some key characteristics of the practice. Immersive works seek to engage more of the audience's senses, challenge common listening formats by experimenting with performance delivery arrangements within a given space, and make the audience feel they are more connected and 'inside the sound'.[23] Immersive music can be a medium that surrounds the listener, with greater emphasis on the spatial qualities of the music and the potential of movement to create unique sonic experiences.[24] Where possible, it is a matter of enveloping the listener to the extent that they feel entirely immersed in sound.

Spatial communication is also a significant factor, placing importance on how visual presentation plays a pivotal role in the way that audiences perceive the work, from a performative and architectural perspective. Thus, the proximity and physical presence of performers, instruments, and other sounding sources to audience members greatly influences the musical experience, fostering an intimacy that makes the sound clearer and larger whilst enhancing engagement.[25] Denis Smalley would argue that the presence of physical action due to our historic and cultural conditioning to expect it and apply it to make sense of our physical world, is of key importance.[26] What is clear from previous works is the need for experimentation and exploration to challenge norms, such as placing the audience in new listening environments whilst simultaneously using technological developments with inventive means. Nonetheless, the audience is always central, and it could be argued that these pioneering works and their approaches have not remained relevant in the popular sphere which lacks an audience-centred attitude.[27] What is further striking, however, is that there are as yet no well-established definitions or strategic approaches in the field of immersive music performance. This paper's primary objective is to consolidate the various theories and practices currently

---

[23] Simon Emmerson, Living Electronic Music (Aldershot: Ashgate, 2007), 103.

[24] William Moylan, 'Considering Space in Recorded Music', Journal on the Art of Record Production (2009), https://www.arpjournal.com/asarpwp/considering-space-in-music/, accessed 12 September 2024.

[25] Natasha Barrett and Marta Crispino, 'The Impact of 3-D Sound Spatialisation on Listeners' Understanding of Human Agency in Acousmatic Music', Journal of New Music Research, 47.5 (2018), 399–415.

[26] Denis Smalley, 'Spectromorphology: Explaining Sound-Shapes', Organised Sound, 2.2 (1997), 107–26.

[27] Angela M Beeching, 'Who is Audience?', Arts and Humanities in Higher Education, 15.3–4, (2016), 395–400, doi.org/10.1177/1474022216647390.

available into potentially usable characteristics for this field. The following taxonomy will therefore endeavour to do just that, by listing key immersive characteristics and provide methodologies that can be applied in practice.

Taxonomy

1. Proximity: The proximity of a listener to sound-sources.

This characteristic refers to the proximity of a listener to any sound-sources, which includes all performing musicians (voices and instruments), and any loudspeaker reproduction systems. The listener's closeness to the sound source enhances intimacy and inclusion, which can produce meaningful qualities from the audience and performers perspective, also known as 'spatial communication'.[28] The sense of 'closeness and involvement with a performance' can help to improve emotional connection to the music, most likely due to our ability to 'discern detail'; inversely, the further we are from sound-sources, the more detached we may feel from it.[29] This can be experienced in very large performance spaces where the distance between performers and the audience is augmented.[30]

Proximity can make the audience feel more connected to the sound and musical content. When you can 'sense body heat and/or cannot avoid contact, (it) may increase listeners' receptivity to intimacy in the music'.[31] There is a connection with the performance that happens beyond listening and falls into the physical realm. The sense of intimacy arises when the audience is able to connect physically and emotionally to the performer, where we can feel the power of their instrument or voice and immediacy of their actions. Similarly, proximity can be produced through other types of sound sources such as the physical presence of loudspeaker systems. Closeness can be understood spatially (physical proximity) and temporally (happening right now), whilst immediacy can imply 'being involved at some level'.[32] The meaning of the projected sound can be decoded much more quickly, whilst the proximate nature ensures better fidelity of the sound, which includes attack and tone before it is coloured by the reverberant

---

[28] Henriksen, 'Space in Electroacoustic Music', 122.

[29] D. Cabrera, Andy Nguyen and Young-Ji Choi, 'Auditory Versus Visual Spatial Impression: A Study of Two Auditoria', Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia, July 2004, 1, http://hdl.handle.net/1853/50760, accessed 12 September 2024.

[30] Larry Austin, 'Sound Diffusion in Composition and Performance: An Interview with Denis Smalley', Computer Music Journal, 24.2 (2000), 10–21, https://doi.org/10.1162/014892600559272.

[31] Henriksen, 'Space in Electroacoustic Music', (121).

[32] Constantin Popp, 'Portfolio of Original Compositions', (PhD diss., University of Manchester, 2014), 1–86, here 9. Available at https://pure.manchester.ac.uk/ws/portalfiles/portal/54553798/FULL_TEXT.PDF, accessed 12 September 2024.

space. Proximity forms an entry point for the listener to the performed material, and hence can increase engagement. In addition to timbral detail, proximity allows us to hear sounds as larger, while the visual aspect helps audiences process and give meaning to performance characteristics.[33]

Proximity as a compositional tool can be dependent on musical style.[34] A large orchestra and choir may sound better from a distance, in which the power of the collective sound is more balanced as it benefits from the reverberant colouration of a larger space. A small folk group however, may better suit a small intimate space, in which the independent gestures of musicians can be better appreciated. This is likely due to the fact that sounds connected to human activity (gesture) are key to expressing closeness and immediacy.[35] When the distance of audience members from sound sources is augmented, it is more likely to create an unbalanced sonic image, impact loudness and reduce intimacy; consequently, there is the loss of detail in the timbral architecture of the instrument and voice inherent in the delivery.[36]

The traditional front–back performance model creates a separation between the performed music (from the stage) and the audience member (in the auditorium), constructing distinct territories with a particular function.[37] This separation is further exaggerated as performance spaces increase in size, where a physical and tangible artform is transformed into a remote and detached experience. In this case, the spectacle changes from the intimate to the grand, reducing proximity and, thus, immersion. The physical space of a performance is established by the physical presence of a live performer, and that human presence in a live performance environment encourages the audience to acknowledge the physical space.[38]

The employment of proximity in a music performance is a form of spatial communication that can produce an intimate and inclusive environment. It can enhance emotional connection to the material on a temporal level due to heighted localization, physical presence and timbral clarity.

---

[33] Moylan, 'Considering Space in Recorded Music'.

[34] Denis Smalley, 'Space-Form and the Acousmatic Image', Organised Sound, 12.1 (2007), 35–58.

[35] Popp, 'Portfolio of Original Compositions'.

[36] Austin, 'Sound Diffusion in Composition and Performance'.

[37] Guy Harries, 'The Electroacoustic and its Double: Duality and Dramaturgy in Live Performance' (PhD diss., City University London, 2011), 108. Available at https://openaccess.city.ac.uk/id/eprint/1118/, accessed 12 September 2024.

[38] Harries, 'The Electroacoustic and its Double'.

2. Envelopment: The practice of surrounding the listener with sound.

The term envelopment refers to the sound fully surrounding the listener, both on the horizontal and vertical planes – a three-dimensional audio environment.[39] The term 'space' can also be referred to as the 'soundfield', i.e., the space in which sound energy is being dispersed. Envelopment is therefore determined by the sound emanating directly from sound sources plus the inherent reverberation characteristics of the room's acoustics.[40]

Envelopment is significant in the study of immersivity and refers to sound being heard from 'all around the listener' with the aim to make the audience feel they are inside the sound.[41] Full immersion is not dominated by any single direction, meaning that localization is not compromised, and therefore a listener may enjoy different 'vantage points' (listening positions).[42] Envelopment is determined by Apparent Source Width (ASW), which is direct sound and early reflections arriving within 80ms, while Listener Envelopment (LEV) are sound reverberations arriving after 80ms.[43] Thus, there are two types of listener envelopment to consider, 1) 'source envelopment', in which we feel surrounded by sound sources, such as instruments or loudspeakers and 2) 'room envelopment', in which we feel surrounded by reflective sounds.[44] The architectural characteristics of any given structure impact the room's acoustics, and therefore impact how we may perceive envelopment within that space.[45] Smaller spaces for example produce both 'louder and more enveloping dynamics'.[46] That means the

[39] Daniel Barreiro, 'Considerations on the Handling of Space in Multichannel Electroacoustic Works', Organised Sound, 15.3 (2010), 290–96.

[40] Hyunkook Lee, 'Apparent Source Width and Listener Envelopment in Relation to Source–Listener Distance', AES 52nd International Conference: Sound Field Control-Engineering and Perception, 2013, www.researchgate.net/publication/286183639_Apparent_source_width_and_listener_envelopment_in_relation_to_source-listener_distance, accessed 12 September 2024.

[41] Chuck Ainlay et al., 'The Recording Academy's Producers & Engineers Wing: Recommendations for Surround sound Production', 2004, chap. 2.1, www2.grammy.com/PDFs/Recording_Academy/Producers_And_Engineers/5_1_Rec.pdf, accessed 12 September 2024.

[42] Smalley, 'Space-Form and the Acousmatic Image', (50).

[43] Lee, 'Apparent Source Width', 1.

[44] Natasha Barrett, 'Ambisonics and Acousmatic Space: A Composer's Framework for Investigating Spatial Ontology', Proceedings of the Sixth Electroacoustic Music Studies Network Conference Shanghai (2010), 1–12 (7), www.natashabarrett.org/EMS_Barrett2010.pdf, accessed 12 September 2024.

[45] Sarah Adair, Michael Alcorn and Chris Corrigan, 'A Study into the Perception of Envelopment in Electroacoustic Music', International Conference on Mathematics and Computing, 2008. Available at https://quod.lib.umich.edu/cgi/p/pod/dod-idx/study-into-the-perception-of-envelopment-in-electroacoustic.pdf?c=icmc&format=pdf&idno=bbp2372.2008.162, accessed 12 September 2024.

[46] Tapio Lokki, Laura McLeod and Antti Kuusinen, 'Perception of Loudness and Envelopment for Different Orchestral Dynamics', The Journal of the Acoustical Society of America, 148.4 (2020), 2137–45, at 2144).

listener's subjective response to the enveloping sound may be determined by their location within that space. Any effective control of envelopment must consider the size of the space and its structural composition, the size and amplitude level of the ensemble and/or speaker-array along with their position within the space, and the placement of the listener and their closeness to the sounding sources.

Thoughtfully assembled multichannel formats accompanied by suitable music for the space and audience, can create full three-dimensional audio 'environments and soundscapes of great immersive impact'.[47] The pursuit of surrounding the user with sound in a manner that emulates the real world is playing a pivotal role in many audio consumer markets. Thus, envelopment has established itself as an immersive characteristic reflected in the growth of immersive technologies and consumer products. At the production stage, computer applications have made the creation of immersive audio much more efficient and practical, while a myriad of consumer products have made it much more accessible. These technologies can be effectively deployed alongside or as an extension to live musicians seamlessly to enhance the immersive nature of the performance – it is a matter of resource, will and ingenuity.

3. Sound Processing – Considering sonic clarity, space, localization and listener perception.

The way in which audiences process and perceive sound is integral to the immersive music performance medium. How we process sound in a live environment is influenced by how the performance space impacts the quality of the sonic image. In addition, listener's experience of musical forms likely predisposes their perception of the work.

Sonic Clarity refers to the listeners' ability to accurately hear sound sources and differentiate timbres within a performance space depending on their position. Sonic clarity is extremely important for any immersive elements to have their intended impact. Sonic clarity must be prioritized so that listener enjoyment is not unintentionally reduced and immersive characteristics are allowed to have their desired effect. Spatial characteristics can provide important sonic clarity, but along with the benefits, there are pitfalls that can result from listening positions and performance coordination errors. This is a fundamental challenge of immersive music performance, where experimental performance and listening arrangements within a given space – with the aim to increase proximity and envelopment – can negatively impact sonic clarity.

---

[47] Barry Truax, 'The Aesthetics of Computer Music: A Questionable Concept Reconsidered', Organised Sound, 5.3 (2000), 119–26, at 122.

The choice of space and its organization is pivotal to this performance medium. There are various arrangement formats that can be employed, and there are no limitations to how one might wish to explore the performance environment. Xenakis provides some graphical examples (Figure 10) [48]:

• *Frontal* – a typical stage situation with a clear division border between the performers and audience
• *Central* – an in-the-round formation where the audience is situated around the performers
• *Sources surrounding the audience* – a common surround-sound setting
• *Sources within the audience* – scattered sound sources including musicians within the audience
• *A narrow or lineal performance space* – as in a procession or parade
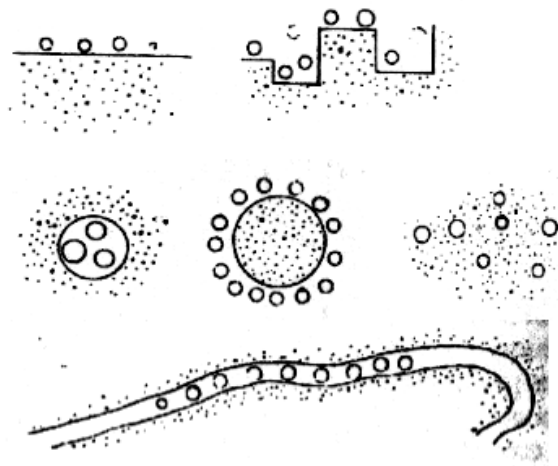• *A hybrid of several of these types*

Figure 10 Graphic description of Xenakis's arrangement types, from Harries,
'The Electroacoustic and its Double', 77–78).

Space itself has its own sonic qualities which cannot be removed, therefore a space which contains music is a musical space, where 'one has limited power to control' the effect of the space on the music.[49] It can be argued that the physical attributes of the building and its environment produce a 'secondary performance of their own'.[50] The listening space and position is vitally important in the way the sound image is perceived, because all movement, that of the body, objects and other listeners, allow us to understand space and spatial

---

[48] Harries, 'The Electroacoustic and its Double', 88.

[49] Austin, 'Sound Diffusion in Composition and Performance', 14.

[50] Harries, 'The Electroacoustic and its Double', 77–78.

behaviours.[51] We are very aware of the performance space, because 'we are constantly aware of personal space within the orbit of our practical daily activities or personal relations'.[52] Thus, immersive performance practitioners must consider the space at both the composition stage and the performance stage.

Music, however, is usually composed in studios which are inherently different sonically than the performance space. There are opportunities to mitigate the effect of the sounding space during rehearsals by making appropriate adjustments in balance and placement. However, the presence of an audience means that the sounding space behaves differently during a concert. Any adjustments, or 'interpretative decisions' during rehearsals are speculative, and most likely require further adjustments during the performance.[53] It is important to listen from various different positions that audience members can occupy when making adjustments prior to a concert. A performer may employ various tactics to mitigate the influence of the sounding space on the work, or alternatively embrace the conditions because 'it is the medium which is fixed, not the music'.[54]

The perception and effectiveness of any performance piece can be said to relate to the listener's position. These can be *fixed* (seated), *variable* (changeable) and *peripatetic* (listening from more than one position).[55] If an audience member is not constrained to a single position and given the opportunity to explore the performance space freely, it may lead to a stronger spatial perception and enhance engagement.[56]

A key immersive factor in music performance is to take advantage of our *localization* mechanism and add *spatialization* techniques. Localization refers to the listener's ability to distinguish the position of a sound-source and any movement of it within the soundfield, including multiple sources simultaneously. The term *spatialization* refers to the technique by which a composer utilizes 'spatial projection, sound location and direction' as important elements in the music [57], whether from a moving musician or the movement of a sound-source

---

[51] Gary Kendall, 'Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music', Organised Sound, 15.3 (2010), 228–38.

[52] Denis Smalley, 'The Listening Imagination: Listening in the Electroacoustic Era', Contemporary Music Review, 13.2 (1996), 77–107, at 91.

[53] Stansbie, 'The Acousmatic Musical Performance', 132.

[54] Stansbie, 'The Acousmatic Musical Performance', 41.

[55] Smalley, 'Space-Form and the Acousmatic Image', 52.

[56] George Klein, 'Site-Sounds: On Strategies of Sound Art in Public Space', Organised Sound, 14.1 (2009), 101–8.

[57] Harley, 'Spatiality of Sound and Stream Segregation', 148.

within a loudspeaker system. A composer can increase the immersive nature of the work by producing sounds that move within the soundfield offering 'physicality and dynamism' to a performance.[58] The use of the full soundfield to project spatially separated sound sources can produce benefits, such as greater clarity of pitch and texture, permitting for more complex music.[59]

Localization results due to the manner in which we process Interaural Time Differences, measured by the time difference the signal arrives at each ear; and Interaural Intensity Differences, measured by the intensity difference at each ear (Figure 11). The IID is affected by 'head-shadowing', in which the head acts as an obstacle for frequencies above 2000Hz (below 2000Hz the 'wavelength can diffract around the obstructing surface').
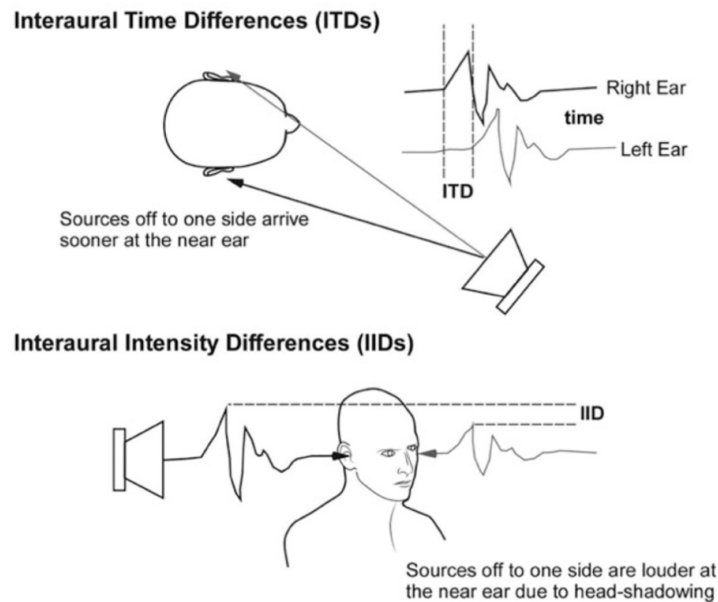

Figure 11 ITD and IID examples, (Roginska and Geluso, *Immersive Sound*, chap., 1).

There are certain qualities in spatial audio that must be considered however, and why conceptual ideas do not always translate in real-time due to nuances in spatial hearing and thinking. Localization is easier for a listener when the sound is moving than when it is stationary.[60] It is also easier to localize sounds that have sharp attack characteristics and those with wide band signals (high frequency content); inversely, sounds with no transient attacks

---

[58] Bates, 'The Composition and Performance of Spatial Music', 4.

[59] Harley, 'Spatiality of Sound and Stream Segregation', 150.

[60] David Malham and Anthony Myatt, '3-D Sound Spatialization using Ambisonic Techniques', Computer Music Journal, 19.4 (1995), 58–70.

and lower frequency content are harder to localize.[61] However, one must be aware of 'localization blur' which can occur due to early-arriving reflections determined by the size and shape of the room.[62] Localization is consequently affected by environmental ambience produced by the room's architectural characteristics.

Another common issue to consider is *precedence*, an auditory mechanism that allows us to determine the direction of sound in complex auditory environments. When a sound is heard directly from a source, in addition to distant sources and reverberation, precedence will be given to the first-arriving direct sound, while later-arriving sounds are suppressed.[63] Precedence therefore effects spatial panning due listener position; only the sweet-spot where all speakers are equidistant, precedence will not take effect.[64] *Image dispersion* explains how even in the sweet-spot spatialization techniques such as circular panning are not as coherent to the side of the head as in-front due to ITDs and IIDs.[65] This phenomenon similarly effects front-back and vertical spatialization. Sound-source localization accuracy is a 'deeply embedded cognitive capacity', and therefore a listener's processing of spatialization techniques will be affected by their experience with the approach.[66] Thus, novice and advanced listeners of immersive music may have very different experiences.

It is therefore important to consider *Listener Perception*, and how much we wish for it to be challenged. Traditional compositional elements such as melody and rhythm are deeply engrained cultural languages, but experimental music practices such as spatialization may require a heavier cognitive load. This becomes particularly pertinent in the performance of new immersive works, in which listener expertise can vary dramatically and will most likely be limited. Nonetheless, audience perception is an essential component in deciphering the success of the musical work against its intentions, particularly in new performance forms.[67] Each listener has predetermined cultural and social expectations, while they are entitled to be critical of art-forms and whether they meet their needs. A composer can misjudge the strategy in which sound-based material is communicated to an audience and lead to the work being

[61] Malham and Myatt, '3-D Sound Spatialization'.

[62] Roginska and Geluso, Immersive Sound, chap. 1.

[63] Gary Kendall and Andres Cabrera, 'Why Things Don't Work: What You Need To Know About Spatial Audio', International Computer Music Conference, 2011, 1–4. Available at https://pureadmin.qub.ac.uk/ws/files/846733/WhyThingsDon_'tWork.pdf, accessed 12 September 2024.

[64] Kendall and Cabrera, 'Why Things Don't Work, 2.

[65] Kendall and Cabrera, 'Why Things Don't Work, 2.

[66] Kendall and Cabrera, 'Why Things Don't Work, 1.

[67] Kendall, 'Spatial Perception and Cognition'.

misunderstood.[68] Immersive performance practitioners must consider whether the audience members have the experience to actively engage with this auditory experience. Do we wish our audience to be entertained naturally through melody, harmony and rhythm, or should we expect them to actively engage in our experimental journey?

The manner in which listeners engage with and perceive sound within any given *performance space* is incredibly complex. Immersive music practitioners should consider the following parameters to improve sound processing:

1) The size and shape of the performance space.
2) The acoustic influence of the performance space.
3) The suitability of the musical material for that space.
4) The suitability and accessibility of the musical material for the audience.
5) The type and arrangement of sound sources within the performance space.
6) The sonic clarity and intelligibility of the performed material.
7) The usage of spatialization techniques including that of physical sound sources.
8) The arrangement of audience members within the space, with the potential for a peripatetic vantage point.

4. Visual Processing – Our ability to visually process any physical actions and movements of sound-sources within the performance space.

The ability for audiences to clearly see and construct musical meaning from any physical action, such as playing, singing and moving, is a visual element of great importance to any live performance.[69] Physical actions are significant in music performance because we make connections between the actions of a musician and the sound that is heard. We intuitively create relationships between what we hear and what we see to help us understand the real world (this is also known as 'source-bonding').[70] When we hear sound in a composition, we use our memory to relate what we have heard to any known sounds we have previously encountered in the real world.[71]

---

[68] Barry Truax, 'Music, Soundscape and Acoustic Sustainability', 2016. Available at https://www.sfu.ca/~truax/Sustainability.pdf, accessed 12 September 2024.

[69] Ed Wright, 'Symbiosis: A Portfolio of Work Focusing on the Tensions between Electroacoustic and Instrumental Music', (2010). Available at www.academia.edu/7451225/Symbiosis_A_portfolio_of_work_focusing_on_the_tensions_between_electroacoustic_and_instrumental_music, accessed 12 September 2024.

[70] Smalley, 'Spectromorphology', 110.

[71] Robin Parmar, 'The Garden of Adumbrations: Reimagining Environmental Composition', Organised Sound, 17.3 (2012), 202–10.

The physical movement of body and limbs are known as *gesture*, which embody expression and act 'as a compositional tool for the artist and a cognitive component for the listener'.[72] Usually, gesture is produced by a performer who acts upon mechanical 'sounding bodies by fingering, plucking, hitting, scraping and blowing'.[73] *Utterance* is a voice generated by a human body, one that acts as a 'vehicle of personal expression and communication', announcing the presence of a human even when used in music which only contains recorded vocals.[74] Historically we have been accustomed to listen to music that offers both gesture and utterance as audio-visual expressions, therefore, both are strongly embedded within us culturally. Furthermore, our experience of listening to instruments is related to years of unconscious audio-visual conditioning, and therefore our understanding of musical sounds and meaning is strongly embedded in these experiences.[75]

> gesture is rooted on archaic principles dating from the very first steps in communication between human beings. It is rooted on the first traces of languages. It is not just a physical gesture. Is a communicative gesture. It is a code. (It) has a meaning. [76]

Physical actions in the traditional sense of music performance are incredibly important in how emotion is conveyed. Physical actions help the audience better understand the emotional intention, and therefore generate engagement; live performances continue to be extremely popular, possibly for this precise reason. These connections cannot be replicated in the same manner via hi-fi systems, headphones or even television, as these media lack the energy conveyed by the physical presence of the musicians and the audience. The *liveliness* of a situation when determined by human physical presence has great influence on audience perception, and encourages the listener 'to acknowledge the physical space'.[77]

It was only with the advent of electroacoustic music performance (one which relies on loudspeakers) that neither the physical representation of vocal utterance nor instrumental

---

[72] Anil Camci, 'A Cognitive Approach to Electronic Music: Theoretical and Experiment-Based Perspectives', International Computer Music Association, 2012, 1–4, at 2, http://hdl.handle.net/2027/spo.bbp2372.2012.001, accessed 12 September 2024.

[73] Smalley, 'The Listening Imagination', 84.

[74] Smalley, 'The Listening Imagination', 86.

[75] Smalley, 'The Listening Imagination', 86.

[76] Alexis Perepelycia, 'Human-Computer Interaction Models in Musical Composition', (2006), 1–98, at 58, www.academia.edu/3183112/Human_Computer_Interaction_Models_in_Musical_Composition, accessed 12 September 2024.

[77] Harries, 'The Electroacoustic and its Double', 88.

gesture were essential. In electronic music, sounds can be produced without the presence of instrumental or vocal performers, and they therefore do not embody the physical actions with which we have become so familiar. Consequently, the listening experience may be perceived very differently in live music that includes performers than it is with electroacoustic music that does not. Music is abstract, intangible, and ethereal; therefore, the visual aspects are necessary for both the audience and the performer to establish and communicate its location and cultural significance within a society. Although trends are certainly changing, historically we have been conditioned to expect music with a certain level of physical action to place meaning against aural events.

> If you take away or weaken the tangibility of the known, visual, gestural model and the direct, universal articulations in utterance, then you undermine the stability of the conscious and unconscious reflexive relationship that the listener seeks. [78]

If for example we have experienced a physical instrument such as a guitar in real world circumstances, we recognize the sound has been made by human physical action, even if we no longer have the visual aid. We can use our imagination to identify the physical actions that occurred to propagate that sound. Furthermore, we instinctively assign physical actions to sounds that we may not even recognize, simply due to this innate conditioning.[79] This is a useful tool in nature, in which we can assign visual cues to non-visible sound objects. Hence, it is not surprising that when listeners have sensory experiences of sound alone, they are able to make sense of these experiences as events that take place with objects, actions and agents, even though these constituent elements may be unobserved, obscure or unknowable.[80]

Physical actions in the music performance medium are any bodily movements that touch, grasp and manipulate physical objects to form sounds that convey musical meaning. This definition is also useful in the context of electronic music, where small physical actions that use MIDI controllers, synthesizers and computers, are still central to the performance. Physical action in electronic music performance is not just the movements of faders and knobs, but they are any movements that can express something.

> Pressing a key or sliding a bow during a performance are movements that hold a meaning in themselves: they establish how a sound will be produced, they determine some characteristics of that sound, they set up connections with previous sonic events,

---

[78] Smalley, 'The Listening Imagination', 96.

[79] Alistair MacDonald, 'Performance Practice in the Presentation of Electroacoustic Music', Computer Music Journal, 19.4 (1995), 88–92.

[80] MacDonald, 'Performance Practice in the Presentation of Electroacoustic Music'.

and, at the same time, they furnish an articulatory path to further sounds. Obviously, they are more than simple movements, they are meaningful gestures.[81]

Due to the absence or reduction of physical action in some types of music performance, there are likely to be 'intrinsic cognitive differences' in both the audience and performer experiences.[82] The development of audio technologies means that there is almost no limitation to what a live electronic music performer is able to generate – offering 'a bewildering sonic array ranging from the real to the surreal and beyond'.[83] It is also worth considering that electroacoustic music is not limited by timbral characteristics of a live performer, or their technical ability, whilst production techniques that were once confined to the studio can now be recreated in real-time.

Electronic music performance, with its limited visual cues, exhibits a dislocation between what is seen and what is heard, because tiny physical gestures can create quite exaggerated sonic changes. This can often become a problem where the audience is unable to make a connection between the performer's physical actions and the resulting sounds. These physical actions are usually minute in comparison to the change in musical sound, such as dynamics, texture etc. A laptop performer can affect the dynamic range of a piece with minimal action, whilst in an ensemble, it requires all of the musicians to adopt a dynamic transformation through very visual gestures. A simple melody such as 'three blind mice' performed by MIDI data (even with the most elaborate programming in place with various expression techniques), cannot compare to the human expression delivered by a violinist, which includes not only gesture, but facial expression.[84] The challenge for electronic performance is how to mirror what is heard with physical actions, especially when more spectacular sounds would necessitate more spectacular gestures (which tend to be quite theatrical) to be proportionally relevant.[85] The electronic performer therefore must make fully concerted efforts to assign meaning to the music in as much detail as possible. This means, that not only do physical actions need to be visible, but so should the electronic equipment used to create and propagate the sounds, whether physically in the space and by mediated means. The performance should also represent some level of emotional interpretation as is expected by more traditional forms of artistic expression. The

---

[81] Fernando Iazzetta, 'Meaning in Music Gesture', Trends in Gestural Control of Music (IRCAM), 1 (2000), 71–87, at 74. Available at www.academia.edu/82379658/Meaning_in_Musical_Gesture, accessed 12 September 2024.

[82] Camci, 'A Cognitive Approach to Electronic Music', 1.

[83] Smalley, 'Spectromorphology', 107.

[84] Dylan Menzies, 'New Performance Instruments for Electroacoustic Music' (PhD diss., University of York, 1999). Available at https://eprints.soton.ac.uk/371730/1/10.1.1.19.7321.pdf, accessed 12 September 2024.

[85] Rodolfo Caesar, 'The Composition of Electroacoustic Music', (PhD diss., University of East Anglia, 1992). Available at www.scribd.com/document/414769344/The-Composition-of-Electroacoustic-Music-pdf, accessed 12 September 2024.

purpose is to allow the audience to construct as much meaning from those devices and agents as that which is attached to traditional instruments and performers.

Visual processing is an essential characteristic in the development of immersive music performance due to the importance physical action plays in the perception of a musical work. The manner in which listeners codify music through visual means is of biological and cultural significance that is essential for audience enjoyment. Composition and performance aspects are entirely interlinked, and this is particularly pertinent in the field of immersive music performance, in which one seeks to exaggerate the movement of sound within a given space. It is possible for movement to occur in three different entities, the performer, the loudspeakers and the audience, meaning that there are more visual cues to process. The most unique and captivating immersive events are likely to arise when both the use of spatial loudspeaker systems and live musicians with the incorporation of spatialization techniques are employed. Multi-disciplinary performances that include live musicians, visual effects and choreography are now common practice, and therefore, the addition of immersive characteristics may further enhance the live music experience.

5. Audience Engagement **–** To engage the audience in the performance.

Whether singing with friends at home, in a venue, or with thousands of strangers in a stadium, there is an emotional connection produced through the act of sharing. When an audience is being engaged it arouses emotions, stimulates physical reactions, taps into memories and fantasies, and triggers a cognitive response.[86] Audience engagement is as an integral feature of all music performance, because the audience is the consumer, and any successful performance must therefore consider an audience centred approach. Audience engagement is multifaceted, which requires an awareness of the listeners' wants and needs to help shape products that meet their expectations, and does not relegate them as merely receiving participants.[87] Engagement can be defined as 'a quality of the user experience that emphasizes the positive aspects of interaction, in particular the fact of being captivated'.[88] Another appropriate definition is the

---

[86] Jennifer Radbourne, Katya Johanson, Hilary Glow and Tabitha White, 'The Audience Experience: Measuring Quality in the Performing Arts', International Journal of Arts Management, 11.3 (2009), 16–29.

[87] Leshao Zhang, Yongmeng Wu, and Mathiue Barthet, 'A Web Application for Audience Participation in Live Music Performance: The Open Symphony Use Case', Proceedings from New Interfaces for Musical Expression, 2016. Available at: www.nime.org/proceedings/2016/nime2016_paper0036.pdf, accessed 12 September 2024.

[88] Najereh Shirzadian et al., 'Immersion and Togetherness: How Live Visualization of Audience Engagement Can Enhance Music Events', Advances in Computer Entertainment Technology: 4th International Conference, ACE 2017, London, UK, December 14-16, 2017, Proceedings, ed. Adrian David Cheok, Masahiko Inami, and Teresa Romão (Cham: Springer, 2018), 488–507, at 492.

ability for entertainment to immerse the consumer in a manner in which they lose awareness of time, lose awareness of the real world, and lose their sense of being in a task environment.[89]

> Immersion is a phenomenon experienced by an individual when they are in a state of deep mental involvement in which their cognitive processes … cause a shift in their attentional state such that one may experience disassociation from the awareness of the physical world.[90]

To increase audience engagement by employing immersive characteristics there are a few factors practitioners should consider. The aspect of *knowledge* recognizes the importance of information which enables audiences to better understand and appreciate the music they are consuming.[91] A music performance which may exhibit various complex compositional elements, may prove challenging for some audience members to process, reducing their overall engagement with the material, leaving them to feel excluded.[92] The knowledge concept is also linked to the visual processing characteristic already discussed, where it is beneficial for an observed *action-to-sound* element to exist, for audiences to fully understand the sound propagated. All this can be particularly true when new forms of music contain progressive and experimental elements. Immersive music performance is arguably one such practice, where abstract and improvisational music seems to be common practice, lacking an audience centred approach. This exact sentiment was noted by Lai and Bovermann who found that their music was too 'improvisational', and that by creating a 'predefined structure' it would better meet audience expectations and therefore enhance engagement.[93] Participants 'feel more immersed in the event … when they *like* the music'.[94] Thus, the audience should be central to the creative and delivery process, not secondary, and certainly not treated as mere receivers. This is perhaps even more significant in the investigation of immersive music performance, which should be broadly more palatable for the general consumer, avoiding elements that required greater knowledge.

---

[89] Shirzadian et al., 'Immersion and Togetherness', 492.

[90] Sarvesh Agrawal et al., 'Defining Immersion: Literature Review and Implications for Research on Immersive Audiovisual Experiences', Journal of Audio Engineering Society, 68.6 (2019), 404–17, 407.

[91] Bonita Kolb, 'You Call This Fun? Reactions of Young First-Time Attendees to a Classical Concert', The Journal of the Music & Entertainment Industry Educators Association, 1.1 (2000), 13–28.

[92] Zhang, Wu, and Mathiue Barthet, 'A Web Application for Audience Participation'.

[93] Chi-Hsia Lai and Till Bovermann, 'Audience Experience in Sound Performance', New Interfaces for Musical Expression, 2013, 170–73, at 173, https://nime.org/proceedings/2013/nime2013_197.pdf, accessed 12 September 2024.

[94] Shirzadian, et al., 'Immersion and Togetherness', (18).

Another factor to consider is *risk* from the audience perspective, which may include *expectational risk* (will this event meet my expectations?), *economic risk* (is this event worth the cost?), *psychological risk* (is this going to challenge me in a way I find uncomfortable?) and *social risk* (is there a social risk in how I want to be perceived by attending this event?).[95] Many patrons of the arts do not want to be entertained in a manner that is perceived as risky for these reasons. It is a matter of satisfaction versus cost, and cultural expectations related to potential socio-economic factors. Nonetheless, these risks can be positively framed as opportunities to experience something new and exciting.[96]

Authenticity is another factor, where 'the greater the authenticity perceived by an audience member … the greater (the) enjoyment of the experience'.[97] This may be a performance containing a certain level of technical standard that is expected with that genre, or whether it meets the audience's subjective perception of *authentic* as something real and believable.[98] This subjectivity notion is so pertinent that some audience members can have a totally different authentic experience than others from the same performance. Immersive music performance as a field is arguably in its infancy, lacking a widely perceived notion of authenticity for the precise reasons at the centre of this paper – a lack of definition which makes it immeasurably more subjective. This hypothesis is tightly linked to the factor of knowledge, in which trained professionals are much more likely to engage with complex art-forms than the general consumer. This is particularly relevant in the field of immersive music performance which requires extensive knowledge, skill and experience. Thus, creators must be acutely aware of this factor, and consciously adopt practices to reduce the need for knowledge with the aim to increase engagement and therefore authenticity as a consequence.

Collective engagement is a factor that refers to an audience being engaged by the performers and other audience members. These are '1) communication between performers and the audience, 2) communication from the audience to the performers, and 3) interaction between audience members'.[99] The key objective is to create a sense of collective belonging whilst delivering a performance that can be considered authentic and of a desired quality.[100] Humans

---

[95] Radbourne, et al., 'The Audience Experience', 20.

[96] Radbourne, et al., 'The Audience Experience', 20.

[97] Radbourne, et al., 'The Audience Experience', 19.

[98] Radbourne, et al., 'The Audience Experience', 20.

[99] Radbourne, et al., 'The Audience Experience', 25.

[100] Radbourne, et al., 'The Audience Experience', 25.

are social beings, our opinions and affected state are greatly influenced by others around us, which equates to the communal feeling of 'togetherness'.[101]

Audience engagement can be experienced directly or indirectly. An example of direct engagement is when musicians directly engage with the audience to elicit a positive response. We may wish to picture Freddy Mercury addressing a full capacity Wembley stadium during Live Aid in 1985, in which a call-and-response segment delivers a rapturous response from the audience. That audience has experienced something powerful as a collective, which is further enhanced due to the exclusivity of that single shared moment.

> Emotions are always unique – what you feel in this moment, you will never feel in this way again. This is why people go to concerts – to feel this uniqueness of the moment.[102]

Indirect engagement may exist in a variety of performative expressions, usually a delivery which captures the essence of the music in that moment. Other indirect factors can be the connection between audience members and their positively shared experience. Another is the manner in which the audience experiences the physical space and its constituent parts, such as its overall architecture, stage presentation, seating arrangement, speaker systems etc.[103]

The matter of audience engagement is multifaceted, encompassing many variables that can co-exist to produce an engaging performance. This paper seeks to emphasize the message that performances which reach out to audiences should 'specifically and directly (be) intended, designed, or meant for audiences.'[104] This is even more pertinent in this field where audiences have less knowledge and experience of the methods employed. The following factors may be useful from a strategic perspective.[105]

1. Create a visually engaging and welcoming performance space.
2. Compose and perform material that contains engaging musical features with clear compositional structures.
3. Perform in a manner that directly and in-directly engages with the audience, using genre-specific communication and information where necessary.

---

[101] Shirzadian, et al., 'Immersion and Togetherness', 494.

[102] Shirzadian, et al., 'Immersion and Togetherness', 503.

[103] Radbourne, et al., 'The Audience Experience', 26.

[104] Stansbie, 'The Acousmatic Musical Performance', (61).

[105] Lai and Bovermann, 'Audience Experience in Sound Performance', 173.

4. Allow the audiences to draw meaning from an observable 'action-to-sound' delivery, by increasing visibility of performance expression.

Much of the discussion surrounding audience engagement further supports the previous immersive characteristics discussed. Proximity, envelopment, sound and visual processing, can all independently and collectively help to enhance immersivity in music performance, and by doing so can improve audience engagement.

Conclusion

The taxonomy serves to realize this paper's key aim – to identify spatial and performative characteristics that can be employed in immersive music performance. It attempts to establish the role of each characteristic whilst providing explanations for their meaningful application in practice. It can potentially inform prospective artists in the design and delivery of their immersive music projects. The characteristics can be applied individually to good effect, but a collective approach may generate more stimulating and unique experiences for the audience. This taxonomy is by no means conclusive, its synthesis underscores a foundational framework which requires further investigation.