

Sound Design for Cinematic Virtual Reality: A State-of-the-Art Review



Hitesh Kumar Chaurasia  and Manoj Majhi

Abstract In the last decade, the rapid development of low-cost 360° cameras, Head Mounted Displays, smartphones, and open-source softwares have created new possibilities of immersive storytelling in virtual reality, popularly known as cinematic virtual reality or virtual reality films. Even in 360° videos, the visual experience remains limited to field of view of head mounted displays. Hence, the spatial audio is considered essential to create ‘Presence’ and ‘Immersion’. The viewer is part of the virtual environment with the freedom to interact, navigate, and choose the viewing direction. This presents challenge of sound recording, designing and playback. The recent development in spatial audio recording, sound synthesis and rendering has addressed some of these challenges. The current study presents the state-of-the-art of sound design in cinematic virtual reality through literature review and textual analysis of the relevant publication in the field since 2015. The objective is to identify further scope of studies. This review could be useful for filmmakers, sound designers, and scholars working in the field of cinematic virtual reality.

Keywords Virtual reality · Immersive storytelling · 360° · Spatial audio · Film Sound design

1 Introduction

Over the years, virtual reality (VR) has slowly and steadily evolved as an independent medium to create simulated experiences in diverse areas such as military training, aviation, health, education, and entertainment. Since 2012 the rapid development of low-cost 360° cameras, Head Mounted Display (HMDs), smartphones and open-source software’s have created new possibilities of immersive storytelling popularly known as cinematic virtual reality (CVR) or 360° VR films. While VR can offer

H. K. Chaurasia (✉)

National Institute of Design, Ahmedabad and Indian Institute of Technology, Guwahati, India

e-mail: hitesh_c@nid.edu

M. Majhi

Indian Institute of Technology, Guwahati, India

a full 360° view, at any given instance the visual experience remains limited to the field of view—maximum visual field a VR headset displays. This leads to fear of missing out (FOMO) concern in viewers [1]. To address this, sound's innate nature of omnipresence can be used to localize the source of events beyond the field of view. Hence, spatial audio is considered essential to create 'Presence' and 'Immersion'. However, there are challenges of sound design as, the visual frame that contextualizes the acoustical events no longer exists. The viewer being part of the virtual environment with the freedom to navigate, interact, and choose the viewing direction further adds to the complexity of sound design as well as playback. The current study presents the state-of-the-art of sound design in cinematic virtual reality through literature review textual analysis of the relevant publication in the field since 2015. CVR is in its nascent stage and sound design practices, process, aesthetics, and technology are still evolving. Hence the objective of this study is to present possibilities and challenges of sound design in CVR and to identify the scope of further studies in the field. This could be useful for filmmakers, sound designers, and scholars working in the field of cinematic virtual reality. Section 2 gives the background of CVR followed by Methodology in Sect. 3, literature review in Sect. 4, discussion in Sect. 5, and conclusion in Sect. 6.

2 Background

Through the advancement in motion picture from the silent era, black and white films to talkies and colour picture to 3D Cinema, the objective has been to make the cinematic representation as close to reality as possible. *Sensorama* developed by Morton Heilig in 1956 is regarded as one of the first attempts to create a simulated environment. By 1965, another inventor, Ivan Sutherland, offered "the Ultimate Display," a head-mounted device that he suggested would serve as a "window into a virtual world" [2]. Since then, with the advancement in computer graphics, processing and visual displays VR has evolved as an independent medium. However, there has been debate over 3D and other VR related developments in terms of control over narrative and perception from the very beginning as discussed in *the cinema of the future* by Heilig [3]. In 2012 Palmer Luckey developed a prototype of the Oculus Rift- a portable VR headset offering a 90° field of vision using the computer's processing. In 2014, Facebook acquired the Oculus VR company for \$2 billion [4] and subsequently Sony, Google and Samsung also launched VR headsets. Google's low-cost mobile-based do-it-yourself Cardboard brought the VR to the consumer domain. The leading entertainment companies such as Walt Disney and 20th Century Fox also started exploring the possibilities of VR as a medium for storytelling. To create an immersive experience, existing spatial audio format such as Binaural and Ambisonic were further developed for VR. The recent development of Dolby Atmos, wave field synthesis and Head-Related Transfer Function has further created possibilities of immersive sound design and storytelling in CVR.

3 Methodology

A detailed literature search was conducted through multidisciplinary databases for keywords- Sound Design, Cinematic Virtual Reality, Immersive Storytelling, Presence, 360° video, Spatial Audio, Localization, Binaural, Ambisonic and 3D-audio. CVR is relatively a new field hence the initial search was limited from 2015 to present-day publications. Though, some widely referred earlier seminal works were included to define the key theories and concepts in the field. Likewise, some articles from the web were also referred to incorporate the filmmaker's views on contemporary practice and technologies. The publications were shortlisted in a three-step manual process based on relevance, starting from title and keyword review followed by abstract review and finally full paper review. Publication in the domain of pure acoustics, engineering or technology were not included as the scope of the study is the creative application of sound design in CVR.

4 Literature Review

4.1 Immersive Storytelling

In VR, the viewer is part of the virtual world with the freedom to navigate, interact, and choose the viewing direction. This creates a sense of immersion, being 'there' or presence for the viewer. This is the biggest advantage of VR over traditional mediums for immersive storytelling. At the same time, it also presents new challenges to filmmakers as they need to let go of the control over the narrative in terms of framing and editing. Hence, filmmaking tools and techniques developed over the last century needs to be reconsidered and a new filmmaking language needs to be developed for effective storytelling in CVR.

4.1.1 Cinematic Virtual Reality: New Cinematic Language

According to Mateer, cinematic virtual reality provides an immersive VR experience where individual users can immerse themselves in synthetic world experience in 360° Videos [5]. However, this definition uses the term *synthetic world* which refers to computer-generated imagery (CGI) and does not consider live-action 360° films. In a fully virtual experience, the viewer can interact and inform the storytelling. Hence, there is debate, if 360° videos can be considered Cinematic VR. There is an argument that the option to look in any direction also allows the audience to interact with the video and create her own space within the storytelling [4]. This study includes 360° videos also in the discussion.

The traditional notion of the fourth wall no longer exists in CVR. As the viewer is part of the virtual cinematic space with the freedom to choose viewing direction,

interact and navigate. Hence, elements of film language as describe by Bordwell and Thompson for screen-based cinema such as cinematography, mise-en-scene, editing, and sound design [6] needs to be reconsidered and a new film language needs to be developed for CVR. Since the viewer is part of the virtual world, her role and the degree of participation in the storytelling needs to be determined before starting the filmmaking process as this shall influence the overall design of the CVR. Dolan and Parets based on their professional experience in the field proposed a four-quadrant system of storytelling based on four viewer types: Observant Active, Observant Passive, Participant active, and Participant passive [7]. On similar lines, Cho et al. also explored the various approaches of viewer's engagement with the story in CVR such as a first-person (the viewer being addressed by a character in the film) and third person (viewer just observes the action) perspective [8]. The level of control/agency given to the viewer also influences the storytelling in VR and conflicts with the filmmaker's control over the narrative. Ruth Aylett defines it as a "narrative paradox" [9]. Godde et al. conducted an empirical study using exiting 360° videos to understand the possible changes required in CVR from screen-based cinema. Some of the key consideration discussed in the study are to define the the role of the viewer as active participant, or passive observer; guiding the viewer's attention; camera placement; Re-thinking framing and Editing [10]. Although this study does not discuss the sound in detail, it's finding such as placement of story elements, the balance of spatial and temporal story density could inform the sound design as well.

In shooting live action films with a 360° camera, the entire sphere is in frame. Hence there are challenges of location selection, framing, lighting, and sound recording. While Computer Graphics Imagery based films are free from these considerations, creating a realistic environment and process sound to reflects the acoustics of the space remains a concern. Some other studies focused on elements of film-making such as editing [11], the interplay between sound design and editing [12] in terms of rhythm, and a framework of screenwriting for an Interactive VR film [13].

4.2 Immersive Sound Design

Humans perceive sound spatially, hence the spatial audio is required to do immersive sound design for CVR as discussed in the following sections.

4.2.1 Perception of Sound

Sound design theories and practices are based on the perception of sound by a human being. Lorenzi explained the way humans localize sound in three dimensions as per psychoacoustics [14]. Further, Kendall demonstrates how sound localization or the sense of direction of a sound is calculated by the brain based on the time difference, phase difference, level difference, and intensity difference between sound signals

arrival at two ears [15]. However, Potisk underlines the limitation of human beings to accurately perceive the distance of a sound through hearing alone and the support of other sensory inputs are required [16]. How the upper human body such as ears, head, torso and shoulders influence the perception of sound in the brain and localize sound is explained by Duda [17]. To compensate for it, in headphone based listening a 3-D audio protocol; Head-Related Transfer Functions (HRTF) captures the transformations of sound wave propagation from the source to the ears [18]. The HRTF is concerned with modelling the physiological responses. These transformations account for the diffractive and reflective influences of the head, pinnae, shoulders and torso, and the combination of two unique HRTFs for the left and right ears satisfactorily addresses the parameters for localization.

4.2.2 Spatial Audio

The human ear can locate source of sound and its direction in three-dimensional space as spatial input. The objective behind the development of a stereo system in the 1930s was to create an impression of a 3D sound environment, surrounding a listener thus simulating auditory reality [19]. For this reason, first Stereo and then multi-channel surround sound system became popular in screen-based cinema even though such format were limited to the horizontal plane. In spatial sound, there are x, y and z-axis for localization. Topping explains that a 3-D audio system can depict a realistic auditory event to an extent that the brain is convinced that a sound originates not from the loudspeakers or headphones, but an arbitrary point in three-dimensional space [20]. In recent times multiple spatial audio formats have been developed. Three widely used spatial audio formats are Binaural, Dolby Atmos and Ambisonic. For the scope of this study, only their main advantages and limitation are summarized in Table 1.

Table 1 Advantage and limitation of three spatial formats

Format	Advantage	Limitation
Binaural (channel based)	<ul style="list-style-type: none"> • Normal stereo-file • Compatibility across platforms • Simple and quick spatial audio solution 	<ul style="list-style-type: none"> • Headphone based • Sounds weird on loudspeaker • Rendered mix, no real-time head-tracking
Dolby atmos (object-based)	<ul style="list-style-type: none"> • Object-based design • Independent of channel/speaker arrangement 	<ul style="list-style-type: none"> • Requires a decoder • Limited distribution and compatibility
Ambisonic (sound field-based)	<ul style="list-style-type: none"> • Compatible with other channel-based formats • Scalable with more channels, for 360° videos 	<ul style="list-style-type: none"> • Requires a decoder • The output varies on different platforms

4.2.3 Sound Design for CVR

The sound recording in live-action 360° is challenging as everything is in the frame, it limits the choice of location sound recording equipment and crew. The CGI based workflow does not have these challenges as voices can be dubbed but to process the voices to match the cinematic space as well as build the soundscape from the scratch remains a challenge. McArthur explored the idea of distance in audio for VR and outlined constraints of sound design in terms of hiding the microphone during a 360° shoot to processing in post-production [21].

Candusso [22] study underlines the difference of workflow from linear media to VR, in terms of 3D sound and its applications [22]. However, the lack of consideration of other elements of cinema and key concepts of VR such as presence and Immersion limits the scope of study and merits further exploration. Through case studies, Erkut et al. proposed a framework of sound recording and design for Virtual Environment known as SIVE (Sonic interaction for the virtual environment [23]. To explore immersive production techniques in cinematic sound design, Downes et al. adopted a scene from the 1981 submarine classic film *Das Boot* in a 9.1 loudspeaker configuration. The study suggests that the context is important to utilise this enhanced spatial format [24]. Milesen et al. found that people sense significant more direction in the sound when they get a monoscopic image with ambisonic sound compared to a stereoscopic image with stereo sound, however, they do not feel significantly more direction in the sound when the imagery is monoscopic with stereo sound or when the imagery is stereoscopic with ambisonic [25]. This underlines the complexities of existing formats for both filmmakers as well as viewer.

Even though output from a multichannel speaker system is possible and a framework for the same has been suggested by Sungsoo and Sripathi [26], headphones remain the preferred choice to view VR content with HMDs. Freeman suggests that numbers of channels do not have a significant impact on presence or immersion [27]. While another study suggest that the increased number of speakers and a wider spatial audio distribution diffused the participant's attention [28]. These studies raise some important questions with regards to the relationship of visual attention and a layered spatial soundtrack as well as the time needed to adapt to such complex conditions. The findings of the study are significant to examine the relationship of presence with spatial audio formats. The relationship between sound design and viewer's emotions and reaction has been discussed in several earlier studies [29–31]. One of the relevant studies in this regard is by Kock and Louven [32]. In this study, a Live action and animated film were produced and analysed with different sound treatments. The findings of the study suggest that non-diegetic sound effects for live-action films and diegetic sound in animation film may enhance the impact of the sound effects and lead to greater viewer immersion or suspense. Even though this study is for screen-based cinema the findings are relevant to understand the different approach of sound design required in CGI and live-action based films in CVR. Further, Boltz examined how the viewer's perception of the speed of visual elements gets influenced by the tempo of diegetic sounds [33]. Walden suggests that details in the audio help the

brain process the visuals faster [34], likewise Cheung has shown that using highly expected sounds increases users' sense of presence [35].

4.3 Immersive Experience for Viewer

The experience of CVR differs from traditional screen-based cinema in two aspects first; (a) viewers use head-mounted-display (HMDs) with headphones and experience the content individually; (b) the viewer is present in the virtual world and experience the content as first-person point of view. Hence the idea of presence and immersion needs to be understood in the context of CVR. Lombard and Ditton defined presence as “the perceptual illusion of non-mediation” [36]. This presence, in which the viewer locates herself in the scene, is one of the greatest strengths, but also one of the biggest challenges for storytelling in VR. For effective immersive storytelling, “presence” is the essential prerequisite. The VRs high chance of creating a feeling of presence has been explored, based on the Two-Level Model of spatial presence, introduced by Wirth et al. [37]. According to Slater and Wilbur, while “presence” is the state of being there in the virtual world, “immersion” describes and assess the character of the technology in terms of systems able to deliver an inclusive, extensive, surrounding, and vivid illusion of virtual environment to a participant [38].

A 2015 experimental study analysed the viewer's experience of cinematic virtual reality with head-mounted displays [39]. For the study, a CVR film *The Prism* was produced and viewers watched it on Oculus Rift with headphone. The findings of the study suggest that even though viewers liked the experience of a VR film and could comprehend the story, being able to look around during the film was a distraction for some viewers. Other studies [40, 41] also examined if freedom to look around in cinematic virtual reality enhances the viewing experience or spoils it. This continues to be a topic of debate amongst scholars to date.

The findings of the study by Brinkman et al. suggest that adding sound to the virtual world has a significant effect on people's experience though no difference in experience between stereo and 3D audio was reported [42]. In the same line, another study by Kobayashi et al. suggest that the sympathetic nervous system activated to a greater extent by the spatialized sound [43]. Ding et al. compared a 2D film against a VR film and the findings suggest that CVR has a significant influence on the emotional processing of the audience in comparison to 2D films [44]. Project Orpheus—a practice-based research- examined the idea of presence and immersion in CVR by making user part of the narrative and the use of sound to guide users' attention [45]. This study presents participants two versions of the same video. In version one, the sound matches the visual while in the experimental version the sound announces the upcoming image/visuals beforehand. The idea behind this experiment was to prompt users about a new event such as characters entry through sound. However, the participants of the study found this version of the sound mix as ‘distracting’ by not finding the corresponding visual of the sound immediately.

4.3.1 Guiding the Attention of Viewer

In 360° CVR often viewers do not know where to look? This results in FOMO (the fear of missing out) complex in viewers [1]. This is a common phenomenon in VR and the storyteller needs to guide the viewer's attention to the key points in the narrative. Rothe et al. summarized the various methods developed in recent years for attention guiding in virtual Reality [46]. Similarly, studies also suggest that depending on the objective, the attention can be space-based (position of an object), feature-based (features of an object) or object-based [47, 48]. The possibility to guide the attention through visual clues is discussed in [49, 50]. These studies also suggest that forced control of user's action may negatively influence presence.

Van der Burg et al. demonstrated through the "Pip and Pop" effect that adding a simple auditory "pip" with a visual colour-change effect "pop" helped in finding the event for viewers [51]. According to Rothe et al. sounds can motivate the user to search for the source of the sound and therefore to change the viewing direction [52]. Binaural audio has a clear advantage to spatial visual processing as it can reduce visual search time [53] However, Mendonca et al. also observed that 3D sound can also have a negative impact on perception [54]. The studies suggest that just like screen-based cinema, in CVR also the sound has to be designed in the context of the narrative along with visuals and the theories cannot be generalized beyond a point.

5 Discussion

The key challenges of sound design in CVR are to define the role of the viewer, to draw the attention of the viewer, and to maintain the sense of presence and immersion. The shooting live-action 360° is challenging as everything is in the frame, it limits the choice of location sound recording equipment and crew. The CGI based workflow does not have these challenges still to re-create voices to match the cinematic space as well as build the soundscape from the scratch remains a challenge. Several studies have shown that sound can be used for guiding attention, but the finding of some studies also suggest it to be distracting as well. The sound design needs careful consideration as any additional layer of sound might be distracting. The same is the case with the use of multi-channel audio output. Though the output through speakers is possible, the preferred mode of audio playback remains headphone, which limits the communal experience of film watching. The advancement in technology such as location microphones, sound field recorders, stand-alone HMDs, HRTF, object-based audio, wave-field synthesis has enhanced the overall sound design for CVR as well as the film viewing experience for the audience. Still the health-related concern of use of HMDs with headphones continues to inform the duration of content as well as its design. This review could be of help to scholars working on the Ergonomics design aspects of HMDs as well. The technologies, processes, and design aesthetics of sound design for CVR is still in flux. The standards are yet to be evolved. In this context, the filmmakers and sound designers need to adapt to the new medium, learn

a few tools and techniques specific to VR. A framework of sound design shall help filmmakers to create immersive experiences.

6 Conclusion

Based on the literature review, it is observed that there is a growing interest amongst scholars in cinematic virtual reality. Still, the studies focused on sound design are rare to find. Simulate reality remains the objective of sound design for filmmakers. Hence it remains the focus of studies as well. While ‘being real’ might be the primary concern in health, aviation, and military training, in cinematic virtual reality, the role of sound design is to support the narrative and enhance the cinematic experience. Hence there is a need to explore sound design beyond just presence. There is debate over the use of sound and the balance between spatiality, density and guiding the viewers’ attention. A detailed study of sound recording, editing, mixing and playback of sound elements i.e., voice, sound effects, ambience, music, and silence is required to develop a comprehensive framework of sound design.

Acknowledgements This study was carried out as part of the PhD program coursework at department of Design, IIT Guwahati. Sincere thanks to course instructors Prof Urmi R. Salve and Prof. Debayan Dhar. The study is supported by Seed Research Grant for Early Career Faculty (SRG-ECF) from NID Ahmedabad.

References

1. De Abreu A, Ozcinar C, Smolic A (2017) Look around you: saliency maps for omnidirectional images in VR applications. In: Ninth international conference on quality of multimedia experience (QoMEX). Erfurt, Germany
2. Mazuryk T, Gervautz M (1999) Virtual reality—history, applications, technology, and future. https://www.researchgate.net/publication/2617390_Virtual_Reality_. Last accessed 20 Mar 2021
3. Heilig ML (1992) El cine del Futuro: the cinema of the future. *Presence Teleoperators Virtual Environ.* 1(3):279–294. <https://doi.org/10.1162/pres.1992.1.3.279>
4. Stuart D (2014) Facebook closes its \$2bn Oculus Rift acquisition. What next? <https://www.theguardian.com/technology/2014/jul/22/facebook-oculus-rift-acquisition-virtual-reality>. Last accessed at 20 Mar 2021
5. Mateer J (2017) Directing for cinematic virtual reality: how the traditional film director’s craft applies to immersive environments and notions of presence. *Media Pract* 18(1). <https://doi.org/10.1080/14682753.2017.1305838>. Last accessed 21 Mar 2021
6. Bordwell D, Kristin T (2012) *Film art: an introduction*, 10th edn. McGraw-Hill Education, Boston
7. Dolan D, Parets M (2021) Medium.com, <https://medium.com/@devon.michael/redefiningthe-axiom-of-story-the-vr-and-360-video-complex-bee3c20d69df#.bwu5blnj>. Last accessed 20 Mar 2021

8. Cho J, Lee T, Ogden J, Stewart A, Tsai T, Chen J, Vituccio R (2016) Imago: presence and emotion in virtual reality. In: ACM SIGGRAPH 2016 VR Village, Anaheim (2016). <http://dx.doi.org/https://doi.org/10.1145/2929490.2931000>. Accessed at 1 Apr 21
9. Aylett R (2000) emergent narrative, social immersion and “storification”. in: narrative and learning environments conference NILE00 Edinburgh, Scotland
10. Gödde M, Gabler F, Siegmund D, Braun A (2018) Cinematic Narration in VR—rethinking film conventions for 360°. In: Virtual, augmented and mixed reality: applications in health, cultural heritage, and industry. VAMR 2018. https://doi.org/10.1007/978-3-319-91584-5_15. Last accessed 10 Mar 2021
11. Brillhart J (2021) medium.com. <https://medium.com/the-language-of-vr/in-the-blink-of-a-mind-prologue-7864c0474a29#v0gfg5v0X>. Last accessed 20 Mar 2021
12. Erkut C (2017) Rhythmic interaction in VR: interplay between sound design and editing. In: IEEE 3rd VR Workshop on sonic interactions for virtual environments (SIVE), pp 1–4. <https://doi.org/10.1109/SIVE.2017.7901611>. Last accessed 10 Mar 21
13. Reyes MC (2017) Screenwriting framework for an interactive virtual reality film. In: Immersive learning research network, Coimbra, Portugal. <http://dx.doi.org/https://doi.org/10.3217/978-3-85125-530-0-15>. Last accessed 10 Mar 21
14. Lorenzi A (2021) cochlea.eu, <http://cochlea.EU/en/sound/psychoacoustics>. Last accessed 10 Mar 2021
15. Kendall GS (1995) A 3-D sound primer: directional hearing and stereo reproduction. *Comput Music J* 19(4):23–46. <http://jstor.org/stable/3680989>. Last accessed 10 Mar 2021
16. Potisk T (2015) Head-related transfer function. Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia, In Seminar Ia
17. Duda RO (2000) 3-D Audio for HCI. UC Davis Electrical and Computer Engineering, University of California. <http://ece.ucdavis.edu/cipic/spatial-sound/tutorial/>. Last accessed 15 Mar 2021
18. Cheng CI, Wakefield GH (2001) Introduction to head-related transfer functions (HRTFs): representations of HRTFs in time, frequency, and space. <http://www.aes.org/e-lib/browse.cfm?elib=10196>. Last accessed 20 Mar 2021
19. Larsson P, Våljamäe A, Västfjäll D, Tajadura-Jiménez A, Kleiner M (2010) Auditory-induced presence in mixed reality environments and related technology. *Human-Comput Interact Ser* 1: 143–163. Springer-Verlag London. http://dx.doi.org/https://doi.org/10.1007/978-1-84882-733-2_8. Last accessed 22 Mar 2021
20. Topping A (2018) A survey of three-dimensional sound and its applications. Capstone Projects and Master’s Theses. 284 California State University, Monterey Bay. https://digitalcommons.csumb.edu/caps_thes_all/284. Last accessed 24 Mar 2021
21. McArthur A, Sandler M, Stewart R (2017) Distance in audio: constraints and opportunities. In: Proceedings of AM ’17. London, United Kingdom. <https://doi.org/10.1145/3123514.3123530>. Last accessed at 1 Mar 2021
22. Candusso D (2017) Designing spatial sound: adapting contemporary screen sound design practices for virtual reality. In: SMPTE17 conference—international convention centre. Sydney, Australia. <https://doi.org/10.5594/M001749>. Last accessed 1 Mar 2021
23. Summers C, Lympouridis V, Erkut C (2015) Sonic interaction design for virtual and augmented reality environments. In: IEEE 2nd VR workshop on sonic interactions for virtual environments (SIVE), Arles, France. <http://dx.doi.org/https://doi.org/10.1109/SIVE.2015.7361290>. Last accessed 05 Apr 2021
24. Downes T, Ronan M (2016) Immersive production techniques in cinematic sound design: context and specialization. In: Audio engineering society 140th convention. Paris France. <https://secure.aes.org/forum/pubs/ebriefs/?elib=18153>. Last accessed 23 Mar 2021
25. Milesen V, Smed MD, Lind B (2017) Quality assessment of VR film a study on spatial features in VR concert experiences. Master Thesis, Alborg University Copenhagen (2017) https://projekter.aau.dk/projekter/files/260078324/Master_MED_2017__2_.pdf. Last accessed 25 Mar 2021
26. Sungsoo K, Sripathi S (2019) Multichannel audio implementation for virtual reality. In: AES international conference on immersive and interactive audio. New York University, New York, NY, USA. <http://www.aes.org/e-lib/browse.cfm?elib=20436>. Last accessed 01 Apr 2021

27. Freeman J, Lessiter J (2021) Here, there, and everywhere: the effects of multichannel audio on presence. In: International conference on auditory display (ICAD). Espoo, Finland. <http://hdl.handle.net/1853/50621>. Last accessed 30 Mar 2021
28. Korshunova V, Remijn G, Elischka S, Mendonça C (2019) The impact of sound systems on the perception of cinematic content in immersive audio-visual productions. In: 12th Asia Pacific workshop on mixed and augmented reality. <http://dx.doi.org/https://doi.org/10.1109/APMAR.2019.8709163>. Last accessed 30 Mar 2021
29. Chion CM (1994) *Audio-Vision: Sound on Screen* (GORBMAN C., Ed.). New York: Columbia University Press
30. Kessling P, Görne T (eds.) (2017) *KLG 2017 (EPiC Series in Technology, vol. 1)*, pp 17–30 (2017). <https://easychair.org/publications/open/xGwB>. Last accessed 12 Mar 2021
31. Hillman, N, Pauletto, S (2014) The craftsman: the use of sound design to elicit emotions. *The Soundtrack* 7(1). http://dx.doi.org/https://doi.org/10.1386/st.7.1.5_1. Last accessed 15 Mar 2021
32. Kock M, Louven L (2018) The power of sound design in a moving picture: an empirical study with emoTouch for iPad. *Empirical Musicology Rev* 13(3–4):132–148. Ohio State University, USA. <https://doi.org/10.18061/emr.v13i3-4.6572>. Last accessed 26 Mar 2021
33. Boltz, MG (2017) Auditory driving in cinematic art. *Music Perception*, University of California, <https://doi.org/10.1525/mp.2017.35.1.77>. Last accessed 05 Apr 2021
34. Walden J (2021) postperspective.com, <http://postperspective.com/vr->. Last accessed 30 Mar 2021
35. Chueng P, Marsden P (2002) Designing auditory spaces to support sense of place: the role of expectation. In: *The role of place in on-line communities workshop, CSCW2002*. New Orleans. https://www.researchgate.net/profile/Priscilla-Chueng-Nainby/publication/272790990_. Last accessed 10 Apr 2021
36. Lombard M, Ditton T (1997) At the heart of it all: the concept of presence. *J Comput-Mediated Commun* 3(2), JCMC321. <https://doi.org/10.1111/j.1083-6101.1997.tb00072.x>. Last accessed 07 Mar 2021
37. Vorderer P, Wirth W, Saari T, Gouveia F, Biocca F, Jäncke L, Böcking S, Hartmann T, Klimmt C, Schramm H, Laarni J, Ravaja N, Gouveia LB, Rebeiro N, Sacau A, Baumgartner T, Jäncke P (2003) Presence: towards a two-level model of the formation of Spatial Presence. In: *Report to the European community, project presence: MEC (IST-2001–37661)*
38. Slater M, Wilbur S (1997) A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. *Presence: Teleoper Vir Environ* 6 (6):603–616. <https://doi.org/10.1162/pres.1997.6.6.603>. Last accessed 1 Mar 2021
39. Syrett H, Calvi L, Gisbergen M (2016) The oculus rift film experience: a case study on understanding films in a head-mounted display. In: *8th international conference on intelligent technologies for interactive entertainment*. Utrecht, Netherlands. https://doi.org/10.1007/978-3-319-49616-0_19. Last accessed 01 Mar 2021
40. Condit (2016) Virtual reality is not the (immediate) future of film. *Engaged*. <http://www.engageadget.com/2015/03/16/virtual-reality-film-gaming/>
41. Franklin-Wallis (2016) Virtual reality will transform cinema in 2016. *Wired*. <http://www.wired.co.uk/news/archive/2016-01/27/virtual-reality-breaks-fourth-wall>
42. Brinkman WP, Hoekstra A, van Egmond R (2015) The effect of 3d audio and other audio techniques on virtual reality experience. *Stud Health Technol Inf* 219. <http://dx.doi.org/https://doi.org/10.3233/978-1-61499-595-1-44>. Last accessed 15 Apr 2021
43. Kobayashi M, Ueno K, Ise S (2015) The effects of spatialized sounds on the sense of presence in auditory virtual environments: a psychological and physiological study. *Presence: Teleoper Vir Environ* 24:163–174. https://doi.org/10.1162/PRES_a_00226. Last accessed 05 Apr 2021
44. Ding N, Zhou W, Fung A (2018) Emotional effect of Cinematic VR compared with traditional 2D film. *Telematics Inf* 35. <https://doi.org/10.1016/j.tele.2018.04.003>. Last accessed 10 Apr 2021

45. Vosmeer M, Schouten B (2017) Project orpheus a research study into 360° cinematic VR. In: TVX '17: proceedings of the 2017 ACM international conference on interactive experiences for TV and online video, pp 85–90. <https://doi.org/10.1145/3077548.3077559>. Last accessed 25 Mar 2021
46. Rothe S, Buschek D, Hussmann H (2019) Guidance in cinematic virtual reality- taxonomy, research status and challenges. *Multimodal Technol Interact* 3(1):19. <https://doi.org/10.3390/mti3010019>. Last accessed 20 Feb 2021
47. Frintrop S, Rome E, Christensen HI (2010) Computational visual attention systems and their cognitive foundations. *ACM Trans Appl Percept* 2010(7):1–39. <https://doi.org/10.1145/1658349.1658355>. Last accessed 15/02/2021
48. Borji A, Itti L (2013) State-of-the-art in visual attention modeling. *IEEE Trans Pattern Anal Mach Intell* 35:185–207. <https://doi.org/10.1109/TPAMI.2012.89>. Last accessed 20/02/2021
49. Nielsen TL, Matias BM, Hartmeyer DS, Ljung T, Nilsson N, Nordahl R, Serafin S (2016) Missing the point: an exploration of how to guide users' attention during cinematic virtual reality. pp 229–232. <https://doi.org/10.1145/2993369.2993405>. Last accessed 02 Apr 2021
50. Danieau F, Guillo A, Dore R (2017) Attention guidance for immersive video content in head-mounted displays. 205–206. 2017 IEEE Virtual Reality (VR). <https://doi.org/10.1109/VR.2017.7892248>. Last accessed 07 Mar 2021
51. Van der Burg E, Olivers C, Bronkhorst A, Theeuwes J (2008) Pip and Pop: Nonspatial Auditory Signals Improve Spatial Visual Search. *Journal of experimental psychology. Human perception and performance* 34. <https://doi.org/10.1037/0096-1523.34.5.1053>. Last accessed 01 Apr 23021
52. Rothe S, Hußmann H (2018) Guiding the viewer in cinematic virtual reality by diegetic cues. In: 23rd ACM symposium on virtual reality software and technology, Switzerland, pp 101–117. <https://doi.org/10.1145/3139131.3143421>. Last accessed 20 Mar 2021
53. Høeg E, Gerry L, Thomsen L, Nilsson N, Serafin S (2017) Binaural sound reduces reaction time in a virtual reality search task. In: 1–4. IEEE 3rd VR workshop on sonic interactions for virtual environments (SIVE). <https://doi.org/10.1109/SIVE.2017.7901610>. Last accessed 25 Feb 2021
54. Mendonça C, Rummukainen O, Pulkki V (2015) 3D sound can have a negative impact on the perception of visual content in audiovisual reproductions. In: 21st international conference on auditory display (ICAD2015). Graz, Austria, pp 1–6. <http://hdl.handle.net/1853/54119>, Last accessed 10 Apr 2021