Leveraging immersive audio production techniques to enhance cultural and historical understandings in a museum setting

Frank Duchêne

MASTER'S THESIS
November 2022

Master in Emerging Media

ABSTRACT

Tampereen ammattikorkeakoulu
Tampere University of Applied Sciences
Master in Emerging Media

FRANK DUCHENE:
Leveraging immersive audio production techniques to enhance cultural and historical understandings in a museum setting

Master’s thesis 58 pages
November 2022

This thesis investigated how audio post-production techniques that have been proven successful for film and TV can be applied to producing VR and AR experiences that use immersive audio to support cultural and historical understandings and promote cultural heritage in a museum setting.

The immersive sound design of two interactive installations at the Royal Museum of Fine Arts in Antwerp, for which the author was responsible, provided the opportunity to explore how traditional audio post-production techniques, commonly used for linear audio-visual storytelling, differ when applied in these scenarios. Furthermore, research was done to map the workflows that allow producing immersive sound without initial access to the VR-playback device and access to the premises and installations of a location-based AR experience.

Applying two workflows for one VR and one AR experience in the Royal Museum of Fine Arts in Antwerp shows that a sound designer can use the more traditional workflows as a basis to create emerging media content. Still, additional productional steps and strategies are necessary to evaluate and optimize the final result. For creating immersive audio content, this research confirms that Ambisonics for VR and channel-based audio for a location-based AR experience are valid choices, even without initial access to the VR-playback device or the possibility of testing the AR installation of a location-based experience. However, in obtaining successful implementation of an AR experience, the environmental conditions play a significant role and should be considered during off-site production.

Keywords: immersive audio, Ambisonics, AR, VR, museum, cultural, post-production techniques, workflow

CONTENTS

ABBREVIATIONS AND TERMS

Audio post-production	Envelops the various processes in which all the audio-related elements for moving pictures are developed. Sound editing, sound design, sound effects, and Foley are part of the post-production practice, depending on the type of project.
Augmented Reality	Augmented reality overlays digitally produced information over physical environments in real time.
AR	See Augmented Reality
dB(C)	dB SPL C-weighted
dB SPL C-weighted	Expressed in decibels; the sound pressure level is an objective measure of the external measured level of a sound. The C defines a form of an electrical filter designed to imitate the relative sensitivity of the human ear to various frequencies at high sound pressure levels.
Headroom	In an audio system, headroom is the ratio between the average level for which the system is designed and the maximum undistorted signal the system can handle.
Head tracking	Head tracking permits an application to distinguish and classify a user's head movements. In audio production, head tracking will steer the rotation of the sound stage, mimicking real-life experiences.
Head-locked	A head-locked audio track is immune to a user’s head movements. This approach is often used for elements one wants to keep static to reduce confusion, e.g., music or voice-over in a VR setting.
Imaging	Refers to the localization of sounds when reproduced in a sound system.
Immersive audio	An audio technology that does not just represent surround sound in the horizontal plane as it also encompasses significant height information. This permits sound to be perceived as being above the listener.
Loudspeaker-agnostic	Indicates that an auditory signal is not restricted to one

	loudspeaker setup but can be played back on a wide variety of setups without losing all of its information.
Pro Tools	A professional digital audio workstation by AVID, commonly used for audio post-production and in music productions.
Transient	A short-duration energy burst, with high amplitude, at the start of a sound.
Virtual Reality	A computer-generated three-dimensional, artificial environment simulating an experience in which users feel immersed, and they can explore by manipulating objects or execute a series of actions.

1 INTRODUCTION

This thesis investigates how immersive audio post-production techniques that have been proven successful for film and TV can be applied in VR and AR experiences that support cultural and historical understandings and promote cultural heritage in a museum setting.

Audio post-production encompasses the various processes in which all the audio-related elements for moving pictures are developed. This includes editing sound and dialogue, sound design, and sound effects. One of the aims of the audio post-production process is stated by Wyatt and Amyes (2004) as “to complete the illusion of reality and perspective through the use of sound effects and the recreation of natural acoustics in the mix, using equalization and artificial reverbs” (Wyatt & Amyes 2004).

The author was approached by the Belgian production company De chinezen (De chinezen NV 2022) in April 2022 to create the sound design, which involves audio post-production, for two location-based immersive installations at the KMSKA, the Royal Museum of Fine Arts Antwerp (KMSKA 2022), five months before its re-opening. These initiatives aim to broaden the public’s cultural and historical understanding of the Flemish Masters on which the museum focuses.

Museums have been using various new forms of storytelling through technology as its entertaining, educational, and engaging capabilities allow them to attract more visitors (Carlsson 2020). This evolution in “offering varied experiences for education, enjoyment, reflection and knowledge sharing” also concurs with the definition for museums, which the International Council of Museums ICOM adopted in August 2022 (ICOM 2022):

> “A museum is a not-for-profit, permanent institution in the service of society that researches, collects, conserves, interprets, and exhibits tangible and intangible heritage. Open to the public, accessible and inclusive, museums foster diversity and sustainability. They operate and communicate ethically, professionally and with the participation of communities, offering varied experiences for education, enjoyment, reflection and knowledge sharing.” (ICOM 2022)

Emerging technologies, such as VR, have quickly gained approval and positive reception in cultural heritage (Carrozzino & Bergamasco 2010). Examples include the virtual reconstruction of a medieval cathedral (Grazioli 2020), and the virtual visits that some museums offer online (Wilson 2020), or onsite (Kennedy 2015). As VR and AR technologies provide a visual component that aims to be as immersive as possible, the audio side can further enhance that experience by convincing the users of being submerged in the auditory realm that supports the images they see. This is where immersive audio production can contribute significantly to making such VR and AR experiences immediate.

One can currently distinguish two approaches in audio productions for VR. The first one is commonly found in VR gaming experiences and uses real-time rendering of audio assets in game engines such as FMOD (FMOD 2022) and WWise (Audiokinetic 2022), allowing interactive audio creation in non-linear storytelling scenarios while a player makes individual choices within a game. Because of its real-time rendering of the audio information, this method is often called an ‘online’ process (Altman, Kraus, Susal & Tsingos 2016). The second approach is a proven traditional linear audio-visual storytelling method, such as film, documentaries, or cinematic content. It can be used for linear, less interactive VR experiences. In this ‘offline’ production process, audio engineers receive video material on which they develop and synchronize their sound design to deliver a final audio mix that is subsequently married with the finished visual production.

This thesis explores how the offline workflow used for linear storytelling can be applied to producing immersive audio for VR and AR experiences in a museum. It analyzes numerous elements in the productional phases, the challenges, and the choices to reach the initial goals and the deadline when producing immersive sound without initial access to the premises and installations.

1.1 Research objectives

This individual thesis's objectives are twofold: to examine how immersive audio production techniques can enhance cultural, historical, and scientific understandings in a museum setting when applied to emerging media while bringing together several parallel specializations the author developed over the years.

The research questions of the thesis are

1. How do the offline workflows, commonly used for linear storytelling, differ from those used for producing immersive audio for VR and AR experiences in the Royal Museum of Fine Arts in Antwerp?
2. What workflows permit achieving the goals and the deadline when producing immersive sound for a VR experience without initial access to the VR-playback device that will be used in the museum?
3. What workflows counter the challenges when producing immersive sound for an AR experience without initial access to the premises and installations of a location-based experience?

Following a practical analysis related to the use of Virtual Reality (VR) and Augmented Reality (AR) technologies in a cultural context, the audio production process of real-life scenarios for the two technologies will be described and analyzed. The various techniques include immersive audio production, sound design, storytelling, education, and research.

The thesis consists of both an artistic and written part. One VR and one AR production at the KMSKA in Antwerp represents the creative artwork. The written portion of the thesis, describing the production process, should help future researchers use a selection of emerging media, focusing on immersive audio production techniques, to convey cultural and historical storytelling.

1.2 Data management and privacy plan

A private, secure server was used to store all the relevant data for his research. All the participants and professionals who were approached for qualitative and quantitative data collection were allowed to remain anonymous for this publication. Each participant received an explicit data privacy statement that indicated that only information essential for the research's purpose would be collected.

When the author wishes to use any data and information identifying the participant, explicit consent will be requested. The author respects all intellectual property rights.

2 BACKGROUND INFORMATION

This chapter presents relevant background information related to the field of audio post-production for linear storytelling and immersive audio.

2.1 Diegetic and non-diegetic sound in audio-visual storytelling

In the production of audio-visual storytelling, the sounds and music used can be categorized into two archetypes. The sounds that are part of the world of the characters and story within the film are diegetic, while the non-diegetic sounds are outside the story space (Önen 2014). Diegetic sound promotes the implied reality of a scene. In contrast, non-diegetic sounds are heard exclusively by the audience, often providing them with more information than the characters, and can promote a sense of fantasy (Beauchamp 2013). The rustling of clothes, the dialogues among the characters, and ambient night noises are illustrations of diegetic sound. The soundtrack of a movie or a voice-over are examples of non-diegetic sounds. Sometimes sounds cannot be posited clearly as diegetic or non-diegetic. These can be labeled as transdiegetic (Jørgensen 2007).

2.2 Immersive audio technologies

Immersive audio is one of the emerging technologies in cinema, TV, VR/AR, videogames, music, and museums to immerse audiences and improve their experiences (Fonseca 2020). While stereo sound lets a listener localize sources as being close or further away, along a line between the two loudspeakers, it cannot convey a sense of height or that sound is coming from behind us (Inglis 2022). Immersive sound, also called 3D Audio, allows sound designers to represent sounds in horizontal and vertical planes to be perceived as being around and above the listener.

Four approaches are at the center of immersive audio technologies. These are

- channel-based audio
- object-based audio
- Ambisonics
- binaural audio

2.2.1 Channel-based audio

In channel-based audio, the relationship between the number of speakers and channels is fixed. One audio channel can only contain the information for one speaker (mono). With two audio channels, a stereo image can be reproduced. Adding speakers and channels can extend the reproduced sound stage and surround the listener. Various surround sound formats have traditionally been channel-based. Figure 1 shows a setup in which five channels carry the signals to five speakers providing channel-based surround sound.

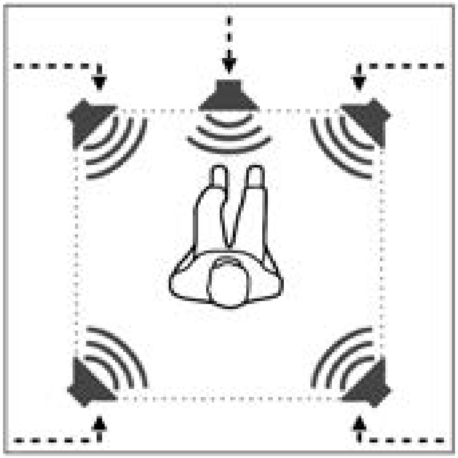


Figure 1. Channel-based audio in 5.0 setup, with five channels providing unique content to five loudspeakers.

The disadvantages of this technology are that the number of speakers dictates the number of channels, and one needs to know precisely the speaker arrangement on which the sound will be reproduced (Fonseca, 2020). The advantage of this approach is that a sound designer can predict how the speakers will reproduce the various sounds to the relevant speakers in the mixing process.

2.2.2 Object-based audio

In object-based audio, the relationship between the channel count and speaker count is no longer fixed because the distribution of all the sounds occurs during reproduction. The layout and number of speakers used in a setup can vary significantly while keeping the same content. During production, the sound designer will identify the localization of a sound using metadata, so an algorithm can use that information to allocate that sound, now called an object, within the setup it controls. An example of an object-based setup in which an algorithm positions the sound upon playback can be seen in Figure 2.

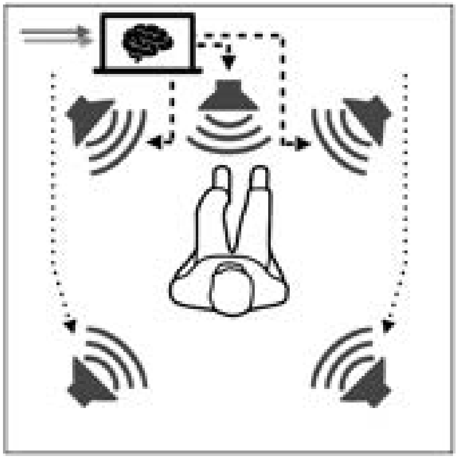


Figure 2. Object-based audio in a 5.0 setup. An algorithm assigns a sound to its location based on metadata during reproduction.

The benefit of providing a single immersive audio mix that can be played back on various speaker setups is clear, on the condition that the algorithms at the reproduction sides can decode the content accurately. But Nicolas Tsingos points out that the file sizes of an object-based audio mix can vary greatly, as this depends on the number and amount of time audio objects are used in a mix (Roginska & Geluso 2018).

2.2.3 Ambisonics

Ambisonics is a spatial audio format used in most 360-degree videos on YouTube and many VR applications. One of its significant advantages lies in its ability to easily rotate the whole sound stage, making it ideal for head-tracking as users turn their heads or when a system, such as VR, requires interactivity upon playback.

Ambisonics is an audio format that carries full-sphere surround sound. The amount of information in the Ambisonics signal dictates the accuracy of the imaging of the sound sources. So-called First-Order Ambisonics (FOA) will produce a sound stage with less spatial resolution than Second-Order (SOA) or Higher-Order Ambisonics (HOA) (Neukom 2006).

Figure 3 represents the full-sphere surround sound that can be captured, created, and reproduced using Ambisonics.

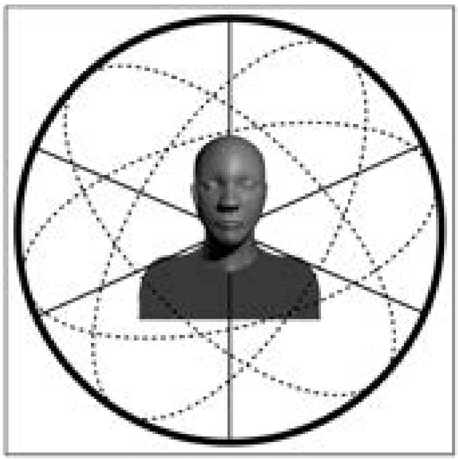


Figure 3. Ambisonics is a full-sphere surround format that needs to be decoded to loudspeakers or headphones.

Depending on the playback system, one decodes the spherical information to the desired speaker layout or headphones. In Ambisonics technology, headroom is limited because of the Omni information represented in the first channel. Peak-level management is, therefore, critical. Nishiyama et al. mention ‘the importance of the head-locked track in VR spatial sound production’ and propose considering additional head-locked channels during production. This hybrid approach offers more loudness than relying entirely on the Ambisonics bus (Nishiyama, Nagata, Ogata, Hasue & Kashiwagi 2018).

2.2.4 Binaural audio

Binaural audio is the reproduction of immersive sound using headphones. Figure 4 shows the auditory illusion that the reproduction of a binaural signal over headphones can create.

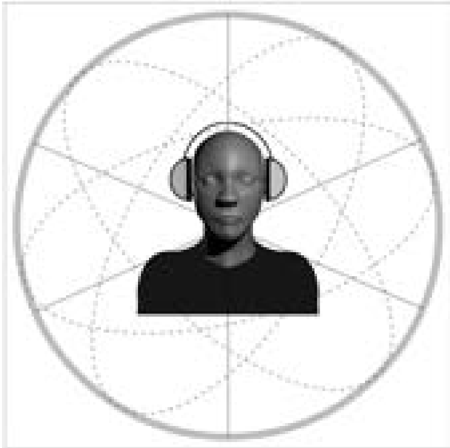


Figure 4. Binaural audio can optimally be experienced over headphones.

Originally, binaural recordings were made using two microphones placed in the ear canals of a so-called dummy head, after which the recording was played back over headphones. These days, using various applications, the three technologies already described, channel-based audio, object-based audio, and Ambisonics, can be converted to binaural. Edgar Choueiri (2018) shows that binaural audio can be reproduced through loudspeakers (Roginska & Geluso 2018). Still, Nuno Fonseca points out that it works best if the signal is reproduced over headphones (Fonseca 2020).

2.3 The importance of sound in VR

While some VR projects develop the visual component only, the audio feature can play a dramatic role in enhancing the overall experience. Staff Technical Writer for the Audio Engineering Society Journal Francis Rumsey notes that the element that creates believability in a VR environment is sound (Rumsey 2016). This view is supported by both audio technology developer Adam Levenson, who states that in VR, a sound has to be credible in the space it is in (Rumsey 2018), and composer and sound designer Anne Jimkes who expresses her belief that sound is fifty percent of the experience in VR (Jimkes 2017). Nishiyama et al. note that video and audio quality are crucial to the immersiveness of virtual reality, as is the interaction between them and the interactivity of the contents (Nishiyama et al. 2018).

Serafin observed that sound design models and ideas for virtual reality simulations could be retrieved from the entertainment industry, as the productions of film and video games understand the significance of sound to add presence and immersion (Serafin 2004).

While a large community of professionals adapts the workflow in this field, it is, in origin, tailored to linear audio-visual storytelling, with minimal interactive possibilities for the viewer. The sound design aims to help the viewer to understand the story and its setting better, pay attention to specific elements, and look in or go in a particular direction. Therefore, it is of interest to investigate how well-established workflows in audio-post production can be applied in the creation of VR and AR experiences.

3 CASE STUDY 1: VR in a museum setting

Through technology, museums have been using new forms of storytelling as its entertaining, educational, and engaging capabilities allow them to attract more visitors (Carlsson 2020). Wyman et al. indicate that in many places, museums now favor multi-faceted experiences, inviting conversation and interaction with their visitors (Wyman et al. 2011).

These technology-driven methods can be experienced at the re-opened KMSKA (KMSKA 2022), the Royal Museum of Fine Arts in Antwerp. It hosts countless Flemish masters, a conservation studio, and a renowned research institution and had been closed for over a decade for a €100 million renovation (Picture 1). Before its re-opening, director Carmen Willems pointed out to journalist Thomas van Huut that museums are often primarily seen as knowledge institutions but that the KMSKA wants to go beyond conveying facts and touch visitors in their emotionality:

> Only when you get to that layer, then the art will stay with you, and you can do something with it in your own life (van Huut 2022)[1].

Picture 1. View on the front entrance of the KMSKA during its renovation, 2022.

[1] "Pas als je op die laag komt, dan gaat de kunst je bijblijven, en kun je er iets mee in je eigen leven" (van Huut 2022)

To reach that goal, the KMSKA commissioned the Belgian production company ‘The Chinese’ (De chinezen NV 2022) to develop a series of experiences for visitors to enhance their visit (Jansen 2022). Two are based on emerging technologies and offer an opportunity to research immersive production techniques. One is focused on a historical recreation of the studio by a 17th-century painter, primarily based on the atelier of Pieter Paul Rubens. A second experience, as described in Chapter 4, consists of large projections of animations inspired by specific parts of four famous paintings to educate the audience about the rich details in works of art that are easily overlooked (van Huut 2022). A dedicated team designed and produced all the graphics, animations, and renderings for these two installations. The author of this thesis was responsible for the audio design and implementation of both projects. Work for producing the immersive audio content commenced in April 2022, five months before the official opening on September 24, 2022.

3.1 The studio of Pieter Paul Rubens in VR

The first experience that allowed the author to research a workflow for linear storytelling in VR consisted of an animated 360-degree recreation of an early 17th-century painter’s studio. This depiction is based on the house Pieter Paul Rubens bought in 1610 (Manderyck, Plomteux, Steyaert & Van Driessche 2019). Using a VR headset, the visitor assumes a position at the master’s desk, overlooking various activities, such as modeling, sketching, grinding pigments, and painting in his studio. Multiple characters with specific roles are animated in a two-minute loop. The whole illustration is in grayscale (Picture 2).

Picture 2. Screenshot from a preliminary version of the animated monoscopic 360-degree video (Courtesy De chinezen).

Eight Pico 2G 4K VR headsets (Pico G2 4K 2022), each containing the same content, are provided by the museum to its visitors in a specific room next to the Rubens' Hall. No one on the team had experience with this particular VR headset, and at the start of the sound design, the specific type of VR headset was unavailable due to supply chain problems.

Furthermore, it is essential to note that the museum's visitors receive minimum information regarding the VR experience and are invited to handle the headsets without additional help and with only light supervision. At the same time, Carrozzino & Bergamasco (2010) argue that it is imperative to immediately engage the users, permitting them to interact with fast learning curves in a natural way. Therefore, three important questions arise. What strategy allows a sound designer to start the production of immersive audio content without access to the final device? How can the final experience be immediate? Moreover, how can the handling of the devices by the public remain as simple as possible?

3.1.1 Research background

The research started in April 2022 and consisted of an initial literature review of technical and historical sources, followed by the soundscape production for the VR experience. Shivappa et al. (2016) indicate that both authentic spatial audio production and the need to track and adapt the audio scene to the listener's head position and orientation are compulsory for creating the auditory illusion of being 'present' in the virtual world, resulting in a fully immersive and convincing VR experience (Shivappa, Morrell, Sen, Peters & Salehin 2016). Altman et al. (2016) argue that creating a captivating mix for VR audio content usually necessitates going beyond reality to provide an enhanced viewpoint on the action. This vision coincides with the suspension of disbelief paradigm (Chandler 1996), which in film production dictates that the sounds one hears do not necessarily have to be authentic in origin but should support the story and the emotions triggered by the action on screen.

Because of the historical setting of this particular VR experience and its intended function to educate the public on the production of paintings in the 17th century, the first step in the research thus focused on gathering historical information on the premises and the activities both in the studio and the city, specifically around its location sometime after its acquisition by Rubens. This approach permits identifying relatable sounds that could be produced and contribute to the overall experience.

Rubens bought his residence in 1610 from the family Gielis-Thys, who fled Antwerp and relocated to Amsterdam because of religious and political tensions in the Low Countries. However, he only moved to the new residence sometime after 1614, and significant renovations (Maclot 2016). This timing was therefore used to reference possible sounds.

According to historian Petra Maclot, Rubens' studio was located just outside the city center (Picture 3) but, at that time, was surrounded by a busy passageway in the north corner and a large garden on the southeast side of the studio (Maclot 2016).

Picture 3. Detail of the map of Antwerp by Virgilius Bononiensis from 1565. The location of the future Rubenshuis can be spotted just above the double tree line. (Courtesy Museum Plantin en Moretus).

The studio had large windows to allow natural light into the workshop. To research these types of windows a building of the same era with original 17th-century windows in the neighboring city of Mechlin was visited. It became apparent that the acoustic insulation of this type of glass is so limited that birdsong can be easily heard inside the room. One can therefore argue that louder outside activities from the city would have bled into the studio as well.

While the building still exists (Picture 4) and the room of the former atelier can be visited, some elements have undergone various changes throughout the centuries. Therefore, the acoustics of the atelier need to be reimagined.

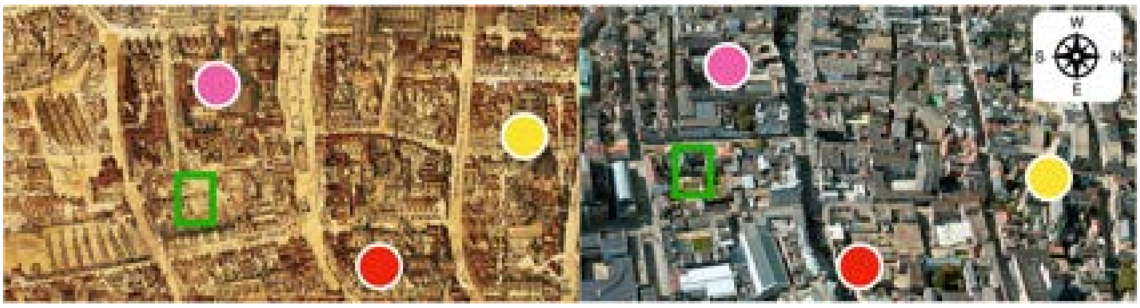
Picture 4. The Rubenshuis in its current state. View from the west side.

Interestingly, the floorplan of the studio has a relatively simple rectangular shape (Picture 5), which corresponds well with most of today’s reverb algorithms that simulate the room acoustics (Rumsey 2018).

Picture 5. Satellite image of the Rubenshuis in 2022. The location of the former studio is marked in green. The top of the picture is the west side. The garden is situated in the east part of the premises (Courtesy of Apple Maps).

In 1625 a Carmelite monastery was built across the street on the west side of the Rubenshuis (Manderyck, Plomteux & Steyaert 1979). Another monastery of cellebroeders (Linters & Plomteux 1979) was located less than 300 meters on the north side of the studio, close to the Saint-James church. Monasteries would use bells to summon the monks to the canonical hours (Forlenza 2017), with the Sext canonical hour being around noon and None around 3 P.M. (Till 2008). The clockwork mechanism of the Saint-James church gave the hour from 1556 until the Germans looted the bells in 1943 (Topa 2021).

Because of the short distance between these locations (Picture 6), and the studio’s large windows, one can presume that the tolling of the bells would be heard in the studio.



Picture 6. The locations of churches and monasteries around the Rubenshuis in the first quarter of the 17th century (left) and today (right). The location of the Rubenshuis is indicated in light green, Saint-James's church in yellow, the Carmelite monastery in pink, and the cellebroeders in red.

As the scene of the painter's workshop is timed sometime in the afternoon, bells, supposedly of the church and the Carmelite monastery signaling the canonical hour of the None, are additional elements used to connect the visible scene with a suggested outside world.

The barking of dogs is also a short animal sound that can imply the outside of the studio. Specific breeds can be seen in several of Rubens' paintings. Although a symbolic value is attributed to these depictions, including loyalty and indicating social class, the approach to using sounds of these specific dogs concurs with the historical approach that defines this project. Both Frans Snijders, who specialized in painting animals and contributed to the Rubens workshop, as the artist himself, have included greyhounds (Sotheby's 2022) in their art (Picture 7); barking sounds of this breed are used as one element of the outside world.

In addition, inspiration for sounds coming from outside the atelier was found in literary descriptions and paintings from the Rubens era. In his book Antwerp: The Glory Years, author Michael Pye mentions the various animals that would be found in the vicinity of even the most prominent houses in the city around the turn of the 16th century. These included pigs, geese, ducks, and stray dogs (Pye 2021). Therefore, these animals offer opportunities for relatable sounds outside of the studio. A blackbird's birdsong and seagulls' call can still be heard today. This provides an opportunity to use these natural sounds, and place them at different azimuth and elevation angles than the animated characters, to expand the immersive soundstage (Mironovs & Lee 2019).

Picture 7. Diana and her nymphs hunting, 1636, Studio of Peter Paul Rubens (courtesy of Museo del Prado).

Hartmann indicates that in a reverberant setting, listeners will localize sounds based on their transients (Hartmann 1993). As the acoustics of the simulated space of the virtual atelier conforms to this condition, this insight could be useful to help the viewer to pinpoint certain noises. The fireplace, grinding of the pigments, the church bells, and the wooden scaffolding squeaking are among the sounds that produce sharp transients and can therefore be used to orientate the observer more easily via auditory means.

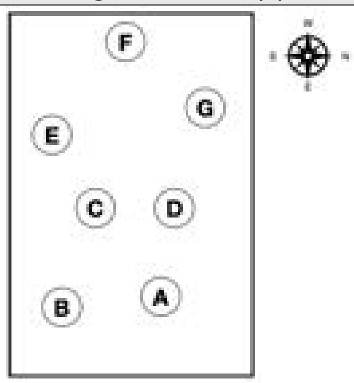
3.1.2 Historical storytelling and sound design

In this VR experience, the storytelling consists of a single scene in the painter's studio at a specific moment where seven characters perform their tasks in a linear sequence. The techniques demonstrated in the video include grinding color pigments, sketching a female model on paper and wood, and painting. An overview is presented in Table 1. These actions were animated using motion capture, yet all characters stay in their respective position. A curious cat dwelling around the studio was added for dramatic purposes. This introduces a familiar domestic setting, nudges the viewer to follow its movements, appreciate the actions of the various characters, and promotes a less static sound design.

There is no storytelling arc, no dialogue, or specific dramatization. However, the VR experience presents the viewer with a historically informed two-minute-long impression of how artists like Rubens organized their workshops, allowing them to produce many works of art.

The animated, stereoscopic scene has no sound. It, therefore, needs tailored sound design, which was obtained by adding all the minute sounds for the animations, a process known as "Foley," after its inventor Jack Foley (Radhika 2015).

Table 1. Overview of the eight characters, their tasks, and positions in the virtual studio.

	
Cat (A)	**Two apprentices sketching on paper and wood (B)**
	
Female model (C)	**Painter looking at model (D)**
	
Painter (E)	**Painter finishing details on large work of art (F)**
	
Assistant grinding color pigments (G)	The positions of the characters in the virtual studio

The approach for this project implies historical storytelling because of the setting and its purpose. The sound design must support the illusion of all the characters' actions. However, it must also be referenced to the sound decor of both the location and the era without losing the entertainment factor. This means that the timbres, reverb characteristics of the atelier, and localization should support the viewers' experience to the maximum amount, but with elements organized in such a way that they avoid becoming predictable but trigger the viewer's imagination throughout the loop. As the technology for recording sound was only invented in the 19th century, no soundscape of Rubens' time exists and needs to be reimagined.

Because the scene can only be experienced from the painter's simulated position in his studio, the result is, in effect, a 360-degree video. This entails that the audio production can be centered around a system with 3DoF (three degrees of freedom), which tracks the pitch, yaw, and roll movements of the user's head, but essentially assumes an otherwise fixed user. 3DoF systems allow the sound designer to prepare particular sounds for everything in the scene while sidestepping the data-driven procedural creation of sound in virtual environments (Hamilton 2015). The audio designer can use Ambisonics' audio processing to place the audio elements in the sound field (Rumsey 2018).

3.1.3 Implementation

A two-minute 360-degree animated video without sound was produced. For the audio post-production, this video was imported as a monoscopic mp4 file into a Pro Tools session, allowing the creation of all the sounds and their localization to match the viewer's experience. Throughout the process, subsequent animation versions were rendered and used for post-production, each time with more visual details but with the main actions and overall length unaltered.

To convey an immersive experience to the viewers when using the VR headset, the audio must have similar characteristics that allow the direction of the reproduced sound to follow the head movements of the viewer. The combination of Ambisonics is the technology most suited for this project because the binaural rendering on headphones and the head-tracking can be integrated with the Pico

2G 4K VR goggles. The team settled on using a nine-channel, Second-Order Ambisonics signal (SOA) instead of the more basic four-channel First-Order Ambisonics approach (FOA) because of its higher localization accuracy. Xie, Mai, and Zhong (2017) have shown that the binaural rendering of an Ambisonics signal with head rotation and tilting also provides a dynamic cue for vertical localization, which can only enhance the immersive audio experience of this historical scene. Sound designer-engineer Glenn Forsythe explains that when viewing the final product of a VR experience, its spatial audio needs to be interactive (Forsythe 2018). Interestingly, Clark (2022) writes that the same audio content has a much more potent influence on listeners' perceptions, decisions, and behaviors when listening on headphones than when hearing on loudspeakers. Researcher On Amir notes that using voice-over in headphones promotes connectivity and empathy towards the speakers and has persuasive power because of the in-head localization phenomenon, which makes a narrator sound as if they are inside your head (Matei 2022).

This VR experience is, however, without any voice-over or music, based on a single setting, and the focus is entirely on recreating the illusion of being in the scene. Therefore, all the sounds are diegetic, without any need for a head-locked track that is immune to the user's head movements.

The Foley mimics the individual characters' actions by handling various materials similar to those found in the workshop, including stone, chalk, wool, sand, and oak. The author recorded and edited most of these sounds in his studio. The materials used for the handling noises were referencing historical ones. The recording was done using a Sennheiser MKH40 microphone and a Sound Devices MixPre-6 mobile recorder for ease of use (Picture 8). Some additional sounds were recorded in a space that has a similar acoustic as Rubens' studio. The sounds for the fireplace, church bells, carts, and horses were procured through commercially available sound libraries.

Picture 8. The recording setup for the custom-made sounds.

The sounds of some indigenous birds, including blackbirds and chickens, were obtained by recording them in and around the city using the mobile setup. A similar approach was used to record a general atmosphere for the atelier, albeit using the setup with a Rode NT-SF1 First-Order Ambisonics microphone.

Two spatialization Ambisonics-compatible software solutions were tested: Audio Ease 360 pan suite (Audio Ease BV 2022) and dearVR PRO (Dear Reality 2022). Both can position individual elements within a 360-degree field and add ambiance through artificial reverb and early reflections, allowing for positioning sound sources and creating virtual acoustics. While 360 Pan Suite relies on a single reverb for all the elements positioned within the room, dearVR PRO allows the adjustment of reverb parameters for every element. The positioning of individual components seemed more intuitive and sounded more convincing with dearVR PRO, possibly because of its tweaking capabilities of the reverb and reflections (Picture 9).

Picture 9. Screenshot of the dearVR PRO-plug-in, which allows positioning individual elements within the Second-Order Ambisonics realm while tailoring acoustic characteristics. These particular settings were used to position the cat on the table, close to the viewer.

One of the alternative plug-ins of Audio Ease offers a unique so-called radar view that accurately visualizes the location of the energy of the various sounds on the 360-degree video (Picture 10).

Picture 10. The 360radar plug-in from Audio Ease visualizes the sound's energy and localization. This screenshot shows the energy generated by the sound of the purring cat on the table.

With the Pico headset not yet available at the start of the production, the evaluation of the sound in relation to the 360-degree video was handled by Audio Ease's 360monitor plug-in. This software displays the 360-degree video on a regular computer screen while decoding the Ambisonics mix to binaural for headphones (Picture 11). The software can communicate via Bluetooth with an additional hardware tracker one attaches to the headphones to permit a dynamic preview and head-tracking. The result is a monitoring system that mimics the interactivity of a VR headset but without using one. The binaural audio was monitored using a Beyerdynamic DT 880 PRO headphones (Beyerdynamic 2022).

Picture 11. The 360monitor simulates the VR experience on a regular computer screen. It uses the movement of the head to look around in the 360-degree video and listen to the immersive audio mix.

A combination of the 360monitor, the 360radar, and the dearVR PRO software proved to be very effective for the production of the initial sound design and moving forward, pending the availability of the Pico headset.

It was only after crafting the first version of a two-minute soundscape within Pro-Tools that the Pico G2 4K headsets became available for testing.

The first problem was identified as it became apparent that this Pico model did not play back spatial audio. Various renderings that tested different formats, including .mp4 and .mkv with Second-Order Ambisonics, failed to offer precise audio playback. This problem was related to the inability of the headset's internal video player to decode any spatial audio content. To circumvent this issue, the VR producer of the team provided a simple test program, designed in Unity and delivered as an Android Package File (APK), allowing accurate playback of the Second-Order Ambisonics files in the headset.

A secondary problem consisted of the location of the loudspeakers on the headset. On this particular Pico model, these are positioned at the front of the goggles, while the successful binaural rendering of spatial audio relies on reproduction at ear level. In addition, this playback of the audio through the Pico headset's small loudspeakers resulted in a loud audio spill, unsuited for playback in a museum. With the deadline in mind, it was decided to research if the speakers could be bypassed and experiment with different additional headphones to see if the experience could be improved. Various headphones were tested, and smaller, over-ear headphones were a good combination regarding comfort, audio quality, and cost. The next step was to investigate a physical solution that blended the Pico headset with the headphones so that visitors would not need additional help using it. A prototype, built with small wooden pieces shaped on an artificial head and using Velcro to attach the construction to the device's headband, proved an approach that was not invasive to the Pico headset, keeping its warranty intact (Picture 12).

Picture 12. The prototype of the Pico G2 4K headset with additional over-ear headphones and the custom-built harness. The procured headphones are Positive Vibration II by House of Marley (House of Marley 2022).

The third problem presented itself as a lack of any instant streaming capabilities from a computer to the headset. While Pico does offer a Streaming Assistant for some of their other headsets, the Pico G2 4K cannot connect to a computer for an immediate evaluation of the audio and video content (Pico XR 2022). This implied numerous file transfers from the computer to the headset for offline assessment, reducing the time efficiency of the production process.

Figure 5 shows how the workflow encompassed the additional stage of rendering, encoding, and transferring the audio mix and preliminary animation to be evaluated on the Pico headset.

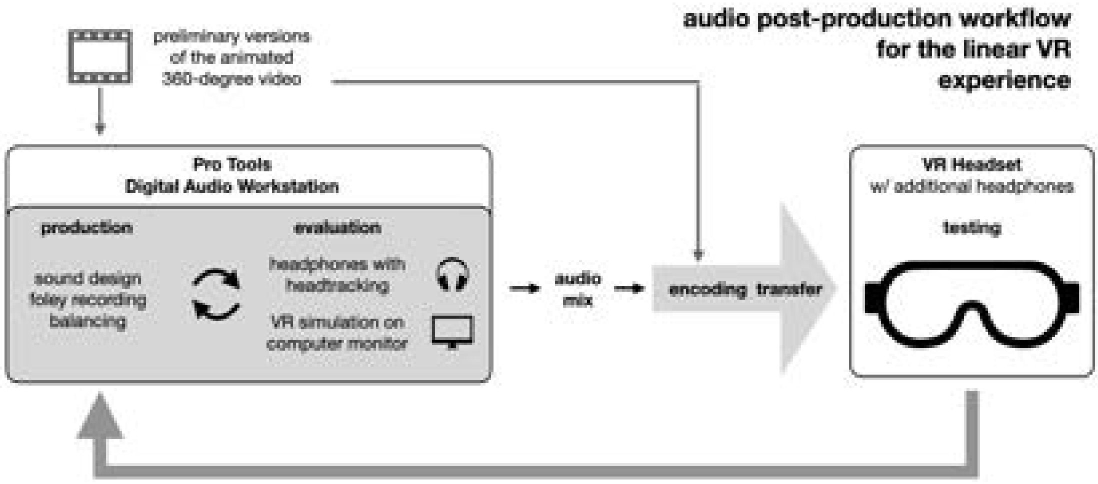


Figure 5. The audio post-production workflow for the linear VR experience on the Pico headset without streaming capabilities.

3.2 Results and conclusions

The final video (Picture 13) and audio versions were rendered two weeks before the deadline. The combined video and audio results were presented to the museum's curators for evaluation in two ways: on one Pico headset with internal speakers and one with additional headphones. No creative adjustments in the sound design had to be made, but because of a much better immersive experience, all parties considered the experience with the headphones superior. For aesthetic reasons, a wooden frame was preferred over the option of using a 3D-printed one. Ten frames with headphones were commissioned for all the headsets, and five were installed at the site, each with a cable attached to provide power and limit the risk of theft.

Picture 13. A monoscopic rendering of the final version of the 360-degree video (Courtesy De chinezen).

Five days before the official opening, a test audience was invited to the museum (Picture 14). The author observed from some distance thirty-three adult visitors who used the headsets without additional help to look for possible difficulties with handling the devices and to listen for unprovoked reactions.

Picture 14. Composite of visitors testing the VR experience in the museum.

The combination of the VR headset with prescription spectacles proved to be a challenge for three people. Two people had difficulties handling the equipment and their glasses simultaneously, and one remarked that his bifocal spectacles were blurring the virtual image. One person tried the VR headset for only a few seconds and expressed unpleasantness when listening to headphones with his hearing device. Three people needed clarification about whether to sit or keep standing when using the headset (Table 1).

Table 1 Observations of the test public, consisting of thirty-three adult visitors

VR headset	**Visitors**		**Handling difficulties (expressed or noticed)**			
	female	male	spectacles	power cable	Confusion Keep standing instead of sitting	Hearing device
	19	14	3	3	3	1

Several visitors spontaneously shared their impressions with each other, with the cat wandering around the studio mentioned to be the favorite character. Eleven visitors were questioned to gauge what presented information resonated with them. Eight referred their experience to Rubens' studio based solely on what they saw. Therefore, most of the testing public reached the initial goal of offering an enlightening and direct VR experience with simple handling of the devices.

The audio-post-production workflow for this VR experience was similar to those for linear cinematographic productions. Additional software was required, and some differences could be observed. In contrast, traditional audio-post production will place sounds near or far from the viewer by balancing the original sound with added spatial information generated by one artificial reverb engine, a far more convincing result was obtained in VR by adjusting reverb parameters for individual sounds, depending on their positions.

When producing an immersive sound design for VR without access to the final headset, a workable strategy can be achieved by evaluating the process on headphones and a computer screen using dedicated software. In the second phase, definitive testing on the VR device remains critical to identify technical and sensorial differences between them. This approach is less instant than if one could evaluate and test simultaneously but does not compromise the artistic results.

Furthermore, compared to traditional delivery of highly normalized formats for TV and film, VR headsets differ significantly, ensuing in a bespoke approach in which the testing phase of the relevant headset becomes a crucial part. This resulted in modifying hardware in the form of wooden frames and additional headphones. While this is clearly outside the scope of traditional audio post-production workflows for cinematographic storytelling, it helped to achieve a final, immediate experience and, with the handling of the VR devices by the public, remained as simple as possible.

The project validates Rumsey's view that sound creates believability in a VR environment (Rumsey 2016) and Levenson's statement that in VR, a sound has to be credible in the space it is in (Rumsey 2018). It also confirms that to create the auditory illusion of being 'present' in the virtual world, authentic spatial audio production and the need to track and adapt the audio scene to the listener's head position and orientation are compulsory (Shivappa et al. 2016). The combination of a high-quality VR headset with the addition of headphones further supports Nishiyama et al. note that video and audio quality are essential to the immersiveness of virtual reality, as is the interaction between them and the interactivity of the contents (Nishiyama et al. 2018).

4 CASE STUDY 2: AR in a museum setting

The second experience at the Royal Museum of Fine Arts in Antwerp is a location-based setup with interactive projections on all four walls in a dedicated room. The result can be considered augmented reality as it blends actual surfaces and surroundings with virtual projected ones. Sixteen individual loudspeakers, mounted near the room's ceiling, allow for multichannel playback and immersive audio.

This case could be used to research workflows that counter the challenges when producing immersive sound for a location-based AR experience without initial access to the premises and installations.

4.1 Projections at the KMSKA

The KMSKA uses a dedicated room in the museum, measuring eighteen by nine meters allowing for large, virtual, and animated curtains to be projected on the four walls with a height of ten meters. In four different scenarios, a virtual curtain is drawn across the room's four walls in forty seconds, after which visitors can interact and trigger extra animated elements. The inspiration for all the images that are shown comes from specific parts of famous paintings in the KMSKA (Picture 15). The main goal of this experience is to educate an audience to pay attention to the rich details one can find in works of art throughout the museum, which can be easily overlooked. The details refer to the following four paintings:

- ‘Madonna Surrounded by Seraphim and Cherubim’ by Jean Fouquet (KMSKA 2021)
- ‘God the Father with Singing and Music-making Angels’ by Hans Memling (KMSKA 2021)
- ‘Scaldis and Antverpia’ by Abraham Janssens I (KMSKA 2021)
- The Fall of the Rebel Angels by Frans Floris I (KMSKA 2021)

Picture 15. Composite of two of the four paintings used as an inspiration for the animated curtains: Fouquet's 'Madonna Surrounded by Seraphim and Cherubim' (left) and Memling's 'God the Father with Singing and Music-making Angels' at the right (Courtesy of the KMSKA).

Each of the four AR scenes, based on a particular painting, consists of four phases. Figure 6 shows the non-interactive and interactive parts of the four AR experiences.

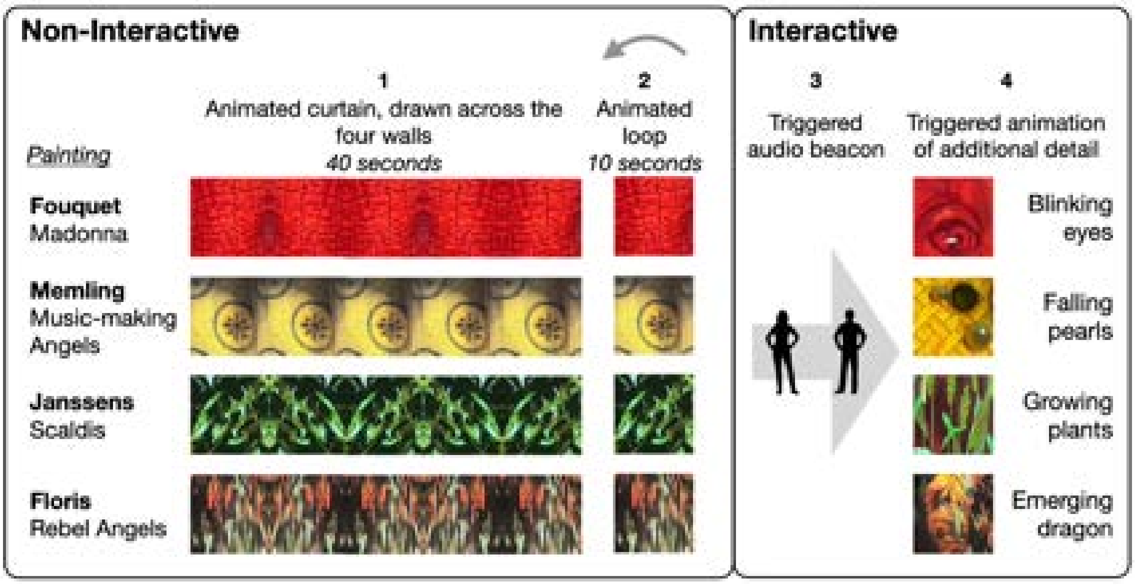


Figure 6. The scenarios and phases for the four AR experiences.

First, an animated curtain appears in one corner and looks drawn around the room, immersing the visitor fully with a 360-degree view of the rich details. Once this action is completed, a ten-second visual and audio loop takes over until the third and fourth phases, in which visitors trigger an audio beacon and subsequent extra animation by standing in one of four locations in the room, marked by a sticker on the floor (Picture 16). This triggered animation, inspired by another painting detail, is projected and disappears after a few seconds. While interactive, the timeline for these scenarios remains linear. Each of the scenes needs custom-made sound.

Picture 16. The AR installation during initial testing with visitors. The person on the left is standing on one of the stickers, indicating a point for triggering additional animations.

4.1.1 Research background

The research started in early May 2022 and consisted of an initial literature review of the location-based use of multimedia in museums. Moreover, several installations that use projections and audio or forms of AR in museums and historical sites were visited in July. This approach allowed for answering two questions: What are audio-visual technologies (AV) and audio-related storytelling techniques currently used for location-specific storytelling with multimedia and AR in museums or historical and cultural sites? Moreover, what are the environmental conditions found in these real-life situations? The various technologies for audio and video playback systems and storytelling techniques of four visited sites were mapped, with the sound pressure level for each one measured using the SPLnFFT app (Kardous & Shaw 2014) to obtain objective information and compare the artistic and technical possibilities with the audio installation at the KMSKA.

Augmented reality has been used in numerous museum exhibitions and cultural heritage sites and can be a catalyst for enjoyable learning surrounding and increase learning motivation (Chen & Lai 2021). Projections to enhance visitors' experience are used in museums worldwide (Arnold-de Simine 2012). Cases include The American National Aeronautics and Space Administration, NASA, offering multimedia and AR experiences as part of their Museum & Informal Education Alliance (NASA 2022), and Scan4Stories developing AR applications to make archived historical books available to visitors in a museum setting (Scan4Stories 2021). An excellent example in Europe of a single non-immersive approach in a museum was found in Lascaux IV in France, the institution devoted to the world-famous cave paintings (Lascaux 2022). In addition to the recreated paintings, a short movie, beamed in a dedicated room by one projector, is used to convey the surroundings of the Cro-Magnon people who inhabited the region some 20.000 years ago and created all the art in the cave (Picture 17).

The non-interactive video is accompanied by a soundscape based on diegetic sounds of animal noises and wind and a male voice-over explaining the site's history. The overall atmosphere in the darkened room is discreet. The sound pressure level of the background noise floor measured around 52 dB SPL C-weighted. Because of this limited noise floor, the audio can be played back at an enjoyable sound pressure level with a high degree of intelligibility.

Picture 17. The projection at the Lascaux IV Museum (courtesy of Lascaux IV International Center).

The second example of a practical, non-interactive approach is the projection used by the Bayeux Museum to support the storytelling of the origin of the Bayeux Tapestry (Bayeux Museum 2021). A fifteen-minute movie is projected in a dedicated darkened theater that combines animation and live-action scenes. Using voice-over and diegetic sounds, the film helps to evocate the Battle of Hastings and the genesis of the Bayeux Tapestry. In contrast to the Lascaux Museum, the theater is equipped with surround sound. The sound pressure level of the background noise floor in the theatre measured around 58 dB SPL C-weighted, still allowing a moderate playback level of the audio content with a high degree of clearness.

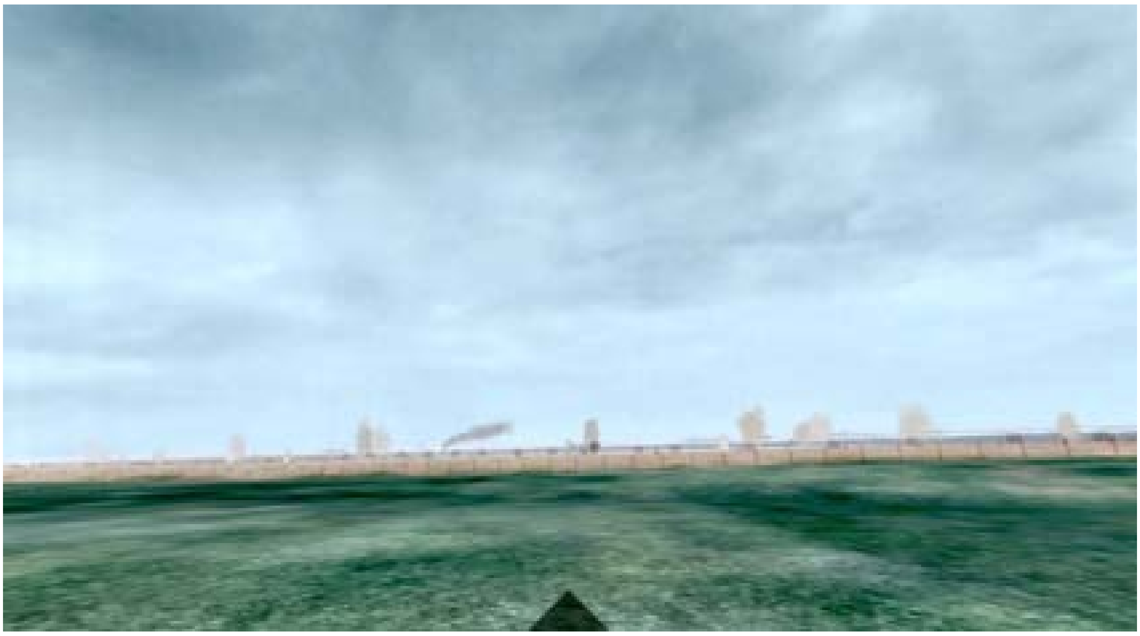
A use case of location-specific AR can be found in the European Parliament's Station Europe in Brussels, where an augmented reality model of the campus, presented through dedicated tablets, informs visitors of the history and activities of the EP (Station Europe sd). While this setup, blending a large-scale model with a virtual world in dimmed surroundings, is highly interactive and technically flawless, it relies only on text and animations to convey the information. It makes no use of any audio (Picture 18).

Picture 18. A tablet with the AR software at Station Europe in Brussels is used to overlay information and animations when pointed toward the real-life model of the European Parliament.

This lack of sound allows visitors to use the tablets in pairs without confusion from other visitors nearby who are simultaneously focusing on different parts of the story of the European Union on their tablets. However, the absence of sound significantly reduces the experience, confirming Nishiyama et al. observation of the quality of video and audio and its importance to experience immersiveness (Nishiyama, Nagata, et al. 2018).

The last example of a location-specific AR experience is the Omaha Beach app, developed by the Dept. of Media & Communication, University of Oslo, to recreate the assault on Omaha Beach in Normandy, France, in the early morning on D-Day in 1944 (Liestøl 2019). This app runs on personal smartphones and is designed to be only used at Omaha Beach WN 62 Strongpoint Fortifications, the precise location of the most potent opposition nest on the Omaha beach shoreline (D-Day-Tours 2013). The app uses diegetic sounds and voice-over and offers interactivity by tracking the phone's orientation.

This choice for location-based storytelling can result in a potentially powerful experience as it forces the user first to make an effort to download the app, physically locate the site and experience the AR at the unique place where the historical event happened (Picture 19).



Picture 19. Screenshot of the Omaha Beach AR app, displaying the attacks on D-Day (courtesy University of Oslo).

However, at the same time, this method can suffer from environmental issues. By testing the app with two different smartphones on 29 July 2022 on location, two real-life problems revealed themselves. At 15:40, the sunlight on the beach was still too strong to allow a good viewing experience on the 2019 Apple iPhone XR and a 2015 iPhone 5S (Picture 20), even when the display with a 1400:1 contrast ratio on the XR or 500:1 on the 5S was set to its maximum brightness (Apple 2022). It is important to note that this problem was acknowledged by the app's developers (Liestøl 2019).

Picture 20. The Omaha Beach AR app in real-life conditions on a 2015 iPhone 5S with its display set to maximum brightness.

The second problem was the background noise level of the sea. The loudness measured with the SPLnFFT app (Kardous & Shaw 2014) was 68.5 dB SPL C-weighted. This background noise, 16.5 dB or almost seven times louder than that at the Lascaux Museum, forces the user to set the volume of the smartphone to its loudest setting when using the internal speaker, resulting in a distorted sound when listening to the voice-over and the diegetic sound of guns and explosions, compromising the experience significantly.

Table 2 shows audio-visual technologies, storytelling techniques, and environmental conditions at the visited sites. None of the sites offer a fully interactive, immersive experience combining video and audio. Diegetic sounds are prominent in multiple cases, independent of the audio technologies used.

Table 2 Overview of the specifics of the location-based multimedia experiences

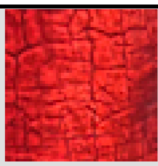
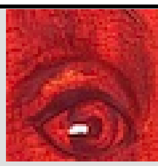
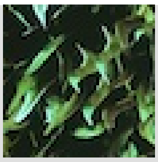
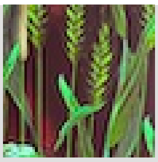
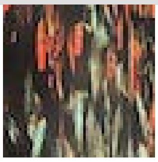
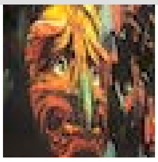
Location	AV-technologies (playback)		Storytelling techniques			Environmental conditions	
	Video	Audio	Type of sounds	Voice Over	interac-tive	Noise dB SPL C	Light
Lascaux IV	projection	stereo	diegetic	yes	no	52	Darkened
Bayeux Museum	projection	surround	diegetic	yes	no	58	Darkened
Station Europe	AR (tablet)	n/a	n/a	no	yes	n/a	Daylight
Omaha Beach app	AR (phone)	stereo	diegetic	yes	yes	68.5	Daylight

It is essential to conclude that environmental conditions play a significant role in the overall success of a location-based experience. This implies that the content should be tailored to specific circumstances to avoid the user's frustration.

4.1.2 Artistic storytelling and sound design

The AR installation at the KMSKA consists of a 360-degree video projection and immersive sound system in a large, darkened room. This installation wants to help visitors to discover their attention to detail in classic works of art and activate their imagination by presenting elements of four paintings of the collection. A virtual curtain is drawn across the room's four walls in forty seconds, after which visitors can interact and trigger additional animated elements. An overview of the various materials, triggers, and types of sounds can be seen in Table 3. The use of voice-over or music was considered counterproductive to this experience, as the exclusive use of diegetic sounds can support the experience's visual side and help the visitor discover details without being too explicit.

Table 3. Overview of the four paintings and corresponding design elements visitors can trigger by standing on one of four locations in the room, indicated by a sticker on the floor.

PAINTING	CURTAIN MATERIAL		ANIMATION TRIGGERED BY THE VISITORS		TYPE OF SOUNDS
Fouquet – Madonna	Leather		Blinking eyes		Diegetic
Memling – Music-making Angels	Gold brocade		Falling pearls		Diegetic
Janssens – Scaldis	Leaves		Growing plants		Diegetic
Floris – Rebel Angels	Feathers		Emerging dragon		Diegetic

Each scene was animated, and the movies were imported into a Pro Tools session, allowing for the creation of all the diegetic sounds. More detailed versions of these animations were provided during the development, with the main events and overall duration unchanged.

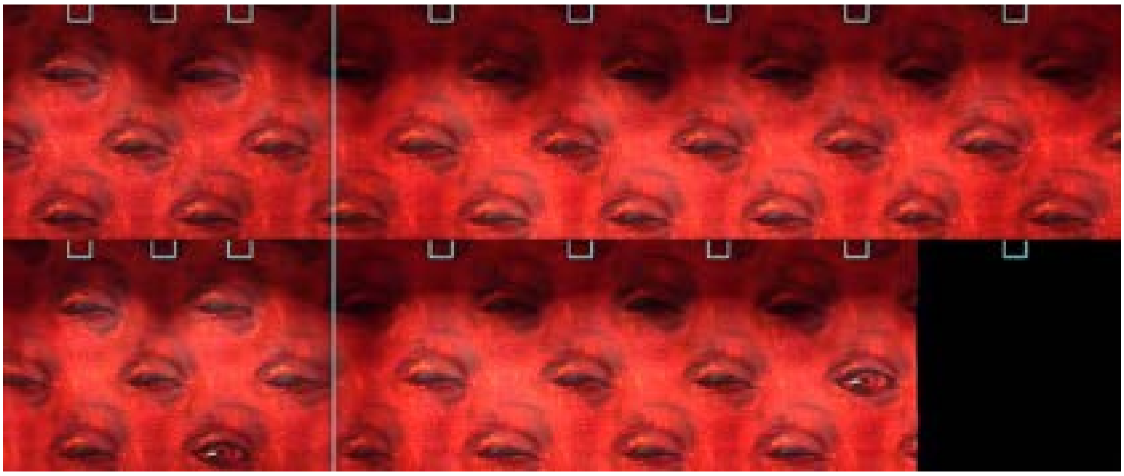
All the sound design for the AR installation at the KMSKA was recorded, created, and mixed in the author's professional studio on a 5.0 surround monitor system. The sixteen channels of the immersive audio mix were routed to the five speakers in this setup to simulate the room at the museum. Combining various Foley techniques (Picture 12) and sounds from commercial sound libraries, the audio for each scene was produced. Inspiration was found in the physical properties of the various curtains as well as the details from the paintings and their origins. Some examples include the manipulation of a small carriage on a model train that formed the basis for the curtain rails. The feather details, taken from The Fall of the Rebel Angels, were sonified by recording the handling of feathers from pigeons.

Audio beacons in museums are used to support visitors in finding their way and as a means of offering comprehensible information on museum content (Landau, Wiener, Naghshineh & Giusti 2005). The short audio beacons for the four scenarios in this project were developed to inform the visitors of the imminent animation they triggered and nudge their attention toward the location of the projections. All were composed with the music creation software Logic Pro (Apple 2022) and referenced the various paintings. Different audio samples of a boys' choir serve as the audio beacon in the Fouquet scenario. These samples are played back at the four positions of the triggers. When activated by the visitors, the result mimics a polyphonic composition, similar to the Cori Spezzati, or split choirs, the practice to place singers at separate locations when performing (Britannica n.d.) a performance style not uncommon in the early sixteenth century in Northern Italy (Schiltz & Ongaro 2018). For the Memling curtain, a lute, one of the musical instruments seen in the original painting, provided the basis for the sonification of the trigger. The growing plants are supported by a rising and descending melody on a harp and the rustling of leaves.

Picture 21. Examples of recordings of real-life materials for the sound design: A model train track formed the basis for the curtain rails (left) and handling bird feathers for The Fall of the Rebel Angels by Frans Floris I (right).

A vital aspect of the experience is the accurate localization of the sounds during all stages of the animations. As the number of speakers is known beforehand, and the installed speaker layout in the museum is fixed, but not referenced to a standardized immersive audio standard, or offering the rendering of object-based content, the choice to use channel-based audio for producing the immersive audio content is a valid one, as this approach does not rely on any additional audio processing at the playback side and allows a clear workflow in which every speaker is fed by its corresponding channel. This entails appropriately routing the various channels to the intended loudspeakers to translate the work from the studio to the dedicated location. Because the installation was not functional at the time of the production, a test file simulating all sixteen channels was generated to allow off-site testing of this approach at the company responsible for the final installation.

A floorplan mapped the exact location of the speakers in the room, and these measurements could be used to construct a blueprint to know how the sound design would be reproduced. Both the art director and graphics designer welcomed the proposal to visualize the location of the speakers in the preliminary videos provided for the sound design (Picture 22). By superimposing small squares that represented the location of the loudspeakers in the room, the accurate positioning of the various sounds during the drawing of the curtain was made feasible.



Picture 22. A screen capture of the curtain with details from Fouquet's painting around the 28-second moment in the animation. The small squares show the positions of the loudspeakers in the installation. The white vertical lines indicate the corners of the room.

This way, the sound design could be created pending a visit to the location. In anticipation, work was already being done on various versions of the animations to further refinement in the process, culminating in a workflow as can be seen in Figure 7.

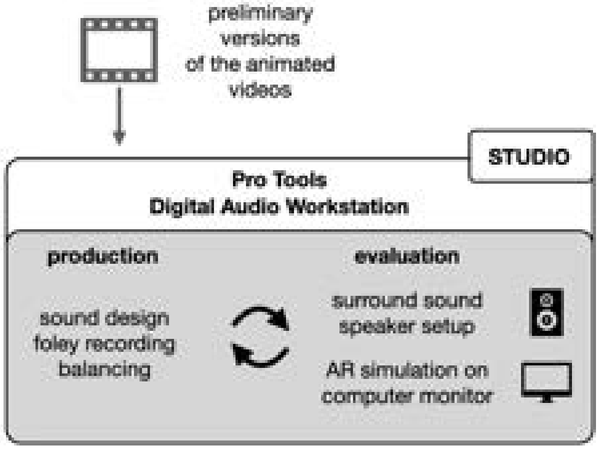


Figure 7. The initial audio post-production workflow in the studio for the linear AR experience.

Once the immersive sound design for the first animations was approved, this approach could serve as a template for the other scenes.

4.1.3 Implementation

The four animated 360-degree videos at the KMSKA are projected on the room's four walls by six large BARCO UDM-4K22 digital large venue projectors (BARCO 2020). All the audio and video content is streamed over IP from a PIXERA Two media server (PIXERA 2022) to four four-channel amplifiers and sixteen loudspeakers by AUDAC (AUDAC 2022).

The installation of the equipment at the museum ran parallel to the production of the video and audio content. This implied a workflow that mainly concentrated on artistic creation before the possibility of checking the real-life conditions at the KMSKA.

When the installation and environmental circumstances could be assessed for the first time (Picture 23), an initial site test that combined preliminary versions of the video and audio content revealed some concerns.

Picture 23. A test projection of one of the scenes at the KMSKA. The loudspeakers can be seen on the walls near the ceiling of the room.

The playback of the video and audio had to be adjusted by the operator of the media server to guarantee a more accurate synchronization between the two. Moreover, because of the impressive size of the projections, it became clear that some animated elements needed a diegetic sound that would correlate more with their largesse. This was fixed in the second phase of the sound design.

The most dramatic problem, however, was the apparent background noise at the location, produced by the cooling fans of the six projectors in the venue. Measured at 68 dB SPL C-weighted in the middle of the room (Table 4), this noise far exceeded the background levels experienced in other museums and masked some of the more subtle aspects of the sound design. This problem with background noise is similar to the experience with the Omaha Beach app (Liestøl 2019).

Table 4. Overview of the specifics of AR installation at the KMSKA

Location	**AV-technologies (playback)**		**Storytelling techniques**			**Environmental conditions**	
	Video	Audio	Type of sounds	Voice Over	interactive	Noise dB SPL C	Light
KMSKA	projection	immersive	diegetic	no	yes	68	darkened

To counter this issue, the background noise was recorded at the site and played back in the studio at the same sound pressure level during the next phase of the production. This approach mimicked the experience of the museum's visitors and aimed for a better translation between the studio's design stage and the room's situation. The mix of the sound design and some elements was then altered to be more audible, dynamic, and less competitive with the noise of the projectors' fans. This resulted in an updated workflow as can be seen in Figure 8.

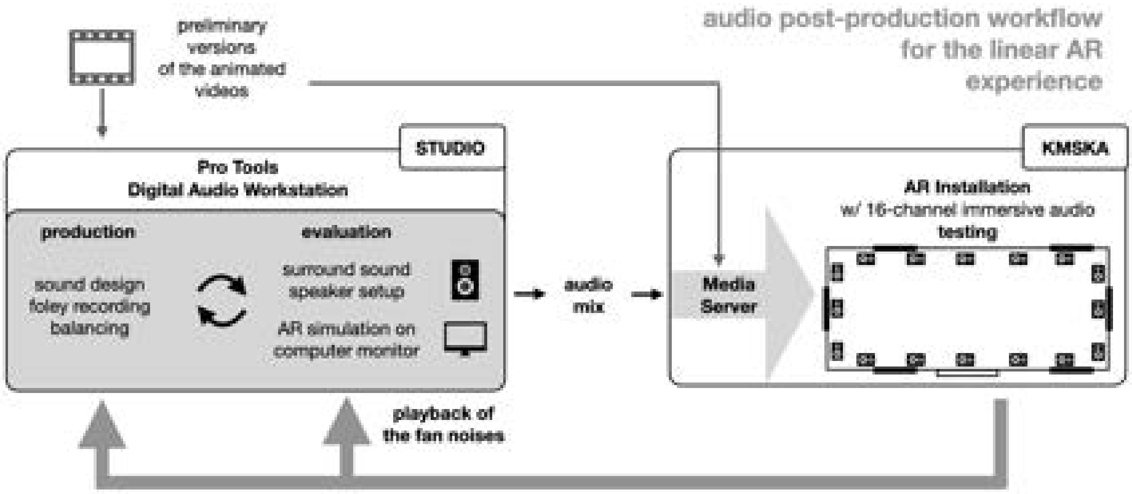


Figure 8 The complete audio post-production workflow for the linear AR experience shows the roadmap between the studio and the installation at the KMSKA and the playback of the recorded fan noise for better evaluation of the sound designs.

4.2 Results and conclusions

The tailor-made sound design for four animated location-based AR experiences was produced using sixteen-channel immersive sound, representing the sound design around and above the listener. As the immersive loudspeaker setup was not referenced to popular immersive audio standards, the choice fell on a channel-based workflow. This method was implemented in Pro Tools, a popular software for audio post-production, allowing for designing the various sounds, and evaluating the immersiveness in a professional audio studio before accessing the premises.

Visits to various location-based installations in France and Belgium demonstrated that environmental conditions can significantly impact visitors' experiences and should be considered in the workflow. Upon visiting the KMSKA, auditory environmental challenges presented themselves as loud fan noises generated by six video projectors, which were recorded and used as a reference while producing the sound design so that balances and the choice of individual sounds could be refined without being masked by this loud background noise. Integrating these conditions into the workflow enhanced the sound design translation from the studio to the museum. Only minimal level changes needed to be made on the media server.

Therefore, the audio-post production for this AR installation followed a similar process to those for linear cinematographic productions. Two additional steps in the process should be noted. Incorporating environmental issues at the location in the production process and using a media server at the implementation stage to handle all interactivity.

5 DISCUSSION

The following questions were addressed by exploring audio post-production workflows for VR and AR experiences in the Royal Museum of Fine Arts in Antwerp.

1. How do the offline workflows, commonly used for linear storytelling, differ from those used for producing immersive audio for VR and AR experiences in the Royal Museum of Fine Arts in Antwerp?
2. What workflows permit achieving the goals and the deadline when producing immersive sound for a VR experience without initial access to the VR-playback device that will be used in the museum?
3. What workflows counter the challenges when producing immersive sound for an AR experience without initial access to the premises and installations of a location-based experience?

The project validates Rumsey's view that sound creates believability in a VR environment (Rumsey 2016) and Levenson's statement that in VR, a sound has to be credible in the space it is in (Rumsey 2018). The combination of a high-quality VR headset with the addition of headphones further supports the proposition by Nishiyama et al. that video and audio quality are crucial to the immersiveness of virtual reality, as is the interaction between them and the interactivity of the contents (Nishiyama et al. 2018).

This research confirms Serafin's suggestion that one can retrieve sound design models and ideas from the entertainment industry for virtual reality simulations to add presence and immersion (Serafin 2004). While the recording of Foley, the selection of diegetic sounds, and general sound design techniques for VR and AR proved to be similar to those in audio-post-production processes for traditional linear audio-visual productions, additional choices in technology and implementing extra productional steps were necessary.

The first step consisted of choosing the immersive technology that suits the project best. Second-order Ambisonics proved to be a solid choice for VR as the immersive sound design can easily be monitored, allows for head-tracking, allows for credible reproduction of spatial content, and can be converted to the desired format.

Working on immersive audio content for VR and location-based AR without access to respectively the playback devices or site poses a series of challenges. For a VR production, this can be somewhat mitigated by using Audio Ease's 360monitor software, which allows the content creator to dynamically interact with the 360-degree video and listen to the immersive mix on headphones during production without the need for a VR headset. The option to use head-tracking further simulates the user's experience on a VR headset. But the visual impression of distance and size can only be evaluated on the preferred VR headset and the limitations of the audio it reproduces. Therefore, an additional step in the workflow needs to be considered in which final necessary adjustments can be made when the result can be checked on the preferred VR headset.

The workflow using artificial reverb to position diegetic sounds for VR content differed significantly from the traditional audio-post-production process. To place sounds near or far from the viewer, the sound designer would commonly use one artificial reverb algorithm that mimics the acoustic behavior of the suggested space and then balance the original (dry) sounds with added spatial information generated by a single setting of the algorithm. In VR, a far more persuasive outcome was obtained by processing individual sounds, each with its own reverb algorithms, and adjusting the parameters slightly depending on their positions. This does increase the processing load on the digital audio workstation during production but offers a far more convincing result upon playback.

The post-production process would be further optimized if the visual content could be streamed in high resolution to the VR headset from the digital audio workstation during the sound design for continuous evaluation. While this option is offered for some VR devices, it was not available for the type of headset used in this production. Ideally, the industry would introduce a multi-platform standard for streaming that would support the creation, evaluation, and consummation of VR content on various VR headsets.

For the location-based AR installation with multiple speakers and no rendering option at the playback side, choosing channel-based audio demonstrated the advantage of a clear workflow during the sound design and the delivery of a final mix compatible with most media servers.

A significant difference was observed. Traditional audio post-production techniques are technically highly normalized because of broadcast, film, and streaming requirements. This ensures that linear audio-visual storytelling workflows have predictable outcomes at the reproduction stage. This research showed that a location-based AR experience could suffer significantly from its environmental conditions, a factor that should be incorporated into the workflow. By identifying, recording, and reproducing the site's noise during the production of the museum's AR experience, the sound design could be tailored so that the result translated well from the studio to the location. Using this method, only minimal adjustments on the overall level of the final mix then had to be made on the media server.

Finally, when producing AR and VR content for museums and heritage sites, one also needs to consider Carrozzino & Bergamasco's (2010) argument that it is imperative to immediately engage the users so they can interact with fast learning curves in a natural way. This forces the production team throughout the process to counter any element that might be detrimental to the experience and can lead to looking for solutions outside the scope of the traditional offline workflow for linear storytelling.

6 REFERENCES

Altman, M., Kraus, K., Susal J. & Tsingos, N. 2016. Immersive Audio for VR. Presented at the Conference on Audio for Virtual and Augmented Reality on 30.9.2016-1.10.2016. Los Angeles.

Apple. 2022. Logic Pro Ridiculously powerful. Seriously creative. Published in 2022. Read 22.8.22. https://www.apple.com/logic-pro

Apple. 2022. iPhone XR - Technical Specifications. Published on 25.7.2022. Read 29.7.2022 https://support.apple.com/kb/SP781

Arnold-de Simine, S. 2012. The ‘Moving’ Image: Empathy and Projection in the International Slavery Museum, Liverpool. Journal of educational media, memory, and society 4 (2), 23-40.

Audio Ease BV. 2022. 360pan suite AudioEase. Published in 2022. Read 3.5.2022. https://www.audioease.com/360/

Audiokinetic. 2022. Wwise, the most advanced, feature-rich interactive audio solution. Published in 2022. Read 7.11.2022 https://www.audiokinetic.com/en/products/wwise

Barco. 2022. UDM-4K22 21,000 lumens, 4K UHD, 3-chip DLP digital large venue projector. Updated on 14.1.2022. Read 12.6.2022. https://www.barco.com/en/product/udm-4k22

Bayeux Museum. 2021. The Bayeux Tapestry The epic adventure of William the Conqueror in 1066. Published in 2021. Read 14.7.2022. https://www.bayeuxmuseum.com/en/the-bayeux-tapestry/

Beauchamp, R. 2013. Designing Sound for Animation. 2nd edition. New York and London: Focal Press. Taylor & Francis Group.

Beyerdynamic. n.d. DT 880 Edition Hi-fi headphones (semi-open). Read 20.7.2022. https://global.beyerdynamic.com/dt-880-edition.html

Britannica. n.d. cori spezzati. Read 24.8.2022 https://www.britannica.com/art/cori-spezzati

Carlsson, R. 2020. Digital ‘Storyteching’ – how museums can use technology to tell their stories. Museum Next. Published on 9.3.2020. Read 5.5.2022. https://www.museumnext.com/article/storyteching-how-museums-can-use-technology-to-tell-their-stories/

Carrozzino, M. & Bergamasco, M. 2010. Beyond virtual museums: Experiencing immersive virtual reality in real museums. Journal of Cultural Heritage 11: 452–458.

Chandler, D. 1996. Coleridge's 'suspension of disbelief’ and Jacob Brucker's 'assensus suspensione.' Notes and Queries 43 (1). Oxford University Press.

Chen, C. & Lai, H. 2021. Application of augmented reality in museums – Factors influencing the learning motivation and effectiveness. Science progress 2021 104 (S3), 1-16.

Clark, C. 2022. A Voice Inside My Head: The Persuasive Power Headphones Have on Listeners. UC San Diego Today. Published on 22.2.22. Read 14.5.2022. https://ucsdnews.ucsd.edu/pressrelease/a-voice-inside-my-head-the-persuasive-power-headphones-have-over-speakers

D-Day-Tours. 2013. Omaha Beach, Normandy. WN 62 Colleville Sur Mer. Published on 20.8.2013. Read 29.7.2022. https://www.d-daytoursnormandy.com/omaha-beach-wn-62-strongpoint-fortifications/

Dear Reality. 2022. dearVR PRO. Published in 2022. Read 25.4.2022. https://www.dear-reality.com/products/dearvr-pro

De chinezen NV. 2022. THE CHINESE. Published in 2022. Read 21.4.2022. https://dechinezen.be/en

FMOD. 2022. Adaptive audio FMOD is a suite of high quality audio solutions for real-time applications. Published in 2022. Read 7.11.2022. https://www.fmod.com/products

Fonseca, N. 2020. All You Need to Know About 3D Audio. Leiria: Sound Particles LDA.

Forlenza, S. 2017. Medieval London Objects: Church Bell. Published in 2017. Read 11.5.2022. https://medievallondon.ace.fordham.edu/exhibits/show/medieval-london-objects/churchbell

Forsythe, G. 2018. Audio Mixing for VR: The Beginners Guide to Spatial Audio, 3D Sound and Ambisonics. Published on 5.2.2018. Read 2.6.2022. https://sonicscoop.com/audio-mixing-for-vr-the-beginners-guide-to-spatial-audio-3d-sound-and-ambisonics/

Grazioli, G. 2020. ArcheoEchi – a virtual reconstruction of a medieval cathedral in Southern Italy. AES 149th Convention on 21-24.10.2020. New York. Requires access right. http://www.aes.org/e-lib/browse.cfm?elib=20904

Hartmann, W. M., 1993. Auditory Localization in Rooms. AES 12th International Conference on 28-30.6.1993. Copenhagen. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=6251

Hamilton, R. 2015. Designing Next-Gen Academic Curricula for Game-Centric Procedural Audio and Music. AES 56th International Conference on Audio for Games on 11-13.02.2015. London. https://ccrma.stanford.edu/~rob/papers/hamilton_aes56.pdf

House of Marley. 2022. Positive Vibration II - Bluetooth. Published in 2022. Read 08 09, 2022. https://marley.nl/positive-vibration-ii-bluetooth.html

ICOM. 2022. ICOM approves a new museum definition. Published on 24.8.2022. Read 25.8.2022. https://icom.museum/en/news/icom-approves-a-new-museum-definition/

Inglis, S. 2022. An Introduction To Immersive Audio. Sound On Sound. Published in 2022. Read 10.10.2022. https://www.soundonsound.com/techniques/introduction-immersive-audio

Jørgensen, K. 2007. On transdiegetic sounds in computer games. Northern Lights: Film and Media Studies Yearbook 1 (1), 105-117.

Jansen, N. 2022. Tussen Hemel & Hel. Arnout Hauben. Published on 5.8.2022. Read 12.8.2022. Requires access rights. https://www.humo.be/tv/arnout-hauben-mochten-we-tijdens-onze-wandeltochten-een-glasbak-hebben-meegezeuld-hij-zou-op-het-einde-meer-dan-overvol-hebben-gezeten~b2b8467d/?referrer=https%3A%2F%2Fduckduckgo.com%2F

Kardous, C.A. & Shaw, P.B. 2014. Evaluation of smartphone sound measurement applications. The Journal of the Acoustical Society of America 135 (4), 186-192.

Kennedy, M. 2015. British Museum uses virtual reality to transport visitors to the bronze age. Published on 4.8.2015. Read 14.6.2022. https://www.theguardian.com/culture/2015/aug/04/british-museum-virtual-reality-weekend-bronze-age

KMSKA. 2021. God the Father with Singing and Music-making Angels. Published in 2021. Read 6.7.2022. https://kmska.be/en/masterpiece/god-father-singing-and-music-making-angels

KMSKA. 2021. Madonna Surrounded by Seraphim and Cherubim. Published in 2021. Read 12.5.2022. https://kmska.be/en/masterpiece/madonna-surrounded-seraphim-and-cherubim

KMSKA. 2021. Scaldis and Antverpia. Published in 2021. Read 15.6.2022. https://kmska.be/en/masterpiece/scaldis-and-antverpia

KMSKA. 2021. The Fall of the Rebel Angels. Published in 2021. Read 15.6.2022. https://kmska.be/en/masterpiece/fall-rebel-angels

KMSKA. 2021. The finest feeling. Published in 2021. Read 14.05.2022. https://kmska.be/en

Landau, S., Wiener, W., Naghshineh, K. & Giusti, E. 2005. Creating Accessible Science Museums with User-Activated Environmental Audio Beacons (Ping!), Assistive technology: the official journal of RESNA 17 (2), 133-143.

Lascaux. 2022. Lascaux IV International Center. Published in 2022. Read 3.7.2022. https://www.lascaux.fr/en#

Linters, A. & Plomteux, G. 1979. Agentschap Onroerend Erfgoed 2022: Meir. Published in 1979. Read 11.5.2022. https://inventaris.onroerenderfgoed.be/themas/11207

Liestøl, G. 2019. Augmented Reality Storytelling – Narrative Design and Reconstruction of a Historical Event in situ. International journal of interactive mobile technologies 13 (12), 196-209. https://doi.org/10.3991/ijim.v13i12.11560

Maclot, Petra. 2016. Rubenshuis. Beheerplan Bijlagen. Bouwhistorisch onderzoek van de Rubenssite in Antwerpen. Antwerpen, 46.

Mandereyck, M, Plomteux, G. & Steyaert, R. 1979. Agentschap Onroerend Erfgoed 2022: Jodenstraat. Published in 1979. Read 10.5.2022. https://inventaris.onroerenderfgoed.be/themas/11156

Manderyck, M., Plomteux, G. & Steyaert, R. 1979. Agentschap Onroerend Erfgoed 2022: Hopland. Published in 1979. Read 10.5.2022. https://inventaris.onroerenderfgoed.be/themas/11150

Manderyck, M., Plomteux G., Steyaert R. & Van Driessche, T. 2019. Agentschap Onroerend Erfgoed 2022: Rubenshuis. Published in 2019. Read 21.4.2022. https://inventaris.onroerenderfgoed.be/erfgoedobjecten/6304

Matei, A. 2022. Listening to podcasts on headphones increases 'perceived intimacy' with host, research finds. Published on 2.3.2022. Read 17.5.2022. https://www.theguardian.com/tv-and-radio/2022/mar/02/podcasts-headphones-perceived-intimacy-host-increased

Mironovs, M. & Hyunkook L. 2019. Localisation of natural sound sources at various azimuth and elevation angles. 146th AES Convention on 20-23.3.2019. Dublin. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=20298

NASA. 2022. Museum & Informal Education Alliance. Updated 4.2022. Read 16.5.2022. https://informal.jpl.nasa.gov/museum/360-video

Neukom, M. 2006. Decoding Second Order Ambisonics to 5.1 Surround Systems. 121st AES Convention on 5-8.10.200. San Francisco. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=13814

Nishiyama, T., Nagata T., Ogata, S., Hasue, T. & Kashiwagi M. 2018. Recording and Mixing Techniques for Ambisonic Sound Production. 2018 AES International Conference on Spatial Reproduction - Aesthetics and Science on 7-9.8.2018. Tokyo. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=19644

Önen, U. 2014. "It Has to Work With the Picture": Audio Education for Film and Media Students. AES 137th Convention on 9-12.10.2014. Los Angeles. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=17402

Pico XR. 2022. Pico G2 4K. VR goes 4K. Published in 2022. Read 11.05.2022. https://www.picoxr.com/us/G2_4K.html

Pico XR. 2022. Pico VR Streaming Assistant. Published in 2022. Read 11.05.2022. https://www.picoxr.com/us/vsa.html

PIXERA. 2022. PIXERA TWO. Published in 2022. Read 12.05.2022. https://pixera.one/en/hardware/pixera-two

Pye, M. 2021. Antwerp: The Glory Years. London: Allan Lane.

Radhika. 2015. Foley: Making Movies Come Alive! Youngzine. Published on 3.7.2015. Read 18.8.2022. Requires access rights. https://link.gale.com/apps/doc/A555116042/ITOF?u=tampere&sid=bookmark-ITOF&xid=41fde471

Roginska, A. & Geluso, P. 2018. Immersive Sound : the Art and Science of Binaural and Multi-Channel Audio. 1st edition. New York: Routledge.

Rumsey, F. 2016. Virtual reality. Mixing, rendering, believability. Journal of the Audio Engineering Society 64 (12), 1073-1077. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=18538

Rumsey, F. 2018. Virtual Reality. Will it be a game-changer? Journal of the Audio Engineering Society 65 (5), 399-402. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=19569

Scan4Stories. 2021. De apotheek van Eléonora Verbeke. Published in 2021. Read 11.11.2022. https://www.scan4stories.be/deapotheekvanel%C3%A9onoraverbeke

Schiltz, K. & Ongaro, G. 2018. Brill's Companions to the Musical Culture of Medieval and Early Modern Europe. A Companion to Music in Sixteenth-Century Venice. Leiden-Boston. BRILL.

Serafin, S. 2004. Sound Design to Enhance Presence in Photorealistic Virtual Reality. Proceedings of the 2004 International Conference on Auditory Display on 6-9.7.2004. Sidney.

Shivappa, S., Morrell, M., Sen, D., Peters, N. & Salehin S.M.A. 2016. Efficient, compelling and immersive VR audio experience using Scene Based Audio / Higher Order Ambisonics. Presented at the Conference on Audio for Virtual and Augmented Reality on 30.9.2016-1.10.2016. Los Angeles. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=18493

Station Europe. n.d. Station Europe. Visiting European Parliament. Read 23.03.2022. https://visiting.europarl.europa.eu/en/visitor-offer/brussels/station-europe

Till, S. 2008. The Canonical Hours. Published on 9/9/2008. Read 11.5.2022. http://steventill.com/2008/09/09/the-canonical-hours

Topa. 2021. Sleutel tot de Sint-Jacobskerk te Antwerpen. De toren, robuust maar gefrustreerd. Published in 2021. Read 2.8.2022. https://topa.be/sint-jacob/beschrijving/toren/

van Huut, T. 2022. Het ‘Rijksmuseum’ van Vlaanderen mikt op de ‘emotionele ervaring’. NRC. Published on 24.8.2022. Read 26.8.2022. Requires access rights. https://www.nrc.nl/nieuws/2022/08/24/het-rijks-van-vlaanderen-mikt-op-de-emotionele-ervaring-a4139757

Wilson, A. 2020. 10 of the world’s best virtual museum and art gallery tours. The Guardian. Published on 23.3.2020. Read 2.6.2022. https://www.theguardian.com/travel/2020/mar/23/10-of-the-worlds-best-virtual-museum-and-art-gallery-tours

Wyatt, H. & Amyes. T. 2004. Audio Post Production for Television and Film: An Introduction to Technology and Techniques. 3rd Edition. New York and London: Focal Press. Taylor & Francis Group.

Wyman, B., Smith, S., Meyers, D. & Godfrey, M. 2011. Digital Storytelling in Museums: Observations and Best Practices: Digital Storytelling in Museums. Curator: The Museum Journal. Curator 54(4), 461–468. https://doi.org/10.1111/j.2151-6952.2011.00110.x

Xie, B., Mai, H. & Zhong, X. 2017. The Median-Plane Summing Localization in Ambisonics Reproduction. Presented at the 142nd AES Convention on 20-23.5.2017. Berlin. Requires access rights. http://www.aes.org/e-lib/browse.cfm?elib=18603