

MUSIC AND SOUND FOR MEDIA

Calum Hailey (2116401)

Introduction	2
Linear vs Non-linear visual media	2
Justifying the use of Music and Sound for Media	2
Case Study – Mark Ware	3
Examples of Music and Sound in visual media	3
User Experience (UX) In Games	3
Historical Audio for UX In Gaming	3
Modern Audio for UX In Gaming	4
Mixing Game Audio with FMOD	4
3d Sound and Immersive Audio	5
History Of 3d Sound - Binaural Audio	5
Binaural Stereo Vs Surround Sound Applications	5
Channel Based Vs Object Based Systems	6
Channel-Based Systems	6
Object-Based Systems	6
Mixing 3d Audio In Logic Pro X	7
Binaural Panning	7
5.1 Surround Panning	7
Dolby Atmos	7
Conclusion	8
References	9

MUSIC AND SOUND FOR MEDIA

INTRODUCTION

This paper investigates how music and sound is used in visual media and how it can influence the user experience. This text aims to: *(A) examine the interaction between humans and exposure to sound and how this may be used to supplement a visual media application. (B) explore historical uses of music and sound in media and outline contemporary developments in the area.*

The exploration of associated technologies and workflows in the area will be discussed in the context of referenced material.

LINEAR VS NON-LINEAR VISUAL MEDIA

The main difference between linear and non-linear visual media is that linear visual media does not have any navigational features. A linear presentation follows a set chronological order and will run continuously from start to finish. One of the most common instances of linear multimedia is a PowerPoint presentation. Other examples include cinema, films, and television shows which are viewed via broadcast or in a controlled manner. Non-linear media, on the other hand, defies this one-way pattern, allowing viewers to explore all multimedia components in any order freely. Typical examples are YouTube and Netflix, where users can freely browse and navigate through media of their choice. (Visual Arts, nd) Linear broadcasts can become non-linear with the use of additional equipment such as TiVo or equivalent recording and playback technology. This allows the user to pause, play, rewind, and fast forward. In the case of TiVo, the interface is accompanied by sound effects to signify media transport, which can enhance the user experience. The impact of music and sound on user experience will be covered later in this paper.

JUSTIFYING THE USE OF MUSIC AND SOUND FOR MEDIA

Sound and vibration are everywhere on earth, and as humans, we hear and feel a response when we touch physical objects. Audio engineers design sounds in virtual environments to mimic this natural response to create a stimulating experience for viewing non-linear visual media. Humans have five senses, but electronic devices like computers generally communicate with us by using predominantly just two – hearing and sight. To bridge the gap of sensory deficit, haptic feedback (vibration) is often used in conjunction with sound and visual artefacts to create a more realistic and ultimately more satisfying experience. A prime example of this is the use of virtual keyboards on touchscreen devices. The design of virtual keyboards uses sound and visual design paired with haptic feedback to mimic the physical response of using a mechanical keyboard. These technologies are also used in games in conjunction with controller feedback and VR headsets. Audio signals from the game trigger motors in the controller, causing vibration, similar to a mechanical loudspeaker driver propagating sound. Using a weapon or being attacked by an enemy is signified by vibration sequences with accompanied aural and visual cues. (Blenkinsopp, R., 2019). Creative industries capitalise on using sound and vibration in media to portray realism, cue emotional responses, and create soundscapes; these elements aid the user experience.

CASE STUDY – MARK WARE

Mark Ware, an audio-visual artist, collaborated with Brighton and Sussex Medical School to test the impact of listening to sounds captured in natural and artificial settings. The collaboration resulted in publishing the paper entitled '*Mind-wandering and alterations to default mode network connectivity when listening to naturalistic versus artificial sounds*'. In the study, the researchers exposed each participant to randomised soundscapes of 5 minutes and 25 seconds. The researchers used MRI Scanners and heart monitors to obtain data from the participants' brain activity and nervous system function. (Gould van Praag et al., 2017). The research concluded that when listening to natural sounds, the human brain activity reflected an outward-directed focus of attention associated with body relaxation. In contrast, when listening to artificial sounds, the brain connectivity reflected an inward-directed focus of attention, associated with anxiety, post-traumatic stress disorder and depression. Using the data from this study, one could argue that music and sound could be used in conjunction with visual media to enhance the user experience significantly. Using audio to trigger varied emotional and physical responses could set the tone and feeling for a visual presentation creating an impactful result.

EXAMPLES OF MUSIC AND SOUND IN VISUAL MEDIA

USER EXPERIENCE (UX) IN GAMES

The effect of user experience in games causes a behavioural response in users. Researchers use PAD theory to study user responses to a gaming experience. The theory asserts that all human emotions are defined by three basic emotions pleasure, arousal, and dominance. (Huang et al., 2017) A successful game must have a good user experience where the interface generates a positive emotional response.

Using audio to supplement the user experience in games can create a memorable experience. Audio implementation in games using soundtrack design, sound effects, and vocal samples can be iconic, and arguably games may not be as successful without these audio features. We can only speculate the success of Super Mario without his "Let's a Go!" catchphrase or Sonic the Hedgehog 3 without Michael Jackson's controversial musical score.

HISTORICAL AUDIO FOR UX IN GAMING

The first commercially available example of video gaming with sound effects was 'Computer Space' in 1971. The game featured crude-sounding audio to represent the movement of a spaceship, including rocket boosters, weapons firing, and explosions. In 2008 Al Alcorn, who helped design the game audio, was interviewed by Cam Shea of IGN and explained that "the sounds were generated using pre-existing tones in the sync generator, which were then gated to separate them and stored on a chip" (Shea, 2008).



FIGURE 1. (SPACEWAR, ND)

Until the introduction of 'Space Invaders' in 1978, game music had consisted mainly of monophonic sounds, loops, and audio phrases between levels or at the start of a new game. The physical limitations of computer memory were simply insufficient to achieve anything else. Like the games before, 'Space Invaders' also featured a simple looped chromatic melody; however, it was dynamic and interacted with the player (Weaver, D. 2019). The melody increased in tempo in line with the game's difficulty level, where the enemies descended on the player at an increasing rate. The introduction of dynamic audio was a turning point for the games industry. Music had become a defining selling point, increasing the user's engagement, enjoyment, and interactivity with the interface.

MODERN AUDIO FOR UX IN GAMING

Today, the industry standard for games defines that, like in real life, anything a player's character can interact with in a game world should have audio feedback. Using audio feedback in games creates realism and relatability; this is vital in a virtual reality context where the user is transported virtually to a new place. If the player interacts with anything on the screen or VR headset, it should be audible, whether it's scrolling through the pause menu, tapping a button, or walking around. (A Studio Lite, 2019). Game designers can implement audio feedback as a tool to help the player; for instance, the sounds could be directly linked to the player's achievements or failures, signify they're going out of bounds or alert the player when something important is about to happen. Audio cues for low health, when an enemy becomes alert, and when a player dies are commonplace and expected as part of the gaming experience.

MIXING GAME AUDIO WITH FMOD

The DAW "FMOD" was created especially for use with game engines, allowing music producers and game developers to quickly adapt an exchange by sharing code between FMOD and Unity or Unreal Engine. It also has a code-based version of the software that can be integrated with customised game engines and runs on C++ or C# versions of the interface. The user can quickly audition changes in the editor or connect to a running game and hear changes in real-time, thanks to FMOD's ability to run in parallel with a game engine (FMOD Studio, nd). FMOD can create dynamic, evolving sound and music suitable for many linear and non-linear applications.



FIGURE 2. (GAMING BOLT, 2015)

3D SOUND AND IMMERSIVE AUDIO

With the recent advancements in three-dimensional visual technologies and virtual reality platforms, there is a growing market demand for immersive audio. This market is apparent in various applications, including digital high-definition television, cinema, video games, and, more recently, smart handheld devices. Immersive audio creates a three-dimensional sound experience for the listener to mimic a natural listening environment. Immersive audio processing envelopes the listener with sound from all sides, above and below, resulting in an entirely new perceptual experience when listening to music, viewing movies, or playing video games. Storm Audio state that using immersive audio "gives the feeling of actually being there.". The effect of immersive audio used in conjunction with visual media can dramatically enhance the user experience. (Storm Audio, nd)

HISTORY OF 3D SOUND - BINAURAL AUDIO

Binaural audio technology has been around for over a century. A binaural audio recording captures sound as you hear it using two microphones arranged to imitate the distance between your ears. When headphones are used, binaural audio creates the impression of being in the exact spot where the recording was made. In 1881, Clement Ader, the inventor of the telephone, pioneered the first binaural audio system, known as the 'Théâtrophone'.

Some consider it the first example of 'binaural' sound since it permitted interaural variations to be communicated to both ears of the listener. The system's Paris debut consisted of an array of paired microphones positioned at the edge of the Opera Garnier stage. Consumers could subscribe to the service where audio from performances was transmitted to their homes via two phone lines to headphones with one line designated for each ear. The device proved a success and was installed in theatres across Europe. (Collins, 2008)

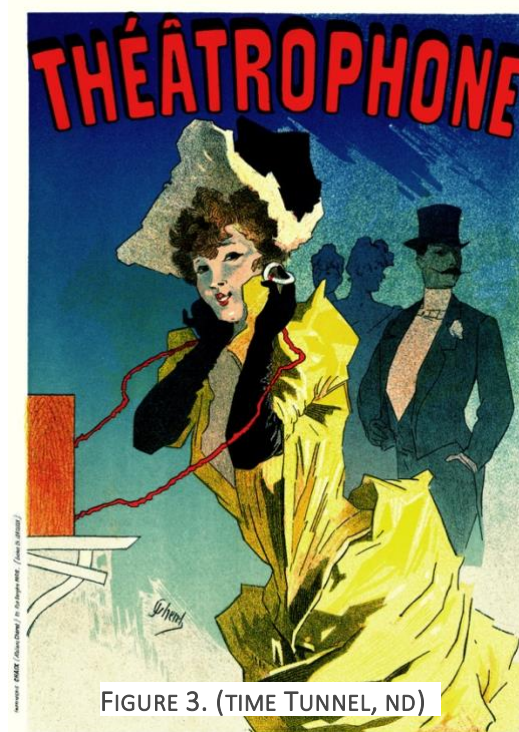


FIGURE 3. (TIME TUNNEL, ND)

BINAURAL STEREO VS SURROUND SOUND APPLICATIONS

As 3D sound becomes a desirable commodity in everyday life, considerations are being made toward the technology people use most. For most people, their go-to media device would be a smartphone or portable device, such as a laptop or netbook. Due to physical limitations of stereo width in smartphones and portable device speaker arrays, the sound produced by these devices is often perceived as monophonic. Binaural processing is often used to mitigate the monophonic audio perception on these devices, effectively broadening the perceived width and depth of the reproduced sound field (Floros and Tatlas, 2011). When listening to sound and music on a larger scale, multichannel surround sound, ambisonics, and wavefield synthesis systems convey 3D Immersive sound. These systems are commonly found in home cinemas, music studios and commercial cinema applications. Due to the scale of these applications, more powerful arrays are

required to fill the field of sound. Compared with binaural processing used on portable devices, these audio solutions can require considerable financial investment, substantial space, and complex loudspeaker configurations to achieve accurate sound reproduction and optimal results.

CHANNEL BASED VS OBJECT BASED SYSTEMS

CHANNEL-BASED SYSTEMS

Traditional surround sound speaker arrays typically consist of several loudspeakers accompanied by a singular low-frequency subwoofer. Some commonly found surround systems include 2.1 (Stereo), 5.1, 7.1, and 9.1; the first number represents the quantity of directional speakers, and the second represents a single low-frequency subwoofer. The most common and arguably most successful is '5.1 Dolby Digital', commonly found in home audio-visual systems. In these traditional systems, the audio path is denoted per channel (centre, left, right, left side and right side). The positioning of the five directional speakers enhances the audio image's directional stability and spatial realism. The 5.1 system achieves this sense of realism by covering a greater area than a conventional two-channel stereo system, with sound projected in front centrally, left, and right, and to the sides of the listener. Although a 5.1 surround system is perceived as more spatially realistic than a two-channel stereo system, it is not perfect. The speaker array is positioned at the listener's ear level, causing the audio image to be restricted to the horizontal plane. Regardless of the limitations of this platform, it is still deemed to be an industry standard with Dolby Digital 5.1 audio encoding present on modern-day BluRay disks (ITU-R, 2015).

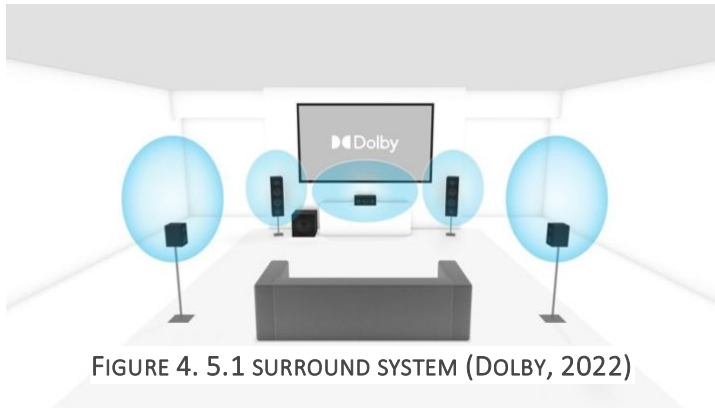


FIGURE 4. 5.1 SURROUND SYSTEM (DOLBY, 2022)

OBJECT-BASED SYSTEMS

In an object-based system, the audio content is projected irrespective of any given loudspeaker array; this is due to the mixing process not having to rely on directional channels like traditional systems. A numbered hierarchy still denotes the size of the speaker array in object-based configurations, but it does not influence the sound's path in the same way. The audio content is represented as objects containing the audio data, combined with metadata that stores parameters such as position, direction, spread or distance (ITU-R, 2015). An example of an object-based array is the Dolby Atmos 9.1.2 array. Where the first number represents the quantity of directional speakers, the second represents a single low-frequency



FIGURE 5. 9.1.2 SURROUND SYSTEM (DOLBY, 2022)

subwoofer, and the third represents the high-mounted speakers. The additional high-mounted speakers offer a distinct advantage of sound in the vertical plane, which adds a higher perception of reality for the listener (Matthes, 2019).

MIXING 3D AUDIO IN LOGIC PRO X

Logic Pro X allows binaural stereo, channel-based surround using its built-in mixing facilities. It is also now possible to perform object-based mixing using a Dolby Atmos plugin in Logic 10.7 onwards.

BINAURAL PANNING

Panning in binaural stereo In Logic Pro X allows the user to view a 3D plane of where the audio will be present within the mix. The user interface features changeable parameters to edit the audio's angle, elevation, depth and spread within the vertical and horizontal planes. It's also possible to edit the size of the virtual environment, which can enhance the effect of space and distance.

5.1 SURROUND PANNING

When mixing in 3D audio in Logic Pro X, the user interface offers many parameters to control the direction, position, and depth of the audio within the horizontal plane. The 'Angle' controls the direction, 'Diversity' controls the depth and 'Spread' controls the width (Spread is only available for stereo audio). Each of the five directional speakers can easily be separated into their respective channels using their dedicated toggle switches (represented as blue speakers) and finely tuned using the 'XF', 'XR' and 'Y' parameters.

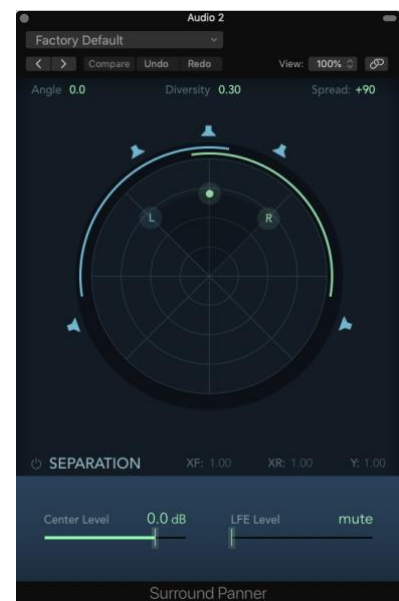


FIGURE 6. 5.1 SURROUND PANNING (APPLE, 2022)

DOLBY ATMOS

The Dolby Atmos plugin features a 3D Object Viewer that depicts the 3D space of your spatial mix. The small spheres serve as a representation of the positions of the object tracks. When selected, these spheres can individually behave in a similar way to the standard binaural pan; they can be moved and manipulated with live feedback in the 3D Object Viewer. Mixing with the 3d viewer is particularly advantageous for audio-visual projects. A producer using this technology could potentially map specific elements of an on-screen production to individual objects, adding to the realism and immersive experience.

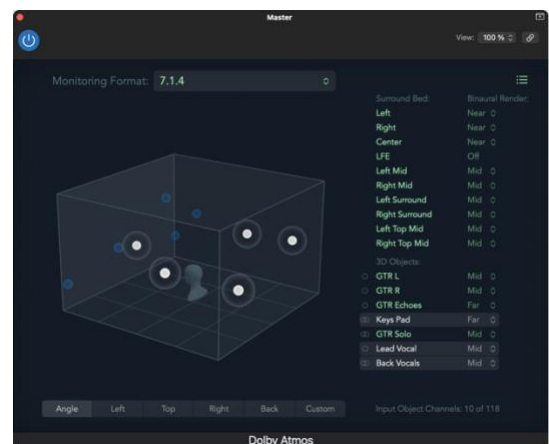


FIGURE 7. DOLBY ATMOS 3D VIEWER (APPLE, 2022)

CONCLUSION

The study was started by defining the differences between linear and non-linear media and, secondly, looking at the relationship between the human sensory system and sound. This approach was undertaken to gain an understanding of why it might be seen as essential to implement music and sound for visual media and games. This initial research strongly suggested that music and sound would dramatically influence human behaviour, ultimately enhancing the user experience and justifying its use. As such discussing UX in games and 3D sound seemed appropriate. The process has covered the origins and basic scientific principles of creating 3D audio and why it is deemed necessary in contemporary applications. The progression in 3D audio is fascinating because it seems like history somewhat repeats itself. Multi-layered binaural audio is at the forefront of immersive audio development and potentially implies the development of channel-based systems was a waste of time. Perhaps this is an example of human development deviating in the wrong direction for many years; it wouldn't be the first time. Seeing as we're utilising concepts developed in the 1800's it does beg the question, could we have achieved large-scale binaural immersive audio systems far earlier?

Regardless of the speculating question, it's clear to see why we utilise sound and music in visual media – it provides us with a sense of realism, something we humans crave.

REFERENCES

5.1 *Virtual Speaker Setup* (no date). Available at:

<https://www.dolby.com/about/support/guide/speaker-setup-guides/5.1-virtual-speakers-setup-guide/> (Accessed: 29 June 2022).

Blenkinsopp, R. (no date) *What is Haptic Feedback? | Ultraleap*. Available at:

<https://www.ultraleap.com/company/news/blog/what-is-haptic-feedback/#how-haptic-feedback-works> (Accessed: 10 June 2022).

Collins, P. (2008b) 'Theatrophone: the 19th-century iPod', *New Scientist - NEW SCI*, 197, pp. 44–45. doi:[10.1016/S0262-4079\(08\)60113-X](https://doi.org/10.1016/S0262-4079(08)60113-X).

Floros, A. and Tatlas, N.-A. (2011) 'Spatial enhancement for immersive stereo audio applications'. doi:[10.1109/ICDSP.2011.6005001](https://doi.org/10.1109/ICDSP.2011.6005001).

FMOD Studio (no date). Available at: <https://www.fmod.com/studio> (Accessed: 29 June 2022).

Gould van Praag, C.D. *et al.* (2017) 'Mind-wandering and alterations to default mode network connectivity when listening to naturalistic versus artificial sounds', *Scientific Reports*, 7(1), p. 45273. doi:[10.1038/srep45273](https://doi.org/10.1038/srep45273).

Huang, M., Ali, R. and Liao, J. (2017) 'The effect of user experience in online games on word of mouth: A pleasure-arousal-dominance (PAD) model perspective', *Computers in Human Behavior*, 75. doi:[10.1016/j.chb.2017.05.015](https://doi.org/10.1016/j.chb.2017.05.015).

Immersive sound (no date) *StormAudio*. Available at: <https://www.stormaudio.com/immersive-sound/> (Accessed: 10 June 2022).

Logic Pro X User Guide - Apple Support – Apple Support (UK) (no date). Available at:

<https://support.apple.com/en-gb/guide/logicpro/welcome/10.7.3/mac> (Accessed: 29 June 2022).

Matthes, A.J. (2019) 'Surround Sound Channels Explained: 2.1, 5.1, 7.1, 9.1, and More!', *The Home Theater DIY*, 28 March. Available at: <https://thehometheaterdiy.com/surround-sound-channels-explained/> (Accessed: 29 June 2022).

'Multichannel sound technology in home and broadcasting applications' (2015). International Telecommunication Union.

Shea, C. (2008) *Al Alcorn Interview, IGN*. Available at:

<https://www.ign.com/articles/2008/03/11/al-alcorn-interview> (Accessed: 10 June 2022).

'SOUND DESIGN IN GAMES: WHY AUDIO MATTERS' (2019) *A Studio Lite*, 5 August. Available at:

<https://astudiolite.com/sound-design-in-games-why-audio-matters/> (Accessed: 13 June 2022).

SPACEWAR! (no date). Available at: <http://andrewjensen.com/CLASSICGAMESSPACEWAR!.HTML> (Accessed: 29 June 2022).

The History of Music and Audio in Video Games | Melbourne (no date) Abbey Road Institute Australia. Available at: <https://abbeyroadinstitute.com.au/blog/history-audio-music-video-games/> (Accessed: 13 June 2022).

Theatrophone - Les Maitre de l’Affiche - 1895 (no date) Vintage Pinup Wall Prints. Available at: <https://www.timetunnel.com/theatrophone-les-maitre-de-laffiche-1895.html> (Accessed: 29 June 2022).