Creating a Homogenous 3-D Auditory Image

DAMIAN CANDUSSO
Creating a Homogenous 3-D Auditory Image

Damian Candusso, Charles Sturt University, Australia

Abstract: With the recent emergence of digital 3-D stereoscopic films, the image has been increasingly the focus of the immersive experience for the cinema audience. With surround sound having been introduced from as early as Disney’s 1940 release of Fantasia (1940) in Fantasound, cinema sound has traditionally been at the forefront in the creation of an immersive cinematic experience, however now appears to be foreshadowed. Over the past 2 years, parallel to the changes in screen technologies, cinema sound has been attempting to provide a true 3-D aural immersive experience. Some in Hollywood argue that cinema sound has been in 3-D since the initial introduction of surround sound, and it is only now, that the film image has become 3-D and has finally caught up. Cinema sound practices have not seen a shift with the introduction of the 3-D imagery, with the exact same soundtrack being shared by both the 2-D and 3-D release of the same film. The introduction of surround sound provided an additional dimension to cinema allowing the sound practitioner to move sounds from apparently within and beyond the visual frame. With the introduction of 3-D visuals, the image is no longer sitting on a 2-D plane, and it can now artificially detach from the screen itself—utilising the z-space of the cinema. However, the z-space is also occupied by the soundtrack, often creating a dislocation between image and sound. Although there is a shift with cinema sound technologies to increase channels and speakers, does this provide an effective 3-D sound solution or does it actually contribute to a breakdown in relationship between the soundtrack and the image?

Keywords: 3-D, Surround Sound, Film, Immersion, Stereoscopy, Film Industry

Introduction

With the recent resurgence of 3-D cinema some 50 years after the short lived boom of the 1950’s, it is yet to be determined how long the current trend of digital 3-D filmmaking will be sustained. Although brought back to increase box-office sales, the validity of this medium as a serious narrative medium is being questioned. Additionally, production practices and the evolution of film technologies are creating challenges, not applicable to 2-D film. Specifically, by allowing the 3-D the vision to occupy the Z-Space, are we dislocating the cohesive relationship between the vision and the soundtrack compared to contemporary 2-D films? The shift to digital cinema is allowing technologies and working methodologies for both the vision and sound to continually evolve in terms of the creation and exhibition of content. With so much emphasis on how “3-D looks,” this attention to the imagery is overshadowing the accompanying soundtrack. To an extent, this is due to the fact that many people believe that cinematic sound was already 3-D.

In an interview with Randy Thom, one of the worlds most well respected sound designers and re-recording mixers with first hand experience in early developments of surround sound including Francis Ford Coppola’s Apocalypse Now (Coppola 1979), he shares the view that “sound has been 3-D since the beginning of surround sound. It is only now, that the film image has become 3-D and has finally caught up with sound” (Thom 2012).

Although agreeing to some extent with Thom, I argue that conventional 5.1 and 7.1 surround sound does add another dimension to sound that emanates from behind the screen, however this does not necessarily define it as being 3-D sound. This is becoming increasingly apparent with new immersive sound formats including Dolby Atmos and Barco’s Auro 3-D marketing themselves as being true 3-D sound as they provide the addition of height channels.

Early Sound for Film

3-D reproduction has existed since 1838, when Sir Charles Wheatstone discovered binocular stereopsis. Wheatstone noted that if we can channel the individual visual information of two
viewpoints of an object separately to each of our eyes, our brain is able to recreate the illusion of
the solid object. This understanding and process of stereography preceded the invention of
motion pictures. Stereo multichannel sound also predates the invention of motion pictures by at
least sixteen years with Alexander Graham Bell’s experiments on two-channel sound transmitted
by two telephones in 1879 (Kerins 2010, 34).

Sound for film has been an integral component of the cinema experience from the early
projection of images accompanied by a pianist, to the first use of synchronized “talky” sound of
The Jazz Singer (Crosland 1927). Since the introduction of synchronous sound, many
conventions have been established. One of the primary conventions is that dialogue is always
played from the speaker behind the center of the screen. It was not long after synchronized
monaural sound that experiments started to take place increasing the number of audio channels.
During the early 1930’s Alan Blumlein began recording using a twin microphone system. This
system was capable of utilizing the same space on the film strip that was occupied by the
monaural optical soundtrack. Disney and RCA were jointly carrying out research into their own
proprietary multichannel audio format that exceeded the previous sound systems. This early
development of multichannel sound was one of the first experiments into surround sound. It
introduced the first use of multichannel sound to the public with its Fantasound, introduced with
Disney’s 1940 release of Fantasia (Algar et al. 1940). Although not used beyond Fantasia, the
Fantasound system became a precursor to some of the modern day cinema sound systems
including stereo and surround sound.

In 1952, the first full color feature length 3-D film was released. Arch Oboler’s Bwana Devil
(Oboler 1952) used NaturalVision, a polarized 3-D system that had two projectors mechanically
locked, each producing a left and right image that complemented the polarized glasses worn by
the audience. Although released with a monophonic soundtrack, it was not long before Warner
Bros. introduced WarnerPhonic, a multichannel sound format to a 3-D film release. Premiering
with House Of Wax (Toth 1953) and also using NaturalVision, the WarnerPhonic sound system
was a four-channel surround system. It had a separate three channel full coat 35mm magnetic
film containing the left, center and right channels that was synchronized to the optical track on
the release print that contained a channel for the surround sound in the auditorium.

With the introduction of surround sound, the sound practitioner was provided with increased
creativity that provided the capacity to pan sounds not only across the screen, but also into the Z-
Space of the cinema. This technique is also used to bring content from beyond the screen, into
the frame. An example is to utilize the surround speakers to introduce sounds from behind the
audience as in the film Australia (Luhmann 2008) where we first hear the sound of wind in the
surround speakers before a stealth-like Japanese Zero bomber appears on screen.

Figure 1: This image shows that the addition of surround speakers allows sounds to emanate from beyond the single
screen plane.
Digital 3-D

Although 3-D has had several resurgences, it has not been until recent years that the latest digital stereoscopic 3-D incarnation appears to have held longer than previous attempts with analogue 3-D. The ground-breaking film *Avatar* (Cameron 2009), although not the first digital 3-D release, certainly set the benchmark as to what can be achieved with this new form of 3-D film making.

The release of James Cameron’s *Avatar* changed the way Hollywood produces feature films. Studios are reporting record-breaking box office sales, and six of the 10 highest grossing movies of all time are 3-D feature films released since and including *Avatar*. Heightened enthusiasm by theater-goers for 3-D demonstrates a willingness to pay premium prices (Altman 2011).

Currently there is very limited documentation on sound design within the 3-D feature film context. As such, there are no rules or standardizations associated with this new practice. Sound designers and film sound mixers continue to use practices associated with 2-D cinema sound, with limited or cautious experimentation and manipulation of spatial sound design within this new 3-D environment. *Gravity* (Cuarón 2013) is an exception as the dialogue pans throughout all speakers, not only the center channel. With such work practices and technology yet to be standardized, and indeed not appearing to be standardized any time soon, now is the time to question contemporary practices of sound design specific to 3-D feature films.

One of the biggest distinctions with contemporary 3-D is the way in which directors and creatives are rethinking the use of the frame of the image. The 3-D image now has an added “depth” value and this is adding to the complexities of framing sound with the image within this additional dimension, and the Z-Space. Not only is there a narrative script, but as noted on *Wreck it Ralph* (Moore 2012) we now have a “depth” script.

The work on the film’s “depth script” resulted in deploying less stereoscopic depth in the emotional scenes and gradually increasing the depth as Ralph, an old-school video-game villain, ventures into more and more advanced games (Hazelton 2012).

There is great debate at present to the validity of this new emergence of digital 3-D filmmaking as a story enhancer. Many state that the 3-D experience is underwhelming, especially considering the premium price of tickets. The 3-D option has been a huge financial boon to the film industry, allowing movie theaters to justify higher ticket prices. In New York, a normal ticket costs $14, while a 3-D film costs $18, or 28 percent more. An IMAX 3-D screening costs $20, a 43 percent premium. Although costing more, certainly it must be said that audiences are still demanding high quality 3-D, with films such as *Avatar* (Cameron 2009), *Hugo* (Scorsese 2011), *Life of Pi* (Lee 2012) and *The Hobbit* (Jackson 2012) all setting new standards in visual possibilities, and all being box office successes. It is hard to argue that these 3-D films have not provided audiences with new immersive experiences superior to that of 2-D.

When talking about the 3-D conversion of James Cameron’s *Titanic*, Barbara Klinger comments on the advantage of the post-converted 3-D version of the film from the 2-D original:

> The word immersion is kind of a cliché, but I think it does have the effect of immersing you more in the action. And I’m not a huge fan of that film, but I admire what he did in the conversion process (Phillips 2013).

*Titanic* (Cameron 1997) the winner of 11 Academy Awards including Best Picture after already making $600 million from an estimated $200 million budget from the original release, did not need a 3-D release to sell tickets.
Where is Cinema Sound Heading in the Short Term

With the release of Toy Story 3 (Unkrich 2010), the first signs of a change in sound for 3-D started to take place. Toy Story marked the first cinematic film release of the Dolby 7.1 surround format. By increasing the speaker channels from the traditional 5.1 (6 channel) format to a 7.1 (8 channel) format, Dolby 7.1 offers both the creators and the audience an increase of an additional two surround audio channels. The increase in channels divides the side walls from the rear walls keeping them discrete from each other, something 5.1 does not allow for. Increasing the surround channels allows for greater directionality and movement of sound from the front of screen to the side walls and to the rear of the cinema.

Consistent with interviews from industry practitioners is that unanimously, all participants agreed that Dolby 7.1 improves the sound design of 3-D films over Dolby 5.1. Dolby 7.1 allows the sound designer and re-recording mixer to have greater flexibility with accuracy to place sounds. This increase also allows a greater fluidity of movement between the screen and the rear of the room. Although being introduced with a 3-D film release, “7.1 is a new technology and can increase the immersive experience of any film” (Nelson 2012).

The Introduction of Immersive Sound

Only two years after the introduction of its 7.1 format, Dolby released what is now known as an immersive sound format, Dolby Atmos. In many ways Dolby Atmos provides a new sound exhibition format for 3-D as it allows up to 64 channels of audio to be discretely positioned around the cinema, including overhead in the ceiling. This differs from earlier sound formats as “objects” can be given directionality and panned around the room providing far greater accuracy than previous surround systems. No longer is a sound bound to a channel, instead it is given a meta data reference to a location in space. The Dolby Atmos system uses a model known as “object based” panning, as individual elements within the soundtrack can be isolated and independently positioned around the cinema. Peter Jackson in a statement to the Hollywood Reporter praised the technology used on The Hobbit (Jackson 2012):

I strive to make movies that allow the audience to participate in the events onscreen, rather than just watch them unfold. Wonderful technology is now available to support this goal: high frame rates, 3-D, and now the stunning Dolby Atmos system (Giardina 2012).

With the introduction of immersive sound technologies in 2012, it is already apparent that a format war is evolving. Competing with Dolby is the Barco Auro 3-D format. Although not an object based system, Auro 3-D also has additional height and a top speaker configuration. By including height information, both Dolby and Auro are claiming to provide 3-D sound, as the height information completes the third-dimension. Although the immersive film sound technologies are marketed as providing 3-D sound, scholarly research outside the film sound discipline state that it is difficult to have true 3-D sound recreation using a sound system that replicates sound through loudspeakers. This is echoed in a recent patent for Object-based audio systems using vector based amplitude panning.

Typical channel-based audio distribution systems are also unsuited for 3-D video applications because they are incapable of rendering sound accurately in three-dimensional space. These systems are limited by the number and position of speakers and by the fact that psychoacoustic principles are generally ignored. As a result, even the most elaborate sound systems create merely a rough simulation of an acoustic space,
which does not approximate a true 3-D or multi-dimensional presentation (Lemieux, Dressler, and Jot 2013).

This raises the question of what determines true 3-D film sound? The argument is for 3-D sound to be true; there must be complete control over the perceived spatial imagery. Acoustics and environment have such an impact on this replication that controlling these variables whilst using speakers within a cinema space is extremely difficult and ultimately impractical. As every seat in a cinema is in a different location Begault states that “it is impossible to predict the position of the listener or of the speakers in any given situation and impossible to compensate for multiple listeners” (Begault 2000, 176).

When listening to the sound depth of a film, the listener must distinguish between the ambience and atmospheric sounds that establish the environment, and the sounds that are tied to visual cues. The atmosphere sounds allow the sound designer to portray almost an infinite amount of depth depending on the story. Throughout Gravity (Cuarón 2013), although the atmospheric sound is sparse, it is the well crafted ambiences within the Z-Space that allows the viewer to be further immersed in the visuals. Dealing with specific individual sound elements bonded to the 3-D images in the negative parallax Z-Space however, becomes difficult. The solution may require the re-recording mixer to utilise the surround channels to “bring the sound off the screen.” Although the surrounds work well with off screen cues, using them to bring the on-screen sound into Z-Space is problematic. The dislocation is also proving to be apparent with the immersive sound formats. Although immersive sound is undeniably superior sounding than previous 5.1 and 7.1 formats, they remain unable to accurately position a sound in the middle of the theatre, in Z-Space. Instead, immersive formats effectively create a highly detailed “canopy” effect.

At present there is debate in industry as to how best approach sound for 3-D. The discussion is further complicated as sound for 3-D films and 3-D sound are not one and the same. The format war between Dolby Atmos and Barco Auro has made many cinema owners nervous as investing in technology is a costly gamble. Some cinema managers have expressed concerns of backing the wrong sound format as not every film is being released in both formats. An Open-Standard MDA (Multi Dimensional Audio) surround format has been proposed to allow greater compatibility of formats.

**Conclusion**

One of the concerns with current working methodologies is, does increasing the number of speaker channels equal a better representation of sound for 3-D film? The commonality of current sound systems is that they all project sound from fixed point speakers placed on the cinema walls and with Dolby Atmos and Auro 3-D, now the ceiling. Do we need to step aside from the contemporary model and begin with a blank canvas in order to better represent 3-D visuals with the soundtrack? How do we make a sound appear to be emitted directly in front of us, just like when we are having an intimate conversation with another person? How can we create the sound of someone whispering in our ear? How can we better represent an arrow flying towards us, splitting the hairs on our head and then landing behind us? Current systems come close to providing these solutions however sound isolation remains the biggest hurdle currently facing 3-D cinema sound. As it stands today, sound for 3-D films has not reached a point of perfection.

Although new immersive sound formats have advanced cinema sound, the fundamental problem with current sound practices for 3-D cinema is the inability to position a single isolated sound within the cinema Z-Space, enabling every audience member to hear exactly the same sound.
REFERENCES


Apocalypse Now, directed by Francis Ford Coppola. (1979, Culver City, CA: Paramount Pictures)

Australia, directed by Baz Luhrmann. (2008, Century City, CA: 20th Century Fox)

Avatar, directed by James Cameron. (2009, Century City, CA: Twentieth Century Fox Film Corporation)

Begault, Durand. 2000. 3-D Sound for Virtual Reality and Multimedia. NASA.


Fantasia, directed by Algar, James, Samuel Armstrong, Ford Beebe, Norman Ferguson, Jim Handley, T Hee, Wilfred Jackson, et al. (1940, Burbank, CA: Walt Disney Productions)


The Jazz Singer, directed by Alan Crosland. (1927, Burbank, CA: Warner Bros.)

Thom, Randy. 2012. Personal Interview: 3D Sound discussion at Skywalker Sound. February 2012.


Toy Story 3, directed by Lee Unkrich. (2010, Burbank, CA: Walt Disney Pictures)


ABOUT THE AUTHOR

Damian Candusso: Senior Lecturer, School of Communications and Creative Industries, Charles Sturt University, Wagga Wagga, New South Wales, Australia
The International Journal of New Media, Technology and the Arts is one of four thematically focused journals in the collection of journals that support the Arts and Society knowledge community—its journals, book series, conference and online community.

The International Journal of New Media, Technology, and the Arts explores technologies of arts production and reproduction old and new, including photography, film, video, multimedia, and the Internet.

As well as papers of a traditional scholarly type, this journal invites presentations of practice—including experimental forms of documentation and exegeses that can with equal validity be interrogated through a process of academic peer review. This, for instance, might take the form of a series of images representing artistic practice, together with explanatory notes that articulate this practice with other, significantly similar or different and explicitly referenced practices.