All-Dimension Surround Sound in the 3D Film Era

By

Yue Lu
2011009434

Magister Artium
Department of Drama and Theatre Arts
University of the Free States

Supervisor: Prof N.J.Luwes

Co-supervisor: Mr D.C. Cloete

STATEMENT

I declare that the dissertation hereby submitted by me for the MA degree at the University of the Free State is my own independent work and has not been previously submitted by me to another university. I furthermore cede copyright of the dissertation in favour of the University of the Free State.

Abstract

3D movies provided an improved immersion in terms of visual perception. But for the associate movie soundtracks, most of the 3D movies were still produced in the conventional format, such as 5.1 or 7.1 surround format. These conventional surround formats were all insufficient in height information of auditory, inaccuracy of spatial localization as well as lack of immediacy. Many sound professionals and auditory technology companies focused on realizing the audio-visual congruence of 3D movie.

In this paper, I would introduce a new auditory concept: ADSS (All Dimensional Surround Sound). By reviewing the history of film sound, I would explore the forming process, theoretical basis, current application and the value of this three-dimensional surround format. With this research, I wanted to raise more filmmakers and film theorists' awareness of the technological development of film sound and utilize it to make more and more blockbusters for the audience.

Index Terms

3D film, three-dimensional sound, reestablishment of sound field, ADSS, Audio-visual integration

Acknowledgements

I am truly thankful to my two supervisors N.J. Luwes and D.C. Cloete at the University of Free State for helping me pursue this research on such a timely research. As few research information could be followed and insufficient of whole introduction of the concept ADSS, they supported me from the very beginning. During the whole research process, they were always patient with this indigestible topic and my different English expression style. They gave me a lot of professional advises and did they best to help me collecting information and data of this totally new technology.

The auditory technicality made this research hard to understand. Here I have to show my deepest gratitude to the members from graduate school of the UFS. They had to discuss the feasibility of this topic and assess the value of this research.

Finally, I also wish to thank all my friends and my family who supported me along the way. In particular, I thank Chen Xu, my best friend, my best family member and my best audient who always ready for helping me. A special thank goes to my parents whose countless sacrifices led me to where I am today.

Contents

Chapter 1

Research proposal

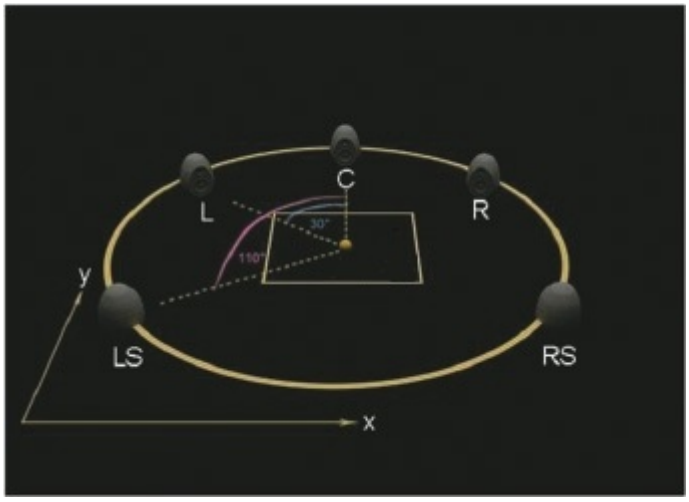
1.1 Introduction

At the end of 2009, a science fiction film *Avatar* (2009) gained world-wide attention (Sina Entertainment, 2010). Until July 2013, *Avatar* was still rated as No.1 box office movie according to Mojo.com (Box Office Mojo, 2013). The amazing 3D visual effects of this film enhanced the visual aesthetic experience. For a while, many 3D films flooded the public's view. It is not the first time the public experience 3D film, but it is the most enthusiastic. 3D films and IMAX films became a new trend in film watching and production. The visual dimensions of film extended to a three-dimensional space, but the sound is still produced in the conventional multi-channel sound format, such as 5.1 or 7.1 surround format. When the audiences watch 3D films, the images are within reach, but the corresponding sounds still come from the fixed positions such as screen or cinema walls. This phenomenon creates a paradox: film sounds are uncoordinated with frames, seriously impacting on the audience's aesthetic experience of 3D film. Sound technicians and film theorists raised the issue: since the visual technology has extended to the three-dimensional space, the sound production should also develop accordingly to achieve the coordination of film frames and sounds.

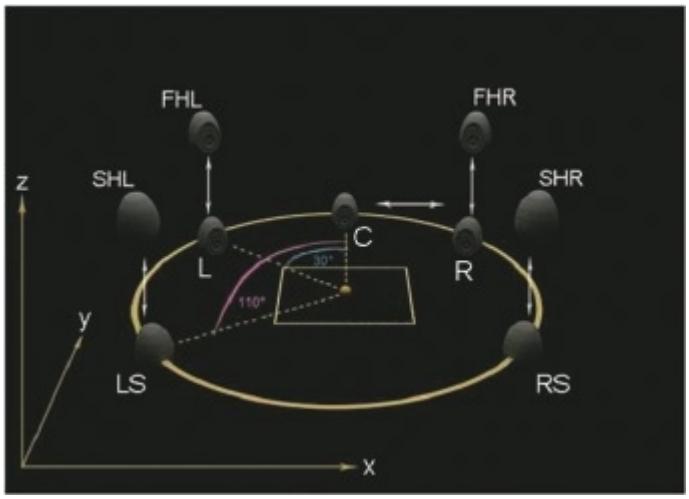
The development process of film frame and film sound is like a chase race: The first film *The Factory gates* (1895) was showed in 1895 (Baidu Library, 2011), but the first "talkie", *The Jazz Singer* (1927), was only released in 1927 (Baidu Library, 2011). From silent film to "talkies", sound was only heard more than 30 years after images in film history. As sound technology kept developing, the invention of Dolby noise reduction and a series of digital sound technologies showed that the sound technology had taken the lead to enter the digital domain, while the picture was still in

the 2D film domain (Dancyger 2010). Soon, the appearance of 3D film showed that frame production had entered into the 3D digital technology era.

In order to remain relevant in terms of the development of film visuals, a new conception of sound reproduction was needed. All dimensional sound (ADSS) emerged. The conventional multi-channel sound format places the loudspeakers all in horizontal planes and they can only express the localization of sound images in two-dimensional space: front and back, left and right. It directly results in two disadvantages: inaccurate sound localization and auditory incoherency. Different from two-dimensional sound, as showed in the illustration below, ADSS can locate sound images in a three-dimensional space.



5.1 surround sound system (two-dimensional space)



ADSS system (three-dimension space)

By supplying the missing sound information in height and realizing a precise sound localization, this new surround sound format creates a perception that film sound comes from almost anywhere in a 360-degree arc at distances from very near to far away and truly provides a fully-immersive sound environment for the audience. Thereby, it realized the coordination of image and sound in 3D film, not only in the technical field, but also in aesthetic creations. In term of technique, as the concept of ADSS appeared, the 3D film realized reproduction in the true three-dimensional space, both the frames and the sound. In term of aesthetic, ADSS realized the coordination of the frames and sound, thus realized the real sense of 3D film.

1.2 Research problem and the objectives

The development of 3D visual technology promotes acoustic development. 3D visual technologies are more technical to be accepted by filmmakers and the general public. In contrast, ADSS is still at its early stage of mainstream appropriation. Due to the specific technical nature and invisible artistic characteristics, professional filmmakers do not know of this technology, let alone the audience.

Therefore, the research problem is: utilizing the theoretical concept of ADSS to realize three-dimensional film sound reproduction, thereby realizing the coordination of vision and audition in 3D film practice.

1.2.1 Primary objective:

I want to explore the concept of All-Dimensional Surround Sound because this concept is relatively unfamiliar to the film theorists, filmmakers and the theatre crew. By exploring the proposition, ADSS can be incorporated by more film professionals and the public.

1.2.2 Secondary objectives

Examine the background and development history of ADSS

Explain the design principles and the specific application of ADSS

Discuss the technological value and aesthetic value of ADSS by analysing the importance / value of the coordination of visual and auditory elements in 3D film production

Provide a possible forecast of the development and trends of ADSS in 3D film era

1.3 Research design and research methodology

1.3.1 Research design

The research design of a study provides the framework for the collection and analysis of data (Quinlan, 2010:420). In this study, qualitative research method will be used to explain the concept: ADSS, and the analysis of current practical applications of this technology and an exploration its artistic value.

1.3.2 Data collection methods

Content study and case analysis will be used in the study using David Bordwell's Historical Poetics approach (David B, 2008).

1.3.2.1 Content study

Primary focus will be kept on gaining a deep, intense overview of content of study and gathering data through documents, photographs, video recordings and other media will be appropriated in the content study.

This study will aim to explain the following three aspects:

- The theoretical basis: ADSS is based on the human ears' physiological structure and reproduction of the auditory environment in film practice.
- The background and development history of ADSS
- The design principles of ADSS

1.3.2.2 Case study

The analysis of the specific applications of ADSS, as currently being applied as Auro-3D, Dolby Atoms and IOSONO 3D

- The practical applications in sound format: Dolby Atmos, Auro-3D, IOSONO 3D
- Current application in cinema: Korean CGV cinema, one of the cinemas that applied ADSS and 3D films based on Auro-3D and Dolby Atmos: *Red Tails* (2012), *Brave* (2012)

1.3.3 Layout and the overview of chapters:

Chapter 1: Orientation and Research Methodology

This chapter provides an overview of this research and introduces the research methodology.

Chapter 2: Introduction

In this chapter, through the analysis of the phenomenon of the disequilibrium between the development of film sound and images, a timeline is created that leads to the concept of ADSS will be created. An introduction to the background and the development history of ADSS will also be provided.

Chapter 3: Exploring the theoretical basis and practical applications of ADSS

This chapter will firstly provide the theoretical basis of ADSS, including the

physiological working principle of the human ear as well as the reproduction of auditory environment in film practice. This theoretical basis will contribute to the exploration of the conception of ADSS. This concept leads to and determines the design principles of ADSS in the view of the researcher. The practical applications of ADSS provide examples to discuss the technological value and film aesthetic value of ADSS. This discussion will extend to the relationship between technology and conception which could be the guide for future development of 3D film.

Chapter 4: Conclusions and recommendations

This part will highlight different production principles and non-uniform project standards in an attempt to forecast the future trends associated with ADSS in 3D film era.

1.4 Value of the research

ADSS is a brand-new field of research in film sound reproduction studies. In terms of technology, it improves the development of sound techniques, promotes the development of film production and satisfies the audience's increasing aesthetic demands. In terms of art, it extends the film space ("fabula" and "syuzhet") for filmmakers. ADSS can express the sound information more precisely and create a truly immersive perception for the audience (Yesky, 2012).

Since ADSS is a such useful and worthy technological concept, my study can help to raise more filmmakers and film theorists' awareness of the technological development of film sound and utilize it to make more and more blockbusters for the audience. This technological is currently only being appropriated in cinemas and theatres, but I believe soon after the film theorists excavate the significance of ADSS. Thereby, the creative concept of film production will also be developed into a 3D field.

New technologies are always a great pull for the box-office. Since downloading films from internet has become an everyday occurrence, 3D film and ADSS will attract the audience come back to cinema again. The 3D film era represents a new interval in technological developments for film sound production and as such provides practitioners with infinite possibilities.

Chapter 2

Introduction

2.1 Preface

After undergraduate, I worked as a sound engineer for over three years. During the three-year working period, I participated in the sound production for many films. Among these films, a thriller film was very impressive in my memory. Thriller films are accentuated in sound effects to enhance the tense atmosphere. In many scenes of this film, there were various fly-over sound effects and a lot of invisible sounds should be located in the designated positions. But the exiting sound-processing techniques at that time were unable to realize such requirements. By then, the relatively advanced sound-processing software was Pro tools HD and for most of cinemas Dolby Digital was the cinema industry standard. The prevalent sound reproduction system was 5.1 or 7.1surround format. Unfortunately, none of these surround sound system would provide a truly immersive audio experience.

2.2 A phenomenon in film production since film *Avatar* (2009) was released

At the end of 2009, most of the major directors and filmmakers took notice of an amazing science fiction film *Avatar* (2009), James Cameron's sci-fi epic 3D film. Depending on the extraordinary digital film technologies, film *Avatar* (2009) created a nearly perfect visual saga. It overtook its predecessor's record of $1843bn at the weekend, after six weeks topping box-office charts around the globe. Using Chinese film market as an example, this film won 1.4 billion yuan box-office (the monetary unit of China), over 2000 million dollar (Sohu Movie Club, 2010). The fantastic story and incredible 3D visual effects deeply impacted all the audience. This motion was even fueled by the mainstream media. More and more filmmakers realized that the film production had developed into a new period of film industry-3D film era. The huge

success of box office aroused a lot of discussions and research about this film. Surprisingly, in the 82th Academy Award, two most technical sound awards, Best Sound Mixing as well as Best Sound Editing, were all granted to another film *The Hurt Locker* (2009) (Douban, 2010).

After reading many related reference and analyzing the film marketing, I considered that the fundamental reason behind the failure in sound award of film *Avatar*'s is that, considering about the essence of film, such impressive visual effect was combined with an inconsistent film sound. More specifically, for the 3D visual effects, film *Avatar* (2009) is beyond comparison. But the audio processing and reproduction of this film are still in a two-dimensional extent. This incompatibility between the film frame and sound is the main reason that resulted in its failure in sound production. Filmmakers and sound technicians realized the incongruence in 3D film production and raised the concept of three-dimensional film sound.

Actually, 3D film and the conception of three-dimensional sound are not new concepts. Since 1980s, when the digital technologies were introduced into film production (360 Beta, 2006), implementing the true surround sound had become a primary objective for sound engineers. But the re-launch of 3D film in the form of huge budget blockbusters, such as *Avatar*, raised the issue again. Since a film cost many hundreds of millions dollars to make, it means there must be vast funds available to produce an appropriate film soundtrack.

3D film is familiar with the traditional 2D film in the basic concept of film production: using light to create film images on the screen and satisfy the audience's aesthetic desired (Mtiome, 2008). For the technological principle of 2D analog films, the way to create images is using total visual elements: shadow, images, lines, color composition and etc (Maigoo, 2012). But for 3D digital films, the production method is utilizing the chromatic aberration of two film projectors to create a three-dimensional film space and the audiences have to rely on the 3D glassed to receive film information and

enjoy the film story (Chen and Chen 1992). That is 3D film employs a distinctly different film projection technique from 2D film. Therefore, for these two different film projection techniques, the creative concept and aesthetic experience of film sound are totally different. In the 2D film era, film sound kept developing: from monophonic recording and reproduction to stereophonic format, and then 5.1, 7.1surround formats. Nowadays, multi-channel surround format had been employed almost in all films. As 3D film production has become a new trend of film production in the future, 3D film will be accepted as a universal film production form for the filmmakers and a common standard for the audience (DOC88.com, 2013). Then, as mentioned above, 2D film and 3D film are two different kinds of film form, the present Multi-channel surround sound system will not be appropriate for 3D film.

The development of film picture and sound is a dynamic development course: from silent film to “talkies”, surround sound film to 3D stereoscopic film, this development law has never changed. But since the 1990s, when commercial digital surround sound format was widely employed in film (Baidu library, 2010), the development of film sound had preceded film frames. However, the amazing stereoscopic film frames made the filmmakers realized that the conventional multi-channel surround sound format, which had been applied in the film production for almost 40 years, only reproduce in horizontal dimensions and it was not a real sense of three-dimension sound. High-fidelity of sound quality, precise sound localization, better originality of sound effects and different sound experience, the characteristics of the corresponding 3D surround sound format for 3D film raised more requirements to the sound professionals: imagination, programming, innovation and setting new producing standards.

ADSS, a brand-new surround sound format, not only remedy the dimensional defect for the existing multi-channel surround system, but also provide a more anticipated probability for sound localization. What is more, it can provide more vivid and precise aesthetic experience of auditory sense.

2.3 The development history of film sound

Film, in fact, is a product of technological invention (Doc88.com, 2009). That is why the development of film production was always followed the development of technology or even technology-centered developed (Baidu Library, 2012). New technologies could always bring about new manifestation modes, themes, contents and theories, which promoted the generation of new aesthetic concepts and new methods for artistic creation, all the more so film production. From silent film to “talkies” (sound film), mono film to stereo film, analog-sound film to digital-sound film, film sound went through three major revolutionary stages. All the stages reflected the technological innovations and important changes of aesthetic concepts.

2.3.1 From silent film to sound film

Since film, as a seventh artistic form (360Beta, 2013), had been invented, it was always a concern in the public eye.

As the limited of technical specifications, the early films could only be produced in a silent form (DOC88.com, 2008). Silent film is a unique film form. It has total different aesthetic characteristics from the subsequent phonofilm (phonofilm is an optical sound-on-film form of film production developed by inventors Lee de Forest and Theodore Case in the 1920s).

In silent film period, there was no film sound along with the film frames. Sometimes, subtitles would be added as a supplement in some plots to explain complicated dialogues, evaluate the plots or hint the following stories. The actors had to use exaggerated facial expressions and body movements in film acting to tell stories and express film characters’ inner thoughts. As for then, this visual method was the clearest expression form for silent films. At that time, comedy films were very popular,

because this over-performing was natural and attractive. But for nowadays, this over-performing form seems somewhat superficial and bumbling (360 Beta, 2013).

In fact, this so called silent film was not totally silent. From an objective perspective, there were some unconscious sounds when silent films were projecting. Nobody could ignore the noises produced by the early film projectors. But it seemed that the audience chose to ignore the noises and did not been disturbed by them at all. They paid more attentions on the constant changed pictures on the screen and tried to catch up with the film stories. This is a phenomenon of psychoacoustics named cocktail party effect, also called as masking effect (360Beta, 2013). This acoustic effect comes from an interesting phenomenon: when two speakers talking in a cocktail party, no matter how noisy the environment is, they can still hear each other's voice and continue the conversation successfully. The humans have this capacity to unconsciously ignore other sounds and choose a specific sound which they want to focus on.

Although, the audience chose to ignore the machinery noises, the filmmakers did not take this in an optimistic way. As the contents of film stories became more and more abundant, the length of films increased correspondingly. The audience could no longer subjectively ignore the machinery noises all the time and their enthusiasm of watching films were also influenced by the annoying noises. In order to cover these machinery noises, filmmakers introduced live music accompaniment into the silent film. Besides film subtitle, live music accompaniment became another supplementary element for silent films. This supplement in auditory sense commendably covered the machinery noises and made the audience distract attentions again to the film stories. Early in 1895, at Lumiere's film premiere, there was a pianist accompanying during the process of film projection (360Beta, 2013). Later, filmmakers found out that, during the film projection, live music could not just simply cover the noises. It could be related to the stories, thereby create a certain atmosphere or motivate the audience's emotion. These features could greatly enhance the appeal of films. Soon afterwards, the

on-site musicians or a whole orchestra successively appeared in cinemas. Live music accompaniment was universally employed in the silent film projection.

Live music accompaniment, after all, was an impromptu performance. The randomness determined that the on-the-spot effects would be different from time to time. Therefore, it would not accurately convey the directors and filmmakers' creative thoughts every time. As the film production became more formal and professional, filmmakers began to tailor music for films: they assembled musicians, string players and conductors to compose specific music for films. In Griffith's epoch-making film *The Birth of a Nation* (1915), all the music appeared in the film were originally composed for it (360Beta, 2013). From then on, original music or special composed music became a popular trend, and soon form a universal film standard.

Meanwhile, technologies were also developing along with the new concepts of film production. These attempts were put into the trial operations for a short while. Early in 1895, Thomas Edison connected his two inventions, the kinetoscope and the phonograph, with a belt to realize the mechanical synchronization of frame and sound. He named it "Kinetophone" (360Beta, 2013). The viewer would look into the peep-holes of the kinetoscope to watch the motion pictures while listen to the accompanying phonograph through two rubber ear tubes connected to the machine. Unfortunately, the sound from the phonograph was not loud enough and hardly to be heard clearly by the viewers. Soon, the defects of this method exposed: this mechanical connection of video and audio devices always resulted in the asynchronization of film frames and the sound. Before long, this method was abandoned on account of the inferior synchronous effect, let alone the inadequate recording quality and amplification.

Then filmmakers realized that, since the films could record the frames, it might be possible to use the films to record the sound. So the technical issue about the asynchronization of pictures and sound was solved in 1907. Dr. Lee De Forest

invented the “audion”, a device capable of amplifying weak electrical signals in a vacuum tube (Baidu library, 2014). “Audion” was the foundation for radio, television and other electronics. Later, it was employed in broadcasting and film sound reproduction. In 1926, BTL (Bell Telephone Laboratory) cooperated with Western Electric and developed “Vitaphone”, a new sound-on-disc system of synchronizing pictures and sound (Wikipedia, 2014). When the filmmakers were shooting a film, they used two different kinds of negative films: one was for recording the film frames and another one was for recording the live sounds. In the post-production period, filmmakers would mix these two records on the official negative film. When projecting the film in the cinema, there would be an extra sound device for the film sound reproduction. What’s more, the speed of film shooting and projecting was changed from 16fps to 24fps (fps: frames per second).

Warner Brothers Studio accepted this system and shot an experimental film *Don Juan* (1926) to try out this new idea (Baidu Library, 2014). There was no recorded dialogue in this film, the narrative being carried on the usual cards of silent films. The soundtrack was music only with a few sound effects such as clashing swords. The audience was greatly impressed by this film and enthusiastic reviews appeared in the following day’s papers. However, most filmmakers and directors still agreed that music was OK but the audience did not want the motion pictures to “talk”.

This opinion did not change until 1927, when Warners released the first real sound film *The Jazz Singer* (1927) in America (Baidu Library, 2014) . It was a basically silent film with several song numbers sung on screen by the comedic star, Al Jolson. The first dialogue in film history was:

“*Wait a minute, wait a minute, you ain’t heard nothin’ yet!*”

This film got an outstanding box office hit and made a widely influence. Cinema goers, who were taken by the snatches of ad-libbed dialogue spoken by Jolson, flocked to see this film. The success of this film, coupled with the cinema-going public’s demand

for “talkies”, caused other film companies as well as film theorists to realize that sound film had arrived.

A new creative conception is always the bone of theoretical research. The appearance of sound film raised a heated discussion (actually, this argument is still carried on in today’s film theoretical research): whether film should be a pure visual art or a combinative art of vision and audition. At that period, many European filmmakers and theories thought that overmuch dialogue might weaken the special artistic value of silent films (Doc In, 2013). In America, sound film became the primary form in film production. On the contrary, silent film was still popular in Japan. Almost all the film theorists in that period, such as Merieux, Griffith, Chaplin, Eisenstein and Murnau, were involved into this discussion. From their creative practice, they accumulated and summarized up a series of mature film theories.

A representative figure who reacted strongly against sound film was famous comedian Chaplin. He insisted that sound would destroy the aesthetics of silent film. Film sound was just a technological creature, not an artwork. In spite of the inevitably historical limitations in his opinions, the real reason was this new film form seriously impacted the box office of his films. Losing economic benefit and being abandoned by the audience made him felt intense repulsion for sound film (Baidu Library, 2014). But ironically, he shot his first sound film *The Great Dictator* (1941) secretly. After all, the audience had their own criterions. Filmmakers and film theorists could only lead the audience’s value orientation, but they could not change it.

Still, there were quite a few film theorists who supported sound film. Eisenstein with Pudomn and Alexandrov, who were famous film artists and theorists at that time, released a joint statement on sound film in 1928 (Baidu Library, 2011). They wrote:

> *“…..Only a contrapuntal use of sound in relation to the visual montage piece will afford a new potentiality of montage development and perfection. The first experimental work with sound must be directed along the ling of its*

distinct non-synchronization with the visual images. And only such an attack will give the necessary palpability which will later lead to the creation of an orchestral counterpoint of visual and aural images……"

Unquestionably, the support of these film theorists was an encouragement for the popularization of sound film. Soon sound film sprang up like mushrooms and every major studio rushed to get on the bandwagon.

Sound, as an indispensable film element, formally joined into the film production. This technological change made film developed from a pure technological invention into an artistic form and then developed into a specific culture.

2.3.2 From mono film to stereo film

Since sound was introduced into film, film was no longer a "dumb". Directors, scriptwriters, actors and filmmakers were set free from the limit of sound expression. They found out that film sound could express more abundant contents and had huge potential for development. Sound film became a mainstream trend of film production. According to information statistics, after the appearance of sound film in Hollywood, the number of the audience soared from 60 million in 1927 increasing to 110 million in 1929, almost a double figure in the two-year time (Dou In, 2011). But the rapid development of sound film was accompanied by a series of relevant problems.

At first, filmmakers did not cognize film sound in terms of artistic creation. They just random added some dialogues to cater to the audience's curious about film sound. The films were always full of meaningless dialogues and inexplicable sound effects. Soon, the audience got bored with the abuse of film sound. They were no longer simply satisfied with "talking" films.

After America stock market crash and economic depression, film production restarted

again. In 1933, RKO (Radio Keith Orpheum Pictures) released the film *King Kong* (1933). This film became a milestone in film sound history. Murray Spivak, an innovative sound technician, creatively devised all of the sound effects for the original *King Kong* motion picture (Baidu Library, 2011). He created King Kong' roar through a series of process: descending the lion's roar by one octave and mixing it with other sounds in the same pitch. It was the first time that filmmakers produced film sound in the artistic way.

Scientific and technological progress and innovation are always the main driving force of film production. During the long-term study and trial, scientists found that human's ear was a complicated auditory system, with the capability not only to distinguish the basic physical characteristics but also to receive the directional and spatial information of sound (Dou In, 2013).

In the early 1930s, BTL, cooperating with the famous conductor Leopald Stokowski, conducted an exploratory stereophonic experiment on "acoustic perspective". Later, in 1935, most of Alan Blumlein's development works on stereo system for cinematic use reached completion and were applied in a few short test films, *Trains at Hayes Station* and *The Walking and Talking Film*. His aim of having film sound associating with the frames and following with the motion pictures were fully realized (BBC News, 2008). Since then, the technological foundations of stereo film production had completed. Filmmakers had a relatively advanced cognition about film sound and some new concepts of film production were in preparation.

In 1940, Walt Disney released film *Fantasia* (1940), the first film produced by multi-channel format (it was named as "Fantasound" at that time). The original recordings were made with 8 push-pull variable-area channels: six were used for close-ups of various orchestral groups. The seventh recorded a combination of all six, while the eighth was for a distant pickup of the orchestra. The original channels were transferred to three tracks. Using optical sound recording system, these, plus a gain

control track for the three stage tracks, were recorded on to a separate 35mm film, run in synchronization with the pictures (Baidu Library, 2013).

In consideration of the huge demand of loudspeakers (54 loudspeakers placed throughout the auditorium) in this system, only two theaters were equipped with the full set of “Fantasound” system. The installations cost up to $85000. On account of the high cost, “Fantasound” was simplified into a “Fantasia Road Shows” system and toured a total 14 theaters, causing quite a stir at that time. The main innovations and reformations that particularly were worth mentioning of this pioneer film:

- The click track
- Dispersion-aligned loudspeaker system with skewed-horn
- The pan-pot
- Control-track level-expansion
- Overdubbing of orchestral parts
- Simultaneous multi-track recording
- A multi-channel surround system

After World War Ⅱ, television production was quickly catching on. In America alone, nationwide surveys revealed that, from 1949 to 1951, the ownership of television sets shot up from 1 million to over 10 million. TV programs offered a wide variety of options for the public. Farces, vaudevilles, cartoons, entertainment programs and Hollywood films were often broadcasted by TV stations. Moving pictures were no longer a patent of film. Comparing with the fixed film schedules in cinemas, watching TV programs at home was more convenient than going to the cinema for the general public.

TV industry became a formidable enemy to film industry. Wide-screen films, as a response to the huge impact of TV production, were expanded rapidly. Film sound, along with the appearance and quick development of magnetic recording medium, developed accordingly. Taking Cinerama system, which was released in 1952, as an example, there were three stereo tracks and a narrower control track. The magnetic

tracks were superimposed on the 35mm film by a striping process. The audience, for the first time, experienced the true Hi-Fidelity stereo sound in the cinema. Unfortunately, Cinerama system was never very popular and doomed from the outset because of its high cost and complexity, but the width-height ratio of wide screen had been retained. This attempt did show the superior ability of film in both vision and audition which would never be achieved by its rival, television (Baidu Library, 2013).

During the same time period, Warner Brothers were experimenting with stereo images, which television could not offer to the audience. Their first trial film *Bwana Devil* (1952), an action-thriller set in Africa, was projected in November (Baidu Library, 2014). This earliest 3D system consisted of two mechanical locked projectors, one for the left eye and another one for the right. When projecting film, two projectors run two copies of the same film and used Polaroid filters on the projectors to complement the Polaroid glasses worn by the audience. Film sound was carried on a separate magnetic film. The channels of the stereo system, recorded on the optical track of the right projector, were left, centre and right with a surround (or effects) channel. A mono optical track was reserved on the left projector as a back-up or for use in the theatres not equipped for stereo. Unfortunately, 3D film did not develop into a standard format for film production at that time. In the after 30 years, 3D film became amusement equipment in playground or the tools of scientific popularization in science museum.

After the heyday of development in fifties, film industry did not offer any new innovation during the 1960s. Few attempts, like IMAX system, were failed to retrieve the dwindling audience and change the status of the general apathy. This situation did not improved until Dolby Laboratories introduced two new concepts to film sound. One was their already successful DNRS (Dolby Noise Reduction System) and another one was matrix audio system (HC360, 2008). Since magnetic recording, as high in quality as it could be, continued to be expensive for film production, quadraphonic matrix audio for stereo optical recordings was heralded as a significant breakthrough for low-cost multi-channel surround sound. The matrix encoded down

the four discrete audio channels to two and recorded them on the film. In the cinema, the stereo optical soundtrack would be decoded back to four. The stereo optical track was fully compatible with the mono one. Among the drawbacks of matrix audio system, the loud noise was the major one. Because the track was split into two and the reduced size gave rise to noise problems. By adding denoiser to magnetic tracks, the dynamic range of medium frequency had been increased and the flux level for on-film recording had been enforced. Although the first Dolby stereo film was *Lisztomania* (1975), it was *Star Wars* (now known as *Star Wars Episode IV: A New Hope*) (1976), a revolutionary film in film sound history, that everyone was aware of this system. With this film, Ben Burtt won the Sound Effects Editing of Academy Award in 1977. The improvement of optical sound revitalized the sound side of film industry. Dolby amended its stereo system after applying DNRS in a few experimental Todd-AO 70mm films: the LE (Left Extra) and RE (Right Extra) speakers were no longer used for full bandwidth soundtracks playback, but for extended bass boxes only. Thereby, the concept of discrete Subwoofer track was formed, which was the predecessor of today's LEF track (low frequency effects). This amendment improved the 70mm sound format to be effectively reduced to four channels plus subwoofer and was firstly employed in Warner's film *Superwoman* (1978) as well as film *Apocalypse Now* (1979), a year later. The popularity of 70mm film waned gradually, however, the concept of adding stereo surround to films became the first step towards the after 5.1 format (with Left, Centre, Right, Left surround, Right surround and LEF).

With a series of technological developments, film sound totally changed conventional concept of film creation. Filmmakers and film theorists began to explore the deeper significance of film sound: the relationship between film frame and the sound and the combination modes of the two film elements.

Since the "great dumb", film, started speaking, the controversy around it had never ceased. At the early period of sound film, some film theorists, represented by Rudolf Arnheim, the famous German film theorist whose major ideology was from the Gestalt

school of psychology, regarded sound film as a continuation of silent film (Baidu Library, 2014). Like the music accompaniment for a concert, sound was just an assistant element for the film vision, while film should be a visual medium. So they raised a fundamental question of about sound film: “whether the existence is reasonable”. He held the important aspects of personal bias, intuition and expression. So that visual perception is what allows us to have a true understanding of experience. With this ideology, the form is greater than the content. When people see something, the brain will reflect the corresponding information. As for sound, it was just a supplement.

Another ideology was from an epistemology of film essence, represented by Siegfried Kracauer, a German film theorist, deemed that camera exceeded sound recorder in contribution (Movie Encyclopedia, 2009). Only the vision was in the primary position (of film) did it conform to film essence. In film, the most distinctive contribution came from the camera rather than the recorder.

Until seventieth, some film theorists still advocated silent film and insisted that film could exist without sound and pictures exceeded sound (Baidu Library, 2012). This transcendental ideology came from the conventional film concepts that film sound was no more than a technique trick added on the silent film to make it more attractive. This school of thought fundamentally confused the means with the results. Film sound and film sound recording are two different concepts. Film sound is an art form that cooperates with film frames. While film sound recording is the means and the process that film sound utilizes to realize its artistic presence and achieve its artistic value. Moreover, even film sound recording itself is not a simple replication process. Every application and process of sound elements is an artistic creation process as well as an aesthetic inspection process.

Film, since its birth, was never silent. As the limitation of technology, sound would not combine with the frames inside the film entity, only be attached during the process of

film project. The concepts of “silent film” or “sound film” are relative to the technologies other than the essence of film. In fact, as a direct extension of acoustic source entity, sound is a physical presence that occupies space and owns the time continuance. It is a concrete sense, as itself, can convey information, express emotions and render atmosphere. So this physical presence, which can be perceived by auditory system, is an independent form of existence. Either in the real world or the virtual film space, the audience desire to acquire both visual images and aural images of objects at the same time. This desire came from the specific functions of film sound. In real world, sound, as a physical presence, people tend to focus on its physical characteristics. While in film space, as a film element, besides its basic physical characteristics, sound is an artistic presence with emotions and colors. It has a dual identity and plays a dual role, not only taking part in setting up the film space, but also carrying an artistic meaning. For example, if we are in the real world, the sound of rain is just a natural phenomenon. But if the same scene appears on the film screen and the director want to utilize the rain to express certain emotions, the viewers will obtain a different perception.

For human being’s cognition of the real world, it is a perceptible space that can be heard as well as seen. Vision and audition give us information about the objects we find in our surrounds. This information is integrated in the brain to form the percept of a single audio-visual object. Human’s eyes, work as a precise optical system, are very sensitive to light. They can quickly collect the major features, such as colors, lines and volume of an object, transmit the information to the brain, and then the brain will form a corresponding image. But the limited of view-angle coverage is 120 degree, while human’s ears are unlimited and can hear all the sound around. The visual and auditory systems interact in human’s brain, but work separately and irreplaceably with each other (Dou In, 2009). Film is precisely based on the working mechanism of our audio-visual system to create a virtual world. The essence of film is recording, thus sounds has significance as well as the pictures. Besides all perspectives (darkroom effects or psychological attributes), film itself requires auditory-visual congruence.

From the perspective of film receiving psychology, the auditory-visual congruence is a stereoscopic mechanism, which is easier to stimulate human's perception.

As in the real world, sounds and pictures are always synchronous. When we see a certain object or be in a certain environment, we do not have to struggle with the tropism of attention, because this synchronization is a natural mechanism (Baidu Library, 2013). But in film space, the auditory-visual congruence is not a simple synchronization. Through artificial process, the different combinations of sound and frames can provide different artistic experience. Generally, the methods of integrating sound and picture are synchronization and contraposition. For the synchronization method, as heard the sound, the audience can find the corresponding frames in the film screen. This combination method accords with human's natural reception mechanism. Contraposition, a concept borrowed from music, means the sound and picture transmit different information at same the time thereby produce a strong impact on the audience (Baidu Library, 2013). The contraposition can be produced as a parallel mode or a separate mode. Parallel mode means the sound and picture go in their own way. For example, a radio is broadcasting in the room. When the picture switches outside, the radio is still on as a background sound. Another mode, as a creative way to process the combination of sound and picture, means that the sound information is contrary with the picture. Then sounds that the audience heard are not from the sound sources in the film screen, or sometimes without sound sources. In film *Apocalypse Now* (1979), the bombing scenes were accompanied by Wagner's epic music. This striking contrast produced a tremendous artistic effect. No matter which mode the directors would take, the concept of audio-visual congruence is always the foundation of film creation.

The audio-visual congruence could be a very nice interpretation for the recession of 3D film in fifties. In the traditional 2D film, the lens of cameras was a frame. All the images had to be designed and presented inside this frame. With the corresponding film sound to cooperate with the images, 2D film utilized its own audio-visual language

to create a virtual three-dimensional space in the 2D frame. While there was no specific frame in 3D film, the images could jump out from the frame of film screen. As opposite to the free images, the film sound still continued the traditional design and format, not to mention the unromantic narrative. The same circumstance happened again in the eightieth. Films like *Friday the 13th* (1980), *Jaws 3* (1982), *Starchaser: The Legend of Orin* (1985), were popular for a time and got a good box office, but soon, 3D films were forgotten as well (Baidu Library, 2014). The fundamental reason for this phenomenon is that, the early 3D films excessive emphasized the visual effects and did not notice the congruence of vision and audition. Like a person with a big head and small body, the imbalance of visual and auditory perception is the main reason that 3D film had never become a trend since it was invented and was abandoned by both audience and filmmakers from time to time.

2.3.3 From analog-sound film to digital-sound film

As going from strength to strength, Dolby's optical stereo system almost monopolized the film sound industry, and undoubtedly became the industry standard. Upgrading cinemas to Dolby stereo had been carried out in a world-wide scale (Baidu Library, 2014). With the development in film production and new technology equipped cinema building program, the audience's appreciation level had correspondingly risen. The new generation of audiences was younger people who were more sensitive to the newest technologies and easily accept them. They were more aware of the importance of sound quality than ever before. Meanwhile, the popularity of Compact Disc (CD) promoted the rising of "home entertainment". The new generation demanded the film could be produced with the same quality sound.

In 1990, Eastman Kodak Company, collaborated with ORC (Optical Radiation Corporation), introduced the world's first digital audio format, CDS (Cinema Digital Sound) for film. In the same year, the film Dick Tracy (1990) was released with this new digital format which was the first time the public heard about digital sound in the

cinema. Unfortunately, it only enjoyed a handful of releases before ORC pulled the format from the marketplace. Actually, the concept of adding the digital sound to films can be traced back to the early eightieth (around 1981 and 1982), when Peter Custer and George Bird invented an audio format named “Digital Fluorescence sound”. The digital sound information was actually printed as fluorescent codes on the frames. When the film was projected in cinema, the fluorescence reader would emit UVA (ultraviolet radiation) to read the digital codes and reproduced 6-channel analogue sound signals without distortion and noise jamming (Baidu Library, 2014).

CDS was a highly engineered technology, incorporating an error detection and correction scheme in the printed bit map to ensure robustness of the digital signal. As the format boasted a very tiny pixel size, audio could be recorded without data compression. But instead of the commonly used 16-bit linear PCM scale, it used a custom 12-bit logarithmic scale that produced excellent sound with a 90 dB dynamic range. But this change brought about one major drawback, that it removed the analogue optical soundtrack and replaced it with their digital one. The film copies were unplayable in a non-digital cinema. In other words, there wasn’t a stereo optical track that could be used as a backup. Moreover, the absence of the backup track directly resulted in the incompatible problem of playback equipment. With so many drawbacks and difficulties, CDS format are still worth to be remembered. As a pioneer for the later digital film sound formats (especially 5.1 surround format), it showed the huge potential of digital film sound formats. What’s more, it was the first time in the film history that the audience would determine which technology they desired. From an artistic point-of-view, discrete digital sound has been a real boon, as it frees the mixers from the limitations and hassle of mixing audio for the matrix. Full range and discrete surround channels have given the mixers the ability to use these channels much more freely (Baidu Library, 2014).

Dolby, however, no longer took the audio industry by storm and was unable to have it all its own way. In Warner Brothers’ film *Batman Returns* (1992), Dolby engineers

cleverly placed audio data between the sprocket holes. This SR/D format, known as Dolby Digital, provided a possible to compress the AC3 audio data for 5.1 audio signals. Learning the mistakes from CDS, Dolby preserved the analog stereo optical recording format on the film alongside the new digital recording. But it was Speilberg's *Jurassic Park* (1993) that established the digital film market. With over 1000 cinemas equipped for this film, DTS (Digital Theater System) took a novel departure from both CDS and Dolby digital. This system placed audio on a separate CD-ROM in a proprietary format and recording only an analog synchronization track on the film. Via a time code on the film, CD-ROM was synchronized to the picture. The DTS had proven to be a very robust method for implementing digital audio on film. Not to be outdone, SONY introduced its SDDS 7.1digital audio format with the dual release of In the *Line of Fire* (1993) and *Last Action Hero* (1993). The different formats brought about a very serious problem, since none of the cinema could afford all three systems, this incompatibility made digital sound very expensive for the cinema. Finally, as DTS begun putting time code onto 70mm film and using the same CD-ROM to generate digital sound for this new format, the problem of incompatibility had been solved.

After ninetieth, as the quick development of technology, new digital audio technology, such as Video Compact Disc (VCD), Digital Video Disk (DVD), Blue-Ray and even Red-Ray, emerged in various types. From 1994, the world had turned on the Internet officially, the Internet suddenly burst and broken the monopolization of conventional film industry. The sharp impact of web download and home theatres became a big blow for cinema industry. Of course, film industry had not been immune to these burgeoning media. Simple system alignment or new format could not withstand the menacing impact of new medium.

Film industry looked for a way out to break away from struggling with attracting the audience back to cinema again. The science fiction film is especially, the most typical genre to meet the audience's reverie of universe, space and the alien. Actually, the public had been in a long favor of science fiction. In 1970s, Dolby surround was

successfully applied in film *Star Wars* (1977), one of the most memorable films in film history, and the amazing box office record was broken by another science fiction film Jurassic Park (1993). More and more films like *The Lord of the Rings* series (2001, 2002 and 2003), *Resident Evil* series (2002, 2004, 2007, 2010 and 2012), *Transformers* series (2007, 2009 and 2011). It is precisely because of the huge box office success of these blockbuster films that filmmakers were keen to shot science fiction films. Meanwhile, many kinds of screen formats, such as huge screen film, high-definition film, circular-screen film, thrill rides, water-screen film and etc, emerged to cater to the increasing demand of this theme. These new film forms enriched the consumer market. But just relying on the change of film theme was not a long-term solution.

Filmmakers thought about 3D film, which had been abandoned many years ago. If computers and software turned the visional imagination into real objects, then 3D film used the phantom reality to create imaginative stories. As mentioned above, 3D film is not a new conception in film history. In fact, 3D film, since it was invented in last century, had intermittently appeared for seven times. As everyone considered that this time, the eighth time of 3D film's return, would be another transient like last seven times, an incredible sci-fi epic film *Avatar* (2009) shocked the whole world. The filmmakers and theorists, who had predicted that 3D films would be short-lived again, now had to reassess its value. This film, after ten year's fermentation and preparation, finally gained a great success, not as a gimmick stunt, but a well-made blockbuster. The director James Cameron told reporters, when he shot *Titanic* (1997), if the 3D techniques was as mature as it was in *Avatar* (2009), he would definitely use it then. "Of course", he added, "I want all my films shot in 3D format" (Tencent Entertainment, 2009). For a short while, many film studios push their 3D films in succession. Some moribund film studios even regarded 3D film as a life-saving straw. Mass-producing 3D films became a global hot spot in film industry. As if overnight, a batch of quick-followed 3D films, produced in all kinds of quality, flooded into cinema. After the fever of enthusiasm about 3D film, the audience calmed down and began to

selectively watch 3D films. Some filmmakers and theorists thought that 3D film would decline again. But another blockbuster *Gravity* (2014), just appeared at the right time, ended the negative speculation of those, who was not optimistic about the future of 3D film. Therefore, after almost a century of its birth, 3D film is finally accepted as a promising film form by the public and film industry. With the amazing visual effects, 3D format was very suitably to be employed into many kinds of films. Science fiction film, horror film, action film, war film are all appropriate genre for 3D film. Film *Avatar* (2009), as a good example, the box office of 2D *Avatar* is 0.5 billion yuan (almost 0.1 billion dollar) in China, but the box office of 3D *Avatar* is over 1.4 billion yuan. Nobody could ignore the triple interests of box office (M1905, 2010).

But the good box office does not mean that *Avatar* (2009) is a perfect film. As mentioned in the beginning of my article, film *Avatar*'s failure of sound award made the filmmakers and sound specialists realized that if they would not improved the audio technology for 3D film, then soon, 3D film would copy the film history again, be forgotten after the transitory enthusiasm. After two years (around 2011), the approaches to realize three-dimensional surround sound came out successively.

2.4 The development process of ADSS technology

ADSS can be interpreted as an extreme real sound reproduction. It is not a specific brand or system, but a general conception. Like digital sound, we can use different digital format to produce sound. For audio companies, they can develop their own ADSS format, or employ other formats to set up an ADSS playback system, or just utilize the conception to produce film sound.

Since the filmmakers learned the lesson from *Avatar*'s fail in sound production, they realized the importance of audio-visual congruence. In fact, early in 2004, the first 3D audio system had appeared in Japan. The NHK (Nippon Hoso Kyokai) technician Kimio Hamasak put forward a concept of 22.2-channel system. The advantage of this

system was the sound could be located in the whole screen scope and the audience would immerse in the all-around sound field. This system was aimed at creating a real 3D sound and downward compatible system. The playback sound channels were divided into three levels as upper layer, middle layer and ground layer. There were 9 channels in the upper layer which were higher than the listening position and one was in the top center. 10 channels were placed in the middle layer, almost as high as the listening position. 3 channels were put under the film screen as the gound layer (Baidu Library, 2014).

The originator of 5.1 surround format, Tomlinson Holman wanted to use fewest sound channels to create a good auditory space. He joint 10.2-channel surround system with USC (University of Southern California), aimed at realizing a double effect as 5.1 format. This system had three features: one is to increase the number of front channel to achieve better sound localization. The second feature is to fill the vacancy between the two surround channels so as to improve the rear sound localization. The third one is to place loudspeakers in the up 45° position to enhance the vertical perception(Baidu Library, 2014).

Adding audio channel was no more than a continuation of conventional 5.1 surround format. So Fraunhoferidmt, a famous German research institution, and IOSONO company lunched IOSONO surround system, which was based on Wave Field Synthesis Principle. This system rebuilt a virtual sound field through spatial sound synthesis. Unlike the traditional stereo system, the locations of virtual sound sources would not change with the listeners' move. Theoretically, this system could effectively solve the problem of sound localization (Baidu Library, 2014).

Dolby, of course, joined in this development trend. Through separating the LRS (Left rear surround) and RRS (Right Rear Surround) from the original LS (Left Surround) and RS (Right Surround) of 5.1 format, Dolby developed an upgraded sound system Dolby Surround 7.1. This system was superior in its compatibility, thus helped the

cinemas and studios avoiding a large-scale upgrade of equipments. But for its essence, Dolby surround 7.1 was still a 2D surround format. This attempt showed clearly that adding the vertical sound was the most effective method to realize 3D surround sound (Baidu Library, 2014).

Barco learned from Dolby Surround 7.1 and avoided the repetition of 2D sound system, lunched its Auro-3D 9.1 surround format, which was based on 5.1 surround format. By adding four extra upper sound channels, Auro-3D 9.1 became the easiest surround system to be installed in cinemas (Baidu Library, 2014).

The thinking and developing about film sound showed that nowadays filmmakers increasingly emphasis the importance of film sound. The new audio technology certainly will trigger a new revolution in film industry and bring about deeper thinking and intense discussion about film sound aesthetics.

Chapter 3

ADSS

By imitating the human's visual and auditory perception, film creates a virtual space (film space or screen space). Filmmakers create in this space and provide aesthetic experience and artistic enjoyment to the audience. The artistic quality and aesthetic features of films are based on reproducing the real-life and processing the stories from real-life (Baidu Library, 2013). Actually, film space consists of two layers:

- Reproduction space

This space reproduces the real-life scenes. For the physical attribute, this space is a two-dimensional space with two directions: the width and the height. But the combination of vision and audition (perspective of images, vertical attribute of objects and the spatial perception of film sound) provides a three-dimensional hallucination for the audience (Baidu Library, 2009).

- Creative space (or narrative space)

Filmmakers utilize film montages to explain the background of the story, describe the characters' inner world and render emotions (especially film music). This space is formed by the audience's psychological effects. It is an assumed space, not a real existence (Baidu Library, 2009).

Human's perception about the real world is a four-dimensional world, with the concepts of space and time. Although humans' brain can process the information in a four-dimensional scope, only three dimensions are controllable for human being. Actually, the concept of space-time is relatively abstract, that based on Albert Einstein's "Relativity Theory". Since human cannot move faster than the speed of light, then time dimension is uncontrollable for the human being. Human can percept the time lapse, but cannot manipulate it as our wishes, like travel back or travel ahead of

time. The other three dimensions are controllable, as we can move up and down, front and back (Baidu Library, 2014).

The conception of dimension is a spatial measurement. “Dimension” means direction. The spatial pattern which is determined by one direction is a one-dimensional space, presents a linear form and be labeled in X-axis of coordinate axis. The typical example in nature is the plant. As the one-dimensional creature, the branches of plants grow in an extended mode (the volume of branches itself can be ignored). Two-dimensional space has two directions: length and width, presents a planarity form. In the coordinate axis it takes X-axis and Y-axis as the directional measurement. Ants are the most typical two-dimensional creature. An interesting experiment about ants’ perception is to lift the food which is in front of the ants. Turns out that the ants just circle around while go nowhere, because in ants’ world, there are only two directions: front-and-back and left-and-right. While the lifted food, in ants’ vision, just mysteriously disappeared. As an advanced creature, human’s world is a three-dimensional space (here, time, as an uncontrollable dimension is not in consideration) with three directions: length, width and height. Human’s whole perceived space presents a stereo state (Baidu Library, 2014).

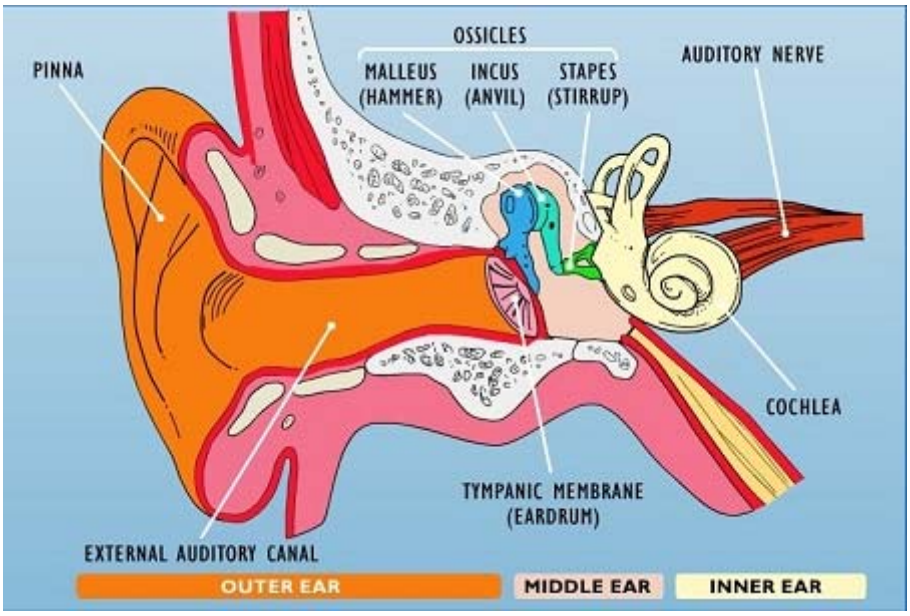
A feature of film is interacting with objects that move in space, particularly objects that move in depth towards the viewer. As mention in the openings, cameras imitate human’s vision and sound recorders imitate human’s audition, so the film should set up a similar space to draw the viewer into the created world. 3D film was just based on human’s perceptive mechanism to create a stereoscopic film world which is closer to real world than 2D film. In cinema, the distance between the film screen and the auditorium is usually more than 10 meters, so the film world of traditional 2D film is always 10 meters away from the audience. This stereoscopic sensation, which created by DOF (depth-of-field) in 2D film, cannot compare with 3D film. The space of 3D film breaks down the boundary of screen frame and extends outside the film screen. In term of film acceptance subjects, the perception that the objects appear to

leap out of the screen towards them creates an immersive reality.

By usefulness, technology should positively contribute to the experience, and in particular to the storytelling. Sound, as one of the indispensable film elements, also participates in establishing the film space. In order to carry out experiments proving the usefulness, it is necessary to have an appropriate 3D film and its corresponding 3D audio soundtrack. The duty of ADSS is to reproduce the real-life sound and create a natural and real acoustic field which can cooperate with 3D visual effect to establish a real sense of 3D film space. To accurately generate a dynamic and rich perception of the three-dimensional perception, the design of such complex sounds should be based on a firm auditory and scientific foundation. Firstly, we have to know human ears' working mechanism.

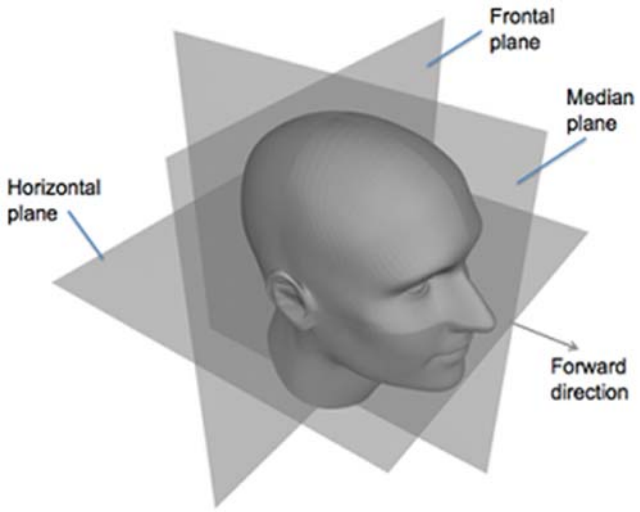
3.1 Human ear's working mechanism and auditory spatial perception

The human (peripheral) auditory system is made up of two fixed ears on each side of the head. They are the primary point of entrance of sound waves in our sensory system. Sound waves travel through the ear channel until they reach the eardrum. The eardrum passes the vibrations through the middle ear bones and onward into inner ear. Here, the vibrations are transformed into electrical impulses which are sent to the brain through hearing nerves. The brain then processes these impulses, tells us that we are hearing a sound and what that sound is (Baidu Library, 2014).



A diagram of the anatomy of the human ear

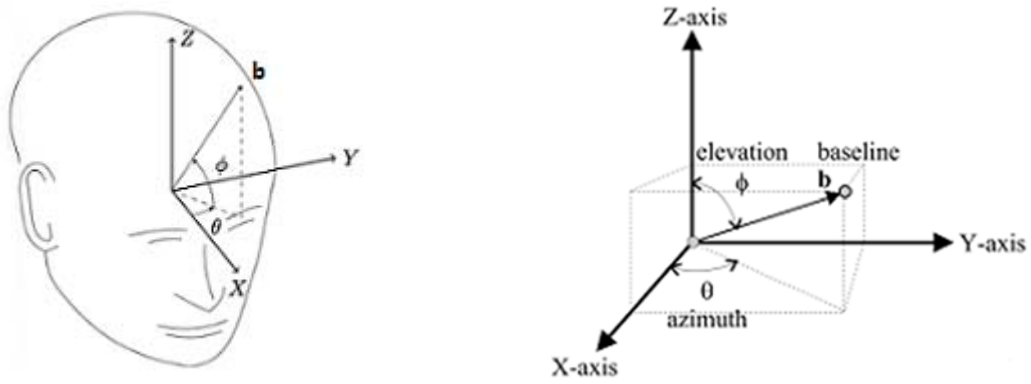
As for human's auditory perception experience, besides the basic physical attributes of sound: loudness, tone, timber, there are two more information that human ear can receive: the direction and distance of the acoustic source. This is human's spatial perception of sound. Actually, with only one ear, we can still hear sound, but can only identify its basic physical characteristics. Human, with two ears working together, are able to determine the spatial information of acoustic source. The location information of sound is extracted by the brain from the two sound-wave signals entering the auditory canals. The direction of sound source is determined solely from the wave pressures in the auditory canals. The figure blow shows the division of space around the head: horizontal, median and frontal planes. Humans can detect angular difference between sounds as close as 1° of azimuth apart on the horizontal plane (Baidu, Library, 2014).



The division of space around the head

We are only sensitive to a part of the spectrum of sound signal called the auditory range. For a young, healthy listener, provided sound reaches us with a sufficient intensity, this range is approximately 20Hz to 20kHz. We can hear sound coming from literally everywhere around us,. At 1kHz, the minimal air pressure that can be detected by the auditory system is 20μPa. This is the reference level for the dB SPL (Sound Pressure Level) units (Moray, 1969).

Because the mechanisms that determine the location of a sound are different in the horizontal plane and the vertical plane, it is customary to express the auditory space in spherical coordinates, with the center of the head at the origin, the X-axis pointing forward (for example, towards the nose), and the Z-axis pointing upward. The Y-axis is defined in such a way that the X, Y and Z-axes from a right-handed coordinate system. The figures below show the axes, the azimuth angle θ and the elevation angle Ø (b with an orientation (θ, Ø). The median plane is the vertical plane XZ passing through the nose, or the plane θ=0. When a sound comes from the side of head, the ear located on the same side of the head as the sound source is the ipsi-lateral ear, and the ear on the opposite side of the head is the contra-lateral ear (Moray, 1969).

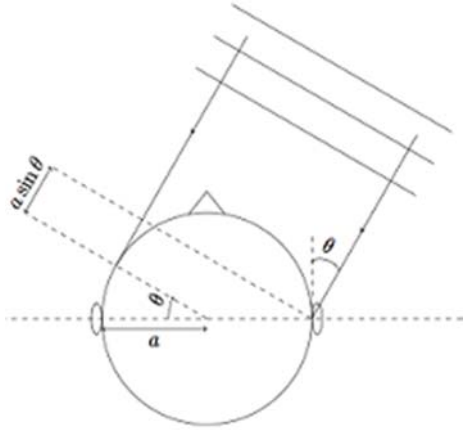


The three dimensioans of human's auditory perception

3.1.1 Sound localization in the horizontal plane and binaural effect

Because of the placements of our ears, a sound coming towards our head must travel further to reach the contra-lateral ear than it has to reach the ipsi-lateral ear. When sounds reach the ears from a particular location in space, the shape of head creates a set of important acoustic cues: Inter-aural Time Difference (ITD) and Inter-aural Intensity Difference (ILD) (Moray, 1969).

ITD: any sound that is closer to one side of the head than the others will reach the closer ear first. By the time it arrives at the farther ear there will be a time delay. Each location in space will create a different set of ITDs. The ITD is illustrated in below figure for the ideal case of a spherical head of radius α.



A (planar) sound wave impinges on an ideal spherical head from a distant source. An inter-aural time delay (ITD) is produced because it takes longer for the signal to reach

the contra-lateral ear (here the left ear). As a first approximation, the ITD is equal to α (sin θ + θ)/c.

Physically, the usefulness of the ITD resulting from simple signals, such as sinusoids, is limited to longer wavelengths, or, equivalently, to low frequencies. At a lower wavelength, when the wavelength becomes on the order of the head size, the sound wave diffracts around the head. In practice, the ITD is an important cue when the signal contains frequencies below around 1.5k Hz. Indeed, at frequencies above around 1.5k Hz, it is not possible to extract the azimuth θ from the phase difference between the sinusoidal signals at the left and right ears (Moray, 1969).

The auditory system can also find an ITD in a frequency rich signal. For example, McFadden and Pasanen distinguish between time delays at onset, the only cue for sounds shorter than about 1ms, and ongoing time delays. These ongoing delays are further divided into two categories: (1) the delays resulting from fine structure analysis of the signal, which correspond to the ITD of and (2) The delays found in the envelope of the signal. The ongoing inter-aural differences have been shown to be much more important than the onset difference for sufficiently long stimuli (longer than about 100ms). Still, the sensitivity to ITDs at high frequencies is lower than the sensitivity to ITDs at low frequencies (Moray, 1969).

When the wavelength is smaller than the diameter of the head, the head becomes an obstacle to the propagation of the wave and the pressure level at the contra-lateral ear is lower than that at the ipsi-lateral ear. The resulting inter-aural level difference (ILD) is an important cue for signals containing frequencies above around 1.5k Hz. The ILD cue becomes negligible at frequencies below 1k Hz (Baidu Library, 2013).

ILD occurs when sound are closer to one side of the head than the other and have a greater intensity compared with the opposite side of head. This auditory effect occurs because the head acts as an “acoustic shadow” that blocks sound waves from

reaching the side of the head that is farther. Although the sound waves reach the other side of head, their intensity is reduced.

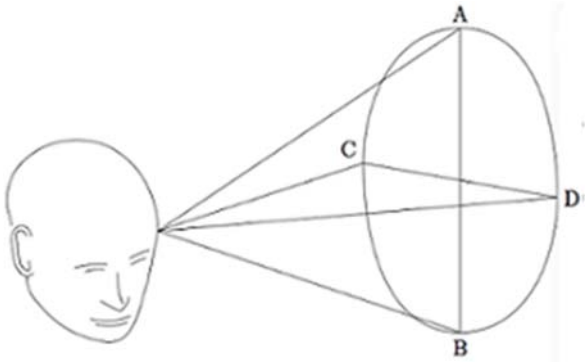
ITD and **ILD** are binaural cues, and together are referred to as the duplex theory of sound localization. For the physical structure of human's head, the ears are separately on either side of head. They can be regarded as two signal receivers with a constant distance. The auditory system calculates ITD and ILD of sounds and determines the location of acoustic sources. Since these acoustic cues come from the comparison of binaural signals' differences, it is very effective for locating acoustic source in the left-and-right positions (the horizontal plane). Despite their importance in horizontal localization, binaural cues are ambiguous for elevation in the median plane, where the ITD and ILD are constant or zero, and on the cone of confusion, the imaginary locus of source positions, centered on the inter-aural axis, which result in constant ITD and ILD cues. The cone of confusion results in front-and-back ambiguities and elevation ambiguities. A listener can easily resolve the front-and-back ambiguities by rotating his/her head. This requires that the stimulus be long enough, but the specialists already obtained a significant reduction in front-and-back errors with a 0.5s-long stimulus (Baidu Library, 2013).

The visible part of the war, or **pinna**, acts as a filter when it collects sound and directs it into the ear canal. Because of the cavities and the convolutions of the pinna, the amplitude of some frequencies is amplified while other frequencies are reduced, depending on the source position. Although this effect is mainly useful in elevation estimation, as we will see later on, it also plays a role in resolving front-and-back ambiguities and a marginal role in horizontal localization, when the stimulus contains high frequencies.

3.1.2 Sound localization in the median plane and pinna effect

The early experiments suggested that participants judged the direction of a one-third

octave noise signal in the median plane on the basis of its center frequency rather than of its physical direction (for example the loudspeaker). The spectrum of a sound signal is largely modified by the presence of the pinna, the head and the torso, which all cause absorption and reflections. These spectral cues are generally considered to be the most important ones to localization in the median plane. However, just as spectral cues secondary cues to horizontal localization, binaural cues serve as secondary cues to vertical localization (Baidu, 2013).



The cone of confusion has its apex at one ear and its axis is the line crossing the two ears. Its basis is circular. Any two sound sources diametrically opposed on a circular cross-section of the cone (such as A and B or C and D) produce the same binaural cues to the listener (Baidu, 2013).

Since localization in the median plane relies primarily on frequency effects, the presence or absence of certain frequency bands has an impact on the performance of the subjects in elevation localization tasks. Progressively reducing the bandwidth of a signal reduces the ability to perceive the elevation of sound and increases the chances of front-and-back confusion and it happens when the frequencies above 2k Hz were removed from the signal. Indeed, at low frequencies, the phase difference is the only remaining cue and this cue is ambiguous for elevation.

By evaluating these spectral characteristics, the brain is able to determine whether sounds arrive from different elevations (up and down), and also whether sounds that have the same angle are in front and back. The ear acts like a comb-filter (especially

for high frequency sound). For each location, the comb-filter lets in more energy at certain frequencies than others. The end result is that the auditory system can use a set of spectral cues to localize sounds in elevation (up and down) or front and back. Thereby, pinna effect solves the problem that binaural effect will confuse the acoustic source localization in front-back and the vertical plane. Actually up-and-down cues are located mainly in the 6-12k Hz band, and front-and-back cues in the 8-16k Hz band. These frequency effects are summarized by the Head Related Impulse Responses (HRIRs), or, in the frequency domain, by the Head Related Transfer Functions (HRTFs), which are the Fourier transforms of the HRIRs. HRTFs consist in a set of two functions, one for each ear, each representing the interaction of sound with the listener's head and torso for that ear. HRTFs depend on the position of the source and vary significantly from one subject to another (Baidu, 2013).

3.1.3 Auditory distance perception

If we limit the analysis to the far-field, then the curvature of the wave-front need not be taken into account. It is generally accepted that the perceived estimated distance D' to the source can be related to the actual distance D by the compressive power function:

$$D' = k D^{\alpha}$$ (k and α are parameters of the fit)

K is close to one ({k} = 1.32) and that α is consistently less than one ({α} = 0.54). This results in overestimation of close distances and underestimation of far distances. A variety of cues serve to the perception of the auditory distance to a sound source. A distinction must be made between a relative cue and an absolute cue. A relative cue allows the listener to compare the differences in distance between sources and the absolute cue allows him or her to give a distance measure for each source. Most often, familiarity with a sound source makes a relative cue become absolute for this particular sound source (Millar, 1994). **Auditory distance cues** including:

Intensity: the acoustic intensity of a sound source gives information on the distance to this source because the intensity decreases with increasing distance to the source.

The precise relationship between intensity and distance depends on properties of the environment and of the source. In the free field, an environment free from any reflective surface in any direction, the intensity of a point-source of fixed power decreases by 6 dB for every doubling of distance. The intensity is a relative cue. It requires familiarity with the source to become absolute.

Air absorption: at large distances, the properties of the air modify the sound spectrum, mainly by attenuating the high frequencies. It is a relative cue.

Motion parallax and variation of loudness with time: when both the observer and the sound source are in translation relative to one another, the motion parallax, that is, the angular movement of the source perceived by the listener, gives a small cue about the distance to the source. Also for a translation movement between the observer and the sound source, the ratio between the intensity and its time-derivative is linked to the time-to-contact, and assuming constant velocity, gives information about the distance to the source.

Direct-to-reverberant ratio: in a reverberant environment, which is contrary to the free field and contains reflective surfaces, part of the sound from a source reaches the listener directly (the primary source), and the rest arrives after interacting with reflecting surfaces (the secondary sources). The ratio of the direct part and the reflected part gives information on the distance to the source. As a first approximation, the (late) reverberant part can be approximated by a diffuse field of constant energy with respect to the position of the secondary source. Therefore, a close source induces a large direct-to-reverberant ratio. Since this cue varies from room to room, it is generally a relative cue. When the room is familiar, the direct-to-reverberation ratio becomes an absolute cue.

3.1.4 Non-acoustic cues

Our perception of the surrounding world is by essence multi-modal, that is, we integrate information across our senses to build an accurate and reliable percept of the world. This holds true of course for our ability to localize sound sources in space. Powerful cues about the location of a sound source are derived from non-acoustic sources of information (Millar, 1994).

Vision Auditory localization can strongly influenced by the presence of a potential visual target. The impact on azimuth evaluation is called ventriloquism, that is, a perceptual phenomenon, where a visual source position captures the position of a discrepant sound source. A visual stimulus can also influence the evaluation of the auditory distance. The addition of a visual stimulus improved the distance judgment accuracy and lowered judgment variability compared to the auditory-only stimulus.

Prior knowledge and familiarity The audience presented with an array of loudspeakers at different distances had to judge which of them was playing an unfamiliar stimulus. At first, the audience were unable to correctly identify the source and consistently associated the source to the closest loudspeaker, but gradually improved their judgment at each trial. This indicates that increasing familiarity can lead to better localization accuracy. Familiarity also turns a relative cue into a absolute one.

Expectation Speech signals are familiar to all the audience. These signals have particular features that one uses, for example, to distinguish between shouting and whispering. In anechoic conditions, the distance to a source of whispered live speech was underestimated while the distance to its shouted equivalent was overestimated. The audience associated whispering with being close to the interlocutor, and shouting with being away from him or her.

3.1.5 Spatial perception

Philosophical discussions of the nature of spatial perception are endless. Attempts to define the terms have been recorded over a span of at least 2000 years, making it one of the longest-running debates in history. This term itself is composed of two elements, which themselves are abstract nouns. The dangers of reifying abstract nouns are well-know, yet the idea of an underlying reality consisting of space which we may objectively measure is so attractive that we may be forgive for assuming that the term space is synonymous with that which can be measured in units. According to this method, perception mirrors reality, and is thus presumed to be similarly measurable in inviolable units. ADSS is just based on this concept and the main aspect is the reproduction of sound field.

3.2 Reproduction of sound field

As mentioned above, the main aim of ADSS is to reproduce the real-life sound and create a natural acoustic field. I use a long length to explain the complex working principle of human ears, then I will utilize these theories to reproduce the auditory environment.

The most direct method to reproduce the auditory environment is to imitate the natural acoustic environment or similar sound field. Thereby, the listener will obtain the same subjective perception as in natural sound field. Another method is to imitate human's bilateral signals by processing sound materials. The former technology aims at the sound production, while the latter one aims at adjusting listeners' objective feeling. These two different targets lead to three methods to rebuild sound field, which will be stated below.

3.2.1 Re-establish sound field by WFS

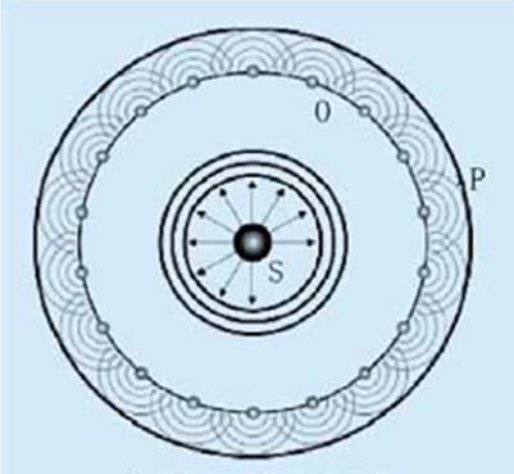
Theoretically, only countless soundtracks and loudspeakers can truly reproduce the sound field and sound localizations of the actual site. But in practical applications, it is

impossible to realize a “countless” reproduction. So from seventieth of last century, sound technicians began to develop surround-sound system: from 4 sound channels to 11 sound channels. All these surround-sound systems have a same feature: the loudspeakers are all located in the same horizontal plane. In other words, they can only express sound localizations in two dimensions: front and back, left and right.

As we know, in physical theory, sound propagates by sound waves through a medium. So sound is a conception in a three-dimensional space. In natural environment, there are three most basic propagations of sound waves: spherical wave, plane wave and cylindrical wave. A certain sound source (for example, a man in a room) propagates as spherical waves. If the propagation distance is not far, then the receiver can find the specific sound source. But if the propagation distance is far away, then sound propagation will present a form of plane waves. Then for the listener in the same sound field, he can only distinguish the direction of sound, but not the specific location. As for cylindrical waves, they are usually formed by the fast-moving objects, like the high-speed rail, racing cars or rockets. So either for the conventional stereo systems or the later surround systems, they can only reproduce the sounds that propagate as spherical waves. But for the speed and distance, they are incapable of action.

WFS technology solves this problem. Wave field is the distribution of sound waves in space. We can use the computer to calculate wave field in a certain space and use loudspeaker array to synthesis and express.

The theoretical basis of **WFS (Wave Field Synthesis)** is **Huygens-Fresnel principle**. Each punctate acoustic source of an advancing wave-front is in fact the center of a fresh disturbance and the source of a new train of waves; additionally, the advancing wave as a whole may be regarded as the sum of all secondary waves arising from punctate acoustic sources in the medium already traversed (Berkhout A.J., 1988).

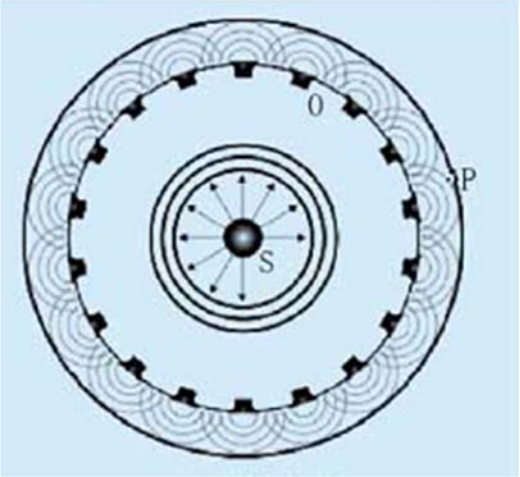


Huygens-Fresnel principle

In the figure above, the spherical wave is generated by a punctate acoustic source S, forms wave-front at P. This wave-front can be regarded as the sum of all spherical waves arising from all punctate acoustic sources on the surface O. Though the acoustic features of wave-front on surface O, we can calculate the acoustic features of wave-front at P.

To put this principle simply: if two rooms are connected by an open doorway and a sound is produced in a remote corner of one room, a listener in the other room will perceives this sound as if it originated at the doorway. As far as the second room is concerned, the vibrating air in the doorway is the source of the sound.

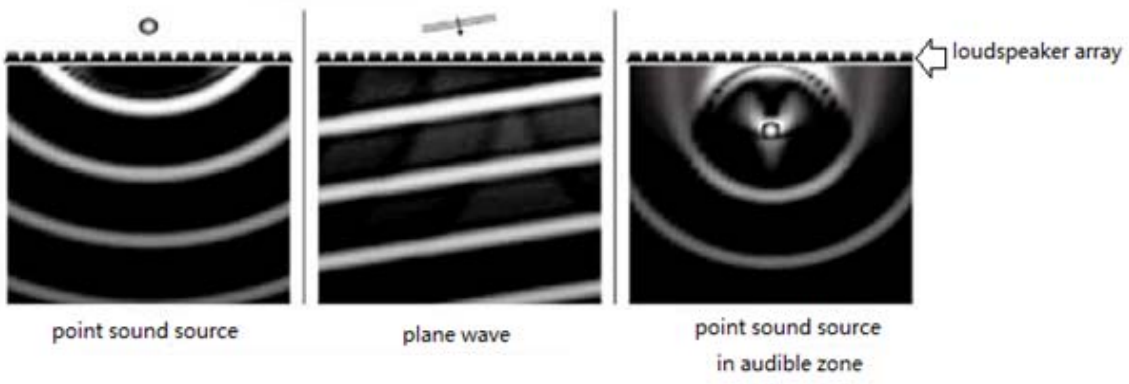
WFS: in the figure below, all the punctate acoustic sources on surface O are replaced by loudspeakers, these loudspeakers also send out spherical waves, so the wave-front at P can be synthesized by sound signals which are sent out by loudspeakers (Start, 1997).



Wave Field Synthesis

WFS, as a powerful theory for sound engineers, is a very advanced technology to reproduce or re-establish sound field. The most prominent advantage is to produce stable and real virtual sound source localization. It can produce three different types of virtual sound sources:

- Virtual punctate acoustic source behind the array of loudspeakers
- Virtual sound source propagates plane waves
- Virtual punctate acoustic source before the array of loudspeakers

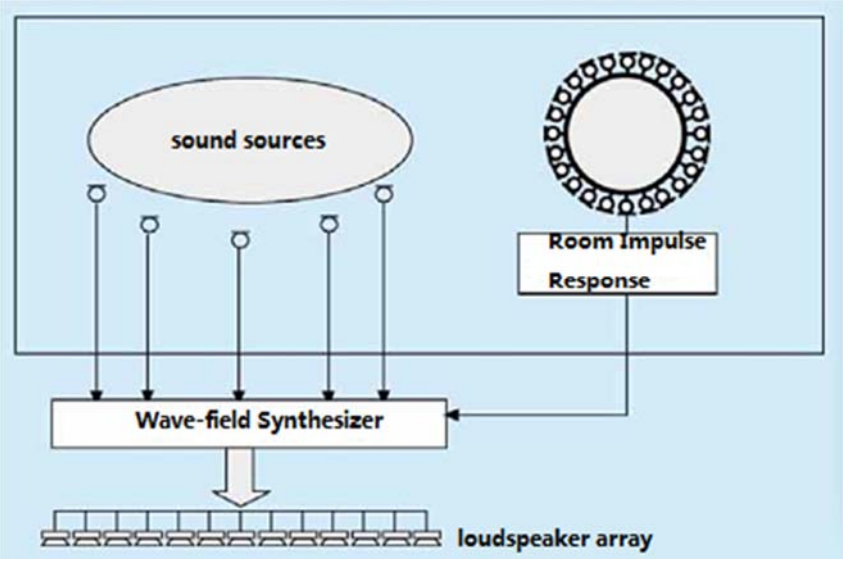


The wave-front of three different virtual sound sources

To compare with the traditional sound reproduction systems, with WFS, the listeners are no longer restricted by “sweet spot”. When the listener walks around in the audible zone, the localization of punctate acoustic source will change synchronously and the

localization of plane waves will also in the same direction. This shows that WFS can provide stable sound localizations. Moreover, when the relative position of the listener or virtual sound source has changed, the listener's auditory perception will correspondingly change with it, similarly to "Parallax Effect". WFS is able to provide authentic sound localizations.

The concrete operational process of WFS is: in sound producing process, we separate the signals of sound sources from its spatial information. The pure sound signals (without any information about spatial information) and its spatial information (the locations of sound sources, the detailed acoustic characteristics of recording studios or natural environment, etc.) of sound sources will be respectively sent to wave-field synthesizer. In wave-field synthesizer, the pure sound signals of sound sources will convolve with room impulse response and sound sources will regain the spatial information. Thereby, we can synthesize the wave field at any time and in any position.



The process of re-establish sound field by WFS

To put it simply, this technology can be regarded as: to separate a sound wave into many sound wavebands and send these information to different loudspeakers to

synthesize and express. It means the more loudspeakers a cinema has and the more compact loudspeaker array is, the more precise the sound localization is.

The precise localization of sound sources in three-dimensional space is the most important feature in ADSS. In traditional sound systems, such as stereo systems and surround systems, all the sounds come from the screen or theater walls. In another words, there is no change of the sound sources locations. That is why there is a “sweet spot” in each cinema. By employing WFS in sound reproduction, sounds can appear in anywhere in the cinema. Sound designers can even locate a sound between two listeners. One listener will feel the sound appears on the left and another will consider it is on the right side.

3.2.2 Rebuild sound field by Ambisonic technology

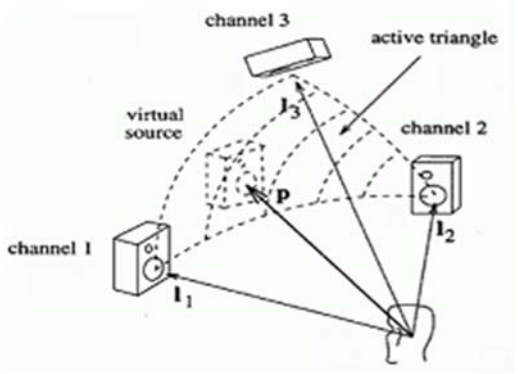
Ambisonic technology is base on spherical harmonics principle. With spherical harmonics principle, we can disassemble and recombine the target sound pressure fields. It is a process of encoding and decoding the spatial parameters of sound fields for multi-channel system. By using spherical harmonics principle, the encoding process is to do spatial sampling for the target sound field and get the Ambisonic parameters of this sound field. The decoding process is: according to the numbers and locations of loudspeakers, we can process the Ambisonic parameters and send them to the corresponding speakers. The encoding process and decoding process are independent (Craven, 2009).

There are two methods to process Ambisonic parameters: one is matrix inversion, another one is vector resultant, based on psychological acoustics: virtual sound synthesis. The precision of Ambisonic technology depends on the installation requirements. The more loudspeakers are used in playback, the more real audiences will percept. Ambisonic technology cannot totally eliminate the “sweet spot”. If the playback system is in a higher standard, the scope of “sweet spot” will be larger.

3.2.3 Regenerate sound field by VBAP (Vector Base Amplitude Panning)

Amplitude panning is a technique in which the same audio signal is reproduced through two or more loudspeakers, with appropriate level differences, so that a phantom virtual source is perceived by the listener in a position between the loudspeakers. VBAP is the result when this technique extended to 3D loudspeakers layouts (Sheng, 2012).

The working principle of VBAP is: we can divide the space where sound sources are located into many triangular spherical spaces. Each space is consisted by three loudspeakers.



Triangular spherical space

Through adjusting the output signals of the three loudspeakers, we can locate the virtual sound source in one vector unit. Then we can regenerate sound field and locate the virtual sound source in anywhere of a specified sound field.

3.2.4 Combination methods

All these techniques aim to reproduce a natural auditory perception. But they all have advantages and disadvantages. WFS can reproduce the naturally spatial perception at the most extent. With it we can make sound appears in anywhere of the cinema space. For example, when we see a plane flies over on the film screen, we can feel

the plane moving along the flight path and even feel the “flows” of the sound. But due to the working principle of WFS, we have to place a great many loudspeakers to achieve a better sound quality. In comparison, Ambisonics technique will use relatively less loudspeakers to reproduce the reverberation and ambient sounds. But the quality of the precisely focused sound sources is barely satisfactory. VBAP works really well for reproducing those sound sources which are in the precise positions, but it is not good at the broad, enveloping sound. Considering about these merit and demerit, some of these technologies can be combined together and integrated into a new approach. Actually, by experiment research, it is relatively easy to combine Ambisonics and VBAP for sound technicians. Another hybrid method is using Ambisonics technique to process the playback signals of existing stereo and 5.1 surround sound system (Laborie et al, 2003).

3.3 The artistic principle of film sound design

The aim of advanced technologies is to serve artistic creation. Before bring in the conception of ADSS, It is necessary to know the principles of film sound design.

Film creation relies on imitating human’s perceptual experience. So the design principles of film sound should firstly agree with human’s auditory experience. In fact, any sound that has been processed by sound equipments is “distorted”. The “reality” which we want to reproduce in the sound production does not refer to objectively copy the real-life sounds. This pursuit of “reality” is based on the audience’s subjectively aesthetic requirements and sound designers’ objectively artistic choice.

Aristotle wrote in his work *Poetic*:

> “*It should be known that the imitation that consists of fables and [historical] narratives is not poetry at all; rather, poetry deals with what is probable or exists by necessity. That would be the case if meter were the only difference between fables and imitations; it is not so, for the speech must*

deal with a thing as it did or did not exist." (Aristotle, 1908)

In this passage above, Aristotle described the difference between poetry and history. By imitating his description form, we can describe the aesthetic design principle of film sound as: the purpose of film sound design is not to reproduce the sounds that already existed, but to consider and design the sounds that may appear or should appear. Thereby, all the film sounds should be designed for the certain purposes and appear as "should appear" or "are expected".

These two kinds of "designed" sounds always appeared in science fiction films. A lot of film scenes in science fiction films are in outer space and besides film dialogues, most of the sound effects are fighting sounds, weapon sounds or sometimes, aliens' strange sounds. But if we seriously think about these "designed" sounds, as for the common sense, we know that sound cannot propagate in the vacuum space. That is to say, in a totally vacuum space, the audience cannot hear any natural sound in outer space. Then what is the basis of these sounds be designed in science fiction films? This kind of sound design which does not conform to the reality comes from a creative sound design called "artistic distortion". Sound designers overturn the audience's aesthetic experience of the real world and create these "consciously designed" sounds which are actually distorted in real life to complete the requirements of film expression. From the early film *Star War* (1977) to nowadays 3D film *Avatar* (2009), these "designed" sounds enriched the film expression and formed a special sound production mode of science fiction films which attracted the audience's attentions for a long time. The reason that audiences never questioned about these distorted sound is because the primary principle of sound design is not to copy the objective reality. It should always agree with the audience's aesthetic experience and emotional imagination. Similarly, some inexistent sound effects in real world appearing naturally in horror films are also consciously designed.

Sound appearing in films, like the invention of film, reflects a historical course. As the

incredibly fast development of technology, technology-based films will be an inevitable trend in the future. As for film sound production, the developments are directly up to the improvements of sound technologies. But the audience's aesthetic experience of film sound is a composite psychological experience with both subjective emotion and objective expression. In this cognition state, film sound production is not just a technological concept for the audience, but an affective enjoyment of aesthetic appreciation.

Film sound, after all, has to serve for the film reality. This film reality derives from an "immersive" film-viewing model. This film-viewing model is more or less like dreaming. When human are in the dreaming state, his mind will leave the realistic world, the uncontrollable emotion will lead he entering into another reality, dream reality. Likewise, when we are in the dark cinema environment, we are cut off from the outside world and immerse ourselves into the film space. Our bodies are relaxed in the confined dark space, which is an advantage for an immersive atmosphere as a wonderful dream state. Digital technologies effectively enhanced this "immersive" perception. Since 3D visual effects have greatly impacted on the audience's visual perception, accordingly, the concept of three-dimensional surround sound comes out within expectation.

3.4 The conception of ADSS

We all know a general physical knowledge: Sound is generated by the vibration of objects and propagates in the shape of sound waves. The nature of sound depends on the frequency and amplitude of sound waves. As for two different sonic phases propagate together, sound waves can be enhanced or counteracted (Baidu Library, 2014).

So if the film sound is produced in mono format, the two sinusoidal waves, with different wavelengths, will superpose during the playback process. Conversely, the

two corresponding sounds may be distorted by hybridizing or offset. As for film sound, if the two kinds of sounds are dialogues and sound effects, then some certain frequencies of the two sounds, with opposite phases, will be offset. It is a disaster for the film narration. From mono sound to stereo sound, the reform was appealed by the audience's demand for a better quality film sound.

Our brains perceive pictures very consciously, whereas sound is rather unconsciously. Therefore, sound is sometimes considered to be less important in the film production (Baidu Library, 2013). As the artistic theory improving constantly, more and more filmmakers realize that sound can create at least as much emotional response as images do. Since film sound is not totally on-the-spot recorded, a mediated acoustic reality, the artistry and creativity become more and more important in sound production. From analog sound to digital sound, the progress was largely pushed by the filmmakers' creates needs.

3.4.1 Current film sound formats

5.1-surround format firstly appeared in 1992 and was invented for the 35mm film. With this format, sound designers could arrange sounds in a certain principle: dialogue was in the center channel, music (or sound effect) was in L/R channel and the sound effect (or music) can be arranged in SL/SR channel (surround left/surround right). With this principle, film sound will be playback in different channels. The audio problems (hybridized or offset) had been solved. Moreover, 5.1 surround is better at sound localization in horizontal plane than the former sound systems. With this technology, we could make a better quality film sound in the cinema. Although 5.1 surround did improve the quality of film sound, the missing dimension was still a big defect of audio-visual congruence. The traditional 2D surround, without the vertical plane, had two disadvantages: inaccurate sound localization and incoherent auditory perception. Explaining with an example that a plane flies from left to right, in 5.1 surround auditory environment, the sound we heard just jumps from the left side to the right. The

situation is better in 7.1 surround auditory environment. But still the sound is incoherent, just jumps from left, rear left, rear right and then to right side. The sound jumps among the loudspeakers but does not “flow” coherently.

Comparing with the frames of 2D film, the most prominent improvement of 3D frames is to produce the vertical-perceptive vision. The realization of this vertical perception is based on the audience’s physiological-visual hallucination and the film projection techniques. By creating a 3D visual phantom, 3D visual effect provided more realistic film frames. For the film production, if there is any change of the film frame, there should be a corresponding change of the film sound. Although 5.1 surround system has a lot of advantages and been world-widely used in most of cinemas, it is still a horizontal sound-reproduction mode. Thereby, adding 3D sound with spatial perception to 3D film became the film industry’s urgent demand. Thanks to the advanced innovative technologies, the process of implementing 3D sound technologies is possible with the help of off-the-shelf processors and DSPs (digital signal processor) at a relatively low cost.

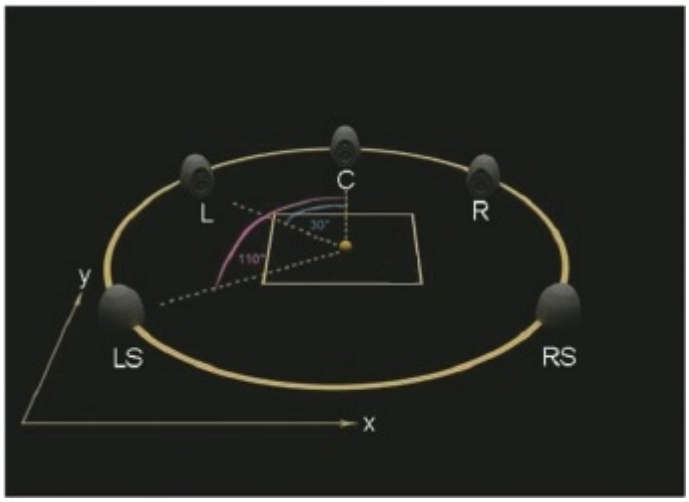
In last decade, many new sound formats have been proposed for reproducing surround sound in the theatrical environments, such as the 10.2 surround format or 22.2 surround format. Most of these attempts existed only in the research labs or a limited extent practice. Although the principle of reproduction surround sound in the horizontal plane had been set up for almost forty years, it is just a few years since this pattern evolution of film sound production had been pushed to improve by the great pressure from development of 3D visual effects.

3.4.2 The conception of ADSS

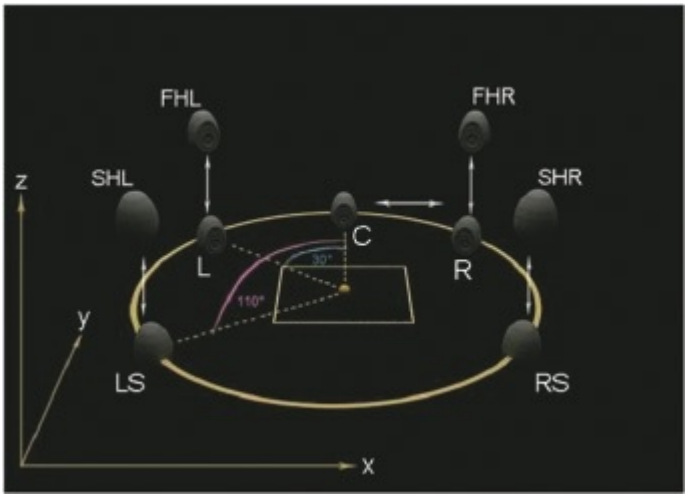
As mentioned many times, by imitating human’s perceptive experience, filmmakers set up a virtual film world. So the auditory conception of ADSS comes from the simulation of our auditory system. In our normal life, our ears are totally exposed in

the three-dimensional sound field. Our auditory system can receive three-dimensional sound information, either the reverberation of a room, the noisy background of the city or the natural ambient sounds. With two ears, we are able to identify the location of sound sources around us and to distinguish sounds coming from left or right, front or back as well as top or bottom. So ADSS format, as a faithful sound reproduction technology, could provide a three-dimensional perception.

Since the conventional multi-channel sound formats, like 5.1surround, can only locate sound in the two-dimensional space: front and back, left and right. It is possible to extend it by adding a third dimension (the vertical plane) to form a new sound format, ADSS (All-dimensional surround sound). The figures below showed the differences between the traditional 2D sound format and 3D sound format in space.



5.1 surround sound system (two-dimensional space)



ADSS system (three-dimension space)

3.4.3 The characteristics of ADSS

We can obviously see from the graphics above, the sound reproduction of the conventional multi-channel surround formats (using 5.1 surround as an example) establishes on the horizontal plane with two spatial axes (X= width, Y= depth). ADSS establishes on the three-dimensional space with three spatial axes (X= width, Y= depth, Z= height).

Actually, focusing on height is not a new attempt. A lot of sound engineers have had some interesting achievements with their “voice of god” installations. The superiority is that we are able to create an acoustic perception that is as closer to as a natural listening experience than ever before. The result is not just the direction and amplitude of a sound event but the relationship between sound events. Of course, the added the Z-axis is not just simply adding height speakers or rearrange the localizations of speakers. It is an enhancement on how the sound is processed for output and provided an optimized ambient sound environment for listeners which is remarkably like the real world. With the supplementary sound information in height, listeners will feel like they are the main observer in any scene. In summary, the advantages of ADSS are as blow:

- Omni-directional localization of sound images

In the three-dimensional virtual film space, the actual sound signals will be localized on the specific locations as if they come from virtual sound sources. So that the audiences can not only identify the sources information by vision, but can also rely on the audition. The auditory system leads the audiences' visual system to search for the sound sources, and then the vision will precisely find the localizations of sound signals. This advantage makes the film sound more similar to the real-life sound.

- Real-time tracking

ADSS can track the changes of virtual sound sources. This real-time tracking includes two aspects: one is that the film sound should coincide to frames. This consistency is conducive to the synchronization of vision and auditory. For example, if lens becomes bigger, then the sound volume should decrease. If the object is moving in a certain direction in the screen, then the sound should also change according to the moving. Another aspect is that if the positions of the audience's heads have changed, then their perceptions about the sound sources locations will also change. This change is consistent to the human's auditory perception.

- The interactivity between the film sound and the audience

This interactivity is mainly reflected in the "immersion" and "participation". The audience can totally immerse themselves in the film space and their reaction will also impact on the film sound. They are no longer just the spectators.

3.5 Current practical applications of ADSS

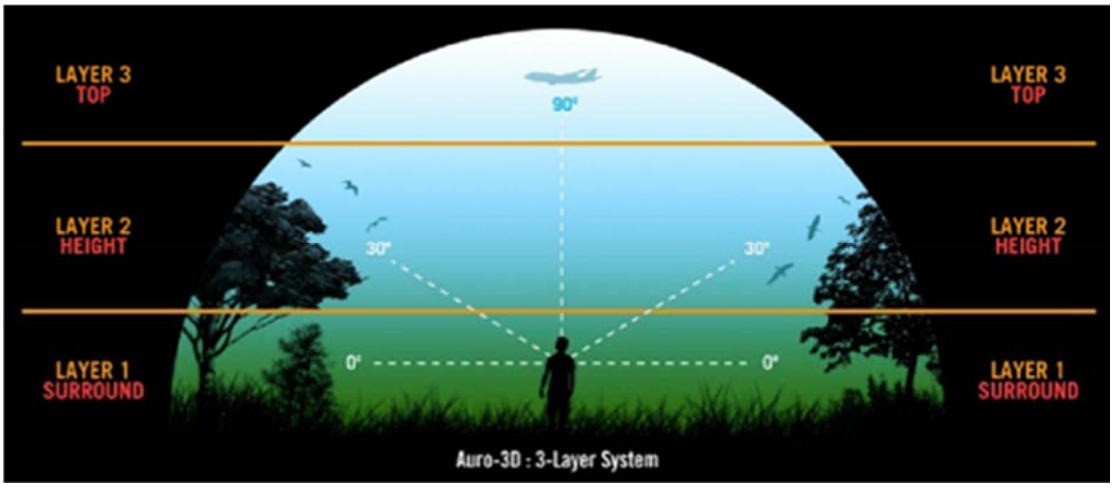
Since the conception of ADSS was raised, it had attracted a world-wide attention. From Europe, America to Japan, many research institutions and commercial corporations aimed at the research and development of ADSS. The development scope covered from digital computation, sound processing to spatial information.

2012, Dolby company released a brand-new solution for 3D surround: Dolby Atmos, in Las Vegas CinemaCon. Besides this, there were many other three-dimensional surround technologies, such as Auro-3D by Barco, Imm Sound 23.1 by Barcelona Media, IOSONO 3D by Germany Fraunhofer and Illusonic 3D by Weiss Medus, etc. Since 2012, Dolby Atoms, Auro-3D 11.1 and Imm sound had been generalized in cinema upgrading and adopted into film production. Films with three-dimensional surround were released successively. IOSONO 3D and Illusonic 3D were relatively quiet. In September 2012, Dolby bought out Imm Sound technology and integrated it into its Dolby Atmos. Until 2013, two ADSS modes: Dolby Atoms and Auro-3D became the most competitive technologies in global film market.

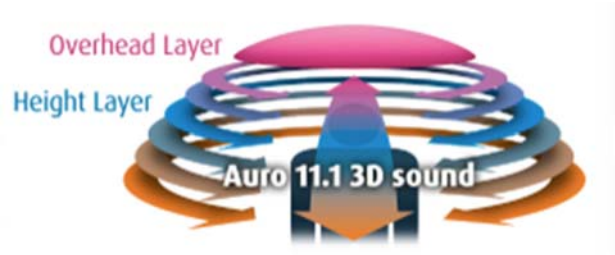
All these new-generation surround technologies are based on the conception of ADSS and aim to provide an immersive auditory experience for the audience. But the technical means and design philosophy of these technologies are widely different. So the practical implications in cinemas are corresponding different. Taking Auro-3D and Dolby Atmos as examples, we can have a further understanding about the practical application of ADSS.

3.5.1 Auro-3D (11.1 or 13.1): channel-based system

Auro-3D, announced by Barco, is a channel-based surround system. For the design pattern, Auro-3D adds two additional layers of loudspeakers (height layer and overhead layer) to the current 5.1 surround: a 5.0 surround system composed of the height speakers and the ceiling overhead channels. It can be regarded as an extended system of 5.1 surround. In term of layout design, this loudspeaker arrangement has a high compatibility. It can totally be compatible with the current 5.1 surround system.



3-layer loudspeaker arrangement



Sound-field diagram of Auro-3D 11.1

In term of function, the height channels and the additional forward channels expand the sound field. The sounds we heard in everyday life originate from the superposition of direct sounds at ear-level and spatial reflected sounds. A large part of spatial reflected sounds come from the overhead space, especially in a hall environment, because the reflected sounds will combine together and form bright reverberation in the high space of a hall. According to this theory, increasing upper reflected sounds is very necessary for the “immersive” perception. In cinema practice, a circle of upper loudspeakers points to the auditorium in order to enhance the spatial experience. This layout is very appropriate for reproducing a wide sound-field with different sound layers, especially the natural environment such like forest, theatre, square and etc. With the ceiling overhead channels, sound engineers can reproduce a clear sound localization and express a smooth and continuous motion trail. For example, to depict fly-over effects, unlike the chaotic effects produced by former surround systems, the

sound information from the ceiling channels can precisely convey the real-time locations and motion trail of the sound sources (Auro-3D,2012).

The cinema practice of Auro-3D

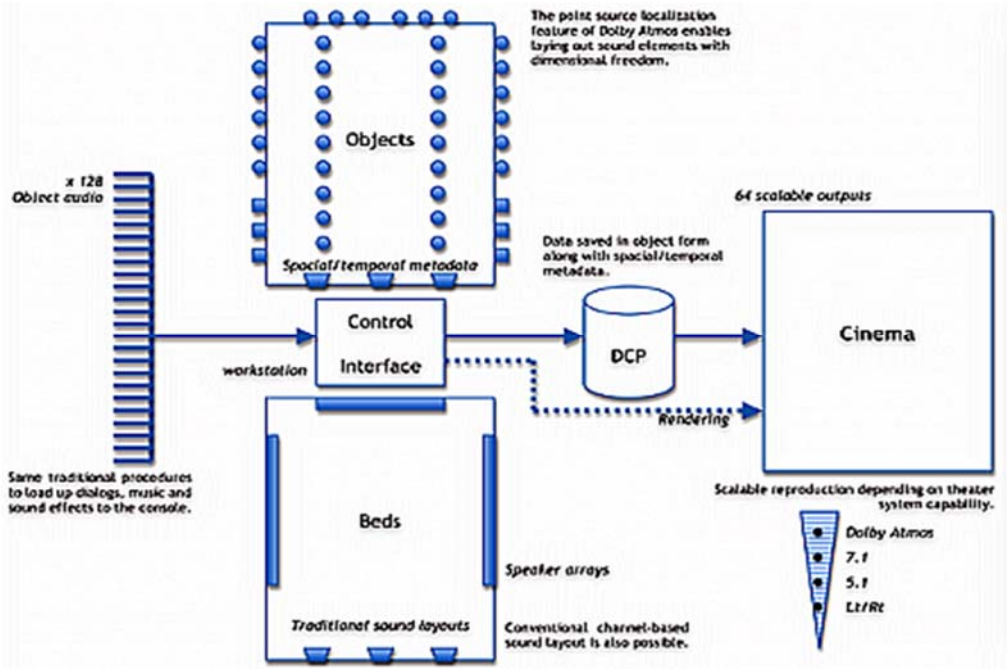
3.5.2 Dolby Atmos: hybrid (channel-and-object-based) system

Comparing with continuing the concept of channels, Dolby Atmos stepped further. 2012, Dolby Laboratories introduced a new concept of “object-based”: Audio objects (either static or mobile) can be considered as groups of sound elements that share the same physical location in the cinema space. Each object has a metadata to describe its position in time. In the practice, a soundtrack can be seen as an audio object and standard panning is similar to metadata. This concept of “object-based” arises from sound engineers and directors’ desire for more precise sound locations rather than their corresponding channels.

Considering that some film sounds, especially many ambient effects or reverberations, work better in channel-based mode, Dolby Atmos retains channel-based function as “bed” in addition to objects. Beds can be created in different channel-based configurations (such like 5.1, 7.1, etc) and mapped to individual loudspeakers or loudspeaker arrays. Objects are rendered through loudspeakers to cover the spherical space.

Instead of having fixed film sound, the process of sound reproduction becomes a 3D map of audio objects moving through a certain space: in the cinema space, all the loudspeakers will be driven independently and then located in an accurate-calculated space model. Through calculating the changes of parameters (beds), the sounds are rendered at the point of delivery (the cinema) and optimized for the particular speaker array and speaker density at the venue, thus realize that the sounds freely move around the space.

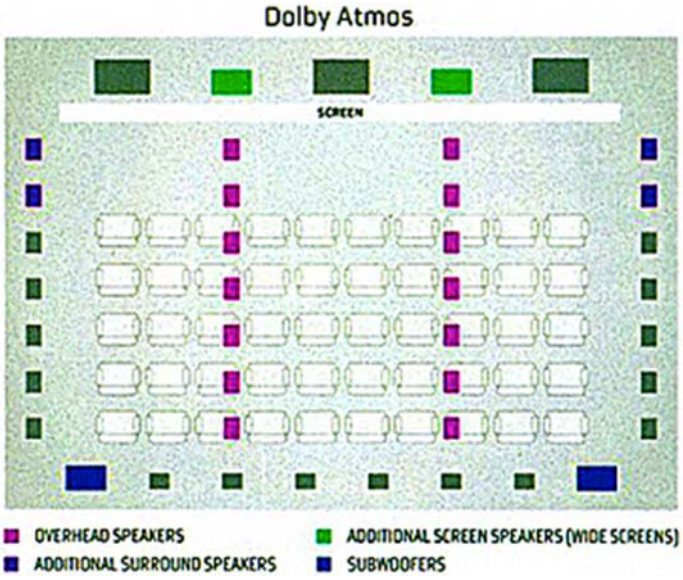
To put it simply, Dolby Atmos can be regarded as a unique hierarchical superposition of sound information: 7.1 surround format + “beds” based on physical channel + objects with metadata data.



Dolby Atmos System Flow

For loudspeaker layout, Dolby Atmos can support up to 64 loudspeaker outputs. As described the design principle above, Dolby Atmos system is very flexible to be extended on the current surround systems (5.1, 7.1 or even 9.1) and the ideal number

of loudspeaker output channels used will vary according to the cinema size. Using 7.1 surround as an example, Dolby Atmos system adds surround loudspeakers (side surround and rear surround), top surround loudspeakers and screen loudspeakers (if the screen is greater than 12m wide) (Dolby, 20.



Dolby Atmos: Loudspeaker locations

3.5.3 IOSONO 3D (holographic sound): object-based system

In the context above, we known the concept of “object” is a technology aimed at processing individual audio object. Dolby Atmos is advanced in its concept of audio objects processing and spatial rendering through loudspeakers. IOSONO deepened the concept of “object” and unveiled a more futuristic approach to realize the sound localizations in three-dimensional space. The sound sources freely move around a certain space on a height, width and depth plane, not just around the audience, but inside or outside the space and overhead in any direction. This unique concept removes the constraint of loudspeaker layout and sweet spot and very audience is able to appreciate his own sonic sphere.

This new surround format is based on the WFS (wave field synthesis) principle. The

WFS principle, which is originated from the Huygens Principle, is a technique to reproduce sound fields corresponding to the sound-wave propagation of real sound sources. The process of sound reproduction can be comprehended like this: split the original sound sources into sound signals and its spatial information. The pure sound signals (without any information about spatial information) and its spatial information (the locations of sound sources, the detailed acoustic characteristics of recording studios or natural environment, etc.) will be respectively sent to Spatial Audio Processor (wave-field synthesizer). The pure sound signals will convolve with the impulse response of a certain space so that sound signals will regain the spatial information and playback by loudspeakers. This production process brings immersive, truly 3D positioning and animation of film sound sources around the audience, within or outside of listening area.

By combing the WFS technology with concept of “object”, IOSONO 3D provides a significantly flexible alternative to any conventional multi-channel format (5.1, 7.1, etc.) and supports scalable loudspeaker types and arrangements. No matter whether there are 5 or 500 loudspeakers, the sound reproduction is optimized for each loudspeaker to create the ideal 3D sound experience (IOSONO 3D, 2012).

IOSONO 3D: Loudspeaker locations in cinema

3.6 The value of ADSS for 3D film

Like sound appeared in the film, 3D sound is not a sudden even that just happened in a particular time or a certain day. It always reflects the development trend of the technology-based film in t film history.

Human society progresses every day, along with the changes of people's lifestyle, emotion and aesthetic standards. Social development is always effected by the science and technology inventions. The aim of technological innovations is always to promote economic development. Since the superstructure is determined by the economic base (360 Beta, 2012), film, as a superstructure form, is inevitably constrained by the economic base. Film is an art form, which combines technological processes and artistic expressions. The expression method of film has to rely on technology. As the seventh art form, film imitates life but higher than life. During the process of recording human's social life, film can also reflect how science and technology influence on human's social life. Its final value has to be embodies as a commercial form.

The development of science and technology afforded a lot of benefits for film production. In term of film frames, it went through black-and-white film, Technicolor and develops into stereoscopic film. In term of recording medium, the traditional cine film has been displaced by digital media. In term of film sound, from silent film to phonofilm, from mono-channel film to multi-channel film, from the initial pursuance of sound articulation to today's requirement of deep artistic expression, film sound started as a technological invention and developed into an artistic creation.

At the end of last century, digital technology swept across the world as an overwhelming momentum. It impacted almost all domains of human society. The applications of digital technology had solved many problems in film production and meanwhile significantly reduce the production cost. The artists' unconstrained

imagination is never confined by logical thought. Digital technology thoroughly emancipated the talented directors and filmmakers to express their artistic emotions as their wills.

Film is aimed at processing human’s social life and recreating a similar film world. Life is always alive with all kinds of sounds. Human’s thought is mainly performed by words. In Bible, there is a doctrine that words rule over everything. For the music, the Pythagorean School considered that it could purify human’s soul. If the film doesn't contain sound as its art element, it will be imperfect. Technology made film “talk”, in turn, film raised claims to technology.

3.6.1 The technological value of ADSS

For the film nature, we know that film is a combination of light, shadow, sound image, etc. From the early research about wave propagation to polarized light principle, which is the core of 3D film, people utilized predecessors’ theoretical achievements to realize 3D vision. Similarly, 3D sound can also be realized. Since Huygens-Fresnel principle fits all kinds of the wave propagation, so either light wave or sound wave should accord with this principle. IOSONO 3D, as a most typical representative, is one of the three-dimensional surround sound solutions based on WFS principle, which originated from Huygens-Fresnel principle. Actually, the significance of all principles is not just to explain the natural law, but to be applied in to human life and promote the development and progress of human society. In turn, the application of principles also reflects human’s speculation about self-development.

The appearance of ADSS makes film closer to the real life. With ADSS, the acoustic effect presents a wider band and lager dynamic range, which can be very close or even exceed the extremity of human audition.

Since film sound finally realizes reproduction in a three-dimensional space, a lot of

new sound recording techniques can be adopted into prophase recording period. Sound reproduction, before ADSS, never broke away from the constraint of loudspeakers. Even Auro-3D and Dolby Atmos have to rely on the layouts of loudspeakers. But in IOSONO 3D surround system, loudspeakers have no longer to be taken into consideration in sound reproduction. The usage of loudspeakers is to realize its physical function: clear sound playback.

3.6.2 The artistic value of ADSS

Like music, film is also a time art. The nature of film (especially 3D film) required film sound has to be coherent and motive. Sound reproduction, for its nature, contains technological production and artistic creation. Acoustic technologies, such as ADSS, give film sound a depth perception. These multi-layered sounds are able to reproduce the natural features of original sound, enlarge audio charm, increase the articulation and render the music. Actually, the main aim of technological means is to realize the creativity of artistic conception of sound design for the films. The varying sounds can create different auditory percepts. Sound production is not just sound modeling or sonic narration, moreover it should be as what the sound may be after be processing.

Audio creation is not imaging without foundation. In fact, it is to create an imaginary audio world, other than to copy an existing one. Moylan discussed about the aesthetic elements of sound recording and production in his book *The Art of Recording*. He put forward that sound has three different dimensions: physical dimensions, psychoacoustic conception and idea. The sound aesthetic feature comes from five physical elements: frequency, amplitude, time, timbre and space. These five acoustic elements interplay and present different acoustic aesthetic forms. The table blow displays the relation between acoustic elements and acoustic aesthetic:

Physical dimensions	Perceptive units (Psychoacoustic conception)	Artistic/ Aesthetic elements
Frequency	High / low pitch	The gradation and relation (Melody, tone, density, range and vibration, etc.)
Amplitude	Loud / low loudness	The layer and relation of sound energy (tune, magnitude, vibrato, etc.)
Time	Time duration (perceiving time)	Rhythm forms (rhythm, time, continuation pattern)
Timbre (dynamic surround, spectrum and spectral range)	Holistic perception of sound quality	Sound sources and sound quality (different groups of sound sources, intensity of sound expression
Space (interaction of sound sources and related surroundings)	Spatial perception (perceiving the interaction of sound sources and related surroundings as well as sound sources and listeners)	Spatial existence (features of locating space, distance, moving status, the reaction of perceiving the space

Obviously, the spatial perception is an aesthetic element of sound. ADSS develops film sound into the three-dimensional space, a more realistic and more accurate space, not only for the in-screen space, but also for off-screen space (especially for off-screen space). The extended dimension may provide more space for sound creation. The immovable sound sources in previous surround configurations, now can move freely around the cinema space.

In term of the dialogues, ADSS is very effectual approach in flashback scenes. In previous surround formats, sound engineers will add appropriate reverberation to

make the sound in flashback different from the dialogues in the normal narrative time. In ADSS format, sound engineers can place this kind of dialogues (flashback scenes) in both front space and upper front space. Unlike the former auditory sensation, ADSS increases the intelligibility and definition of dialogues. The audio atmosphere will full of acoustic tension. Another situation is two persons with different statures. The height difference will be well expressed by ADSS format.

In term of sound effects, ADSS is good at describing motion trail and sound details. For example, sound source moves from front via overhead to rear, from LF (left front) via overhead to RR (right rear), from RF (right front) via overhead to LR (left rear) or revolves around the space. The lateral and overhead sound localizations are conducive to expand the off-screen space. The big scenes such like battle or war scenes and fireworks display. For the battle or war scenes, in previous sound production, all the sound effects are sent into main loudspeakers. In ADSS format, when sound engineers are producing the bomb-detonating, they should consider about the relevant sound effects, such as the sand or pebbles flying up. To produce the sound effects for scraps flying up, sound engineers can design motion trails from ground to upper space. Moreover, the timbre of bomb-detonating on the ground is usually medium and low frequency and scraps is medium and high frequency. The auditory sensation of bomb-detonating should contain a transition of timbre, which is closer to real scenes. For the fireworks display, with conventional surround format, the sound effects of fireworks dispersing usually be located in front channels and other relevant sound effects are in rear surround channels. There is a big gap between front and rear channels, so the sound effects are disconnected and fragmented. ADSS surround format can located fireworks overhead and the relevant sound effects in surround channels. The sound effects are is full of layering and dispersing fireworks surround the audience.

In term of the ambient sound, ADSS can easily create a nature immersive audio perception. We can always hear some natural ambient sounds in films, like birds

chirping, sheep bleating, cattle lowing, wind blowing and leaves rustling. These ambient sounds can be arranged in the upper space in ADSS format. This sound design is very accord with human's natural auditory sensation. Another example is the sound of church belling, which is a very good example as off-screen spatial expression. Such kind of sound effects constitutes a stereo sonic plane and gives depth for film ambient. In the meanwhile, to elevate these ambient sounds into upper space can effectively avoid generating the tuning or masking effect with other sounds. In disaster films, scenes of extreme weather and climate always appear on the screen, like tempest, thunderstorm, avalanche, earthquake, landslide, tsunami, etc. To process such shocking sound effects, sound engineers can not only elevate the sound effects to the upper space, but also can utilize non-direction of LFE (low-frequency effects) to create the near-ground sound effects, such like earthquake. There is enough spatial space for these sound elements to express energy range and create immersive perception. Moreover, this sound design is conform to the laws of nature.

In term of film music (cine-music), ADSS offer a total envelopment for the audience and a complete creative space, which is almost same as the real world, for the film composers. It is because film music, unlike normal music, has to participate in film narration. Film composers use "space" as a musical parameter, with objects changing their aspect or location with respect to listener, or even by the depiction of entire "unlikely" or "unfamiliar" environments. A very simple example, a music paragraph is the sacred religious music in a church. Sound engineer can draw out the pipe organ and solo bass voice from the music layers and put them in the upper space. This sound process will produce a sacred atmosphere. The film will be enveloped in a mystical keynote.

3.6.3 The value of 3D Audio-visual congruence

ADSS surround system realizes producing a 3D sound scene spatially coherent with the visual content of 3D film (truly 3D film). In Zettl's book *Sight*, *Sound*, *Motion*, he

mentioned about aesthetic factors of sound. He considered that everyone, even not a sound specialist, can perceive the three basic aesthetic factors: Figure-ground, Sound perspective and Sound continuity.

Figure-ground means that sound states a spatial relation with the picture. As the lens panning in or out, the auditory sensation will present a distance perception, close or far away. For example, sound processing for two kinds of scenes, close-up shots and full-length shots, are different. Depending on the Figure-ground, sound perception can show the sound quality and depth. On the one hand, sound should reasonably go with the picture. On the other hand, sound engineers have to know what kind of sound they want the audience heard. About Sound continuity, it involved sound volume, quality and the sound connecting between frames.

In process of film production, all the film elements or aesthetic elements, either sounds or frames, are interdependency and impact the whole film presentation together. Some of the filmmakers and film critics insist that, the nature of film is “content”, in another word the film story, but not the audiovisual effects. It is true that the nature of film is “content”. The artists appeal their emotions to film creation. The films are the emotional carriers. By this emotional expressing form film can arouse the audience’s response with the directors. If the film story can be seen as “the essence of things”, then audiovisual effects can be taken as “the appearance of things”. From the epistemology of view, the process of cognizing everything is “seeing through the appearance to perceive the essence”. Just imagine a film with poor audiovisual effects, fuzzy images and untextured sounds. It is impossible to do the further observation by such indistinct appearance. If it is impossible to adequately cognize the appearance of things, let alone the essence of things. Thereby, it would be meaningless to improve the quality of film frame and film sound. To accept the artists’ emotion behind the artistic appearance is the objective, enjoying films is the process. Objective is important, but the process should not be ignored. Modern people pursue practicality, emphasize utility, value finalism and always despise the process. The deep-seated

reason is lack of rational sense. Using Einstein's theory of relativity as a simple example, relativity means relation and its essence is to reveal the universal relation between substances and motion, substances and time or space. The result derived by Lorentz coordinate transformation is that time and space is relative. That is why we call it the theory of relativity. Universal relation is the essence, and relatedness is the result. If we just focus on the result but ignore the process, this theory will be cognized in a contrary way, which thoroughly exposes the lack of rational sense.

But here, no matter how great promotion and convenience ADSS could make for film production, human's emotion is irreplaceable by any technology. Applying immature technologies in film production will result in a waste of resources and desecration of art. Film *Simone* (2002) is a very typical example. Chinese big-budget film *Master of the Crimson Armor* (2005) played with the profound dialogue, obscure symbol and meaningless scene which draw on a lot of critics. Advanced technology is not a "panacea" but a key to success.

3.7 Deeper thinking about ADSS: the relationship between technology and artistic concept

Since digital technology was introduced into film field, film production entered into a "freewheeling" era. People never paid such great attention on film sound than ever before. In the public's viewpoint, the appearance of ADSS not only temporized on the requirement of high audio quality but also developed in accordance with the development of film vision development. ADSS has a further significance for artistic theory. It represents the most advanced audio technology. Reviewing through film history, he development of film technology are always associated with the development of film art.

3.7.1 The development of film sound technologies promoted the evolvement of film conceptions

Since sound appeared as an expressive element in film, there were a series of changes about theconception of film sound from boycott, doubt, accept to extensive applications. All these changes have reasons. The most pivotal reason is the development of sound recording and reproduction technologies.

3.7.1.1 The development of technology changed outdated concept

The outdated technologies restrict filmmakers' conceptions and sometimes would lead to one-sided or even wrong theses. As the technology keeping developing, new possibilities of technologies prompted people to rethink the old theories.

There are three most important stages of audio technologies in film sound history: sound appearing in film, the applications of magnetic sound recording and multi-channel surround sound in sound reproduction. As these audio technologies appeared in succession, filmmakers gradually changed old aesthetic conception (see section of history of film sound).

First of all, film sound made people had a *new cognition about film space*. In the **silent film era**, the filmmakers' cognition about film space was restricted to the visual frame. Some of the filmmakers tried to explore the space out of the pictures. For example, in the film *The Birth of a Nation* (1914), Griffith used an arm which was hugging around the son to imply the mother's existing. Similarly, in the film *Battleship Potemkin* (1925), Eisenstein used a shadow of the Tsar army to imply tsar army which was out of the picture. Later, filmmakers improved this method: the actors going into or out the frame, the sight that pointed to the outside space of picture, for example, using an incomplete body to imply the outside space of picture. All these attempts and visual designs were based on the pictures. As the appearance of phonofilm, this visual implying method was no longer the only way to imply the outside space of film picture.

Sound would offer a possibility for limitless space outside the picture and brought about a new concept of film space: film space is a balance space of vision and audition. Sound and picture interact with each other. It means that the film space is constituted with the space inside and outside the picture. The old conception about film that film picture only existed in the film screen in the silent film era had been broken. Sound made film space extended out to outside space. Things could move freely around the inside and outside picture, the camera could focus on anything even this interest point was outside the picture. What's more is that the film information could be delivered by sound.

Secondly, the magnetic recording technology made *people rethink the cognition about ambient sound*. At the beginning of the phonofilm period the quality of audio equipment is bad, in order to aviod influence dialogue, sound engineers'main purpose was to reduce the ambient sound. ambient sound was considered as a kind of noise and should not be a part of film sound. Although, in the film *Applause* (1929), Mamoulian made a breakthrough try about ambient sound which was considered as violate the principle of film production.

Another reason that ambient sound was repelled by film production was, at that time, the sound technology and device could not reproduce sound veritably. Only when the film sound could express the material which could tell the sound source, like wood, stone and metal, or process of sound occurring, like friction, blowing, shake and vibration, then this sound would be more realistic. So a broad frequency range and a clear sound-occurring process were of great concern. But before 1950, phonofilm adopt optic sound recording. The frequency range of sound was limited below 10000Hz. This restriction resulted in the limit application of ambient sound. Without high frequency sound such like, insects chirp, footstep on the cobbled road and crunching of doors would be vacuous, inanimate or even unacceptable. That was why most of films were repelled these meaningless and unrelated sounds.

Until 1950, filmmakers began to adopt magnetic sound recording technology. The frequency range, *dynamic range and separation degree had been great improved*. Film sound could finally reproduce the lifelike ambient sound. Then people gradually paid attention to ambient sound and the status of ambient sound changed increasingly. Later, the appearance of magnetic tape recorder provided a much easier way for people to do outdoor synchronous recording. Ambient sound began to be widely used and people also began to notice the function of it. The ambient sound had a strong ability of creating film space. In 1961, Robert Rowthorn created an exquisite off-screen space by using ambient sound in the film *Cheater (1961)*. Filmmakers kept innovating in their practice and gave a lot of new connotations of ambient sound. If the filmmakers could process and arrange sound in an appropriate way, even noise could convey some information or a certain theme. In Luis Buñuel's film *Le Charme Discret de la Bourgeoisie* (1972), the sound effects, like typing or jet plane taking off, far exceeded what they were in the real life. They carried more profound meanings namely .

Another reason was, thanks to the multi-channel surround sound, people changed their notion about the attribute of film sound. Although early in nineteen thirties, sound engineers had demonstrated the perspective of sound and introduced the sound-field reconstruction. These experimental demonstrations did not attract many theorists' attention. Béla Balázs had a certain understanding about sound space.

> *Every sound has a space-bound character of its own. The same sound sounds different in a small room, in a cellar, in a large empty hall, in a street, in a forest or on the sea. But meanwhile, he also pointed out that,sound has the same character in one part of space as in any other; it can only be louder or softer, closer or more distant and mixed with other sounds in differing ways. . . . We can see the dimension of space and see a direction in it. But we cannot hear either dimension or direction…the direction of their source cannot be discerned even by a perfect ear, ifno visual impression is present to help* (Balázs, 2011).

We can see from above, in that time filmmakers' conception about sound space was only in a reverberation level. The filmmakers had no idea about the deeper attributions of sound like position of sound sources, which was involved with human ears.

The reason that might cause this one-sided conception was that in the past, film artists did few research about abstract electro-acoustic and had less apperception about real life. The mono-sound technology would not reproduce sound space was a very powerful proof for their theories. Film sound in mono-sound film was recorded on one sound track. But in the real life, sound spreads in a three-dimensional space, recording (one channel) and reproduction technology of mono-sound made sound lost parts of directional sense, left only reverberation-sense. Then, of course, peop le could not 'hear' sound space in the films.

Since film sound entered into stereo era (dual channels), the physical property of film sound, three-dimension, could been adequately reproduced. Béla Balázs's argument 'people could not distinguish the direction of sound', had been thoroughly refuted. He tried to use the different direction-distinguishing ability between vision and audition to improve the importance of vision. But this argument was short of scientific evidence. However, his theory about 'sound has the same character in one part of space' was reasonable. But he ignored the audience's feelings about sound. It is because that, even in the same space, the audiences who sit in different locations will have different feelings about sound.

In multi-channel surround sound system: left, centre, right, surround, ULF (ultra-low frequency), all these channels together made sound-direction much clearer, people had a new cognition about the physical property of sound, especially the space property of film-sound. Three-dimensional sound, as a ubiquitous substance, made the film world became a quasi-real existence form between hallucination and real material: the screen space and the film world are hallucinatory, but auditory

three-dimension space is a real physical presence.

3.7.1.2 New abilities expand the aesthetic function of sound

Since sound involved into film, the process that filmmakers' exploring film sound could be considered as the process of phonofilm aesthetics developing, improving and maturity. New technology not only changes the behindhand aesthetic conceptions, but also bring about infinite surprise and variety of aesthetic functions.

Since stereo technology and digital technology were successively invented, sound showed a strong expression of film space. Early in nineteen forties, people had realized that film-sound had a powerful ability of film-space expression which could be used into practice and made the phonofilm represented a new state. Using sound (ambient sound) to express screen space and using voice-over to express off-screen space was a very important improvement. In Orson Welles' film *Citizen Kane* (1941), he used sound to account for the off-screen space for many plots which broke the limited of picture frame. Since this milestone film, filmmakers began to explore the ability of sound in space expression.

Mono-sound technology made film sound loose parts of sound space (especially directional sense), but other features like reverberation, dynamic range were well retained and produced by audio devices. Therefore, people realized that the important of the spatial features and its powerful effect of spatial expression, they began to utilize feature to create film-space. In the film *Citizen Kane* (1941), sound professionals used a reverberation to create a huge stone palace with a deserted empty echo to depict the protagonist's vacuous life. This try also developed a new way to produce the sound (especially ambient sound) in the post-production: it is not necessary to record film sound in the real shooting site. By means of changing the volume, tone, definition and reverberation of sound, some sound effects which had different spatial features could be reproduced in the post-production process.

In the mono-sound period, people's impression of sound functions remained in the superficial, aesthetic layer of experience: participating in narrative, expressing emotions and showing the off-screen space. In the stereo-sound period, along with the reproduction of the three-dimension properties of sound, the pursuing of special perception in the mono-sound period turned out to be very easy to mimic. On the one hand, stereo sound provided a new ability for off-screen activity. Sound could imply the existence of off-screen space, which could also express by different position. For example, a door opened in off-screen space: the audience could deduce where the door was through sound source placement without actor's sight or movement. Visual factors would be liberated through sound expression. On the other hand, people's requirements for film-sound no longer just remained in the clear and natural domains. The advanced technologies provided a wide space for people to establish new consciousness about environment. On account of noise reduction technology in stereo-sound system, sound professionals could use numerous microphones to record ambient sound in full detail and produce these sound materials into multi-layered and exquisite ambient sound in the post-production, which would increase the sense of space and reality. For example, in the film *Apocalypse Now* (1979), excluding dialogue, all the sounds were produced in the post-production stage. In a shot of 'air forces' the sound effects were produced with 200 sound materials in 127 sound tracks. By massively and particulously producing these sound materials sound engineers created a stereo sound space. Without stereo-sound technologies, especially noise reduction technology, these sound effect could not be achieved.

Analog stereo-sound technology reproduced the physical property of sound.Three-dimensional soundmade the orientation perception of sound clearer. Digital stereo-sound technology went a further step: It realized the discrete surround sound which enhanced rear sound space. The off-screen space extended to the audience's back. The surround-sound channels, which are located in the back part of the cinema were the audience's second pair of eyes. The audience could distinguish

the directions of objects through the surround sound. This provided a new possibility for sound expression. After the 90s, in some films, especially the war films and science fiction films (which placed emphasis on audiovisual effect), sound effects had developed from pursuing the orientation perception to all-around immersion. For example, in film *Saving Private Ryan* (1998), bullets fire past the soldiers' ears and in the film *Pearl Harbour* (2001), the plane flew behind the audience. In the film *Stars Wars Episode I*(1999), the digital stereo-sound effect of flying saucers' circling and leaping pushed the expression of sound movement to a new level. Moreover, the feature that digital multi-channel surround sound system had a large dynamic range, wide frequency band and low noise which could show more dynamic fidelity and details. The audience would not only be touched by the muttering (of what?), but also cheer for the landslide victory of Anacin Skywalker. Due to the increase of available sound tracks, the viewing space (cinema) became lively.

About film music, technologies and new devices solved many troublesome issues and paved the way for a better expression of film music. 'In a sense, piano was the beginning of film sound.' 1896, Lumière brothers showed their film L'arrivée d'un train à La Ciotat (The Arrival of the Mail Train). It was accompanied by piano. About this phenomenon that early silent films were always accompanied by live music could continue and develop, Sierfried Kracauer thought that:*'From the very beginning people seem to have felt that the sheer presence of music considerably increases the impact of the silent images. Film music was originally an ingredient of film performances rather than an element of film itself. Its vital function was to adjust the spectator physiologically to the flow of images on the screen.'* 1908, famous French composer Camille Saint-Saens composed the film music for film *Assassinat du duc de Guise, L'* (*The Assassination of the Duke de Guise*). It consisted of an overture and five scene songs. He became the first film composer. 1925, when Eisenstein's film *Броненосец Потёмкин* (*Battleship Potemkin*) was released in Germany, Edmund Meisel's music enhance the film by the bright revolutionary nature and high passion. The Germany government evaluated: Without music, film is art; with music, film is a

little demonstration of political dynamite. But a huge orchestra was an unbearable economic burden. The performers' musicianship and playing skills were different from each other, so the performance style for same song was hard to be unifiedand the music expression was affected a lot.

In August 1926, Warner Communications shot the first Vitaphone film Don Juan. Since then, various styles of music successively appeared in films with dialogue and sound effects. But during the period of shooting this film, sound was recorded on the phonograph record. Sound and picture of this film could not totally synchronize with each other momently. So, at that time, music could only render some scenes with few movements.

Only after the phototube was introduced into film field, this dis-coordination of sound and picture had been solved. Finally, music would cooperate with picture. In Soviet film *Мы из Кронштадта* (*We are from Kronstadt*) (1936), sound professionals utilized music ingeniously. They brought music into story successfully. For example, when Red Army men were fighting in the position, there was no directly fighting scene. Instead, the director used the downsizing of musical instruments to show the brutality of war: the sounds of musical instruments drop off over time till only the sound of trumpet and drum are left. This expressive technique was very imaginative. But in the 30's and 40's, these successful explorations were minority, 'the main function of film music in the early phonofilm did not change a lot: to tell stories and explain pictures. Sometimes, it just used for filling the interval between dialogues. Most film composers followed classical post-romanticism style, 'their composition did not follow the feature of films but structure of dramatic films' to give a leading motive to characters: following Wagner's leading motive method to compose sorts of symbolic songs for picture. 'film music was directly used for announcing characters' motives'. This application disturbed the narrative and destroyed the suspense. In the meanwhile, because of the outdated technologies and concepts, there was no cognition about ambient sound. The audience thought the film sound was only consisted of dialogue and music. So

music became the main method to fill in the blank space of film sound. In around 25 years after phonofilm appeared, few films used the music shorter than one third of films' length. Sometimes, the music would last half time of the film-length or even more.

As the widely used of montage, the films' pacing speeded up and the inapplicability of leading motive music came out instantly. Music theoretician Eisler considered that, leading motive is appropriate for the long and consecutive part of film, not for interlaced short part.'Then people began to explore the variety of audio-visual counterpoint. But the main method for recording music was phonogram which was hard to do film editing. It limited the recreation of film music.

After World War II (Second World War), the application of the tape recorder (magnetic strip) solved this problem. Tape recorder replaced the position of phonogram and became the main recording method by its virtue of erasure and reusability. It provided a possibility for music editing. In addition, the large frequency range of tape recorders could express more dynamic tones. Many new combinations, like counterpoint and synchronization of sound and picture, enhanced the expression force of music. Filmmakers sum up the functions as: applying background atmosphere, expressing actors' feelings, deepening theme, film reviewing, linking up two scenes and other functions (describing scenes, implying a certain space and time).

At the end of 60's, the musical usage became more rational. Between 30's and 40's, film music was very popular. For example, in the film *Gone with the Wind* (1939), music lasted for 134 minutes, accounted for 60 percent of the whole film. To the 60's, such excessive, epic music was seldom used in film. The reason for this change was dual. In technology respect, at the beginning period of phonofilm, technology limited film sound design. A lot of ambient sound and sound effects were inanimate, so they were excluded from consideration. To fill the blank of dialogue, the only thing that

sound professionals could do was to use a mass of music. After 60's, thanks to the development as sound technology, the quality and the vraisemblance of ambient sound and sound effect had been improved a lot. Since the expressive force of film music had been found constantly, it did not have to execute a duty of replacement. In art respect, filmmakers were fully alive to the important aesthetic function of film music in the exploring process. Film music should not have to totally cater to the movements and emotions or to be the underlay of background. It should be an independent film element which contains a certain emotion or symbolic meanings. Only when film music was considered with other sound elements, it could fulfill the expression in films.

Film music developed together with film for over a half century. In the 1960s, some filmmakers thought that there was no space for improvement of film music, because music was more mature and perfect than other sound elements. But the development of technology always surprised us. When film music developed into the next decade, people realized that there was more space for it.

In 1970s and 1980s, sound technology made an astounding advance. The development of stereo-sound technology was a new dynamism for film. Stereo-sound technology restored the spatial feature of film sound. Of course, film music benefited from this new technology. In stereophonic film, music layerings were clear-cut. Music had a more accurate directivity which could express depth information. Early in the Cinerama stereo period, in American film *Julius Caesar* (1953), filmmakers made a try in film music: they used stereo loudspeakers to express the spatial information, like direction perception of music.

The improvement of sound technology made filmmakers paid more attention on the density, gradation and direction. Music, as a part of film sound, since it was included into film, it was defined on the relatively suitable gradation. Because, people realized that 'if there were many sound elements appeared in a same scene, these sound

elements must be arranged orderly in distinct gradations from the requirement of artistic expression. Stereo-sound technology and multi-channel technology leaded people to a practical period of film music creating. By means of advanced technology, filmmakers could arrange film music in a more exquisite way. Then film music could changed from light to heavy, featured a certain musical instrument or voice, or even formulated the spatial character of music.

3.7.1.3 The technological development brought in new creating conditions, methods as well as trains of thought

If the creators have to solve the problem about technologies or machinery all the time when they are creating, the quality of the works will be influenced inevitably. The artistic conceptions will be difficult to be put into practice. Sometimes, they even have to concede to the basic technology problems. So the development of technology will directly influence artistic creation. Advanced technology will provide an more efficient and convenient creating environment. Creators can concentrate on creating and new creating methods will broaden creators' scope of mind.

Actually, ambient sound was mostly influenced by technological factors. In the early films, all the ambient sound was recorded during the shoting site period. But because of the low SNR (Signal to Noise Ratio), the environment of the shooting site was hard to control. So the ambient sound was always fuzzy and discursive. Sometimes it would be interfused staff's voice.

By the 1960s, along with the improvement of microphone and the application of magnetic sound recording and tape recorder, the creating condition of ambient sound had been improved a lot. In the aspect of technology, 'from condition of acoustic field (sound recording place) to the advanced degree of electronic sound equipment and recording technique, the definition of sound could be the same as picture'. In the aspect of art, filmmakers had a relatively clear thought and deep cognition about

time-space relationship of sound and picture, functions of film music and lifestyle dialogue. ‘At this time, both technical specification and artistic practice were qualified enough to produce films with balanceable audio-visual factors. What is more, in the auditory range, actors’ voice, natural sound and music were also in a balanced state.’

However, when all the film sound (dialogue, sound effect and music) were considered in the same status, people found that, on account of masking effect, they could not adjust all parts of sound as what they expected, especially the quality of ambient sound. The frequency which could guarantee the articulation of sound was around 3~5KHz. So if dialogue was mixed with music or ambient sound which had the main energy around this frequency, the articulation of dialogue would reduce a lot, unless the volume of other soundswas very low. Experiment showed that, other sounds in this frequency (music, ambient sound etc), should be 5~10% lower than dialogue in total, then they would not destroy the articulation and comfort degree of dialogue. But the frequency of most ambient sound would conflict with dialogue. In order to ensure the articulation of dialogue, filmmakers could not help but reduce the creative application of ambient sound. So this became the main problem that filmmakers urgently wanted to solve. Unquestionably, the appearance of multi-track recorder and multi-channel stereo sound system brought about a new way to create.

On the one hand, ambient sound could be recorded in different tracks by the multi-track recorder. During the post-production, sound editors could adjust the volume and frequency of ambient sound according to actual demand. They could also do noise reduction when it was necessary. Through this, sound professionals could ensure the high quality and coordination of ambient sound in whole film. On the other hand, since we had multi-channel stereo system, the main dialogue could be arranged in the centre loudspeaker, ambient sound and other sounds could be arranged in the left , right and rear loudspeakers. By doing this, when other sounds kept a high volume, they would not disturb dialogue. This production method provided a realer, more exquisite and more infectious ambient sound.

After 90s, digital technology provided a more flexible convenient creating method for filmmakers. When producing ambient sound, it would take a great many of soundtracks and need repeatedly re-recording. Sometimes, it would take over 100 soundtracks, so how to keep sound quality was a pivotal. Digital stereo sound system solved a lot of problems which were hard to deal in analogue technology period, especially the loss of signal during the re-recording period. When doing sound recording, analogue system produces a series of consecutive electrical signal. Every time when doing mixing and re-recording, the quality of this sound signal will be reduced. But digital system changed sound information into a series of on-off pulse electrical signal, this kind of signal would not be reduced in sound mixing or re-recording, so the sound quality would not be reduced. Moreover, many problems, like the limited capacity of information recording, cross talk, background noise of recording radium, were solved by digital system. Since sound professionals got rid of troubles of hardware and technologies, they did not have to calculate the exact soundtracks or re-recording times to avoid loss of signal. What is more, ‘digital editing produces the most essential part of film. It intensified sound materials and provides a powerful control, sound editors do not have to worry about the quality of materials.

Music creating benefited a lot from the development of technology. At first, film music was recorded on the phonogram which was very hard to editing. So before editing, sound professionals had to follow the film and directors’ request to calculate the exact length and appearance time and during the recording period, it should not make any mistake. Otherwise, when filmmakers found out that sound did not match picture, everything had to be repeated from the very beginning. If the director changed his mind and wanted to modify film music, even it was just a little change, the sound had to be recorded from beginning. This kind of recording has no flexibility at all.

After 50’s, the application of tape recorder made a radical change in film music creating. Music was recorded on magnetic tape, so it could be editing willfully. The

flexibility of music creating was improved greatly. Composers could choose a certain part of music on tape to musicalize film. In the post-production, if filmmakers wanted to add or delete some parts of music, it was also very convenient to manipulate.

In 70's, Tascam Company launched a four-channel tape recorder. This recorder's main feature was to recording multi-instrument music or multi-channel music. For a certain song, different instruments could be recorded on different soundtracks. In the post-production period, they could be separated to debug the volume or other indexes until all the voice parts achieved a perfect frequency balance. Compare with the old recording method that all the sounds had to be recorded in the shooting site, n-track recording surely improved work efficiency and music quality.

After 70's, more and more electronic music instruments were put into use for film music. In order to realize the connection of different instruments, the standard of MIDI (Musical Instrument Digital Interface) was set up. MIDI would coordinate the operation of electronic synthesizers, electronic drum and other electronic instruments.

So composers would use MIDI to control all instruments. At the end of 80's, MIDI could connect with SMPTE (Society of Motion Picture and Television Engineers) time code. Then cyber-synth music could easily fuse with images. In the past days, composers always composed by piano. Different instruments or voice parts could only exist in composers' minds and soundless musical notations. Before orchestra performing it, the effect of instrumental ensemble was uncertain. But digital electronic synthesizer broke the routine. It could imitate various instruments which would set composers free from paper into practice. MIDI technology assisted composers in playing different electronic instruments with facility. When composers were creating music on electric synthesizer, meanwhile, they could also use MIDI to coordinate and control the performing status of other instruments. The production of film music became more intuitive and more efficient.

The main purport of technology development is always to create a higher quality, more convenient and efficient artistic creation environment. Since 90's, film music production presented a characteristic of digitization in all directions: design in early stage, recoding period and post-production stage.

Usually, in early-stage design, film composers had to compose according to film copy which had already been edited. Obviously, they could not mark on video tapes. When composers created music, they had to refer to rhythm manual to confirm numbers of music bars. When they filled the time list, they had to mark the time code of music entering point and the length of it, etc. But these cumbersome works were all replaced by computer software. With the help of computer software, composers would throw off all these headache timing works and focus on composing.

During the creating period, music composition software would inspire composers now and then. When finish demo, according to composers' information input, computer software could imitate all kinds of instruments and do real-time performance. This kind of music production software emerged in endlessly. From the early software: Cakewalk and Sibelius to the more advanced one: Pro Tools series, cyber-musical software helped composers to realize the expectation that inputting, modifying and reviewing music in real time. According to composers' requirement, some musical software could even generate an off-the-cuff music which could inspire them a lot. Creating by software also had an unparalleled advantage in storage capacity. Usually, a three-minute song would take about 15M hard-disk space, but a MIDI-format music would reduce a thousand times space which would save plenty of hard-disk space.

In music recording period, digital multi-track recorder made recording process much more accurate and flexible: different instruments could be recorded in different time. What is more, it could record or copy in the breakpoint which could save a lot of human and material resources. Along with the birth of multi-track recorder with hard-disk, the storage capacity of recorder, read-write speed and stability would be

enhanced drastically.

In post-production period, the powerful digital audio workstation greatly simplified the working process. It could modify music accurately: ‘in point’ and ‘out point’ of music, fade in and fade out, the volume, acoustic image, reverberation, or effect of each soundtrack or even minor mistakes. As the technology kept developing, digital audio workstation already developed into a powerful digital audio processing system consists of sound recorder, sound console and effector, etc. These years, more and more films were produced by audio processing system. Films like *Gladiator* and *Crouching Tiger, Hidden Dragon* stolen the show in 2001Oscar Awards. They were all produced by Pro Tools audio processing system.

When the incidental music was completed, the music file would be sent to anywhere in the world. Composers could summarize all advices to modify the music. They could even convoke a real-time meeting via the internet.

About dialogue, soon after film appeared, people began to record sound and picture separately. As the beginning of phonofilm period, there were a lot of problems in recording live sound. For example, it was hard to record high-fidelity sound in the open air and hard to arrange the position of microphones in long shot, etc. So film dubbing emerged. But the saptial expression was restricted by technical conditions. Therefore, the main purport of dubbing in this period was to make the dialogue smooth and clear.

After 50’s, the application of magnetic recording recorder and the invention of Nagra recorder leaded dialogue production enter into a new stage. The coverage of optical sound recording would not extend to the low frequency of human voice. So it was hard to avoid voice distorted. Magnetic recording extended frequency range of sound, thus improving the voice distorted. In the same time, Nagra recorder increased the flexibility of live recording.

At the end of 70’s, advanced sound recording devices and multi-channel stereo system improved SNR (Signal to Noise Ratio) of film sound. For purpose of shortening production cycle and enhancing the sense of reality, more and more films adopted live sound recording. The most important part of live sound recording was dialogue. That is because if it was necessary, other sound elements, like ambient sound could be modified in the post-production period. But it would be difficult to record live dialogue again. Since microphone was the crux of dialogue quality, so high-sensitivity microphones were indispensable. Early in 30’s, people fixed microphones on booms to collect sounds. This method could retain most sound information. However, this method could do nothing about long-distance sound. The farther talkers were, the worse sound quality would be. In 70’s and 80’s, the invention of directional microphone solved this problem. Even the sound source was far away from microphone, it could also be recorded clearly. Other microphone types were also widely used in different environments. Among these types, with a better directivity, cardioid microphone and hyper-cardioid microphone were the edge tools of recording live sound. Some of other microphones which had voice attenuators or switch of CTM (closed talking microphone), could used for special dialogue recording.

Although the quality of live recording was improved continually, but dubbing had been retained till now for various reasons and this technology had a new term: ADR (Additional Dialogue Recording). Compare with the early-stage dubbing, ADR could ensure the articulation of dialogue. What is more, ADR could solve some problems, like lip-synchronization and the sound spatialization which were insoluble problems before. ADR is no longer a technology which filmmakers are forced to be chose. Conversely, it is an advanced technology that can make the film more detailed today.

After 90’s, digital revolution also influenced dialogue production. Firstly, DAT (Digital Audio Tape) took place of Nagra recorder. DAT had many advantages: compact and lightweight, easy to carry, high quality, convenient to read and write, high capacity,

super-strong EC (Error-correcting Capability). It became the main device for film dialogue recording after Nagra analog-recording. DAT was used in the famous film *Jurassic Park*(1993).

Secondly, with the help of DAW (Digital Audio Workstation), the process of post-production for film dialogue had also changed. When sound editors edited live sound materials, it was very easy to achieve sound-and-picture synchronization. Filmmakers could compare several shots at the same time and pick up the best one. The recording technology of ADR was also different from traditional method. In the past, only after the editing of pictures, sound professionals could edit sound. Now, sound professionals could begin to edit sound materials earlier through DAW sharing system. Most of DAW systems had the auto-cycle function and recording tools. So the automaticity of ADR was improved.

In addition, the quick popularization of network technology and high network transmission speed paved the way for realizing remote ADR. In producing film *Searching for Bobby Fischer* (1993), Ednet provide a network channel between Santa Monica and London. So director Steven Zaillian who was in California could guide actors who were in UK to do ADR. If the actors flied to California for one day ADR, it would take 8000~10000 Dollar for each person. But using Ednet, two person's ADR could be finished in an afternoon time and only this method only cost 2500 dollar.

3.7.1.4 The development of technology facilitated multivariate styles of musical creating

About music style, since composer Camille Saint-Saens composed film *The Assassination of the Duke de Guise*, early-stage film music was doomed to inherit classical post-romanticism. Composers all adopted this form. Wagner, Puccini, Verdi's operatic form were highly regarded and this style was brought into full play.

From 30's and 40's, composers were no longer satisfied with this serious style. Film music appeared diversified development trends and the appearance of electronic musical instruments provided a possible for this transition. 1934, Ondes Martenot was used in film *L'idée* (*The Idea*) which was a pioneer of application of electronic music. 1941, Waksman used electronic violin in film *Suspicion*, Bernard Herrmann used unconventional musical instrument in ensemble. But Robert Wise's *The Day the earth Stood Still* (1951) and Fred Wilcox's *Forbidden Planet* (1956) were the earliest film which caused a sensation in using electronic instrument. The composers of these two films: Bernard Herrmann, Louis and Bebe Barron used a shrill synthetic sound. Although the sound was uncomfortably, they effectively enhanced the dangerous atmosphere of these two science fiction films.

In early 60's, integrated circuit was widely used. People invented a new generation of sound recording devices: electronic musical instruments with integrated circuit and analog synthesizer, which were much easier to operate and more reliable. By electric synthesizer, sound professionals could modify the exiting tones and synthetic totally new tones. The film music styles became diversified: jazz, rock, electronic music, folk, etc. These music styles all could be found in the films. In this period, there were a number of popular and unfailing theme songs: jazz style song *The Panther* from film *The Pink Panther* (1963), folk style song *The Sound of Silence* from film *The Graduate* (1967) and Beatles' rock style song from the film *A Hard Day's Night* (1964).

Advanced technologies were the best tools for film composers. Because of invention of digital synthesizer, digital multi-track recorder and DAW, the 90's film music was with good acoustic quality and creative production. On the one hand, the digital creating process extremely extended the frequency and dynamic range, improved the musical quality. So film music not only strengthened artistic appeal and expression, but also surpassed the film story and became the Hi-Fi (High Fidelity) classic works. *Ghost* (1990), *Batman Returns* (1992), *The lion King* (1994), almost every successful film would be followed with a classic original soundtrack album. For example, film

Titanic (1997) leaded a great enthusiasm for the original soundtrack album. On the other hand, since composers could produce different types of music at the same time, the 90's music style was very bold. Different music types could appear in a same film and they were perfect combined with film. Since 1991, various kinds of music types showed up on screen: Rap, Reggie, New Age. In director Oliver Stone's film*Natural Born Killers* (1994), the music style changed among opera, Reggie, Bruce, heavy metal rock, electronic music, classic music, jazz, etc. This bold attempt created an unprecedented audio-visual effect. In Ang Lee's film *Crouching Tiger, Hidden Dragon* (2001), Chinese classic musical instruments and western instruments combined to create a new audio-visual experience.

3.7.1.5 Forming new conceptions

Since multi-channel stereo system came out, people realized that film sound should have an ability to reproduce a whole sound environment. Film sound, especially ambient sound was brought into consider in the early-stage design. The conception of film sound changed from 'recording sound' to 'produce sound'. People present a conception of 'acoustic thought'. Other sounds, except music and dialogue, were no longer been simply considered as sound effect but more abundant and comprehensive. These complex sound materials were not just the result of recording. They should be the result of thinking: what kind of sound could ensure that it retains the spatialization and localization of real space; what kind of sound could deepen the theme of film and what kind of sound could keep the continuity of film and not result in vapidity. All these questions should be taken into consideration. Otherwise, film sound was just a copy.

If Ben Burtt did not do such an excellent job in *Star Wars* (1977), then robot R2-D2 would not be so deeply into the audience's minds and Jedi Knight's light saber would lose magic power. Film *Star Wars* made a great success, not just in the visual effect and fantastic story, but also in the creative sound production. People began to

recognize these sound professionals' contribution. They were entitled as sound designer. 'Sound designers are responsible for the sound of whole film, from the beginning to the end. They expound the directors' intention, conceive the sound of script, coordinate composers and sound editors, participate in re-recording and even have to consider about the acoustics of theaters. Several decades ago, they were just called as sound editing management, but this salutation reduced their job nature. They are aural artists. Sound designers not only have to read scripts, communicate with directors, make a plan for recording the anticipated sound materials, but also have to think imaginatively or even create some nonexistent sounds and make these unreal sounds sound real in theaters. In post-production, they have to adjust sound layers to match the emotion changes of film. Today, sound designers have already become a pivotal role in film production.

After several decades of exploration and development, in the middle of 70's, a new conception gradually was formed: Sound Ambience or Sound Architecture. This conception was widely used in all fields of film, like film sound design, theatre acoustics design, etc. The appearance of this conception indicated that people had realized the special spatial attribute and interrelation of sound and picture. This was the second revolution of film sound.

3.7.2. Advanced conception is a guide for the development of technology

Technology is a key role of conception evolution, but not the only one. From film history, we know that although there were many conceptions came out after homologous technology, there were still some filmmakers who had farsighted thoughts and made some inconceivable attempts. Their thoughts received far less attention than they should be, but these advanced ideas were the clue for after development of technology.

3.7.2.1 The attempt for reproducing sound localizations in 30's~40's and the

development of stereo technology

Multi-channel stereo system was getting improved and developed well in 70's.Besides the research 'perspective of sound' in 30's, there were many films had attempted the practice about film space. These were all the practical precedents for its future perfection.

In film *Applause* (1929), director Mamoulian returned to moving camera shoot and reproduced abundant ambient sound in a simple and crude condition. However, this technical breakthrough did not receive public identification. Only after 70's, when sound technology and sound equipment were able to realize this artistic conception, people began to accept the conception of using ambient sound to create film space.

In film *Citizen Kane* (1941), director Orson Welles used his successful experience of radio play to create a three-dimension sound space with distance and atmosphere by changing the attribute of sound. This creative attempt was worth to be high commendable before stereo system was invented. What is more, different from other films, in this film, filmmakers created an infinite continuous space which was an ideological basis for the future conception of ambient sound.

3.7.2.2 The breakthrough of dubbing in 40's and ADR technology in 80's

Actually, the dubbing technique was a compelled choice under the excessively crude sound recording condition. Since dubbing would make films lose some reality, filmmakers timely found out this problem and tried their best to make up this defect in process of dubbing. When James G Stewart recalled his recording process of Orson Welles' film *The Magnificent Ambersons* (1942), he mentioned that, in order to recorded clear dialogue in the carriage, he recorded the main dialogue ahead in the still carriage. But Welles was not satisfied with this recording. He said that, he could not feel actors were talking in a moving carriage. There was no sense of movement of

the conversation and it was very rigid. He wanted the sound could be stained and on the move. As in so doing, the sound would be more real. So James let the actors dub as they imitating bumpy movement and he received a satisfactory result. His conception of pursuing sound reality directly influenced the future development of ADR.

For a quite long time, film dubbing lacked reality, not mention the spatialization of sound. So when live sound recording rose again in 60's, film dubbing was almost like an incurable disease of film sound production. But in practical application, people found that it was not easy to abandon it. Shooting a film was a complex process full of unknown and variation. No matter how advanced the devices were, there were always various problems. For example, actors from different countries or regions could not communicate with each others, the shooting site would not be totally quiet, directors changed dialogue during the post-production or even used a totally different emotion to express the dialogue. So inspired by Orson Welles and other directors' film, people began to explore a new dubbing technology to approach the real-life sound. Then ADR was invented. ADR could solve a lot of technology problems, such like, lip-synchronization and electrical level. But different from the early dubbing: actors had to get realistic feelings by personally practicing, relying on the advance technology, ADR could provide a relaxed environment for actors and directors.

3.7.3 The relationship between technology and concept

From the above two aspects of discussion, we can draw a conclusion: the development of technology largely decided and improved the formation and development of film conception. Of course, advanced conceptions can also direct the future development direction of technology. But about the relationship between technology and conception, there were always some different opinions came up by film theorists and filmmakers. Some of them considered it in different dimensions, some kept different theory. For example: some film theorists thought the source of

these limitations about film conceptions mainly lay in out-dated technology. The limitations of film conception would disappear as the development of technology. Some technologists believed that as long as filmmakers applied advanced technologies, film conceptions would definitely improve. Some filmmakers considered that, since we were in the digital age, it was possible to abandon some out-dated film techniques (like film dubbing). And some film aesthetician insisted the over applications of technologies (especially digital technologies) would result in the recession of film art.

3.7.3.1 Is it the out-dated technology that limits the film conceptions?

In the first couple of years of phonofilm, “dialogue recording was like Samusel Johnson’s dog, people never cared whether it could walk well or not, if only it could walk.” Camera stayed still; the visual expression of film was greatly reduced; because of the low sensitive of early microphone, actors could hardly perform and their jobs were only to enunciate clearly. So, the first batch of phonofilm seemed like a pile of farcical counterfeit.

Some film theorists considered this countermarch of art as a result of incomplete of sound technology. They insisted if the technology improved, this limitation of conception would disappear with it. However, just five years after phonofilm was appeared, film sound technology developed rapidly. Directional microphones were invented, sound and picture could be recorded separately and the camera could move around. But still, some films were stick to the aesthetic principle of silent film. A lot of stereotyped dramatic dialogue was used in films. No matter what kind of perspective, the camera shot from front side or back side, the scenes were always actors’ talking.

So, the main constraint of film was not technology, but conception. Eisenstein and others sharply pointed out that, if the directors abused “this new technical discovery may not only hider the development and perfection of the cinema as an art but also

threaten to destroy all present formal achievements." Fortunately, filmmakers realized the problem about theatrical film in time and corrected the old conception.

The first German phonofilm *Der Blaue Engel* (*The Blue Angel*) (1930) came out in 1930. Although the technical merit was backward, it still attracted the audient's attention. The film dialogue was concise and kept the visual expression methods of silent film. The English-and-German mixed dialogue did not affect the audience's understanding about this film. Orson Welles was one of the directors who used realistic dialogue in the film. In the film *Citizen Kane* (1941), the conflicted dialogue, interruptions and incomplete sentences brought in an unprecedented reality and vitality. It totally changed the conception of film dialogue. These breakthroughs made a great influence on the after film aesthetics.

This showed that, the development of technology could improve the film conceptions, but whether the conceptions and expression techniques were advanced or backward was not only related to technology. Actually, minded filmmakers could create miracles by poor devices but bad expressions were hopeless.

3.7.3.2 Does new technology surely improve film conceptions?

From the development history of film sound technology, we knew that, advanced technology could change filmmakers' thoughts patterns and improve development of film aesthetics. Then some filmmakers often overjoyed with new inventions. They considered that the application of new technologies must provide a good direction for film conception. But as a matter of fact, there is no clear causal relationship between them. And the abuse of new technology would result to the regression of film conception.

In 30's, technician invented boom microphone to liberate actors from the restriction of performing area. But the defect was it was hard to arrange the position of boom

microphones. They appeared easily in the pictures and caused goofs. So in 60's, relied on the good hidden property and convenience, wireless microphones were widely used in dialogue recording. But different from boom microphones, wireless microphones lost spaciousness and distance perception of sound. When the actors were talking while they walked away from the camera, the voice did not away. In film *The Graduate* (1967), this goof happened.

This wireless microphone recording method was questioned by a lot of sound professionals and film critics. This recording method raised a discussion and thinking: must the dialogue exactly match the characters' moving scenes? Through repeated experimentations and comparisons, filmmakers realized: the spaciousness of dialogue should be consistent with characters' moving. But when some pictures were switched, this consistent match would result in sound skip. So a better solution was: no matter choosing what kind of microphone, no matter using live recording or dubbing, sound engineers should balance all the sounds in post-production. They should not only keep the spaciousness of sounds and make appropriate adjustment, but also consider about the connection between two scenes.

The application of wireless microphone made filmmakers rethink about the relationship between scenes and dialogue and found out an appropriate solution. But its negative impact on film sound continued to today. Some sound professionals used wireless microphones for convenience, but did not do anything for reproducing the spaciousness of sound in post-production. The dialogue of the whole film was meaningless and tedious. As more and more filmmakers realized the importance of film sound, this indifference was restrained. But there were still some shoddy films were produced for low cost.

Every time when we put a new technology to use, we should consider about these issues: How will this technology effect on film production? What result will this technology bring in for film production? Illustrate wireless microphone as an example.

It could bring great convenience to live recording, but increase the burden for post-production. Filmmakers should give sufficient full consideration to the advantages and disadvantages of a new technology and take full advantage of it. Otherwise, even if it is a new technology, it can be a fatal error of a film.

3.7.3.3 Does the digital sound make films too noisy?

At the beginning of phonofilm, some filmmakers, represented by Rudolf Arnheim, were against sound appearing in the film. They insisted that film was a pure visual art and sound might destroy the artistry of film. In the after decades, filmmakers paid attention to “the reform and innovation of film sound, but the conceptions about film sound were far away from its nature. Film sound was always in a subordinate position of picture. ” Until 60’s, when filmmakers had the relatively clear perception and cognition about the relationship between sound and picture, then the theory of film sound step to a new stage. Film sound began to be considered in the equal status as picture. But since 90’s, as digital sound was applied in film sound, large dynamic sound (over 100db) greatly impact the audience’s ears, the filmmakers and film artists raised an issue: Does digital sound make films too noisy?

In fact, every time when a new technology was released, it would experience a process: be sought after, be criticized and be applied rationally. When magnetic sound recording gave prominence to function of ambient sound, some filmmakers dropped film music. But later, they found that, the endless ambient sound would exhaust the audience’s freshness and interest for films. Then film music returned back to film. When multi-channel stereo sound technology appeared, filmmakers tried their best to show off this technology in films. Then the application of this new technology returned to rational consideration combined with aesthetic. So it was understandable for filmmakers and film artists to doubt about this large dynamic sound. However, till today, there were still quite a number of films filled with breathtaking sound effects but few changes about the volume and rhythm of the sound. It was too much for the

audience's ears. For example, in American war film *Pearl Harbor* (2001), the sound design of this film was affirmed by sound professionals, but the old Hollywood love story was very boring. The only information that the audience could obtain from the sound effects were: "on the left side, there was a helicopter flying past", or "a bullet flies from back", etc.

The main problem did not lie in digital sound itself. It was not DTS, SDDS, nor broadband and large dynamic, nor digital sound devices that made film so noisy. It was because the filmmakers did not realize the key role of digital sound in the film. Actually, it was still a misunderstanding of conception. The film *Saving Private Ryan* (1998) was not successful by the amazing sound effects of the war. *Titanic* (1997) did not deeply move the audience by the shocked roar of ship sinking. Digital technology just provided more possibilities for sound expressions. Even the film was explicitly shot for sensory entertainment, it could not just produced by technology. Without moving stories, no matter how advanced the technology was, the audience just felt bored.

About the relationship between technology and creation, Bassil considered that there were issues to be clarified, the nowadays filmmakers relied too much on advanced equipments. The most filmmakers' misunderstood was they had to utilize technologies to create. Many filmmakers lost their imagistic perception. But in fact, the creative thinking was always, and it should be generated before utilizing technologies. A good work was always formed and accomplished by creative imagination. Digital technology was not the source of a good work. It just provided a possibility.

Chapter 4

Conclusions and Recommendations

4.1 The comparisons of the current 3D surround formats

Obviously, as introducing in chapter 3, these three surround formats are all aimed at adding spatially accurate sound rendering (immersive perception) to 3D films and providing a perceptually matched sound source at the position of every object producing sound in the three-dimensional film space, thus realize the coordination of video and audio in 3D films. But these three formats are widely divergent on both rationales and implementation modes. Auro-3D excels at its continuation of channel-based concept and intuitive loudspeaker layout. Dolby Atmos develops on an overwhelming brand-superiority. Each has its own competitive advantages. IOSONO 3D presents an extremely advanced solution that is based on WFS.

4.1.1 Rationale acceptability

According to the concept of ADSS, adding a third dimension on the conventional two-dimensional (horizontal plane) surround formats to describe the sound characteristics in the vertical dimension is the fundamental objective of ADSS.

By arraying two additional layers of speakers above the audience, Auro-3D adds the next, third dimension (height) on the current surround systems. This added dimension of sound produces an authentic, full immersive auditory experience. In term of rationale acceptability, this additive process is very easy to be accepted by sound engineers, filmmakers as well as the audience. Moreover, the production workflow is more readily achieved by continuing in a channel-based environment than the object-based systems.

By comparison, object-based and hybrid approaches may not come so easily. In term of comprehension, the concept of object is more or less as a process of splitting sound elements and the WFS principle seems like a recombination of sound information. But for the specific theories, the involved academic scope requires the sound engineers and filmmakers have a certain relevant knowledge reserves about physics and advanced mathematics, including auditory principle of human ear, acoustical principle, psychoacoustics, sound reproduction, indoor acoustic design, the integration of acoustic space and architectural space, calculus, convolution, etc. Of course, a large amount of calculations is processed by acoustical processors and workstation, but only if the actual operating personnel understand the corresponding principles (why they do it), they could be able to direct and purposeful produce and operate in practice.

4.1.2 Market share and Brand awareness

The name Dolby Laboratories has long been associated with innovative and world-leading sound reproduction. Taking advantage of highest dominance in audio production, distribution and projection, Dolby invested market promotion of Dolby Atmos, from the film sources to film release, at a steep financial, material and human cost. So far, at least 10 film production company support Dolby Atmos, including Skywalker Sound, Todd-AO, Walt Disney Pictures, 20th century Fox, etc. Up to June 2013, 29 released films had adopted Dolby Atmos, such as *Brave* (2012), *Life of Pi* (2012), *The Hobbit: An Unexpected Journey* (2012), *Taken 2* (2012), *Trance* (2013), *Oz: the Great and Powerful* (2013), *The Croods* (2013), *Iron Man 3* (2013), etc. What's more, over 64 cinemas had been upgraded into Dolby Atmos-equipped surround. Such high market share and brand awareness provides a good development foundation.

Competing with such formidable rival, Auro-3D chose to cooperate with Barco, a brand which is famous for its film projector, and released a joint project of integrative

audio-visual solution for the digital cinema. This surround system was supported by famous director George Lucas. In 2012, the first film *Red Tails* (2012) with Auro-3D surround format was produced and released by Skywalker. Unfortunately, this film bombed at the box office and did not propel its market influence. Till now, there are up to 46 cinemas adopted Auro-3D. Moreover, Dream Works Animation SKG promised its following 15 cartoon films would be produced in Auro-3D surround format.

IOSONO 3D was the earliest 3D surround system that launched on the film market. Early in November 2011, film *Immortals* (2011) had been released in IOSONO 3D surround format. Two Korean films *My Way* (2011) and *R2B: Return to the Base* (2012), following closely, were released successfully in Asia. The application of IOSONO 3D, a holographic sound system with the ability of supremely accurate acoustic localization, has already stepped out of the cinema domain and extended to opera, planetarium, theme parks, art installation, etc.

4.1.3 Audio quality

After all, all these audio technologies are invented for 3D film production. So the key requirement is to producing a 3D soundscape spatially coherent with the 3D vision of 3D film.

As a representative of channel-based 3D surround, the biggest advantage of Auro-3D is rendering the surroundings and atmosphere with ambient sounds as well as music. Object-based approach is especially powerful for describing the precise localizations of sound sources. However, in real world, sounds are always superposed by reverberation or reflected sounds or accompanied by ambient sounds. Thus, Dolby Atmos retained the concept of channel as its “bed” rendering.

Although, channel-based surround system can create a relatively rich environmental atmosphere, it still weak at depicting the details of individual sound source, such as

volume, source type, accurate position, etc. The film sounds are still distributed via the certain audio channels as well as reproduced by pre-determined loudspeakers. Of course, the three-dimensional loudspeaker configuration (especially the overhead layer) will help to set up a spatial experience for film sound. But still, for a single soundtrack (such like dialogue soundtracks or sound effects soundtracks), all the sound elements are mixed and playback together. It still lacks the “granular” perception (with details or characteristics) of every sound element. For the localizations of sound sources, channel-based approach is able to express them in the corresponding ranges, other than the punctiform localizations.

But accurate localization may result in a disperse perception, sound elements detached from one another, and reduce the coherence of film sound. Moreover, overemphasizing the independence of sound elements may result some of the sound elements “jump out” from the holistic sound atmosphere and divert the audience’s attention away from the screen, thereby reduce the immersive experience.

Either channel-based or hybrid approaches, as long as they depend on channel-based mode to playback film sound, cannot keep away from “sweet spot” and usually there is just one “sweet spot” at which the spatial effect of the mix is exactly as expected. On the contrary, IOSONO 3D is based on the WFS principle, so that the spatial information of sound elements gets rid of the channel concept and be “tailored” by the specific spatial parameters of the playback space.

Although all the three 3D surround formats above support the downward compatibility, they still have to be upgraded and reconstructed to meet the demands of 3D surround format. The list below is the standards of loudspeaker layout:

Auro-3D: from 9.1 up to 11.1 (base on 5.1 surround)

- 9.1 Auro-3D: + 4 height channels (one above each corner speaker/ array)

- 10.1 Auro-3D: + 5 height channels (+ TS= “Voice of God” channel)
- 11.1 Auro-3D: + 6 height channels (+ Height Center)

Auro-3D: from 12.1 and 13.1 (base on 6.1 or 7.1 surround)

- 12.1 Auro-3D: 11.1 Auro-3D + Rear Center (6.1)
- 13.1 Auro-3D: 12.1 Auro-3D + Height Rear Center

Dolby Atmos: the composition is as the following, at a minimum

- Three screen speakers
- Surround speakers: individually driven or addressed with traditional 5.1 or 7.1 surround system, including 6 left/right symmetrical zones:
  Left/right side surround (Lss, Rss): on the wired side wall
  Left/right rear surround (Lrs, Rrs): on the rear wall
  Left/right upper surround (Lts, Rts): on the ceiling (overhead)
- Screen or LFE (low-frequency effects) subwoofer (s)

The specific number of loudspeakers depends on the actual size of certain cinemas.

IOSONO 3D: from 5 to 500 loudspeakers

There is no strict requirement for the number of loudspeakers. Only depends on what kind of 3D film sound quality is required. Usually, for a spatial reproduction involving a 360° panorama, covering a much greater area than just a “sweet spot”, around 20 loudspeakers need to be arranged in a circle.

As a matter of fact, no matter which 3D surround format above is chosen, the more accurate the sound localization is, the more loudspeakers will be required. This is an immutable principle. Although, apparatus had already helped us to complete a large part of perceptual work in sound production, separating, intensifying and analyzing

original sound, still they are just a perspective, a sample or a unscramble of the original sounds.

4.2 The significance of ADSS

We have concluded that, the desire to produce 3D film sound is synonymous with film frames. The concept of ADSS was originally proposed for reducing the error between 2D images and spatially accurate sound reproduction. 3D audio has just been integrated into the professional working-flow and is now beginning to reach the audiences. Being a new surround format, the introduction usually pushes for the development of a new aesthetic language, which depends on its peculiarities and the effects it provokes on the audiences. Firstly, 3D audio can provide an “immersive” perception for the audiences, in the sense that it brings more involvement to the listener. Secondly, 3D audio has the ability to perceive better certain separate sources when they are spread in the 3D panorama.

Although the field of ADSS is based on old and well-known principles, now that it is reaching widespread diffusion it is being reviewed with a fresh approach and many of its facets are being discovered and polished. The technology is maturing and the implemented solutions are growing fast. At the same time, the artistic potential is promising but still largely hidden and unexplored.

ADSS provides a huge creative space for all the filmmakers and film theorists. Technologies are always the impetus for the development of film production and film theories. I believe that, in the near future, this new technology could be well and widely applied in film production. Meanwhile, the audience could achieve a better film-viewing experience.

Bibliography

360Beta. 2013. “Cocktail Party Effect”. [Online]
http://baike.so.com/doc/6242214.html
[Retrieved on 12.11.2013]

360 Beta. 2006. “Digital Film technology”. [Online]
http://baike.so.com/doc/6802050.html
[Retrieved on 23.07.2013]

360Beta. 2012. “Economics”. [Online]
http://baike.so.com/doc/5087180.html
[Retrieved on 18.04.2014]

360Beta. 2013. “Louis Lumière”. [Online]
http://baike.so.com/doc/5425254.html
[Retrieved on 15.11.2013]

360Beta. 2013. “Silent Film”. [Online]
http://baike.so.com/doc/1071561.html
[Retrieved on 30.10.2013]

360Beta. 2013. “The Birth of a Nation”. [Online]
http://baike.so.com/doc/5459274.html
[Retrieved on 29.11.2013]

360Beta. 2013. “The Seventh Art”. [Online]
http://baike.so.com/doc/3337369.html
[Retrieved on 20.09.2013]

360Beta. 2013. “Thomas Alva Edison”. [Online]
http://baike.so.com/doc/5368085.html
[Retrieved on 20.12.2013]

Aristotle. 1908, *Poetics*. Oxford: Clarendon Press. pp52.

Auro-3D. 2012. Auro-3D. [Online]
http://www.auro-3d.com/
[Retrieved on 14.04.2014]

Baidu Library. 2009. “Film Space”. [Online]
http://baike.baidu.com/view/357787.htm
[Retrieved on 30.03.2014]

Baidu Library. 2010. “Digital Technology in Film Production”. [Online]
http://wenku.baidu.com/view/2ff6f6323968011ca30091cd.html
[Retrieved on 30.07.2013]

Baidu Library. 2011. “The Development history of Sound Film”. [Online]
http://wenku.baidu.com/view/07d71f7b31b765ce050814cd.html
[Retrieved on 15.07.2013]

Baidu Library. 2011. “The Notion and Method of Film Sound Design”. [Online]
http://wenku.baidu.com/link?url=hHKNvkfLNq2snOZd8OX-rvixSAmlOAksB3IYbbVShXVX2c-CD4TP2GGG6zvmcU2XxoLrHNawy-6oWwgJNWo_eF-wUvO1s8TFjixqrxzMS-G
[Retrieved on 04.03.2014]

Baidu Library. 2011. “The First Film in the World”. [Online]

http://wenku.baidu.com/view/adcb8f46a8956bec0975e332.html
[Retrieved on 14.07.2013]

Baidu Library. 2012. “A Brief history of film production technology”. [Online]
http://wenku.baidu.com/view/3f8f2a83b9d528ea81c779c9.html
[Retrieved on 26.08.2013]

Baidu Library. 2012, “Film: 1960s-1970s”. [Online]
http://wenku.baidu.com/link?url=ucfvNmEedFcLjGqfz4Cdele0hL-tIgpHoj0tJjBbISAfvMBB_5mDACNj7Pblah-J5VD8RkY09Ej2BYAipK4SFqQijWtrJJZwfdGYgAml_Xi
[Retrieved on 20.03.2014]

Baidu, Library. 2013. “Binaural Effect”. [Online]
http://baike.baidu.com/view/416257.htm
[Retrieved on 03.04.2014]

Baidu Library. 2013. “Cochlear Effect”. [Online]
http://baike.baidu.com/view/1285291.htm?fromtitle= 耳 廓 效 应 &fromid=10737504&type=syn
[Retrieved on 05.04.2014]

Baidu Library. 2013, “Cognitive Psychology”. [Online]
http://baike.baidu.com/view/453940.htm
[Retrieved on 22.03.2014]

Baidu Library. 2013. “Fantasia 2000”. [Online]
http://baike.baidu.com/link?url=KJgXISeJEMtyalokWnPR808emXRWZMbtbOdSlspsnb_JC3Vodqy9ocy2_fm7U3kJxWR9-2VGt8WWA71HlVN2ca
[Retrieved on 10.03.2014]

Baidu Library. 2013. “Human’s Perception”. [Online]
http://baike.baidu.com/link?url=YK4ZfzBnLhaJMjXHdPO9ijUNJruts1r65GlMdgQHz0Yai1Iq-zDOWuJbTXJE0TCelanjYZhj0VjRHIs4K8aj3K
[Retrieved on 30.07.2013]

Baidu Library. 2013. “Movie”. [Online]
http://baike.baidu.com/subview/2382/7943995.htm?fr=aladdin
[Retrieved on 30.03.2014]

Baidu Library. 2013, “The Methods of Integrating Sound and Picture”. [Online]
http://baike.baidu.com/view/425099.htm?fr=aladdin
[Retrieved on 23.03.2014]

Baidu Library. 2013. “The Invention and Popularization of Television”. [Online]
http://wenku.baidu.com/link?url=tdZ61qQuFP-uyGYuNQ-gr0NcmGAbRlJ3O1Jf3kX8OsDTGWbp2DQeoMrcj5EWzL9NcOdQSaPdSnjLhdLZbbMwfa6fK04rQpJDqVb5YSYyLiO
[Retrieved on 11.03.2014]

Baidu Library. 2014. “Alcatel-Lucent Bell Labs”. [Online]
http://baike.baidu.com/view/37115.htm?fr=aladdin
[Retrieved on 03.02.2014]

Baidu Library. 2014, “CD”. [Online]
http://baike.baidu.com/link?url=g9GVAFXvRXvMWJrUQ0pvppjNDe3tP2USyZQ9k50tN0dgk7rQY96Q4NtN93WJm8oT
[Retrieved on 24.03.2014]

Baidu Library. 2014. “Charlie Chaplin”. [Online]
http://baike.baidu.com/subview/29591/12123994.htm?fr=aladdin

[Retrieved on 11.02.2014]

Baidu Library. 2014. “Digital Film”. [Online]
http://baike.baidu.com/link?url=VZAnknrdcbshQ2z2g0issW0HZhu9C6LKXoLut-773Ac8SzUJDm62HYe7BqrhZ9aC
[Retrieved on 23.03.2014]

Baidu Library. 2014. “Film and Literary Appreciation”. [Online]
http://wenku.baidu.com/link?url=KSL-pFnYI73lyG0A38djDcorp19qjInCDvv9xcd5qP62WH4FnMUWBziqFHzr_20M4fxUCaUYz_VbyTAi5GdBSAZD-lZp1O3MqC2kxawF4EC
[Retrieved on 12.02.2014]

Baidu Library. 2014. “Perceptual Dimensions of Human Being”. [Online]
http://baike.baidu.com/view/132974.htm
[Retrieved on 30.03.2014]

Baidu Library. 2014, “Rudolf Arnheim”. [Online]
http://baike.baidu.com/link?url=SY-i2riGDEizjef1USZK1VWNaFMucuz5FGs4vekx8PXXWn_WrTTiFFjz3dlRqFJo-UNlyzOtaOE07bizs7qEX_
[Retrieved on 13.03.2014]

Baidu Library. 2014. “Sound and hearing”. [Online]
http://baike.baidu.com/view/261454.htm
[Retrieved on 02.04.2014]

Baidu Library. 2014. “Spatial Dimension”. [Online]
http://baike.baidu.com/view/820048.htm
[Retrieved on 30.03.2014]

Baidu Library. 2014. “The History of 3D film”. [Online]
http://baike.baidu.com/link?url=nULgkRA_7LeLT9HzAlk5u9IFC80hCGcCFvU_qd0wifyXqrDXZO4UCY3jjYVBHqRV
[Retrieved on 12.03.2014]

Baidu Library. 2014, “The History of Dolby”. [Online]
http://baike.baidu.com/view/75049.htm?fr=aladdin
[Retrieved on 23.03.2014]

Baidu Library. 2014. “The Propagation of Sound”. [Online]
http://baike.baidu.com/subview/27063/8076671.htm?fr=aladdin
[Retrieved on 13.04.2014]

Baidu Library. 2014. “Warner Bros Pictures”. [Online]
http://baike.baidu.com/view/3900714.htm?from_id=1931797&type=search&fromtitle=华纳兄弟&fr=aladdin
[Retrieved on 04.02.2014]

Balázs, B. 2011. *Early Film Theory: Visible Man and The Spirit of Film*. Oxford: Berghahn Books Press.

BBC News. 2008. “Celebrating a stereo pioneer: Alan Blumlein”. [Online]
http://news.bbc.co.uk/2/hi/technology/7538152.stm
[Retrieved on 06.03.2014]

Berkhout A.J. 1988. *A Holographic Approach to Acoustic Control, Journal of the Audio Engineering Society*, Vol.36, No.12 December 1988, pp977-995.

Box Office Mojo. 2013. “Avatar”. [Online]
http://www.boxofficemojo.com/movies/?id=avatar.htm

[Retrieved on 3.07.2013]

Chen,F. &Chen, X.X. 1992. *An analysis of 3D Technique of Digital Film*. Advanced Motion Picture Technology.

Craven, P., Law, M., and Travis, C. 2009. *Preceedings of the Ambisonics Symposium*. Microphone arrays using tangential velocity sensors.

Dancyger, K. 2010. *The Technique of Film and Video Editing: History, Theory and Practice*. Focal Press.

David B. 2008. *Poetics of Cinema*, Routledge Chapman & Hall, pp.256-300

DOC88.com. 2008. “Silent Film, technical Summary”. [Online]
http://www.doc88.com/p-608927682768.html
[Retrieved on 26.10.2013]

DOC88.com. 2009. “Comprehend Film”. [Online]
http://www.doc88.com/p-486330621818.html
[Retrieved on 26.08.2013]

DOC88.com. 2013. “3D film: a Real Future”. [Online]
http://www.doc88.com/p-153267246384.html
[Retrieved on 25.07.2013]

Dolby. 2012. “Dolby Atmos”. [Online]
http://www.dolby.com/cn/zh/professional/technology/cinema/dolby-atmos.html
[Retrieved on 15.04.2014]

Douban. 2010. “The 82th Annual Academy Awards List”. [Online]

http://www.douban.com/note/62743134/
[Retrieved on 20.07.2013]

Dou In. 2009, “Perceiving the World”. [Online]
http://www.docin.com/p-20502460.html
[Retrieved on 21.03.2014]

Dou In. 2011. "The Realism Tendency of Hollywood Films”. [Online]
http://www.docin.com/p-164263084.html
[Retrieved on 03.03.2014]

Dou In. 2013. “Ear Recognition Based on the Edge Information of the Auricle Contour”. [Online]
http://www.docin.com/p-733210829.html
[Retrieved on 06.03.2014]

Doc In. 2013. “The Invention and Development of the Film and Television Technology”. [Online]
http://www.docin.com/p-705089172.html
[Retrieved on 05.02.2014]

HC360. 2008, “Dolby Noise Reduction”. [Online]
http://info.audio.hc360.com/2008/06/19102252495.shtml
[Retrieved on 12.03.2014]

IOSONO. 2012. “IOSONO 3D”. [Online]
http://www.iosono-sound.com/home/
[Retrieved on 15.04.2014]

Laborie, A., Bruno, R., Montoya, S. 2003. *A new Comprehensive Approach of*

Surround Sound Recording. 114th Conv. Audio Eng. Soc., Amsterdam, The Netherlands, Mar. 2003, Preprint 5717.

M1905. 2010, “The Chinese Box Office of *Avatar*”. [Online]
http://www.m1905.com/news/20100402/339429.shtml
[Retrieved on 30.03.2014]

Maigoo. 2012. “Film Production Process”. [Online]
http://zhishi.maigoo.com/79733.html
[Retrieved on 25.07.2013]

Millar, S. 1994. *Understanding and Representing Space: Theories and Evidence from Studies with Blind and Sighted Children*. Clarendon Press, Oxford, p15, p262.

Mtimes. 2008. “The Invention Principle of Film Production”. [Online]
http://group.mtime.com/yaocinema/discussion/192806/
[Retrieved on 23.07.2013]

Moray, N. 1969. *Listening and Attention*. Penguin Books, Harmondsworth.

Movie Encyclopedia. 2009, “Siegfried Kracauer”. [Online]
http://www.cinepedia.cn/w/Siegfried_Kracauer/
[Retrieved on 15.03.2014]

Quinlan C. 2011. *Research Methods*, Cengage Learning, Tomas Rennie, pp. 420

Sheng, X.W. 2012. *Research of Technologies of Sound Analysis/Synthesis and Simulation for Flight Simulator*. PHD thesis, Book118.

Sina Entertainment. 2010. “Avatar is sweeping the world”. [Online]

http://ent.sina.com.cn/f/m/avatar/
[Retrieved on 27.06.2013]

Sohu Movie Club. 2010. “The discussion about film *Avatar*”. [Online]
http://club.yule.sohu.com/r-zz0170-356686-0-0-900.html
[Retrieved on 20.07.2013]

Start, E.W. 1997. *Direct Sound Enhancement by Wave Field Synthesis*, thesis, T.U.D.

Tencent Entertainment. 2009, “*Avatar* will be released in January, Director Cameron Talked About this Film”. [Online]
http://ent.qq.com/a/20091106/000109.htm
[Retrieved on 30.03.2014]

Wikipedia. 2014. “American Film”. [Online]
http://zh.wikipedia.org/wiki/AmericanFilm
[Retrieved on 05.01.2014]

Yesky. 2012. “A new era for 3D film and 4K image”. [Online]
http://dh.yesky.com/59/33973059.shtml
[Rerieved on 18.07.2013]