Wave field synthesis: The future of spatial audio

RISHABH RANJAN AND WOON-SENG GAN

We all are used to perceiving sound in a three-dimensional (3-D) world. In order to reproduce real-world sound in an enclosed room or theater, extensive study on how spatial sound can be created has been an active research topic for decades. Spatial audio is an illusion of creating sound objects that can be spatially positioned in a 3-D space by passing original sound tracks through a sound-rendering system and reproduced through multiple transducers, which are distributed around the listening space. The reproduced sound field aims to achieve a perception of spaciousness and sense of directivity of the sound objects. Ideally, such a sound reproduction system should give listeners a sense of an immersive 3-D sound experience. Spatial audio can primarily be divided into three types of sound reproduction techniques, namely, loudspeaker stereophony, binaural technology, and reconstruction using synthesis of the natural wave field [which includes Ambisonics and wave field synthesis (WFS)], as shown in Fig. 1(a).

The history of spatial audio dates back to the late 1800s, with the very first invention being the gramophone used in sound recording. As shown in the timeline in Fig. 1(b), there have been major advancements in terms of both technical and perceptual aspects in the last century. Spatial sound systems have evolved over the years from a two-channel stereo system to a multichannel surround sound system. These surround systems are not only limited to cinemas and auditoriums but are also being adapted in home entertainment systems. Conventional headphones, which employ a pair of small emitters, aim to produce high-quality sound close to the ears, and they do not need to account for inaccuracies due to surroundings in contrast to loudspeakers. Nowadays, multiple emitters are embedded inside the ear cup to create a virtual surround sensation in 3-D surround headphones. Modern electroacoustic systems have improved significantly with new functionalities to adapt or correct the sound field in a given room acoustic. Toward the end of the 19th century, new reproduction techniques like Ambisonics and WFS [see Fig. 1(b)], which use the principle behind physical sound wave propagation in air and thus provide true sound experience in any environment, were introduced to overcome the limitations of stereo systems.

Two-channel stereophony is the oldest and simplest audio technology, which has been progressively extended to multichannel stereophony systems, through 5.1, 7.1, 10.2, and 22.1 surround sound systems. [Note that in the $x.y$ representation, x indicates the number of full

Digital Object Identifier 10.1109/MPOT.2012.2212051
Date of publication: 26 March 2013

0278-6648/13/$31.00©2013IEEE

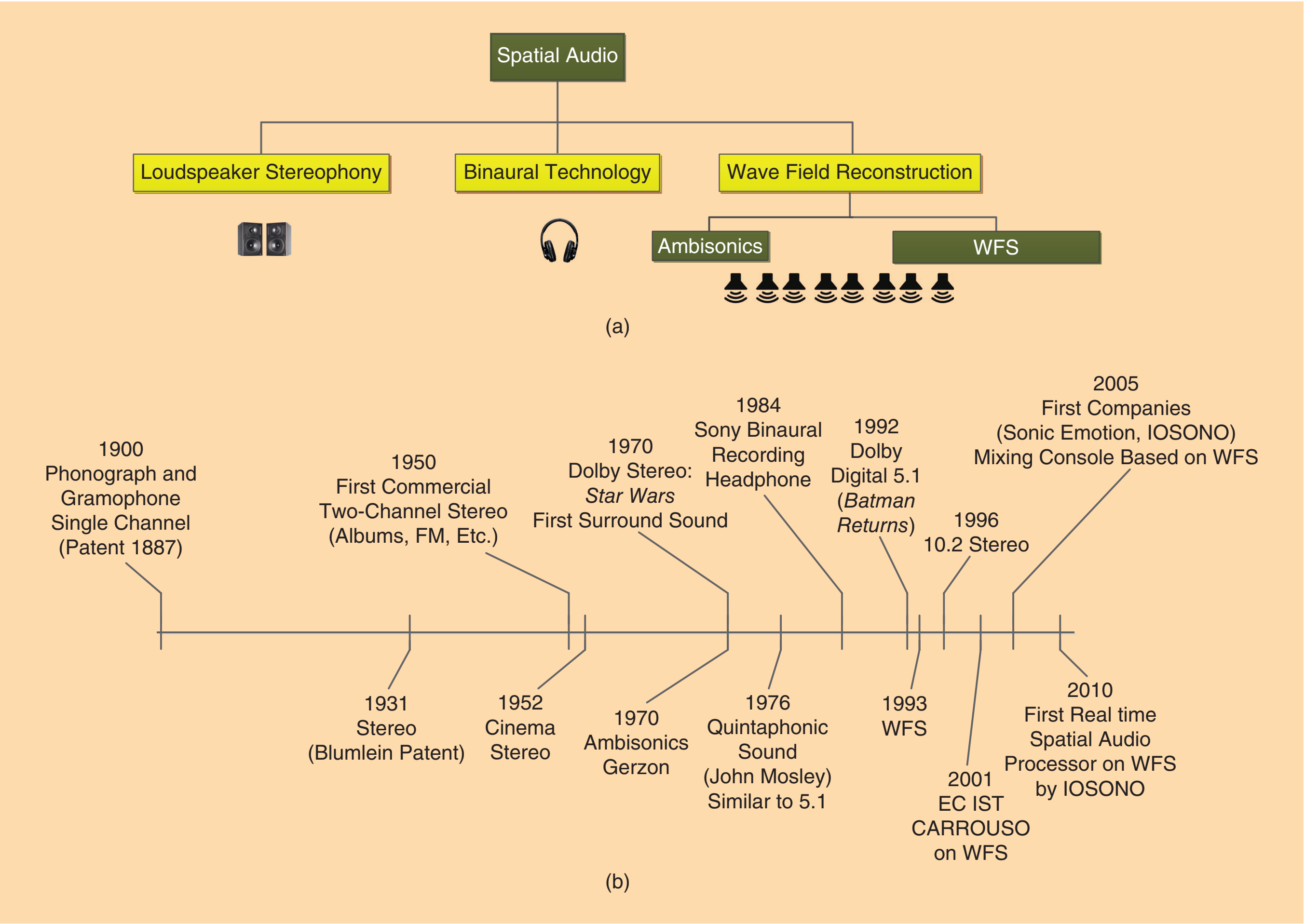

Fig. 1 (a) The classification of spatial audio and (b) the timeline of evolution of spatial audio.

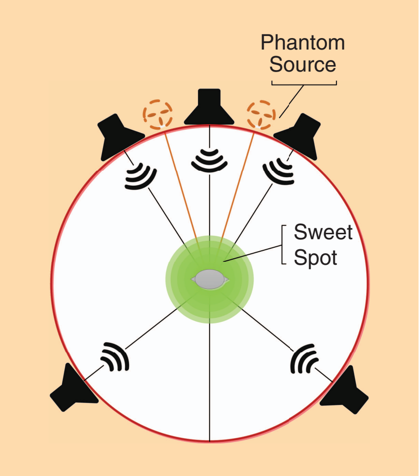

Fig. 2 A typical 5.1 stereo.

bandwidth channel and y indicates the number of low frequency channels, known as low frequency effects (LFE) subchannel.] These multichannel systems have been widely used in cinema, home entertainment, and gaming to create an immersive surround sound experience. Figure 2 shows a typical setup of a 5.1 stereo system with three front and two rear loudspeakers. It uses rear speakers to enhance the ambient sound quality and a center speaker to enhance the frontal perception. The disadvantages of a multichannel stereophony system are the localization of phantom sources and the sweet spot. In other words, a phantom source can only be located along the lines connecting two loudspeakers, and listeners will only be able to experience the best surround sound effect at the sweet spot or focal point of all multichannel speakers.

Binaural technology is another approach to reproduce sound signals naturally. Binaural technology consists of the recording, as well as the reproduction, of natural sound scenes at the two ears. Sound signals are recorded using a pair of microphones positioned inside the ears of a dummy head or inside the ear canal of an actual human listener. A binaural recording set up using a dummy head is shown in Fig. 3.

Recorded sounds can be reproduced accurately at the ears by filtering the source signals using acoustic transfer functions between the source location and both ears, which are popularly known as the head-related transfer function (HRTF). HRTF contains three important cues of interaural time differences (ITDs), interaural level differences (ILDs), and spectral cues (SCs). These cues are essential for us to correctly localize and perceptually visualize the sound scenes. As a result, binaural reconstruction can produce excellent spatial awareness and sound color under given circumstances. Binaural signals can be played back via a loudspeaker or headphones. Direct reproduction of binaural signals through loudspeakers suffers from the problem of crosstalk between the left and right ear signals. A crosstalk cancellation system must be inserted between the loudspeakers and binaural processing in order to achieve an accurate 3-D audio display. Binaural reproduction using headphones is the most efficient way, as the signals are correctly reproduced at each ear and do not suffer from any distortion due to environments. But there have been several inherent limitations of binaural sound reproduction through headphones, which includes front-back

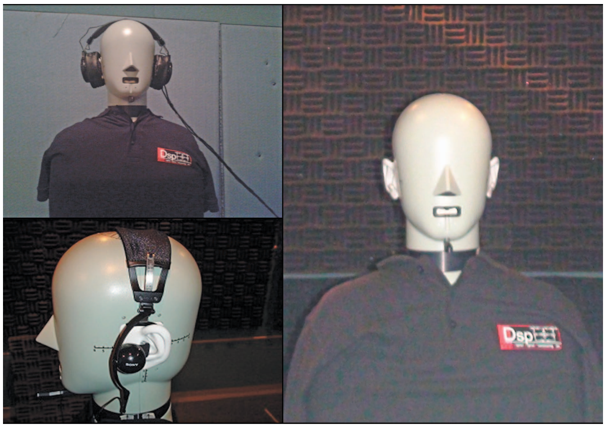

Fig. 3 A binaural recording system at Nanyang Technological University, Singapore.

WITH THE HELP OF LOUDSPEAKER ARRAYS, BOTH AMBISONICS AND WFS ARE ABLE TO SYNTHESIZE A NATURAL SOUND ENVIRONMENT IN AN ENLARGED LISTENING AREA WITH PERFECT SOUND SOURCE LOCALIZATION.

confusions, in-head localization, and incorrect perception of the elevation of virtual sound sources. These limitations of binaural technology are due to inabilities of human beings to disambiguate between ITD, ILD and SC cues.

Both multichannel stereophony and binaural technology are widely used in cinemas, auditoriums, home entertainment systems, and headphones playback. However, their inherent limitations, which rely mainly on psychoacoustic principles in creating a fully immersive environment, have inspired researchers to look into more natural ways of reproducing 3-D sounds. The two technologies, which utilize the concept of natural propagation of wave fields, are Ambisonics and WFS. In contrast to stereophony and headphones playback, the wave-field-based approach uses holographic principles to synthesize a true sound field rather than relying on psychoacoustic principles for recreating sound scenes. With the help of loudspeaker arrays, both Ambisonics and WFS are able to synthesize a natural sound environment in an enlarged listening area with perfect sound source localization.

Ambisonics was first proposed by Gerzon in 1970, while Berkhout invented WFS in 1988. Although both approaches follow the same basic principles, the difference lies in the detailed mathematical derivations. The main advantage of Ambisonics is that it can synthesize a sound field for any number of loudspeakers arranged in an arbitrary shape. According to Francis Rumsey, "Ambisonics is mainly a collection of elegant principles and signal representation forms rather than a particular implementation." However, it has yet to gain acceptance in the commercial field and extensive research is being carried out on various derivatives of higher-order Ambisonics to improve its commercial feasibility.

Principle of wave field synthesis

A WFS-based reproduction system aims to accurately synthesize the sound field within the entire listening space,

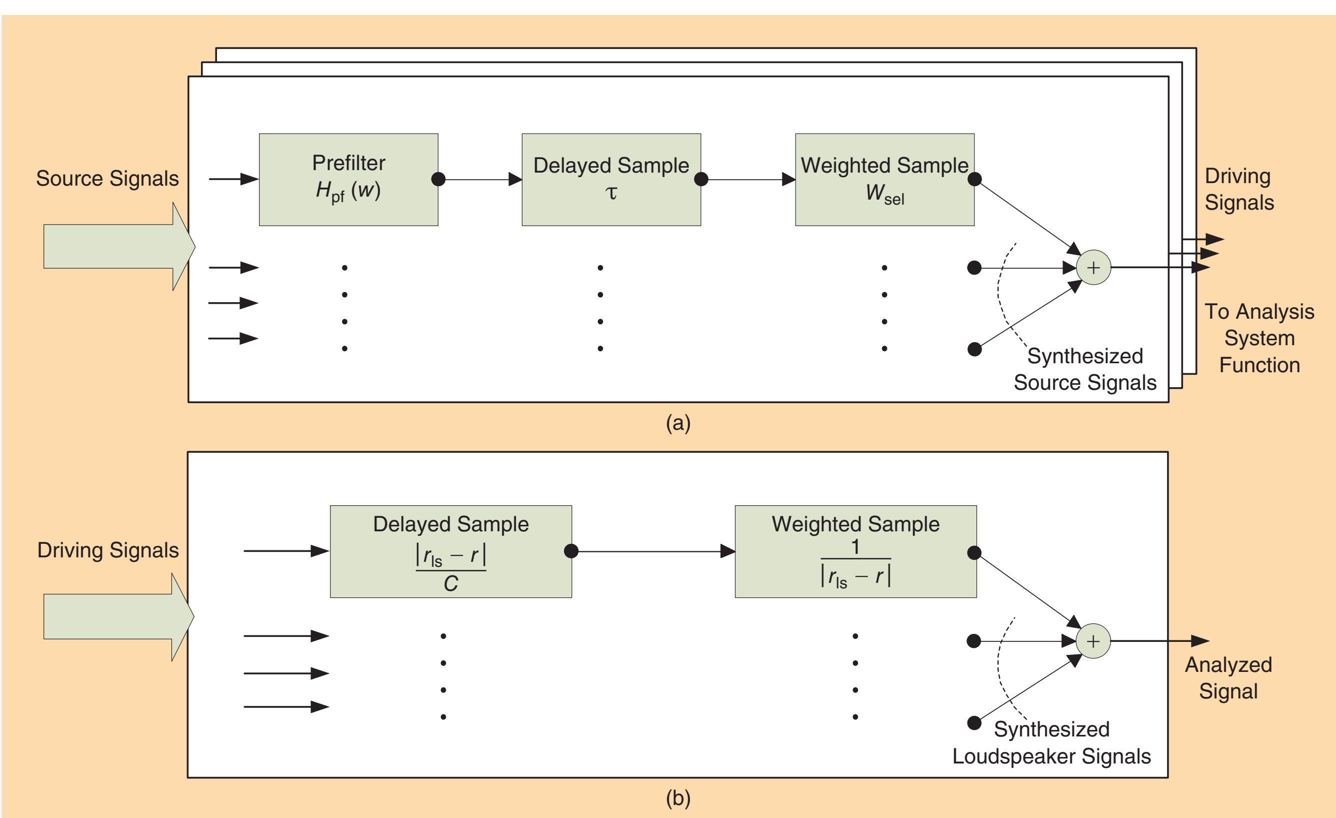

Fig. 4 A block diagram of a WFS reproduction system. (a) Synthesis system function. (b) Analysis system function.

i.e., the sweet spot is everywhere in the listening room. WFS duplicates the sound field generated by primary sources (real sources that produce the sound) with the help of loudspeaker arrays acting as secondary sources (sources responsible for reproduction of sound produced by primary sources) in an enlarged listening area. Source localization is possible anywhere in the physical space, which is only limited by the extent of the visible area covered by the configuration of loudspeaker arrays and the listener. Also, unlike binaural listening, virtual source does not move with the listener's movement. Listeners feel as if they are in a real environment and sounds appear to come from where they are meant to be. The main objective in developing a WFS reproduction system is to obtain driving signals (loudspeaker signals) where the primary source signals are processed using wave propagation theory. A typical WFS system can be divided into two subsystems: a synthesis system function and an analysis system function, as shown in Fig. 4. Source signals are filtered, delayed, and weighted to compute all the driving signals. This process of computing driving signals from primary source signals is termed "synthesis system function." These driving signals act as inputs to the loudspeaker array. The subsequent process of the WFS system is termed "analysis system function," which analyzes the reproduced sound field at different listening positions by computing the sound pressure signals due to contributions from weighted and delayed driving signals. A detailed explanation of the two subsystems is provided later in this article.

The fundamental theory of WFS has been derived from the concept of Huygens principle. This principle states that a spherical wave front (radiated by a primary source) is formed by continuous infinite secondary sources, the source strength of which determines the successive wave front and so on. In WFS, these secondary sources are replaced by loudspeaker arrays, which eventually reproduce the replica of the original sound field retaining the physical properties (temporal and spatial) of sound waves. The Kirchhoff-Helmholtz (KIH) integral forms the mathematical basis for WFS, which applies the Huygens principle. The KIH integral states that a sound field at any point can be calculated if the pressure and pressure gradient (due to primary or virtual sources) at the boundary of the source free volume enclosed by a surface are known. In other words, 3-D enclosed volume is surrounded by an infinite

THERE ARE TWO REAL-LIFE SCENARIOS WHERE SOUND REPRODUCTION IS EMPLOYED: 1) REAL OR PRIMARY SOURCES EXIST, FOR EXAMPLE, IN LIVE PERFORMANCES, AND 2) PRIMARY SOURCES SIGNALS ARE RECORDED FOR REPRODUCTION IN FUTURE, LIKE IN CINEMAS, TELEVISION, OR RECORDED EVENTS.

number of monopole and dipole secondary sources on its surface, which in turn reproduce the original sound field.

Two assumptions that have been made to arrive at the KIH integral are inhomogeneous media (like air) and free field condition for wave propagation. Rayleigh proposed two modifications to the KIH integral, for it to be applied to real scenarios.

- It is practically impossible to have an infinite continuous array of loudspeakers on the surface of enclosed volume. It is shown by Rayleigh that the surface can be degenerated to a plane of loudspeakers separating the listening area from the source area. Later, it was proposed that it is also possible to move virtual sources in front of the loudspeaker array.
- In the KIH integral, the combined monopole and dipole secondary sources cancel out the undesirable wave fields propagating outside the enclosed surface but sum up inside the enclosed space. Since we intend to correctly reproduce a sound field inside the volume, either monopole or dipole can be eliminated from the KIH integral. The above process results in a non-zero sound field outside the listening volume.

The above propositions led to the introduction of two famous Rayleigh integrals I and II, which state that sound pressure at any point on one side of the loudspeaker plane can be synthesized from sources on the other side of the loudspeaker plane. Rayleigh integral I uses monopole loudspeakers as secondary sources with pressure gradient as signal strength (of primary source) at surface of the plane. Similarly, the Rayleigh II integral uses dipole loudspeakers as a

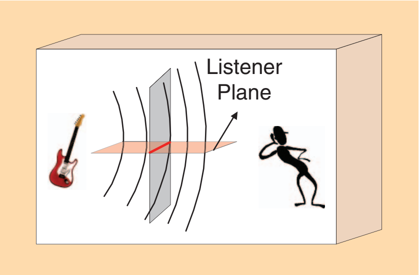

Fig. 5 WFS in a practical scenario.

secondary source with pressure as the signal strength.

Berkhout introduced the concept of a virtual source, i.e., a source that is perceived by a listener when the sound is reproduced. There are two real-life scenarios where sound reproduction is employed: 1) real or primary sources exist, for example, in live performances, and 2) primary sources signals are recorded for reproduction in future, such as in cinemas, television, or recorded events. In the former scenario, the virtual source coincides with the real source, and in the latter, WFS aims to reproduce the virtual source as close as possible to the real source. Furthermore, two approximations have been proposed to practically realize the WFS reproduction system. First, the loudspeaker plane is reduced to a line array configuration on the horizontal plane, which is due to the infeasibility of covering the entire vertical plane with loudspeakers. Second, since loudspeakers cannot be infinite in numbers and are always discrete, the continuous line array is reduced to a finite discrete array with uniform spacing between the loudspeakers. The above approximations are shown in Figures 5 and 6, which describe the two-dimensional (2-D) reproduction of

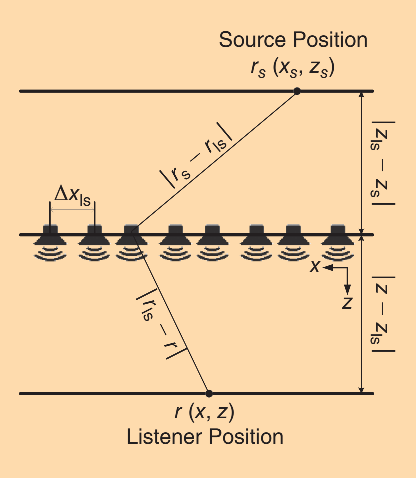

Fig. 6 Geometry used in WFS formulations.

the sound field that can be perfectly reproduced on the listener plane.

Driving signals are derived using stationary phase approximation from Rayleigh integral I by Vogel in 1993 with monopoles forming the secondary source array, i.e., 3-D to 2-D approximation. The equation for driving signal is shown in Fig. 7(a). Driving signals can be calculated in the time domain by pre-filtering the source signal followed by weighted and delayed sample of the filtered source signal. The frequency-dependent prefilter term and the distance-dependent correction factor term are crucial in driving signal equation and compensate for planar-to-linear array reduction. Spors further modified the equation for driving signal to introduce a selection criterion or window function to suppress the undesired reflections from the side of the loudspeaker array so as to minimize the error in the reproduced wave field. Pressure at the listener position is given by the equations shown in Fig. 7(b). Green function in the figure represents the radiation of a monopole source and, thus, is used in determining sound pressure in analysis system function. Pressure at the listener position is due to the contribution from delayed and weighted samples of all the driving signals. The geometry for the equations is shown in Fig. 6.

It should also be noted that since driving signal depends on perpendicular distance between the loudspeaker array and the listener, reproduced sound is accurate only on a reference line, usually chosen in the center of the listening area and parallel to the loudspeaker array (also called the "sweet line" by de Vries). Synthesis system function requires source parameters (source signals, positions, and orientations) and loudspeaker parameters (loudspeaker positions, orientation, and number of loudspeakers) as inputs. Similarly, analysis system function requires driving signals, listener position, and loudspeaker parameters as input for analyzing a sound field in the listener space. A block diagram for synthesis system function and analysis system function is shown in Fig. 4.

$$D_{2.5D}(r_{ls}, w) = \underbrace{\left(S(w)\cdot\sqrt{\frac{jk}{2\pi}}\right)}_{\substack{\text{Prefiltering}\\ H_{pf}(w)}} \cdot \underbrace{\sqrt{\frac{|z - z_{ls}|}{|z - z_s|}}}_{\substack{\text{Correction}\\ \text{Factor}\\ f_{cf}}} \cdot \underbrace{w_{sel}(x, z)}_{\substack{\text{Weighting}\\ \text{Factor}\\ W_{sel}}} \cdot \underbrace{e^{-jw\tau(|r_s - r_{ls}|)}}_{\text{Delaying}}$$

(a)

$$P(r, w) = -\sum_{n_{ls}} D_{2.5D}(r_{ls}, w) \cdot \underbrace{\frac{e^{-jw|r_{ls} - r|}}{|r_{ls} - r|}}_{\text{Green Function}} \cdot \Delta x_{ls}$$

$$P(r, t) = -\frac{1}{4\pi}\sum_{x_0} \underbrace{\frac{1}{|r_{ls} - r|}}_{\substack{\text{Weighting}\\ \text{Factor}}} \cdot \underbrace{d_{2.5D}\left(r_{ls}, t - \frac{|r_{ls} - r|}{c}\right)}_{\text{Delayed Samples}} \cdot \Delta x_{ls}$$

(b)

Fig. 7 (a) The driving signal equation for synthesis system function and (b) the sound pressure equation for analysis system function.

Practical constraints and solutions

As discussed in the previous section, it is not practically realizable and cost effective to place loudspeakers everywhere in a closed space. Moreover, the computational power of a typical WFS processing engine is largely dependent upon the number of loudspeakers and the complexity of auditory scenes.

WFS formulations for driving a signal equation work well only for the reproduction on the horizontal plane (listener plane) because of the approximation to the line array. Since the two ears are located on the horizontal plane, it would be sufficient to assume that the sound perceived will be natural to us. A reproduced sound field is accurate only at the sweet line and thus resulting in amplitude error, which is measured as deviation from sound pressure on the sweet line (in dB). Because of the 2-D reproduction on the horizontal plane, virtual sources are not correctly perceived in the vertical plane. But with the advent of 3-D audio-visual contents, like gaming, videos, and 3-D movies, where elevation perception is of utmost importance, it is required to find a solution that emulates the 3-D plane reproduction. Recently, Montag proposed a multiple array line of loudspeakers in the vertical plane to extend the traditional WFS to 3-D reproduction.

Mathematically, we can derive a driving signal for any arbitrary configuration of closely spaced loudspeaker array, but in reality it is near to impossible to have a spacing less than 1 cm. This is due to the reduction of an infinite continuous line array to a finite discrete array and suffers perceptual quality degradation to some extent. As a result of the reduction to a finite continuous line array, diffraction effects and additional trailing waves (also called "shadow signals" by de Vries) are observed in the sound field derived using analysis equation. This effect is also known as the truncation effect, which originates from loudspeakers at the extremes of an array. Perceptually, this can cause a slight coloration effect and echo perception depending on the time difference between desired response and shadow signals. Tapering is a technique used to smoothen the truncation effects, where loudspeakers positioned on the edges are given less weighting but at the cost of a reduced effective array length. An effective array length can be increased by using N-shape arrays, with tapering applied at the two extremes, like shown in Fig. 8.

A finite continuous array is reduced to a finite discrete array by applying sampling in spatial domain resulting in

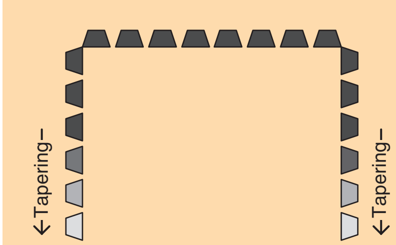

Fig. 8 Tapering applied at extremes to reduce diffraction effects.

spatial aliasing, which is similar to aliasing in the frequency domain. It is easier to analyze the aliasing artifacts in the wave number domain, which is obtained by taking Fourier transform of sound signals in spatial domain. However, WFS is correctly achieved only up to a corner frequency known as "spatial aliasing frequency." In spite of the inaccurate synthesis of the sound field, it has been found that a reasonable amount of deviation from aliasing criteria does not significantly degrade the perceptual quality. It has been experimentally verified that a separation of 10–30 cm between the loudspeakers is appropriate for reproduction purposes. Spatial aliasing is the most critical of all the WFS artifacts as it leads to a distortion in frequency response and the physical sound field. A number of methods have been proposed in literature to minimize spatial aliasing effects. Spatial bandwidth reduction uses the notion of directive sound sources to reduce the interference between loudspeakers. Another method is to randomize the high-frequency content over the loudspeaker array and, thus, reduce the periodicity in spatial aliasing artifacts. In recent years, researchers have analyzed the spatial aliasing artifacts by deriving several aliasing criterions, which not only depend on the spacing between loudspeakers but also on source directivity and listener positions. In one of the recent works by Corteel, spatial aliasing frequency is increased with the help of dynamic selection of the subpart of the loudspeaker array to target reproduction within a preferred listening area.

Another limitation of Rayleigh theory, which states that the source (nonfocused source) can only exist behind the loudspeaker array, has been resolved by the introduction of focused sources. A focused source can be perceived in front of the array, i.e., in the listener space. The only constraint with focused source reproduction is that the listener area is reduced and the listener is not permitted to sit in between the array and the focused source. Both focused and nonfocused sources are crucial in recreating the immersive sound field around the listener. The listener can feel the depth of the source, but it requires the entire listener space to be surrounded by a closed configuration of loudspeaker arrays.

All loudspeakers in practice possess some directivity pattern in contrast to the ideal monopole secondary sources. This implies that the conventional driving signal equation holds only for the ideal monopole conditions. De Vries derived that the driving signal for a linear array can be adapted to loudspeakers with arbitrary directivity characteristics. It should also be noted that traditional equations for WFS assume free-field conditions, and room reflections must be accounted while dealing with real room simulations. A mirror image source model is commonly used for the analysis of room reflections.

AS A RESULT OF THE INCREASED INTERESTS AND PARTICIPATIONS IN WFS, SEVERAL RESEARCH GROUPS AND R&D LABS COLLABORATED TO STANDARDIZE A WFS FORMAT, WHICH LED TO THE START OF THE EUROPEAN UNION INFORMATION SOCIETY TECHNOLOGIES "CARROUSO" PROJECT IN 2001 AND COMPLETED BY 2003.

Evolution of WFS

Since the introduction of WFS by Berkhout in 1988, WFS has come a long way over the last two decades and is now playing a vital role in spatial audio reproduction technology. Berkhout started the research on the WFS-based system at Delft University and laid the foundation for further developments. He was supported by fellow researchers, in particular, De Vries, Vogel, Start, and others, in the following years. The first WFS-based practical laboratory setup, which consists of 48 channels with DSP processors, was developed at Delft University in 1993 and later extended to the university's auditorium.

Berkhout's work was followed up by many other prominent research groups and many WFS-based setups were installed in various places, including cinemas, lecture halls, and concert halls. Until the late 1990s, most of the research was carried out at universities, mainly focused on developing mathematical formulations of WFS equations and also practical measurements based on various configurations (linear, circular, and rectangular) of loudspeaker arrays. As a result of the increased interests and participations in WFS, several research groups and R&D labs collaborated to standardize a WFS format, which led to the start of the European Union (EU) Information Society Technologies (IST) "CARROUSO" ("Creating, Assessing, and Rendering in Real Time of High-Quality Audio-Visual Environment in MPEG-4 Context") project in 2001 and completed by 2003. The main goal of the CARROUSO project was to develop a new technology that can record, encode, transmit, and reproduce the sound field recorded at a virtual or a remote place. This project prompted researchers as well as commercial markets based on spatial audio in different parts of the world to focus on WFS with the goal of creating potential applications in spatial audio systems that could potentially replace multichannel surround sound systems placed in cinemas, live events, or home theater systems. The successful completion of the CARROUSO project led to the emergence of two new companies, IOSONO and Sonic Emotion, which are aimed to provide the services and solutions for installations of 3-D audio systems based on WFS. IOSONO is supported by Fraunhofer IDMT Research Institute, while Sonic Emotion is cofounded by Renato Pellegrini. Both of these companies have played a significant role in the success of the CARROUSO project. They are now the major providers of WFS-based products for the consumer market, as well as research applications in spatial audio systems.

IOSONO recently launched a spatial audio processor to control any kind of loudspeaker arrangement, room geometry, and listener numbers. Sonic Emotion has manufactured 3-D audio chips based on their own patented technology employing WFS, psychoacoustics, and others. In 2011, Haier launched a 3-D sound bar using this chip, which claims to create the unique sound experience that can replace the current home theater systems.

In 2008, Spors and his team at Deutsche Telekom in Berlin revisited WFS theory and also proposed modified driving equations addressing arbitrarily shaped loudspeaker arrays for three-dimensional sound reproduction. They also installed a practical WFS setup of 56 channels of circular loudspeaker array. Furthermore, they have developed a generic spatial audio renderer framework for real-time audio processing, which is very useful for sound reproduction in real time. This versatile software

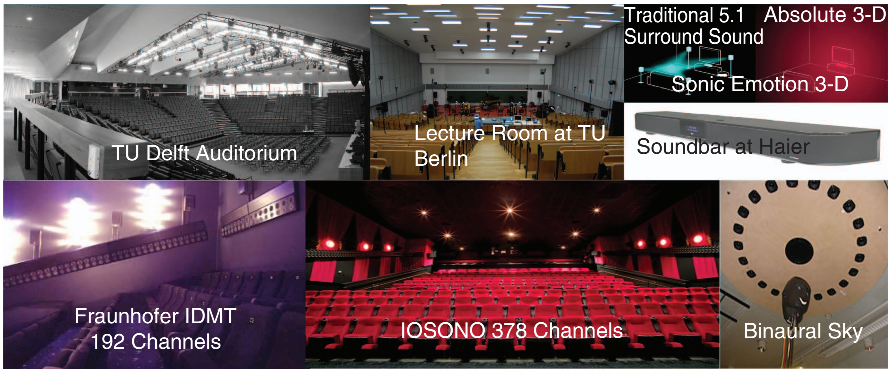

Fig. 9 A look at various WFS developments.

allows the rendering of several rendering modules like WFS, binaural, Ambisonics, and virtual amplitude-based panning. Researchers at IRT, Germany, developed a novel system known as the "Binaural sky," which uses WFS technology for binaural sound reproduction. The latter system consists of an overhead circular array of loudspeakers and synthesizes focused sources using a head tracking system. Fig. 9 shows various WFS set ups installed at various universities and auditoriums.

Future trends

In the last few years, WFS has increasingly become more popular in commercial deployment. WFS-based reproduction systems are now readily accepted as the most optimal way of reproducing spatial sound. Several companies have already started the mass commercialization of WFS installations in public places. Recently, the Game of Life Foundation developed the world's first transportable WFS setup and demonstrated at Amsterdam in 2011. A similar setup was earlier demonstrated at the 124th AES convention on the eve of 20 years of WFS in 2008. Until now, we have mainly seen large-scale installations of WFS in large public places. Many people now appreciate the immersive environment reproduced by a WFS-based system in such places. In recent years, researches have started to look into the small-scale applications of WFS, i.e., targeting a small group of audience. Some small-scale WFS applications include virtual reality, 3-D gaming environments, and video conferencing. WFS may eventually replace the current surround systems in home entertainment in the near future.

A major hurdle for the use of WFS technology in such small-scale applications is that these systems require a minimal number of loudspeakers but at the same time, the sweet spot needs to be maximized. Since a typical WFS system requires a large and costly set up of loudspeaker arrays, it is an open research problem to devise a trade-off between the number of loudspeakers and the size of the sweet spot area. Corteel from Sonic Emotion has recently proposed a new methodology to employ fewer loudspeakers while increasing the spatial aliasing frequency using the focused sound reproduction in a "preferred listening area." But for WFS to enter into our homes, all the recording should be carried out in WFS compatible format (as explained in CARROUSO project) before distributing them. Only then, we will be able to take full advantages of WFS in immersive 3-D sound reproduction.

WFS-BASED REPRODUCTION SYSTEMS ARE NOW READILY ACCEPTED AS THE MOST OPTIMAL WAY OF REPRODUCING SPATIAL SOUND. SEVERAL COMPANIES HAVE ALREADY STARTED THE MASS COMMERCIALIZATION OF WFS INSTALLATIONS IN PUBLIC PLACES.

In this article, we provied an overview of the principles of WFS and presented some of the key research work and commercial products from the past two decades. We also highlighted some practical limitations and technical challenges of WFS-based sound reproduction systems. Increasingly, we are witnessing WFS as one of the key spatial audio technologies in next-generation home entertainment systems.

Acknowledgments

This work was supported by the Singapore Ministry of Education Academic Research Fund Tier-2, under research grant MOE2010-T2-2-040.

Read more about it

• M. A. Gerzon, "Periphony: With-height sound reproduction," *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 2–10, 1973.

• A. J. Berkhout, "A holographic approach to acoustic control," *J. Audio Eng. Soc.*, vol. 36, no. 12, pp. 977–995, 1988.

• F. Rumsey, *Spatial Audio*. Woburn, MA: Focal Press, 2001, ch. 1.

• S. Spors, R. Rabenstein, and J. Ahrens, "The theory of wave field synthesis revisited," in *Proc. 124th AES Conv. Audio Engineering Society*, 2008, vol. 124, pp. 17–20.

• M. Montag and C. Leider, "Wave field synthesis by multiple line arrays," in *Proc. 131st AES Conv. Audio Engineering Society*, 2011.

• D. de Vries. (2009). *Wave Field Synthesis* (AES Monograph). [Online]. Available: http://www.aes.org/publications/monographs/

• E. W. Start, "Direct sound enhancement by wave field synthesis," Ph.D. dissertation, Department of Imaging Science and Technology, Delft Univ., Techol., Delft, The Netherlands, 1997.

• D. de Vries, E. W. Start, and V. G. Valstar, "The wave field synthesis concept applied to sound reinforcement: Restrictions and solutions," in *Proc. 96th AES Conv. Audio Engineering Society*, Amsterdam, The Netherlands, 1994.

• P. Vogel, *Application of Wave-Field Synthesis in Room Acoustics*, Ph.D. dissertation, Department of Imaging Science and Technology, Delft Univ. Technol., Delft, The Netherlands, 1993.

• E. Corteel and R. Pellegrini, "Wave field synthesis with increased aliasing frequency," in *Proc 124th Conv. Audio Engineering*, 2008, pp. 225–226.

About the authors

Rishabh Ranjan (rishabh001@ntu.edu.sg) is currently pursuing his Ph.D. degree in electrical and electronic engineering at Nanyang Technological University.

Woon-Seng Gan (ewsgan@ntu.edu.sg) is an associate professor of electrical and electronic engineering at Nanyang Technological University. He is a Senior Member of the IEEE.