inter.noise
HAMBURG 2016

Application of Wave Field Synthesis in Virtual Acoustic Engineering

Jakob BERGNER[1,2]; Tobias CLAUSS [2]; Albert ZHYKHAR[2]; Christoph SLADECZEK[2]; Sandra BRIX[2]

[1] Ilmenau Universtiy of Technology, Ilmenau, Germany
[2] Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

ABSTRACT

State-of-the-art product design processes are driven by virtual reality (VR) technologies. However VR technologies these days are often limited to visualization only. Due to the lack of robust psychoacoustic models that predict parameters like pleasantness of audio signals, a plausible auralisation of product sound is mandatory.

Modern sound reproduction techniques, such as wave field synthesis (WFS), can help us to embed an appropriate acoustical environment in virtual engineering. Possible use cases are noise reduction, sound design, sound branding, product presentation as well as soundscape planning. WFS is a sound reproduction technique for physical synthesis of a virtual sound field. In contrast to stereo or surround sound, it is possible with WFS to overcome the "sweet spot" problem which is essential for interactive multi-user VR systems. Currently this technology is mainly used in entertainment applications.

This paper introduces a concept and a prototypical implementation of an object-based acoustical environment for virtual engineering. It is designed for the auralisation of both single sources as well as complex sound scenes by means of up-to-date wave field synthesis technologies. The presented system covers latest developments in spatial audio reproduction, e.g. auralization of directional sources, interactive real-time room acoustic simulation and an intuitive user interface.

Keywords: auralization, virtual engineering, I-INCE classification: 73.5

1. INTRODUCTION

Today, the process of product development is more and more shifted towards virtual prototyping. Using modern CAD (Computer Aided Design) and CAE (Computer Aided Engineering) tools enables engineers to validate multiple design alternatives or explore product properties on virtual prototypes, to replace physical products in order to reduce time and design costs [1].

The methods and tools of virtual engineering are centered on the integration of its components within a computer-generated environment. However, this user-centered virtual environment so far is mainly focused on using visualization aspects. Depending on the task to be solved, e.g. design decisions, structural dynamic or heat distribution simulations all resulting data is visualized using stereoscopic techniques.

Analysis of market trends in industrialized countries show that the number of competing products which becoming more similar or comparable is increasing. That means a larger product choice exists for a potential consumer making the decision for a specific product more difficult. Often such a decision is based on subtle cues. One of these major cues is sound which means that for a manufacturer the acoustic behavior of a product becomes more important [2]. This aspect is twofold:

- The industry is concerned to control and manage noise emission from products
- Product designers pursue the goal of creating brand sound

Looking at the acoustic product development process today, it is obvious that a lot of companies already take care of acoustic properties of their products. However, the development of acoustic treatment is often done by trial and error which always requires to build real prototypes. On the other side a lot of techniques for vibro-acoustics and airborne sound behavior simulation exist [3], which open possibilities for virtual acoustic product engineering. One of the main obstacles so far was the lack of sound reproduction technology that allows a natural auralization

[1] jakob.bergner@tu-ilmenau.de
[2] givenname.familyname@idmt.fraunhofer.de

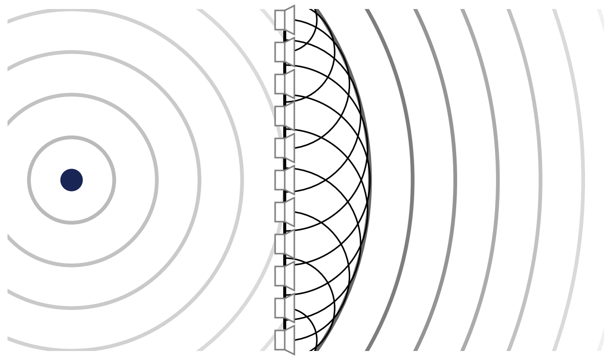

Figure 1: Concept of the wave field synthesis technology. An array of loudspeakers is individual driven so that the superposition of all sound waves radiated by the speakers generates the sound field of a virtual source.

of simulated or recorded sound fields. As product engineering is always a collaborative process, also the sound reproduction technology need to support multiple users [4].

2. WAVE FIELD SYNTHESIS

For existing sound reproduction technologies like stereo or surround sound, an essential disadvantage is the existence of a so-called "sweet spot". This term describes a single listening position for which the spatial reproduction is correct. If the user is moved to a different position the perception of the audio scene will change or completely flip to one of the speakers. In this case all audio sources seems to be placed inside that loudspeaker. The reason for this is, that these techniques are based on psychoacoustic models which use simple loudness panning to create the illusion of a phantom source. The speaker signals can only be adjusted for a single loudspeaker setup. If the audio scene mixed for a dedicated setup is reproduced on another setup where the loudspeakers are positioned differently, the whole spatial scene is perceived incorrectly. As this approach requires to store ready-mixed loudspeakers signals, the technology is called channel-based audio reproduction.

Wave field synthesis is a spatial sound reproduction technology invented at Delft University of Technology (Netherlands) in the 1990s [5]. The basic concept of this approach is to use an array of loudspeakers for the physical creation of virtual source sound fields. For this each speaker of the array is driven with an individual signal, so that a correct sound field of a virtual source is generated by superposition of the sound waves radiated from the speakers. This concept is depicted in Figure 1. In contrast to the channel-based audio reproduction this approach is called object-based. While the channel-based approach only requires the playback of ready-mixed loudspeaker signals, for the object-based approach the loudspeaker signals have to be calculated in real-time using a mathematical model of the sound source object.

The synthesis of a virtual source sound field using a linear array of loudspeakers can be mathematically described by integrating over a plane surface S of monopole sources each driven by a signal $Q(r,\omega)$. This is given by the first Rayleigh integral yielding:

$$P(\vec{r}_R,\omega) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \underbrace{\frac{1}{2\pi}\left(-\vec{n}\cdot\vec{\nabla}P(\vec{r},\omega)\right)}_{Q(\vec{r},\omega)} \frac{e^{-jk\Delta r}}{\Delta r} dxdz, \tag{1}$$

with the corresponding geometry depicted in Figure 2 [6]. As $Q(r,\omega)$ directly depends on the normal pressure gradient the mathematical model of the sound object influences the complexity of synthesis operator. As the loudspeaker signals need to be determined interactively commercial available wave field synthesis systems often only allow the reproduction of simple sound source types like point sources or planar waves [7].

2.1 Virtual sound sources with complex radiation patterns

Point sources and planar waves are only of theoretical nature, but nevertheless they can be used to approximate the behavior of real sound sources. In reality the intensity of the acoustical field radiated by a source is direction-dependent. As this directional behavior has a major influence on its perception, the goal is to enable the reproduction of directional sound sources [8, 9].

Different approaches exist for the reproduction of directional sources using WFS depending on the type of sound sources under consideration (e.g point source, cylindrical source), the geometry of the reproduction setup, required processing power or psychoacoustic relevance [10–14].

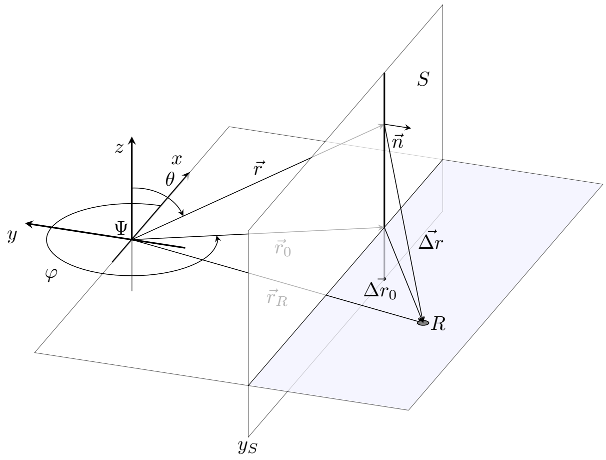

Figure 2: Geometry used for the mathematical description of wave field synthesis using a planar array of monopole sources.

2.2 Room acoustical simulation

Room acoustics has a great effect on the perceived properties of a sound source. It affects perceptual parameters like apparent source width or depth, or distance to the source. The acoustic behavior of a room can be represented by a room impulse response, which can be interpreted as the fingerprint of a room for a dedicated source and receiver position [15]. The state of the art is to use convolution-based room acoustic simulation devices, where the impulse response of a room is used as a filter for an audio input signal. Once the (”dry” – no room acoustic information is included) input signal is convoluted with the filter, it will sound like recorded in the room described by the filter. This approach is used in the field of spatial sound reproduction but is mainly limited to stereo or 5.1 surround sound.

Applying the convolution-based auralization approach to acoustic holography allows new applications for real-to-nature room simulation in wave field synthesis. The basic principle is to measure or simulate the acoustical behavior of a room on a dedicated aperture of a microphone array, and then using a virtual acoustic model to extrapolate this behavior to the desired positions in the room [16]. For this, the measured sound field is decomposed into directional components. A very interesting link between the theory of acoustic array processing and wave field synthesis is, that using the basic sound source types (plane wave, point source) is sufficient to reconstruct the sound field measured in a room at another place where a WFS system is installed [17]. To do so, the directional parts resulting from the wave field decomposition are used as room-impulse response filters of special sound source apertures. The superposition of all sound source signals will create the complex room acoustical sound field inside the listening area.

3. AURALIZATION AS PART OF VIRTUAL ENGINEERING

Virtualization is one of the key aspects of modern research and development strategies. With the aid of computer based systems and tools (CAx) the process of prototyping has been moved increasingly from the physical to a virtualized world. The use of computer aided systems began decades ago with the design of geometric surfaces, shapes and solids (CAD). Modern vector-based tools in two and three dimensions provide possibilities to construct e.g. prototypical machine parts and composite constructions. However isolated sole CAD models only virtualize the visual geometry of a specific shape.

The successive extension of sole CAD models in terms of the dynamic behavior due to physical real world properties is introduced by computer aided engineering systems (CAE). CAE provides simulations of a CAD component’s reactions to any type of physical force including kinematics, thermodynamics, fluid mechanics and stress analysis. This dynamic virtualization of components and assemblies in mutual dependency to the laws of physics is an eminent progress in the virtual reproduction of physical processes.

Acoustics as an interdisciplinary science with links to above mentioned physical disciplines has therefore also its role in virtual engineering and CAE systems. Using tools like finite and boundary element methods (FEM, BEM) acoustic sound propagation of vibrating structures can be simulated. Those physical simulations are mostly represented by sound pressure values either as time domain signals, impulse responses or as complex spectra in the frequency domain. Evaluation of a component’s acoustic properties therefore are often restricted to quantitative data based analysis. In real world applications acoustic evaluation often demands qualitative ratings based

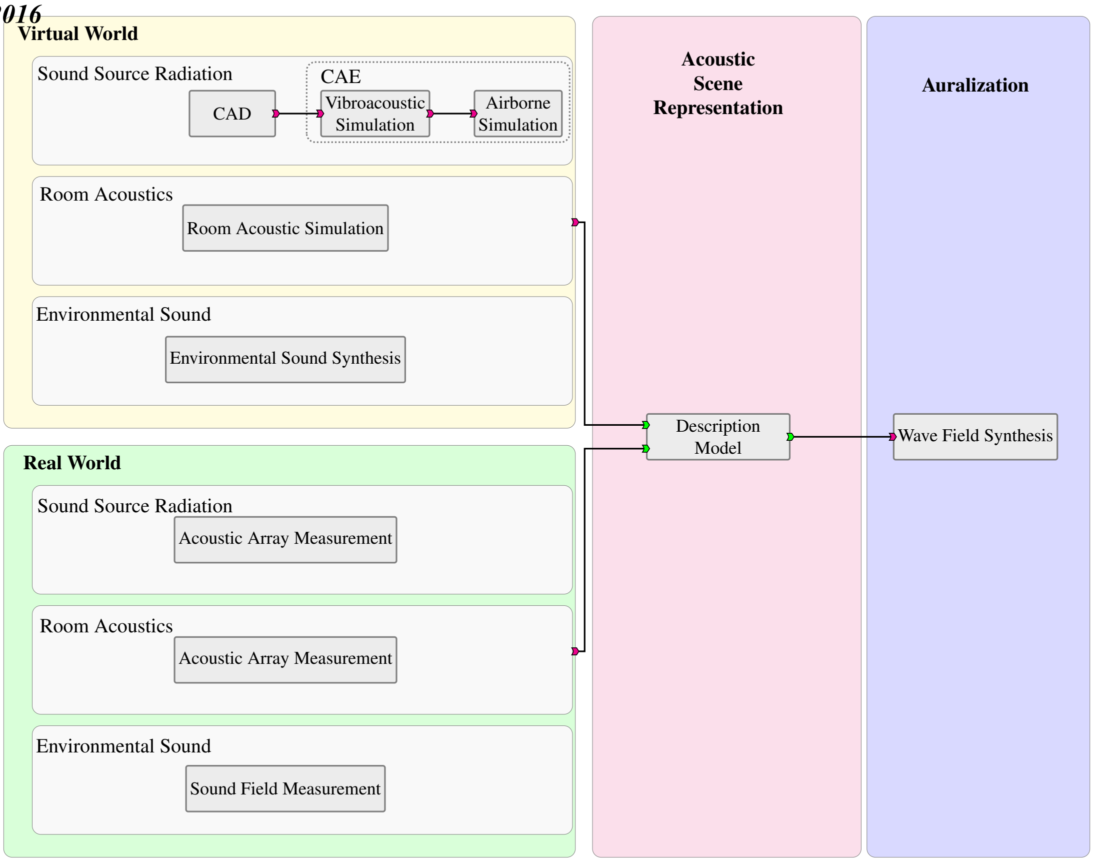

Figure 3: General concept of virtual acoustic signal acquisition. The representation of the acoustic scene is saved in a description model. The description model can be fed from simulation data (*Virtual World*) as well as from measurements and recordings (*Real World*).

on human perception of sound. One theoretical way to overcome this dilemma is to translate the physical sound pressure values into established psychoacoustic parameters like loudness, sharpness or roughness [18]. These parameters were developed on basis of empiric, qualitative listening tests and represented a mathematical description for physical properties according to psychoacoustic effects. The result again is a number whose interpretation is ambiguous, as the psychoacoustic models have not yet been developed to this extent. An explicit question of a sound developer could be: *How does this machine with amplitude response* x*, loudness* y*, and roughness* z *actually sound like?*

This question proves that a demand for plausible auralization of virtually modeled components and assemblies exists. Due to the lack of established psychoacoustic parameters, a visual analysis of the sound is not possible. Application examples are sound design and noise reduction during the product development process. The benefit of listening to machine sound before it is physically build is eminent. Prototype iterations can be decreased whereby time and money is saved. On the other hand new aspects of factory planning can be considered. The possibility arise to simulate the expected noise by virtually placing different machines at individual positions and being able to auralize those setups with suitable room acoustical treatments.

4. VIRTUAL ACOUSTIC DATA ACQUISITION

For the virtual acoustic product development process, a concept for sound acquisition is needed. A general concept of this is depicted in Figure 3. The acoustic description of the virtual scene is stored in a description model. This models contains all the necessary information e.g. radiation behavior of the source of interests, room acoustic properties of the surrounding area as well as environmental sounds. These data can be generated using established simulation algorithms (*Virtual World*) or be recorded from real sources (*Real World*).

4.1 Sound Source Model Representation

To obtain a structured auralization system it is necessary to define the representation of individual sound source models which can be transferred to a wave field synthesis renderer. These sound source models may have their origin in both CAE systems and real measurements. This flexibility is mandatory to evaluate simulated sources with their physical equivalent. The description model for a sound source suitable for WFS auralization contains two main parts: the actual audible signal and the source's propagation characteristics, i.e. the directivity. In terms of signal and system theory it is very suitable to keep up with this distributed structure. However when measuring a real non-trivial sound source like an engine, the sound propagation is intrinsically tied to the source's directivity. And also for simulated CAE source models this separation is not always feasible or reasonable. This ambiguity in representation demands at least two types of acoustic sound source models. For clarity the representations are distinguished as *distributed* and *combined* models. The distinction between these two types is directly dependent on the method used for radiation pattern representation as discussed in section 2.1. While the decomposition of a source's radiation pattern into spherical harmonics (cf. [10, 11]) is to be assigned to distributed models, data based approaches like [13] belong to combined source models. Both models have in common that several states of the sound source are stored. This is motivated by sound propagation that is not time-invariant in a strict definition. For example the sound propagation of an engine will differ with varying revolutions per minute rpm. This and other cases is coped by switchable state-dependent filter sets or signal sets respectively.

4.2 Influence of Room Acoustical Characteristics

Both CAE models and processed airborne sound source models consider the sound propagation in a non-reflecting free field environment. For analysis of the mere sound source characteristics this is not only sufficient but mandatory. Nevertheless room acoustical properties have a strong influence on the perceived sound characteristic. Thus when aiming for plausible, close to reality auralization, the application driven scenario and environment has to be modeled as well. A large-scale machine for example that is placed on different positions and orientations in a machine hall will affect the overall sound propagation and thus the physical and perceived noise pollution at points or areas of interest. Accordingly the auralization of room acoustical influences demand a flexible room simulation regarding the physical room properties themselves as well as adaptation to sound source movements. Fo this a parametric room acoustical model has to be defined consisting of direct sound, discrete early reflections and plausible reverberation. An established method for this is a hybrid model of defined mirror sound sources and diffuse reverberation [19].

4.3 Acoustic Scene Description Model

As defined in ISO 12913-1 an acoustic environment is *sound at the receiver from all sound sources as modified by the environment* [20]. In order to describe an acoustic scene the properties of all available sound sources, the room acoustical characteristics and possible interaction between the receiver and the acoustical environment have to be defined in advance. On the one side the acquisition of all those relevant assets can be conducted during the virtual acoustic engineering process as depicted in Figure 3 and described in the previous sections. On the other hand side the incorporation of real world measurements is not only an additional feature but an essential property, because during the development of a plausible holistic auditive system, the comparison between simulated and measured acoustic scenes is a mandatory task.

For recordings of diffuse sound fields, well-known microphone-techniques like INA5 configuration, IRT cross, Hamasaki square [21] can be included as well as techniques optimized for WFS like circular microphone arrays [22]. Referring to psychoacoustic properties like immersion and localization the so-called "Lindberg-Cube" [23] shows best results within an informal listening test even though spatial artifacts may arise to a certain extend [24]. Small (in terms of size) sound sources can be recorded with common close-mixing-techniques where microphone directivity should be taken into account to avoid undesired crosstalk. For larger sound sources conventional stereo microphone techniques like XY, AB or ORTF can be used with appropriate accurateness [25].

5. AUDIOVISUAL VIRTUAL ENGINEERING SYSTEM

To integrate the above described methods in a holistic audiovisual virtual engineering system, it is necessary to define certain prerequisites. These include the aimed usage and user experience, the resulting demands for inter-connectivity as well as psychophysical evaluation processes.

5.1 Usage and User Experience

The target field of application defined in section 1 dictates the desired usage of the audiovisual system. Keeping that in mind it becomes necessary to design the usage for both expert and non-expert users. Although the background processes of CAE and auralization involve complex mathematical and physical relationships, the actual usage should be as intuitive as possible. Consequently the need arises for an implementation of a certain intelligence to control the system according to the user's demands. The desired application environment complies with existing visual virtual reality systems, i.e. either a mobile system with head mounted display and headphones or a stationary system with stereoscopic projection (VR cave) and loudspeaker arrays [26]. The presented approach concentrates on latter system. Established visual virtual reality environments typically consist of a position tracking system and some sort of haptic controllers for user inputs to interact with the virtual scene. These components

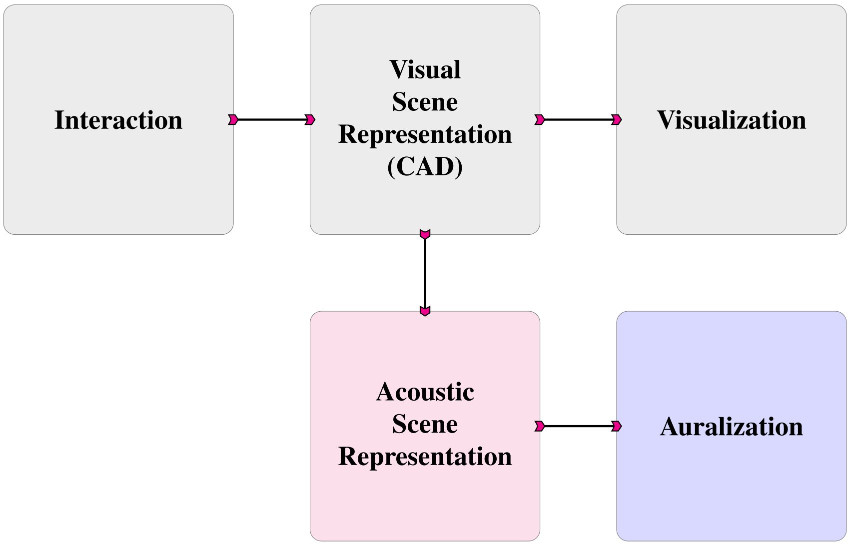

Figure 4: Detailed block diagram based on user interaction. A visual scene representation is interactively controlled, providing real-time (stereoscopic) visualization. As the visual scene is connected to the acoustic representation, user interaction directly controls the real-time auralization.

should be used as well. The resulting meta and control data regarding point of view and scene interaction can directly be used for interacting with the auralization part. Detailed description of the data handling is discussed in section 5.2. For evaluation purposes both visual and auditive systems can be evaluated individually. Standardized methods like ITU-R BS.1534-3 [27] for audio quality evaluation or ITU-T P.910 [28] for subjective video quality assessment are provided for testing these sub-systems. Holistic evaluation of user experience on the other side, can be evaluated with methods of 'quality of experience testing' [29].

5.2 System Design

After defining the demands for innovative auralization techniques in section 1, relevant methods and tools of the CAx data basis in section 3, appropriate modeling concepts in sections 4.1 and 4.2, the desired user experience in section 5.1 and the available state-of-the-art implementations of wave field synthesis as technology of choice in section 2 a holistic system can be designed. As the process of connecting CAx methods and tools with a WFS auralization system is still mostly unknown, novel connection structures and strategies have to be developed and implemented. Due to intrinsic proximity and the availability of commercial systems of CAD, CAE and corresponding user interfaces the technologies could be connected easily. CAVE-like systems that handle this kind of data are already in commercial use. Therefore the main task will consider interfaces and interpreters for connecting the CAE processes with the auralization approach. Doing this, a separation is introduced between application data (i.e. geometric, behavioral and acoustic source models, room models) and meta/control data like type, positioning or orientation. Figure 4 shows a schematic block diagram of the inter-connectivity between the individual assets. The connection between visual and acoustic scene representation is conducted by binding objects and environments from both domains with each other. The presented object-based approach allows to control both representations with appropriate user inputs. The visual scene representation that can be controlled directly by the user, affects the acoustic scene representation and consequently the auralization system. This is done by processing the application data with the corresponding meta data in order to conform with the WFS reproduction. There are three different main assets provided to describe an acoustic scene: the acoustic environment which contains room acoustical properties, the sounding objects themselves which are identified by a sound source model and relevant meta data and the point of view of the user within the audiovisual scene. The presented approach's scene design has a nested structure of environments and acoustic objects to provide possibilities for room-in-room acoustical scenes. The main task to connect visual and acoustic scene representation is to collect, process and transfer these data according to the WFS rendering's properties, specifically (in order of execution):

1. Arranging sounding objects relatively to the point of view regarding positioning and orientation
2. Decision whether sounding objects are audible depending on distance top point of view, type and state of object and possible interference with other sounding objects
3. Assigning sound source model data to actual state of sounding objects

4. Assigning room acoustical filtering according to current environment settings

5. Transfer relative positioned audio objects with according filter settings to WFS reproduction.

6. CONCLUSION

In this paper an approach was presented to enhance virtual engineering processes by connecting the auralization of the acoustical behavior of virtual products during the early design process. Innovative auralization techniques and modeling concepts were shown, the possibilities and requirements for a holistic audiovisual virtual environment were defined and finally a concept for the connection of visualization and auralization systems was developed.

ACKNOWLEDGMENT

The presented concepts were developed in the course of two interdisciplinary research projects funded by the ProExzellenz initiative of the Free State of Thuringia called VISTA4F and by the Federal Ministry of Economic Affairs and Energy (BMWi) called AVP^3.

REFERENCES

[1] Ovtcharova J. Virtual Engineering - Virtuelle Produktentstehung. Springer, Berlin; 2016.

[2] Lyon RH. Designing for Product Sound Quality. MARCELL DEKKER, New York; 2000.

[3] von Estorff O, Markiewicz M, Zaleski O. Validation of Numerical Methods in Acoustics: What can we expect? F Magouls (Ed): Computational Methods for Acoustics Problems, Saxc-Coburg Publications, Dun Eaglais, Station Brae Kippen, Stirlingshire, Scotland, 2008. 2008;Mubve.

[4] Brix S, Brix T, Sladeczek C. Wave Field Synthesis for Virtual Prototyping in VR Systems. In: The 17th International Conference in Engineering Design. Stanford, CA, USA: Center for Design Research Stanford University; 2009. .

[5] Berkhout AJ. A Holographic Approach to Acoustic Control. Journal of the Audio Engineering Society (JAES). 1988;36(12):977–995.

[6] Williams EG. Fourier Acoustics - Sound Radiation and Nearfield Acoustical Holography. London, UK: ACADEMIC PRESS; 1999.

[7] Verheijen ENG. Sound Reproduction by Wave Field Synthesis; 1998.

[8] Giron F. Investigations about the Directivity of Sound Sources; 1996.

[9] Meyer J, Hansen U. Acoustics and the Performance of Music: Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instrument Makers. Modern Acoustics and Signal Processing. Springer New York; 2009. Available from: `https://books.google.de/books?id=Mlkut4PAAiUC`.

[10] Corteel E. Synthesis of Directional Sources Using Wave Field Synthesis, Possibilities, and Limitations. EURASIP J Appl Signal Process. 2007 January;2007:188–188. Available from: `http://dx.doi.org/10.1155/2007/90509`.

[11] Ahrens J, Spors S. Implementation of Directional Sources In Wave Field Synthesis. In: Applications of Signal Processing to Audio and Acoustics, 2007 IEEE Workshop on; 2007. p. 66 –69.

[12] Melchior F, Sladeczek C, de Vries D, Fröhlich B. User-Dependent Optimization of Wave Field Synthesis Reproduction for Directive Sound Fields. In: 124th Audio Engineering Society Convention; 2008. .

[13] Baalman M. On Wave Field Synthesis and electro-acoustic music, with a particular focus on the reproduction of arbitrarily shapes sound sources; 2007.

[14] Sladeczek C, Zhykhar A, Brix S. Synthesis of directional sound sources with complex radiation patterns using a planar array of loudspeakers. In: Audio Engineering Society Conference: UK 25th Conference: Spatial Audio in Todays 3D World; 2012. Available from: `http://www.aes.org/e-lib/browse.cfm?elib=18115`.

[15] Kuttruff H. Room acoustics. London, New York: Spon Press; 2000. Available from: `http://opac.inria.fr/record=b1097189`.

[16] Hulsebos E. Auralization using Wave Field Synthesis; 2004.

[17] Melchior F. Investigation on spatial sound design based on measured room impulse responses; 2011.

[18] Fastl H, Zwicker E. Psychoacoustics- Facts and Models. Berlin: Springer; 2013.

[19] Melchior F, Sladeczek C, Partzsch A, Brix S. Design and Implementation of an Interactive Room Simulation for Wave Field Synthesis. In: 40th AES International Conference. Tokyo, Japan; 2010. .

[20] ISO 12913-1:2014 - Acosutics - Soundscape-Part 1. Geneva; 2014.

[21] Theile G, Dickreiter M, Graul W, Camerer F, Spikofski G. Tonaufnahme und Tonwiedergabe. In: Handbuch der Tonstudiotechnik. 8th ed. Berlin: De Gruyter Saur; 2014. .

[22] Spors S, Ahrens J. Generation of Highly Immersive Atmospheres for Wave Field Synthesis Reproduction. In: 116th Audio Engineering Society Convention. Berlin; 2004. .

[23] Grewe Y. Vergleich von immersiven Hauptmikrofonverfahren Gegenüberstellung von drei Aufnahmetechniken in Praxis und Simulation [Bachelor's thesis]. Hochschule Offenburg. Offenburg; 2015.

[24] Clauß T. Entwicklung eines emotionsbasierten Lärmemissionstest zum Beispiel der Geräuschbewertung von hydraulischen Bauteilen. Ilmenau; 2014.

[25] Ballou G, Ciaudelli J, Schmitt V. Microphones. In: Handbook for Sound Engineers. Oxford: Focal Press; 2008. .

[26] Springer JP, Sladeczek C, Scheffler M, Hochstrate J, Melchior F, Frohlich B. Combining Wave Field Synthesis and Multi-Viewer Stereo Displays. In: IEEE Virtual Reality Conference (VR 2006); 2006. p. 237–240.

[27] Recommendation ITU-R BS.1534-3: Method for the subjective assessment of intermediate quality level of audio systems MUSHRA;.

[28] ITU-T P.910: Subjective video quality assessment methods for multimedia applications;.

[29] Möller S, Raake (ed) A. Quality of Experience. Berlin: Springer; 2014.