Publications in Telecommunications Software and Multimedia
Teknillisen korkeakoulun tietoliikenneohjelmistojen ja multimedian julkaisuja
Espoo 1999 TML-A3

Modeling Techniques for Virtual Acoustics

Lauri Savioja

TEKNILLINEN KORKEAKOULU
TEKNISKA HÖGSKOLAN
HELSINKI UNIVERSITY OF TECHNOLOGY

Publications in Telecommunications Software and Multimedia
Teknillisen korkeakoulun tietoliikenneohjelmistojen ja multimedian julkaisuja
Espoo 1999 TML-A3

Modeling Techniques for Virtual Acoustics

Lauri Savioja

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission for public examination and debate in Auditorium T1 at the Helsinki University of Technology (Espoo, Finland) on the 3rd of December, 1999, at 12 o'clock noon.

Helsinki University of Technology
Department of Computer Science and Engineering
Telecommunications Software and Multimedia Laboratory

Teknillinen korkeakoulu
Tietotekniikan osasto
Tietoliikenneohjelmistojen ja multimedian laboratorio

Distribution:
Helsinki University of Technology
Telecommunications Software and Multimedia Laboratory
P.O. Box 5400
FIN-02015 HUT
Tel. +358-9-451 2870
Fax. +358-9-451 5014

ISBN 951-22-4765-8
ISSN 1456-7911

Picaset
Espoo 1999

Author Lauri Savioja
Title Modeling Techniques for Virtual Acoustics

The goal of this research has been the creation of convincing virtual acoustic environments. This consists of three separate modeling tasks: the modeling of the sound source, the room acoustics, and the listener. In this thesis the main emphasis is on room acoustics and sound synthesis.

Room acoustic modeling techniques can be divided into wave-based, ray-based, and statistical methods. Accurate modeling for the whole frequency range of human hearing requires a combination of various techniques, e.g., a wave-based model employed at low frequencies and a ray-based model applied at high frequencies. An overview of these principles is given.

Real-time modeling has special requirements in terms of computational efficiency. In this thesis, a new real-time auralization system is presented. A time-domain hybrid method was selected and applied to the room acoustic model. Direct sound and early reflections are computed by the image-source method. They are individually auralized, i.e., rendered audible with the addition of late reverberation, generated with a recursive digital filter structure. The novelties of the system include implementation of a real-time image-source method, parametrization of auralization, and update rules and interpolation of auralization parameters. The applied interpolation enables interactive movement of the listener and the sound source in the virtual environment.

In this thesis, a specific wave-based method, the digital waveguide mesh, is discussed in detail. The original method was developed for the two-dimensional case, which suits, e.g., simulation of vibrating plates and membranes. In this thesis, the algorithm is generalized for the N-dimensional case. In particular, the new three-dimensional mesh is interesting since it is suitable for room acoustic modeling. The original algorithm suffers from direction-dependent dispersion. Two improvements to the original algorithm are introduced. Firstly, a new interpolated rectangular mesh is introduced. It can be used to obtain wave propagation characteristics, which are nearly independent of the wave propagation direction. Various new interpolation strategies for the two-dimensional structure are presented. For the three-dimensional structure a linearly interpolated structure is shown. Both the interpolated mesh, and the triangular mesh, have dispersive characteristics. Secondly, a frequency-warping technique which can be used to enhance the frequency accuracy of simulations, is illustrated.

UDK 534.84, 681.327.12, 621.39
Keywords virtual reality, room acoustics, 3-D sound, digital waveguide mesh, frequency warping

This research was carried out in the Laboratory of Computer Science, Laboratory of Acoustics and Audio Signal Processing, and in the Laboratory of Telecommunications Software and Multimedia, Helsinki University of Technology, Espoo, during 1992-1999.

I am deeply indebted to Prof. Tapio Takala, my thesis supervisor, and Prof. Matti Karjalainen for their encouragement and guidance during all phases of the work. I also wish to thank Prof. Reijo Sulonen for giving me the idea of researching computational modeling of room acoustics.

Special thanks go to Dr. Vesa Välimäki and Dr. Jyri Huopaniemi for fruitful and innovative co-operation, inspiring discussions, and their patience teaching me the basics of digital signal processing. Mr. Tapio Lokki is thanked for fluent co-operation in the fields of auralization and interactive virtual acoustics.

I want to express my gratitude to Prof. Julius O. Smith III for several discussions on various issues related to digital waveguide meshes.

I would like to thank my other co-authors, Mr. Tommi Huotilainen, Ms. Riitta Väänänen, and Mr. Timo Rinne. I am grateful to all the people who have contributed to the DIVA system, especially Mr. Rami Hänninen, Mr. Tommi Ilmonen, Mr. Jarmo Hiipakka, and Mr. Ville Pulkki. Mr. Aki Härmä is acknowledged for insightful discussions on frequency warping.

I wish to express my gratitude to Mr. Nick Zacharov for his help in improving the English of this thesis.

Finally, my warmest thanks to my wife Minna for her love and patience during this work. Thanks also to our daughters, Hanna and Kaisa, for their ability to efficiently keep my thoughts out of room acoustics (excluding noise reduction).

This thesis summarizes the following articles and publications, referred to as [P1]-[P10]:

[P1] L. Savioja, J. Huopaniemi, T. Huotilainen, and T. Takala. Real-time virtual audio reality. In *Proc. Int. Computer Music Conf. (ICMC'96)*, pages 107–110, Hong Kong, 19-24 Aug. 1996.

[P2] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen. Virtual environment simulation - advances in the DIVA project. In *Proc. Int. Conf. Auditory Display (ICAD'97)*, pages 43–46, Palo Alto, California, 3-5 Nov. 1997.

[P3] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen. Creating interactive virtual acoustic environments. *J. Audio Eng. Soc.*, 47(9):675–705, Sept. 1999.

[P4] L. Savioja, T. Rinne, and T. Takala. Simulation of room acoustics with a 3-D finite difference mesh. In *Proc. Int. Computer Music Conf. (ICMC'94)*, pages 463–466, Aarhus, Denmark, 12-17 Sept. 1994.

[P5] L. Savioja and V. Välimäki. Improved discrete-time modeling of multi-dimensional wave propagation using the interpolated digital waveguide mesh. In *Proc. Int. Conf. Acoust., Speech, Signal Processing (ICASSP'97)*, volume 1, pages 459–462, Munich, Germany, 19-24 April 1997.

[P6] J. Huopaniemi, L. Savioja, and M. Karjalainen. Modeling of reflections and air absorption in acoustical spaces — a digital filter design approach. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'97)*, Mohonk, New Paltz, New York, 19-22 Oct. 1997.

[P7] L. Savioja. Improving the three-dimensional digital waveguide mesh by interpolation. In *Proc. Nordic Acoustical Meeting (NAM'98)*, pages 265–268, Stockholm, Sweden, 7-9 Sept. 1998.

[P8] L. Savioja and V. Välimäki. Reduction of the dispersion error in the interpolated digital waveguide mesh using frequency warping. In *Proc. Int. Conf. Acoust., Speech, Signal Processing (ICASSP'99)*, volume 2, pages 973–976, Phoenix, Arizona, 15-19 March 1999.

[P9] L. Savioja and V. Välimäki. Reduction of the dispersion error in the triangular digital waveguide mesh using frequency warping. *IEEE Signal Processing Letters*, 6(3):58–60, March 1999.

[P10] L. Savioja and V. Välimäki. Reducing the dispersion error in the digital waveguide mesh using interpolation and frequency-warping techniques. *Accepted for publication in IEEE Transactions on Speech and Audio Processing*, 1999.

A	gain coefficient
$A(z)$	allpass transfer function
$A_k(z)$	air absorption filter
c	speed of sound
c_θ, c_ϕ	HRTF interpolation coefficient
D	delay
$D_k(z)$	source directivity filter
f	temporal frequency
f_s	sampling frequency
$F_k(z)$	listener model filter block
g_k	distance attenuation gain
$g(\xi_1, \xi_2)$	spectral amplification factor
E	relative frequency error
h	weighting coefficient
$h_A, \ldots, h_E$	HRTF filter coefficients
H	number of wave propagation directions
k	integer variable
$k(\xi_1, \xi_2)$	dispersion factor
k_{DC}	dispersion factor at zero frequency
L	listener
M_k	propagation delay
$M(i, j)$	visibility matrix
n	integer variable
N	integer constant
p	sound pressure at a junction
P	reflection path
$P(n, \xi_1, \xi_2)$	Fourier transform of p
r	reflection coefficient
$R_k(z)$	reflection filter
RT_{60}	reverberation time
$s(n)$	signal samples
$s_w(n)$	warped signal samples
S	sound source
t	time variable
T	sampling interval
$T_k(z)$	auralization filter block
z	z-transform variable
w_{ratio}	warping ratio
$y(n)$	output signal

α	angle
$\delta(n)$	unit impulse
Δx	spatial sampling interval
λ	warping factor
ϕ	elevation angle
τ_1, τ_2	interpolation coefficients
θ	azimuth angle
ξ	spatial frequency
ω	angular frequency

2-D	Two-dimensional
3-D	Three-dimensional
BEM	Boundary Element Method
BRIR	Binaural Room Impulse Response
DIVA	Digital Interactive Virtual Acoustics
DSP	Digital Signal Processing
ETC	Energy-Time Curve
FIR	Finite Impulse Response
FDTD	Finite Difference Time Domain
FEM	Finite Element Method
FFT	Fast Fourier Transform
GUI	Graphical User Interface
HRTF	Head Related Transfer Function
IIR	Infinite Impulse Response
ILD	Interaural Level Difference
ITD	Interaural Time Difference
MIDI	Musical Instrument Digital Interface
RFE	Relative Frequency Error
VBAP	Vector Base Amplitude Panning

Virtual acoustics is a broad topic including modeling of sound sources, room acoustics and the listener. In this thesis an overview of these topics is given. Room acoustic modeling techniques are focused upon for application to interactive virtual reality.

1.1 Background

Traditionally, sound reproduction has been either monophonic or stereophonic. For a listener this means that the sound appears to come from one point source or from a line connecting two point sources, if the listening environment is anechoic. To create more realistic soundscapes, a technique is required in which sounds can emanate from any direction. This can be created with current algorithms using multichannel or headphone reproduction, or to a certain degree with only two loudspeakers. In general, these latter techniques are referred to as *sound spatialization* or *three-dimensional (3-D) sound*.

Only recently has spatial sound gained the interest it deserves. The spatialization of sound provides the possibility of creating fully *immersive* three-dimensional soundscapes. This is an important enhancement to virtual reality systems, in which the main focus has traditionally been on visual immersion.

In general, virtual acoustics has a wide range of application areas related to virtual reality. Nowadays, the most common use is entertainment, in which 3-D sound is used widely in applications varying from computer games to movie theaters. It is even possible to buy a sound card for a PC, capable of creating rudimentary three-dimensional soundscapes. Other application areas include for example tele- and videoconferencing, audio user interfaces for blind people and aeronautical applications [1].

1.2 Modeling of Virtual Acoustics

There are three separate parts to be modeled in a virtual acoustic environment [1][P3]:

- Sound sources
- Room acoustics
- The listener

These items form a source-medium-receiver chain, which is typical for all communication models. This is also called the *auralization* chain as illustrated in Fig. 1.1 [2][P3].

Sound source modeling consists of following two items:

- Sound synthesis
- Sound radiation characteristics

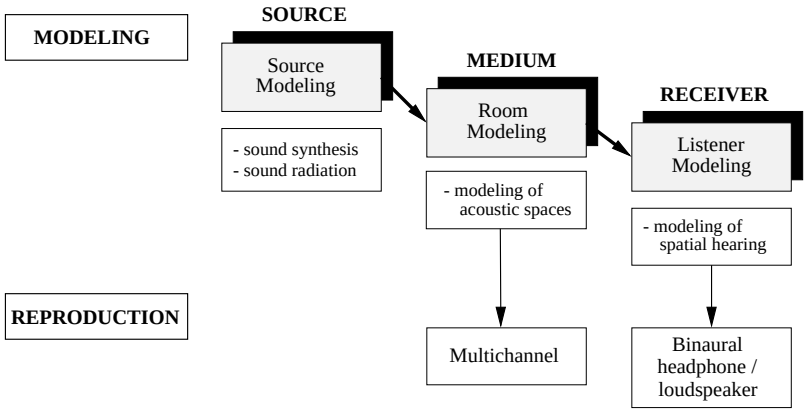

Figure 1.1: The process of implementing virtual acoustic environments consists of three separate modeling tasks. The components to be modeled are the sound source, the medium, and the receiver [2][P3].

Sound synthesis has been studied in depth, and there are various approaches to this topic. The main application has been musical instruments and their sound synthesis. For this thesis, the most interesting technique is physical modeling, which imitates the physical process of sound generation [3]. The technique is eligible for sound source simulation in virtual acoustic environments [4]. The most commonly employed physical modeling technique is the digital waveguide method. It is capable of real-time sound synthesis for one-dimensional instruments, such as strings and woodwinds [5, 3, 6, 7].

The simplest approach is to assume the sound source to be an omnidirectional point source. Typically, sound sources have frequency dependent directivity and this has to be modeled to achieve realistic results [4][P3]. For example, typical musical instruments radiate most of the high frequency sound to the front hemisphere of the musician and at low frequencies they are omnidirectional. Another important aspect of sound radiation is the shape of the source. Most sources can be modeled as point sources, but, for example, a grand piano is so large, that it cannot be modeled as a single point. Comprehensive studies on the physics of musical instruments and their sound radiation are presented, e.g., in [8, 9].

In room acoustic modeling (Fig. 1.1) the sound propagation in a medium, typically air, is modeled [10, 11]. This takes into account propagation paths of the direct sound and early reflections, and their frequency dependent attenuation in the air and at the boundaries. Also, the diffuse late reverberation has to be modeled.

In multichannel reproduction the sound field is created using multiple loudspeakers surrounding the listener. A similar effect can also be created with binaural reproduction, but this requires modeling of the listener (Fig. 1.1), in which the properties of human spatial hearing are considered. A simple means of providing a directional sensation of sound is to model the interaural level and time differences (ILD and ITD) as frequency independent gain and delay differences. For high-quality auralization we also

need head-related transfer functions (HRTF) which model the reflections and filtering of the head, shoulders and pinnae of the listener as well as frequency dependence of ILD and ITD [12, 11, 13, 14].

1.3 The DIVA System

The three modeling phases of Fig. 1.1 enable creation of realistic sounding virtual acoustic environments. One such system has been implemented at Helsinki University of Technology in the DIVA (Digital Interactive Virtual Acoustics) research project [15][P3]. The aim in the DIVA project has been to create a real-time environment for full audiovisual experience. The system integrates the whole audio signal processing chain from sound synthesis through room acoustics simulation to spatialized reproduction. This is combined with synchronized animated motion. A practical application of this project is a virtual concert performance [16, 17].

In Fig. 1.2 the architecture of the DIVA virtual concert performance system is presented. There may be two simultaneous users in the system, namely, a conductor and a listener, who may both interact with the system. The conductor wears a tailcoat with magnetic sensors for tracking. Through movement the orchestra may be conducted, and it may contain both real and virtual musicians [18, 19].

In the DIVA system, animated human models are placed on stage to play music from MIDI files [20]. The virtual musicians play their instruments at a tempo and loudness guided by the conductor.

At the same time a listener may freely fly around within the concert hall. The graphical user interface (GUI) sends the listener position data to the auralization unit which renders the sound samples provided by physical models and a MIDI synthesizer. The auralized output is reproduced either through headphones or loudspeakers. The developed room acoustic model can be used both in real-time and non-real-time applications. The real-time system has been demonstrated at several conferences [16, 21, 22]. The DIVA virtual orchestra has even given a few public performances in Finland. Non-real-time modeling has been applied to make a demonstration video of a planned concert hall [23], for example.

1.4 Scope of the Thesis

The virtual acoustic modeling techniques presented in this thesis contain both real-time and non-real-time algorithms. They contribute both to sound source modeling and to room acoustic modeling.

The main algorithms discussed in this thesis are:

- Real-time room acoustic modeling and auralization based on geometrical room acoustics.
- Digital waveguide mesh, which can be used to model both sound sources and room acoustics.

The real-time auralization part discusses technical challenges concerning implementation of comprehensive interactive systems such as the DIVA system.

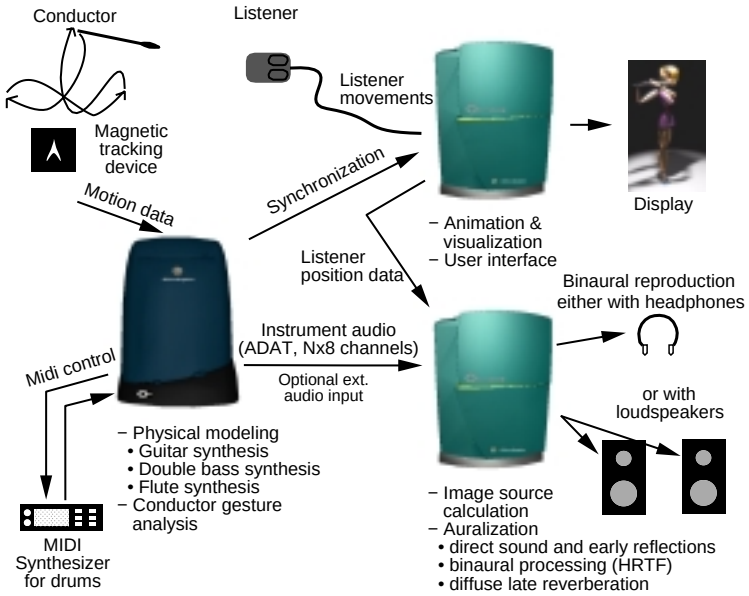

Figure 1.2: In the DIVA system a conductor may conduct musicians while a listener may move inside the concert hall and listen to an auralized performance [P3].

The digital waveguide mesh method is a modeling technique typically applied at low frequencies and it also includes diffraction modeling. The research presented in this thesis concentrates on improving the frequency accuracy of the method using interpolation and frequency-warping techniques.

1.5 Contents of the Thesis

This thesis is organized as follows. Chapter 2 gives an overview of room acoustic modeling techniques. In Chapter 3, real-time modeling techniques and their special requirements are presented. Chapter 4 discusses the digital waveguide mesh method. Chapter 5 concludes the thesis.

Computers have been used for over thirty years to model room acoustics [24, 25, 26], and nowadays computational modeling together with the scale models are a relatively common practice in acoustic design of concert halls. A good overview of modeling algorithms is presented in [27, 11].

Figure 2.1 presents a simplified room geometry, and propagation paths of the direct sound and some early reflections. In the figure all the reflections are supposed to be specular. In real reflections there exists always a diffuse component too (see, e.g., [10] for more on basics of room acoustics).

An impulse response of a concert hall can be separated into three parts: direct sound, early reflections, and late reverberation. The response illustrated in Fig. 2.2(a) is a simplified one, in which there are no diffuse or diffracted reflection paths. In real responses there would also be a diffuse sound energy component between early reflections as shown in Fig. 2.2(b).

2.1 The Main Modeling Principles

Mathematically the sound propagation is described by the *wave equation*, also known as the *Helmholtz equation*. An impulse response from a source to a listener can be obtained by solving the wave equation, but it can seldom be performed in an analytic form. Therefore, the solution must be approximated and there are three different approaches in computational modeling of room acoustics as illustrated in Fig. 2.3 [P3]:

- Wave-based methods
- Ray-based methods
- Statistical models

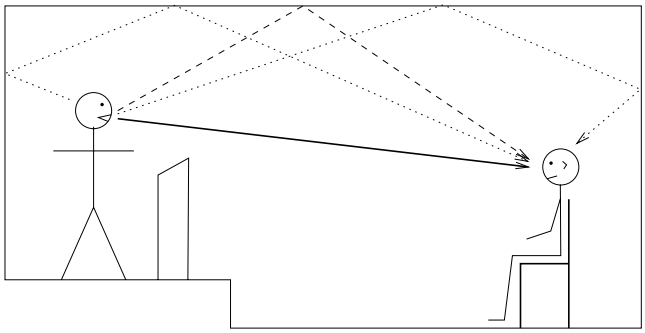

Figure 2.1: A simple room geometry, and visualization of the direct sound (solid line) and one first-order (dashed line) and two second-order (dotted line) reflections.

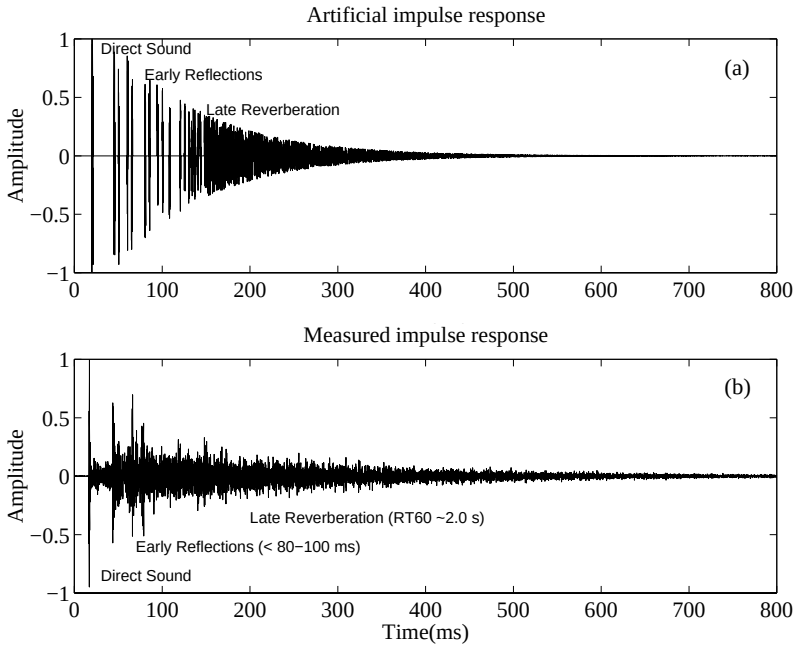

Figure 2.2: (a) An imitation of an impulse response of a concert hall. In a room impulse response simulation, the response is typically considered to consist of three separate parts: direct sound, early reflections and late reverberation. In the late reverberation part the sound field is considered diffuse. (b) A measured response of a concert hall [P3].

The ray-based methods, the ray-tracing [24, 28] and the image-source methods [29, 30], are the most often used modeling techniques. Recently, computationally more demanding wave-based techniques such as the finite element method (FEM), boundary element method (BEM) and finite-difference time-domain (FDTD) methods have also gained interest [27, 31, 32][P4]. These techniques are suitable for simulation of low frequencies only [11]. In real-time auralization the limited computation capacity calls for simplifications, by modeling only the direct sound and early reflections individually and the late reverberation by recursive digital filter structures.

The statistical modeling methods, such as the statistical energy analysis (SEA) [33], are mainly applied in prediction of noise levels in coupled systems in which sound transmission by structures is an important factor. They are not suitable for auralization purposes since typically those methods do not model the temporal behavior of a sound field.

The goal in most room acoustics simulations has been to compute an energy time curve (ETC) of a room (squared room impulse response), from which room acoustical attributes, such as reverberation time (RT_{60}), can be derived (see, e.g., [10, 34] for more about room acoustical attributes). The aim of this thesis is auralization, i.e., listening to the modeling results [11].

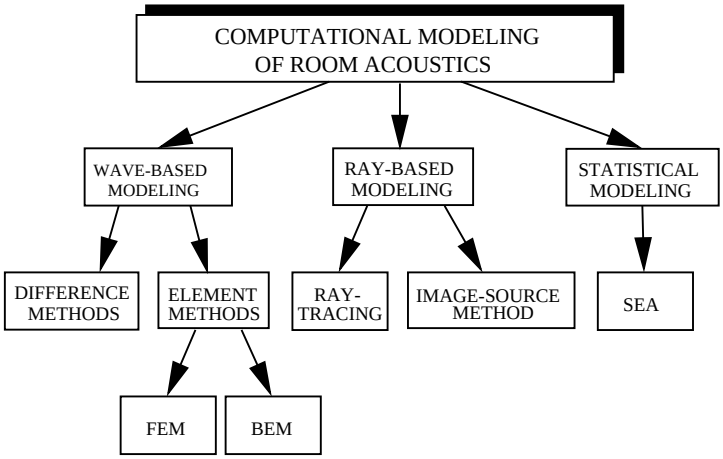

Figure 2.3: Principal computational models of room acoustics are based on sound rays (ray-based) or on solving the wave equation (wave-based) or some statistical technique. Different methods can be employed together to form a valid hybrid model [P3].

2.2 Wave-based Methods

The most accurate results can be achieved with the wave-based methods. An analytical solution for the wave equation can be found only in rare cases such as a rectangular room with rigid walls. Therefore numerical wave-based methods must be applied. Element methods, such as FEM and BEM, are suitable only for small enclosures and low frequencies due to heavy computational requirements [11, 35]. The main difference between these two methods is in the element structure. In FEM, the complete space has to be discretized with elements, but instead in BEM only the boundaries of the space are discretized. In practice this means that matrices used by a FEM solver are large but sparsely filled, whereas BEM matrices are smaller and denser.

Finite-difference time-domain (FDTD) methods provide another possible technique for room acoustics simulation [31, 32]. The main principle in the technique is that derivatives in the wave equation are replaced by corresponding finite differences [36]. The FDTD methods produce impulse responses better suited to auralization than FEM and BEM, which typically calculate frequency domain responses [11]. One FDTD method, the digital waveguide mesh, is presented in detail in Chapter 4. Another similar method can be obtained by using multidimensional wave digital filters [37] which have been used, e.g., to solve partial differential equations in general.

The main benefit of the element methods over FDTD methods is that one can create a denser mesh structure where required, such as locations near corners or other acoustically challenging places. Another advantage of the element methods is the ease of making coupled models, in which various wave propagation media are connected to each other.

In all the wave-based methods, the most difficult part is definition of the boundary conditions. Typically a complex impedance is required, but it is hard to find that data from existing literature.

2.3 Ray-based Methods

The ray-based methods of Fig. 2.3 are based on geometrical room acoustics, in which the sound is supposed to act like rays (see, e.g., [10]). This assumption is valid when the wavelength of sound is small compared to the area of surfaces in the room and large compared to the roughness of surfaces. Thus, all phenomena due to the wave nature, such as diffraction and interference, are ignored.

The results of ray-based models resemble the response in Fig. 2.2(a) since the sound is treated as rays with specular reflections. Note that in most simulation systems the result is the ETC which is the square of the impulse response.

The most commonly used ray-based methods are the ray-tracing [24, 28] and the image-source method [29, 30]. The basic distinction between these methods is the way the reflection paths are typically calculated. To model an ideal impulse response all the possible sound reflection paths should be discovered. The image-source method finds all the paths, but the computational requirements are such that in practice only a set of early reflections is computed. The maximum achievable order of reflections depends on the room geometry and available calculation capacity. In addition, the geometry must be formed of planar surfaces. Ray-tracing applies the Monte Carlo simulation technique to sample these reflection paths and thus it gives a statistical result. By this technique higher order reflections can be searched for, though there are no guarantees that all the paths will be found.

Ray-tracing Method

Ray-tracing is a well-known algorithm in simulating the behavior of an acoustic space [24, 28, 38, 39]. There are several variations of the algorithm, which are not all covered here. In the basic algorithm the sound source emits sound rays, which are then reflected at surfaces according to specular reflection and the listener keeps track on which rays have penetrated it as audible reflections. The way sound rays are emitted can either be predefined or randomized [28]. A typical goal is to have a uniform distribution of rays over a sphere. By use of a predefined distribution of rays a superior solution can be achieved with fewer rays.

The specular reflection rule is most common, in which the incident angle of an incoming ray is the same as the incident angle of the outgoing ray. More advanced rules which include for example some diffusion algorithm have also been studied (see, e.g., [40, 41]).

The listeners are typically modeled as volumetric objects, like spheres or cubes, but the listeners may also be planar. In theory, a listener can be of any shape as far as there are enough rays to penetrate the listener to achieve statistically valid results. In practice a sphere is in most cases the best choice, since it provides an omnidirectional sensitivity pattern and it is

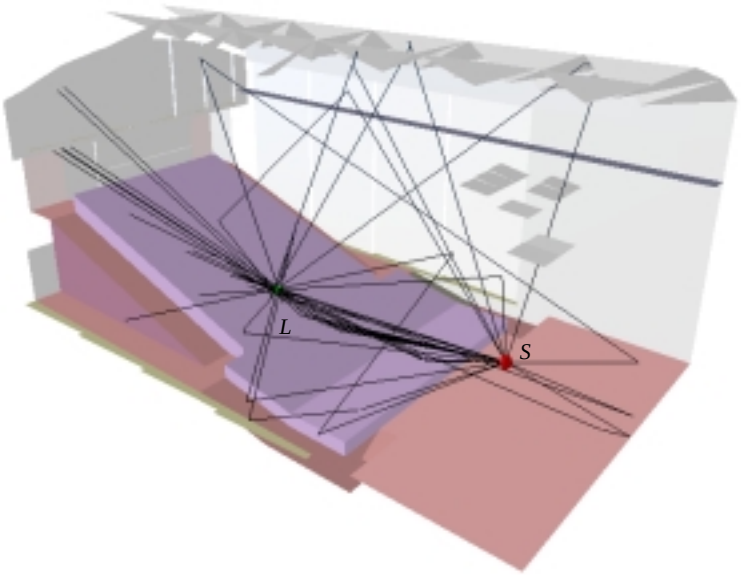

Figure 2.4: The direct sound and first- and second-order reflection paths in a concert hall obtained by a ray-tracing simulation. The source and the listener are denoted by S and L, respectively.

easy to implement.

Figure 2.4 presents a model of a concert hall with direct sound and all the first- and second-order reflection paths obtained by ray-tracing. The geometrical model of the hall contains ca. 300 polygons and 40,000 rays were emitted uniformly over a sphere. The model represents the Sigyn concert hall in Turku, Finland [42].

Image-source Method

From a computational point of view the image-source method is also a ray-based method [43, 29, 30, 44, 45, 46]. The basic principle of the image-source method is presented in Fig. 2.5. Reflected paths from the real source are replaced by direct paths from reflected mirror images of the source.

Figure 2.5 contains a section of a simplified concert hall consisting of a floor, ceiling, back wall, and balcony. Image sources S_c and S_f represent reflections produced by the ceiling and the floor. There is also a second-order image source S_{fc} which is the reflected image of S_f with respect to the ceiling. After finding the image sources a visibility check must be performed. This indicates whether there is an unoccluded path from the source to the listener. This is done by forming the actual reflection path (P_c, P_{fc} and P_f in Fig. 2.5) and checking that it does not intersect any surface in the room. In Fig. 2.5, the image sources S_c and S_{fc} are visible to the listener L whereas the image source S_f is hidden under the balcony since the path P_f is intersecting it. It is important to notice that locations of the image sources are not dependent on the listener's position and only

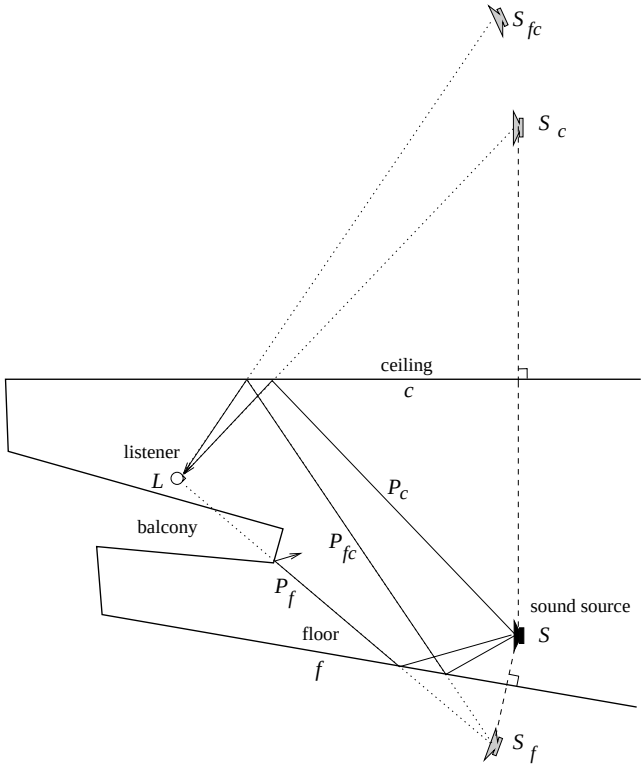

Figure 2.5: In the image-source method the sound source is reflected at each surface to produce image sources which represent the corresponding reflection paths. The image sources S_c and S_{fc} representing first- and second-order reflections from the ceiling are visible to the listener L while the reflection from the floor P_f is obscured by the balcony [P3].

the visibility of each image source may change when the listener moves.

Figure 2.6 illustrates the first- and second-order image sources in a concert hall. The simulation setup is similar to the one presented in the previous section concerning ray-tracing (see Fig. 2.4).

There are also hybrid models, in which ray-tracing and image-source method are applied together [47, 48, 41]. Typically early reflections are calculated with image sources due to its accuracy in finding reflection paths. The number of image sources grows exponentially as a function of order of reflections, and it is computationally inefficient to use the image-source method to find the higher order reflections. Therefore, later reflections are handled with ray-tracing. Hybrid methods can be extended also to contain some wave-based methods for low frequencies [49, 32].

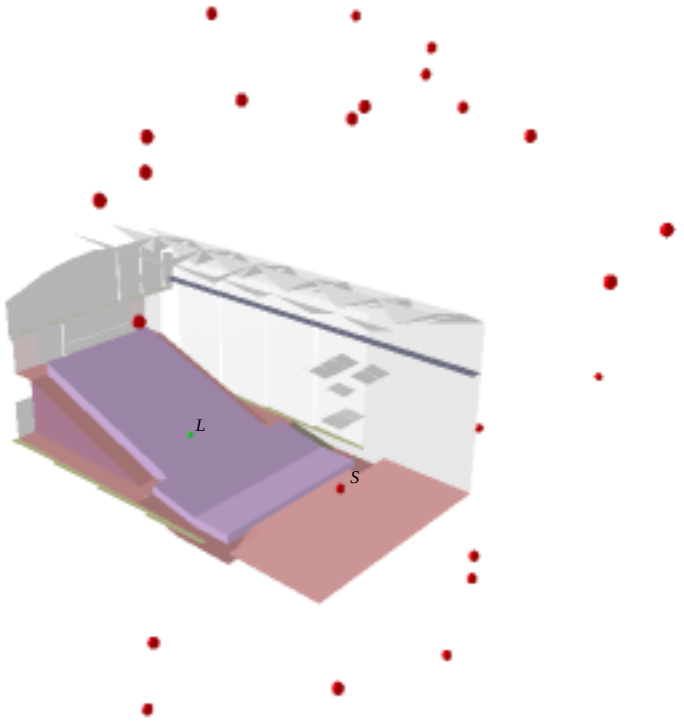

Figure 2.6: The computed image sources in a concert hall. All the visible first-order and second-order image sources are shown as spheres. The source and the listener are denoted by S and L, respectively.

The acoustic response perceived by the listener in a space varies according to the source and receiver positions and orientations. For this reason an interactive auralization model should produce an output which depends on the dynamic properties of the source, receiver, and environment. In principle, there are two different ways to achieve this goal. The methods presented in the following are called *direct room impulse response rendering* and *parametric room impulse response rendering* [P3].

The direct room impulse response rendering technique is based on binaural room impulse responses (BRIR) which are obtained a priori, either from simulations or from measurements. This method is suitable for static auralization purposes, but it cannot be used for interactive dynamic simulations. By interactive it is implied that users have the possibility to move around in a virtual hall and listen to the acoustics of the room at arbitrary locations in real-time.

A more robust way for dynamic real-time auralization is to use a parametric room impulse response rendering method. In this technique the BRIRs at different positions in the room are not pre-determined. The responses are formed in real-time by interactive simulation. The actual rendering process is performed in several parts. The initial part consists of direct sound and early reflections, and latter part represents the diffuse reverberant field.

The parametric room impulse response rendering technique can be further divided into two categories: the physical and the perceptual approach [50]. The physical modeling approach relates acoustical rendering to the visual scene. This involves modeling individual sound reflections off the walls, modeling sound propagation through objects, simulating air absorption, and rendering late diffuse reverberation, in addition to the 3-D positional rendering of source locations. The perceptual approach [50] investigates the perception of spatial audio and room acoustical quality.

In this chapter the auralization unit of the DIVA system (see, Fig. 1.2) is discussed [15][P1,P2,P3]. It is based on parametric room impulse response rendering using the physical approach. In the DIVA system, various real-time and non-real-time techniques are used to model room acoustics as illustrated in Fig. 3.1 [P3].

Performance issues play an important role in a real-time application. Therefore there are quite few alternative modeling methods available. The real-time auralization algorithm of the DIVA system uses the image-source method to calculate the early reflections and an artificial late reverberation algorithm to simulate the diffuse reverberant field. The image-source model was chosen since both the ray-tracing and the digital waveguide mesh method are too slow for real-time purposes.

The artificial late reverberation algorithm is parametrized based on room acoustical attributes obtained by non-real-time room acoustic simulations or measurements. The non-real-time calculation method is a hybrid technique containing digital waveguide mesh for low frequencies and ray-tracing for higher frequencies.

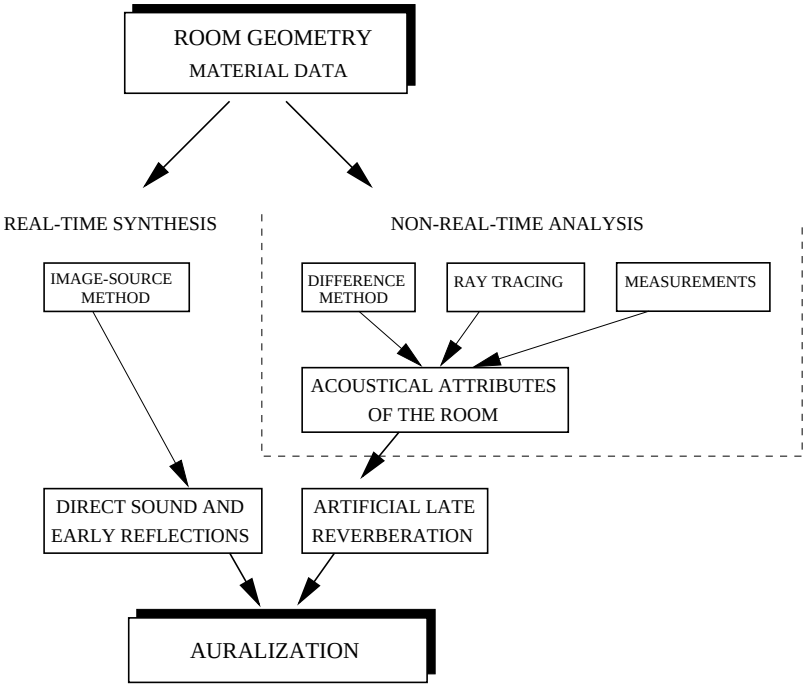

Figure 3.1: Computational methods used in the DIVA system. The model is a combination of real-time image-source method and artificial late reverberation which is parametrized according to room acoustical parameters obtained by simulations or measurements [P3].

3.1 Image-source Method

The implemented image-source algorithm is quite traditional and follows the method presented earlier in Section 2.3. However, there are some enhancements to achieve a better performance level [P2,P3].

In the image-source method, the number of image sources grows exponentially with the order of reflections. However, only some of them actually are effective because of occlusions. To reduce the number of potential image sources to be handled, only the surfaces that are at least partially visible to the sound source are examined. This same principle is also applied to image sources. The traditional way to examine the visibility is to analyze the direction of the normal vector of each surface. The source might be visible only to those surfaces which have the normal pointing towards the source [30]. After that check there are still unnecessary surfaces in a typical room geometry. The novel method to enhance the performance further is to perform a preprocessing run with ray-tracing to statistically check the visibilities of all surface pairs. The result is a Boolean matrix where item $M(i, j)$ indicates whether surface i is at least partially visible to surface j or not. Using this matrix the number of possible image source can be remarkably reduced [51].

One of the most time consuming procedures in an interactive image-

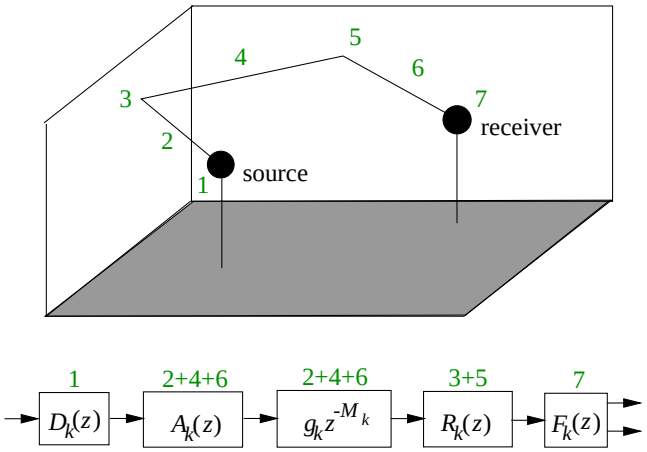

Figure 3.2: An example of a second-order reflection path [54, 55].

source method is the visibility check of image sources which must be performed each time the listener moves. This requires a large number of intersection calculations of surfaces and reflection paths. To reduce this, some advanced data structures can be applied. As a new improvement the author has chosen to use a geometrical directory EXCELL [52, 53]. The method is based on regular decomposition of space and it employs a grid directory. The directory is refined according to the distribution of data. The addressing of the directory is performed with an extendible hashing function [51][P3].

3.2 Auralization

In the auralization, the direct sound and early reflections are handled individually. To obey the law of nature, the direct sound and each reflection have to have appropriate distance attenuation ($1/r$-law), air absorption, and absorption caused by reflecting surfaces. Typically, all these effects can efficiently be modeled by low-order digital filters.

Figure 3.2 illustrates one second-order reflection path and the applied auralization filters. All such filters are linear in nature, such that they can be performed in an arbitrary order. The best performance is obtained when all the similar filters are cascaded. In the case presented in Fig. 3.2 the filters for reflection path k are:

- Directivity of the source 1, $D_k(z)$
- Air absorption, $A_k(z)$ for propagation path $2 + 4 + 6$
- $1/r$-law distance attenuation and propagation delay, $g_k z^{-M_k}$ for propagation path $2 + 4 + 6$
- A second-order wall reflection, $R_k(z)$ from surfaces 3 and 5
- Listener model $F_k(z)$ for two-channel reproduction

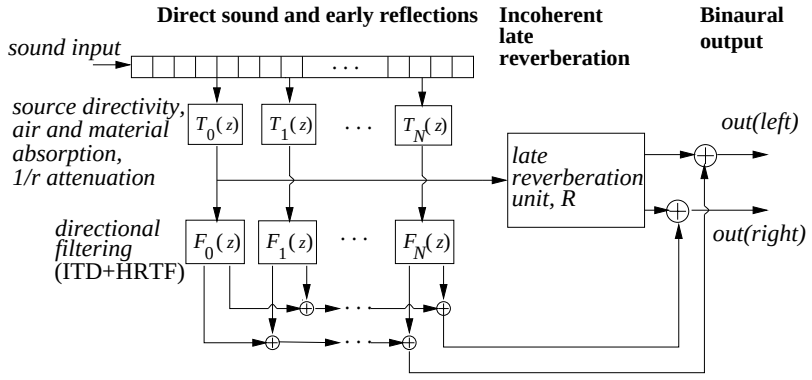

Figure 3.3: Structure of the auralization process in the DIVA system for headphone listening [P3].

The structure of the auralization process in the DIVA system is presented in Fig. 3.3. The audio input is fed to a delay line. This corresponds to the propagation delay from the sound source and each image source to the listener. In the next phase all of the sources are filtered with filter blocks T_k, where $k = 0, 1, 2, ..., N$. Each T_k contains the filters $D_k(z)$, $A_k(z)$, $g_k z^{-M_k}$, and $R_k(z)$ described above, except that $R_k(z)$ is not applied to the direct sound.

The signals produced by $T_k(z)$ are filtered with listener model filters $F_k(z)$, which create the binaural spatialization. This is summed with the output of the reverberator R. The listener model is not applied to the multichannel reproduction [P3].

Air Absorption

The absorption of sound in the transmitting medium (normally air) depends mainly on the distance, temperature, and humidity. There are various factors which participate in absorption of sound in air [10]. In a typical environment the most important is the thermal relaxation. The phenomenon is observed as an increasing low-pass filtering as a function of distance from the sound source. Analytical expressions for attenuation of sound in air as a function of temperature, humidity and distance have been published in, e.g., [56].

Based on the standardized equations for calculating air absorption [57], transfer functions for various temperature, humidity, and distance values were calculated, and second-order IIR filters were fitted to the resulting magnitude responses [P6]. Results of modeling for six distances from the source to the receiver are illustrated in Fig. 3.4(a). In Fig. 3.4(b), the effect of distance attenuation (according to $1/r$-law) has been added to the air absorption filter transfer functions.

Material Reflection Filters

The problem of modeling the sound wave reflection from acoustic boundary materials is complex. The temporal or spectral behavior of reflected sound as a function of incident angle, the scattering and diffraction phe-

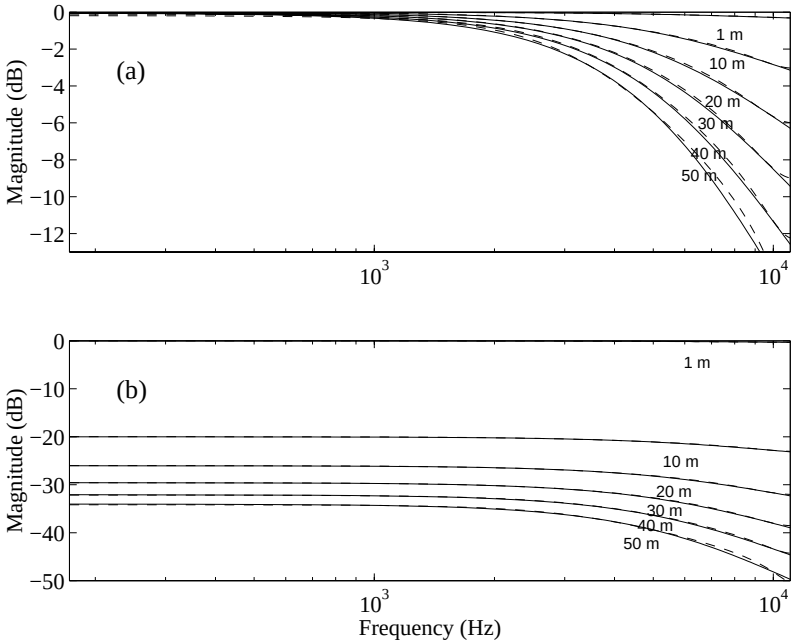

Figure 3.4: (a) Magnitude of air absorption filters as a function of distance (1m-50m) and frequency. The continuous line represents the ideal response and the dashed line is the filter response. For these filters the air humidity is chosen to be 20% and temperature 20°C. (b) Magnitude of combined air absorption and distance attenuation filters as a function of distance (1m-50m) and frequency [P6,P3].

nomena, etc., make it impossible to develop numerical models that are accurate in all aspects. For the DIVA system, computationally simple low-order filters were designed. Furthermore, the modeling was restricted to only the angle independent absorption characteristics [P6].

As an example, two reflection filters are depicted in Fig. 3.5. The magnitude responses of first-order and third-order IIR filters designed to fit the corresponding target values are shown. Each set of data is a combination of two materials (second-order reflection): a) plasterboard on frame with 13 mm boards and 100 mm air cavity [58], and glass panel (6+2+10 mm, toughened, acousto-laminated) [42], b) plasterboard (same as in previous) and 3.5-4 mm fibreboard with holes, 25 mm cavity with 25 mm mineral wool [58].

Late Reverberation

The late reverberant field of a room is often considered nearly diffuse and the corresponding impulse response exponentially decaying random noise [59]. Under these assumptions the late reverberation does not have to be modeled as individual reflections with certain directions. Instead, the reverberation can be modeled using recursive digital filter structures, whose

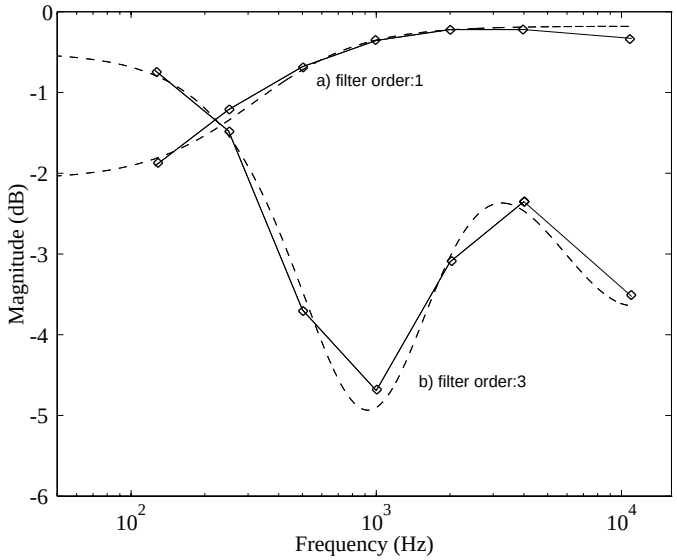

Figure 3.5: Two different material filters are depicted. The continuous lines represent the target responses and dashed lines are the corresponding filter responses. In (a) first-order and in (b) third-order minimum-phase IIR filters are designed to match given absorption coefficient data [P6,P3].

response models the characteristics of real room responses, such as the frequency dependent reverberation time. Producing incoherent reverberation with recursive filter structures has been studied, e.g., in [59, 60, 61, 62]. A good summary of reverberation algorithms is presented in [63].

The following four aims are essential in late reverberation modeling [59]:

1. The impulse response should be exponentially decaying with a dense pattern of reflections, to avoid fluttering in the reverberation.

2. Frequency domain characteristics should be similar to a concert hall which ideally has a high modal density especially at low frequencies. No modes should be emphasized more than the others to avoid coloration of the reverberated sound, or ringing tones in the response.

3. The reverberation time has to decrease as a function of frequency to simulate the air absorption and low-pass filtering effect of surface material absorption.

4. The late reverberation should produce partly incoherent signals at the listener's ears to attain a good spatial impression of the sound field.

In the DIVA system, a reverberator containing 4 to 8 parallel feedback loops is used [64]. This is a simplification of the *feedback delay networks* [61], and it fulfills all the afore mentioned criteria [P3].

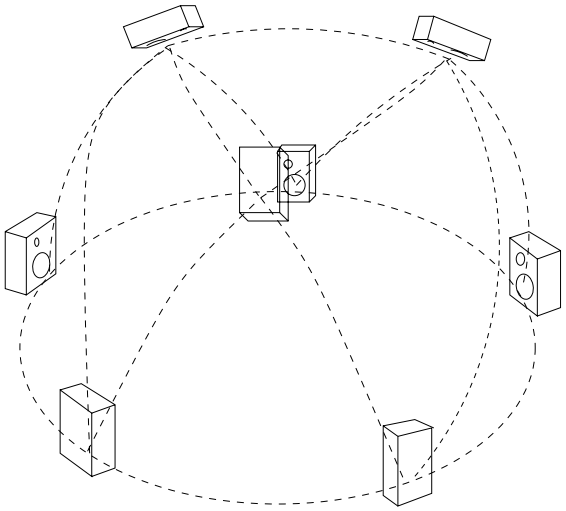

Figure 3.6: A typical 8-loudspeaker setup for VBAP reproduction. The listener can be in any place inside the hemisphere surrounded by the loudspeakers [68].

Reproduction

Reproduction schemes of virtual acoustic environments can be divided to the following categories [11]: 1) binaural (headphones), 2) crosstalk canceled binaural (two loudspeakers), and 3) multichannel reproduction. Binaural processing refers to three-dimensional sound image production for headphone or two-channel loudspeaker listening. For loudspeaker reproduction of binaural signals, a cross-talk canceling technology is required [65, 63]. The most common multichannel 3-D reproduction techniques applied in virtual reality systems are Ambisonics [66, 67] and vector base amplitude panning (VBAP) [68].

The DIVA system is capable of both binaural and multichannel loudspeaker reproduction. The binaural reproduction is based on head-related transfer functions (HRTF) [69, 70, 14], which are individual transfer functions for free-field listening conditions [12]. The binaural listening has been implemented for both headphones and two loudspeakers.

The multichannel reproduction in the DIVA system employs the vector base amplitude panning (VBAP) technique [68, 71]. This method enables arbitrary positioning of multiple loudspeakers and it is computationally efficient. Figure 3.6 [68] illustrates one possible loudspeaker configuration for eight channel reproduction.

3.3 Auralization Parameters

In the DIVA system the listener's position in the virtual room is determined by the graphical user interface (GUI, see Fig. 1.2). The GUI sends the movement data to the room acoustics simulator, which calculates the visible image sources in the space under study. To calculate the image sources the model needs the following information [P1]:

- Geometry of the room
- Materials of the room surfaces
- Location and orientation of each sound source
- Location and orientation of the listener

The orientations of the listener and source in the previous list are relative to the room coordinate system. The image-source model calculates positions and relative orientations of real and image sources with respect to the listener. Data of each visible source is sent to the auralization process. This novel set of auralization parameters is [15][P1,P3]:

- Distance from the listener
- Azimuth and elevation angles with respect to the listener
- Source orientation with respect to the listener
- Set of filter coefficients which describe the material properties in reflections

In the auralization process, the parameters affect coefficients of filters in filter blocks T_k and F_k and pick-up point from the input delay line in Fig. 3.3.

The number of auralized image sources depends on the available computing capacity. In the DIVA system typically parameters of 10-20 image sources are passed forward.

The novel strategy for handling multiple simultaneous sound sources is presented in [P2]. The basic idea is to handle direct sound from each source individually. The image sources are formed by summing multiple sources to a single source if the sources are close to each other. Otherwise image sources of each real source are treated individually.

Updating the Auralization Parameters

The auralization parameters change whenever the listener moves in the virtual space. The update rate of auralization parameters must be high enough to ensure the quality of auralization is not degraded. According to Sandvad [72] rates above 10 Hz can be used. In the DIVA system a rate of 20 Hz is typically applied.

In a changing environment there are a few different situations which may cause recalculation of auralization parameters [P1,P3]. The main principle in the updating process is that the system must respond within a tolerable latency to any changes in the environment. That is reached by gradually refining calculation. In the first phase only the direct sound is calculated and its parameters are passed to the auralization process. If there are no other changes waiting to be processed, first-order reflections are calculated and then second-order, and so on.

In Table 3.1, the different cases concerning image source updates are listed. If the sound source moves, all image sources must be recalculated. The same also applies to the situation when reflecting walls in the environment move. Whenever the listener moves the visibilities of all image

	Recalculate locations	Recheck visibilities	Update orientations
Change in the room geometry	X	X	X
Movement of the sound source	X	X	X
Turning of the sound source			X
Movement of the listener		X	X
Turning of the listener			X

Table 3.1: Required recalculations of image sources in an interactive system [P3].

sources must be validated. The locations of the image sources do not vary and therefore there is no need to recalculate them. If the listener turns without changing position there are no changes in the visibilities of the image sources. Only the azimuth and elevation angles must be recalculated.

During listener or source movements there are often situations where some image sources abruptly become visible while some others become invisible. This is due to the assumption that sources are infinitely small points and also due to the lack of diffraction in the acoustic model. The changes in visibilities must be auralized smoothly to avoid discontinuities in the output signal, causing audibly disturbing clicks. The most straightforward method is to fade in the new sources and fade out the ones which become invisible.

Lower and upper limits for the duration of fades are determined by auditory perception. If the time is too short, the fades are observed as clicks. In practice the upper limit is dictated by the rate of updates. In the DIVA system, the fades are performed according to the update rate of all auralization parameters. In practice, 20 Hz has been found to be a good value.

Interpolation of Auralization Parameters

In an interactive simulation the auralization parameters change whenever there is a change in listener's location or orientation in the modeled space. There are various methods of how the changes can be utilized. The topic of interpolating and commuting filter coefficients in auralization systems is discussed, for example, by Jot et al. [73]. The methods described below are applicable if the update rate is high enough, e.g., 20 Hz as in the DIVA system. Otherwise more advanced methods including prediction should be employed in order to keep latencies tolerable.

The interpolation strategy of the DIVA system [P3] is based on ideas presented by Foster et al. [74] and Begault [1]. The main principle in all the parameter updates is that the change must be performed so smoothly that the listener cannot perceive the exact update time.

Updating Filter Coefficients

In the DIVA system, coefficients of all the filters are updated immediately each time a new auralization parameter set is received. The filters for each image source include:

- Sound source directivity filter
- Air absorption filter
- HRTF filters for both ears

For the source directivity and air absorption the filter coefficients are stored with such a dense grid that there is no need to interpolate between data points. Instead, the coefficients of closest data point are utilized. The HRTF filter coefficients are stored in a table with grid of azimuth $\theta_{grid} = 10°$ and elevation $\phi_{grid} = 15°$ angles. This grid is not dense enough that the coefficients of nearest data point could be used. Therefore the coefficients are calculated by bilinear interpolation from the four nearest available data points [1]. Since the HRTFs are minimum-phase FIRs this interpolation can be applied [2]. The interpolation scheme for point E located at azimuth angle θ and elevation ϕ is:

$$\begin{aligned} h_E(n) = \quad & (1-c_\theta)(1-c_\phi)h_A(n) + c_\theta(1-c_\phi)h_B(n) + \\ & c_\theta c_\phi h_C(n) + (1-c_\theta)c_\phi h_D(n) \end{aligned} \tag{3.1}$$

where h_A, h_B, h_C, and h_D are h_E's four neighboring data points as illustrated in Fig. 3.7, n goes from 1 to the number of taps of the filter, and c_θ is the azimuth interpolation coefficient $(\theta \bmod \theta_{grid})/\theta_{grid}$. The elevation interpolation coefficient is obtained similarly $c_\phi = (\phi \bmod \phi_{grid})/\phi_{grid}$.

Interpolation of Gains and Delays

All the gains and delays are linearly interpolated and changed at every sound sample between two updates. These interpolated parameters for each image source are:

- Distance attenuation gain ($1/r$-law)
- Fade-in and fade-out gains

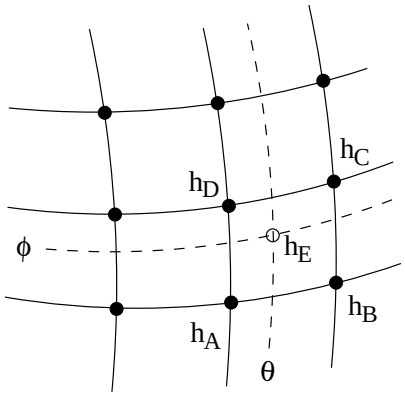

Figure 3.7: HRTF filter coefficients corresponding to point h_E at azimuth θ and elevation ϕ are obtained by bilinear interpolation from measured data points h_A, h_B, h_C, and h_D [P3].

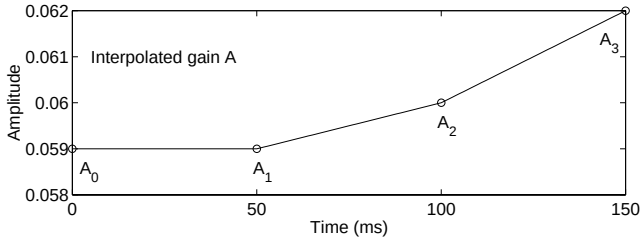

Figure 3.8: Interpolation of the amplitude gain A. Interpolation is done by linear interpolation between the key values of gain A which are marked with circles [P3].

- Propagation delay
- ITD

Interpolation in different cases is illustrated in Figs. 3.8, 3.9, and 3.10. In all of the examples the update rate of auralization parameters and thus also the interpolation rate is 20 Hz, i.e., all interpolations are done in a period of 50 ms. The linear interpolation of the gain factor is straightforward. This technique is illustrated in Fig. 3.8, where the gain is updated at times 0 ms, 50 ms, 100 ms, and 150 ms from value A_0 to A_3.

Interpolation of delays, namely the propagation delay and ITD, deserve a more thorough discussion. Each time the listener moves closer to or further from the sound source the propagation delay changes. In terms of implementation, it means a change in the length of a delay line. In the DIVA system, the interpolation of delays is performed in two steps. The applied technique is presented in Fig. 3.9. The figure represents a sampled signal in a delay line. The delay is linearly changed from value D_1 to the

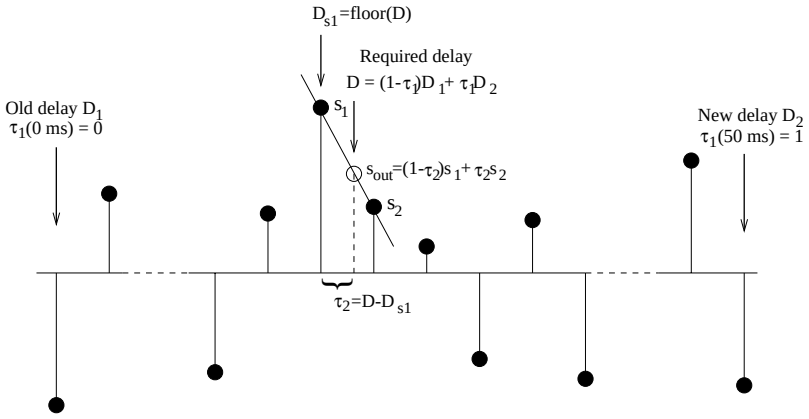

Figure 3.9: In the interpolation of delays a first-order FIR fractional delay filter is applied. In the figure there is a delay line that contains samples (...,s_1,s_2,...) and output value s_{out} corresponding to the required delay D is found as a linear combination of s_1 and s_2 [P3].

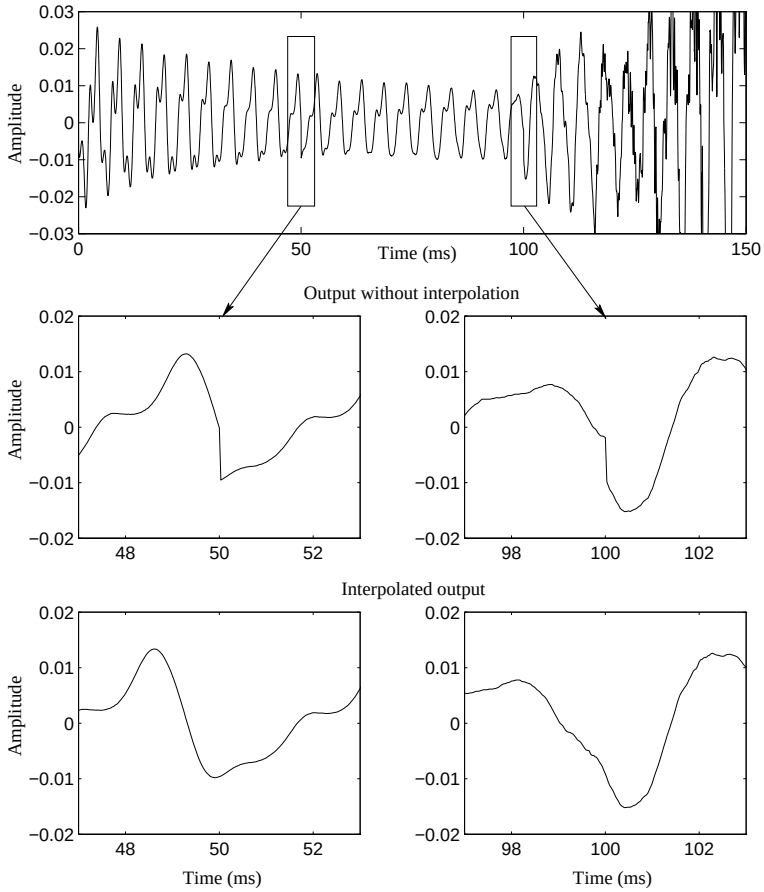

Figure 3.10: The need for interpolation of delay can be seen at updates occurring at 50 ms and 100 ms. In the first focused set there is a clear discontinuity in the signal at those times. The lower figure is the continuous signal obtained with interpolation [P3].

new value D_2 such that interpolation coefficient τ_1 linearly goes from 0 to 1 during the 50 ms interpolation period and D represents the required delay at each time. In the first step the interpolated delay D is rounded so that two neighboring samples are found (samples s_1 and s_2 in Fig. 3.9). In the second step a first-order FIR fractional delay filter with coefficient τ_2 is employed to obtain the final interpolated value (s_{out} in Fig. 3.9).

Accurate implementation of fractional delays would need a higher order filter [75]. The linear interpolation is found to be good enough for the purposes of the DIVA system, although it introduces some low-pass filtering. To minimize the low-pass filtering the fractional delays are applied only when the listener moves. At other times, the sample closest to the exact delay is used to avoid low-pass filtering. This same technique is applied with the ITDs.

In Fig. 3.10 there is a practical example of interpolation of delay and gain. There are two updates at times 50 ms and 100 ms. By examining the waveforms one can see that without interpolation there is a discontinuity in the signal while the interpolated signal is continuous.

The applied interpolation technique results in Doppler effect at fast movements [76]. Without the interpolation of delay each update would probably result in a transient sound. A constant update rate of parameters is essential to produce a natural sounding Doppler effect. Otherwise some fluctuation to the perceived sound is introduced. This corresponds to a situation in which the observer moves at alternating speed [P3].

3.4 Latency

The effect of the update rate of auralization parameters, latency, and spatial resolution of HRTFs on perceived quality have been studied by, e.g., Sandvad [72] and Wenzel [77, 78]. From the perceptual point of view, the most significant parameters are the update rate and latency. The two are not completely independent variables, since a slow update rate always introduces additional time lag. The above mentioned studies are focused on the accuracy of localization, and Sandvad [72] states that the latency should be less than 100 ms. In the DIVA system, the dynamic localization accuracy is not a crucial issue, the most important factor being the latency between visual and aural outputs when either the listener or some sound source moves thus causing a change in auralization parameters. According to observations made with the DIVA system this time lag can be slightly larger than 100 ms without noticeable drop in perceived quality. Note that this statement holds only for updates of auralization parameters in this particular application, in other situations the synchronization between visual and aural outputs is much more critical such as lip synchronization with facial animations or if more immersive user interfaces, such as head-mounted displays, are used (see, e.g., [79, 80]).

The major components of latency in the auralization of the DIVA system are the processing and data transfer delays and bufferings. The total accumulation chain of these is shown in Fig. 3.11. The latency estimates presented in following are based on simulations made with a similar configuration as in Fig. 1.2 containing one Silicon Graphics O2 and two Octanes connected by 10Mb Ethernet. The numbers shown in Fig. 3.11 represent typical values, not the worst case situations.

Delays in Data Transfers

There are three successive data transfers before a user's movement is heard as a new soundscape. The transfers are:

- GUI sends data of user's action to image-source calculation
- Image-source calculation sends new auralization parameters to the auralization unit
- Auralization sends new auralized material to sound reproduction

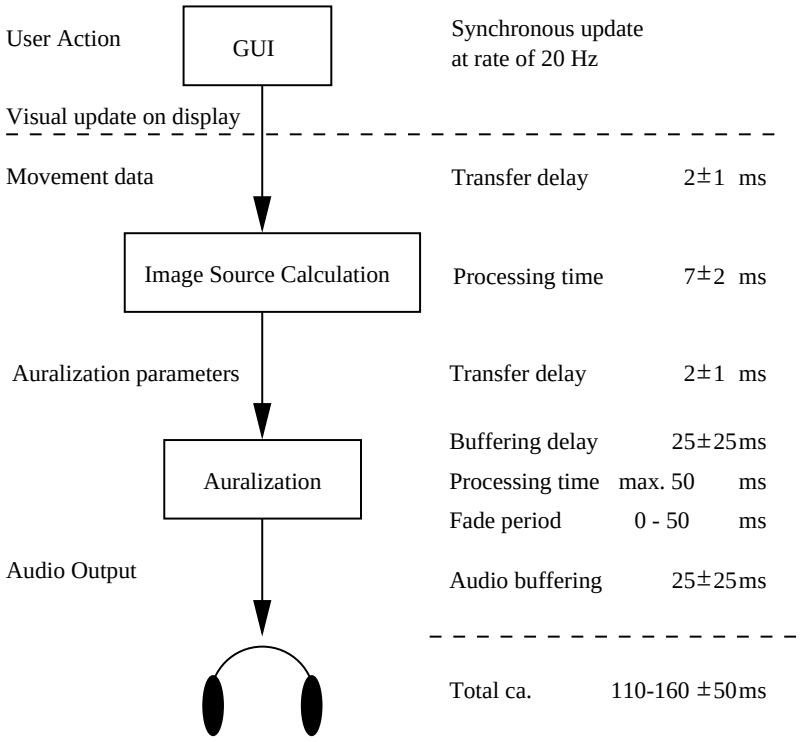

Figure 3.11: There are various components in the DIVA system which introduce latency to the system. The most significant ones are caused by bufferings [P3].

The two first data transfers are realized by sending one datagram message through the Ethernet network. A typical duration of one transfer is 1-3 ms. Occasionally, much longer delays of up to 40 ms may occur. Some messages may even get lost or duplicated due to the chosen communication protocol. Fortunately, in an unoccupied network these cases are rare. The third transfer is implemented as a memory to memory copy instruction inside the computer and the delay is negligible.

Buffering

In the auralization unit, the processing of audio is buffered. The reason for such audio buffering lies in the system performance. Buffered reading and writing of audio sample frames is computationally cheaper than performing these operations sample per sample. Another reason is the UNIX operating system. Since the operating system is not designed for strict real-time performance and the system is not running in a single-user mode, there may be other processes which occasionally require resources. Currently, an audio buffering for reading, processing and writing is done with an audio block size of 50 ms. The latency introduced by this buffering is between 0 ms and 50 ms due to the asynchronous updates of auralization parameters.

In addition to buffered processing of sound material, the sound repro-

duction must also be buffered due to the same reasons described above. Currently an output buffer of 100 ms is used. At worst this can introduce an additional 50 ms latency, when processing is done with 50 ms block size.

Delays Caused by Processing

In the DIVA system, there are two processes, the image-source calculation and auralization, which contribute to the latency occurring after the GUI has processed a user's action.

The latency caused by the image-source calculation depends on the complexity of room geometry and number of required image sources. As a case-study, a concert hall with ca. 500 surfaces was simulated and all the first order reflections were searched. A listener movement causing visibility check for the image sources took 5-9 ms depending on the listener's location.

The processing time of one audio block in auralization must be on average less than or equal to the length of the audio block, otherwise the output buffer underflows and this is heard as a disturbing click. Thus the delay caused by actual auralization is less than 50 ms in the DIVA system.

In addition, the fades increase the latency. Each change is fully applied only after the complete interpolation period of 50 ms in the DIVA system.

Total Latency

Altogether in a typical situation the DIVA system runs smoothly and produces continuous output with $110 - 160 \pm 50$ ms latency as illustrated in Fig. 3.11. However, in the worst case the latency is more than 200 ms which may cause a buffer underflow resulting in a discontinuity in the auralized sound. If less latency is required, the buffers, which cause most of the delays, can be shortened. However, then the risk of discontinuous sound is increased.

The latency depends much on the underlying hardware platform including both computers and the network. With dedicated servers and network the system can run with remarkably shorter delays, since the audio buffers can be kept shorter [P3].

The digital waveguide mesh method is a variant of the FDTD methods, but it originates in the field of digital signal processing (DSP) [81, 82]. It is designed for efficient implementation and it also provides the possibility to expand the method with other DSP techniques.

In this chapter, the basics of digital waveguide mesh method are presented and the accuracy of various mesh topologies is analyzed. An optimally interpolated structure is introduced as well as a frequency-warping technique which enhances the frequency accuracy of the method. Also the issue of boundary conditions is discussed. The discussion mainly considers the two-dimensional mesh, which can be employed, e.g., for the simulation of vibrating plates. The three-dimensional version is suitable for room acoustic simulations, for example. The author has applied the digital waveguide mesh method for simulation of a loudspeaker enclosure [83], in the design of a listening room [84, 85], and in simulation of various diffuser structures [86]. Other application areas of digital waveguide meshes include, e.g., simulation of musical instruments [87, 88, 89, 90], and usage of a 2-D mesh as a multi-channel reverberator [91, 92, 93].

4.1 Digital Waveguide

The main contributor to the theory of digital waveguides has been Smith [94, 5, 95]. The method was first applied to artificial reverberation [96], although nowadays the main application is simulation of one-dimensional resonators [5, 97, 3, 6, 7].

A digital waveguide is a bidirectional digital delay line as illustrated in Fig. 4.1. The state of the system is obtained by summing the two wave components traveling to opposite directions

$$y(x,t) = y^+(x,t) + y^-(x,t). \tag{4.1}$$

It can be shown that the digital waveguide is an exact solution to the wave equation up to the Nyquist limit.

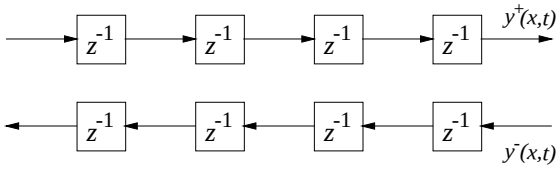

Figure 4.1: A digital waveguide is formed of two delay lines with opposite wave propagation directions. The sum of the two components $y^+(x,t)$ and $y^-(x,t)$ represents the state of the digital waveguide at location x at time t.

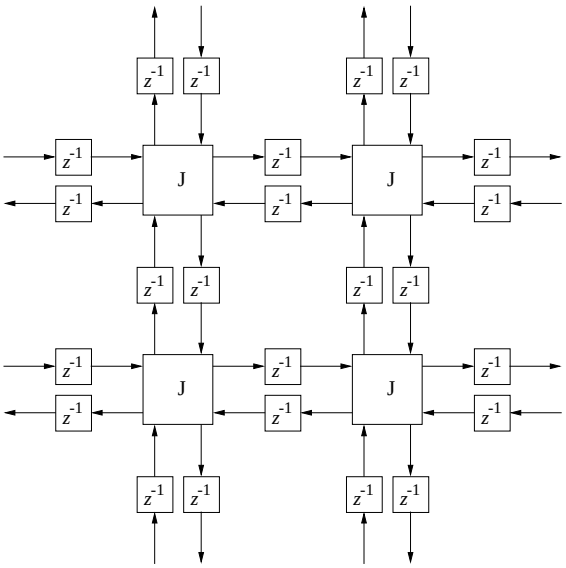

Figure 4.2: In the original 2-D digital waveguide mesh each node is connected to four neighbors with unit delays [82][P10].

4.2 Digital Waveguide Mesh

A digital waveguide mesh [81, 82] is a regular array of discretely spaced 1-D digital waveguides arranged along each perpendicular dimension, interconnected at their intersections. A two-dimensional case is illustrated in Fig. 4.2. The resulting mesh of a 3-D space is a regular rectangular grid in which each node is connected to its six neighbors by unit delays [81, 98][P4].

Two conditions must be satisfied at a lossless junction connecting lines of equal impedance [81]:

1. Sum of inputs equals the sum of outputs (flows add to zero).
2. Signals at each intersecting waveguide are equal at the junction (continuity of impedance).

Based on these conditions the author has derived a difference equation for the nodes of an N-dimensional rectangular mesh [P4]:

$$p_k(n) = \frac{1}{N}\sum_{l=1}^{2N} p_l(n-1) - p_k(n-2) \tag{4.2}$$

where p represents the sound pressure at a junction at time step n, k is the position of the junction to be calculated and l represents all the neighbors of k. This equation is equivalent to a difference equation derived from the Helmholtz equation by discretizing time and space. The update frequency of an N-dimensional mesh is:

$$f_s = \frac{c\sqrt{N}}{\Delta x} \approx \frac{588.9}{\Delta x} Hz \tag{4.3}$$

where c represents the speed of sound in the medium and Δx is the spatial sampling interval corresponding to the distance between two neighboring nodes. The approximate value stands for a typical room simulation ($c = 340m/s, N = 3$). That same frequency is also the sampling frequency of the resulting impulse response.

An inherent problem with the digital waveguide mesh method is the direction dependent dispersion of wavefronts [81, 99, 88]. High-frequency signals parallel to the coordinate axes are delayed, whereas diagonally the waves propagate undistorted. This effect can be reduced by employing structures other than the rectangular mesh, such as triangular or tetrahedral meshes [99, 87, 88, 89], or by using interpolation methods [100][P5,P7,P10].

The accuracy of the digital waveguide mesh depends primarily on the density of the mesh. Due to the dispersion, the original model is useful in high-accuracy simulations only at frequencies well below the update frequency f_s of the mesh. In the estimation of room acoustical attributes phase errors are not severe and results may be used up to one tenth of the update frequency. In that case, there are at least six mesh nodes per wavelength. For auralization purposes the valid frequency band is more limited.

4.3 Mesh Topologies

Rectangular Mesh Structure

The rectangular mesh structure, illustrated in Fig. 4.2, is the original topology proposed for a digital waveguide mesh [81, 82].

The main problem with this structure is the direction-dependent dispersion which increases with frequency. The dispersion error of the original 2-D digital waveguide mesh has been analyzed in [81, 82]. The main principle in the applied Von Neumann analysis [36] is the two-dimensional discrete-time Fourier transform of the difference scheme with sampling interval T. In the transform, the point (ξ_1, ξ_2) in the two-dimensional frequency space corresponds to the *spatial frequency* $\xi = \sqrt{\xi_1^2 + \xi_2^2}$. After the Fourier transform, the scheme can be represented by means of the *spectral amplification factor* $g(\xi_1, \xi_2)$. The stability and dispersion characteristics of the scheme may be analyzed when the amplification factor is presented as $g(\xi_1, \xi_2) = |g(\xi_1, \xi_2)|e^{-j2\pi\xi c'(\xi_1,\xi_2)T}$, where $c'(\xi_1, \xi_2)$ represents the phase velocity in the direction $\alpha = \arctan(\xi_2/\xi_1)$. For the original difference scheme, the Fourier transform is as presented by Van Duyne and Smith [82]:

$$P(n, \xi_1, \xi_2) = b(\xi_1, \xi_2)P(n-1, \xi_1, \xi_2) - P(n-2, \xi_1, \xi_2) \tag{4.4}$$

where b is

$$b(\xi_1, \xi_2) = \tfrac{1}{2}(e^{j\omega_1 cT} + e^{j\omega_2 cT} + e^{-j\omega_1 cT} + e^{-j\omega_2 cT}) = \cos(\omega_1 cT) + \cos(\omega_2 cT) \tag{4.5}$$

in which $\omega_1 = 2\pi\xi_1$ and $\omega_2 = 2\pi\xi_2$. The same equation can be expressed by means of the amplification factor $g(\xi_1, \xi_2)$ as follows:

$$g(\xi_1, \xi_2)^n P(0, \xi_1, \xi_2) = b(\xi_1, \xi_2)g(\xi_1, \xi_2)^{n-1}P(0, \xi_1, \xi_2) - g(\xi_1, \xi_2)^{n-2}P(0, \xi_1, \xi_2) \tag{4.6}$$

The amplification factor thus is

$$g(\xi_1, \xi_2) = \frac{1}{2} b(\xi_1, \xi_2) \pm \frac{1}{2} j \sqrt{4 - b(\xi_1, \xi_2)^2} \tag{4.7}$$

The difference scheme (4.2) is stable since

$$|g(\xi_1, \xi_2)| = \sqrt{\frac{1}{4} b(\xi_1, \xi_2)^2 + \frac{1}{4}[4 - b(\xi_1, \xi_2)^2]} = 1 \tag{4.8}$$

as shown in [82]. The wave propagation speed of the scheme can now be calculated from the equation

$$2\pi\xi c'(\xi_1, \xi_2) T = \angle g(\xi_1, \xi_2) = \arctan \frac{\sqrt{4 - b(\xi_1, \xi_2)^2}}{b(\xi_1, \xi_2)} \tag{4.9}$$

The desired wave propagation speed in the mesh is $c = 1/(\sqrt{2}T)$, so that waves propagate one diagonal unit in two time steps. The ratio of the actual speed to the desired speed is represented by the dispersion factor [89]

$$k(\xi_1, \xi_2) = \frac{c'(\xi_1, \xi_2)}{c} = \frac{\sqrt{2}}{2\pi\xi} \arctan \frac{\sqrt{4 - b(\xi_1, \xi_2)^2}}{b(\xi_1, \xi_2)}. \tag{4.10}$$

The dispersion factor is a function of spatial frequency. However, it is presented in Fig. 4.3 as a function of the *normalized temporal frequency* f, since the resulting audio signals are also available as a function of f. In the figure, the center point $(0, 0)$ represents the DC component, and the distance from the center is directly proportional to the actual temporal frequency such that $f = c\xi$, and $\angle(\xi_1, \xi_2)$ determines the propagation direction.

There is no dispersion in diagonal directions ($\alpha_{min,rect} = \angle(\xi_1, \xi_2) = \pi/4 \pm k\pi/2$, where $k = 0, 1,$ or 2), but in all other directions there is some dispersion, as illustrated in Fig. 4.3. The maximal error is obtained in the axial directions ($\alpha_{max,rect} = \angle(\xi_1, \xi_2) = \pm k\pi/2$, where $k = 0, 1,$ or 2). The response of the rectangular digital waveguide mesh begins to repeat at normalized temporal frequency 0.25, represented by an emphasized circle in Fig. 4.3(b) [81, 99]. Figure 4.4(a) illustrates maximal and minimal relative frequency errors (RFE) in the original structure. The RFE is obtained as follows:

$$E(\xi_1, \xi_2) = \frac{k(\xi_1, \xi_2) - k_{DC}}{k_{DC}} \cdot 100\% \tag{4.11}$$

where $k_{DC} = \lim_{\xi_1, \xi_2 \to 0} k(\xi_1, \xi_2)$. Note that the result is presented as a function of normalized spatial frequency $f = c\xi$.

Triangular Mesh Structure

Alternative sampling lattices have been studied to obtain more uniform wave propagation characteristics in all directions. When the sampling of the surface is hexagonal such that $N = 6$ in (4.2), the triangular digital waveguide mesh is obtained [87, 88, 89], as illustrated in Fig. 4.5. This structure has better dispersion characteristics than the rectangular mesh. The same dispersion analysis as presented for the rectangular mesh is valid

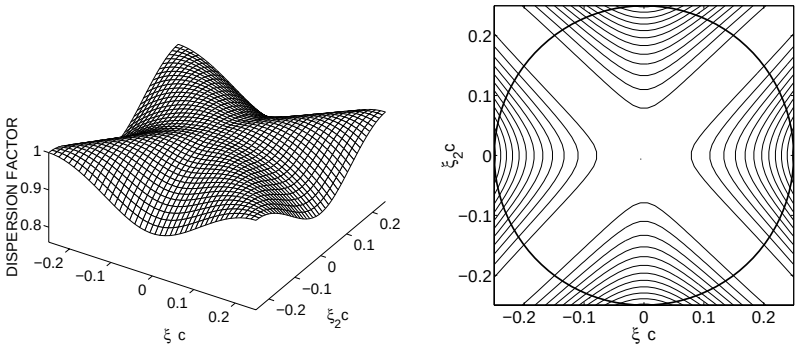

Figure 4.3: The wave propagation speed in the rectangular digital waveguide mesh such that the distance from the center is directly proportional to the frequency, and the ratio of ξ_1 and ξ_2 (with the sign) determines the propagation direction. In (a) the value 1 represents the ideal speed which is achieved in diagonal directions at all frequencies, but in all other directions there is dispersion. The equal dispersion factor contours in (b) descend from the center point in increments of 1% [P10].

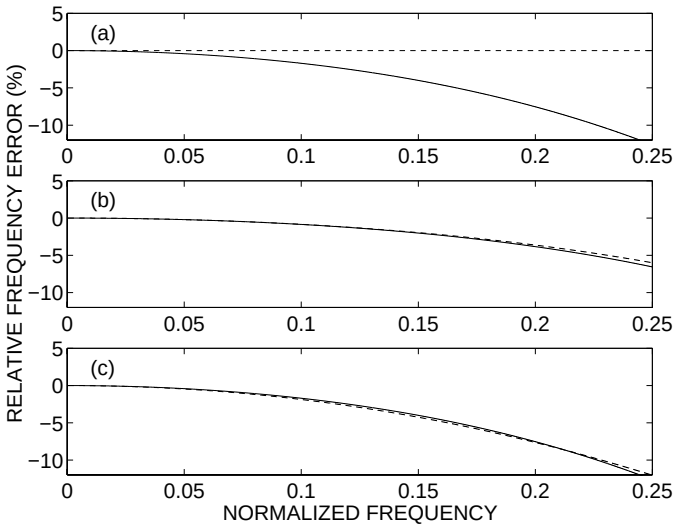

Figure 4.4: Maximal and minimal relative frequency errors are represented by the solid and dotted lines, respectively, for the (a) rectangular, (b) triangular, and (c) optimally interpolated digital waveguide mesh structures as a function of normalized temporal frequency, where 0.5 corresponds to the Nyquist limit. Note that in the original rectangular structure there is no error in the diagonal direction.

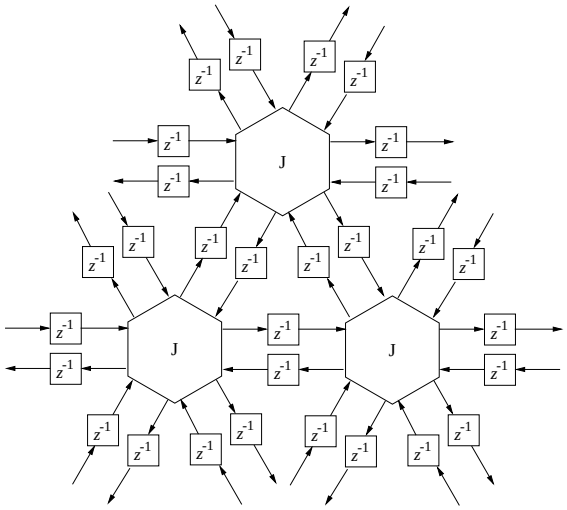

Figure 4.5: In the triangular digital waveguide mesh each node is connected to six neighbors with unit delays [P10].

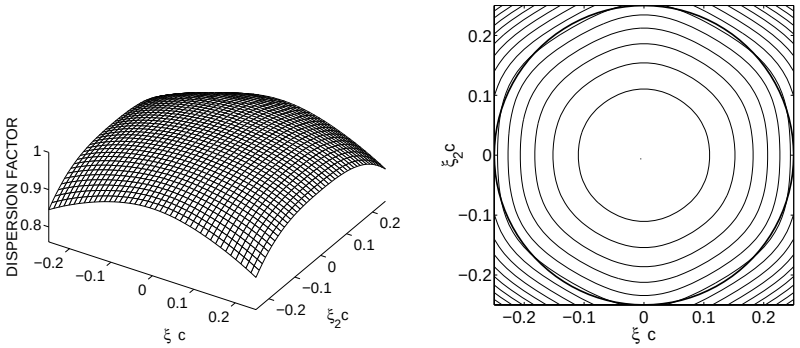

Figure 4.6: Wave propagation speed in the triangular digital waveguide mesh [P10].

for the triangular mesh. The only difference is that b in (4.10) obtains a new formulation:

$$b(\xi_1, \xi_2) = \tfrac{2}{3}[\cos(\omega_1 cT) + \cos(\omega_1 cT/2 + \sqrt{3}\omega_2 cT/2) + \cos(\omega_1 cT/2 - \sqrt{3}\omega_2 cT/2)] \tag{4.12}$$

The resulting dispersion factor presented in Fig. 4.6 shows that there is some dispersion in all directions, but now it is nearly independent of propagation direction. The maximal and minimal RFEs illustrated in Fig. 4.4(b) are obtained in directions $\alpha_{max,tri} = \angle(\xi_1, \xi_2) = \pm k\pi/3$ and $\alpha_{min,tri} = \angle(\xi_1, \xi_2) = \pi/6 \pm k\pi/3$, respectively, where $k = 0, 1, 2$, or 3.

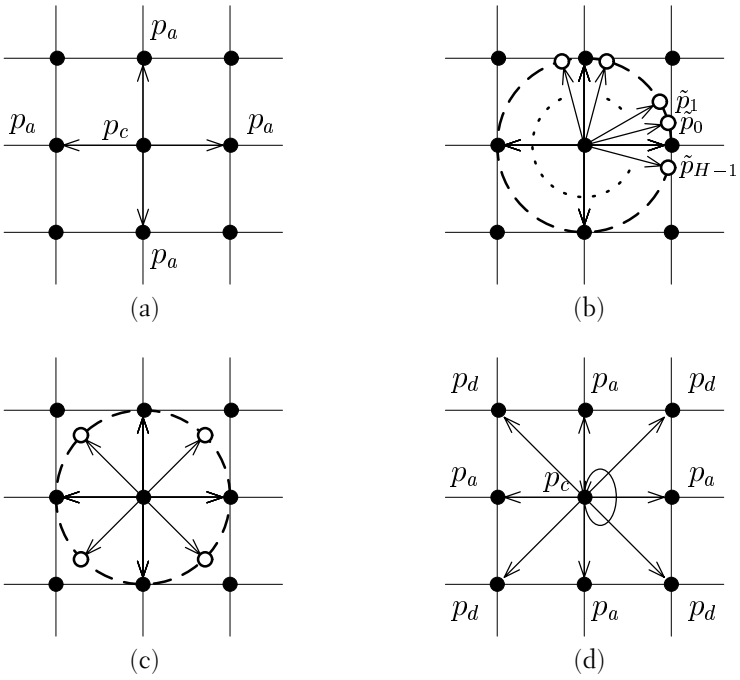

Figure 4.7: The construction phases of the interpolated 2-D digital waveguide mesh structure: (a) the original mesh with 4 directions, (b) the hypothetical version with H propagation directions, (c) the case $H = 8$ for which the best results have been obtained, and (d) the new interpolated waveguide mesh, in which the new nodes are spread onto the neighboring nodes by deinterpolation [101, 97, 100][P5,P10].

Interpolated Rectangular Mesh Structure

In the 2-D rectangular digital waveguide mesh algorithm, each node has four neighbors, two on both axes (points labelled p_a in Fig. 4.7(a)). These are connected by unit delay elements. The author has extended this scheme to an arbitrary number of neighbors by adding unit delay lines from a node into other directions (hypothetical points labelled $\tilde{p}_0 \ldots \tilde{p}_{H-1}$ in Fig. 4.7(b)) such that there are altogether H propagation directions [P5].

Although it is not possible to exactly implement the hypothetical structure with H directions in the time domain, the two-dimensional Fourier transform may be computed, that is

$$P(n, \xi_1, \xi_2) = \frac{2}{H} \sum_{i=0}^{H-1} e^{j(\omega_1 \cos\alpha_i + \omega_2 \sin\alpha_i)cT} P(n-1, \xi_1, \xi_2) - P(n-2, \xi_1, \xi_2) \tag{4.13}$$

where α_i are the directions of delay lines. The equation for the speed ratio

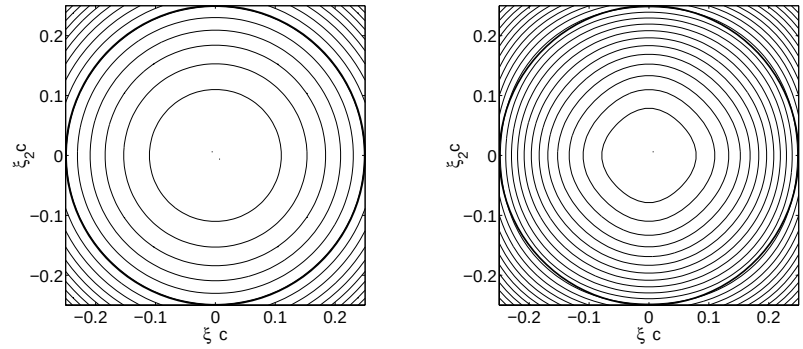

Figure 4.8: The dispersion factor (a) in the hypothetical 8-directional, and (b) in optimally interpolated digital waveguide mesh structures. The equal dispersion factor contours descend from the center point in increments of 1% [P10].

is the same as (4.10), where b is

$$b(\xi_1, \xi_2) = \frac{2}{H} \sum_{i=0}^{H-1} e^{j(\omega_1 \cos\alpha_i + \omega_2 \sin\alpha_i)cT} \tag{4.14}$$

The contour plot in Fig. 4.8(a) shows the speed ratio as a function of frequency and direction of an 8-directional structure corresponding to the situation in Fig. 4.7(c). In this structure there is some dispersion, but now it is practically independent of wave propagation direction. The 8-direction structure was chosen since the addition of delay lines in more than 8 directions does not introduce much improvement [P5].

This new structure is called the interpolated rectangular mesh [100][P5]. It can be implemented by using various techniques. For first-order and second-order interpolation schemes, the difference equation is

$$p_c(n) = \frac{2}{H} \sum_{l=1}^{3} \sum_{k=1}^{3} h_{l,k} p_{l,k}(n-1) - p_c(n-2) \tag{4.15}$$

where $p_{l,k}$ represents p_c and all its neighbors p_a and p_d (see Fig. 4.7(d)), and $h_{l,k}$ are the weighting coefficients of each node. Due to symmetry, the diagonal neighbors have the same weighting coefficient $h_d = h_{11} = h_{13} = h_{31} = h_{33}$, and all axial nodes have the coefficient $h_a = h_{12} = h_{21} = h_{32} = h_{23}$. The center node p_c is weighted by $h_c = h_{22}$.

The two-dimensional Fourier transform for the interpolated structure (4.15) is

$$\begin{gathered} P(n, \xi_1, \xi_2) = \\ \tfrac{2}{H} P(n-1, \xi_1, \xi_2)[h_a(e^{j\omega_1 cT} + e^{j\omega_2 cT} + e^{-j\omega_1 cT} + e^{-j\omega_2 cT}) + \\ h_d(e^{j\delta_+ cT} + e^{j\delta_- cT} + e^{-j\delta_- cT} + e^{-j\delta_+ cT}) + h_c] - P(n-2, \xi_1, \xi_2) \end{gathered} \tag{4.16}$$

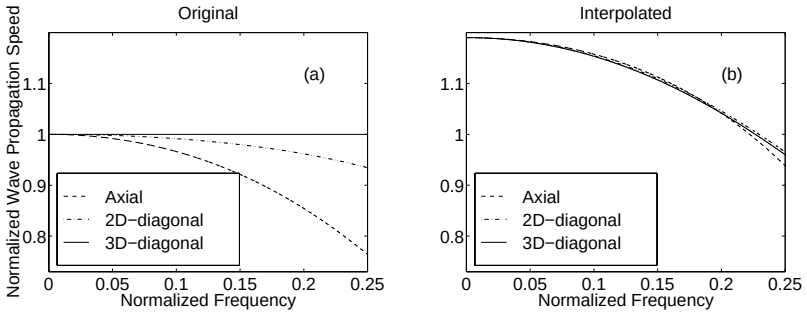

Figure 4.9: The relative wave propagation speed in axial, 2D-diagonal and 3D-diagonal directions (a) in the original and (b) in the new interpolated three-dimensional structure.

where $\delta_+ = \omega_1 + \omega_2$ and $\delta_- = \omega_1 - \omega_2$. For the speed ratio (4.10), b is

$$b(\xi_1, \xi_2) = \tfrac{4}{H}[h_a(\cos\omega_1 cT + \cos\omega_2 cT) + h_d(\cos\delta_+ cT + \cos\delta_- cT) + \tfrac{h_c}{2}] \tag{4.17}$$

Both the first-order and second-order linear interpolation can be applied to find the weighting coefficients h_a, h_d, h_c, but the best result is obtained by optimized interpolation coefficients as shown in [P10]. The resulting weighting coefficients are

$$h_d = 0.375930, h_a = 1.24814, h_c = 1.50372, \tag{4.18}$$

when $H = 8$.

Figure 4.8(b) presents the wave propagation characteristics of this optimized interpolated scheme. In this mesh, the wave propagation speed is nearly independent of direction, although the speed is lower at high frequencies. The result is better than the ones obtained using first- or second-order Lagrange interpolation, but not as good as the triangular mesh, which is seen by comparing maximal and minimal RFEs presented in Figs. 4.4(b) and 4.4(c) [P10].

The same interpolation technique can also be applied in the 3-D case. Some results of using linear interpolation in three dimensions to obtain the weighting coefficients are presented in [P7]. The best result was achieved by adding delay lines to 2-D and 3-D diagonal directions resulting in $H = 26$ propagation directions. In Fig. 4.9 the relative wave propagation speeds in three different directions are illustrated. Figure 4.9(a) depicts the original structure. Fig. 4.9(b) illustrates the linearly interpolated structure in which the wave propagation characteristics are nearly independent of direction. In the future, similar coefficient optimization as in the 2-D case [P10] should be conducted.

4.4 Reduction of the Dispersion Error by Frequency Warping

In the triangular and interpolated rectangular mesh structures the dispersion has been rendered practically independent of propagation direction.

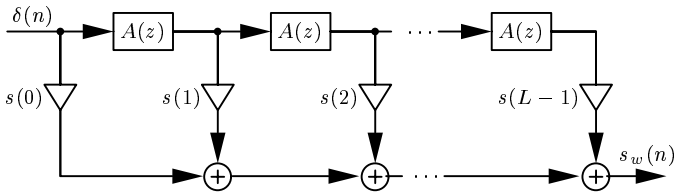

Figure 4.10: The allpass-filter structure which is used in the warping of a simulation result of the interpolated digital waveguide mesh [P8].

In addition, the frequency error curves are smooth. Due to these reasons, the dispersion error can be reduced remarkably by a novel use of a frequency warping technique [P8,P9].

The frequency warping is applied to the input and output signal of the mesh using a warped FIR filter [102, 103]. This filter shifts frequencies to compensate for the dispersion error. For this purpose a first-order allpass warping is suitable.

A warped FIR is a FIR filter, in which each unit delay element has been replaced with a first-order allpass filter having the transfer function:

$$A(z) = (z^{-1} + \lambda)/(1 + \lambda z^{-1}) \tag{4.19}$$

The resulting structure is illustrated in Fig. 4.10.

The filter coefficient λ determines the extent of warping, and the same value of λ is used for all the allpass filters in the chain. The warped FIR is used by setting the tap coefficients equal to signal samples $s(n)$ and then a unit impulse is fed into the structure. The output signal $s_w(n)$ is the frequency-warped version of the original signal.

There are various different strategies for finding the optimal value of λ [P8]. One practical choice is to minimize the maximal error in a frequency band $[0, 0.25 f_s]$. This results in $\lambda = -0.327358$ for the interpolated mesh, and for the triangular mesh $\lambda = -0.109540$. The RFE after warping is obtained as follows:

$$E_{warped}(\xi_1, \xi_2) = \frac{k(\xi_1, \xi_2) \cdot w_{ratio}(\lambda, \xi \cdot k(\xi_1, \xi_2)) \cdot D - k_{dc}}{k_{dc}} \cdot 100\% \tag{4.20}$$

where

$$w_{ratio}(\lambda, \xi) = 1 + \frac{1}{\pi\xi} \arctan \frac{\lambda \sin 2\pi\xi}{1 - \lambda \cos 2\pi\xi} \tag{4.21}$$

returns the ratio of warping of given frequency ξ [104], and

$$D = \frac{1 - \lambda}{1 + \lambda} \tag{4.22}$$

is the phase delay at low frequencies caused by warping [75, eq. 86].

The RFE of each structure with warping is presented in Fig. 4.11. The corresponding RFEs without warping are presented in Figs. 4.4(a), 4.4(b), and 4.4(c). Note that in the original rectangular mesh there is no reasonable way to apply frequency warping since the RFEs in the axial and diagonal directions are so different. Of these structures, the triangular mesh is

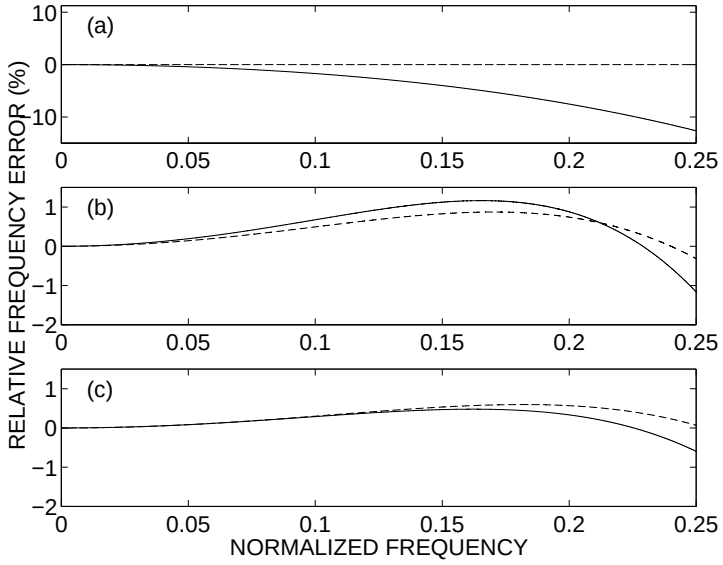

Figure 4.11: Maximal and minimal relative frequency errors represented by solid and dashed lines, respectively, in the (a) rectangular, (b) warped interpolated ($\lambda = -0.327358$), and (c) warped triangular ($\lambda = -0.109540$) digital waveguide mesh structures. The warping coefficients were optimized to minimize the maximal error. Note the different scale in (a) [P10].

the most accurate having the maximal error of 0.60% (Fig. 4.11(c)). The interpolated rectangular mesh is nearly as good as the triangular mesh resulting in a 1.2% maximal error (Fig. 4.11(b)). Both of these structures are more accurate than the original structure in which the maximal error is 13% (Fig. 4.11(a)) [P10].

4.5 Boundary Conditions

The boundary conditions describe the acoustic properties of surface materials and they have a remarkable contribution to the acoustics of a space.

The most straightforward way to set boundary conditions in the digital waveguide mesh, is the structure in which each boundary node has only one neighbor.

The simplest boundary condition is the reflection coefficient. The difference equation for such boundary node is [98]:

$$p(n) = (1 + r)p_{opp}(n - 1) - rp(n - 2) \tag{4.23}$$

where r is the reflection coefficient, p is the boundary node and p_{opp} is its neighbor.

In general, the boundary condition can be replaced by any digital filter as shown by the author in [P6]. Figure 4.12 [P6] represents the numerical simulation results of a 2-D digital waveguide mesh with two different boundary filters and incident angles of the reflected plane wave. Figures

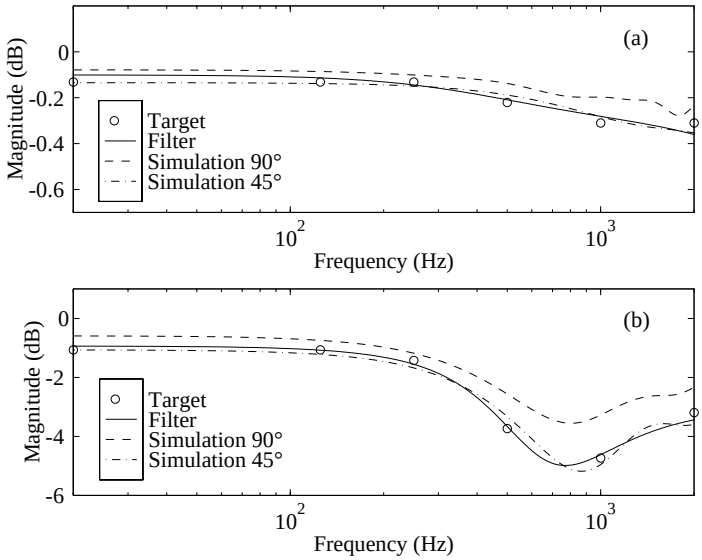

Figure 4.12: Reflection responses of two 2-D digital waveguide mesh simulations. In (a) the wall is hard and the simulation is close to the target response. In (b) the wall is more absorbing, and the difference between the incident angle of 45° and the perpendicular reflection is remarkable [P6].

show the magnitude attenuation of boundary filter and simulation results. The target response is also plotted. In Fig. 4.12(a), the wall is hard and target absorption is low and both simulation results follow the filter response reasonably well. In Fig. 4.12(b) the boundary material is more absorbing. The simulation result of 90° incident angle is not as good as in the upper figure, but the shape of the attenuation response is correct. Simulation of incident angle of 45° is correct. Note that the same material filters can be utilized both in the DIVA system (see Section 3.2) and in the digital waveguide mesh.

The results presented in Fig. 4.12 are obtained using the original rectangular mesh in which the wave propagation characteristics depend on the propagation direction. The same also applies to the reflection characteristics. To obtain direction independent reflection properties, use of a larger number of neighborhood nodes at the boundaries is required. This can be implemented using similar interpolation such as in the interpolated rectangular mesh, but it has not been studied yet. One method to implement diffuse reflections in a digital waveguide mesh is presented in [105].

5.1 Main Results of the Thesis

The main results of the thesis can be summarized as follows:

- The parametric room impulse response rendering is found to be an efficient method for real-time auralization.
- The image-source method with an advanced visibility checking technique is an efficient method for computing early reflections in an interactive auralization system.
- Natural sounding auralization was achieved within the DIVA system employing auralization parameters applied to the direct sound and early reflections. These include distance, direction and orientation, and acoustic properties of materials which have reflected the sound.
- For an interactive virtual acoustic environment to be realistic, all the changes in gains and delays must be smooth. This is achieved by employing interpolation. With fast movements this creates a natural Doppler effect if the auralization parameters are updated at a constant rate.
- For an immersive virtual orchestra application, 150 ms latency between visual and aural updates is not disturbing according to preliminary observations obtained with the DIVA system.
- The optimally interpolated digital waveguide mesh performs almost equally well as the triangular mesh, but is easier to implement and rectangular tessellation is more simply performed.
- Frequency warping remarkably increases the frequency accuracy of off-line digital waveguide mesh simulations.

5.2 Contribution of the Author

The author of this thesis is the primary author of all the publications [P1]-[P10], with the exception of [P6]. None of these publications have previously formed a part of another thesis. The scientific contribution of the articles is following:

[P1] The real-time auralization system, DIVA, is presented. Novelties of the system are: real-time image-source method, auralization parameters including their update and interpolation systems, and distributed implementation.

[P2] Efficient implementation of the image-source method for a real-time application is presented. The basic principle of multiple simultaneous sound sources in a virtual acoustic environment is presented.

[P3] The article discusses in detail the DIVA system, the applied auralization parameters and their interpolation and updates. Implementation of a real-time image-source method in a room with arbitrary geometry employing an efficient visibility checking strategy is described. Analysis of the latency of the DIVA system is also provided.

[P4] The article presents an N-dimensional digital waveguide mesh. A new three-dimensional rectangular mesh is especially interesting since it suits the application of room acoustic simulation.

[P5] The bilinearly interpolated rectangular digital waveguide mesh is presented.

[P6] The article describes how arbitrary filters can be used as boundary conditions in a digital waveguide mesh.

[P7] A new three-dimensional interpolated rectangular digital waveguide mesh structure is introduced.

[P8] The novel use of frequency warping to reduce the dispersion error in the interpolated rectangular digital waveguide mesh is shown.

[P9] The frequency warping technique is applied to the triangular digital waveguide mesh to enhance the frequency accuracy.

[P10] The author shows new interpolation techniques for the rectangular digital waveguide mesh structure. The article also compares the frequency accuracy of various structures.

In article [P1] the author has written Section 2 (except for 2.4), and one half of Sections 1, 4, and 5. In article [P2] the author has written 60% of the text. In article [P3] the author has written Section 2 (except for 2.5-2.7). Of Sections 0, 4, 5, and 7-9 the author has written 50% of the text.

The presented new ideas in [P1]-[P3] are the result of co-operation of the author with Dr. Huopaniemi. The author has implemented the ray-tracing and image-source methods as well as the base of the auralization system described in this thesis.

In Articles [P4],[P5], [P8]-[P10] the author has provided all the simulations, created all the figures, and written 60% of the text. Of Article [P6] the author has written Section 4 and a half of Section 5. The author is the sole author of [P7].

The ideas of the interpolated rectangular mesh and the use of frequency warping in the digital waveguide mesh are results of co-operation with Dr. Välimäki.

5.3 Future Work

In the future, the computational efficiency of the DIVA system should be increased. This can be done by perceptual optimization based on listening tests. The importance of the following factors should be evaluated:

- Reflection filters

- Number of auralized early reflections
- Air absorption filters
- Directivity filtering of auralized reflections

In this thesis, the results of the digital waveguide mesh contribute mainly to the 2-D case. By applying these improvements also to the 3-D case, the method could also be a practical tool for room acoustic simulations. The most important things which should be studied are:

- Optimally interpolated rectangular three-dimensional digital waveguide mesh
- Direction independent boundary conditions using an interpolation technique

[1] D. Begault. *3-D Sound for Virtual Reality and Multimedia*. Academic Press, Cambridge, MA, 1994.

[2] J. Huopaniemi. *Modeling of Human Spatial Hearing in the Context of Digital Audio and Virtual Environments*. Licentiate Thesis, Helsinki University of Technology, June 1997.

[3] V. Välimäki and T. Takala. Virtual musical instruments – natural sound using physical models. *Organised Sound*, 1(2):75–86, 1996.

[4] J. Huopaniemi, M. Karjalainen, V. Välimäki, and T. Huotilainen. Virtual instruments in virtual rooms – a real-time binaural room simulation environment for physical models of musical instruments. In *Proc. Int. Computer Music Conf. (ICMC'94)*, pages 455–462, Aarhus, Denmark, 12-17 Sept. 1994.

[5] J. O. Smith. Physical modeling using digital waveguides. *Computer Music J.*, 16(4):74–87, 1992 Winter.

[6] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy. Physical modeling of plucked string instruments with application to real-time sound synthesis. *J. Audio Eng. Soc.*, 44(5):331–353, May 1996.

[7] J. O. Smith. Principles of digital waveguide models of musical instruments. In M. Kahrs and K. Brandenburg, editors, *Applications of Digital Signal Processing to Audio and Acoustics*, chapter 10, pages 417–466. Kluwer Academic, Boston, MA, 1997.

[8] J. Meyer. *Acoustics and the Performance of Music*. Verlag das Musikinstrument, Frankfurt/Main, 1978.

[9] N. H. Fletcher and T. D. Rossing. *The Physics of Musical Instruments*. Springer, New York, NY, 1991.

[10] H. Kuttruff. *Room Acoustics*. Elsevier Applied Science, London, UK, 3rd edition, 1991.

[11] M. Kleiner, B.-I. Dalenbäck, and P. Svensson. Auralization – an overview. *J. Audio Eng. Soc.*, 41(11):861–875, 1993 Nov.

[12] H. Møller. Fundamentals of binaural technology. *Applied Acoustics*, 36(3-4):171–218, 1992.

[13] J. Blauert. *Spatial Hearing. The Psychophysics of Human Sound Localization*. MIT Press, Cambridge, MA, 2nd edition, 1997.

[14] J. Huopaniemi. *Virtual Acoustics and 3-D Sound in Multimedia Signal Processing*. Doctoral thesis, Helsinki University of Technology, Lab. of Acoustics and Audio Signal Processing, Report 53, 1999.

[15] T. Takala, R. Hänninen, V. Välimäki, L. Savioja, J. Huopaniemi, T. Huotilainen, and M. Karjalainen. An integrated system for virtual audio reality. In *the 100th Audio Engineering Society (AES) Convention, preprint no. 4229*, Copenhagen, Denmark, 11-14 May 1996.

[16] DIVA Group: J. Hiipakka, R. Hänninen, T. Ilmonen, H. Napari, T. Lokki, L. Savioja, J. Huopaniemi, M. Karjalainen, T. Tolonen, V. Välimäki, S. Välimäki, and T. Takala. Virtual orchestra performance. In *Visual Proceedings of SIGGRAPH'97*, page 81, Los Angeles, CA, 1997. ACM SIGGRAPH.

[17] T. Lokki, J. Hiipakka, R. Hänninen, T. Ilmonen, L. Savioja, and T. Takala. Real-time audiovisual rendering and contemporary audiovisual art. *Organised Sound*, 3(3):219–233, 1998.

[18] T. Ilmonen. *Tracking Conductor of an Orchestra Using Artificial Neural Networks*. Master's thesis, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, 1999.

[19] T. Ilmonen and T. Takala. Conductor following with artificial neural networks. In *Proc. Int. Computer Music Conf. (ICMC'99)*, pages 367–370, Beijing, China, 22-27 Oct. 1999.

[20] R. Hänninen. *LibR – An Object-oriented Software Architecture for Realtime Sound and Kinematics*. Licentiate thesis, Helsinki University of Technology, Telecommunications Software and Multimedia Laboratory, 1999.

[21] T. Lokki, L. Savioja, J. Huopaniemi, R. Hänninen, T. Ilmonen, J. Hiipakka, V. Pulkki, R. Väänänen, and T. Takala. Virtual concerts in virtual spaces – in real time (invited paper). 14-19 March 1999. Paper presented in the ASA/EAA Joint Meeting, Berlin, Germany.

[22] T. Lokki, J. Hiipakka, and L. Savioja. Immersive 3-D sound reproduction in a virtual room. In *Proc. AES 16th Int. Conf. on Spatial Sound Reproduction*, pages 172–177, Rovaniemi, Finland, 10-12 April 1999.

[23] T. Takala, E. Rousku, T. Lokki, L. Savioja, J. Huopaniemi, R. Väänänen, V. Pulkki, and P. Salminen. Marienkirche - a visual and aural demonstration film. In *Electronic Art and Animation Catalogue (SIGGRAPH'98)*, page 149, Orlando, FL, 19-24 Jul. 1998. Presented in SIGGRAPH'98 Computer Animation Festival (Electronic Theater).

[24] A. Krokstad, S. Strom, and S. Sorsdal. Calculating the acoustical room response by the use of a ray tracing technique. *J. Sound Vib.*, 8(1):118–125, 1968.

[25] M. R. Schroeder. Digital simulation of sound transmission in reverberant spaces. *J. Acoust. Soc. Am.*, 47(2 (Part 1)):424–431, 1970.

[26] M. R. Schroeder. Computer models for concert hall acoustics. *Am. J. Physics*, 41:461–471, 1973.

[27] H. Kuttruff. Sound field prediction in rooms. In *Proc. 15th Int. Congr. Acoust. (ICA'95)*, volume 2, pages 545–552, Trondheim, Norway, June 1995.

[28] A. Kulowski. Algorithmic representation of the ray tracing technique. *Applied Acoustics*, 18(6):449–469, 1985.

[29] J. B. Allen and D. A. Berkley. Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.*, 65(4):943–950, 1979.

[30] J. Borish. Extension of the image model to arbitrary polyhedra. *J. Acoust. Soc. Am.*, 75(6):1827–1836, 1984.

[31] D. Botteldooren. Finite-difference time-domain simulation of low-frequency room acoustic problems. *J. Acoust. Soc. Am.*, 98(6):3302–3308, 1995.

[32] L. Savioja, J. Backman, A. Järvinen, and T. Takala. Waveguide mesh method for low-frequency simulation of room acoustics. In *Proc. 15th Int. Congr. Acoust. (ICA'95)*, volume 2, pages 637–640, Trondheim, Norway, June 1995.

[33] R. Lyon and R. DeJong. *Theory and Application of Statistical Energy Analysis.* Butterworth-Heinemann, Newton, MA, 2nd edition, 1995.

[34] L. Beranek. *Concert and Opera Halls – How They Sound.* Acoustical Society of America, New York, NY, 1996.

[35] A. Pietrzyk. Computer modeling of the sound field in small rooms. In *Proc. AES 15th Int. Conf. on Audio, Acoustics & Small Spaces*, pages 24–31, Copenhagen, Denmark, 31 Oct. - 2 Nov. 1998.

[36] J. Strikwerda. *Finite Difference Schemes and Partial Differential Equations.* Chapman & Hall, New York, NY, 1989.

[37] T. Schetelig and R. Rabenstein. Simulation of three-dimensional sound propagation with multidimensional wave digital filters. In *Proc. Int. Conf. Acoust., Speech, Signal Processing (ICASSP'98)*, volume 6, pages 3537–3540, Seattle, WA, 12-16 May 1998.

[38] A. Kulowski. Error investigation for the ray tracing technique. *Applied Acoustics*, 15(4):263–274, 1982.

[39] D. van Maercke and J. Martin. The prediction of echograms and impulse responses within the Epidaure software. *Applied Acoustics*, 38(2-4, Special Issue on Computer Modelling and Auralisation of Sound Fields In Rooms):93–114, 1993.

[40] H. Lehnert and J. Blauert. Principles of binaural room simulation. *Applied Acoustics*, 36(3-4):259–291, 1992.

[41] G. M. Naylor. ODEON – another hybrid room acoustical model. *Applied Acoustics*, 38(2-4, Special Issue on Computer Modelling and Auralisation of Sound Fields in Rooms):131–143, 1993.

[42] T. Lahti and H. Möller. The Sigyn hall, Turku – a concert hall of glass. In *Proc. Nordic Acoustical Meeting (NAM'96)*, pages 43–48, Helsinki, Finland, 12-14 June 1996.

[43] B. M. Gibbs and D. K. Jones. A simple image method for calculating the distribution of sound pressure levels within an enclosure. *Acustica*, 26(1):24–32, 1972.

[44] H. Lee and B.-H. Lee. An efficient algorithm for the image model technique. *Applied Acoustics*, 24(2):87–115, 1988.

[45] R. Heinz. Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail. *Applied Acoustics*, 38(2-4, Special Issue on Computer Modelling and Auralisation of Sound Fields in Rooms):145–159, 1993.

[46] B.-I. Dalenbäck. *A New Model for Room Acoustic Prediction and Auralization*. PhD thesis, Chalmers Univ. of Tech., Gothenburg, Sweden, 1995.

[47] D. van Maercke. Simulation of sound fields in time and frequency domain using a geometrical model. In *Proc. 12th Int. Congr. Acoust. (ICA'86)*, volume 2, paper E11-7, Toronto, Ont., Canada, July 1986.

[48] M. Vorländer. Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm. *J. Acoust. Soc. Am.*, 86(1):172–178, 1989.

[49] E. Granier, M. Kleiner, B.-I. Dalenbäck, and P. Svensson. Experimental auralization of car audio installations. *J. Audio Eng. Soc.*, 44(10):835–849, Oct. 1996.

[50] J.-M. Jot. Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces. *Multimedia Systems, Special Issue on Audio and Multimedia*, 7(1):55–69, 1999.

[51] L. Savioja. *Computational Modeling of Room Acoustics*. Licentiate Thesis, Helsinki University of Technology, 1995, (in Finnish).

[52] M. Tamminen. The EXCELL method for efficient geometric access to data. *Acta Polytechnica Scandinavica, Mathematics and Computer Science Series*, (34), 1981.

[53] H. Samet. *The Design and Analysis of Spatial Data Structures*. Addison-Wesley, 1990.

[54] H. Strauss. Implementing doppler shifts for virtual auditory environments. In *the 104th Audio Engineering Society (AES) Convention, preprint no. 4687*, Amsterdam, the Netherlands, 16-19 May 1998.

[55] J. Huopaniemi, V. Pulkki, and T. Lokki. Psychophysics and technology of virtual acoustic displays, Nov. 1998. Tutorial given at *the Int. Conf. on Auditory Display (ICAD'98)*.

[56] H. Bass and H.-J. Bauer. Atmospheric absorption of sound: Analytical expressions. *J. Acoust. Soc. Am.*, 52(3):821–825, 1972.

[57] Standard ISO 9613-1. *Acoustics — Attenuation of Sound During Propagation Outdoors — Part 1: Calculation of the Absorption of Sound by the Atmosphere.* 1993.

[58] G. Naylor and J. Rindel. *Odeon Room Acoustics Program, Version 2.5, User Manual.* Technical Univ. of Denmark, Acoustics Laboratory, Publication 49, Copenhagen, Denmark, 1994.

[59] M. R. Schroeder. Natural-sounding artificial reverberation. *J. Audio Eng. Soc.*, 10(3):219–223, 1962.

[60] J. A. Moorer. About this reverberation business. *Computer Music J.*, 3(2):13–28, 1979.

[61] J.-M. Jot. *Etude et réalisation d'un spatialisateur de sons par modèles physique et perceptifs.* PhD thesis, l'Ecole Nationale Supérieure des Télécommunications, Télécom Paris, Sept. 1992.

[62] J. Stautner and M. Puckette. Designing multi-channel reverberators. *Computer Music J.*, 6(1):569–579, 1982.

[63] W. Gardner. Reverberation algorithms. In M. Kahrs and K. Brandenburg, editors, *Applications of Digital Signal Processing to Audio and Acoustics*, pages 85–131. Kluwer Academic, Boston, MA, 1997.

[64] R. Väänänen, V. Välimäki, J. Huopaniemi, and M. Karjalainen. Efficient and parametric reverberator for room acoustics modeling. In *Proc. Int. Computer Music Conf. (ICMC'97)*, pages 200–203, Thessaloniki, Greece, Sept. 1997.

[65] M. Schroeder and B. Atal. Computer simulation of sound transmission in rooms. *IEEE Conv. Record*, 11(7):150–155, 1963.

[66] M. Gerzon. Periphony: With-height sound reproduction. *J. Audio Eng. Soc.*, 21(1/2):2–10, 1973.

[67] D. Malham and A. Myatt. 3-D sound spatialization using ambisonic techniques. *Computer Music J.*, 19(4):58–70, 1995.

[68] V. Pulkki. Virtual sound source positioning using vector base amplitude panning. *J. Audio Eng. Soc.*, 45(6):456–466, Jun. 1997.

[69] J. Huopaniemi and J. O. Smith. Spectral and time-domain preprocessing and the choice of modeling error criteria for binaural digital filters. In *Proc. AES 16th Int. Conf. on Spatial Sound Reproduction*, pages 301–312, Rovaniemi, Finland, 10-12 April 1999.

[70] J. Huopaniemi, N. Zacharov, and M. Karjalainen. Objective and subjective evaluation of head-related transfer function filter design. *J. Audio Eng. Soc.*, 47(4):218–239, April 1999.

[71] V. Pulkki and T. Lokki. Creating auditory displays with multiple loudspeakers using VBAP: A case study with DIVA project. In *Proc. of the Int. Conf. on Auditory Display (ICAD'98)*, Glasgow, UK, 1-4 Nov. 1998.

[72] J. Sandvad. Dynamic aspects of auditory virtual environments. In *the 100th Audio Engineering Society (AES) Convention, preprint no. 4226*, Copenhagen, Denmark, 11-14 May 1996.

[73] J.-M. Jot, V. Larcher, and O. Warusfel. Digital signal processing issues in the context of binaural and transaural stereophony. In *the 98th Audio Engineering Society (AES) Convention, preprint no. 3980*, Paris, France, 1995.

[74] S. Foster, E. Wenzel, and R. Taylor. Real-time synthesis of complex acoustic environments. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'91)*, Mohonk, New Paltz, NY, 1991.

[75] T. I. Laakso, V. Välimäki, M. Karjalainen, and U. K. Laine. Splitting the unit delay – tools for fractional delay filter design. *IEEE Signal Processing Magazine*, 13(1):30–60, Jan. 1996.

[76] T. Takala and J. Hahn. Sound rendering. *Computer Graphics*, SIGGRAPH'92(26):211–220, 1992.

[77] E. Wenzel. Analysis of the role of update rate and system latency in interactive virtual acoustic environments. In *the 103rd Audio Engineering Society (AES) Convention, preprint no. 4633*, New York, NY, 26-29 Sept. 1997.

[78] E. Wenzel. Effect of increasing system latency on localization of virtual sounds. In *Proc. AES 16th Int. Conf. on Spatial Sound Reproduction*, pages 42–50, Rovaniemi, Finland, 10-12 April 1999.

[79] Deutsche Telekom. Investigations on tolerable asynchronism between audio and video, April 1995. Document 11A/DTAG1, Question ITU-R 35-2/11.

[80] M. P. Hollier and A. N. Rimell. An experimental investigation into multi-modal synchronisation sensitivity for perceptual model development. In *the 105th Audio Engineering Society (AES) Convention, preprint no. 4790*, San Francisco, CA, 26-29 Sept. 1998.

[81] S. Van Duyne and J. O. Smith. Physical modeling with the 2-D digital waveguide mesh. In *Proc. Int. Computer Music Conf. (ICMC'93)*, pages 40–47, Tokyo, Japan, Sept. 1993.

[82] S. Van Duyne and J. O. Smith. The 2-D digital waveguide mesh. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'93)*, Mohonk, New Paltz, NY, Oct. 1993.

[83] M. Karjalainen, V. Ikonen, A. Järvinen, P. Maijala, L. Savioja, A. Suutala, J. Backman, and S. Pohjolainen. Comparison of numerical simulation models and measured low-frequency behavior of a loudspeaker. In *the 104th Audio Engineering Society (AES) Convention, preprint no. 4722*, Amsterdam, the Netherlands, 16-19 May 1998.

[84] A. Järvinen, L. Savioja, H. Möller, V. Ikonen, and A. Ruusuvuori. Design of a reference listening room—a case study. In *the 103th Audio Engineering Society (AES) Convention, preprint no. 4559*, New York, NY, 26-29 Sept. 1997.

[85] L. Savioja, A. Järvinen, K. Melkas, and K. Saarinen. Determination of the low frequency behavior of an IEC listening room. In *Proc. Nordic Acoustical Meeting (NAM'96)*, pages 55–58, Helsinki, Finland, 12-14 June 1996.

[86] A. Järvinen, L. Savioja, and K. Melkas. Numerical simulations of the modified Schroeder diffuser structure (abstract). *J. Acoust. Soc. Am.*, 103(5):3065, May 1998. Paper presented in the ICA/ASA Joint Meeting, Seattle, WA, Jun. 1998.

[87] F. Fontana and D. Rocchesso. A new formulation of the 2D-waveguide mesh for percussion instruments. In *Proc. XI Colloquium on Musical Informatics*, pages 27–30, Bologna, Italy, 8-11 Nov. 1995.

[88] S. Van Duyne and J. O. Smith. The 3D tetrahedral digital waveguide mesh with musical applications. In *Proc. Int. Computer Music Conf. (ICMC'96)*, pages 9–16, Hong Kong, 19-24 Aug. 1996.

[89] F. Fontana and D. Rocchesso. Physical modeling of membranes for percussion instruments. *Acustica united with Acta Acustica*, 84(3):529–542, May/June 1998.

[90] J. Laird, P. Masri, and C. N. Canagarajah. Efficient and accurate synthesis of circular membranes using digital waveguides. In *Proc. IEE Colloquim on Audio and Music Technology: The Creative Challenge of DSP (Ref. No. 1998/470)*, pages 12/1–12/6, 18 Nov. 1998.

[91] D. Murphy, D. Howard, and A. Tyrrell. Multi-channel reverberation for computer music applications. In *Proc. 1998 IEEE Workshop on Signal Processing Systems*, pages 210–219, Boston, MA, 8-10 Oct. 1998.

[92] D. Murphy and D. Howard. Modelling and directionally encoding the acoustics of a room. *Electronics Letters*, 34(9):864–865, 1998.

[93] D. Murphy. The WaveVerb multi-channel room acoustics modelling system. In *Proc. Int. Computer Music Conf. (ICMC'99)*, pages 472–475, Beijing, China, 22-27 Oct. 1999.

[94] J. O. Smith. *Music Applications of Digital Waveguides.* Stanford, CA, Stanford University, Dept. of Music, Center for Computer Research in Music and Acoustics (CCRMA) Tech. Report no. STAN-M-39, 27 May 1987.

[95] J. O. Smith. Physical modeling synthesis update. *Computer Music J.*, 20(2):44–56, 1996 Summer.

[96] J. O. Smith. A new approach to digital reverberation using closed waveguide networks. In *Proc. Int. Computer Music Conf. (ICMC'85)*, pages 47–53, Vancouver, BC, Canada, Aug. 1985.

[97] V. Välimäki. *Discrete-Time Modeling of Acoustic Tubes Using Fractional Delay Filters.* Doctoral thesis, Helsinki University of Technology, Lab. of Acoustics and Audio Signal Processing, Report 37, 1995. Available at http://www.acoustics.hut.fi/~vpv/publications/vesa_phd.html.

[98] L. Savioja, M. Karjalainen, and T. Takala. DSP formulation of a finite difference method for room acoustics simulation. In *Proc. IEEE Nordic Signal Processing Symp. (NORSIG'96)*, pages 455–458, Espoo, Finland, 24-27 Sept. 1996.

[99] S. Van Duyne and J. O. Smith. The tetrahedral digital waveguide mesh. In *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'95)*, Mohonk, New Paltz, NY, Oct. 1995.

[100] L. Savioja and V. Välimäki. The bilinearly deinterpolated waveguide mesh. In *Proc. IEEE Nordic Signal Processing Symp. (NORSIG'96)*, pages 443–446, Espoo, Finland, 24-27 Sept. 1996.

[101] V. Välimäki, M. Karjalainen, and T. I. Laakso. Fractional delay digital filters. In *Proc. IEEE Int. Symp. on Circuits and Systems (ISCAS'93)*, volume 1, pages 355–358, Chicago, IL, 3-6 May 1993.

[102] A. Oppenheim, D. Johnson, and K. Steiglitz. Computation of spectra with unequal resolution using the Fast Fourier Transform. *Proc. IEEE*, 59(2):299–301, Feb. 1971.

[103] U. K. Laine, M. Karjalainen, and T. Altosaar. Warped linear prediction (WLP) in speech and audio processing. In *Proc. Int. Conf. Acoust., Speech, Signal Processing (ICASSP'94)*, volume 3, pages 349–352, Adelaide, Australia, 19-22 April 1994.

[104] A. V. Oppenheim and D. H. Johnson. Discrete representation of signals. *Proc. IEEE*, 60(6):681–691, June 1972.

[105] J. Laird, P. Masri, and N. Canagarajah. Modelling diffusion at the boundary of a digital waveguide mesh. In *Proc. Int. Computer Music Conf. (ICMC'99)*, pages 492–495, Beijing, China, 22-27 Oct. 1999.

Publication [P4]

Page 4, section 5, first paragraph: "... one half of ..." should be "... one third of ...". As a consequence, following corrections should be made:

- First equation: "$c = \frac{\sqrt{3} \times d}{2 \times T}$" should be "$c = \frac{\sqrt{3} \times d}{3 \times T}$"
- Second equation: "$f = \frac{1}{2 \times T} = \frac{2 \times 343}{2 \times \sqrt{3} \times d} Hz$" should be "$f = \frac{1}{2 \times T} = \frac{3 \times 343}{2 \times \sqrt{3} \times d} Hz$"
- "... limit frequency is approximately 2 kHz. About 8000 time ..." should be "... limit frequency is approximately 3 kHz. About 12000 time ..."

Publication [P7]

Page 3, section 3, first paragraph: "... $d_{3d} = \frac{1}{2}$ ..." should be "... $d_{3d} = \frac{1}{\sqrt{3}}$ ..." As a consequence, following corrections should be made:

- "$h_a = 1 + 4(1 - d_{3d})^2 d_{3d} + 4 d_{2d}(1 - d_{2d}) = 2\sqrt{2} - \frac{1}{2}$" should be "$h_a = 1 + 4(1 - d_{3d})^2 d_{3d} + 4 d_{2d}(1 - d_{2d}) = -\frac{11}{3} + 2\sqrt{2} + \frac{16}{3\sqrt{3}} \approx 2.241$"
- "$h_{2d} = 2(1 - d_{3d}) d_{3d}^2 + d_{2d}^2 = \frac{3}{4}$" should be "$h_{2d} = 2(1 - d_{3d}) d_{3d}^2 + d_{2d}^2 = \frac{7}{6} - \frac{2}{3\sqrt{3}} \approx 0.7818$"
- "$h_{3d} = d_{3d}^3 = \frac{1}{8}$" should be "$h_{3d} = d_{3d}^3 = \frac{1}{3\sqrt{3}} \approx 0.1925$"
- "$h_c = 12(1 - d_{2d})^2 + 8(1 - d_{3d})^3 = 19 - 12\sqrt{2}$" should be "$h_c = 12(1 - d_{2d})^2 + 8(1 - d_{3d})^3 = 34 - 12\sqrt{2} - \frac{80}{3\sqrt{3}} \approx 1.6334$"
- Figures 2b and 3b are computed using erroneous interpolation coefficients. Fig. 4.9 in this thesis represents a new version of Fig. 2. A new version of Fig. 3b is so similar to the presented one that it is not inserted here.