EXTENDED SCATTERING DELAY NETWORKS INCORPORATING ARBITRARILY SHAPED ROOMS AND EDGE DIFFRACTION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

SERKAN PEKÇETIN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
GAME TECHNOLOGIES

JULY 2016

Approval of the thesis:

EXTENDED SCATTERING DELAY NETWORKS INCORPORATING ARBITRARILY SHAPED ROOMS AND EDGE DIFFRACTION

submitted by **SERKAN PEKÇETIN** in partial fulfillment of the requirements for the degree of **Master of Science in Modelling and Simulation Department, Middle East Technical University** by,

Prof. Dr. Nazife Baykal
Director, **Graduate School of Informatics Institute, METU** _____

Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu
Head of Department, **Modelling and Simulation, METU** _____

Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu
Supervisor, **Modelling and Simulation, METU** _____

Examining Committee Members:

Prof. Dr. Yasemin Yardımcı Çetin
Information Systems, METU _____

Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu
Modelling and Simulation, METU _____

Assoc. Prof. Dr. Ahmet Oğuz Akyüz
Computer Engineering Department, METU _____

Prof. Dr. Haşmet Gürçay
Computer Engineering Department, Hacettepe University _____

Assoc. Prof. Dr. Banu Günel Kılıç
Information Systems, METU _____

Date: **18 July 2016**

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: SERKAN PEKÇETIN

Signature :

ABSTRACT

EXTENDED SCATTERING DELAY NETWORKS INCORPORATING ARBITRARILY SHAPED ROOMS AND EDGE DIFFRACTION

Pekçetin, Serkan

M.S., Department of Modelling and Simulation

Supervisor : Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu

July 2016, 53 pages

In order to achieve immersion in computer games, room acoustics models are often employed. Due to highly computational cost involved in full-scale room simulation, artificial reverberators are commonly used. Scattering Delay Network is such a reverberator that can run at interactive rates. The reverberator relies on geometrical acoustics to render first-order reflections correctly while it offers a convincing approximation in higher order reflections. The original model was exclusively proposed for shoebox type of room settings. The aim of this study is to extend the reverberator by allowing it to work in arbitrarily shaped rooms. Besides, enhanced perception of the virtual environment is proposed by approximating first-order edge diffraction model. The extended algorithm uses visibility checks by ray casting in order to determine which surfaces have an energy transaction. Edge diffraction component of the room impulse response is then approximated by using a best-fit approach. Decay parameter for the inverse exponential function is obtained by using ternary search. Scope of the original algorithm is extended at the expense of minimal computational complexity.

Keywords: Room Acoustics, Artificial Reverberation, Edge Diffraction

ÖZ

SAÇILIMLI GECİKMELİ AĞLARA İSTEĞE GÖRE ODALAR VE KÖŞE KIRINIM MODELİ EKLENEREK GENİŞLETİLMESİ

Pekçetin, Serkan
Yüksek Lisans, Oyun Teknolojileri Bölümü
Tez Yöneticisi : Doç. Dr. Hüseyin Hacıhabiboğlu

Temmuz 2016 , 53 sayfa

Bilgisayar oyunlarında tutulurluğu sağlamak için sıklıkla oda akustiği modelleri kullanılır. Yüksek hesaplama maliyetlerinin tam ölçekli oda benzetimlerine izin vermemesinden dolayı, yapay yankışım modellerine başvurulur. Saçılımlı Gecikmeli Ağ, etkileşimli hızlarda çalışabilen bir yapay yankışım modelidir. Bu model, birinci dereceden yansımaları geometrik akustik kurallarına bağlı olarak hesaplarken, yüksek derece yansımalar için ise tatminkar bir benzerlik sunar. Özgün model sadece ayakkabı kutusu geometrisine sahip odalar için geçerlidir. Bu çalışmanın amacı, yankışım modelinin genişletilerek isteğe bağlı oda geometrilerinde çalışabilmesini sağlamaktır. Bunun yanında, sanal ortam algısının birinci derece kırınımın da yaklaşık olarak modellenerek artırılması hedeflenmiştir. Genişletilmiş algoritma ışın saçma yöntemi ile görünürlük denetlemeleri yaparak, güç transferi olan yüzeyleri belirler. Oda akustik cevabının kırınım kısmı ise en uygun fonksiyona yaklaştırılır. Ters üstel fonksiyonunun eksilme parametresi üçlü bölerek arama yöntemi ile bulunur. Sonuç olarak özgün algoritmanın kapsamı, asgari hesap yükü harcanarak genişletilmiştir.

Anahtar Kelimeler: Oda Akustiği, Yapay Yankılama, Köşe Kırınımı

To those who missed the starting gun...

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor, Assoc. Prof. Dr. Hüseyin Hacıhabiboğlu for his invaluable guidance and support. Expertise and patience from his side helped me a lot in learning during my time in his "Analysis-based synthesis for spatial audio reproduction" project, which was funded by TÜBİTAK. Being part of the Spatial Audio Research Group under his wings prepared me for the further challenges in academia.

I would also like to thank to my friends Deniz Uğurca and Tarık Kaya. They were the people who motivated me into programming and I have spent some great time with them. The crunch mode and competitive mindset when we were together was a unique experience. Countless hours at labs, libraries and offices right next to them molded persistence into my character. Selim Nar, who is another source of inspiration, provided me with helpful tips during implementation. I also learned a lot from Dr. Erdal Yılmaz, whom we worked together in both academic and professional settings. Having a chance to work under him at Argedor was a rewarding experience. Presence of Murat Karataş and Oğuz Çetin were the proof of best friends know no distance. We backed up each other at our hardest times no matter what, which is not so common nowadays. Misery loves company, so I would like to thank the inhabitants of Z-12; especially Mehmetcan Fal and Tuğçe Nur Bozkurt.

Prof. Dr. Veysi İşler was more than helpful and motivating during my studies at METU Game Technologies. Besides, he looked after our efforts in establishing our audio lab in MODSIMMER during his period as the director. I remind my very first game, which I developed upon his encouragement during a course he taught.

Prof. Dr. Mehmet Çalışkan and his colleagues at Mezzo Stüdyo generously provided me with tools and resources which were essential to conduct my research. I regret spending little time beyond my thesis work with them.

Most credit goes to my family. Nothing could have been achieved without their unconditional love and endless support. It gives me that warm and fuzzy feeling when they sound delighted at even my slightest accomplishments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

GPU	Graphical Processing Unit
DWN	Digital Waveguide Network
DWM	Digital Waveguide Mesh
RIR	Room Impulse Response
FDN	Feedback Delay Network
DSP	Digital Signal Processing
RT60	Reverberation Time
EDT	Early Decay Time
SDN	Scattering Delay Network
ISM	Image Source Method
BTM	Biot-Tolstoy-Medwin Diffraction Approximation
UTD	Uniform Geometrical Theory of Diffraction
EDC	Energy Decay Curve
ITU	International Telecommunication Union
PEAQ	Perceptual Evaluation of Audio Quality
ODG	Objective Difference Grade
DI	Distortion Index
IIR	Infinite Impulse Response

CHAPTER 1

INTRODUCTION

Most of the computer games provide audio renderers to achieve maximum immersion and realism. One of the key features in these audio renderers is the ability of reverberation. Reverberation schemes allow developers to convey significant characteristics of a given room. Accurate room acoustics simulations thus provide the player with valuable cues for perception of the 3D scene [1]. Although there are models for full-scale room acoustics simulation, they are not suitable for real-time applications. In order to achieve interactive frame rates, artificial reverberators are used. Artificial reverberators often sacrifice realism in exchange for lower computational complexity. Therefore, design of a high-performance yet accurate artificial reverberator still remains as a challenge.

Selection of an artificial reverberation technique depends on the trade-off between the intended accuracy and computational load. One can simply use primitive DSP effects which are very fast and easy to implement. However, the modeling capability of the technique is minimal. One can measure statistical properties such as RT60 or EDT and use it to convey room acoustics via these properties. This methodology lacks the unique properties of a given room as different room geometries may share the same parameters.

FDN reverberators use unitary, energy-preserving matrices in order to give the sense of room acoustics. Finding the matrix coefficients that simulates a given room is not a trivial task. One shall exhaustively plug in distinct values that would give the closest approximation to RIR. Synthesized reverberation suffers from the parameter selection problem as well. Another solution could be to convolve real RIR recordings, which is not feasible for virtually defined geometries.

Ray tracing methods yield accurate results for any geometry that can be defined. Yet, it is infeasible for real-time applications as complexity increases according to the order of reflections, which is essential for accurate simulations. DWN works by placing mesh junctions with bidirectional delay lines in order to calculate RIR. Memory requirements for such a system depends on the amount of nodes within the mesh which could be prohibitively high if one aims for accuracy. This characteristic of the methodology allows DWN to be suitable for parallel processing. Given that GPU resources are prioritized exclusively for graphics rendering, DWN is not a popular method in

computer games and interactive simulations.

SDN provides accurate first-order RIR with plausible higher-order approximation [2]. Computational complexity and memory footprint is minimal. Although SDN models provide a reasonable trade-off between accuracy and computational cost, the modeling capability is limited with shoe box-type room geometries. Beyond any geometry that is not a rectangular prism, accuracy of the SDN model decreases.

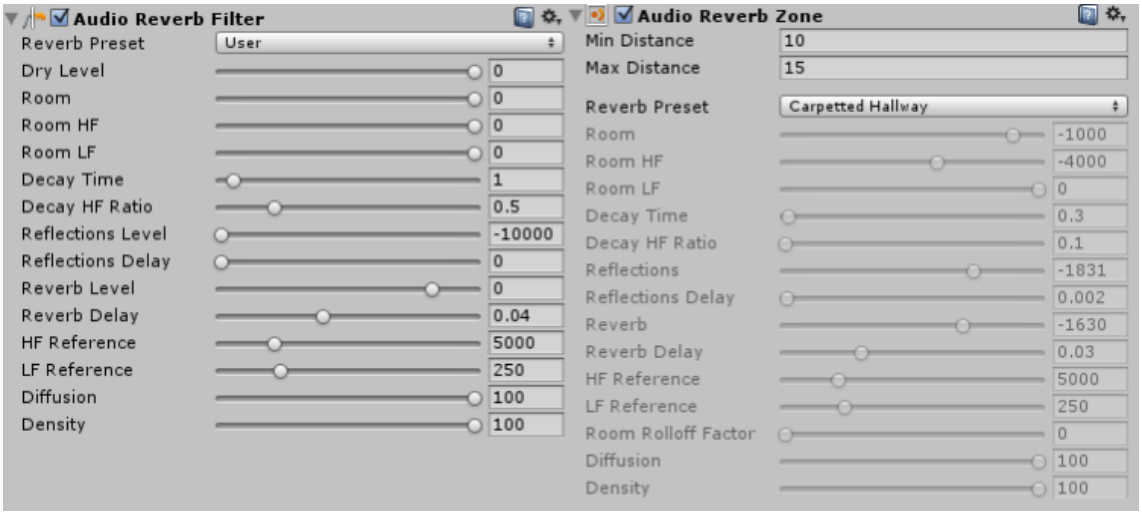

Figure 1.1: Unity3D: Reverb filter and reverb zone parameters.

Some of the reverberation schemes mentioned above are utilized in game development software. Freeverb and nReverb are two implementations of digital artificial reverberators that utilize combinations of various DSP filters [3] [4]. Similar approaches found their use in commercial software: Commonly used game engines such as Unity3D and Unreal Engine provide DSP-level reverberation schemes out of the box [5] [6]. Besides the reverb effect itself, implementations of so-called audio reverb zones allow sound designers to define the regions and their corresponding reverb parameters so that a primitive acoustics simulation can be achieved. Figure 1.1 shows the parameters supported by Unity3D. The user can also define reverb zones and place them around the scene. It is allowed to save preset parameters for reverb filter for future use.

Third-party audio middleware software solutions such as Wwise and FMOD provide options for reverberation. Wwise offers convolution and matrix reverb, which allows developers to choose between high performance or realistic acoustical simulation [7]. FMOD similarly features DSP-level reverb and and convolution reverb [8]. These middleware software have extensions and plug-ins that allow them to integrate with most of the game engines.

1.1 Contributions

Although SDN is very fast and it yields a convincingly accurate output, it can only model a small set of geometries because of its by-definition limitations. In this study, a method for incorporating arbitrarily shaped rooms to original SDN model is proposed. Furthermore, an approximation for modeling edge diffraction is also explained in detail. Performance characteristics of these methods are analyzed and discussed.

In an overview, contributions of this work to original SDN model is as follows:

- Visibility and validity checks for image sources are integrated to extend the model to work with arbitrarily shaped rooms.
- Visibility checks between source, node and receiver pairs are proposed.
- First-order edge diffraction modeling is added in order to handle rooms with finite wedges.
- An algorithm is proposed for faster simulation of first-order edge diffraction which allows the model to run at interactive rates.
- All of the items above are implemented as a robust, interactive application that runs an acoustical simulation of a given room to ensure that the model runs in real-time with the extensions.
- Performance benchmarks for CPU run times and memory utilization are presented.

CHAPTER 2

BACKGROUND

In this section, background information regarding modeling room acoustics is given. Physical and artificial reverberators are explained in detail. Models of edge diffraction are also discussed.

2.1 Modeling Room Acoustics

In order to simulate room acoustics, both physical and computational models can be employed. From these models, one can obtain the RIR.

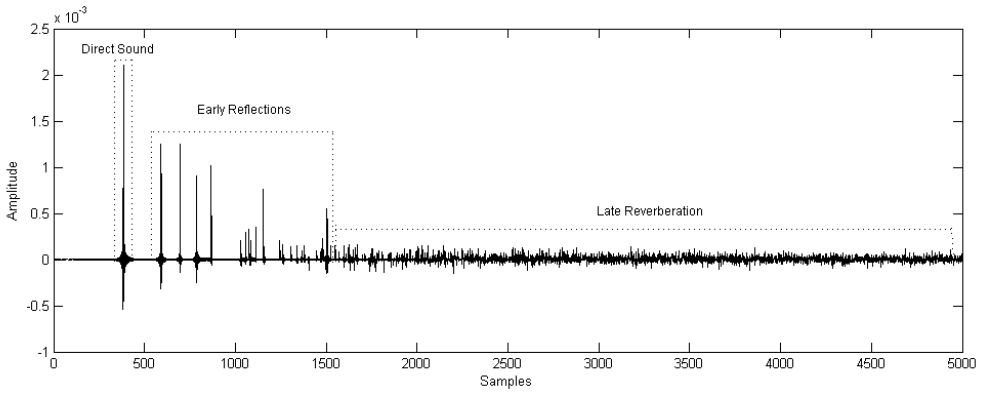

Figure 2.1: RIR of a sample room and its components.

As indicated in Figure 2.1, RIR have three basic components; direct sound, early reflections and late reverberation. Direct sound is directly related to positions and directivity patterns of source and receiver. Early reflections are affected by the room geometry, specular and diffuse reflections as well as the materials of the reflective surfaces. Late reverberation holds higher-order reflections alongside with diffuse reflections. Accuracy of a room acoustic simulation can be assessed on its ability to reproduce RIR for a given geometry. Using the RIR for a given room, one can reproduce the sound as it is heard within that room by convolving dry recordings with corresponding RIR. RIR also gives an idea about how to treat certain issues regarding acoustic properties of the room. Common methods used in room acoustics modeling are:

2.1.1 Scale Physical Models

This is a physical method in which a scale physical model of the room is built. It is intended that the scale model has similar geometry to the original hall. The propagation of sound waves are then predicted by utilizing ultrasound waves which have higher frequencies and shorter wavelengths. Therefore, ultrasound wave method compensates for scaling of the sound waves as well [9].

As several propagation properties are similar for different type of waves, some other alternatives can be employed. In a few extreme cases, it is possible to use light as a substitute to sound waves. Changes in the model should be made accordingly so that the material are selected according to their absorption of light. Modeling of diffraction is not possible due to relatively small optical wavelengths.

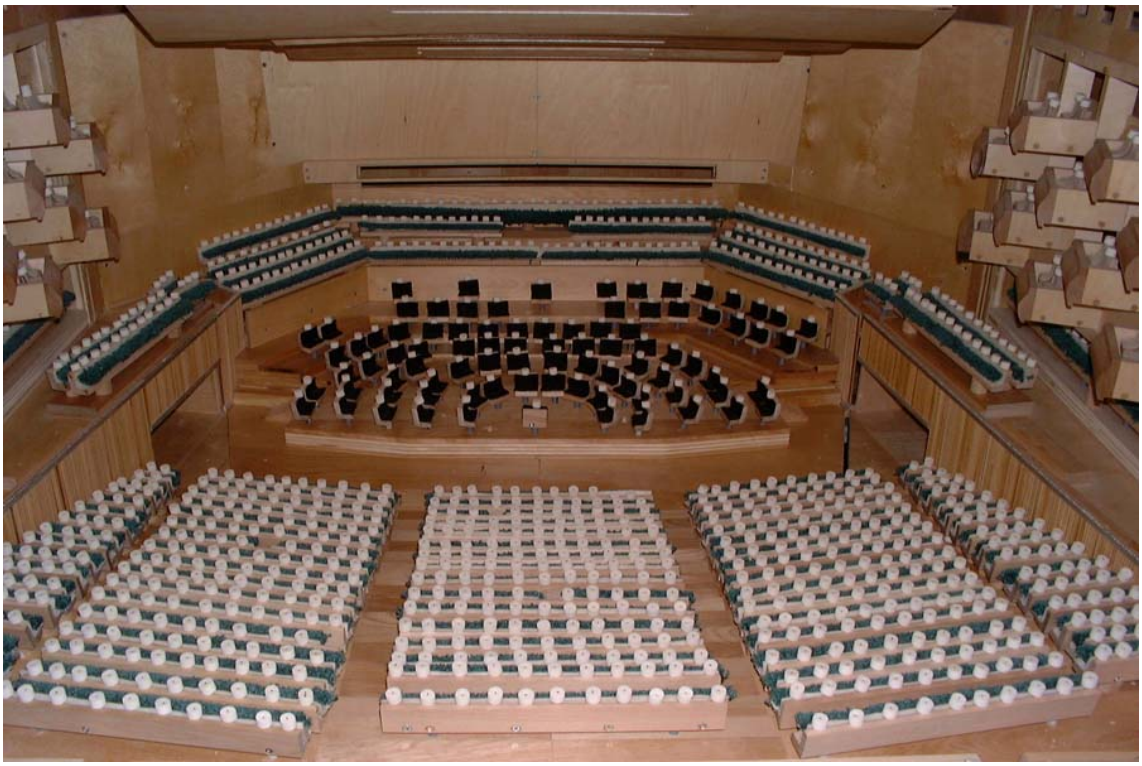

Figure 2.2: Scale model of Royal Festival Hall, London. Image courtesy of University of Salford.

Figure 2.2 shows a scale model of Royal Festival Hall in London [10]. For physical models, variations in room geometry and materials are possible with relatively little expense. However, it should be noted that the scale model itself has a fixed cost as well. Building scale physical models is mostly deprecated after the introduction of computational methods.

2.1.2 Image Source Method

ISM is built upon the principle that reflective surfaces of a room would act as mirrors [11]. Reflective surfaces therefore generate first-order image sources. They contribute to RIR just as the direct source in terms of delay and attenuation. To obtain complete RIR, higher-order image sources are located according to their lower-order parents. As a method of geometrical acoustics, ISM treats acoustical energy as path of rays.

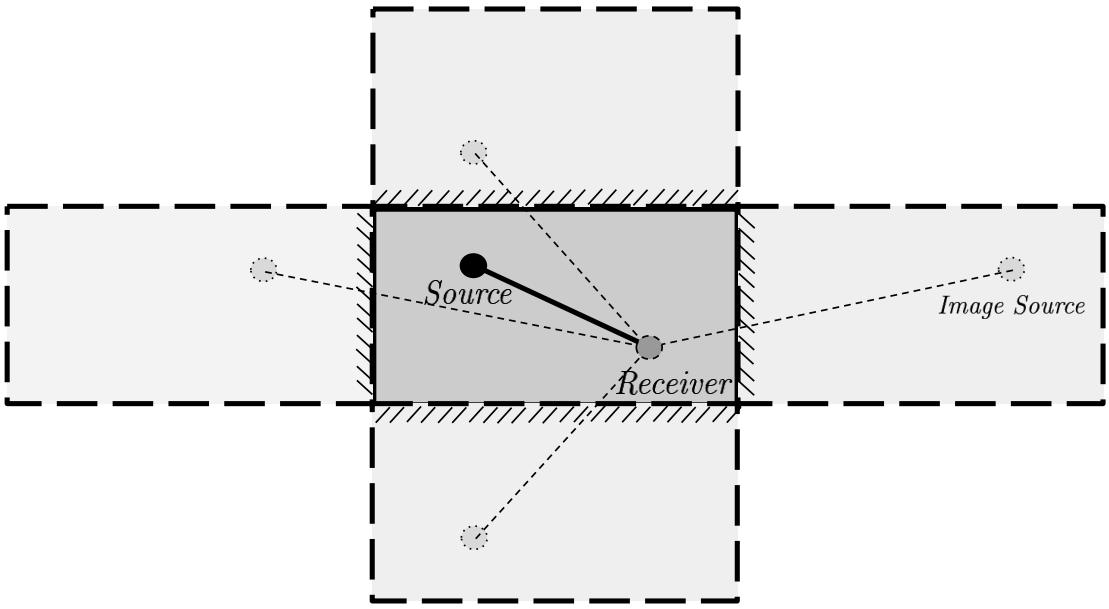

Figure 2.3: Image Source Method: Top view from a rectangular room.

Image sources in rectangular rooms can be calculated in a fairly straightforward way. Figure 2.3 shows the scenario for a rectangular room. A source-receiver pair and reflective surfaces create the first-order image sources. Higher order image sources could be found by reflecting them through other surfaces. All paths between receiver and image sources contribute to impulse response with respective attenuations.

However, for a scenario in which the room geometry has re-entrant angles, additional visibility tests are required. Visibility checks are to ensure that a given image-source is visible by the receiver. If the image-source is not valid, its contribution is neglected. For a given image-source, visibility is not enough. Another validity check ensures that the image-source is generated from the reflective side of a surface.

A suitable scenario for ISM could be a rectangular room whose room impulse is dominated by the first-order terms. Simplicity of the model allows very fast computation if the room is not an arbitrarily shaped polyhedra. Figure 2.4 shows the side view from a rectangular room and its corresponding first-order image sources. For a shoe box-type of room with 6 sides, there are same number of virtual image sources. Shaded, grey room in the middle

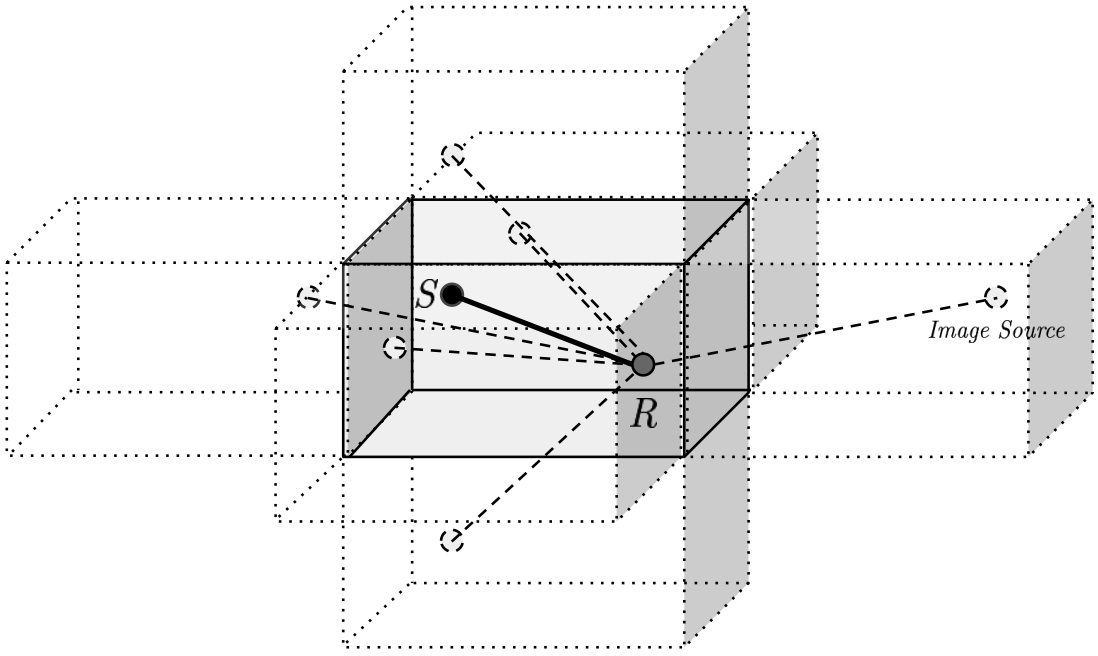

Figure 2.4: Image Source Method: Side view from a rectangular room.

is the original enclosure whereas rooms with dotted outlines are the virtual rooms which are formed by the reflective side of walls. In order to calculate second or higher order image sources, one can apply the same procedure iteratively by starting from the first-order image sources and so on. Despite its advantages, ISM can not model diffraction and diffuse reflections. Besides, the number of image-sources grow cubically as the reflection order increases. Therefore, modeling late reverberation in real-time remains as a challenge [12]. ISM is useful in approaches where calculations of the early reflections are required. SDN reverberators use a similar approach to ISM to calculate node positions within a given room.

2.1.3 Ray Tracing

Similar to its use in computer graphics domain, ray tracing allows computation of both specular and diffuse reflections [13]. The method traces all ray paths from source to microphone. In ray tracing method, predefined number of rays are cast from the source. In order to account for directivity, one can assign weights to the amplitudes of rays. Each reflection of a particular ray and the path it traverses until it hits the receiver contributes as an impulse. For intercepting the rays, receiver is defined in finite sizes as shown on Figure 2.5. For each point of incidence on the surface where specular reflection is located, diffuse reflections can be modeled by casting additional rays.

Although the method does not require any visibility or validity check, accuracy of the model depends heavily on the number of rays which is tightly coupled with the computational cost. Calculation of edge diffraction is not

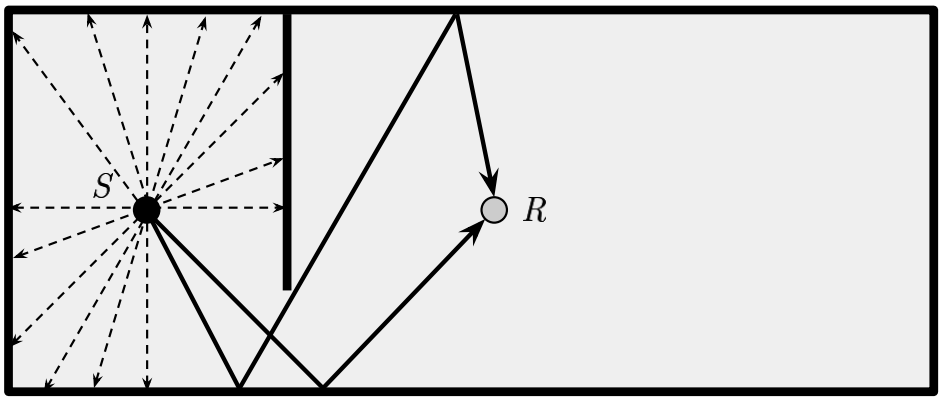

Figure 2.5: Ray tracing: Path from each ray between source and receiver is calculated.

possible with ray tracing and receiver is not a point source. These characteristics decreases the accuracy even further.

Figure 2.5 shows a top-view from a rectangular room with an obstacle between source S and receiver R. A fixed number of rays are emitted from the source and while some of them reach the receiver, the remaining ones are pruned before they reside on a path to receiver. Late reverberation and overall accuracy of the model depends on the number of rays emitted and number of reflections to be calculated before a particular ray is pruned.

2.1.4 Beam Tracing

In an analogous approach to ray tracing, beam tracing method casts beams of rays [14]. Rather than individual rays, beams of rays are emitted and and intersecting polygons are determined. Then a transmission beam is formed which mirrors over the intersecting polygon's surface to create a reflection beam. Underlying assumption here is that a sound field within a region is spatially coherent unless it encounters a reflector surface.

As shown in Figure 2.6, beam of rays are cast to the edges of reflective boundaries. The reflective surface a creates the virtual source S'. The virtual source S' then forms the region defined by R_a. Each and every point in the region R_a is visible and valid from the virtual source S'. Then, for the second-order reflections, one should only consider the polygons which lie in the region R_a. Therefore, for such a closure with arbitrary shape, pruning of specular reflections can be triggered earlier comparing to other methods. [15]

Due to its deductive approach, real-time applications that utilize beam tracing exist. The model has several advantages over ray tracing and image source methods. Since a beam is covering the entire 2D space, the accuracy does not depend on sampling rate. In ray tracing approach, since receiver has a finite volume, some rays do not intersect with the bounding sphere of the receiver even though the ray travels along the path to the receiver. Beam tracing

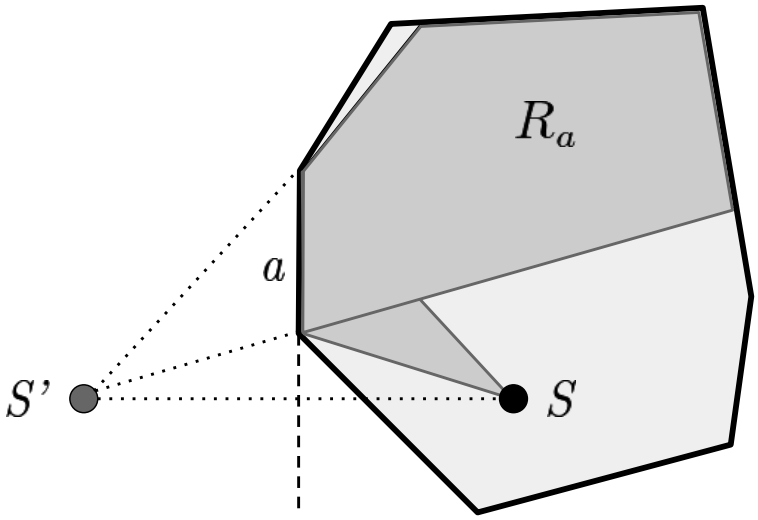

Figure 2.6: Beam tracing: Beam of rays are cast from the virtual source.

prunes the invalid virtual sources earlier comparing to image source method, which makes beam tracing preferable over image source method when the room is not a shoe box-shaped one. A disadvantage of the model is mainly the difficulty of tracing the beams through curved surfaces and indents. Besides, similar to ray tracing and image source methods, edge diffraction is not modeled in beam tracing models.

2.1.5 Digital Waveguide Mesh

DWM relies on wave equation and bidirectional delay lines in order to simulate room acoustics [16]. Each node on a DWM holds a scattering matrix. As waves are incoming to the junctions, they are multiplied by the corresponding scattering matrix thus generating outgoing wave variables.

DWM carries out its operation in two steps; the scattering pass and propagation step. These steps are conducted for each and every node in the network. In scattering pass, wave variables move on bidirectional delay lines and they form incoming wave variables as shown in Figure 2.7. In Figure 2.7, for the DWM junction i, the pressure which is incoming from j is indicated as p_{ji}^{-}. Outgoing pressure from node i to node j is p_{ij}^{+}. Arrival time depends on the delay between nodes, z^{-1}.

During propagation step, outgoing wave variables are added to bidirectional delay lines. Pressure in any node is then given by the sum of incoming and outgoing wave variables scaled by the factor of 2/N, where N is the number of nodes in the network.

In this methodology, as sampling rate increases, the resolution would also increase. This yields high number of nodes which introduces a steep memory usage which may not be feasible. Due to low computational complexity during the scattering passes and propagation steps, this methodology is suitable

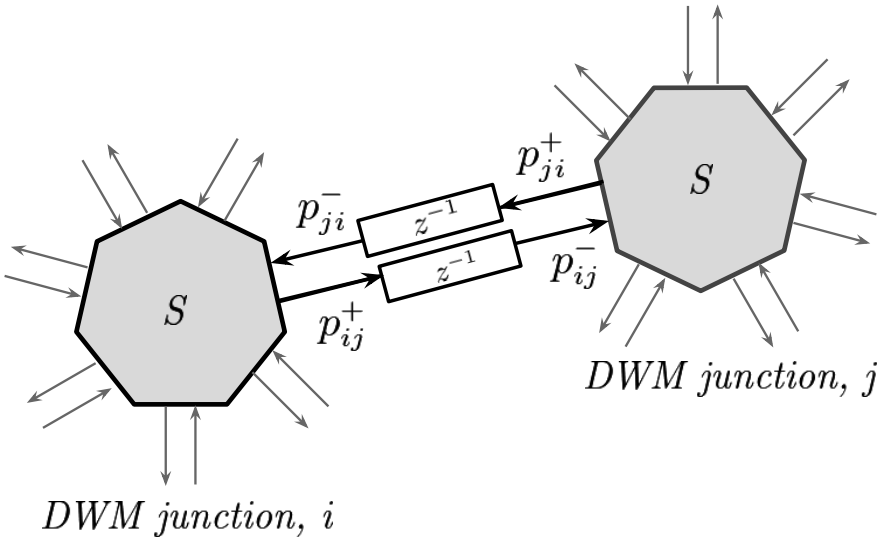

Figure 2.7: Two DWM junctions, with incoming and outgoing waves.

for parallel processing. However, given that GPU resources are prioritized to achieve better graphics quality, auditory rendering in real-time systems might not be possible. One of the building blocks of the original SDN model relies on the DWM solution.

2.2 Artificial Reverberators

Artificial reverberators can be either physical or digital. Even though physical models are mostly dominated by their digital counterparts, they can still be found in usage for certain settings. Digital reverberators allow users to manipulate fundamental parameters regarding the room with less effort comparing to physical ones [17].

2.2.1 Physical Artificial Reverberators

Reverberators of these type solely rely on physical models and mechanisms. They utilize electronic devices such as pickups, circuits, as well as mechanic apparatus.

2.2.1.1 Reverberation Chambers

These chambers are built by placing reflectors around the room, as shown in Figure 2.8 [18]. The reflectors are convex and although reverberation chambers are considered as artificial reverberators, output they produce have very little or no artifacts. One can play dry recordings in a reverberation chamber to obtain reverberated components. Capabilities of manipulating the chamber to simulate an arbitrary geometry is not possible as it may require a com-

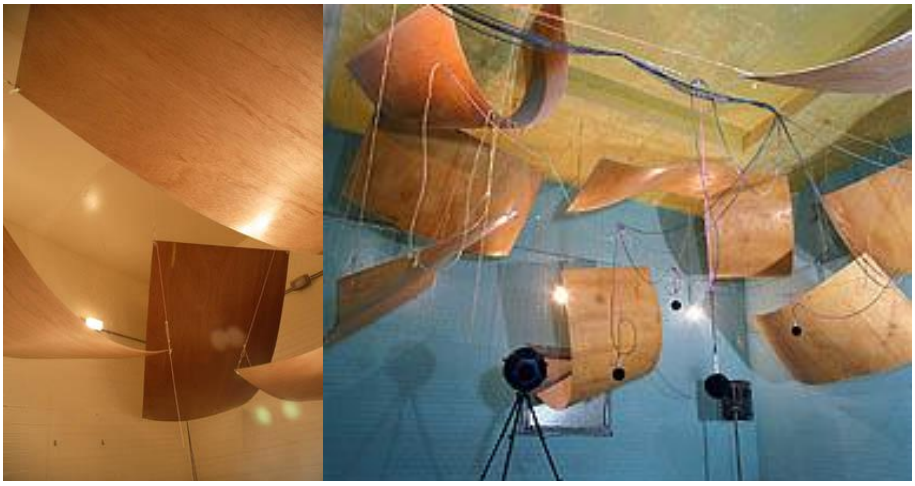

Figure 2.8: Reflective plates and a reverbertaion chamber. Image courtesy of University of Salford.

plete reconstruction of the chamber. Therefore, it could find a use in a scenario that the user only wants to add an arbitrary reverberation on a dry recording.

2.2.1.2 Spring Reverberators

Spring reverberators are the mechanisms that have two electromagnets and numerous springs on them. A coil is connected to the input signal, which then creates a magnetic field according to it. The vibration caused by the magnets moves the springs that are connected to them. Mechanical wave which is present on springs triggers the movement on the output coil. The current generated by the mechanism is the reverberation output [19]. Figure 2.9 demonstrates the parts that constitute a spring reverberator [20]. While outer steel, inner aluminium and support springs are utilized to hold parts together as a unit as they vibrate, transducers and transmission springs are the parts where reverberation is generated.

2.2.1.3 Plate Reverberators

In plate reverberators, a transducer excites a metal plate. A damping pad can be used to control the reverberation time. According to number of pickup transducers, one can obtain an output that has either one or two channels. Output of the system has similar reverberation characteristics when it is compared with reverberation of a rectangular room. Figure 2.10 is an example plate reverberator mechanism. The reverberator holds on to a frame. An audio signal then feeds the driver, which causes vibration on the plate. Unlike the spring reverberator in which the vibrations are one-dimensional, in plate reverberator, since the vibrating object is a plate, the vibrations are now in two dimensions. The output can then be gathered via pickups. If the reverberator have one pickup would create a mono output is produced. For stereo output, two pickups should be employed. While plate reverberators

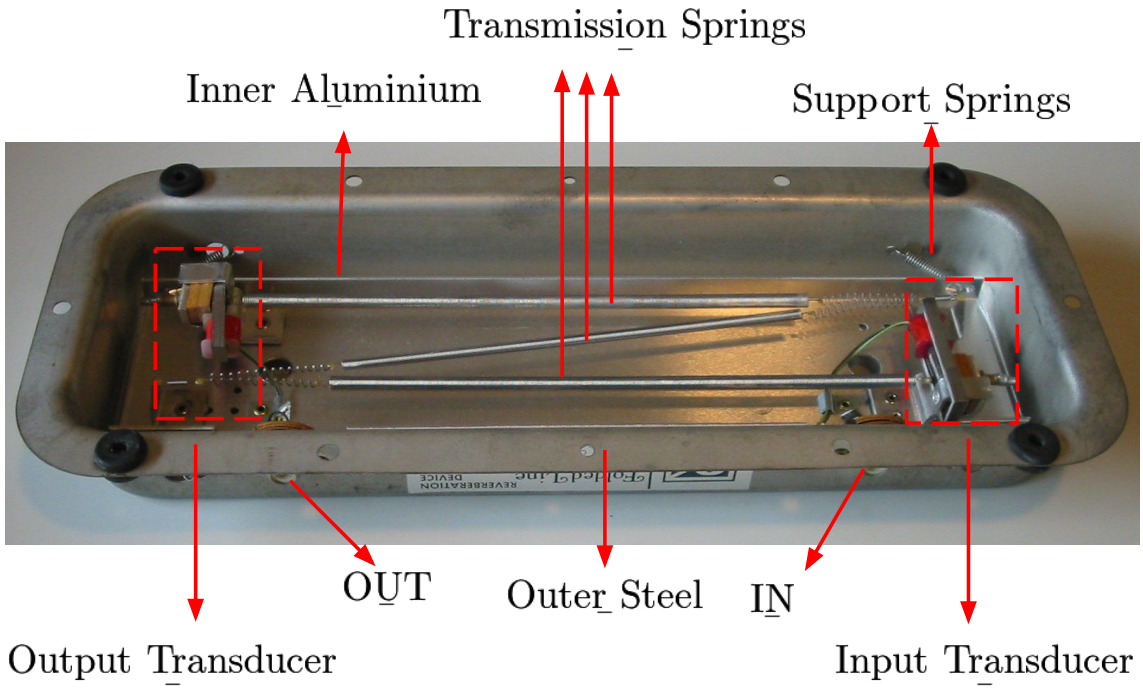

Figure 2.9: Spring Reverberator. Image modified from the original to highlight the components.

offer two-dimensional vibration, they are physically heavy and cumbersome comparing to spring reverberators.

2.2.2 Digital Artificial Reverberators

In order to assess digital artificial reverberators, some criteria should be satisfied. One expects the output to sound natural. Audible artifacts should not be present. Operating at interactive rates requires low computational cost. The model should allow changing of decay characteristics while RIR should be as accurate as possible.

2.2.3 Schroeder's Reverberator

Schroeder's artificial reverberator employs comb filters and all-pass filters [21]. Modal density is enhanced by parallel comb filters in this reverberation scheme. Series of all-pass filters are used to increase the reflection density.

Schroeder suggests that delays should be chosen as co-prime values in order to minimize the interference between frequency modes. As a rule of thumb, ratio of the largest-to-smallest delay should be 1.5 or any other value that is close to it.

There are several flaws of Schroeder's reverberator. The output does not

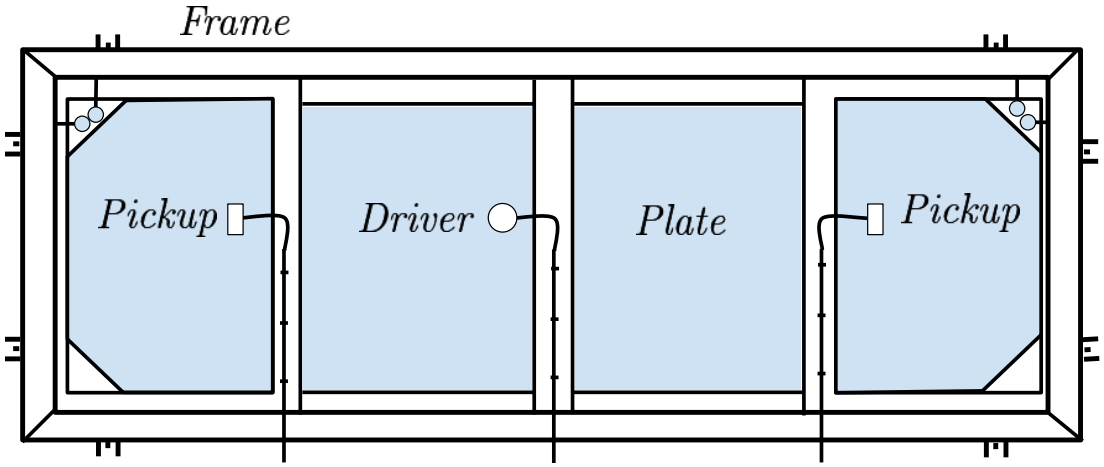

Figure 2.10: Plate reverberator mechanism.

sound natural, giving metallic sounding artifacts. Impulsive sounds are problematic and differences in frequency are not accounted.

However, implementations of Schroeder reverberators introduce little computational cost. A popular to-day implementation of Schroeder reverberator is Freeverb. Freeverb reverberator uses four all-pass filters in series which are combined with eight comb filters. This scheme can be repeated as much as the number of desired channels, so Freeverb is a multi-channel artificial reverberator. Careful selection of filter parameters prevents artifacts to an extent.

2.2.4 Moorer's Reverberator

Moorer's Reverberator is an extension to Schroeder's reverberator. There is only one all-pass filter and Moorer employs exactly six comb filters. It also proposes use of low-pass filters. The parameters for tapped-delay lines, gain factors and filter coefficients were presented for Boston Symphony Hall in Moorer's original work [22].

The output sounds not metallic and it is relatively free of artifacts comparing to Schroeder's reverberator. However, Moorer's reverberator still lacks the frequency dependent reverberation. An improvement to Schroeder's reverberation scheme is that Moorer accounts for the effects of air absorption.

2.2.5 FDN

FDN, which is proposed by Stautner and Puckette, is a multi-channel reverberator that employs four channels [23]. FDN utilizes parallel delay lines which are recursively connected over a unitary feedback matrix. Selection of a unitary matrix ensures that the system is preserving energy. One can pick any unitary matrix so that the gain factors could be manipulated. Low-pass filters

could be connected to delay lines in order to achieve frequency-dependent reverberation.

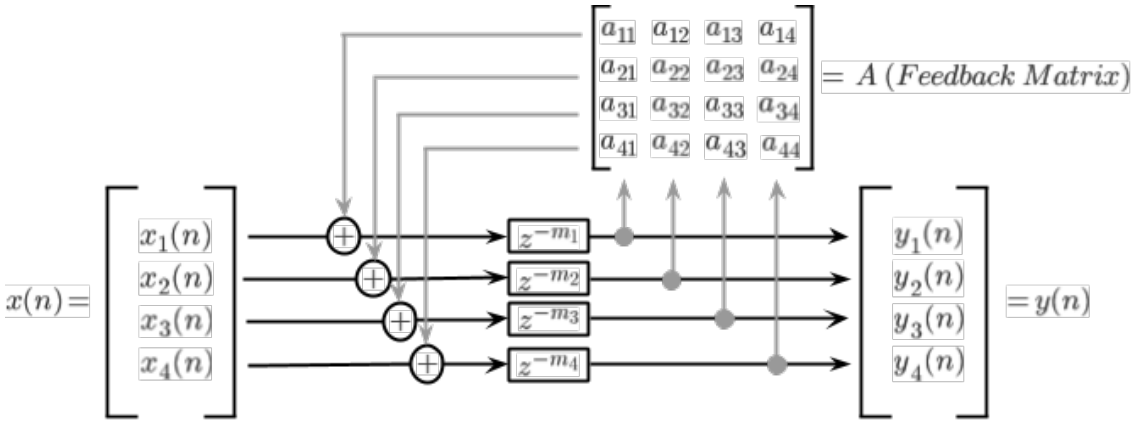

Figure 2.11: FDN reverberator.

In Figure 2.11, the feedback matrix A should be a unitary matrix in order to ensure stability of the system. Four parallel delay lines are recursively connected via the feedback matrix.

A popular derivation of FDN reverberator is Jot's reverberator [24]. Jot exclusively defined feedback matrices to be Householder matrices. However, Jot's reverberator provides only single channel of output.

2.2.6 DWN

DWN consists of many individual DWM nodes that build a closed network. Similarly to DWMs, DWN junctions have bidirectional delay lines between each other. A trade-off of energy loss can be exploited to achieve richer reverberation. One can employ many DWM junctions to form a DWN network [25]. DWN junctions operate similarly to their DWM counterparts. From any node, input could be fed or output could be obtained.

A cubic DWM mesh which is demonstrated in Figure 2.12 can be fit into a shoe box-type of room to simulate its acoustical properties. In Figure 2.12, light blue nodes are the individual junctions, whereas lines are the bidirectional delay line elements.

2.3 Edge Diffraction

A room geometry with re-entrant angles introduce the phenomenon called edge diffraction. Acoustic edge diffraction occurs when the traveling wave is obstructed by an impedance discontinuity in a part of its propagation path. Edge diffraction is especially significant in settings in which source and receiver are not visible to each other. Similarly to specular reflections, edge

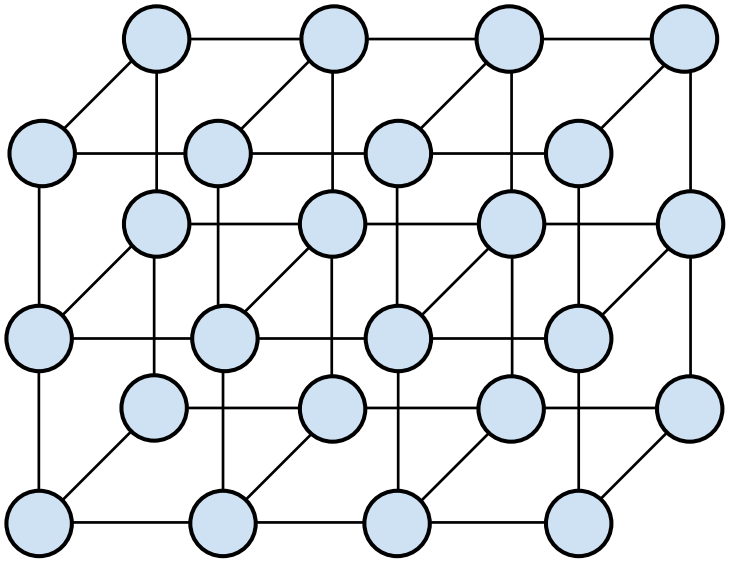

Figure 2.12: A cubic DWM mesh.

diffraction terms can have first and higher-order components.

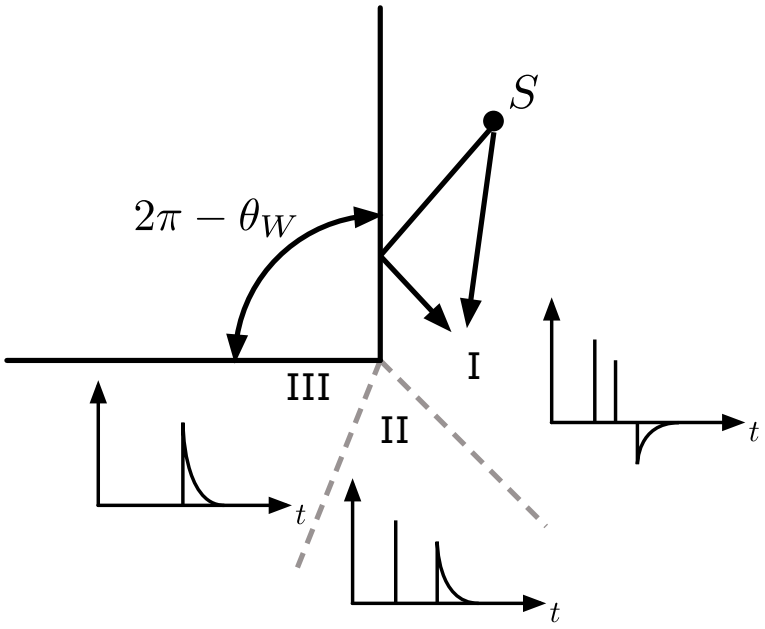

Figure 2.13: A wedge which divides the room to three distinct regions.

Phase, attenuation and delay caused by the edge diffraction depends on parameters such as wedge angle, the distance between source and receiver and their corresponding angle difference with the wedge itself [26].

Figure 2.13 show the contribution of edge diffraction to RIR in three subsections of the room. In the first region(I), source and receiver are visible to each other and a first-order specular reflection is also present, as it is on the RIR graph. In this case, phase of the diffraction term is negative. If the receiver is in the second region(II), there is no specular reflection and the phase of the

diffraction term is now positive. The third region(III) represents the scenario in which source and receiver are not visible to each other. Only positive-phase edge diffraction term contributes to the RIR.

Geometrical acoustics models can not account for edge diffraction, as there can be geometries with no visible image-sources. In such cases when there is no direct sound and specular reflections, RIR terms would be all set to zero. Within the third region in Figure 2.13, nothing is audible although one can hear the diffracted terms in a real setting. Therefore, in order to calculate edge diffraction terms, some methods are proposed.

2.3.1 Kirchoff Diffraction Approximation

This approximation method can be used in both time and frequency domains. Although, Kirchoff diffraction approximation had use in both seismology as well as acoustics, BTM expressions shown that the Kirchoff approximation yields large errors in both low and high frequencies [27].

2.3.2 Uniform Theory of Diffraction

UTD holds the high frequency assumption. Therefore, at low frequencies, UTD leads to errors. It found use in optics and electromagnetic scattering problems as well as in acoustics, until BTM model proposed a general solution [28]. In UTD, diffracting edges are assumed to be in infinite length. Besides, source and receiver are assumed to be in far field. Therefore, modeling capabilities of UTD model are short-handed when simulating room acoustics as in room acoustics, source and receiver are usually in near-field and wedges have finite length. However, UTD does not integrate along paths which makes it useful to incorporate with existing geometrical acoustics models by assigning delay lines for each path between source and receiver along the infinite wedge.

2.3.3 Biot-Tolstoy-Medwin Technique

In order to model a wedge and its contribution in diffraction, one should determine the angle between source and the wedge, θ_r, the angle between receiver and the wedge θ_s and finally the wedge angle θ_w. Geometry of the wedge could then be expressed in cylindrical coordinates; by defining the distances between source-wedge r_s and receiver-wedge r_r. Then the IR contribution h_d for the rigid wedge could be expressed as:

$$h_d(\tau) = -\frac{cv}{2\pi} \frac{\beta(\tau)}{r_s r_r \sinh \eta(\tau)} H(t-\tau) \tag{2.1}$$

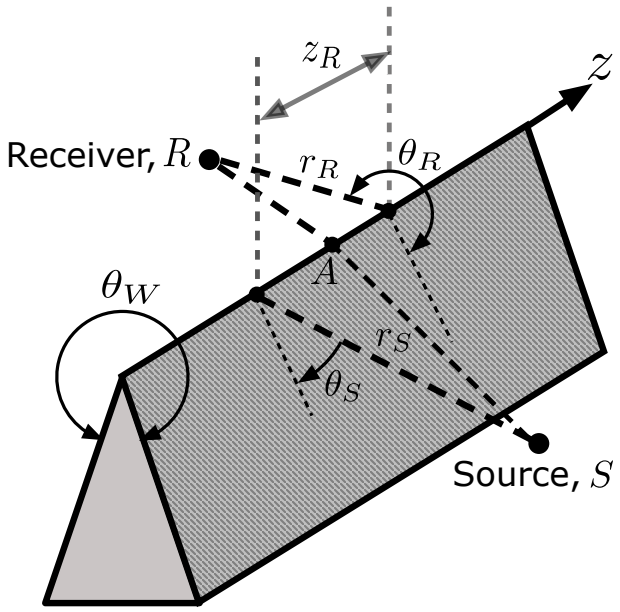

Figure 2.14: Wedge geometry according to BTM model.

whereas $\beta(\tau)$ and $\eta(\tau)$ are the following expressions:

$$\beta(\tau) = \beta_{++}(\tau) + \beta_{+-}(\tau) + \beta_{-+}(\tau) + \beta_{--}(\tau), \tag{2.2}$$

$$\beta_{\pm\pm}(\tau) = \frac{\sin[v(\pi \pm \theta_s \pm \theta_r)]}{\cosh[v\eta(\tau)] - \cos[v(\pi \pm \theta_s \pm \theta_r)]} \tag{2.3}$$

$$\eta(\tau) = \cosh^{-1}(\frac{c^2\tau^2 - (r_s^2 + r_r^2 + z_r^2)}{2r_s r_r}) \tag{2.4}$$

In the equations above, c is the speed of sound, $\boldsymbol{v}$ is the wedge index $\frac{\pi}{\theta_w}$, H($\tau - \tau_0$) is the Heaviside unit step function. The time τ_0 equals to the distance L_0 divided by c, where L_0 is the shortest path distance between source and receiver which connects along a point on the wedge.

A wedge divides a room into three subsections as it is demonstrated in Figure 2.13. Within each section, contribution from specular and diffracted components may differ. Besides, phase of the diffraction contribution may change according to wedge setup. In Equation 2.1 the phase difference, directivity patterns and attenuations are all accounted by the expression.

Although BTM edge diffraction model is similar to UTD, it does not hold the high frequency assumption. Therefore, it is employed in acoustics problems as it yields correct solution for lower frequency terms as well [29]. However, a straightforward approach to integrate the diffraction components over a given wedge is computationally costly therefore it may not be feasible for an interactive application. Another major issue is the singularity of the function at τ=τ_0. The singularity is also present when the receiver moves between boundary zones shown in Figure 2.13 [30]. Therefore, straightforward im-

plementation of an edge diffraction model built upon BTM was not possible until secondary source models are proposed.

2.3.4 Secondary Source Model

Admitting BTM model proposes a solution for the case where there is an infinite wedge, finite wedges require special treatment. One should note that expression in Equation 2.1 is explicit, and it is for a single point on the infinite wedge . It does not hold any information regarding the attenuation as well. Such properties prevent BTM model to be directly used in digital artificial reverberators.

Svensson et al. proposes the impulse response contribution a single point on a finite wedge [31]. They also suggest that by integrating over the wedge, one can obtain the diffracted contribution to impulse response. For room acoustics computations, each point over the wedge are assumed to be acting as a secondary source. Thus, computational complexity and accuracy depends on the resolution of the secondary source points. Efficient secondary source subdivision strategies were proposed by Calamia et al [32]. Moreover, an efficient auralization technique for edge diffraction is proposed by Lokki et al [33]. However, in SDN setup, given that bidirectional delay lines should be introduced for each and every secondary source, an approximation should be done in order to speed up the method. Another way of speeding up the diffraction calculations is to employ pruning methods, which culls higher-order diffraction terms from the RIR [34].

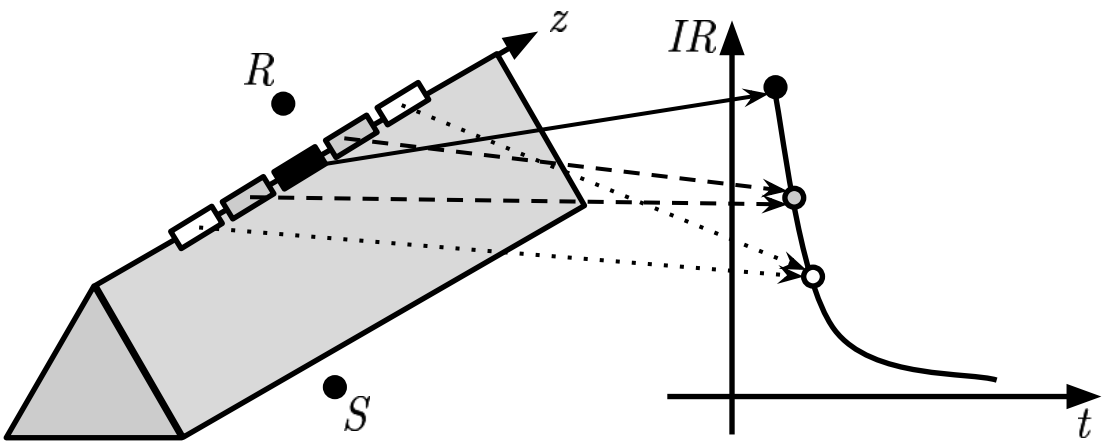

Figure 2.15: Secondary sources and their contribution to RIR.

In Figure 2.15, contribution from the apex is highlighted with black dot. Besides, contribution from the two symmetrical sides of the wedge is shown with grey and white dots. Since they are symmetrical, the distance between source-wedge point-receiver is the same for both points. This leads to equivalent attenuation, phase and delay. So they add up to one component in their corresponding delay time.

Similar to BTM model, the secondary source model also employs polar coordinates of source and receiver: (θ_s, r_s, z_s) and (θ_r, r_r, z_r), which could be interpreted as the terms in Figure 2.14. The impulse response $h(\tau)$ is the integral of the expression:

$$h(\tau) = -\frac{v}{4\pi} \sum_{i=1}^{4} \int_{z1}^{z2} \delta(\tau - \frac{m+l}{c}) \frac{\beta_i}{ml} dz \tag{2.5}$$

where v is the wedge index $\frac{\pi}{\theta_w}$, c is the speed of sound, m and l are the distances from source and receiver respectively while β_i is defined as:

$$\beta_i = \frac{\sin(v\varphi_i)}{\cosh(v\eta) - \cos(v\varphi_i)} \tag{2.6}$$

In Equation 2.7, φ_i is the angle combinations $\pi \pm \theta_s \pm \theta_r$ as it is shown in Equation 2.2 whereas η is the auxiliary function:

$$\eta = \cosh^{-1}(\frac{ml + (z - z_S)(z - z_R)}{r_S r_R}) \tag{2.7}$$

Integrating along the wedge endpoints ***z1*** and ***z2***, one can find the impulse response contribution for the finite wedge. The memory and CPU constraints of taking many points along the wedge enforce users to sacrifice between performance and accuracy.

CHAPTER 3

PREVIOUS WORK

Among the artificial reverberators that have been proposed for simulating room acoustics, SDN offers a reasonable trade-off between accuracy and computational complexity. This section discusses and reviews the SDN model.

3.1 SDN Model Overview

SDN models are proposed as a mid-point between full-scale room simulations and artificial reverberators. It is considered as one of the significant models proposed within the last fifty years of artificial reverberation [35]. Working principle of SDN model is similar to FDN and DWN approaches as the model employs scattering nodes and bidirectional delay lines. Geometrical acoustics properties such as source, receiver positioning, directivity and frequency-dependent wall absorption are supported as well. SDN guarantees accurate first-order reflections while it offers plausible results in higher-order reflections. Parameters related to room acoustical measurements such as RT60 and EDT are also closely approximated. Informal listening tests on the original model suggest that SDN models produce natural sounds without noticeable artifacts. A frequency-domain approach for SDN networks are also possible. [36].

SDN features scattering wall nodes that are placed on the intersection between the wall surface and the line defined between receiver and first-order image source. These scattering nodes are shown in Figure 3.1 with letter **S**. Bidirectional delay lines are employed in between scattering nodes, which are indicated with arrows in Figure 3.1. Source-to-node, node-to-receiver and source-to-receiver delay lines are mono-directional and are shown with dotted and dashed lines in Figure 3.1.

3.2 SDN Definitions

SDN uses source, receiver and room definitions. A source is defined by its position, directivity function and its acoustical axis. A receiver is defined as

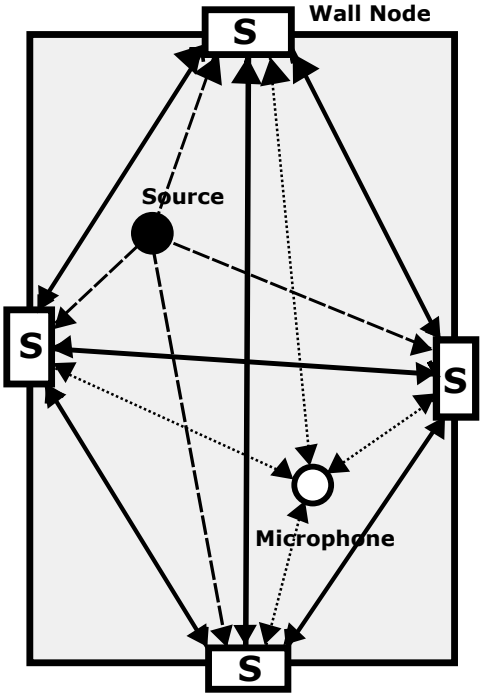

Figure 3.1: Overview of the SDN model.

same as the source. The room geometry is defined by the wall surfaces and their corresponding frequency-dependent wall absorption coefficients.

3.3 Calculations Involved

Given the source, receiver and room definitions; SDN model first calculates the first-order image sources. For each image source, it then calculates the line connecting the image source and receiver node. The intersection point of the line with the wall surface plane is the position where SDN node would be placed. The process is illustrated in Figure 3.2.

In order to design the wall filters for modeling frequency-dependent absorption, tabulated absorption coefficients are employed. Then minimum-phase wall filters are defined. For each frequency band, separate networks could be run to get results in desired octave bands.

3.4 Source to Receiver Path

Directivity is accounted by finding the angles between source and receiver nodes. The propagation delay and 1/R attenuation law are also considered. This path which simulates the direct sound does not require any special treatment and the methodology is similar to ISM or any other geometrical acoustics solution.

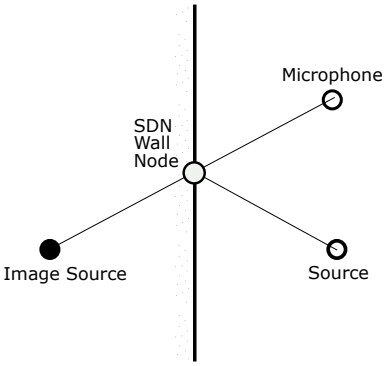

Figure 3.2: Finding position of a SDN wall node.

3.5 Source to Wall Node Paths

Similar to source to receiver approach, angles between source and wall nodes are also found in order to calculate directivity pattern. The incoming samples from the source are added to node interconnections after the propagation delay. 1/R attenuation is present between source to wall paths.

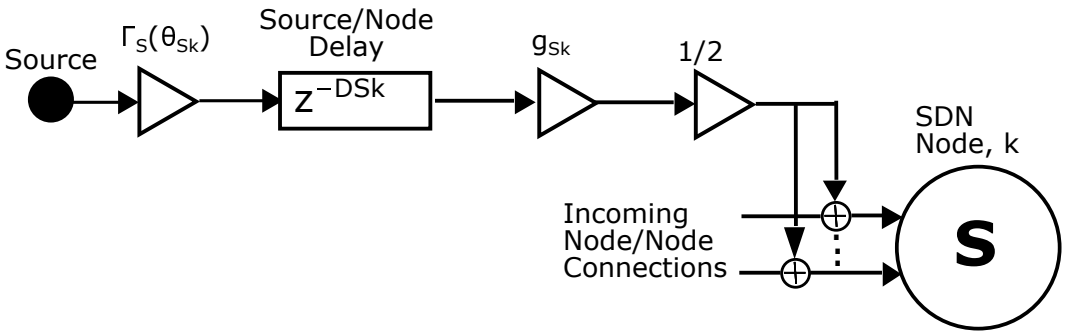

Figure 3.3: Connection between a wall node and source.

The block diagram in Figure 3.3 represents directivity pattern with $r_S(\theta_{Sk})$. Attenuation is g_{Sk} and output from each node is multiplied with 1/2 before they are added in an SDN node, similar to the DWM method. The propagation delay between source and the wall node is indicated with $z^{-D_{Sk}}$.

3.6 Node to Node Paths

Node interconnections are obtained by bidirectional delay lines. These delay lines also simulate the energy lines between walls of the room. In order to simplify the model, a single gain element $H_k(z)$ could be substituted with the absorption filters. While propagation delays $z^{-D_{Sk}}$ are calculated according to node to node distances, 1/R distance attenuation law is not accounted deliberately for these paths.

The block diagram in the Figure 3.4 shows that SDN nodes act similar to a

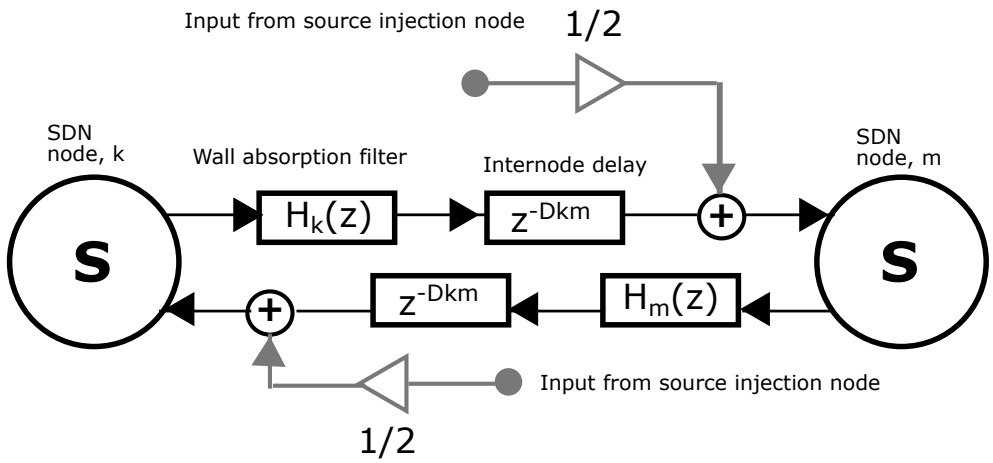

Figure 3.4: Node to node connections.

DWM node. The wave arriving to each node propagates back to all other nodes. A lossless scattering matrix is employed as in the DWM model. Reverberator is then defined by the nodes and their interconnections. One shall also note that directivity is not accounted between two nodes.

3.7 Node to Receiver Paths

Node to receiver path similarly account for directivity by finding the angles between nodes and the receiver. Propagation delay and 1/R attenuation are employed as well.

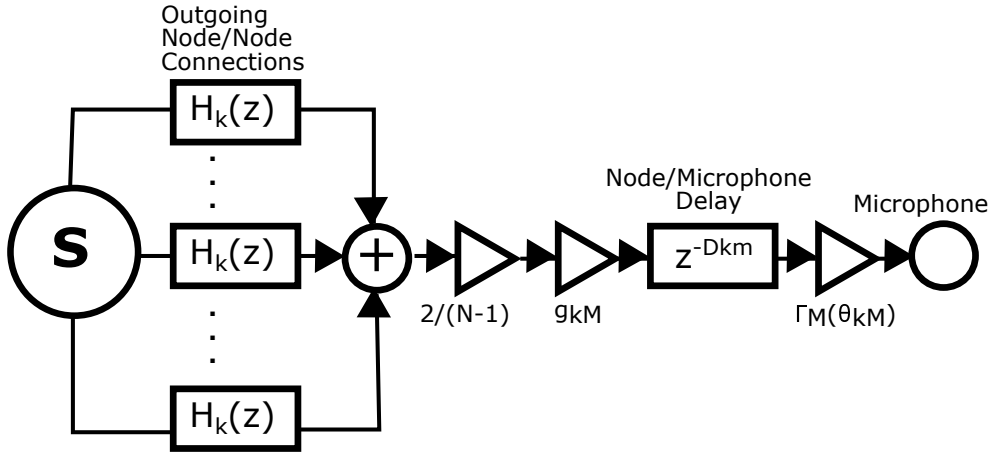

Figure 3.5: Connection between a wall node and receiver.

The block diagram in Figure 3.5 suggest that contribution from all outgoing node-to-node interconnections are summed up firstly. This summation is then multiplied by the coefficient, 2/(N-1), where N is the number of nodes in the network. The distance attenuation between node k and the receiver is indicated as g_{kM}. Directivity between node k and the receiver is then $r_M(\theta_{kM})$. Lastly, the propagation time delay between node k and the receiver is defined

with $z^{-D_{kM}}$.

3.8 Interactivity

If source or receiver moves, state of the SDN network changes. Since, prior action would result in new image source positions, SDN wall node positions are updated accordingly. This would then lead to changes in the delay line lengths and their contents. Attenuation coefficients should be calculated as well since the distances between components change.

One exception to these calculations is the case when only rotation of the source or microphone changes. In this case, update of whole network is not necessary as the distances between components do not change. Delay line lengths and their contents are not manipulated. Therefore, calculating the new directivity vectors suffice.

3.9 Recording Techniques

In order to get multi-channel output, the network does not need duplication. Instead, for near-coincident coincidence microphone settings, one can only change directivity functions, which would reduce complexity and memory footprint. This approach can be used to apply panning. Binaural output is also possible. However, in binaural setting, for each ear, an HRTF should be employed. This procedure does not require duplication of the network as well although updating and maintaining a HRTF filter bank would be a computational challenge.

CHAPTER 4

PROPOSED METHOD

As described in prior chapters, this study aims to allow SDN model to work in arbitrarily shaped rooms. Since SDN relies on first-order image sources in order to determine scattering node positions, a scheme that would propose image sources in an arbitrarily shaped room is required.

A method to extend the image-source model to arbitrarily-shaped polyhedra is proposed by Borish [37]. According to Borish, visibility and validity checks should be performed in order to determine image-sources that would contribute to RIR. If a surface has re-entrant angles and the surface fails to generate a valid and visible image-source, then the surface is a wedge. A wedge does not contribute to geometrical acoustics terms. Thus, the wedge contributes to RIR with edge diffraction terms.

Svensson et al. proposed a model for edge diffraction models in room acoustics simulations [38]. Their methodology involving secondary sources require delay lines for each secondary source defined over the wedge. This characteristic of the edge diffraction model requires prohibitively high memory requirements. Instead of accounting for each and every point, an approximation method derived.

4.1 Visibility Checks

Borish suggests that in order to a image source to be visible, the line formed from the image source to receiver should only intersect with the surface that is generating it. In case the line intersects with another surface within the room, then the image-source is obstructed and there is no direct energy transfer between receiver and the image source.

Figure 4.1 demonstrates the visibility requirement. In order to image source to be visible to the receiver, the receiver must be in the shaded region. This region is defined by the beams covered by the endpoints of the reflecting surface. This requirement also tells whether there is a specular reflection between the surface and the receiver is present.

For the cases in which the receiver is not visible by the image source, SDN

node is placed to the centroid of the reflecting surface. Since lack of specular reflections suggest that there is no first-order reflection as well for a given surface, the centroid assumption does not violate the accuracy of the SDN model for the early reflections.

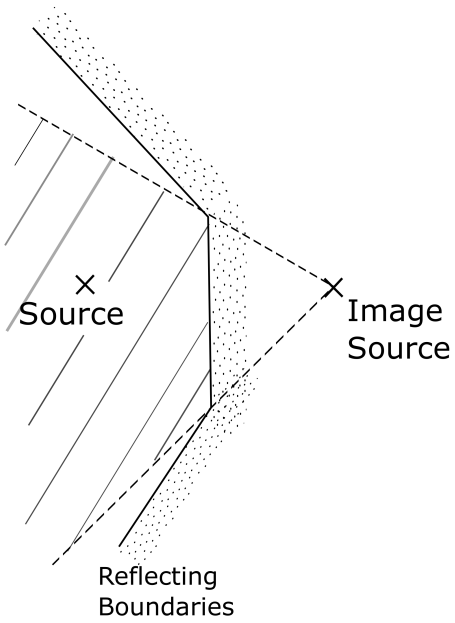

Figure 4.1: The region for the image source to be visible.

4.2 Validity Checks

An image source is considered valid if it reflects from the non-reflective side of the boundary. In order to a scattering node to be generated, the image source related to that node must hold both visibility and validity properties.

An example scenario with valid and invalid image sources is shown in Figure 4.2. A valid first-order image source S_1 is first created by the reflective side of the boundary (1). It is valid and visible. The second-order image source derived from S_1 is S_{12} and it is generated by the reflective side of the boundary (2). However, the second-order image source S_{12} is not visible since it is obstructed by the boundary (1). A similar situation occurs for the second-order image source S_{21}, which is generated by the first-order image source S_2 and boundary (1). S_{21} is not visible as it is obstructed by boundary (2). Therefore, S_{12} and S_{21} are not valid image sources and they should be pruned from the further calculations.

In order to incorporate extension proposed by Borish to original SDN model,

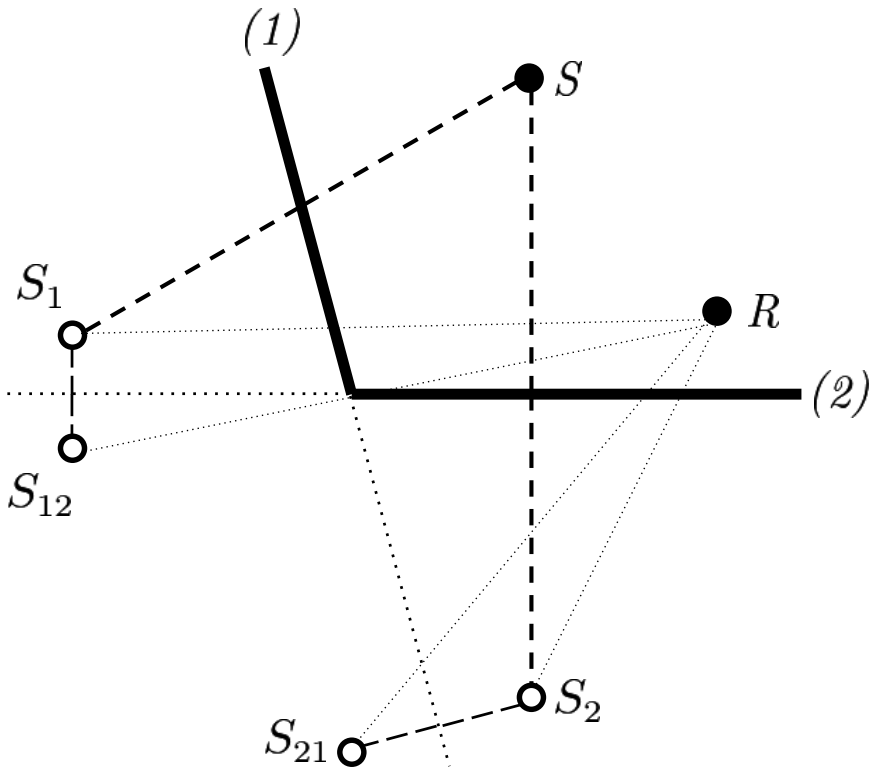

Figure 4.2: Case of a valid and invalid image sources.

criteria for visibility and validity is imposed on source-to-node, node-to-node, source-to-receiver and node-to-receiver connections. Node-to-node interconnections also require a back-face detection as nodes should be visible from front in order to have energy exchange between each other.

Finding SDN node positions then require intersection tests. These intersection tests should determine that there is only reflecting surface lies between image source and the receiver. One must ensure that this intersecting surface belongs to the reflective boundary. A pseudo-code of this routine is as follows: In order to determine the intersection between the line and a polygon, the problem is first deducted to the 2D case by finding the intersection point of the plane defined by the polygon surface and polygon position vector. If the angle between intersection point and two consecutive vertices sum up to 2π radians for the whole polygon, then the point is inside the polygon. 2D

Require: Wall polygons w[i], source s, receiver r
  for all i in w[i] **do**
    Find the image source point s'
    Form a line between s' and r
    if Number of intersections with walls = 1 and intersected wall is **i then**
      Place SDN node at intersection with w[i]
    else
      Place SDN node at the centroid of w[i]
    end if
  end for

Figure 4.3: Algorithm 1, determines SDN nodes and performs validity checks.

```
Require: Wall polygons w[i], source s and receiver r
  for s and r do
    intersections ← 0
    for wall[i] do
      if wall[i] intersects s or r then
        intersections = intersections + 1
      end if
    end for
    if intersections = 0 then
      gain = 1
    else
      gain = 0
    end if
  end for
```

Figure 4.4: Algorithm 2, performs visibility checks.

point-in-polygon strategies are further surveyed thoroughly by Haines [39].

Visibility checks in Algorithm 2 are to ensure that the source-to-node and receiver-to-node paths are not obstructed by any walls. If they are obstructed, they do not contribute as expected since their gain factors are set to 0. For each wall, the source or the receiver is tested once. Therefore, if there are n walls, the complexity is then $O(n)$. Same checks should be made for the line formed by each node-to-node connection and each and every other wall in the network. The complexity becomes $O(n^2)$.

Another visibility check should be performed between source and receiver as well. Similarly to Algorithm 2, if there is a wall detected between the source and the receiver, gain factor of the direct sound component should be set to 0. The line between source and the receiver is tested for each wall. So, if there are n walls, this time the complexity is $O(n)$. However, the overall complexity of visibility and validity checks are dominated by the $O(n^2)$ term introduced by the node-to-node visibility checks.

4.3 Edge Diffraction

In order to account for the cases where edge diffraction is present, an accurate yet efficient enough method to run at interactive rates should be derived. Given that even for subdivided edges, an SDN model would require bidirectional delay lines between source-receiver and the points along the divided wedge, memory costs could become a bottleneck. Maintaining these delay lines while source and receiver is moving would become another challenge.

Given that bidirectional delay lines should be introduced for each and every secondary source, an approximation is proposed in order to speed up the method. The edge diffraction impulse response could be approximated by an exponential function. An exponential function with a known decay parameter can be implemented as a first-order IIR order.

When tested against a range of decay parameters, the sum of squared errors between the exact calculation from secondary source model for edge diffraction and becomes apparent. The log-log plot in Figure 4.5 suggests that the function of sum of squared errors over different decay parameters alpha has a global minimum, which the unique mode of the function. This observation is exploited in order to prune the exhaustive integration steps along the wedge which degrades the number of delay lines to one.

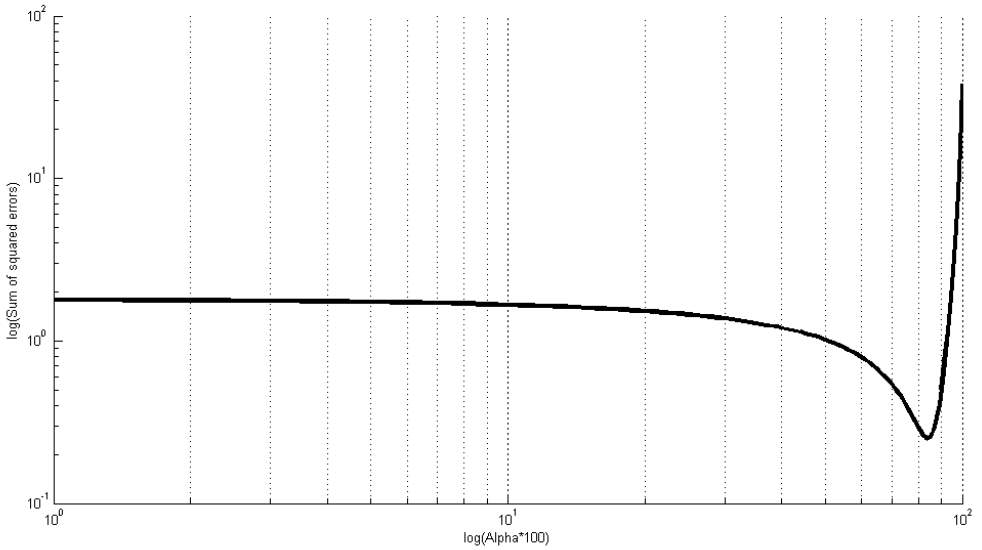

Figure 4.5: Log-log plot for sum of squared errors vs. decay parameters.

Formally, global minimum of the sum of squared errors is sought and this translates into the optimization problem:

$$min_{\alpha} f(\alpha) = \sum_{n} (h_{diffr}(n) - (\alpha)^{n-1})^2 \quad \alpha \in [0, 1] \tag{4.1}$$

In Equation 4.1, the term h_{diffr} is the normalized diffraction terms along the finite wedge. The normalization is carried out such that the impulse response contribution from the apex is equal to 1. α^{n-1} is the corresponding estimation for a particular point on wedge, n. Therefore, it is ensured that first term of the approximation is also equal to 1. Subtraction between normalized diffraction term and the approximation term is then squared in order to find the squared error term. For each point on the finite wedge, this would then correspond to sum of squared errors. Global minimum is the point where sum of squared errors is minimized. Thus, corresponding α would yield the best-fit decay parameter for the finite wedge impulse response approximation.

For this one-dimensional optimization problem, several approaches could be implemented. A naive approach would be to compute all error terms for $\alpha \in [0, 1]$ and perform a linear search to find the global minimum. However, this would require calculating all possible squared error functions for a given precision level, which is inefficient in terms of computational load and mem-

ory. Given that, source and receiver may move constantly in an interactive setting, a faster approach is more desirable. Binary search is a fast algorithm with worst-case $O(\log n)$ complexity, but the algorithm requires sorted input. Figure 4.5 demonstrates that the error function is by no means sorted. Therefore, it is not possible to utilize binary search as well.

Another possible approach could be the gradient descent method which is extensively used in finding local minimum or maximum of a multi-dimensional function. Since Equation 4.1 defines a unimodal function, finding local minimum would actually yield the global minimum. Therefore, gradient descent is effective in finding the optimal alpha. For our setting, gradient descent would work as follows:

$$\alpha_0 = \alpha_0 - \beta \frac{d}{d\alpha_0} f(\alpha_0) \qquad \alpha_0 \in [0, 1] \tag{4.2}$$

In Equation 4.2, the process is repeated until convergence. Step size is denoted as β for clarity. At each step, α_0 converges to the local minimum by the rate of β times the slope of the function $f(\alpha_0)$ at the point α_0, which is equal to the first derivative.

Here, one should carefully select the step size β. A small value of β might yield exceptionally slow running times. This is due to the fact that slope of the function converges to 0 as it reaches to the a local minimum. Smaller step sizes combined with descending slope would require more processing time until convergence to the local minimum. Contrarily, a higher value of β would imply wild jumps at each iteration, so it may terminate before it can find the local minimum. Yet, in order to calculate slope, additional calculations for each α_0 term are required. Therefore, even though gradient descent is useful in our case, the ternary search algorithm would not require us to decide on a step size.

Thus, in extended SDN model, impulse response gains for a manageable number of points are calculated firstly. The squared error between edge diffraction response and exponentially decaying function is a unimodal function as it can be observed from Figure 4.5. Therefore, one can exploit the process by employing ternary search, as the algorithm is guaranteed to find local minimum or maximum, depending on how one interprets it.

Gain of the shortest path between source and receiver yields the highest gain. This gain acts as an attenuation factor for the wedge. First-order IIR filter that yields the decay parameter then approximates the sample to the least squared-error function. Thus, instead of exhaustively integrating over the wedge, one can obtain audible results in a few steps.

An earlier scheme for ternary search is proposed by Dijkstra [40] and our implementation is based on Hesselink's prior work [41]. Ternary search is a very fast algorithm with $O(\log n)$ complexity. However, at each step, decay parameter alpha is tested against existing values for the normalized edge diffraction, so real complexity is closer to $O(n)$ where n is the number of calculated

diffraction points. The pseudo-code of the edge diffraction approximation routine is then as follows:

Algorithm 3
Require: Wedge parameters in Equation 2.1, sampling rate f_s, maximum time (1sec), empty array for comparison comp[i]
period = 1 / f_s
$\tau = \tau_0$
while τ < maximum time **do**
Calculate h_d according to Equation 2.1, store it in comp[i]
τ += *period*
end while
max = *comp[0]* {Normalization of array *comp[i]*}
for each element i in *comp* **do**
comp[i] /= *max*
end for
left alpha = 0.01
right alpha = 0.99
while true **do**
left error = 0, right error = 0
if abs(right alpha - left alpha) < ϵ **then**
return best alpha = (right alpha + left alpha) / 2
end if
left third = left alpha + (right alpha - left alpha) / 3
right third = right alpha - (right alpha - left alpha) / 3
left value = *comp[0]*
right value = *comp[0]*
for each element i in *comp* **do**
left error += squared error of left alpha
right error += squared error of right alpha
left value *= left third
right value *= right third
end for
if left error < right error **then**
right alpha = right third
end if
if left error >= right error **then**
left alpha = left third
end if
end while

Figure 4.6: Pseudocode of Edge Diffraction Approximation Routine.

The accuracy of the decay parameter depends on the time interval in which diffraction terms are calculated. As time step t between each τ increases, less points are needed. However, this would lead to less accuracy. As a rule of thumb, we propose that it should be $1/f_s$ seconds where f_s is the sampling rate of the system. Since diffraction terms are only calculated once, the complexity of calculations are amortized in the long run. Depending on the rever-

beration time of the room, the wedge simulation may run for at most a couple of seconds. Since Equation 2.1 is an explicit expression, the results should be normalized and the points which singularities occur should be pruned.

CHAPTER 5

RESULTS AND DISCUSSION

In this section, results from extended SDN method are analyzed. Firstly, the edge diffraction approximation is tested against Svensson's Edge Diffraction Toolbox (EDB2). Then, RIRs for three room settings are compared with the results obtained from ODEON acoustical simulation software.

5.1 Analysis of Edge Diffraction Approximation

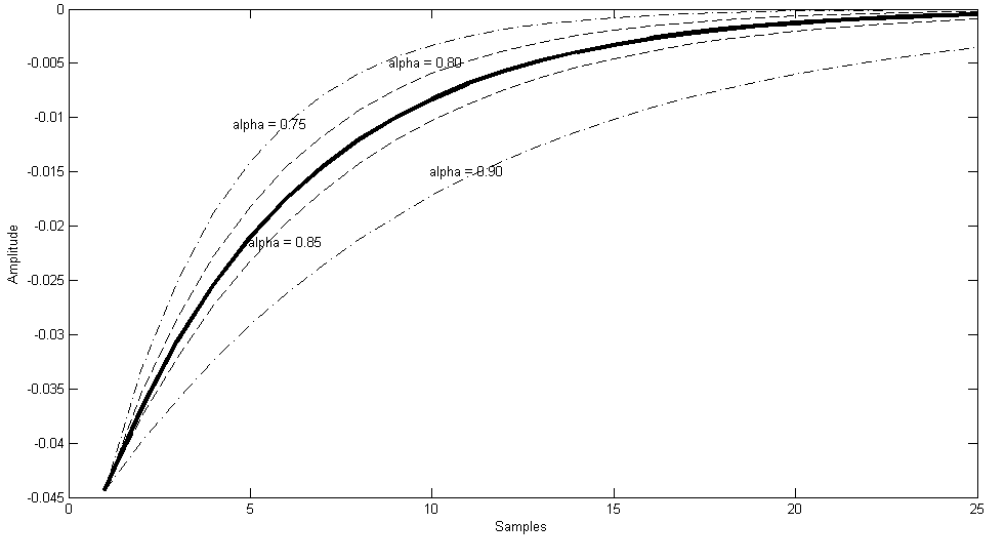

Figure 5.1: Comparison of edge diffraction terms and alpha approximation.

An analysis of the functions obtained from different alpha values and the exact diffraction components for a sample wedge is given in Figure 5.1. The diffraction components were obtained from Edge Diffraction Toolbox (EDB2) of Svensson [42]. The toolbox utilizes secondary source model to obtain edge diffraction RIR, which is addressed in prior chapters.

The bold line is the result obtained from EDB2 and the dashed lines are exponential functions with various decay parameters, alpha. It is noticeable that

for this case, the least-squared error alpha lies between 0.85 and 0.80. Instead of calculating edge diffraction for each and every sample, the ternary search finds a result between 0.85 and 0.80, depending on the ε given. The ε defines the precision and 0.01 would yield results that are plausibly close to original RIR obtained from EDB2.

5.2 Comparison With Exact Simulation Methods

In this section, an impulse response and an audio output of the algorithm would be compared with an exact ray-tracing and diffraction method.

For the purposes of testing, an acoustic simulation tool named ODEON is used. ODEON is an industry-standard software and it includes a combined method of ray-tracing and diffraction in order to carry out its calculations. RIR for three distinct room geometries were obtained by both models. For the sake of brevity and clarity, all walls have absorption coefficients of (α=0.2).

The informal listening tests were already carried out with previous SDN model and it is shown to have no audible artifacts.

5.2.1 Regular Rectangular Room

In this example, a rectangular, shoe box-type of room shown in Figure 5.2 with width of 9m, length of 7m and height of 5m is considered. In Figure 5.2, side view of the room is on the left. Source and receiver positions are marked from a top view on the right side. The source is placed to the middle of the room (4.5m, 3.5m, 2.5m) while the microphone is at position (2m, 2m, 1.5m).

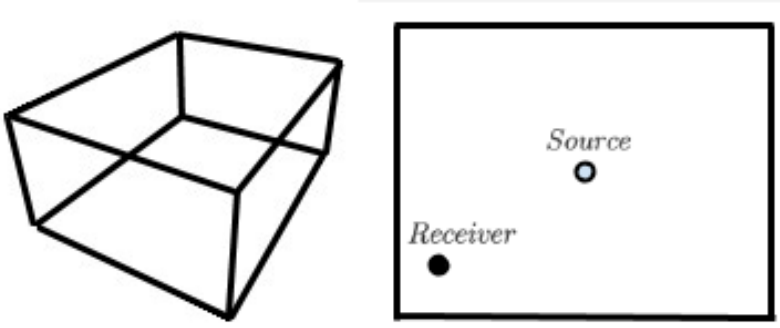

Figure 5.2: Geometry of the shoe box-type room.

The RIR obtained from SDN model and ODEON are as shown in Figure 5.3 and Figure 5.4 respectively:

For the regular room setup, SDN output shows that it could render early reflections and direct sound close to the ODEON output. One should note that, the richer late reverberation tail introduced by ODEON output is due to the

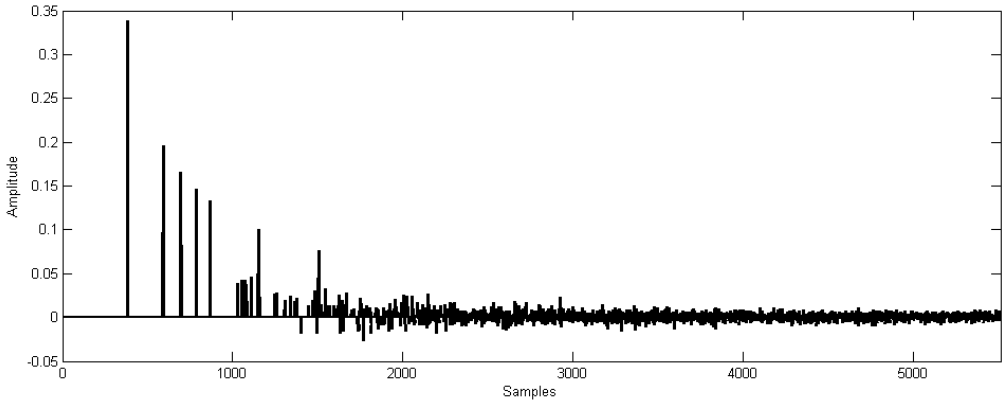

Figure 5.3: RIR output from SDN for regular room.

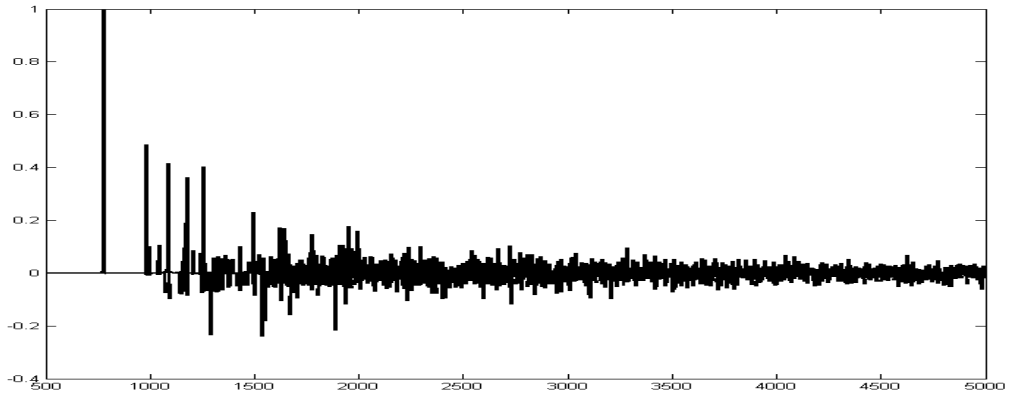

Figure 5.4: RIR output from ODEON for regular room.

fact that SDN renders the room with only 6 nodes in order to achieve real-time rates. In this case, 120 seconds of audio is processed in 22.847 seconds with SDN. In our case, regular rectangular room uses the same approach in the original model. However, our aim was to show that extended version does not affect the performance and regular rectangular room still runs fast as it is demonstrated by the run-time.

5.2.2 Trapezoidal Room

Another example is a trapezoidal room. This is similar to a regular room, but it introduces the visibility and validity checks. For this scenario, one of the walls within the room fails to generate a valid first-order image source. Therefore, an SDN node is placed in the centroid. The room geometry is as in Figure 5.5. Source position is (2m, 1m, 1m) and the microphone is placed at (3.5m, 3.5m, 0.5m). In the Figure 5.5, side view is to the left and to the right shows the source and receiver positions.

The RIR obtained from SDN model and ODEON are as shown in Figure 5.6 and Figure 5.7 respectively:

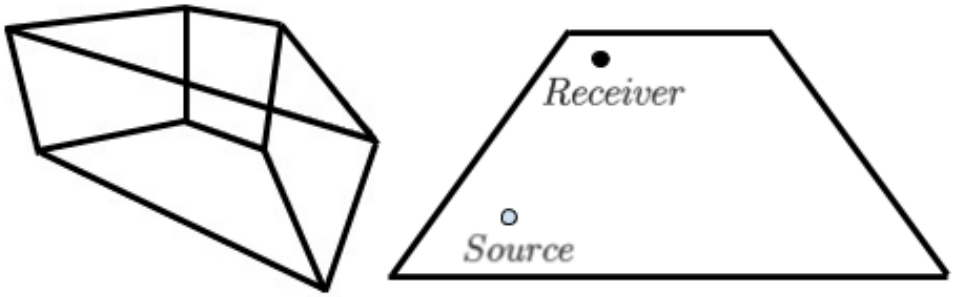

Figure 5.5: Geometry of the arbitrarily shaped trapezoidal room.

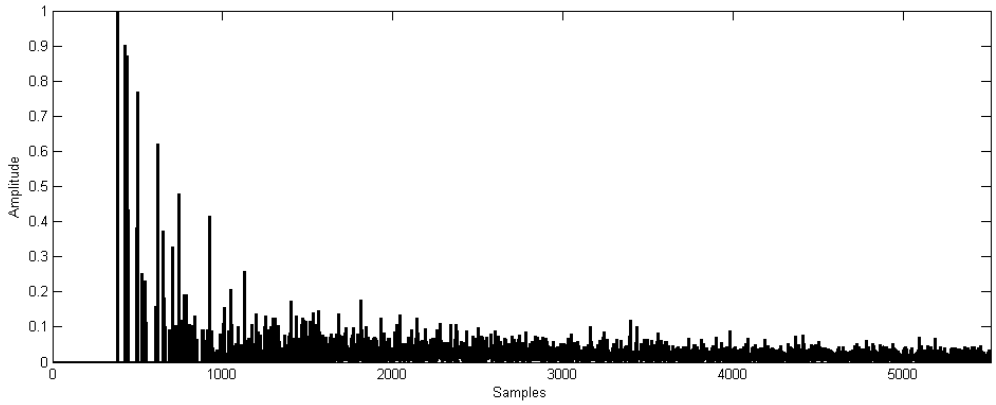

Figure 5.6: RIR output from SDN for arbitrarily shaped trapezoidal room.

Since this is a regular room turned into a trapezoid, the room is slightly more reverberant and it shows in both RIR outputs. The geometry is deliberately manipulated in a way that one of the walls do not produce a valid first order image source. Here, SDN output overly-estimates a first-order reflection by putting one of its scattering nodes to the centroid of the wall. However, higher order reflections and late reverberation compensates for the overestimation in the long run because of the richer reverberation tail introduced by ray-tracing solution. As demonstrated in Figure 5.7, the reverberation tail is slightly more powerful in ODEON output as expected.

According to ISO 3882-1 standards, some parameters are defined to measure room acoustics [43]. Early decay time (EDT), T_{20}, T_{30} and T_{40} were the parameters of our interest.

Table 5.1: Room acoustical measurements for the trapezoidal room.

	eSDN	**ODEON**
EDT	1.0244 s	0.7986 s
T20	0.5298 s	0.5685 s
T30	0.4942 s	0.5199 s
T40	0.4884 s	0.5066 s

In Table 5.1, extended SDN gives similar results except EDT in comparison with the ODEON output. The trapezoidal room is a rather small room with a

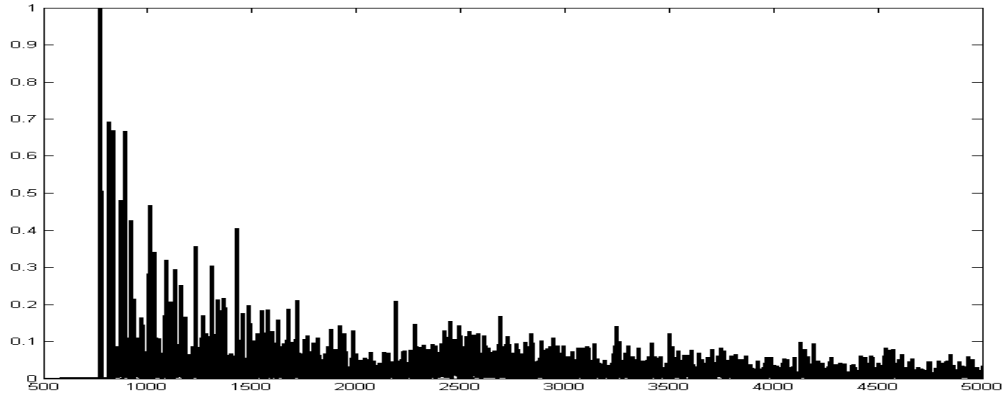

Figure 5.7: RIR output from ODEON for arbitrarily shaped trapezoidal room.

volume of $96m^3$ and the node placed on the centroid did not prevent the energy transfer. The unusual EDT value of extended SDN is caused by the steep curve caused by the early reflections. Since ODEON model relies on a hybrid ray tracing approach, from the early reflections, energy decay is downwards sloping as expected. However, in SDN model, the simulation is run by a limited number of nodes and this causes almost flat slopes and sharp decreases around early reflections.

5.2.3 L-shaped Room

The last example consists of an L-shaped room, This room introduces wall surfaces that are not directly visible to source and receiver. Therefore, their SDN nodes will be placed on centroid of the corresponding walls. In the L-shaped room, the wall nodes that are on centroids do not show up on first-order reflections at all. However, since they are connected through the node-to-node network, the energy transfer are expected to show up in higher-order reflections.

Geometry of the L-shaped room is shown in Figure 5.9. This room is similar to regular room, with a rectangular indent. Left side of the Figure 5.9 demonstrates the side view of the room whereas right side shows the top view. Source is positioned at (4.5m, 3.5m, 2m) while the receiver is at (2m, 2m, 1.5m).

RIR obtained from SDN model and ODEON is shown in Figure 5.9 and Figure 5.10. Since the walls which are not visible are only connected to the network through node-to-node network, late reverberation tail becomes significantly weak for the SDN model. Lower order terms are matching with each other.

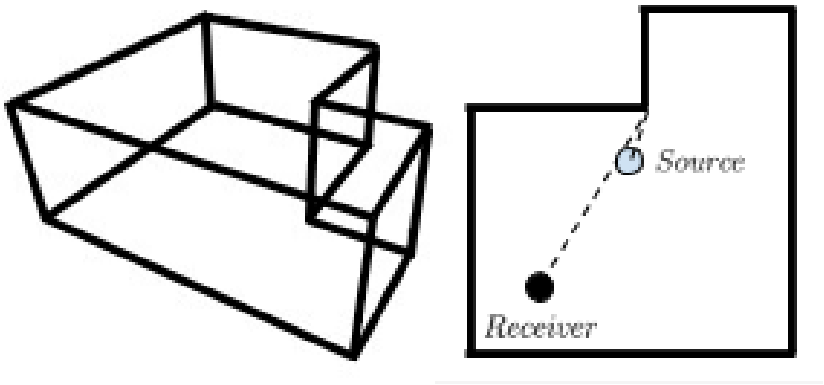

Figure 5.8: Geometry of the L-shaped room.

Table 5.2: Room acoustical measurements for the L-shaped room.

	eSDN	**ODEON**
EDT	0.9608 s	0.8714 s
T20	0.5065 s	0.5567 s
T30	0.4717 s	0.5141 s
T40	N/A	0.5003 s

In Table 5.2, the loss of energy transfer caused by the visibility checks is apparent. Since the nodes which are not directly visible to the source and receiver pair may only contribute at higher order reflections, energy decay is fast starting at early reflections. This is caused by the limited number of nodes contributing for the first-order reflections. Extended SDN can not intercept many higher-order reflections due to its limitations by-design.

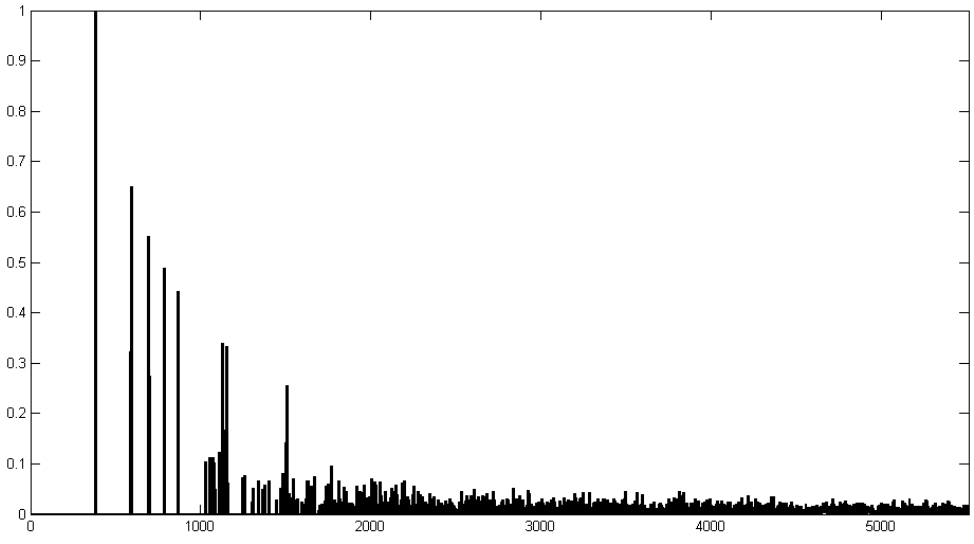

Figure 5.9: RIR output of SDN for the L-shaped room.

For the case of L-shaped room, since there are walls that are not connected by first-order terms even through polygon centroids, the energy decay curve

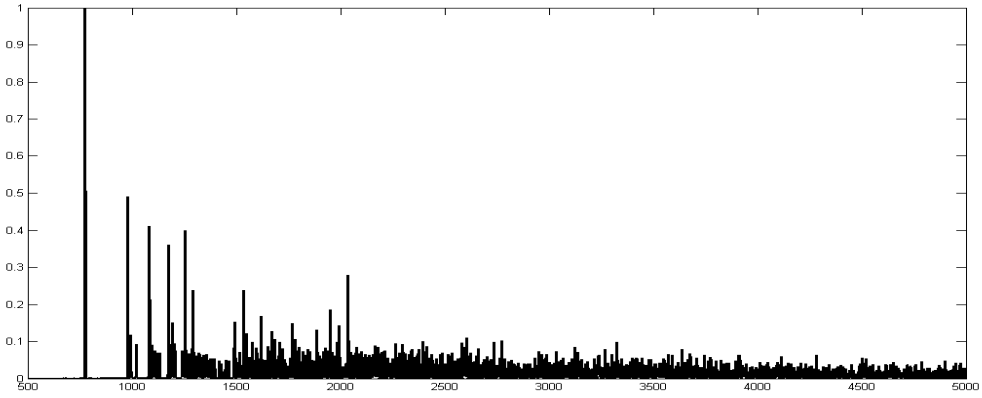

Figure 5.10: RIR output from ODEON for L-shaped room.

of the RIR generated from the SDN model should be double sloped. Energy decay curve is an indicator of remaining energy in the reverberator [44]. The double-slope suggest that a portion of the energy is lost in the parts of the room that is not visible to the receiver. EDC is defined as the integration in Equation 5.1. EDC for the L-shaped room is shown in Figure 5.11.

$$EDC(t) = \int_t^\infty h^2(\tau) d\tau \tag{5.1}$$

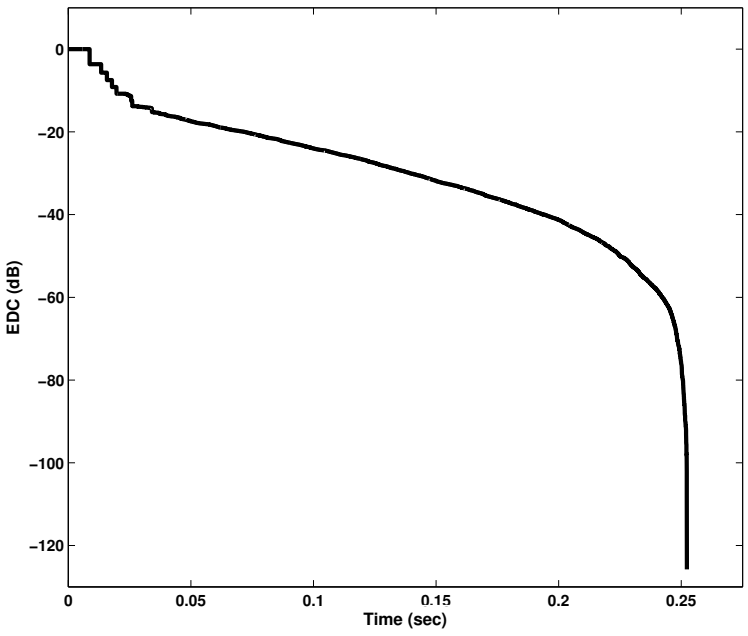

Figure 5.11: Double-sloped EDC for L-shaped room.

The spectrograms of extended SDN and ODEON outputs for trapezoidal and

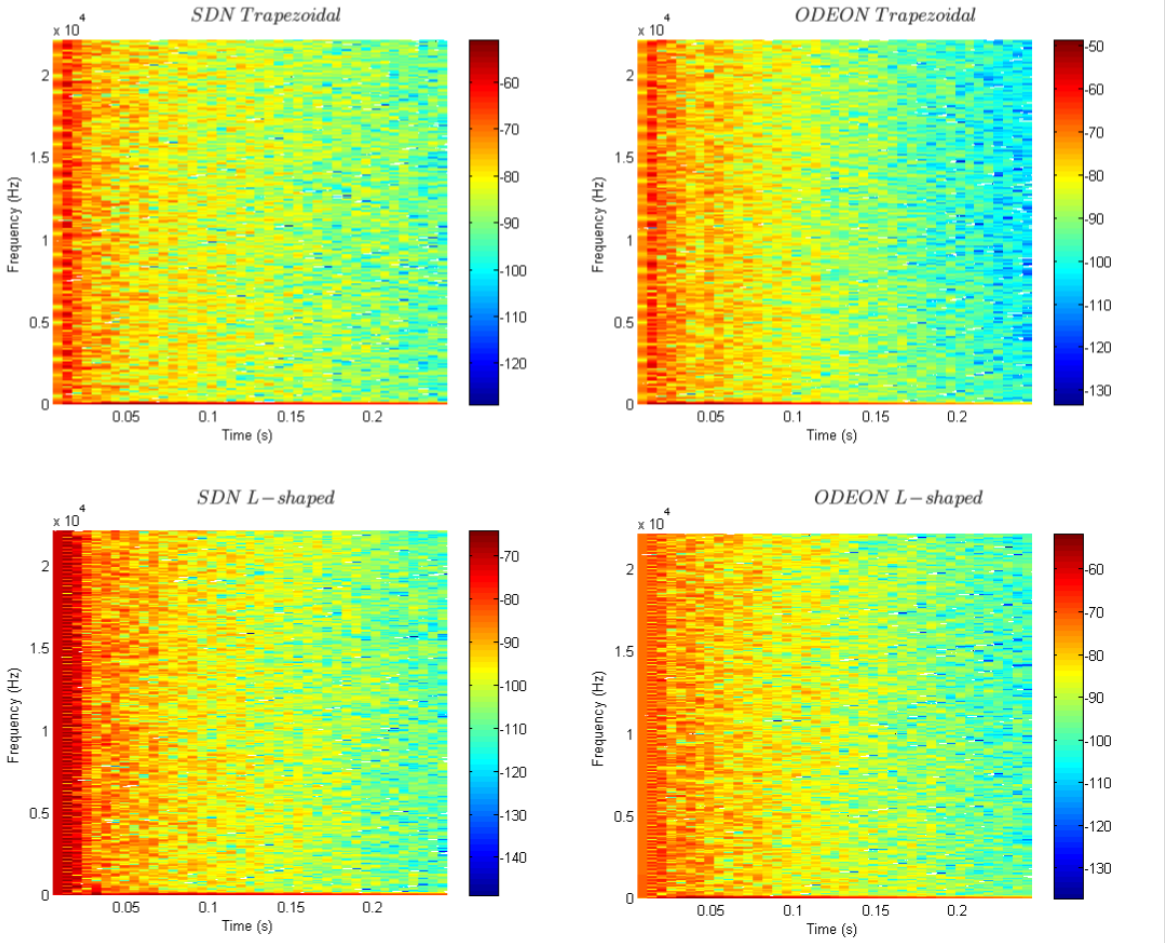

Figure 5.12: Spectrograms for extended SDN and ODEON output.

L-shaped room are demonstrated in Figure 5.12. The absorption coefficients are uniform for all octave bands as it is evident from the spectrograms.

5.3 Objective Evaluation of Audio Quality

Algorithms for objective assessment of audio quality allows flexibility in comparison to subjective evaluations and are in use recently. In order to produce meaningful statistics out of subjective tests, there are multiple criteria to be satisfied. It is recommended to have many subjects and they should be tested under same environmental conditions. The availability of objective assessment models according to corresponding ITU recommendations allowed us to perform objective evaluation. "Perceptual Evaluation of Audio Quality", namely PEAQ is used to assess the audio quality from the extended SDN model. PEAQ method requires 10-20 seconds of 16-bit PCM audio files with 48kHz sampling rate. When PEAQ finishes, an objective difference grade score in the range [-4, 0] is given as an output [45]. Each side of the interval corresponds to a impairment scale which between 'imperceptible' and 'very annoying'. Table 5.3 shows the impairment scale which applies for

PEAQ [46]. For lower signal qualities, another output parameter, Distortion Index (DI), is generated as well. DI is another complementary quality factor and its range is between [-12, 3]. For both ODG and DI, larger values indicates indicate better audio quality audio quality and both scales are continuous within their respective intervals. When ODG points to annoying impairments, DI is utilized. However, ODG of one system should not be compared with DI of another system. Differences between the curves of ODG and DI is shown at Figure 5.13, which is adapted from [47].

Table 5.3: Impairment scale according to ITU-R BS.1387.

Grade	Impairment	ODG
5.0	Imperceptible	0
4.0	Perceptible but not annoying	-1
3.0	Slightly annoying	-2
2.0	Annoying	-3
1.0	Very annoying	-4

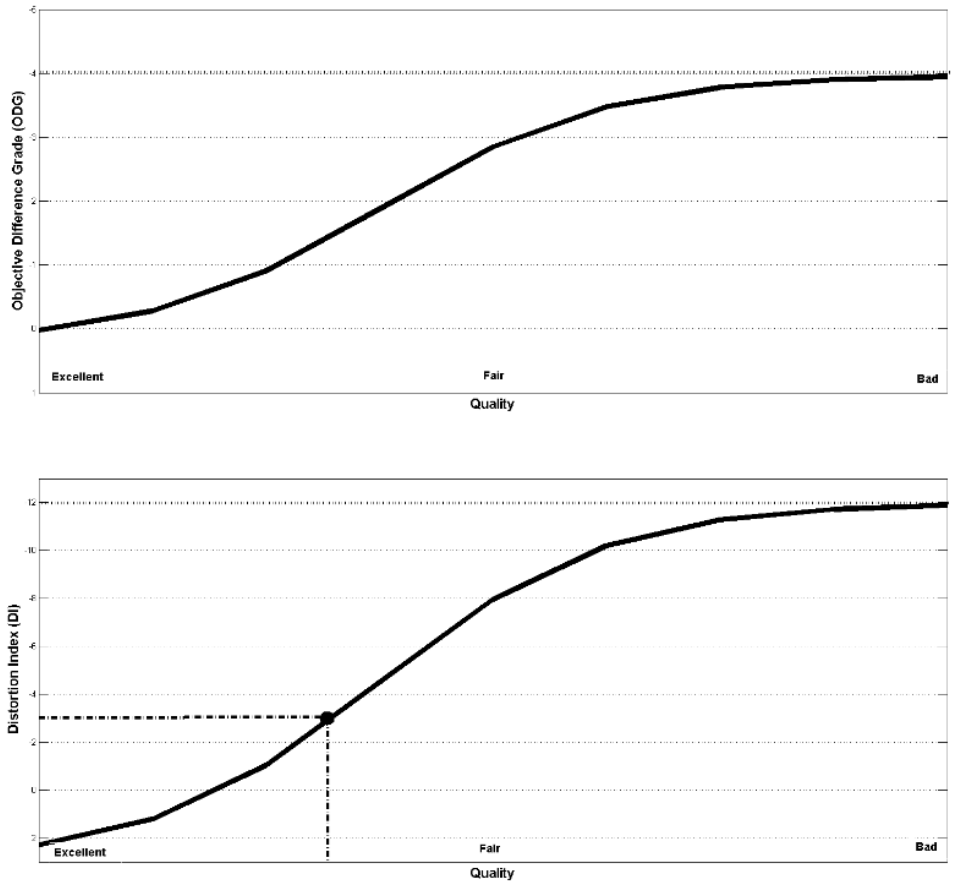

Figure 5.13: ODG (left) and DI (right).

In order to benchmark, we convoluted impulse response from the ODEON output for the L-shaped room with a dry recording. The output is normalized

to ensure that resulting wave file has all samples between [-1,1]. Same dry recording is also convoluted with the corresponding extended SDN model impulse response output and same normalization procedure is applied. The ODEON output was used as a reference file and extended SDN model output was used as the test file. The PEAQ implementation we employ were the MATLAB routines that Kabal laid out [48]. The implementation is the 'Basic' version, which is suitable for assessment of audio algorithm implementations [46]. The output variables from the code are shown in Table 5.4. These output variables are constitute two output values from the system, ODG and DI. According to the impairment scale in Table 5.3, the objective difference grade of -3.773 corresponds to a difference in perceptible audio quality which is between "Annoying" and "Very annoying".

It should be noted that during initial informal listening tests, such impairments were not reported. However, since ODG points to slightest impairments by definition, differences in reverberation tail and other components of RIR are considered as an impairment by the procedure and are evaluated as impairments. Therefore, in order to have a discrete perspective about objective evaluation of the audio quality, DI value could be utilized. The DI value of -2.958 suggests that although impairments are perceptible, the audio quality is fair when distortions are taken into account. The corresponding point on the DI graph is marked in Figure 5.13, which shows that evaluation score is on the fair side when DI is taken into account.

Table 5.4: PEAQ Model Output Variable (MOVs) values and respective descriptions.

Output Variable	Description	Value
ODG	Objective Difference Grade	**-3.773**
DI	Distortion Index	**-2.958**
BandwidthRefB	Reference Signal Bandwidth	672.36
BandwidthTestB	Test Signal Bandwidth	664.573
Total NMRB	Noise-To-Mask Ratio	-3.07722
WinModDiff1B	Windowed Modulation Difference	36.8493
ADBB	Average Block Distortion	2.03882
EHSB	Harmonic Structure of the Error	0.957994
AvgModDiff1B	Average Modulation Difference	38.9054
AvgModDiff2B	Average Modulation Difference	57.0214
RmsNoiseLoudB	Loudness of Distortion	1.83601
MFPDB	Max. Filtered Probability of Detection	1
RelDistFramesB	Relatively Distributed Frames	0.880086

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Discussion

The extended SDN is useful in acoustics simulation where the room is arbitrarily shaped. For games, virtual reality applications and simulations, extended SDN introduces little computational cost. Yet, the results are more accurate comparing to original SDN model. The algorithm is most advantageous when the computational cost is favored against computational complexity.

Table 6.1: Extended SDN run times and memory usage.

Nodes	Run time	Visibility Checks	Memory
10	0.146 s	0.001 s	26.8 MB
50	0.219 s	0.034 s	40.3 MB
100	0.545 s	0.314 s	58 MB
150	1.26 s	0.954 s	75.8 MB
200	2.261 s	1.889 s	95.1 MB
250	4.346 s	3.917 s	114.8 MB
300	6.597 s	6.09 s	136.4 MB
350	10.934 s	10.19 s	157.6 MB
400	15.585 s	14.571 s	179.1 MB
450	21.193 s	20.371 s	204.9 MB
500	29.308 s	27.943 s	216.7 MB
1000	225.792 s	221.116 s	524.8 MB

In Table 6.1, run times and memory utilizations for a single-threaded CPU implementation of extended SDN is presented. The runs were taken on an Intel i7-4500HQ computer with 16GBs of RAM. As number of nodes grow larger, visibility checks dominate the computational complexity. However, in a GPU programming environment, the naive visibility checks can be made embarrassingly parallel since solutions for each node pair does not depend on other nodes. If the geometry of the room is assumed to be fixed during run time, visibility check results could be performed offline and cached for later use which can result in dramatic increase in performance as it can be observed in Table 6.1. Still, a single-threaded CPU version can run at interactive rates

when the room size is small: 0.219 seconds of run time which corresponds to 50 nodes would have an equivalent frame rate of 45.6 frames per second. Although is tempting to create the visibility graph offline and store it for later use in a look-up table, exhaustive combinations of source and receiver positions prohibits this approach. Creating the collections of delay lines and ring buffers were included in run times and they will have fixed length throughout the run time. This would imply that an overall improvement in visibility checks would have a direct impact on running times and simulations with many nodes are possible. Capability of modeling 50 nodes at interactive rates allow modeling of even complex geometries as their simplified versions are shown to preserve their acoustical properties in room acoustics modeling by Siltanen et al [49]. The manual of the benchmark software, ODEON, suggests that most of the concert halls can be modeled with 100-2000 surfaces [50].

Table 6.2: Extended SDN run times and memory usage for x50 scaled room.

Nodes	**Run time**	**Visibility Checks**	**Memory**
10	0.407 s	0.001 s	69.4 MB
50	2.173 s	0.036 s	322.8 MB
100	4.264 s	0.247 s	637.8 MB
150	6.84 s	0.806 s	953 MB
200	10.034 s	1.896 s	1238 MB
250	13.798 s	3.579 s	1542 MB
300	18.883 s	6.26 s	1846 MB

For a room acoustics simulation model, size of the room does not solely depend on number of surfaces or in our case; number of nodes. Dimensions and volume of the room should be considered as well as these properties would have a direct impact on delay line lengths. A wider enclosure implies longer delay lines, thus higher memory allocation. Table 6.2 demonstrates run times and memory utilization when the room volumes are 125,000 times greater than the hypothetical rooms considered in Table 6.1. Even though run times for visibility checks are more or less the same as the problem size for visibility checks depend on number of nodes, an overall increase in run times can be noticed. This is due to the fact that allocation of larger delay lines. Considering that corresponding data structures for ring buffers are allocated only once per node when the model is initialized, the increase in room size has an impact on initialization time and visibility checks constitute the significant bottleneck for the interactive system.

6.2 Contributions

In this study, SDN for room acoustics simulation is presented and extensions to this technique are proposed. With these propositions, visibility and validity checks for first-order image sources are presented in order to account for arbitrarily shaped rooms. Original SDN model relies on image source method to compute node positions and work of Borish is intended to find im-

age sources in an arbitrarily shaped polyhedra [37]. The edge case when there is no visible or valid image source for a surface is handled by putting the SDN node to centroid of the corresponding surface. This allows us to account for higher-order reflections caused by that surface.

For the cases in which the direct view between source and receiver are obstructed, an approximate method for edge diffraction is suggested. Accuracy is traded off for better computational complexity. A best-fit decay parameter for the diffraction function is calculated via ternary search. Unimodal property of the function which calculates sum of squared errors between the implicit solution and our exponential decay approach is exploited to achieve reliability and better worst-case complexity in comparison to other methods such as linear search or the gradient descent method. Instead of storing all the integration values along the edge within secondary source and assigning distinct delay lines to them, edge diffraction is handled by employing a simple first-order IIR filter. Overall, at the expense of a reasonable computational cost which does not compromise interactivity of the system for a small room, capabilities of the SDN model is extended.

6.3 Future Work

The bottleneck of the current system lies on the visibility and validity checks, as each wall surface is checked against other. Suitable data structures which could allow faster visibility checks could be designed. These checks may involve space partitioning algorithms for rooms with larger geometries. Octree data structures are known to provide fast collision detection in 3-D environments and they are suitable for parallel processing. An algorithm that builds upon a linear octree, which could be stored in a list data structure, is demonstrated to solve the collision detection problem between all pairs in a virtual environment. The algorithm proposed by Lucchesi has the complexity of $O(n \log(n)/k)$, where n is the problem size and k is the independent clusters dedicated for parallel processing [51]. For complex geometries, top-down octree approaches that uses CPU efficiently exist as well [52].

The process could be parallelized in GPU given the chance that graphics and audio renderer could work on same geometry. An existing GPU-based geometrical acoustics application exploited this approach and successfully calculated image sources up to third order at interactive rates. However, it does not handle diffraction and third order image sources could simulate the impulse response up to 500 miliseconds, which could over-simplify large, reverberant enclosures [53]. Such pre-computations on room geometry is shown to allow movement of dynamic objects in complex scenes [54]. Similar computations and mesh simplification approaches are expected to overcome current shortcomings of the extended SDN model. Original SDN model restricts the geometrical model to rectangular walls and in our extended version we hold the same assumption. With the incorporation of efficient triangulation methods, wall surfaces could be given as arbitrary polygons as well. Furthermore, the room partitions could then be treated as a network of coupled rooms. Since

each coupled room is expected to have distinct acoustical properties, it is possible to solve each system separately in parallel threads and link intermediate results via linking nodes between each enclosure.

For rooms with huge volumes, it is possible that there can be delay lines that do not significantly contribute to the acoustical properties of the room. A culling algorithm could be employed to determine and remove such delay lines in order to simplify the model. Higher-order diffraction estimation is possible by connecting the diffraction network to the network of specular nodes in SDN model. Yet, it would not significantly contribute to the room impulse response as most of the first-order diffraction components are usually culled in simulations due to their minimal contribution to impulse response.

REFERENCES

[1] Jyri Huopaniemi, Lauri Savioja, and Tapio Takala. DIVA Virtual Audio Reality System. 1996.

[2] Enzo De Sena, Hüseyin Hacihabiboğlu, and Zoran Cvetković. Scattering Delay Network: An Interactive Reverberator for Computer Games. In *Audio Engineering Society Conference: 41st International Conference: Audio for Games*. Audio Engineering Society, 2011.

[3] Julius O. Smith. *Physical Audio Signal Processing*. `http://-ccrma.stanford.edu/~jos/pasp/`, accessed 04/2016. online book, 2010 edition.

[4] nreverb : Orchestra Opcodes and Operators. `http://www.csounds.com/manual/html/nreverb.html`. accessed: 04/2016.

[5] Unity Technologies. Audio SFX Reverb Effect. `http://docs.unity3d.com/Manual/class-AudioReverbEffect.html`. accessed: 04/2016.

[6] Epic Games Inc. Audio System Overview. `https://docs.unrealengine.com/latest/INT/Engine/Audio/Overview/index.html`. accessed: 04/2016.

[7] Audiokinetic Inc. Wwise Source/Effect Plug-ins. `https://www.audiokinetic.com/library/2015.1.6_5553/?source=Help&id=wwise_source_effect_plug_ins`. accessed: 04/2016.

[8] FireFlight Technologies. Firelight Technologies FMOD Studio API. `http://www.fmod.org/documentation#content/generated/FMOD_DSP_CONVOLUTION_REVERB.html`. accessed: 04/2016.

[9] Heinrich Kuttruff. *Room Acoustics*. Spon Press, London, New York, 2000.

[10] Science & Engineering The Acoustics Research Centre at University of Salford School of Computing. Concert Hall Acoustics: Art and Science. `http://www.acoustics.salford.ac.uk/acoustics_info/concert_hall_acoustics/?content=testing`. accessed: 04/2016.

[11] Jont B. Allen and David A. Berkley. Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*, 65(4):943–950, 1979.

[12] Gary Kendall, William Martens, Daniel Freed, Derek Ludwig, and Richard Karstens. Image Model Reverberation from Recirculating Delays. In *Audio Engineering Society Convention 81*. Audio Engineering Society, 1986.

[13] Asbjørn Krokstad, S. Strom, and Svein Sørsdal. Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibration*, 8(1):118–125, 1968.

[14] Thomas Funkhouser, Nicolas Tsingos, Ingrid Carlbom, Gary Elko, Mohan Sondhi, James E. West, Gopal Pingali, Patrick Min, and Addy Ngan. A beam tracing method for interactive architectural acoustics. *The Journal of the acoustical society of America*, 115(2):739–756, 2004.

[15] Thomas Funkhouser, Ingrid Carlbom, Gary Elko, Gopal Pingali, Mohan Sondhi, and Jim West. A beam tracing approach to acoustic modeling for interactive virtual environments. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 21–32. ACM, 1998.

[16] Lauri Savioja, Timo J. Rinne, and Tapio Takala. Simulation of Room Acoustics with a 3-D finite difference mesh. In *Proceedings, Volume 1994*, pages 463–466. International Computer Music Association, 1994.

[17] William G. Gardner. Reverberation algorithms. In *Applications of digital signal processing to audio and acoustics*, pages 85–131. Springer, 2002.

[18] Science & Engineering The Acoustics Research Centre at University of Salford School of Computing. Reverberation Chambers. `http://www.salford.ac.uk/acoustics-testing/labs/reverberation-chambers`. accessed: 04/2016.

[19] Julian Parker and Stefan Bilbao. Spring Reverberation: A Physical Perspective. In *Proceedings of the 12th International Conference on Digital Audio Effects (DAFx'09)*, pages 416–421, 2009.

[20] Wikimedia Commons. Folded Line reverberation device. `https://commons.wikimedia.org/wiki/File:Reverb-2.jpg`. accessed: 04/2016.

[21] Manfred R. Schroeder and Benjamin F. Logan. 'Colorless' Artificial Reverberation. *Audio, IRE Transactions on*, (6):209–214, 1961.

[22] James A. Moorer. About This Reverberation Business. *Computer Music Journal*, pages 13–28, 1979.

[23] John Stautner and Miller Puckette. Designing Multi-Channel Reverberators. *Computer Music Journal*, 6(1):52–65, 1982.

[24] Jean-Marc Jot and Antoine Chaigne. Digital Delay Networks for Designing Artificial Reverberators. In *Audio Engineering Society Convention 90*, Feb 1991.

[25] D.T. Murphy and D.M. Howard. Modelling and directionally encoding the acoustics of a room. *Electronics Letters*, pages 864–865, 1998.

[26] U. Peter Svensson, Roger I. Fred, and John Vanderkooy. An analytic secondary source model of edge diffraction impulse responses. *The Journal of the Acoustical Society of America*, 106(5):2331–2344, 1999.

[27] John Vanderkooy. A Simple Theory of Room Diffraction. In *Audio Engineering Society Convention 92*. Audio Engineering Society, 1992.

[28] Herman Medwin. Shadowing by finite noise barriers. *The Journal of the Acoustical Society of America*, 69(4):1060–1064, 1981.

[29] Maurice Anthony Biot and Ivan Tolstoy. Formulation of Wave Propagation in Infinite Media by Normal Coordinates with an Application to Diffraction. *The Journal of the Acoustical Society of America*, 29(3):381–391, 1957.

[30] U. Peter Svensson and Paul T. Calamia. Edge-Diffraction Impulse Responses Near Specular-Zone and Shadow-Zone Boundaries. *Acta Acustica united with Acustica*, 92(4):501–512, 2006.

[31] Paul T. Calamia and U. Peter Svensson. Fast Time-Domain Edge-Diffraction Calculations for Interactive Acoustic Simulations. *EURASIP Journal on Advances in Signal Processing*, 2007(1):1–10, 2006.

[32] Paul T. Calamia and U. Peter Svensson. Edge subdivision for fast diffraction calculations. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2005.*, pages 187–190. IEEE, 2005.

[33] Tapio Lokki, U. Peter Svensson, and Lauri Savioja. An Efficient Auralization of Edge Diffraction. In *Audio Engineering Society Conference: 21st International Conference: Architectural Acoustics and Sound Reinforcement*, Jun 2002.

[34] Paul T. Calamia, Benjamin E. Markham, and U. Peter Svensson. Diffraction Culling for Virtual-Acoustic Simulations. *Acta Acustica United with Acustica*, 94(6):907–920, 2008.

[35] Vesa Välimäki, Julian D Parker, Lauri Savioja, Julius O. Smith, and Jonathan S. Abel. Fifty Years of Artificial Reverberation. *Audio, Speech, and Language Processing, IEEE Transactions on*, 20(5):1421–1448, 2012.

[36] Hüseyin Hacihabiboğlu, Enzo De Sena, and Zoran Cvetković. Frequency-Domain Scattering Delay Networks for Simulating Room Acoustics in Virtual Environments. In *Signal-Image Technology and Internet-Based Systems (SITIS), 2011 Seventh International Conference on*, pages 180–187. IEEE, Nov 2011.

[37] Jeffrey Borish. Extension of the image model to arbitrary polyhedra. *The Journal of the Acoustical Society of America*, 75(6):1827–1836, 1984.

[38] Rendell R. Torres, U. Peter Svensson, and Mendel Kleiner. Computation of edge diffraction for more accurate room acoustics auralization. *The Journal of the Acoustical Society of America*, 109(2):600–610, 2001.

[39] Eric Haines. Point in Polygon Strategies. *Graphics Gems IV*, 994:24–26, 1994.

[40] E.W. Dijkstra. Constructing the Binary Search once more. EWD1293. `http://www.cs.utexas.edu/users/EWD/`, 1999.

[41] Wim H. Hesselink. Ternary search. http://wimhesselink.nl/pub/whh303.pdf, 2007.

[42] U. Peter Svensson. Edge Diffraction Toolbox. http://www.iet.ntnu.no/~svensson/software/index.html, 2015.

[43] C. C. J. M. Hak, R. H. C. Wenmaekers, and L. C. J. van Luxemburg. Measuring Room Impulse Responses: Impact of the Decay Range on Derived Room Acoustic Parameters. *Acta Acustica united with Acustica*, 98(6):907–915, 2012.

[44] Philip McCord Morse, Acoustical Society of America, and American Institute of Physics. *Vibration and sound*, volume 2. McGraw-Hill New York, 1948.

[45] Amaro A. de Lima, Fabio P. Freeland, Rafael A. de Jesus, Bruno C. Bispo, Luiz W.P. Biscainho, Sergio L. Netto, Amir Said, A. Kalker, R. Schafer, Bowon Lee, et al. On the quality assessment of sound signals. In *IEEE International Symposium on Circuits and Systems, 2008. ISCAS 2008.*, pages 416–419. IEEE, May 2008.

[46] Recommendation BS ITU-R. 1387,"Method for objective measurements of perceived audio quality". *International Telecommunication Union, Geneva, Switzerland*, Nov 2001.

[47] Klaus Schiffner. Psychoacoustic Audio Quality Measurements Using R & S ®UPV Audio Analyzer. https://cdn.rohde-schwarz.com/pws/dl_downloads/dl_application/application_notes/1ga49/1GA49_0E.pdf, Apr 2009.

[48] Peter Kabal. An Examination and Interpretation of ITU-R BS. 1387: Perceptual Evaluation of Audio Quality. *TSP Lab Technical Report, Dept. Electrical & Computer Engineering, McGill University*, pages 1–89, 2002.

[49] Samuel Siltanen, Tapio Lokki, Lauri Savioja, and Claus Lynge Christensen. Geometry Reduction in Room Acoustics Modeling. *Acta Acustica united with Acustica*, 94(3):410–418, 2008.

[50] Claus Lynge Christiensen and Georgios. Koutsouris. ODEON Room Acoustics Program. http://www.odeon.dk/pdf/ODEONManual12.pdf. accessed: 05/2016.

[51] Benjamin J. Lucchesi. A Parallel Linear Octree Collision Detection Algorithm. Master's thesis, University of Nevada Reno, 2002.

[52] Jorge Revelles, Carlos Urena, and Miguel Lastra. An Efficient Parametric Algorithm for Octree Traversal. In *Journal of WSCG*, pages 212–219. Václav Skala-UNION Agency, 2000.

[53] Zuofu Cheng. *Design of a real-time GPU accelerated acoustic simulation engine for interactive applications*. PhD thesis, Sep 2014.

[54] Nikunj Raghuvanshi, John Snyder, Ravish Mehra, Ming Lin, and Naga Govindaraju. Precomputed Wave Simulation for Real-time Sound Propagation of Dynamic Sources in Complex Scenes. *ACM Trans. Graph.*, 29(4):68, July 2010.