Acronyms

VAE Virtual Acoustic Environment 2
VSS Virtual Singing Studio 4
RIR Room Impulse Response 2
ISM Image Source Method 2

Background

The project described in this paper aims to extend the functionality of a currently implemented system called the Virtual Singing Studio situated in the listening space of the Audio Lab at the University of York. The system allows a user to hear themselves as though they are present in another acoustic environment. This project looks into extending this functionality to allow the user to move themselves around a digitally modelled version of Hendrix Hall, one of the lecture halls on the Universities campus by generation a large grid of synthetic room impulse responses, using room acoustic simulation software Odeon.

The following section covers material that forms the basis of the system produced as a result of this project.

2.1 - Virtual Acoustic Environments

Virtual acoustics has been previously described [1] as follows: *"Virtual acoustics is a general term for the modelling of acoustical phenomena and systems with the aid of a computer"*

By this definition, a Virtual Acoustic Environment (VAE) can be thought of as an environment (such as a room) for which the acoustical phenomena have been either recreated or synthesised. To produce a VAE, prior knowledge regarding the room which is to be acoustically recreated must be known; how do all audible frequencies propagate around the room for a set sound source location and receiver location?

This information can be gathered by taking a Room Impulse Response (RIR) and used to recreate the acoustics of a room for the set sound source and receiver location.

2.2 - Room Impulse Responses

In order to reproduce the acoustical phenomena of a room, an RIR must be obtained. This is done by exciting all audible frequencies within the room by using a sound source such as a loudspeaker, and recording the result using a receiver microphone.

There are a number of techniques used for exciting all audible frequencies. These include an **impulse** (such as a starter pistol) or an **exponentially swept sine** for which a sine wave is exponentially increased in frequency over a fixed period of time. Using a starter pistol means that no post processing is required as all the frequencies are excited at the same time and the impulse recorded at the receiver position can be used for convolution with an audio source. Using an exponentially swept sine requires post processing in order to time align all of the frequency dependent room reflections, thus producing an impulse response through the use of a deconvolution algorithm, however this method produces a greater signal to noise ratio thus is the desired method [2].

Though using an omni-directional microphone to record an RIR is a set standard [3], it is also possible to record RIR's using techniques such as Ambisonics to capture a three dimensional sound field.

2.3 - Synthetic Room Impulse Responses in ODEON

A method for measuring RIR's has been discussed, however there is also a way to synthesis RIR's by using room acoustic simulation software such as Odeon [4].

Odeon was designed to provide reliable predictions of room acoustics by using a hybrid of two geometric acoustic models, Ray-tracing and the Image Source Method (ISM) to synthesis RIR's. These methods model how sound would propagate around a room as though it were a straight line (ray). This inherently neglects wave phenomena such as phase and diffraction, properties that are negligible at high frequencies, however they are fundamental in describing low frequency wave behaviour [5]. Therefore geometrical methods are not accurate at modelling sound propagation for low frequency waves.

The reason for Odeon using a hybrid of the two method comes from the inherent problems encountered in each method.

2.3.1) Ray-Tracing

The ray-tracing method imitates a sound source by emitting a large number of particles in various directions from a single point [6]. These particles are then traced around the room, losing energy each time the particle encounters a surface according to the absorption coefficient assigned to that surface. The angle at which the particle is then reflected is determined by the scattering coefficient assigned to the surface of contact, ranging from a specular reflection to a completely random reflection [7]. For a specific receiver position, an volume around said point is defined in which rays are collected and used to calculate the results.

Using ray-tracing, it is a risk that some rays may not pass close enough to the receiver and therefore will not contribute to the final result.

2.3.2) Image Source Method

The ISM represents specular reflections from surfaces as its own source, mirroring that of the original sound source, creating what is known as “image sources” [6] and can be used to find all possible specular reflection paths. Figure 1 illustrates this concept. The advantage of the ISM is that each surface can be modelled by an image source which provides a great deal of data regarding the contribution to the received sound making this method more accurate than ray-tracing.

2.4 - Ambisonics

Ambisonics is a technique used to encode and decode three dimensional spatial audio information using just four audio channels. A three dimensional sound field can be recorded using a microphone known as a Soundfield microphone shown on the left in figure 2. These microphones contain four coincident capsules, one of which is an omni-directional capsule (W) and the rest of which are figure of 8 capsules used to record sound in the X (front and back), Y (left and right) and Z (up and down) direction illustrated on the right in figure 2. By combining these signals in what is known as a B-Format audio file, the sound field surrounding the microphone can be captured.

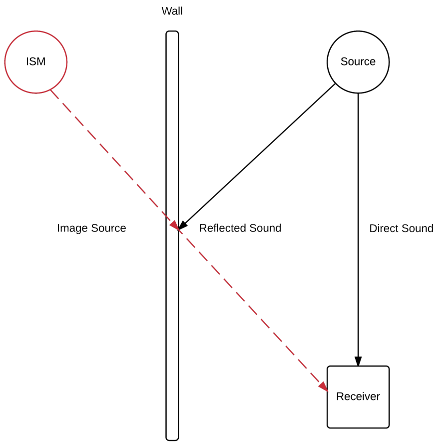

Figure 1: Illustration of the image source method (by the author)

By using a system specific Ambisonic decoder, the soundfield can be accurately reconstructed by replaying the B-Format file over a spherical loudspeaker array of arbitrary size.

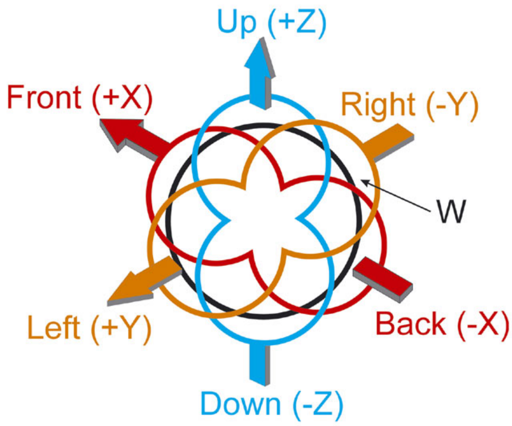

Figure 2: **Left**: Picture of a Soundfield microphone with coincident capsules exposed **Right**: Soundfield microphone polar pattern [8]

2.5 - The Virtual Singing Studio

The Virtual Singing Studio (VSS) is a loudspeaker based room acoustics simulator used as a tool for analysing the correlation between room acoustic characteristics and vocal performance parameters as part of Dr JudeUeretons PhD Thesis [9]. It is comprised of a head-mounted microphone used to capture a real-time audio input from a singer, a software patch that convolves the audio signal with a number of Ambisonic B-Format RIR's and finally a spherical array of 16 loudspeakers for which the convolved audio signal is decoded and fed to. In addition, a head-tracking device (an Oculus Rift [10]) is also used to track which direction the user is facing in the virtual space. A flow diagram of the system can is shown in figure **??**.

The VAE used in the VSS was initially the National Centre for Early Music, a space in York frequently used for musical performance. Using a Soundfield microphone, 16 Ambisonic B-Format RIR's were captured. Four positions were chosen and four RIR's facing in four directions (front, left, back, right) were recorded in each. These four directional RIR's are used to approximate the room acoustic phenomena that would occur if the user were projecting in that direction. This is done by using the data from the head-tracking device to amplitude pan the convolved signals before sending them to the spherical speaker array. Though this is a crude method for achieving such a state, it provides enough variation between the two directions to be considered convincing.

2.6 - Project Motivation

The VSS addressed a problem faced when trying to research how musicians preform in different acoustic environments: having to travel to each performance space with musicians and researchers. This also indirectly provides a solution for performers wanting to rehearse is spaces that are often inaccessible and would otherwise be expensive to book and travel to. By obtaining an RIR of the desired location, the only time travelling will be necessary is to initially obtain said RIR. However, one limitation of using the VSS is the restriction of position. If a performer wanted to try and sing at another point in the room, an RIR would have to be taken in that position too. This could be done initially, taking a range of RIR's in a number of positions, however it can not be guaranteed that all locations desirable to the performer will be available.

Therefore, this project aims to implement a system that will allow the user of the VSS to move around a VAE freely.

Previous projects have addressed mobility within VAE's before, such as [11]. The paper initially looks at two methods for providing mobility: *direct room impulse response rendering* and *parametric room impulse response rendering*. Direct room impulse response rendering consists of obtaining a set number of RIR's in a grid and interpolation between them to synthesis the users position. This method suffered from the fact that a large number of RIR's are required, thus the requirement for a large amount of storage space.

Parametric room impulse response rendering actually synthesises RIR's in real time for the given position of the user in the VAE, making them more accurate for the given location. This method avoids the need for a large amount of storage space and a system to retrive the correct files, however, as will be seen through the rest of this report, post-processing of the obtained RIR's is required, which would not be possible with real time rendering.

2.7 - Similar Projects

2.7.0.1 Geometrical Methods

References

[1] J. Huopaniemi, L. Savioja, T. Lokki, and R. Väänänen, "Virtual acoustics - applications and technology trends," *Proc. European Signal Proc. Conf. (Eusipco'2000)*, pp. 2201–2208, 2000.

[2] G. Stan, J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," *Journal of the Audio Engineering Society*, vol. 50, no. 4, pp. 249–262, 2002. [Online]. Available: http://orbi.ulg.ac.be/handle/2268/34825

[3] Iso 3382-1, "Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces," vol. 3, 2009.

[4] (2016) Software. Odeon. [Online]. Available: http://www.odeon.dk/development-room-acoustics-software

[5] S. Siltanen, T. Lokki, and L. Savioja, "Rays or Waves? Understanding the Strengths and Weaknesses of Computational Room Acoustics Modeling Techniques," *Proceedings of the International Symposium on Room Acoustics, ISRA 2010*, no. August, pp. 1–6, 2010. [Online]. Available: http://www.acoustics.asn.au/conference{_}proceedings/ICA2010/cdrom-ISRA2010/Papers/O5a.pdf

[6] J. H. Rindel, "Computer Simulation Techniques for Acoustical Design of Rooms," *Acoustics Australia*, vol. 23, pp. 81–86, 1995. [Online]. Available: http://www.odeon.dk/pdf/AustralAc_1995_Rindel.pdf

[7] C. L. Christensen and G. Koutsouris, "ODEON User's Manual," 2015.

[8] P. White. (2016) Recording a live choral performance. Website. sos. [Online]. Available: http://www.soundonsound.com/sos/jun04/articles/liveconcert.htm

[9] J. S. Brereton and B. A. Hons, "Singing in Space (s): Singing performance in real and virtual acoustic environments — Singers ' evaluation , performance analysis and listeners ' perception ." no. August, 2014.

[10] (2016) Website. Oculus. [Online]. Available: https://www.oculus.com/en-us/

[11] L. Savioja, J. Huopaniemi, T. Lokki, and R. Vaananen, “Creating Interactive Virtual Acoustic Environments,” 1999.