



University of HUDDERSFIELD

University of Huddersfield Repository

Johnson, Dale and Lee, Hyunkook

Taking advantage of geometric acoustics modeling using metadata

Original Citation

Johnson, Dale and Lee, Hyunkook (2016) Taking advantage of geometric acoustics modeling using metadata. In: Interactive Audio Systems Symposium 2016, 23rd September 2016, University of York.

This version is available at <http://eprints.hud.ac.uk/id/eprint/29718/>

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

<http://eprints.hud.ac.uk/>

Taking advantage of geometric acoustics modeling using metadata

DALE JOHNSON AND HYUNKOOK LEE

Applied Psychoacoustics Laboratory, The University of Huddersfield, Huddersfield, UK
e-mail: Dale.Johnson@hud.ac.uk, H.Lee@hud.ac.uk

September 23rd 2016

Abstract

Virtual reality has recently experienced resurgence in popularity, and thanks to the advances in processor technology over the last decade it has become feasible to use geometric acoustics as a method of simulating the acoustics of a virtual environment using the same scene model data. Usually, the highest priority is realism, however this is not always favourable, especially for spatial audio. The acoustics of a virtual environment may not be pleasing. This paper discusses the development of a custom geometric acoustics program, a data structure (Raw Impulse Vector) produced by the program that provides open access to all of the captured rays including metadata about each ray, and finally the types of processing that can be performed on this data in order to perceptually optimise the acoustics. This can be of great benefit to interactive audio systems that employ a geometric acoustics algorithm as a method of artificially simulating reverb where the characteristics of the reverb can be optimised based on the listener's preferences.

1 Introduction

In recent years, virtual reality has regained momentum as a new form of entertainment. This is thanks to the advances in computer technology where graphics processing units (GPUs), are now able to render highly detailed, virtual environments. Also, it is now possible to perform high performance, general purpose computing on a GPU, opening the gateway to on GPU virtual acoustics modelling. This is often coupled with virtual reality (VR) for games, movies, or other virtual environments in order to immerse the listener and enhance the experience. The model data that is used to generate the virtual scene can also be used by geometric acoustics to simulate the acoustics of the scene.

VR focuses on creating a convincing experience, however a convincing experience may not always be a pleasant one. Accurate rendering may result in unwanted artefacts such as comb filtering and distracting flutter echoes, or poor intelligibility and poor source localisability. In real life, these effects are dependent on the listener's surroundings, however in VR, the effects could be altered in order to improve the listening experience by the perceptual optimisation of the acoustic parameters. For example, what if the apparent source width could be widened or narrowed depending on the listener's preference, whilst at the same time retain a pleasing sense of localisability?

The merit of geometric acoustics is that data generated during the rendering stages can be taken advantage of, allowing for such perceptual optimisation to be performed. This means producing and manipulating data that is much more advanced than a discrete impulse response. This paper provides a brief background to geometric acoustics and how they work, the development of a custom algorithm, a data structure utilised by the said algorithm, and finally how the data can be used to perform perceptual optimisation.

2 Geometric acoustics

Geometric acoustics is a term used to describe algorithms that artificially model room acoustics using geometry. The most popular methods are ray tracing, a method pioneered by [1][2] and [3], the image source method, or the ISM, developed by [4] and later improved by [5] and [6], and finally beam tracing [7]. The ISM in particular has proven to be a rather popular method of real-time auralisation, for example [8] combined the ISM with a feedback-delay network reverberator to create an interactive, binaural simulation of a rectangular room.

Briefly, ray tracing simulates acoustics by quantising a sound wave into a finite number of rays. These rays are uniformly emitted from a point source in all directions, traced and reflected around the room model, losing energy each time they reflected off a

surface. The tracing process is terminated once the energy of all of the rays is negligible. Capturing rays that cross over a spherical receiver forms the RIR. The receiver can be thought of as a virtual microphone. The ISM on the other hand operates by mirroring the source through each surface to create virtual images, shown in Figure 1. These images are recursively mirrored to produce higher order images. Back-tracing a reflection path from the receiver to the source via each image forms the RIR. The path is valid if it is reflected by the surfaces that create the images and is not occluded by any surface not belonging to the reflection sequence.

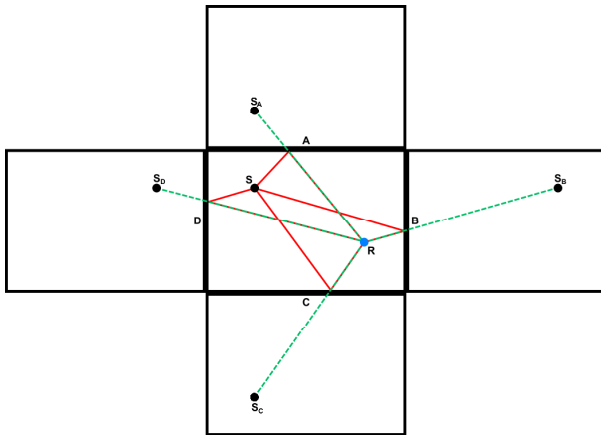


Figure 1 – First order ISM.

2.1 Custom Algorithm

For the purposes of this study, a custom, 3D geometric acoustics program was developed in C++. The algorithm generates a room impulse response, or RIR, using a hybrid ISM and ray tracing method. The ISM is used to render the early reflections whilst ray tracing renders the remaining late reflections. The reason for this hybridisation is to counteract the limitations and exploit the advantages each method has to offer. The ISM is very efficient for rendering low order early reflections accurately yet becomes computationally intensive for high order reflections. The ISM is also unable to model diffusion. On the other hand, ray tracing is far less computationally expensive for rendering late reflections and is capable of modeling diffusion, yet is not a highly efficient or accurate

algorithm for early reflection modelling.

The program is capable of modeling arbitrarily shaped rooms and simulates both surface material absorption and air absorption. Using multiple receivers multichannel RIRs can be rendered. The receivers themselves can be freely positioned and are able to emulate microphone polar patterns allowing for 3D microphone arrays to be simulated. A binauralisation function is currently being developed.

2.2 Raw Impulse Vector

The process the program takes to render a RIR is shown in Figure 2. A key part of the process is the “Raw Impulse Vector” (RIV). The RIV is a data structure that contains all of the rays captured by a receiver, along with metadata about each ray. This metadata includes:

- Delay time
- Direction of arrival
- Octave-band energy
- Reflection order
- Reflection history

The RIV gives access to individual rays, which of course allows for precise manipulation of the rays to be performed, or processes not normally possible with a RIRs.

3 Perceptual optimisation

Caching of metadata opens up geometric acoustics to applications beyond artificial RIR generation. As mentioned earlier, the application most relevant to interactive audio is perceptual optimisation. Perceptual optimisation can be defined as controlling the characteristics of information received by the senses in a way that is optimised for the mechanics of those senses. In the case of spatial audio and room acoustics, this means controlling the perceived spatial impression of a room, whether to simply alter an effect or remove an unwanted side effect.

Spatial impression can be described as two sub-paradigms: apparent source width (ASW) and listener envelopment (LEV). Briefly, ASW is an apparent broadening of the sound source due to early lateral reflections that arrive at the listener between 0

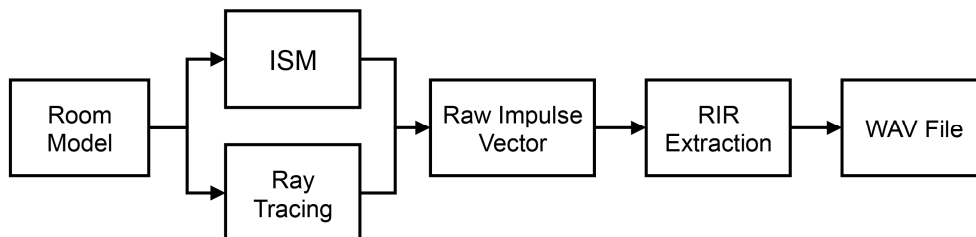


Figure 2 – Rendering process of the custom algorithm

and 80ms from onset. According to [9][10][11] and [12], ASW depends on the temporal and angular distribution, the lateral fraction, and inter-aural cross-correlation (IACC) of early reflections between 0 and 80ms after onset. LEV can be described as the sense of being enveloped by the reverb. [13][14] and [15] have found this is caused by late reflections arriving from all around the listener from 80ms to infinity.

ASW in particular another perceptual attribute, localisability. Localisability is the accuracy the human hearing is able to locate the position of an auditory source. ASW tends to operate against localisability [16] since ASW is an increase in source width, and so it becomes increasingly difficult to accurately localise a sound source. This effect depends on the nature of reflections and so is dependent on the room geometry. Furthermore, spatial impression itself is influenced by the nature of the reflections, therefore to perceptually control the spatial impression, the parameters of the reflections that contribute to each paradigm need to be altered.

3.1 Methods

From an architectural point of view, one method would be to alter the geometry of the room model, however this would not be an ideal solution. First, the way the geometry would be altered is ambiguous. Whilst surfaces that affect rays that contribute to certain effects could be identified using the reflection history of each ray, the way that the surfaces need to be altered is unclear. This leads onto the second problem; this would affect the entire reverb and thus the spatial impression. Lastly, virtual models are usually designed to appear a certain way to the user, the acoustics of the models are not taken into consideration, and therefore it would not be possible to modify the geometry in order to alter spatial impression. A more favourable solution would be to directly manipulate the rays that contribute to each effect. The RIV provides precise, almost ‘granular’ control over each ray. Rays could be identified and grouped according to what effect they contribute to based on their metadata. This creates an interesting concept where perceptual maps can be applied to the RIV depending on the desired optimisation.

For example, using map that describes how ASW operates, rays that contribute to ASW and affect localisability can be identified based on their arrival time, direction and octave-band energy. Then by using a reverse engineered mode, the rays can be manipulated to in order to improve localisability whilst retaining ASW. One method could involve attenuating or removing rays that arrive between 0 to 20ms and 60 to 90 degrees from the front of the receiver. This type of

processing would be very difficult to perform on a RIR as it requires analysis that is much precise and advanced.

In summary, the goal is to create a perceptual control stage in the existing algorithm that can be used to perform perceptual optimisation in order to either remove and unwanted effect, or emphasise an effect independent from others, whilst retaining a plausible sense of spatial impression.

4 Conclusion

This paper has shown that geometrical acoustics can be used to create an advanced data structure that is more than just an artificial room impulse response (RIR). Whilst the applications are hypothetical and are the main goals of the first author’s work, they hope to highlight that providing open access to data produced by the algorithm can lead to applications not possible with a traditional RIR.

The main application of the Raw Impulse Vector (RIV) described here is perceptual optimisation. Individual rays and their associated metadata can be manipulated prior to the RIR conversion, something that cannot be easily achieved in the real world or with traditional yet artificially generated RIRs. Different optimisations can be applied to the RIV depending on the desired outcome e.g. if locatedness is to be improved, reflections that disturb locatedness could be removed whilst retaining reflections that contribute to the overall spatial impression.

To achieve perceptual optimisation, a perceptual model must first be created. This will involve testing the mechanics of each paradigm. Focusing on ASW and locatedness improvement, an evaluation the current model described by [9][10] and [11] must be performed. The model will then be improved and mapped onto a sphere, which in turn will be used to analyse rays captured by a receiver. The analysis will identify the most significant rays and perform necessary locatedness improvements whilst preserving an adequate sense of spatial impression.

Such an algorithm would be useful for interactive audio applications such as: virtual concert hall simulations; gaming where the locatedness of certain auditory events such as speech, vehicle noises or gunshots could be improved dynamically whilst retaining the spatial characteristics of the scene.

References

- [1] A. Krokstad, S. Strøm & S. Sørsdal, “Calculating The Acoustical Room Response By The Use Of A Ray Tracing Technique”, *Journal of Sound and Vibration* vol 8, no. 1, pp. 118 –125 (1968).

- [2] M. R. Schroeder, "Digital Simulation of Sound Transmission In Reverberant Spaces", *Journal of the Acoustical Society of America* vol 47, no. 2, pp. 424 – 431 (1970).
- [3] J. Borish, "An Auditorium Simulator for Domestic Use", *Journal of the Audio Engineering Society* vol 33, no. 5, pp. 330 – 341 (1985).
- [4] J. B. Allen & D. A. Berkley, "Image method for efficiently simulating small-room acoustics", *Journal of the Acoustical Society of America* vol 65, no. 4, pp. 943 – 950 (1979).
- [5] J. Borish, "Extension of the image model to arbitrary polyhedral", *Journal of the Acoustical Society of America* vol 75, no. 6, pp. 1827 – 1836 (1984).
- [6] M. Vorländer, "Simulation of the transient and steady-state sound propagation in rooms using a new combined ray-tracing/image-source algorithm", *Journal of the Acoustical Society of America* vol 86, no. 1, pp. 172 – 178 (1989).
- [7] T. Funkhouser, I. Carlborn, G. Elko, P. Gopal, M. Sondhi & J. West, "A Beam Tracing Approach to Acoustic Modeling for Interactive Virtual Environments", *Proceedings of the 25th annual conference on computer graphics and interactive techniques*, ACM, pp. 21-32 (1998).
- [8] T. Wendt, S. Van De Par & S. D. Ewert, "A Computationally-Efficient and perceptually-Plausible Algorithm for Binaural Room Impulse Response Simulation", *Journal of the Audio Engineering Society* vol 62, no. 11, pp. 748 – 766 (2014).
- [9] M. Barron & A. H. Marshall, "Spatial impression due to early lateral reflections in concert halls: the derivation of a physical measure", *Journal of Sound and Vibration* vol 77, no. 2, pp. 211 – 232 (1981).
- [10] J. Blauert & W. Lindemann "Auditory spaciousness: Some further psychoacoustic analyses", *Journal of the Acoustical Society of America* vol 80, no. 2, pp. 533 – 542 (1986).
- [11] M. Morimoto & Z. Maekawa, "Effects of low frequency components on auditory spaciousness", *Acta Acustica* vol 66, no. 4, pp. 190 – 196 (1988).
- [12] T. Hidaka, L. Beranek & T. Okano, "Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls", *Journal of the Acoustical Society of America* vol 98, no. 2, pp. 988 – 1007 (1995).
- [13] J. S. Bradley & G. A. Soulodre, "Objective measures of listener envelopment", *Journal of the Acoustical Society of America* vol 98, no. 5, pp. 2590 – 2597 (1995).
- [14] M. Morimoto, K. Iida & K. Sakagami, "The role of reflections from behind the listener in spatial impression", *Applied Acoustics* vol 62, no. 2, pp. 109 – 124 (2001).
- [15] G. A. Soulodre, M. C. Lavoie & S. G. Norcross, "Objective Measures of Listener Envelopment in Multichannel Surround Systems", *Journal of the Audio Engineering Society* vol 51, no. 9, pp. 826 – 840 (2003).
- [16] W. M. Hartmann, "Localization Of Sound In Rooms", *Journal of the Acoustical Society of America* vol 74, no. 5, pp. 1380 – 1391 (1983).