Procedural 3D Audio for AR Applications

Angeliki Skandalou
John Jeremy Ireland

Supervisor:
Michael Rose

Kongens Lyngby 2017

Abstract

This is the first paragraph

THis is the second

THird paragraph of abstract

Four paragraphs is enough I guess

Acknowledgements

We would like to express our gratitude and appreciation to our supervisor for his support and guidance throughout this thesis work. Several discussion sessions and advice helped us take the most out of this project and make this study possible.

We would like to express special thanks also to the rest of our classmates who did their thesis at the same time under the same supervisor and offered us their advice. And last but not least to our family and friends whose support throughout this thesis was invaluable.

Contents

Abbreviations

ASW Apparent Source Width.

AVIL Audio Visual Immersion Lab.

BRIR Binaural Room Impulse Response.

CS Compressive Sensing.

DOA Direction of Arrival.

ERB Equivalent Rectangular Band.

HATS Head And Torso Simulator.

HOA Higher Order Ambisonics.

HRTF Head-Related Transfer Function.

IACC Inter-Aural Cross Coherence.

ILD Inter-Aural Level Difference.

ITD Inter-Aural Time Difference.

STFT Short-Time Fourier Transform.

WFS Wave Field Synthesis.

Nomenclature

$\mathbf{\Omega}_{LS}$ Vector containing directions of Loudspeakers in reproduction.

$\mathbf{\Omega}_L$ Grid of directions used for the CS algorithm.

$\mathbf{\Omega}_s$ Subvector of $\mathbf{\Omega}_L$ containing only the prominent directions after CS processing.

$\breve{\mathbf{H}}$ Combined transfer matrix for mixed-norm problem.

$\breve{\mathbf{p}}$ Combined measurement pressure vector for mixed-norm problem.

$\mathbf{x}$ Combined amplitude.

ℓ_p Norm-p.

$\mathbf{H}$ Transfer Matrix for plane waves impinging on rigid sphere.

$\mathbf{p}$ Measurement vector for the pressure on the spherical array.

$\mathbf{x}$ Amplitude vector for plane waves impinging on the sphere.

$\widetilde{\mathbf{p}}$ Pressure vector reconstructed from prominent plane waves.

B_n^m Ambisonics coefficients.

L Number of plane waves in a discrete grid of directions.

LS Number of Loudspeakers in reproduction.

N Truncation order for the spherical Harmonic Functions.

P_n^m The associated Legendre polynomials of the first kind.

Q Number of sampling points on the spherical microphone array.

R_0 Radius of reproduction area.

Y_n^m Spherical harmonic Functions.

Ω Angular Dependency on both azimuth and inclination angle.

λ Regularization factor for natural field HOA processing.

$\mathbf{B}_N$ Ambisonics coefficients vector truncated at order N.

$\mathbf{S}$ Loudspeaker signals resulting from HOA decoding.

$\mathbf{W}$ Vector containing radial functions W_n.

$\mathbf{Y}_N(\mathbf{\Omega}_L)$ Spherical harmonics vector truncated at order N for all measurement angles in vector $\mathbf{\Omega}_L$.

$\mathbf{p}'$ Residual pressure.

ε Noise parameter for Compressive Sensing Algorithm.

a Radius of microphone array.

"w/ Residual" Exploiting the residual pressure (full implementation of signal path in Figure **??**).

"w/o Residual" Residual pressure is neglected (only upper path in Figure **??**).

Chapter 1

Introduction

Immersion and all these stuff that makes our thing good. Why we are doing it and what do we want to give to the community?

Why is our method better that others? (eg wavetable)? And why we think this is the future of the audio in video games?

Chapter 2

Theoretical Background

This is a way to link to explanations Direction of Arrival (DOA)

THis is a todo:

THis is smth done:

2.1 State-Of-The-Art

2.2 Modal Analysis

$$\frac{1}{r^2}\frac{\partial}{\partial r}(r^2\frac{\partial p}{\partial r}) + \frac{1}{r^2 \text{sin}\theta}\frac{\partial}{\partial \theta}(\text{sin}\theta\frac{\partial p}{\partial \theta}) + \frac{1}{r^2\text{sin}^2\theta}\frac{\partial^2 p}{\partial \phi^2} - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0. \tag{2.1}$$

[43].

2.2.1 Features Extraction

2.3 Modal Synthesis

2.3.1 Sinusoidal Additive Synthesis

2.3.2 Filter-based Modal Synthesis

Chapter 3

Method

A combination of the methods described in Chapter 2 is proposed in the present study.

3.1 Chuck language

Modal features extraction code

3.2 PureData

Resynthesis patches

3.3 Heavy Compiler

3.4 Unity

3.5 Overview

CHAPTER 4

Measurements

Here we can describe the audio recordings and put pictures

Chapter 5

Implementation

Here we can put pictures and codes snippets

5.1 Impact Sounds

5.1.1 Sinusoidal Additive Synthesis

5.1.2 Filter-based Modal Synthesis

5.2 Rolling Sounds

5.3 Scratching Sounds

5.4 User Interface

Chapter 6

Results & Discussion

6.1 Which Synthesis Method Is Better?

6.2 Did we manage to achieve what we wanted?

6.3 How can we improve our work?

Chapter 7

Conclusion

This is the conclusion
4-5 paragraph approx

Appendix A

Results of tests to users

Appendix B

User Guide to our product

Bibliography

[1] Marshall Day Acoustics. Iris. `http://www.iris.co.nz/`.

[2] Jens Ahrens and Sascha Spors. An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions. *Acta Acustica united with Acustica*, 94(6):988–999, 2008.

[3] J.B. Allen and D.A. Berkley. Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65(4):943, 1979.

[4] Amir Avni and Boaz Rafaely. Interaural cross correlation and spatial correlation in a sound field represented by spherical harmonics. In *Conference paper presented at the Ambisonics Symposium Graz*, 2009.

[5] Amir Avni and Boaz Rafaely. Sound localization in a sound field represented by spherical harmonics. In *International Symposium on Ambisonics and Spherical Acoustics*, 2010.

[6] Augustinus J Berkhout. A holographic approach to acoustic control. *Journal of the audio engineering society*, 36(12):977–995, 1988.

[7] Augustinus J Berkhout, Diemer de Vries, and Peter Vogel. Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America*, 93(5):2764–2778, 1993.

[8] Stéphanie Bertet, Jérôme Daniel, and Sébastien Moreau. 3d sound field recording with higher order ambisonics - objective measurements and validation of spherical microphone. In *Audio Engineering Society Convention 120*, May 2006.

[9] Christopher M Bishop. Pattern recognition. *Machine Learning*, 128, 2006.

[10] Jens Blauert. *Communication acoustics*, volume 2. Springer, 2005.

[11] Emmanuel J. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346:589–592, 2008.

[12] Jérôme Daniel. Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format. In *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*. Audio Engineering Society, 2003.

[13] Jerome Daniel, Sebastien Moreau, and Rozenn Nicol. Further investigations of high-order ambisonics and wavefield synthesis for holophonic sound imaging. In *Audio Engineering Society Convention 114*, Mar 2003.

[14] Jérome Daniel, Jean-Bernard Rault, and Jean-Dominique Polack. Ambisonics encoding of other audio formats for multiple listening conditions. In *Audio Engineering Society Convention 105*, Sep 1998.

[15] Larry S Davis, Ramani Duraiswami, Elena Grassi, Nail A Gumerov, Zhiyun Li, and Dmitry N Zotkin. High order spatial audio capture and its binaural head-tracked playback over headphones with hrtf cues. In *Audio Engineering Society Convention 119*. Audio Engineering Society, 2005.

[16] P Dietrich, M Guski, M Pollow, B Masiero, M Müller-Trapet, R Scharrer, and M Vorländer. Ita-toolbox–an open source matlab toolbox for acousticians. *Fortschritte der Akustik–DAGA*, pages 151–152, 2012.

[17] Efren Fernandez Grande and Angeliki Xenaki. *Sparse acoustic imaging with a spherical array*. 2015.

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

[19] Michael A Gerzon. General metatheory of auditory localisation. In *Audio Engineering Society Convention 92*. Audio Engineering Society, 1992.

[20] Bradford N. Gover, James G. Ryan, and Michael R. Stinson. Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array. *Journal of the Acoustical Society of America*, 116(4):2138–2148, 2004.

[21] Michael Grant and Stephen Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. `http://stanford.edu/~boyd/graph_dcp.html`.

[22] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, March 2014.

[23] Per Christian Hansen. Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems, March 2008.

[24] Per Christian Hansen and IMM. Regularization tools, version 4.1. `http://www.netlib.org/numeralgo/na4-matlab7.tgz`, 2007.

[25] Finn Jacobsen and Peter Moller Juhl. *Fundamentals of General Linear Acoustics*. John Wiley & Sons, 2013.

[26] D. P. Jarrett, E. A. P. Habets, M. R. P. Thomas, and P. A. Naylor. Simulating room impulse responses for spherical microphone arrays. *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2011.

[27] D. P. Jarrett, E. A. P. Habets, M. R. P. Thomas, and P. A. Naylor. Rigid sphere room impulse response simulation: algorithm and applications. *Journal of the Acoustical Society of America*, 132(3):1462–1472, 2012.

[28] D. Malioutov, M. Cetin, and A. S. Willsky. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Transactions on Signal Processing*, 53(8):3010–3022, Aug 2005.

[29] Marton Marschall, Torsten Dau, Ewen MacDonald, and Jorg Buchholz. *Capturing and reproducing realistic acoustic scenes for hearing research*. PhD thesis, 2014.

[30] Márton Marschall, Sylvain Favrot, and Jörg Buchholz. Robustness of a mixed-order ambisonics microphone array for sound field reproduction. In *Audio Engineering Society Convention 132*, Apr 2012.

[31] Toshiyuki Okano, Leo L Beranek, and Takayuki Hidaka. Relations among interaural cross-correlation coefficient (iacce), lateral fraction (lfe), and apparent source width (asw) in concert halls. *Journal of the Acoustical Society of America*, 104(1):255–265, 1998.

[32] J. Pätynen, S. Tervo, and T. Lokki. Analysis of concert hall acoustics via visualizations of time-frequency and spatiotemporal responses. *J. Acoustical Society of America*, 133(2):842–857, Feb. 2013.

[33] Boaz Rafaely. *Fundamentals of Spherical Array Processing.* Springer, 2015.

[34] Wookeun Song, Wolfgang Ellermeier, and Joergen Hald. Psychoacoustic evaluation of multichannel reproduced sounds using binaural synthesis and spherical beamforming. *Acoustical Society of America. Journal*, 130, 2011.

[35] Wookeun Song, Wolfgang Ellermeier, and Jørgen Hald. Using beamforming and binaural synthesis for the psychoacoustical evaluation of target sources in noise. *The Journal of the Acoustical Society of America*, 123(2):910–924, 2008.

[36] Sascha Spors and Jens Ahrens. A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling. In *Audio Engineering Society Convention 125.* Audio Engineering Society, 2008.

[37] Sascha Spors, Rudolf Rabenstein, and Jens Ahrens. The theory of wave field synthesis revisited. In *124th AES Convention*, pages 17–20, 2008.

[38] Sascha Spors, Hagen Wierstorf, Alexander Raake, Frank Melchior, Matthias Frank, and Franz Zotter. Spatial sound with loudspeakers and its perception: A review of the current state. *Proceedings of the IEEE*, 101(9):1920–1938, 2013.

[39] Two!Ears Team. Two!ears auditory model 1.2. Doi: `10.5281/zenodo.47487`, 2015.

[40] S. Tervo, J. Pätynen, and T. Lokki. Acoustic reflection localization from room impulse responses. *Acta Acustica united with Acustica*, 98:418–440, 2012.

[41] Juha Vilkamo, Tapio Lokki, and Ville Pulkki. Directional audio coding: Virtual microphone-based synthesis and subjective evaluation. *J. Audio Eng. Soc*, 57(9):709–724, 2009.

[42] Darren B Ward and Thushara D Abhayapala. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Transactions on speech and audio processing*, 9(6):697–707, 2001.

[43] Earl G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography.* Academic Press, 1999.

[44] Yan Jennifer Wu and Thushara D Abhayapala. Theory and design of soundfield reproduction using continuous loudspeaker concept. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(1):107–116, 2009.

[45] Angeliki Xenaki and Klaus Mosegaard. Compressive beamforming. *Acoustical Society of America. Journal*, 136(1):260–271, 2014.

[46] Dmitry N Zotkin, Ramani Duraiswami, and Larry S Davis. Rendering localized spatial audio in a virtual auditory space. *IEEE Transactions on Multimedia*, 6(4):553–564, 2004.

[47] Franz Zotter and Matthias Frank. All-round ambisonic panning and decoding. *Journal of the Audio Engineering Society*, 60(10):807–820, 2012.