Spherical harmonic domain parametric methods for 3D audio production enhancement(check capitalization)

subtitle Required?

Author: Andrés Pérez López

TESI DOCTORAL UPF / year 2020

THESIS SUPERVISORS

Dr. Emilia Gómez GutiérrezDr. Adan Garriga Torres

Department of Information and Communication Technologies


upf.
Universitat
Pompeu Fabra
Barcelona

Write here your dedication dedication

Thanks thanks to.... thanks

Abstract

This is the abstract of the thesis in English. Please, use less than 150 words. abstract in english

Resum

Vet aquí el resum de la tesi en català. Si us plau, utilitzeu menys de 150 paraules. abstract in catalan

Preface **is that really needed?**

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Motivation

1.1.1 3D arrays

Sound propagates in 3D: need for 3D mic arrays to capture spatial properties

1.1.2 spherical microphone arrays

- even distribution of capsules
- mathematical convenience: spherical harmonics

1.1.3 ambisonics

advantages on the vr/ar context

- device independent
- intermediate storage format
- signal-independent transformations are easy
- de-facto standard for vr

1.1.4 Current limitations of vr/ar production

1.2 Goals

Research question: How can we exploit the characteristics of ambisonic recordings in order to manipulate them more adequately?

1.3 Context

Different levels of application/contribution:

- Acoustic Parameter Estimation (low level, audio2data)
  - Direction of Arrival estimation
  - Coherence analysis
  - Acoustic description (RT60, etc)
  - Source counting
- Signal Enhancement (high level, audio2audio)
  - Source Separation
  - Dereverberation / denoising
  - IR estimation
- Scene Description (high level, audio2data)
  - Event Detection
  - Acoustic Scene Classification

introduce rest of chapters?

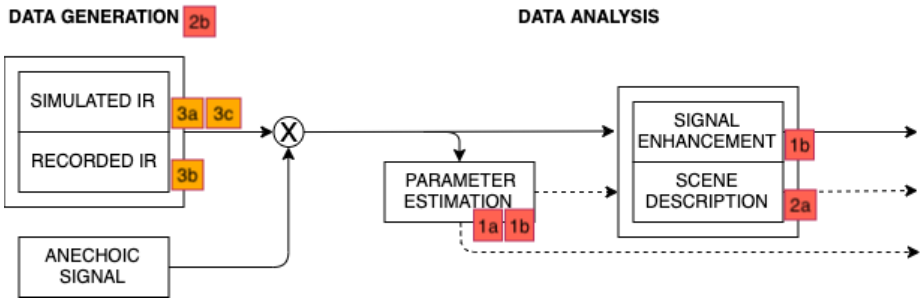


Figure 1.1: todo caption

Chapter 2

SCIENTIFIC BACKGROUND

2.1 State of the Art

Chapter 3

COHERENCE ESTIMATION

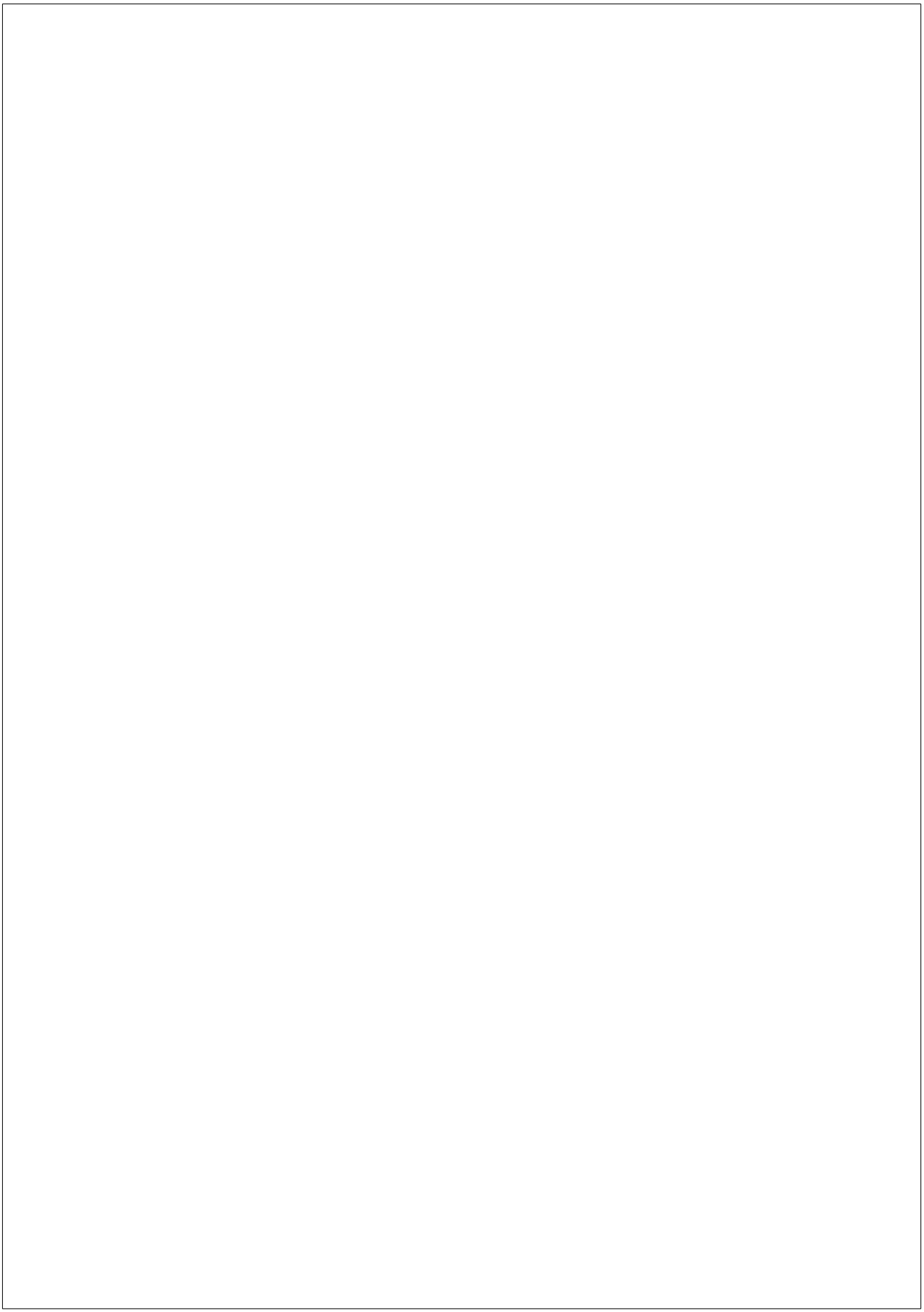
explain paper

Chapter 4

AUTOREGRESSIVE IMPULSE RESPONSE MODELS

Chapter 5

SOUND EVENT LOCALIZATION AND DETECTION



Chapter 6

DATA GENERATION AND STORAGE

Explain about mono files plus ambisonics IRs.

6.1 Recorded IRs

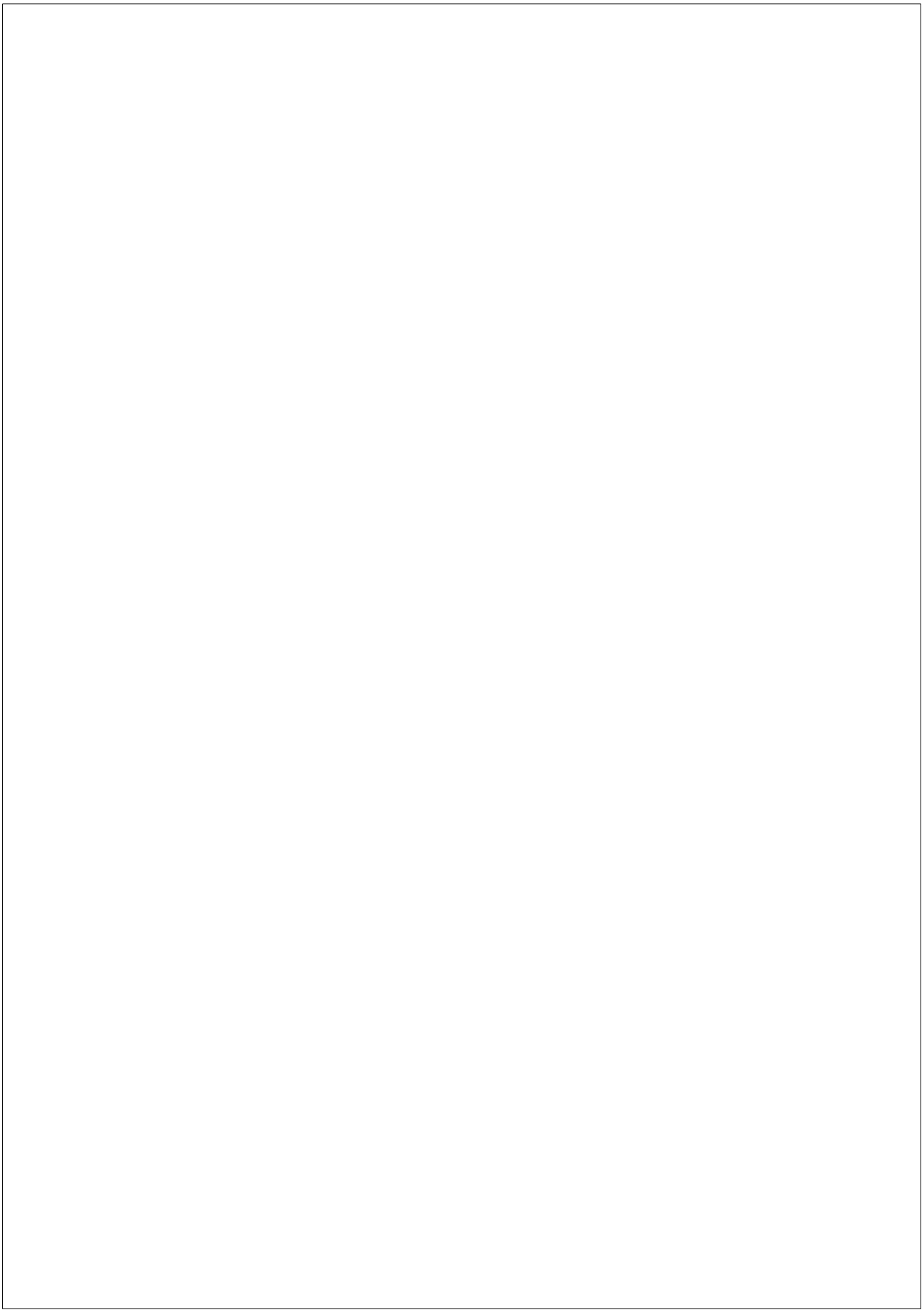
AmbisonicsDRIR: explain problematics, standard, etc. also: pysofaconventions

6.2 Simulated IRs

explain different methods and libraries. tell about masp.

6.3 High-level scene description

ambiscaper



Chapter 7

CONCLUSIONS

7.1 Summary of Contributions

- Academic Contributions

  1. Analysis of spherical isotropic noise fields with an A-Format tetrahedral microphone ref
     - **Parameter estimation**: Contribution to the characterization of coherence with tetrahedral microphones (the most common spherical arrangement)
  2. Autoregressive B-Format Late IR Estimation ref
     - **Parameter estimation**: Novel technique for RT60 estimation from autoregressive models (subproduct of dereverberation)
     - **Signal enhancement**: Novel methodology to re-reverberate sound scenes (include new elements in the scene using the reverb of the recorded scene)
  3. A hybrid parametric-deep learning approach for sound event localization and detection ref
     - **Scene Description**: ovel State-of-the-Art methodology for Sound Event Localization and Detection

- Software Contributions

    1. Ambiscaper: A Tool for Automatic Generation and Annotation of Reverberant Ambisonics Sound Scenes ref
        - **Data Generation**: Novel tool for reverberant ambisonic dataset generation
    2. Ambisonics Directional Room Impulse Response as a New Convention of the Spatially Oriented Format for Acoustics ref
        - **Recorded IRs**: File standard/convention proposal for storage of recorded ambisonic IRs
    3. Multichannel Array Signal Processing library ref
        - **Simulated IRs**: Library for acoustic simulation (IR generation, microphone array simulation, etc)
    4. pysofaconventions ref
        - **Recorded IRs**: implementation of SOFA for python

7.2 List of related publications

Bibliography