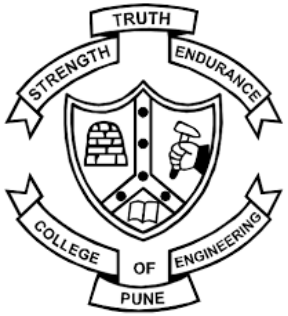



College of Engineering, Pune

Department of Computer Engineering and Information Technology

# Binaural Rendering of Ambisonics B-format

Btech Project Presentation by Anupam Godse (111408016)

Project Guide:- Dr. V.K. Pachghare

Project in collaboration with NVIDIA

# Introduction

- About Ambisonics
  - Stores directional information
  - No predefined speaker layout required
  - Usage for real time applications like gaming
- Ambisonics B-format
  - Widely used to store first order ambisonics information
  - 4 channels
    - W : Omni directional sound pressure.
    - X : Front-Back direction with respect to the listener
    - Y : Left-Right direction with respect to the listener
    - Z : Up-Down direction with respect to the listener

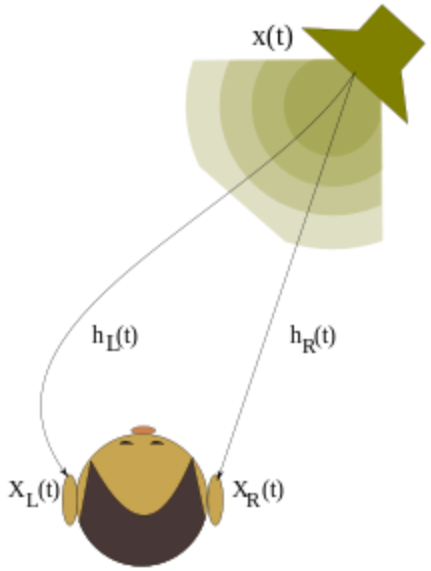
# Introduction

- Encoding Ambisonics
  - Using special surround sound microphone (tetrahedral with 4 cardoids)

- Storage
  - Widely used Ambisonics B-format
- Decoding Ambisonics
  - Using decoder to render it over any on the fly defined speaker configuration

# Introduction

- Binaural Rendering
  - Converting ambisonics decoded output into headphone outputs by using Head Related Transfer Functions.



# Introduction

- Head Related Transfer functions
  - To adjust Interaural time and intensity differences.
- CIPIC HRTF Database
  - 45 Subjects
    - 25 Azimuths
    - 50 Elevations

# Traditional Approaches

- Required predefined layout of speakers
  - Information was stored in form of channels
  - For ex 5.1, 7.1
- No consideration of height dimension
- Poor audio effects
- Listener needed to be at a particular position to get intended effects
  - Sweet spot

# Advantages of Ambisonics Approach over traditional approaches

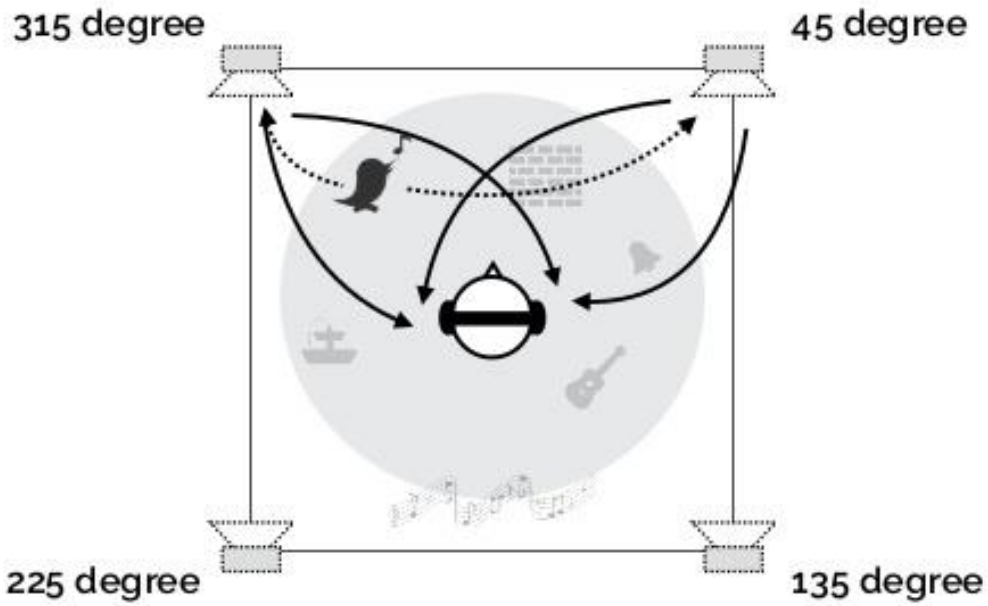
- No predefined speaker layout needs to be defined
  - Ambisonics stores the directional information of sound and not the channel information
- Listener need not to be at particular place
  - Head tracking can be used to track users head movements
- Can be used in variety of applications
  - High end real time gaming
  - 360 degree videos
  - Other VR applications

# Problem Defination

- Design a system which takes ambisonics B-format (4 channels) as an input and generates output for headphones i.e. binaural output (2 channels). For real time applications, it should take information from the head tracking devices and generate the output accordingly.

# Channel to Binaural Rendering

Set loudspeaker layout as the source positions and binaural rendering



# Literature Survey

- Traditional Approaches
  - Michael Gerzon has criticized the traditional surround sound approaches and also has given the criteria for the design of the surround sound systems [1].
  - The traditional quadraphonic systems never gave the optimum results. The aim of these systems were to duplicate the effect of 'original' 4 track tape, but they failed to do so [3].
  - Peter Fellgett said that the existing techniques were inadequate in number of ways like they restricted to fixed number of speakers and the production needs of 4 channels to be available [3].
  - They suffered from 'hole in the middle' effect and if the situation is less ideal it becomes unusable. For example when the room is non-square or when the listener is not at the sweet spot [5].

# Literature Survey

- An Ambisonics approach
  - While traditional technique of surround sound had its limitations and disadvantages, Ambisonics, developed in the early 1970s by Peter Fellget [3] and Michael Gerzon [4] is a way of recording and reproducing surround sound in both horizontal and vertical surround, which gave more immersive experience to the listener and provided full upward compatibility to any number of loudspeakers in the user defined configuration.

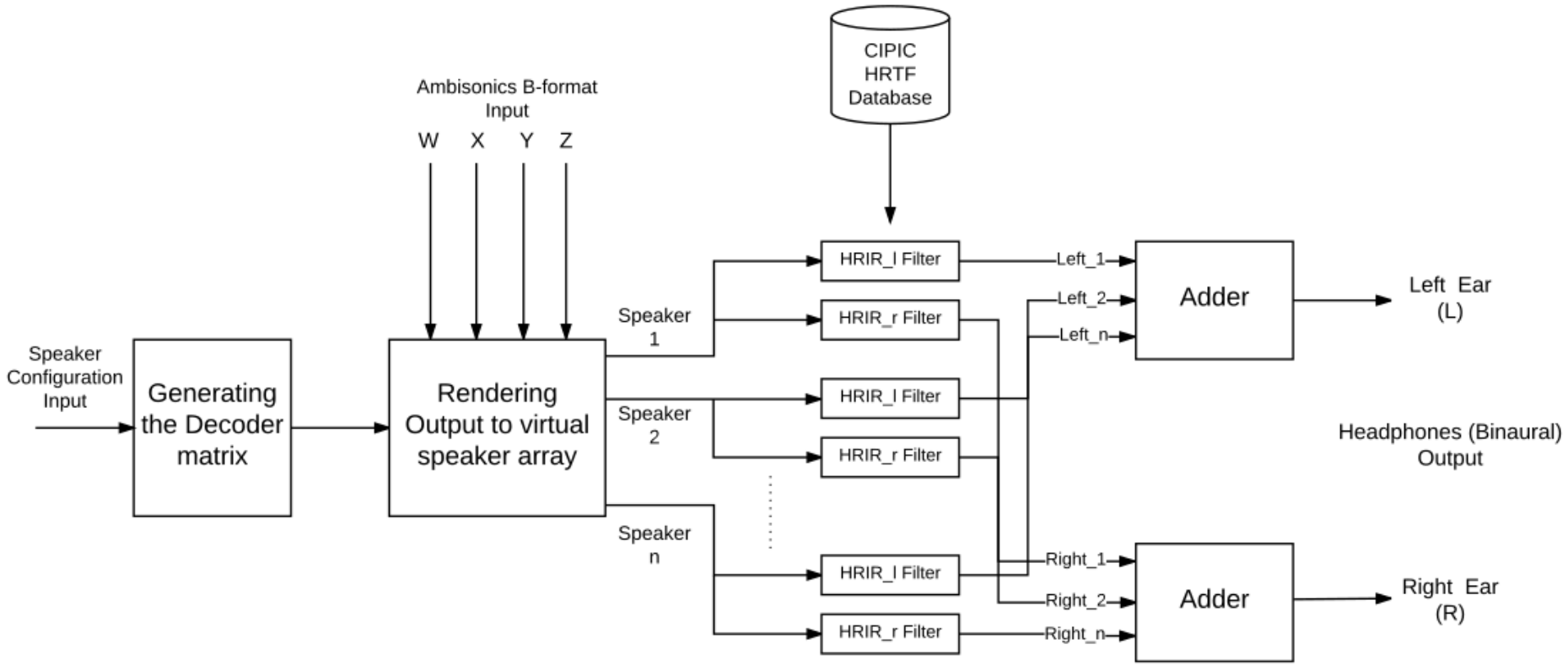
# Literature Survey

- Virtual Speakers array
  - Convincing binaural sound reproduction requires to filter the sound sources with the HRTFs. Moreover incorporation head-tracking further improvements in localization.
  - Also increasing large number of virtual sound sources helps [6]. Using the finite number of speakers gives good approximation of the original sound field over a finite area.

# Literature Survey

- Development of Binaural System
  - After rendering over the virtual speaker array HRTFs are used to filter these signals of each speaker and converting mono to left and right signals for each speakers.
  - When these all signals are superimposed we get a single left and single right (Binaural) headphone output signals.
  - Further using head tracking to take into account the head rotation will further improve the effects [6 7 10].

# Proposed System Architecture:Diagram



# Proposed System Architecture: Explaining basic blocks

- **Generating the Decoder matrix:**
  - The speker configuration (the azimuth and evevation for each speaker) will be taken as input and this function will generate a decoder matrix.

- **Rendering output to virtual speaker array:**
  - This function will take the 4 channels
  - (W, X, Y and Z) and the decoder matrix as inputs and generate a mono output for each speaker of the speaker array.

# Proposed System Architecture: Explaining basic blocks

- **The Cipic HRTF Database:**
  - This is the Database which has the HRIR pairs(left and right) for the range of azimuth and elevation pairs for each of the speakers.

- **Adder: Superimposing**
  - This is a simple adder which will add all the outputs from each HRIR-l and HRIR-r filters and generate a single left and right final binaural audio.

# Requirements Specification

- **Functional Requirements**
  - The system must take an ambisonics B-format (4 channel) file as input and give 2 channel output.
  - The head-tracking device inputs should be taken into consideration for real time applications and output should be generated accordingly.
  - It should work for 360 degree videos, real-time high-end gaming and other Virtual Reality applications.

# Requirements Specification

- **Non-Functional Requirements**
  - **Usability:** The system should be easy to integrate with any Virtual reality application.
  - **Performance:** The system should be efficient to provide the output by reducing the number of computations.
  - **Latency:** The latency must be very low to be suitable for real time applications.

# Requirements Specification

- **Hardware Requirements**
  - **RAM:** System should have minimum 2 GB RAM.
  - **Hard Disk:** System should have minimum 40 GB of free space.
  - **Head tracking device:** The user should have a head tracking device to connect to be system to give head rotation inputs for real time applications.
  - **Graphics Card:** System should have a sufficient graphics card to support the required Virtual Reality Application.

# Requirements Specification

- **Software Requirements**
  - **C++ Platform**
  - **MATLAB:** 5x or higher version
  - **CIPIC HRTF Database**

# Conclusion

- Plenty of advantages of using ambisonics approach
- Can be used for real time applications
- Better immersive audio effects
- Eliminating hardware dependency on loudspeaker requirements
- Usable for mobile applications

# Timeline

| Stage | Expected Completion |
|---|---|
| Finalizing Project Topic | Mid September |
| Literature Survey | Mid October |
| System Architecture | Mid November |
| Proof of Concept for static applications | November End |
| Working System for Static Applications | December End |
| Proof of Concept for Real-time Applications | Mid-January |
| Working System for Real-time Applications | Mid February |
| System Optimizations | February End |

# References


1. Michael Gerzon ‘Surround Sound Psychoacoustics’WirelessWorld, December 1974, pp 483-6, www.ai.sri.com/ajh/ambisonics/wireless-world-gerzon-12-1974.pdf.
2. Michael Gerzon ‘What’s Wrong with Quadraphonics?’ Studio Sound May 1974
3. Peter Fellgett. ’Ambisonics Part one: General system description’, August 1970, pp338-342
4. Gerzon, M.A. ’Ambisonics Part two: Studio Techniques’, STUDIO SOUND, August 1975. pp24 - 30, https://www.michaelgerzonphotos.org.uk/articles/Ambisonics
5. Gerzon M.A. ’Ambisonics in multichannel broadcasting and video.’ J. Audio Eng. Soc. 1985, 33 (11) pp.859-8711549-4950
6. Noisternig M., Musil T., Sontacchi A., Holdrich R. ‘A 3D Real Time Rendering Engine for Binaural Sound Reproduction’, Proceedings of the 2003 International Conference on Auditory Display, Boston, MA, USA, 6-9 july, 2003

# References


7. Cedric Yue, Teun de Planque, ‘3-D Ambisonics experience for Virtual Reality’, Stanford University, 2009
8. Noisternig M., Musil T., Sontacchi A. Robert Holdrich, ‘3D Binaural Sound Reproduction using a Virtual Ambisonic Approach’, VECIMS 2003 - lntemafional Symposium on Virtual Environments, Human~Compuler Interfaces, and Measurement Systems Lugann, Switzerland, 27-29 July 2003, pp 174-178, IEEE, Washington (2003)
9. Algazi V.R., Duda R.O., Thompson D.M., Avendano C. ‘The ClPlC HRTF Database’, in Proc IEEEWorkhop on Applicalions of Six. Proc lo Audio and Elecfrmcou~ficpsp, 9-102, NY, 2001

# Thank You!

# Questions?