



The Evolution of Spatial Sound for Consumers

Pingkang Chen

MAT200A



(Pierre Boulez, Répons - Ensemble intercontemporain - Matthias Pintscher)

1: Spatialization in Classical Music

Historical Examples of Sound Spatialization

Spatialization in Classical Music:

- In the score of opera Don Giovanni (1787), Mozart wrote parts for three separate orchestras: one in the pit, one onstage, and one backstage.
- Hector Berlioz's Requiem (1837): Four orchestras are positioned in different parts of the concert venue.
- Stockhausen's "Gruppen" (1955-1957): Karlheinz Stockhausen's composition for three orchestras requires them to be positioned around the audience, with sounds moving dynamically between groups. Link: https://youtu.be/34_SfP7ZCXA?si=egHyb6UOZ5KrV8TC





(The Acousmonium, a multichannel spatializer)

2: Spatialization in 20th-Century Electroacoustic Music

Spatialization in 20th-Century Electroacoustic Music:

- Karlheinz Stockhausen's "Gesang der Jünglinge" (1956): Projected over five groups of loudspeakers in a concert hall.
- Edgard Varèse's "Poème Électronique" (1958): Used over 400 loudspeakers in the Philips Pavilion, dynamically moving sound around the audience.
- Salvatore Martirano's "Sal-Mar Construction" (1971): Used 250 loudspeakers suspended in concert hall ceilings.
- The Acousmonium, a multichannel spatializer designed by the Groupe de Recherches Musicales (GRM), installed in Olivier Messiaen concert hall, Maison de Radio France, Paris, in 1980. Projecting sound over eighty loudspeakers played through a 48-channel mixer, the Acousmonium achieves a complexity of sound image rivaling that of an orchestra. It lets a composer reorchestrate an electronic composition for Acousmonium spatial performance. Link: <https://youtu.be/NqRHUPFt5iw?si=yOEjXcJ5-0P58SxJ>

.....



(5.1 surround sound system)

3: Channel based audio

Channel based audio

"Channel-based audio" refers to a system where each individual sound source is assigned to a specific speaker channel (like left, right, center, surround), meaning the audio is mixed and encoded for a predetermined speaker layout (like 5.1 or 7.1), and each channel feeds a dedicated speaker, creating a spatial sound experience based on the speaker configuration

Common formats: mono, stereo, 4.0, 5.1, 7.1, 22.2, etc.

Channel-based workflow

***The downside of the channel-based audio format is that it is locked into a specific channel layout.



Audio



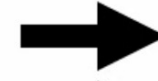
Console - Panner



Quadraphonic
(4.0 Surround)

Loudspeakers:

Left Front



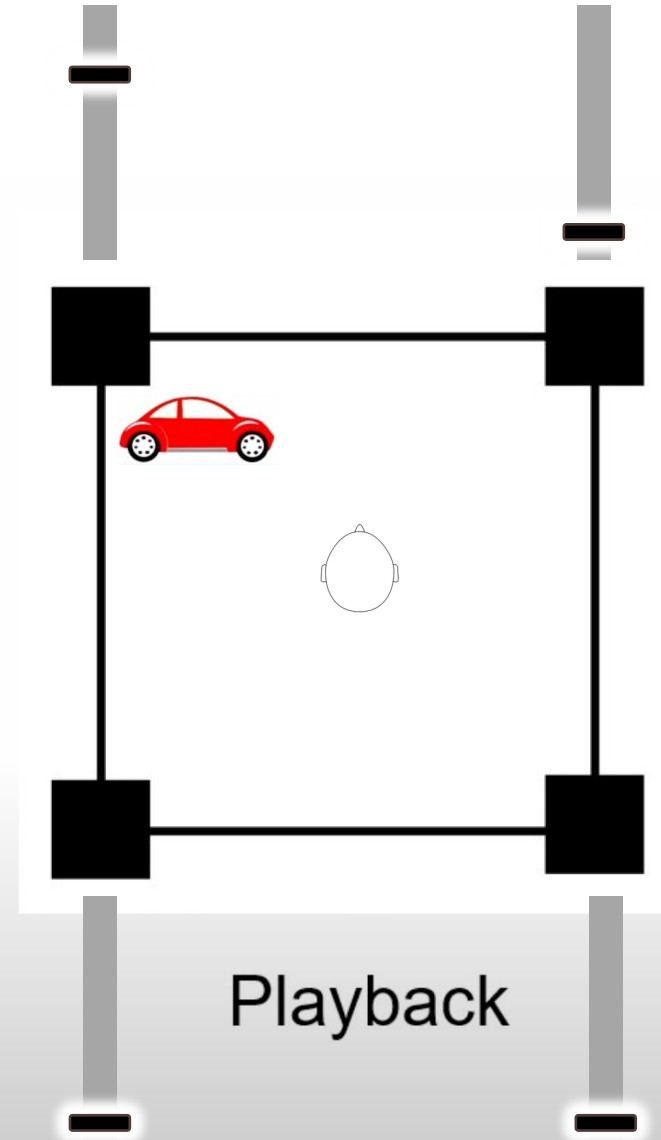
Right Front



Left Rear



Right Rear



Playback

Example



Believer - Imagine Dragons (7.1 surround mix)

Link: <https://youtu.be/74d2Gxum9mI?si=pXxMAOhPzqE7onqp>

This famous song was remixed into a 7.1 surround version, which can be streamed with seven speakers: left front, center, right front, left surround, right surround, left rear, right rear, and a subwoofer.



(Dolby Atmos Theater)

4: Object based audio

Object based audio

Object-based audio is a technology that creates immersive sound experiences by treating each sound source as an individual object with its own metadata (x,y,z coordinates data).

Common formats: Dolby Atmos, Sony 360 Reality Audio, DTS:X, MPEG-H, AuroMax

Commonly used in Movie, Broadcast, Music, Video Games.

Object-based workflow

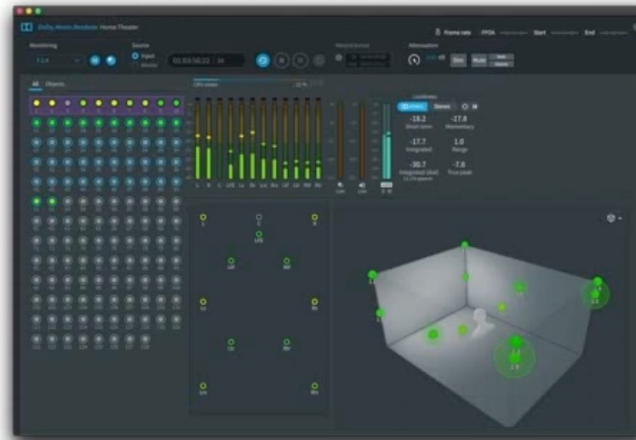
***Object-based audio is scalable, adaptable

the volume ratios are dynamic and computed in real-time based on metadata.

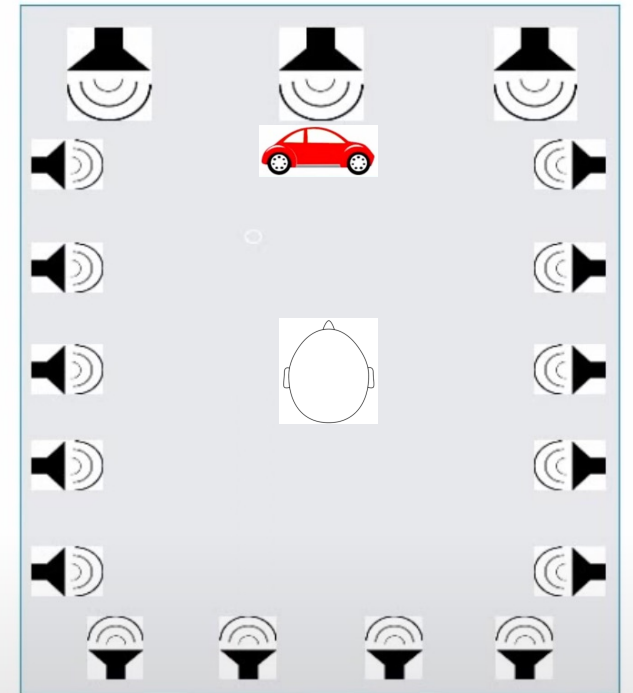


X Y Z coordinates

Audio + metadata



Renderer



Playback



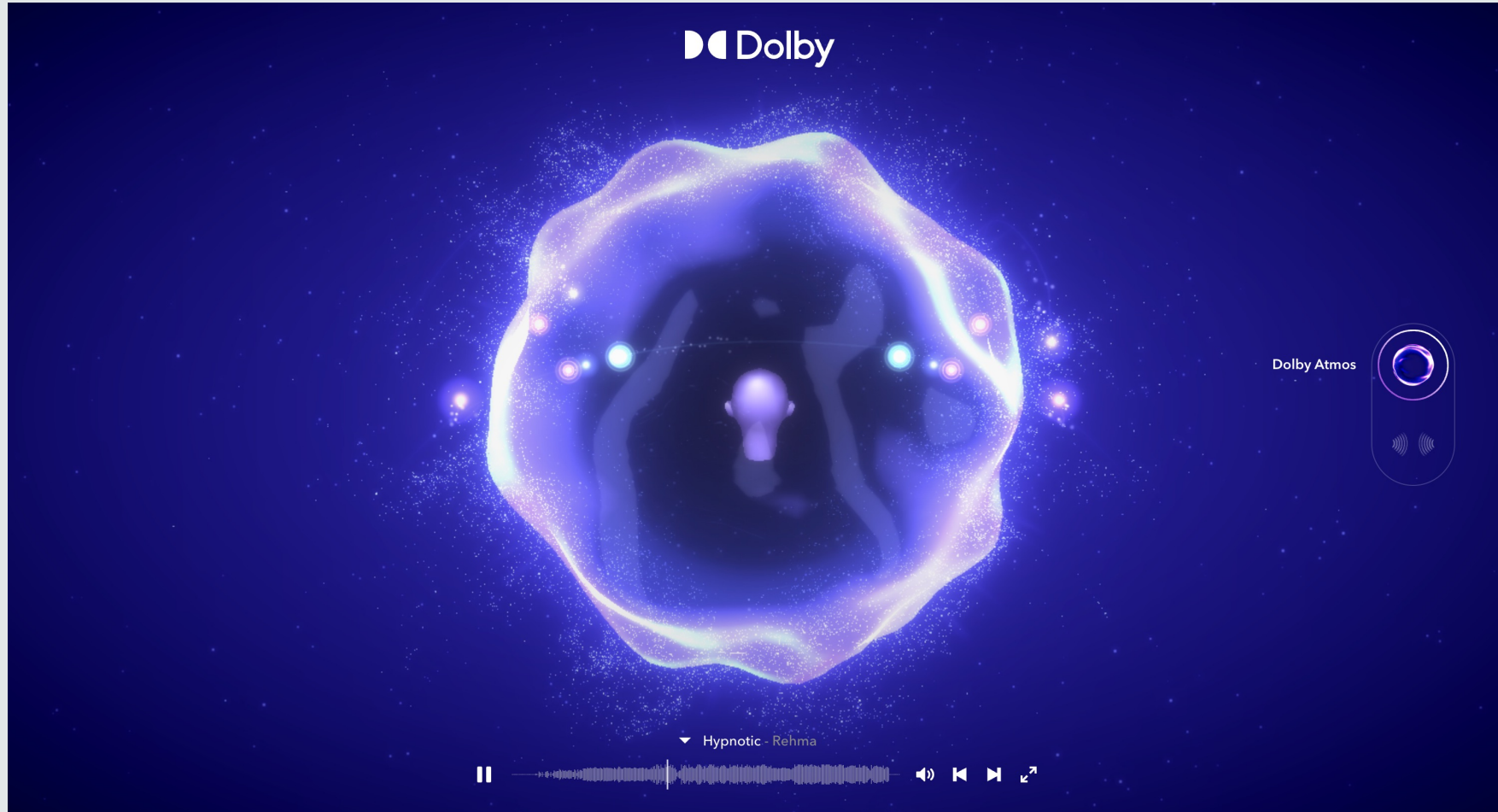
Sony 360 Reality Audio

Dolby Atmos

The author of the chapter 'Spatial Music Composition' in the book '3D Audio' notes, 'object-based approaches allow us to create spatial music without needing to specify, nor in fact consider, the loudspeaker configuration at the outset. As composers, we are liberated from 'loudspeakers'.

QUOTE

Example



Link: <https://www.dolby.com/atmos-visualizer-music/>

This webpage, called Dolby Atmos Visualizer, allows you to switch between stereo and its Dolby Atmos mixed versions, vividly showcasing the artistry of objects moving around the listener's head and reflecting the magic of object-based audio, providing a greater sense of immersion.



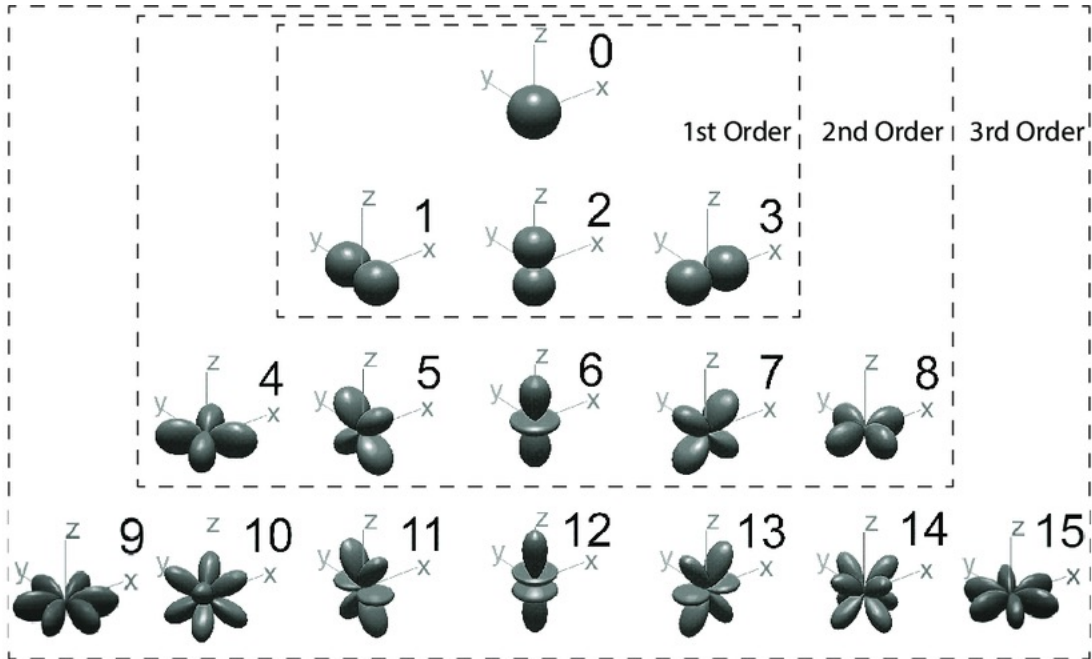
5: Headphone-based audio



Ambisonics + Binaural audio (Headphones)

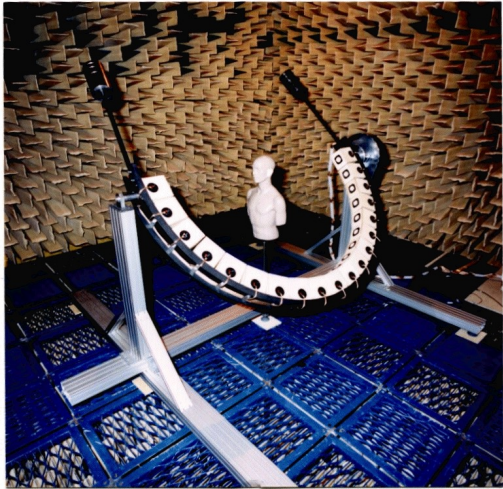
With advances in spatial signal processing, the rise of headphone audio streaming, and the growth of VR, ambisonics + binaural audio was employed, applied in VR/AR, 360° Video, 360° Live-Streaming, etc.

Ambisonics is a spatialization technology that captures 360° sound in spherical harmonics, and projects the sound field over a loudspeaker array specified in a decoder. Each spherical harmonic is represented in an audio channel, and the number of spherical harmonics determines the maximum precision at which the signal can be recreated.

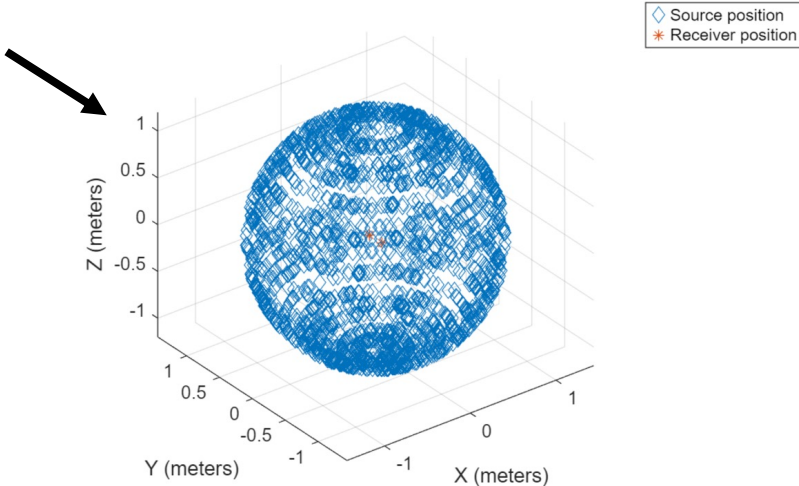


Ambisonics spherical harmonics

Binaural signals involve two channels that simulate how sound enters a listener's left and right ears.



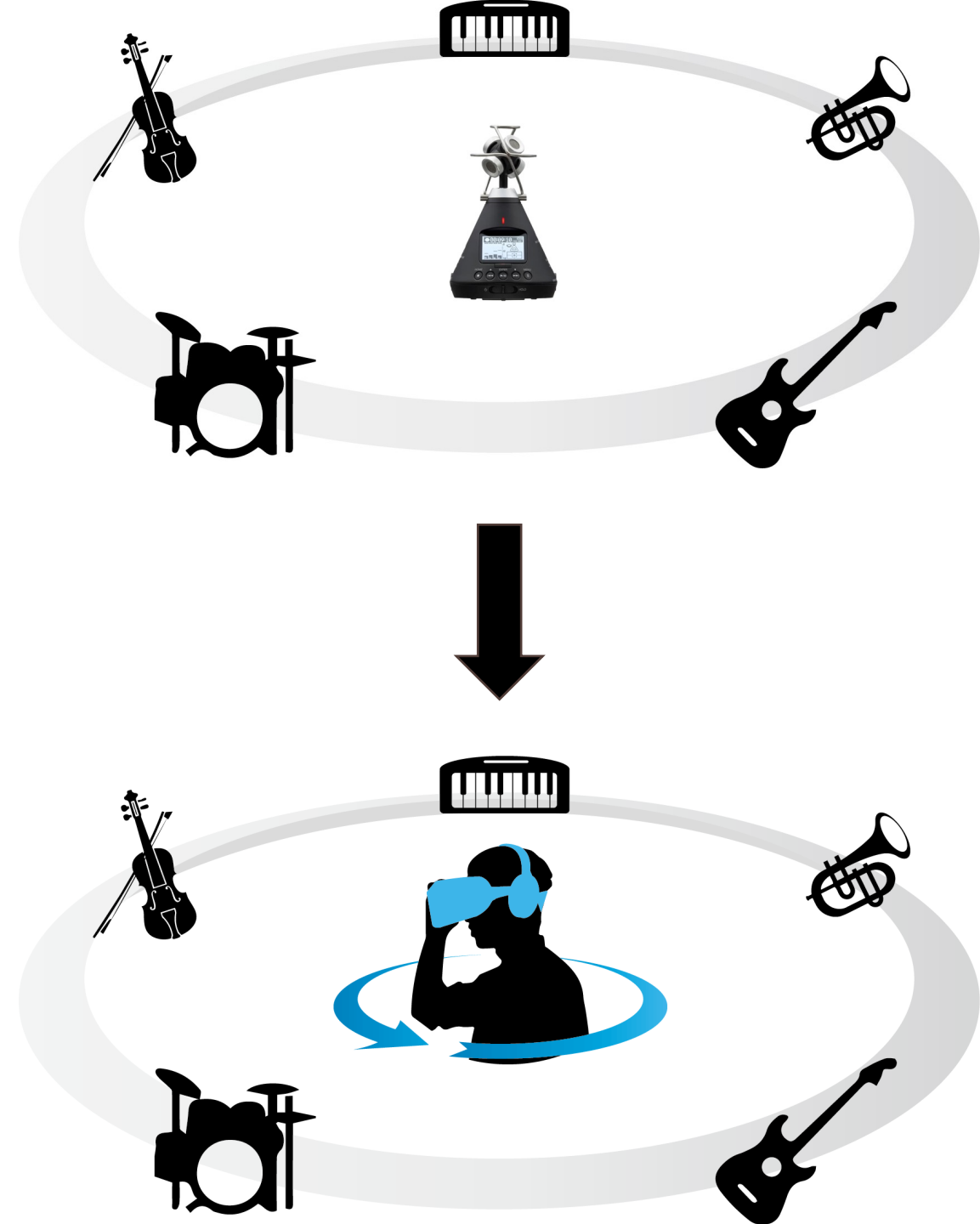
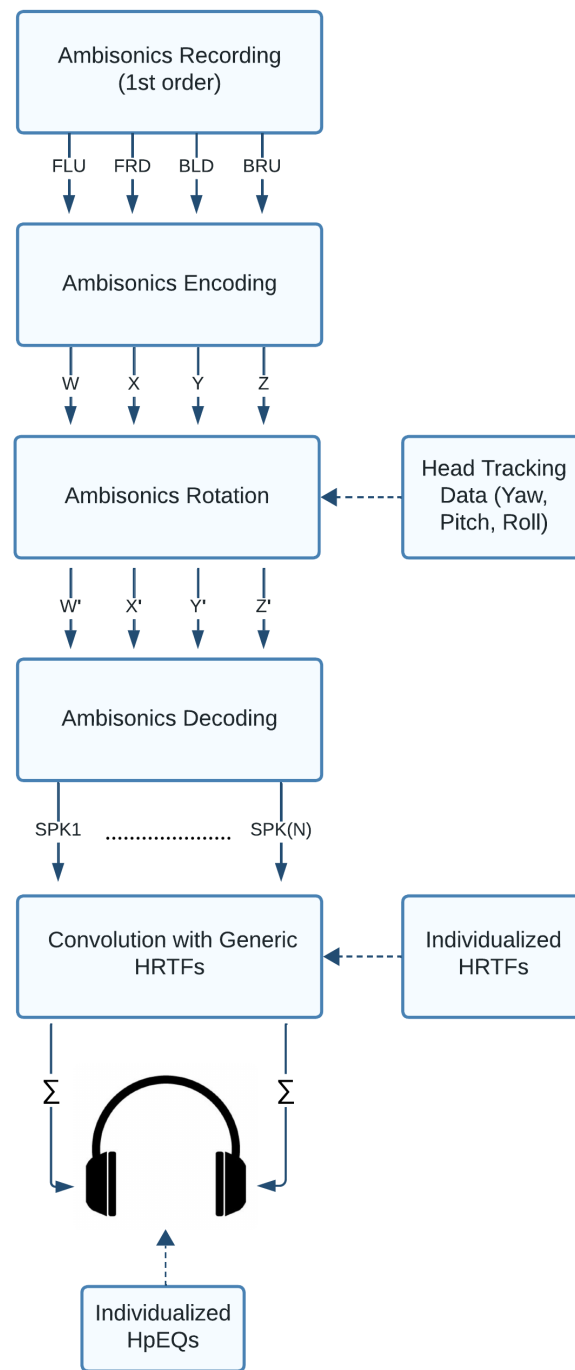
HRTF (Head-related transfer function) measurement using dummy head microphone



Measurement Geometry of source locations

Production of 3DoF head tracked binaural audio

The audio signal chain for 1st-order Ambisonics recording to head-tracked binaural audio over headphones under 3Dof.



Example



Link: <https://aes2.org/publications/elibrary-page/?id=22702>

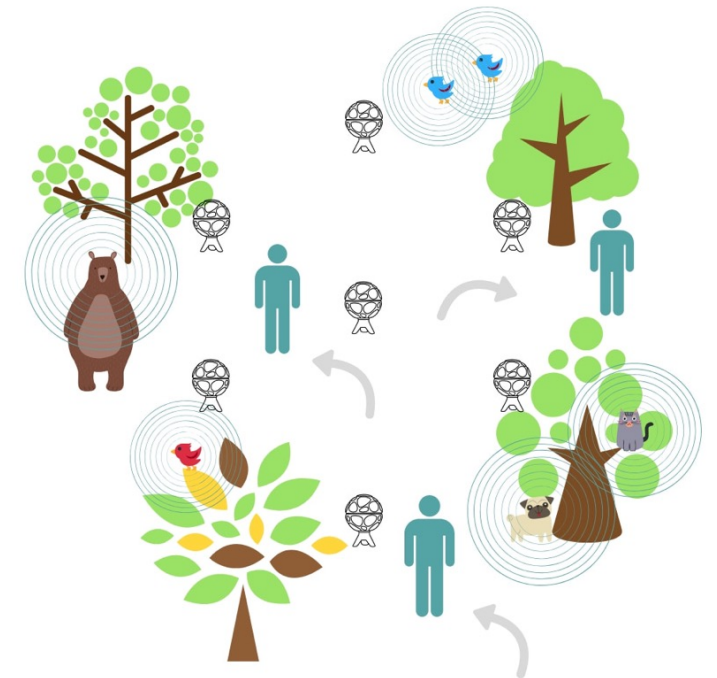
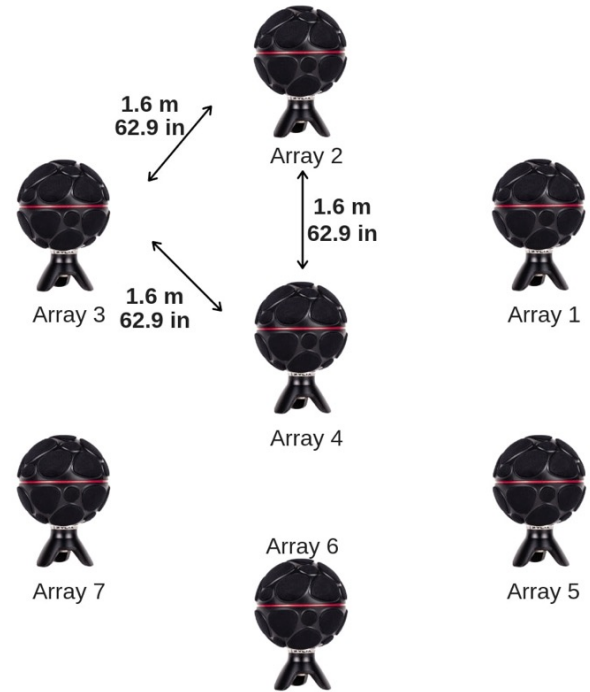
This paper outlines a process for achieving a head-tracked binaural audio system designed to enhance realism for players of digital pianos. Using an Ambisonic microphone to sample an acoustic piano, followed by leveraging off-the-shelf equipment, the system allows players to experience changes in the sound field in real-time as they rotate their heads while wearing headphones under three degrees of freedom (3DoF).

6 Degree of Freedom binaural Audio

- Capture Ambisonics sound fields of multiple points in the recorded scene simultaneously.

- Interpolation Algorithm for obtaining the In between ambisonics sound fields

- Listener can move inside the captured Ambisonics sound fields with 3 rotational (roll, pitch and yaw) and 3 translational (up-down, left-right and forward-back) movements.



Example

ZYLIA 6 Degrees of Freedom Navigable Audio

Link:

<https://youtu.be/m5YgcCSvLp4?si=PNJa0GAMsycma-XA>

This app 'ZYLIA Concert Hall' can allow listeners navigate smoothly around the scene and experience the 6Dof binaural sound from anyplace – close perspective, backstage, or the audience.



ZYLIA Concert Hall

***Downloadable on Meta Store

A decorative graphic consisting of two concentric, light gray circular outlines that are partially visible on the left and right sides of the frame, framing the central text.

END