

SPATIAL AUDIO INTERFACES

Manohar Swaminathan

Microsoft Research



MULTIMODAL INTERFACES

- Input

- Type/text
- WIMP
- Speech
- Gaze
- Gestures
- Touch
- Muscle
- Body movement

- Output

- Text, graphics, animation, drawing, video
- Audio
- Speech
- Haptics
- (smell/taste)
- Muscle stimulus
- Pressure suits

SPATIAL AUDIO

(SPEECH AND NON-SPEECH)

- What is Spatial Audio
- Why spatial audio
- Technologies for spatial audio
- Usage scenarios
- Research Directions
- Research Questions



WHAT IS SPATIAL AUDIO?

- Audio is one of the two far-senses that humans have evolved
- Vision is the other far-sense
- Both essential for survival
- Ability to 'precisely' locate the audio source in 360-degree space
 - Among a sea of other sources
- 'Hearing Detects and Vision Locates'



WHY SPATIAL AUDIO?

- Second highest b/w input to the brain
- The human brain has evolutionary structures
- Complex audio environments can be interpreted
- Present with all humans with unimpaired hearing
- Demonstrated capability in everyday experience
 - Cocktail party effect
 - Surviving in Bangalore traffic
- Currently underutilized in HCI

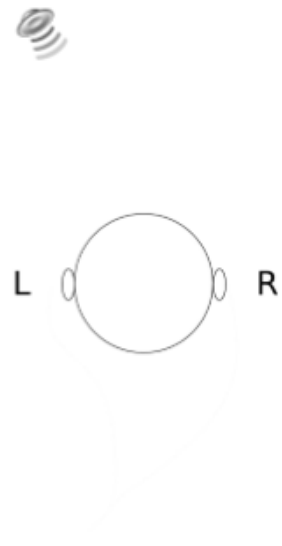


Definition: Spatial audio

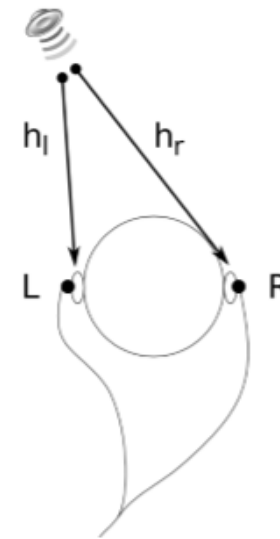
- Within audible frequency and dynamic range
 - Delivered to one or both ears
 - Contains auditory localisation cues:
 - Interaural time and level differences
 - Spectral cues
 - Reverberation
 - Dynamic and multimodal cues
 - (expectation and experience)
- } Head-related transfer function (HRTF)



Spatial audio rendering



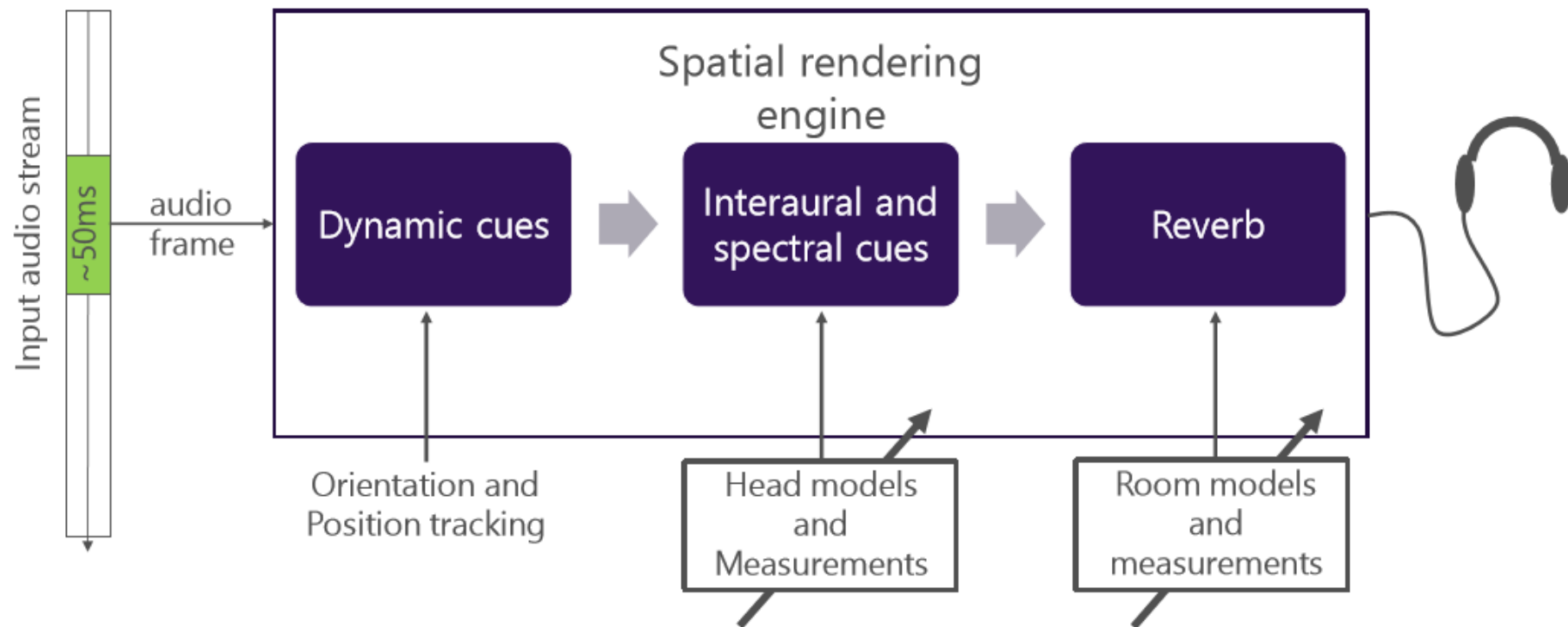
Real sound source



Spatial audio rendering via head-related transfer functions (HRTFs)



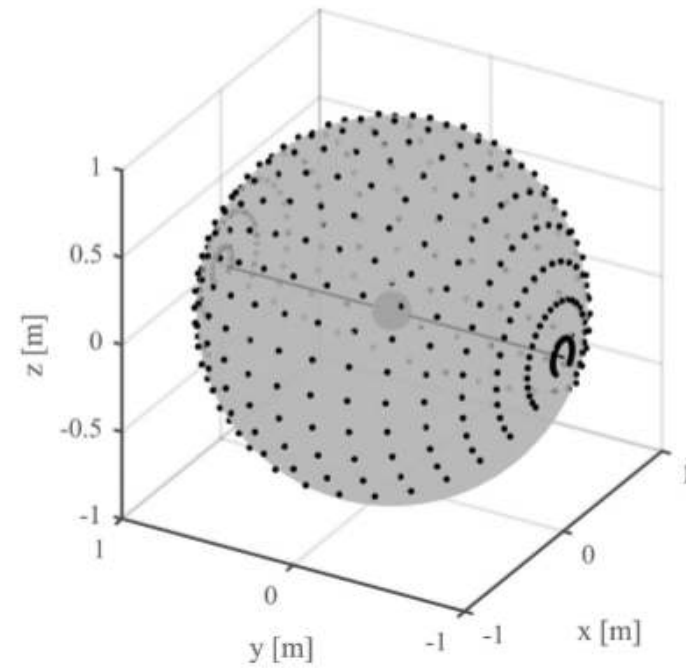
Spatial audio rendering



HRTF measurement



HRTF measurement rig



Measurement locations



SPATIAL AUDIO TECHNOLOGIES

- Head tracking is important to provide accurate spatial audio
- Custom HRTF is needed for each individual for top quality, but approximate HRTFs will be useful in many situations
- Standard 3d audio library in Windows
- Availability of commercial head-tracked headphones like the Jabra



USAGE SCENARIOS

- Games
- Immersive VR experiences
- 360 degree video
- Accessibility



SPATIAL AUDIO FOR ACCESSIBILITY

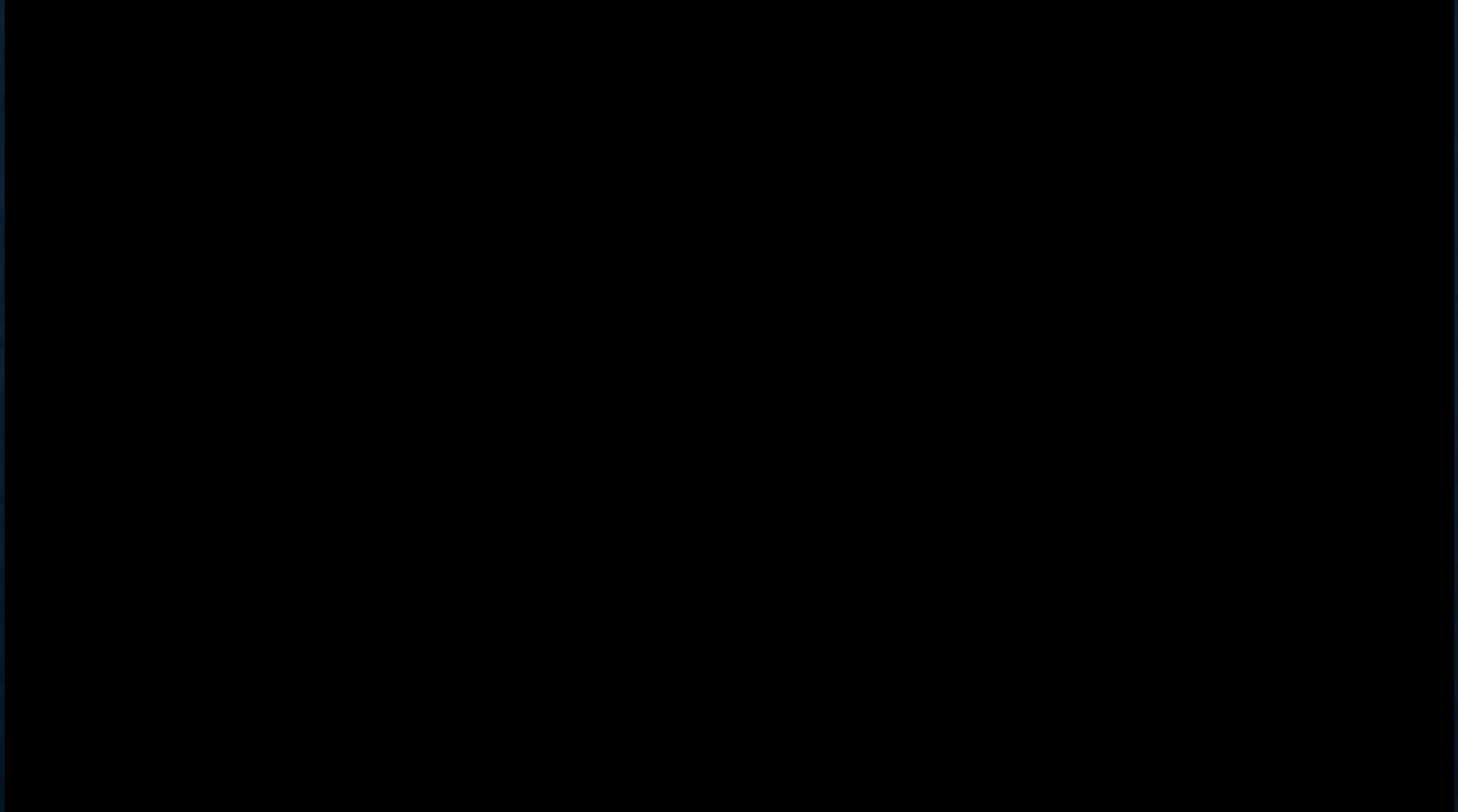
- For making mainstream video games and virtual environments accessible to gamers without sight.
- Extract information about the world around in the form of a spatialized audio cloud: speech, sounds, earcons and audicons to describing diverse scene elements
- A set of tools to extract information from this spatial audio cloud



SPATIAL AUDIO INTERACTION TOOLS

- Viewdio
- View Scan
- Sound Sabre
- Audio compass
- Select and reach
- Bump noise, material specific





BodyScan

On selection of BodyScan, the player gets spatialized audio playback of all objects that intersect a virtual frustum emanating outwards from the player's body. Every object that intersects this frustum is sonified with spatial audio, using text to speech of the object's name.

Select and Reach

During BodyScan, the user can activate the Select and Reach tool to select one of the objects being listed. On selection, a spatialized siren, with modulation in volume and pitch, guides the gamer to reach the selected object.

Watch Tower

Watch Tower

Fence

Entrance

Fence

You are facing an entry point to an abandoned.....

Player Position

Viewdio

Viewdio is a text description can be as brief or as detailed to convey a crisp summary of the scene as seen by a sighted player from that location and view direction. It has a view vector and a field of view which can be controlled by the game designer



UNITY PLUGIN FOR DEVELOPERS

- Responsive Spatial Audio for Immersive Gaming
- Available for download at the Unity Store
 - Our accessibility tools are bundled in a 14MB Unity package
 - Adds all accessibility tools to the editor,
 - Makes building accessible games a drag-drop operation





VR/MR games for people with vision impairment



SPATIAL AUDIO FOR ACCESSIBILITY IN THE REAL WORLD

- [Microsoft Soundscape App for iOS Helps the Visually Challenged Navigate Cities With Audio Cues | Technology News \(ndtv.com\)](#)



FURTHER READING AND RESEARCH

- HRTF personalization: [Head-related transfer function personalization for the needs of spatial audio in mixed and virtual reality \(microsoft.com\)](#)
- Podcast on Hearing in 3D by Ivan Tashev: [Hearing in 3D with Dr. Ivan Tashev - Microsoft Research](#)
- Video gaming for the vision impaired: [Video Gaming for the Vision Impaired | Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility](#)
- Unity asset store: [Responsive Spatial Audio for Immersive Gaming, a Microsoft Garage project | Game Toolkits | Unity Asset Store](#)



QUESTIONS AND DISCUSSION