



Auditory Reality

<https://auditory-reality.vercel.app>

Fang, Xiaohai, Yuchen, Rui



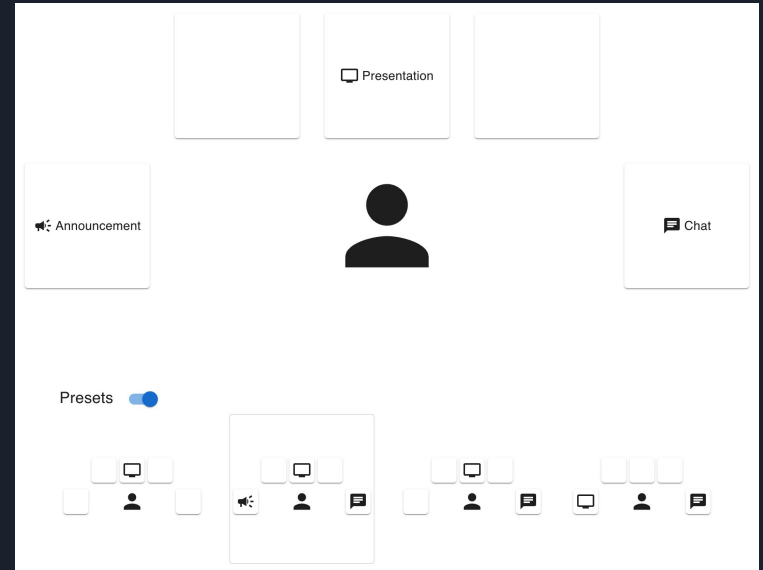
User Needs

- People who are blind or have low vision that use screen readers during video conferences (e.g. Zoom) often are **overwhelmed** by multiple sound sources **played simultaneously**
- They want to **retrieve information** from different sound sources without **auditory overload**

Solution

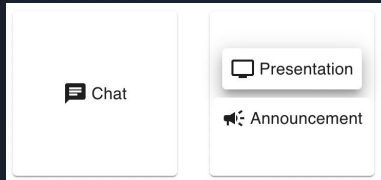
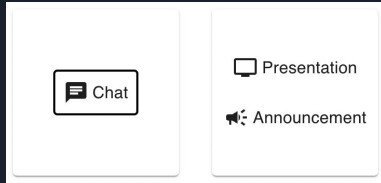
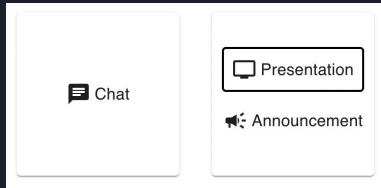
Create a virtual conferencing reality
using **spatial audio**

- Separate sound sources to different positions
 - Presentation
 - Chats
 - Announcements
- Control sound source positions
 - Move sound like a chess piece
 - Adjust layout in real time

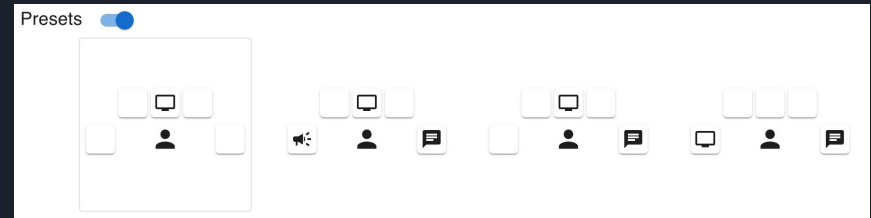


Solution: Control

Manual mode / Preset mode for sound source positions



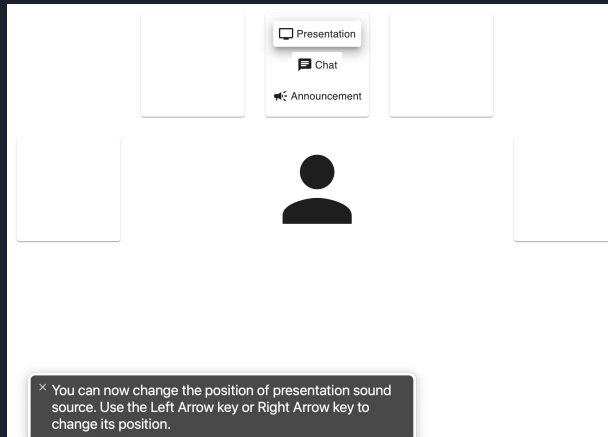
Cycle through sound sources and select one to move around



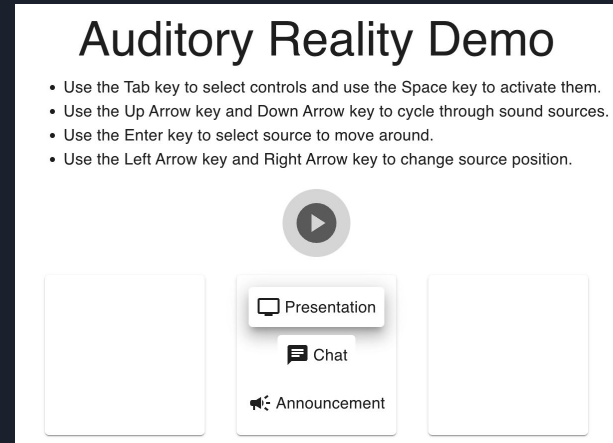
Directly select a preset for sound source positions

Solution: Accessibility

Screen reader / keyboard navigation compatible



Every visual cues are accompanied by audio cues understood by screen readers



Use Tab, Space and direction keys to achieve any functionality



Result

Spatial audio helps information retrieval

- It helps users to distinguish among different sound sources
- It helps users to focus on one or more sound sources
- It increases the understanding of content

“It does make a difference. Everything is separated now.” - P2

“It is easier to listen to one side when you put the presentation on the left and other things on the right.” - P3

Mono channel 1.83



Spatial audio 4.17

1 “I don’t understand at all”

5 “I understand everything”



Result

Spatial audio reduces
information overload

- It helps users to focus
- It makes users more relaxed and calm and reduces the overwhelmed feeling

"I'm not overwhelmed. I can pay attention to where I want to be." - P1

"Very relaxed. Very easy to pick up events." - P2

Mono channel 4~5



Spatial audio 1~2

5 "very overwhelmed"

1 "relaxed and calm"



Result

Timeliness can be sacrificed to reduce distraction

“In the meeting, use sound instead of words for chat. If you want to read it, you can; if you don't, you don't have to.”
- P1

“Typically while chat is interesting, but not mandatory, I don't need to hear all that necessarily, unless it's a direct message to me.” - P3

- Users prefer to know something is happening in a subtle way and come back to it at their own convenience
- Users prefer signaling sound like beeps instead of full announcements
- Users want to control the volume of sound sources and even mute them



Result

Spatial audio builds social simulation and mental maps

“Presentation needs to be up front so I can focus.” - P1

“I want to look at you, so you are in the front. If the presenter is on my left, I need to look left. I still try my best for eye contact.” - P2

- Users prefer the main presentation in the front, so they could face the presenter as if in a real social setting.
- Users talked about two mental maps
 - “Auditorium”: they are audience
 - “Table”: they push unneeded items off to the side but know where to retrieve them

Video Demo

