### Abstract

IEAR is a flexible spatial audio rendering tool for use with photorealistic monoscopic and stereoscopic panoramas across various display systems. IEAR allows users to easily present multichannel audio scenes over variable speaker arrangements, while maintaining tight integration with the corresponding visual elements of the display media. Built in the Mac/MSF Audio Programming Environment, IEAR utilizes well-established panning methods to accommodate a wide range of speaker configurations. Audio scene orientation is tied to the visual scene using an OSC connection with the visualization software, allowing users to render and spatialize multichannel environmental audio recordings in tandem with the changing perspective in the visual scene.

### The Problem

Capture and display of stereoscopic panoramic imagery has become increasingly efficient. With technologies such as the CavorCam, stereo capture of spherical panoramas takes only minutes in the field [1]. These panoramas can then be rendered by large-scale stereo VR systems, such as the CAVE, NextCave (see Figure 1), and CAVE2, or any stereo-enabled display technology [2], [3], [4]. The user becomes immersed in this photorealistic rendering of the space and can utilize the experience for research, presentation, and experimentation. Similarly, capture of monoscopic panoramas has become simple, thanks in part to the ubiquity of mobile-device cameras and panoramic image software. Google Maps’ street view layer delivers a photorealistic VR environment on an unprecedented scale using monoscopic panoramas.

While display workflow and associated technologies have developed considerably in this field, yielding impressive results in the visual domain, the lack of a purpose-built tool for rendering immersive audio within these frameworks has limited users’ experience in terms of presence. The benefits of an audio-enabled display environment on a user’s experience are well documented [5], [6], [7]. Many tools exist for the development of real-time spatial audio rendering (OpenAL, SuperCollider, and Max/MSP, for example), but they are generally low-level and can only be deployed efficiently by audio development specialists. There are also a number of platforms for the simultaneous presentation of immersive VR and audio, but their implementation for the purpose of enhancing photorealistic VR with audio would be impractical for most facilities, given the complexity and the specific needs of these systems exceeding [8], [9], [10]. A purpose-built tool that can quickly tie in to any existing framework is required.

### Solutions:

- Output to n channels
- Simple, intuitive interface for system setup
- Joun format conversion file
- OSC/CUPS offering with visualization application for orientation
- Up to 8 simultaneous sources/source channels
- Binaural Mode for facilities with limited loudspeaker coverage

### iEar

**Speaker Setup**

- **Number of Loudspeakers:**
  - L: Loudspeaker Channel: Output Setting
  - R: Location of Loudspeakers: Call Location

**Presetbilization Method**

- **Master Level:**
  - Play
  - Stop

**Source Setup**

- **Number of Sources:**
  - 0: Location of Sources: Call Location

**Network**

- **Select Remote:**
  - Play
  - Stop

### Future Work

- Additional panning algorithms to offer more idiographic rendering
- Integration of decoration option for sound design use case
- Automatic computation of loudspeaker delays
- Expansion of supported network protocols
- Support for collection of panoramas (intelligent scene morphing and orientation metadata handling)

### Stereoscopic Panoramas

Stereoscopic panoramas are displayed and explored quite frequently at KAUST, whether for the purpose of demonstrating the quality of the VR system for visitors, or for the exploration of archeological and historical sites. The primary system on which these panoramas are displayed is the NextCave: a hemispherical array of passive-stereo LCD panels with infrared head and hand tracking, and an atypical 5.1 speaker arrangement (see Figures 1 and 4). The displays are driven by 11 render nodes and 1 headnode, all running Scientific Linux 6, tied together via a 10Gb network infrastructure.

The software used to display the stereoscopic panoramas on KAUST’s VR systems is CaVR [12]. Within this framework we have developed a custom plugin that, via OSC protocol, communicates the user’s orientation within the virtual scene to our audio server, along with information about the particular scene being viewed.

Because IEAR receives filename information from the visualization software, users need only load an audio-enabled panorama via CaVR and they become enveloped by a vivid, multimodal scene.

### References


### Acknowledgement

The authors would like to thank the staff of the Visualization Lab at King Abdullah University of Science and Technology with special thanks to Rob Collins and Neil Smith for their editorial assistance.