Pigments of Imagination: An A-Frame WebVR Project

Pigments of Imagination

An A-Frame WebVR Project

GABRIEL ZALLES BALLIVIAN

UC SAN DIEGO, GZALLES@UCSD.EDU

TIM GMEINER

UC SAN DIEGO, TGMEINER@UCSD.EDU

EITO MURAKAMI

UC SAN DIEGO, EMURAKAM@UCSD.EDU

Additional Keywords and Phrases: web-xr, spatial audio, multimedia-arts, ambisonics, FOSS.

ACM Reference Format for the Audiovisual Symposium:

Gabriel Zalles Ballivian, Tim Gmeiner, Eito Murakami. 2022. Pigment of Imagination: An A-Frame WebVR Project. In Audiovisual Symposium Notes '22, November 25, 2022, Dalarna University, Falun, Sweden.

1 INTRODUCTION

Pigment of Imagination is a popular music track touching on themes of imagination and space-travel. The song is composed and produced by Timothy "Ill Poetic" Gmeiner and features vocals by both Ill Poetic and Nick Tolford, as well as flugelhorn by renowned jazz artist Stephanie Richards, and additional synthesizer parts by DJ/Producer King Britt. The stems from the song were used to create a special HOA (Higher Order Ambisonic) 3D spatial audio mix, which you can hear online via the A-Frame project we created. In addition to the HOA, the A-frame project features a custom movie that provides dynamism to the experience. During our experiments, we found that our HOA files did not load on most mobile devices, thus, alternate mixes in FOA (First Order Ambisonic) and static binaural are also made available. Currently hosted on: https://gabrielzalles.neocities.org

2 AESTHETIC DISCUSSION

A-frame and WebVR were selected due to their interoperability between hardware devices and because it is free and open source. Anybody with a desktop computer or phone can view our work, it is not necessary to have a VR (virtual reality) headset. We are interested in the problem of socio-economic access in spatial audio compositions, a dynamic which has led to a technocratic society which values channel-count over artistry. Instead, here we focus, both in process as well as product, on the accessibility of our components, from the economic perspective.

Ambisonic has an isotropic characteristic which means it considers all directions as equally important in the reproduction and capturing of audio. This method of spatial audio synthesis is characterized by its use of basis functions in lieu of channel-based information. In surround sound we usually have 6 channels corresponding to left, right, center, left surround, right surround, and subwoofer. Ambisonics is not “channel-based” rather the four audio channels in FOA correspond to information in the X, Y, and Z axis, as well as a W channel which corresponds to an omnidirectional signal (e.g., sound arriving from all directions equally).

Ambisonics is very flexible because with a HOA or FOA mix, one can reproduce the material over any number of speakers from 1 to infinitum. Furthermore, ambisonics counts with linear transformations such as rotation, which allow us to seamlessly reproduce audio binaurally. In binaural reproduction we defined a number of virtual loudspeakers and play the audio through these consequently filtering the sound with HRTFs (head-related transfer functions). These HRTFs contain the time/gain/frequency differences for sounds played from various positions, for both ears. In other words, they characterize the effect our heads and ears have on sound.

This project is part of a larger initiative called SElectOr which seeks to bring undergrads, graduate students and faculty together in collaborative music-making projects. In addition to these three main groups we also invite alumni, staff, and external members to contribute. The aim is to broaden access to the tools and skills required to create modern multi-media works and forge community in the process. The group has been operating since 2018.

3 TECHNICAL SETUP

This piece does not require a traditional multi-channel audio system or projection equipment. All one needs to properly experience this work is a Google Cardboard device (and an Android phone). Naturally, given the binaural nature of one would also need a pair of headphones. Another option is to set it up on a computer, although we find the Cardboard experience to be slightly more immersive. The project is currently not operational on HMDs, such as Oculus or Vive. Unfortunately, none of the creators has access to one of these devices, so we were only able to test on Cardboard and Desktop.

This document is meant to provide greater insight into the creation process behind this project. Our hope is that the reader will gain insight into the tools and methods which we believe are essential for ambisonic music today. These comprise open source VST libraries for producing the sound, JS systems to deliver them, and, optionally, media belonging to the commons which can be used to create satisfying visual accompaniments. Showcasing this same artwork in a black box theatre might be prohibitively expensive for some institutions, this project is an example of a low-cost alternative which exploits FOSS and accessible materials.

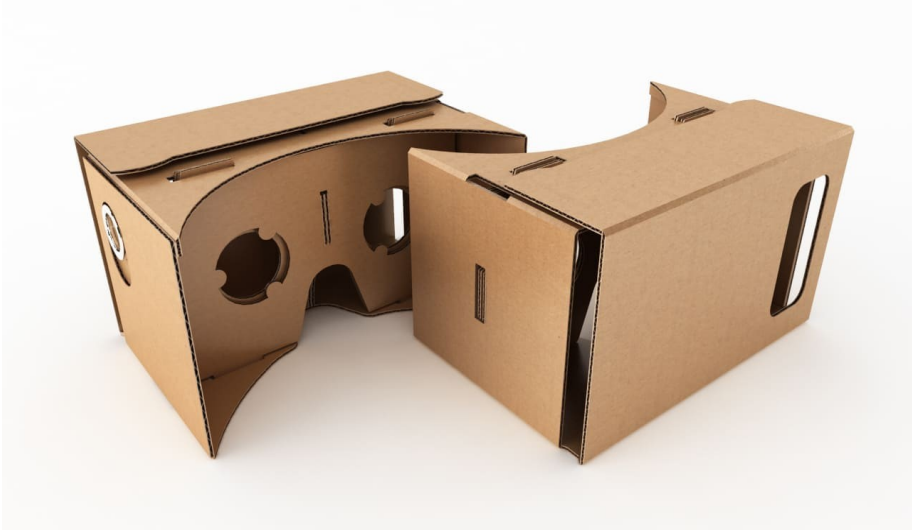

Figure 1: Google Cardboard. Photograph by derealidadvirtual.site.

4 DESCRIPTION

The main element allowing this work to be possible is the A-Frame Ambisonic component, which is an A-Frame component providing a high-level abstraction to the Omnitone API. The Omnitone project is responsible for the binaural decoding of HOA/FOA sound files which were created in Reaper, a cross-platform DAW, using the IEM Plug-In Suite (image below of the stereo encoder).

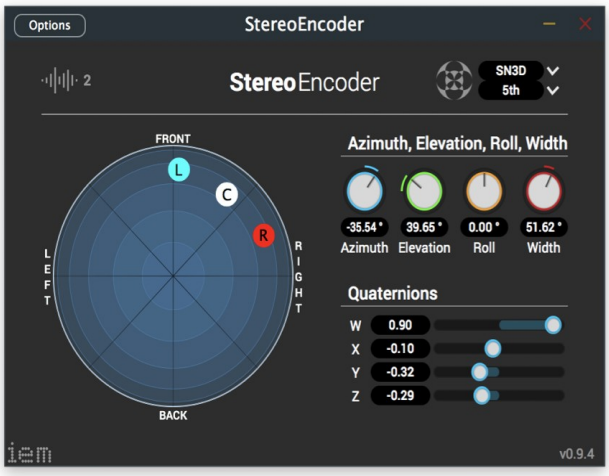

Figure 2: Stereo Encoder GUI. Image by plugins.iem.at.

Following the creation of our raw HOA sound file, with 16 channels, in uncompressed format, the sound files were imported into Audacity for compressing and splitting into two 8-channel files, as per the provided example. The OGG codec was chosen, also as per the example, and different quality settings were tested to determine a compromise between load-time and over-compression artifacts. Due to the length of the song, and the number of channels, we used a timeout feature in A-Frame which gives the listener's connection 10 second to download the 67MB of audio required of the project (as well as other assets). We ultimately selected a quality factor of 5 in the OGG compression, out of the possible ten quality levels.

As per the recommendation of the A-Frame documentation, we opted for the server solution provided by neocities.org (which is open-source). A movie created using open source software and media from the commons was created to accompany the music. A particle system component was adopted to give the piece some additional movement. Several spheres were textured and set to rotate around the listener serving the theme of outer space. A 360 degree photograph from the European Southern Observatory was used to texture our sky element, after it had been resized to a lower resolution (4K) in order to reduce its size. Some GLTFs (3D models with low-polygon count) were also imported into the scene which were collected from Google Poly.

In contrast to real-world decoding of the ambisonic signal, where audiences would sit inside an array of speaker, all spectators in this scene are perfectly located in the sweet spot, which is an argument for this type of reproduction. Figure 3 shows the audio pipeline for ambisonic binaural rendering. One of the benefits of ambisonics is that these orthonormal basis functions can easily be transformed using non-destructive linear functions. The most common function is rotation, which allows embedded gyroscopes in mobile devices to send head orientation information to the decoder, which adjusts the playback accordingly. In other words, when the listener moves his or her head, the sound field is rotated to simulate spatial audio using dynamic digital filtering.

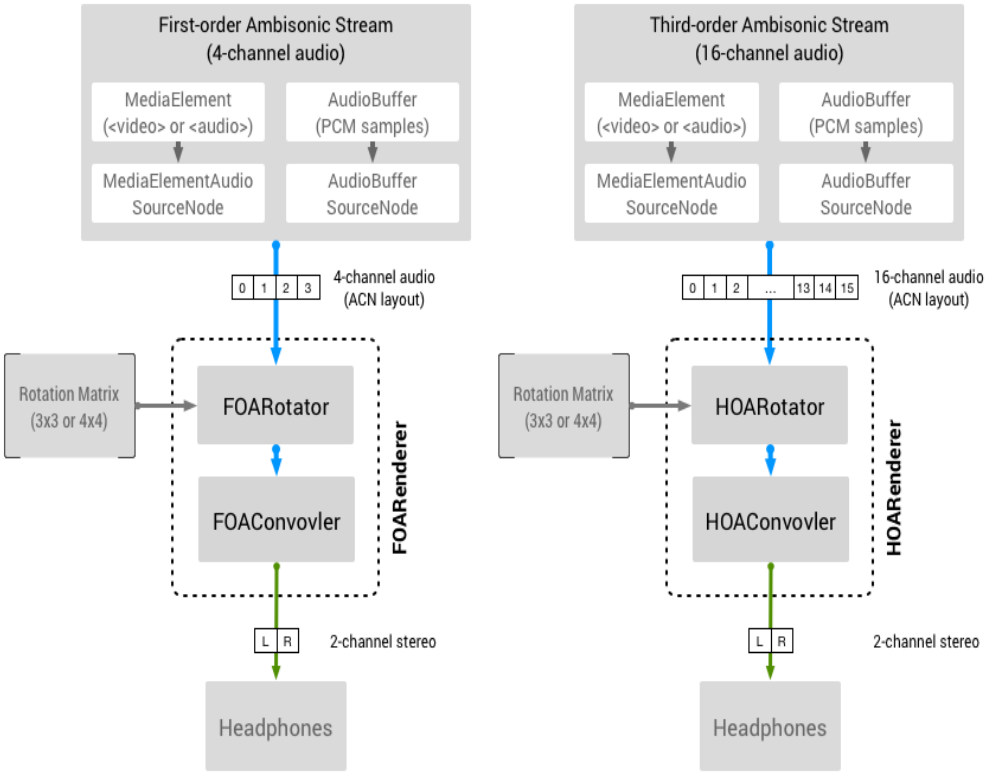

Figure 3: Omnitone Diagram. Image by github.com.

The project was tested using several browsers and mobile phones with various internet connections, all in the San Diego area. While the assets loaded properly with some connections, a longer timeout was needed on other networks. The preload feature of A-Frame has not yet been implemented into our project for HOA. On mobile devices, we were unable to get Omnitone to playback HOA, so alternate versions in FOA and binaural were created to satisfy the needs of each user, based on their connection.

We believe the HOA version on Desktop browsers such as Chrome and Firefox sounds really fantastic, and the virtualization of moving sources is clear and distinct. The vocalists' stems, which are fixed at the front sonically, clearly change in timbre as one changes perspective. In future projects we are interested synchronizing the bimodal elements of the audio-visual experience to further facilitate the localization of sound sources – animations of flying birds for example could be attached to sound sources representing these objects.

5 COMPOSITIONAL PROCESS

"Pigments of Imagination" presents the creative process as a narrative relating the neurological inner-workings of synesthesia to the universe around us through vocals and instrumentation. It is an exploration of the imagination as told through the story of a small child's journey to the moon. The musical aspects of this project include studies of

composition, piano, theory, acoustics, sound-design, audio engineering/production, screenplay drafting, and visual communication by way of 3D texturing and world-building.

Highlights of the music-creation process relative to this project and conference theme include:

- Flugelhorn processed through granular synthesis.
- Vintage synthesizers patched together and captured as a single live performance.
- Live field recording of birds digitally manipulated through granular and wavetable synthesis. Stretching and other processing techniques were used to pitch live recordings to harmonic frequencies. The result was used as a sonic bed for instrumentation and vocals.
- Sounds of the Mars Rover run through granular processing to increase the stereo field of the composition, and create an atmosphere for the music and narrative.
    - All collaborative elements of the song were recorded in environments native to each artist.
- Various forms of delays, reverb and convolution reverb were then used for processing while mixing the song to accentuate musical spaciousness prior to HOA mixing.

Tim Gmeiner has continued working on the POI experience using Unreal Engine. The project showcased here is an early prototype of that larger project focusing on socio-economically accessible technologies and WebVR. The Unreal version of the project can be seen here: https://www.soundrzn.com/school-portfolio-poi.

ACKNOWLEDGMENTS

We would like to thank all the musicians involved in this project, IEM and Google for the FOSS which enabled this work and the DAVAMOT conference for the opportunity to share this experience more broadly.