Audio Mixing in Virtual Reality

Avrosh Kumar; *Department of Music Technology, Georgia Institute of Technology*

Georgia Tech School of Music, College of Architecture

Georgia Tech Center for Music Technology, College of Architecture

Motivation

The interface for mixing and mastering audio is heavily derived from its hardware counterparts like knobs and faders. While the interfaces have become more and more detailed the Digital Audio Workstations (DAW) have also become more and more cluttered. This perhaps raises a question whether we are using the same interfaces designed years ago because they are effective or they haven't yet undergone and major redevelopment.

In this project, Virtual Reality (VR) is proposed as a direction to explore and redefine the process of mixing. Perhaps, a knob can be represented by what it represents and not just by its value as a parameter. The space can be used not just to add elements to control but to describe audio. VR with which provides a wide field of vision may then be used to unclutter the DAW interface.

This project doesn't propose a replacement of a DAW by a Virtual Reality application but an addition to it that an audio engineer can use to get a different perspective on the mix that is not available in the conventional DAW. One the things that conventional DAWs are bad at doing is presenting an overview of the entire mix. It is hard to know which track is panned where, which audio track is taking filling which spectral space in a single view without switching between individual tracks. In VR a song and its tracks can be represented in space giving an overview, for example, a mixing engineer can keep track of what's clipping in a single view.

This application also tries to approach the process of mixing differently from where it is mostly laying tracks and setting fader and knob values. Perhaps, an intuitive space where you place objects in the space and they map themselves to a panning direction and loudness position respective to the listener, may be used in a educational context to teach audio mixing.

Background

The task of mixing audio involves starting with stem files of audio and layering them all together to form a cohesive and 'good' sounding recordings for the consumers and vied variety of speakers that it will be played on. This process involves finding space for each source of audio, accentuate them or make them subtle in how far they seem form the listener's point of view; finding the space in directional sense from where the sounds seem to emanate (panning); and finding the space for each source in the spectral domain (equalizing) so that each source is audible and doesn't build up in the `lows' or the 'highs' or at specific frequencies which can sound annoying. Audio mixing and mastering also involves adding effects and modifying the sounds sources to suit the listeners style or the genre of the song being mixed.

Audio mixing process is carried out on mixing boards with signal chains to affect, route and control audio coming through each channel. This process has been simplified by the use of Digital Audio Workstations (DAW) by replacing a lot of the hardware requirements and create virtual audio channels.

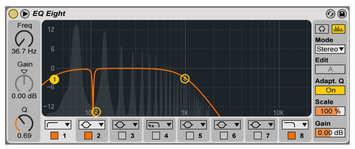

Fig.1 An EQ plugin to see the spectrum of the audio track

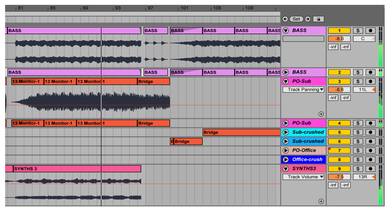

Fig.2 Virtual audio channels represented as tracks in a DAW

Interface Design (I)

Audio has properties that can be described in 3D space. In this representation, each audio track is described by a thin vertical tube in 3D space. The height of the tracks shows the spectrum of the audio. The left and the right edges describe the left and the right of the track if it is stereo and so one can choose to compare both.

In this representation, the spectrum moves in z-axis representing time. One also gets to see the overview of the mix, how the tracks are panned and which track occupies what space in the spectrum; highs or the lows which is useful information for mixing engineer to EQ the tracks.

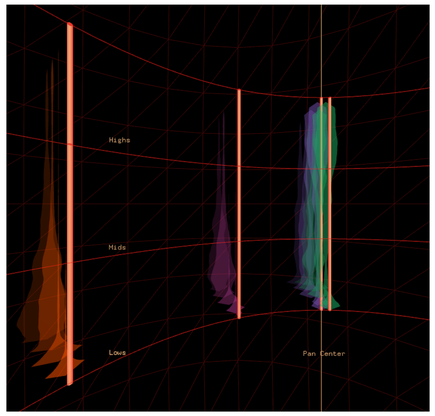

Fig.3 Representation of tracks that show the spectrograms has height of the tubes and move in time over z-axis

In this application, one can also interact and move the objects to change the panning positions of the tracks. Interaction with the EQ, although not implemented in this prototype, should allow the mixings engineer to move points in space to define the EQ parameters.

Interface Design (II)

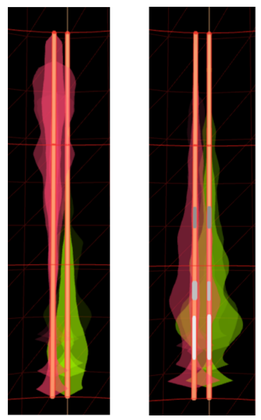

Fig.4 Shows overlaps in spectra of two different tracks.

Another use case where this representation comes handy is while mixing two tracks that build up in the same frequency bands. For example, a kick and bass tend to build up in the similar low frequency bands. This view shows the overlap regions of the two tracks where this build up happens.

This kind of representation is not easily seen in the conventional DAWs. There are some plugins "Voxengo SPAN" that allow the users to see the overlaps but require routing of audio tracks. This representation can be seen as the more intuitive form of looking at the overlaps in the regions while also being presented with the overview of the track.

Visualizing compression, reverb, automation are further problems that were explored in this project. It is not possible to visualize 3rd party plugins since we cannot access their data but a shift in this direction may give rise to VR 3d audio plugins that will allows us to do that

System Description

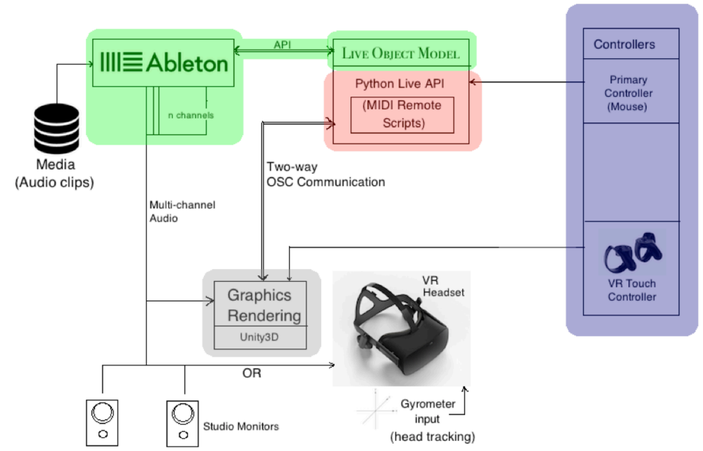

Fig.4 Block diagram describing how Ableton Live can be used to visualize a mix in VR

Object Oriented programming in Python	Live Object Model	openFrameworks (C++)	Unity 3D Engine
MIDI Remote Scripts (Ableton Push Scripts)	**Ableton DAW**	**Graphics**	Oculus LibVR
OSC over network	**Interactions & Communication**	**Audio**	(DirectSound/ NAudio) in Windows
Collision Detection in UI	UserInterface Controls in VR	CoreAudio on MacOS	Virtual Audio drivers

Fig.6 Technical descriptions of the components required for this prototype

Discussion and Conclusion

This prototype maybe used in the educational context as tool to see see and affect an audio mix in the virtual reality. Currently, this prototype seems short of being able to be used in a professional context because it is hard to access multichannel audio data at the rate at which it can be represented accurately.

Recommendations and Future Work

This work was possible due to Ableton's Live API which allows us to access the DAW's controls. In future, as DAWs open up more control to allow access to multichannel audio information of the DAW in should become easier to represent the audio.

The interfaces explored in this project fall short of interactive points and controls. There is lot of potential in exploring and redefining way to show audio parameters more intuitively.

Contact Information

Avrosh Kumar
avroshk@gatech.edu

Georgia Tech