Performance Evaluation of Audio SDK's in Virtual Reality

[1]Vishal Chiluka, [2]Prof. Varshapriya J. N., [3]Ambrish Dhantrey
[1]Mtech(NIMS), [2]Professor, [3]Manager
[1]Computer Department,
[1]Veermata Jijabai Technological Institute, Mumbai, India

***Abstract*— Virtual reality (VR) is a real-time and interactive technology. It means that the computer can detect user inputs and modify the virtual world instantaneously. Interactivity and its captivating power contribute to the feeling of being the part of the action on the environment that the user. Sound in Virtual Reality can be used to ease people into the experience, what we also call "on boarding." It can be used subtly and subconsciously to guide viewers by motivating them to look in a specific direction of the virtual world, which surrounds them. Various audio SDK's and sdks for enhancing realism and immersive experience of audio in virtual reality are available and various parameters are taken into consideration while judging the performance of any SDK. The proposed system is then used to monitor and evaluate an in-depth study of 3 different audio sdks which are available for virtual reality i.e. NVIDIA's VrWorks audio, steam's photon audio sdk and oculus audio sdk.**

***Keywords*— Virtual Reality, Oculus Rift, HTC Vive, Performance, FPS, Audio/Sound in VR, SDK's**

I. INTRODUCTION

Virtual reality (VR) is a computer technology that uses Virtual reality headsets, sometimes in combination with physical spaces or multi-projected environments, to generate realistic images, sounds and other sensations that simulate a user's physical presence in a virtual or imaginary environment. A person using virtual reality equipment is able to "look around" the artificial world, and with high quality VR move about in it and interact with virtual features or items. VR headsets are head-mounted goggles with a screen in front of the eyes. Programs may include audio and sounds through speakers or headphones.

Oculus Rift, HTC Vive, Google Daydream, PlayStation VR and many more are the Virtual Reality headsets available nowadays which are often referred to as HMDs (Head mounted displays). Even with no audio or hand tracking, holding up Google Cardboard to place your smartphone's display in front of your face can be enough to get you half-immersed in a virtual world.

Virtual reality demands a new way of thinking about audio processing. In the many years of history of games and video on at screens, the standard of realism for audio rendering has remained relatively low, especially when compared against contemporaneous advances in graphics and cinematic video rendering. Although hearing is inherently a 3-dimensional sense, audio for at screen games and cinema/video has usually minimized the use of 3D and other advanced audio rendering technologies, for the simple reason that all of the graphics and video is in front of you. With a 2D screen in front of you, if you hear a sound of a voice or an object behind you and turn toward it, all you will see is perhaps a speaker or the walls of your apartment.

With the exception of some FPS games where 3D audio provides a tactical assist that players learn to use to their advantage, too much realism in audio for a at screen game or movie can sometimes be a distraction, especially if it is inconsistent with the visual experience. A head-mounted display changes everything. Now the user can turn in any direction and see a continuous visual scene, and with recent advances, walk independently through a virtual world. In virtual reality applications, according to the visual real-time graphics, proper audio must be rendered at right time. Today's conventional audio basically works. It could be positional but its directional audio comes directly from whatever that direction is, but it doesn't interact with the environment.

II. RELATED WORK

Virtual reality has been notoriously difficult to define over the years. Many people take "virtual" to mean fake or unreal, and "reality" to refer to the real world. This results in an oxymoron. The actual definition of virtual, however, is "to have the effect of being such without actually being such". The definition of "reality" is "the property of being real", and one of the definitions of "real" is "to have concrete existence". [4][5]

Using these definitions "virtual reality" means "to have the effect of concrete existence without actually having concrete existence", which is exactly the effect achieved in a good virtual reality system. There is no requirement that the virtual environment match the real world. Inspired by these considerations, for the virtual wind tunnel, the following definition is adapted: "Virtual reality is the use of computer technology to create the effect of an interactive three-dimensional world in which the objects have a sense of spatial presence."[5][7]

In this definition, "spatial presence" means that the objects in the environment effectively have a location in three-dimensional space relative to and independent of your position. Note that this is an effect, not an illusion. The basic idea is to present the correct cues to your perceptual and cognitive system so that your brain interprets those cues as objects "out there" in the three-dimensional world. These cues have been surprisingly simple to provide using computer graphics: simply render a three-dimensional object (in stereo) from a point of view which matches the positions of your eyes as you move about. [1][3]

If the objects in the environment interact with you then the effect of spatial presence is greatly heightened. Note also that we do not require that the virtual reality experience be "immersive". While for some applications the sense of immersion is highly desirable, we do not feel that it is required for virtual reality. The main point of virtual reality, and the primary difference between conventional three-dimensional computer graphics and virtual reality is that in virtual reality you are working with: “Things as opposed to Pictures of Things.” [9]

There are various audio sdks available which improvise and enhance audio effects in virtual reality to give realism and immersion effects to the user. Various parameters are considered while developing these audio sdks. Hence there is a scope to do a comparative performance analysis of these audio sdks and evaluate them on what basis, they are good at. [11]

A software development kit (SDK or devkit) is typically a set of software development tools that allows the creation of applications for a certain software package, software framework, hardware platform, computer system, video game console, operating system, or similar development platform. To enrich applications with advanced functionalities, advertisements, push notifications and more, most app developers implement specific software development kits. Some SDKs are critical for developing a platform-specific app. For example, the development of an Android app on Java platform requires a Java Development Kit, for iOS apps the iOS SDK, and for Universal Windows Platform the .NET Framework SDK. There are also SDKs that are installed in apps to provide analytics and data about activity. Prominent examples include Google, InMobi and Facebook. [16]

It may be something as simple as the implementation of one or more application programming interfaces (APIs) in the form of some libraries to interface to a programming language or to include sophisticated hardware that can communicate with a particular embedded system. Common tools include debugging facilities and other utilities, often presented in an integrated development environment (IDE). SDKs also frequently include sample code and supporting technical notes or other supporting documentation to help clarify points made by the primary reference material. [11]

Audio sdks include api’s which are directly imported to the application or game code and integrated to have features like sound Spatialization, sound reverbs and occlusion. Various audio sdks are developed based on various parameters which bring realism and immersiveness in the application when integrated with the game or application.

Key Terminologies with audio in Virtual Reality:

- **Audio/Audio Signal:** An audio signal is a representation of sound, typically as an electrical voltage. Audio signals have frequencies in the audio frequency range of roughly 20 to 20,000 Hz (the limits of human hearing). Audio signals may be synthesized directly, or may originate at a transducer such as a microphone, musical instrument pickup, phonograph cartridge, or tape head. Loudspeakers or headphones convert an electrical audio signal into sound. [7]
- **Virtual Reality:** Virtual reality (VR) is a computer technology that uses Virtual reality headsets or multi-projected environments, sometimes in combination with physical spaces, to generate realistic images, sounds and other sensations that simulate a user's physical presence in a virtual or imaginary environment. [5]
- **Spatialization:** Spatial music is composed music that intentionally exploits sound localization. Though present in Western music from biblical times in the form of the antiphon, as a component specific to new musical techniques the concept of spatial music. [6]
- **Reverb:** Reverberation is the persistence of sound after a sound is produced. A reverberation, or reverb, is created when a sound or signal is reflected causing many reflections to build up and then decay as the sound is absorbed by the surfaces of objects in the space – which could include furniture, people, and air. This is most noticeable when the sound source stops but the reflections continue, decreasing in amplitude, until they reach zero amplitude. [8]
- **Occlusion:** It’s the obstruction of sound when sound source is behind or covered by some material or object and there is no direct line pf path in between source and receiver. Sound occlusion can be felt when sound source of receiver moves and there is an objection in the path. [9]
- **Path Tracing:** Path tracing is a computer graphics Monte Carlo method of rendering images of three-dimensional scenes such that the global illumination is faithful to reality. Fundamentally, the algorithm is integrating over all the illuminance arriving to a single point on the surface of an object. [11]
- **Direct Path:** Direct path deals with no obstruction in between sound source and receiver. Sound waves are travelled directly from source to receiver. [12]
- **Indirect Path:** Indirect path deals with obstruction in between sound source and receiver. Sound waves are travelled by reflecting from different directions from source to receiver. [12]
- **Attenuation:** Attenuation is a telecommunications term that refers to a reduction in signal strength commonly occurring while transmitting analog or digital signals over long distances. [12]
- **HRTF (Head Related Transfer Function):** A head-related transfer function (HRTF) is a response that characterizes how an ear receives a sound from a point in space. As sound strikes the listener, the size and shape of the head, ears, ear canal, density of the head, size and shape of nasal and oral cavities, all transform the sound and affect how it is perceived, boosting some frequencies and attenuating others. The HRTF boosts frequencies from 2 - 5 kHz with a primary resonance of +17 dB at 2,700 Hz. But the response curve is more complex than a single bump, affects a broad frequency spectrum, and varies significantly from person to person. [13]
- **Material Reflection Co-efficient:** Acoustic Reflection refers to the process by which a material, structure, or object reflects out sound energy when sound waves are encountered, as opposed to absorbing the energy. Part of the energy is transformed into heat and part is transmitted through the absorbing body, rest is reflected to the space. [15]

- **Material Absorption Co-efficient:** Acoustic absorption refers to the process by which a material, structure, or object takes in sound energy when sound waves are encountered, as opposed to reflecting the energy. Part of the absorbed energy is transformed into heat and part is transmitted through the absorbing body. The energy transformed into heat is said to have been 'lost'. [15]
- **Material Transmission Co-efficient:** Acoustic transmission is the transmission of sounds through and between materials, including air, wall, and musical instruments. The degree to which sound is transferred between two materials depends on how well their acoustical impedances match. [15]

Game Engine – Audio Plugin SDK interaction:

Game Engine interacts with Audio plugin and sends audio meta-data to the audio plugin. Based on the meta-data different filters are generated which in turn are applied to sound sources in the game. Following are the steps which describe how audio plugin interacts with game engine: [14]

Step 1: Game interacts with game engine by sending and receiving data (e.g. audio, video, input).
Step 2: Audio plugin takes all the information about the audio i.e. audio source id, audio geometry, audio source position and orientation etc. and generate filters which gets all the effects (e.g. Spatialization, occlusion, reverbs etc.).
Step 3: Various effects are also supported by game engine can be applied based on user settings or can be overridden by plugin's effects.
Step 4: All audio data is passed through the filters which applies all effects to the audio.
Step 5: Wet Audio (i.e. Audio with all effects) is then passed to sound system for playback on headphones or speaker.

III. PROPOSED SYSTEM

Performance analysis of audio SDK's in virtual reality is done based on different parameters. Unreal Engine is a game development engine by epic games. A Game/Application (e.g. unreal tournament) can be used as a sample application for measurement of various parameters which determine their performance in virtual reality environment. Analysis of 3 different SDK's (NVIDIA's VrWorks Audio SDK. Phonon Steam Audio SDK and Oculus Audio SDK) are considered. These 3 SDK's are integrated into unreal tournament in unreal engine.

The parameters which are considered for performance analysis are as follows:

- **Frames Rate:** Indicates the current number of frames being rendered per second and the amount of time taken to render the frame in milliseconds (ms).
- **CPU Usage:** Amount % of CPU in use
- **GPU Usage:** Amount % of GPU in use
- **CPU cycles:** Total number of CPU cycles taken
- **Audio Update Time:** Updating overall time taken by the audio in Audio Device.
- **Gathering Wave Instances:** Collecting, sorting and prioritizing all potential audio sources.
- **Processing Sources:** Starting and/or updating any audio sources that have a high enough priority to play
- **Submit Buffers:** Number of times Submitting audio buffers to the system.
- **Source Create:** Number of times sound sources are created.
- **Source Initialization:** Number of times sound sources are initialized.
- **Buffer Creation:** Number of times audio buffers are created
- **Updating Sources:** Updating the sound sources (position, velocity, volume, pitch, etc.).
- **Updating Effects:** Number of traces and number of times filters are generated and audio effects (reverb, LPF, EQ) are applied.
- **Prepare Audio Decompression:** Initializing of sound decompression in general (currently, only vorbis data).
- **Decompress Vorbis:** Decompressing vorbis data.
- **Audio Memory Used:** Total memory used for all sound assets. Does not include any workspace or internal buffers.
- **Audio Buffer Time:** Total time of stored sound data in seconds.
- **Audio Buffer Time (w/ Channels):** Total time of stored sound data channels in seconds. Each channel of a multi-channel sound is added up to get this value.
- **Audio Components:** Total active sounds being maintained for processing (ambient sounds all over the map, player sounds etc).
- **Audio Sources:** Total audio components that are audible and high enough priority to hear.
- **Wave Instances:** Total sound waves being processed.
- **Wave Instances Dropped:** Total sounds not allocated to a sound source (as they are inaudible).
- **Audible Wave Instances Dropped:** Total sounds not being played but were audible.

IV. CONCLUSION

Performance evaluation of different audio sdks will lead to a comparative analysis of the same based on different parameters to judge which sdk is more suitable in which environment for appropriate decision making. Audio in virtual reality is an emerging domain and hence a comparison has to be carried out in order to classify and portray the features of these SDK's to improve realism and immersiveness of audio in virtual reality. These audio sdks are developed using different functions and values which needs to be compared for better analysis and parameters are considered in such a way which analyze evaluation effectively.

REFERENCES

[1] Bastug, E., Bennis, M., Médard, M. and Debbah, M., 2017. Toward Interconnected Virtual Reality: Opportunities, Challenges, and Enablers. IEEE Communications Magazine, 55(6), pp.110-117
[2] Mazuryk, T. and Gervautz, M., 1996. Virtual reality-history, applications, technology and future.
[3] NVM Express Inc., “NVM-Express Explained Whitepaper”, http://www.nvmexpress.org/wp-content/uploads/2013/04/NVM-whitepaper.pdf
[4] Audio in Virtual Reality (http://www.audiomediainternational.com/news/how-to-producing-immersive-audio-for-vr/06432)
[5] Begault, D.R. and Trejo, L.J., 2000. 3-D sound for virtual reality and multimedia.
[6] Importance of audio in virtual reality (http://www.immerseuk.org/newblog/2017/7/19/the-importance-of-audio-in-augmented-and-virtual-reality)
[7] Youngblut, C., 1998. Educational Uses of Virtual Reality Technology (No. IDA-D-2128). INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA VA.
[8] Trends in Virtual Reality (http://www.androidauthority.com/virtual-reality-work-702049/)
[9] Design Infrastructure for Measurement & Analysis of Parameters that affect VR Performance & User Experience (https://www.ijircce.com/upload/2016/december/19_355_Design.pdf)
[10] Frame Rate (https://en.wikipedia.org/wiki/Frame_rate)
[11] Principles of Designing in VR (http://www.uxness.in/2015/08/ux-virtual reality.html)
[12] Steam Audio (https://valvesoftware.github.io/steam-audio/)
[13] Oculus Audio (https://developer.oculus.com/documentation/audiosdk/latest/concepts/book-audio-intro/)
[14] NVIDIA's VrWorks Audio (https://developer.nvidia.com/vrworks-audio-sdk-depth)
[15] Unreal Engine (https://en.wikipedia.org/wiki/Unreal_Engine)
[16] Software Development Toolkit (https://en.wikipedia.org/wiki/Software_development_kit)