Virtual Reality and Auditory Environments

Akarsha Sehwag
Roll no. 2015010

***Abstract*—In this Technical Report, I will present an overview of the technologies used to create a virtual environment and how auditory environments can be set up using similar tools and technologies as that of Graphics Rendering.**

***Index Terms*—Virtual Environment, Scientific Visualization, Virtual Reality**

I. Introduction

Virtual Reality or VR is a computer simulated 3-D environment which provides the user with visual experiences and in some cases, additional sensory information; involving the incorporation of 3D sound, smell and force-feedback (provides touch sensation). The term Virtual Reality was first used by Founder of VPL Research, Jaron Lanier. Initially it referred to Immersive Virtual Reality, in which the user gets fully-immersed into the artificial world. In full-immersion, user is made to wear a head mounted display to experience 3D audio and visual information. Today, Virtual reality is used in a variety of ways; including partial immersion levels by projecting computerized visual on large monitors and head-up displays.

In this report, Ill introduce you to Virtual Reality and further talk in detail about Virtual Auditory Environments. Lets start with the basic maths behind VR displays, Stereographic Projections.

II. Technology Used

A. Stereographic Projection

By definition: Stereographic projection is a particular mapping (function) that projects a sphere onto a plane. It provides the extension of field of view coupled with stereographic display to give us the perception of 3-D depth of a particular object in a scene.

Two basic types of depth perception cues used by eye-brain system are: Monocular and Binocular; signifies if they are apparent when one or two eyes are used. Main monocular cues are: Perspective; overlaps (closer objects overlap the far ones); Movement parallax (closer objects appear to move faster than far objects); Relative size of objects; shadows and highlights; ability of the eye to resolve fine details in near objects but not in the far ones. And main binocular cues are: Retinal disparity(difference in distance from the eye). Monocular cues create weak perceptions of 3-D depth whereas the binocular cues create strong 3D depth perceptions since the brain combines two separate images into a single image.

Stereographic images have characteristics similar to true binocular vision. There are two methods which use filters to insure reception of different images (by right and left eye) one of which uses polarizing filters and other uses color filters to view the two images created by both the eyes where the left eye can only see the left image and right eye can only see right image; Its then combined into a single 3-D image by the eye-brain system. Both require projection of a single point `P` on a plane `z=0` from two centres of projection(right and left eye) at $E_L(-e,0,d_e)$ and $E_R(e,0,d_e)$. Figure 1 shows the translated projection of left eye so that it lies on the z-axis.

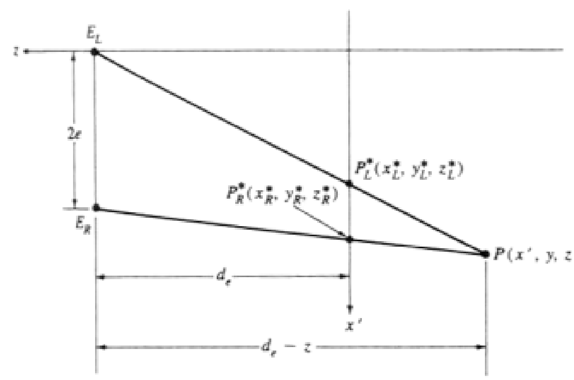

Fig. 1. Stereo Projection for Left Eye

Using triangle properties, we can infer the following:

$$\frac{{x_{\mathrm{L}}}^{*'}}{d_{\mathrm{e}}} = \frac{x'}{d_{\mathrm{e}} - z} \tag{1}$$

Figure 2 shows the translated projection of right eye to make it lie on z-axis.

Similar inferences can be taken for the stereo projection for the right eye.

$$\frac{{x_{\mathrm{R}}}^{*''}}{d_{\mathrm{e}}} = \frac{x''}{d_{\mathrm{e}} - z} \tag{2}$$

The transformations are given by the following matrices:

$$[S_{\mathrm{L}}] = [Tr_{\mathrm{E_L}}].[P_{\mathrm{rz}}] \tag{3}$$

and

$$[S_{\mathrm{R}}] = [Tr_{\mathrm{E_R}}].[P_{\mathrm{rz}}] \tag{4}$$

Where $\mathrm{Tr}_{\mathrm{E_L}}$ and $\mathrm{Tr}_{\mathrm{E_R}}$ are translations of centre of projections for the left eye and the right eye respectively. We can obtain a stereographic projection by transforming the scene using equation 3 and equation 4 displaying the two images.

III. VR SYSTEMS

- **CAVE:** It stands for CAVE Automatic Virtual Environment; a scientific visualization system. User needs to wear a LCD shutter glasses and is surrounded by video projection displays along with head tracking and it projects stereographic images on the walls of the room. It has superior resolution and wider field of view unlike HMDs. CAVE libraries automatically generate the user centred perspective for each screen.

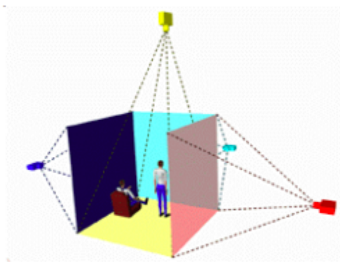

Fig. 2. CAVE

- **ImmersaDesk:** Its a drafting table format VR display. This also uses shutter glasses to get accurate stereoscopic images. Its logistically simpler and smaller to CAVE. And applications developed for CAVE can run on ImmersaDesk without any code changes.
- **Infinity wall:** A larger system was designed around basic resources as CAVE, but its intended for a large group of people; as in a classroom setting. It comprises of a 9x12 foot screen with one or two SGI onyxes to drive the display. Shutter glasses are used here also as in the other two technologies. Here, the stereo video formats are of lower pixel resolution than monoscopic video formats.

IV. ARCHITECTURE OF VR SYSTEMS

The VR system includes an input processor, world database, simulation processor and rendering processor. Input processor controls the devices to input the information. It inputs the coordinate data to the system without any lag; Simulation processor is the process of imitating real things in an virtual environment. It takes the user inputs and decide the actions that will take place in the virtual reality; Next comes in Rendering processor produces the sensation that are seen by the user. Different rendering processors are used for visual, auditory and other sensory systems. Each renderer derive the world stats from the world database for each time-step. The world database contains the objects that user will see virtually; like if we want to experience a classroom, the world database must have table, chair, board etc.

We'll shed some more light on how we incorporate sound in Virtual Environments. It's analogous to Graphics Rendering when creating virtual auditory environments. This is just an extension of Image rendering to show how other sensory environments can be created.

V. VIRTUAL AUDITORY ENVIRONMENTS

In images or graphics, we calculate light distribution to create an image. Methods such as radiosity and ray tracing are concepts taken from physics of light. On the same pattern, sound rendering is based on laws of sound reflection and propogation. In this section, we create a mapping between sound rendering and visual image rendering to describe how to perform sound rendering based on source and destination knowledge. Auralization – making audible – means visualization in sound rendering.

A. Approaches

We use methods such that they enable dynamic rendering, in which the positions of the objects are not constant and are subject to change during the rendering process; it would be difficult to process though, but has a wider scope of applicability. eg. virtual concert recordings. In most applications, User Interaction is an essential feature and thus dynamic rendering must be performed in real-time.

Based on physics of sound or human perceptions, we can add cues like sound source location to the signal. Mostly, we take original sound as the essential content and rendering is only to add special features as effects. For this, a 3D model is not required but can use parameters such as brilliance to produce a desired auditory sensation. This approach is more suited for postprocessing music performances where we need to create room acoustic effects; physics based approach is preferred when we need the acoustic environment rather than the sound.

B. Interactive Sound Rendering

For sound rendering we need to collect information of sound propogation from source to desired destination in impulse response containing information about the radiation and room's reverberation. Best way is through convolution with stimulus signal which is generally anechoic sound – free from reverberation and echo – this way, we add information to sound.

While dynamically rendering sound, the impulse response in convolution is parameterized to save computational load. Here we'll write a three-level approach to parameterization as follows:

☐ **Defining a scene**
Acoustic model of a scene contains 3D geometry. Polygonal modelling is sufficient where each polygon is associated with a material defined by its absorption and diffusion coefficients; they depend upon direction and frequency. Apart from this, the model contains orientation, direction and location of the source as well as the listener; both of them have directivity characterstics.

☐ **Real-time accoustic simulation**
There are several ways – ray-based and wave-based – to calculate sound propogation and reflections in space.

Ray-based methods are generally used for real-time simulations whereas the image-source methods are better suited for calculating early reflections. People tend to assume that later reverberation is direction independent and exponentially decaying; this is to let us use the computationally effiient recursive structures which can be controlled by statistical late reverberation parameters such as reverberation time and energy.
Reflecting process can be used recursively to find higher order image sources. Image source method is ray-based in which every reflection path can be replaced by image source. In a room, most of the higher-order image sources are invalid since the reflection path is not realizable and some of them are invisible due to occlusion. Thus, the reflection path through all the reflecting surfaces is reconstructed in the model. Then the occlusions are calculated and checked if all the reflecting points lie in the surface boundaries.

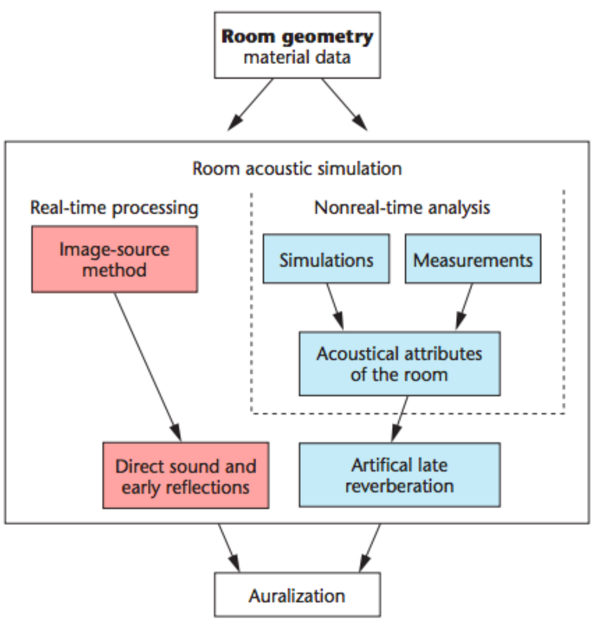

Fig. 3. Parameterization levels and hybrid method for sound rendering

For image-source calculation, the input data is the room geometry and material data. Based on the orientation of sound sources, we can calculate information of audible specular reflection. The parameters of image source are as follows: - orientation of sound source
- order of reflection
- distance from listener
- incoming direction of sound.
As in figure 3, we calculate the parameters of late reverberation offline and lets us tune the reverberation rime and other features. In our model, we don't construct the impulse response for convolution–auralization–process. Rather we use a signal processing and with the image source method, we calculate the parameters.

☐ **Auralization and signal processing**
Auralization parameters provided by image source calculation are converted to signal-processing parameters. This is because the auralization parameters are not updated for every single audio. So for efficiency, we take up the signal processing parameters from the pre-calculated tables or created them using interpolation.

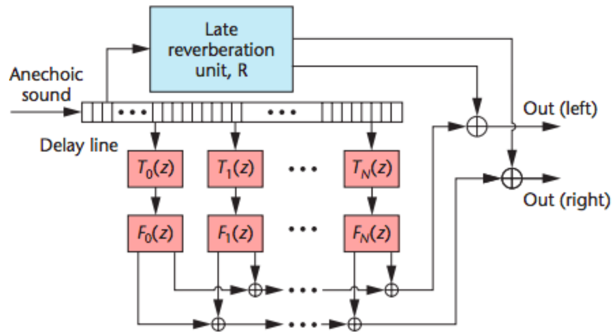

Fig. 4. Auralization signal processing structure

Final Auralization process is implemented as a signal processing structure as given in Figure 4. These are pre-calculated and stored in data structures for easy accessibility. Figure 4 shows a long delay line in which anechoic sound is fed. The distance of the image source defines the point to the filter blocks $T_{0..N}(z)$ where N is the Image source ID. These blocks modify the sound signal with distance dependent gains and other material filters. These outputs are then superimposed to finally sum up with the outputs of a complex recursive algorithm R.

This section was to gain an insight about how audio can be rendered in a similar pattern as the image rendering. Now we move on to the applications of Virtual Reality in our lives.

VI. Applications of VR

It is desirable that Immersed Virtual reality should allow applications and users(networked around the world) to share a common virtual world and interact with it. The devices help us achieve that level of sophistication. Common applications of Virtual Reality are:

- **Architecture Walkthrough:** Main idea is to explore the already designed but not-yet-constructed buildings to get a visualization of the building and thus allow the architect to prototype the target and create a walk-through the virtual building.
- **Medical:** It's used for training and education and surgical simulation for diagnosis. It's also been widely used to treat phobias and post traumatic stress disorder.
- **Simulation and Training:** It's one of the most common applications of interactive virtual Simulation. It can reduce the training costs and also create a safe virtual environment. Eg. Driving or flying simulations. They

Fig. 5. VR used in military training

provide user expertise of driving a vehicle without the consequences of making a mistake.

- **Engineering:** It's used in collaborative design and engineering, maintainance analysis, sales, marketing, education, digital prototyping etc to improve efficiency and user-experience.
- **Scientific Data Analysis:** Used to visualize data coming from simulations in fluid mechanics, molecular modelling etc. It's a virtual world which represents the generated data during the simulation for better analysis of the data. Virtual Wind tunnel is a nice example of pipeline which converts data into image for scientific reasons.

VII. Challenges and Future work

Psychologists believe that these immersions in VR can possibly affect a user psychologically. Major challenges in this field is to improve the tracking systems, and finding better and more natural ways for user to interact with Virtual Environment. Other than these, we need to decrease the time needed to build convincing virtual spaces. It currently takes a lot of time to build a realistic environment. It can take an year for a team of programmers to duplicate a real room with accuracy.

Also, it is important to create a Virtual System which avoids bad ergonomics. Most systems rely on hardware that restricts the user by physical means. Without well designed hardware, the user can experience trouble like cyber-sickness, with symptoms like nausea and disorientation.

In future well be using virtual phones to communicate. Japan Phone companies are developing a system to allow a person to see the 3-D image of the other person. Its future depends upon how we address the scaling factor in Virtual environments. So as the computers process faster, theyll be able to create better simulations. Thus, the benefits will remain immeasurable in future.

References

[1] Gilson Giraldi, Rodrigo Silva, Juavane C de Oliveira ”Introduction to Virtual Reality” , LNCC Research Report 06/2003, National Laboratory for Scientific Computation, 2003

[2] Sharmistha Mandal ”Brief introduction to Virtual Reality and its challenges” International Journal of Scientific Engineering Research, Volume 4, Issue 4, April-2013 304 ISSN 2229-5518

[3] T.Lokki,L. Savioja, R.Vaananen ”Creating interactive virtual auditory environments”, IEEE Computer Graphics and Applications (Volume: 22, Issue: 4, Jul/Aug 2002)

[4] https://www.slideshare.net/Sumitsharma443/virtual-reality-report