Pontificia Universidad Javeriana Cali
Facultad de Ingeniería.
Ingeniería de Sistemas y Computación.
Anteproyecto de Grado.

Accurate spatialization of sound sources in video games

Juan Camilo Arévalo Arboleda

Director: Dr. Jullian Villegas

May 30, 2017

Santiago de Cali, May 30, 2017.

Sirs
Pontificia Universidad Javeriana Cali.
Dr. Andres Navarro Newball
Director of the Systems Engineering and Computation program.
Cali.

I hereby inform you that the Ssystems Engineering and Computation student Juan Camilo Arévalo Arboleda (cod: 199384) works under my supervision in the thesis titled "Accurate spatialization of sound sources in video games".

Sincerely,

Dr. Jullian Villegas

Santiago de Cali, May 30, 2017.

Sirs
Pontificia Universidad Javeriana Cali.
Dr. Andres Navarro Newball
Director of the Systems Engineering and Computation program.
Cali.

Hereby I present the preliminary draft of the thesis called "Accurate spatialization of sound sources in video games" in order to meet the necessary requirements of finalizing my engineering degree in Systems Engineering and Computation.

By signing here, I give my faith that I undestand and know the guidelines for presentetion of a thesis, approved by the Engineering faculty on november 26^{th}, 2009. Said rules stipulate the deadlines and norms for the development of the project and final thesis.

Sincerely,

Juan Camilo Arévalo Arboleda
Code: 199384

Abstract

Tis project aims at connect systems of spatialisation for three-dimensiona spatialization of sounds sources on Pure Data with a video game engine called Unity. The spatializers that are actually used in video games cannot be accurate in location of sources sounds, this spatializers propose here, focus in veracity of virtual sound in near field (ergo, distances equals or minor that stature of user), and in three-dimensions movement (azimut, elevation and distance). ll ofa this with the propose of giving better experiences to users in a virtual reality environments.
Keywords: Audio Spatializer, Binaural hearing, Unity, Pure Data, Framework, Video games.

Contents

Introduction

Chapter 1

Description of the problem

1.1 Problem statement

The concept of human-computer interaction(HCI) refers to different methods that humans can use to exchange information with computer system. These methods may involve the use of different divices such as display, joysticks, keyboard, tocuh-sensitive displays, etc. In the area of video games, It is becoming a fundamental challenge to attain a greater interaction with the enviroment. This goal is being achieved by using multimedia content in the form of so called virtual reality(VR), which seeks to use the previously mentioned devices to provide implicit interaction with the environment, as well as sensory immersion in real time.

For example, during the development of the so called "first generation" of VR, 3D visualisation was greatly developed. One instance includes a mountain climbing game in first person, which uses the Oculus Rift as an interface for immersion. The Oculus Rift allow the user to have a 3D view of the environment, depending on the inclination and direction of the head of the user. In the game, the player is equipped with two ice axes and has to climb to the highest height of different peaks in each continent [DPC+14]. So far, although video games have focused on making the visual part as reliable as possible, true immersion requires the encompassing of all the different senses of human beings.

In addition to visual content, audio is another important factor in virtual immersion, since accompanied by a visual component, it can give a better spatialization of the user in virtual world. In audio, the most used to generate a immersion scene in sound as for cinema and video games, is surround sound 7.1, which is based on giving a sound quality of an audio source with additional channels coming of devices that emit sound in a horizontal plane with a radius of 360. In the case of cinema, these additional sound channels are adding in editing of movie and are constant, unlike a VR scenario, these additional channels must be dynamic and respond to the interactions that the user has in the virtual environment. In addition, surround systems require specialized hardware to be able to have sound reliability. In video games there is software that emulates 7.1 surround sound, such as the Razer Surround Personalized 7.1 Gaming Audio Software [1].

Although, it is possible to emulate or have surround sound 7.1, the result is not precise enough to give the exact location of an element in the virtual world, in other words, it cannot define the elevation, distance and azimuth from which the source of sound with accuracy, all of this in order to

[1] https://www.razerzone.com/surround

obtain of a true 3D spatialization of audio. However, There are some 3D audio spatializers such as the Oculus spatializer plugin for digital audio workstation and earplug (developed in Pure Data), which we will discuss later. Video game lack such spatializers and the technologies they use to emulate them fail to provide an accurate spatialization. Therefore, in order to endow video games with with this auditory quality, it is necessary to include audio spatialization in video game engines, making it easier for developers to implement such characteristics in games. The previous, paired with a vastly developed graphic part, will enrich the user's experience in terms of interaction and immersion.

1.1.1 Formulation

How to integrate a three-dimensional location system for audio sources to a video game engine, aiming to achieve a better immersion in virtual environments?

1.1.2 Systematization

- How to include the functions of a sound spatialiser in a video game engine?
- How communicate the necessary elements for the spatialisation sound generated by objects in the video game engine to a spatial software implemented in Pure Data?
- How to evaluate the results of the sound spatialisation in a video game to verify the correct functioning of the spatialization and the user's immersion?

1.2 Objectives

1.2.1 General objectives

Develop an interface to use a 3D sound spatialization in a video game engine to improve immersion in a virtual scene.

1.2.2 Specific objectives

(a) Analyse how to incorporate a 3D sound spatial tool into a video game engine.

(b) Analyse, design and implement a framework that allows the spatialization of sound sources considering a listener in a video game engine.

(c) Develop a functional prototype with spatial sound integration tool into a video games engine using a framework developed.

(d) Perform evaluations of the implemented system with users and compare them with conventional sound spatialization systems.

1.3 Justification

A VR environment requires various components and technologies to achieve an effective immersion between the user and the virtual scene. In this case, the intention is to emphasize the importance of sound in virtual environments to improve the VR experience, since in the past, sound has often been a secondary consideration in visually intensive enviroments, such as VR systems and computer games[MN10] , unlike their visual counterparts, which have been extensively studied.

Considering that in video game engines there are different implementations of audio spatializers, to achieve a better sound immersion, we require additional devices, or do not focus on the veracity of sound. This proposed project that will be helpful for the developer or researcher providing the tools to be able to achieve spatialisation of sound sources in a three-dimensional space, in order to get the user to have a greater immersion in the scenarios modeled by a video game engine.

1.4 Scope

Since the aim is primarily to integrate three-dimensional space sound spatialization in video game engine, the project will focus on the perception that the user would have of the location of the sound with respect to the orientation and position of his head in The virtual world. Likewise, the project does not seek to focus on the location of sound sources or the directionality of sound in the virtual world, nor does it simulate the medium in which sound sources are found in the virtual environment, ie emulate the reverberation, object dimensionality (length, width, depth), etc.

To use well-known technologies with a considerable amount of users, the selected video game engine for this project will be Unity [CBM08]. Also, as mentioned before, the audio spatialization tool used in the project has been developed in Pure data. Integration with a VR systems is out of the scope of this project.

1.4.1 Deliverables

- Prototype framework for Unity with which you can spatialize 3D audio source in a game.
- Publication with the results obtained from the research.
- Manuscript that reports our findings.

Chapter 2

Project development

2.1 State of the art

2.1.1 Thematic areas

- H.5.1 Multimedia Information Systems
- I.3.7 Three Dimensional Graphics and Realism
- D.2.2 Design Tools and Techniques

2.1.2 Theoretical framework

VR is an immersive multi-sensory experience, it is the illusion of participation in a synthetic environment rather than external observation of such an environment, VR is also referred to as virtual environments, virtual worlds or micro worlds. To achieve this, VR relies on three-dimensional space scope, stereoscopic, head tracked, hand/body tracking and binaural sound[Ear14]. Binaural sond is the auditory illusion perceived when two waves are presented to a listener in a dichotic manner, that is, one to each of the ears[GA14], and in way to obtain this in VR, is necessary a system spatialization of 3D sound, that the idealistic goal of this system designer ought to involve the notion of complete control and manipulation of the spatial auditory perception of a listener using processes that either complement or replace spatial attributes that existed originally in association with a given sound source[BT00].

The taxonomy of spatial perception involves an egocentric frame of reference, and measurements and orientation of sound images are given from the listener position in polar coordinates about direction and position of listener head[Mil01], that show in Fig.2.1, given as follow:

- **Azimuth:** Is the angle described in degrees ($0^o - 360^o$), from direction between the position of listener face and the position of sound source in a circular horizontal plane around the listener head.
  This perception is very important, since human ears are located at almost opposite position on either side of the head, favoring hearing of the relative angle of sound sources on a plane parallel to the surface of the ground[BT00].

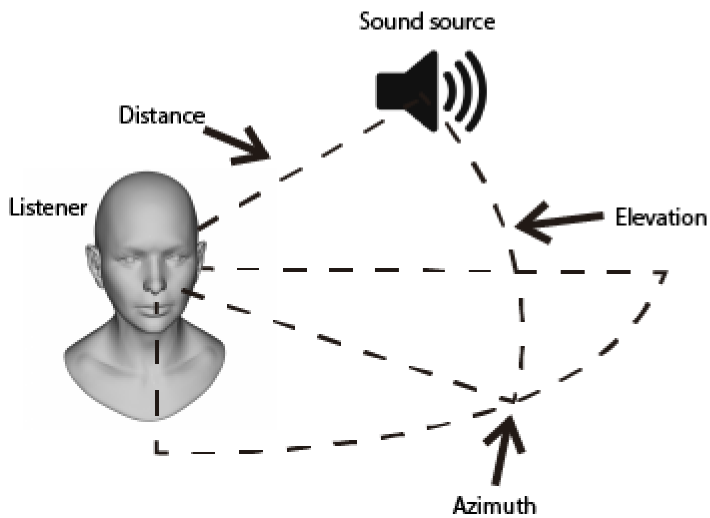

Figure 2.1: Taxonomy of spatial sound in a three-dimensional scope with measurements in polar coordinates.

- **Elevation:** Is the angle in degrees ($-90^o - 90^o$) as the direction between the position of listener face and the position of sound source in a plane orthogonal to the azimuth, taking 0 angle as the intersection of these planes.

- **Distance:** Is the magnitude of the direction vector between the listener and the sound source given in centimetres.

Other consideration in VR is synthetic environments in VR, since the more complex and detailed the visual representations become, the more elaborated and intricated the sonic attributes need to be in order to match the uesrs expectations. This can pose many problems for sound designers in terms of realistic acoustic simulation and sound source emission because they can be very varied and sometimes very complex[MN10].

The principal effects refers a environmental contexts is reverberation, caused by repeated reflections of sound source from the all surfaces of an scene, and each of these surfaces, depending on their material, can give a distinct reverberation, these can potentially cause a significant effect on how a sound source is perceived. Only imagination and the tools at hand pose limits to the spatial manipulations possible for a user to make with existing technology[BT00], taking into account all its variables, however, we can approach by creating a real-time system with which we can locate a sound source with great accuracy in a three-dimensional space presentend by Julián Villegas [Vil15].

2.1.3 Related works

In general, the concern about audio in virtual environments has been the audio quality, D. Brandon Lloyd et al. [LRG11] propose an interactive system for synthesising high quality , physically based audio on current video game consoles which use to synthesise variations of the sound on the fly, that show for many sounds greater quality is obtained by using the amplitude envelopes of the extracted modes directly rather than fitting the envelopes to the standard exponential decay model.

Although the quality of the audio is important, it should be taken in consideration, to have goos acustic properties, that it is elementary to be able to provide immersion in the scene environment, Brent Cowan et al.[CK11] proposed an acoustical diffraction's using GPU hardware, given the widespread use and ability of these component of computer graphics, GPU have been successfully applied to other, non-graphics applications including audio processing and acoustical diffraction modeling. The idea is include the addition of acoustical textures that map audiobased parameters (e.g., specular reflection coefficient, a parameter that denotes the fraction of sound blocked by the object, and the pre-calculated ambient occlusion value) over the surface of an object. Using such acoustical textures allows these audio properties to vary across the surface of an object ultimately providing more flexibility while increasing running time requirements very minimally. The method is computationally efficient allowing it to be incorporated into real-time, dynamic, and interactive virtual environments and video games where the scene is arbitrarily complex.

Also, in order to get an immersion of a virtual scene, the sound spatializer is very important, Meghan Deutscher et al.[DHTF05] created an interactive sound and video for participants play with directional sounds initiated by the playful movement of Beluga whales In the water captured with a single live webcam provided by the Vancouver Aquarium, this method of spatialization is used in the 8-channel panner. The panner accepts up to eight channels of input sound signals that can each exist in a unique sound scope position at the same time using four of these, one for each sound signal. The panner accepts x and y coordinates for each sound signal. But this method focuses on a surround sound system, David Doukhan et al.[DS09] proposed a Earplug, that is a binaural filter for sound spatialization in Pure Data, this tool do real time interpolation of HRTFs, thus allowing a high quality perception of the position(azimuth, elevation) of various sound sources. Even so, Earplug have an limitations with a distance of sound sources, cant give a accurate of distances[Vil15].

2.2 Methodology

2.2.1 Type of study

This project is part of an exploratory study scheme, since sound spatialisation technology is under development, this project seeks a way to design, implement and test an audio spatialisation tool for video games in order to improve the immersive user experience.

2.2.2 Activities

1. **Work related to the thesis document**
   (a) Write the document.
   (b) Meetings with director and assessor of thesis.
2. **Tools integration (Objective a)**
   (a) Seek and explore tools to integrate an audio spatial into video game engine.
   (b) Test compatibility and usability of these tools.
   (c) Select and proof the algorithms of relevant tool to will integrate into video game engine.
3. **Develop framework (Objective b)**
   (a) Analyse and define the way in which the components of the video game engine will share the necessary elements for the spatialization of the sound emitted by an object in the virtual scene.
   (b) Design and implement a framework to spatialize sound in a video game engine.
   (c) Make unity proof in algorithms of these frameworks.
   (d) Make a user manual of this framework.
4. **Develop functional prototype (Objective c)**
   (a) Design a prototype that allows to test the audio specialization compared with the native spatialization of the video game engine.
   (b) Implement prototype to show for users.
5. **Validate user experience (Objective d)**
   (a) Make a surveys to identify strengths, weaknesses and performance for framework.
   (b) Validate and consolidate the information given for users.

2.3 Expected results

By the end of this project, besides the dissemination of results, we expect to build a sound spatializer prototype for Unity, where the users would be able to hear multiple sounds coming from arbitrary directions and distances, corresponding to sound sources in the virtual world using headphones, to demonstrate feasibility and benefits of near-field spatialization.

Bibliography

[BT00] Durand R Begault and Leonard J Trejo. 3-d sound for virtual reality and multimedia. 2000.

[CBM08] Jeff Craighead, Jennifer Burke, and Robin Murphy. Using the unity game engine to develop sarge: a case study. In *Proceedings of the 2008 Simulation Workshop at the International Conference on Intelligent Robots and Systems (IROS 2008)*, 2008.

[CK11] Brent Cowan and Bill Kapralos. Gpu-based acoustical occlusion modeling with acoustical texture maps. In *Proceedings of the 6th Audio Mostly Conference: A Conference on Interaction with Sound*, AM '11, pages 55–61, New York, NY, USA, 2011. ACM.

[DHTF05] Meghan Deutscher, Reynald Hoskinson, Sachiyo Takashashi, and Sidney Fels. Echology: An interactive spatial sound and video artwork. In *Proceedings of the 13th Annual ACM International Conference on Multimedia*, MULTIMEDIA '05, pages 937–945, New York, NY, USA, 2005. ACM.

[DPC+14] Tristan Dufour, Vincent Pellarrey, Philippe Chagnon, Ahmed Majdoubi, Théo Torregrossa, Vladimir Nachbaur, Cheng Li, Ricardo Ibarra Cortes, Jonathan Clermont, and Florent Dumas. Ascent: A first person mountain climbing game on the oculus rift. In *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-human Interaction in Play*, CHI PLAY '14, pages 335–338, New York, NY, USA, 2014. ACM.

[DS09] David Doukhan and Anne Sédès. Cw_binaural˜: A binaural synthesis external for pure data. In *Puredata International Convention Proceedings*, page 78, 2009.

[Ear14] Rae A Earnshaw. *Virtual reality systems.* Academic press, 2014.

[GA14] Robert Gilkey and Timothy R Anderson. *Binaural and spatial hearing in real and virtual environments.* Psychology Press, 2014.

[LRG11] D. Brandon Lloyd, Nikunj Raghuvanshi, and Naga K. Govindaraju. Sound synthesis for impact sounds in video games. In *Symposium on Interactive 3D Graphics and Games*, I3D '11, pages 55–62 PAGE@7, New York, NY, USA, 2011. ACM.

[Mil01] Joel David Miller. Modeling interaural time difference assuming a spherical head, 2001.

[MN10] David Murphy and Flaithrí Neff. Spatial sound for computer games and virtual reality. *Game sound technology and player interaction: concepts and developments*, pages 287–312, 2010.

[Vil15] Julián Villegas. Locating virtual sound sources at arbitrary distances in real-time binaural reproduction. *Virtual Reality*, 19(3):201–212, 2015.