International Journal of Advanced Smart Convergence Vol.5 No4 15-20 (2016)
https://doi.org/10.7236/IJASC.2016.5.4.15

IJASC 16-4-3

Listener Auditory Perception Enhancement using Virtual Sound Source Design for 3D Auditory System

Cheol Yong Kang[1], Vinayagam Mariappan[2], Juphil Cho[3], Seon Hee Lee[4†]

[1]*Graduate of Nano IT Design Fusion, Seoul National University, of Science and Tech., Seoul, Korea*
e-mail : doumceo@daum.net
[2]*Media IT Engineering, Seoul National University, of Science and Tech., Seoul, Korea*
e-mail : vinayagam_m@hotmail.com
[3]*Dept. Of Integrated IT & Communication Eng., Kunsan National University, Kunsan, Korea*
e-mail : stefano@kunsan.ac.kr
[4]*Dept. Of Electronics and IT Media Eng., Seoul National University, of Science and Tech., Seoul, Korea*
e-mail : seonhee@seoultech.ac.kr

Abstract

When a virtual sound source for 3D auditory system is reproduced by a linear loudspeaker array, listeners can perceive not only the direction of the source, but also its distance. Control over perceived distance has often been implemented via the adjustment of various acoustic parameters, such as loudness, spectrum change, and the direct-to-reverberant energy ratio; however, there is a neglected yet powerful cue to the distance of a nearby virtual sound source that can be manipulated for sources that are positioned away from the listener's median plane. This paper address the problem of generating binaural signals for moving sources in closed or in open environments. The proposed perceptual enhancement algorithm composed of three main parts is developed: propagation, reverberation and the effect of the head, torso and pinna. For propagation the effect of attenuation due to distance and molecular air-absorption is considered. Related to the interaction of sounds with the environment, especially in closed environments is reverberation. The effects of the head, torso and pinna on signals that arrive at the listener are also objectives of the consideration. The set of HRTF that have been used to simulate the virtual sound source environment for 3D auditory system. Special attention has been given to the modelling and interpolation of HRTFs for the generation of new transfer functions and definition of trajectories, definition of closed environment, etc. also be considered for their inclusion in the program to achieve realistic binaural renderings. The evaluation is implemented in MATLAB.

Key words: *Virtual Sound Source, 3D Auditory System, Binaural Synthesis, HRTFs, Music Widening Algorithm, Sound Image.*

Manuscript Received: Sep. 15, 2016 / Revised: Sep. 30, 2016 / Accepted: Oct. 10, 2016
Corresponding Author: seonhee@seoultech.ac.kr
Tel: +82-02-970-6426, Fax: +82-02-974-6123
Dept. Of Electronics and IT Media Eng., Seoul National University, of Science and Tech., Seoul, Korea

1. Introduction

The sound system technologies initially uses the simplest, a monophonic sound system, is incapable of reproducing the spatial characteristics of sound. Two canal stereo sound systems are by far superior, enabling the reproduction of sound image that are spatially distributed between two loudspeakers. The capabilities of stereo sound systems can be augmented by adding additional speakers to the sides or rear of the listener. The resulting surround systems are generally able to reproduce sound images anywhere in the horizontal plane surrounding the listener.

A virtual acoustic display or 3-D audio system is a system capable of rendering sound images positioned arbitrarily around a listener. The binaural technology has become enabling technology with a significant impact on various fields, such as information technology, hearing aids and advanced sound measurement techniques. There are three aspects, which comprise a binaural system. They are physical, psychoacoustic and the psychological aspect. The physical aspect is concerned with the way of the source signals before they reach the inner ears. The Psychoacoustic Aspect deals mainly with the signal processing in the subcortical auditory system and the Psychological Aspect focuses on cognitive brain function.

The binaural technology rests mainly on knowledge of the physical aspect, with an increasing use of psychoacoustics. This paper deals with the aspects of physics and psychoacoustics as the subcortical part is not that important for the implementation of the 3-D system. Firstly, the effects of lateralization have been evaluated in terms of the Head Related Transfer Function (HRTF) and howthey affect the perception of a source that is being placed at a different angle around the listener. This first approach emulates what happens when a sound source is displaced in space under near listening conditions.

2. Human Auditory System

The human auditory system consists of the outer ear, middle ear, and inner ear built with pinna, eardrum, ossicles, and basilar membrane shown in Figure 1. The outer ear consists of the pinna, the auditory canal, and the tympanic membrane (eardrum).The pinna collects sound and funnels to the ear canal which acts as a filter helping us localize sounds.

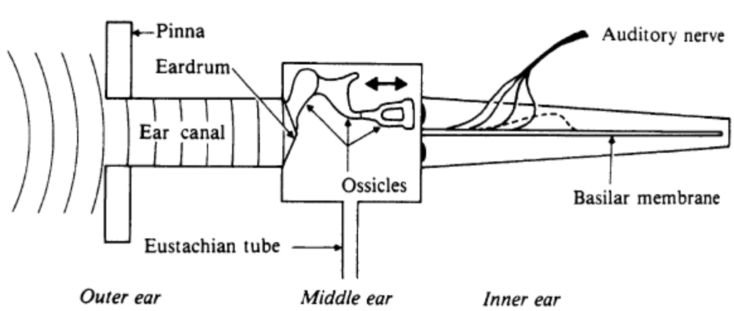

Figure 1. Human Auditory System [1]

The auditory canal acts as an acoustic tube closed at one end and boosts hearing sensitivity in the range 2000-5000Hz. The middle ear consists of the eardrum, to which three small bones, called the ossicles, are attached. The eardrum changes pressure variations of incoming sound waves into mechanical vibrations which are then transmitted via the ossicles to the inner ear. The Inner Ear consist of cochlea and basilar membrane. The cochlea transforms pressure variations into properly coded neural impulses. The basilar membrane stops just short of the end of the cochlea to allow the fluid to transmit pressure waves around the end of the membrane. The fluid pressure causes frequency dependent vibration patterns of the basilar membrane (H) within the inner ear, which causes numerous fibers producing from auditory hair cells. These activate electrical action potentials within the neurons of the auditory system, which are combined and processed at higher levels with information from the opposite ear.

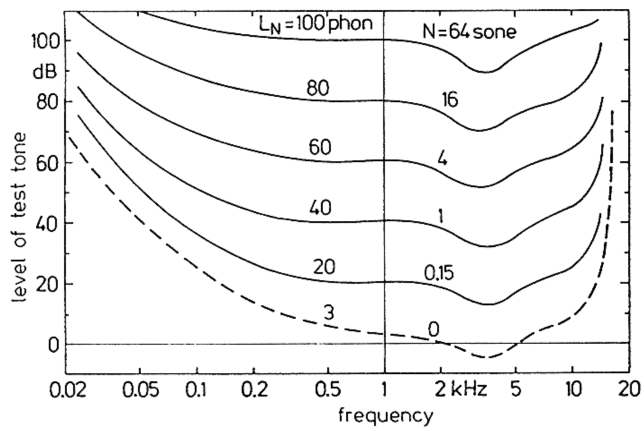

Figure 2. Human Auditory System Spectral Response [1]

The human auditory system is a quite complex and amazing device which transduce fluctuations in ambient atmospheric pressure into electrical signals that's processed by the brain and perceived as sound by the humans. The dynamic range of a human auditory system describes the difference between the highest level that can pass through the system and the noise floor [1]. The human auditory system threshold of human hearing is around 0.00002 Pascals (Pa) at mid frequencies and can withstand peaks of up to 200 Pa at mid frequencies. The human auditory system spectral response shown in Figure 2. The spectral responses shows sensitivity is high to low frequency ranges and the sensitivity is low at higher at low frequencies.

3. Virtual Sound Sources

Virtual Sound Source reduction, which means the reproduction of a desired sound source over a finite area by a loudspeaker array, can deliver spatial sound impressions to multiple listeners. There are two types of virtual sources, one is a virtual source outside, which is located outside an enclosed array, and the other is a virtual source inside, which is located inside an array, that is, on the same side of the array as the listener's position.

Head-related transfer functions (HRTFs) used to create virtual sound sources. The HRTFs works by filtering a signal with a left/right pair of HRTFs from a certain direction. When the resulting left and right ear signals are played through hearing system, the listener perceives the sound as coming from a point in space that corresponds to the chosen direction. By selecting another pair of HRTFs one can change the position of the virtual sound source in any direction or even create a moving sound source.

HRTFs are influenced by directional and nondirectional components. Directional components are the torso, shoulder reflection, head diffraction and reflection and pinnae. Nondirectional components are the cavum conchae dominant resonance, the ear canal and eardrum impedance. A structural model of these components is shown in figure 3.

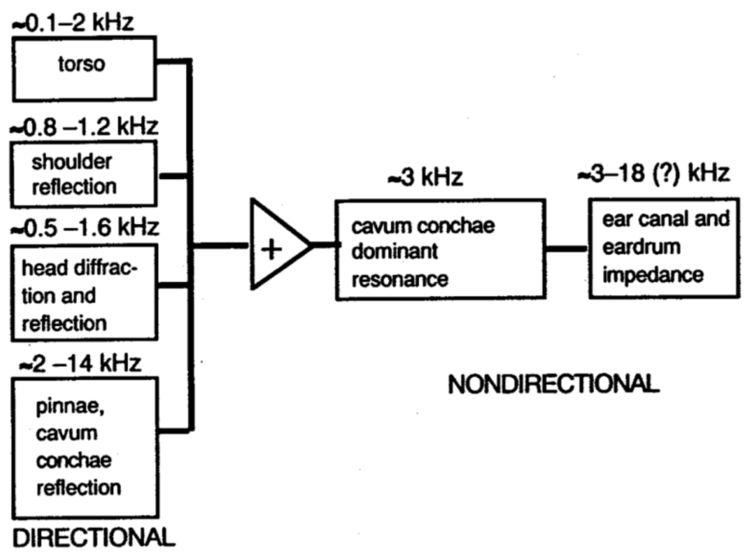

Figure 3. HRTF Structural Model [3]

The influence of the upper body and shoulders become apparent in the range of 100 Hz- 2 kHz. The shoulder influence is about ±5dB and the influence of the torso about ±3dB. The ear canal can be seen as an acoustic transmission line between the eardrum and the outer ear. With a length of 2.5 cm and a diameter of 7-8 mm, it has a significant resonance at 3-4 kHz.

4. 3d Auditory System Binaural Synthesis

The 3D auditory system binaural synthesis, the room impulse response is combined with HRTFs to simulate a virtual sound source in a room. The binaural synthesis is implemented by replacing every sound contribution uses the direct sound as well as every individual reflection by a pair of HRTFs, since each contribution comes from a different angle. The particular pair of HRTFs depends on the direction from which the sound arrives at the listener. The directions and times of arrival of the reflections can be calculated by means of a mirror-image model of the room. So, every reflection is slightly different at the two ears, in that the times and amplitudes are as found in the particular HRTFs for the corresponding direction. Furthermore the reflections are gradually more damped as they hit more walls. Implementing this binaural room impulse response ensures that the acoustical cues at the two ears are correct.

The room reflections make a positive contribution to the perception of the sound source. Notice that the direct sound and the reflections can come from all directions around the listener. There are several factors analyzed when facing the problem of simulating the human peripheral auditory system. The HRTFs help in characterizing in a highly accurate way the effects of the interaural differences used to locate where a sound is coming from. The structure model for binaural synthesis is shown in Figure 4.

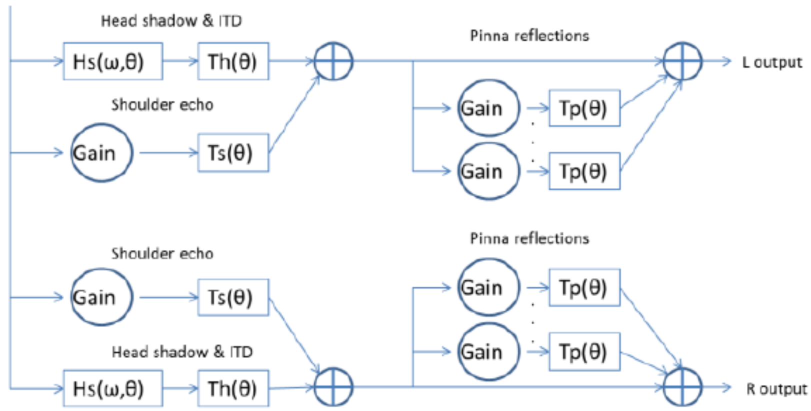

Figure 4. Structural Model of Binaural Synthesis [1]

Ultrasound is a mechanical acoustic wave with the frequency range from roughly 10 kHz to 20 MHz. It imparts high energy to reaction medium by cavitation and secondary effects. Microbubbles containing solvent vapors are generated that grow and undergo radial motion as acoustic energy propagates through the liquid medium.

5. Simulation and analysis

The proposed Virtual Sound Source Design for 3D Auditory System method simulated using binaural synthesis to evaluation directional and non-directional localization evaluation in a room with virtual measurement space as shown in Figure 5. The performance evaluation method analysis implement on x64 Windows 7 PC using MATLAB.

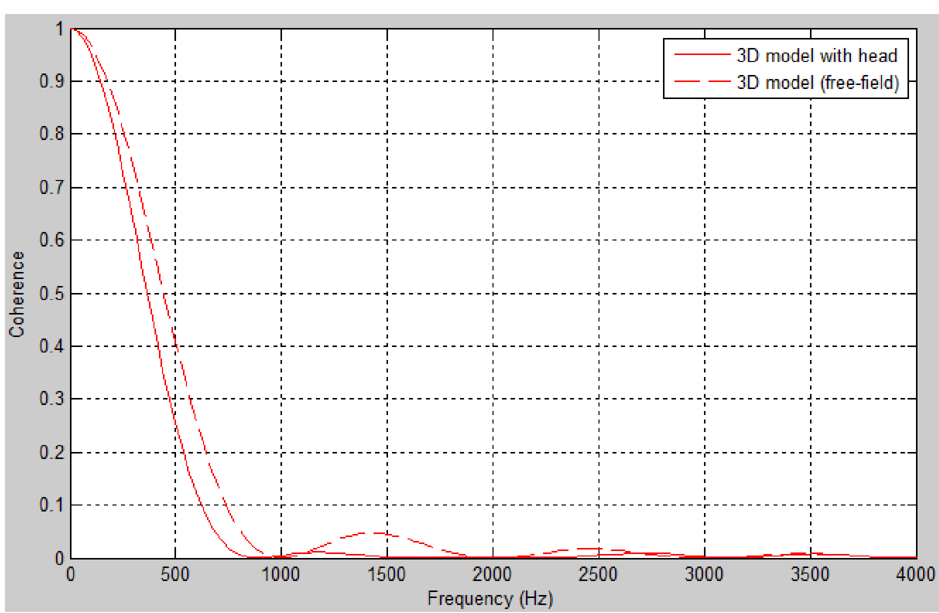

Figure 5. Binaural Synthesis Simulation

The directive sound sources were placed in virtual space and the response characteristic of each speaker was measured at the listening position, when the listening position deviates from the directivity characteristic of the speaker, the response characteristic was significantly lower than that in the directivity characteristic. To evaluate the performance, the coherence is measured in the 1 kHz ~4 kHz band from the listener position and the graphical representation of characteristic influenced on the clarity on directionality of the source.

6. Conclusions

The Auditory Perception Enhancement algorithm was developed to improve the poor sound image that can result from listening to music on from speaker using Virtual Sound Source Design for 3D Auditory System Music. A simple mirror-image model of a room is used to calculate the contributions of the direct sound and the room reflections. By replacing each contribution with a virtual sound source, it is possible to create the impression of sound sources in a simulated room. This sound image is externalized, i.e., perceived to be outside the head. This is in contrast to traditional stereo, where the sound image can be lateralized, but is still perceived as being inside the head. The results from the listening experiment indicate that applying this processing to a streamed music signal can dramatically improve the spatial impression.

Acknowledgement

This study was supported by the Research Program funded by the Seoul National University of Science and Technology

References

[1] Glen Ballou, "Handbook for Sound Engineers" , 2nd Edition Boston : Focal Press, Page 35
[2] J. Blauert J: Spatial Hearing: The Psychophysics of Human Sound Localization. Rev. ed. Cambridge, MA: MIT Press, 1997.
[3] Møller H: Fundamentals of binaural technology. Applied Acoust 1992; 36:171–218.
[4] Helmholtz HLF: On the Sensations of Tone, 2nd English ed. translated from the 4th German ed., 1863. New York: Dover Publications, 1954.
[5] Kleiner M, Dalenbäck B-I, Svensson P: Auralization: An overview. J Audio Eng Soc 1993; 41(11):861–875.
[6] Nielsen SH: Auditory distance perception in different rooms. J Audio Eng Soc 1993; 41(10):755–770.
[7] Beranek L: Concert Halls and Opera Houses: Music, Acoustics, and Architecture, 2nd Ed. New York: Springer, 2003.
[8] Kahrs M, Brandenburg K: Applications of Digital Signal Processing. Boston: Kluwer Academic Publishers, 1998.
[9] Bradley RA, Terry MA: Rank analysis of incomplete block designs: The method of pair comparisons. Biometrica 1952; 39:324–345.
[10] Luce RD: Individual Choice Behavior. New York: Wiley, 1959.