THREE MAIN SPATIAL TECHNIQUES OF SPATIAL MIXING SYSTEM

Sangkyou Park 420155724

Unit of Study Name, DESC9115, Semester 1 2014
Graduate Program in Audio and Acoustics
Faculty of Architecture, Design and Planning, The University of Sydney

ABSTRACT

This written review is to have a deep review for three techniques to realize spatial effects, which are Head Related Transfer Functions, Image Source Method, and Schroeder Reverberation, utilized in Spatial Mixing System, 2nd lab report.

1. INTRODUCTION

From 1960s, there have been a number of researches to implement spatial effects. Representative techniques could be "Head-Related Transfer Functions (HRTFs)", "Image Source Method (ISM)", "Schroeder Reverberation (SR)". Although each technique has a unique system to reproduce a spatial sound effect [1], it could lead better spatial experiences with the combination of these techniques based on a research from Gary et al which is about the image source model and Schroeder reverberation [6]. Chapter 2 present the Head-Related Transfer Functions. This section mainly introduces HRTFs including Duplex Theory, measurement procedures, and the structure of HRTFs. Image Source Model is explained in Chapter 3. Schroeder Reverberation is presented in Chapter 4 in terms of pros and cons.

2. REVIEW FOR THE TECHNIQUES

2.1. Head Related Transfer functions

Duplex Theory

The Duplex theory is a simple model introduced by Lord Rayleigh to explain directional hearing in the azimuthal direction [5] with two binaural cues: inter-aural time difference (ITD), defined as the difference in arrival times at both left and right ears, and inter-aural intensity difference (IID), defined as the amplitude difference generated by a sound between left and right in the free field (See Fig.1 and Fig.2). [23]

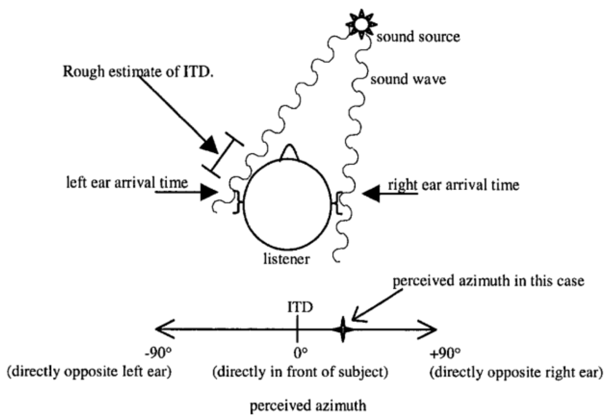


Figure 1 Basic concept of the Inter-aural Time Differeces (ITD)

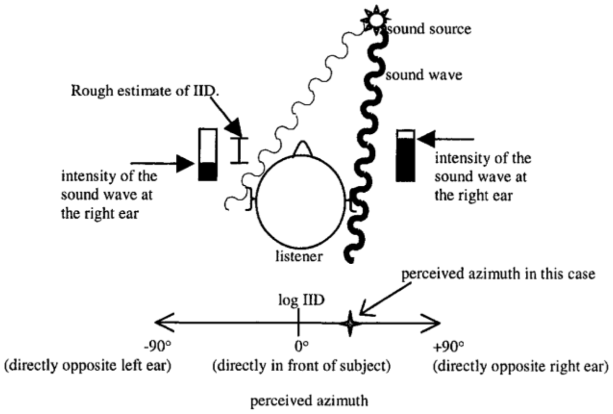


Figure 2 Basic concept of the Inter-aural Intensity Differeces (IID)

ITD cues for azimuth become ambiguous above 1500 Hz because the wavelength of a sinusoid becomes comparable to the diameter of the head. At this frequency and above, ITD may correspond to distances that are longer than one wavelength. Thus an aliasing problem occurs above 1500 Hz and the difference in phase no longer provides a unique spatial location (See Fig 3).

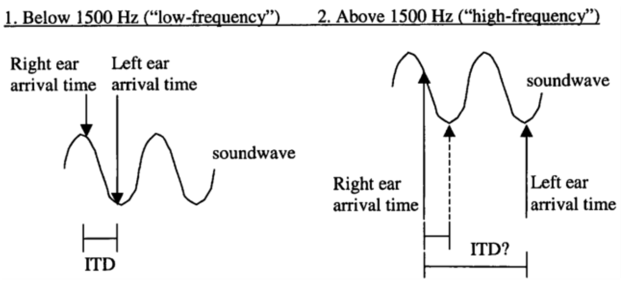


Figure 3 The ambiguity of ITD's in determining lateral position for higher frequencies

Since this theory only explains the perception of azimuth with limitation of ITD and IID cues, there are a number of locations along curves of equal distance. This set of points is called the "Cone of confusion" because the location of all sound originating from points on this cone are indistinguishable (See Fig. 4).

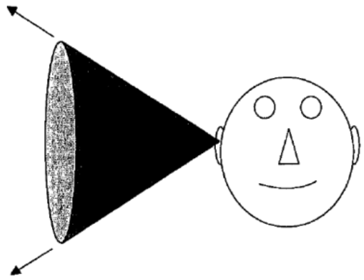


Figure 4 The cone of confusion

HRTFs

Formally, a single HRTF is defined to be a specific individual's left or right ear far-field frequency response, as measured from a specific point in the far-field to a specific point in the ear canal. Typically, several different azimuths and elevations are considered with both left and right ear for the HRTF measurement (See Fig.5) [1].

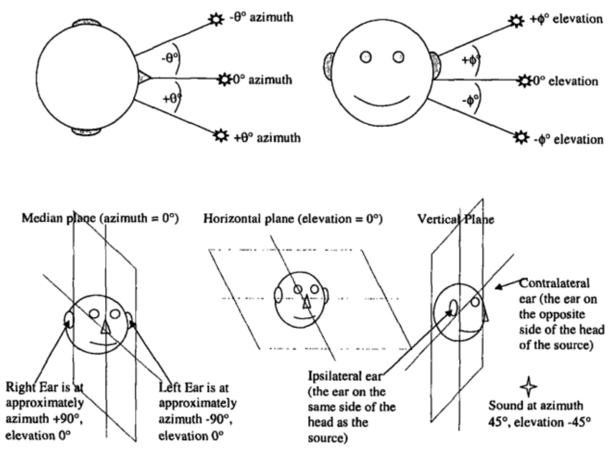


Figure 5 Spatial coordinate system and terminology used in HRTF measurement

Measurement of HRTFs

The basic procedure of measuring HRTF is to insert probe tube microphones partially into ears, and then to perform a simple form of system identification by playing spectrum stimulus through a speaker which is located in a specific position with azimuth(θ), elevation($\emptyset$) and distance from a subject [7].

Issues with HRTF-based synthesis

Simple HRTF-based spatial effects algorithms like Fig. 6 could not always provide a sound with appropriate spatial feelings. When azimuth is 0 degree, localizing the sound at the front is one of the issues from the HRTF-based sound synthesis system [8][9]. Each person has different size of ear, pinna, canal, and head size so it could be almost impossible to generalize the spatial sound which might provide wrong position about the sound source [10][11].

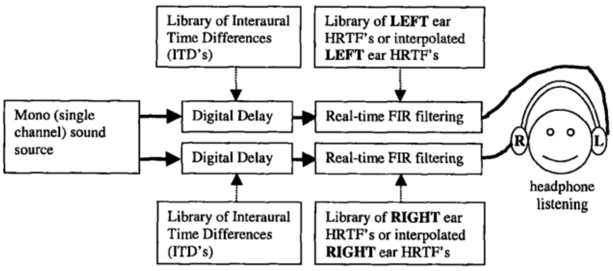


Figure 6 Block diagram of a simple HRTF-based spatial sound Synthesis system

Structure of HRTF data

Frequency domain representation

There are a number of benefits from frequency domain-based HRTFs data with principle components analysis [12], pole-zero modeling [13] and interpolation of HRTFs [14]. Specifically, major differences between theoretical and measured HRTF data, elevation effects and diffraction effects can be monitored based on frequency domain representation.

Time domain representation

The time domain HRTF data is called Head-Related Impulse Responses (HRIRs) [15]. This form of data can be inserted directly into sound spatialization system such as the one shown in Fig. 6.

Spatial domain representation

HRTFs can be represented in the spatial domain in different ways: plotting ITD and IID as a function of azimuth and elevation [25] and plotting frequency domain of HRTFs with a common azimuth or elevation as a surface where has highlights patterns effectively, called Spatial Frequency Response Surfaces (SFRSs). [23][24].

2.2. Image Source Modeling

Basic concept

Image Source Method (ISM) is a ray-based method with Ray-tracing. The basic distinction between the two methods is the way to find the reflection paths. The basic concept of the image source method is presented in Fig. 7. To find reflection paths from the sound source to the listener, the source is reflected against all surfaces in the room. [1][6][22]

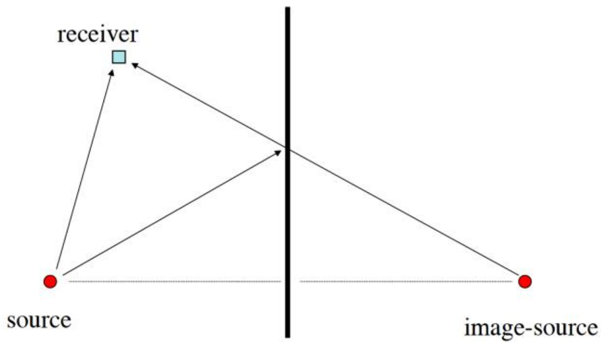


Figure 7 Basic concept of image source model

The sound source and reflected sound can be presented in Fig. 8. Blue line shows the direct sound between source and listener and green line represents the path from image source.

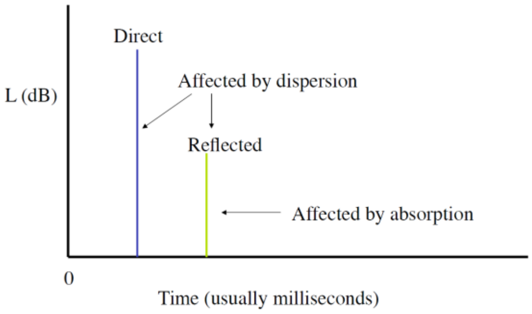


Figure 8 Direct sound and reflected sound from the image source modeling

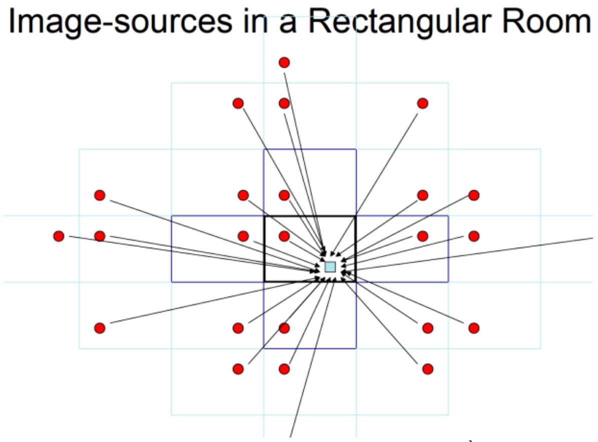
The number of image sources is proportional to the amount of time. Hence, the volume of the wave moving from the source can be calculated with the followed equation,

$$V_{wave} = \frac{4}{3}\pi c^3 t^3,$$

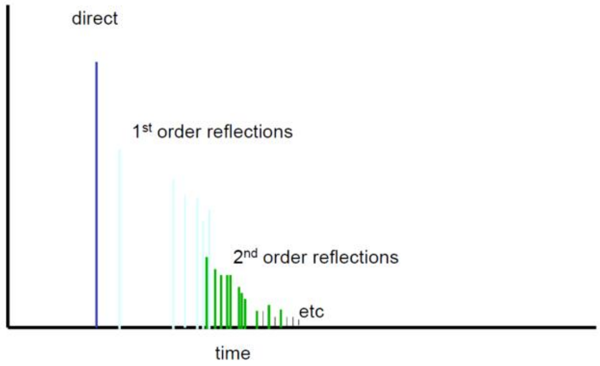
as the volume of a sphere is 4/3*pi*r^3. With the equation and the volume of a rectangular room, the number of image sources is

$$n = \frac{4\pi c^3 t^3}{3L_x L_y L_z},$$

where L_x is length, L_y is width and L_z is height of the room so L_x*L_y*L_z is the volume of the room.



(a) Image source of 1st order and 2nd order.



(b) Level difference between direct sound, 1st order, and 2nd order reflection sound.

Figure 9 Example of Image source modeling

The higher number of order means late reverberation and the intensity of each impulse could be affected by dispersion, absorption, dissipation and scattering aspects. As for dispersion, based on the inverse square law, it is possible to estimate the effect of dispersion over time approximately.

$$I = \frac{P}{4\pi c^2 t^2}$$

As for the relationship with absorption, the average distance that sound travels between reflections can be calculated from the ratio of room volume to room surface area.

$$l_m = \frac{4V}{S}$$

Therefore, the intensity of an impulse due to multiple reflections is

$$I = I_0(1-\alpha)^{\frac{ctS}{4V}},$$

with assuming a fraction of the sound intensity which could be absorbed. The sound intensity could be estimated from above two equations.

$$I = \frac{P(1-\alpha)^{ctS/4V}}{4\pi c^2 t^2}$$

From this equation, sound intensity level is estimated from

$$L_I = L_W + 10\log\left(\frac{(1-\alpha)^{ctS/4V}}{4\pi c^2 t^2}\right).$$

Implementation

From the image source modelling, the coordinates, azimuth, elevation, path length (delay time), the intensity (gain) and distance of image sources can be calculated.

1) The coordinates of image sources

Based on the number of order for the modelling, the coordinates of the sound sources can be determined with the size of room and the source location.

2) Azimuth, Elevation and Distance

With the coordinates of image sources(x, y, z), these values are available. Firstly, path legnth (r) is

$$r = \sqrt{(x-x_r)^2 + (y-y_r)^2 + (z-z_r)^2}$$

where x_r, y_r, z_r are the coordinates of receiving position (listener).

Azimuth is $\tan^{-1}(\frac{Z}{r})$ and Elevation is $\sin^{-1}(\frac{Y}{X})$.

3) Intensity (gain)

The only change is to count the exact number of reflections which is used for the exponent of the reflection coefficient.

$$I = P * \frac{(1-\alpha)^{(n_x+n_y+n_z)}}{4*\pi*r^2}$$

$$L_I = L_W + 10 * \log(\frac{(1-\alpha)^{(n_x+n_y+n_z)}}{4*\pi*r^2})$$

2.3. Schroeder Reverberation

In terms of artificial reverberation, Manfred Schroeder introduced a reverberator with recursive comb filter and delay-based all pass filter. [1][16~21] Followed equation and figure 1 show the concept of all pass filter.

$$y(n) = -g \cdot x(n) + x(n-m) + g \cdot y(n-m),$$

where m is the length of the delay in samples and described in Fig.2

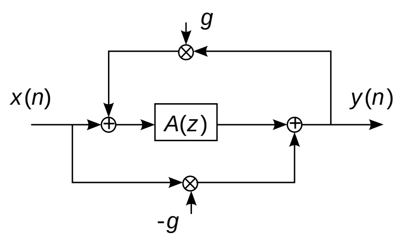


Figure 10 Structure of all pass filter [1]. Where A(z) is replaced by the delay line

The reason this is notable is that real rooms and acoustic spaces have an echo density that increases with time, while the parallel combs and series all-pass reverberators have a constant echo density.

3. THE CONNECTIONS IN SPATIAL MIXING SYSTEM

The fundamental idea originated from the previous research [6]. One of the significant concepts of the paper is to connect the Schroeder reverberator and the image model to create “spatial reverberation”. From this research, binaural technique could provide different sound source localization with the spatial reverberation. Fig. 11 represents the signal flow for the Spatial Mixing System.

Figure 11 Basic signal flow of the Spatial Mixing System

4. DISCUSSION AND CONCLUSION

Although each techniques has an innovative idea to create spatial ascetics, merging different systems could bring far more effective spatial emotions for ordinary people. As a future work, the Spatial Mixing System should consider to accept another virtual effects for multichannel loudspeaker system like Vector base amplitude planning. Furthermore, Moorer reverberator and Feedback delay networks could provide a broaden spatial experience. Therefore, based on the previous studies, it could be possible to have a creative ideas to lead a virtual hearing space.

5. REFERENCES

[1] Udo Zo ̈lzer, DAFX: Digital Sound Effects second edition, 2011.

[2] A. Bee, C.D. Player, and X. Lastname, “A correct citation,” in *Proc. First Int. Conf. (IC)*, Helsinki, Finland, June 2001, pp. 1119–1134.

[3] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*. Springer-Verlag, Heidelberg, Germany, 1990.

[4] M.R. Smith, *A good journal article*, J. Acoust. Soc. Am., vol. 110, no. 3, pp. 1598–1608, Mar. 2001.

[5] Rayleigh, L., *On our perception of sound direction*, Philosophical Magazine 13:1907.

[6] K. Garydall, M. William, F. Daniel, L. Derek, and K. Richard, Image Model Reverberation from Recirculating Delays, The Journal of AES, 81th Convention, 1986.

[7] Wightman, Frederic L. and Kistler, Doris J. , *Headphone simulation of free-field listening. I: Stimulus synthesis*, Journal of the Acoustical Society of America, 85(2): February 1989.

[8] Wightman, Frederic L. and Kistler, Doris J. , *Headphone simulation of free-field listening. II: Psychophysical validation*, Journal of the Acoustical Society of America, 85(2): February 1989.

[9] Griesinger, David., *Objective Measures of Spatiousness and Envelopment*, Proceedings of the 16th Audio Engineering Society (AES) International Conference on Spatial Sound Reproduction, Rovaniemi, Finland:1999.

[10] Pralong, Dani61e, and Carlile, Simon. , *The role of individualized headphone calibration for the generation of high fidelity virtual auditory space*, Journal of the Acoustical Society of America, 100(6): December 1996.

[11] Wenzel, Elizabeth M. et al., *Localization using nonindividualized head-related transfer functions*, Journal of the Acoustical Society of America, 94(1): July 1993.

[12] Kistler, Doris J. and Wightman, Frederic L. *A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction*, Journal of the Acoustical Society of America, 91(3): March 1992.

[13] Blommer, Michael Alan., *Pole-zero Modeling and Principal Component Analysis of Head-Related Transfer Functions*., Ph.D. Dissertation for the University of Michigan, Dept. of Electrical Engineering and Computer Science, Systems division, Ann Arbor, Michigan: 1996.

[14] Hartung, Klaus et al. *Comparison of Different Methods for the Interpolation of Head-Related Transfer Functions*., Proceedings of the 16 thAudio Engineering Society (AES) International Conference on Spatial Sound Reproduction, Rovaniemi, Finland: 1999.

[15] Wu, Zhenyang et al., *A time domain binaural model based on spatial feature extraction for the head-related transfer function*., Journal of the Acoustical Society of America, 102(4): October 1997.

[16] J. Blauert. Spatial Hearing. *The Psychophysics of Human Sound Localization*, 2nd edition MIT Press, 1997.

[17] M. R. Schroeder. *Improved quasi-stereophony and "colorless" artificial reverberation*. J. Acoust. Soc. Am., 33(8): 1061–1064, 1961.

[18] M. R. Schroeder. *Natural-sounding artificial reverberation*. J. Audio Eng. Soc., 10(3): 219–223, 1962.

[19] M. R. Schroeder. *Digital simulation of sound transmission in reverberant spaces*. J. Acoust. Soc. Am.,47(2 (Part 1)): 424–431, 1970.

[20] M. R. Schroeder. *Computer models for concert hall acoustics*. Am. J. Physics, 41: 461–471, 1973.

[21] E. J. Macaulay, W. M. Hartmann, and B. Rakerd. *The acoustical bright spot and mislocalization of tones by human listeners*. J. Acoust. Soc. Am., 127: 1440, 2010.

[22] J.H. Rindel, *The use of Computer Modeling in Room Acoustics*, Journal of Vibroengineering, 2000, No 3-4, p.41-72.
Corey I. Cheng and Gregory H. Wakefield, *Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space*, The Journal of AES, 107th Convention, 1999.

[23] Cheng, Corey L and Wakefield, Gregory H. ,*Spatial Frequency Response Surfaces: An Alternative Visualization Tool for Head-Related Transfer Functions (HRTF's)*., Proceedings of the 1999 International Conference on Acoustics, Speech, and Signal Processing (ICASSP99), Phoenix, Arizona: 1999.

[24] Cheng, Corey L and Wakefield, Gregory H., *Spatial Frequency Response Surfaces (SFRS's): An Alternative Visualization and Interpolation Technique for Head-Related Transfer Functions (HRTF's)*., Proceedings of the 16th Audi Engineering Society (AES) international Conference Spatial Surround Reproduction Rovaniemi Finland: 1999.

[25] Carlile, S. and Pralong, D., *The location-dependent nature of perceptually salient features of the human head-related transfer functions*., Journal of the Acoustical Society of America, 95(6): June 1994.