HUMAN COMPUTER INTERFACE VIRTUAL ACOUSTIC DISPLAY USING GYROSCOPIC SENSORS

Megha Sunny, Ayse Kalkan-Savoy and Charles Thompson

University of Massachusetts Lowell
Center for Advanced Computations and Telecommunications
Department of Electrical and Computer Engineering
1 University Ave, Lowell, MA 01854

ABSTRACT

In the paper the spherical-head model for the HRTF is developed and is used in conjunction with yaw angle tracking provided using software developed for a WiiRemote controller to simulate sound radiated from a point source as a function of head position. The yaw rotation angle is evaluated using gyros supplied in the Wii motion plus extension device. The real-time data obtained is used to control the parameters of the head related transfer function. This work aims to bring affordable experimental platforms for educating students in physical acoustics and signal processing.

***Index Terms*—** HRTF, position tracking, yaw tracking, Wii, Wiimote

1. INTRODUCTION

The physical and psychological processes governing spatial hearing in humans is an area of continuing research study. The characteristic sound received at the ears of a listener is the result of a signal filtering process. The acoustic interactions along the path between the source and the receiver such as scattering from the head, torso, pinna, and ear canal are manifest in the filtered signal. The amplitude and time delay of the received signal may be attenuated or reinforced and will depend on the position of the ears relative to the position of the source. It is the time delay and the amplitude differences experienced that allows one to determine the location of a sound source.

In recent years researchers have actively pursued models and systems for the accurate reproduction of sound using headphones and loudspeakers. Enhancements in sound reproduction using headphones is focused on developing electronic means to simulate the natural process of hearing. The Head-Related Transfer Function(HRTF) has been an important step to this end.

In headphone based sound reproduction the HRTF is central to the development of spatial hearing systems. The investigations of Blauret [1] [2] and Wightman and Kistler [3] [4] detail the spatial synthesis experiments using HRTF data collected from the subjects in anechoic environments and synthesized over headphones.

The connection between sound localization and the scattering of acoustic waves was first made by Lord Rayleigh[5] nearly 100 years ago. His observations form the basis of the duplex theory which attributes the human sound localization to the Interaural Time Difference (ITD) and Interaural Level Difference (ILD) of the sound pressure between the two ears. The ITD is a measure of the arrival time difference between the left and right ears and ILD is the level of the intensity of sound received by each ear. These two cues are important for the virtual acoustic localization. Numerous investigators [6] [7] [8] have used the spherical-head model to explain how acoustic scattering of head affects the received signal.

In the paper the spherical-head model for the HRTF is developed and is used in conjunction with yaw angle tracking provided using software developed for a WiiRemote controller to simulate sound radiated from a point source as a function of head position. The HRTF allows one calculate the time-delay and the scattering wave amplitude as a function of frequency and yaw angle. This work aims to bring affordable solution towards educating students in acoustics and signal processing.

2. PROBLEM FORMULATION

A head related transfer function relates the pressure at a single ear to the pressure at the source. The signal path is depicted in Figure (1), where x_L and x_R denote the left and right source signals respectively. These signal are filtered the left and right HRTF to yield the received signals y_L and y_R. The pressure at each ear is a function of the lateral distance and the angular orientation of the head relative to the source. For this reason two HRTFs are required and the tracking of the head position is critical. Spatial information regarding the head location is

---

This work was supported by the NSF, grants CCF0649235 and DGE0841392.

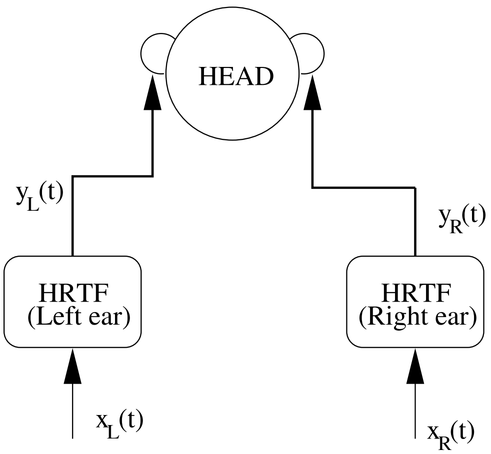

Fig. 1. Head phone sound synthesis.

used to control the signal characteristics of each head related transfer function and in turn control localization. A Nintendo Wii's video game controller (WiiRemote) is used to track the position and orientation of the head. The headphone signal is generated via a sound card equipped computer. Adjustment of the location of the sound source relative to the listener position is software controlled.

The WiiRemote is equipped with a three-axis accelerometer, infrared camera and a Bluetooth wireless functionality. The WiiRemote can be connected easily to an PC using a Bluetooth adapter. The basic WiiRemote is only able to sense the angular acceleration in pitch and roll orientations making it unacceptable for a head-tracking application. However since its original release, Wii has introduced an extension device [9][10] for the WiiRemote that is equipped with a two-axis tuning fork gyroscope. The extension enables one to measure the angular acceleration in the yaw direction. The yaw angular acceleration feature will be used exclusively in the work presented here.

3. METHODOLOGY

3.1. Theory

This study uses a rigid spherical model for HRTF, and excludes the diffraction and reflection of sound from any other parts of the body. The sound source is assumed to be a point source. The filter model includes computation of sound arrival times for left and right ear, and scattering spherical wave effects and sound pressure. The pressure is calculated as a function of frequency, source location and yaw angle. Under the assumption that the typical time scale of the head motion is long when compared to the impulse response duration the HRTF may be consider to behave as a linear time invariant system. In such a case the spatial location effects can be evaluated by the convolution of the input signal with the head related impulse response.

The classical model for the ITD is based on the time delay between the center of the head and left/right ear by trying to solve the problem for line-of-sight and non line-of-sight geometry. Figure (2) shows the case of non line-of-sight geometry. The circle represents the head, and the ear position is at $\underline{e}$ vector location. It is assumed that the wave length is much smaller than the diameter of the scatter. The source is located at position $\underline{r}$. The distance between ear and source is $|\underline{d}_1| + |\underline{d}_2|$. For case of the line-of-sight between the source and ear, $\underline{d}_2$ does not exist therefore distance vector $\underline{d} = \underline{d}_1$.

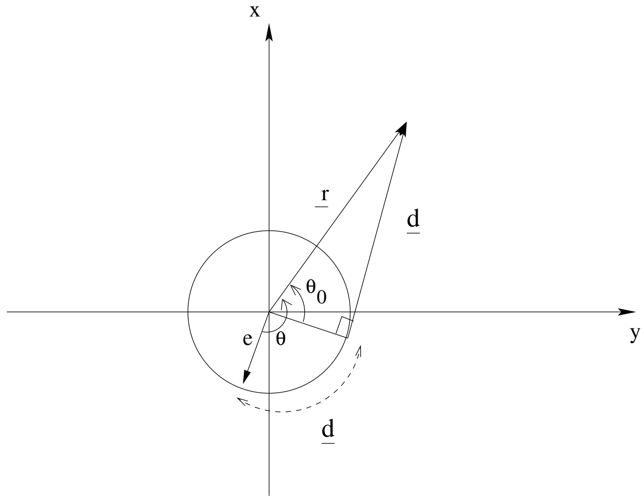

Fig. 2. Non line-of-sight geometry.

The sound arrival time to position $\underline{e}$ is equal to

$$t_{arrival} = \Big[|\underline{d}_1| + |\underline{d}_2|\Big]/c \tag{1}$$

where c is the speed of sound. Hence the time delay scaled on the characteristic time e/c is

$$\Delta t\frac{c}{e} = \begin{cases} \sqrt{1+\rho^2-2\rho cos\theta} - \rho & \text{if } 0 \leq \theta \leq \theta_0 \\ \sqrt{\rho^2-1} + (\theta - \theta_0) - \rho & \text{if } \theta_0 \leq \theta < \pi \end{cases} \tag{2}$$

where $\rho = r/e$, $r = |\underline{r}|$, $e = |\underline{e}|$, and $\theta_0 = cos^{-1}\frac{1}{\rho}$. On the other hand to determine the ILD one must consider the scattered and diffraction of the incident wave by the head. When wavelength is large compared the typical head diameter weak scattering ensues. This case be approximated using the Born approximation. For short wavelength diffraction effect can be modeled using a geometrical optics approximation.

The HRTF can be directly computed as the solution of the Helmholtz equation for the case where the head is modeled simply as a rigid sphere. The objective is to determine the sound pressure on the surface of the sphere as a function of the source location. The geometry of the problem is shown in Figure 3. The position $\underline{r}_s$ is the the location of a harmonically

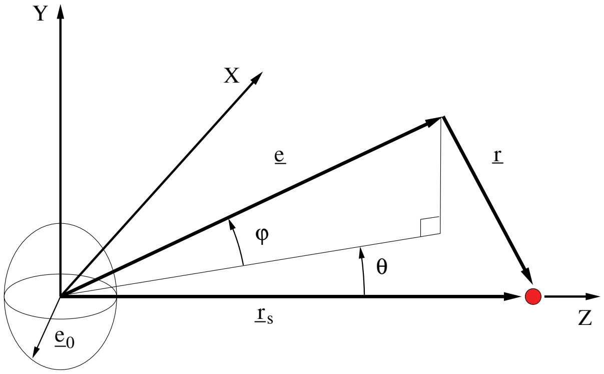

Fig. 3. Geometry of the source and the scatterer.

time varying point source with time dependence $e^{-i\omega t}$. The total pressure is the sum of the incident and scattered pressure.

$$P = P_i + P_s = ikh_0(kr) + \sum_{m=0}^{\infty} a_m h_m(ke) L_m(\eta) \tag{3}$$

where $k = \omega/c$ is the wavenumber, h_m is the m-th order spherical Hankel function, L_m is the Legendre function of order m and $\eta = cos\theta$. For a rigid sphere one determines the coefficients a_m by imposing the condition that the normal component of the particle velocity is equal to zero on surface $e = e_0$. In such a case

$$a_m = -ik(2m+1)\int_{-1}^{1} \frac{h_0{}'(kr_0)}{h_m{}'(ke_0)} \frac{e_0 - r_s\eta}{2r_0} L_m(\eta) d\eta \tag{4}$$

where $()'$ is used to denote differentiation with respect to the argument. The distance between the source and an annular ring of the surface of the sphere is given by $r_0 = \sqrt{e_0^2 + r_s^2 - 2e_0 r_s \eta}$. The fact that pressure is constant to variation in φ results in ambiguity in source localization with changing elevation. For this reason the HRTF should be updated based on the the head yaw angle.

Upon expressing the incident pressure at $e = e_0$ in terms of a Legendre function expansion, the pressure on the surface of the sphere is can be written as

$$P(e_0, \rho, \mu, \theta) = \frac{-1}{e_0\mu} \sum_{m=0}^{\infty} \frac{h_m(\mu\rho)}{h_m{}'(\mu)} L_m(\eta)(2m+1) \tag{5}$$

where $\mu = ke_0$, $\rho = r_s/e_0$. The HRTF is obtained by multiplying P by the distance between the source to the origin thereby removing the effect of spherical spreading. The HRTF will be defined as

$$H(\theta, \rho, \mu) = -\frac{\rho}{\mu} \sum_{m=0}^{\infty} \frac{h_m(\mu\rho)}{h_m{}'(\mu)} L_m(cos\theta)(2m+1) \tag{6}$$

The nondimensional group delay distance d_g for the spherical-head model can be expressed in terms of the phase of H given by the function γ as

$$d_g = \frac{d}{d\mu}\gamma(\theta, \rho, \mu) \tag{7}$$

To recover the time-delay one needs to multiply d_g by $e_0 * c$.

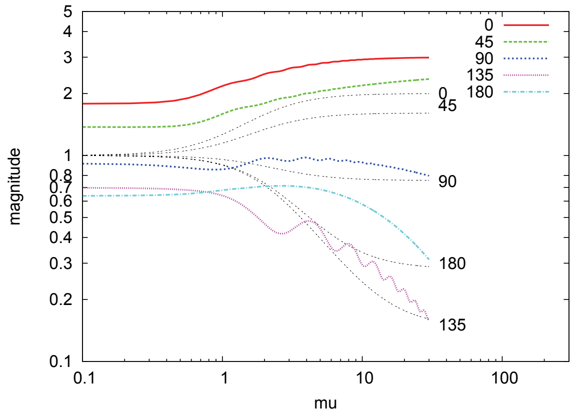

Fig. 4. Magnitude of H and H_{Brown}.

In Figure 4, the comparison of magnitude of spherical model and that given by Brown-Duda[11] for ρ=3 is presented. The Brown-Duda result is given by the black lines. The angle θ between the source and the observer is given in degrees and written at the right side of the curve. In both models the response decreases with the angle for the angular values between 0 to 150 degrees. For angles greater than 150 degrees the response of spherical model is greater than that obtained by Brown-Duda model. The lack of amplitude agreement is the result of the large ρ assumption used in the Brown-Duda model.

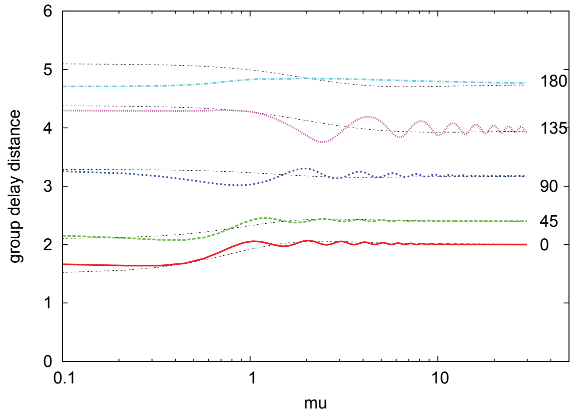

Fig. 5. Group delay distance d_g and d_{Brown}

In Figure 5, the group delay distance for the two model is compared for ρ=3. Here d_g is the group delay distance for the spherical model. The group delay distance d_{Brown} for the Brown-Duda model is given by the black lines. The group delay at low frequency under NLOS conditions is over estimated by the Brown-Duda model. This is expected since the delay in the Brown-Duda model is evaluated using the geometric optics approximation. The group delay distance[12] in the limit of μ approaching zero for the spherical head model is

$$\lim_{\mu->0} d_g = \rho - \frac{3}{2}cos\theta + \frac{1}{\rho} \tag{8}$$

3.2. Software implementation

The software program was written in C/C++ on a Linux Ubuntu system. The open-source nature of the system allows one to customize the system libraries. The screen display of user interface was comprised of the listener's head at the center, and a graphical representation of sound source. The head has the left and right ears, and the nose. The nose specifies the center of face and the location of zero yaw angle. The spatial input of the system is the yaw angle. The distance between the user and sound source is a fixed parameter. For simplicity, the display maps the observer's angle and distance data as angle and location changes to the sound source, rather then the observer. For each new sound source location, the HRTF filters are updated to accommodate the spatial hearing changes.

Using HRTF given in Eqn (6), we constructed a HRTF filter. HRTF filter output is the convolution of sampled HRTF impulse response and input data. Overlap-Add Method with FFT is used to evaluate the block convolutions. WiiRemote and computer communication is established using a Bluetooth adapter. BlueZ is bundled in Ubuntu system, and is the official Linux Bluetooth stack. The Bluetooth management on our systems is handled by BlueZ software.

WiiRemote Motion Plus communication and data collection are done using libcwiid development library. The most recent version of this library contains calls to receive WiiRemote Motion Plus data. The information and download can be found at http://abstrakraft.org/cwiid/. Cwiid download package contains libcwiid, and sample application programs that use this library[13]. We modified this library by adding an initialization program. We implemented an interface module between the simulation and the library. The simulation program calls the interface functions to connect to WiiRemote, to initialize it, and to collect data. The received data is yaw angular velocity which is converted to angle value. A scale factor is defined to map the angle value to yaw value (θ). The data is collected into a buffer at certain time intervals. The data collection process is controlled by a timer.

4. RESULTS

Our initial procedure to test the system was done against the horizontal rotations of WiiRemote controller on a flat surface. During these experiments, real-time plots were generated to measure error. After responses of the system were proved to be in an acceptable range of error, we tested the continuous yaw angle change under a user's control.

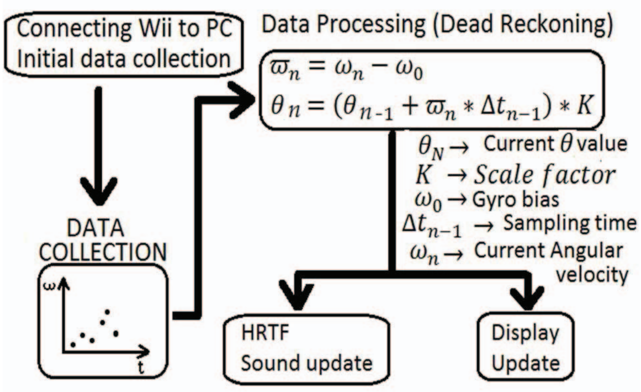

Fig. 6. Flow chart of the system.

Our tests showed that HRTF simulation could receive reliable real-time yaw inputs using WiiRemote Motion Plus controller. The accuracy can be and has to be improved in next phases. The stability of the system depends very much on the initial state. It is required that the simulation gets activated only when the WiiRemote is at the initial state of rest.

5. CONCLUSION

We are interested in designing an affordable headphone system that can be used to demonstrate spatial localization and can serve to create a realistic virtual environment. We implemented a head tracking interface using WiiRemote Motion Plus, which controlled the HRTF filter parameters.

The head tracking interface worked and allowed us to stage a real-life application. In our future work, we will work on accuracy of yaw data. The long run of the system causes accumulation of data errors. An additional feedback system to offset error using error estimation, or/and periodic calibration of WiiRemote is necessary. We will continue our work to decrease the error rate.

6. REFERENCES

[1] Blauert, J. P. (1969). *"Sound localization in the median plane,"* Acustica, 22, 205-213.

[2] Blauert, J. P. (1997). *"Spatial Hearing,"* MIT Press, Cambridge, MA, pp. 50-93,373-374.

[3] Wightman, F. L., and Kistler, D. J. (1989). *"Headphone simulation of free-field listening. I: Stimulus synthesis,"* J. Acoust. Soc. Am. 85, 858-867.

[4] Wightman, F. L., and Kistler, D. J. (1997). *"Factors affection the relative salience of sound localization cues,"*

Binaural and Spatial Hearing in Real and Virtual Environments, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum, Mahwah, NJ), pp. 1-23.

[5] Rayleigh, L. (1907). *"On our perception of sound direction,"* The London, Edinburgh and Dublin Philosophical Magazine and J. of Science, 6, 214-232

[6] Hartley, R.V.L and Fry, T.C. (1921). *"The binaural localization of pure tones,"* Phy. Rev., 18,431-442.

[7] Kuhn, G.F. (1977). *"Model for interaural time-differences in the azimuthal plane,"* J. Acous. Soc. Am.,18,157-167.

[8] Brungart, D.S. and Rabinowitz, W.R. (1999). *"Auditory localization of nearby source. Head-related transfer functions,"* J. Acous. Soc. Am., 106, 1465-1479.

[9] Wiibrew (2009). *"Wiimote/Extension Controllers,"* http://wiibrew.org/wiki/Wiimote/Extension_Controllers.

[10] Random Hacks of Boredom/Genious (2009). *"Motion Plus and Nunchuck together on the Arduino,"* http://randomhacksofboredom.blogspot.com/2009/07/motion-plus-and-nunchuck-together-on.html

[11] Brown, C.P and Duda, R.O. (1998). *" A structural model for binaural sound synthesis,"* IEEE Trans on Speech and Audio Processing,6, 476-488.

[12] Patel, H. (2004). *"A rigid sphere model for the Head Related Transfer Function and its Application to Audio Signal Processing,"* Master's Thesis, UMASS Lowell. Available from Dissertations and Theses database. (UMI No. 1421864)

[13] Ubuntu Documentation(2009). *"CWiiD"*. https://help.ubuntu.com/community/CWIID