한국소음진동공학회 2005년도 추계학술대회논문집, pp. 616~619.

가상입체음향 시스템 개발을 위한 새로운 머리전달 함수(HRTF) 소개 및 응용

New HRTFs (Head Related Transfer Functions) and Applications to the Virtual Acoustic Imaging Systems

김 영 태† ·김 선 민*
Youngtae Kim, Sunmin Kim

Key Words : HRTF(Head Related Transfer Function; 머리전달함수), Virtual Acoustic Imaging System(가상입체음향 시스템)

ABSTRACT

An extensive data base of HRTFs (Head Related Transfer Functions) has been established in order to work with high qualities of 3D acoustic appliances. The basic specifications of the measurement presented are that a spatial resolution of 10° in elevation angles (ranging from -40° to 90°) and uniform spatial resolution of 5° in azimuth angles. The distance from the measurement sources to the centre of the dummy head is 2m and the sampling frequency is 48 kHz and the quantisation depth is 24-bits. The data is presented for three arrangements of pinna models (large, small and no pinna) which were combined with the open and blocked ear canal cases to give a total of 6 sets of measurements. The data base may contribute to show promise of providing useful applications of 3D sound.

1. Introduction

A virtual acoustic imaging system attempts to generate an illusion in a listener of being in a virtual acoustic environment which is entirely different from that of the space in which the listener is actually located. In order to produce the virtual environment to a listener, binaural technology [1] is often used. The principle of this technology is to control the sound field at the listener's ears so that the reproduced sound field coincides with what would be produced when the listener is in the desired real sound field. One way of achieving this is to use a pair of loudspeakers at different positions in a listening space with the help of signal processing to ensure that appropriate binaural signals are obtained at the listener's ears. In such cases, knowledge of HRTFs is essential in the synthesis of binaural signals, such as those used in virtual reality applications.

Previous measurements of the HRTFs have been published in the literatures [2-3] such as those reported by MIT (Massachusetts Institute of Technology, USA) Media Lab and CIPIC (the Center of Image Processing and Integrated Computing) Interface Lab (UC Davis, USA). The MIT measurements consist of the left and right ear impulse responses from a loudspeaker mounted 1.4 m from the KEMAR (Knowles Electronics Mannequin for Acoustics Research) dummy head. Total 710 different positions were sampled with 44.1 kHz. The CIPIC database measurement was performed on 45 subjects at 25 different azimuths and 50 different elevations (1250 directions). The sampling frequency they used is 44.1 kHz. The measurements were performed for the small and large pinnae at a distance of 1m from the loudspeaker.

Recently, HRTF measurements were performed by SAIT (Samsung Advanced Institute of Technology)/ISVR (Institute of Sound and Vibration Research, University of Southampton, England) Joint Lab.[4]. In this paper, the features of Samsung HRTF are presented and some practical possibilities will be introduced for the virtual acoustic imaging systems.

2. Measurement apparatus and procedures for new HRTFs

The measurement for Samsung HRTFs was performed on a KEMAR dummy head at 72 different azimuth and 14 different angles of elevation giving a total of 1008 directions. The azimuth angle resolution is 5° and the elevation angle resolution is 10°. The sampling frequency used is 48 kHz and the quantization depth is 24 bits Pink noise was used to obtain HRIRs (Head Related Impulse Responses). The KEMAR model was equipped with model DB 65/DB66 for large pinnae or DB 60/DB61 for small pinnae. In order to

† 삼성종합기술원(SAIT)
E-mail : yt65.kim@samsung.com
Tel : (031) 280-8192, Fax : (031) 280-9207
* 삼성전자(주)

give a total of 6 sets of the measurements (see Table 1), three arrangements of pinna models were combined with the open and blocked ear canal cases. In the measurement for the open ear canal, the calibrated microphone was mounted on the Zwislocki coupler [5] used for simulate the terminating impedance of the eardrum.

Measurement	Pinna	Open/blocked
1	Small	Open
2	Large	Open
3	No	Open
4	Small	Blocked
5	Large	Blocked
6	No	Blocked

Table 1 The 6 sets of measurements

All measurements were undertaken in the ISVR anechoic chamber with dimensions 9.15m by 9.15m by 7.32m. As depicted in Figure 1, the inner radius of the new half-circular metal frame was designed to be 2m from the centre to the frontal surface of loudspeakers. 14 loudspeakers were arranged on the half circle from -40° below the horizontal plane with respect to the centre up to 90° directly overhead at every 10°. Instead of spinning the heavy circular rig, the KEMAR model can be rotated through 360°. The characteristic responses of the loudspeakers are for the 14 loudspeakers appear reasonably identical up to the high frequency range (10~15 kHz) with a deviation of less than 2dB. All the measurement procedures and processes were automatically controlled by a computer with the measurement system. Once the measurement equipment is calibrated in the anechoic chamber, the rotating turntable for KEMAR model and the switch box are initialized, the KEMAR faces straight towards the array of loudspeakers (0° azimuth), and the pink noise is played from the loudspeaker at the lowest elevation (-40° elevation). So, impulse responses with KEMAR model for all elevations can be measured at an initial azimuth angle. And then impulse responses will be measured at each azimuth angle with the KEMAR model mounted on the turntable that rotates clockwise from 0° to 355° in -5° decrements (for totally 72 azimuth angles). This procedure depicted in Figure 2 is repeated for 6 cases listed in Table 1.

In order to obtain the response functions of KEMAR model only, the free-field transfer functions measured will be used to the deconvolution process (which is also called the free-field equalization process). The free-field equalisation is undertaken by simply dividing the HRTF at each angle with the free field transfer function at the same angle (obtained by taking an FFT (Fast Fourier Transform) of the free field responses. Another main feature of the Samsung HRTF database is non-minimum phase version. In real time filter applications with HRTFs, it is generally recommended that the system designer takes into account for the non-minimum phase characteristics. This will be particularly relevant in the case of designing inverse filters for example crosstalk cancellers.

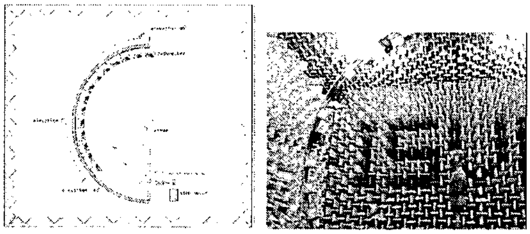


Figure 1 A schematic diagram of the measurement rig

3. Validation of the measurements

3.1 Validation of the results

This section presents the results of the measurement conducted at the SAIT/ISVR Joint Lab. A large amount of data was collected in this measurement and a summary is presented in this section.

Figure 2(a) shows sound that comes directly in front of the listener in the horizontal plane (azimuth = 0° and elevation = 0°). As would be expected, it can be found that the sound levels and the arrival times at both of the ears are about the same. Figure 2(b) shows sound that comes from the left side in the horizontal plane (azimuth = 270° and elevation = 0°). The attenuation due to the head shadowing effect is clearly seen at both of the two ears. It is also observed that the signal at the left ear arrives slightly before zero. This is due to that the left ear is closer to the sound source than the middle of the head. The right ear signal arrives later than zero, and can be measured to be approximately 0.7 ms after the signal arrived to the left ear (0.7 ms corresponds to approximately half the perimeter of the KEMAR dummy head). The frequency response at the left ear is higher than at the right ear, due to the shadowing effect of the head. The response at the left ear has a gain above 0 dB at most frequencies, which indicates a pressure build-up due to reflections from the head and the pinna. The pinna notch is also clearly visible at approximately 9 kHz for the left ear. Figure 2(c) presents the result when sound comes from the right side in the horizontal plane (azimuth = 90° and elevation = 0°). This is very similar result with that in Figure 2(b).

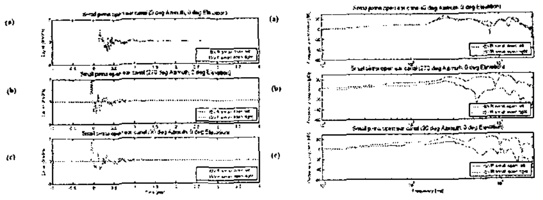
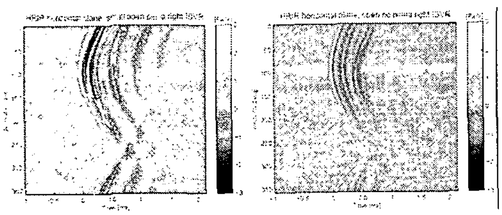


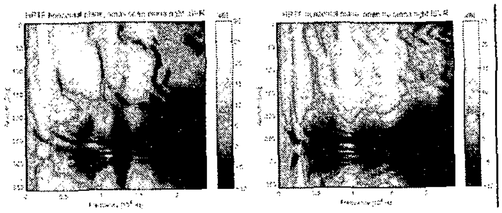
Figure 2 Time and frequency domain responses of the small pinna with an open ear canal for three different azimuth angles at 0° elevation: (a) 0°, (b) 90° and (c) 270°

Images representation of the measured HRIRs in the horizontal plane is shown in Figure 3. The images show the impulse response of the right ear as a function of azimuth angle and time, where the strength of the response is represented by brightness. It is seen that the sound is strongest and arrives soonest when it is coming from the right side (azimuth = 90°) and similarly it is weakest and arrives latest when it is coming from the left side (azimuth = 270°). The difference between the shortest and longest arrival time is about 0.7 ms. In Figure 3(a), it is possible to see the initial sequence of fluctuations due to pinna reflections compared with the image of Figure 3(b). The response when the source is in front is similar to the response when the source is at the back. In the open ear canal measurements the ear canal resonance is present. It is possible to see that the wavelength of the ear canal resonance is approximately 0.5 ms, which corresponds to 2 kHz. The images in Figure 4 show the frequency response for the right ear of the KEMAR. The frequency response is shown for all measured azimuths in the horizontal plane. As expected, the response is greatest when the source is at 90° and the directed into to the right ear. The response is weakest when the source is at 270° on the opposite side of the head. The peak around 2-3 kHz is due to the ear canal resonance..



(a) (b)

Figure 3 HRIR in the horizontal plane for the right (a) small open (b) no pinna (time domain responses).



(a) (b)

Figure 4 HRTF in the horizontal plane for the right (a) small open (b) no pinna (frequency domain responses).

3.2 Characteristics comparison between large, small and no pinnae

In the previous section, the basic features of HRTFs and HRIRs have been presented. This section shows characteristic comparisons between small and large pinnae and no pinna characteristic for the HRTFs.

In Figure 5 and 6, the small and the large open pinna ear canal measurements are compared. It is seen that the frequency response is very similar up to approximately 7 kHz. It can be found in the figure the different characteristics of the two small/large pinna models above 7 kHz for the left side and above the 10 kHz for the right side. This will be presented that the pinna is only associated with high frequency features in the HRTFs.

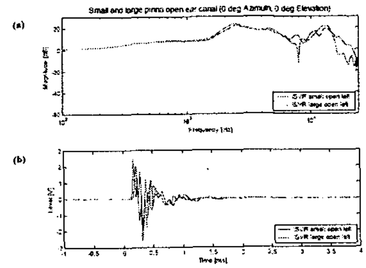


Figure 5 HRIR (b) and HRTF (a) the left small open pinna and the large open pinna.

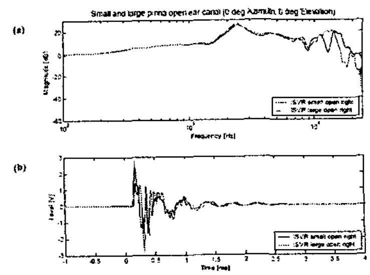


Figure 6 HRIR (b) and HRTF (a) the right small open pinna and the large open pinna.

In Figure 7, the small open ear canal measurement is compared to the no pinna open ear canal measurement. It is seen that the frequency response is very similar up to approximately 1 kHz and that it deviates significantly above that frequency. It is also seen that the pinna results in an amplification of the sound level in the region from 1kHz to 8kHz. In the time domain it is seen that the impulse response without the pinna is shorter than with the small pinna and thus the secondary wave due to reflections between 0.2 ms to 0.6 ms should be related to the presence of the pinna.

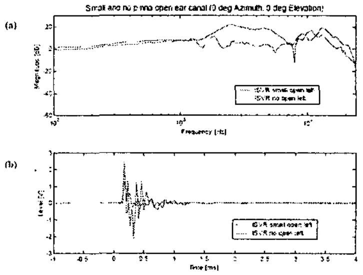


Figure 7 HRIR (b) and HRTF (a) for the left small open pinna and the no open pinna.

4. 3D sound associated with the HRTF

The goal of the 3D sound generation system designer is to ensure that the reproduced signals in the ears of a listener, through sound generation devices, are equivalent to those detected under real listening conditions. For example, in such cases the direction of arrival of the sound can be controlled by filtering the original signals through an appropriate set of HRTFs. In recent years a number of applications associated with the HRTF have grown rapidly due to the advances in computing power and the possibilities to implement digital filters. The 3D sound techniques associated with the HRTF are possible to be applied to various kinds of the digital media platforms including mobile phone. For example in the binaural synthesis, HRTFs are used to convolve with the input audio data for producing the virtual sound image at arbitrary position. The crosstalk canceller using by HRTFs gives one complete control over the sound field at a number of target positions for the loudspeaker applications [6]. Also, the HRTFs is widely used for stereo enhancement applications [7-8] and 2 channel virtualizer [9-10] in order to provide improved virtual sound images.

5. Conclusions

The basic specifications of the measurement presented are that a spatial resolution of 10° in elevation angles (ranging from -40° to 90°) and uniform spatial resolution of 5° in azimuth angles. The distance from the measurement sources to the centre of the head is 2m and the sampling frequency is 48 kHz and the quantisation depth is 24-bits. To conclude, the KEMAR HRTF database measurement conducted by the SAIT/ISVR Joint Lab have resulted in the most extensive KEMAR HRTF database yet available. The result for ITDs and ILDs are presented for the 6 sets of measurements in the horizontal plane. The results in the horizontal plane show features such as the head shadowing effect, arrival time difference, pinna and shoulder reflections. The no pinna measurement is a new measurement, which has not been performed before and it aims to be a starting point for future modelling of pinna responses.

References

(1) D. R. Begault, 1994, "3-D sound for virtual reality and multimedia," AP Professional, Cambridge, MA.

(2) W. Gardner, K. Martin, 1994, "HRTF measurement of a KEMAR Dummy-Head Microphone," Technical report #280, MIT Media Lab.

(3) V. Algazi, R. Duda, D. Thompson, 2001, "The CIPIC HRTF Database," New Paltz, New York.

(4) P. Mannerheim, M. Park, T. Papadopoulos, P.A. Nelson, 2004, "The measurement of a database of Head Related Transfer Functions," ISVR Contract report No.04/07, University of Southampton, England.

(5) J. Zwislocki, 1970, "An acoustic coupler for earphone calibration," Special report LCS-S-7, Syracus University, New York, USA.

(6) P. A. Nelson, 1994, Active Control of Acoustic Fields and the Reproduction of Sound. Journal of Sound and Vibration, vol. 177(4), pp. 447-477.

(7) Stereo enhancement system, 1997, US5661808, SRS Labs, Inc..

(8) Method and system for sound expansion, 1999, US5974153, Q Sound Labs.

(9) System and method for localization of virtual sound, 2001, US6307941 B1, Desper products, Inc.

(10) Sound recording and reproduction systems, 1997, WO 97/30566, Adaptive audio Ltd..