SIMULATION OF INDIVIDUAL HEAD-RELATED TRANSFER FUNCTIONS USING AN IMPROVED HEAD-NECK-SHOLDIER MODEL

Tingwei Yin, Guangzheng Yu, Yingyang He
Department of Physics, School of Sciences, South China University of Technology, Guangzhou, China

e-mail:scgzyu@scut.edu.cn

For the head-related transfer functions (HRTFs) below about 5 kHz, the spectral cues caused by torso (shoulder) reflections have been simulated using simple geometric models that consist of spherical head and torso. However, the spherical head and torso are different from that of human subject. In order to more accurately approximate the HRTFs at low frequencies, the simple geometric model of head-neck-shoulder (HNS) consisting of an ellipsoidal head, a cylindrical neck and a semi-ellipsoidal torso is proposed in this work. The structural parameters of HNS model are simulated from the scanned KEMAR model with shoulder. For comparison, the scanned KEMAR without pinnae and the head-neck-torso (HNT) model consisting of spherical head and torso are also adopted. Then, the HRTFs of three models are calculated using the boundary elementary method (BEM). The effective frequency limit is still below 5 kHz. Results indicate that the HRTF magnitude spectra of the HNS model are more similar to that of the KEMAR model without pinnae, especially below about 2.0 kHz.

1. Introduction

The sound wave propagation from a point source to two ears of human subject in free field can be regarded as a linear time-invariant system and described by head-related transfer functions (HRTFs). In theory, HRTF is the function of direction, distance and frequency of the sound source, and also related to subject’s anatomical parameters. Because HRTFs contain localization cues of the sound sources, they are critical data-basis of binaural auditory and virtual auditory display. In practice, HRTF data can be acquired through experimental measurements. Recently, some research groups have established the HRTF databases for human subjects or KEMAR. However, it is time-consuming to measure and set up a complete HRTF database. As an alternative, HRTFs at low frequencies can be approximately obtained using simple geometric models of the head and torso.

In previous work, the spherical head model and ellipsoidal head model have been first used to approximate subject’s head. In order to introduce the effects of torso-reflections, the simple geometric models of head-and-torso (HAT) and head-neck-torso (HNT) have been adopted. Then, the elevation localization cues of HRTFs at low frequencies caused by torso-reflections have been confirmed. However, the head and torso of human subject are not spherically symmetric. Comparing to the spherical torso, the real torso will provide more significant shoulder reflections in the left and right sides, but do weaker reflections in the front and back directions.
As an improvement, the simple geometric model of head-neck-shoulder (HNS) consisting of an ellipsoidal head, a cylindrical neck and a semi-ellipsoidal torso is proposed. In addition, the appropriate offset of the ellipsoidal head centre relative to the coordinate origin is also compensated, which is similar to the pinna offset method described in Ref. 7.

For the sound source at high elevations, the semi-ellipsoidal shoulder may enhance the reflections in the left and right sides, but weaken the reflections in front and back. For comparison, HRTFs of HNT model, HNS model, and KEMAR model without pinnae are calculated using boundary element method (BEM). In order to improve computation efficiency, the fast multipole BEM (FMBEM) and acoustic reciprocity principle are adopted too. The frequency limit is still blow 5 kHz.

2. Method

2.1 Calculation model

The origin $O$ of the coordinate system is set at the midpoint between two ear canal entrances. The point source locates at the spherical coordinate $(r_0, \theta_0, \phi_0)$ or position vector $r_0$, as shown in Fig. 1. $r_0$ is the distance from point source to the head centre; $\phi_0$ represents the sound elevation, its bounds is $-90^\circ \leq \phi_0 \leq 90^\circ$; $\theta_0$ denotes the sound source azimuth, its bounds is $0^\circ \leq \theta_0 \leq 360^\circ$. Three planes used in the following include the horizontal plane ($z_0 = 0; \phi_0 = 0^\circ$), median plane ($y_0 = 0; \theta_0 = 0^\circ$ or $180^\circ$), frontal plane ($x_0 = 0; \theta_0 = 90^\circ$ or $270^\circ$), respectively. Using the spherical coordinate $(r_0, \theta_0, \phi_0)$ in Fig. 1(a), HRTFs can be defined by,

$$
\begin{align*}
H_L (r_0, \theta_0, \phi_0, f, a) &= \frac{P_L (r_0, \theta_0, \phi_0, f, a)}{P_0 (r_0, f)} \\
H_R (r_0, \theta_0, \phi_0, f, a) &= \frac{P_R (r_0, \theta_0, \phi_0, f, a)}{P_0 (r_0, f)}
\end{align*}
$$

(1)

where, $P_L$ and $P_R$ are complex sound pressures at the left and right ears, respectively; $P_0$ is complex sound pressure at the head centre with the subject absent; $f$ represents the sound source frequency; $a$ denotes the anatomical parameters. The diagrams of HNS model in front and right view are shown in Fig. 1(b) and 1(c), respectively. Based on the scanned KEMAR in Fig. 2(c), all parameters are measured and shown in Tab. 1.

![Figure 1](image1.png)

Figure 1. (a) Space coordinate system; (b) the HNS model in front view; (c) the HNS model in right view.

The three calculation models, including the HNT, HNS, and scanned KEMAR without pinnae are shown in Fig. 2. The parameters of the HNS model in Fig. 2(b) are indicated in Table 1. For example, the semi-ellipsoidal shoulder is determined by radii of $a_8, a_9/2, a_{10}/2$. The ellipsoidal head is described by radii of $a_1/2, a_2/2, a_3/2$. The center of the ellipsoidal head locates at the offset coordinate $(x = 0, y = a_5, z = a_4)$. The HNT model in Fig. 2(a) consists of a spherical head with radius of
21st International Congress on Sound and Vibration (ICSV21), Beijing, China, 13-17 July 2014

87.5 mm, a cylinder neck with radius of 60 mm, and a spherical torso with radius of 169 mm. In addition, the scanned KEMAR model without pinnae is indicated in Fig. 2(c).

### Table 1. Parameters of HNS model, and the measured results based on the scanned KEMAR model.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Symbol</th>
<th>Results (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head width</td>
<td>$a_1$</td>
<td>152</td>
</tr>
<tr>
<td>Head height</td>
<td>$a_2$</td>
<td>191</td>
</tr>
<tr>
<td>Head depth</td>
<td>$a_3$</td>
<td>224</td>
</tr>
<tr>
<td>Head centre offset down</td>
<td>$a_4$</td>
<td>23</td>
</tr>
<tr>
<td>Head centre offset back</td>
<td>$a_5$</td>
<td>1.5</td>
</tr>
<tr>
<td>Neck height</td>
<td>$a_6$</td>
<td>54</td>
</tr>
<tr>
<td>Neck diameter</td>
<td>$a_7$</td>
<td>60</td>
</tr>
<tr>
<td>Shoulder height</td>
<td>$a_8$</td>
<td>105</td>
</tr>
<tr>
<td>Shoulder depth</td>
<td>$a_9$</td>
<td>198</td>
</tr>
<tr>
<td>Shoulder width</td>
<td>$a_{10}$</td>
<td>432</td>
</tr>
</tbody>
</table>

Figure 2. The calculation models: (a) HNT model; (b) HNS model; (c) KEMAR model without pinnae.

### 2.2 Solution

For all calculation models in Fig. 2, their HRTFs can be obtained via numerical simulation methods, such as boundary element method (BEM). In order to calculate the HRTFs as defined by Eq. (1), the binaural pressures should be solved first. The pressure at left or right ear is the special situation of sound pressure at any field point outside of the model. If the point source locates at spatial vector $r_0$, the pressure $P(r, r_0, f)$ at any field point $r$ satisfies the Helmholtz equation,

$$\nabla^2 P(r, r_0, f) + k^2 P(r, r_0, f) = 0$$

(2)

where, $k = 2\pi f / c$ represents the wave number, $c$ is the sound speed in the air. If the surfaces of the calculation models can be regarded as the acoustic rigid boundary, the sound pressure $P(r, r_0, f)$ satisfies the expression as follow,

$$\left[ \frac{\partial P(r, r_0, f)}{\partial n} \right]_g = 0$$

(3)

In infinite region, the sound pressure $P(r, r_0, f)$ also satisfies the Sommerfeld radiation condition,

$$\lim_{r \to \infty} \left[ \frac{\partial P(r, r_0, f)}{\partial r} + jk P(r, r_0, f) \right] = 0$$

(4)

The surface sound pressures $P(r, r_0, f)$ of the model are first obtained via submitting the boundary conditions Eq. (3) and (4) into the initial condition Eq. (1), then, the sound pressure at any field point including the left or right ear can be derived from the surface pressures of the model, as described in Ref. 11. In order to improve the computation efficiency, the principle of acoustic reciprocity and fast multipole boundary element method (FMBEM) are also adopted in this work.
3. Results

In this work, the distance dependence of HRTFs is not considered, and the far-field HRTFs at source distance of 2.0 m are used to investigate the influences of various shoulder reflections. The sound source directions are distributed in three planes, including the horizontal plane, median plane, frontal plane, as defined in section 2.1. Because the influences of shoulder reflections on HRTFs are more significant in the left and right sides, the left ear HRTFs with sound source pointing to the target ear are compared first in section 3.1. Then, the HRTF variation trends influenced by various shoulder reflections are studied in detail in section 3.2.

3.1 Comparison in a single direction

When the sound source locates at the coordinate \((r_0 = 2.0 \text{ m}, \theta_0 = 270^\circ, \phi_0 = 0^\circ)\), the logarithmic magnitude spectra of left ear HRTFs are shown in Fig. 3. The HRTFs of KEMAR, HNT and HNS are denoted by solid line, dotted line and solid line with mark “*”, respectively. Results indicate that the first magnitude peaks below about 2.0 kHz are similar between the HRTFs of HNS and KEMAR, and more than 6.0 dB. However, the peak of HNT appears below about 1.0 kHz, and less than about 4.5 dB. Therefore, the HRTFs of HNS are more similar to that of KEMAR without pinnae. However, as the frequency increases, the similarity becomes insignificant, which could be caused by the detailed shape differences between HNS and KEMAR.

![Figure 3. The left ear HRTFs of three models under the conditions of \((r_0 = 2.0 \text{ m}, \theta_0 = 270^\circ, \phi_0 = 0^\circ)\).](image)

3.2 Comparisons in multi-direction

The left ear HRTFs in horizontal plane, median plane and frontal plane are separately arranged in three rows, as shown in Fig. 4. The results of HNT, HNS and KEMAR are shown in Fig. 4(a), 4(b) and 4(c), respectively. Overall, below about 2.0 kHz, the logarithmic magnitude spectra of HNS are similar to that of KEMAR. The detailed discussions are given as follows.

In the horizontal plane, for both HNS and KEMAR, the first magnitude peak in ipsilateral HRTFs \((180^\circ \leq \theta_0 \leq 360^\circ)\) arises from about 0.5 kHz to 2.0 kHz. However, for the HNT model, the first peak appears only between 0.5 kHz and 1.0 kHz. These results are consistent with that in section 3.1. In addition, the “bright spot” effect of HNT disappears in the contralateral HRTFs of HNS and KEMAR. For the HNS model, the disappearance of “bright spot” could be caused by the...
center coordinate offset of the ellipsoidal head, which results in the unsymmetrical propagation paths of sound waves to the contralateral ear and thereby improves the HRTF simulation.

In the median plane, the logarithmic magnitude spectra of HNT present small variations as the elevation increases. However, for HNS and KEMAR, the magnitude spectra are weaken significantly at low elevations below about −40° due to the stronger shadow of shoulder (relative to the spherical torso), and enhanced at elevations from about −40° to 0° due to the weaker shadow of shoulder (relative to the spherical torso).

In the frontal plane, comparing with the HRTFs of HNT, the ipsilateral HRTFs of HNS at elevations around horizontal plane are enhanced due to the stronger reflections from the ellipsoidal shoulder; the contralateral HRTFs of HNS on the ipsilateral HRTFs are weakened due to the greater shadow effects of the ellipsoidal head. In addition, the results also reveal that the HRTFs of HNS and KEMAR without pinnae are more similar.

Therefore, the HRTFs of HNS are more similar to that of KEMAR at low frequencies, and the HNS model consisting of an ellipsoidal head, a cylinder neck and a semi-ellipsoidal shoulder can be looked as the improved model of HNT.

![Figure 4](image-url)

**Figure 4.** The logarithmic magnitude spectra of the left ear HRTFs. Results in horizontal plane, median plane and frontal plane are shown in the first, second, and third rows, respectively. In columns (a), (b), and (c), the results in various planes are acquired from HNT, HNS, and KEMAR without pinnae, respectively.

### 4. Conclusions

In this work, the improved head-neck-shoulder (HNS) model consisting of an ellipsoidal head, a cylindrical neck and a semi-ellipsoidal shoulder is proposed to approximate the HRTFs at low frequencies. The parameters of the HNS model are measured from the scanned KEMAR model with shoulder. Overall, the logarithmic magnitude spectra of HRTFs are similar between the HNS model and KEMAR without pinnae, especially below about 2.0 kHz. Therefore, the HNS model can
be regarded as the improved version of the previous HNT model, and breeds more accurate HRTFs at low frequencies. In practice, the HNS model is easily constructed once the individual anatomical parameters are obtained, which contributes to the personalized approximation of HRTFs at low frequencies. In the future, the psychoacoustic experiments could be used to evaluate the influence of the improved HNS model on the elevation localization cues.

REFERENCES