Audio Engineering Society
Convention Paper

Presented at the 115th Convention
2003 October 10–13 New York, New York

VIRTUAL ACOUSTIC SYSTEM WITH A MULTICHANNEL HEADPHONE

Ingyu Chun, P. A. Nelson

ISVR, University of Southampton, Highfield, Southampton, SO17 1BJ, United Kingdom

ABSTRACT

The performance of current virtual acoustic systems is highly sensitive to the geometry of individual ear at high frequencies. The objective of this paper is to study a virtual acoustic system which may not be sensitive to individual ear shape. The incident sound field around ear is reproduced by using a multi-channel headphone. The results of computer simulations show that the desired sound pressure at the eardrum can be successfully replicated in a virtual acoustic environment by using a multi-channel headphone.

1. INTRODUCTION

In everyday life, we can easily locate and interact with auditory events in three-dimensional (3-D) space. 3-D sound systems or virtual acoustic systems try to control and manipulate spatial auditory perception of a listener. There are many applications related to virtual acoustic systems. For example, interactive virtual reality systems combine 3-D audio technology with 3-D video technology [1]. Teleconferencing systems also use virtual reality systems with a limited bandwidth of signals for the transmission of human voices [2]. Architectural acousticians often use auralization systems to hear the acoustics of designed rooms or auditoria. Virtual acoustics can be useful for home theatre systems, HDTV (High Definition Television), and 3-D games.

The goal of virtual acoustic systems is to improve the ability of audio systems to produce virtual acoustic environments such that listeners cannot tell the difference between real sound images and the virtual sound images that are produced by such systems. The major cues for human sound localisation are provided by the interaural time difference (ITD) cue, the interaural level difference (ILD) cue, and the spectral cue. The ITD cue is dominant at low frequencies below about 1.5kHz [3]. The spectral cue is dominant at high frequencies. The spectral cue depends on the size and geometrical shape of the listener's pinna at high frequencies [4]. However, the performance of the current virtual acoustic systems is highly sensitive to geometry of individual pinnae at high frequencies. Therefore, an individualized Head-Related-Transfer-Function (HRTF) should be used to produce a virtual acoustic environment. However, each person has such different individualized HRTF and the calculation of

the individualized HRTF is a very time-consuming and expensive process. This is one of the most critical problems associated with virtual acoustic systems. The objective of this study is to investigate virtual acoustic systems with multi-channel headphones, which are not sensitive to the geometry of individual ear, even at high frequencies.

2. BOUNDARY SURFACE CONTROL

If there is no source in a given volume V bounded by a surface S, the solution of the inhomogeneous wave equation in a single frequency sound field reduces to the Kirchhoff-Helmholtz integral equation that is given by [5]

$$C(\mathbf{x})\,p(\mathbf{x}) = \int_S \left(g(\mathbf{x}|\mathbf{x}_s)\frac{\partial p(\mathbf{x}_s)}{\partial n} - p(\mathbf{x}_s)\frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n} \right) dS \quad (1)$$

where $\mathbf{x}$ is a position vector, $\mathbf{x}_s$ is a position vector on the boundary surface S, $\mathbf{n}$ is the unit outward normal vector on S, p is the complex acoustic pressure, $g(\mathbf{x}|\mathbf{x}_s)$ is the free space Green function, and $C(\mathbf{x})$ is equal to one if $\mathbf{x}$ is within V, zero if $\mathbf{x}$ is outside V, and 0.5 if $\mathbf{x}$ is on a smooth boundary S. This integral equation can be solved if the boundary conditions on the boundary surface S are given.

The Kirchhoff-Helmholtz integral equation can be interpreted in terms of the following boundary surface control principle [6]: the pressure field within the volume V can be controlled by controlling the pressure and its gradient on the surface S. In this case, the Green function and its gradient can be regarded as constants determined by the boundary shape. A sound field reproduction system based on the boundary surface control principle tries to reproduce the sound pressure and its gradient on the boundary surface S enclosing the controlled volume V in the secondary field so that these variables are identical to those in the primary field. In practice, the control surface S is divided into N control points $\mathbf{x}_i$ (i=1...N). The pressure gradient at $\mathbf{x}_i$ can be approximately calculated from the two point pressures at $\mathbf{x}_i$ and $\mathbf{x}_i + c\mathbf{n}_i$ where $\mathbf{n}_i$ is the corresponding normal vector and c is coefficient that is small compared to the wavelength, which can be given by

$$\frac{\partial p(\mathbf{x}_i)}{\partial n} = \frac{p(\mathbf{x}_i + c\mathbf{n}_i) - p(\mathbf{x}_i)}{c}. \quad (2)$$

Therefore, the sound pressures at the $2N$ control points are recorded in the primary field and reproduced in the secondary field. This method can also solve the nonuniqueness problem of the Kirchhoff-Helmholtz integral equation [7].

Now assume that the sound source generates single frequency sound for simplicity. In the primary field, the sound pressures on the boundary surface control points can be given by

$$\mathbf{p}_p = \mathbf{g}\, q_p \quad (3)$$

where $\mathbf{g}$ is the acoustic transfer impedance vector relating the complex sound pressure vector $\mathbf{p}_p$ on the boundary surface control points to the strength q_p of the real source in the primary field. The number of elements of both vectors $\mathbf{g}$ and $\mathbf{p}_p$ equals the number of control points on the boundary surface. The vector $\mathbf{p}_p$ can be recorded with a given q_p in the primary field.

In the secondary field, the sound pressures on the boundary surface control points can be written as

$$\mathbf{p}_s = \mathbf{G}\,\mathbf{q}_s \quad (4)$$

where $\mathbf{G}$ is the acoustic transfer impedance matrix relating the complex sound pressure vector $\mathbf{p}_s$ on the boundary surface control points to the strength vector $\mathbf{q}_s$ of secondary sources. If the number of control points is L and the number of secondary sources is M, then $\mathbf{G}$ is an $L \times M$ matrix. The mth column of the matrix $\mathbf{G}$ can be calculated by recording the vector $\mathbf{p}_{sm}$ when only one secondary source produces sound with a given mth element q_{sm} of the vector $\mathbf{q}_s$ in the secondary field. Then the matrix $\mathbf{G}$ is thus given by:

$$\mathbf{G} = [\mathbf{p}_{s1}/q_{s1} \quad \mathbf{p}_{s2}/q_{s2} \quad \cdots \quad \mathbf{p}_{sM}/q_{sM}]. \quad (5)$$

To replicate the primary sound field in the secondary sound field, $\mathbf{p}_s$ must be equal to $\mathbf{p}_p$. If L equals M, we can find the exact $\mathbf{q}_s$ to match $\mathbf{p}_s$ with $\mathbf{p}_p$ by inverting the matrix $\mathbf{G}$. However, in practice, if we try to reduce the number of sources, M is smaller than L and the matrix $\mathbf{G}$ is not invertible. In this case, we can find the optimal $\mathbf{q}_s$ to minimize the following cost function:

$$J = \sum_{i=1}^{L} \left| p_{p,i} - p_{s,i} \right|^2 = (\mathbf{p}_p - \mathbf{G}\,\mathbf{q}_s)^{\mathrm{H}} (\mathbf{p}_p - \mathbf{G}\,\mathbf{q}_s) \quad (6)$$

where the subscript i denotes the ith element of pressure vector and the superscript H means the Hermittian transpose. The optimal strength $\mathbf{q}_{so}$ of the secondary sources that minimize J is given by [5]

$$\mathbf{q}_{so} = \left(\mathbf{G}^{\mathrm{H}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{H}}\mathbf{p}_p \qquad (7)$$

and the minimum value of J corresponding $\mathbf{q}_{so}$ is given by

$$J_o = \mathbf{p}_p^{\mathrm{H}}\left(\mathbf{I} - \mathbf{G}\left(\mathbf{G}^{\mathrm{H}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathrm{H}}\right)\mathbf{p}_p . \qquad (8)$$

Ise [6] suggested the development of a virtual acoustic system based on the boundary surface control principle, which controls the pressure and its gradient on the boundary surface in the secondary field so that they are identical to those in the primary field. The secondary field can be reproduced by multiple secondary sources located at arbitrary positions outside the controlled volume. The performance of this system is independent of a listener inside the controlled volume. However, a great number of loudspeakers are needed if we want to control a sound field at high frequencies. Virtual acoustic systems based on the boundary surface control principle may be a robust system for which performance is not dependent upon the listener. However, the number of sources is a critical problem in practice. To reduce the required number of sources, the control region of the system should be as small as possible. If we confine the controlled volume of such a virtual acoustic system to sound fields around the ear, the number of control points on the boundary surface are reduced and then the number of loudspeakers can be reduced. If just a few loudspeakers were enough to control the sound field around ear using the boundary surface control principle, the virtual sound field could be reproduced by using a multi-channel headphone.

When we hear sounds, incoming sound waves are modified by the listener's head, torso and pinna. The closer the control point is located to the listener's ear, the more sensitive is the individual difference of the ear to the sound field. However, if the incident sound field on the ear is reproduced, which is independent of the geometry of the ear, a performance of the virtual acoustic system may be independent of the listener's ear. The incident sound field on the ear consists of the direct waves from the source, the waves scattered by objects and walls in a room, and the waves scattered by the listener's head and torso. When the incident sound field on the ear is reproduced, the listener hears virtual sound through the listener's own pinna filter. This critical problem of the virtual acoustic systems based on HRTF technology may therefore be solved by reproducing the incident sound field using a multi-channel headphone.

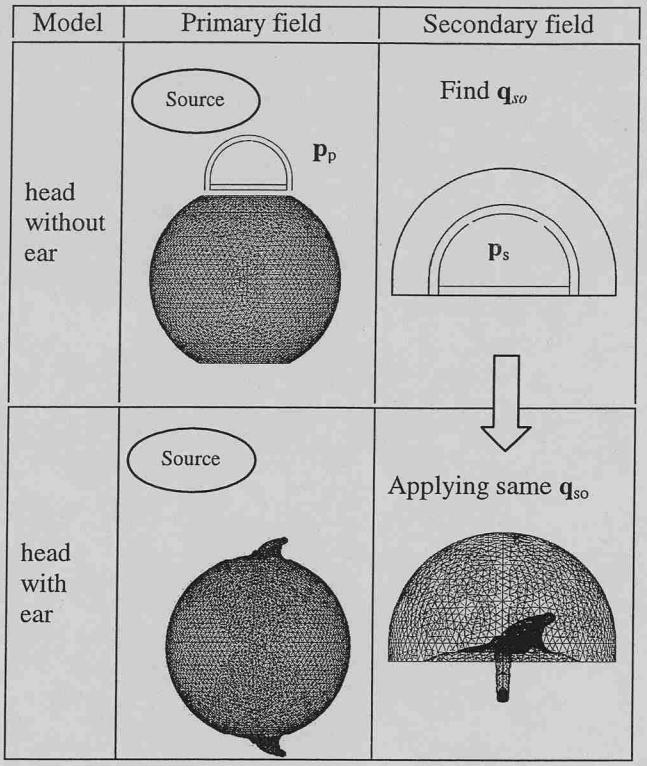

Figure 1: Incident sound field reproduction method

Figure 1 shows the concept of the incident sound field reproduction method based on the boundary surface control principle. First assume that a listener perceives an auditory event in a free-field environment and that his head is in a fixed position. In the primary field, a double layer boundary control surface is put on a head model without the ear. Then the sound pressures $\mathbf{p}_p$ on the boundary surface control points, which is caused by the source, are recorded. In the secondary field, the same double layer boundary control surface is put on a headphone model without the ear. Then the acoustic transfer impedance matrix $\mathbf{G}$ relating the sound pressures $\mathbf{p}_s$ on the boundary surface control points to the headphone source strengths $\mathbf{q}_s$ is measured. Then the optimal headphone source strengths $\mathbf{q}_{so}$ can be calculated using the least square method explained above. The same headphone source strength can be applied to the headphone model with the ear.

The objective of this system is to reproduce the same sound pressure at a listener's eardrum in the secondary field as the sound pressure at the eardrum in the primary field. If the sound pressures at the control points are recorded in the primary field and accurately reproduced in the secondary field, the sound pressure at the eardrum can be exactly

replicated. This hypothesis will be verified in this paper. The detailed theory of the incident sound field reproduction method is explained in the next section.

3. THEORY OF THE INCIDENT SOUND FIELD REPRODUCTION METHOD

3.1. Determination of the required secondary source distribution

To study the theory of the incident sound field reproduction method, simplified head models are used. The case of the human head is similar to the case of the two scattering bodies and acoustic sources in a free field. In this case, the surface of the human head is divided into two surfaces, an ear surface and a head surface excluding the ears. Each surface acts as a scattering body.

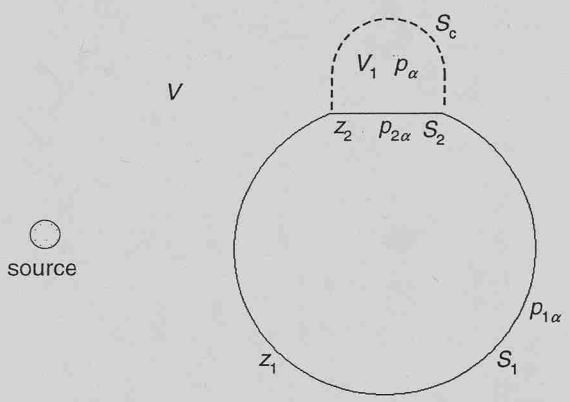

Figure 2: Primary sound field α in the case of a human head without ears

Figure 2 shows the primary field produced by a sound source or sources and a human head without the ears in a free field, which simulates the real acoustic environment. This is called the sound field α for simplicity in this paper. The acoustic source strength distribution Q_{vol} in an unbounded acoustic domain V outside the scattering bodies is assumed to be known. The bounding surface of the head denotes S_1 and S_2. The surface S_1 is the surface of the head with the ear excluded. The surface S_2 is the flat surface of the dummy head, which simulates the ear. All surfaces throughout this paper are assumed to be locally reacting surfaces [8]. The specific acoustic impedance z_1 of the surface S_1 and z_2 of the surface S_2 are also assumed to be known.

Now consider the control volume V_1 bounded by the surface S_c and the surface S_2. The surface S_c is a transparent imaginary surface as depicted in Fig. 2. The primary sound sources are assumed to be located outside the volume V_1. When the vector $\mathbf{x}$ is in the volume V_1, p_α denotes the sound pressure in the sound field α. Since the volume V_1 is inside the volume V, the sound pressure p_α inside the volume V_1 at a single frequency can be written as one of the following equations:

$$C(\mathbf{x})\,p_\alpha(\mathbf{x}) = p_{in}(\mathbf{x}) - \int_{S_1} H_1(\mathbf{x}|\mathbf{x}_s)\,p_{1\alpha}(\mathbf{x}_s)\,dS - \int_{S_2} H_2(\mathbf{x}|\mathbf{x}_s)\,p_{2\alpha}(\mathbf{x}_s)\,dS \tag{9}$$

or

$$C(\mathbf{x})\,p_\alpha(\mathbf{x}) = p_{in}(\mathbf{x}) - \int_{S_1} H_1'(\mathbf{x}|\mathbf{x}_s)\,v_{n1\alpha}(\mathbf{x}_s)\,dS - \int_{S_2} H_2'(\mathbf{x}|\mathbf{x}_s)\,v_{n2\alpha}(\mathbf{x}_s)\,dS \tag{10}$$

where

$$p_{in}(\mathbf{x}) = \int_{V_1} Q_{vol}(\mathbf{x}_v)\,g(\mathbf{x}|\mathbf{x}_v)\,dV, \tag{11}$$

$$H_1(\mathbf{x}|\mathbf{x}_s) = \frac{j\omega\rho_0 g(\mathbf{x}|\mathbf{x}_s)}{z_1(\mathbf{x}_s)} + \frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n} \quad (z_1(\mathbf{x}_s)\neq 0), \tag{12}$$

$$H_1'(\mathbf{x}|\mathbf{x}_s) = j\omega\rho_0 g(\mathbf{x}|\mathbf{x}_s) + z_1(\mathbf{x}_s)\frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n} \quad (z_1(\mathbf{x}_s)<\infty), \tag{13}$$

the vector $\mathbf{x}$ is in the volume V_1, the vector $\mathbf{x}_s$ is on the surface S_1 or S_2, the vector $\mathbf{x}_v$ is in the volume V, the pressure $p_{1\alpha}$ is the sound pressure on the surface S_1, ω is the angular frequency, ρ_0 is the mean density of the fluid, the velocity $v_{n1\alpha}$ is the normal particle velocity on the surface S_1, the impedance z_1 is the specific acoustic impedance z_1 of the surface S_1, and $p_{2\alpha}$, $v_{n2\alpha}$, z_2, H_2 , H'_2 are the corresponding values of these variables on the surface S_2. The sound field $p_{\text{in}}(\mathbf{x})$ denotes the incident sound field that can be interpreted as the sound field in the absence of the scattering body. That is, the incident sound field emitted from the source is not changed by the presence of scatterers and other sources. The second and third term on the right hand side of Eq. (9) or (10) is the a scattered sound field that is produced by the interaction of the incident sound field with the scatterers. Equation (9) can be used when the specific impedance is not equal to zero, that is, the scattering surface is not a soft boundary. Equation (10) can be used when the specific impedance is finite, that is, the scattering surface is not a rigid boundary. Throughout this paper, the type of transfer

function H such as those appearing in Eq. (12) will be used, since all surfaces are assumed not to be of the soft boundary type.

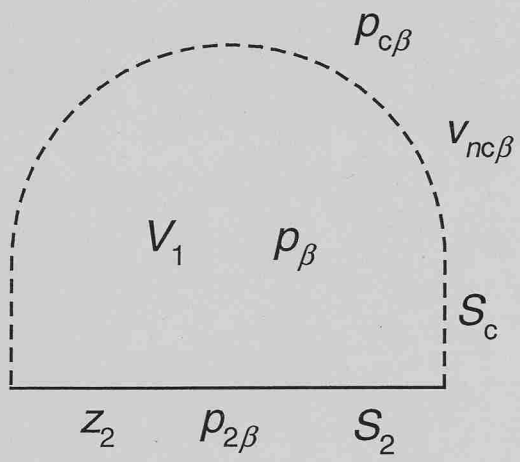

Figure 3: Secondary sound field β in the case of a human head without ears

Figure 3 shows the secondary field produced by different sound sources, which simulates the virtual acoustic environment assuming the same scattering surface S_2 as that in the sound field α. This is called the sound field β in this paper. When the vector $\mathbf{x}$ is in the volume V_1, p_β. denotes the sound pressure in the sound field β. If the continuous transparent monopole and dipole source layers are placed on the surface S_c, the sound pressure p_β in the volume V_1 in a single frequency sound field can be written

$$C(\mathbf{x})\,p_\beta(\mathbf{x}) = -\int_{S_2} H_2(\mathbf{x}|\mathbf{x}_s)\,p_{2\beta}(\mathbf{x}_s)\,dS - \int_{S_c}\left(j\omega\rho_0 v_{nc\beta}(\mathbf{x}_s)\,g(\mathbf{x}|\mathbf{x}_s) + p_{c\beta}(\mathbf{x}_s)\frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n}\right)dS \quad (14)$$

where the pressure $p_{c\beta}$ and the velocity $v_{nc\beta}$, are on the surface S_c. The distribution of the monopole source strength on the surface S_c is given by $q_{c\beta}(\mathbf{x}_s)=-v_{nc\beta}(\mathbf{x}_s)$ and the distribution of the dipole source strength on the surface S_c is given by $\mathbf{f}_{c\beta}(\mathbf{x}_s)=p_{c\beta}(\mathbf{x}_s)\mathbf{n}$. The sound pressure $p_{c\beta}(\mathbf{x}_s)$ and the normal velocity $v_{nc\beta}(\mathbf{x}_s)$ on the surface S_c should be obtained to reproduce the same sound field as the sound field α inside the control volume V_1. If the sound field β is the same as the sound field α in the volume V_1, that is, $p_\alpha(\mathbf{x})=p_\beta(\mathbf{x})$ where the vector $\mathbf{x}$ is in V_1 and $p_{2\alpha}(\mathbf{x}_s)=p_{2\beta}(\mathbf{x}_s)$ where the vector $\mathbf{x}_s$ is on the surface S_2, then the following equation results from subtracting Eq. (14) from Eq. (9):

$$p_{in}(\mathbf{x}) + p_{sc1}(\mathbf{x}) = -\int_{S_c}\left(j\omega\rho_0 v_{nc\beta}(\mathbf{x}_s)\,g(\mathbf{x}|\mathbf{x}_s) + p_{c\beta}(\mathbf{x}_s)\frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n}\right)dS \quad (15)$$

where $p_{sc1}(\mathbf{x})$ is given by

$$p_{sc1}(\mathbf{x}) = -\int_{S_1} H_1(\mathbf{x}|\mathbf{x}_s)\,p_{1\alpha}(\mathbf{x}_s)\,dS\,. \quad (16)$$

This equation shows that the strengths of the continuous monopole and dipole source layers on the surface S_c can be obtained in order to reproduces both the incident sound field $p_{in}(\mathbf{x})$ produced by the source in the sound field α and the scattered sound field $p_{sc1}(\mathbf{x})$ produced by the head surface excluding the ear surface in the sound field α. If this scattered sound field $p_{sc1}(\mathbf{x})$ is regarded as another incident sound field in the volume V_1, the secondary sources reproduce only the total incident sound field. This means that the scattered sound field from the ear surfaces does not need to be reproduced because it is determined by only the total incident sound field including the head scattered sound field and it is produced by the ear surface itself.

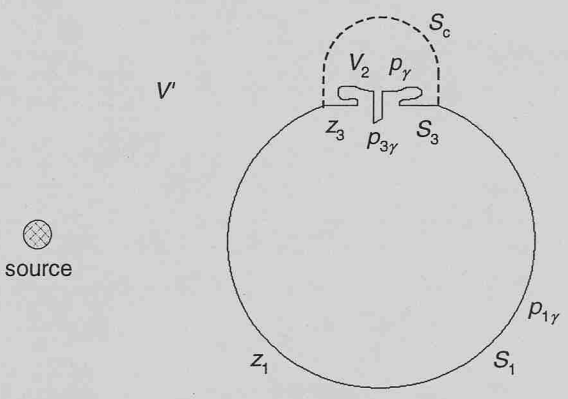

Figure 4: Primary sound field γ in the case of a human head with an ear

3.2. Conditions for production of the virtual sound field

Figure 4 shows another primary field produced by the same sound source distribution and the same human head surface as that in the sound field α but with a different ear surface. In this case, a realistic human ear surface S_3 is used. This is called the sound field γ in this paper. The specific acoustic impedance z_3 of

the surface S_3 is assumed to be known. Now consider the control volume V_2 bounded by the surface S_c and the surface S_3. When the vector $\mathbf{x}$ is in the volume V_2, p_γ denotes the sound pressure in the sound field γ. The sound pressure p_γ in the volume V_2 in a single frequency sound field can be written as

$$C(\mathbf{x})\,p_\gamma(\mathbf{x}) = p'_{in}(\mathbf{x}) - \int_{S_1} H_1(\mathbf{x}|\mathbf{x}_s)\,p_{1\gamma}(\mathbf{x}_s)\,dS - \int_{S_3} H_3(\mathbf{x}|\mathbf{x}_s)\,p_{3\gamma}(\mathbf{x}_s)\,dS \quad (17)$$

where $p_{1\gamma}$ is the sound pressure on the surface S_1, $p_{3\gamma}$ is the sound pressure on the surface S_3, $p'_{in}(\mathbf{x})$ is the same as Eq. (11) except for the integral over V_2 and $H_3(\mathbf{x}|\mathbf{x}_s)$ is given by

$$H_3(\mathbf{x}|\mathbf{x}_s) = \frac{j\omega\rho_0 g(\mathbf{x}|\mathbf{x}_s)}{z_3(\mathbf{x}_s)} + \frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n} \quad (18)$$

The incident sound field is again the sound field in the absence of the scattering body and so is not influenced by the scattered sound field or the scattering body. That means the incident sound field in the sound field α is same as that in the sound field γ even though the scattered sound field in the sound field α is different from that in the sound field γ. Therefore, the incident sound field $p'_{in}(\mathbf{x})$ and $p_{in}(\mathbf{x})$ can be extended to the domain in which the scattering body is located. Then the incident sound field $p'_{in}(\mathbf{x})$ is equal to $p_{in}(\mathbf{x})$ where the vector $\mathbf{x}$ is in the union $V_1 \cup V_2$. Equation (17) can be rewritten as

$$C(\mathbf{x})\,p_\gamma(\mathbf{x}) = p_{in}(\mathbf{x}) + p_{sc2}(\mathbf{x}) - \int_{S_3} H_3(\mathbf{x}|\mathbf{x}_s)\,p_{3\gamma}(\mathbf{x}_s)\,dS \quad (19)$$

where

$$p_{sc2}(\mathbf{x}) = -\int_{S_1} H_1(\mathbf{x}|\mathbf{x}_s)\,p_{1\gamma}(\mathbf{x}_s)\,dS \quad (20)$$

The scattered sound field $p_{sc2}(\mathbf{x})$ is produced by the head surface excluding the ear surface in the primary sound field γ.

Figure 5 shows another secondary field produced by different sound sources and the same scattering surface S_3 as that in the sound field γ. This is called the sound field δ in this paper. When the vector $\mathbf{x}$ is in the volume V_2, p_δ denotes the sound pressure in the sound field δ.

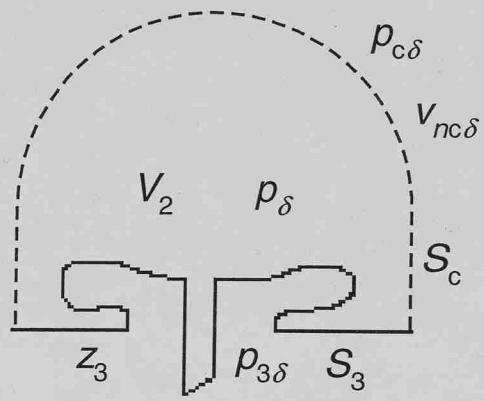

Figure 5: Secondary sound field δ in the case of a human head with an ear

If the continuous transparent monopole and dipole source layers are placed on the surface S_c, the sound pressure p_δ in the volume V_2 in a single frequency sound field can be written

$$C(\mathbf{x})\,p_\delta(\mathbf{x}) = -\int_{S_3} H_3(\mathbf{x}|\mathbf{x}_s)\,p_{3\delta}(\mathbf{x}_s)\,dS - \int_{S_c}\left(j\omega\rho_0 v_{nc\delta}(\mathbf{x}_s)\,g(\mathbf{x}|\mathbf{x}_s) + p_{c\delta}(\mathbf{x}_s)\frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n}\right)dS \quad (21)$$

where the position vector $\mathbf{x}$ is in the volume V_2. If the same monopole and dipole source strengths obtained in the sound field β are applied to the sound field δ, that is, $p_{c\delta}(\mathbf{x}_s) = p_{c\beta}(\mathbf{x}_s)$ and $v_{nc\delta}(\mathbf{x}_s) = v_{nc\beta}(\mathbf{x}_s)$ where the vector $\mathbf{x}_s$ on the surface S_c, the following equation results from Eq. (15).

$$\begin{aligned} &-\int_{S_c}\left(j\omega\rho_0 v_{nc\delta}(\mathbf{x}_s)\,g(\mathbf{x}|\mathbf{x}_s) + p_{c\delta}(\mathbf{x}_s)\frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n}\right)dS \\ &= -\int_{S_c}\left(j\omega\rho_0 v_{nc\beta}(\mathbf{x}_s)\,g(\mathbf{x}|\mathbf{x}_s) + p_{c\beta}(\mathbf{x}_s)\frac{\partial g(\mathbf{x}|\mathbf{x}_s)}{\partial n}\right)dS \\ &= p_{in}(\mathbf{x}) + p_{sc1}(\mathbf{x}) \end{aligned} \quad (22)$$

where the vector x is in the intersection $V_1 \cap V_2$. This shows that the sound field reproduced by the secondary surface sources in the sound field δ is the same total incident sound field in the sound field β as that in the case of the flat surface S_2 in place of the ear surface S_3. Equation (21) can be rewritten as

$$C(\mathbf{x})\,p_\delta(\mathbf{x}) = p_{in}(\mathbf{x}) + p_{sc1}(\mathbf{x}) - \int_{S_3} H_3(\mathbf{x}|\mathbf{x}_s)\,p_{3\delta}(\mathbf{x}_s)\,dS \quad (23)$$

By subtracting Eq. (23) from Eq. (19) , the following equation results

$$C(\mathbf{x})\{p_\gamma(\mathbf{x}) - p_\delta(\mathbf{x})\} = -\int_{S_3} H_3(\mathbf{x}|\mathbf{x}_s)\{p_{3\gamma}(\mathbf{x}_s) - p_{3\delta}(\mathbf{x}_s)\}\,dS + p_{sc2}(\mathbf{x}) - p_{sc1}(\mathbf{x}) \quad (24)$$

where the vector $\mathbf{x}$ is in the intersection $V_1 \cap V_2$. Strictly speaking, $p_{sc1}(\mathbf{x})$ is different from $p_{sc2}(\mathbf{x})$, that is, the sound field scattered from the surface S_1 in the sound field α is different from that in the sound field γ. This occurs since the geometry and boundary condition of the other scatterer, in this case ear surface, in the sound field α is different from that in the sound field γ. That is, $S_2 \neq S_3$ and $z_2 \neq z_3$. This is because the sound field produced by the one scatterer can be influenced by the sound field produced by the other scatterer. However, $p_{sc2}(\mathbf{x}) - p_{sc1}(\mathbf{x})$ can be assumed to be zero if one of the following requirements is met.

First, the geometry and boundary condition of the ear surface in the sound field α is similar to that in the sound field γ. That is, $S_2 \approx S_3$ and $z_2 \approx z_3$. At low frequencies, the detailed shape of the ear is not significant so this requirement can be fulfilled. However, at high frequencies, individual differences of the ear shape are significant so this requirement cannot be fulfilled.

Second, the pressure $p_{1\alpha}(\mathbf{x})$ is similar to the pressure $p_{1\gamma}(\mathbf{x})$ on the head surface S_1. If the scattered sound pressure produced on the head surface by different ear surfaces are similar, this requirement can be fulfilled. On the other hand, if the scattered sound pressure produced on the head surface by the ear surface is sufficiently weak compared to the pressure on the head surface, this requirement can also be fulfilled. At strong resonances of ear, the scattered sound field from the ear surface may be influential to the surface pressure on the head surface near the ear.

Third, the both $p_{sc1}(\mathbf{x})$ and $p_{sc2}(\mathbf{x})$ are much less than the other terms in Eq. (19) and Eq. (23). That is, the scattered sound field from the head surface S_1 is much weaker than the incident sound field from the sources or much weaker than the scattered sound field from the ear surface in the control volume. This requirement can be fulfilled when the source in the primary field is placed on the same side of the ear for example.

Thus it may reasonably be assumed that $p_{sc1}(\mathbf{x})$ is same as $p_{sc2}(\mathbf{x})$, that is:

$$\int_{S_1} H_1(\mathbf{x}|\mathbf{x}_s)\,p_{1\alpha}(\mathbf{x}_s)\,dS = \int_{S_1} H_1(\mathbf{x}|\mathbf{x}_s)\,p_{1\gamma}(\mathbf{x}_s)\,dS \quad (25)$$

where the vector $\mathbf{x}$ is in the intersection $V_1 \cap V_2$, and the vector $\mathbf{x}_s$ is on the surface S_1. This is the important assumption. This means that the head scattered sound field is not influenced by the ear scattered sound field. Because these head scattered sound fields can be regarded as other "incident" sound fields which are not influenced by the scattering body, $p_{sc1}(\mathbf{x})$ and $p_{sc2}(\mathbf{x})$ can be therefore extended to the domain in which the scattering body is located. Then the sound field $p_{sc1}(\mathbf{x})$ is equal to $p_{sc2}(\mathbf{x})$ where the vector $\mathbf{x}$ is in the union $V_1 \cup V_2$. Equation (24) is valid where the $\mathbf{x}$ is in the union $V_1 \cup V_2$ under this assumption. When the vector $\mathbf{x}$ is on the surface S_3, equation (24) can be rewritten as

$$\frac{1}{2}\{p_{3\gamma}(\mathbf{x}) - p_{3\delta}(\mathbf{x})\} = -\int_{S_3} H_3(\mathbf{x}|\mathbf{x}_s)\{p_{3\gamma}(\mathbf{x}_s) - p_{3\delta}(\mathbf{x})\}\,dS \quad (26)$$

This can be the sound field δ when $p_{c\delta}(\mathbf{x}_s)=v_{nc\delta}(\mathbf{x}_s)=0$. Therefore, $p_{3\delta}(\mathbf{x}_s)=p_{3\gamma}(\mathbf{x}_s)$ where the vector $\mathbf{x}_s$ on the surface S_3. Then, equation (24) is equal to zero and thus

$$p_\gamma(\mathbf{x}) = p_\delta(\mathbf{x}) \quad (27)$$

where the position vector $\mathbf{x}$ is in the volume V_2. When the same source strengths are applied, which are those obtained in the case of sound field β, the secondary sound field δ turns out to be the same as the primary sound field γ inside the control volume even though the geometry and boundary condition of the scattering body inside the control volume is different.

This theory demonstrates that if any sound field inside the control volume produced in the primary field is exactly reproduced in the secondary field, and all the conditions outside the control volume in both the primary and secondary fields remain the same, the secondary sound field is always same as the primary sound field. This is regardless of the geometry and boundary condition of the scattering

body inside the control volume and relies on the reasonable assumption that the sound field outside the control region is not changed irrespective of the scattering body within the control volume.

4. NUMERICAL MODELLING

The human head with or without the ear and headphones were modelled numerically to simulate an incident sound field reproduction system. Numerical meshes of the following models were generated by using the ANSYS software package [9]. The numerical human head was modelled as a sphere having a radius of 87.5 mm that is the average head radius for a number of individuals [10]. Although the head size and shape vary substantially within the population, the radius of a spherical head model is fixed throughout the numerical simulations for simplicity and consistency. The DB65 pinna was modelled, which is the artificial ear of the KEMAR dummy head. The numerical meshes of the DB65 pinna model were generated by using a laser scanning technique [11]. The numerical model of the ear canal was made by using the data of Johansen [12]. The canal length is 26 mm and the eardrum occupies a portion at the innermost 4 mm of the canal length. Figure 6 shows the numerical mesh of the DB65 pinna model with the ear canal model.

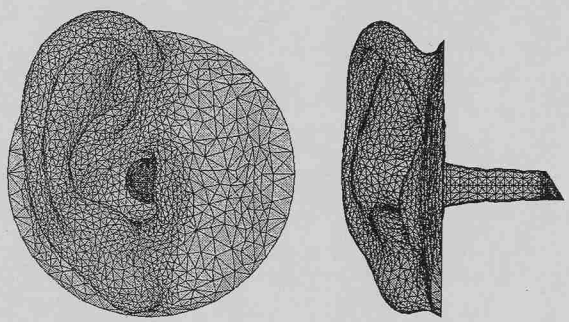

Figure 6: The DB65 pinna model with the ear canal model

The spherical head without the ear was modelled. Flat surfaces were placed at the locations where ears were supposed to be placed. Figure 7 shows the spherical head model with the flat surfaces. In this model, the flat surface having a radius of 30.5 mm is placed at the height of 82 mm from the center of the sphere. Then the spherical head with the ear was modelled. The DB65 ear is placed at a height of 82 mm from the center of the sphere. Figure 8 shows the spherical head model with the DB65 ear model.

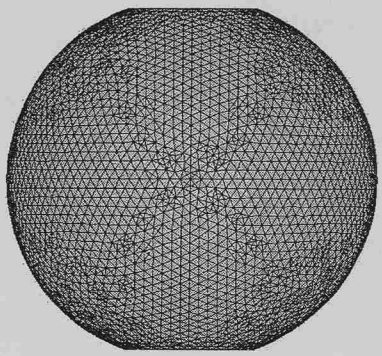

Figure 7: The spherical head model with the flat surfaces

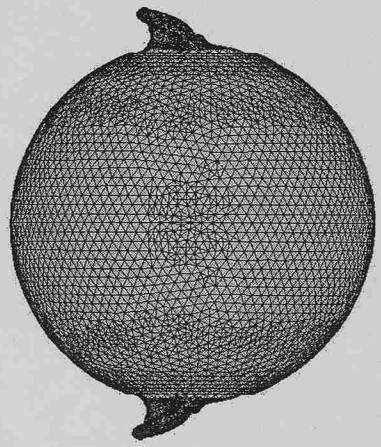

Figure 8: The spherical head model with the DB65 ear model

The headphone without the ear was also modelled. The headphone was modelled as a hemisphere having a radius of 60 mm and a cylinder with 10 mm width and the same radius attached to the hemisphere. The bottom of the headphone should contain the head model where the headphone is supposed to be located, so the flat surface is placed on the bottom of the headphone. Figure 9 shows the headphone model with the flat surface for the small control field. This model contains part of the spherical head model between the height of 77 mm and 82 mm from the center of the sphere. The headphone sources are modelled as piston-like vibrating surfaces having a radius of 5mm. Figure 10 shows 45 headphone sources that are distributed evenly, so two adjacent headphone sources spread at an angle of about 22.5°. Then the headphone with the ear was modelled. Part of the spherical head model with the DB65 ear model where the headphone was supposed to be located was attached to the bottom of the headphone model. Figure 11 shows the headphone model with the DB65 ear model with the same 45 headphone sources.

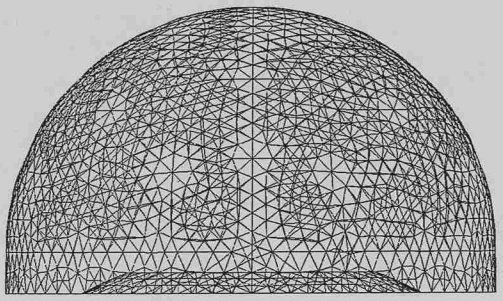

Figure 9: The headphone model with the flat surface

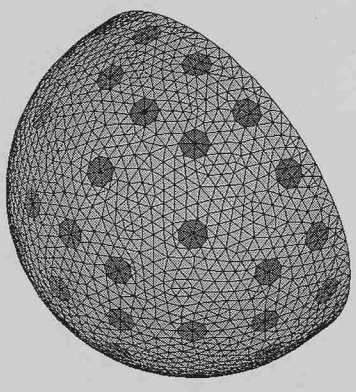

Figure 10: 45 headphone sources

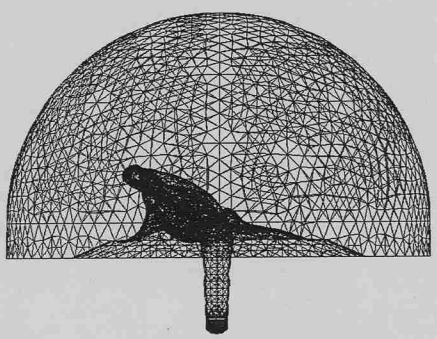

Figure 11: The headphone model with the DB65 ear model

The control field where the sound pressures were recorded and reproduced was modelled as double layers to measure sound pressure and its normal gradient. The distance between the inner and outer layer is 3.4 mm which is one-tenth of the wavelength at 10kHz. The control field is modelled as a hemisphere and the radius of outer layer is 30 mm. The number of points of the control field is 84.

The boundary of the models is discretized into a number of three-node triangular linear elements. The number of elements per wavelength depends on the required accuracy, the frequency range of interest and the calculation time. When the mesh resolution increases, both the accuracy of the numerical model and calculation time increase. At least six linear elements per wavelength are usually required to model acoustic wave propagation phenomena accurately by using the boundary element methods, and thus to identify the locations of peaks and troughs of sound pressure with reasonable accuracy [13]. The frequency range of interest is set to be 0 ~ 10kHz. Since the wavelength of a 10kHz sinewave is 34 mm when the speed of sound is 340 meter per second, the mesh resolution is set to be about 5 mm for the headphones and the spherical heads without the ear. Since the shape of ear is complicated and the measurement point is located at the eardrum, the mesh resolution around the eardrum is set to be 0.5 mm and that for the ear canal and pinna is set to be 2 mm.

5. NUMERICAL SIMULATION

To produce the virtual acoustic fields, the procedure used in the numerical calculations is as follows. First, the sound pressure on the control field is measured in the primary sound field when a sound source and the spherical head model without the ear are placed in a free field. Then the transfer impedance matrix between headphone sources and control points is calculated in the secondary sound field when the headphone model without the ear is placed in the free field. The optimal strengths of the headphone sources are calculated by multiplying the inverse transfer impedance matrix by the matrix of the desired sound pressure on the control field, which is measured in the primary sound field. The sound pressure at the eardrum is measured in the primary sound field when the same sound source as that in the first step and the spherical head model with the ear are placed in a free field. The sound pressure at the eardrum is measured in the secondary sound field when the headphone model with the ear is placed in the free field, and the same strengths of the headphone sources are applied. These are calculated in the third step. If the sound pressure at the eardrum measured in the fifth step is same as that in the fourth step, the virtual acoustic field is successfully produced. Sound pressures around numerical models are calculated by using the SYSNOISE software package that uses the boundary element method [13]. The optimal strengths of the headphone sources are calculated by using the MATLAB software package [14].

Human skin is nearly rigid but hair is absorbent material [15]. To design the head model properly, the boundary condition of head model should be

considered. However, the boundary surface of the spherical head model both with and without the ear is assumed to be rigid for simplicity. The ideal monopole sound source is used in the primary sound field, which is placed 1 meter away from the center of the head and its pressure amplitude is 90.97 dB ref 2×10^{-5} Pa at all frequencies. The sound pressures are calculated from 200Hz to 10kHz with constant frequency increments of 200Hz by using the direct boundary element method. In this simulation, the boundary condition of the headphone including the surface of the headphone sources is set to be perfectly absorbent, that is, of which the specific acoustic impedance is 416.5 Pa s m^{-1}. Note that the flat surface or the ear on the bottom of the headphone is still rigid. The monopole sound source in the primary sound field is put on the horizontal plane with constant increments of 15 degrees azimuth angle. The number of the overdetermination points is 50 up to 4kHz and 130 from 4kHz to 10kHz.

Because the number of the headphone sources is less than the number of control points, there will be some control error when the optimal strengths of the headphone sources are calculated. To assess this control error, the following average sound pressure level difference at the control points between the desired pressure and the reproduced pressure can be used:

$$E_{AD} = \frac{1}{L}\sum_{i=1}^{L}\left(20\log_{10}\frac{\left|\mathbf{p}_p\right|_i}{\left|\mathbf{G}\mathbf{q}_{so}\right|_i}\right) \quad [\text{dB}]. \qquad (28)$$

In this expression, $\mathbf{q}_{so}$ is the optimal headphone source strength in the secondary field, $\mathbf{G}$ is the acoustic transfer impedance matrix, $\mathbf{p}_p$ is the complex sound pressure at the control point measured in the primary field and L is the number of control points. Figure 12 shows the average sound pressure level difference in this case. This figure shows the control error is not significant in this case, for example, the average sound pressure level difference is below 0.3 dB up to 10kHz.

Figure 13 shows the sound pressure level measured at the eardrum of the DB65 ear when the primary source is at 60° azimuth, and Figure 14 shows that of the DB65 ear when the primary source is at 240° azimuth. The solid line represents desired sound pressure level measured in the primary sound field, and the dashed line represents reproduced sound pressure level measured in the secondary sound field.

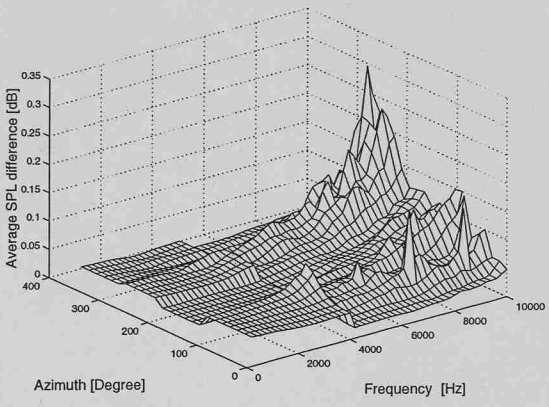

Figure 12: The average sound pressure level difference

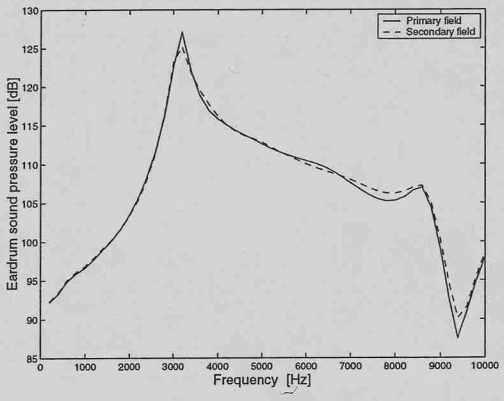

Figure 13: The sound pressure level measured at the eardrum of the DB65 ear when the primary source is at 60° azimuth

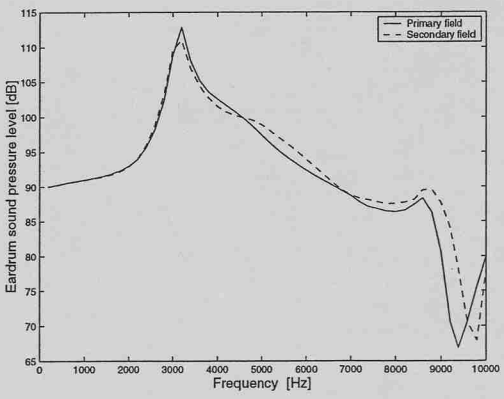

Figure 14: The sound pressure level measured at the eardrum of the DB65 ear when the primary source is at 240° azimuth

Some measure is needed to assess the effectiveness of the virtual acoustic system. Measures of eardrum sound pressure level difference to evaluate reproduction accuracy are suggested by

$$\text{Eardrum sound pressure level difference} = \left| 20\log_{10}\left(\frac{|p_{ep}|}{|\mathbf{g}^T \mathbf{q}_{so}|} \right) \right| \quad [\text{dB}] \tag{29}$$

where p_{ep} is the complex sound pressure measured at the eardrum in the primary field and $\mathbf{g}$ is the acoustic transfer impedance matrix between headphone source strengths and sound pressure at the eardrum in the secondary field. This is the difference of sound pressure levels measured at eardrum between desired sound pressure level in the primary field and reproduced sound pressure level in the secondary field. Zero means a perfect reproduction. Figure 15 shows this eardrum sound pressure level difference when the source for every 15° azimuth is on the horizontal plane in the primary field. Figure 16 shows the eardrum sound pressure level difference averaged over frequencies. Those figures show the eardrum sound pressure level difference throughout most horizontal angles up to 10kHz is below 1dB, which can be regarded as successful reproduction, except when the source is between 180° azimuth and 270° azimuth from 9kHz.

If a different ear model is used, a different frequency response will be obtained. However, this simulation shows the incident sound reproduction system can reproduce the sound field with any ear successfully because the optimal strengths of the headphone sources are obtained without the ear.

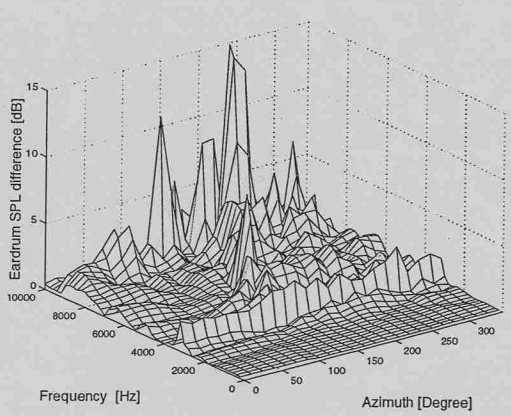

Figure 15: The eardrum sound pressure level differences

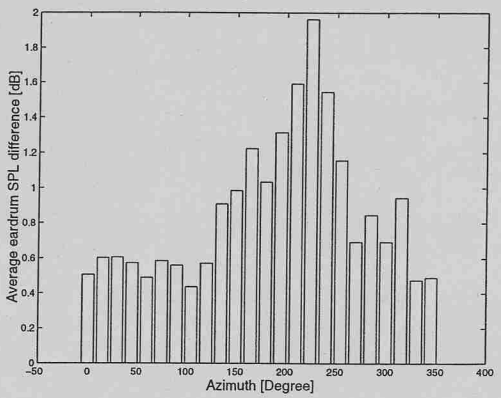

Figure 16: The average eardrum sound pressure level difference

6. DISCUSSION

Two things can be required to assess virtual acoustic systems; The difference between the eardrum sound pressure in the real environment and the eardrum sound pressure in the virtual environment, that is, reproductive accuracy should be small. When the ear is changed, the eardrum pressure difference should be still small. That means the system is robust. If the incident sound reproduction system for one arbitrary ear shows good reproductive accuracy, the system is robust because the optimal source strengths are obtained from the system without the ear.

There are the following three main sources of error and they are interdependent. Control errors are caused by inverting a non-square matrix. Numerical errors are caused by numerical calculation. Physical errors are caused for acoustical reasons.

Control errors mainly depend on the number and position of headphone sources and the size and shape of the control surfaces. If we put as many transducers on the headphone as there are control points, we can make nearly perfect reproduction. However, in practice, we can put just a few transducers on the headphone. Because perfect reproduction is not possible in practice, the optimal model of the headphone should be investigated. To reduce control errors, the number of headphone sources should be as many as possible and the number of control points should be as few as possible. The complexity of the incident sound field can raise the control error.

Numerical errors also arise from false numerical modelling of real geometry. This depends on mesh resolution, boundary condition and the shape of the model. It also depends on the numerical modelling

method used. If the mesh resolution is increased, the numerical error can be reduced but calculation time is increased. Numerical errors are increased as frequency is increased due to the decreased number of elements per wavelength and the non-uniqueness problem. To compare the various results of the numerical simulation, the modelling method is unified as the direct boundary element method and same numerical models are used consistently.

Even though control errors and numerical errors are not significant, reproductive accuracy can be poor. If the boundary condition of headphone surface is not perfectly absorbent, the scattered sound field from the headphone surface can make errors. If the angle of the incident sound wave matches the angle of the headphone source, the physical error can be reduced. As we discussed, if the assumptions are not valid, the reproduction accuracy can be degraded.

There are some parameters that can influence the performance of the incident sound field reproduction system. The number of headphone sources is critical. The more the headphone sources, the better the performance of the system, but the more difficult it becomes to make a practical system. The position of headphone sources determines range of good performance of the system. As the angle of the incident sound wave approaches the angle of the headphone source, the performance of the system improves. To cover all angles of an incident sound wave, the headphone sources should be distributed evenly. But, for example, if we focus on only the horizontal plane, the headphone sources should be placed on the horizontal plane. The boundary condition of headphone surface should be considered. The inner surface of the ideal headphone model should be perfectly absorbent to remove room modes inside the headphone and absorb the scattered sound field from the ear. However, practical headphones cannot have a perfectly absorbent boundary. The size and shape of the control field can be influential in the system performance. The bigger the control field is, the greater is the system robustness for various size of pinna and placements of the headphone on the head. However, the bigger the control field is, the worse the performance of the system is when the number of headphone sources is fixed. The size and shape of the headphone can also be influential. The size and shape of the headphone depends on the size and shape of the control field. If the headphone is too big, it is impractical to use. If the headphone is too small, that is, the distance from the headphone sources to the control points is too close, the performance of the system can be degraded when the headphone surface is not perfectly absorbent. Those parameters are interdependent, so optimisation of those parameters is needed to make an effective headphone system using as small a number of transducers as possible.

7. CONCLUSION

The most critical problem of the virtual acoustic imaging system is the high sensitivity of the performance of the system for the individual differences of the frequency response of the pinna. To solve this problem, a new virtual acoustic imaging system has been proposed, which uses a multi-channel headphone. This system is based on the boundary surface control principle derived from the Kirchhoff-Helmholtz integral equation. If all the sound field outside the control region is assumed to be the incident sound in the primary field, which is not changed regardless of the scattering body within the control volume, the secondary sound field is always the same as the primary sound field. This is regardless of the geometry and boundary condition of the scattering body within the control volume and is achieved by reproducing exactly the primary sound field with any scattering body in the control volume. Therefore, this analysis implies that the virtual sound field can be created inside the headphone regardless of the geometry and boundary condition of the ear by reproducing the sound field that is recorded using a dummy head without ears.

To minimize the difference between the desired sound field and the reproduced sound field, the least squares solution is used. The results of the numerical simulation using the boundary element method imply that we can make the robust virtual acoustic system with multi-channel headphone by using the incident sound field reproduction method. The design of the optimal multi-channel headphone will focus on the investigation of the optimal number of transducers and the optimal boundary condition of the headphone.

8. REFERENCES

[1] J. Blauert, H. Lehnert, J. Sahrhage, and H. Strauss, "An interactive virtual-environment generator for psychoacoustic research. I: architecture and implementation," *Acustica*, vol. 86, 94–102 (2000).

[2] D. R. Begault, "Virtual acoustic displays for teleconferencing: intelligibility advantage for telephone-grade audio," *J. Audio Eng. Soc.*, vol. 47, 824–828 (1999).

[3] F. L. Wightman, D. J. Kistler, "Factors affecting the relative salience of sound localization," in R. H. Gilkey and T. R. Anderson (eds), *Binaural and*

Spatial Hearing in Real and Virtual Environments (Lawrence Erlbaum Associates, New Jersey, 1997).

[4] E. A. G. Shaw, "Acoustical features of the human external ear," in R. H. Gilkey and T. R. Anderson (eds), *Binaural and Spatial Hearing in Real and Virtual Environments* (Lawrence Erlbaum Associates, New Jersey, 1997).

[5] P. A. Nelson, and S. J. Elliott, *Active Control of Sound*, (Academic Press, London, 1992).

[6] S. Ise, "A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems," *Acustica*, vol. 85, 78–102 (1999).

[7] S. Takane, Y. Suzuki, and T. Sone, "A new method for global sound field reproduction based on Kirchhoff's integral equation," *Acustica*, vol. 85, 250–257 (1999).

[8] A. D. Pierce, *Acoustics - An Introduction to Its Physical Principles and Application*, (The Acoustical society of America, New York, 1989).

[9] ANSYS 6.1., *Users manual*, (ANSYS Inc., Canonsburg)

[10] V. R. Algazi, C. Avendano, and R. O. Duda, "Estimation of a spherical-head model from anthropometry," *J. Audio Eng. Soc.*, vol. 49, 472–479 (2001).

[11] Y. Kahana, *Numerical Modelling of the Head Related Transfer Function*, (Ph.D. thesis, ISVR, University of Southampton, Southampton, 2000).

[12] P. A. Johansen, "Measurement of the human ear canal," *Acustica*, vol. 33, 349–351(1975).

[13] SYSNOISE 5.5., *Users manual*, (LMS International, Leuven).

[14] MATLAB 5.3., *Users manual*, (The MathWorks Inc., Natick).

[15] B. F. G. Katz, "Acoustic absorption measurement of human hair and skin within the audible frequency range," *J. Acoust. Soc. Am.*, vol. 108, 2238–2242 (2000).