ADVISE: A NEW METHOD FOR HIGH DEFINITION VIRTUAL ACOUSTIC DISPLAY

Shouichi Takane[*], *Yôiti Suzuki*[†], *Tohru Miyajima*[‡] *and Toshio Sone*[*]

[*] Faculty of Systems Science and Technology, Akita Prefectural University
84-4 Ebinokuchi, Tsuchiya, Honjo, Akita, 015–0055 Japan
E-mail: takane@bip.eis.akita-pu.ac.jp
[†] Research Institute of Electrical Communication
and Graduate School of Information Science, Tohoku University
2-1-1 Katahira, Aoba-ku, Sendai, Miyagi, 980-8577 Japan
[‡] Institute of Technology, Shimizu Corporation
3-4-17 Etchujima, Koto-ku, Tokyo, 135-8530 Japan

ABSTRACT

A method for high definition virtual acoustic display (VAD) based on a model called the "virtual sphere model" is introduced in this paper. This method is named "ADVISE" (Acoustic Display based on the VIrtual SpherE model). In *ADVISE*, a virtual sphere-shaped boundary is defined around a listener, and the sound transmission for the sound source in the original field to the entrances of the listener's ears is divided into two parts. One consists of Head-Related Transfer Functions (HRTFs) corresponding to the points on the boundary of the virtual sphere, and the other consists of Room Transfer Functions (RTFs) from the sound source to the points on the boundary. Then, these two kinds of transfer functions are convolved in real-time, taking into account dynamic changes in these functions due to the listener's head movement. In this paper, the theory of *ADVISE* is described in detail.

1. INTRODUCTION

Virtual acoustic displays (VADs) have been one of the strong and promising means to realize auditory display. Numerous studies on the methods for VADs exist[1, 2], and those methods are classified into four approaches.

The first approach is so called the multi-channel stereophonic system, the main aim of which is to simulate certain parts of subjective impression of the original sound field, such as presence and spaciousness. This approach, however, seems to be theoretically weak and inexact, and therefore, not useful for high definition auditory display.

In the second approach, the direct and/or reflected sound waves existing in a room are simulated[3]. A practical way to realize this approach is to distribute multiple loudspeakers around a listener. Each loudspeaker reproduces direct and reflected sounds from the corresponding directions. Reverberation is usually reproduced by exponentially decaying impulse responses uncorrelated among the loudspeakers. This type of auditory display is usually combined with the CAD system in architechtural acoustics[4].

In the third approach, the theory of sound field reproduction based on the Huygen's principle [5] or Kirchhoff-Helmholtz boundary integral equation is applied. Some researchers [7, 8, 9, 10] reported that the theoretically precise reproduction of sound field can be achieved with this type of VAD. However, a disadvantageous feature of the second and the third approaches is their large-scale hardware, which detracts portability and thus makes practical implementation difficult.

The fourth approach is based on the binaural reproduction of sound waves reaching the entrances of the listener's external ears. This type of VAD reproduces a sound field as a set of sound images located in specific directions. This can be formed by a relatively small number of loudspeakers, or by a headphone [1, 2, 11, 12]. Hence, a portable system, compared with the second and third approaches, can be realized with this method. As efficient tools for human acoustic communication, many VADs based on this approach have been realized not only in laboratory but also in commercial applications[2]. However, the practical application of this type of VAD also involves some problems as follows:

1. Difficulty in synthesis of Room Transfer Functions (RTFs) which involve direct sound and multiple reflections,
2. Difficulty in dealing with dynamic cues in sound localization, such as listeners' head movements[15],
3. Variation in Head-Related Transfer Functions (HRTFs) caused mainly by individuality.

As stated above, each type of conventional VAD has both advantages and disadvantages. To overcome the disadvantages of the conventional methods, a concept named the "Virtual Sphere (VS) model" was proposed by the authors[13, 14]. This VAD, named *ADVISE* (Acoustic Display based on the VIrtual SpherE model), can realize precise synthesis of HRTFs and RTFs theoretically with a practical scale of hardware. In this paper, our proposed theory of the virtual acoustic display based on the virtual sphere model is extended to the reproduction in more practical acoustic environments.

2. THEORY OF *ADVISE*

2.1. Sound field reproduction based on the Kirchhoff-Helmholtz boundary integral equation

Figure 1 conceptually illustrates the sound field to be controlled. A closed region Ω is established with its boundary Γ. When no sound source exists in Ω, sound pressure at the point $\boldsymbol{r}_P$ in Ω can be expressed by internal Kirchhoff-Helmholtz boundary integral equation as follows:

$$P(\boldsymbol{r}_P,\omega)=\int_{\Gamma(\boldsymbol{r}_q)}\left\{G_{\mathrm{F}}(\boldsymbol{r}_P,\boldsymbol{r}_q,\omega)\frac{\partial P(\boldsymbol{r}_q,\omega)}{\partial n_q}-P(\boldsymbol{r}_q,\omega)\frac{\partial G_{\mathrm{F}}(\boldsymbol{r}_P,\boldsymbol{r}_q,\omega)}{\partial n_q}\right\}d\Gamma, \qquad (1)$$

where G_{F} is the Green function of the Helmholtz equation. This indicates that the original sound field in Ω can be reproduced if the sound pressure and its derivative on the boundary Γ are equalized with those in the original field. If the boundary Γ is divided into N small elements $\Gamma_j(j=1,\cdots,N)$, and sound pressure as well as its derivative are assumed not to change on these elements, the boundary integral equation can be modified as the equation

$$P(\boldsymbol{r}_P,\omega)\approx\sum_{j=1}^{N}\left\{\frac{\partial P(\boldsymbol{r}_j,\omega)}{\partial n_j}\int_{\Gamma_j(\boldsymbol{r}_j)}G_{\mathrm{F}}(\boldsymbol{r}_P,\boldsymbol{r}_j,\omega)d\Gamma-P(\boldsymbol{r}_j,\omega)\int_{\Gamma_j(\boldsymbol{r}_j)}\frac{\partial G_{\mathrm{F}}(\boldsymbol{r}_P,\boldsymbol{r}_j,\omega)}{\partial n_j}d\Gamma\right\}. \qquad (2)$$

In the right hand side of Eq. (2), the first term inside the summation can be regarded as the sound pressure from a monopole source distributed on Γ_i whose strength is equal to $\partial P(\boldsymbol{r}_j,\omega)/\partial n_j$, and the second term indicates the sound pressure from a dipole source on Γ_i whose strength is equal to $P(\boldsymbol{r}_j,\omega)$.

It is well known that the Kirchhoff-Helmholtz boundary integral equation is a variation to express the Huygen's principle[16]. The principle for sound field reproduction based on the Kirchhoff-Helmholtz boundary integral equation was proposed by Ise[9], and its improvement by Takane *et al.*[10] followed.

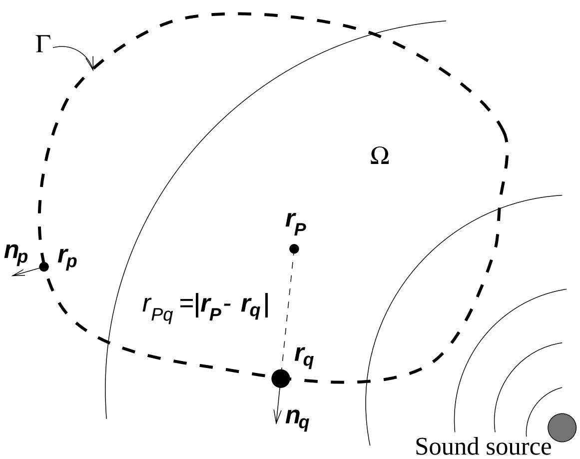

Figure 1: Conceptual illustration of sound field to be controlled

2.2. Virtual Sphere (VS) model

The abovementioned principle enables theoretically precise reproduction of the sound field in an arbitrary enclosed space Ω, under the condition that there is no source inside Ω. The method of *ADVISE* is introduced based on this principle.

As shown in Figure 2 and Figure 3, a virtual boundary Γ is set around the listener. While shape of this boundary can be arbitrarily set, spherical shape is chosen since this shape is convenient for the processings relating to the transfer functions. This boundary is called a Virtual Sphere (VS) boundary hereafter. Sound transmission system from the sound sources to the listener is divided into two parts. One is the transmission system involving the transfer functions from the sound sources to the points on the VS boundary. These transfer functions reflect the acoustic properties of the sound field to be reproduced. They are called Room Transfer Functions (RTFs) here. The other part is the transmission system involving the Head-Related Transfer Functions (HRTFs) corresponding to the points on the VS boundary. Sound pressure at the the entrance of the listener's ears is reproduced by convolving the corresponding RTFs and HRTFs for each direction and then by summing the results for all directions around the listener.

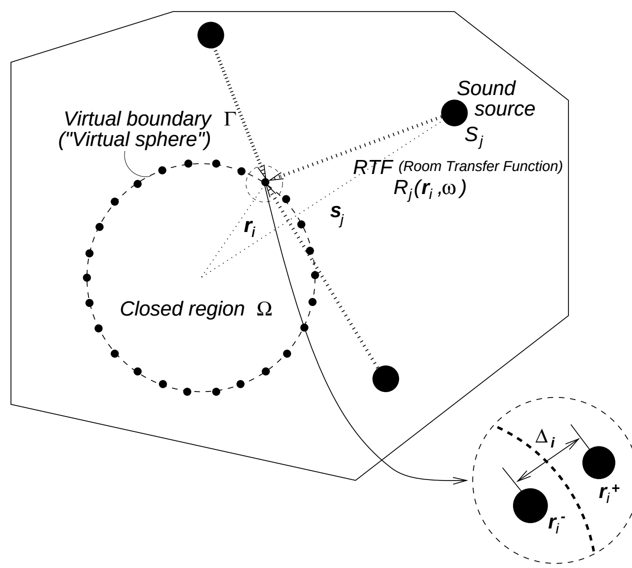

Figure 2: Sound transmission in a room

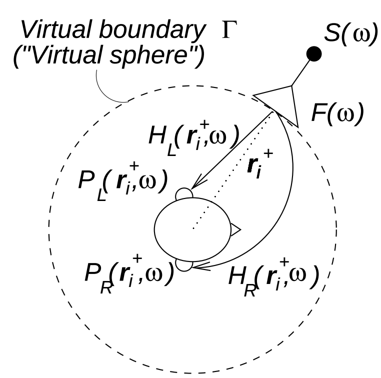

Figure 3: Sound transmission around a listener

2.3. Synthesis of transfer functions

When there exist the M sound sources outside the virtual sphere boundary, sound pressure at the point $\boldsymbol{r}_i \in \Gamma$ can be expressed as follows:

$$P(\boldsymbol{r}_i,\omega)=\sum_{j=1}^{M} S_j(\omega)R_j(\boldsymbol{r}_i,\omega), \tag{3}$$

where $S_j(\omega)$ and $R_j(\boldsymbol{r}_i,\omega)$ indicate the input signal to the j-th sound source, and RTF (Room Transfer Function) from the j-th sound source to the point $\boldsymbol{r}_i$, respectively. The center of the listener's head is assumed to be the origin of the coordinates. Sound pressure at the listener's position $\boldsymbol{r}$ (corresponding to that of listener's head) inside the VS boundary is approximated from the Kirchhoff-Helmholtz boundary integral equation as the following:

$$P(\boldsymbol{r},\omega)\approx\sum_{i=1}^{N}\left\{A_i(\boldsymbol{r}_i,\omega)\frac{\partial P(\boldsymbol{r}_i,\omega)}{\partial n_i}-B_i(\boldsymbol{r}_i,\omega)P(\boldsymbol{r}_i,\omega)\right\}, \tag{4}$$

where $\partial/\partial n_i$ indicates the partial derivative in direction outward from Γ_i. $A_i(\boldsymbol{r}_i,\omega)$ and $B_i(\boldsymbol{r}_i,\omega)$ are expressed as:

$$A_i(\boldsymbol{r},\omega)=\int_{\Gamma_i(\boldsymbol{r}_i)} G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i,\omega)d\Gamma, \tag{5}$$

$$B_i(\boldsymbol{r},\omega)=\int_{\Gamma_i(\boldsymbol{r}_i)} \frac{\partial G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i,\omega)}{\partial n_i}d\Gamma, \tag{6}$$

where $G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i,\omega)$ is the Green function of the Helmholtz equation in free field. As stated in **2.1**, Eq. (4) indicates that the sound field inside Ω is reproduced precisely by the phantom monopole and dipole sources distributed on Γ. Hence, the sound pressure at the listener's ears can be reproduced by convolving the listener's HRTFs to each of the phantom sources.

For convenience of implementation, $A_i(\boldsymbol{r},\omega)$ and $B_i(\boldsymbol{r},\omega)$ are approximated as follows:

$$A_i(\boldsymbol{r},\omega)\approx G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i,\omega)\Delta S_i, \tag{7}$$

$$B_i(\boldsymbol{r},\omega)\approx\frac{1}{\delta_i}\left\{G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i^{+},\omega)-G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i^{-},\omega)\right\}\Delta S_i. \tag{8}$$

In Eqs. (7) and (8), the integral operation is approximated by the values of the integrand at the center of each element times the area of the i-th element ΔS_i, and the derivative operation $\partial G_{\mathrm{F}}/\partial n_i$ in Eq. (6) is approximated by the difference of G_{F} between two closely located points at $\boldsymbol{r}_i$, $\boldsymbol{r}_i^{+}$ and $\boldsymbol{r}_i^{-}$, of which interval is δ_i. This corresponds to the approximation of the dipole source to the two closely located monopoles. The sound pressure at the listener's left and right ears, $P_{\mathrm{L}}^{(\mathrm{O})}(\omega)$ and $P_{\mathrm{R}}^{(\mathrm{O})}(\omega)$, can be synthesized by using HRTFs for the sound source at $\boldsymbol{r}_i$, $\boldsymbol{r}_i^{+}$ and $\boldsymbol{r}_i^{-}$, when the listener exists at the center of the VS. From Eq. (4) and the definition of HRTF[1], $P_{\mathrm{L}}^{(\mathrm{O})}(\omega)$ can be expressed as follows:

$$\begin{aligned}P_{\mathrm{L}}^{(\mathrm{O})}(\omega)=\sum_{i=1}^{N} C_i\Big[&\Delta P(\boldsymbol{r}_i,\omega)H_{\mathrm{L}}(\boldsymbol{r}_i,\omega)G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i,\omega)-\\ &-P(\boldsymbol{r}_i,\omega)\left\{H_{\mathrm{L}}(\boldsymbol{r}_i^{+},\omega)G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i^{+},\omega)\right.\\ &\left.-H_{\mathrm{L}}(\boldsymbol{r}_i^{-},\omega)G_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i^{-},\omega)\right\}\Big],\end{aligned} \tag{9}$$

where H_{L} indicates the left-ear HRTF for the sound source on the VS boundary, $\Delta P(\boldsymbol{r}_i,\omega)=P(\boldsymbol{r}_i^{+},\omega)-P(\boldsymbol{r}_i^{-},\omega)$, and $C_i=\Delta S_i/\delta_i$. $P_{\mathrm{R}}^{(\mathrm{O})}(\omega)$ is obtained by changing H_{L} to H_{R} in Eq. (9).

2.4. Tracking of the listener's movement

From Eq. (9), it is shown that the change in $P_{\mathrm{L}}(\omega)$ due to the listener's movement can be traced by changing only the HRTFs from the phantom sources distributed on the VS boundary. The boundary of a spherical shape helps tracing such a change effectively and the system size ought to be smaller than that for compensating the whole sound transmission paths from each sound source to the listener's ears. It should be emphasized as an advantageous feature in *ADVISE* that the sound pressure in a certain sound field can be simultaneously presented to different listeners with different movement if the set of HRTFs for each listener may be provided. This is because the sound transmission paths from the sound sources to the listeners are expressed by cascading RTFs and HRTFs in *ADVISE*.

Moreover, listener may move around inside the VS if the HRTFs for the phantom sources can be obtained. Listener's position and head rotation can be traced by using magnetic and/or gyro sensors[2]. Since the VS boundary is used only for dividing the sound transmission system into RTFs and HRTFs, size of the VS boundary may be set arbitrarily. However, there exist some limitations. This is discussed in another paper also presented in the *ICAD2002* by the authors[23].

2.5. Discussion on the application of *ADVISE* to the practical environment

Considering the practical implementation of the VAD based on the abovementioned theory, two sorts of problems arise. One is the synthesis of sound when the sound sources are inside the VS boundary Γ. Typical example of this case is to synthesize the listner's own voice. This is impossible within the abovementioned theory, since all sound sources must be outside Γ. In order to discuss this problem, sound pressure at the listener's left ear is divided into two components as follows:

$$P_{\mathrm{L}}(\omega)=P_{\mathrm{L}}^{(\mathrm{O})}(\omega)+P_{\mathrm{L}}^{(\mathrm{I})}(\omega), \tag{10}$$

where $P_{\mathrm{L}}^{(\mathrm{O})}(\omega)$ and $P_{\mathrm{L}}^{(\mathrm{I})}(\omega)$ represent the sound pressure generated by the sound sources outside and inside the VS boundary, respectively. Substitution of the letter L for R in Eq. (10) yields the expression for sound pressure at the listener's right ear. $P_{\mathrm{L}}^{(\mathrm{O})}(\omega)$ and $P_{\mathrm{R}}^{(\mathrm{O})}(\omega)$ can be synthesized by the application of the theory stated above.

The other problem is the synthesis of sound from the moving sound sources. The synthesis of sound from the moving sources according to the introduced theory requires real-time changing of both the RTFs and the HRTFs.

2.5.1. Synthesis of sound from the sources inside the VS boundary

Synthesis of sound pressure at the listener's ears can be achieved based on the method mentioned in **2.3**, when the sound sources are outside Ω. However, it is impossible to reproduce the sound pressure by the sound sources in Ω. For the sound sources of this kind, transfer functions from them to the listener's ears must be obtained specifically. If this transfer function from the i-th source in Ω is obtained as $F_{\mathrm{L},j}(\omega)$, $P_{\mathrm{L}}^{(\mathrm{I})}(\omega)$ can be synthesized by computing the following equation:

$$P_{\mathrm{L}}^{(\mathrm{I})}(\omega)=\sum_{j=1}^{M_{\mathrm{I}}} F_{\mathrm{L},j}(\omega) S_j^{(\mathrm{I})}(\omega), \tag{11}$$

where $S_j^{(\mathrm{I})}(\omega)$ indicates the source signal of the j-th source, and M_{I} is the number of the sources in Ω. $P_{\mathrm{R}}^{(\mathrm{I})}(\omega)$ can be computed in the same way. If the size of the VS is small, number of such sources may be relatively small in practical implementation, thus the consideration of the sound sources inside the VS may not have the serious increase in processings.

2.5.2. *Synthesis of sound from the moving sources*

The transfer functions from the moving sources to the listener change dynamically by the movement of both the listener and the sources themselves. If the RTFs of the moving sources outside Γ can be obtained with taking their dynamic changes into account, it is possible to synthesize the component of the moving sources in terms of the theory stated in **2.1**$\sim$**2.3**. As for the sources inside the VS boundary, $F_{\mathrm{L},j}(\omega)$ and $F_{\mathrm{R},j}(\omega)$ $(j=1,\cdots,M_I)$ may also change dynamically due to the movement of both the source and the listener. Synthesis of sound from the sound sources moving across the VS boundary, for example, may be difficult to deal with in the VS model because the processing of those sources is different, and the processing must be altered depending on their position. When the original sound field includes such sources, transfer functions from these sources to the listener could be synthesized independently of the other sound sources.

2.6. Reproduction system

If the listener's position is traced and the dynamic change in HRTFs due to the listener's movement can be processed in real-time, the sound pressure at the listener's ears may be properly synthesized by using not only binaural but also transaural systems[17, 18], though the crosstalk cancellation is needed in the transaural systems. It should be noted that when the cancellation is executed, the dynamic change in the transfer functions from the loudspeakers to the listener's ears caused by the movement of the listener in the reproduction environment should also be traced in real-time[17, 19].

3. OVERVIEW OF *ADVISE*

3.1. Formulation using the vector and matrix notations

To express the overall algorithm of *ADVISE*, the processings introduced in the previous section is describged in terms of the vector and matrix notations.

1. Source signals of the static sources and the corresponding RTFs from them to the points on the VS boundary are convolved. Here, the vector consisting of the source signals $S_j(\omega)(j=1,\cdots,M)$, and the matrix filter of RTFs whose component consists of $R_j(\boldsymbol{r}_i,\omega)$ are respectively expressed as $\mathbf{s}$ and $\mathbf{R}$. The matrix filters of RTFs whose component involves $R_j(\boldsymbol{r}_i^+,\omega)$ and $R_j(\boldsymbol{r}_i^-,\omega)$ are also needed to compute the derivatives of the sound pressure on the VS boundary. These are expressed by $\mathbf{R}^+$ and $\mathbf{R}^-$, respectively. When $\mathbf{P}$, $\mathbf{P}^+$ and $\mathbf{P}^-$ are used for the vectors respectively consisting of $P(\boldsymbol{r}_i,\omega)$, $P(\boldsymbol{r}_i^+,\omega)$ and $P(\boldsymbol{r}_i^-,\omega)$, the following equations are obtained:
$$\mathbf{P}=\mathbf{R}\mathbf{s},\quad \mathbf{P}^+=\mathbf{R}^+\mathbf{s},\quad \mathbf{P}^-=\mathbf{R}^-\mathbf{s}. \tag{12}$$
2. Eq. (9) can be described in terms of the vector expressions as follows:
$$P_{\mathrm{L}}^{(\mathrm{O})}(\omega)=\Delta\mathbf{P}^{\mathrm{T}}\mathbf{h}_{\mathrm{L}}\mathbf{g}^{\mathrm{T}}-\mathbf{P}^{\mathrm{T}}\left\{\mathbf{h}_{\mathrm{L}}^+(\mathbf{g}^+)^{\mathrm{T}}-\mathbf{h}_{\mathrm{L}}^-(\mathbf{g}^-)^{\mathrm{T}}\right\}, \tag{13}$$
where $\Delta\mathbf{P}=\mathbf{P}^+-\mathbf{P}^-$, $\mathbf{h}_{\mathrm{L}},\mathbf{h}_{\mathrm{L}}^+,\mathbf{h}_{\mathrm{L}}^-,\mathbf{g},\mathbf{g}^+$ and $\mathbf{g}^-$ are vectors consisting of $H_{\mathrm{L}}(\boldsymbol{r}_i,\omega)$, $H_{\mathrm{L}}(\boldsymbol{r}_i^+,\omega)$, $H_{\mathrm{L}}(\boldsymbol{r}_i^-,\omega)$, $C_iG_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i,\omega)$, $C_iG_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i^+,\omega)$ and $C_iG_{\mathrm{F}}(\boldsymbol{r},\boldsymbol{r}_i^-,\omega)$, respectively. The similar expression for the sound pressure at the listener's right ear is given by
$$P_{\mathrm{R}}^{(\mathrm{O})}(\omega)=\Delta\mathbf{P}^{\mathrm{T}}\mathbf{h}_{\mathrm{R}}\mathbf{g}^{\mathrm{T}}-\mathbf{P}^{\mathrm{T}}\left\{\mathbf{h}_{\mathrm{R}}^+(\mathbf{g}^+)^{\mathrm{T}}-\mathbf{h}_{\mathrm{R}}^-(\mathbf{g}^-)^{\mathrm{T}}\right\}. \tag{14}$$
In Eqs. (13) and (14), the components of the vectors $\mathbf{h}_{\mathrm{L}}$, $\mathbf{h}_{\mathrm{L}}^+$, $\mathbf{h}_{\mathrm{L}}^-$, $\mathbf{h}_{\mathrm{R}}$, $\mathbf{h}_{\mathrm{R}}^+$, $\mathbf{h}_{\mathrm{R}}^-$ are dynamically changed due to the listener's movement.
3. The following vector product yields the sound pressure at the listener's ears generated by the static sound sources inside the VS boundary:
$$P_{\mathrm{L}}^{(\mathrm{I})}(\omega)=\mathbf{f}_{\mathrm{L}}^{\mathrm{T}}\mathbf{s}^{(\mathrm{I})}, \tag{15}$$
$$P_{\mathrm{R}}^{(\mathrm{I})}(\omega)=\mathbf{f}_{\mathrm{R}}^{\mathrm{T}}\mathbf{s}^{(\mathrm{I})}, \tag{16}$$
where $\mathbf{f}_{\mathrm{L}}$ and $\mathbf{f}_{\mathrm{R}}$ are vectors whose components are $F_{\mathrm{L},j}(\omega)$ and $F_{\mathrm{R},j}(\omega)$, and $\mathbf{s}^{(\mathrm{I})}$ is a vector whose component is $S_j^{(\mathrm{I})}(\omega)$ $(j=1,\cdots,M_{\mathrm{I}})$. Vectors $\mathbf{f}_{\mathrm{L}}$ and $\mathbf{f}_{\mathrm{R}}$ may be varied due to the listener's movement.
4. Total sound pressure at the listener's ears is obtained by summing the components computed in 2. and 3., *i. e.*, Eq. (10).

Althogh the formulation in this paper was made in the frequency domain, the algorithm listed above can be implemented in the time domain by taking inverse Fourier Transforms of the tranfer functions and the source signals, and computing the convolutions of them instead of the products in the frequency domain.

3.2. Block diagram of *ADVISE*

According to the processings described in **3.1**, block diagram of *ADVISE* is illustrated in Figure 4. In this figure, the source signals are assumed to be divided into two classes according to whether the sources inside or outside the VS boundary. The VS model is adopted in the processings for the sources outside the VS boundary. This means that $P_{\mathrm{L}}^{(\mathrm{O})}(\omega)$ and $P_{\mathrm{R}}^{(\mathrm{O})}(\omega)$ are synthesized from the sound pressure and its derivative at the points on the VS boundary, no matter how many sources are distributed. Compensation of the HRTFs corresponding to the points on the VS boundary is needed. Compensation of the change in RTFs due to the movement of the listener and the sources is also required.

When the number of the sources outside the VS boundary is small comparing with that of the phantom sources distributed on the VS boundary, the processing cost in *ADVISE* is estimated much higher than those in the conventional auditory displays, in which the transfer functions from those sources to the listener are directly convolved to the source signals. However, it is difficult and not practical to obtain the set of transfer functions in the original sound

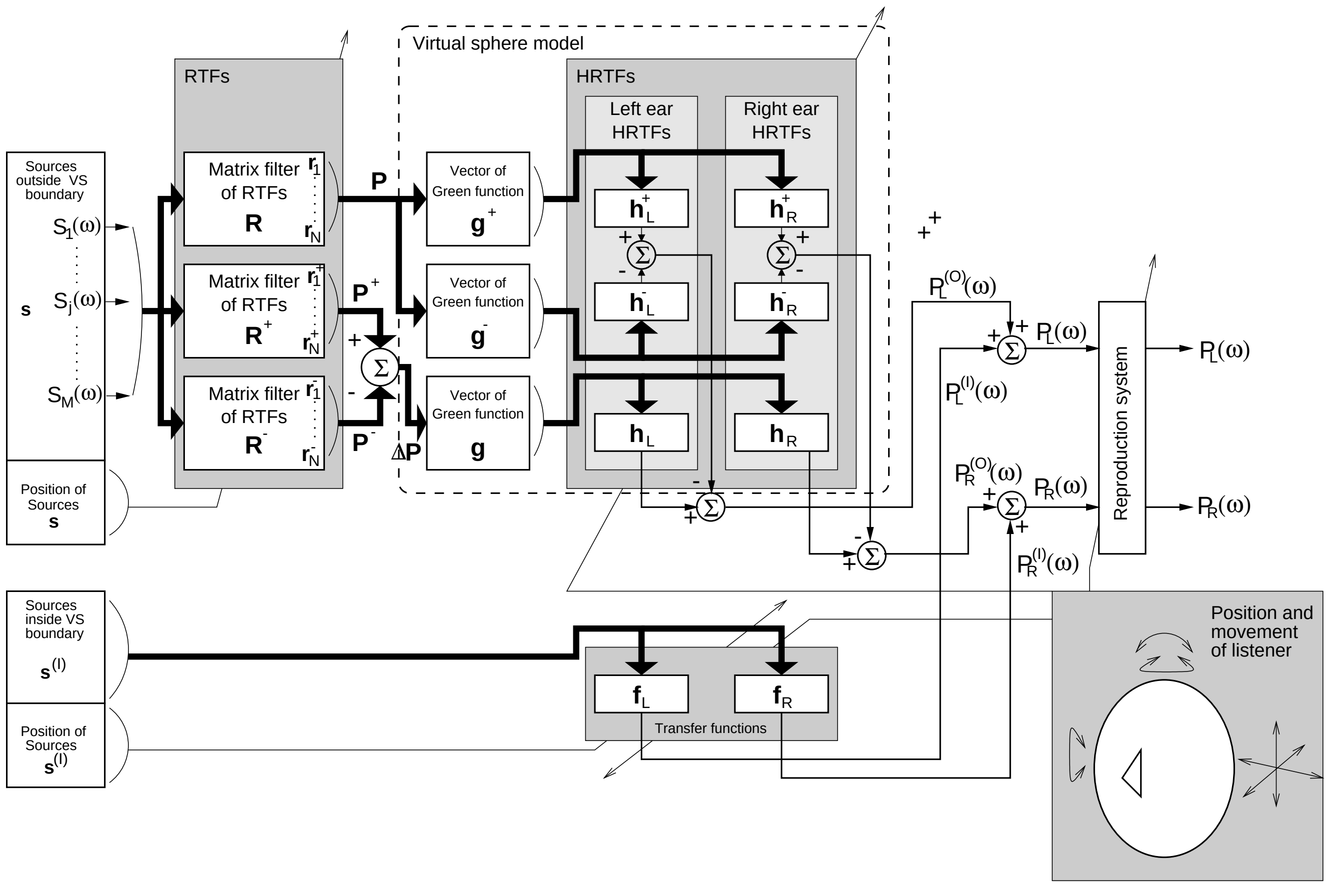

Figure 4: Block diagram of *ADVISE*

field for the individual listener. Moreover, the RTFs can change abruptly by a slight movement of the listener's positions[20]. In contrast, the virtual sphere model of *ADVISE* can separate the synthesis of the transfer functions into RTFs and HRTFs, and the real-time processing corresponding to the listener's movement is required only for the synthesis of the HRTFs for the points on the VS boundary.

As for the sound sources inside the VS boundary, change in the transfer functions from these sources to the listener must be compensated for. The number of sources of this kind may be small comparing with the sources outside the VS boundary, thus the processings needed in the synthesis of the sound from those sources may not seriously affect the total processing costs.

3.3. Other processings in *ADVISE*

There exist some other processings required in *ADVISE* which are not explicitly appeard in Figure 4. To compensate for the change in the tranfer functions due to the movements of the listener and the moving sources, the transfer functions must be changed in real-time. Since these transfer functions cannot be obtained as their functional forms, their changes are compensated by the interpolation with the functions measured (or computed) at discrete points. As for the HRTFs for the points on the VS boundary, some databases of HRTFs open in public are available[21, 22]. However, not only $\mathbf{h}$ but $\mathbf{h}^+$ and $\mathbf{h}^-$ are also needed in the VS model. How to get approximations of $\mathbf{h}^+$ and $\mathbf{h}^-$ from $\mathbf{h}$ is discussed in the companion paper[23]. The interpolation of the transfer functions must be processed both spatially and temporalily, but this can be made in the same way as that used in the conventional auditory displays[2].

4. CONCLUDING REMARKS

In this paper, a new method for VAD named *ADVISE* was introduced. It was theoretically shown that accurate reproduction of sound waves reaching the entrances of the listener's ears is realized with *ADVISE*.

A binaural VAD has been practically developed by using five DSPs (Texas Instruments TMS320C6701) hosted by a personal computer. This will also be introduced in the companion paper at *ICAD2002*[23].

5. REFERENCES

[1] J. Blauert: *Spatial Hearing* (MIT Press, Cambridge, 1983).

[2] D. R. Begault: *3-D SOUND for virtual reality and multimedia* (AP Professional, Cambridge, 1994).

[3] E. Meyer, W. Burgtorf, and P. Damaske, "Eine Apparatur zur electroakustischen Nachbildung von Shallfeldern. Subjektive Hörwirkungen beim Übergang Kohärenz-Inkohärenz," *Acustica*, **15**, 339-344(1965).

[4] For example, Shimizu Corporation, "Audible room acoustics simulator," *J. Acoust. Soc. Jpn.*, **48**(4), 263-265(1992) (in Japanese).

[5] M. Camras, "Approach to recreating a sound field," *J. Acoust. Soc. Am.*, **43**(6), 1425-1431(1967).

[6] K. Furuya and Y. Ichinose, "Sound field control in a closed space using Boundary-Pressure-Method," *Tech. Rep. IEICE*, EA90-15, 599-606(1990) (in Japanese).

[7] A. J. Berkhout, D. de Vries and P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, **93**(5), 2764-2778(1993).

[8] S. Ise, "Theory of the sound field reproduction based on the Kirchhoff-Helmholtz integral equation," *Proc. 15th ICA*, Vol. I, 85-88(1995).

[9] S. Ise, "A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems," *Acustica united with Acta Acustica*, **85**(1), 78-87(1999).

[10] S. Takane, Y. Suzuki and T. Sone, "A new method for global sound reproduction based on Kirchhoff's integral equation," *Acustica united with Acta Acustica*, **85**(2), 250-257(1999).

[11] M. R. Schroeder and B. S. Atal, "Computer simulation of sound transmission in rooms," *IEEE Inter. Conv. Rec.*, Part 7, 150-155(1963).

[12] Y. Ando, S. Shidara, Z. Maekawa and K. Kido, "Some basic studies on the acoustic design of room by computer," *J. Acoust. Soc. Jpn.*, **29**(3), 151-159(1973) (in Japanese).

[13] S. Takane, T. Miyajima, Y. Yamada, D. Arai, Y. Suzuki and T. Sone, "An auditory display based on virtual sphere model," *Proc. ASVA*, 379-383(1997).

[14] S. Takane, Y. Suzuki, T. Miyajima and T. Sone, "Analysis of accuracy of sound reproduction by using binaural auditory display based on the virtual sphere model," *Proc. 7th Western Pacific Regional Acoustics Conference*, **1**, 301-304(2000).

[15] J. Kawaura, Y. Suzuki and T. Sone, "Sound localization in headphone reproduction by simulating transfer functions from the sound source to the external ear," *J. Acoust. Soc. Jpn.(E)*, **12**(5), 203-216(1991).

[16] P. M. Morse and K. U. Ingard, *Theoretical acoustics* (Princeton University Press, New Jersey, 1968).

[17] W. G. Gardner, *3-D audio using loudspeakers* (Kluwer Academic Publishers, Boston, 1998).

[18] L. Savioja, J. Huopaniemi, T. Loppi and R. Väänänen, "Creating interactive virtual acoustic environment," *J. Audio Eng. Soc.*, **47**(9), 675-705(1999).

[19] J. Garas, *Adaptive 3D sound systems* (Kluwer Academic Publishers, Massachusetts, 2000).

[20] M. Tohyama, H. Suzuki and Y. Ando, *The nature and technology of acoustic space* (Academic Press, London, 1995).

[21] `http://sound.media.mit.edu/KEMAR.html`

[22] S. Takane, D. Arai, T. Miyajima, K. Watanabe, Y. Suzuki and T. Sone, "A database of Head-Related Transfer Functions in whole directions on upper hemisphere," *Acoust. Sci. Tech.*, **23**(3), 160-162(2002).

[23] Y. Suzuki, S. Takane, S. Takahashi and T. Miyajima, "A preliminary development of high definition virtual acoustic display based on *ADVISE*," to be presented in *2002 International Conference on Auditory Display*(2002).