MULTI-CHANNEL INVERSE FILTERING WITH SELECTION AND ENHANCEMENT OF A LOUDSPEAKER FOR ROBUST SOUND FIELD REPRODUCTION

Shigeki Miyabe, Masayuki Shimada, Tomoya Takatani, Hiroshi Saruwatari and Kiyohiro Shikano

{shige-m, sawatari, shikano}@is.naist.jp
Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama-cho, Ikoma-shi, Nara, 630-0192, JAPAN

ABSTRACT

This paper describes a new sound field reproduction strategy, where the system can give accurate sound images if a user is at a specific position, and still provides the direction of the primary source if the user moves. The existing methods do not take into account the accurate reproduction outside the specific control points, and if the user moves from the control points, he cannot feel the accurate sound image. To solve this problem, we propose a novel design algorithm of inverse filters that make a secondary source in the direction of the primary sound source have the largest power. In the proposed method, the user can feel the sound image toward the enhanced secondary source even around the control points. Simultaneously the accurate reproduction at the control points can be achieved as well as the conventional method. The subjective evaluation shows that the proposed method is more robust against the user's move compared with the conventional method.

1. INTRODUCTION

Sound field control/reproduction is an requisite technology for constructing a basis of audio virtual reality system, which requires prompt attention. To realize three dimensional auditory display, a lot of approaches have been attempted in many fields, i.e., perception, reproduction, architecture and so on, for many years, even from earlier than the 20th century [1].

From the viewpoint of transducer devices, sound field reproduction can be classified into two groups, namely, using headphones and loudspeakers. By reproducing the signals observed at microphones set on human's (or a dummy head's) ears, called binaural signals, a user can listen to almost the same sound as that the user listened to when he/she was in the recorded environment [1]. However, the above-mentioned method has a fatal drawback that the headphone wearing compels listener into bodily constraint. Therefore in this paper, we will mainly deal with only loudspeaker reproductions.

Loudspeaker reproduction can be classified into two groups again; whether to compensate impulse responses from the loudspeakers to the listener's ears or not. The systems without the compensation are called discrete surround system including the most popular stereophonic reproduction and 3/2 format used in Dolby Digital system. The idea is so simple that the sound intensities and phases are just panned to multiple loudspeakers surrounding the listener. In particular, the intensity panning has an advantage that it is robust against the shift of the listener's position even though its precision of reproduction is limited.

By fixing the listener's position and compensating the impulse responses around the listener's ears, binaural signals can be reproduced with loudspeakers. Such systems are called *transaural systems* [2, 3, 4]. Since the loudspeaker reproduction has crosstalk components, they have to be removed for the reproduction of binaural signals. Crosstalk canceller realizes this by using inverse filter of transfer functions between the loudspeakers and the listener's ears in an anechoic environment, called head related transfer functions (HRTFs). However, they can not provide strict reproductions of the original binaural signals because the reproduced signals are distorted by the reverberation of the listening environment. Therefore, we must compensate the impulse responses of the user's ears including the reverberation, called binaural room impulse responses (BRIRs). In order to obtain the accurate inverse filter of BRIRs which are in general non-minimum phase systems, Miyoshi *et al.* have proposed multiple input/output inverse theorem (MINT) utilizing more loudspeakers than control points (the listener's ears) [5].

There is a problem in the conventional transaural systems using inverse filter of BRIRs. Since these methods considers only the accuracy of reproduction at the control points (*sweet spot*), the directional cues are not held on the other outer areas. As for the crosstalk canceller, a method to expand the sweet spot towards the front and the back of the listener, called stereo-dipole system, has been proposed [6]. However, for the strict reproduction at the control points using inverse filters of BRIRs without microphones at the user's ears just used in [3], no method has been proposed for mitigating the effect of the listener's movement.

In this paper, we propose an algorithm to design an inverse filter to alleviate the sweet spot problem. We design an inverse filter whose intensity is weighted to the loudspeaker in the direction closest to the source's DOA, by finding the closest filter matrix to the one which only uses single loudspeaker. With this method, we can reproduce the binaural signals at the control points with almost the same accuracy as that of the conventional MINT, while the directional cues are held even outside the sweet spot. The efficiency of the proposed method is ascertained in a subjective evaluation experiment where the subjects move their heads.

2. CONVENTIONAL SOUND FILED REPRODUCTION USING INVERSE FILTER

2.1. Principle

In the transaural system, we must reproduce binaural signals at the fixed control points which are arranged at the listener's ears. This can be realized with an inverse filter of BRIRs. Although BRIRs are non-minimum phase systems, it is proved in [5] that there exists an inverse filter of BRIRs by using more loudspeakers than control points. Hereafter we address the problem to reproduce N input signals at N control points C_n $(n = 1, \ldots, N)$

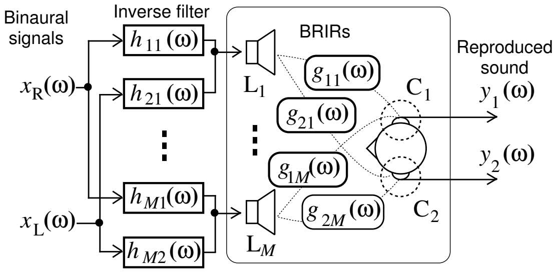


Figure 1: *Configuration of a transaural system with two control points and M loudspeakers.*

(commonly single listener is assumed and $N = 2$ to control the sound pressures at both of the ears) with M loudspeakers L_m ($m = 1, \ldots, M$). We show the configuration of the transaural system with 2 control points and M loudspeakers in Fig. 1.
We designate the signals to be reproduced at control points C_n as $\boldsymbol{x}(\omega) = [x_1(\omega), \ldots, x_N(\omega)]^{\mathrm{T}}$, where ω denotes an angular frequency and $\{\cdot\}^{\mathrm{T}}$ denotes transposition. We measure all $N \times M$ impulse responses between L_m and C_n, denoting them as $g_{nm}(\omega)$. We define an $N \times M$ matrix $\boldsymbol{G}(\omega) = [g_{nm}(\omega)]_{nm}$, where $[a]_{ij}$ represents a matrix which includes the entry a in the i-th row and the j-th column. We design an $M \times N$ inverse filter matrix defined as $\boldsymbol{H}(\omega) = [h_{mn}(\omega)]_{mn}$ to satisfy the following condition

$$\boldsymbol{G}(\omega)\boldsymbol{H}(\omega) = \boldsymbol{I}, \tag{1}$$

where $\boldsymbol{I}$ denotes an identity matrix. When we output $\boldsymbol{H}(\omega)\boldsymbol{x}(\omega)$ from L_m, i.e., input signals $\boldsymbol{x}(\omega)$ filtered by the inverse filter $\boldsymbol{H}(\omega)$, signals at control points $\boldsymbol{y}(\omega) = [y_1(\omega), \ldots, y_N(\omega)]$ satisfy the condition $\boldsymbol{y}(\omega) = \boldsymbol{G}(\omega)\boldsymbol{H}(\omega)\boldsymbol{x}(\omega) = \boldsymbol{x}(\omega)$. Therefore, input signals $x_n(\omega)$ are reproduced at the control points.

2.2. Inverse Filter Design Based on Least Norm Solution

As shown in Eq. (1), $\boldsymbol{H}(\omega)$ is a generalized inverse filter of the matrix $\boldsymbol{G}(\omega)$. Since $M > N$, the solution is indefinite. To decide $\boldsymbol{H}(\omega)$, adoption of Moore-Penrose generalized inverse matrix which gives least norm solution (LNS) is proposed. Using the LNS, a total gain of the inverse filter is minimized and its control becomes robust against the error.
At first, to obtain Moore-Penrose generalized inverse matrix, the singular value decomposition (SVD) is applied to $\boldsymbol{G}(\omega)$. In the case that $\boldsymbol{G}(\omega)$ is N-full-rank, SVD can be written as

$$\boldsymbol{G}(\omega) = \boldsymbol{U}(\omega)\underbrace{[\boldsymbol{\Gamma}(\omega), \boldsymbol{O}_{N,M-N}]}_{N\times M}\boldsymbol{V}^{\mathrm{H}}(\omega), \tag{2}$$

where $\{\cdot\}^{\mathrm{H}}$ denotes conjugate transposition, $\boldsymbol{U}(\omega) = [\boldsymbol{u}_1(\omega), \ldots, \boldsymbol{u}_N(\omega)]$, $\boldsymbol{V}(\omega) = [\boldsymbol{v}_1(\omega), \ldots, \boldsymbol{v}_M(\omega)]$, $\boldsymbol{\Gamma}(\omega) = \mathrm{diag}[\gamma_1(\omega), \ldots, \gamma_N(\omega)]$, $\mathrm{diag}[x_1, \ldots, x_N]$ denotes $N \times N$ diagonal matrix whose n-th diagonal element is x_n, $\gamma_n(\omega)$ is the n-th largest singular value of $\boldsymbol{G}(\omega)$, N-dimensional vectors $\boldsymbol{u}_n(\omega)$ and M-dimensional vectors $\boldsymbol{v}_n(\omega)$ for $n = 1, \ldots, N$ are eigenvectors corresponding to singular values $\gamma_n(\omega)$, M-dimensional vectors $\boldsymbol{v}_m(\omega)$ for $m = N+1, \ldots, M$ are unit vectors which span the nullspace of $\boldsymbol{G}(\omega)$, and $\boldsymbol{O}_{i,j}$ denotes an $i \times j$ zero matrix. Note that $\boldsymbol{U}(\omega)$ and $\boldsymbol{V}(\omega)$ are unitary matrices. Then generalized inverse matrix of $\boldsymbol{G}(\omega)$, denoted by $\boldsymbol{G}^{-}(\omega)$, can be written as

$$\boldsymbol{G}^{-}(\omega) = \boldsymbol{V}(\omega)\underbrace{\begin{bmatrix}\boldsymbol{\Lambda}(\omega)\\ \boldsymbol{\Pi}(\omega)\end{bmatrix}}_{M\times N}\boldsymbol{U}^{\mathrm{H}}(\omega), \tag{3}$$

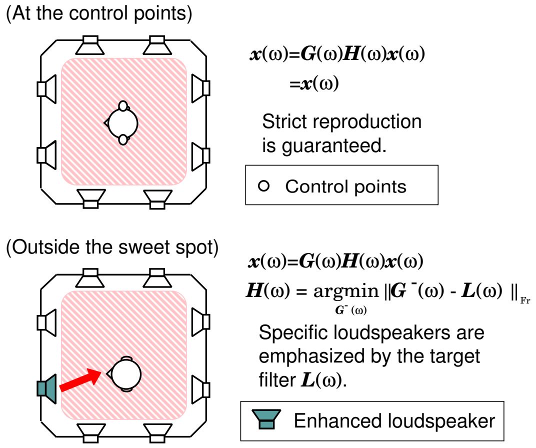

(At the control points)
$\boldsymbol{x}(\omega)$=$\boldsymbol{G}(\omega)\boldsymbol{H}(\omega)\boldsymbol{x}(\omega)$
=$\boldsymbol{x}(\omega)$
Strict reproduction
is guaranteed.
○ Control points
(Outside the sweet spot)
$\boldsymbol{x}(\omega)$=$\boldsymbol{G}(\omega)\boldsymbol{H}(\omega)\boldsymbol{x}(\omega)$
$\boldsymbol{H}(\omega) = \underset{\boldsymbol{G}^{-}(\omega)}{\mathrm{argmin}} \|\boldsymbol{G}^{-}(\omega) - \boldsymbol{L}(\omega)\|_{\mathrm{Fr}}$
Specific loudspeakers are
emphasized by the target
filter $\boldsymbol{L}(\omega)$.
Enhanced loudspeaker

Figure 2: *Strategy of the proposed approach.*

$$\Lambda(\omega) = \mathrm{diag}\left[\frac{1}{\gamma_1(\omega)}, \ldots, \frac{1}{\gamma_N(\omega)}\right], \tag{4}$$

where $\boldsymbol{\Pi}(\omega)$ is an arbitrary $(M-N) \times N$ matrix. Here Moore-Penrose generalized inverse matrix $\boldsymbol{G}^{+}(\omega)$ can be obtained by the substitution $\boldsymbol{\Pi}(\omega) = \boldsymbol{O}_{M-N,N}$ as

$$\boldsymbol{G}^{+}(\omega) = \boldsymbol{V}(\omega)\begin{bmatrix}\boldsymbol{\Lambda}(\omega)\\ \boldsymbol{O}_{M-N,N}\end{bmatrix}\boldsymbol{U}^{\mathrm{H}}(\omega). \tag{5}$$

Then we use $\boldsymbol{G}^{+}(\omega)$ as an inverse filter; $\boldsymbol{H}(\omega) = \boldsymbol{G}^{+}(\omega)$.

3. PROPOSED METHOD: INVERSE FILTER WITH SECONDARY SOURCE SELECTION AND ENHANCEMENT

3.1. Approach

We depict the basic strategy of our approach in Fig. 2. Since the conventional LNS-based inverse filter designing considers only the reproduction at the specific control points, the directional cues cannot be presented outside the sweet spot. Though strict reproduction of primary sound field in a large area is difficult, it should be worthwhile that the listener perceives the correct DOAs outside the sweet spot. Therefore, in this section we propose an inverse filter design method to satisfy both of the following requirements as;

(R1) the strict reproduction is guaranteed at the control points,

(R2) robustness of the DOAs perceived outside the sweet spot.

One of the way to satisfy the condition (R2) is to output the signals only from a loudspeaker in the direction of the source. When sound is outputted from a specific loudspeaker, the listener perceives the source along the direction of this loudspeaker. This configuration is robust against movement of the listener but cannot reproduce the sources precisely. To satisfy both (R1) and (R2), we design an inverse filter whose output gain of the loudspeaker at the target direction is enhanced. Firstly, we design a multi-channel filter $\boldsymbol{T}(\omega)$ which has full bandpass and linear phase property for the loudspeaker in the source direction, and has zero gain for the other loudspeakers.
Secondly, we compute the closest inverse filter $\boldsymbol{H}(\omega)$ to $\boldsymbol{T}(\omega)$ according to a given norm. In the following discussion, we will call $\boldsymbol{T}(\omega)$ a *target filter.* Though single source is assumed in this paper due to the limited space, we can also deal with multiple sources. At first, we separate the binaural signals into each

of the sources by using blind source separation, and estimate their DOAs. Then, we design the proposed filters for each of the sources, and impose outputs of them.

3.2. Design of Target Filter

In the next section, we minimize the distance between the inverse filter and the target filter which is described in this section. To make the output of the resultant inverse filter natural, we must compensate the difference of the gains and delays between the target filter and the LNS inverse filter.

To make the difference of delay to a minimum, we synchronize the peak of the target filter and the LNS inverse filter $\boldsymbol{G}^{+}(\omega)$. At first we obtain the time delay τ when the impulse response of the inverse filter has the largest amplitude in time domain. Then we give the target filter linear phases with the delay of τ. If the k-th loudspeaker is to be emphasized, the $M \times N$ target filter matrix $\boldsymbol{T}(\omega) = \left[T_{mn}(\omega)\right]_{mn}$ has nonzero gains and delay of τ in the components corresponding to the k-th loudspeaker, and has zero gains in the other components, as;

$$T_{mn}(\omega) = \begin{cases} s(\omega) \cdot e^{-j\omega\tau} & (\text{if} \quad m = k) \\ 0 & (\text{otherwise}), \end{cases} \tag{6}$$

for $n = 1, \ldots, N$, where $s(\omega)$ is a constant to decide the gain of $\boldsymbol{T}(\omega)$. Then we decide $s(\omega)$ to compensate the difference of gain. For this compensation, we give $\boldsymbol{T}(\omega)$ the equal total gain to the LNS inverse filter $\boldsymbol{G}^{+}(\omega)$ as

$$\|\boldsymbol{T}(\omega)\|_{\text{Fr}} = \|\boldsymbol{G}^{+}(\omega)\|_{\text{Fr}}, \tag{7}$$

where $\|\cdot\|_{\text{Fr}}$ denotes Frobenius norm; a Frobenius norm of an $I \times J$ matrix $\boldsymbol{X} = [x_{ij}]_{ij}$ is defined as $\|\boldsymbol{X}\|_{\text{Fr}} = \sqrt{\sum_{i=1}^{I} \sum_{j=1}^{J} |x_{ij}|^2}$. From Eq. (7), $s(\omega)$ can be obtained as $s(\omega) = \|\boldsymbol{G}^{+}(\omega)\|_{\text{Fr}} / \sqrt{N}$. Therefore, for $n = 1, \ldots, N$, $\boldsymbol{T}(\omega)$ can be given by

$$T_{mn}(\omega) = \begin{cases} \dfrac{\left\|\boldsymbol{G}^{+}(\omega)\right\|_{\text{Fr}} \cdot e^{-j\omega\tau}}{\sqrt{N}} & (\text{if} \quad m = k) \\ 0 & (\text{otherwise}). \end{cases} \tag{8}$$

3.3. Minimization of Distance from Target Filter

Here we discuss the minimization problem of a distance between the generalized inverse matrix $\boldsymbol{G}^{-}(\omega)$ shown in Eq. (3) and the target filter $\boldsymbol{T}(\omega)$ in Eq. (8). In this problem we apply Frobenius norm as a distance measure of matrices. Therefore, our objective is to obtain an inverse filter $\boldsymbol{H}(\omega)$ which has minimum Frobenius norm to $\boldsymbol{T}(\omega)$ as

$$\boldsymbol{H}(\omega) = \underset{\boldsymbol{G}^{-}(\omega)}{\operatorname{argmin}} \left\|\boldsymbol{G}^{-}(\omega) - \boldsymbol{T}(\omega)\right\|_{\text{Fr}} \tag{9}$$

From Eq. (3), the square of Frobenius norm for $\boldsymbol{G}^{-}(\omega) - \boldsymbol{T}(\omega)$, denoted by $F(\omega)$, can be written as

$$\begin{aligned} F(\omega) &= \left\|\boldsymbol{G}^{-}(\omega) - \boldsymbol{T}(\omega)\right\|_{\text{Fr}}^2 \\ &= \left\|\boldsymbol{V}(\omega) \begin{bmatrix} \boldsymbol{\Lambda}(\omega) \\ \boldsymbol{\Pi}(\omega) \end{bmatrix} \boldsymbol{U}^{\text{H}}(\omega) - \boldsymbol{T}(\omega)\right\|_{\text{Fr}}^2 . \end{aligned} \tag{10}$$

Here it is notable that $\boldsymbol{U}(\omega)$ and $\boldsymbol{V}(\omega)$ are unitary matrices as described in Eq. (2). Since multiplication of a unitary matrix does not change the Frobenius norm, Eq. (10) can be rewritten as

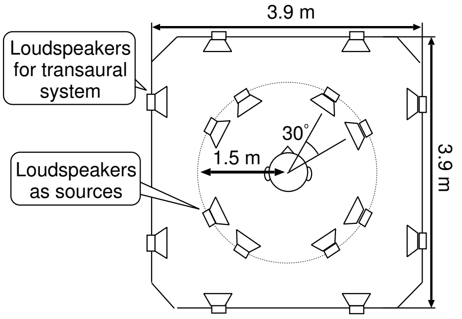


Figure 3: *Experimental conditions.*

$$\begin{aligned} F(\omega) &= \left\|\boldsymbol{V}^{\text{H}}(\omega)\left(\boldsymbol{G}^{-}(\omega) - \boldsymbol{T}(\omega)\right)\boldsymbol{U}(\omega)\right\|_{\text{Fr}}^2 \\ &= \left\|\begin{bmatrix} \boldsymbol{\Lambda}(\omega) \\ \boldsymbol{\Pi}(\omega) \end{bmatrix} - \boldsymbol{V}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega)\right\|_{\text{Fr}}^2 \\ &= \left\|\begin{bmatrix} \boldsymbol{\Lambda}(\omega) - \boldsymbol{V}_{\text{span}}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega) \\ \boldsymbol{\Pi}(\omega) - \boldsymbol{V}_{\text{null}}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega) \end{bmatrix}\right\|_{\text{Fr}}^2 \\ &= \left\|\boldsymbol{\Lambda}(\omega) - \boldsymbol{V}_{\text{span}}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega)\right\|_{\text{Fr}}^2 \\ &\quad + \left\|\boldsymbol{\Pi}(\omega) - \boldsymbol{V}_{\text{null}}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega)\right\|_{\text{Fr}}^2 , \end{aligned} \tag{11}$$

where $\boldsymbol{V}_{\text{span}}(\omega)$ is a truncated matrix of $\boldsymbol{V}(\omega)$ and is composed of eigenvectors which span row space of $\boldsymbol{G}(\omega)$ as $\boldsymbol{V}_{\text{span}}(\omega) = [\boldsymbol{v}_1(\omega), \ldots, \boldsymbol{v}_N(\omega)]$. Similarly, $\boldsymbol{V}_{\text{null}}(\omega)$ is a truncated matrix of $\boldsymbol{V}(\omega)$ and is composed of unit vectors which span null space of $\boldsymbol{G}(\omega)$ as $\boldsymbol{V}_{\text{null}}(\omega) = [\boldsymbol{v}_{N+1}(\omega), \ldots, \boldsymbol{v}_M(\omega)]$. In Eq. (11), the term $\left\|\boldsymbol{\Lambda}(\omega) - \boldsymbol{V}_{\text{span}}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega)\right\|_{\text{Fr}}^2$ cannot be changed because $\boldsymbol{\Lambda}(\omega)$ is fixed to satisfy the generalized inverse matrix of $\boldsymbol{G}(\omega)$. On the other hand, $\boldsymbol{\Pi}(\omega)$ is arbitrary and the term $\left\|\boldsymbol{\Pi}(\omega) - \boldsymbol{V}_{\text{null}}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega)\right\|_{\text{Fr}}^2$ can be minimized to zero by a substitution

$$\boldsymbol{\Pi}(\omega) = \boldsymbol{V}_{\text{null}}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega), \tag{12}$$

then $F(\omega)$ is minimized. Therefore, substituting Eq. (12) in Eq. (3), the optimal inverse filter can be obtained as

$$\boldsymbol{H}(\omega) = \boldsymbol{V}(\omega) \begin{bmatrix} \boldsymbol{\Lambda}(\omega) \\ \boldsymbol{V}_{\text{null}}^{\text{H}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega) \end{bmatrix} \boldsymbol{U}^{\text{H}}(\omega) . \tag{13}$$

4. EXPERIMENTS AND DISCUSSIONS

4.1. Comparison of Reproduction Performance at Control Points

To verify the accuracy of the reproduction at the control points, we have conducted a subjective evaluation experiment comparing the proposed method with the conventional LNS inverse filter. The experiment was conducted via eight loudspeakers for reproduction, in a room of 3.9 m×3.9 m with the reverberation time of 160 ms. We used two music sources which consist of piano and drums musical performance, respectively, with sampling frequency of 48 kHz. The positions of the sound sources are set at 1.5 m apart from the user and their directions are ±30°, ±60°, ±120° and ±150° clockwisely, where the direction in front of the user is set to be 0°. The loudspeakers for reproduction were set on the same directions as the sound sources with different distance from the user. The passband frequency was 150–4000 Hz.

We made 48 patterns of signals to be reproduced in simulations, i.e., 16 combinations of the eight positions of the sources and the two sources for each of three methods; the proposed method, true sound source and the conventional LNS inverse filter. For each source, at first we presented the subjects to the sounds using two inverse filter methods in random order after presenting the sound from true source. Then we let them answer which of the latter two is close to the first. The subjects were organized with nine males and one female in their 20th.

The scores of the conventional method and the proposed method were 50.6% and 49.4%, respectively. We can say that there is no significant difference between them. Therefore, it is ascertained that the proposed method does not degrade the reproduction performance when the listener is at the sweet spot.

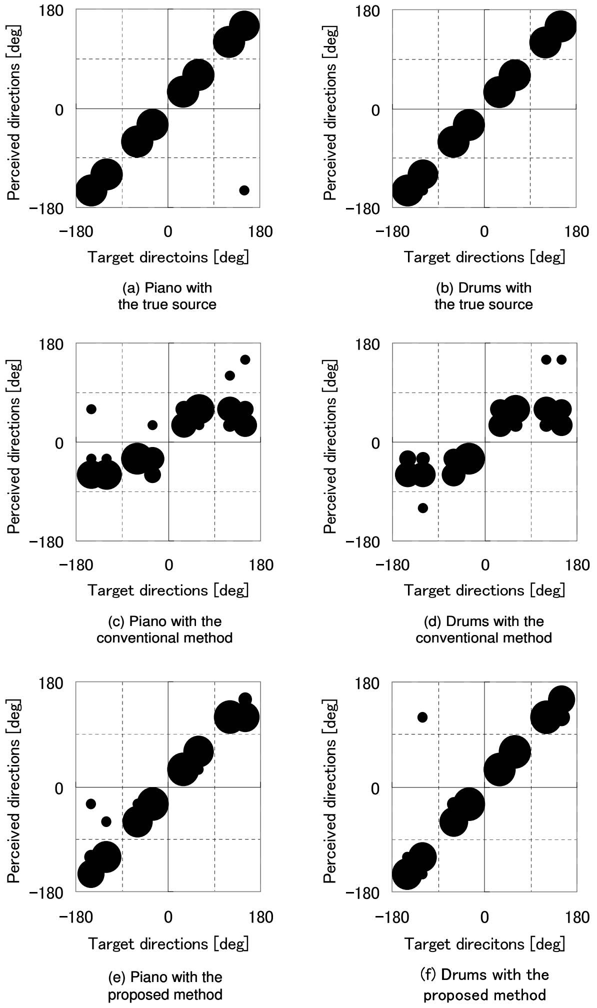
4.2. Comparison of the Source Image Apart from the Sweet Spot

To examine at which directions the listener perceives the source, we performed a subjective evaluation. The subjective experiment was conducted in the same room described at Sect. 4.1. The sound was played back in a random order. The duration of all the signal to be reproduced were 15 seconds. The sweet spot was set on the ears when the listener sits on a chair stood in the center of the room and set his/her head on a headrest of the chair. To prevent the listener from listening to the reproduced sound on the sweet spot, we let the subjects sit on the chair but detach their head from the headrest and move their heads freely. We gave eight candidate directions and they are enforced to choose one direction from which the sources arrive. The sound and the subjects are the same as those in Sect. 4.1.

We show the results of the experiment in Fig. 4. In the figure, (a) and (b) show the results using the true sources, (c) and (d) are the results for the conventional method, (e) and (f) are the proposed method. The results of piano source are shown in (a), (c) and (e), and drums source in (b), (d) and (f). In these figures, the horizontal axes show the true DOAs of the sources in the reproduced signals, the vertical axes show the directions answered by the subjects, and the diameters of the circles show the frequency of the answer. While the conventional method fails to localize sources in the back, the true source and the proposed method could present the source directions to the listeners successfully for both the piano and drums. Therefore it is proved that the proposed method has a faculty to present the source direction even out of the sweet spot.

5. CONCLUSIONS

We proposed an inverse filter design method which is robust against changes of the listening position in the neighborhood of the sweet spot. The proposed inverse filter has minimum distance from the filter to use a specific loudspeaker, and has the largest gain in the channel of the loudspeaker close to the source's direction. The results of subjective experiments showed the efficiency of the proposed method.



(a) Piano with the true source

(b) Drums with the true source

(c) Piano with the conventional method

(d) Drums with the conventional method

(e) Piano with the proposed method

(f) Drums with the proposed method

Figure 4: *The answered directions.*

6. REFERENCES

[1] J. Blauert, *Spatial Hearing,* MIT Press, Cambridge, MA, 1983.

[2] M. R. Schroeder, and B. S. Atal, "Computer simulation of sound transmission in rooms," *IEEE Conv. Rec.,* vol.7, pp.150–155, 1963.

[3] P. A. Nelson, H. Hamada, and S. J. Elliott, "Adaptive inverse filters for stereophonic sound reproduction," *IEEE Transactions on Signal Processing,* vol.40, no.7, pp.1621–1632, 1992.

[4] Y. Tatekura, S. Urata, H. Saruwatari, and K. Shikano, "Online relaxation algorithm applicable to acoustic fluctuation for inverse filter in multichannel sound reproduction system," IEICE Trans. Fundamentals, vol.E88-A, no.7, pp.1747–1756, 2005.

[5] M. Miyoshi, and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust. Speech Signal Process,* vol.36, no.2, pp.145–152, 1988.

[6] P. A. Nelson, O. Kirkeby, T. Takeuchi, and H. Hamada, "Sound fields for the production of virtual acoustic images," *J. Sound Vib.,* vol.204, no.2, pp.386–396, 1997.