J. Acoust. Soc. Jpn. (*E*) **20**, 4 (1999)

Representation of sound localization transfer function and psychoacoustical evaluation

Ikuichiro Kinoshita and Shigeaki Aoki

NTT Cyber Space Laboratories,
3-9-11, Midori-cho, Musashino, 180-8585 Japan

(*Received 26 May 1998*)

A method is proposed for determining representatives of the sound localization transfer function (SLTF) aiming at a compromise between feasibility and localization accuracy. This method is based on principal components analysis (PCA). The magnitudes of SLTF measured at both ears of 56 subjects for 24 targets were modelled by linear combination of six principal components (PC). For each target and ear, the representative was determined to be the SLTF whose weights of PC contributing to the magnitude best approximated the average over subjects. The effectiveness of the representatives was evaluated both statistically and psychoacoustically. The statistical analysis indicates that the individual distribution of the weights of PC is assumed to be a six-dimensional normal distribution that is densest around the average. The discrepancy of the representatives from the average is within 1.5 percent of the individual accumulation. The psychoacoustical evaluation for 28 of the subjects indicates that the increase in the front/back confusion rate in the use of the representatives is significant for frontal targets, but not for other targets. The lateral judgement is robust against SLTFs for all targets tested. However, considerable individual dispersion of the front/back confusion rate hindered from determining the allowance of the discrepancy between individual SLTFs and the representatives for acceptable localization accuracy.

Keywords: Binaural, Rransfer functions, Localization, Representation, Principal components

PACS number: 43. 66. Mk, 43. 66. Qp

1. INTRODUCTION

The application of virtual acoustics can greatly improve the perceived quality of virtual reality and multimedia applications (Begault, 1994, *etc.*). Virtual acoustics are intented to manipulate spatial perception by replicating stimuli from the target source at both ears of a listener. This replication can be implemented by filtering a sound source with filters constructed from target-to-eardrum transfer functions or head-related transfer functions (HRTFs) (Blauert and Laws, 1973, *etc.*). The use of listener's individual sound localization transfer functions (SLTFs), which are a kind of HRTFs, provides precision in the replication (Shimada and Hayashi, 1992). Although the precise replication gives cues for accurate localization at the target position, the need to prepare the listener's individual SLTFs hinders its feasibility (Wenzel *et al.*, 1993). It is more practical to use a pair of SLTF representatives common to potential listeners, even if the replication of stimuli is not precise at both ears.

Conventional means of collecting the SLTF representatives are either (a) measurement with an artificial head as a subject, (b) measurement with a human subject, who can accurately localize a sound image to the target position (Wenzel *et al.*, 1993), or (c) preparation of multiple pairs of SLTF candidates measured with many human subjects (Shimada *et al.*, 1994). However, the effectiveness of approaches (a) and (b) can be evaluated solely on a subjective basis. In contrast, approach (c) is based on clustering the interaural differences in the frequency characteristics of many subjects. However, the clustering ignores monaural frequency characteristics, which is assumed to provide cues to front/back judgement (Blauert, 1982). In addition, the preparation of multiple pairs of SLTF candidates compels the listener to select one single pair for each target position.

In an attempt to make it feasible and effective to use a single pair of SLTF representatives, we propose a method for determining the single pair based on principal components analysis (PCA) (Kinoshita and Aoki, 1995). PCA is an effective method for reducing the complexity in an SLTF by modelling the magnitude, which is generally composed of a large number of variables (typically over hundreds) (Martens, 1987; Kistler and Wightman, 1992; Takahashi and Hamada, 1994). Therefore, the distribution of SLTFs over subjects can be measured with less difficulty in computation and it is helpful in determining the representatives objectively. It should be noted, however, that failure to replicate the stimuli from the target position precisely is one of the factors that deteriorate localization accuracy, *e.g.*, frequent front/back confusion. To confirm the effectiveness of this approach, we conducted a statistical analysis of the distribution of SLTFs and a psychoacoustical evaluation by comparing the SLTF representatives and the listener's individual SLTFs. The results are discussed here in terms of the discrepancy between the both classes of SLTFs.

2. COLLECTION OF SOUND LOCALIZATION TRANSFER FUNCTIONS

The Sound Localization Transfer Function (SLTF) $H_{jkr}(f)$ is a kind of head-related transfer function, defined as a target-to-eardrum transfer function $H_{jkr}(f)$ modified with an headphone-to-eardrum transfer function $E_{jk}(f)$ and the input-output characteristics of the sound driver used for the measurement $V(f)$ (Shimada and Hayashi, 1992). This can be expressed as

$$H_{jkr}(f)=G_{jkr}(f)/(E_{jk}(f)\cdot V(f)). \quad (1)$$

$G_{jkr}(f)$ represents the transfer characteristics from a source at a target position to the eardrums of the listener in a virtual field. As illustrated in Fig. 1, filtering the source signal with the SLTF for each channel replicates the stimuli from the source at the target position at both left and right ears. This replication enables the listener to perceive the sound at the target position (Wightman and Kistler, 1989a; Wightman and Kistler, 1989b; Pralong and Carlile 1996; *etc.*).

Fifty-six subjects participated in the collection of SLTFs. As shown in Fig. 2, the measurements were conducted for each of the subjects in a room 4.4 m deep, 4.7 m wide and 2.5 m high with a reverberation of 200 ms at 500 Hz. The room was acous-

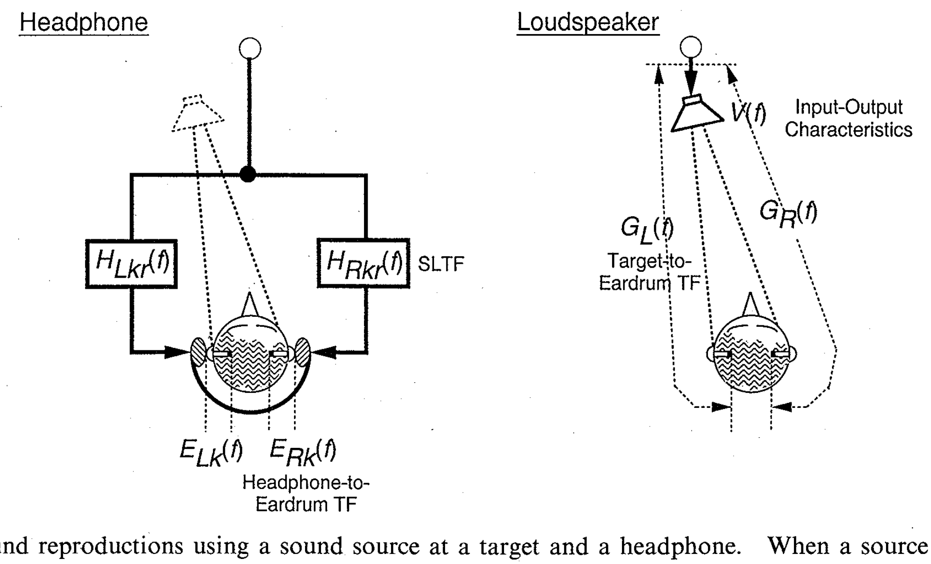

Fig. 1 Sound reproductions using a sound source at a target and a headphone. When a source signal with SLTFs is filtered from the target to each ear, stimuli can be replicated at both ears even when a headphone is used.

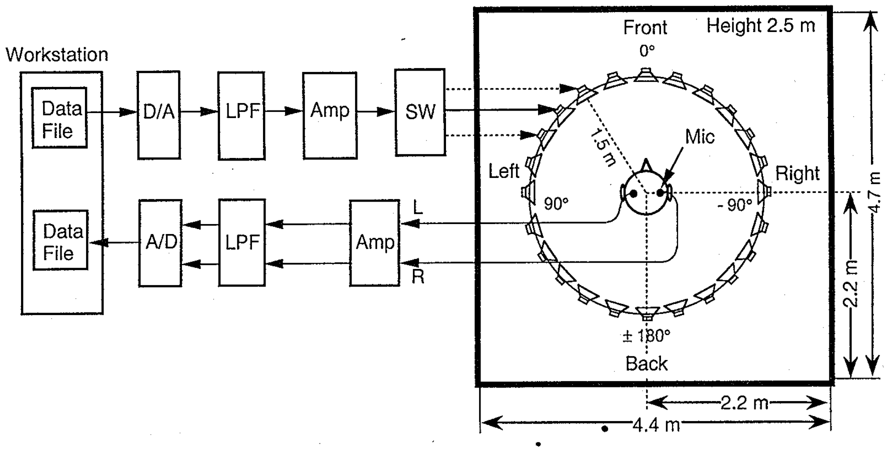

Fig. 2 Illustration of setup for measuring target-to-eardrum transfer functions, composed of a workstation, D/A and A/D converters, filters, power amplifier; 24 locations of loudspeakers and the subject were contained in an anechoic room.

tically insulated from outside with a transmission loss of 90 dB at 500 Hz. Twenty-four loudspeakers (AURATONE 5PSC) were assigned in a horizontal circle of a 1.5m-radius around the subject at intervals of 15 degrees, with its center axis at the same height as the subject's ears. The azimuth of the target position was set counter-clockwise to 0 degree in front of the subject.

Each of the loudspeakers was actuated to present a stimulus for measurement of the target-to-eardrum transfer functions $G_{jkr}(f)$. Selection of the loudspeaker to be actuated was manually controlled with a multi-channel switch. The fast method proposed by Alrutz (1981) was employed to obtain the target-to-eardrum transfer functions $G_{jkr}(f)$ as an impulse response with a length of 2,048. An M-sequence of the 12th order generated in a workstation (SUN ELC) was converted by using a 16-bit digital-to-analog converter (SBS DASBOX) into an analog signal at a sampling frequency of 48.0 kHz. The analog signal was filtered through a low-pass filter (NF P-88) with a cut-off frequency of 20.0 kHz, and then the sound driver was actuated. The response signal was received by means of a miniature probe microphone (RION UC-92A), which was inserted 25 mm deep into each ear canal. The received signal was also enhanced by using an amplifier in a noise level meter (RION NA-20) and filtered through a low-pass filter (NF P-88) with a cut-off frequency of 20.0 kHz. Both channels of the filtered signal were converted into a digital sequence using a 16-bit analog-to-digital converter (SBS DASBOX) at a sampling frequency of 48.0 kHz. The sequence was used to obtain the $G_{jkr}(f)$ as an impulse response.

Likewise, the headphone-to-eardrum transfer function $E_{jk}(f)$ was measured using a headphone (SONY DR-200) as the sound driver. The input-output characteristics of the sound driver $V(f)$ had been measured in advance by using a condensor microphone (B&K 4133) 5 cm off the front center of the driver (Shimada and Hayashi, 1992) in advance.

3. STATISTICAL ANALYSIS AND REPRESENTATION OF SLTFS

3.1 Modelling of SLTF Magnitude based on Principal Components Analysis

The magnitude of the SLTF $|H_{jkr}(f)|$ was modelled to be a linear combination of principal components (PCs) $\boldsymbol{u}_l$ as basic spectra (Afifi and Azen, 1972; Martens, 1987). That is,

$$\boldsymbol{H}_{jkr} = \boldsymbol{U}\boldsymbol{w}_{jkr}\,, \qquad (2)$$

where $\boldsymbol{U}$ is the set of the PCs $[\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_m]$, $\boldsymbol{H}_{jkr}$ is the set of magnitudes of SLTF $[|H_{jkr}(f_1)|, |H_{jkr}(f_2)|, \cdots, |H_{jkr}(f_p)|]^{\mathrm{T}}$ at each frequency f_i $(1 \leqq i \leqq p)$. p denotes the number of elements in a PC $\boldsymbol{u}_l$ and that in a set of SLTF magnitudes $\boldsymbol{H}_{jkr}$. Based on the frequency range 200 Hz to 15.0 kHz, the sampling frequency of 48.0 kHz and the length of 2,048, the p was set to be 632. The indices j, k and r specified the ear, subject and azimuth of the target position, respectively. The PCs $\boldsymbol{u}_l$ were obtained as eigenvectors of a variance-covariance matrix $\boldsymbol{S}$, that can be written as

$$S=\sum_{ijk}\boldsymbol{H}_{jkr}\boldsymbol{H}_{jkr}^{H}/(M-1)\,, \tag{3}$$

where M is the total number of SLTFs (2688, *i.d.* 56 subjects × 2 ears × 24 target positions). The order l of principal components $\boldsymbol{u}_l$ was determined in a decreasing order of the eigenvalues λ_l of the matrix $\boldsymbol{S}$, which represents the contribution to the sum of all squared magnitudes of the SLTF.

We call the set of weights $\boldsymbol{w}_{jkr}=[w_{jkr1}, w_{jkr2}, \cdots, w_{jkrm}]^{\mathrm{T}}$ the weighting vector. To reduce the complexity of the SLTF, the set of weights $\boldsymbol{w}_{jkr}=[w_{jkr1}, w_{jkr2}, \cdots, w_{jkrm}]^{\mathrm{T}}$ was computed as

$$\boldsymbol{w}_{jkr}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{H}_{jkr}\,. \tag{4}$$

The number of elements m was determined according to the cumulative contribution P_m from the first PC $\boldsymbol{u}_1$ up to the m-th PC $\boldsymbol{u}_m$ over 90%. The P_m can be calculated as follows:

$$P_m=\sum_{l=1}^{m}\lambda_l/\sum_{l=1}^{P}\lambda_l\,. \tag{5}$$

At the cumulative contribution of 90%, the use of the reconstructed transfer functions enabled the listener to localize as accurately as with the use of the original transfer functions (Kistler and Wightman, 1992; Takahashi and Hamada, 1994). Therefore, we considered that the magnitude of the reconstructed SLTFs could be approximated to those of the original SLTFs on a subjective basis.

3.2 Method of SLTF Representation

For each ear j and each azimuth r, the SLTF representative $H_{jr}^{*}(f)$ was determined to be the SLTF whose weighting vector $\boldsymbol{w}_{jkr}$ best approximated the average weighting vector $<\boldsymbol{w}_{jr}>$

$$<\boldsymbol{w}_{jr}>=\sum_{k=1}^{hs}\boldsymbol{w}_{jkr}/n_{\mathrm{s}}\,, \tag{6}$$

where n_{s} was the number of subjects. That is the distance between the weighting vector $\boldsymbol{w}_{jr}^{*}$ for the SLTF representative $H_{jr}^{*}(f)$ and the average weighting vector $<\boldsymbol{w}_{jr}>$ should be minimum over the subjects. The distance between the weighting vector $\boldsymbol{w}_{jkr}$ and $<\boldsymbol{w}_{jr}>$ were represented by Mahalanobis' generalized distance (MGD) D_{jkr}, which is defined as follows:

$$D_{jkr}{}^{2}=(\boldsymbol{w}_{jkr}-<\boldsymbol{w}_{jr}>)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{w}_{jkr}-<\boldsymbol{w}_{jr}>)\,, \tag{7}$$

where $\boldsymbol{\Sigma}^{-1}$ is the inverse of variance-covariance matrix $\boldsymbol{\Sigma}$

$$\boldsymbol{\Sigma}=\sum_{k-1}^{ns}(\boldsymbol{w}_{jkr}-<\boldsymbol{w}_{jr}>) \times(\boldsymbol{w}_{jkr}-<\boldsymbol{w}_{jr}>)^{\mathrm{T}}/(n_s-1)\,. \tag{8}$$

As expressed in Eqs. (7) and (8), the MGD is an Euclidean distance that is normalized by variances and covariances among the weights of PC over the subjects. This normalization removes the differences of the contribution among variances and covariances between weights of PCs. Thus, the MGD is expected to model the cumulative individual distribution from the average.

3.3 Results and Discussions

3.3.1 PCA Results

We analyzed the SLTFs those were collected by the procedure described in section 2. According to the PCA, the cumulative contributions P_m from the first PC $\boldsymbol{u}_1$ up to the m-th PC $\boldsymbol{u}_m$ were 60.2, 80.3, 84.5, 86.9, 88.9, 90.5% *etc.* According to the 90% criterion, the number of elements m in the weighting vector $\boldsymbol{w}_{jkr}$ was determined to be 6.

An analysis of variance shows a significant effect of azimuth r on weights w_{jkrm} for each PC $\boldsymbol{u}_m$ [$F(23,1265)=1661.2$, 325.4, 54.9, 54.2, 20.0, 49.1; $p<0.001$ (listed in an increasing order of m)].

3.3.2 SLTF representatives

Figure 3 shows the Mahalanobis' generalized distance (MGD) D_{jr}^{*} between the weighting vector of the SLTF representative $\boldsymbol{w}_{jr}^{*}$ and the average weighting vector over subjects $<\boldsymbol{w}_{jr}>$ as a function of azimuth for each ear. The D_{jr}^{*} were confined within 1.0 for both ears.

A histogram of weighting vectors $\boldsymbol{w}_{jkr}$ and that of the chi-square distribution of the sixth order are shown as functions of MGD D_{jkr} in Fig. 4. In this

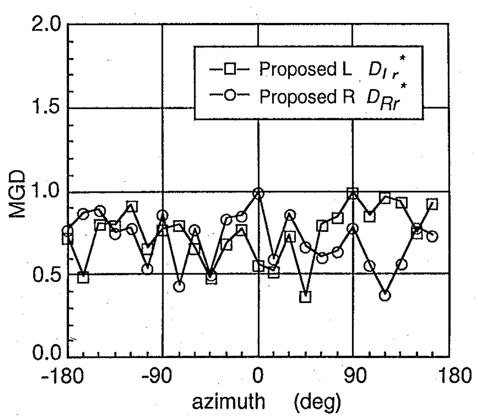

Fig. 3 MGD as a function of azimuth of target position for each class of SLTFs and each ear.

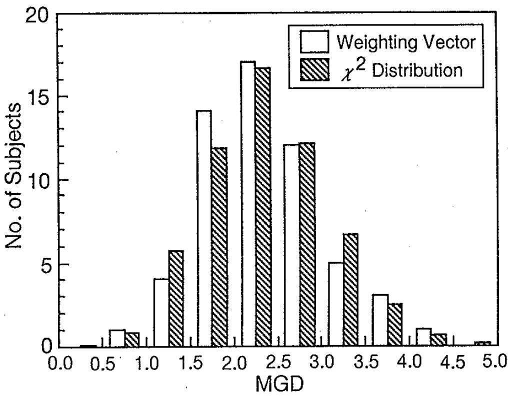

Fig. 4 Histogram of weighting vectors and that of the chi-square distribution of the sixth order as functions of MGD for azimuth 0 degrees.

Table 1 $\chi_9{}^2$ for goodness-of-fit test of histogram of weighting vectors to the sixth order of the chi-square distribution.

azimuth (deg.)	$\chi_9{}^2$ L	$\chi_9{}^2$ R	azimuth (deg.)	$\chi_9{}^2$ L	$\chi_9{}^2$ R
−180	0.236	0.616	0	1.855	2.449
−165	2.054	1.088	15	2.980	1.320
−150	0.870	2.335	30	1.343	2.410
−135	2.658	1.793	45	2.884	1.926
−120	2.276	1.630	60	2.338	1.521
−105	3.055	2.126	75	2.739	2.544
−90	2.678	1.253	90	0.559	1.781
−75	1.136	0.369	105	2.655	3.112
−60	2.069	3.203	120	1.194	2.215
−45	1.498	1.852	135	0.677	2.555
−30	2.283	1.595	150	1.125	2.949
−15	1.095	2.615	165	1.838	2.237

example, the weighting vectors $\boldsymbol{w}_{jkr}$ are based on the SLTFs for the left ear and an azimuth of 0 degrees. The chi-square distribution is scaled by the number of subjects 56. A goodness-of-fit test shows that the histogram of weighting vectors $\boldsymbol{w}_{jkr}$ can approximate the histogram of the chi-square distribution ($\chi_9{}^2=1.855$, $p>0.99$). Also for other azimuths, the result of the goodness-of-fit test is as listed in Table 1. No significant differences were shown between the histogram of weighting vectors $\boldsymbol{w}_{jkr}$ and that of the chi-square distribution for all azimuths tested. This implies that the distribution of the weighting vectors $\boldsymbol{w}_{jkr}$ can be regarded as a six-dimensional normal distribution that is densest around the average.

Since the distribution of the MGD D_{jkr} is assumed to be a chi-square distribution of the sixth order, the cumulative distribution of subjects $P(D_{jkr})$ can be approximated as follows.

$$P(D_{jkr})=\int_0^T \chi_6{}^2(t)\,dt\,, \tag{9}$$

where $T=D_{jkr}{}^2$.

According to the MGDs D_{jr}^* mentioned above, the cumulative distributions $P(D_{jr}^*)$ are as follows. For the maximum D_{jr}^* of the SLTF representatives 1.0, $P(1.0)=0.0144$. According to the $P(D_{jr}^*)$, the magnitude of the SLTF representative $\boldsymbol{H}_{jr}^*$ is in the vicinity of the average $<\boldsymbol{H}_{jr}>$, where the distribution of the individual SLTFs was densest. This result suggests that the use of the SLTF representatives $H_{jr}^*(f)$ can minimize the proportion of potential listeners with poor localization accuracy to the target position. It is based on the assumption that deterioration of localization accuracy is due to distance between the spectrum of the SLTF representative $H_{jr}^*(f)$ and that of listener's individual SLTF $H_{jkr}(f)$.

4. PSYCHOACOUSTICAL EVALUATION

4.1 Experiment

In order to confirm the effectiveness of the SLTF representative $H_{jr}^*(f)$, we conducted listening tests. Deterioration in the localization accuracy was assessed in comparison with the subject's individual SLTFs $H_{jkr}(f)$. Therefore, 28 of the 56 subjects with whom the transfer functions had been measured, participated in the experiment. All the subjects were of the ages of 24 to 48 and had normal hearing.

4.1.1 Setup

Tests were conducted in a room in which all the transfer functions $E_{jk}(f)$ and $G_{jkr}(f)$ had been measured. As shown in Fig. 5, stimuli were presented through a headphone (SONY DR-200), which was used for the measurement of the headphone-to-eardrum transfer functions $E_{jk}(f)$. The rest of the setup was located outside of the room.

Source signals for the stimuli were produced by using a broadband noise convolved with an impulse response of individual SLTF $h_{jkr}(n)$ or that of the SLTF representative $h_{jr}^*(n)$ by means of a convolver (NIPPON KYASTEM Sim**2) for each channel. The source signal was a quasi-random digital sequence generated in another computer (HP-755)

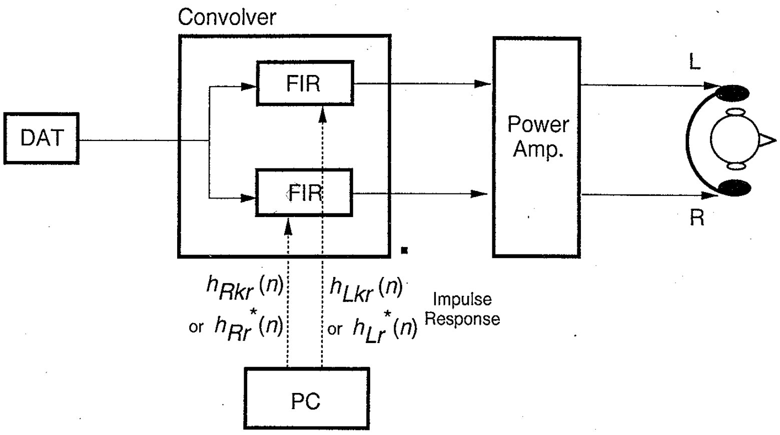

Fig. 5 Illustration of setup for listening test.

and recorded by digital audio tape recorder (SONY 77ES). The playback signal was converted to an analog signal using a digital-to-analog converter (SBS DASBOX 200) at a sampling frequency of 48 kHz and passed through a low-pass filter with a cutoff frequency of 20 kHz (NF P-88). The level of the stimuli was set to be 70 dB (A-weighted) at the miniature microphone that was used for measurement of the transfer functions $E_{jk}(f)$ and $G_{jkr}(f)$.

4.1.2 Procedure

The SLTF representative $H_{jr}^{*}(f)$ and the individual SLTFs $H_{jkr}(f)$ for the 24 azimuths of target position were investigated for each subject in 480 trials. The sequence of the 48 conditions (2 classes × 24 target positions) of SLTFs $H_{jr}^{*}(f)$ or $H_{jkr}(f)$ was randomly set in advance. That is, the number of trials for each condition N was 10. The experiment with 480 trials was managed for each trial block with 120 trials in the random sequence. The subject was not informed with the condition of SLTFs to be tested.

In each trial, the stimulus was binaurally presented for 2 seconds. The subject was supposed to sit on a stool located at about the center of the room. The subject was instructed to respond to the apparent position of the sound by clicking a mouse on a diagram (Fig. 6) displayed on a monitor (Macintosh SE) immediately after presentation of the stimulus.

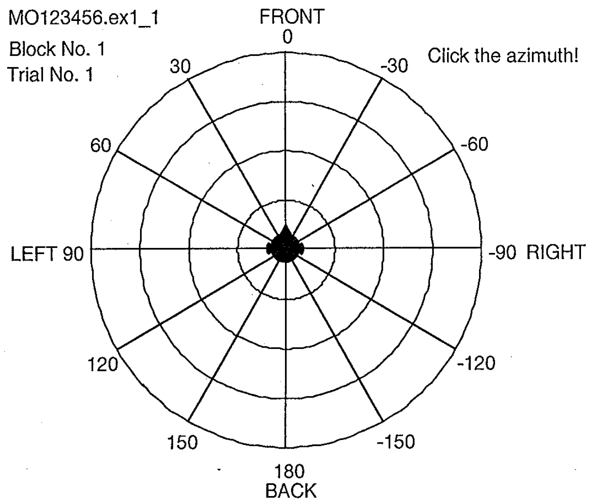

Fig. 6 Diagram for responding the apparent position.

4.2 Results and Discussions

4.2.1 Measures of deterioration

The deterioration in localization accuracy for SLTF representatives $H_{jr}^{*}(f)$ was scaled for each subject k by using the differential front/back confusion rate ΔN_{FBk} and the error angle $\Delta\Theta_k$. The differential front/back confusion rate ΔN_{FBk} was defined as

$$\Delta N_{FBk}=(N_{FB}^{*}-N_{FBk})/N\,, \tag{10}$$

where N_{FB}^{*} and N_{FBk} are the numbers of trials for front/back confusion when the SLTF representatives $H_{jr}^{*}(f)$ were used and when the individual SLTFs $H_{jkr}(f)$ were used, respectively. The increase in the differential front/back confusion rate ΔN_{FBk} indicates the deterioration of the accuracy in the front/back judgement for the SLTF representatives $H_{jr}^{*}(f)$.

The error angle $\Delta\Theta_k$ was calculated as the mean difference between the azimuth of the apparent position for the SLTF representatives $H_{jr}^{*}(f)$ and that for the individual SLTFs $H_{jkr}(f)$ for each azimuth of target position and subject k. That is,

$$\Delta\Theta_k=\sum_{s=1}^{N}(\theta_{ks}^*-\theta_{ks})/N\,, \tag{11}$$

where the θ_{ks}^* and the θ_{ks} are the apparent azimuth for the SLTF representatives $H_{jr}^*(f)$ at a trial s and that for the individual SLTF $H_{jkr}(f)$, respectively. Prior to calculation, the apparent positions were modified such that the front/back judgement was resolved. Therefore, the greater the discrepancy in lateral judgement between the two classes of SLTFs is, the error angle $\Delta\Theta_k$ should increase.

4.2.2 Differential front/back confusion rate ΔN_{FBk}

Figure 7 shows the average of the differential front/back confusion rate over subjects $<\Delta N_{FB}>$ as a function of the azimuth of the target position. $<\Delta N_{FB}>$ for frontal azimuths apparently exceeds that for backward azimuths. The use of the SLTF representatives $H_{jr}^*(f)$ contributes to the excess, given the average front/back confusion rates that resulted: 0.45 for maximum $<N_{FB}^*/N>$ (azimuth$=$15°), 0.04 for minimum $<N_{FB}^*/N>$ (azimuth$=-150$°) when SLTF representatives $H_{jr}^*(f)$ were used; 0.16 for maximum $<N_{FBk}/N>$ (azimuth$=$15°), 0.03 for minimum $<N_{FBk}/N>$ (azimuth$=$165°) when individual SLTFs $H_{jkr}(f)$ were used. Based on an analysis of variance based on the N_{FB}^* and the N_{FBk}, the non-zero significance of $<\Delta N_{FB}>$ was tested. As listed in Table 2, the $F(1,27)$ values indicate the non-zero significance for azimuths of -60 to 60 degrees [$F(1,27)>4.21$, $p<0.05$], but not for other azimuths, including the negative values for backward azimuths of -150°, -135°, -120° and 165°.

Figure 8 shows an example of the differential front/back confusion rate ΔN_{FBk} as a function of modified MGD D_{kr}' between the individual SLTF and SLTF representative for an azimuth of 45 degrees. ΔN_{FBk} is plotted with circles for each subject k. The modified MGD D_{kr}' is defined as follows:

$$D_{kr}'=\sqrt{D_{Lkr}{}^2+D_{Rkr}{}^2}\,, \tag{12}$$

In this example, the differential front/back confusion rate ΔN_{FBk} appears to disperse independently of D_{kr}'. The nonzero significance of the correlation between the ΔN_{FBk} and the D_{kr}' was evaluated by using the value $t(26)$ for each azimuth. According to the $t(26)$ values listed also in Table 2, the correlation between the ΔN_{FBk} and the D_{kr}' was

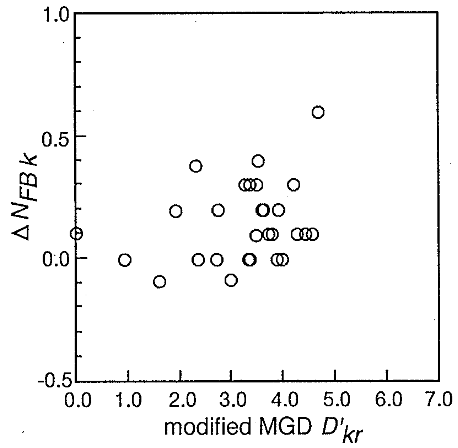

Fig. 8 Differential front/back confusion rate ΔN_{FBk} as a function of modified MGD D_{kr}' between the subject's individual SLTFs and the SLTF representatives (azimuth$=$45°).

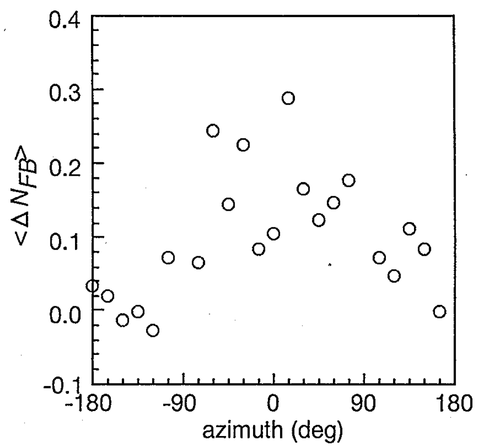

Fig. 7 Average of differential front/back confusion rate $<\Delta N_{FB}>$ as a function of azimuth of the target position for a subject.

Table 2 $t(26)$ and $F(1,27)$ for the differential front/back confusion rate ΔN_{FBk} as a function of azimuth.

azimuth (deg)	$F(1,27)$	$t(26)$	azimuth (deg)	$F(1,27)$	$t(26)$
−180	1.91	−1.13	0	7.76	1.30
−165	3.24	−0.35	15	5.26	2.70
−150	0.81	0.72	30	10.75	2.25
−135	0.69	−0.74	45	17.06	2.36
−120	0.80	0.24	60	11.20	0.82
−105	1.30	−0.15	75	4.12	1.67
−90	—	—	90	—	—
−75	2.15	−0.45	105	1.28	−0.29
−60	8.29	1.69	120	1.77	0.63
−45	5.02	1.22	135	2.19	0.43
−30	11.23	2.42	150	2.79	0.91
−15	5.94	2.30	165	0.87	−0.65

significantly nonzero for frontal azimuths of target positions of −30, −15, 15, 30 and 45 degrees but not for other azimuths [$t(26)>2.056$, $p<0.05$]. These results indicate that the larger the discrepancy between the magnitudes of each class of SLTFs is, the more frequently errored front/back judgement tends to be for frontal target positions. The $<\Delta N_{FB}>$ was maximized to be 0.289 at an azimuth of 15 degrees. For other azimuths of target position, front/back judgement was found to be insensitive to the discrepancy in magnitude.

4.2.3 Error angle $\Delta\Theta_k$

Figure 9 shows the average of the error angle over subjects $<\Delta\Theta>$ as a function of azimuth of the target position. The $<\Delta\Theta>$ for frontal azimuths between −60 and −15 degrees and that for 30 degrees apparently exceeded those for other azimuths. However, statistical analyses on the apparent azimuth for two classes of SLTFs indicated that the $<\Delta\Theta>$ was not significantly non-zero for all azimuths tested [$F(1, 27)<4.21$, $p>0.05$; see Table 3.]

Table 3 $t(26)$ and $F(1, 27)$ for the error angle $\Delta\Theta_k$ as a function of azimuth of the target position.

azimuth (deg)	$F(1,27)$	$t(26)$	azimuth (deg)	$F(1,27)$	$t(26)$
−180	0.99	0.77	0	3.68	−0.81
−165	0.58	−0.04	15	0.64	−1.58
−150	0.83	0.44	30	2.57	−0.91
−135	2.99	0.35	45	0.77	0.59
−120	0.75	0.03	60	0.57	0.52
−105	0.62	−0.47	75	1.57	0.41
−90	0.82	−1.17	90	0.99	0.39
−75	0.50	1.26	105	0.19	0.23
−60	3.96	−0.10	120	2.21	0.28
−45	1.33	−0.15	135	2.64	−0.48
−30	2.15	0.50	150	2.38	0.58
−15	2.54	1.98	165	2.32	−0.43

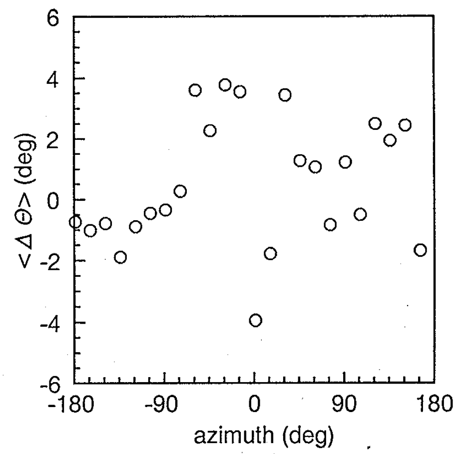

Fig. 9 Average of error angle $<\Delta\Theta>$ as a function of azimuth of the target position for a subject.

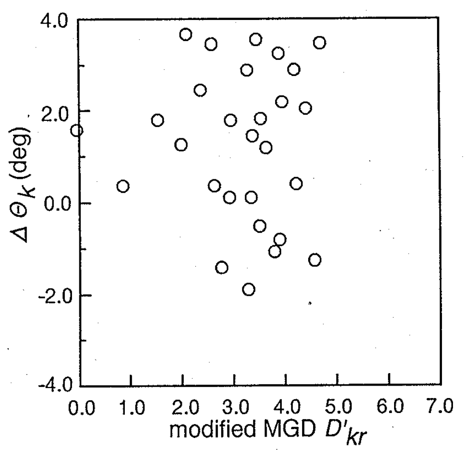

Fig. 10 Error angle $\Delta\Theta_k$ as a function of modified MGD D_{kr} between the subject's individual SLTFs and the SLTF representatives (azimuth=45°).

Figure 10 shows the error angle $\Delta\Theta_k$ plotted as a function of modified MGD D_{kr}' between the two classes of SLTFs for an azimuth of 45 degrees, with no systematical variation in the $\Delta\Theta_k$ over D_{kr}'. The significance of correlation with the $\Delta\Theta_k$ and the D_{kr}' was tested for each azimuth. Table 3 also lists the $t(26)$ for the error angle $\Delta\Theta_k$. In relation to the $t(26)$, the correlation with the D_{kr}' over subjects is insignificant for any azimuths [$t(26)>2.056$, $p<0.05$]. The insignificant correlation originates not only from the difference of $\Delta\Theta_k$ among subjects but also from that of the apparent azimuth among trials. The $\Delta\Theta_k$ ranges within the variance of the apparent azimuth among trials, typically 5 degrees. As regards the variation in the apparent azimuths, the insignificance in the $\Delta\Theta_k$ revealed the robustness of lateral judgement regardless of the class of SLTFs.

4.3 Discussions

We now further discuss the results in subsections 4.2.2 in connection with the individual distribution of MGDs. Recall that the magnitude of the SLTF representative $\boldsymbol{H}_{jr}^{*}$ are in the vicinity of the average $<\boldsymbol{H}_{jr}>$, where the distribution of the $\boldsymbol{H}_{jkr}$ is densest. With respect to the density, the magnitude of the SLTF representative $\boldsymbol{H}_{jr}^{*}$ may approximate those of the majority of potential listeners $\boldsymbol{H}_{jkr}$. This approximation reduces the overall distortion of spectrum of SLTF, which gives cues to front/back judgement.

However, front/back confusion was frequently observed in the use of the SLTF representative $H^*_{jr}(f)$ for frontal azimuth of target position. In addition, the differential front/back confusion rate $\Delta N_{\mathrm{FB}k}$ was found to be individually dispersed, irrelevantly to MGD. These results indicates that the sensitivity to spectral distortion, that is assumed to distract front/back judgement (Begault, 1991; *etc.*), depends on listener.

Owing to the considerable dispersion in the differential front/back confusion rate $\Delta N_{\mathrm{FB}k}$ among subjects, we were unable to determine the limit of allowable discrepancy between the individual SLTF $H_{jkr}(f)$ and the SLTF representative $H^*_{jr}(f)$. For example, at an azimuth of 45 degrees, the interval of confidence in the correlation ρ between the $\Delta N_{\mathrm{FB}k}$ and the modified MGD D_{kr}' was found to be $0.36 \leqq \rho \leqq 0.84$ $[p<0.95]$ (Afifi and Azen, 1972). When the $\Delta N_{\mathrm{FB}k}$ is allowed to be up to 0.25, the modified MGD D_{kr}' can range between 0.288 for ρ of 0.36 and 0.471 for ρ of 0.84. In relation to the modified MGD D_{kr}' of 2.88 and 4.71, the probability integral over subjects $P_{12}(D_{kr}')$ are respectively 28 and 97%. The $P_{12}(D_{kr}')$ was defined as follows:

$$P_{12}(D_{kr}')=\int_0^T \chi_{12}{}^2(t)\,dt\,, \qquad (13)$$

where $T=D_{kr}'2$ and $\chi_{12}{}^2(t)$ is the twelfth order of chi-square distribution.

The broad range of probability integral over subjects $P_{12}(D_{kr}')$ may well deprive the allowance limit of significance.

5. CONCLUSION

We have proposed a method of representating SLTFs based on PCA, in an attempt to find a compromise between the feasibility and localization accuracy. For each ear and position, the representative was determined to be the SLTF whose weights representing the contribution of each basic spectrum best approximated the average weights.

According to the statistical evaluation, the spectrum of the representative can approximate that of the SLTFs for the majority of potential listeners. This notion is supported by the following facts: (1) The distribution of weights of SLTF among subjects is assumed to be a multi-dimentional normal distribution that is densest around the average. (2) The weights of the representative can approximate the average, compared with the distribution of weights of SLTF among subjects.

Psychoacoustical evaluation of the effectiveness of SLTF representatives was considered in terms of the significance of increase in front/back confusion rate and that of lateral judgement error. Compared with the subject's individual SLTFs, increases in the front/back confusion rate and the lateral judgement errors were found to be insignificant in most cases. In other words, localization accuracy to the target position was maintained even in the use of SLTF representatives. In contrast, the front/back confusion rate increased for frontal azimuths of target position, compared with that for other azimuths. However, a limit of allowable discrepancy between the SLTF representatives and the individual SLTFs could not be determined due to the considerable dispersion in the front/back confusion rate among subjects.

REFERENCES

Afifi, A. A. and Azen, S. P. (1972). *Statistical Analysis: A Computer Oriented Approach* (Academic Press, New York).

Alrutz, H. (1981). "Ein neuer Algorithmus zur Auswertung von Messungen mit Pseudorauschsignalen," Fortschr. Akust. DAGA '81, Berlin, 525–528.

Begault, D. R. (1994). *3D SOUND* (AP Professional, Cambridge, MA).

Begault, D. R. (1991). "Challenges to the successful implementation of 3-D sound," J. Audio Eng. Soc. **39**, 864–870.

Blauert, J. (1982). *Spatial Hearing* (MIT Press, Cambridge, MA).

Blauert, J. and Laws, P. (1973). "Verfahren zur orts- und klanggetreuen Simulation von Lautsprecherbeschallungen mit Hilfe von Kopfhörern," Acustica **29**, 273–277.

Kinoshita, I. and Aoki, S. (1995). "Distributions and sound localization transfer functions," Proc. Autumn Meet. Acoust. Soc. Jpn. (I), 523–524 (in Japanese).

Kistler, D. J. and Wightman, F. L. (1992). "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," J. Acoust. Soc. Am. **91**, 1637–1647.

Martens, W. L. (1987). "Principal components analysis and re-sysnthesis of spectral cues to perceived direction," Proc. International Computer Music Conf., 274–281.

Pralong, D. and Carlile, S. (1996). "The role of individualized headphone calibration for the generation of high fidelity virtual auditory space," J. Acoust. Soc. Am. **100**, 3785–3793.

Shimada, S. and Hayashi, S. (1992). "Stereophonic sound image localization system using inner earphones," Proc. FASE (Federation of the Acoustical Societies of Europe) Congr. 1992, 157-160.

Shimada, S., Hayashi, N., and Hayashi, S. (1994). "A clustering method for sound localization functions," J. Audio Eng. Soc. **42**(7/8), 577-584.

Takahashi, M. and Hamada, H. (1994). "A study of head-related transfer function model (2nd Report)," Proc. Autumn Meet. Acoust. Soc. Jpn., Vol. 1, 659-660 (in Japanese).

Wenzel, E. M., Arruda, M., Kistler, D. J. and Wightman, F. L. (1993). "Localization using nonindividualized head-related transfer functions," J. Acoust. Soc. Am. **94**, 111-123.

Wightman, F. L. and Kistler, D. J. (1989a). "Headphone simulation of free-field listening. I: Stimulus synthesis," J. Acoust. Soc. Am. **85**, 858-867.

Wightman, F. L. and Kistler, D. J. (1989b). "Headphone simulation of free-field listening. II: Psychophysical validation," J. Acoust. Soc. Am. **85**, 868-878.