【カテゴリーⅡ】
日本建築学会環境系論文集 第88巻 第806号，243-254，2023年4月
J. Environ. Eng., AIJ, Vol. 88, No. 806, 243-254, Apr., 2023
DOI https://doi.org/10.3130/aije.88.243

VERTICAL AND HORIZONTAL SPREADS OF A SOUND IMAGE INDUCED BY EARLY REFLECTIONS FROM THE LATERAL OR VERTICAL DIRECTION

側方あるいは上下方向から到来する初期反射音による上下及び水平方向への音像の拡がり

Yuki TAKADA [*1] *and Makoto OTANI* [*2]
高 田 有 紀，大 谷 　 真

This study experimentally investigates the effects of early lateral and vertical reflections on the horizontal spread (HS), that is auditory source width (ASW), and vertical spread (VS) of a sound image. A psychoacoustical experiment employed early lateral energy fractions L_f and early vertical energy fractions V_f for controlling auditory stimuli. HS increases with a larger L_f. In contrast, the effect of V_f on HS is limited, although HS is larger when early vertical reflections are present. Results also reveal that VS increases with a larger L_f or V_f, but the effect of L_f on VS is greater than V_f.

Keywords: *Concert hall acoustics, Sound image, Horizontal spread, Vertical spread, Early reflection*
ホール音響，音像，水平方向への拡がり，上下方向への拡がり，初期反射音

1. Introduction

Spatial impression (SI) is a critical perceptual measure for evaluating acoustics in a concert hall. SI is comprised of two perceptual measures: auditory source width (ASW) and listener envelopment (LEV) [1),2)]. ASW is defined as the width of a sound image fused temporally and spatially with a direct sound image. LEV is defined as the degree of fullness of sound images around the listener, excluding the sound image that composes ASW [1)].

Many researchers have investigated ASW and proposed various relevant physical measures. The physical factors affecting ASW include the sound pressure level, spectral contents, interaural cross-correlation coefficient (IACC), and early lateral energy fraction L_f [3)]. Among them, $IACC_{E3}$ [4),5)], that is, IACC for early reflections averaged among 500 Hz, 1 and 2 kHz octave bands, and L_f are important measures in acoustic design of concert halls as $(1 - IACC_{E3})$ is proportional to ASW [4),5)] and L_f correlates with ASW [3)]. Here, an early reflection is generally defined as a reflected sound arriving between 0 and 80 ms after direct sounds. Overall, ASW generally increases as L_f becomes larger or as IACC becomes smaller for early reflections. In addition, late reverberation, that is, reflected sounds arriving after 80 ms, affects LEV [6),7)].

On the other hand, a sound image spreads not only in the horizontal direction but also in the vertical direction [8),9)]. Hereafter, horizontal and vertical spreads of a sound image induced by early reflections are called HS and VS, respectively. HS is identical to ASW. In contrast, very little is known about VS.

Furuya *et al.* [10)] performed experiments to investigate the effects of early reflections from the upper direction on SI. They defined an early reflection as a reflected sound arriving between 0 and 200 ms after a direct sound. They found that VS increases as the delay time between the direct and reflected sounds increases from 10 to 40 ms when a single reflected sound arrives from 50° in the upper direction of the median plane. Additionally, LEV is induced by a reflected sound arriving from non-lateral directions within 200 ms after the direct sound. They defined V_{f200}, which is a quantitative parameter for the energy of reflections arriving from the upper direction, as the ratio of the

[*1] Grad. Student, Graduate School of Engineering, Kyoto University, M. Eng. 京都大学大学院工学研究科　大学院生・工修
[*2] Assoc. Prof., Graduate School of Engineering, Kyoto University, Dr. Eng. 京都大学大学院工学研究科　准教授・工博

vertical energy to the total energy of reflections arriving within 200 ms of the direct sound. Both the upper and lateral reflections may contribute on LEV. Although they do not result in different SIs, they do have differing degree of influence to LEV.

Hanyu *et al.* [11] conducted an experiment to examine if the late reverberation from the upper and lateral directions lead to different SIs by defining two parameters, $Lf_{T-\infty}$ and $Vf_{T-\infty}$, which represent the lateral energy ratio and the vertical energy ratio, respectively. SIs induced by an upper or lateral late reverberation can be perceptually discriminated, indicating that the upper late reverberation induces a distinct auditory effect from that induced by the lateral late reverberation.

To summarize, it has yet to be revealed how early reflections from vertical (upper or lower) directions affect HS or VS and how their effects differ from those by early lateral reflections. Furuya *et al.* mentioned that further studies on the effects of early reflections on VS are necessary [10]. In addition, how early reflections from lateral directions affect VS remains unclear. This study performs a psychoacoustical experiment by employing two physical parameters, L_f and the vertical energy fractions, V_f, to clarify how early reflections from vertical or lateral directions induce HS and VS.

2. Experiment

A psychoacoustical experiment was performed to investigate the effects of early reflections arriving from the lateral or vertical direction on HS and VS. Three loudspeakers were used to simulate a direct sound and two early reflections. An auditory stimulus consisted of a direct sound presented from a loudspeaker located in front of the participant and two reflected sounds presented from a pair of loudspeakers located on the left and right for lateral reflections or upper and lower for vertical reflections. The participants listened to auditory stimuli. They were instructed to denote the position of the center of the sound image and edges of the sound image in the left, right, upper, and lower directions. Two physical parameters, L_f and V_f, were employed to control the lateral and vertical energy ratios compared to the direct sound energy of the auditory stimuli, respectively.

2.1. Physical Parameters

Two physical parameters, L_f and V_f, were used to control the auditory stimuli presented to participants in the psychoacoustical experiment. L_f and V_f are respectively defined as

$$L_f = \frac{\int_{5\,\mathrm{ms}}^{80\,\mathrm{ms}} p_{\mathrm{L}}^2(t)dt}{\int_{0\,\mathrm{ms}}^{80\,\mathrm{ms}} p^2\,(t)dt} \tag{1}$$

and

$$V_f = \frac{\int_{5\,\mathrm{ms}}^{80\,\mathrm{ms}} p_{\mathrm{V}}^2(t)dt}{\int_{0\,\mathrm{ms}}^{80\,\mathrm{ms}} p^2\,(t)dt}, \tag{2}$$

where $p(t)$ denotes an omnidirectional impulse response at a receiver. $p_{\mathrm{L}}(t)$ and $p_{\mathrm{V}}(t)$ indicate the impulse responses measured by a bidirectional microphone whose null direction is directed to the sound source and the maximum sensitivity direction is directed to the lateral and vertical direction, respectively.

2.2. Experimental Condition

This experiment employed three loudspeakers. The loudspeaker located in front of a participant radiated a sound corresponding to the direct sound. Two loudspeakers located in the horizontal or median plane radiated two sounds corresponding to the reflected sounds. The participants were asked to listen to the auditory stimuli and then indicate the positions of edges and the center of the perceived sound image.

2.3. Auditory Stimuli

The auditory stimulus signal was a pink noise (duration: 3 s, sampling frequency: 44.1 kHz, quantization: 16 bit). Figure 1 illustrates the loudspeaker arrangement. Here, the five loudspeakers located in front, +30° and −30° of azimuth, and +30° and −30° of elevation are labeled as F, L, R, U, and D, respectively. In each condition, loudspeaker F presented a sound corresponding to the direct sound. Sounds corresponding to the first and second reflected sounds were presented with one of the L-R, R-L, U-D, and D-U pairs. Sound pressure levels were about 50 dB (A-weighted) at the center of the participant's head when the direct and two reflected sounds were presented.

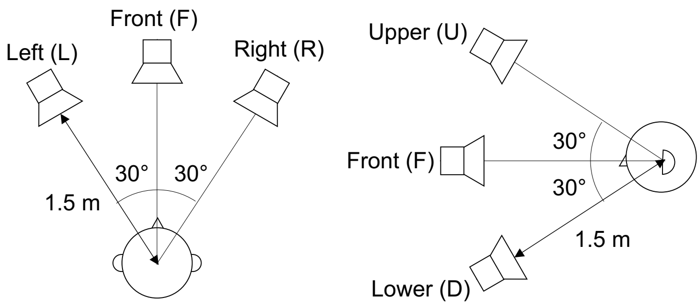


Fig. 1 Loudspeaker arrangement.

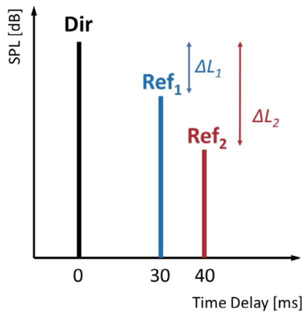


Fig. 2 Schematic of the temporal structure of the impulse responses used as auditory stimuli.

Table 1 Experimental conditions of auditory stimuli

Stimulus ID	Loudspeakers Dir	Ref$_1$	Ref$_2$	ΔL_1 [dB]	ΔL_2 [dB]	L_f	V_f	1−IACC$_{E3}$
1	F	L	R	−12	−12	0.0428	-	0.0947
2	F	L	R	−9	−9	0.0730	-	0.1671
3	F	L	R	−6	−9	0.0916	-	0.2206
4	F	L	R	−6	−6	0.1180	-	0.2735
5	F	L	R	−3	−6	0.1414	-	0.3342
6	F	L	R	−3	−3	0.1738	-	0.4033
7	F	R	L	−12	−12	0.0428	-	0.0948
8	F	R	L	−9	−9	0.0730	-	0.1673
9	F	R	L	−6	−9	0.0916	-	0.2209
10	F	R	L	−6	−6	0.1180	-	0.2739
11	F	R	L	−3	−6	0.1414	-	0.3346
12	F	R	L	−3	−3	0.1738	-	0.4039
13	F	U	D	−12	−12	-	0.0428	0.0004
14	F	U	D	−9	−9	-	0.0730	0.0005
15	F	U	D	−6	−9	-	0.0916	0.0005
16	F	U	D	−6	−6	-	0.1180	0.0006
17	F	U	D	−3	−6	-	0.1414	0.0005
18	F	U	D	−3	−3	-	0.1738	0.0006
19	F	D	U	−12	−12	-	0.0428	0.0004
20	F	D	U	−9	−9	-	0.0730	0.0005
21	F	D	U	−6	−9	-	0.0916	0.0006
22	F	D	U	−6	−6	-	0.1180	0.0006
23	F	D	U	−3	−6	-	0.1414	0.0007
24	F	D	U	−3	−3	-	0.1738	0.0006
25	F	-	-	-	-	0	0	0.0004

Figure 2 depicts a schematic of the temporal structure of an impulse response used to synthesize the auditory stimuli. The delays of the first and second reflected sounds, labeled Ref$_1$ and Ref$_2$, respectively, to the direct sound, labeled Dir, Δt_1 and Δt_2 were fixed to 30 and 40 ms, respectively, while their relative sound pressure levels ΔL_1 and ΔL_2 were varied (Table 1). For comparison, the experiment included the condition where only a direct sound was presented. Table 1 shows the values of L_f and V_f derived for each condition by measuring p, p_L, and p_V using a first-order microphone array (AMBEO VR mic, Sennheiser) with the conversion from A-format to B-format [12]. For reference, the

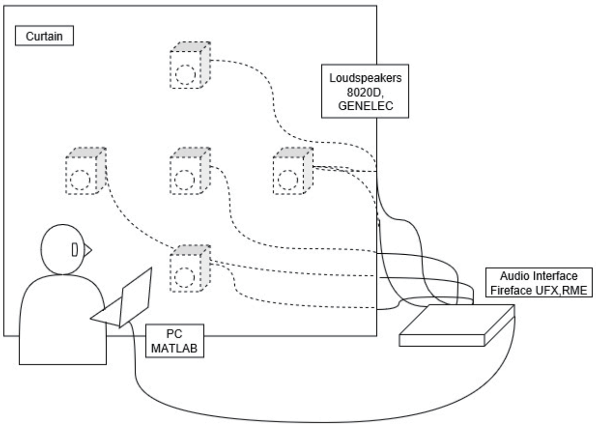


Fig. 3 Experimental setup.

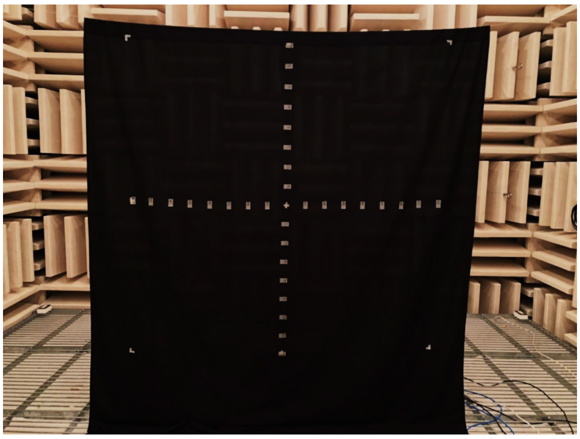


Fig. 4 Reference markers on the black curtain.

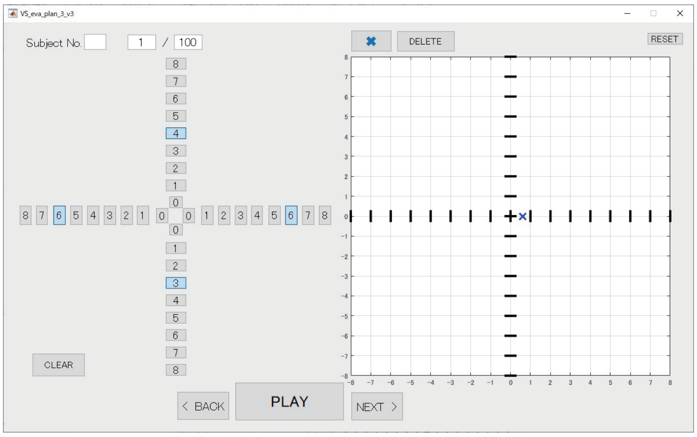


Fig. 5 Graphical user interface (GUI) to indicate the center and edge positions of a sound image.

value of 1–IACC$_{E3}$ is also shown for each stimulus by calculating impulse responses at both ears as a sum of head-related impulse responses (HRIRs) for the three point sources corresponding to a direct and two reflected sounds. The HRIRs were numerically simulated by using the boundary element method and a computer model of a head and torso simulator (4128C, Brüel & Kjær)[13]. The 1–IACC$_{E3}$ values highly correlate with L_f (L-R: linear regression $y = 2.2611x - 0.0032$, $R^2 = 0.9984$) as suggested in the literature[3]. Slight differences in the 1–IACC$_{E3}$ values between L-R (stimulus ID: 1-6) and R-L (stimulus ID: 7-12) are due to a left-right asymmetricity of HRIRs. On the other hand, the 1–IACC$_{E3}$ values for U-D and D-U (stimulus ID: 13-24) are approximately zeros because the signals at both ears are almost identical for sounds coming from the median plane.

2.4. Experimental Setup

The experiment was performed in an anechoic chamber (background noise level: 14.3 dB (A-weighted)) in Katsura Campus, Kyoto University. Figure 3 illustrates the experimental setup. All stimuli were generated in a personal computer (PC) (Probook 430G3, HP) and presented to the participants via loudspeakers (8020D, Genelec) through an audio interface (Fireface UFX, RME). To eliminate visual effects, an acoustically transparent black curtain was placed 1 m from the center of the participant's head. The curtain contained reference markers: one for the center of the curtain, corresponding to the frontal loudspeaker (F); eight markers in 10 cm intervals the left, right, upper, and lower directions numbered from 1 to 8 where the outermost (No. 8) marker's position in each direction corresponds to the loudspeaker L, R, U, or D (Fig. 4). During the experiment, the lighting in the anechoic chamber was dimmed to the extent that the markers on the black curtain, which the participant referenced in their responses, were visible.

2.5. Procedures

Nine participants between 22-25 years of age (mean age: 23.0 ± 1.07 [standard deviation] years). All had normal hearing sensitivity. The participants' responses were collected using a graphical user interface (GUI) operating with the PC and MATLAB (Mathworks) (Fig. 5).

An auditory stimulus, which consisted of a direct sound and two reflected sounds, was presented after the participant clicked the "PLAY" button on the PC display. Participants could repeat the stimulus as many times as they desired. They were instructed to indicate the positions of the sound image edges in the left, right, upper, and lower directions by clicking one of the numbered buttons in each direction in the left panel of the GUI. The buttons corresponded to the numbered markers on the curtain. The button with "0" corresponded to the center of the curtain, which could be selected as the sound image edge. The participants were also instructed to indicate the center of the sound image by clicking the corresponding position in the coordinate system displayed in the right panel of the GUI. The markers also corresponded to those on the curtain. The next trial started after indicating all the positions and then clicking the "NEXT" button.

Prior to the experiment, each participant listened to 13 representative conditions. To familiarize themselves with the auditory stimuli, they listened to each one once and completed the response procedure on the GUI. The experiment presented each condition four times. Thus, each participant completed 100 (25 conditions × 4 repetitions) trials. To maintain concentration, a break was taken during the trials. The auditory stimuli were presented randomly.

3. Results

The result of the one participant who perceived two separate sound images in the upper and lower directions for VS were excluded in the following analysis. All of other participants perceived one sound image in upper and lower directions for VS. For each trial and participant, the angles of the center of the sound image and the edges in the left, right, upper, and lower directions were derived from the pixel coordinates obtained from the participants responses. This data was used to calculate the visual angle of the sound image spread in the horizontal direction, namely HS, as the sum of the absolute angles of the edges in the left and right directions. Similarly, the visual angle of the sound image spread in the vertical direction, namely VS, was calculated as a sum of the absolute angles of the edges in the upper and lower directions. A mean value was used as representative value for each participant.

Two-way repeated-measures analysis of variance (ANOVA) was applied to the mean visual angles of HS and VS as well as the mean angles of the center and the edge in each direction. The within-participant factors were the values of L_f or V_f (0, 0.0428, 0.0730, 0.0916, 0.1180, 0.1414, 0.1738) and the loudspeaker pair (L-R and R-L, U-D and D-U). Post-hoc analyses (Ryan's method) were applied if the main effects of L_f or V_f were identified as significant ($p < .05$) in the ANOVAs.

3.1. Horizontal and Vertical Spreads of Sound Image

3.1.1 Effects of Early Lateral Reflections

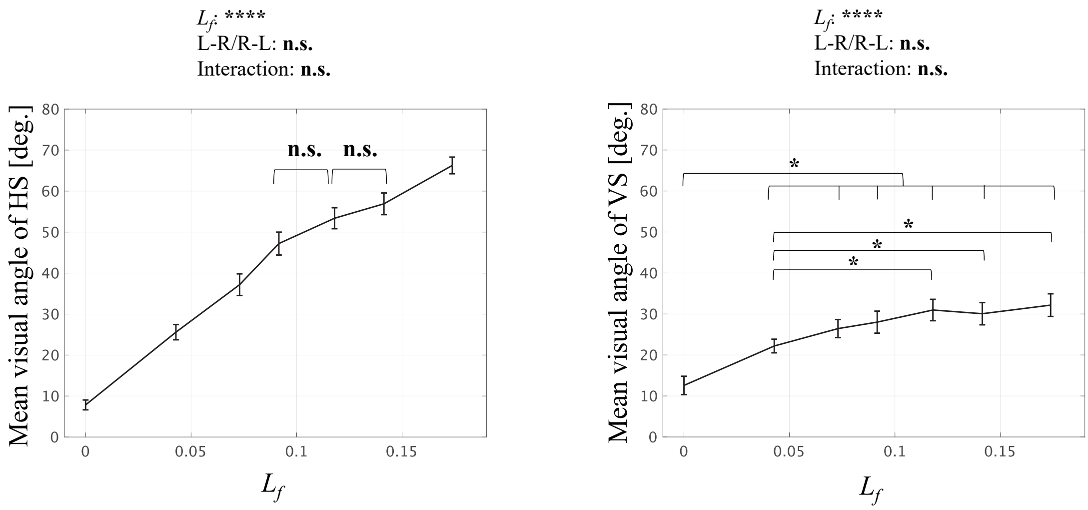
The ANOVA revealed a significant main effect of L_f on HS in L-R and R-L ($F_{6,42} = 81.156$, $p < .001$), but not for the loudspeaker pair ($F_{1,7} = 1.919$, $p = .201$) or the interaction effect ($F_{6,42} = 0.232$, $p = .968$). The ANOVA revealed a significant main effect of L_f on VS in L-R and R-L ($F_{6,42} = 15.938$, $p < .001$), but not for the loudspeaker pair ($F_{1,7} = 0.096$, $p = .766$) or the interaction effect ($F_{6,42} = 0.461$, $p = .833$).

Figures 6(a) and (b) illustrate the mean visual angles of HS and VS when reflected sounds arrive from the lateral directions, in which Ref_1 and Ref_2 arrived from L and R (L-R) and R and L (R-L), respectively. Because the ANOVAs did not indicate a significant main effect of the loudspeaker pair (L-R/R-L) or a significant interaction effect, the results for L-R and R-L are not separately shown. Namely, for each L_f value, the mean visual angles and standard errors of HS and VS were calculated from 16 responses, consisting of 8 for L-R and 8 for R-L. The abscissa represents L_f, ranging between 0 and 0.1738, where 0 denotes the condition with direct sound only. The ordinate represents the mean visual angles [deg.] of HS in Fig. 6(a) and VS in Fig. 6(b). The error bars indicate the standard errors of the means. In Fig. 6(a), all L_f pairs, except for the two indicated by **n.s.** (not significant), differed significantly ($p < .05$). In Fig. 6(b), an asterisk indicates significantly different L_f pairs. Hence, L_f significantly affected both HS and VS. Both HS and VS increased as L_f increased, but the rate of increase was smaller for VS.

3.1.2 Effects of Early Vertical Reflections

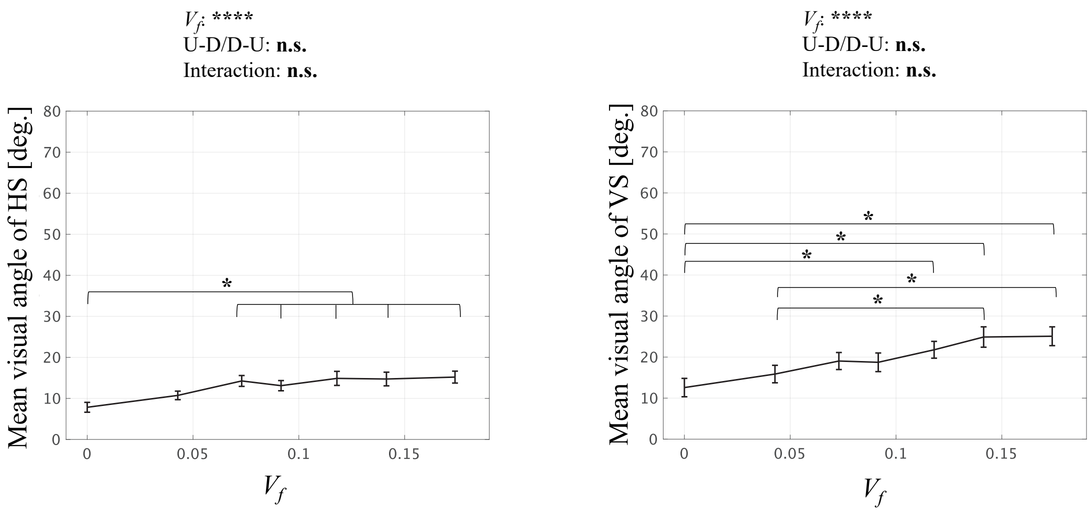
The ANOVA revealed a significant main effect of V_f on HS in U-D and D-U ($F_{6,42} = 6.561$, $p < .001$), but not for the loudspeaker pair ($F_{1,7} = 0.169$, $p = .697$) or the interaction effect ($F_{6,42} = 1.674$, $p = .151$). A significant main effect of V_f was observed on VS in U-D and D-U ($F_{6,42} = 8.305$, $p < .001$), but not for the loudspeaker pair ($F_{1,7} = 2.327$, $p = .171$) or the interaction effect ($F_{6,42} = 0.999$, $p = .439$).

Figures 7 (a) and (b) illustrate the mean visual angles of HS and VS when the reflected sounds arrive from the vertical directions, in which Ref_1 and Ref_2 arrived from U and D (U-D) or D and U (D-U), respectively. Because the ANOVAs did not indicate a significant main effect of the loudspeaker pair (U-D/D-U) or a significant interaction effect, the results in U-D and D-U are not separately shown. Namely, for each V_f value, the mean visual angles and standard errors of HS and VS were calculated from 16 responses, consisting of 8 for U-D and



(a) Effect of L_f on HS (b) Effect of L_f on VS

Fig. 6 Mean visual angles of the horizontal spread (HS) and vertical spread (VS) of a sound image with variations of L_f. ANOVA results (main effects of L_f and the loudspeaker pair (L-R/R-L), and interaction effect) are indicated above each panel (****: $p < .001$, ***: $p <.005$, **: $p < .01$, *: $p < .05$, **n.s.**: not significant). Error bars indicate standard errors. Asterisks denote L_f pairs identified as significant ($p < .05$) by the post-hoc analysis. Note that, in (a), all the L_f pairs except those indicated by **n.s.** (not significant) are significant.



(a) Effect of V_f on HS (b) Effect of V_f on VS

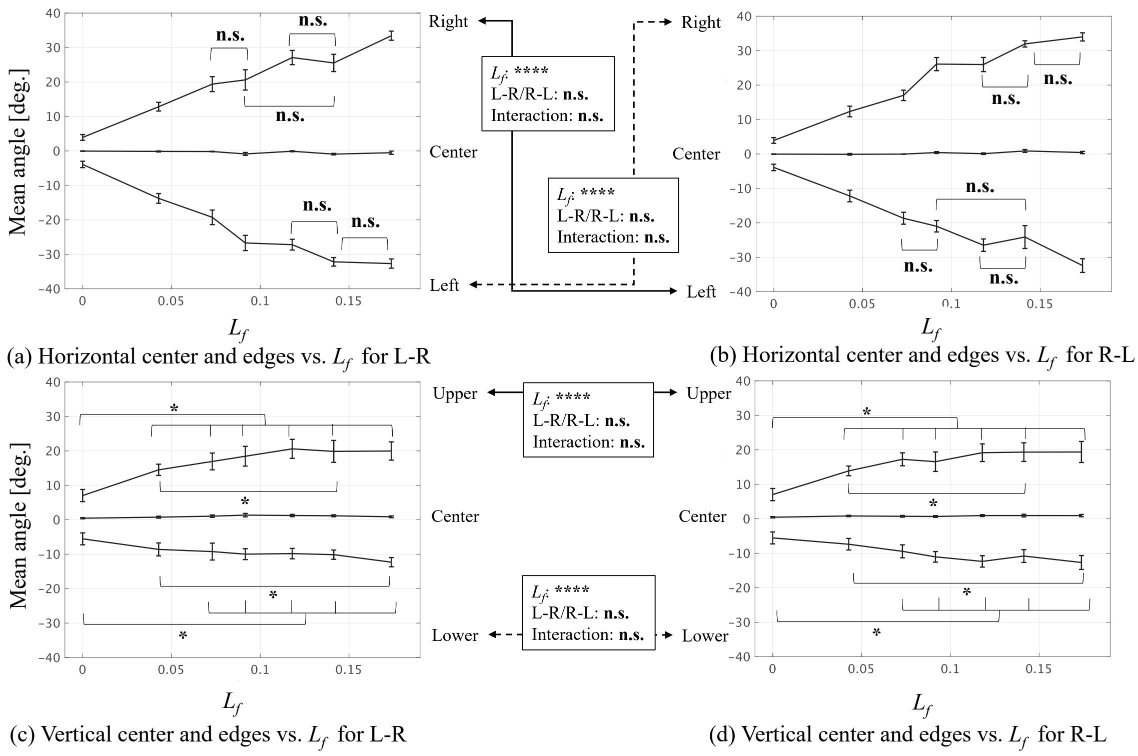
Fig. 7 Mean visual angle of the horizontal spread (HS) and vertical spread (VS) of a sound image with variations of V_f. ANOVA results (main effects of V_f and the loudspeaker pair (U-D/D-U), and interaction effect) are indicated above each panel (****: $p < .001$, ***: $p <.005$, **: $p < .01$, *: $p < .05$, **n.s.**: not significant). Error bars indicate standard errors. Asterisks denote V_f pairs identified as significant ($p < .05$) by the post-hoc analysis.

8 for D-U. The abscissa represents V_f, ranging between 0 and 0.1738. V_f significantly affected HS when the V_f pairs included $V_f = 0$ (Fig. 7(a)). In contrast, V_f significantly affected VS (Fig. 7(b)). As V_f became larger, VS increased.

3.2. Sound Image Spreads in Each Direction

3.2.1. Effects of Early Lateral Reflections

This subsection discusses the effects of L_f and V_f on the sound image spreads in the left, right, upper, and lower directions. The two-way repeated-measures ANOVAs were applied to the two conditions in which the symmetricity in left/right or upper/lower was expected. Namely,



(a) Horizontal center and edges vs. L_f for L-R

(b) Horizontal center and edges vs. L_f for R-L

(c) Vertical center and edges vs. L_f for L-R

(d) Vertical center and edges vs. L_f for R-L

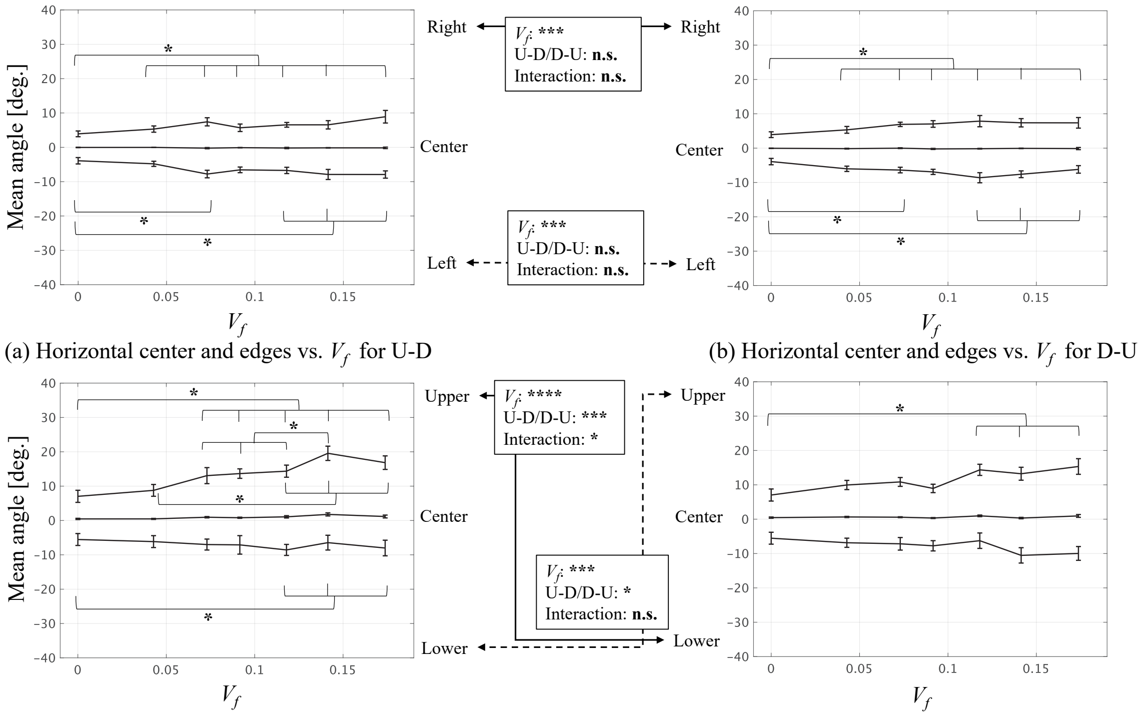
Fig. 8 Mean angles of the center and edges of sound images in the horizontal and vertical directions with variations in L_f for L-R and R-L. ANOVA results (main effects of L_f and the loudspeaker pair (L-R/R-L), and interaction effect) are shown between the panels (****: $p < .001$, ***: $p < .005$, **: $p < .01$, *: $p < .05$, **n.s.**: not significant). Error bars indicate standard errors. Asterisks denote L_f pairs identified as significant ($p < .05$) by the post-hoc analysis. Note that, in (a) and (b), all the L_f pairs except those indicated by **n.s.** (not significant) are significant.

for the effects of L_f on the horizontal spread (left and right edges), the ANOVA was applied to the results consisting of the left edge in the L-R condition and the right edge in the R-L condition and *vice versa*, while the ANOVAs were applied to the results consisting of the upper edges in the L-R and R-L conditions and those consisting of the lower edges in the L-R and R-L conditions. Similarly, for the effects of V_f on the vertical spread (upper and lower edges), the ANOVA was applied to the results consisting of the upper edge in the U-D condition and the lower edge in the D-U condition and *vice versa*. The ANOVAs were also applied to the results consisting of the left edges in the U-D and D-U conditions and those consisting of the right edges in the U-D and D-U conditions.

A significant main effect of L_f on the left edge in L-R and the right edge in R-L ($F_{6,42} = 117.789$, $p < .001$) was observed, but not for the loudspeaker pair ($F_{1,7} = 1.290$, $p = .294$) or the interaction effect ($F_{6,42} = 1.019$, $p = .426$). For the effects of L_f on the right edge in L-R and the left edge in L-R, the ANOVA revealed. A significant main effect of L_f on the right edge in L-R and the left edge in R-L ($F_{6,42} = 44.712$, $p < .001$) was observed, but not for the loudspeaker pair ($F_{1,7} = 0.258$, $p = .627$) or the interaction effect ($F_{6,42} = 0.104$, $p = .996$). A significant main effect of L_f on the upper edges in L-R and R-L ($F_{6,42} = 13.081$, $p < .001$) was observed, but not for the loudspeaker pair ($F_{1,7} = 3.741$, $p = .094$) or the interaction effect ($F_{6,42} = 0.390$, $p = .882$). For the effects of L_f on the lower edges in L-R and R-L, the ANOVA revealed. A significant main effect of L_f was observed on the lower edges in L-R and R-L ($F_{6,42} = 6.003$, $p < .001$), but not for the loudspeaker pair ($F_{1,7} = 1.258$, $p = .299$) or the interaction effect ($F_{6,42} = 2.266$, $p = .055$).

Figures 8 (a) and (b) show the mean angles of the horizontal center and edges with varying L_f in the L-R and R-L conditions. The abscissa represents L_f. The ordinate represents the mean angles [deg.] of the directions of the center and edges of a sound image in the horizontal direction, where the positive (negative) values indicate the R (L) directions. All the L_f pairs, except those indicated by **n.s.** (not significant), differed significantly in the post-hoc analyses ($p < .05$). The mean angles of both the left and right edge directions increased as L_f became larger in both the L-R and R-L conditions. Because the main effect of the loudspeaker pair and the interaction effect were not significant, the horizontal spreads of a sound image for L-R and R-L were symmetric with respect to the median plane.

Figures 8 (c) and (d) show the mean angles of the vertical center and edges with varying L_f in the U-D and D-U conditions. The abscissa



(a) Horizontal center and edges vs. V_f for U-D (b) Horizontal center and edges vs. V_f for D-U

(c) Vertical center and edges vs. V_f for U-D (d) Vertical center and edges vs. V_f for D-U

Fig. 9 Mean angles of the center and edges of the sound images in the horizontal and vertical directions with variations in V_f for U-D and D-U. ANOVA results (main effects of V_f and the loudspeaker pair (U-D/D-U), and interaction effect) are indicated between each panel (****: p < .001, ***: p <.005, **: p < .01, *: p < .05, **n.s.**: not significant). Error bars indicate standard errors. Asterisks denote V_f pairs identified as significant (p < .05) by the post-hoc analysis.

represents L_f. The ordinate represents the mean angles [deg.] of the directions of the center and edges of a sound image in the vertical direction, where the positive (negative) values indicate U (D) directions. An asterisk (*) denotes a significant difference in the L_f pairs ($p < .05$) by the post-hoc analyses. The mean angles of both the upper and lower edge directions increased as L_f became larger in both the U-D and D-U conditions. However, the L_f pairs with significant differences were limited to those that included $L_f = 0$, indicating that L_f had a limited effect on the vertical spread.

3.2.2. Effects of Early Vertical Reflections

The ANOVA revealed a significant main effect of V_f on the left edges in U-D and D-U ($F_{6,42} = 4.669$, $p = .001$), but not for the main effect of the loudspeaker pair ($F_{1,7} = 0.00$, $p = .996$) or the interaction effect ($F_{6,42} = 2.173$, $p = .065$). Similarly, the ANOVA revealed a significant main effect of V_f on the right edges in U-D and D-U ($F_{6,42} = 3.887$, $p = .004$), but not for the main effect of the loudspeaker pair ($F_{1,7} = 0.332$, $p = .582$) or the interaction effect ($F_{6,42} = 1.202$, $p = .324$). The ANOVA revealed significant main effects of V_f on the upper edge in U-D and the lower edge in D-U ($F_{6,42} = 9.985$, $p < .001$) and the loudspeaker pair ($F_{1,7} = 20.276$, $p = .003$), and a significant interaction effect ($F_{6,42} = 2.486$, $p = .038$). Simple main effect tests revealed that V_f significantly affected the upper edge in U-D ($F_{6,84} = 10.945$, $p < .001$) but not the lower edge in D-U ($F_{6,84} = 2.065$, $p = .066$). The ANOVA revealed significant main effects of V_f on the lower edge in U-D and the upper edge in D-U ($F_{6,42} = 4.489$, $p = .001$) and the loudspeaker pair ($F_{1,7} = 9.388$, $p = .018$), and no significant interaction effect ($F_{6,42} = 2.010$, $p = .086$).

Figures 9(a) and (b) show the mean angles of the horizontal center and edges for various V_f in the U-D and D-U conditions, respectively. The abscissa represents V_f. The ordinate represents the mean angles [deg.] of the directions of the center and edges of a sound image in the horizontal direction, where positive (negative) indicates the R (L) direction. Asterisks denote V_f pairs identified as significant ($p < .05$) by the post-hoc analyses. V_f significantly affected the mean angles of the right and left edges, but significant differences were identified only for the V_f pairs that included $V_f = 0$, suggesting that the horizontal edges have a negligible change with $V_f > 0$, although the early vertical reflections affect the horizontal edges.

Figures 9(c) and (d) show the mean angles of the vertical center and edges for various V_f in the U-D and D-U conditions, respectively. The abscissa represents V_f. The ordinate represents the mean angles [deg.] of the directions of the center and edges of the sound image in the vertical direction, where positive (negative) indicates the U (D) direction. Asterisks denote V_f pairs identified as significant ($p < .05$) by the post-hoc analyses. V_f significantly affected both the lower edge in D-U and the upper edge in U-D. However, significant differences were only observed for the V_f pairs that included $V_f = 0$, suggesting that the lower edge in U-D and the upper edge in D-U have a negligible change with $V_f > 0$. Furthermore, the significant main effect of the loudspeaker pair (U-D/D-U) indicate an asymmetric response, *i.e.*, the magnitudes of the effect of V_f on the lower edge in U-D and the upper edge in D-U differed. As for the upper edge in U-D and the lower edge in D-U, the effects of V_f differed between them. The upper edge in U-D was sensitive to V_f, whereas significant differences were not identified for any V_f pair for the lower edge in D-U. Hence, the responses were asymmetric. These results suggest that the vertical spread of a sound image induced by the early vertical reflections is not upside-down reversed when the early vertical reflections are upside-down reversed.

4. Discussion

L_f has a significant effect on HS, increasing with larger L_f (Fig. 6(a)), confirming the results of a previous study [3]. The results also reveal that L_f has a significant effect on VS, increasing with larger L_f, while the rate of increase for VS is smaller than that for HS (Fig. 6(b)). Thus, a larger L_f leads to larger HS and VS where VS is less sensitive to L_f. Furthermore, V_f has a significant effect on VS, increasing with larger V_f (Fig. 7(b)). In contrast, V_f has a significant effect on HS, but the difference is only significant for some of the V_f pairs that include $V_f = 0$. Compared to presenting direct sound only, HS becomes larger when presenting both the early vertical reflections and the direct sound. However, HS is insensitive to V_f. Interestingly, comparisons between VS induced by the early lateral reflections (Fig. 6(b)) and that by the early vertical reflections (Fig. 7(b)) reveal that the early lateral reflections induce a larger VS than the early vertical reflections. That is, the early lateral reflections lead to more than 30 deg. of VS for $L_f > 0.1180$, whereas the early vertical reflections did not for the same values of V_f, at least for the conditions examined in this study. In addition to HS, VS is affected more by the early lateral reflections than by the early vertical reflections when their energies are equal (*i.e.*, $L_f = V_f$).

As demonstrated in Figs. 6(a) and (b), HS is larger than VS, namely, the perceived sound images are horizontally wide, when early lateral reflections are presented. Furthermore, larger L_f, or $1-\mathrm{IACC}_{E3}$, result in both larger HS and VS with maintaining the relation HS > VS. Therefore, it is likely that the early lateral reflections produce a horizontally-long ellipsoidal sound image; the size of ellipsoidal sound image increases, with roughly keeping the ratio of major (horizontal) axis length to minor (vertical) axis length, as L_f, or $1-\mathrm{IACC}_{E3}$ increases, which reasonably explains how the early lateral reflections lead to not only larger HS but also larger VS. On the other hand, when presenting early vertical reflections, the perceived sound images are vertically long, namely VS > HS (Figs. 7(a) and (b)). In this case, VS is roughly proportional to V_f whereas HS does not prominently vary with V_f. It should be noted that the values of $1-\mathrm{IACC}_{E3}$ are almost zero regardless of the values of V_f (Table 1), indicating that $1-\mathrm{IACC}_{E3}$ cannot account for the sound image spreads induced by early vertical reflections. This suggests that the production mechanism of VS induced by the early vertical reflections are different from that by the early lateral reflections.

Regarding the precedence effect when single reflected sound is presented from a vertical direction, Ege *et al.* [14] reported that the time delays of a lagging sound, which result in the precedence effect, summing localization, or split of sound image, are different between when the lagging sound comes from lateral and vertical directions. Their results indicate that the summing localization, not the precedence effect, occurs when presenting a lagging sound from a vertical direction with 30-40 ms after a leading sound, which corresponds to the stimulus condition employed in the current study. Considering their results, the current results indicate that the early vertical reflections lead to the summing localization and vertically long sound images.

In addition to Ege *et al.* [14], Takada and Otani [15] also reported an experimental result on the sound images when single reflected sound is presented from a vertical direction while the time delay was limited between 0.5 and 10 ms. Their results showed that most of the participants perceived single sound images as a result of summing localization, which is consistent with Ege *et al.* [14]; however, some participants reported a split of sound image, instead of summing localization, in some trials. Similarly, also in the current experiment, it is likely that the participant excluded from the analysis in this study perceived a split of sound image, instead of summing localization, although it is not clear why the participant perceived a split of sound image while others did not. Exploring this issue would necessitate another experiment that clarifies a sound image shape, not only sound image spreads, *e.g.*, by asking the participants to draw the shape of sound images.

The analysis of the spread in each direction shown in Figs. 8 and 9 confirm the results from Figs. 6 and 7. For the early lateral reflections with a given L_f, the mean angle of the right edge of a sound image for L-R is almost equal to that of the left edge for R-L, and *vice versa* (Figs. 8(a) and (b)). Namely, the HSs of sound images for L-R and R-L are in a left-right reversed manner. In contrast, for a given V_f, the mean angles of the upper edge for U-D and D-U are not equal to those of the lower edge of a sound image for D-U and U-D, respectively

(Figs. 9(c) and (d)). This demonstrates that the VSs of sound images for U-D and D-U do not have an upside-down reversed relation because the sound image spread toward the lower direction is generally smaller than that in the upper direction for D-U where ΔL_1 (Ref$_1$: D) is equal to or greater than ΔL_2 (Ref$_2$: U). Namely, VS is asymmetric while HS is symmetric. The reason for such asymmetricity of VS between U-D and D-U can be explained by an upward bias of sound image spreads. The sound image spreads are biased upward when early vertical reflections are presented, as shown in Figs. 9(c) and (d). Furthermore, even when presenting the early lateral reflections, the sound image spreads are also biased upward (Figs. 8(c) and (d)). Namely, the sound image spreads are biased upward in all the stimulus conditions examined, regardless of the direction from which the early reflections arrive. Such an upward bias of the sound image spreads can account for the asymmetricity of VS between U-D and D-U observed in this study. However, it is yet to be revealed why the sound image spreads are biased upward, which would require further investigation in a future work.

V_f has a significant effect on HS when the V_f pairs include $V_f = 0$ (Fig. 7(a)), but not when V_f pairs include $V_f > 0$. However, it is possible that the early vertical reflections with a much larger energy, which were not examined in this study, may lead to a larger HS.

These results indicate that early vertical reflections produce smaller effects on the spatial spread of a sound image, compared to the early lateral directions. This implies that the conventional acoustical design policy for concert halls, which emphasizes early lateral reflections, is reasonable. However, the early vertical reflections also significantly affect both HS and VS although the effect is smaller than that of the early lateral reflections. In fact, Furuya *et al.* reported that unpreferable auditory impressions are observed for under-balcony seats in a concert hall that lack early upper reflections like "listening through a window" [16),17)]. The effect of the early vertical reflections on spatial spread of a sound image may explain this degraded impression from a fundamental psychoacoustical viewpoint.

One limitation on the current study is that the loudspeakers for reflected sounds were located only ±30° both in azimuth and elevation. This was because even a single set of loudspeaker arrangement required lots of trials, thereby making it difficult to conduct an experiment with other loudspeaker arrangements. In addition, Morimoto *et al.* [18)] suggested that the ASWs, *i.e.*, the HSs in the current study, do not vary with the number and the direction of arrival of reflected sounds if IACC is constant. Therefore, other sets of loudspeaker arrangement were not examined in this study. From the same reasons, variations in temporal structure of reflected sounds, *i.e.*, other than a pair of 30 ms and 40 ms, and their intensities, such as ones where ΔL_2 is larger than ΔL_1, were not investigated in this study. However, as discussed above, the sound image spreads induced by the early vertical reflections cannot be accounted for by the values of IACC, and then the underlying perceptual mechanisms of sound image spread are likely different between the early lateral and vertical reflections, indicating that Morimoto *et al.*'s suggestion [18)] does not hold for the VS for the early vertical reflections. Therefore, clarifying the details of effects of reflected sounds, especially from vertical directions, on the HS and VS would necessitate further experiments considering the variations in the direction of arrival, temporal structure, and intensity of reflected sounds, which should be addressed in future studies. Furthermore, in real acoustic environments, both early lateral and vertical reflections arrive to a listener. The current study presented early lateral or vertical reflections separately. Therefore, future studies should include experiments that simultaneously present both reflections to further elucidate their effects on HS and VS.

5. Conclusions

This study performed a psychoacoustical experiment to explore the effects of early lateral and vertical reflections on the horizontal (HS) and vertical spreads (VS) of a sound image. In the experiment, the energy ratio of the early lateral or vertical reflections to the direct sound was controlled by two physical parameters L_f and V_f.

L_f significantly affects both HS and VS. The larger L_f, the larger the values of HS and VS. However, the rate of increase is greater in HS than in VS. Furthermore, V_f also significantly affects both VS and HS, but the effect of V_f on VS is smaller than that of L_f on VS. A larger V_f does not prominently increase HS. On the other hand, early vertical reflections result in a larger VS compared to when only a direct sound is presented.

Further analysis of the sound image spread in each direction indicates that HSs are perceived in a left-right reverse manner, or as mirrored sound images, when the early lateral reflections are left-right reversed. However, when the early vertical reflections are upside-down reversed, VSs are not perceived in an upside-down reversed manner due to the smaller spread of the sound image in the lower direction than in the upper direction even when the lower reflection has a greater sound pressure level than the upper reflection. Such asymmetricity of VS is explainable by the upward-bias of the sound image spreads.

L_f correlates with ASW [3)] corresponding to HS in the current study. In addition, the current results provide new knowledge on HS and VS induced by early lateral and vertical reflections and their variations with L_f and V_f. Early vertical reflections contribute to the production of both HS and VS, and this influence should be considered in concert-hall acoustics.

Acknowledgments

This work was partly supported by Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (Nos. JP19H04153 and JP19H04145).

Reference

1) Morimoto, M., Fujimori, H. and Maekawa, Z.: Discrimination between auditory source width and envelopment, Journal of Acoustical Society of Japan, Vol.46, pp.449-457, 1990.
(DOI: https://doi.org/10.20697/jasj.46.6_449)
2) Bradley, J.S. and Soulodre, G.A.: Objective measures of listener envelopment, Journal of Acoustical Society of America, Vol.98, pp.2590-2597, 1995.
(DOI: https://doi.org/10.1121/1.413225)
3) Barron, M. and Marshall, A.H.: Spatial impression due to early lateral reflections in concert halls – The derivation of a physical measure, Journal of Sound and Vibration, Vol.77, pp.211-232, 1981.
(DOI: https://doi.org/10.1016/S0022-460X(81)80020-X)
4) Hidaka, T., Beranek, L.L. and Okano, T.: Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls, Journal of Acoustical Society of America, Vol.98, pp.988-1007, 1995.
(DOI: https://doi.org/10.1121/1.414451)
5) Okano, T., Beranek, L.L. and Hidaka, T.: Relations among interaural cross correlation coefficient ($IACC_E$), lateral fraction (LF_E), and apparent source width (ASW) in concert halls, Journal of Acoustical Society of America, Vol.104, pp.255-265, 1998.
(DOI: https://doi.org/10.1121/1.423955)
6) Barron, M.: The subjective effects of first reflections in concert halls – The need for lateral reflections, Journal of Sound and Vibration, Vol.15, pp.475-494, 1971.
(DOI: https://doi.org/10.1016/0022-460X(71)90406-8)
7) Bradley, J.S. and Soulodre, G.A.: The influence of late arriving energy on spatial impression, Journal of Acoustical Society of America, Vol.97, pp.2263, 1995.
(DOI: https://doi.org/10.1121/1.411951)
8) Otani, M., Yamazaki, K., Toyoda, M., Hashimoto, M., and Kayama, M.: Relation between frequency bandwidth of broadband noise and largeness of sound image, Acoustical Science and Technology, Vol.38, pp.35-37, 2017.
(DOI: https://doi.org/10.1250/ast.38.35)
9) Otani, M., Yamazaki, K., Toyoda, M, Hashimoto, M., and Kayama, M.: Largeness and shape of sound images captured by sketch-drawing experiments: Effects of bandwidth and center frequency of broadband noise, Acoustical Science and Technology, Vol.38, pp.157-160, 2017.
(DOI: https://doi.org/10.1250/ast.38.154)
10) Furuya, H., Fujimoto, K., Takeshima, Y., and Nakamura, H.: Effect of early reflections from upside on auditory envelopment, Journal of Acoustical Society of Japan (E) Vol.16, pp.97-104, 1995.
(DOI: https://doi.org/10.1250/ast.16.97)
11) Hanyu, T., Hoshi, K., and Sato, R.: Spatial impression from late overhead reflections in concert halls, Journal of Acoustical Society of Japan, Vol.69, pp.7-15, 2013.
(DOI: https://doi.org/10.20697/jasj.69.1_7)
12) Farrar, K.: Soundfield microphone, Wireless World, Vol.85, pp.48-51, 1979.
13) Otani, M. and Ise, S.: Fast calculation system specialized for head-related transfer function based on boundary element method, Journal of Acoustical Society of America, Vol.119, pp.2589-2598, 2006.
(DOI: https://doi.org.10.1121/1.2191608)
14) Ege, R., van Opstal, A.J., Bremen, P., and van Wanrooij, M.M.: Testing the precedence effect in the median plane reveals backward spatial masking of sound, Scientific Reports, Vol.8, 8670, 2018.
(DOI: https://doi.org/10.1038/s41598-018-26834-2)
15) Takada, Y. and Otani, M.: Precedence effect and related phenomena in the median plane for short time delay, Acoustical Science and Technology, Vol.43, No.1, pp.50-56, 2022.
(DOI: https://doi.org/10.1250/ast.43.50)
16) Furuya, H., Fujimoto, K., Takeshima, Y., and Nakamura, H.: Acoustical characteristics of early reflections and subjective impression of envelopment under the balcony in auditoria, Journal of Architecture, Planning, and Environmental Engineering, (Transactions of AIJ), Vol.486, pp.1-7, 1996.
(DOI: https://doi.org/10.3130/aija.61.1_8)
17) Furuya, H., Fujimoto, K., Toyomura, A., Ohnishi, K., and Nakamura, H.: Relationship between the balcony configuration and acoustical characteristics of

early reflections in auditoria, Journal of Architectural, Planning, and Environmental Engineering (Transactions of AIJ), Vol.498, pp.15-21, 1997.
(DOI: https://doi.org/10.3130/aija.62.15_6)

18) Morimoto, M., Iida, K., and Furue, Y.: Relation between auditory source width in various sound fields and degree of interaural cross-correlation, Applied Acoustics, Vol.38, pp.291-301, 1993.
(DOI: https://doi.org/10.1016/0003-682X(93)90057-D)

（2022 年 10 月 4 日原稿受理，2023 年 1 月 6 日採用決定）