PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

The role of median plane reflections in the perception of vertical auditory movement

Florian WENDT, Matthias FRANK, Robert HÖLDRICH

Institute of Electronic Music and Acoustics
University of Music and Performing Arts Graz, Austria
Email: {wendt, frank, hoeldrich}@iem.at

Abstract
This contribution considers how delay alterations of a median plane reflection influence the vertical auditory movement. Depending on the excitation signal we could show two partly contrasting effects dominating our perception: For broadband noise signals a delay modification between the leading direct sound and the lagging reflection yields spectral cues and the auditory movement is explained by the pitch-height effect. In contrast, results obtained for speech signals indicate that other localization cues induced by the interference pattern of direct sound and reflection influence our perception. Respective movement directions are deduced from the movement of a physical sound source, suggesting that the auditory system learns to associate interference patterns to sound source movements.
Keywords: localization, movement perception, median sagittal plane, floor reflection;

1 INTRODUCTION

Changes in level, time of arrival, and spectrum of the ear signals are cues for the perception of sound movement [1]. The relative importance of these time-variant cues is different for the perception of vertical and horizontal movement. In the horizontal plane, the use of inter-aural level and time differences (ILDs and ITDs) as localization cues is proven. In the median plane, binaural cues are essentially zero and spectral properties play an important role in contrast to inter-aural differences.

Sound rendering systems typically use level panning techniques to move virtual sound sources along the horizontal plane. At low frequencies these level differences yield inter-aural time differences resembling those of a real source [2]. Regarding vertical localization, studies could show a similar relation between inter-channel level differences and perceived elevation [3, 4]. Thus, variations in amplitude between vertical loudspeakers are used in the same manner to create respective auditory movements, e.g. in Vector-Base Amplitude Panning (VBAP) [3] or Ambisonics [5]. Studies on the use of time differences on the other hand did not show stable localization curves [6] and they are mostly disregarded in common panning techniques.

The direct sound of an emitting sound source is typically followed by attenuated reflections. Lateral reflections contribute mainly to qualitative attributes of the perceived sound [7] without affecting the localization [8]. In contrast, studies reported vertical shifts in localization for conditions with an additional attenuated sound instance in the median plane [9, 10, 11].

This contribution studies how median plane reflections can influence the perception of vertical movement. Two listening experiments are conducted with an experimental setup simulating direct sound and median plane reflections. Experiment 1 considers the influence of continuous delay alterations of reflections from floor and ceiling. Based on the results two signal-dependent perceptual effects are deduced. The movement perception of noise signals can be explained by spectral cues eliciting the pitch-height effect. For speech signals there are indications that binaural cues, caused by the fusion of direct sound and reflection, influence the movement perception. Experiment 2 examines this latter effect in greater detail. The impact of loudspeaker setup and signal character is studied and results are compared to vertical amplitude panning.

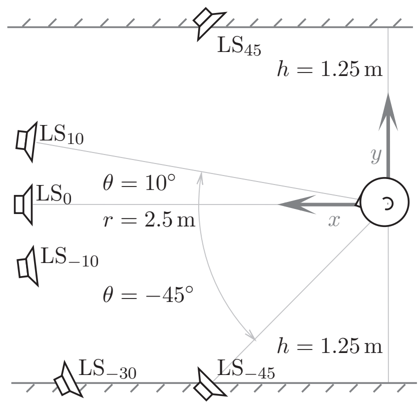


Figure 1. Sketch of the experimental setup in the anechoic laboratory. Loudspeakers $\mathrm{LS}_{(10,0,-10,-30)}$ are used to play back the direct sound and loudspeakers $\mathrm{LS}_{(45,-45)}$ simulate an attenuated specular reflection.

2 EXPERIMENTAL SETUP

All experiments took place in the anechoic laboratory at our institute. The layout consisted of six vertically arranged Genelec 8020A loudspeakers in the median plane at elevation angles $\theta = (45^\circ, 10^\circ, 0^\circ, -10^\circ, -30^\circ, -45^\circ)$ denoted as LS_θ, cf. Figure 1. Loudspeakers $\mathrm{LS}_{(10,0,-10,-30)}$ were placed at a distance of $r = 2.5\,\mathrm{m}$ and loudspeakers $\mathrm{LS}_{(45,-45)}$ at $1.8\,\mathrm{m}$ to the listening position and each was level- and delay-compensated. The listener's ears were adjusted to $h = 1.25\,\mathrm{m}$ above floor loudspeakers $\mathrm{LS}_{(-30,-45)}$. During the experiments, listeners were requested to face the 0° direction and to minimize body movement, but small head movements were tolerated. In this way binaural cues were mostly lacking, but not completely absent.

Loudspeakers $\mathrm{LS}_{(10,0,-10)}$ were used to supply the leading direct sound at angle θ_{dir} and loudspeakers $\mathrm{LS}_{(45,-30,-45)}$ were used for simulating a lagging specular reflection from ceiling or floor at θ_{refl}. The level of reflections was attenuated by $\Delta L = 4.4\,\mathrm{dB}$ compared to the direct sound, simulating a θ_{dir}-dependent absorption coefficient. Hence, the presented studies focus on the influence of temporal properties of the reflection.

3 EXPERIMENT 1: Delay-based perceptual phenomena for vertical sound movement

The spectral coloration evaluated by the auditory system for vertical sound localization is mainly caused by the directional filtering of pinna, head, and torso [12].

Consider a source and a receiver at a certain distance and at equal height above the floor. The vertical movement of the source along the median plane changes delay and level of the floor reflection. Studies have shown an influence of the reflection pattern on timbre and thus on the spectrum of the perceived sound [13] providing additional information on the height of the source.

The kind of localization cues that are introduced by median plane reflections is examined by means of a listening experiment. Experiment 1 studies how the time delay of a median plane reflection can influence the perceived height of a sound source.

3.1 Listening test procedure

Starting from the delay T of the lagging sound simulating the specular reflection, the listeners task was to adjust both an upper and lower limit of the delay, denoted as $T_+ = T + \Delta t_+$ and $T_- = T - \Delta t_-$, yielding the most shifted virtual sound source evoked by a continuous vertical movement.

The upper and lower limits of the delay $\Delta t_\pm$ were determined separately, using the same task. Using a fader, the delay initiated with $\Delta t_\pm = 0\,\mathrm{ms}$ was increased as long as a monotone vertical source movement was perceived.

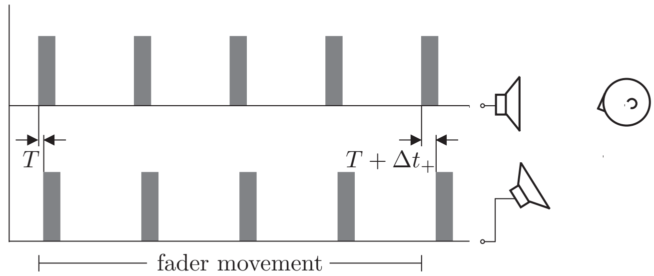
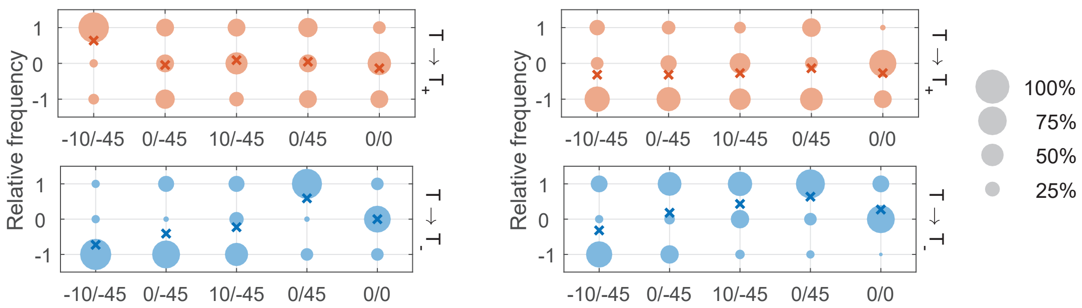


Figure 2. Exemplary illustration of the sequence repetition for a condition with increasing delay T_+ and noise bursts. At the beginning of direct sound and floor reflection are separated by the specular delay T. By moving the fader, the listener increases the delay of the lagging reflection until reaches $T_+ = T + \Delta t_+$ after 4 noise bursts.

Once the listener determined the fader position yielding the most vertically deflected source, they responded as to the direction in which the sound source was perceived to move, i.e. *up* or *down*, and the adjusted delay $\Delta t_\pm$ was stored by pressing a button. The maximum allowable delay alteration was fixed with $\Delta t_\pm^{\max} = 2\,\text{ms}$, which was determined by informal listening of the authors. If no or only unstable movements were perceived, listeners could also answer with *no*. After logging in the answer, the next sample was loaded and the motorized fader jumped back to the zero position $\Delta t_\pm = 0\,\text{ms}$ and the next condition started.
Three conditions $-10/-45$, $0/-45$, and $10/-45$ were tested consisting of direct sound and reflection played back at $\theta_{\text{dir}}/\theta_{\text{refl}}$ with corresponding specular delays $T = (2.1, 3.0, 3.9)\,\text{ms}$. Additionally, a control condition with $\theta_{\text{dir}} = \theta_{\text{refl}} = 0°$ and $T = 3\,\text{ms}$, denoted as 0/0, was included in the test in order to evaluate the influence of pure comb filtering. Tested signals were chosen to investigate the influence of spectro-temporal properties (envelope, bandwidth, etc.) and familiarity to the effect: *female* speech and two 100 ms long pink noise bursts of which *noise1* had onset and release times of 10 ms and *noise2* of 2 ms and 98 ms, respectively. The noise samples were followed by a pause time of 60 ms, whereas for the 7 s long speech sample no pause was added. For each condition, the samples were played back in a loop at 70 dB(A) until the listeners logged in their answers. Figure 2 shows an exemplary illustration of the sequence repetition for a condition with increasing delay $T \to T + \Delta t_+$.
The test sequence for every listener was an individual random permutation of the entire set yielding 3 sounds $\times$ 4 conditions $\times$ 2 repetitions $\times$ 2 delay alterations $= 48$ adjusted delays and 48 movement directions per listener. Twelve experienced listeners participated in the experiment (all male; age 24 to 54) and all of them reported normal hearing acuity. Before conducting the experiment listeners familiarize with the signals in a training. Therefore, single loudspeakers were used for playback and listeners were aware which loudspeaker is active.

3.2 Results

The listeners' consistency is obtained by comparing their answers concerning the direction of perceived movement of both repetitions. Overall the consistency is relatively poor and after excluding one listener (consistency 40%), it accounts 56% on average (min. 50%, max. 67%).
A pairwise statistical analysis of *noise1* and *noise2* reveals no significant difference of the signals for any condition ($p > 0.10$). Therefore further analysis is done by pooling the data of both noise signals.
Figure 3 shows histograms of obtained answers concerning the direction of perceived movement. Since the distance between each item category (*up-no, no-down*) on the ordinal scale is equivalent, mean values are obtained by assigning a value of $+1$ for *up* responses, -1 for *down* responses, and a 0 for *no* responses. The standardized difference between computed mean values of the incrementing/decrementing (inc./dec.) delay in each condition corresponds to the effect size expressed as Cohen's d. Respective effect sizes for conditions of Experiment 1 are listed in Table 1.



(a) Speech (b) Noise

Figure 3. Histogram of obtained answers concerning the direction of perceived movement when increasing (red) and decreasing (blue) the delay T for speech (left) and noise signals (right). The value $+1$ corresponds to *up* responses, -1 to *down* responses, and 0 to *no* responses. Corresponding means are indicated as $\times$.

Table 1. Effect size expressed as Cohen's d for Experiment 1. Values are color coded according to the classification of [14] with $|d| = 1.2$ defining a very large, $|d| = 0.8$ a large, and $|d| = 0.5$ a medium effect.

cond.	$-10/-45$	$0/-45$	$10/-45$	$0/45$	$0/0$
T	2.1 ms	3.0 ms	3.9 ms	3.0 ms	3.0 ms
female	1.88	0.43	0.42	-0.65	-0.19
noise	0.00	-0.68	-1.00	-0.84	-0.93

A Mann-Whitney U test reveals the *signal type* to be a (weakly) significant factor for most conditions ($p \leq 0.08$), except for incrementing conditions 0/45, $0/-45$ and decrementing condition 0/0. The factor *alteration* of the delay (incrementing vs decrementing) is highly significant for conditions $-10/-45$ and 0/45 ($p < 0.02$) for speech, while for noise it is highly significant for all other conditions ($p \leq 0.01$). This agrees with the effect size d listed in Table 1. Interestingly, the effect direction of conditions with a floor reflection ($\theta_{\text{refl}} < 0°$) is opposite for the two signal types and we deduce two different perceptual phenomena. First for noise, increasing the delay of the reflection yields downwards movements, and vice versa. This result, which is obtained similarly for the control condition 0/0, is attributed to the *pitch-height effect* [1, 15], as incrementing (decrementing) the delay results in a time-variant comb filtering. Several listeners mentioned that they heard comb filtering for broadband noise signals, perceived as an downward (upward) glissando. Based on respective effect sizes in Table 1, the lower limit of the effect is reached with condition $0/-45$ and $T = 3.0$ ms. An upper limit was not reached with tested delays.

For the speech signal on the other hand the effect direction depends on the direction of the reflection. Increasing the delay of floor reflections yields upwards movements, and vice versa. For a reflection from the ceiling an inverse relation is found. However, this effect is not active for the control condition. A statistical analysis of test conditions for both delay alterations separately with speech using the Friedman test reveals the *condition* to be significant when the delay is increased ($p_{\text{inc}} < 0.01$) and weakly significant when it is decreased ($p_{\text{dec}} = 0.08$). From this analysis it is not clear whether the parameter T or θ_{dir} is influencing the effect size as both parameters are varied across conditions. Though answers for decrementing speech conditions ($0/-45$, $10/-45$, 0/45) in Figure 3 are more consistent than respective incrementing conditions, which is why we assume the delay T to be the effect parameter.

Table 2 lists adjusted delays as mean values and standard deviations for both alterations. Delays are highly subjective yielding high standard deviations. Nevertheless, mean values of conditions are similar with overall means for incrementing and decrementing conditions of $\Delta\bar{t}_{\pm} = 1\,\mathrm{ms}$. Such a delay alteration of a physical floor reflection is achieved by shifting the height of an emitting source at $\theta = 0°$ and $r = 2.5\,\mathrm{m}$ by approximately $\Delta\theta_{\pm} = 10°$.

Table 2. Mean values and standard deviations of adjusted delays $\Delta t_{\pm}$ for speech and noise.

condition	Δt_{+} in ms		Δt_{-} in ms	
	female	*noise*	*female*	*noise*
$-10/-45$	1.2 ± 0.6	1.0 ± 0.5	1.0 ± 0.5	0.8 ± 0.4
$0/-45$	1.1 ± 0.5	1.2 ± 0.6	1.2 ± 0.5	1.1 ± 0.5
$10/-45$	1.1 ± 0.6	0.8 ± 0.5	1.1 ± 0.6	1.1 ± 0.4
$0/45$	0.9 ± 0.6	1.1 ± 0.7	0.6 ± 0.4	0.8 ± 0.6
$0/0$	1.2 ± 0.7	1.0 ± 0.7	1.5 ± 0.6	1.2 ± 0.5

4 EXPERIMENT 2: Temporal cues induced by a floor reflection

In Experiment 1 two different effects were shown to influence the movement perception, depending on the excitation signal. The pitch-height effect is well-known in localization studies. Results obtained by speech indicate the existence of other localization cues induced by a median plane reflection. To the authors' knowledge, such an effect was not studied in a listening experiment of the literature so far and the few studies investigating the influence of temporal cues on vertical localization focus on the precedence effect, e.g. [17, 18].

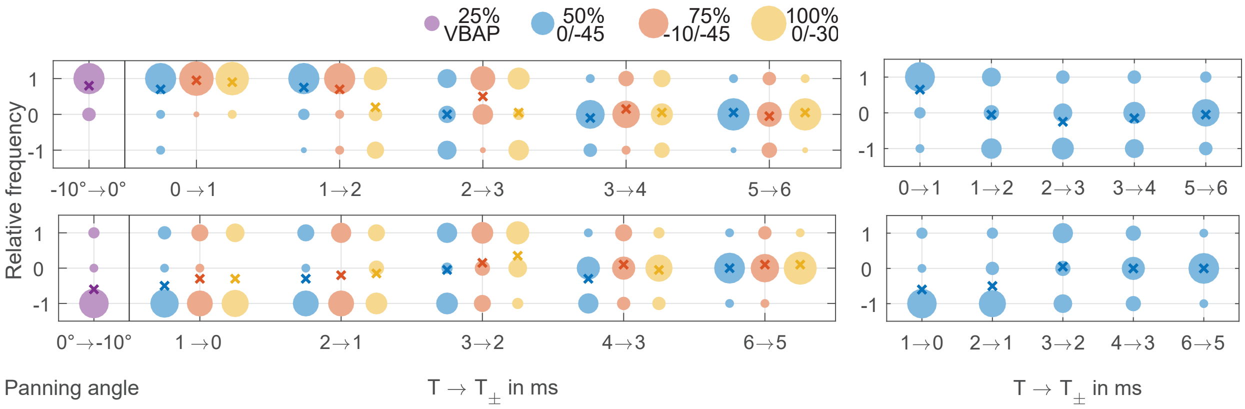
Experiment 2 studies the effect obtained for speech more closely. Possible influences of the pitch-height effect are minimized by restricting the study on floor reflections, for which effect directions of obtained effects were shown to be contrasting. The listening experiment was carried out to assess the influence of the overall delay T and loudspeaker angles $\theta_{\mathrm{dir/refl}}$.

4.1 Listening test procedure

In Experiment 2 listeners were asked to indicate perceived source movements for continuous alterations of the delay from T to T_{+} and T to T_{-}. Based on the results from Experiment 1 delay alterations of $\Delta t_{\pm} = 1\,\mathrm{ms}$ were constant for all conditions and covered a range of $T = 0...6\,\mathrm{ms}$. To keep the testing time decent, a gap between $4\,\mathrm{ms}$ and $5\,\mathrm{ms}$ was inserted yielding five delay increments $T \rightarrow T_{+} = (0 \rightarrow 1, 1 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 4, 5 \rightarrow 6)\,\mathrm{ms}$ and corresponding delay decrements $T \rightarrow T_{-} = (1 \rightarrow 0, 2 \rightarrow 1, 3 \rightarrow 2, 4 \rightarrow 3, 6 \rightarrow 5)\,\mathrm{ms}$. To study directional dependence, two different angles of the direct sound were used in combination with two angles of the reflection, yielding conditions $0/-45, -10/-45$, and $0/-30$. Tested sounds were 3 s long samples of *male* speech and *congas*. These non-stationary signals were chosen to prevent the pitch-height effect.

In the experiment, delays of the reflections were automatically increased or decreased between 0.5 s to 2.5 s of the sample lengths and listeners had to rate the perceived movement of the sound source with *up*, *down*, or *no* continuous movement. The speech sample was tested with all conditions, whereas the conga sample was tested with condition $0/-45$ only. Additional control conditions included dynamically panned male speech signals using VBAP between LS_0 and LS_{-10} and vice versa. Just like the delay alterations, the panning started at 0.5 s and ended at 2.5 s of the sample length.

The test sequence consisted of [(3 speech conditions + 1 conga condition) $\times$ 5 delays + 1 VBAP] $\times$ 2 repetitions $\times$ 2 alterations = 84 samples per listener. Ten experienced listeners with normal hearing acuity participated in the experiment (all male, age 24 to 54), of which eight already participated in Experiment 1. The training conducted before evaluation was similar than for Experiment 1 and included additional VBAP conditions.



(a) Speech (b) Conga

Figure 4. Histograms of obtained answers concerning the direction of perceived movement when increasing and decreasing T by 1 ms for tested signals *speech* (left) and *conga* (right). Delay conditions are color coded (blue, red, yellow). Additionally for *speech*, auditory movement is studied using VBAP (violet).

Table 3. Color-coded effect size expressed as Cohen's d for Experiment 2: very large, large, and medium.

signal	cond.	$0 \rightleftharpoons 1$ ms	$1 \rightleftharpoons 2$ ms	$2 \rightleftharpoons 3$ ms	$3 \rightleftharpoons 4$ ms	$5 \rightleftharpoons 6$ ms
speech	$0/-45$	1.45	1.14	0.05	0.30	0.11
speech	$-10/-45$	1.35	0.90	0.40	0.07	-0.27
speech	$0/-30$	1.23	0.40	-0.40	0.17	-0.16
conga	$0/-45$	1.66	0.59	-0.34	-0.21	-0.11

4.2 Results

The listeners consistency is monitored by comparing answers of both repetition and the mean consistency of 58% (min. 50%, max. 65%) is just above the value found for Experiment 1. The data of all listeners is shown in Figure 4 as histograms and mean values. A statistical analysis of perceived movement direction for both speech and conga with condition $0/-45$ reveals the *signal type* not to be significant for any of the delay alterations (Mann-Whitney U: $p \geq 0.72$). For all conditions, movement directions resemble those obtained by speech in Experiment 1 and increasing the delay tends to be perceived as upwards movement, while a decrease is perceived as a downward-moving source.

While distinct movements are achieved at short delays, the effect reduces with increasing the delay T. The Friedman test reveals the *delay* to be a weakly significant factor for both incrementing and decrementing delays with speech ($p_{\text{male}} \leq 0.09$) and strongly significant with the conga signal ($p_{\text{conga}} \ll 0.01$). For the latter the factor *alteration* is highly significant only for the shortest delay ($0 \rightleftharpoons 1$ ms, $p \ll 0.01$). The same statistical significance is obtained for speech with condition $0/-30$, whereas for conditions $0/-45$ and $-10/-45$ the significant range extends to the two shortest delays ($0 \rightleftharpoons 1$ ms, $1 \rightleftharpoons 2$ ms; $p \ll 0.01$). Based on the effect size of speech condition in Table 3, the effect seems to be dependent on the aperture angle between direct sound and reflection, with wider angles yielding a stronger effect.

To compare the obtained effect with the VBAP condition ($\theta_{\text{VBAP}} = -10° \rightleftharpoons 0°$), speech conditions $-10/-45$ and $0/-45$ with shortest delays are considered ($-10/-45$: $0 \rightarrow 1$ ms; $0/-45$: $1 \rightarrow 0$ ms). The difference of respective means is similar to the difference obtained by VBAP, cf. Figure 4. Consecutively, the effect size expressed as Cohen's d, describing the standardized difference between means, reveals similar values, i.e. $d_{\text{delay}} = 1.75$, $d_{\text{VBAP}} = 1.86$, and according to the classification of [14] a *very large* effect is obtained with both methods.

5 POSSIBLE EXPLANATION OF EFFECTS

We have shown that by a continuous alteration of the delay between a vertically arranged lead and attenuated lag in the range of ± 1 ms the auditory movement can be influenced. We believe, that depending on the excitation signal, different cues are evaluated by the auditory system.

5.1 Spectral cues

For stationary broadband noise, incrementing the delay of a lagging sound instance yields the perception of downward movements. This can be explained by the pitch-height effect as time-variant comb filtering is perceived as downward glissando. Thus, this effect is independent of the direction of lagging reflections and it can be concluded that reflections from the ceiling support the localization of a moving physical sound source, whereas floor reflections yield localization cues that do not comply with the physical movement.

5.2 Binaural cues

For time-variant signals with multiple narrowband components, e.g. speech, a different relation is obtained. A continuous increase of the delay T between the direct sound and the floor reflection yields an upward movement in the perception, whereas for the ceiling reflection the perceived movement is opposed. Thus, the perception is in compliance with the movement of a physical sound source.

A possible explanation of this effect can be found in [19], which studies the impact of early reflections on binaural cues by means of numerical models. They could show that an early reflection produces a discontinuity in the inter-aural phase difference at the interference frequency f, which can be arbitrarily large and yields an ITD that can take any value between $\pm 1/(2f)$. Based on this insight, the authors in [19] conclude that floor reflections with short delays provide reliable cues for elevation estimation by changing ITDs. It is possible that they are learned in the same way as pinna cues and small angles of the head are sufficient to give a directional sense. Although the delay of floor reflections mostly exceed the active range found in Experiment 2 ($T \leq 2$ ms), short delays occur if we consider reflections from the torso or other obstacles, or far away sources (> 10 m).

6 CONCLUSIONS

A method is presented to control the perceived vertical movement of a virtual sound source with loudspeakers resembling a listening situation consisting of direct sound (lead) and a median plane reflection (lag). Two listening experiments were conducted and it is shown that a simple delay modification in the range of ± 1 ms between leading and lagging sound is sufficient to evoke an auditory movement along the median plane.

Results of Experiment 1 show a signal dependency of the movement direction and we assume two partly contrasting perceptual effects dominating the perception of auditory movement. For noise signals, continuous delay alterations yield time-variant comb filtering and movement perception is governed by the pitch-height effect. This spectral effect, obtained when the overall delay exceeds 3 ms, appears independently of directions of incoming sounds. For speech signals the auditory movement is deduced from the movement of a physical sound source and the direction of the lag specifies the direction of movement.

In Experiment 2 the latter effect is studied more closely and compared to amplitude panning. We could prove it for continuous delay alterations with speech and conga samples. It was found to be strongest for short delays (overall delay $T \leq 2$ ms) with a wide loudspeaker spacing of lead and lag. For such a condition the effect size was found to be comparable to vertical VBAP. Agreeing with [19] we believe that median plane reflections produce localization cues which are encoded by the auditory system.

ACKNOWLEDGMENTS

Our research was partly funded by the Austrian Science Fund (FWF) project nr. AR 328-G21, Orchestrating Space by Icosahedral Loudspeaker (OSIL). We thank all listeners taking part in the experiment.

REFERENCES

[1] J. Blauert. *Spatial hearing - the psychophysics of human sound source localization*. The MIT Press, 1983.

[2] K. Wendt. *Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie*. PhD thesis, RWTH Aachen, 1963.

[3] V. Pulkki. Localization of Amplitude-Panned Virtual Sources II: Two- and Three-Dimensional Panning. *Journal of the Audio Engineering Society*, 49(9):753–767, 2001.

[4] F. Wendt, M. Frank, and F. Zotter. Panning with height on 2, 3, and 4 loudspeakers. In *2nd International Conference on Spatial Audio (ICSA)*, Erlangen, 2014.

[5] M. A. Gerzon. Periphony: With-Height Sound Reproduction. *Journal of the Acoustical Society of America*, 21(1):2–10, 1973.

[6] R. Wallis and H. Lee. The effect of interchannel time difference on localization in vertical stereophony. *Journal of the Audio Engineering Society*, 63(10):767–776, 2015.

[7] M. Barron. The subjective effects of first reflections in concert halls-The need for lateral reflections. *Journal of Sound and Vibration*, 15(4):475–494, 1971.

[8] T. Okano. Image shift caused by strong lateral reflections, and its relation to inter-aural cross correlation. *The Journal of the Acoustical Society of America*, 108(5):2219–2230, 2000.

[9] T. Somerville, C. L. S. Gilford, N. F. Spring, and R. D. M. Negus. Recent work on the effects of reflectors in concert halls and music studios. *Journal of Sound and Vibration*, 3(2):127–134, 1966.

[10] R. Guski. Auditory localization: effects of reflecting surfaces. *Perception*, 19(6):819–830, 1990.

[11] F. Wendt, R. Höldrich, and M. Frank. The Influence of the Floor Reflection on the Perception of Sound Elevation. In *Fortschritte der Akustik*, pages 767–770, 2017.

[12] E. H. A. Langendijk and A. W. Bronkhorst. Contribution of spectral cues to human sound localization. *The Journal of the Acoustical Society of America*, 112(4):1583, oct 2002.

[13] P. Rubak and L. G. Johansen. Coloration in Natural and Artificial Room Impulse Responses. *23rd International AES Conference*, 2003.

[14] S. S. Sawilowsky. New Effect Size Rules of Thumb. *The Journal of Modern Applied Statistical Methods*, 8(2):597–599, 2009.

[15] D. Cabrera and S. Tilley. Vertical localization and image size effects in loudspeaker reproduction. *AES 24th Int. Conf. on Multichannel Audio, The New Reality*, 2003.

[16] ITU. ITU-R BS.1534-1: Method for the subjective assessment of intermediate sound quality (mushra), 2006.

[17] M. Y. Agaeva. The precedence effect in the horizontal and vertical planes in experiments with a moving lagging signal. *Human Physiology*, 37(5):545–549, 2011.

[18] A. Tregonning and B. Martin. The vertical precedence effect: Utilizing delay panning for height channel mixing in 3D audio. *Audio Engineering Society Convention*, 2015.

[19] B. Gourévitch and R. Brette. The impact of early reflections on binaural cues. *The Journal of the Acoustical Society of America*, 132(1):9–27, 2012.