7 - 12 SEPTEMBER 2014, KRAKÓW
FORUM ACUSTICUM

Preliminary study on the perception of orientation-changing directional sound sources in rooms

Franz Zotter[1,2], Matthias Frank[1,2], Andreas Fuchs[2], and Daniel Rudrich[2]
[1]Institute of Electronic Music and Acoustics,
[2]University of Music and Performing Arts, Graz, Austria.

Summary
Recently, how to reproduce the direction-dependent sound of musical instruments was investigated using spherical loudspeaker arrays, with the idea to increase naturalness in playback. As an interesting abstraction, such devices are effectively employed in spatial computer music using adjustable, artificial directivities (sound beams). When used in a room, this produces interesting sound objects outside the spherical loudspeaker array whenever the sound from wall reflections exceeds the echo threshold of the direct sound in its level. Nevertheless, our understanding of the nature and thresholds of the perceived spatial sound objects is only vague. Therefore, our contribution shows results of an initial investigation of artificial-directivity sound sources with variable orientation. To simplify repeatability, experiments were done in a semi-anechoic chamber. In the test setup, direct sound and first-order reflections of a shoebox room arriving at the listener were all simulated using a surrounding full-range loudspeaker setup by directional mapping to the closest loudspeaker. Hereby algorithm- and size-related properties of a particular spherical array prototype as well as the properties of a particular physical room are avoided. The results confirm that the orientation of directional sound sources can be heard, and that it causes localization deflected from the direct path.

PACS no. 43.55.Hy, 43.66.Qp, 43.30.Zk

1. Introduction

In rooms, what we hear from a source comes to our ears on various paths of propagation (direct sound, early reflections, diffuse sound). The signal carried by each of these paths is individual in strength, direction of radiation, direction of arrival, arrival time, coloration. Modifying the direction-dependent loudness or timbre of an instrument, speech, or any other sound source will therefore change what is being heard or received [1]. Most naturally, a particular modification of this kind occurs whenever a sound source changes its orientation, cf. [2]. However, it appears that we hardly know anything about the information that is accessible to our perception. To give some examples:

- *Is it possible to hear the orientation of a talker in a room?*
- *Can a source of adjustable directivity be used to produce "phantom sources"?*
- *Can we model auditory localization with various arrival times, directions, and magnitudes?*

After naming these examples it does not appear too blunt to the authors to state: Any scientific answer, psychoacoustic experimental evidence, or perception model is still only speculation, rare, or inexistent. Although recent work gives good hope for new answers, most notably [3, 4, 5], new experimental studies are required before models can be developed.

We present results of initial listening experiments concerning the afore-mentioned questions. Our first experiment deals with how good subjects can aurally align source orientation to given directions, and the second experiment deals with localization and so-called "phantom source" due to the reflected directional sound. A binaural and a geometric localization model are finally compared with the data.

2. Experimental setup

The experimental setup consists of a virtual 2D rectangular shoebox room of $8 \times 5\,\mathrm{m}$ with the listener in the middle. With regard to the listener, there is a virtual directional sound source placed at the polar coordinates $(r, \varphi)_A = (1.1\,\mathrm{m},\ 45°)$ and $(r, \varphi)_B = (1.8\,\mathrm{m},\ 65°)$, here in a clock-wise definition of the polar angle, cf. Fig. 1.

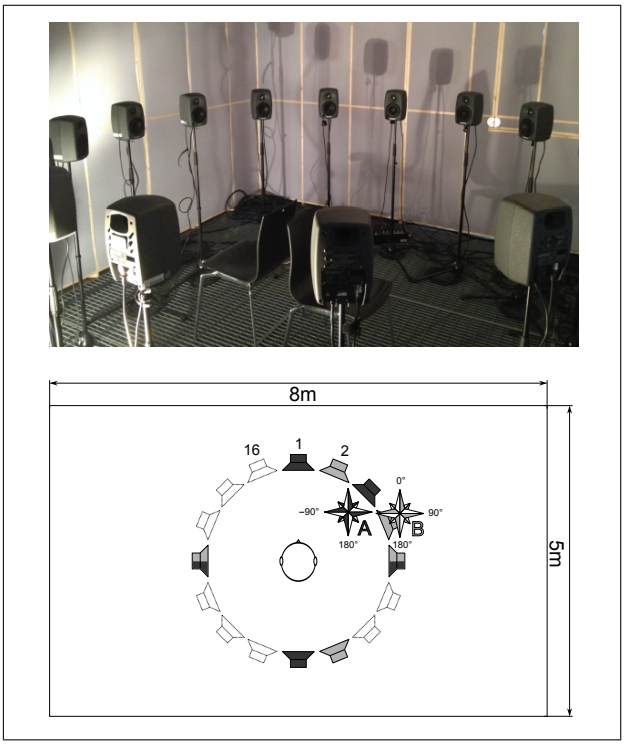

Figure 1. Experimental setup at IEM for 16 loudspeakers at 1.2 m radius to auralize direct and first-order reflected sound of a directional virtual source at position A or B in a 5×8 m virtual room using the gray-shaded loudspeakers.

The first-order wall reflections of the directional source were simulated by assuming fully reflecting walls (zero absorption). The simple auralization model used an attenuation $1/r$ depending on the distance, a delay of r/c, c is the speed of sound, and directional amplitude patterns $g(\varphi)$ for direct sound and suitably mirrored for image sources [6].

To represent the auralized result, a 2D loudspeaker-based room auralization setup was used, cf. Fig. 1. Loudspeaker-based auralization was done applying delays and amplitude weights and by directional mapping of every acoustic path arriving at the listener to the closest loudspeaker, similarly as in LoRA [7].

Three directional amplitude patterns were employed, the third of which being frequency-dependent:

$$g_1(\varphi) = \frac{1+\cos\varphi}{2} \quad \ldots \text{first-order} \tag{1}$$

$$g_2(\varphi) = \frac{1+\cos^3\varphi}{2} \quad \ldots \text{third-order} \tag{2}$$

$$g_3(\varphi,\omega) = \frac{[1.4+0.6\cos^3\varphi]\, H\left(\frac{\omega}{\omega_\mathrm{u}(\varphi)}\right)}{2}, \tag{3}$$

$$\omega_\mathrm{u}(\varphi) = 2\pi(20\mathrm{kHz} - 17.5\mathrm{kHz}|\sin\frac{\varphi}{2}|^{0.6})$$

$$H(\Omega) = \frac{1}{1+\mathrm{i}\Omega} \ldots 1^{\mathrm{st}} \text{ order low pass.}$$

All nine participants of the listening experiment were experienced (4 electrical engineering - audio engineering students at University of Music and Performing Arts and Technical University Graz, 5 Institute of Electronic Music and Acoustics employees).

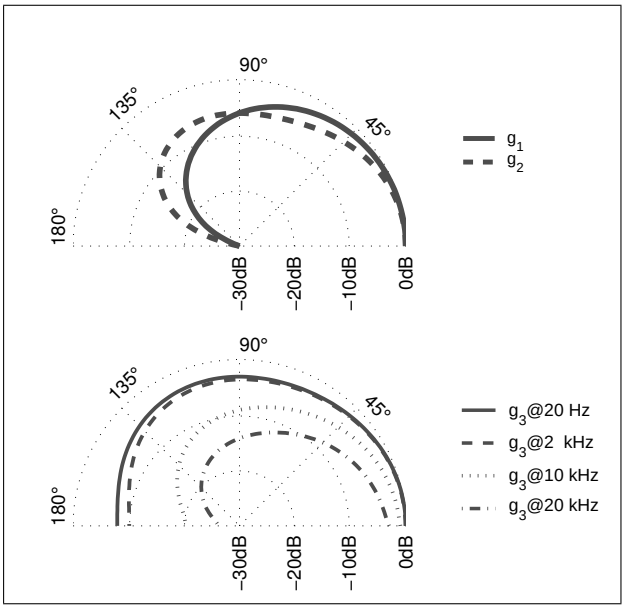

Figure 2. Broad-band directivity patterns (top): first-order $g_1(\varphi)$ (solid) and third-order $g_2(\varphi)$ (dashed). Frequency-dependent directivity pattern third-order $g_3(\varphi)$ at various frequencies (20 Hz, 2 kHz, 10 kHz, 20 kHz).

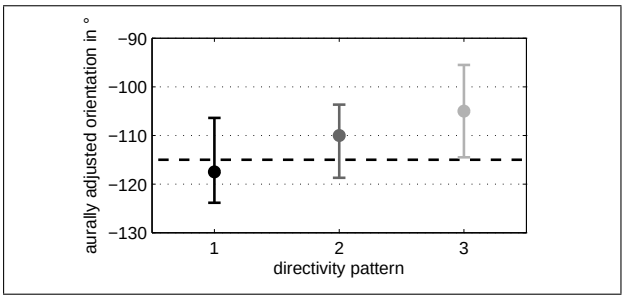

Figure 3. The desired orientation of the virtual source is 115° when it points at the listener (dashed line). Medians and 95% confidence intervals of the familiarization task describe responses of subjects who had to aurally find this orientation for a randomly oriented source.

The participants were given an PreSonus® *FaderPort* to change the orientation of the virtual source with the pan knob during the listening tasks. Per detent of the knob, a clockwise turn resulted in a 5° clockwise turn of the orientation.

3. Familiarization

As a familiarization task, participants of the listening experiment were asked to rotate the directional virtual source at the location B by turning an knob until they could hear the virtual source directivity exactly facing their listening position. The start orientation was selected randomly from $\{-20, 110, 45\}$°. In this task using speech, participants could familiarize with orientation-changing directional sound sources.

Fig. 3 shows that the participants could respond to the task with high accuracy for any of the source types. Their responses' confidence interval of about $\pm 10°$ mostly enclosed the correct orientation.

4. Orientation-changing directional sound source

The more difficult task of our preliminary study deals with the question: *Is it possible to hear the orientation of a talker in a room?*

Prior to experimentation, tests showed that it was hardly possible to hear static orientation. So participants were asked to rotate the directivity of the virtual source at B to face a target angle relative from the source that was $\{0, 90, 180, 270\}°$, i.e., front, right, back, left. The start orientation was $-115°$ (facing the listener).

Fig. 4 shows the results, in which a bias of $+45°$ is clearly visible. It might be due to the spatial configuration of the listening room, which might have appeared to be optically skew, and due to the lateral position of the source. For any of the directivities, the confidence intervals appear to be quite small although we expected larger intervals. Differences between the subjects appear to dominate the standard deviation.

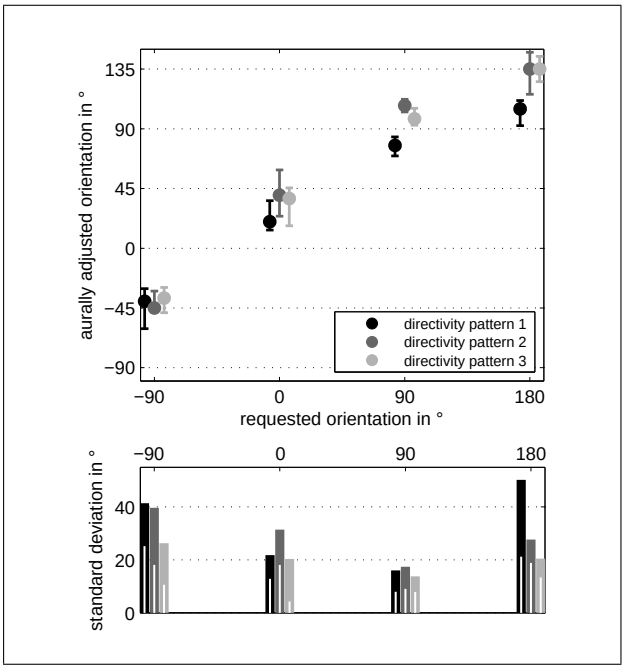

Figure 4. Top: Confidence intervals indicating that except for a bias of $+45°$, subjects of the listening experiment are able to adjust the orientation of a directional source in the room. Bottom: Inter-subjective (bar) and average intra-subjective (line) standard deviation.

5. Localization of a directional sound source

The third part of our preliminary study deals with two questions that are particularly interesting for artistic applications as discussed in [8]: *Can a source of adjustable directivity produce "phantom sources" and how to model them?*

In the task, labels $(1, 2, 3, \ldots)$ were sticked to the loudspeakers and participants were asked to give numbers (with half steps allowed) according to the position where they localized sound. Prior to the experiment, it turned out that the frequency-dependent directivity g_3 was not useful to show pronounced effects. It was therefore excluded from the experiment. Tested stimuli used three short pink noise bursts (100 ms for each: attack, sustain, decay, pause) that were played back to each participant from the sets of 2 different directivities $\{g_1, g_2\}$, 10 pre-defined orientations, for each of 2 virtual source positions {A, B}. Every participant had to localize each of the examples twice, whereby the repetition of the stimuli was left-right mirrored to avoid fatigue. The playback order of the $2 \times 10 \times 2 \times 2$ trials was random. The pre-defined orientations were selected to cover interesting effects before conducting the experiment.

Figs. 5, 6 show the histograms of the perceived direction responses in terms of black bubbles whose diameter enlarges by the number of responses. Diagrams in Fig. 5 show the results for the directivities $\{g_1, g_2\}$ at position A, Fig. 6 contains diagrams for position B. All diagrams depend on a set of virtual source orientations φ_S,
A: $\{-35, -15, -5, 35, 60, 65, 70, 75, 80, 105\}°$,
B: $\{-40, -10, 25, 35, 45, 65, 80, 95, 110, 125\}°$.
Along their horizontal axis (source orientation), radiation angles matching the existing propagation paths are marked by the corresponding label (fr, di, ri, ba, le). Along the vertical axis (perceived direction), arrival angles matching the simulated propagation paths are also marked correspondingly.

Position A. According to Fig. 5, both directivity patterns $\{g_1, g_2\}$ yield an expectable dominant localization of the direct sound for the source orientations $\varphi_S = \{-35, 105\}°$.

Table I. Distances of the direct sound and the 4 first-order reflections (front, right, back, left) for the source positions A and B in terms of radius r_l, time of arrival τ_l, and polar angles $\varphi_{S,l}$ leaving the source and $\varphi_{R,l}$ arriving at the listener.

A	fr	di	ri	ba	le
r_l in m	4.3	1.1	7.3	5.8	8.8
$\varphi_{R,l}$ in °	0.0	45.0	90.0	180.0	-90.0
$\varphi_{S,l}$ in °	-10.4	-135.0	96.1	-172.3	-95.1
B	fr	di	ri	ba	le
r_l in m	4.5	1.8	6.4	6.0	9.7
$\varphi_{R,l}$ in °	22.5	67.5	90.0	157.5	-90.0
$\varphi_{S,l}$ in °	-21.0	-115.0	96.8	-164.2	-94.5

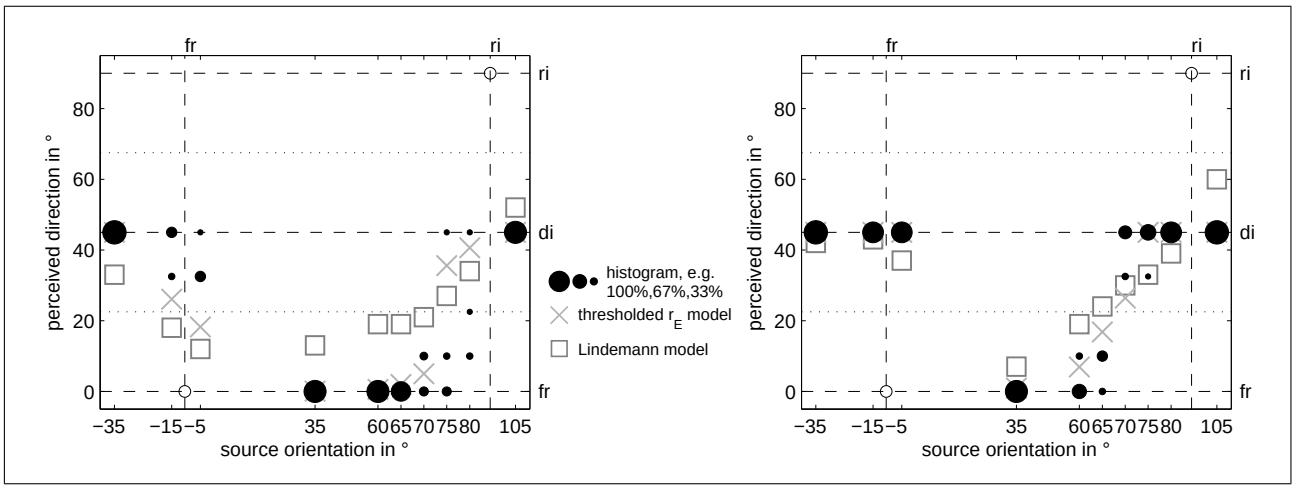

Figure 5. Histogram of perceived directions of virtual source of directivity g_1 (left) and g_2 (right) at position A under variation of its orientation (black bubbles). Squares (Lindemann) and crosses (thresholded $\boldsymbol{r}_\mathrm{E}$) show predicted directions.

By contrast, for the orientation $\varphi_\mathrm{S} = +35°$, localization fully flips to the frontal loudspeaker at $\varphi_\mathrm{R} = 0°$. Tab. I shows that direct sound leaves the source at $-135°$. Minus the source orientation of 35° to evaluate both directivities, this yields a strongly attenuated direct sound: $g_1(170°) = -42\,\mathrm{dB}$, $g_2(170°) = -33\,\mathrm{dB}$. The strongly suppressed direct sound obviously allows the frontal wall reflection to dominate for this source orientation.

As the source orientation φ_S changes from 35° to 105°, the transition between localization of direct sound and front wall reflection is similarly pronounced for both directivities, but slightly better controlled for g_2, Fig. 5. It is less controlled for an orientation change between $-35°$ to 35°, where only the broad directivity g_1 yields a stable "phantom source" between the auralization loudspeakers, i.e. acoustic paths.

In comparison, it seems that localization would normally flip sharply between the dominant direct sound and the temporarily dominant frontal reflection. Only the presence of the right wall reflection seems to support stable phantom sources in between the directions front and direct. The times of arrival $[\tau_\mathrm{fr}, \tau_\mathrm{di}, \tau_\mathrm{ri}]$ are $[13, 3, 21]$ ms. The right wall reflection is present in both transitions, but it seems to be too weak for $\varphi_\mathrm{S} = \{-15, 5\}°$ in g_2 to support a stable "phantom source" between the front wall reflection and the direct sound.

Position B. Fig. 6 indicates a very similar behavior for position B, however with the tendency of the localization being pulled to the right. This is not surprising in position B as the virtual source is shifted to the right, supporting the right wall reflection in both higher relative level and relatively earlier arrival time with regard to other propagation paths. The interpretation of the supporting character of the right wall reflection seems to apply for position B as well, and an overall better control of the "phantom source" localization is observed.

6. Modeling variable-directivity-induced localization

The above discussion of the results already gives some interpretation that raises the question, whether there can be a model for the auditory localization in case of variable-orientation directional sources. Its answer appears to be difficult, as the reflection paths are differently delayed in time. Reflections at different time scales, and different mechanisms in the precedence effect might need to be considered. *How could we still roughly model variable-directivity-induced localization in 2D?*

6.1. Binaural model (Lindemann)

The aim is to predict the perceived angles of the experimental conditions by a binaual lateralization model after Lindemann [9, 10], which is part of the Auditory Modeling Toolbox[1]. Table I shows for every acoustic path: the angle of radiation $\varphi_{\mathrm{S},l}$ and arrival $\varphi_{\mathrm{R},l}$, the acoustic path lengths r_l and thus the delays $\tau_l = r_l/c$. With any source directivity from Eq. (1)(2) rotated to the variable orientation φ_ori and evaluated at the radiation angles $g(\varphi_{\mathrm{S},l} - \varphi_\mathrm{ori})$, and the freely available dataset [11] of head-related impulse responses (HRIRs) evaluated at the arrival angles and times $h_\mathrm{LR}(\varphi_{\mathrm{R},l}, t - \tau_l)$, we are able to model the binaural impulse responses (BRIRs)

$$b_\mathrm{LR}(t) = \sum_l \frac{g(\varphi_{\mathrm{S},l} - \varphi_\mathrm{ori})\; h_\mathrm{LR}(\varphi_{\mathrm{R},l}, t - \tau_l)}{r_l}. \quad (4)$$

Characterizing the stimuli of the experiment, the BRIRs are convolved with 80 ms of pink noise and are then fed to the input of the lateralization model. The model divides the signals into 36 ERB frequency

[1] freely available on `amtoolbox.sourceforge.net/`

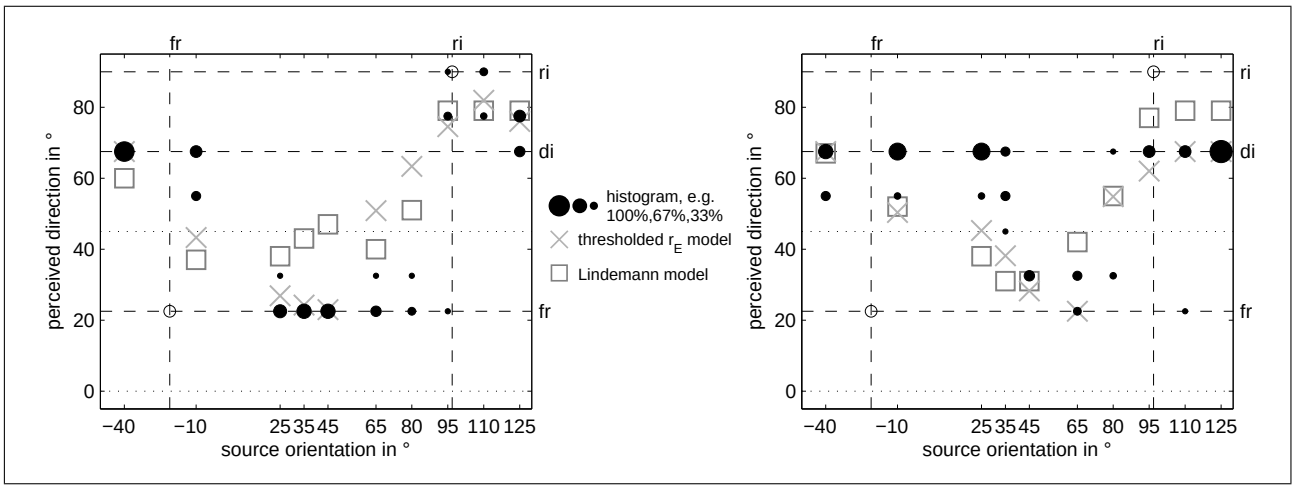

Figure 6. Histogram of perceived directions of virtual source of directivity g_1 (left) and g_2 (right) at position B under variation of its orientation (black bubbles). Squares (Lindemann) and crosses (thresholded $\boldsymbol{r}_\mathrm{E}$) show predicted directions.

bands and models the auditory nerve by half-wave rectification and low-pass filtering at 800 Hz. In each band, the inter-aural time difference (ITD) is computed as the centroid of the inter-aural cross correlation function [12], yielding one ITD value for each frequency band. The model also detects the inter-aural level-difference (ILD) and considers it in contra-lateral inhibition that modifies the ITD. Within each frequency band, we compare the ITD value with a lookup table. The pre-calculated values of this table are the ITD values of all individual HRIRs. The direction corresponding to the table entry with the best match estimates the direction of the auditory event in each frequency band. As a final outcome, the median value across all frequency bands gives us a single overall estimate of the perceived direction. The best fit to the median experimental responses has been achieved when using 17 frequency bands covering the range from 124 Hz ... 1753 Hz.
The model cannot distinguish between front and back, but also there were no responses from the back in the experiment. Hence, the ITD lookup was restricted to cover only directions between ±90°.

6.2. $\boldsymbol{r}_\mathrm{E}$ model with echo threshold

With the success of an energy interpretation in the work of Robinson et al. on the localization of diffuse reflections [4] and the success of the "energy vector" models in [13], we also attempt to model the experimental responses by squared signal components arriving from the given directions. In 2D, an "energy vector" model can be briefly written as the angle of a complex scalar:

$$\hat{\varphi}_\mathrm{S} = \angle \left\{ \sum_l w_l^2 \, e^{\mathrm{i}\varphi_{\mathrm{R},l}} \right\}. \tag{5}$$

Here, w_l denotes the amplitude factor of the l^th signal arriving from the angle $\varphi_{\mathrm{R},l}$. The above equation does not yet dependent on the arrival time $\tau_l = r_l/c$ in each individual propagation path l, but it can be easily modified to consider precedence: This is achieved by running the summation only over selected propagation paths whose amplitude w_l lies above the echo threshold $w_\mathrm{thr}(\tau)$ at the time they arrive, i.e. the indices of the summation are restricted to

$$\forall l : w_l > w_\mathrm{thr}(\tau_l). \tag{6}$$

The threshold discovered in [4, 5] only regards signal amplitudes whose dB level is larger than of every earlier signal minus $\frac{1}{4}$ dB/ms, weighted by the delay in ms elapsed after the respective earlier signal:

$$w_\mathrm{thr}(\tau) = \max_{\forall l:\tau_l<\tau} \left\{ w_l \, 10^{\frac{(\tau_l-\tau)/\mathrm{ms}}{4\cdot 20}} \right\}. \tag{7}$$

The weights w_l themselves are similarly composed as before. The directivity g from Eq. (1)(2) is evaluated at each angle of radiation minus the adjustable source orientation, which is then all divided by the path length r_l and weighted by the broad-band directivity of the human ear g_ear [13]

$$w_l = \frac{g\left(\varphi_{\mathrm{S},l} - \varphi_\mathrm{ori}\right) \, g_\mathrm{ear}(\varphi_{\mathrm{R},l})}{r_l}. \tag{8}$$

The broad-band human ear directivity g_ear can be found in [13, 14, 15], and its dB values can be expressed by the cosine series coefficients $(-2.9,\ 3.5, -1.5, -0.3)$ dB. With this, the $\boldsymbol{r}_\mathrm{E}$ estimate yields a fair average absolute difference to the median-experimental data of 11°. Further optimizing the cosine series coefficients with MATLAB's lsqnonlin yields a 6.7° average absolute difference with

$$20 \lg g_\mathrm{ear}(\varphi) = \begin{pmatrix} -4.4 \\ 4.5 \\ -0.3 \\ -1.5 \end{pmatrix}^\mathrm{T} \begin{pmatrix} 1 \\ \cos\varphi \\ \cos 2\varphi \\ \cos 3\varphi \end{pmatrix} \text{ in dB}. \tag{9}$$

Surprisingly, the resulting solid curve Fig. 7 is similar to the ear directivity for the 4.5 kHz third-octave, $g_\mathrm{ear}(\varphi) = \sqrt{\int \left[h^2_{4.5\,\mathrm{kHz},\mathrm{L}}(\varphi,t) + h^2_{4.5\,\mathrm{kHz},\mathrm{R}}(\varphi,t) \right] \mathrm{d}t}$.

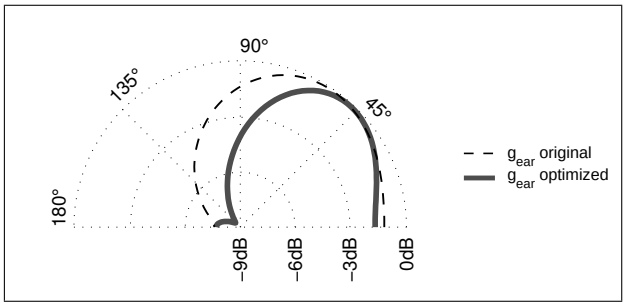

Figure 7. Overall, 3rd-order decomposed broad-band energy directivity of both ears together (dashed) and the ear directivity optimal for the localization model (solid).

6.3. Discussion

The two models presented above deliver estimates that are correlated with the experimental data: (i) the binaural localization model yields a coefficient of determination of $R^2 = 0.65$ and an average absolute difference of 12° to the median responses; (ii) the echo-thresholded $\boldsymbol{r}_\mathrm{E}$ model yields $R^2 = 0.81$ and an average absolute difference of 6.7° to the median responses.

7. CONCLUSIONS

We could give initial experimental results in this contributions showing that

- listeners are able to hear the orientation of directional sound source under certain circumstances, although its angle may be biased,
- directional sources of adjustable orientation can produce an audible "phantom source" under certain circumstances, in particular near orientations suppressing the direct sound,
- a binaural model is able to coarsely model localization, even if the sound arrives at different arrival times and magnitudes,
- and the "energy" vector model with echo threshold and ear directivity was less coarse in modeling localization.

Despite this study already providing some strong answers, it is still only at the beginning of the investigations. Primarily, the experimental setup and selection of directivities can easily be improved and made more general. Moreover, models are still only coarse and need refinement, also with aspects of apparent source width or stability and position dependence. Finally, we hope that results expected from current investigations on precedence, which seem to be a strong effort that is being undertaken in the psychoacoustic world, will help to improve our models.

Acknowledgement

The authors thank the subjects for participation. Experiments were carried out in the seminar *Algorithms in Computer Music and Acoustics* 2013/2014. The work was partly supported by ASD, a project funded by the Austrian ministries BMVIT, BMWFJ, the Styrian Business Promotion Agency (SFG), and departments 3 and 14 of Styria. The Austrian Research Promotion Agency (FFG) conducted funding under the Competence Centers for Excellent Technologies (COMET, K-Project).

References

[1] B.-I. Dalenbäck, M. Kleiner, P. Svensson: Audibility of changes in geometric shape, source directivity, and absorptive treatment-experiments in auralization. J. Audio Eng. Soc **41** (1993) 905–913.

[2] L. M. Ronsse, L. M. Wang: Effects of room size and reverberation, receiver location, and source rotation on acoustical metrics related to source localization. Acta Acustica u. Acustica **98** (2012) 768–775.

[3] R. Mehra, L. Antani, S. Kim, D. Manocha: Source and listener directivity for interactive wave-based sound propagation. IEEE Transactions on Visualization and Computer Graphics **20** (2014) 495–503.

[4] P. W. Robinson, A. Walther, C. Faller, J. Braasch: Echo thresholds for reflections from acoustically diffusive architectural surfaces. J. Acoust. Soc. Am. **134** (2013) 2755–2764.

[5] B. Rakerd, W. M. Hartmann, J. Hsu: Echo suppression in the horizontal and median sagittal planes. J. Acoust. Soc. Am. **107** (2000) 1061–1064.

[6] J. B. Allen, D. A. Berkley: Image method for efficiently simulating small-room acoustics. J. Acoust. Soc. Am. **65** (1979) 943–950.

[7] S. Favrot, J. Buchholtz: Lora: A loudspeaker-based room auralization system. Acta Acustica u. Acustica **96** (2010) 364–375.

[8] G. K. Sharma, F. Zotter, M. Frank: Orchestrating wall reflections in space by icosahedral loudspeaker: findings from first artistic research exploration. ICMC-SCM Athens, September 2014.

[9] W. Lindemann: Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals. J. Acoust. Soc. Am. **80** (1986) 1608–1622.

[10] W. Lindemann: Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front. J. Acoust. Soc. Am. **80** (1986) 1623–1630.

[11] H. Wierstorf, M. Geier, A. Raake, S. Spors: A free database of head-related impulse response measurements in the horizontal plane with multiple distances. 130th Convention of the Audio Engineering Society, May 2011.

[12] L. A. Jeffress: A place theory of sound localization. Journal of comparative and physiological psychology **41** (1948) 35–39.

[13] M. Frank: Phantom sources using multiple loudspeakers in the horizontal plane. Dissertation. University of Music and Performing Arts Graz, 2013.

[14] V. P. Sivonen, W. Ellermeier: Directional loudness in an anechoic sound field, head-related transfer functions, and binaural summation. J. Acoust. Soc. Am. **119** (2006) 2965–2980.

[15] G. Jahn: Zum Unterschied zwischen einohrigem und zweiohrigem Hören. Hochfrequenztechnik und Elektroakustik **72** (1963).