Auralization of Nearby Sound Sources - Part 2: Evaluation

Christian Borß, Jens Blauert
Institute of Communication Acoustics, Ruhr-Universität Bochum, 44780 Bochum, Germany, Email: christian.borss@rub.de

Abstract

In order to evaluate the near-field model as proposed in the previous part [1], we conducted a Multi-Stimulus test with Hidden Reference and Anchor (MUSHRA). The participants in the listening test were asked to rate the difference of a given reference and multiple stimuli using a novel category-rating scale, specifically designed for small to moderate perceptual differences. The reference stimuli were processed with Binaural Room Impulse Responses (BRIRs) that were rendered with Head-Related Transfer Functions (HRTFs) measured in the near-field. In each test, the reference was compared with the results of two models: a conventional model that only relies on far-field HRTFs and the proposed model, which predicts near-field HRTFs from far-field HRTFs. The test results show that the proposed model improves the auralization result for nearby sources, that is, provides more reliable distance cues [2].

Stimuli

The stimuli which we used in our listening test were generated by convolving a 14 s long segment of a dry vocal recording with a set of synthetic BRIRs. As BRIRs also include early reflections and late reverberation, they contain one of the main distance cues, the direct-to-reverberant ratio. Replacing the impulse for the direct sound by the various HRTFs described in [1] results in BRIRs which differ only by one property i.e., if and how the near-field effect is taken into account. Such BRIRs were synthesized for a $6.65 \times 4.90 \times 3.50$ m^3 sized room with a frequency-dependent reverberation time listed in Table 1. In total 7 test conditions were generated: a human speaker at 40 cm distance at $-110°$, $-30°$, $0°$, $30°$, and $110°$ and a human speaker at 80 cm distance at $-110°$ and $110°$. The image sources were rendered using a set of human HRTFs which were previously measured on a full sphere [3, Sec. 6.2]. The late reverberation was synthesized by a statistical time-frequency model [2]. For the generation of the lower anchor [4], we used a BRIR where the a sound source was placed at a doubled distance (i.e., 80 cm source distance for the 40 cm test cases and 1.6 m source distance for the 80 cm test cases).

f/Hz	125	250	500	1000	2000	4000	8000
$T_{60}(f)$/s	0.60	0.49	0.49	0.49	0.45	0.36	0.27

Table 1: Reverberation times of the synthetic BRIRs used for the listening test

Category-rating Scale

In order to gain an insight into the quantitative benefit of the proposed near-field model, we conducted an adapted "MUlti Stimulus test with Hidden Reference and Anchor (MUSHRA)" [4]. The categories of the MUSHRA test could be used to rate the degree of similarity between a reference stimulus and slightly different test stimuli. A pilot study, however, revealed that the test subjects rather judged the *dissimilarity* than the degree of similarity. Judging the perception of dissimilarity means specifying a distance in the multi-dimensional perceptual feature space. Hence, asking the subjects to rate the difference between the reference and the stimuli should result in a smaller variance, because it is in better accordance with the subject's course of action.

If the subjects are asked to rate the difference instead of the degree of similarity, then the conventional MUSHRA scale will be no longer applicable. For the formal listening test, a new category rating scale for small to moderate differences was developed. The rating scale uses the following verbal anchors in the style of Borg's CR10 scale [5]:

> A *"just noticeable difference" (0.5)* is so extremely small that it is barely audible i.e., something that is on the limit of what is possible to perceive.
>
> A *"very small difference" (1)* requires concentration but not your full concentration in order to notice the difference.
>
> A *"small difference" (2)* is a clearly perceivable difference which does not require much concentration in order to notice it.
>
> A *"moderate difference" (3)* is clearly audible but only significant in direct comparison.
>
> A *"big difference" (5)* describes a dissimilarity which does not require a direct comparison in order to perceive it.

In addition to the printed test instructions, a hardcopy of these definitions together with the scale was placed directly in front of the subjects during the listening test.

Listening Test

Our listening test was realized by a modified version of the publicly available MUSHRAM software [6]. To encourage the subjects not to limit their ratings to the values specified above, the user interface which is depicted in Fig. 1 showed tick marks at intervals of 0.2.

In total 37 subjects participated in the listening test. From these 37, the results of 2 subjects were not consid-

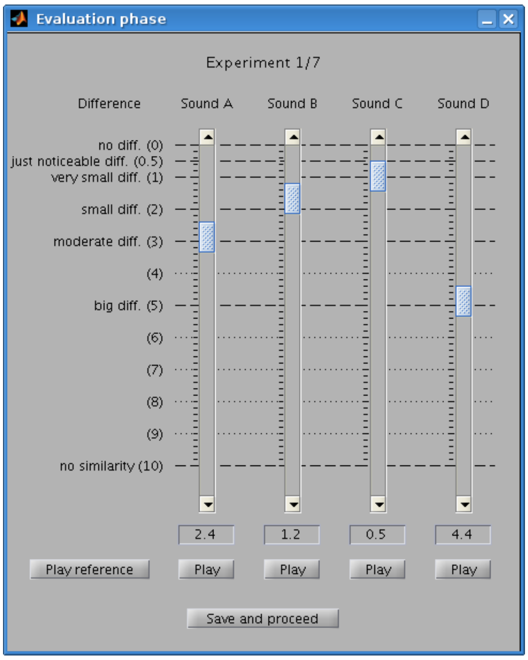

Figure 1: User interface of the test software

ered due to detected hearing impairments. To motivate the subjects to listen closely to the audio examples during the training phase, the subjects were instructed to describethe perceived differences. In contrast to the actual test phase, the order of the stimuli during the initial training phase was not randomized. The training phase contained two test conditions: the stimuli for a sound source at $-110°$ (left) at 40 cm distance and for a sound source at 30° (right) at 40 cm distance.

Figure 2 displays the result for one of the test conditions, that is a sound source at 40 cm distance on the left hand side at $-110°$ (see [2] for more results). In this figure, the cumulative distributions of the listening test results CR for the hidden reference (green), the proposed near-field model (solid black), the conventional method without near-field-equalization (dashed black), and the lower anchor (red) are shown. The median values and quartile ranges are indicated by the horizontal dotted lines. The observable improvement of the near-field-equalization method becomes even more apparent

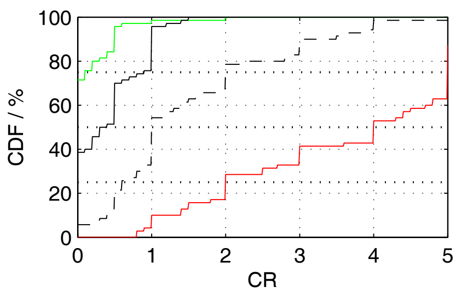

Figure 2: Cumulative distribution of the listening test results for a sound source at 40 cm distance at -110° (green: hidden reference; solid black: proposed near-field model; dashed black: without near-field-equalization; red: lower anchor)

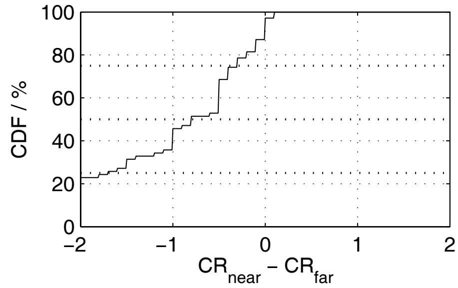

Figure 3: Cumulative distribution of the difference between the listening test results for the stimuli with and without near-field-equalization

in Fig. 3, where the cumulative distribution of the difference is shown. About 90% of the participants rated the near-field-equalized stimulus to be more similar to the reference than the stimulus without the near-field-equalization.

Conclusions

The results of the conducted listening test documents the quantitative benefit of the near-field-equalization method that we discussed in the previous part [1]. The steps being observed in the cumulative distribution of the test results can be attributed to the verbal anchors of the category rating scale that we designed for small to moderate perceptual differences.

Acknowledgements

The authors would like to thank Prof. Rainer Martin, Institute of Communication Acoustics, Ruhr-Universität Bochum, for his valuable comments and suggestions, which helped to improve this article.

References

[1] Borß, C., "Auralization of Nearby Sound Sources - Part 1: Near-Field Model," in *Proc. Annual Meeting of the German Acoust. Society (DAGA)*, Düsseldorf, Germany, Mar. 2011.

[2] Borß, C., *An Improved Parametric Model for the Design of Virtual Acoustics and its Applications*, Ph.D. thesis, Institute of Communication Acoustics, Ruhr-Universität Bochum, Bochum, Germany, to appear 2011.

[3] Silzle, A., *Generation of Quality Taxonomies for Auditory Virtual Environments by Means of Systematic Expert Survey*, Ph.D. thesis, Institute of Communication Acoustics, Ruhr-Universität Bochum, Bochum, Germany, 2007.

[4] ITU-R, *Rec. ITU-R BS.1534-1, Method for the Subjective Assessment of Intermediate Quality Level of Coding Systems*, 2003.

[5] Borg, G. and Borg, E., "A New Generation of Scaling Methods: Level-anchored Ratio Scaling," *Psychologica*, vol. 28, pp. 15–45, 2001.

[6] *MUSHRAM 1.0 User Guide*, Nov. 2005, URL: `http://www.elec.qmul.ac.uk/digitalmusic/downloads/index.html#mushram`.