Experimental investigations into the distance perception of nearby sound sources: Real vs. WFS virtual nearby sources

Stefan Kerber[1], Helmut Wittek[2], Hugo Fastl[1], Günther Theile[2]

[1] *AG Technische Akustik, MMK, TU München, D-80290 München, Germany, Email: kerber, fastl@ei.tum.de*
[2] *Institut für Rundfunktechnik, D-80939 München, Email: wittek,theile@irt.de*

Introduction

It is known, that there are a number of cues (like intensity, direct-to-reverberant energy ratio, spectral cues, ...) that serve us to estimate the distance of a sound source. One feature for perception of near sources is the interaural level difference (ILD), which increases with approaching distance. Brungart and Rabinowitz ([1]) have shown that even in anechoic environment it is possible to estimate distance by evaluating this cue. It is also known that in principle WFS is able to produce an exact spatial copy of a natural wavefield. Combining these two facts leads to the conclusion that distance perception for close sources should lead to more accurate estimations, when sound is reproduced by WFS focused sources rather than played back by conventional reproduction techniques. Therefore some theoretical simulations were done by Wittek, presented in a companion paper ([2]), and a listening tests was accomplished to check this assumption. The test was split in two parts, due to interferences in the designs. The first part examined the distance perception for WFS focused, the second the distance perception for real sources. This test is the topic of the present paper.

Experimental setup

Basic setup

Both listening tests took place in the anechoic chamber of the IRT in Munich. A schematic setup for the WFS experiment can be seen in figure 1. No figure is given for the setup of the other experiment, because it is the same with a movable loudspeaker hanging from the ceiling which replaces the WFS array. The test persons were placed in the room, so that the (real and virtual) sound sources lay on their interaural axis in seven distances between 25 cm and 190 cm. The correct positions of the WFS array (16 IRT custommade speakers, loudspeaker interspacing 17 cm) and the loudspeakers (ELAC Type 301) which acted as real sources, have been hidden by an acoustically transparent curtain. A movable loudspeaker (not connected, just as a visual help) on a cable car in front of the subjects was used to show the apparent distance of the sounds (six pink noise pulses of duration 1 s, pause 0.4 s, ramp 0.1 s) perceived by the subjects. The function of this display has been verified in a preliminary test with 18 test subjects. On the subjects' left hand side three speakers have been placed. They were used for a training before the actual listening test startet. All sources were measured and linearised before the experiments started, to avoid distance estimates because of differences in sound colour.
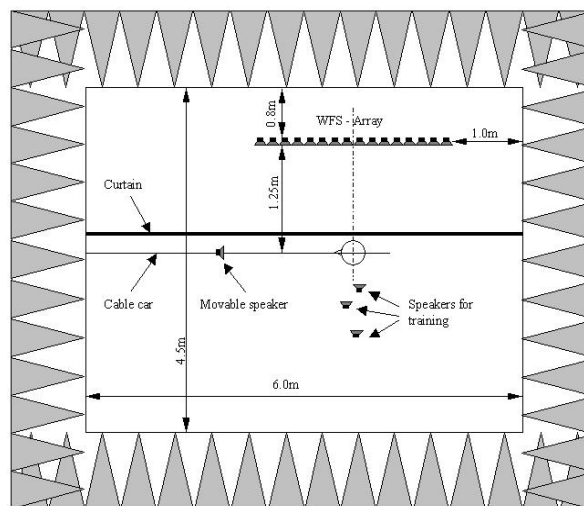

Figure 1: Setup for the WFS experiment in the anechoic chamber of IRT. The setup for the other experiment is identical with the WFS array replaced by three movable loudspeakers.

Subjects

As test subjects acted staff members of IRT, department Audiosystemtechnik. There were eight persons for the test with real sources and eleven persons for perception of virtual sources. The average time for one listening session was about 20 minutes. All participants possessed normal hearing, and had experience in former listening tests. This was important due to the difficulties of these experiments.

"Conflicting-" and "Not conflicting" cues

In case of playing back all sources with a constant source level during the experiment, listeners would only have estimated distance by evaluating the perceived loudness, because loudness is a strong cue and therefore able to overwrite other distance cues. To avoid this effect the concept of "conflicting" and "not conflicting" cues was introduced. During a listening session sources were played back two ways. One way was the "not conflicting" way, what means a constant source level, which causes farer sources to be quieter at the receiver position. The other method, refered to as "conflicting cues", means that source levels were not kept constant. Here listeners were forced to evaluate other distance cues than loudness to give their distance estimations. These levels were derived by permutation of the natural receiver levels following a fixed scheme (Table 1).

Procedure

Each session started with a short training for the subjects, so the subjects got used to make distance estimates under these

difficult circumstances. After the training session the persons started with the experiment. A sound source was played, and after that the person had to position the cable car in the distance he / she assumed the sound source to be. This distance was annotated by the experimenter, and a new run was started. The listening tests were split in four sessions, two for perception of real sources, and another two for distance estimation of virtual sources, which all took place on different days. The order of the played sound sources was random.

"not conflicting cues"	distance	"conflicting cues"
±0.0 dB(A)	190 cm	+5.0 dB(A)
±0.0 dB(A)	150 cm	+13.8 dB(A)
±0.0 dB(A)	110 cm	+7.5 dB(A)
±0.0 dB(A)	85 cm	- 4.6 dB(A)
±0.0 dB(A)	65 cm	- 9.1 dB(A)
±0.0 dB(A)	45 cm	- 3.4 dB(A)
±0.0 dB(A)	25 cm	- 9.2 dB(A)

Table 1: Relative sound source levels for "conflicting" and "not conflicting" cues.

Results

The evaluation was done separately for conflicting and not conflicting cues and for real versus virtual sound sources. The results for conflicting cues are displayed in figure 2, which shows the mean distance estimations for all test subjects with 95% - confidence intervals. Additionally the frequency of occurence for discrete answers can be seen in the graph.
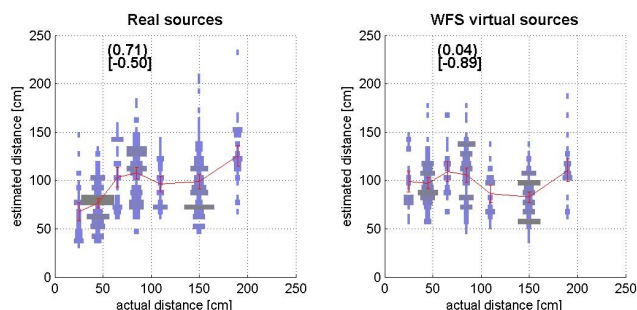

Figure 2: Results for conflicting cues: Distance perception for real sources in the nearfield is more acurate than for virtual sources. The numbers on the top left corners show Spearman - correlation between estimated and real distance () and estimated distance and level [].

Results plotted in figure 2 suggest that distance perception only works for small distances (< 1 m) with real sound sources. This is the area where Brungart and Rabinowitz ([1]) reported the existence of ILDs which can act as distance cues. The correlation between actual and estimated distance (significance 5%) also verifies this assumption. It can also be seen that small distances get overestimated, large ones get unterestimated, which was also found in older studies ([3]). Taking a closer look at the results for WFS virtual sources reveals a strong correlation (-0.89, significance 1%) between level and estimated distance. These two facts lead to the following conclusion: Dry focused WFS sources are not able to produce the same

distance perception as real sound sources do. This might be caused by the lack of other cues than level in this experiment. Therefore a distance estimate without the presence of correct level is not possible in this experiment.

This also holds when taking a closer look at the results of the individual subjects. Six out of eleven persons reach a significance level for the correlation level with distance, which is better than 5%.

When presenting sound sources played back with natural level, the above argument is again approved (Figure 3).
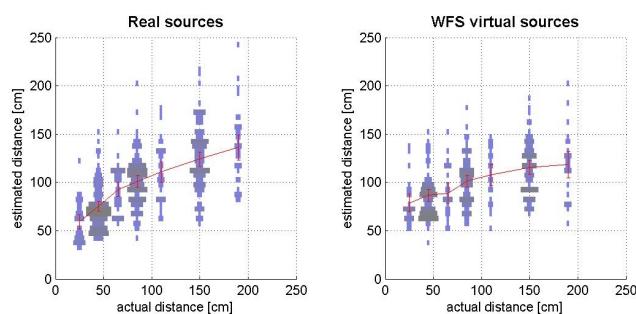

Figure 3: Results for not conflicting cues: Distance perception for real sources in the nearfield is again more acurate than for virtual sources.

The distance perception now works over the whole observed area for both rendering techniques. Notice the more accurate slope and the smaller over- respectively underestimation of the distances for real sources. This also leads to the conclusion of an improved perception for real sources, maybe due to the existence of correct natural ILDs. Calculation of the Spearman correlation reveals a strong dependency between actual and estimated distances for both, WFS and real sources over all subjects.

Conclusion

The experiments compared the near field distance perception for real versus dry focused WFS virtual sources. As shown, the distance perception is worse compared to real sources, maybe due to lack of other important distance cues like ILDs. A next possible follow up step could reveal, how distance perception for focused WFS sources can be improved by adding other selected cues to the signals.

References

[1] D. S. Brungart and W. M. Rabinowitz. Auditory localization of nearby sources. Head-related transfer functions. *J.Acoust.Soc.Am.* **106**: 1465-1479, (1999)

[2] H. Wittek. Spatial perception in WFS rendered sound fields: Distance of real and virtual nearby sources. Companion paper, CFA/DAGA 04, Strasbourg, March 2004.

[3] P. Zahorik. Auditory Distance Perception: A Literature Review, August 1996, URL:
`http://www.waisman.wisc.edu/~zahorik/papers/dist.html`