

The Perceptual Effect of Reflective Surfaces on Acoustic Crosstalk Cancellation Using an 8-Channel Linear Loudspeaker Array

Karim BAHRI; Jens AHRENS¹

Audio Technology Group, Division of Applied Acoustics
Chalmers University of Technology, 412 96 Gothenburg, Sweden

ABSTRACT

We recently presented a user-tracked crosstalk cancellation system based on a linear array of loudspeakers. The system uses Ambiphonics in the low frequency range and a superdirective nearfield beamformer in the mid and higher frequency range. Efficient real-time execution is achieved by parameterization of pre-computed beamformer weights. In this contribution, we investigate the perceptual effect of different reflecting surfaces that are apparent in the reproduction room on binaural content that is presented through our system. Our user study shows that a floor reflection does not have an audible effect. An increasing amount of reverberation from the playback room makes perception more spacious and pleasant. Contrary to previous studies, even a strong lateral reflection from a close wall did not impair localization significantly.

Keywords: Crosstalk cancellation, binaural audio, beamforming

1. INTRODUCTION

Crosstalk cancellation (CTC) has been pursued since the 1960s [1]. Most of the early implementations employed a pair of loudspeakers with a loudspeaker positioned ipsilateral to each of the ears of the listener, respectively. The crosstalk of each loudspeaker to the corresponding contralateral ear was estimated by different means, and a cancellation signal was emitted by the other loudspeaker. Numerous authors have contributed to the field. We refer the reader to the references in, for example, [2-4] for a non-exhaustive list. Starting in the 1990s, researchers have been evaluating the capabilities of loudspeaker arrays to perform acoustic CTC [5] with just as many contributors as to two-channel CTC. We again refer to the references in, for example, [2-4].

CTC systems are usually designed assuming free-field conditions. The effect of reverberation on channel separation was studied instrumentally in [6]. The effect of a low number of lateral early reflections on localization of sound sources in binaural content was studied in [7] using a 2-loudspeaker setup operating in an anechoic chamber. The reflections were produced using flat rigid surfaces that were installed in the chamber. It was found that such isolated early reflections can cause a localization bias towards the incidence direction of the reflection. Front/back confusions were higher in the non-anechoic conditions with a bias towards localization in the front hemisphere where the loudspeakers were located. The effect of the floor reflection was not studied.

It was found in [8], on the other hand, that the room response did not affect localization significantly when simulated with an image-source model.

Simple room simulations as the one in [8] do not account for relevant acoustic effects and can sound artificial when rendered spatially, and the manual handling of reflective surfaces as in [7] limits the range of acoustic conditions that can be tested. We therefore use measured impulse responses from the loudspeakers of a CTC system to the ears of a dummy head in different acoustic environments in the present study to auralize the CTC system. This allows for switching between environments by the click of a button and puts no limits on the range of acoustical conditions that can be covered.

¹ jens.ahrens@chalmers.se

2. SYSTEM DESIGN

The present study employs the system that we presented in [2-4]. It is a linear equi-spaced 8-channel loudspeaker array as depicted in Fig. 1. The geometry of the scenario under consideration is depicted in Fig. 2. The core of the system is a convex superdirective nearfield beamformer that directs a beam to one of the ears of the listener and produces a null at the other (contralateral) ear. Assuming that the loudspeakers emit ideal spherical waves, then the transfer functions from one of (the two) input channels of the system to the ears of the listener that is assumed to be located centrally at 1 m distance to the array show a channel separation of at least 20 dB over the vast part of the audible frequency range as depicted in Fig. 3.



Figure 1 – Example prototype using 8 Neumann KH 80 DSP loudspeakers with a spacing of 154 mm

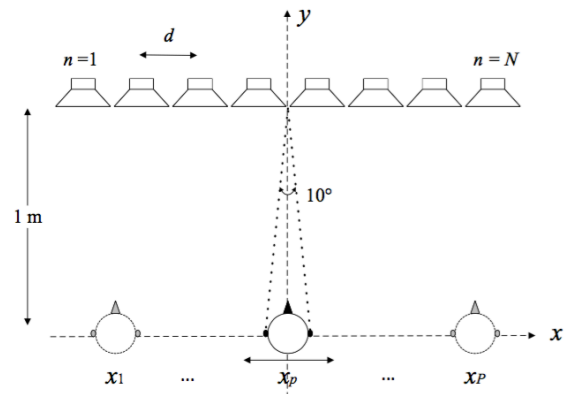


Figure 2 – System geometry; the enabled listening locations are along a line parallel to the array

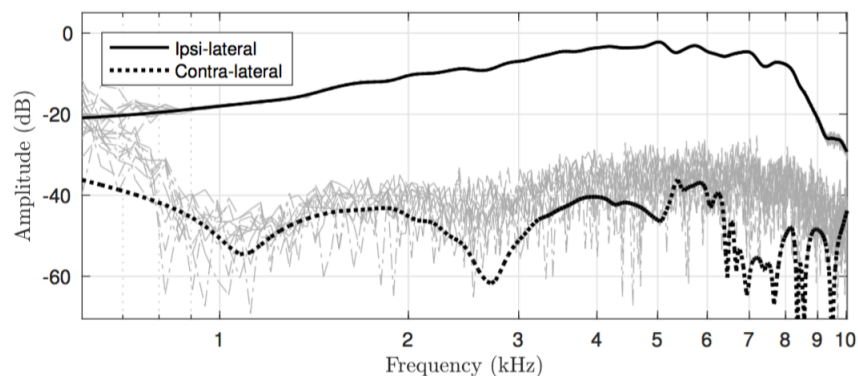


Figure 3 – System transfer function to the two ears of the listener under ideal conditions (black lines) as well as with simulated loudspeaker mismatch (gray lines) (data from [2])

Fig. 3 reveals that the channel separation can drop significantly at frequencies below 1 kHz as soon as some amount of mismatch of the sensitivity of the loudspeakers and uncertainties in the loudspeaker placement are included in the simulation.

This reduction of the channel separation at low frequencies is not surprising as we are looking at a frequency range in which the two control points, the ears, are separated by less than a wavelength. We therefore chose to use the beamformer only above 800 Hz and recursive ambiophonic crosstalk elimination (RACE) [9], which is a simple two-channel CTC method that has shown to be effective and degrade gracefully, below 800 Hz.

Translations of the listener are enabled by pre-computing the beamformer weights for a set of listener positions and subsequent parameterization as presented in [3,4].

Measurements on the prototype under anechoic conditions show that a channel separation of around 20 dB is apparent over most of the frequency range independent of the listener position [4] (cf. also Fig. 6). We demonstrated in [10] that the design of the beamformer based on the measured directivity of the loudspeakers increases the CTC only in a limited frequency range around 2 kHz. The data from Fig. 6 do not take the loudspeaker directivity into account.

3. MEASUREMENT DATA AND INSTRUMENTAL EVALUATION

To assess the perceptual influence of reflecting surfaces and reverberation on crosstalk canceled binaural content, we measured impulse responses from each of the loudspeakers of the array to the ears of a KEMAR dummy head (DH) in different environments and for different head orientations in steps of 1°. Automatic rotation was performed using VariSphear [11]. This allows for performing binaural head-tracked auralization of the virtual array as well as switching between different environments by the click of a button. The following environments were measured:

1. Anechoic chamber
2. Anechoic chamber with a floor reflection
3. Small dry laboratory room
4. Small dry laboratory room with a reflective side wall and the DH located in the center of the room
5. Small dry laboratory room with one reflective side and rear wall and the DH located in the center of the room
6. Small dry laboratory room with a reflective rear wall and with the DH located 1 m away from a highly reflective lateral wall (to the right of the DH)

Fig. 4 shows photographs of some of the setups. Fig. 5 depicts sample ear impulse responses. The resulting channel separation for the environments *anechoic* (1), *anechoic plus floor reflection* (2), and *dry room* (3) is depicted in Fig. 6. We chose not to show data on channel separation for the other (more reverberant) conditions as it is not straightforward to differentiate direct sound and room reverberation, which may both trigger different perceptual mechanisms. Interpreting frequency-domain channel separation can be misleading in these cases. It is evident from Fig. 6 that the floor reflection does not have a significant effect on channel separation.

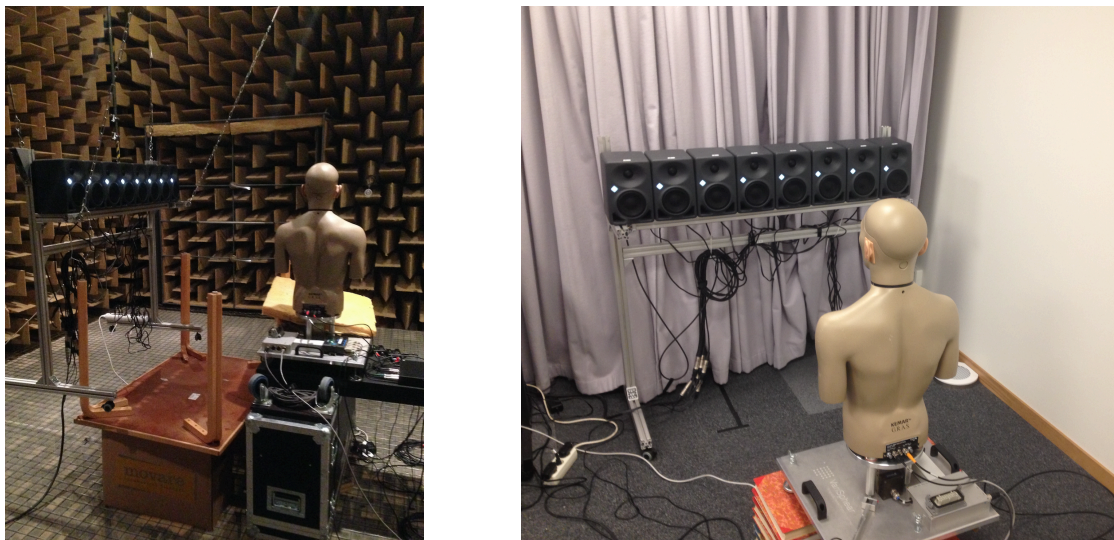


Figure 4 – Photographs of example measurement setups:
environment 2 (left); environment 6 (right)

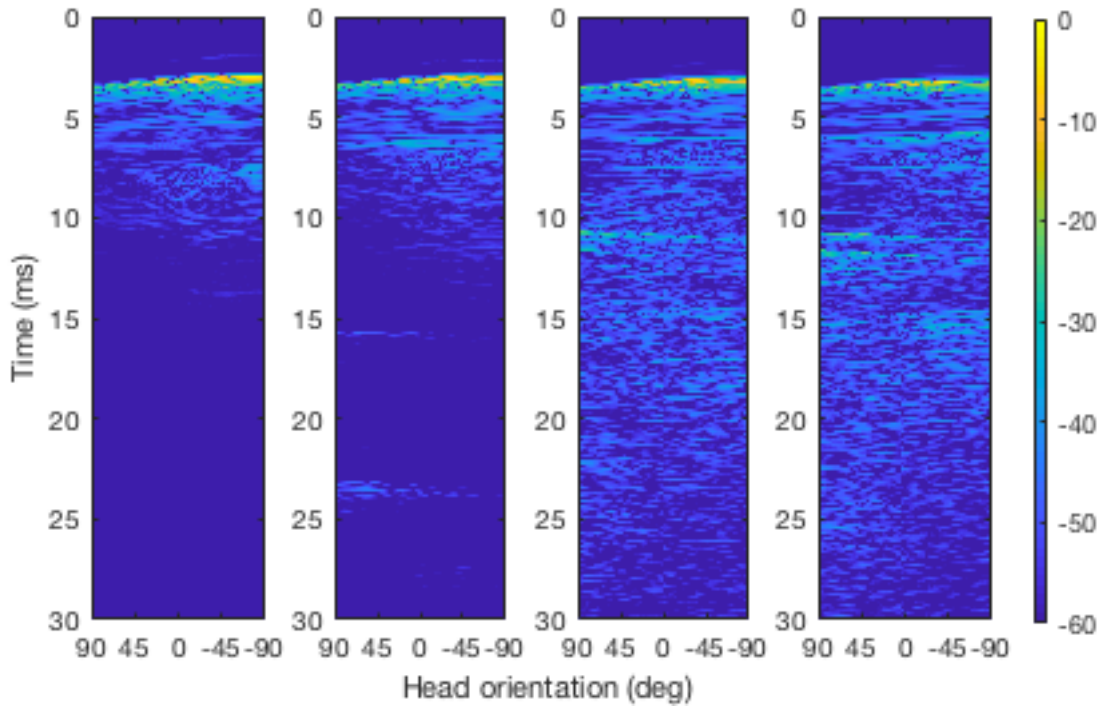


Figure 5 – Impulse responses from loudspeaker 4 to the left ear of the KEMAR manikin on a logarithmic scale (from left to right): environment 1, environment 2, environment 5, environment 6

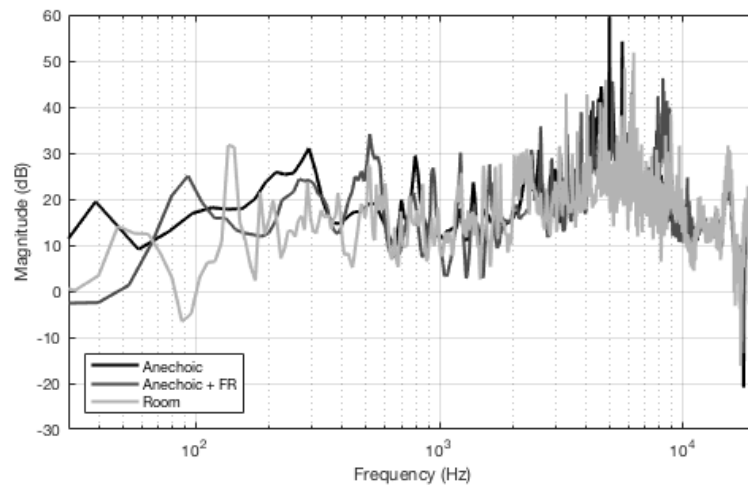


Figure 6 – Channel separation at the ears of the DH for three of the measured conditions; no separation of direct sound and reverb was performed

4. PERCEPTUAL EVALUATION

4.1 Experiment Design

We auralized the CTC array in all 6 environments from Sec. 3 by means of head-tracked binaural synthesis of the DH measurement data and using the SoundScape Renderer (SSR) [12,13]. The subjects were instructed not to move their head excessively as the CTC method does not account for this. We nevertheless employed head tracking so that the auralization accounts for small head movements, which can reduce distortion of the spatial perception [14].

The virtual loudspeaker array was playing binaural content, which consisted of anechoic male speech spatialized in the directions straight ahead (0°), 30° to the left, and 90° to the left by means of KEMAR HRTFs. The listener was positioned at 1 m distance from the array and symmetrically w.r.t. the array. The virtual orientation of the listener was normal towards the array. Note that we found that making the listener look away from the array in normal direction can significantly reduce front/back confusions [4]. We chose not to investigate front/back confusions in the present study as the phenomenon reported in [4] has not been decoded yet.

The subjects were provided with a graphical interface with which they were able to switch seamlessly between the conditions while the speech signal was playing continuously. A total of 18 different conditions were present (6 rooms times 3 virtual source positions).

We found in a pilot study that the perceptual differences between the auralizations of the different room conditions can span a broad range from hardly or not perceptible to clear multidimensional differences. We therefore chose to run the study as an interview. The subjects were comparing the 6 different conditions for each of the 3 virtual source positions separately and reported in free speech to the experimenter what differences they were hearing. The subjects did not have further information on what it was that they were listening to.

4.2 Results and Discussion

9 grown-up subjects with self-reported normal hearing participated in the experiment. The average duration was 23 min.

We distilled the following from the subjects' responses:

- The floor reflection has no audible influence (environment 1 vs. 2).
- The lateral virtual source positions were perceived more spacious.
- The effect of the dry room (environment 3) is minor. The perception is very similar to (environments 1 and 2).
- An increasing amount of reverberation increases externalization. Internalization can occur in environments 1 and 2.
- An increasing amount of reverberation makes spatial perception more plausible in general, for example, in terms of localization accuracy, locatedness, and source width.
- Front/back confusions occurred mostly in environments 1 and 2.
- Environments 4 and 5 evoked the most plausible and pleasant perception.
- Approx. half of the subjects perceived the virtual source as slightly elevated. Reverberation mitigated this effect.
- The strong lateral reflection in environment 6 was perceived disturbing by most of the subjects.
- No major effect of the strong lateral reflection in environment 6 on localization was reported.

Our results suggest that there is a range of room acoustic conditions that influence the presentation of binaural content only to a minor extent (environments 1-5). This confirms the results from [8], which were obtained based on a simple room simulation. Strong and isolated lateral reflections caused localization bias in the experiment presented in [7]. We cannot confirm this observation. Environment 6 in our study comprised a similarly strong reflection but embedded in the room reverberation. This reflection did have a negative effect, however, not primarily on localization but rather on a more general level. It seems that embedding a reflection in reverberation mitigates the perceptual impairment.

5. CONCLUSIONS

We auralized an 8-channel crosstalk cancellation array based on manikin data measured in different environments to assess the perceptual influence of the acoustic response of the environment on

the perception of binaural content. We found that the floor reflection does not have a noticeable influence. An increasing amount of reverberation and early reflections support spaciousness and reduce front/back confusions. Even excessive reverberation yielded acceptable results.

ACKNOWLEDGEMENTS

We thank Georgios Zachos for building the arrays.

REFERENCES

- [1] B. B. Bauer, "Stereophonic Earphones and Binaural Loudspeakers," *J. Audio Eng. Soc.*, vol. 9, no. 2, pp. 148–151 (1961).
- [2] C. Hohnerlein, J. Ahrens, "Perceptual evaluation of a multiband acoustic crosstalk canceler using a linear loudspeaker array," presented at the 2017 IEEE ICASSP, pp. 96–100 (2017 March)
- [3] X. Ma, C. Hohnerlein, J. Ahrens, "Listener-Position Adaptive Crosstalk Cancellation Using A Parameterized Superdirective Beamformer," in *Proc. of IEEE SAM* (2018 July).
- [4] X. Ma, C. Hohnerlein, J. Ahrens, "Concept and Perceptual Validation of Listener-Position Adaptive Superdirective Crosstalk Cancellation Using a Linear Loudspeaker," in *JAES* (2019) (submitted)
- [5] J. Bauck, D. H. Cooper, "Generalized Transaural Stereo and Applications," in *JAES* 44(9) pp. 683-705 (1996)
- [6] D. B. Ward, "On the performance of acoustic crosstalk cancellation in a reverberant environment," in *JASA* 110(2) (2001 August)
- [7] A. Sæbø, *Influence of Reflections on Crosstalk Cancelled Playback of Binaural Sound*, Doctoral Dissertation, NTNU Trondheim (2001)
- [8] D. Kosmidis, Y. Lacouture-Parodi, and E. A. P. Habets, "The Influence of Low Order Reflections on the Interaural Time Differences in Crosstalk Cancellation Systems," in *IEEE ICASSP*, Florence, Italy (2014 May)
- [9] R. Glasgal, "360 Localization via 4.x RACE Processing," presented at the 123th Audio Engineering Society Convention (2007 Oct)
- [10] J. Ahrens, "The Effect of Loudspeaker Radiation Properties on Acoustic Crosstalk Cancellation Using a Linear Loudspeaker Array," in *Proc. of DAGA*, Rostock, Germany, (2019 March)
- [11] B. Bernschütz, C. Pörschmann, S. Spors, and S. Weinzierl, "Entwurf und Aufbau eines variablen sphärischen Mikrofonarrays für Forschungsanwendungen in Raumakustik und Virtual Audio," in *Proc. of DAGA*, Berlin, Germany (2010 March)
- [12] Geier, M., Spors, S., and Ahrens, J. (2008). "The SoundScape Renderer: A unified spatial audio reproduction framework for arbitrary rendering methods," in 124th Convention of the AES, p. 7330, the software can be downloaded from <http://spatialaudio.net/ssr/> (Last viewed 3/27/2019).
- [13] Geier, M., and Spors, S. (2010). "Conducting psychoacoustic experiments with the SoundScape Renderer," in 9. ITG Fachtagung Sprachkommunikation, Bochum, Germany, pp. 1–4.
- [14] Begault, D.R., Wenzel E.M., and Anderson M.R., "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," in *JAES* 49(10), p. 904-16 (2001 Oct)