Evaluation of the Impact of Spatial Aliasing on Perceived Spaciousness in Wave Field Synthesis

Jens Ahrens

Audio Technology Group, Chalmers University of Technology, Gothenburg, Sweden,
Email: jens.ahrens@chalmers.se

Introduction

The theory of sound field synthesis is well understood [1]. Under certain prerequisites – one of which is the availability of a continuous distribution of secondary sources – a given desired sound field can be synthesized with perfect accuracy. Practical implementations always use a finite number of discrete loudspeakers, which causes spatial aliasing above a given frequency (the *spatial aliasing frequency*). Spatial aliasing typically manifests itself as spurious wave fronts that follow the desired wave front as illustrated in Fig. 1. We focus on the approach of wave field synthesis (WFS) in this paper, which is what is shown in this figure.

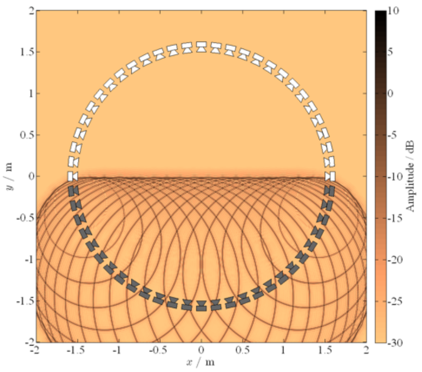

Figure 1: Snapshot of the sound pressure of a sample virtual planar wave front propagating into positive y-direction and carrying a time-domain impulse; a cross-section through the horizontal plane is shown; the wave is synthesized by a circular 56-channel wave field synthesis array; the spatial aliasing frequency is approx. 1,700 Hz; all wave fronts other than the leading straight wave front are spatial aliasing; gray loudspeaker symbols represent active loudspeakers; white loudspeaker symbols represent inactive loudspeakers

The effect of spatial aliasing on various perceptual dimensions has been studied in a number of works. It was shown in [2] that spatial aliasing increases the apparent source width when single wave fronts are synthesized.

The effect on timbre under free-field conditions was investigated in [3]. In [4], numerical simulations showed that the listening room (the room in which the loudspeaker system is installed) has the potential of evening out alterations of the magnitude spectrum of the synthesized sound field that occur due to spatial aliasing.

Simulations on the interaction of loudspeaker array and listening room based on the image source method in [5] suggested that in the present situation – a room-in-room presentation (a virtual room created inside a real room) – the energy decay of the evolving sound field is indeed altered compared to the virtual room.

Previous attempts show only partial success in measuring the presence of spatial aliasing as well as its effect instrumentally. It was found that the interaural cross-correlation (IACC) is indeed lower for a synthetic wave front compared to a single loudspeaker [2]. Measuring the reflection (or wave front) density based on the method from [6] was not successful [7].

In the present work, we present a user study that investigates the effect of spatial aliasing as well as the effect of the listening room on the perceived spaciousness when virtual rooms are synthesized using WFS.

All data in this work were processed with/based on the Sound Field Synthesis Toolbox [8].

Virtual reverberation in wave field synthesis

For convenience, we assume in this work that reverberation is composed of discrete early reflections that impinge from various directions, and which become gradually denser in time. Finally, after the so-called mixing time, late reverberation is apparent which exhibits an approximately exponential decay. Refer also to Fig. 2.

A first outline of the process of creating artificial reverberation for WFS (or large-scale loudspeaker arrays in general) can be found in [9] where a two-stage implementation is described. Early reflections are generated using a mirror image model and late reverberation is generated using signals with appropriate statistical parameters. A perceptual evaluation of the complete system is not available.

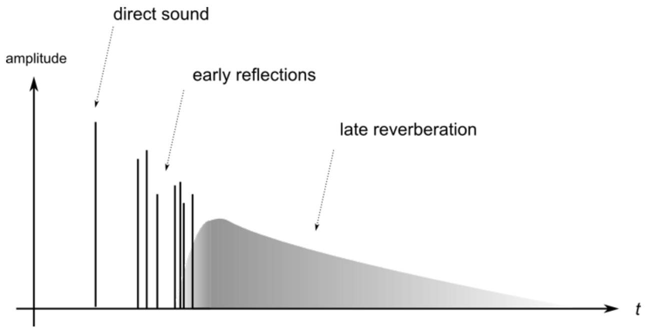

Figure 2: The model of reverberation that is assumed in this paper

Stimulus creation

It becomes evident from the sparse literature on creating reverberation in WFS that no standard approach exists neither for the rendering of acoustical room models nor for other means of creating artificial reverberation. In order to minimize the uncertainties, we work based on the binaural impulse responses of real rooms in this paper. The stimulus creation comprised the following steps:

1) Measure the binaural room impulse responses of suitable rooms for different head orientations; these data will serve as the ground truth in the user study
2) Create a parametric model of the measured room impulse responses by manual tuning of a set of filtered head-related impulse responses (HRIRs) for the early reflections and filtered Gaussian noise for the late reverberation; successful completion of this step assures that plausible parameters are being used for the reverberation
3) Measure the binaural room impulse responses (BRIRs) of each of the loudspeakers of a WFS system to different head orientations; this allows for simulating the loudspeaker system over headphones with head tracking; we measured the BRIRs of the WFS system that is installed at the Quality and Usability Lab at Technische Universität Berlin for different head orientations in steps of 1°. The parameters of the WFS system are identical to those in Fig. 1.
4) Play back the parametric room model with the headphone simulation of the WFS system

A perceptual comparison of the data from 1) and 2) allows for validating the parametric model, which is the prerequisite for investigating the spaciousness when the parametric model is rendered in WFS. A perceptual comparison of the data from 2) and 4) allows for determining the perceptual impact of the WFS system. The impact of the WFS system can be two-fold:

Firstly, the system produces spatial aliasing, which may have an effect on the perceived spaciousness because of the additional wave fronts that occur. Secondly, WFS systems do typically not operate under free-field conditions. It has been noted in [4,5,7], that the impact of the listening room may be considerable. In order to separate the effect of spatial aliasing from the effect of the listening room, we identified the first reflection after the direct sound in the BRIRs of the loudspeaker system and faded out the BRIRs right before this reflection in order to create virtual free-field conditions. See Fig. 3 for an illustration.

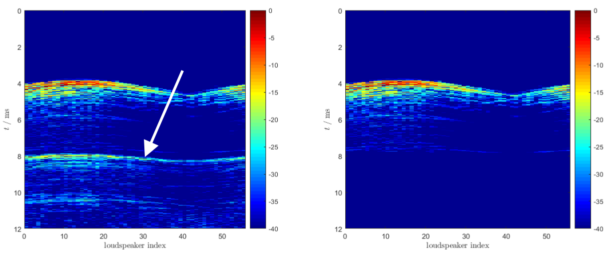

Figure 3: Sample BRIRs of the 56-channel circular array installed at the Quality and Usability Lab at Technische Universität Berlin for a given head orientation including the room response (left) and with the room response suppressed (right); the first occurring room reflection is marked by the white arrow

Method

3 female and 9 males subjects were recruited. All stimuli were presented over headphones with head tracking applied. The processing was performed by the BRS renderer of the SoundScape Renderer (SSR) [10]. A 2-minute long loop of male speech was playing continuously, and seamless switching between different stimuli was implemented by muting/un-muting the according virtual sources in SSR. This creates a smooth crossfade between the signals without audible artifacts.

The subjects used a computer mouse to operate the graphical interface. The paradigm was as follows: The subjects were presented with triads of stimuli. The task was to rank the stimuli of a given triad according to the perceived amount of spaciousness. A stimulus was defined to be more spacious than another one if:

- The reverberation was stronger
- The room sounded larger
- The sound source appeared farther away
- The sound source appeared larger

The ranking was performed based on a dedicated set of buttons as can be deduced from Fig. 4. The subjects were only able to move forward to the next triad if all 3 stimuli of the current triad were reflected in the ranking. They were instructed to give an arbitrary response if they did not perceive a difference in the spaciousness or if they were perceiving contradictory cues as per the above list.

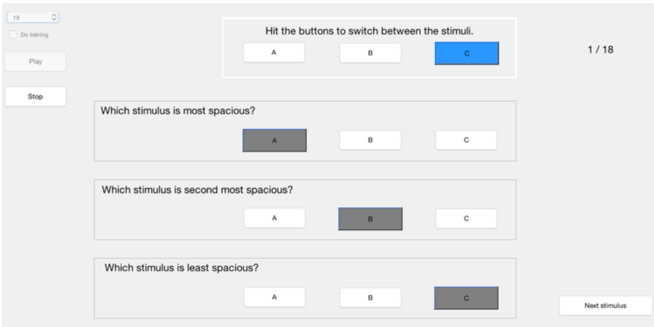

Figure 4: Screenshot of the user interface

After written instructions, the subjects completed a training of 4 triads followed by 18 triads of experiment. The complete process took about 20 minutes.

The experiment consisted of 9 different triads each of which occurred twice in the experiment. The order of triads was randomized, so was the assignment of stimuli to buttons. The triads were conceptually grouped into 3 different sets:

- Set 1: 4 triads of different combinations of a single loudspeaker of the system and a virtual plane wave synthesized by the entire system (recall Fig. 1) both under free-field conditions as well as including the listening room; these stimuli were chosen for being able to identify the relative strength of the effect of spatial aliasing compared to that of the listening room;
- Set 2: 3 triads containing combinations of the BRIRs of a small and dry laboratory room ("Calypso", $t_{60} = 200$ ms; see [7] for more details; these data are the ground truth), its parametric binaural model (to validate the model), and the parametric model rendered in WFS in a free-field as well as in a room with and without manual thinning out of the virtual reflection pattern; it was suggested in [7] that it might be favorable to render fewer virtual reflections than required as spatial aliasing adds wave fronts;
- Set 3: 2 triads similar to the Calypso data but for a mid-size meeting room ("Sputnik", $t_{60} = 500$ ms; see [7] for more details)

Results and Discussion

Tab. 1 presents the results for stimulus set 1: Remarkably, the ranking according to spaciousness given in Tab. 1 is reflected by all subjects' responses. This is particularly remarkable because the differences between the stimuli were rather subtle.

Two important conclusions may be drawn from the data in Tab. 1: Spatial aliasing does indeed increase the perceived spaciousness (WFS-FF is considered more spacious than iLS-FF). The effect of the listening room is stronger than the effect of spatial aliasing (iLS-R is considered more spacious than WFS-FF).

Table 1: One individual loudspeaker (iLS) vs. a virtual plane wave (WFS) in a free-field (FF) or in the room (R); consistency is the percentage of corresponding rankings performed by the subjects

Rank	Condition	Consistency
1	WFS-R	100 %
2	iLS-R	100 %
3	WFS-FF	100 %
4	iLS-FF	-

Tab. 2 presents the results for stimulus set 2: The only 100-%-consistent observation is that the WFS system in the listening room is considered more spacious than everything else. This suggests a strong effect of the listening room. The circumstance that there are discrepancies in the subjects' rankings for the other stimuli suggests that the differences are very small or that there are contradictory cues. We may cautiously conclude that the parametric model for Calypso's reverberation is validated and that the parametric model rendered in WFS-FF evokes a spaciousness that is comparable to the original measured data. This suggests that the effect of spatial aliasing is not considerable when it comes to such scenarios that are more plausible than the single synthetic wave front in stimulus set 1.

The effect of the thinning out of the parametric reverberation is unclear at this point. Subjects reported that the resulting reverberation sounded impaired so that is approach does not seem to be relevant.

Table 2: Calypso data (*synth.* represents the parametric binaural model; *meas.* is the measured BRIRs)

This …	…is considered more spacious than this.	Consistency
Synth. BRIR	Meas. BRIR	75 %
WFS-FF	Meas. BRIR	30 %
WFS-R	*Everything else*	100 %
WFS-FF (thinned)	WFS-FF	67 %
WFS-R (thinned)	WFS-FF	96 %

Finally, Tab. 3 presents the results for stimulus set 3: All stimuli seem to exhibit a very comparable amount of spaciousness. This suggests that the effect of the listening room – which was very strong for Calypso – vanishes when more reverberant virtual rooms are rendered in WFS.

Table 3: Sputnik data

This …	…is considered more spacious than this.	Consistency
WFS-FF	Synth. BRIR	71 %
WFS-R	WFS-FF	79 %
WFS-R (thinned)	WFS-FF	12 %
WFS-R	WFS-R (thinned)	67 %

We had performed informal pilot studies with virtual rooms that are more reverberant than the ones employed in the formal study. Our observation was that it is becoming progressively more difficult to identify differences in the spaciousness of the different conditions.

The increase of the challenge of ranking the stimuli with stronger virtual reverberation is also supported by the observation that the average time that it took the subjects in the formal study to establish a ranking is lower for the conditions from Tab. 1 and 2 (32 s, 33 s) than for the conditions from Tab. 3 (43 s).

Summary

We found that the chosen test paradigm produces very consistent responses even for stimuli with rather small differences. Spatial aliasing in WFS does increase the perceived spaciousness in that it increases the perceived source width (recall [2]) under specific and rather artificial situations. This effect disappears as soon as virtual reverberation is rendered.

The listening room produces a measureable increase of the perceived spaciousness when very dry virtual rooms are rendered. This effect disappears as soon as the virtual reverberation is moderate or stronger than that.

Literatur

[1] J. Ahrens. *Analytic Methods of Sound Field Synthesis.* Springer-Verlag, 2012.

[2] E. W. Start, "*Direct sound enhancement by wave field synthesis*," PhD thesis, Delft University of Technology, 1997.

[3] H. Wierstorf, "*Perceptual Assessment of Sound Field Synthesis*," PhD thesis, University of Technology Berlin, 2004.

[4] V. Erbes und S. Spors, "Influence of the Listening Room on Spectral Properties of Wave Field Synthesis," in *DAGA*, Kiel, Germany, 2017.

[5] V. Erbes, S. Spors, and S. Weinzierl, "Analysis of a spatially discrete sound field synthesis array in a reflective environment," in *Euronoise*, Maastricht, The Netherlands, May/June 2015.

[6] J. Abel and P. Huang, "Robust Measure of Reverberation Echo Density," in *121st Convention of the AES*, San Francisco, CA, Oct. 2006.

[7] J. Ahrens, "On the Generation of Virtual Early Reflections in Wave Field Synthesis," in *DAGA*, Aachen, Germany, Mar. 2016.

[8] H. Wierstorf, S. Spors, "Sound Field Synthesis Toolbox," in *Proceedings of the 132nd Convention of the Audio Engineering Society*, Budapest, Hungary, 2012.

[9] D. de Vries, A. J. Reijnen, and M. A. Schonewille. The wave field synthesis concept applied to generation of reflections and reverberation. In *96th Convention of the AES*, Amsterdam, The Netherlands, Feb./Mar. 1994.

[10]M. Geier, S. Spors, J. Ahrens, "The SoundScape Renderer: A Unified Spatial Audio Reproduction Framework for Arbitrary Rendering Methods," in *124th Conv. of the AES*, Amsterdam, The Netherlands, May 17–20, 2008.