Signal Processing for Spatial Sound Reproduction with Wave Field Synthesis

Rudolf Rabenstein

University Erlangen-Nuremberg
Erlangen, Germany
`rabe@LNT.de`

Sascha Spors

Deutsche Telekom Laboratories
Berlin, Germany
`Sascha.Spors@telekom.de`

Conventional systems for the reproduction of spatial audio are mainly based on intensity panning techniques. They adjust the contributions from the different loudspeaker channels in such a way that their superposition produces the required intensity levels at the listeners ears. Consequently, the spatial reproduction quality is only guaranteed in the vicinity of the targeted listener position.

To enlarge the preferable listening area, several novel audio reproduction techniques have been suggested. They can be roughly categorized into advanced panning techniques, Ambisonics systems, and wave field synthesis. Advanced panning techniques enlarge the listening area by panning between an increased number of loudspeakers. Ambisonic systems represent the sound field in an enclosure by an expansion into low-order three-dimensional basis functions of the acoustic wave equation. Wave field synthesis techniques are formulated in terms of the acoustic wave equation and the description of its solutions by Greens functions. They use loudspeaker array technology to reproduce sound fields in enclosures.

The digital signal processing tasks for these spatial audio reproduction techniques have different levels of complexity. Two-channel stereo requires basically no signal processing on the reproduction side and performs well even with purely analog equipment. The established standards for five and seven channel stereophony call for digital storage of coded signals for each channel, e.g. on DVD. The reproduction equipment must provide the capability for decoding of multi-channel audio in real time. After equalization and amplification, the decoded signals are suitable for driving the loudspeakers.

However, for the novel audio reproduction techniques described above there is no one-to-one correspondence between audio channel and loudspeaker driving signal. This is especially true for wave field synthesis with typically some ten to a few hundred loudspeakers. Here, a different approach has to be taken: The audio channels for transmission and storage contain the different tracks or voices of the sound objects. The composition of these voices to the desired audio scene requires advanced digital signal processing steps which can be formulated as a multiple-input, multiple-output system.

This contribution gives an overview on advanced spatial audio rendering systems with a special emphasis on wave field synthesis. The essential steps from the physical description in terms of the acoustic wave equation and the Kirchhoff-Helmholtz-integral to the computation of the loudspeaker driving signals are presented. Finally, examples of implemented spatial audio rendering systems are shown.

An implementation of a wave field synthesis system with a circular loudspeaker array is shown below. A total of 48 two-way loudspeakers are mounted on a circle with 3 m diameter and with a spacing of about 20 cm. The driving signals for the loudspeakers are computed by fast convolution techniques in real-time on a personal computer. The system described here is located at the Telecommunications Laboratory (Multimedia Communications and Signal Processing) of the University of Erlangen-Nuremberg in Germany (`http://www.LNT.de/LMS`).