

Editorial

Spatial Sound and Virtual Acoustics

Ville Pulkki,¹ Christof Faller,² Aki Härmä,³ Tapio Lokki,⁴ and Werner de Bruijn³

¹Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology, P.O. Box 3000, 02015 TKK, Finland

²Laboratoire de Communications Audiovisuelles 1, Institut de Systèmes de Communication, Faculté Informatique et Communications, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

³Digital Signal Processing Group, Philips Research Labs, 5656 Eindhoven, The Netherlands

⁴Telecommunications Software and Multimedia Laboratory, Helsinki University of Technology, P.O. Box 5400, 02015 TKK, Finland

Received 22 February 2007; Accepted 22 February 2007

Copyright © 2007 Ville Pulkki et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Spatial sound is a field which investigates the influence of room or whatever acoustical environment to the presented sound. It gathers different techniques to capture and (re-)produce sound in a space for different purposes, and all issues concerning how humans perceive different attributes of sound which change depending on spatial characteristics of source, space, and listener.

The audio technology has evolved to such level that many monaural characteristics of sound, such as frequency spectrum and temporal structure, can be reproduced with current microphones and loudspeakers with such good quality that a human observer cannot perceive difference between original and reproduction. However, many spatial characteristics cannot be reproduced transparently or be synthesized with perfect quality. The most important spatial characteristics are the direction and the distance of sound sources and the attributes of the room perceivable by humans. In this issue, some papers present new results in this field.

A topic in technologies related to spatial sound has been of interest to some of the authors of this issue. When sampling spatial sound field using more than one microphone, more information can be extracted from differences between microphone channels. For example, capturing the sound with narrow spatial selectivity and estimation of direction of sound are common topics in this research.

Virtual acoustics can be seen as a subtopic of spatial sound, and also as a subtopic of traditional acoustics. In it, the acoustics of an imaginary or real space is modeled computationally, and this model is further used to make the acoustics of the modeled room audible to a human listener. In interactive virtual acoustics, the listener can even change

his listening position and/or the modeled sound sources may change their position during listening.

The topics of the papers in this special issue can be grouped roughly into five categories, which are presented next.

Simulation and modeling of room acoustics

One of the research topics relevant to spatial audio is the estimation of properties of the room. The paper by M. Kuster and D. de Vries proposes a method for the acoustic imaging of the shape of a wall based on acoustic transfer functions measured between a source and a large number of microphone positions close to the wall.

M. Karjalainen and T. Paatero propose a novel method to tackle a well-known equalization problem in sound reproduction. In their paper “Equalization of loudspeaker and room responses using Kautz filters: direct least squares design,” they equalize combined loudspeaker-room responses with Kautz filters which allow the allocation of frequency resolution freely. The method is evaluated with several interesting case studies.

Beamforming and sound source localization

The ultimate reference for spatial audio is the desired sound field in an extended listening area. The plane wave decomposition is an efficient analysis tool for multidimensional fields, particularly well fitted to the description of sound fields. The paper by M. Guillaume and Y. Grenier describes a method to estimate the plane wave decomposition of sound fields by means of beamforming.

Spatial audio for communication applications

One important area for spatial audio is speech communications. The capture of speech in car environment is a challenging topic due to the high noise levels. The basic methods for speech capture and enhancement require the positioning of the talkers. Jwu-Sheng Hu et al. introduce an algorithm to find the locations of active talkers using a Gaussian mixture model.

The paper by Shigeki Miyabe et al. studies the problem of echo cancellation in spoken dialog systems where the double-talk (or barge-in) is even a more severe problem than in human-to-human communications. Their method is based on spatial cancellation of the sound field at microphone positions using the reproduction with an array of loudspeakers instead of using more common approach of cancelling the unwanted from the microphone signal.

Spatial sound recording and reproduction

In the paper “3D-audio matting, postediting, and re-rendering from field recordings,” Emmanuel Gallo et al. present an approach towards the segmentation of field recordings made with, to some extent, arbitrary spatial distributions of multiple microphones into individual auditory components with corresponding source locations. They show that within certain assumptions, analysis and storage of a spatial sound field in such a way can be highly efficient and enable reconstruction of the original spatial sound field, or if desired manipulation of the original spatial source distribution, over a wide range of loudspeaker setups.

The paper “Extraction of 3D information from circular array measurements for auralization with wave field synthesis” by Diemer de Vries et al. presents two methods to extend the reproduction of room impulse responses measured with horizontal microphone arrays over reproduction systems capable to reproduce 3D sound scene. The focus is in wave field synthesis, which is a technique which uses a large number of loudspeakers and aims to reproduce the sound field correctly.

Virtual acoustics

The paper “Virtual reality system with integrated sound field simulation and reproduction” by Tobias Lentz et al. presents an integrated real-time audio rendering system which contains both room acoustics modeling software and binaural reproduction tools for headphone-free reproduction. The main contribution of the paper is to present the fluent interaction of all involved subsystems. The presented system is one of the first complete sound rendering systems for virtual reality applications.

For a convincing interactive simulation of the acoustics of a complex environment, it is important to include the contribution of edge diffractions to the total sound field, in addition to the specular reflections that can be modeled

in a relatively simple way. In the paper “Fast time-domain edge-diffraction calculations for interactive acoustic simulations,” P. T. Calamia and U. P. Svensson present a computationally efficient approach towards inclusion of these edge-diffraction effects using an edge-subdivision strategy, which offers a tradeoff between computation time and accuracy and enables implementation in interactive simulation applications.

A specific topic inside virtual reality, the synthesis of the directivity of sound sources using a wave field synthesis, is presented by E. Corteel in his paper “Synthesis of directional sources using wave field synthesis, possibilities, and limitations.” Corteel presents a theory how to take the directivities into account using spherical harmonics, and also presents solutions how to overcome the artifacts generated by wave field synthesis.

*Ville Pulkki
Christof Faller
Aki Härmä
Tapio Lokki
Werner de Bruijn*

Ville Pulkki received the M.S. and D.S. (Tech.) degrees from Helsinki University of Technology in 1994 and 2001, respectively. He majored in acoustics, audio signal processing, and information sciences. Between 1994 and 1997, he was a Full-Time Student in the Department of Musical Education at the Sibelius Academy. In his Doctoral dissertation, he developed vector base amplitude panning (VBAP), which is a method to position virtual sources to any loudspeaker configuration, and studied its performance with psychoacoustic listening tests and with modeling of auditory localization mechanisms. The VBAP method is widely used in multichannel virtual auditory environments and in computer music installations. His research activities cover methods to reproduce spatial audio and methods to evaluate quality of spatial audio reproduction. He has also worked on diffraction modeling in interactive models of room acoustics. He enjoys being with his family (wife and two children), playing various musical instruments, and singing.



Christof Faller received an M.S. (Ing.) degree in electrical engineering from ETH Zurich, Switzerland, in 2000, and a Ph.D. degree for his work on parametric multichannel audio coding from EPFL Lausanne, Switzerland, in 2004. From 2000 to 2004, he worked in the Speech and Acoustics Research Department at Bell Laboratories, Lucent Technologies and Agere Systems (a Lucent Company), where he worked on audio coding for digital satellite radio, including parametric multichannel audio coding. He is currently a Part-Time Postdoctoral Employee at EPFL Lausanne. In 2006, he founded Illusonic LLC, an audio and acoustics research company. He has won a number of awards for his contributions to spatial audio coding, MP3 surround, and MPEG surround. His main current research interests are spatial hearing and spatial sound capture, processing, and reproduction.



Aki Härmä was born in Oulu, Finland, in 1969. He received the Master's and Doctor's degrees in electrical engineering from the Helsinki University of Technology, Espoo, Finland, in 1997 and 2001, respectively. In 2000-2001, he was a Consultant at Lucent Bell Laboratories and the Media Signal Processing Research Department of Agere Systems, Murray Hill, NJ. In 2001, he returned to the Laboratory of Acoustics and Audio Signal Processing, Helsinki University of Technology (TKK), Finland. Since 2004, he has been with the Digital Signal Processing Group of Philips Research, Eindhoven, The Netherlands. His research interests are mainly in social presence technology, spatial sound, and acoustic signal processing.



Tapio Lokki was born in Helsinki, Finland, in 1971. He has studied acoustics, audio signal processing, and computer science at the Helsinki University of Technology (TKK) and received an M.S. degree in electrical engineering in 1997 and a D.S. (Tech.) degree in computer science and engineering in 2002. At present, he is a Lecturing Researcher with the Telecommunications Software and Multimedia Laboratory at TKK. In addition, he is an Adjunct Professor at Laboratory of Acoustics and Audio Signal Processing at TKK. His research activities include 3D sound, room acoustics, virtual acoustic environments, auralization, and virtual reality. He is a Member of the Audio Engineering Society, the IEEE Computer Society, and the Acoustical Society of Finland.



Werner de Bruijn was born in Tilburg, The Netherlands, in 1973. In 1998, he received the M.S. degree in applied physics from Delft University of Technology. His graduation work was carried out in the Laboratory of Acoustical Imaging and Sound Control, from which the concept of wavefield synthesis (WFS) originates, and was concerned with the investigation of techniques for the recording and reproduction of reverberation and reflections for WFS and other multichannel sound reproduction systems. In 2004, he received the Ph.D. degree in applied physics for work in the same group, which concerned the application of WFS in life-size videoconferencing, with a focus on investigation of the audio-visual interaction effects that occur in such systems that combine 2D video with audio that includes a realistic reproduction of depth. Since 2003, he has been working as a Senior Scientist in the field of acoustics and sound reproduction in the Digital Signal Processing Group of Philips Research, Eindhoven, The Netherlands, where his main research focus is on loudspeaker arrays.

