AES 28th International Conference

June 30 – July 2, 2006
Piteå, Sweden

Jan Berg
conference chair

Francis Rumsey
papers chair

Neil Gilchrist
AES president

Roger Furness
AES executive director

The Future of Audio Technology–
Surround and Beyond

Regular AES conference-watchers might be forgiven for wondering what it is about the Arctic Circle in Scandinavia that seems to attract the Society to it. The 19th International Conference explored Father Christmas territory in Rovaniemi, Finland, bang on the arctic circle in deep-freeze conditions during the first part of 1999, and the 30th Conference will explore even more northerly lands when it takes place in March 2007 in Saariselkä, Finland, north of the Arctic Circle. Preferring 24-hour sunlight to the bracing Scandanavian winters, the 28th International Conference, *The Future of Audio Technology—Surround and Beyond*, was held recently in Piteå, Sweden, just 200 km south of the Arctic Circle. Surprisingly warm weather, with bright sunshine and an almost Mediterranean climate, greeted those who had ventured north to the area known as the Swedish Riviera.

The School of Music, an outpost of Luleå University of Technology in the small town of Piteå, is a remarkable music and media education center that combines a traditional (and not-so-traditional) music conservatory with just about every facet of modern media technology. This conference, chaired by Jan Berg, a member of the School of Music academic staff, was an opportunity to share the latest research in all things to do with spatial audio. Berg and his conference committee, together with a support crew of students and staff, had put together a three-day program of papers, tutorials, discussions, and demonstrations that covered the spatial audio signal chain all the way from source to perception.

DAY ONE

The first day of the conference kicked off with a two-paper session emphasizing recording techniques, chaired by Diemer de Vries. The first paper, by Shintaro Hosoi and his colleagues from Pioneer, dealt with the possibilities for bass cancellation between main channels and LFE channel in 5.1 surround, both at the recording and reproduction ends of the signal chain. By means of careful processing they showed that it is possible to reduce or eliminate this

problem in reproduction systems provided that care is taken with the recording chain. A Pioneer demo that ran throughout the conference provided delegates an opportunity to listen to the effects of such corrections. Toru Kamekawa of the Tokyo University of Fine Arts and Music, then proceeded in the next presentation to describe changes in spatial impression resulting from different configurations of three front microphones in multichannel stereo recording.

During the remainder of the first morning, conference delegates could choose between attending a tutorial on basic psychoacoustics of 5.1 surround recording, given by Francis Rumsey, or attending some of the demonstrations. Rumsey explained how the concepts of 2-channel stereo had been extended to 5-channel surround. He warned that users should be aware of the phantom imaging limitations of the format, such as difficulties in side panning and rear imaging. Loudspeakers in the standard locations used for 5-channel surround, however, have been shown to be adequate for delivering the impression of a diffuse soundfield.

The afternoon session of the first day, chaired by Ville Pulkki, provided an opportunity to hear about spatial soundfield analysis and synthesis. In the first presentation in the session, Gregory Pallone of France Telecom described a spatialized soundfield environment synthesizer. Next, Hyunjoo Chung of Seoul National University in Korea described a grouped-reflections algorithm for soundfield processing in home theater systems. His idea was that since accurate localization of, for example, side reflection images in 5-channel surround was not possible, such reflections could be analyzed and grouped into locations around the nearest loudspeaker. Chung's colleague Hwan Shim followed by presenting a paper dealing with the design of an artificial reverberator to control the location of sound sources for surround audio. The session concluded with a presentation by Mathieu Guillaume of Telecom Paris, concerned with the use of a microphone array for experimental 3-D soundfield analysis.

Integration of sound and picture topped off the papers menu on Friday with four presentations on different aspects of the subject; this session was also chaired also by Ville ➡

Authors

Terence Caulkins

Marie-Josefin Meindl

Ulrich Reiter

Sascha Spors

Kimio Hamasaki

Pulkki. Kimio Hamasaki of NHK showed how the presence of a picture, when using NHK's super-high-definition television system with 22.2-channel sound, reduced some perceived aspects of surround sound perception, particularly that of naturalness. Lower-resolution sound systems were more tolerant of the presence of a picture. The use of wavefield synthesis in conjunction with augmented reality systems was the main theme of a presentation by Frank Melchior of Fraunhofer IDMT. The last two presentations in this session were given by Ulrich Reiter of the Technische Universität Ilmenau. Reiter dealt first with the subjective assessment of the optimum number of loudspeaker channels in interactive audiovisual applications using large screens, and subsequently with some modifications that are proposed for the MPEG-4 AABIFS perceptual approach. This approach to describing spatial audio scenes uses high-level descriptors rather than physical models or parameters to describe the spatialization of sound sources, and this paper looked at it in the context of interactive audiovisual application systems.

No more than an hour from Piteå, up the Piteälven (the Piteå River), are to be found some of the most impressive waterfalls in Europe—the Storforsen. The committee had planned an excursion here on the first evening of the conference, providing a powerful experience of the natural resources of northern Sweden. Millions of gallons of water per second poured down from the Norwegian mountains and the noise of this roaring tumult was likened by many audio engineers to a giant pink noise generator. Much of the power for the region is generated up here. This is a place where local people like to come for picnics and barbeques in the summer; there are barbeque pits with benches among the rocks of the original course of the waterfall where fires can be built for outdoor cooking. A delightful buffet dinner in a large banquet hall close to the river rounded off a memorable evening.

DAY TWO

Day 2 was largely dedicated to the theme of spatial rendering techniques, demonstrating a wide range of methods for generating spatial sound. The first session, chaired by Anders Ågren, was strongly concerned with issues relating to wavefield synthesis (WFS), starting with a paper from Terence Caulkins and his colleagues from IRCAM and Trinnov Audio. In this presentation Caulkins described how a high-resolution microphone array using directional beams formed by means of signal processing could be used to characterize the early reflections generated by a WFS loudspeaker array. Diemer de Vries, one of the original developers of wavefield synthesis, also described the further optimization of multiactuator panel loudspeakers for reproducing WFS signals, showing one possible solution to the problem of how to implement this technology in a practical fashion. Sascha Spors of Deutsche Telekom contributed to this session by presenting a paper on the evaluation of circular harmonic decomposition for WDAF-based based listening room compensation; WDAF stands for wave domain adaptive filtering. James Hall described a process for recording and replaying the sound from acoustic spaces using variations on a novel microphone array originally proposed by James "JJ" Johnston. Concluding the morning session, Damian Murphy presented a poster promoting the work of a new UK research network for spatial audio known as SpACE-Net, the Spatial Audio Creative Engineering Network, which is designed to bring together composers and researchers in the field.

An impromptu organ recital by Francis Rumsey, given at lunchtime on Saturday, provided delegates with a chance to hear one of the lovely neoclassical organs installed at the School of Music, where organ performance and church music are an important study topic. As Rumsey explained, the organ is a highly spatial instrument and the "Dorian"

Fakheredine Keyrouz

Sungyoung Kim

Ville Pulkki

Time followed during which delegates could sample the demonstrations offered in adjacent rooms, which included one of Directional Audio Coding (DirAC), given by Ville Pulkki of the Helsinki University of Technology. Pulkki's system enables the decomposition and reconstruction of directional sound fields, separating the direct and diffuse components so as to enable their separate enhancement and rendering. In a loudspeaker array that included a height component, Pulkki showed how his novel coding approach compared with unprocessed signals, and it was possible to hear a striking change in openness of the reverberation, depth of soundfield, and image stability when moving around or sitting off center. Further demonstrations were available including both AAC Plus and MPEG Surround from Coding Technologies and Livesound's novel loudspeaker with an extremely long line-source ribbon driver for 5.1 reproduction in high-quality live sound applications. Pioneer's system for eliminating bass cancellation was also being shown, as mentioned above.

The area around Piteå is a popular holiday locale with Swedes and also with Norwegians who come down from the mountains for the sun and sandy beaches on the coast. At the Piteå Havsbad, a seaside resort complex out of the town where many of the conference delegates stayed, a buffet banquet was served on Saturday evening, preceded for those in the mood by some gentle swimming or a spell in the steam room. Many stayed late into the night, taking advantage of the midnight sun that is so striking in this area of northern Sweden.

DAY THREE

First thing on Sunday morning, Ville Pulkki gave a paper on his directional audio coding system, described above, in a session chaired by Francis Rumsey. He also explained how this might be used for various forms of upmixing. This led into two sessions, taking up the remainder of the day, concerned with the perception of spatial sound—the ultimate finishing point of the signal chain. Chaired by Russell Mason, the first session opened with a presentation by Tim Brookes of the University of Surrey dealing with a means of controlling the perceived spatial orientation of a reproduced human voice source. Brookes explained how different acoustical factors change as a human talker rotates her head. He described an experiment designed to discover which of these factors is influential on the perception during reproduction. He found that in some cases listeners could tell better what direction the talker was facing when using synthetic cues than when using real stereo recordings of a human talker, but this was thought to be due to the relatively crude simulations adopted.

Håkan Ekman of the host organization, Luleå Univer- ➡

Toccata and Fugue by J. S. Bach he played showed off the spatial distribution of the divisions and pipes, both in the height and the width dimensions. He finished with Bach's Prelude and Fugue in G major to round off this entertaining interlude in the proceedings.

Part 2 of the session on spatial rendering techniques, chaired by Kimio Hamasaki, included two papers on head-related stereophony. The first, by Fakheridine Keyrouz of the Technische Universität München, was concerned with interpolation of head-related transfer functions, while the other, by Florian Koenig of Ultrasone, involved a discussion of some issues affecting the design of surround sound headphones. Alex Southern of the University of York showed how a room-modeling application known as the digital waveguide mesh, based on a similar physical model to that used in string synthesis, can be combined with a form of spatial encoding with some success. Some conventional measures of spatial performance such as lateral fraction and lateral gain were shown to work in such a model, suggesting it was performing in a predictable way. Yoshnori Takehashi rounded off this session with a presentation on rendering spatial reverberation and the evaluation of perceived source distance.

Toward the end of the afternoon, a short session on spatial audio coding, also chaired by Kimio Hamasaki, provided an opportunity for Jürgen Herre to explain the principles of the forthcoming ISO standard for spatial audio coding, known as MPEG Surround. This format enables surround sound to be encoded in the form of backward-compatible main channels plus additional information representing interchannel differences in level and coherence that enable the surround information to be reconstructed. Concluding the afternoon, Nikolaus Färber of Fraunhofer IIS talked about Internet Radio Surround in a particular implementation involving MPEG Surround delivered over ISMA Ultravox.

Francis Rumsey gave an impromptu organ concert during the lunch break on Saturday.

28th Conference Committee: from left, Francis Rumsey, papers chair; Hans Vesterberg, conference vice chair; Jan Berg, conference chair; Lars Hallberg, facilities chair; and Anders, Ågren, papers vice chair.

Theater-style seating in the lecture room afforded good visibility to all.

Panel discussing the future of enhanced spatial audio: from left, moderator Francis Rumsey, Florian Camerer, Jeff Levison, Kimio Hamasaki, and Søren Bech.

Based on the high quality of the papers in the conference proceedings, Søren Bech, chair of the AES Publications Policy Committee, recommended that authors should consider updating and revising their papers for submission to the *Journal.*

sity of Technology, discussed experiments he had carried out into the perceived distances to recorded musical instruments. He compared the responses of musicians and sound engineers and found that, in general, sound engineers perceived instruments closer than musicians by about 0.7 m, but musicians preferred a greater distance to sources than sound engineers. This may have something to do with the way sound engineers like to listen analytically to sound balances. In a similar vein, Marie-Josefin Meindl, of the University of Music and Performing Arts in Vienna, described the effects of surround microphone setup on room perception. She showed how different left-right and front-back spacings, as well as alternative polar patterns, of a pair of surround microphones could change the perceived room size and envelopment, for example.

Following the lunch break, Tim Brookes chaired the next session, with the first two presentations given by Sunyoung Kim from McGill University, Montreal. Kim first examined the effect of presentation order on the preference choices of listeners for different surround recordings of a piano. When

different piano pieces were intermixed during the comparisons of microphone technique, preferences were affected by musical selection order, but when the musical items were presented as a block it enabled comparisons between techniques to be made more consistently. In his second presentation he described how the preferred microphone technique is quite strongly affected by musical selection.

The papers session on the final afternoon concluded with a presentation from Russell Mason of the University of Surrey. He introduced delegates to the principles of a binaural model he had developed, which was designed to evaluate perceived width of sources or a reverberant environment in an objective fashion. He showed both animated and static plot outputs from the model, in which it is possible to observe changes in the location and width of signals at different frequencies.

Wrapping up the 28th Conference were two panel discussions, the first was chaired by Nyssim Lefford of the Interactive Institute. The panelists—Johnny Wingstedt, a composer associated with the Interactive Institute, Gregory Pallone of ➡

where audio comes alive!

MOSCONE CONVENTION CENTER
SAN FRANCISCO CALIFORNIA USA

121st
AES
Convention

exhibits: October 6-8, 2006
conference: October 5-8, 2006

AUDIO A ES®
Audio
Engineering
Society

master classes • training sessions
live sound seminars • live sound symposium
broadcast events • tours • exhibits • exhibitor seminars
technical papers • educational events • workshops

28th AES International Conference

On Saturday evening attendees enjoyed a banquet at seaside resort Piteå Havsbad (below) with music provided by students (left) from the School of Music.

Student volunteers—from left, Håkan Ekman (coordinator), Jens Andersson, Markus Hammerin, Tyra von Sydow, and Pontus Svensson—provided support to the conference committee. Technical support group members, Ulf Olofsson and Tommy Lundström, missed this photo; they provided behind-the-scenes preparations of the technical facilities.

Demo rooms ran throughout the conference during coffee and lunch breaks: clockwise from above, Coding Technologies' was on aacPlus and MPEG Surround; Ville Pulkki's looked at directional audio coding; and Pioneer's was on bass cancellation. Livesound also had a demo on surround with ribbon line source speakers.

Friday evening attendees went on an excursion to Storforsen, a nature preserve with these spectacular rapids on the Piteå River.

France Telecom, Jürgen Herre of Fraunhofer, and Linus Feldt, a developer of innovative computer games for children—gave their views about ways in which sound, interactivity, and games could work together in the future. They agreed that games are a major growth area and that the development of immersive sound and adaptive music that can stimulate emotional responses to scenes and characters is going to have a significant effect on the market.

A final panel discussion led by Francis Rumsey included Søren Bech of Bang & Olufsen, Florian Camerer of ORF, Jeff Levison of DTS, and Kimio Hamasaki of NHK. Rumsey asked them what the most important application areas for enhanced spatial audio would be, and how they would spend a substantial research budget in this field. Some felt that personal spatial audio is most important, particularly home environments that take into account the fact that people move around and between rooms and might want sound to follow them. Others were keen to ensure the most accurate and lifelike representation of musical concerts, while there was also a plea for concentrating on learning how to use current systems in the most appropriate way. Even if classical recording is not a major sector of the market, other fields can learn a lot from discoveries in this area.

The conference closed with thanks from AES President Neil Gilchrist to the entire conference committee: Jan Berg, chair; Hans Vesterberg, vice chair; Lars Hallberg, facilities chair; Francis Rumsey, papers chair; Anders Ågren, papers vice chair; and the team of student technical assistants led by Håkan Ekman. The committee delivered on the promise of the conference theme: *The Future of Audio Technology—Surround and Beyond*, making this the third successful AES conference devoted to surround. The conference proceedings and CD-ROM can be purchased online at www.aes.org/publications/conf.cfm.
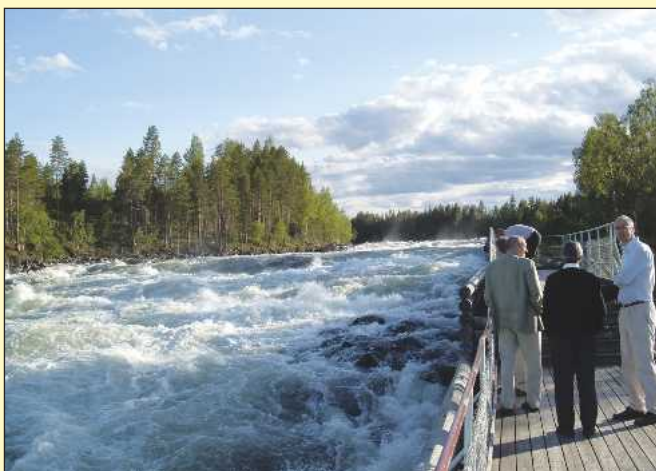
Editor's note: A two-day preconference on education was held before the technical conference. See Education News in an upcoming issue for information on the preconference.