AES 52ⁿᵈ International Conference
Sound Field Control
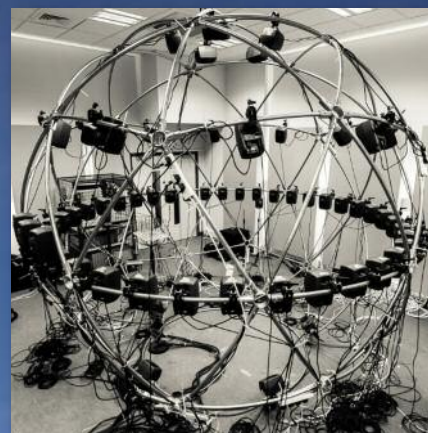
2–4 September 2013
University of Surrey
Guildford, UK

Sponsored by

⊛YAMAHA

ca
commercial audio

CONFERENCE REPORT

Sound zone demonstrations in the sphere loudspeaker array

The Austin Pearce Lecture Theatres, where the main sessions were held

Conference chair Francis Rumsey welcomes delegates to Surrey.

Presenters and delegates from all over the world gathered at the University of Surrey in the UK for a three-day conference on sound field control, chaired by Francis Rumsey. The University of Surrey is internationally known for its teaching and research in sound recording and psychoacoustics, as well as its work in telecommunications, machine listening, and signal processing. Based in Guildford, close to London's major international airports, its facilities provided an excellent base for three days of papers, posters, workshops, and demonstrations related to this growing field of audio research.

Introducing the conference, Francis Rumsey explained that sound field control enables the active management of audio delivered in an acoustical environment. Sophisticated signal processing and reproduction tools increasingly enable the engineer to tailor the sound field for specific applications, occupancy, or listeners' requirements. This can include the creation of independent sound zones in listening spaces, the active control of noise, personal communication systems, the electroacoustic manipulation of auditorium acoustics, and the generation of complex spatial sound fields using multichannel audio systems. Sound field control can be used in automotive audio, consumer entertainment systems, mobile devices, aircraft interiors, concert halls, museums, and other public venues. All this raises questions such as how sound fields can be controlled without detriment to sound quality, what the perceptual effects of different methods of control might be, and how to optimize systems for a specific quality of experience.

The 52nd Conference brought together engineers and perceptual scientists from the four corners of the earth to share their research in the field, and to discuss the numerous interactions between acoustics, signal processing, psychoacoustics, and auditory cognition in this fast-moving field. It is likely that we will see a number of real products based on these principles appearing in our lives over the coming years, as sound becomes ever more tailored to the complicated environments we live in.

During the opening session, the chair introduced his committee, all of whom had done sterling work in bringing the event to fruition. Søren Bech and Filippo Fazi had assembled a bumper papers program, including a number of prominent invited speakers from the world research community. At over 350 pages, the conference proceedings offers an impressive collection of papers on the latest research in sound field control, demonstrating what a productive area of activity this is today. Philip Jackson was in charge of workshops and demonstrations, showing how sound field control technology can be applied and what its perceptual implications might be. Chris Hummersone, as publicity officer, had ensured that all the right people found out about the event, using the latest social media, while Khan Baykaner made sure that conference facilities were coordinated and working properly. Russell Mason acted as the conference treasurer. Roger Furness, AES Deputy Director, and Heather Lane from the UK office were on hand to ensure that delegates could register quickly and easily, and the conference ran very smoothly as a result of their excellent work both before and during the event. Three student helpers from the university were on hand to assist with the organization and the sound sphere demonstrations—Alice De Oliveira, Lucy Kolodynska, and Jack Wensley.

Francis Rumsey (left) with Heather Lane and Roger Furness of AES HQ

The 52nd Conference committee: from left, Filippo Fazi (papers cochair), Khan Baykaner (facilities), Chris Hummersone (publicity), Russell Mason (treasurer), Francis Rumsey (chair), Søren Bech (papers cochair), and Philip Jackson (workshops and demonstrations).

CAREFUL KEYNOTES

Three influential keynote speakers had been invited to begin the proceedings of each day with inspiring overviews of the main themes of the conference. The first day dealt with engineering principles, the second with perception, and the third with creative applications.

Getting the conference off to a good start, Steve Elliott of the Institute of Sound and Vibration Research gave a fascinating introductory keynote tutorial on the principles behind sound field control. He looked back at the early days of sound field rendering with Steinberg and Snow's wall of loudspeakers, prefiguring wavefield synthesis by many decades. In a clever synthesis of the histories and development of both active noise control and sound reproduction systems, he showed both how they are similar and

Steve Elliott

how they are different. Although the principles are very similar in many ways, in active noise control, he pointed out, most of the summation and cancellation of waves that gives rise to the perceived result takes place in the acoustic domain. Steve concluded by examining the applications of sound field control technology in personal audio systems, sound zones, and automotive audio, setting the scene for the packed program of more narrowly focused papers to follow.

Professor Armin Kohlrausch of Philips Group Innovation and Eindhoven University of Technology brought the audience up to

Armin Kohlrausch

speed with trends in psychoacoustic modeling. In particular he posed a question about how far one can get with perceptual models in relation to spatial audio and sound field control. Perceptual modeling, he showed, had enabled restaurant owners to arrange tables optimally in crowded and acoustically poor spaces, for example. Models had shown that depending on the position and head orientation of diners, significant differences in speech intelligibility could be observed. Including head movements in models is difficult, though, Kohlrausch explained. When you make such movements yourself you can interpret the resulting changes in auditory cues, but getting models to do the same thing without voluntary control over the movements is much more challenging. Elevation cues are more idiosyncratic than horizontal ones, which presents another challenge to auditory models. In an attempt to address this, research has taken place into template-based comparison of spectra to determine the locations of sources. Armin concluded by pointing to a very useful public collection of auditory modeling tools, recently made available at http://amtoolbox.sourceforge.net.

Opening the third day with a stimulating discussion of the creative applications of sound field control, Frank Melchior of BBC R&D showed how the manipulation of sound objects was set to revolutionize broadcasting production and consumer entertainment. "What new user experiences can we deliver?" he asked, considering the gradual move toward object-based broadcasting. The idea behind this is to deliver media assets separately to the audience, allowing adaptation, rendering, and interaction at the user end. Audio content elements are transmitted separately from metadata describing them. It's not just an engineering exercise, he said, but it's really happening in applications such as an object-based radio drama and football match broadcast. Echoing a theme developed quite strongly at the 52nd Conference, Frank emphasized that success here is not about accurate physical correctness in representing original sound fields, but about plausible experience. It's not

Frank Melchior

necessarily about recreating reality but about designing a listener/viewer experience that is adapted to their context, task, and environment. Listeners might be allowed to choose their seat in a concert hall, for example. How should one adapt reproduction for tablets, headphones, mobile devices? These are current avenues for his research. Creative people want independent control over sound field parameters that are not tied to physical models, it was proposed. They might want to change the distance of a source without changing the reverberation, or change the Doppler shift without changing the location, or almost any other seemingly odd combination one could think of. There has been a lack of innovation in mixing user interfaces in particular, even in software, said Frank, and there is definitely room for change here. One needs to be able to envisage an object-based mixing console with automation of object trajectories. Although gesture-based mixing interfaces have not proved particularly practical to date, he explained, there is certainly some mileage to be had out of novel sound controllers, one of which he demonstrated.

WONDERFUL WORKSHOPS

Frank Melchior was also the chair of the first workshop entitled "The Interplay between Engineering and Perception in the Design of Sound Systems for Listeners." On the panel were Jung-Woo Choi, of the Korea Advanced Institute of Science and Technology (KAIST), Glenn Dickins of Dolby Laboratories in Sydney, and Armin Kohlrausch, keynote speaker of day two. Frank questioned each speaker about their experience of the interplay between engineering and perception. Glenn Dickins noted that as time went on sometimes he felt he was becoming more of a conjuror or an illusionist with sound, because he had discovered that engineering

Panelists for the first workshop: from left, Armin Kohlrausch, Jung-Woo Choi, Glenn Dickins, and Frank Melchior (moderator).

does not trump perception. The psychology of the moment, he emphasized, often overwhelms the engineering solution.

Peter Lennox, one of the conference authors, suggested that gaining greater clarity about what a sound field is in natural listening might help us understand better what it should be in artificial terms. Learning spatial perception is important, it was suggested, and artificial or hyperreal constructs can also be learned. One of the problems in conducting good experiments in this multifaceted field, suggested Armin Kohlrausch, is knowing whether one can reasonably ask subjects to pay attention to multiple constructs in one test. There is also the challenge of how to weigh the results to arrive at overall ratings or evaluations of the experience in question. In answer to this, Glenn Dickins suggested that waiting for the answer about accurate multidimensional weightings of the elements of experience to build a complete model was something of a waste of time. It's better, he suggested to "grab some low hanging fruit" and move forward in small steps by trying things in subdomains. Predictive models should try to help the engineer to make small, informed steps forward. Focus groups and feedback from them can be useful from a practical point of view, as opposed to formal experiments.

Boaz Rafaely, Ben-Gurion University of the Negev, moderated a workshop entitled "Emerging Techniques, Applications, and Opportunities for Sound Field Control" on the second day. He was joined on the panel by Alain Berry of the Université de Sherbrooke and McGill University, Karlheinz Brandenburg of Fraunhofer IDMT and Ilmenau University of Technology, Gavin Kearney of the University of York, and Emanuël Habets of the International Audio Laboratories, Erlangen.

Alain Berry pointed out that it is no longer possible to evaluate the acoustic performance or quality of a dishwasher, say, using a number. Sound field control and evaluation allows us to do this in a more sophisticated way using subjective evaluation. We can then relate computational acoustics to sound field control, evaluating the quality of sound objects. When approaching complicated problems such as how turbulent boundary layers create noise in an aircraft, he said, we can now use arrays of loudspeakers and DSP to simulate this rather than having to build accurate physical models and measure real sound fields.

Cinema owners now have the ability to decouple the sound reproduction array from the film sound format, said Gavin Kearney, allowing mixing engineers not to have to worry about that any longer. With so many loudspeakers in modern installations, conventional panning tools are no longer practical, and material can be rendered over numerous different layouts. If you want to do third-order Ambisonic mixing under such circumstances then conventional digital audio workstations can't cope in terms of channel count. The lack of standardization of a common intermediate sound field representation format is hampering progress here, he proposed.

Echoing a theme that had been building from a number of people during the conference, Sascha Spors pointed out that adding lots more loudspeakers in sound reproduction systems had not necessarily made things a lot better in

Panelists for the second workshop: from left, Alain Berry, Gavin Kearney, Emanuël Habets, and Karlheinz Brandenburg

Boaz Rafaely chairs the workshop on emerging techniques.

terms of quality. Karlheinz Brandenburg disagreed, suggesting that some demonstrations showed clearly noticeable phenomena with large arrays that simply can't be achieved with something like 5.1 surround. The issue seemed to center on how to more forward in sophisticated sound rendering methods without sacrificing some of the simple, good features of sound that had been experienced in earlier days. There was a vigorous debate on the merits of reproducing sound fields with high physical accuracy versus the art of "plausible experience" as Frank Melchior had put it.

In an attempt to move the debate out of the realm of the familiar and into areas of new possibility, Søren Bech pointed to a need for detailed sound field control in novel fields such as medical applications. We need to be able to keep patients isolated from the upsetting noises in hospitals, for example. Houses are becoming more reverberant and need communications solutions. It was an important call to consider how this new area of technology can be applied to improve the quality of life for people rather than simply to entertain them.

DRAMATIC DEMONSTRATIONS

Set up in the studio facilities at Surrey were a number of demonstrations relating to the conference theme, provided by authors from the conference and commercial sponsor, Yamaha. These helped to show how sound field control can be applied in a range of interesting roles for personal listening, room acoustics control, and to aid the hard of hearing.

The Surrey Sound Sphere was a geodesic metal structure of radius 1.9m used to support 72 Genelec 8020 loudspeakers. These were individually controlled by software via MOTU soundcards. The loudspeaker arrangement was the combination of a regular 60-element circular array mounted on the sphere's equator and a 22.0 surround format, which includes additional loudspeakers at elevations of approximately −30°, +30° and +90°. The system was installed in the relatively dead Studio 2 for sound field control demonstrations of several different methods of forming personal sound zones.

A compact 16-channel loudspeaker array had been developed in collaboration between the University of Southampton and the University of California, San Diego. The real-time DSP engine allowed for multizone sound reproduction and multichannel crosstalk cancellation for binaural audio and for beamforming with listener tracking.

A superdirective loudspeaker array for improving TV audio for the hearing impaired had been developed at the Institute of Sound and Vibration Research, University of Southampton. This device had been designed to generate a highly directive sound beam for spatially
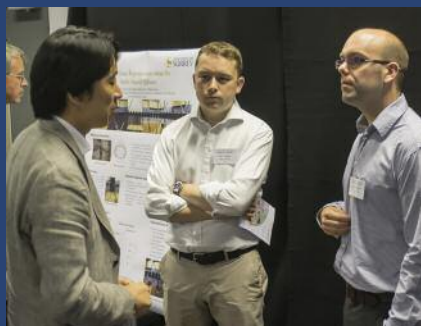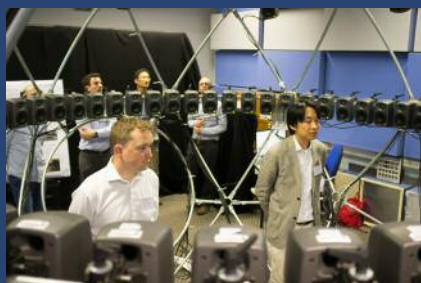
Top row, from left, listening to sound zones in the sphere; recreating a multisource spatial sound scene with a GPU accelerator; a real-time listening test on sound zone interference. Bottom row, from left, delegates discuss their experience in the sphere; Ron Bakker of Yamaha explains reverberation enhancement with Tim Harrison (inset) playing various musical instruments.

DEMONSTRATIONS

localized audio enhancement, improving the intelligibility of audio material delivered to listeners with reduced hearing capabilities.

In addition to the engineering demonstrations, there was also the opportunity to participate in perceptual evaluation of interfering sound zones, such as might occur during simultaneous reproduction of television and radio program material within a single acoustic environment. This included demonstrations of listening tests and the opportunity to listen to a range of stimuli that were representative of interfering sound zone situations. It allowed attendees to experience the potential problems that might be perceived when using these systems, and how various factors such as program material choice, system performance, and listener task affect the perceived result.

Yamaha Commercial Audio had set up a complete reverberation enhancement system in Studio 1, and during an interesting sponsor seminar Ron Bakker explained the history of reverberation enhancement. He showed how the early Royal Festival Hall system, known as Assisted Resonance, amplified the natural reverberant field of the hall by using narrow-band resonators housing microphones, connected to loudspeakers strategically placed. Other in-line systems used artificial reverberation between microphones and loudspeakers. Yamaha's AFC3 is the 3rd generation of hybrid regenerative acoustic enhancement systems since 1987, using FIR filtering and spatial averaging techniques to achieve system stability with a small amount of independent channels. The hybrid part of the AFC3 system consists of a 4-channel convolution reverberator that is used to adjust the existing acoustic response of the room,

rather than replacing it, which is the domain of the 'in-line' systems. The system shown in PATS Studio 1 was a small single-module system that enhanced the diffuse reverberation field in the audience area as well as on stage, using an AFC3-FIR DSP core with four DPA 4060 omnidirectional microphones, twelve IF2108 loudspeakers, and Dante based audio distribution. The system was tuned by Takayuki Watanabe of Yamaha's Spatial Audio System Group in Japan and was visited by almost all the delegates throughout the conference for a closer look and listen.

POWERFUL PAPERS

The main backbone of the conference was provided by nine papers sessions on key aspects of sound field control. These were complemented by a busy poster session preceded by a preview during which authors had a chance to promote their poster to the assembled audience. Overall, 37 papers were given at the 52nd Conference, bringing together some of the most notable researchers in the topic area.

The sessions on sound field control theory and applications included papers on source-width extension, the design of source arrays, determination of sound field control, reproduction of flight recordings, scene analysis from compact microphone arrays, and acoustic element approaches. In the poster session that followed topics ranged far and wide, including real-time sound field transmission, multizone audio reproduction, the uncanny valley of spatial voice, and the relaxation effects of binaural phenomena.

On the second day, two sessions on psychoacoustics followed

Peter Lennox on cognitive maps in spatial sound

Armin Kohlrausch's keynote. Here we learned, among other things, about the perceptual optimization of loudspeaker selection for the creation of personal sound zones, as well as about the prediction of acceptability of auditory interference. Peter Lennox took the audience into new territory by considering cognitive maps in spatial sound, proposing that it's important to learn about the way that the brain constructs its perception of sound fields. Anthony Tucker added to the controversy about whether one should always aim for physical accuracy in sound field reconstruction by revealing the "dirty little secret" that it often helps to introduce some errors rather than control the field too precisely.

Systems to create sound zones were described in five papers after lunch, looking into subjects such as planarity, the effect of reflections, scattering with a head and torso simulator, control strategies for a car cabin system and the design of a superdirective array. This was followed by a session on transducers, array design, and beam forming, including Mark Poletti's presentation on the design of a prototype variable-directivity loudspeaker for improved surround reproduction, and Jiho Chang's

Mark Poletti on variable directivity loudspeakers

presentation on the advantages of double-layer arrays.

Wednesday morning brought the topic of sound field control for multichannel audio, following Frank Melchior's keynote lecture. Mincheol Shin spoke about the control of velocity for sound field reproduction, and Angelo Farina introduced "Spatial PCM Sampling" as an alternative method for sound recording and playback. After the break we got into room acoustics control, looking at a Danish low-frequency test facility and the design of active acoustic absorbers. Boaz Rafaely presented an invited paper together with Hai Morgenstern and Noam Shabtai, on

Richard Furse discusses a poster.

Joe McCabe of Bose asks a penetrating question.

Jordan Cheer waxes lyrical on car cabin personal audio systems.

sound field control in enclosures by spherical arrays.

To wrap up the proceedings the final afternoon session on wavefield synthesis was opened by Karlheinz Brandenburg, who spoke on the future of intelligent multichannel signal processing in audio reproduction systems. Dylan Menzies followed by introducing efficient 2.5D driving functions for quasi-wavefield synthesis, and finally Keunwoo Choi discussed the process of multichannel to WFS upmixing using sound source separation.

SPECTACULAR SOCIALIZING
One of the main reasons people go to conferences is to meet others working in the same field, and the 52nd provided plenty of opportunities for good social interactions between the delegates. Many were staying on site in university student accommodation, so there was a good sense of community during the three days of the event.

A buffet supper on the first evening followed seamlessly from the poster session, allowing an informal atmosphere in which people could either renew old friendships or forge new ones, continuing the discussions they might have started around one of the posters. The conference dinner on the second night was held in Guildford Cathedral Refectory, during which an excellent three-course meal was served. The outstanding weather, which was warm and sunny, allowed delegates to enjoy the outside air before the dinner and in-between conference sessions, taking their lunch outside or just relaxing before another heavy period of concentration. There was an almost endless supply of coffee, tea, juice and cakes, courtesy of the university's catering operation, to keep everyone well fueled for the day's work. Many returned home a few pounds heavier, it seems likely.

SUMMARY
After a busy three days the assembled company departed with numerous new ideas and a much better overview of the topic of sound field control. Many said that it had enabled them to see beyond their own narrow slice of the subject, and to make connections between the engineering and perceptual domains. Of the key themes raised during discussions at the conference, the chair summarized two as most important. First, that context is everything—it is possible to do almost anything you like with an array of loudspeakers and some clever signal processing, but the measure of success depends on the task in hand and what is appropriate for one situation may be entirely inappropriate for another. Second, the boundary between illusion and physical accuracy is a hard one to cross successfully. The closer we get to

physical accuracy in sound field control, the more in danger we may be of falling into the "uncanny valley" where the incongruities between perception, expectation, environment, and acoustics become more keenly noticed.

Editor's note: You can download papers from this conference from the AES e-library at http://www.aes.org/e-lib/

Francis Rumsey hands out certificates of appreciation to various participants, together with bottles of local English beer for them to sample.

Guildford Cathedral refectory, location of the conference dinner

Enthusiastic delegates swap business cards during a break.

Delegates enjoy a splendid meal in Guildford Cathedral Refectory.