

AES 49th International Conference Audio for Games

London, UK
6–8 February 2013

CONFERENCE REPORT

The AES 49th International Conference, *Audio for Games*, was held in London this year from February 6th through 8th. This was the third AES conference devoted to game audio, held every two years in London. The conference is one of the largest conferences in the world focused specifically on sound and technical art of interactive audio. Looking to bring together the working practices, discussions, innovations, and diversity in interactive audio across professional game development and academic research, the AES *Audio for Games* conference helps bridge the gap between works done across different disciplines in interactive audio. The inaugural AES Audio for Games conference in 2009 (the 35th) ushered in the beginning of the current video game console generation. As the industry stands on the cusp of another transition to new technology, the importance

of knowledge sharing between theoretical research into interactive audio and the practical application of techniques is more critical than ever.

The facilities at 2 Carlton House Terrace, where the conference was held, provided an intimate and flexible space to accommodate the free flow of ideas throughout the week. The main presentation hall played host to over fifteen speakers, and another auditorium hosted the paper and poster sessions. When this second auditorium was not in use, a live video feed from the main presentation hall was projected, allowing for overflow and to help inspire further conversation. In addition to these spaces there were three smaller rooms that ran short sponsored educational sessions throughout the week.

The central London location proved the perfect stage for the three days of conference proceedings in



Michael Kelly, Conference Chair

Sponsored by





The impressive Mall to the rear of Number Two Carlton House Terrace, where the conference took place.

an inspiring program of sessions, workshops, and tutorials centered on game audio. Split between the main conference room, an auxiliary auditorium, and several small breakout rooms, the space was large enough in scope with added diversity to suit the different program materials. While all of the sessions, along with the keynote, were held in the main room, papers were presented downstairs, along with sponsored tool tutorials and demonstrations in the break-out rooms. The diversity of presentations allowed the opportunity to fine-tune the focus of one's experience by easily mixing between the main track, technical papers, and tools/tutorial discussions.

The first day of the conference started with short introductions that found conference-goers in a crystalline state of anticipation that shaped the next day's events. Attendees had come from all over the world to participate in focused presentations, technology, and spirited discussions centered around interactive, and specifically, game audio. The palpable excitement of these first moments, punctuated by stories of commutes to the conference on the ubiquitous London Underground, found everyone quickly warming to the idea of being surrounded by colleagues and old friends. The best of the game audio community was out representing legendary London studios such as Sony Computer Entertainment Europe, Codemasters, Lionhead, and EA in addition to European favorites Guerrilla, IO Interactive, Wargaming, and North America's finest from Microsoft, Audiokinetic, and a contingent of developers from Japan and Australia.

One of the pervasive aspects apparent at this year's conference was the presence of dynamic mixing and loudness trends. These aspects were pervasive throughout several talks and panels over the course of the three days. It would seem that with the escalating voice count and increased complexity of dynamic systems that techniques and tools are finally emerging to address this formidable task. From Garry Taylor's talk on loudness standards, to a panel of game audio experts discussing the disparity between film and game mixing, a thread of best practices was discussed that further overflowed into a panel discussing the future of game audio. Meanwhile Xavier Buffoni's talk on HDR Audio exposed a topic that has been circling around in proprietary toolsets for years, and is now poised to become accessible to users in the forthcoming generation of middleware. Coupled with the Atmos demonstration during Dolby's sponsored event, dynamic mixing



Robin Rimbaud (aka "Scanner") delivers an inspirational keynote.

was a topic of conversation for many. It became clear through the week that it is a topic at the forefront of people's thinking.

KEYNOTE

This year's keynote featured British composer Robin Rimbaud (aka Scanner) who chronicled his career as a fearless designer of sounds across many media. While he's not been involved with video games directly, his work has often touched on real-time collaborative interactions, be it with people directly or within a space. His rapid-fire explanation of various works assured that there was something that everyone could connect with—an understanding of the overlap and a feeling of inspiration from his work creating unique audio experiences.

In one example he told of a bus ride through London where microphones attached to the outside and speakers on the inside provided a dynamic and randomized sample-set that he mixed in real-time during the course of each sojourn. Guaranteed to be different every time, the audience was treated to an aural view of London outside their usual experience (and in one case, he explained, a very unusual experience!). Another project that provided direct interaction between participants was an installation that consisted of a large geometric cube housed outdoors in Brooklyn, NY, that could be walked through. A series of triggers controlled sample playback and lighting effects throughout the structure, providing an interactive soundtrack based on movement throughout the space.

As Michael Kelly explained in his introduction, Robin was a risky choice as a keynote speaker because it was unknown how well his career would translate to the discipline of interactive audio and specifically audio for games. Throughout his presentation he was able to cite examples that bridged disciplines and some projects which would be equally comfortable within an interactive gaming experience. It was clear through conversations with conference-goers after his presentation that his work was inspirational and helped set a tone of exploration outside the boundaries of typical interactive sound.



SOME DAY 1 SPEAKERS

Scott Selfon (top left), Xavier Buffoni (top right), Martin Walsh (bottom left), and Stephan Schutze (bottom right)

DAY 1

After introductions on Wednesday morning, sessions began with Scott Selfon (Microsoft Advanced Technology Group) presenting a talk entitled “Crossing the Streams.” His talk aimed to increase awareness of areas where audio has an ability to interoperate between a game and an audio engine. Kicking things off with a series of comedic videos, the presentation helped to communicate a common language of the subtle intricacies at work in games that are harnessing dynamic audio. Some of the specific examples that broke out of this paradigm included “You Don’t Know Jacks” implementation of specific callbacks to audio regarding day-of-the-week information to day/time-specific VO and a comparison of the PS2 version of SSX, which modulated a low-pass filter when in air, versus the more recent implementation of EA’s RUMOR technology, which enabled on-beat landing (where an animation will wait for music-specific timing before finishing a jump). Throughout the presentation, Scott showed the power of two-way communication between the game and audio engine and encouraged developers to leverage the inherent interactivity to extend the capabilities of audio to reflect gameplay.

Xavier Buffoni dug in deep to expose the “Myths, Facts, and Techniques behind HDR Audio” in a presentation that provided an overview of the known facts of HDR from DICE’s public presentations on HDR Audio using their Frostbite engine. He went on to describe the process of Audiokinetics’ own research into the application of HDR Audio and shared some of their findings. Fundamental to the presentation was the question: “Why map sounds to HDR audio?” His simple answer being: in order to represent real life dynamics and greater dynamic range. Referencing HDR photography, he went on to compare this technique with its use in HDR mixing with the focus on loudness as an approximation of the importance of sound in time. He then presented examples that helped to illustrate and frame decisions and workflow choices involved with implementing HDR Audio

within a game audio pipeline using the Wwise authoring application. Key to these examples was the ability to visualize the application of an HDR Audio system working in different ways within Wwise to illustrate the resulting affect.

Martin Walsh overviewed a “A (Fairly) Brief History of 3D Audio in Games” pulling from his past experience. As someone who has been working in the industry for the last 20 years, pushing the forefront of 3D audio, he provided a valuable perspective on the development of technology that has lived on the cutting edge during his time at Creative Labs and other associated companies. His personal history has been skewed toward the PC platform, and his experience stems back to the definition of 3D audio. This included how spatialization was approached through technology, including positional audio, dynamic reverb “morphing,” and HRTFs. His history covered the inception of audio APIs for handling 3D audio, through DirectSound 3D, A3D, and beyond. His deep involvement as one of the major players in the development of early 3D audio lent a “fly-on-the-wall” perspective that resulted in many inside anecdotes and stories from the bygone days when the battle for 3D audio raged between different developers. His presentation provided more than just a history and helped shed light on the behind-the-scenes machinations at work during the upheaval in hardware versus software interactive audio development on the PC.

Tireless audio advocate Stephan Schutze took the stage with a confidence that was infectious. His talk on “How Sound Affects Realities: Enhancing Narrative Through Audio” was a borderline sermon on the ability of sound to communicate to the player in a way that visuals are sometimes unable. He used a sound sample of an airplane and asked conference attendees to listen and connect with the sound through their own experiences. He then went on to detail an instance where the authenticity of this sound was more than just an airplane; for him it held very specific connotations that transcend the simple definition of the sound. As creators of experiences, he urged the audience to respect the use of sound and what it can convey in a given situation, outside of the direct need to reflect visuals. His example of the use of airplane sounds, detached from any physical representation in a game, showed the value that contextual sound can bring to an experience.

DAY 2

Garry Taylor kicked off Day 2 with his talk on “Measuring Loudness in Interactive Entertainment.” Garry has worked tirelessly at SCEE to assemble and analyze data related to loudness in games as part of an initiative within Sony, and he shared this experience with attendees. As part of his findings he was finally able to share the standard that Sony has adopted, based on the ITU-R BS.1770. The process they have adopted includes audio from a title measured as a whole; not separated by music, sound, or dialog, for a minimum of 30 minutes. This measured content should be representa-



Olivier Thereaux and Jory Prum after their presentation on a new standard for web audio.



SOME DAY 2 SPEAKERS

From top left: Garry Taylor, Adele Cutting, Mike Caviezel, Mikkel Christiansen and Franz Galschiøt Quaae



A hands-on session about FMOD Studio on Thursday afternoon



John Broomhall chairs a workshop on "Theoretical, Technical, and Practical Frameworks for Interactive Mixing: A Moderated Panel Discussion."

tive of the different types of gameplay. Their standard is being enforced through quality assurance testing as part of platform requirements for their first-party Sony titles. This standard also forms the basis for a recommendation for all titles released on Sony platforms.

Jory Prum and Olivier Thereaux gave an overview for a "New Standard for Web

Audio" by educating conference attendees in the history of web audio, current activity in W3C Audio Working Group and WebAudioAPI, as well as providing examples of the current implementations and applications. The general feeling was that web audio has established a substantial foundation for the future of interactive audio in the browser. The crowd response was favorable to this presentation; many spoke of the potential being established by web audio. While the effort is still relatively new, it also affords great possibilities for crosscollaboration between game and web audio in order to educate and inform the direction of future efforts. As such, the potential for game audio profes-

sionals to assist in establishing these standards and best practices is greatly needed in order to leverage the experience of the game audio community.

Adele Cutting moderated a panel called "The Future of Game Audio: A Retrospective" that took predictions of the future of game audio from the AES 35th Audio for Games conference and reflected on where the industry at the time thought we would be by now. The first topic addressed was the use and development of the mix process and comments on the rise of dynamic mixing and readily available tools to accomplish this. Tom Colvin (Ninja Theory) advocated the continued rise of standardization in game audio and the addition of a postproduction phase for mixing. This postproduction mix was seconded by Thomas Bärtschi (IO Interactive) who was able to achieve this nearing the end of *Hitman: Absolution*. Storytelling and narrative were seen as an area in games that required work to achieve a greater quality, and the panel felt like in the interim years there had been good progress across all disciplines to increase the ability of games to tell stories with emotions with the help of sound. A question about how to accomplish better dialog in games turned into a general advocacy for the increased communication between different departments as a way to improve the quality of a title overall.

DAY 3

Mike Caviezel, long time racing game sound specialist and now Microsoft SoundLab manager, addressed the technical and emotional task of implementing vehicle sound systems for games in a presentation titled "Planes, Trains, and Automobiles: Creating and Implementing Vehicle Sound Systems for Games." One part science and another part rock and roll, the presentation walked a fine line between automotive specifics and the passionate delivery of the speaker. The components' combustible engines were discussed in an accessible way, allowing for the presentation of interactive audio systems that are able to support both the accuracy of today's physics simulations, as well as the emotive needs specific to games. Mike's experience and personality shed light on the complexities involved with recreating these systems both technically and in a way that appropriately amplifies the drama expected from games as opposed to strict simulations.

Mikkel Christiansen and Frans Galschiøt Quaae from IO Interactive gave an overview of "Behind the Mix—An In-Depth Look at the Audio Engine in Hitman: Absolution." Their overview focused on the music and dynamic crowd systems and their implementation using their internally developed Glacier 2 engine, which extended the functionality of the FMOD Ex audio engine. Their

tool-side integration built upon the framework implemented throughout their game engine pipeline and enabled, through the use of a node graph interface, the complex scripting of audio behavior based on other aspects of the game development pipeline. Apart from establishing the building blocks of their audio pipeline within the game editor, on most occasions the team was able to make decisions about the behavior of sound without involving programmers. Different systems can be used as templates containing complex structures that can be reused throughout the project. The dynamic crowd system in *Hitman: Absolution* constantly adjusts the vocal reactions of the nearby crowd based on game-state information about the player's action. The four states derived from the behavior of the Hitman that modify the vocal reactions of the crowd are: Ambient, Alerted, Scared, GetDown. These states are used to change volume between different layers of crowd ambience, which correspond to the mood that should be expressed by the crowd. Additionally, the size of the crowd and other factors can contribute to determining which ambient sounds are played.

PAPERS

The true success of the papers presented at this year's conference was not only their diversity in subject matter and depth of research, but in their relevance to development of audio for games. A diverse selection of topics assembled by Damian Murphy illustrated the significance of educational research into the reproduction of real-time audio. It could be said that by looking to the work being done within the educational incubators, one can catch a glimpse of the future. Some ideas and methodologies presented relied heavily upon deep-thinking, theory, and shared a utopian vision for interactive audio, while others sought answers to questions that can better enable the greater understanding of our relationship to audio in games.

"Modeling and Real-Time Generation of Pen Stroke Sounds for Tactile Devices..." approached the task of delivering dynamic audio feedback through touch interfaces by breaking down gestural interface interactions and modeling them in conjunction with real-time resampling of prerecorded audio content. The paper identified the main sound components of the physicality of pen stroke interaction and built an audio system to support these components for dynamic interaction with tactile devices.

Seeking to expose the effectiveness of player interaction with specific types of sound, Niels Bøttcher presented on Friday morning the findings of his paper titled "Can Interactive Procedural Audio Affect the Motorical Behavior of Players in Computer Games with Motion Controllers?" Detailing his use of procedurally generated swordswings, in conjunction with the Nintendo Wii remote, the presenter established a test to determine whether the procedural and dynamic nature of synthesized sound would be perceived differently from prerecorded samples

using a simulated swordfighting game. The test findings showed that 16 of the 40 participants showed a difference in physical interaction between procedural and prerecorded swordswing techniques.

Also of note was a paper from David M. Young titled "The Future of Adaptive Game Music: The Continuing Evolution of Dynamic Music Systems," which established the current state of music for games and speculates on what is yet to come. In addition to the technological advancements allowing for greater control of authoring musical content for games, the author envisions a future where biofeedback is used to adapt the soundtrack based on the player-state awareness enabled by future "emotion detectors."

POSTERS

Posters were displayed in the second-floor atrium during the final day of presentations and served as a counterpoint to the discussions surrounding the last day's speakers. Attendees were able to speak directly with representatives and ask questions based on the information displayed.

In a poster presented by Tapani Pihlajamaki entitled "Plausible Mono-to-Surround Synthesis in Virtual-World Parametric Spatial Audio," a process was presented to allow for the synthesizing of multi-

channel output from a mono signal based on an estimation of reverberant energy in a signal. Simply put, the paper suggests a way of estimating the positional spread of a sound based on various factors as a way to "fill in the gaps in the sound scene to make it sound surrounding and immersive to the listener." This paper was presented in conjunction with "Modular Architecture of Virtual-World Parametric Spatial Audio Synthesis," which



SELECTED PAPER PRESENTERS

Niels Bøttcher (top left), Erik Sikström (top right), Søren K. Jacobsen (bottom left), Tapani Pihlajamaki (bottom right)



Intense discussion about posters in the library



Roberto Pomoni takes a look at "Global Game Audio Production: Where Are We Going?"

restructures some of Directional Audio Coding (DirAC) work that is at the core of their proposed system for parametric spatial audio coding.

A paper from Philip Jackson and Naveen Desiraju attempted the "Use of 3D Head Shape for Personalized Binaural Audio" in order to custom tailor the spatialization based on the head shape of the listener. Using a stereo-vision face-capture system comprised of Kinect camera hardware and proprietary capture system, a model of the face and head was captured for use in creating the HRTF (or HRIR) dataset. Listening tests for subjects who participated in personal face and head modeling were able to better locate sound sources than participants that used a dummy-head HRIR.

A tool was suggested for "Geometric and Wave-Based Acoustic Modeling using Blender" by Jelle van Mourik and Damien Murphy as a way to automate reverberation based on 3D-modeled spaces. Presented as a "means of real-time sound rendering" the paper goes on to detail the process of developing acoustical models using Blender and comparison between other models of reverberation. Findings have shown that the performance of Blender as a "sound scene" renderer is a viable way to generate impulse responses that show similarities to other methods of modeling the acoustics of spaces.

While there exist other opportunities to submerge oneself in game audio at other events, the AES Audio for Games conference allows a rare opportunity to rub shoulders with creatives and academics alike. It's this interaction between researchers, scientists, professionals, and students working with interactive audio that makes the intersection of knowledge sharing at the confer-



The gathered conference congregation, led by secretary Becky Stewart, chair Michael Kelly, and papers chair Damian Murphy.

ence such a unique experience. The exposure to the different disciplines and approaches within audio leads to rich dialogs, smart questions, and opportunities to educate across all levels of expertise. It's hard to see what the direct result of this may be in the future, but the opportunity to be able to get everyone on the same page inspires new ways of thinking. It is this inspiration that AES Chair Michael Kelly was hoping to cultivate when he set it as the theme for this year's conference, and judging by the acclamation from colleagues during the event it can safely be said that he succeeded.

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