

# AES 56<sup>th</sup> International Conference Audio for Games London, UK 11–13 February 2015

## CONFERENCE REPORT

The 56th International Audio Engineering Society Conference on Game Audio was held in stunning central London surroundings in February at Burlington House. Two rooms at The Royal Institute of Chemistry in Piccadilly hosted the gathering of audio engineers from all corners of the globe. Presentations were delivered in the Science Room on the ground floor, papers on the first floor in the Council Room, and tea and biscuits enjoyed in the somewhat curiously-named but splendid Fish Room.

With nearly 100 delegates with varying interests and backgrounds, the conference provided a forum for presenters to share ideas, visions, techniques, and tips for the future of computer game audio. Developing intricate and fresh approaches to sound design, multichannel audio, and utilizing creative compositional techniques within the confines of technology became a reoccurring theme over the three days.

### PRESENTATIONS

Chair Michael Kelly began proceedings, gave a warm welcome and introduction, and set the stage for the days ahead.

The first two sessions were perfectly pitched to invigorate those wearied by an early Monday morning commute into central London. Presented by Scott Selfon, principal



Michael Kelly, Conference Chair



Sponsored by





Scott Selson, principal software engineering lead at Microsoft, delivered a presentation on “Game Audio Implementation: Past, Present and Future”

software engineering lead at Microsoft, and Stefan Schutze, director at Sound Librarian, each session was both energetic and inspiring.

Scott began by giving a brief overview of Game Audio, past and present, before moving on to look at what audio frontiers and barriers lay ahead of sound designers in the coming years. He highlighted the fact we are now in the third year of what was coined the next generation, and Scott asked “what comes next?” The answer—“more!” More voices, more fidelity, and more DSP. This became the theme for the next hour and the rest of the conference.

Using examples including *Braid*, which used an early audio scrubbing control to rewind music as the game played in reverse, *Star Wars X-Wing* that utilized music hinting at knowledge of the future, and *Peggle 2*, that uses pizzicato notes and chord progressions to encourage game progression, Scott walked us through examples that increased awareness of the in-game environment through accurate reinforcement and detailed, realistic audio experiences.

Scott then reflected on the current state-of-the-art techniques for areas ranging from spatial simulation and acoustic modeling, to evolving and dynamic mixing, audio as a feedback mechanism, and highly personalized audio experiences within gaming. This included using sound to reinforce situation, location, and environment through sound propagation in relation to distance, temperature, humidity, altitude, and airflow. Also, improving game awareness by creating dynamic, interactive, responsive, and informative soundscapes, and the use of spatial sound and binaural processing.

Looking to the future, Scott highlighted games that are responsive to external audio, games with interactive dialogue, and the development of diegetic and nondiegetic audio.

Stefan Schutze delivered a fast-paced presentation on techniques he has developed using F-Mod Studio middleware (there was also a fantastic interactive demonstration from Stefan on this software). Stefan’s energetic approach to game audio has been inspired by the desire to move away from simply handing over a file of sound effects to underwhelmed audio programmers, and towards programming audio in real time alongside the game in development, essentially enabling the sound designer to work on the game right up until release.

Pushing boundaries certainly appeared to be Stefan’s impetus. After claiming that AAA projects can be way too safe when it comes to game audio, he went on to use his current project, *Defect Spaceship Destruction Kit*, to describe his techniques, which reportedly took six months of planning before a single sample was programmed.

The idea of the game, in Stefan’s words, is to “design a spaceship, take it out, and fight with it.” Each part of each ship has a dedicated musical motif. So players are essentially developing a specific musical score as defined by the components of their ship. Set chord progressions then develop, depending on gameplay—attacking, flying, and so forth. Musical parts and parameters also respond to game play. Music is triggered by time and script.



Stefan Schutze—The Sound Librarian

By linking FMod to Unity, the sound designer can program notes while playing the game. Keeping most elements largely percussive means the audio is able to change on a beat. Although some more musical elements take longer, everything is quantized to time, resulting in barely noticeable transitions.

Using Fmod in a similar way to a standard DAW, Stefan has developed techniques using EQ and dynamic processing to change the sound of a single sample for different applications within the same game; this “pseudo-velocity” is massively efficient memory-wise. Stefan explained that he liked to aim for the top sonically before peeling back layers to achieve the desired CPU load. Designing random macros and using output busses for effects helped keep CPU usage manageable.

Following a tasty lunch and yet more coffee, the conference settled back into its seats in the Science Room to hear Ilmari Hakkola, the head of audio at Rovio Entertainment, speak on audio branding. Ilmari joined Rovio as a graphic designer initially but developed strong ideas on how audio should influence gameplay, affect game design, tell a story, and influence player decisions by bringing extra information to the player about the virtual world itself.

Ari Pulkkinen’s familiar theme for *Angry Birds* is based on native Finnish folk or Umpah music, and its quirky and distinctive style was used as a foundation for Rovio’s “audio brand.” Using Slovakian



Ilmari Hakkola, head of audio at Rovio Entertainment, brought some “rock star” to the conference.



**AN ATTENTIVE AUDIENCE LEADS TO QUESTIONS**

and Balkan influences, Ilmari and his team went on to produce variations on the Angry Birds theme with the intention of making music that was “more fool than cool.” Using instrumentation that sonically represented elements of the game (low-range brass instruments for pigs, for example), modal variations were incorporated to represent different geographical locations, and popular themes worked into the original compositions to represent franchise versions. The Death March, for example, was incorporated into Angry Birds Star Wars.

Ilmari’s key point was that consistent audio created an audio brand. Rather than having audio bolted onto a creative project, audio designers should push to be part of the creative development process, developing an audio brand within the corporate structure. This was something that was strongly echoed by the keynote speaker, Joanna Orland.

Audio branding is not limited to the score but to dialog too. Those familiar with the Angry Birds brand will know that there is no specific dialog but the characters do speak fluent gibberish. Original approaches to recording and animating were undertaken and characters developed key sounds to identify specific emotions. This technique required animators, sound designers, and actors to work together creatively, connecting the audio department across the company.

Ilmari ended by discussing ownership and rights for musicians and producers working within game audio. Obviously it would be financially beneficial for musicians and producers to move away from the traditional audio buyout scenario, and games companies are never keen to license tracks or work with third-party music. Ilmari’s advice? If you can’t beat them join them! Consider becoming a publisher, develop a brand, and use it to exploit your work.

Martin Loxton of Frostbite drew the first day to a close reviewing new approaches to structuring audio content to quickly and effectively

build game audio, without compromising on quality due to memory and performance restraints. He discussed the technologies developed by Frostbite to support rapid content creation using modular sound design coupled with data inheritance, allowing for a succinct separation between the modeled behavior and the rendered result.

Using guns and cars as his prime examples, Martin observed that traditional methods of sound design for AAA games had reached their limit. Creating sample data was taking too much time, sample storage and recall were becoming problematic, and maintaining consistency of sounds throughout development of game was an issue. The solution? Palette or modular sound design. These techniques are based around using a pool of individual samples (an exhaust, a manifold, air intake, for example) and then combining these sounds in different ways as multilayered samples for different scenarios. Gameplay then decides how the sounds are played. Integrated nodes affect the sounds in different ways, incorporating effects, pitch, and filters. The game then mixes and renders the sounds in runtime.

Creating new sounds in this way is efficient, as new data is not needed for each new sample. The sound designer mixes and matches data from samples that are already available to create new sounds. Martin admitted that initially there can be a lower quality bar, but this improves quickly as the designer becomes more proficient. Each voice is a component; each component can be controlled by logic. The sound is a result of combining the components. Separate the data from the structure and the sound patch becomes a template. The different data results in a different sound.

Once the final session was complete a cavalcade of cabs arrived to ferry delegates to Dolby’s headquarters in Soho for a stunning demo of their Atmos system.

Day Two was kicked off by Amaury La Burthe, CEO of AudioGaming, France. Amaury continued to discuss methods of improving sound design technique, and delivered an overview of experiments around the creation of a procedural audio framework. The challenges he highlighted were twofold: the need for an efficient framework to generate low-level optimized code, and developing an approach for designing procedural audio models.

Although not a replacement for interactive sound and game audio engines, game engines have overcome the inherent linearity of pre-recorded audio and so procedural audio has many advantages. It saves memory by using code instead of sample data, it can lighten



Conference chair, Michael Kelly (right), with Loro Solomon at the Dolby reception.



Natanya Ford, conference secretary, enjoys a slice of pizza at the Dolby reception with Dan Peters.



Bill Foster, conference treasurer

the burden of tedious tasks (sounding hundreds of interacting objects or animations), and it brings flexibility to the pipeline (because sound is directly linked to game parameters, changes will modify audio generation accordingly to keep it synced, avoiding the need for sound redesign).

**KEYNOTE**

Canadian-born sound designer Joanna Orland was invited to deliver the conference keynote. Her presentation on the creative and original approach to her work as senior sound designer at Sony Games, was incredibly refreshing. She has worked on titles including *Book of Spells*, *Diggs Nightcrawler*, *God of War 3*, *The Deep*, and *Street Luge* demos for Project Morpheus.

Joanna was another passionate speaker who has a desire to put an end to treating audio as an afterthought in game production. She outlined five key points that she believes will ensure audio is integrated earlier in to the game production cycle, resulting in a strong audio identity. She asserted that while research and knowledge are important, imagination matters just as much if not more. Experimentation leads to innovation.

Researching media in the genre in which you are working, then analyzing the patterns that emerge from the research to develop a particular audio style, formed the first two key points. Using the horror and fantasy genres as examples, Joanna described how she discovered, through research, that the majority of magic spell sounds are based in nature, thus reinforcing the impression that the fantasy is in fact reality. In the horror genre, realistic character sounds are placed in hyper-real surroundings, which creates tension and fear.

As a result of research and analysis, an audio statement of intent should naturally develop, which in turn gives the audio team and the development team a common language in which to discuss sonic aspects of the game.

With the start of the second afternoon of the conference, it was time for Dan Finnegan’s presentation on audio-only games. The charismatic entrepreneur and research engineer at Somethin’ Else Sound Directions kept the conference intrigued for the entire hour on the development of his team’s audio-only iOS game release, *Audio Defence Zombie Arena*.

Surrounded by zombies, the player cannot see and must rely on what he hears to aim and fire. It’s all facilitated by binaural sound and immersion. According to Dan, it is truly unlike any other zombie game. The zombies and playing environment are all rendered in high-

fidelity binaural sound. Using only ordinary headphones, because the sound is binaural, the player actually hears the zombies in the world all around, and is situated right in the middle of the game. The game has also been engineered to be entirely playable by the visually impaired. Even the game menus are designed to be read by those suffering from color blindness.

Dan and his team have already won a BAFTA, a Tiga games industry award, and an Applevis award.



Dan Finnegan discusses audio-only games.

**PAPERS**

The paper sessions, chaired by Jez Wells (organized by Damien Murphy), were delivered on Thursday upstairs in the grand, pillared Council Room.

The first paper looked at using game audio programming techniques in university curricula. As part of a team delivering an undergraduate sound engineering degree, intent on delivering transferable engineering skills across all disciplines this was an interesting and relevant session by Rob Hamilton.

Rob’s presentation detailed a multitiered set of technologies and workflows appropriate for the introduction and exploration of beginner, intermediate, and advanced procedural audio and music techniques. Initially, engaging students who may not have had a direct interest in game audio through traditional channels like musicianship or foley work. Then steadily introducing and encouraging them to learn specific systems and work flows for rapid game-audio prototyping, real-time generative audio and music systems, as well as performance optimization through low-level code generation.

Three papers based on generative music systems followed. Alper Gungormusler from Ireland kicked off with barelyMusician, an interactive and comprehensive music composition tool capable of generating and manipulating audio samples and musical parameters in real time in order to create smooth transitions between musical patterns. Alper was able to demonstrate his software in the session, portraying varying emotional states and moods that may be evident during gameplay.

Duncan Williams presented “Dynamic Game Soundtrack Generation in Response to a Continuously Varying Emotional Trajectory.” A paper that addressed some of the practical and aesthetic challenges in game audio programming. The team from The University of Reading and Plymouth University developed a system that can respond in real time to emotional trajectories derived from two-dimensions of affect on the circumplex model (arousal and valence), which are mapped to five musical parameters.

The final paper of Session One described a system for integrating user playlists, what has been termed musical UGC (User Generated Content), into games. The system, Veemix, allows for the simultaneous integration of music into games.

The three papers that followed were based around game music systems. Leonard Paul from the School of Video Game Audio in Vancouver gave an overview of the generative music process used in the programming of the educational game, *Sim Cell*. The audio was



Joanna Orland, keynote speaker, makes a plea for audio not to be an after-thought in game production



**FOUR PAPER PRESENTERS**

Clockwise from top left, Alper Gungormusler, Leonard Paul, Aristotel Digenis, and Duncan Williams



Stately corridors of the Royal Society of Chemistry

**AES THANKS THE CONFERENCE COMMITTEE**

Chair: Michael Kelly  
 Secretary: Natanya Ford  
 Treasurer: Bill Foster  
 Papers Chair: Damian Murphy  
 Content Team: Scott Selfon, Steve Martz and Damian Kastbauer  
 Volunteer Coordinator: Brecht De Man  
 Webmaster: Carl Bussey

**Volunteers**

Heather Lane,  
 Simon-Claudius Wystrach,  
 Joseph Reese-Jones,  
 Dave Moffat,  
 Rod Selfridge,  
 Christian Heinrichs

possible to enhance the perception of fear based on the direction of the source in a 3D environment.

**CONCLUSION**

Clear themes around developing technologies and techniques were obvious staples of this year's conference, but creativity and the desire to change the perception of audio in games and the practices of sound designers really shone through. Whether programming multichannel audio for AAA games, or designing audio motifs for iOS apps, sound designers are pushing the boundaries more than ever in pursuit, not of what is required, but of what is possible. As sound designers gain more DSP and more storage and with the help of more education, inspiration, and imagination they will constantly be pushing for more.

produced using a generative music system utilizing the open-source visual-programming language Pure Data (PD).

Richard Stevens, Dave Raybould, and Danny McDermott from Leeds Beckett University presented observations on the over-reliance of visual user interface in games.

The paper identified the limitations of current binary threshold approaches to audio feedback and put forward music as a potential vehicle for providing richer data to players.

Researchers from Queen Mary University, London, presented a project that aims to implement and assess sound-synthesis models with dynamic Level of Audio Detail (LOAD). With synthesis techniques often being too computationally complex, the research aims to modify existing models to achieve a more or less complex implementation while still retaining the perceptual characteristics of the sound.

Three final papers were delivered on spatial audio in games. The first was from Aristotel Digenis of FreeStyleGames, UK, who is interested in gamers who play wearing headphones and in developing specific audio settings for this community. Joseph Sinker and Jamie Angus from the University of Salford discussed the use of head-related transfer functions (HRTFs) to convert audio from conventional loudspeaker formats to binaural format. Finally Samuel Hughes and Gavin Kearney of the University of York described how it is



The assembled company poses for a group picture at the end of a successful conference.