

AES 41st INTERNATIONAL CONFERENCE

Audio for Games

London, UK

February 2–4, 2011

Technical Sessions

This preliminary program is accurate as of press time. See updates at www.aes.org/events/41

Wednesday, February 2

INTRODUCTION / WORKSHOPS

W-1	Introduction to Games and Sound	9:00
W-2	Interactive Music Tutorial	10:30
W-3	Interactive Music Tutorial	13:15
W-4	Interactive Music Tutorial	14:30

Thursday, February 3

WORKSHOP

Mobile Game Audio	9:00
-------------------	------

KEYNOTE ADDRESS 10:30

POSTER SESSION 1 12:15

P1-1 Application of HRIR Factorization to Game Audio
—Gavin Kearney, Claire Masterson, Marcin Gorzel, Henry Rice, Frank Boland, Trinity College Dublin, Dublin, Ireland

In this paper we present a strategy for simplifying real-time audio processing for binaural rendering in game audio. The strategy employs factorization of HRIR datasets into directional dependent and independent subsystems and further decomposition of the datasets into spherical harmonic basis functions. Decomposition is evaluated up to 3rd Order Ambisonics, with factorized minimum-phase HRIR datasets yielding as low as 12 sample runtime filters. An objective evaluation of the method is made using psychoacoustic evaluation criteria, and a subjective listening test is implemented to verify the methodology.

P1-2 Audio Transformation Technologies Applied to Video Games—Oscar Mayor, Jordi Bonada, Jordi Janer, Universitat Pompeu Fabra, Barcelona, Spain

In this paper we present a set of spectral domain audio transformation technologies and some existing and proposed applications of these technologies in video games. The technologies include real-time voice transformation and tempo/pitch transformation for music. Voice Transformation is mainly used to create new character's voices modifying existing recorded material but also for transforming the player's voice in real-time in a singing game. Tempo/Pitch transformations are more focused in transforming background music to create different sound atmos-

phere for different scenes in a video-game. Audio Transformation techniques are very useful to reduce sound creation/recording costs but also for reducing the size of the game without compromising the quality of the sound. Moreover, these real-time techniques offer new attractive sound features to games.

P1-3 Providing an Immersive Gaming Experience Using Wireless Low-Latency Coded Audio Streaming—David Trainor,¹ Gary Spittle²

¹Cambridge Silicon Radio, Belfast, Northern Ireland, UK
²Cambridge Silicon Radio, Cambridge, UK

Modern gaming environments contain numerous peripheral devices that enhance the gaming experience by streaming audio in real time to and from the principal gaming platform. Such peripherals include hi-fi equipment, surround loudspeaker systems, specialist gaming headsets, and specialist game controllers. Gamers are increasingly demanding the convenience of wireless communications between these devices, without compromising audio quality or significantly adding a sense of delay or lack of synchronization to the gameplay. This paper discusses how Bluetooth communications technology, coupled with appropriate audio coding and processing techniques, may be configured and deployed to satisfy these stringent demands. Furthermore, the additional advantages of Bluetooth integration throughout the gaming ecosystem are described.

P1-4 Location-Aware Interactive Game Audio—Natasa Paterson, Katsiaryna Naliuka, Tara Carrigy, Mads Haahr, Fionnula Conway, Trinity College Dublin, Dublin, Ireland

With the affordability and pervasiveness of GPS enabled smart phones, downloadable location-aware applications have become increasingly popular. These applications include the overlay of local information onto physical sites to multiplayer games and tourism. However, many of these applications rely predominantly on graphical overlays and simple user interface audio. For narrative led location aware experiences, audio can play an important role in immersing the player in the story line and physical location. Coupled interactivity and adaptability of audio to player movements encourage the user to remain engaged with the physical location. This paper describes the theory and implementation of interactive audio for a location-aware tourism application. The aim of the application is to create an immersive gaming experience where the audio supports the historical setting and given narrative.

P1-5 Sound Effects Processing and Manipulation with Wavelet Packet Transform—*Bruno Silva, Rafael Santos Mendes*, State University of Campinas (UNICAMP), Campinas, Sao Paulo, Brazil

It is proposed a new method for more realistic and less repetitive sound effects for video games. Preserving the main characteristics of the sound, the method synthesizes and manipulates sound based on the wavelet packet transform of the signal. The coefficients of the transform can be used for manipulations in the frequency and temporal spectra, leading to different sounds. They can also be shuffled in time in a continuous synthesis creating a continuous sound texture.

P1-6 Granulation of Sound in Video Games—*Leonard J. Paul*, Vancouver, BC, Canada

Granulation of sound for games is becoming a more viable tool for sound artists as the processing power of game consoles continues to improve. Granulation is a relatively recent method of sound generation that allows for sampled sound to be modified in real-time to allow pitch to change independently of tempo among other audio effects. It has not been commonly used with the previous iterations of game platforms due to its relatively costly DSP overhead and number of voices required. With the current generation of game platforms, these limitations have been relaxed allowing for the increased possibility of real-time granulation in games. The combination of granulation and the output of a game's physics engine is a promising relationship as the physics engine can supply the large amount of real-time control data for the grain parameters that granulation requires. This paper investigates some current research in sound granulation and demonstrates some effective methods of utilizing granulation for games.

PAPER SESSION 1: EDUCATION AND STANDARDIZATION

13:15

1-1 Integrating "Audio for Games" into the Modern Audio Production Curriculum—*Mark J. Sarisky*, The Art Institute of Austin, Austin, TX, USA

The Electronic Game industry is becoming a bigger part of the audio production world every day. Proposals for stand-alone degree programs for colleges, schools, and universities have appeared recently, but this paper proposes that electronic game audio be an integral part of a general curriculum in audio production, along with established components such as sound for film and studio recording.

1-2 Audio Air Hockey: A Pilot Study in Using Audio-Based Games for the Measurement of Loudspeaker Placement Preferences for Smart Tables—*Karen Collins,¹ Bill Kapralos,² Alexander Hodge,¹ Andrew Hogue²*

¹University of Waterloo, Waterloo, Ontario, Canada

²University of Ontario Institute of Technology, Oshawa, Ontario, Canada

Modern game hardware technology (e.g., mobile, touchpad, and smart table) requires new considerations for mixing and surround sound due to changing viewer angles. Can audio-based games provide us with a useful subjective measure of player preference of different loudspeaker placements? To determine an answer to this question, our method was to measure subjective responses to two different quadraphonic loudspeaker placements with a smart table using a custom-built

audio-based game. Results showed that the majority of players had a preference for speaker placement, and that the placement was perceived to impact their gameplay. Counter to our hypothesis, results showed that players did not prefer the standard quadraphonic set-up.

PAPER SESSION 2: MUSIC, SYNTHESIS, AND SPEECH PROCESSING

15:00

2-1 An Online Platform for Interactive Soundscapes with User-Contributed Content—*Jordi Janer, Stefan Kersten, Mattia Schirosa, Gerard Roma*, Universitat Pompeu Fabra, Barcelona, Spain

A principal role of audio in virtual environments is the creation of a sound ambiance or soundscape. Current developments (e.g., online communities, web and mobile technologies) might require new paradigms of soundscape collaborative design and interaction. We propose an online platform that aims at simplifying the authoring process, but offering at the same time, a realistic and interactive soundscape. Graph models, representing virtual acoustic sources, control the sequencing of events (samples) using concatenative synthesis. Samples are retrieved from a user-contributed audio repository, which integrates a content-based search engine and an ecological acoustics taxonomy. A server architecture allows the synthesis engine to support several independent listeners simultaneously. Client applications are responsible for sending position updates, receiving the soundscape as a web stream.

2-2 Server-Based Pitch Detection for Web Applications—*Sascha Grollmisch, Christian Dittmar, Estefanía Cano, Karin Dressler*, Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

Pitch detection algorithms find use in many different applications ranging from automatic music transcription, music classification and analysis, to sound separation. Most recently music games and karaoke applications have also incorporated pitch detection algorithms in their gameplay. Music and audio games pose different technical requirements in terms of the implementation and efficiency of such algorithms. In this paper we describe the development of a server-based pitch detection algorithm intended to work in web applications. Real time performance, algorithm robustness, and efficiency were critical aspects of this work.

2-3 Music from the Environment: Perception of Music Created from Sounds of a Video Game Environment—*Mario Cáceres*, Instituto Profesional Duoc UC, Santiago, Chile

In the past few years, experimental and theoretical research has been conducted to define the concept of immersion and test the different levels of involvement a player may achieve experiencing a video game. This paper describes the creation of a video game music soundtrack constructed from sounds of the environment of the game itself and the testing of the perception of this new form of game music in a preliminary experimental research with a determined number of subjects. The participants tested a stage of the OpenArena 0.7.6 game in three versions: without music, with music created using standard timbres, and with the same music but created from sounds extracted from the effects of the video game. The results suggest a slight tendency to a high level of immersion with the use of the music from the environment.

2-4 Language Scrambling for In-Game Voice-Chat Applications—*Nicolas Tsingos, Charles Robinson, Dolby Laboratories, San Francisco, CA, USA*

We propose a solution to enable speech-driven alien language synthesis in an in-game voice-chat system. Our technique selectively replaces the users' input speech with a corresponding alien language output synthesized on-the-fly. It is optimized for a client-server architecture and uses a concatenative synthesis framework. To limit memory requirements, as well as preserve forwarding capabilities on the server, the concatenative synthesis is performed in the coded domain. For gaming applications, our approach can be used to selectively scramble the speech of opposing team members in order to provide compelling in-game voice feedback without exposing their strategy. The system has been implemented with multiple alien races in a virtual environment with effective, entertaining results.

Friday, February 4

WORKSHOP

9:00

Education in Game Audio

PAPER SESSION 3: GAME REVERB

10:30

3-1 Scattering Delay Network: An Interactive Reverberator for Computer Games—*Enzo De Sena,¹ Huseyin Hacihabiboglu,^{1,2} Zoran Cvetkovic¹*
¹King's College London, London, UK
²METU, Ankara, Turkey

Many 3-D computer games incorporate audio renderers simulating room acoustics to provide the user with a high level of immersiveness and realism. Full-scale interactive room auralization systems are impractical for use in computer games due to their high computational cost. As a low-cost alternative, artificial reverberators can be used. This paper is concerned with the design of a scalable interactive reverberator inspired by digital waveguide mesh (DWM) models and feedback delay networks (FDN). This reverberator is by construction tightly linked to the acoustics of the enclosure that it simulates. Simulation of unequal and frequency-dependent wall absorption, as well as directional sources and microphones can also be incorporated. It is shown that the response of the proposed reverberator accurately renders the early reflections and room modes, as well as providing RT60 values consistent with Sabine and Eyring equations.

3-2 GPU-Based Acoustical Diffraction Modeling for Complex Virtual Reality and Gaming Environments—*Brent Cowan, Bill Kapralos, University of Ontario Institute of Technology, Oshawa, Ontario, Canada*

Despite the importance of acoustical diffraction in our natural environment, modeling of such effects is complex and computationally expensive for all but trivial environments and therefore, typically ignored in virtual reality and gaming applications altogether. Driven by the gaming industry, consumer computer graphics hardware and the graphics processing unit (GPU) in particular, has greatly advanced in recent years, outperforming the computational capacity of central processing units (CPUs). Given the widespread use and availability of computer graphics hardware, GPUs have been successfully applied to other, non-graphics applications including audio processing and acoustical diffraction modeling. Here we build upon an existing GPU-based acoustical occlusion/diffraction mod-

eling method that can become problematic when the sound source and the listener are in separate rooms. The proposed method approximates acoustical occlusion/diffraction effects for complex, multi-room environments. The method is computationally efficient allowing it to be incorporated into real-time, dynamic, and interactive virtual environments and videogames where the scene is arbitrarily complex.

3-3 OpenAIR: An Online Auralization Resource with Applications for Game Audio Development—*Simon Shelley, Aglaia Foteinou, Damian T. Murphy, University of York, Heslington, York, UK*

Game developers are now turning to convolution reverb effects in order to offer an immersive and realistic audio experience. This requires impulse response data that is either recorded or generated artificially and that ideally matches the in-game environment. To help this process the Open Acoustic Impulse Response (OpenAIR) Library has recently been made available. It consists of an online resource designed to facilitate the sharing of acoustic impulse responses and related information. Open-source software tools are also provided, including a library for the Pure Data (PD) programming environment that implements real-time convolution. Details of this tool is described in this paper along with an interactive example that uses computer generated impulse response data to auralize a space in real-time.

POSTER SESSION 2

12:15

P2-1 On the Perception of Dynamic Sound Sources in Ambisonic Binaural Renderings—*Marcin Gorzel, Gavin Kearney, Henry Rice, Frank Boland, Trinity College Dublin, Dublin, Ireland*

One of the requirements for a fully immersive gaming experience is the correct and visually coherent reproduction of location of sounding objects. Limitations of human spatial resolution in this regard have been known for decades and are commonly taken into account when designing audio reproduction systems. However, mechanisms responsible for the perception of static and dynamic sound sources are not the same and less attention has been devoted to the problem of perception of dynamically relocated sound sources, omnipresent in video games. This paper explores the human ability to follow moving sound sources when presented as First and Higher Order Ambisonic renderings over headphones and aims at finding the optimal, psychoacoustically justified parameters that could significantly reduce computational requirements of audio engines.

P2-2 UGen++ —An Audio Library: Teaching Game Audio Design and Programming—*Martin Robinson, University of the West of England, Bristol, UK*

Teaching sound designers and teaching audio programmers are clearly different activities reflected by these distinct roles in the games industry. Existing middleware and APIs such as FMOD and Wwise offer good experiences in terms of the process of communication and collaboration between audio programmers and sound designers. This paper presents UGen++, an audio library with both a high level and low level interface. The high level interface is simple yet powerful enough for sound designers to learn synthesis and audio processing concepts while giving them an awareness of the requirements of the programming role. The low level system is extensible enough to teach fundamental audio

programming concepts and software design while giving an awareness of the sound designer role.

- P2-3 Songs2See: Learn to Play by Playing—** *Sascha Grollmisch, Estefanía Cano, Christian Dittmar*, Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany

Songs2See is a web-based application designed to assist music students in their practice time. Its main features are: use of real musical instruments instead of midi or game-controllers, immediate feedback and assessment, fingering suggestions for each instrument, entertaining ways of displaying content, users possibility to generate their own game content out of audio tracks and user selected options like tempo and difficulty. To provide the user with such a system, current audio processing technologies for pitch tracking, sound source separation, music transcription, and rhythm analysis have been included. We aim to take music gaming into a new direction where learning, playing and having fun can finally converge.

- P2-4 Visualizing and Controlling Sound with Graphical Interfaces—** *Liam O'Sullivan, Frank Boland*, Trinity College Dublin, Dublin, Ireland

Developments in abstract representations of sound from the field of computer music have potential applications for designers of musical computer games. Research in cognition has identified correlations in the perceptions of visual objects and audio events—experiments show that test subjects associate certain qualities of graphical shapes with those of vocal sounds. This “sound symbolism” has been extended to non-vocal sounds, and this paper describes attempts to exploit this and other phenomena in the visualization of audio. The ideas are expanded upon to propose control for sound synthesis through the manipulation of virtual shapes. Mapping between parameters in the auditory and visual feedback modes are discussed. An exploratory user test examines the technique using a prototype system.

- P2-5 Multiplatform Audio Game—** *Jaroslav Beksa, Krzysztof Majewski, Rafal Sadowski*, Orange Labs, Polish Telecom, Warsaw, Poland

The goal of this project is to verify the audio games concept: is it possible to play computer games based on sound only? Can such games be attractive for both sighted and visually impaired players? Is it possible to control such games using only a touchscreen? And finally, can audio games attain commercial success and become a new trend in electronic entertainment? In this paper we disclose the work we did over the last two years creating a commercial multiplatform audio game.

- P2-6 GSound: Interactive Sound Propagation for Games—** *Carl Schissler, Dinesh Manocha*, University of North Carolina, Chapel Hill, NC, USA

We present a sound propagation and rendering system for generating realistic environmental acoustic effects in real time for game-like scenes. The system uses ray tracing to sample triangles that are visible to a listener at an arbitrary depth of reflection. Sound reflection and diffraction paths from each sound source to the listener are then validated using ray-based occlusion queries. Frame-to-frame caching of propagation paths is performed to improve the consistency and accuracy of the output. Furthermore, we present a flexible framework, which takes a small

fraction of CPU cycles for time-critical scenarios. To the best of our knowledge, this is the first practical approach that can generate realistic sound and auralization for games on current platforms.

WORKSHOP 13:15

Sound Internationalization

PAPER SESSION 4: SPATIAL AUDIO 14:15

- 4-1 Synchronizing Auditory and Visual Sensation in Computer-Simulated Environments—** *Sentagi S. Utami, Mojtaba Navvab, Jason Corey*, University of Michigan, Ann Arbor, MI, USA

Application of auditory rendering in simulation and gaming currently focuses in providing information of object category detection such that user's auditory experience during the virtual interaction seems unrealistic. In this paper auralizations of small and large, open and closed spaces with the impact of a variety of architectural elements are synthesized using common techniques in computer-simulation. Virtual spaces were modeled based on localization of critical elements and visualization of the sound field obtained from field-measurement. Noticeable differences of the auditory sensation are measured through the use of subjective evaluation. This integrated method enables to indicate to what extent the simulation modeling is required to create a realistic auditory rendering that enhance the visual sensation.

- 4-2 Surround Sound with Height in Games Using Dolby Pro Logic IIz—** *Nicolas Tsingos,¹ Christophe Chabanne,¹ Charles Robinson,¹ Matt McCallus²*
¹Dolby Laboratories, San Francisco, CA, USA
²Red Storm Entertainment, Cary, NC, USA

Dolby Pro Logic IIz is a new matrix encoding/decoding system that enables the transmission of a pair of height channels within a conventional surround sound stream (e.g., 5.1). In this paper we provide guidelines for the use of Pro logic IIz for interactive gaming applications including recommended speaker placement, creation of elevation information, and details on how to embed the additional height channels within a 5- or 7-channel stream. Surround sound with height is already widely available in home-theater receivers. It offers increased immersion for the user and is a perfect fit for 2-D or stereoscopic 3-D video games.

- 4-3 A Perspective on the Adoption of Ambisonics for Games—** *Andrew J. Horsburgh,¹ Kenneth McAlpine,² D. Fraser Clark¹*
¹University of West of Scotland, Paisley, Scotland, UK
²University of Abertay, Dundee, Scotland, UK

In this paper the history of Ambisonics in computer gaming is presented, then described with a discussion of its current adoption and implementation, concluded by an examination of the role Ambisonic technology might have in the future of games audio. Surround formats, such as Dolby 5.1 Surround, have been used effectively within the computer games sector for over 10 years, with format used almost exclusively in some genres. However, recent developments in processing power and advancements in game-audio engines (such as OpenAL incorporation) it has become possible to use Ambisonics as an alternative surround audio format in all video games. There are currently only a limited number of game titles which use Ambisonic materials, and these have not yet been systematically catalogued.