


AES 35th INTERNATIONAL CONFERENCE

Audio for Games

	WEDNESDAY, 11TH FEBRUARY	THURSDAY, 12TH FEBRUARY	FRIDAY, 13TH FEBRUARY
08:00	Registration (08:15–18:00)	Registration (08:15–18:00)	Registration (08:15–17:00)
09:00	9:00 Introduction to Sound Design for Games	9:00 Paper Session 1: Sound Recording and Foley	9:00 Paper Session 2: Audio Codecs
09:20			
09:40	10:00 Introduction to Music for Games	9:00 Special Event: Techniques for Building Virstual Worlds	
10:00		11:00 Coffee Break	10:00 Demonstration: DirAC-Based Audio Engine
10:20	10:00 Graduate Student Forum		
10:40	11:00 Coffee Break		11:00 Coffee Break
11:00	11:40 Taking Up Space: Understanding and Implementing 3-D Audio	11:40 Keynote Address: Steve Root	
11:20		12:40 Lunch Break	11:40 Special Event: Tools, Techniques, and Tech from Treyarch's "Call of Duty: World at War"
11:40	12:40 Lunch Break		
12:00		14:00 Paper Session 5: Advanced Spatial Audio and Reverb	12:40 Demonstration: Advanced Spatial Audio with OpenAL
12:20	14:20 Tutorial: Configuring 7.1 Studios		
12:40		15:20 Panel Discussion: Education and Training in Game Audio	15:20 Demonstration: DirAC-Based Audio Engine
13:00	15:40 Coffee Break		
13:20	15:20 Audio for Games: An Interactive Tutorial	17:00 Coffee Break	16:00 Paper Session 6: Game Music Systems
13:40			
14:00	16:20 Coffee Break		
14:20	16:40 Dummy's Guide to DSP for Games	17:20 Special Event: Sonic Boom: Discussions into the Future of Audio in Games	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.
14:40			
15:00	17:40 Special Event: Sonic Boom: Discussions into the Future of Audio in Games		
15:20	18:00 Drinks Reception (Open to all conference attendees)		
15:40	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
15:40	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
16:00	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
16:20	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
16:40	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
17:00	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
17:20	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
17:40	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
18:00	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
18:30	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		
19:00	18:30 Social Event: Please join us for food, drinks and gaming at NAMCO STATION. Attendance in this event is free to all participants with a two or three day conference pass.		

AES 35th INTERNATIONAL CONFERENCE

AUDIO FOR GAMES



London, UK • February 11–13, 2009

Wednesday, 11th February

08:15 Registration

INTRODUCTION / TUTORIAL DAY

09:00 Room F4

T-1 Introduction to Sound Design for Games

Presenter: Dan Bardino, SCEE, London, UK

For audio designers and engineers of traditional media, sound for games can often seem like a strange and impenetrable world. This introductory talk will attempt to clarify and demystify the basics of making interactive sound by looking at the basic principles of designing interactive sound for games; both the technical challenges involved and the conceptual creative challenges they bring. From alleviating repetition to creating interactive car engines to understanding the fundamental differences between linear and interactive media and the role sound plays in both.

10:00 Room F4

T-2 Introduction to Music for Games

Presenter: Adam Levenson, Activision, Santa Monica, CA

During the past two decades, original music for videogames has evolved into an art form rivaling or surpassing scoring quality in film and TV. This tutorial provides an overview of the history of game scores, evaluates where we are today, and describes exciting future possibilities for music in interactive media. The discussion will include relevant aspects of original music composition, production, asset management, implementation, the business of game music, and the technology behind the scenes.

COFFEE BREAK

11:40 Room F4

T-3 Taking Up Space: Understanding and Implementing 3-D Audio

Presenter: Scott Selfon, Microsoft, Redmond, WA, USA

High fidelity, nonrepetitive sound effects, and ambience are only the beginning of a compelling in-game sound experience. Titles are increasingly exploring sophisticated real-time sound manipulation to enhance both realism and immersion. This talk will explore these aspects as they relate to the “virtual world” of the game’s universe—dynamic simulation of position, distance, interaction with game geometry, and environmental reverberation techniques will be discussed. This talk will additionally cover simulation challenges and aesthetic considerations particular to interactive gameplay.

LUNCH

14:20 Room F4

T-4 Introduction to Synthesis for Interactive Applications

Presenter: Leonard Paul, Lotus Audio, Vancouver, BC, Canada

There are many promising synthesis methods that have become practical for use on modern consoles due to the advanced processing power of the Xbox 360 and PlayStation 3. Methods such as physical modeling, modal synthesis, granulation, and others are surveyed with examples to give an insight into the leading methods for the generation of sound for today’s games.

15:20

Room F4

T-5 Audio for Games: An Interactive Tutorial

Presenter: Richard Stevens, Dave Raybould, Leeds Metropolitan University, Leeds, UK

This interactive tutorial builds on the introductory lectures earlier in the day and demonstrates the principles of audio for games and illustrates their technical implementation within the Unreal 3 engine. The general principles explored will cover the basics and discuss the implications and challenges of interactivity including repetition, variation, and memory usage. The tutorial also explores methods for approaching musical interactivity. Demonstrations include reactive and adaptive audio; dynamic switching of sound; algorithmic and generative sound; processing, mixing, and controlling sound for aesthetic effect. A demonstration of a project built using these techniques is available for attendees to explore.

COFFEE BREAK

17:00

Room F4

T-6 Dummies Guide to DSP for Games

Presenter: Jason Page, SCEE, London, UK

Moving from hardware to software-based audio in game development has presented a new working paradigm for game developers. This is a shift that Sony Computer Entertainment had to respond to very quickly when creating PlayStation 3 MultiStream Audio Engine. This talk reviews the techniques and optimizations used to implement DSP algorithms for real-time in-game use. The audio team spent a lot of the development time planning ahead for the future needs of game developers. This talk also takes a retrospective look at how game developers are making use of the new tools they have, and examines what lessons can be learned as software audio engines evolve.

DRINKS RECEPTION

Thursday, 12th February

08:15 Registration

PAPER SESSION 1: SOUND RECORDING AND FOLEY

09:00 Room F4

1-1 Generating Meaningful Sound: Quantifying the Affective Attributes of Sound Effects for Real-Time Sound Synthesis—*Karen Collins, University of Waterloo, Waterloo, Ontario, Canada*

Much research has been undertaken to discover what parameters in a musical composition carry emotional meaning. We now take for granted that harmonic content, instrumentation, tempo, timbre, pitch range, and dynamics (etc.) all play some role in music’s affective abilities. However, there has been little research into similar aspects of affect when it comes to sound effects. Though many audio synthesis methods strive for greater realism, realism is not always the most believable sound in multimedia situations. This paper seeks to explore a methodology for research into the affective attributes of sound effects. An understanding of these affective elements can lead to more advanced real-time sound synthesis methods for audio-visual media.

1-2 Localization Cues Affect Emotional Judgments— Results from a User Study on Scary Sound—

Inger Ekman, Raine Kajastila, Helsinki University of Technology, Espoo, Finland

The current paradigm for creating emotional impact in game sound is to carefully choose which sounds to play. This paper takes an alternative approach, suggesting that emotional impact of sounds can be affected by choosing how to play those sounds. We describe a novel concept for emotional sound design—emotional fine-tuning—and show how it is possible to systematically influence the emotional impact of a single sound sample. A controlled user study with eight subjects confirmed that changing the reproduction of a sample so that source localization of the sound is challenged will increase its perceived scariness compared to the same sound with clearly detectable source. The work extends experimental research on emotion perception in sound. It has practical implications for sound design in games and other interactive media.

1-3 Avoiding Tedium—Fighting Repetition in Game

Audio—*Jean-Frederic Vachon*, Artificial Mind & Movement (A2M), Montreal, Quebec, Canada

Historically, one of the biggest problems facing game audio has been the endless repetition of sounds. From sound bites that play constantly, to repeated sound effects to a limited music selection that loops endlessly, players have had every reason to be annoyed at game audio. Despite increased memory budgets on modern consoles, this problem is still relevant. This paper will examine the pros and cons of various approaches used in game audio, as well as the various technologies and researches that might eventually be applied to the field.

PAPER SESSION 2: AUDIO CODECS

09:00

Room F1

2-1 SAOC for Gaming—The Upcoming MPEG Standard on Parametric Object Based Audio Coding—

Leonid Terentiev, Oliver Hellmuth, Johannes Hilpert, Cornelia Falch, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany; *Werner Oomen*, Philips Applied Technologies, Eindhoven, The Netherlands; *Jonas Engdegard*, Dolby Sweden AB

Following the current trend of employing parametric enhancement tools for increasing coding and spatial rendering efficiency, the MPEG audio group performs the standardization activities on Spatial Audio Object Coding (SAOC) technology. The SAOC system extends the MPEG Surround standard by exploiting its rendering capability and decoding efficiency. Moreover, the SAOC technology introduces user-controllable rendering functionality together with flexible choice of various playback configurations. These aspects are of potential interest for a large range of gaming applications that will benefit from the efficient coding and interactive rendering. Although SAOC targets many different application scenarios, this paper will describe the basic SAOC architecture and the manifold of enhancement tools, with a specific focus on the relevance for gaming applications.

2-2 Viable Distribution of Multichannel Audio-over-IP for Live and Interactive “Voice Talent”-Based Gaming Using High-Quality, Low-latency Audio Codec Technology—

Gregory Massey, APT Ltd., Belfast, UK

The delivery of multichannel audio—from monophonic to surround-sound—in real-time over public IP networks for the purpose of interactive, multi-participant, on-line game entertainment presents a serious design engineering challenge to game developers, console architects, content distribution networks, and ISPs. Leveraging audio engineering expertise gained in professional broadcasting and recording studio postproduction, APT has developed a robust and scalable audio codec technology that enables on-line gaming sys-

tems to realize real-time distribution of high-quality audio for immersive, instantaneous audio experiences in multi-player scenarios of the kind where mass audiences engage with voice talent: talking and singing.

2-3 An Analysis of Differences in Playback Levels Among In-Game Sound Levels, Game Start-up Sounds, and Reference Levels of Game Production Studios—

Masataka Nakahara, SONA Corp./ONFUTURE Ltd., Tokyo, Japan; *Kazutaka Someya*, DiMAGIC Co., Ltd., Tokyo, Japan; *Steven P. Martz*, THX Ltd., San Rafael, CA, USA

The audio experience can vary greatly when playing a video game. After establishing a playback level for a reproduction device, the game may still be too loud or too soft. Average reproduction levels can fluctuate from title to title and even within a game. It is important for game production studios to carefully consider how they configure their systems and the effects it may have on consistency in their games. In this paper the authors survey: (1) the in-game recorded levels of 55 opening cut-scenes from video game titles, (2) the output levels of start-up sounds for four kinds of game consoles, and (3) the current reference levels of sound production facilities for video games. As a result, we conclude: (1) that -20.4 dBFS is the average RMS value per channel of in-game sound, so that -20 dBFS is considered to be the useful reference level for video game production; (2) that -16.4 dBFS is the average Leq(A) value from the summed output of in-game sound, and that it has a good relationship with the RMS value; (3) that there is little difference among the average levels in games across genres and across game consoles, and that there are large differences among production companies; and (4) that the level of start-up sounds that are implemented into game consoles are approximately 10 dB lower than the level of in-game sounds in opening cut-scenes.

PAPER SESSION 3: SPATIAL AUDIO

10:00

Room F4

3-1 Building an OpenAL Implementation Using Ambisonics—

Richard Furse, Blue Ripple Sound, London, UK

OpenAL allows 3-D spatialization to be provided from outside game code; new implementations can be “plugged in.” Ambisonics separates an abstract sound field representation from actual loudspeaker feeds and allows decoding to large number of loudspeakers, including arrangements in 3-D. We present the design of an Ambisonic OpenAL driver implemented in three segments: an “environment” managing state, a “renderer” using this state to generate an abstract Ambisonic sound field representation, and a “decoder” that applies whatever decoding algorithm is desired and feeds the available hardware. We also describe techniques to decode for surround stereo and headphones.

3-2 3-D Sound for 3-D Games—Beyond 5.1—

Simon Goodwin, Codemasters Software Company, Southam, UK

7.1, HDMI 1.3, and 3-D loudspeaker configurations could potentially improve game audio beyond the scope of 5.1 cinema surround, but there are practical issues including ergonomics, compatibility, and the lack of standard layouts. The paper explores possibilities and trade-offs and outlines a basis for standardization. It seeks to justify the need for collaboration across the game and audio industry so that consumers benefit, rather than just get more confused. While there are still details that need to be agreed, the paper illustrates a set of approaches that give high quality 3-D with equipment that is fast becoming the new consumer standard, while remaining closely compatible with existing CD and DVD mixes.

3-3 Binaural Reproduction Over Loudspeakers Using in-situ Measurements of Real Rooms: A Feasibility Study—

Simeon Delikaris-Manias, Paul Gillieron Acoustic

Design, London, UK; *Timos Papadopoulos*, Institute of Sound and Vibration Research (ISVR), Southampton, UK

Binaural reproduction over loudspeakers represents a very promising technique for PC and console gaming applications. Its performance could, however, significantly improve by the incorporation of in-situ plant measurement in its implementation. We present such in-situ plant measurement results for various real rooms, and we give a detailed analysis of the optimal choice of parameters for their inversion and of the expected improvements in the performance of the system.

3-4 Localization Quality Assessment in Source Separation-Based Up-Mixing Algorithms—*Dan Barry*, Dublin Institute of Technology, Dublin, Ireland; *Gavin Kearney*, Trinity College, Dublin, Ireland

In this paper we explore the source localization accuracy and perceived spatial distortion of a source separation-based up-mix algorithm. Unlike traditional up-mixing techniques, source separation-based techniques allow individual sources to be separated from the mixture and repositioned independently within the surround sound field. Generally, spectral artifacts generated during the source separation process are masked when the up-mixed sound field is presented in its entirety; however, this can lead to spatial distortion and ambiguous source localization. Here, we use subjective testing to compare the localization perceived on a purposely generated 5.1 presentation and an up-mix (2.0 to 5.1) of the same source material using a source separation based up-mix algorithm.

PAPER SESSION 4: SPEECH PROCESSING AND ANALYSIS

10:00

Room F1

4-1 Vowel-Based Voice Conversion and its Application to Singing-Voice Manipulation—*Yuri Yoshida*, *Ryuichi Nisimura*, *Toshio Irino*, *Hideki Kawahara*, Wakayama University, Wakayama City, Japan

A novel and light-weight voice conversion method is applied to manipulate a singer's identity and singing style in real time. The proposed method is based on a nonlinear spectral morphing method that uses proximity information for vowel templates of the source and the target singing materials. The proposed method is based on the STRAIGHT speech analysis, modification, and resynthesis system, and it yields highly natural manipulated sounds. To deal with the difficulties in applying our vowel-based voice conversion method to singing voices, singular-value decomposition and robust statistical measures are introduced to handle the huge variability of vowel spectra and fundamental frequencies in singing voices. Distance measures for preparing vowel templates and calculating proximity information are designed based on a psychophysical frequency scale, the equivalent rectangular band, ERB_N rate.

4-2 Fast and Reliable F0 Estimation Method Based on the Period Extraction of Vocal Fold Vibration of Singing Voice and Speech—*Masanori Morise*, Kwansei Gakuin University, Nishinomiya Hyogo, Japan; *Hideki Kawahara*, Wakayama University, Wakayama-City, Japan; *Haruhiro Katayose*, Kwansei Gakuin University, Nishinomiya Hyogo, Japan

A fast and reliable fundamental frequency (F0) extraction method is proposed for real-time interactive applications using a singing voice. It is based on a period detection of the vocal fold vibration, so it does not require expensive computation such as STFT or autocorrelation. Parallel processing architecture and a new cost function made this simple idea competitive with state-of-the-art F0 estimation methods. A series of tests using publicly accessible F0 reference databases revealed that the proposed method supersedes conventional methods in terms of speed and accuracy. Finally,

comparative tests using artificial test signals with fast and deep vibrato were conducted to demonstrate the effectiveness of the proposed method in interactive real-time applications for singing sounds.

4-3 Kaleivoiceope: Voice Transformation from Interactive Installations to Video Games—*Oscar Mayor*, *Jordi Bonada*, *Jordi Janer*, Pompeu Fabra University, Barcelona, Spain

A real-time voice transformation technology and its applications are presented in this paper. The technology allows the transformation of a human voice, such as changing gender from male to female or transforming a teenager to an old woman. More exotic transformations are also possible, for instance robotizing the voice or giving the voice an alien character as it was taken from a science fiction film. The technology has been already used for real-time installations in museums and in postproduction applications. Now, it's being adapted to interactive videogames to transform the voice of the user or any of the game characters.

4-4 Natural Transformation of Type and Nature of the Voice for Extending Vocal Repertoire in High-Fidelity Applications—*Snorre Farnar*, *Axel Roebel*, *Xavier Rodet*, Ircam, Paris, France

Natural voice transformation will reduce the need for authentic voices in many situations, ranging from vocal services via education and entertainment to artistic applications. Transformation of one voice to correspond to that of another person has been studied for decades but still suffers from limitations that we propose to overcome by an alternative approach. It consists in modifying pitch, spectral envelope, durations, etc., in a global way. While it sacrifices the possibility to attain a specific target voice, the approach allows the production of new voices of a high degree of naturalness with different sex and age, modified vocal quality (soft, breathy, and whisper), or another speech style (dullness and eagerness). The transformation of sex and age has been evaluated by a listening test.

COFFEE BREAK

KEYNOTE ADDRESS BY STEVE ROOT, CODEMASTERS

11:40

Rooms F1 & F4

Striking a Chord—Collaborative Solutions for Innovative Game Audio

Steve will explore the broad range of disciplines that contribute to game audio, from teaching and R&D through audio middleware and programming to sound design and composition. Using examples, he'll explain how each discipline plays a critical part in making games sing. He will examine some techniques and ideas which he believes could revolutionize the way games sound, and argue that innovation can only truly succeed when the whole band is playing in harmony.

Stephen Root trained at Goldsmith's College and earned his chops arranging and recording with artists including Mike Rutherford, BA Robertson, Burt Bacharach, and Dave Stewart. His composition work on BBC Sci-fi productions in the 1990s, and his admiration for the music of film composers like Jerry Goldsmith and James Horner, led to Steve writing the soundtracks for early PlayStation classics such as Fox Interactive's Alien Trilogy. As Head of Audio at Acclaim Entertainment, he also penned music for Die Hard Trilogy, Fantastic Four, and Forsaken. In 2000 Root joined Criterion Games, an EA Studio based in the UK. Over 8 successful years as Senior Audio Director at EA he built up one of the most highly regarded audio teams in the industry, delivering multi award-winning and critically acclaimed soundtracks for AAA franchises Burnout and Black. He is currently Director for Audio at Codemasters Software Limited. He relishes the challenges of leading a multi-site, multi-discipline team striv-

ing to create innovative and exceptional audio for a wide variety of games on many different platforms.

LUNCH

DEMONSTRATION: ADVANCED SPATIAL AUDIO WITH OPENAL

12:40

Room G1

Presenter: **Richard Furse**, Blue Ripple Sound

This demonstration provides a practical view of the techniques and methods described in the corresponding paper (Session 3-1). This interactive demo will be repeated a number of times so all conference attendees can have the opportunity to visit the demo. Attendees can sign-up for their preferred slot at the conference registration desk throughout the conference.

PAPER SESSION 5: ADVANCED SPATIAL AUDIO AND REVERB

14:00

Room F4

5-1 Precomputing Geometry-Based Reverberation Effects for Games—*Nicolas Tsingos*, Dolby Laboratories, San Francisco, CA, USA

Current games either pre-render reverberation effects into the sound effects or implement them at run-time using artificial reverberation filters. While interactive geometrical approaches can be used for more accurate acoustical modeling, the increased authoring complexity and the additional cost of geometrical calculations still appears to overshadow their potential benefits. This paper presents solutions to integrate off-line geometrical acoustic modeling in game environments. By precomputing image-source gradients for early reflections and directional decay profiles, we can generate location-dependent reverberation effects without storing or accessing the actual geometry at run-time. We render such reverberation effects using a frequency-domain scalable processing approach. In this context, we introduce an efficient prioritization scheme and evaluate alternative transforms for late reverberation processing. Our pipeline enables fine-grain rendering of distance and surface proximity effects and modeling of both outdoor and coupled indoor spaces with arbitrary reverberation decay profiles.

5-2 Generating a Spatial Average Reverberation Tail Across Multiple Impulse Responses—*Rebecca Stewart, Mark Sandler*, Queen Mary, University of London, London, UK

Auralization using impulse responses rendered or measured in non-real-time is often limited to static sound source and receiver positions. An increasingly common approach is to use a large database of previously rendered impulse responses to simulate moving through a modeled space. Described here is a method to remove redundant information from that database of impulse responses. In particular, the late reverberation tail is analyzed and a generalized version is synthesized.

5-3 Acoustic Impulse Response Interpolation for Multichannel Systems Using Dynamic Time Warping—*Claire Masterson, Gavin Kearney, Frank Boland*, Trinity College Dublin, Dublin, Ireland

We present a method of interpolation of room acoustic impulse responses for reducing measurement sets in multichannel convolution systems. The method employs the use of Weighted Dynamic Time Warping for the synthesis of early reflections and critical band analysis for diffuse decay decorrelation. An objective study of the reduced data sets in comparison to full spatial resolution measurements taken in a reverberant environment is presented through binaural

measurements. Listening tests are conducted under controlled conditions for investigation into the perceptual attributes associated with the synthesized audio.

5-4 An Improved Parametric Model for Perception-Based Design of Virtual Acoustics—*Christian Borss, Rainer Martin*, Ruhr-Universität Bochum, Bochum, Germany

In this paper we present an architecture for the generation of artificial acoustics that can be used for plausible as well as artistic virtual environments. In our approach, we create synthetic acoustics derived from few acoustical design parameters like the room geometry, the frequency dependent reverberation time, and the echo density profile. We describe how the intended design parameters are mapped to our signal generation architecture and how audio signals can be rendered for presentation with an arbitrary number of loudspeakers or with headphones.

PANEL DISCUSSION

15:20

Room F4

Education and Training in Game Audio: A Curriculum for a Degree in Audio Design for Game Development

Chair: **Mark Sarisky**, University of Texas, Austin, TX, USA

Panelists: **Sarah LeMarie**, SCEE; **Karen Collins**, University of Waterloo; **Richard Stevens**, **Dave Raybould**, Leeds Metropolitan University

This panel discussion session focuses on education issues related to game-audio.

Topics covered include development of suitable course content to meet the needs of industry and how universities and game developers can develop stronger relationships to build research collaborations. The panel consists of educators from a number of well-known institutions as well as industry representatives who work with academia. The workshop begins with a number of presentations highlighting these issues and then opens into a discussion with the audience about the best ways that we may take this into the future. Pragmatic as well as theoretic aspects of education in the Audio For Games industry will be discussed. Audience participation is strongly encouraged.

DEMONSTRATION: DIRAC-BASED AUDIO ENGINE

15:20

Room G1

Presenter: **Ville Pulkki**, Helsinki University of Technology

Directional audio coding (DirAC) is a method for spatial sound processing, which is in this case applied for spatial audio synthesis, coding, and reproduction for virtual worlds, as presented in accompanying poster. The on-line and off-line demonstrations will show the use of DirAC as audio engine. Sound sources are positioned into virtual world, and DirAC is used to control the auditorily perceived direction, distance and extent of them.

COFFEE BREAK

SPECIAL EVENT

17:20

Rooms F1 & F4

Sonic Boom: Discussions into the Future of Audio in Games

Presenter: **Masaya Matsuura**, NanaOn-Sha, Tokyo, Japan

PaRappa creator Masaya Matsuura brings together a broad range of trends with thought provoking examples to inspire thought on where game audio may head in the near to distant future.

In April 1983 Masaya formed the band PSY'S (pro-

nounced “Size”), and the band pushed the frontiers of computer music for the time. After nine hit albums, PSY’S disbanded in August 1996. That same year, Masaya took Japan by storm with the release of “PaRappa The Rapper” for PS1. The game won the 1996 CECA Award, the Japan Software Award, and was also named Japan’s “Game of the Year” for 1997. Masaya’s 1999 release, “vib-ribbon” (PS1), was another gaming revolution, generating gameplay from the player’s own favorite music CDs. In 2003, Masaya produced and composed sounds for the new Aibo “ERS-7” robot dog, and also released “mojibribbon” (PS2), a game unique in both gameplay and visual style.

2005 saw the release of “Tamagotchi Connection: Corner Shop” (Nintendo DS), which sold more than a million copies in Japan alone. The game’s runaway success led to acclaimed sequels in 2006, 2007, and 2008. Continuing his desire to broaden the appeal of games, in 2007 Masaya also masterminded the release of “musika” for the iPod. He is currently working on the highly anticipated music game “Major Minor’s Majestic March” for the Wii. He is also the first and only Japanese Emeritus member of the Advisory Board for the Game Developer’s Conference (GDC) in the United States.

Friday, 13th February

08:15 Registration

SPECIAL EVENT: WORKSHOP

09:00 Room F4

Techniques for Building Virtual Worlds: Perspectives from Leading Japanese Game Studios

Co-chairs: Steven P. Martz, THX; Bike H. Suzuki, Sona

Presenting Authors: Tetsukazu Nakanishi, Kanako Kakino, Ryuichi Takada, Namco-Bandai; Kazuya Takimoto, Capcom; Eiji Nakamura, Chiharu Minekawa, Kazutaka Someya, DiMAGIC

This workshop evaluates current methods of game sound production in Japan. Panelists will demonstrate techniques and systems used in creating the audio for several cutting-edge Japanese video game titles through a progressive look at the sound design process. Presenters will explain the production flow used in the making of a cut scene and introduce their postproduction strategies in construction of the sonic landscape. The cut scenes will show the progression from a dry mix (no reverb, no processing, etc.) to a fully immersive environment through the viewpoints of a sound programmer, a sound mixing engineer, and an acoustician. Their results will be shown with explanations of the methods used.

DEMONSTRATION: DIRAC-BASED AUDIO ENGINE

10:00 Room G1

Presenter: Ville Pulkki, Helsinki University of Technology

Directional audio coding (DirAC) is a method for spatial sound processing, which is in this case applied for spatial audio synthesis, coding, and reproduction for virtual worlds, as presented in accompanying poster. The on-line and off-line demonstrations will show the use of DirAC as audio engine. Sound sources are positioned into virtual world, and DirAC is used to control the auditorily perceived direction, distance and extent of them.

GRADUATE STUDENT FORUM

10:00 Room F1

Presenter: Rebecca Stewart, Queen Mary, University of London, London, UK

A session for graduate students in research relating to audio for games to meet. It will allow students to discuss their work

in the broader context of their thesis. Supervisors and industry members are welcome to participate in this session.

COFFEE BREAK

DEMONSTRATION: 3-D SOUND FOR 3-D GAMES—BEYOND 5.1

11:00 Room G1

Presenter: Simon Goodwin, Codemasters

This demonstration provides a practical view of the techniques and methods described in the corresponding paper (Session 3-2: “3-D Sound for 3-D Games—Beyond 5.1”). This interactive demo will be repeated a number of times so all conference attendees can have the opportunity to visit the demo. Attendees can sign-up for their preferred slot at the conference registration desk throughout the conference.

SPECIAL EVENT

11:40 Room F4

Tools, Techniques, and Tech from Treyarch’s “Call of Duty: World at War”

Presenters: Blair Bitonti, Adam Levenson, Activision; Stephen McCaul, Treyarch, Santa Monica, CA, USA

This presentation is an overview of the development philosophy, work pipeline, and technological developments for sound in “Call of Duty: World at War.” This includes in-game editing, mixing, voice management, preemptive filtering to prevent masking, simple interactive music, occlusion, and use of surround. This project was marked by the interdisciplinary outlook of the sound department. There were no pure sound designers and even the programmer contributed directly to creative and aesthetic aspects of the game. We feel strongly that sound designers need to take responsibility of their sounds from conception to final in game implementation for the greatest quality of work. This includes scripting, which is often barred from sound designers. Conversely the engineers need to have deep understanding (not just programming) and concern with the sound of the game.

LUNCH

POSTERS

13:00 Room F1

P1-1 An Efficient Implementation of 3-D Audio Engine for Mobile Devices—Frederic Amadu, Jean Michel Raczinski, Arkamys, Paris, France

This paper presents a generic and customizable 3-D audio engine, which has been specially designed for gaming on low-end mobile devices. The engine is based on source and listener 3-D positioning for headphone playback. Distance attenuation, Doppler effect, and reverberation can be added to fit JSR234 specifications. In order to address platform diversity, we have developed a PC application to easily design the best 3-D audio engine in accordance with processor capabilities. Standard HRTF-based processes have been simplified to obtain a limited number of fixed-point IIR filters, which have been successfully implemented on several platforms. Then, objective and subjective validation methods allow us to certify the quality of the porting.

P1-2 Efficient Spatial Sound Synthesis for Virtual Worlds—Ville Pulkki, Mikko-Ville Laitinen, Cumhur Erkut, Helsinki University of Technology, Espoo, Finland

Directional audio coding (DirAC) is a frequency-band processing method for spatial audio based on psychophysical assumptions and on energetic analysis of sound field. The applications of DirAC in spatial sound synthesis for virtual

worlds are presented in this paper. The techniques are independent of the sound reproduction method, which can be any loudspeaker setup or headphones. It is shown that DirAC can be used to position and to control the extent of virtual sound sources, and also to generate reverberation efficiently in virtual worlds.

P1-3 Augmented Reality Audio for Location-Based Games—*Mikko Peltola, Tapio Lokki, Lauri Savioja*, Helsinki University of Technology, Espoo, Finland

Location-based games, such as pervasive games and geocaching, could benefit from the use of audio, in particular spatial sound. In this paper we present a binaural recording and rendering system that is capable of both including location and orientation information to audio files and playing audio content related to location. Altogether we introduce a platform for building augmented reality audio applications suitable for outdoor use. The Audiomemo application is presented to highlight the possibilities of location-based spatial sound for games.

P1-4 How Players Listen—*Simon Goodwin*, Codemasters Software Company, Southam, UK

The games industry currently lacks detailed understanding of the audio configurations that their listeners use to play games. Codemasters surveyed players to investigate the listening systems and configurations they had available and those they preferred. The results of this survey have implications for the way audio assets are prepared, rendered, and mixed in games. Future consumer research is proposed in light of new platforms and audio interfaces.

P1-5 Acoustic DDR: An Automated Test Tool for 3-D Sound Perception Evaluation with Visually Impaired Users—*Hector Szabo, Philippe Mabillean, Bessam Abdulrazak*, Universite de Sherbrooke, Sherbrooke, Quebec, Canada

Common off-the-shelf 5.1 acoustic systems for PC computers are nowadays an affordable price, turning attractive the production of inclusive acoustic games and simulators, oriented to bound blind and non-blind users. The use of contextual 3-D sound beacons as a navigational aid allows user orientation and more sophisticated environments. However, predicting 3-D contextual sound usability and/or playability as navigational aid with COTS equipment constraints can be difficult. We present our experimental setup using COTS 5.1 acoustic and standard PC and early findings, using a 3-D acoustic “Dance Dance Revolution” game to record user response time, precision, and torso/head position when presented to 3-D acoustic stimuli. User performance indicates that higher frequency broadband sounds enhances user aiming to sound virtual azimuth and exposes the need of tactile guides for her or his orientation.

P1-6 A Precise Sound Image Panning Method for Side Areas Using 5.1 Channel Audio Systems—*Keita Tanno, Akira Saji, Shinya Ito, Jie Huang*, University of Aizu, Aizu-Wakamatsu City, Fukushima, Japan; *Wataru Hatano*, Tamura Corp., Tokyo, Japan

5.1 channel home theater systems have been widely used for home audio systems and also for high reality game audio systems. We have conducted two experiments to improve the precision and clarity of sound image creation and reproduction in the side areas. The experimental setup was to change the intensity ratios between the two left loudspeakers, L and SL, and ask the listeners about the directions and clarity of the sound images. From the results, we found the traditional amplitude panning method in the side areas is not linear and asymmetrical, and the motion is almost obtained on the middle range of intensity ratios. Based on the localization curve obtained in this experiment, we can compensate the nonlinearity and asymmetry of sound panning. We also added the frequency characteristics of HRTF to the sound signals assigned to L and SL speakers by amplitude panning method.

These changes of frequency characteristics can increase the reality of the sound signals to the near ear and improve the precision and clarity of sound images in the side areas.

TUTORIAL: CONFIGURING 7.1 STUDIOS

14:20

Room F4

7.1 Game Studio Set-Up and Design

Presenters: Steve Martz, THX, San Rafael, CA, USA; *Dirk Noy*, WSDG, Basel, Switzerland

As game audio hardware has improved over the years, so have the facilities. And if it's time to build a multichannel studio, it likely means 5.1 or 7.1 channels of playback. With so much on the line for each title, this can be a daunting task. This tutorial will address the challenges and decisions that need to be made when creating these environments. We will provide guidelines to practical studio design and proper setup for multichannel configurations. Topics covered include: acoustics, room geometry, subwoofer and surround placement, signal path, reference levels, and bass management.

TUTORIAL: MODAL SYNTHESIS

14:40

Room F1

An Innovative Approach for Game Audio: Using Modal Synthesis to Reduce Memory Consumption and Increase Sound Variations

Presenter: Francois Thibault, AudioKinetic, Montreal

Game developers are constantly faced with having to make sound design trade offs when dealing with sound variety to reduce repetition, and the size of run-time memory available for audio data. Modal synthesis provides a high-performance solution to these very common problems while accurately modeling a wide range of resonant impact type sounds. The main advantages of using modal synthesis are: (1) synthesis transformations allows for a wide range of variations to be created; (2) production costs are reduced, as no sound variations need to be recorded; (3) audio memory footprint is highly reduced, since no dedicated audio files are used as variations and modeled audio data can be compressed more effectively than the original sounds; (4) quality and performance of the synthesis can be scaled dynamically at run-time; and (5) increase interactivity by controlling synthesis parameters based on context-specific (game engine/ physics) run-time parameters.

This tutorial will detail the main components of modal synthesis, which consist of the signal analysis phase, transformation, and the synthesis. This presentation will also include demonstrations of compelling sound transformation examples while exposing pitfalls and advantages of different implementations. Attendees will leave with a clear understanding of the real practical benefits of using modal synthesis as a tool for game audio production. From a technical perspective, attendees will also leave with a good understanding of the inner workings of real-time modal synthesis techniques and offline modal parameter estimation techniques based on linear predictive coding. This session will demonstrate that modal synthesis is realistic and ready for next generation gaming platforms.

COFFEE BREAK

PAPER SESSION 6: GAME MUSIC SYSTEMS

16:00

Room F4

6-1 Building Interactive Networked Musical Environments Using q3osc—*Robert Hamilton*, Stanford University, Stanford, CA, USA

Interactive networked musical gaming environments designed as control systems for external music and sound programming languages can be built using the q3osc Quake

III/quake3 gaming mod. Bidirectional support for the Open Sound Control (OSC) messaging protocol compiled into the game engine allows for the real-time tracking, sonification, spatialization, and third-party control of game entities, clients, and environmental parameters. Reactive audio environments ranging from abstract multi-user musical performance spaces to representative acoustic models of physical space can be constructed using either a standard user-centric audio perspective or a potentially more immersive and inclusive space-centric perspective. Issues of space and perspective are discussed as related to the distribution of performance space and sonified environment across both local and wide-area networks.

6-2 Approaches to Creating Real-Time Adaptive Music in Interactive Entertainment: A Musical Perspective—

Kenneth McAlpine, Matthew Bett, James Scanlan,
University of Abertay Dundee, Dundee, UK

In this paper we discuss the different roles that music plays in an interactive entertainment title, suggesting both creative and procedural approaches to its creation and execution, and in particular, highlighting the importance of procedural music engines to support creative activity. We suggest further, the role that algorithmic and procedural generation routines may have in creating music for interactive entertainment titles in the future and the role that human composers might play in the next-generation game soundtrack.

6-3 Mapping Sounds into Commands—*Giordano Cabral, Roberto Cassio Silva Jr.,* MusiGames Studio, Recife, Brazil

Recent years have witnessed the boom of musical games. These games associate the commands of the player with musical or sonorous events. While the creation and edition of these associations remain a key factor for the musical game industry, digital signal processing techniques continue evolving, providing very useful information about songs. Even though these techniques cannot provide (nowadays) a perfect transcription of songs, they can be successfully mapped into game commands if a proper strategy is applied. This paper discusses some strategies used by MusiGames to answer questions like, “How can the information automatically retrieved be used in a game?” More specifically, “How can this information be used to determine which commands should a player hit on specific moments of a game?”

PAPER SESSION 7: REAL TIME SYNTHESIS

16:00

Room F1

7-1 Design and Evaluation of Physically Inspired Models of Sound Effects in Computer Games—*Niels Bottcher, Stefania Serafin,* Aalborg University, Ballerup, Denmark

A simple audio only game was designed in order to test the possibilities of different interactive sound synthesis techniques simulating aerodynamic sword-like sounds. The Wii remote was used as the controller for generating the sound and this device was connected to the Max/MSP sound synthesis engine. In order to gain a user centered perspective on the potential of the diverse synthesis techniques, in terms of sound quality, interactivity and entertainment value, several techniques to synthesize real time interactive sword-like sounds were implemented. In this test a sample based model was compared to physically inspired modal synthesis, purely perceptually modeled subtractive synthesis and granular synthesis.

7-2 Game Audio Lab - An Architectural Framework for Nonlinear Audio in Games—*Kees Went, Sander Huiberts, Richard van Tol,* Utrecht School of the Arts, Utrecht, The Netherlands

Nonlinear and adaptive systems for sound and music in games are gaining popularity due to their potential to enhance the game experience. This paper is about the Game Audio Lab: a framework for academic purposes that enables research and rapid prototyping of nonlinear sound for games. It enables researchers and designers to map composite variables and adapt sound and music design in real-time during active game play.

7-3 Retargeting Example Sounds to Interactive Physics-Driven Animations—*Cecile Picard, Nicolas Tsingos,* INRIA Sophia-Antipolis, Sophia-Antipolis, France; *Francois Faure,* INRIA Rhone-Alpes, Grenoble, France and *Universite de Grenoble and CNRS,* Grenoble, France

This paper proposes a new method to generate audio in the context of interactive animations driven by a physics engine. Our approach aims at bridging the gap between direct playback of audio recordings and physically-based synthesis by retargeting audio grains extracted from the recordings according to the output of a physics engine. In an off-line analysis task, we automatically segment audio recordings into atomic grains. The segmentation depends on the type of contact event and we distinguished between impulsive events, e.g., impacts or breaking sounds, and continuous events, e.g., rolling or sliding sounds. We segment recordings of continuous events into sinusoidal and transient components, which we encode separately. A technique similar to matching pursuit is used to represent each original recording as a compact series of audio grains. During interactive animations, the grains are triggered individually or in sequence according to parameters reported from the physics engine and/or user-defined procedures. A first application is simply to reduce the size of the original audio assets. Above all, our technique allows to synthesize non-repetitive sounding events and provides extended authoring capabilities.

Sponsored by



audiokinetic

BINARY
SONORI

GENELEC®

SCEE^{OXO}
Research & Development

SpACE-Net
Spatial Audio Creative Engineering Network