I3DL2 and Creative®EAX

Jussi Mutanen
Jussi.Mutanen@hut.fi

Abstract

I3DL2 3D audio rendering guidelines gives the minimum rendering requirements for the 3D audio developers, renderer's, and vendors. I3DL2 defines how the 3D technology is applied to the current PC systems. Creative EAX reverb API is based on these guidelines and hardware acceleration is supported under OpenAL and Microsoft DirectSound3D™. EAX Advanced HD is the next generation of 3D audio rendering API and today only a few hardware and software applications support it.

1 Introduction

Today the objective in a computer game industry is to produce a realistic 3D environment for the player. The visual component of this 3D world has developed in very rapid pace and it is very common to have computers displaying 3D graphics with texture mapping, shading, multiple light sources, but only two tiny speakers playing binaural sound. Virtual acoustic environments can be rendered analogous to visual environments. Ray tracing and radiosity methods are based on the physics of the light propagation and in the same way, sound rendering is based on physical laws of sound propagations and reflection. (Savioja *et al.* , 1999) (Lokki *et al.* , 2002)

3D Application Programming Interfaces (APIs) are basically a set of instructions for game developers for positioning sound sources in 3D space and specifying the environments where they should be played. This paper presents I3DL2 and EAX 3D sound rendering APIs to the reader, what kind of similarities they have, how they could be implemented, and what kind of hardware and software requirements they need to enable realistic 3D audio gaming experience.

2 I3DL2

3D Audio Rendering Guidelines - Level 1.0 (3D, 1998) specification was defined by 3D Audio Working Group (3DWG) of the Interactive Audio Special Interest Group (IASIG) of the MIDI Manufacturers Association in 1998. The goal was to improve the development of platforms for interactive multimedia applications especially in the area of 3D audio.

The guideline was not intented to be Operating System (OS) or API specific, but however, it has some portions referring to specific API, i.e. particularly Microsoft. Level 1.0 guideline presents the general concepts of three-dimensional audio technology, the basics of human hearing, and provides knowledge for evaluating performance of 3D audio. Level 2.0 (3D, 1999) specification is build upon that work and adds more advanced models for environment reverberation, distance, occlusion, and obstruction. The specification is also known by the acronym 'I3DL2'. These guidelines define the minimal rendering requirements, lexicon of terms, and evaluation guidelines for the 3D audio developers, renderer's, and vendors.

2.1 Environment reverberation model

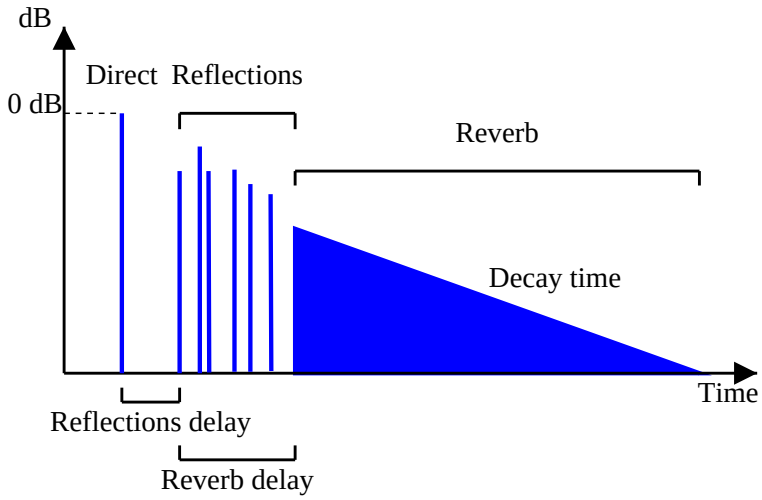

Figure 1: Reverberation response model (3D, 1999).

Reverberation response model The reverberation response model senses the space where the listener is located. This model is divided into three temporal sections: `direct`, `reflections`, and `reverb`. The model is presented in figure 1 (3D, 1999) and it assumes that the listener and the source are in the same environment. Thus all the reflections and the reverberation refer to the listener's room properties.

The reverberation response is characterised by the following parameters:

- Energy in each of three sections: `direct`, `reflections`, and `reverb`.
- Reflection and reverb delays.
- "Direct filter": a low-pass filter for direct component to reduce its energy at high frequencies.
- "Room filter": a low-pass filter for reflections and reverb components to reduce their energy also at high frequencies.

- Decay time at low and high frequencies.
- Diffusion and density of the late reverberation, which is assumed to be diffuse.

The I3DL2 guideline takes also under consideration distance effects and defines notions of `minimum distance`, `maximum distance` and `rolloff factor`. The renderer applies an attenuation to the source signal according to source-listener distance. If `distance` equals `minimum distance`, the attenuation is 0 dB. If `minimum distance <= distance <= maximum distance` and `rolloff factor` is set to 1 the attenuation is -6 dB for each doubling of distance. For values different from 1, `rolloff factor` acts as a multiplier applied to the source-listener distance.

I3DL2 specification assumes that the reflected intensity also decays with increasing distance but not so much as the direct intensity. It can be thought that the reflected sound acts like a plane wave and attenuation occurs only if the intensity is small enough. Thus the level-2.0 guideline has a mean to attenuate the reflected sound intensity according to the source-listener distance in addition to the attenuating the direct component as in Level-1.0. A new parameter for this is used, `room rolloff factor`. Default value for `rolloff factor` is 1, while the default value for `room rolloff factor` is 0. This implies that the reflected intensity does not vary with distance.

EAX 2.0 specification provides also an additional parameter, `air absorption HF` which is expressed in dB per meter, to control attenuation at high frequencies. For example, when increasing `air absorption HF` higher humidity in the air could be simulated.

Spectral effects used in these guidelines are based on the following assumptions:

1. All the effects are specified as simple low-pass filters.
2. A particular filter topology is not used so the filters can be designed with one pole, two pole, etc topology.
3. Content authors should always assume that with a given parameter set the same effect is produced on all end-user platforms.

The guideline interface allows to adjust the cutoff frequency of the low-pass filters, which allows the hardware or software vendors to choose what low-pass filter topology to use. All spectral effects in EAX are controlled in the above manner by specifying an attenuation at a reference high frequency of 5 kHz. It enables cascading several low-pass filter effects and still the resulting attenuation at 5 kHz is predictable.

Environment definition and presets The environment reverberation response contains the all sound sources in the `environment`. The `environment` is characterised by the all reverberation response parameters when `distance` equals `minimum distance`. This is called as `environment preset`.

Only the energies in the three temporal sections `direct`, `reflections`, and `reverb` varies with distance. At the `minimum distance` the intensity level of the direct component is the same as the source signal, i.e. no attenuation has occurred respect to the source signal. The reverberation response model also assumes that `decay time`, defined as the time when the reverberation has decayed 60 dB from its onset, is the same at all the frequencies.

Diffusion, density, and room size The output of different reverberation units are usually highly vendor or algorithm dependent. Thus the use of parameters like `density`, `diffusion`, or `room size` are very problematic. However, these parameters could not be left out because it would then significantly reduce the functionality offered to the content creator. The guideline includes these parameters as normative properties, expressing the values in percents. Absolute values for the echo density or modal density are not given. The guideline suggest that `diffusion` of 100% produces a smooth reverberation decay and a reduced `diffusion` makes the reverberation more "grainy". With 100% `density` a natural-sounding reverberation timbre is produced. In the same way a reduced `density` makes a much more "coloured" sound. The renderer should ensure that the scale 0 - 100% maps lineary to the perceived quality. The one problematic parameter, `room size` affects to the all other parameters so its definition remains highly vendor dependent.

2.2 Interaction between sound sources and environment

Source processing model The processing model for each sound is presented in figure 2 (3D, 1999). For each sound source an attenuation and low-pass filter are applied for the direct and for the reflected part separately.

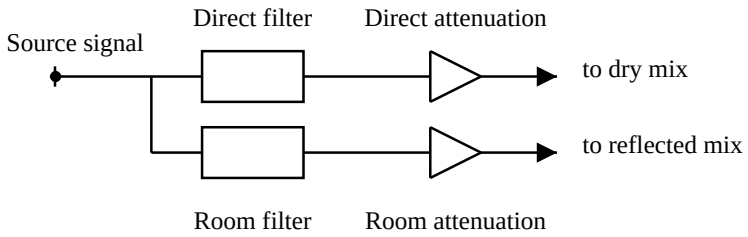

Figure 2: Source processing model (3D, 1999).

Orientation and directivity A sound with no orientation has the same intensity at the same distance in every direction. A directive sound source can be presented with a concept of *Sound Cones*, see figure 3 [1]. The sound source radiates its maximum intensity to the inner cone and minimum intensity to the outer cone. Between these two

[1] http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dx8_vb/directx_vb/htm/_dx_sound_cones_dxaudio.asp

cones there is a transition zone where the intensity decreases when the angle increases. The sound source directivity can be altered by changing the angles of the inner and outer sound cones.

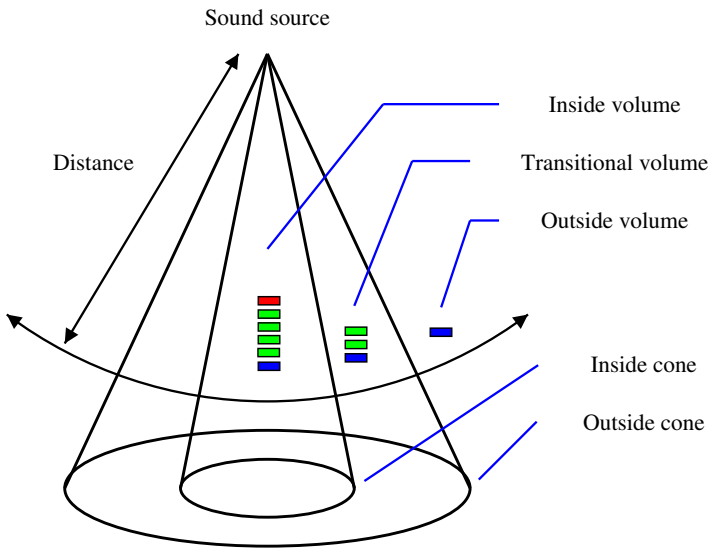

Figure 3: Concept of *Sound Cones*.

Occlusion and obstruction The occlusion and obstruction properties enable to reproduce sound effects from the different kind of obstacles between the source and the listener, see figure 4. The sound from the source can be altered as it sounds from the other room, behind the door, around the corner, or just muffled by an obstacle. Diffraction is also possible. Diffraction occurs when the sound travels around corners or through openings and a low-pass attenuation is applied. The amount of the diffraction depends on the size of the obstacle and the wavelength of the sound. As a rule of thumb, the greater the angle of the diffracted sound the greater the applied high-frequency attenuation.

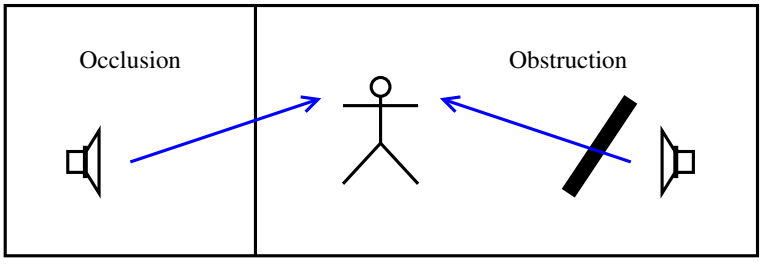

Figure 4: Occlusion and obstruction

In generally when the sounds are transmitted through material structures a frequency dependent low-pass attenuation occurs. The amount of attenuation and the

slope of the high-frequency roll depend on the used material, its thickness, and the overall construction. A set of predefine material presets help the programmer to setup proper occlusion and obstruction properties. Table 1 (3D, 1999) introduces all the eight predefined material presets.

Table 1: Predefined material transmission presets.

Material	Attenuation (mB)	Low and high frequency ratio
Single window	-2800	0.71
Double window	-5000	0.40
Door, thin	-1800	0.66
Door, thick	-4400	0.64
Wooden wall	-4000	0.50
Brick wall	-5000	0.60
Stone wall	-6000	0.68
Curtain	-1200	0.15

The occlusion and obstruction effects are of similar nature because in both cases the sound undergoes a variable amount of low-pass attenuation. However, in the case of obstruction the reflected sound remains unaffected while it is affected in the case of occlusion. The both effects are implemented using simple low-pass filters combined with frequency independent attenuation. In the other words the low-frequency ratio is relative to the main attenuation value. The biggest advance in this kind of solution is that the both occlusion and obstruction filters can be controlled via a single knob. An example of the two topology occlusion filter specified by I3DL2 (3D, 1999) is shown in figure 5 (3D, 1999).

2.3 Minimum requirements

This section presents the minimum rendering requirements of the 3D positional audio. The following functions has to be supported if classified as an IASIG Interactive 3D Audio renderer.

- Playback of simultaneous sources
  * Level-1.0: 8, recommended 16. Level-2.0: 16, recommended 32.
  - a minimum sample rate of 22050 Hz at 16-bit resolution.
  - object and listener 3D position (x, y, z).
  - object and listener velocity.
  - listener orientation.
  - object orientation and radiation pattern (*Sound Cone*).
  * Only in Level-2.0:

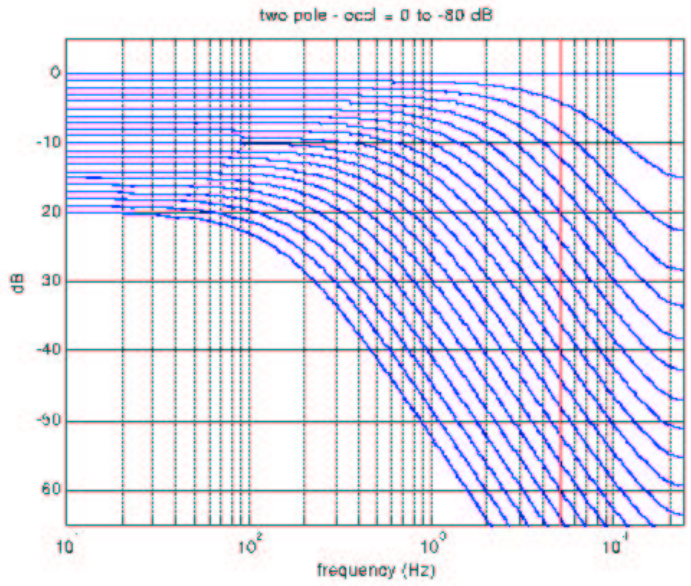

Figure 5: Two pole implementation of the occlusion filter (3D, 1999).

- object and listener dynamic reverberation effects.
- individual low-pass filter and attenuation applied to direct path sound for each sound source.
- individual low-pass filter and attenuation applied to reflected sound for each sound source.

- Rendered effects
  - distance (attenuation and reverberation).
  - Doppler.
  - Cartesian 3D positions (x, y, z).
  - radiation model.

  * Only in Level-2.0:

  - room effect including early reflections and surrounding reverberation matching the reverberation response model.
  - occlusion and obstruction.

- All rendering is done in real-time and interactive without audible artifacts and with low latency.

3 Creative®EAX

Creative®'s Environmental Audio Extensions (EAX) is an open standard for 3D sound reproduction and it takes advantage of any hardware accelerated sound card with help of EAX Unified Interface (see section 3.3). EAX adds enhanced 3D reverberation models to the DirectSound™component of DirectX or to OpenAl.

3.1 EAX 1.0/2.0

EAX 1.0/2.0 specifications are based on IASIG 3D Audio Rendering Guidelines. Microsoft DirectSound™release 5.0 of DirectX (DS3D) [2] can be extended by using additional Property Sets. EAX 2.0 (Creative, 2001c) comprises two different property sets, one for the listener and one for the sound source. It provides relatively simple APIs to the programmer for creating sound sources, adjust sound source positions and velocities in 3D world. These kind of environment audio parameters can be set through specific properties. The environment audio parameters are divided into two separate layers, a high-level and a low-level, which provides customised control of the aural sensation. To help the implementation of EAX technologies and functionalities an Environmental Audio Graphical Librarian Editor (see section 3.4) (EAGLE) was developed for the programmers.

High-level control With the high-level properties excellent environmental audio effects can be implemented with only very little programming effort. For example, with `environmental preset` property a type of room or hall could be easily chosen. When the room has been chosen and specified the EAX engine automatically sets all the necessary low-level properties. The low lever parameters specify the exact room size, amount of reflections, amount of reverberation, and so on. All the predefined `environmental presets` are listed in table 2 (3D, 1999).

Table 2: Predefined environmental presets.

Default	Arena	Plain
Generic	Hanger	Parking lot
Cell, padded	Carpet hallway	Sewer pipe
Room	Hallway	Underwater
Bathroom	Stone corridor	Small room
Living room	Alley	Medium room
Stone room	Forest	Large room
Auditorium	City	Medium hall
Concert hall	Mountains	Large hall
Cave	Quarry	Plate

[2] http://www.microsoft.com/windows/directx/default.asp

Low-level control With low-level control functionality a specific property can be exactly adjusted and controlled. Thus a complete control of environmental audio parameters is achieved. Reverberation decay time, reflection and reverberation delays, and reverberation diffusion can be set by the low-level properties. The low-level properties are usually automatically controlled and set by the high-level properties, e.g. with use of `environmental presets`.

3.2 EAX Advanced HD™

EAX Advanced HD™(Creative, 2002) is the third generation EAX 3D audio rendering API providing an increased 3D audio rendering performance, but requiring also an efficient and powerful audio processing architecture, e.g. Creative's Sound Blaster®Audigy 2[3]. EAX Advanced HD™provides all the environment parameters for reflection and delay control, environment size adjustment, and obstruction and occlusion effects just like in previous versions of EAX but also a couple of new features.

Multi-Environment Both the local and as well as all the surrounding environments affect to the listening experience. With the multi-environment technology multiple audio environments can be rendered in real-time enabling rich and complex audio landscapes.

Environment Filtering The environment filtering enhances the sound propagation model presented in EAX 2.0 with high-pass filter components. With high-frequency filters and sound processing more detailed simulations of acoustics spaces are possible and the determination of the source distance becomes more accurate.

Environment Morphing The environment morphing simulates a continuous and seamless audio transition when listener moves from one environment to other. EAX 1.0 and 2.0 handled this kind of transition referring all the reflection and reverberation parameters to the listener's room. If the source is located in a different environment than the listener realism and interactivity is limited because reverberation from the source has to be ignored or pre-rendered to the source. The environment Morphing interpolates each value of any two environment's properties producing a gradual transition.

Environment Panning The environment panning is related to the sound source localisation and position in the 3D scene. In the environment panning, not only the sound sources but the whole environments can be steered and positioned in multi environment landscapes. This concretely means that echoes of a train coming from a tunnel could be heard so that the acoustic properties of the tunnel is correctly positioned in the 3D space.

[3] http://www.soundblaster.com/products/audigy2/

Environment Reflections The environment reflections simulates early reflections and echoes from the walls, floors, and other surfaces. It gives more details and realism in 3D gaming. For example, it enables effects such as sound of a car engine bouncing off adjacent walls so that the primary sound and the latter decays can be heard.

Music Technologies EAX Advanced HD^{TM}also introduced a couple of new technologies to improve the media playback of all types of media. Audio Clean-up supports noise reduction and click removal from the digital audio content. Audio Clean-up improves the playback quality of the old vinyl recordings, tapes, and LPs.

Creative Multi-Speaker Surround 3D (Avendano & Jot, 2002) (CMSS 3D^{TM}) is a technology for up-mixing stereo content into the multi-channel surround. EAX Advanced HD^{TM}improves CMSS deployed with SoundBlaster Live![4]. CMSS 3D is capable of up-mixing stereo content up to 6.1 surround. CMSS 3D places the center channel signals like vocal and speech dialogue to the center channel and separating the realistic front channels to the main speakers leaving the surround signals for the rest of the speakers. C. Avendano presented CMSS up-mixing algorithms in the Audio Engineering Society (AES) 22nd international conference on virtual, synthetic, and entertainment audio. (Avendano & Jot, 2002)

Smart Volume Management (SVM^{TM}) uses a dynamic compression algorithm to equalise loudness differences within songs. This means that quiet parts of the song are made more audible while the louder parts less dynamic. Also the speed of the music playback can be altered with Time Scaling without affecting the pitch. For example Time Scaling enables to slow down a composition so that the difficult solo can be learnt note by note.

3.3 EAX Unified Interface

The EAX Unified Interface (Creative, 2001b) ensures that the EAX application properties can be supported on any PC platform. The Unified Interface searches for a previous version of EAX or a compatible standard and makes the translation of EAX function calls to the most capable interface found. It tries to preserve as much as possible of the EAX features. The current version of the interface provides the translation between EAX 2.0 and EAX 1.0.

3.4 EAGLE 2.0

EAGLE (Creative, 2001a) is a tool for sound designers for creating realistic 3D aural worlds for games. With EAGLE design and audition of environmental models, obstacle behaviours, and sound source position can be made intuitive. Ready made 3D geometry levels can be imported in 3D Studio Max, Microsoft DirectX, and Lightwave 3D Object formats, which makes the 3D landscape generation easy. A custom plugins

[4] http://www.soundblaster.com

for importing 3D scenes can be also written. The imported level geometry is then divided into separate acoustical sections, where all the environment and obstacle models are parsed together and finally saved in an Environmental Audio Library (.EAL) file.

4 3D sound and reverb engines

The sound card uses a 3D sound and reverb engine[5] for positioning the sound sources and creating the reverb. The used sound engine determines what APIs (e.g. DS3D, EAX 1.0, EAX 2.0) to support. CPU usage, maximum number of 3D streams, and the overall sound quality varies according to the used sound card/chip. However, the positioning of the sound sources and the quality of the reverb remains the same even if the sound cards are changed if the same 3D sound engine is used.

Nowadays there are three 3D sound engine companies (QSound, Sensaura, and Creative Labs) in the PC market. QSound's Q3D and Sensaura's S-3DPA sound engine can be licenced by any chip manufacturer and used in their products. Creative Labs's 3D sound engine is available as a part of a drivers for their products. Previously there was a one 3D sound engine by Aureal but this engine is now a property of Creative Labs.

Creative Labs Creative has separate 3D engines for the SoundBlaster, SoundBlaster Live!, and Audigy sound cards. EAX 1.0 and 2.0 hardware acceleration is supported under OpenAL and Microsoft®'s DirectSound3D™API. SoundBlaster 16, 64, and 128 PCI sound cards are based on ES137x chip, whereas EMU10K1 audio DSP (Digital Signal Processor) is the heart of the SoundBlaster Live! sound cards. I3DL2 is not supported and likely it won't never going to be a issue. Audigy sound engine is the only one to support EAX Advanved HD™.

Linux drivers for the SoundBlaster Live! and Audigy cards can be found from the Creative's open source web site[6]. EMU10K1 drivers for Alsa are also under development Alsa[7]. SoundBlaster Live! card is also available for the Apple Macintosh computers with EAX 1.0 and 2.0 support.

QSound Q3D QSound[8] differs from the other sound engines that it does not use Head Releated Transfer Functions (HRTF) or cross-talk cancellation methods for two speaker playback. Instead of HRTF and cross-talk cancellation method QSound developed specific audio filters. It took more than 550 000 listening tests before the filters were finally chosen. QSound reverb engine, QEM, supports EAX 1.0 and 2.0 but lacks the support for I3DL2 and EAX Advanced HD.

[5] http://www.3dss.com/features/articles/3DSoundEngines/3DSoundEngines.html
[6] http://www.opensource.creative.com
[7] http://www.alsa-project.org
[8] http://www.qsound.com

Sensaura S-3DPA Sensaura's[9] S-3DPA is not the most famous engine but it is widely used. Several top of the line sound cards use it, e.g. Turtle Beach's Santa Cruz and Terratec's Six Pack 5.1+. Sensaura's 3D engine offers support for I3DL2, EAX 1.0 and 2.0 APIs. It also lacks the support for EAX Advanced HD. Currently, it is difficult to tell how major issue this is, because only a few games with EAX Advanced HD support have been released. For example, games such as Mafia and Soldier of Fortune II: Double Helix support EAX Advanced HD. Sensaura's GameCODA is a multi-platform, Sony PlayStatuib®2, Nintendo GameCube, Microsoft Xbox, and any PC running Microsoft Windows OS, audio solution with a low level game sound API.

5 Conclusion

3DWG of IASIG specified I3DL2, 3D Audio Rendering Guidelines for 3D audio developers. I3DL2 guideline defines the minimal rendering requirements and it has models for environment reverberation, distance effects, occlusion, and obstruction. EAX is an implementation of these guidelines made by Creative Labs. Today EAX 2.0 is widely supported by all the 3D engines in the market. Both IASIG and Creative Labs realised that without the multi-environment morphing the listening experience in the border of two differently behaving virtual environments is unrealistic. Creative EAX Advanced HD is today the only 3D sound rendering API to support multi-environment morphing but only a few newest games support it.

REFERENCES

3D Working Group. 1998. Interactive 3D Audio Rendering Guidelines, Level 1.0. *Interactive Audio Special Interest Group*, June, 27. revision 1.0, http://www.iasig.org/wg/closed/3dwg/3dl1v1.pdf [Referenced 21.1.2003].

3D Working Group. 1999. Interactive 3D Audio Rendering Guidelines, Level 2.0. *Interactive Audio Special Interest Group*, Sept., 25. revision 1.0a, http://www.iasig.org/wg/closed/3dwg/3dl2v1a.pdf [Referenced 21.1.2003].

Avendano C., & Jot J-M. 2002. Frequency Domain Techniques for Stereo to Multichannel Upmix. *22nd AES International Conference on Virtual, Synthetic and Entertainment Audio*, July 15-17, 121–130.

Creative. 2001a. EAGLE 2.0 Tutorial. Version 2.0, http://developer.creative.com/games/files/EAGLE2Tutorial.zip [Referenced 21.1.2003].

Creative. 2001b. The EAX Unified Interface. Version 1.0, http://developer.creative.com/games/files/eaxunified.exe [Referenced 21.1.2003].

[9] http://www.sensaura.com

Creative. 2001c. Environmental Audio Extensions: EAX 2.0. Version 1.3, http://developer.creative.com/games/files/EAX20SDK.exe [Referenced 21.1.2003].

Creative. 2002 (Nov.). *Homepage of Creative EAX*. http://eax.creative.com.

Lokki Tapio, Savioja Lauri, Väänänen Riitta, Huopaniemi Jyri, & Takala Tapio. 2002. Creating Interactive Virtual Auditory Environments. *IEEE Computer Graphics and Applications*, **22**(4), 49–57.

Savioja Lauri, Huopaniemi Jyri, Lokki Tapio, & Väänänen Riitta. 1999. Creating Interactive Virtual Acoustic Environments. *J. of the Audio Engineering Society*, **47**(9), 675–705.