Knud Bank Christensen and Thomas Lund
TC Electronic A/S
DK-8240 Risskov, Denmark

Presented at
the 107th Convention
1999 September 24-27
New York

AUDIO
ES

AES

This preprint has been reproduced from the author's advance manuscript, without editing, corrections or consideration by the Review Board. The AES takes no responsibility for the contents.

Additional preprints may be obtained by sending request and remittance to the Audio Engineering Society, 60 East 42nd St., New York, New York 10165-2520, USA.

AN AUDIO ENGINEERING SOCIETY PREPRINT

Room Simulation
for Multichannel Film and Music

KNUD BANK CHRISTENSEN AND THOMAS LUND

TC Electronic A/S
Sindalsvej 34, DK-8240 Risskov, DENMARK

To fully exploit new Reverb and Spatial algorithms for 5.1 and 7.1 environments in music and film, effect mixing procedures may have to be changed slightly.

The paper will describe production techniques for positioning of sources and listener in virtual rooms, and an algorithm structure to achieve this.

0. INTRODUCTION

Typical production techniques involve only one or two inputs to the room simulation system, thereby limiting the precision of source positioning to be only a matter of send level differences and power panning. This "one source - one listener" model is not very satisfying when producing for mono or stereo, but even worse when the reproduction system is multichannel.

Multichannel recording and reproduction is an opportunity for the production engineer to discriminate deliberately between scenes or instruments heard from a distance, and sources directly engaging the listener.

For film work, engaging audio has a very pronounced effect for stimulating the viewer emotionally, and may therefore significantly add to the illusion presented by the picture.

In the search of more authenticity in artificial room generation, long term studies of natural early reflection patterns have led us to propose new production and algorithm techniques. Using ray tracing in conjunction with careful adjustments by ear, we have achieved simulation models with higher naturalness and flexibility, which is the basis of true source positioning.

The paper will discuss two aspects of precise room simulation for multi source, multichannel environments to cover distant and engaged listening:

- Present different production techniques
- Describe an algorithm structure to achieve the objectives

1

1. SINGLE SOURCE REVERB

By having only one or two inputs in a room simulator, the rendering is based upon multiple sources sharing the same early reflection pattern, and therefore it is not really convincing.

In the real world, all actors or instruments are not piled up on top of each other.

1.1 Music Production

In many studios, one good reverb is used to render the basic environment of a particular mix. One aux send, set at different levels on the different channels, is used to obtain depth and some complexity in the sound image.

To obtain a sound image of a higher complexity and depth, several auxes and reverbs have normally been used. Tuning of the levels, pans and reverb parameters in such a setup may be very time-consuming.

For effect purposes, anything goes, but if the goal is a representation of a natural room or a consistent rendering of a virtual room, it may be hard to achieve using conventional reverbs.

1.2 Film and Post Production

For applications where picture is added to the sound, several psychological studies have proven audio to be better at generating entertainment pleasure and emotions than visual inputs. When it comes to counting neurological synapses to the brain, vision has long been known to be our dominant input source. However, a study by Karl Küpfmüller [4] has suggested, that stimulation of even our conscious mind is almost equally well achieved from visual compared to auditive inputs.

Sense	No of Synapses	Conscious input, bps
Eye	10.000.000	40
Ear	100.000	30
Skin	1.000.000	5
Smell	100.000	1
Taste	1.000	1

Stimulation of conscious mind [4]

Realism in audio is just as important when it is accompanied by picture.

2

In multichannel work for film, several reverbs configured as mono in - mono out are often used on discrete sources. By doing so, the direct sound and the diffused field are easy to position in the surround environment. The technique is therefore especially effective for point source distance simulation.

As an alternative, several stereo reverbs are used on the same sources to achieve a number of de-correlated outputs routed to different reproduction channels.

With both approaches, adjustments can be very time-consuming, and a truly engaging listening experience is difficult to achieve.

2. MULTIPLE SOURCE ROOM SIMULATION

To obtain the most natural sounding and precise room simulation, an artificial reverb system should be based upon positioning of multiple sources in a virtual room. Each source should have individual early reflection properties with regards to timing, direction, filtering and level.

We have found this to be true for both stereo and multichannel presentations.

If the target format is 5.1, at least two directional configurations should exist in the room simulator, namely for home (110 degree surround speakers) and theatre (side array surround speakers) reproduction.

The room simulator should also be flexible enough to easily adopt to new multichannel formats, e.g. the Dolby EX scheme.

By changing the production technique slightly, multiple sends from e.g. the Auxes, Group busses or Direct outs of the mixing console can be used to define several discreet positions as inputs to the room simulation system.

From a production point of view, multiple source room simulation can be configured two ways, as described below. Any large scale console build for stereo production can adapt to both routing schemes.

3

2.1 The Additive Approach

The conventional approach to reverb is additive. Dry signals are fed to the reverb system, and wet-only signals are returned and added at the mixer.

With a multiple input room simulator, this configuration works much better than with an single source reverb, because at least each source can be approximated to fit the nearest position rendered. However, normal power panning still needs to be applied in the mixer.

An even more precise rendering can be achieved using the integrated approach described below.

2.2 The Integrated Positioning Approach

The sources in a mix needing the most precise positioning and room simulation, should be treated this way:

The source is completely positioned and rendered into a precise position by passing the dry signal through the simulation system, from which a composite output from a number of source positions are available.

XY positioning to any target format, stereo or multichannel, will be rendered as a best fit. The positioning parameters (replacing conventional power panning) can be controlled from a screen, a joystick or discrete X and Y controls.

With all positioning done in the room simulator, consoles made for stereo production may thereby overcome some of their limitations.

3. ALGORITHM STRUCTURE

This part of our paper describes a generic algorithm currently in use for Multichannel Room Simulator development. It is not a description of any particular present or future product, but rather a presentation of the framework and way-of-thinking that has produced our latest Room Simulation products and is expected to produce more in the future.

3.1 Design conditions

The overall system requirements can be stated as follows:

- The system must be able to produce a natural-sounding simulation of a number of sources in acoustic environments ranging from "phone-booth" to "canyon"

- The system should not be limited to simulating natural acoustics: Often quite unnatural reverb effects are desired, e.g. for pop music or science fiction film effects.

- The system should be able to render the simulation via a number of different reproduction setups, e.g. 5.1, 7.1, stereo etc.

- The system should be modular so that new rooms, new source positions in existing rooms, new source types or new target reproduction setups can be added with minimal change to existing elements.

- The system should be easily tuneable: In our experience, no semi-automatic physical modeling scheme, however elaborate, is likely to produce subjective results as good as those obtained by skilled people tuning a user-friendly, interactive development prototype by ear.

Fortunately there are a few factors that make the job easier for us:

- There are no strict requirements for simulation accuracy: Certainly not physical accuracy (the sound field around the listener's head), and not even perceptual accuracy (the listener's mental image of the simulated event and environment). The listener has no way of A/B switching between the simulation and the real thing, so only **credibility** and **predictability** counts: The simulation must not in any way sound artificial, unless intended to, and the perceived room geometries and source positions should be relatively, but not absolutely, accurate.

- Moore's Law is with us. The continual exponential growth in memory and calculation capacity available within a given budget frame has two effects: It constantly expands the practical limits for algorithm complexity, and it makes it increasingly feasible to trade in a bit of code overhead for improved modularity, tuneability, etc.

- There are physical modeling systems readily available, which may provide a starting point for the simulation.

3.2 Block diagram

The overall block diagram of the Room Simulator is shown in fig. 1. As often seen, the system is divided into two main paths: An early reflections synthesis system consisting of a so-called Early Pattern Generator (EPG) for each source and a common Direction Rendering Unit (DRU) that renders the early reflections through the chosen reproduction setup. And a Reverb system producing the late, diffuse part of the sound field. Note that – contrary to what is normally the case – there is no direct signal path. The dry source signals are merely 0th order reflections produced by their respective EPGs. In the following, a more detailed description of the individual blocks is given.

3.3 Early Pattern Generators

Each EPG takes one dry source input and produces a large set of early reflections, including the direct signal, sorted and processed in the following "dimensions"

Level
Delay
Diffusion
Color
Direction

The Level and Delay dimensions are easily implemented with high precision, the other 3 dimensions are each quantized into a number of predefined steps, for instance 12 different directions. Normally, the direct signal will not be subjected to Diffusion or Color. The quantization and step definition of the Direction dimension must be the same for all sources, because it is implemented in the common Direction Rendering Unit. Physical modeling programs such as Odeon [1] may provide an initial setting of the EPG.

3.4 Direction Rendering Unit

The purpose of this unit is to render a number of inputs to an equal number of different, predefined subjective directions-of-arrival at the listening position via the chosen reproduction setup, typically a 5-channel speaker system. Thus, the DRU may be a simple, general panning matrix, a VBAP [2] system or an HRTF- or Ambisonics-based [3] system.

3.5 Reverb Feed Matrix

The reverb feed matrix determines each source's contribution to each Reverberator input channel. Besides gain and delay controls, some filtering may also be beneficial here.

3.6 Reverberator

To ensure maximum de-correlation between output channels, each has its own independent reverb "tail" generator. Controllable parameters include:

Reverberation time as a function of frequency $T_r(f)$
Diffusion
Modulation
Smoothness

We take particular pride in the fact that our "tail" can achieve such smoothness in both time and frequency, and that modulation may be omitted entirely. This eliminates the risk of pitch distortion and even the slightest Doppler effect, which tends to destroy focus of the individual sources in a multichannel room simulator.

Again, an initial setting of $T_r(f)$ may be obtained from Odeon.

3.7 Speaker Control

This block is by default just a direct connection from input to output. But it may also be used to check the stereo- and mono compatibility of the final simulation result by applying a down-mixing to these formats. Also it provides delay- and gain compensation for non-uniform loudspeaker setups, which may also - as a rough approximation - be used the other way around to emulate non-uniform or misplaced setups and thus check the simulation's robustness to such imperfections.

4. CONCLUSION

The system described above is evidently a very open system under continual development. At the time of writing these words, our test system is running in real time on a multiprocessor SGI server with an 18-window graphical user interface providing interactive access to approximately 2000 low- and higher-level parameters. However, this is not the time or place to go into more details. When this paper is presented at the 107th AES Convention in about 4 months, we will have more real life experience with the system.

If integrated positioning is used with multi-source room simulation, our experiments have already shown how much there is to gain in terms of realism and working speed. But even with the less radical additive approach, virtual rooms may be rendered more convincingly with multi-source simulators.

For applications where picture is added to the sound, the most stimulating source will be one, where audio and video are treated with equal attention to quality and detail.

The new possibilities available from multi-source room processors may be exploited to generate a real quality improvement at the end listener, especially when his reproduction system is multichannel.

More convincing sound generates more convincing picture.

REFERENCES

[1] http://www.dat.dtu.dk/~odeon/

[2] Ville Pulkki: "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", JAES Vol. 45, No. 6, pp. 456, 1997.

[3] Jérôme Daniel, Jean-Bernard Rault & Jean-Dominique Polack: "Ambisonics Encoding of Other Audio Formats for Multiple Listening Conditions", AES Preprint no. 4795, 1998.

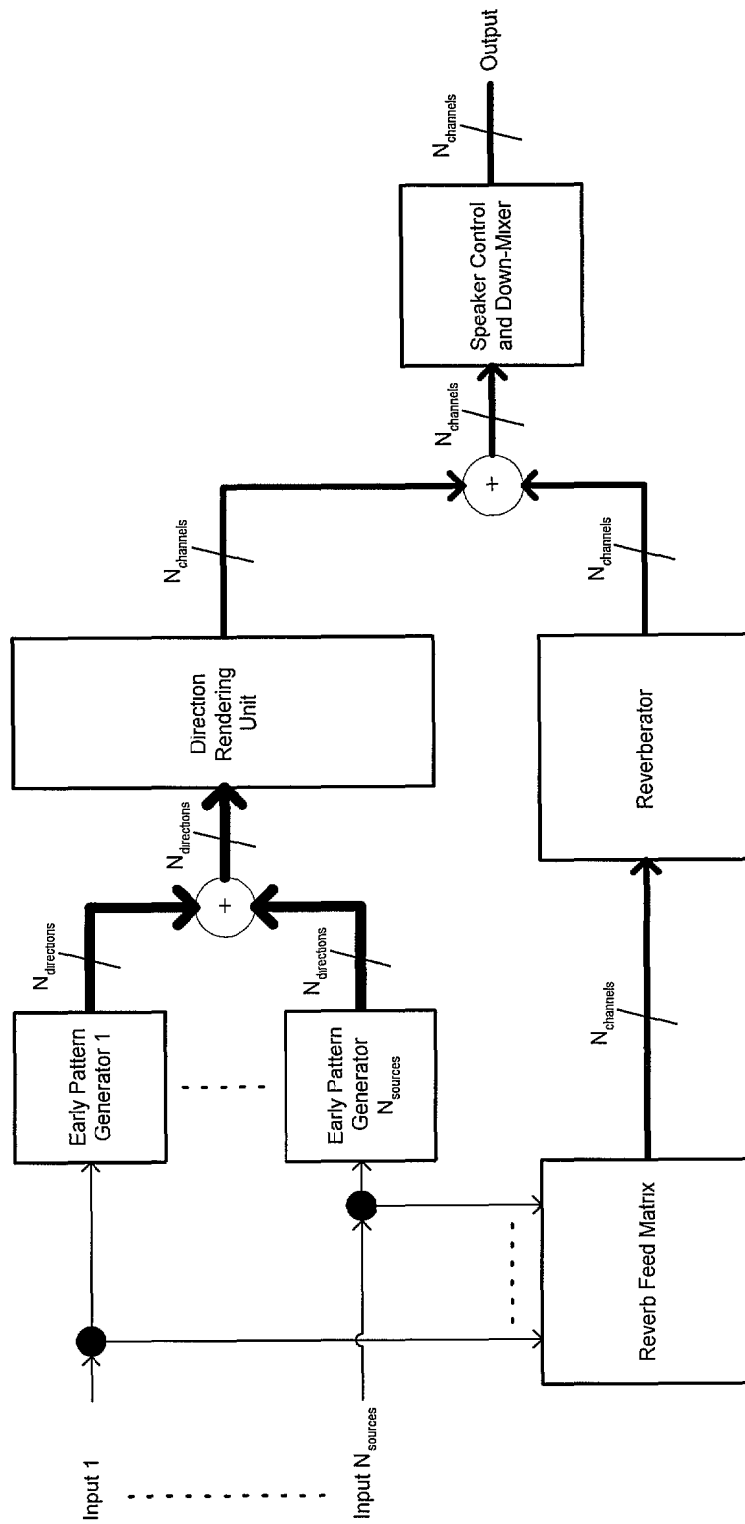[4] Küpfmüller, Karl: "Nachrichtenverarbeitung im Menschen", University of Darmstadt, 1975.

Fig 1

Overall block diagram of Room Simulation Algorithm.