Presented at the 96th Convention of Audio Engineering Society,
Amsterdam 1994 February 26 - March 01 Preprint 3860

THE USE OF A DIGITAL AUDIO MAINFRAME FOR ROOM ACOUSTICAL AURALIZATION

Jens Holger Rindel, Claus Lynge
The Acoustics Laboratory
Technical University of Denmark, Lyngby, Denmark

Graham Naylor
Oticon a/s Eriksholm
Snekkersten, Denmark

Kim Rishøj
T.C.Electronic
Brabrand, Denmark

Abstract

A new user-friendly method for auralization has been developed by interfacing a digital audio mainframe and a room acoustical computer model. Instead of using a convolution technique, early reflections are created by time delays and late reflections by an electronic reverberation processor. The system is verified through measurements of a number of room acoustical parameters on the terminals of the mainframe.

0 INTRODUCTION

In recent years many researchers have been working on techniques for auralization of the calculation results from room acoustical computer models. An overview has been presented by Kleiner [1], and it appears that methods for binaural presentation considered until now are based on the convolution of room impulse responses with time sequences of anechoic recordings. This implies in principle that the individual head-related transfer functions must be known at a large number of angles of incidence, and the corresponding impulse responses for the two ears should be convolved with the room impulse response. An overview of the binaural technique has been presented by Møller [2].

Drawbacks of the convolution technique are, that special and expensive hardware is needed or - if less expensive hardware is used - the calculation is very time consuming. In general it is a problem to handle very long impulse responses and very long sound signals, and the quality of the sound is sometimes disappointing.

The present work suggests a new alternative technique for auralization, which avoids the convolution bottleneck. An interface has been developed between a digital audio mainframe and a room acoustical computer model. This means that auralization can follow immediately after the room acoustical calculation in a receiving point, and there are no limitations on length of the source signal. The early reflections and the late reverberant reflections are treated by two different techniques, which is also the case in modern room acoustical computer models of the hybrid type like the ODEON program used here.

Early reflections are represented by time delay and energy level and each reflection is divided into a left ear reflection and a right ear reflection with separate settings. Up to 160 early reflections are possible, but in most well-behaved rooms a number less than 40 seems to be sufficient for obtaining a realistic audible impression of the simulated room. Late reflections are described statistically by calculated reverberation time as a function of frequency. The settings of an electronic reverberation processor are adjusted according to the calculated values in a number of frequency bands.

As an objective verification test the reverberation time, Clarity and other room acoustical parameters have been measured on the terminals of the digital audio mainframe, and the results have been compared with the values from the room acoustical computer model. In an earlier study calculated values were tested against measured values in real rooms, ref. [3].

In addition to the interaural time delay used for early reflections it is possible to activate a number of built-in filters, which are simplified Head-related Transfer Functions for a number of different directions of incident sound. The work with these filters is not finished at the moment of writing, but they are expected to improve the localization of sound sources and the spaciousness. However, the aim of the present method for auralization has not been to simulate every detail of the sound field with high precision, but to offer a sufficient degree of realism for auralization to be used for evaluation of different situations in a room. The sound quality obtained by this method is high enough for the listener to evaluate even minor differences in the room responses, e.g. different source or receiver positions or orientations, different room geometries or absorption distribution. The switching between different settings to be compared is easy and fast and may be done while the sound signal continues.

1 THE HYBRID COMPUTER MODEL

The ODEON model used in this work belongs to the 'hybrid' type of models, which contain elements of both ray tracing and image source models [4-5]. It is based on the idea that an efficient way to find branches of the image source tree having high probabilities of containing visible image sources is to trace rays from the source and note the surfaces they hit. The reflection sequences thus generated are then tested as to whether they give a contribution at the chosen receiver position, in line with image source theory. The finite number of rays used places an upper limit on the length of accurate reflectogram obtainable. Thereafter, some other method has to be used to generate a reverberant tail.

1.1 Ray tracing from sources

For the purpose of ray tracing the following assumptions are made: Any room is composed of plane surfaces, which reflect rays specularly until some criterion is met (e.g. reflection order, path length, hitting a predefined 'diffusor'). After such a criterion has been fulfilled by a ray, it may be reflected into non-specular paths, with probabilities determined by the values of 'diffusion coefficients' assigned to surfaces. This 'diffusion' functions primarily to avoid gross errors in the late reverberation decay rate, which can occur if the ray tracing process is insufficiently 'mixing'. Rays are sent out from the source and followed around the room as they become reflected, and the data thus produced is stored for use later in the determination of reflections received at a point.

1.2 Early reflections at a receiver

Every time a ray is reflected at a surface the position of an image source lying behind that surface is found. These image sources are sources of potential 'early reflections' and undergo rigorous checking according to Image Source Theory. For each potential early reflection image source, it has to be determined whether that source can in fact be 'seen' from the receiver position. Once an image is found to be valid, the laws of Geometrical Acoustics are applied: The level of the corresponding reflection in the reflectogram is thus simply the product of the energy reflection coefficients of the walls involved in generating the image and the level of the source in the relevant direction of radiation. The reflection's arrival time is given by the distance to the image source and its direction is that of the image source seen from the receiver. A typical calculated reflectogram of early reflections is shown in Figure 1.

1.3 Late reflections at a receiver

At a certain reflection order there is a transition from early to late reflections, and the rays are treated as transporters of energy rather than explorers of the geometry. When a ray hits a surface, a secondary source is generated at the collision point. Each secondary source is considered to re-radiate energy into a hemisphere inside the room and the intensity reaching a receiver from a secondary source is proportional to the projected area of a notional surface element at the secondary source position, as seen from the receiver. The time of arrival of a late reflection is determined by the sum of the path lengths (i) from the primary source to the secondary source via any intermediate reflecting surfaces and (ii) from the secondary source to the receiver. The direction of arrival is given by the direction from the secondary source to the receiver. A secondary source only contributes a reflection to the response if it is visible from the receiver, thus non-homogeneous reverberant fields can be modelled. An example of a calculated decay curve is shown in Figure 2.

2 THE DIGITAL AUDIO MAINFRAME

The auralization is realised with the M5000 Digital Audio Mainframe from t.c.electronic. It should be equipped with one AD-DA card and two DSP cards, one of which must be equipped with extra RAM. The M5000 is interfaced to the PC through a RS232 remote cable, and from the built-in floppy disk drive it must be loaded with the special ODEON-SOUND algorithm. The main auralization features are:

- one or two input channels
- early reflections can be set for left and right output separately, up to 160 early reflections (up to 80 reflections with two input channels)
- the reverberation tail can be set in three frequency bands.
- early reflections may be assigned to eight different groups, and each group may have a special filter setting, which is used for presentation through headphones.

2.1 Early reflections

For discrete reflections, the ODEON-SOUND algorithm for the M5000 can in total handle 160 double taps which are distributed with 80 to the left output channel and 80 to the right output channel. A double tap can be set to a gain and two different time delays. The gain value represents sound energy and it is linear from 0 to 32767. The time delay can be set with a resolution of 23 µs in the range from 0 to 1485 ms (up to 743 ms in case of two input channels). Each tap can be programmed to read from either of the input channels.

2.2 Reverberation tail

The reverberant part of the ODEON-SOUND algorithm for the M5000 can handle decays with Reverberation Time (RT) from 0.3 to 30 s with a resolution of 0.1 s. RT can be set separately in three different frequency bands. The time delay before onset of the reverberation tail can be set from 0 to 200 ms with a resolution of 1 ms. The level of the tail is controlled in steps of 0.5 dB from 0 to 40 dB attenuation, and in steps of 3 dB from 40 to 97 dB attenuation.

2.3 Technical specifications

The M5000 has a lot of input/output connections and other facilities which are not mentioned here. Some of the specifications of importance for the auralization are summarized in Table 3.

Table 1 Selected technical specifications for M5000 Digital Audio Mainframe.

Sampling rate	44.1 kHz (32 kHz or 48 kHz are also possible)
Frequency response	10 Hz - 20 kHz, ±0.5 dB
Harmonic distortion	< 0.006% 1kHz, 0 dBm
Inter modulation distortion	< 0.006%
Dynamic range	> 100 dB
Crosstalk	< -80 dB @ 1 kHz
Group delay linearity	< 5 µs
Phase linearity	Better than 5°

3 THE INTERFACE SOFTWARE

A program called ODSOUND has been developed to allow the M5000 to be controlled according to the calculation results from ODEON. This program is described in the following.

3.1 Room acoustical data for auralization

The file created for auralization as one of the results of a source-receiver response calculation in ODEON contains data for identification, data on early reflections and values of some room acoustical parameters. The file name is normally also identification of the room name, and for each name there may be up to 100 different auralization files.

Data for identification include a text string with optional description, position of source/sources, position of receiver and direction of receiver.

Early reflections are described by arrival time, strength at mid frequencies (500 and 1000 Hz), asimuth and elevation angle.

Statistical parameters for setting the reverberation tail are octave band values of RT, Early Decay Time (EDT), Centre Time and Total Sound Pressure (integrated squared impulse response).

3.2 Setting the early reflections

The early reflections are very important for obtaining a realistic auralization. For presentation through headphones the following three methods are used in order to obtain localization outside the head:

- Interaural time difference. Each reflection is divided into a left ear and a right ear reflection, and the interaural time delay is calculated as the free field time delay between two points at the ears of the receiver according to the angle of incidence relative to the receiver. This is the dominant cue for localization of broad band sound in the horizontal plane, see Wightman and Kistler [6].

- Interaural intensity difference. The signal to the ear in the direction of the incident reflection is raised with 6 dB times the cosine of the angle of incidence relative to the direction of a line through the two ears. This is a simplified representation of the reflection effect of the head relative to free field.

- Spectral cues. The spectral peaks and notches due to the outer ear are roughly simulated by filters. Although this is known to be the main cue for elevation, the intention at this stage has not been to create a localization for different elevation angles, but rather to avoid the front-back confusion and to improve the out-of-the-head localization.

The most difficult part is the use of filters with spectral cues. The question how many directional sectors are needed has been studied by Mommertz and Aprath [7]. They also considered the elevation, but if we only look at localization in the horizontal plane the result was that 12 sectors and 36 sectors were equally good, whereas four sectors are clearly insufficient. For the reverberation tail they found that six horizontal sectors are sufficient.

For the present work it was decided to use eight sectors in the horizontal plane. Each sector represents different filters for left and right ear, but due to symmetry it is sufficient to divide the reflections into only five groups. The filters in each group reproduce what are thought to be the important peaks and notches, which can be identified on measured head-related transfer functions; this includes a low-pass filter to represent the shadowing effect of the head. Such measurements are reported by several researchers, see e.g. Hammershøi et al. [8].

The spectral cues are known to be slightly different from one person to another, but in order to avoid the front-back confusion the cues must be very accurate. Thus it is necessary to have the possibility for individual calibration of the filters, and this is done in a relatively simple two-step procedure. The individual calibration is saved for later use, so it need only to be done once for each person.

In future improvements of the system it might be possible to increase the number of directional sectors in order to obtain a better localization as a function of elevation angle. However, for most uses of the auralization technique this is thought to be of minor importance. The aim of the present work has been to obtain a realistic sounding impression of the room of sufficient quality to allow comparisons between situations of acoustic relevance.

3.3 Setting the reverberation tail

The decay of the reverberation tail is set with a RT equal to the value calculated from the -5 dB point to the -35 dB point. Finally, the level of the tail is determined in such a way that the total level is correct, i.e. from the total energy minus the summed energy of the early reflections.

The problem is to determine the time delay of the tail in order to obtain the most natural sounding result. The method developed here implies that the reverberation tail has an overlap with some of the early reflections, and so the method is similar to the calculation method in ODEON described earlier. It is obvious that the start of the reverberation tail must be somewhere between the two time limits shown in Figure 2, i.e. after the arrival of the first late reflection and before the last early reflection. If the delay is too short the overlap will lead to a local top in the decay, but if the delay is too long there will be a valley instead. After many experiments a method has been developed, which is based on EDT and Centre Time.

3.4 Operating the ODEON-SOUND system

Before starting to listen to the auralization the left input channel on the M5000 is fed with a program of anechoic recordings and the remote control cable is connected to the computer running the ODEON and ODSOUND programs. Both input channels may be used as described later.

When a number of source-receiver response results have been saved as auralization-files from ODEON, the ODSOUND program may be called. A listening session is prepared by creating a list of auralization-files of interest. For the two-channel option each selection contains a pair of files, typically representing different source positions in the same room. The list may contain any number of files.

When auralization is selected a screen display for operating the M5000 pops up, see Figure 3. With the up/down cursor a file is selected from the upper window with a scrolling list of auralization-files. Each line displays the number of the file, the file name, the identification text and a level setting. The lower window shows 'No Source' for the usual case of one input channel. For the two channel option a pair of files is selected, and the lower window shows information for the companion file. When a file is selected for auralization the M5000 is reset accordingly after a few seconds. The anechoic input signal may continue while shifting between the different auralization situations. The first ten files in the list can be selected directly by a single key stroke.

A realistic level setting is very important for serious use of auralization. For this purpose the 'Direct' function may be selected. This removes the reverberation tail and all reflections except the first one, which is supposed to represent the direct sound in the actual situation. Knowing the distance from source to receiver and the recording level of the anechoic signal it is now possible to adjust the level. The room reflections can be put back by selecting the same file again.

The function 'Filter Off' removes the directional filters and the interaural intensity difference, which are only relevant for presentation through headphones. Only the interaural time difference is kept. This setting should be activated when headphones are not used, e.g. presentation through loudspeakers.

3.5 Possibilities with two input channels

The uses of auralization as described in the known literature is based on one single mono recording of anechoic music or speech. The auralization system described here has the possibility of two different input channels and each of the early reflections can be assigned to either channel. The reverberation tail, however, must be common to the two channels, which is no major limitation as the reverberation tail represents the room where the listener is given the impression to be present. Each sound channel may be represented in the model by one or several source positions; in ODEON 2.4 up to twenty sources with separate level, time delay and directivity may be used. Examples of use include:

- Two speakers in a dialogue represented in the model by two different source positions.
- An orchestra recorded stereophonically and represented in the model by two different groups of source positions.
- A singer and an accompanying orchestra recorded in two separate channels and represented in the model of an opera house by one source position on the stage and a group of source positions in the pit.
- The second channel can be used for background noise. One example is a subway station with a speaker announcement from a group of loudspeakers and train noise represented through a line of source positions above the tracks.

The balance between the two channels is controlled by the level settings in the ODSOUND program. These examples illustrate that the two channel feature of the auralization system opens up a new range of acoustical situations, which can be studied with this new technique.

4 OBJECTIVE VALIDATION OF THE AURALIZATION SYSTEM

For a number of examples the setting of the M5000 was measured on the terminals with the B&K 2231 Sound Level Meter with the Room Acoustics Module BZ7109. During the measurements the 'Filter Off' function was used, so omnidirectional characteristic was simulated for source and receiver. Measured room acoustic parameters RT, EDT, Clarity (C80) and Centre Time (T_s) are compared to the values calculated by ODEON, see Table 2.

Table 2 Examples of calculated ODEON values and measured M5000 values of room acoustical parameters at the 1kHz octave band.

Room model	Number of early refl.	RT (s) calc.	RT (s) meas.	EDT (s) calc.	EDT (s) meas.	C80 (dB) calc.	C80 (dB) meas.	T_s (ms) calc.	T_s (ms) meas.
LISTEN	207	0.40	0.53	0.44	0.42	11.6	12.0	26	31
RFH (S1R1)	41	1.62	1.45	1.23	1.14	4.8	3.0	63	82
RFH (S1R2)	69	1.59	1.60	1.26	1.30	2.8	2.7	87	95
RFH (S1R3)	64	1.50	1.49	0.86	1.24	3.5	4.2	67	76
RFH (S1R4)	88	1.57	1.50	1.21	1.60	1.1	1.5	95	103
RFH (S1R5)	111	1.53	1.53	1.08	1.49	0.3	0.5	97	120

LISTEN is a model of an IEC listening room with relatively dead acoustics. The RFH room is a model of the Royal Festival Hall, London, which was described in more detail in ref. [3]. Results are shown for five receiver positions. Early reflections were calculated up to reflection order 4 (RFH) or 5 (LISTEN). The number of early reflections found by ODEON is shown, but no more than 160 were used by the M5000. It must be emphasized that the results are preliminary, but especially the behaviour of C80 within a wide range seems to indicate a satisfactory control of the auralization system. A subjective validation of the system including the directional filters remains to be done.

5 CONCLUSION

The auralization system presented here is not perfect. One drawback is that the frequency variation of the reflections can only be roughly modelled in the reverberation tail, and not at all in the early reflections, which are set according to the mid frequency values. This may be improved in future versions, however, for many applications it seems to be a minor problem.

The method for mixing of the reverberation tail with early reflections was found very critical, but a satisfactory solution has been achieved. During the development it has been a great help that the system is fast and easy to use, thus allowing a large number of experiments to be carried through within a limited time.

6 REFERENCES

[1] M. Kleiner, B.-I. Dalenbäck and P. Svensson, "Auralization - An Overview," J. Audio Eng. Soc., vol. 41, pp. 861-875 (1993).

[2] H. Møller, "Fundamentals of Binaural Technology," Appl. Acoust., vol. 36, pp. 171-218 (1992).

[3] G. Naylor and J.H. Rindel, " Predicting Room Acoustical Behaviour with the ODEON Computer Model," Presented as paper 3aAA3 at the 124th Acoustical Society of America Meeting, New Orleans (1992).

[4] G. Naylor, "ODEON - Another Hybrid Room Acoustical Model," Appl. Acoust., vol. 38, pp. 131-144 (1993).

[5] G. Naylor, "Treatment of Early and Late Reflections in a Hybrid Computer Model for Room Acoustics," Presented as paper 3aAA2 at the 124th Acoustical Society of America Meeting, New Orleans (1992).

[6] F.L. Wightman and D.J. Kistler, "The dominant role of low-frequency interaural time differences in sound localization," J. Acoust. Soc. Am., vol. 91, pp. 1648-1661, (1992).

[7] E. Mommertz and S. Aprath, "Zur Quantisierung der Schalleinfallsrichtungen in der binauralen Simulation," DAGA'93 (1993).

[8] D. Hammershøi, H. Møller, M.F. Sørensen and K.A. Larsen, "Head-Related Transfer Functions: Measurements on 24 Human Subjects," Presented as Preprint 3289 at the 92nd Convention of the Audio Engineering Society, Vienna, (1992).

Figure 1: A reflectogram of early reflections in ODEON. The vertical line marked 'Start of reverb' indicates the time of arrival of the earliest late reflection.

Figure 2: An energy response histogram and reverse-integrated decay curve in ODEON.

```
 Info  Files  Calibration  Auralisation  Setup  Help
╔[■]════════════════════════ Selected Files List ═══════════════════════════════╗
║ Calibrated for                                    Date       adj  1   adj  2   ║
║  No Name                                        1993 10 22    0.000    0.000   ║
║                                                                                ║
║ No   Source 1     Description                                      Level SNo  ║
║    3 RFH3.S02     Royal Festival Hall, S1R3                          0.0  1    ║
║    4 RFH3.S03     Royal Festival Hall, S1R4                          0.0  1    ║
║    5 RFH3.S04     Royal Festival hall, S1R5                          0.0  1    ║
║    6 BOSTONS2.S01 S1R2, Centre Source, Stalls                        0.0  1    ║
║    7 BOSTONS2.S04 S2R2, Right Source, Stalls                         0.0  1    ║
║                                                                                ║
║   Set Level 1        Set Source 1       Set Level 2        Set Source 2       ║
║                                                                                ║
║ No   Source 2     Description                                      Level SNo  ║
║    3 No Source                                                       0.0  1    ║
║    4 No Source                                                       0.0  1    ║
║    5 No Source                                                       0.0  1    ║
║    6 No Source                                                       0.0  1    ║
║    7 BOSTONS2.S05 S3R2, Left Source, Stalls                          0.0  1    ║
║                                                                                ║
║   Leave         M5000         Direct         Filter Off          Remove       ║
║                                                                                ║
╚════════════════════════════════════════════════════════════════════════════════╝
 Alt-X Exit  F1 Help  F2 Select Files  F3 Calibration List  F4 Auralisation
```

Figure 3: Screen menu for operation of ODEON-SOUND.