AN INVESTIGATION ON THE TRANSITION FROM EARLY REFLECTIONS TO A REVERBERATION TAIL IN A BRIR

Kittiphong Meesawat

Department of Acoustics, Aalborg University,
Fredrik Bajers Vej 7, Aalborg Ø, 9220, Denmark
ktphong@acoustics.auc.dk

Dorte Hammershøi

Department of Acoustics, Aalborg University,
Fredrik Bajers Vej 7, Aalborg, 9220 Ø, Denmark
dh@acoustics.auc.dk

ABSTRACT

Reflections in a binaural room impulse response (BRIR) can be classified into 2 groups: the early reflections and a reverberation tail. The aim of this work is to determine at which time a transition from early reflections to late reverberation can be said to occur. In the literature several suggestions exists for the determination of a such transition time, based on objective room parameters. However, these lead to different values, and none may relate generally to a perceptually justified transition time. In this investigation BRIRs are measured in a room with a listener and a loudspeaker in different positions. The BRIRs are subsequently cut at 20, 60, 10 and 140 ms and different "heads" and "tails" are combined. The modified BRIRs are compared with the original BRIRs in a 3AFC experiment to reveal the audibility of any possible position specific information in the tails. Only preliminary results from a pilot experiment exists at the time of writing.

1. INTRODUCTION

A binaural room impulse response (BRIR) can be categorized into 3 parts; the direct sound, a number of early discrete reflections, and the reverberation tail. It is possible to identify the partition between the direct sound and early reflections by visual inspection of a reflection diagram of a BRIR. In contrast, the partition between early reflections and the reverberation tails is not so easy to define. Determining the transition time in a BRIR may be useful for many acoustical studies, e.g. when determining the influence of the reverberation tail in the spatial hearing, or the acoustical characteristics of the reverberation tail in a BRIR, besides the great potential in binaural synthesis, just to name but a few. There are several objective criteria that suggest a transition time in a BRIR but these criteria lead to different values, and they may not be generally applicable, when it comes to the perceptual significance in any given situation (e.g. a simulation). In this investigation the different transition times suggested by existing criteria is compared to the transition time, which is obtained subjectively (from listening experiments). In Section 2, the definition of the reverberation tail is reviewed and the focus of the reverberation tail on the current work is also described. A number of criteria suggesting the transition time are reviewed in Section 3. Section 4 presents the method of measurement and signal pre-processing. The methods used in the experiment are described in Section 5 and the preliminary results are given in 6.

2. REVERBERATION TAIL

2.1. The traditional description of the reverberation tail

When a room is excited by an impulsive sound signal, the sound in the room decays as a function of time. The decay occurs because the room boundaries absorb a portion of sound energy on each reflection. After a large number of reflections the sound in the room may be assumed to have become diffuse: the average energy density is the same throughout the volume of the enclosure, and all directions of propagation are equally probable (Kinsler et al. [1]). Though the model over-simplifies the behavior of sound in a room, it is correct that the diffusion of the sound field is increasing with time. After some time, the diffusion of the sound field is high enough so that the individual characteristic of each reflection can not be detected. For the BRIR, this part is called the reverberation tail. In other words, the reverberation tail starts when the angle of incidence of the sound wave to the listener can be represented by a random process with a uniform probability distribution function, and the sound energy is equally distributed throughout the room.

2.2. The focus of the reverberation tail in the current work

The ideally diffused sound field does not exist in ordinary rooms. Therefore it is impossible to locate the transition time of a BRIR when a part of it meets the diffuse condition. In this study, the dimension of human sound perception is taken into the consideration. The reverberation tail starts when sound energy arrives at a listener with random incidence with a such high number of reflections so that the listener can not hear out the individual reflections or detect a difference due to any other characteristics. Moreover, the reverberation tail will not sound different if the listener is in a different location in the same room with the same source-receiver distance. The focus of the reverberation tail in this study is the sound component in the BRIR that is perceptionally independent from the source-receiver alignment. The transition time is defined as a point in time domain that the reverberation tail starts.

3. CRITERIA REVIEWS

In the literature a number of criteria have been proposed for determining a transition time. These criteria are based on different definitions of the reverberation tail. This section reviews some of them. These criteria are 1) fixed valued parameters, 2) reflection order 3) mean free path, 4) reflection density, and 5) room volume.

3.1. Fixed value parameters

Some studies used an objective parameter to locate the transition, i.e. at the time of 50 or 80 milliseconds after the direct sound (e.g. Begault [2], and Bradley and Soulodre [3]). This criteria do not refer to the transition time between early reflections and the reverberation tail directly, but rather to classify early part from the late arrival sound energy. The fixed value was suggested by classical architectural acoustics. The criteria discards all room acoustical parameters, and is the same for a highly reverberant room as for a dry room.

3.2. Reflection order

Using of reflection order to specify the transition time in a BRIR is a widely used criteria in BRIR synthesis. This is equivalent to specifying a transition time that depends on the room size. With the same transition reflection order, the transition time for a larger room is longer than the one for a smaller room. The criteria does not state the transition point explicitly. It says only that after the transition point it is possible to use a reverberation model for the tail of a BRIR. There is no standard for such a transition order. Naylor and Rindel [4] suggest the 4^{th} order for a maximum transition time while Martin et al. [5] propose 2 stages of transition points, from early reflections to the beginning of the reverberation process and then to the statistical part of the reverberation at the 4^{th} to the 6^{th} order and at the 10^{th} to the 15^{th} order respectively. Different binaural synthesis studies use different transition orders.

3.3. Mean free path

By an expression of the mean free path $\bar{l}$ in Equation 1 (Kuttruff [6]), the mean free path can be interpreted as a mean distance of a sound ray between 2 reflections in a room. In Equation 1, c refers to the speed of sound, t is the observation time interval, N refer to the total number of reflections that occur in the interval t, and $\overline{n}$ refer to the average number of reflections occurring per second.

$$\bar{l} = \frac{c \cdot t}{N} = \frac{c}{\overline{n}} \tag{1}$$

The mean free path of a room is determined as by

$$\bar{l} = 4 \cdot \frac{V}{S}, \tag{2}$$

where V is the room volume and S is the room surface area. This gives an exact value for a room, which is not the case for the reflection order criteria. This difference may be large if the room dimension ratio is uneven. Rubak and Johansen [7] suggest that the transition time for a synthesized BRIR is approximately 4 times the time a sound particle spends for a distance of the mean free path. For Equation 2, the walls are assumed to be diffusely reflecting (Kuttruff [6]).

3.4. Reflection density

Reflection density or temporal density of echoes can be expressed as

$$\frac{dN}{dt} = \frac{4\pi c^3}{V} \cdot t^2 \tag{3}$$

from Kuttruff [6]. Kuttruff also mentions that this formula is valid not only for a rectangular shape room but also for a room with an arbitrary shape. The reflection density does not determine the transition time directly, but rather acts as an indicator for, whether the sound field is diffuse or not. There are several suggestions on the reflection density required for the sound field to be considered diffuse. Schroeder [8] suggested that 1000 echoes per second was sufficient to sound indistinguishable from diffuse reverberation while Griesinger [9] has suggested that 10000 echoes per second is required. In Gardner [10], Griesinger is quoted for stating that this value is a function of the bandwidth of the system. Rubak and Johansen [7] suggest that 4000 echoes per second would be necessary.

3.5. Room volume

There is also a suggestion of determining the transition time by the room volume:

$$t_{mixing} \approx \sqrt{V}\ (milliseconds) \tag{4}$$

This value was proposed by Reichardt and Lehmann [12] (according to Jot et al. [11]) as a reasonable approximation for the transition time between early reflections and late reverberation.

4. METHOD OF MEASUREMENT AND PRE-PROCESSING

This section describes the measurement setup, measurement system description, and signal pre-processing methods used in the experiment.

4.1. Measurement setup

Figure 1 shows a conceptual drawing of the measurement scheme.

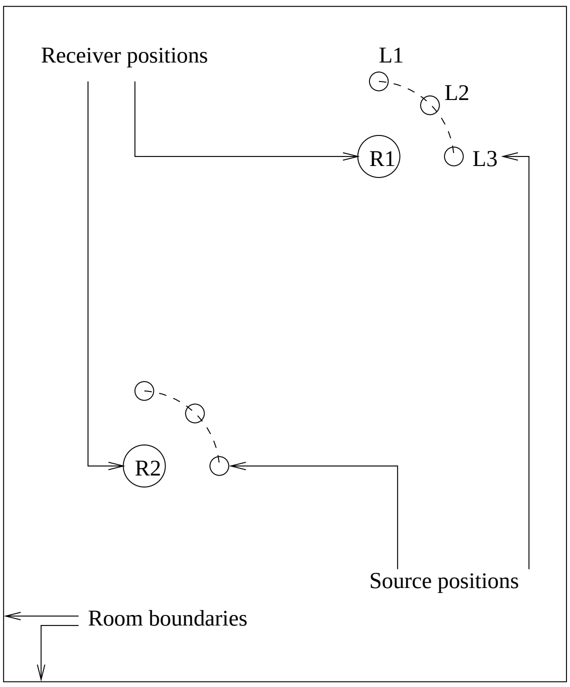

Figure 1: *A conceptual drawing of the BRIR measurement scheme.*

A number of BRIRs were measured in the room with the following description. The room used for the measurement is a medium size lecture room at Aalborg University with the size of $8.7 \times 7.1 \times 3.0\ m^3$. The overall reverberation time of the room is 0.7 seconds. Two receiver positions in the room were chosen. At each receiver position, BRIRs were measured for three locations of the sound source, one directly in front of the receiver, one at 45^o to the right, and one at 90^o to the right, at a distance of 2.66 meters. This gave a total of 6 set of BRIR measurements.

4.2. Measuring system descriptions

The BRIR measuring system is shown in Figure 2. The system consists of a computer with a multichannel soundcard, a power amplifier, a loudspeaker, two measuring microphones in two ears of an artificial head (experimental head developed at the Department of Acoustics, see e.g. Christensen and Møller [13]), and a microphone pre-amplifier. The measurement system is based on an experimental measurement system developed by Olesen et al. [14].

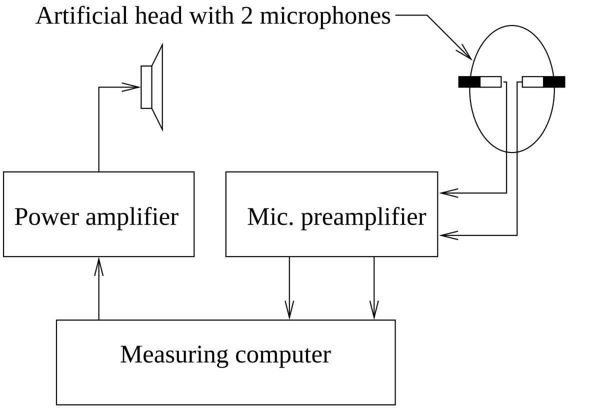

Figure 2: *Diagram of the measuring chain.*

The measuring signal is a maximum length sequence of 16^{th} order, generated by the computer. The signal is sent to the loudspeaker through the soundcard and the power amplifier. The two measuring microphones record the sound pressures at the ears of the artificial head and feed them back to the computer via the microphone amplifier and the soundcard. The measurements were done with 4 pre-averages and 4 averages in order to improve noise immunity. The BRIR is computed by the computer.

4.3. Signal enhancement

Though the MLS method is generally–and in particular as implemented by Olesen et al. [14]–noise immune, noise remains that is induced in electrical system and ambient noise in the room. As the tails of the BRIRs will be interchanged (see Section 4.4), the noise level will be altered. The method of signal concatenation includes signal scaling to ensure the amount of energy emitted to the listener is preserved. This scaling process may increase the noise level for which reason the BRIR is degraded by the amplified noise. Consequently, the listener may potentially hear the presence of background noise in the tail portion of a BRIR, and falsely detect the interchange of the tails by the noise level and/or the noise charateristics.

A signal enhancement was applied to suppress the background noise on the tail of the impulse response. The method of noise suppression is to force the tail to continue the natural exponential decay beyond the transition from the true impulse response into the noise region, with a decay rate determined from an average of the decay rate computed from the early part of the impulse response. An example of the noise suppression is shown in Figure 3.

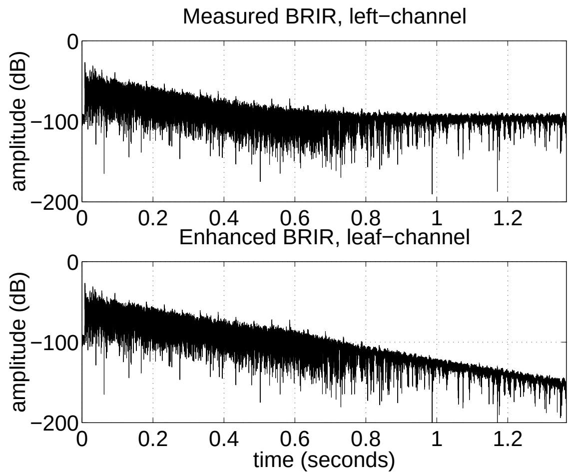

Figure 3: *Signal enhancement by tail suppression.*

4.4. BRIR concatenation process

The BRIRs used in the listening experiment are either the original set of BRIRs, or a set of BRIRs that has the same beginnings as the original BRIRs but tails taken from another set of BRIRs. The concatenation procedure includes energy preservation and cross-fading at the junction of the concatenation. This is done for each side of the BRIRs separately.

4.4.1. Energy preservation

In the concatenation the level of the tail to substitute the original may be much different from the level in the original tail. Simply concatenating the early part of the original IR with the alternative tail may thus lead to excessive or inadequate of sound energy. Energy preservation ensures that the total energy of the IR after the concatenation is equal or close to its original energy level. This can be done by scaling the tail either up or down depending on its energy level compared to the energy level of the original tail. Equation 5 shows the calculation for scaling factor where t_x refers to the transition time, T refers to the length of the IR, p_1 is the original IR, p_2 is the IR from which the tail will be taken, and a is the scaling factor. An example of the energy ratio for modified IRs and the original IR in various combinations is shown in Figure 4. The transition time of the IR in this example is 20 milliseconds.

$$a = \frac{\int_{t_x}^{T} p_1^2(t)dt}{\int_{t_x}^{T} p_2^2(t)dt} \qquad (5)$$

The combination number in Figure 4 refers to different combinations of the concatenation pairs. In this experiment, 6 BRIR measurements from the room are used, and that makes 30 possible concatenation combinations for each value of the transition time. The energy ratio reflects the error in the amount of sound energy

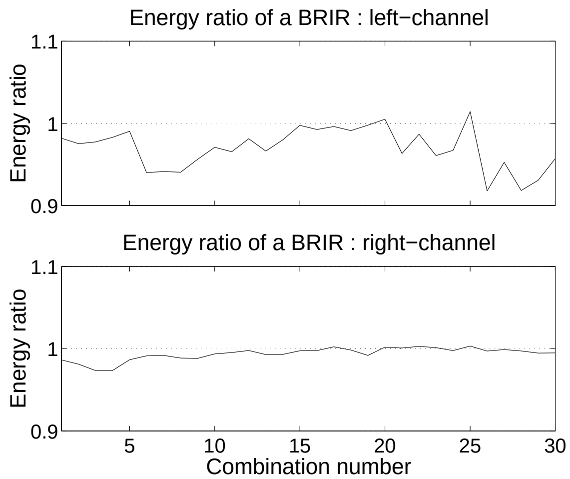

Figure 4: *Ratio of energy of a modified IR with transition time of 20 milliseconds to an original IR.*

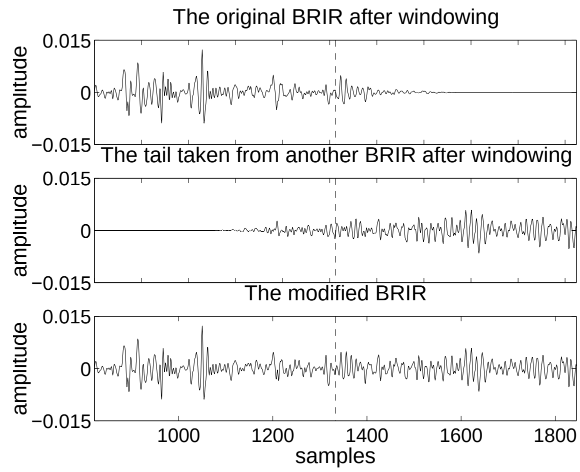

Figure 5: *Cross-fading of "head" and "tail" (left ear BRIR at 20 ms).*

of the modified IR, e.g. energy ratio of 0.95 means that the sound energy of the modified IR is 5% lower than the original. Notice that some of the percentage error of the left channel are larger than that of the right channel. This is for measurement positions, where the sound source is closer to the right ear than to the left ear. A small error on the left channel can therefore make a large percentage error.

4.4.2. Cross-fading

An abrupt change may occur at the transition from the earlier part of an original IR to the tail taken from another IR. This can be prevented by using cross-fading or interpolation of the IRs. The cross-fading consists of 2 one-side triangular windows each of length 512 samples. A right half window is applied to the end of the original IR while a left half window is applied to the beginning of the other IR with 512 samples overlap. Equation 6 describes the cross-fading method.

$$p_m[i] = (w[j] \cdot p_1[i]) + (w[k] \cdot p_2[i]), \qquad (6)$$

where w is a triangular window and the index variable, i is limited to the transitional region which is ± 256 samples from the transition index, i_x.

$$i_x - 256 \le i \le i_x + 256 \qquad (7)$$

The window indexes j and k determine whether the left side or the right side of the window is applied. The indexes are calculated from equation 8 and 9.

$$j = i - i_x + 256 \qquad (8)$$

$$k = i - i_x - 256 \qquad (9)$$

An example of the cross-fading is shown in Figure 5. The dashed lines indicate the transition index.

5. EXPERIMENTAL DESIGN

A listening experiment was designed to determine the transition time of a BRIR from its early reflections to late reverberation. With the focus on the reverberation tail defined in Section 2.2, this can be done by finding a concatenation time between the early part and the later part of a BRIR that the listener could not hear, whether the later part taken from the same BRIR as the early part or is taken from another BRIR measured in the same room.

Candidates of the transition time to be examined were 20, 60, 100, and 140 milliseconds. These values of the transition time are suggested by the objective criteria mentioned in Section 3 and by a preliminary experiment. The transition times suggested by the criteria are listed in table 1.

Criteria	Suggested transition time (ms)
Fixed value [2], [3]	80.0
Four times of mean free path [7]	39.6
Reflection density of 1000 [8]	19.1
Reflection density of 4000 [7]	38.2
Reflection density of 10000 [9]	60.5
Square root of room volume [12]	13.6

Table 1: Transition time suggested by criteria from the literature

There are two types of BRIRs to be used in the experiment. The one type is the original BRIRs which is one of six measured pairs of BRIRs without modification. The other type is the modified BRIRs which are the original BRIRs truncated at either 20, 60, 100, or 140 ms, and concatenated with the later part from another pair of BRIRs (using the concatenation process described in Section 4.4).

5.1. Playback system description

In the listening experiment all stimuli will be presented to the listener via a playback system. The playback system consists of a computer, a soundcard, and a set of headphones. The headphones were equalized. The stimuli were generated by the computer and were amplified by the soundcard. The amplified stimuli were presented to the listener by the headphones.

5.2. Procedure

The listener is presented a series of stimuli, each consisting of two presentations of a test sound convolved with the original BRIR and one presentation of a test sound convolved with the modified BRIR in a random order. The listener is asked to indicate which of stimuli that differed, and is forced to answer. The stimuli presentation may be repeated up to 3 times on the subject's request.

Within one comparison, the test stimuli may be presented either as the first stimulus, as the second or as the third. There are 6 BRIR measurements, which gives a total of 30 early-late combinations. Each combination exist for each of the 4 values of the transition time chosen. This leads to a total of 360 comparisons for one complete listening experiment.

The comparisons are distributed into 18 sessions with 20 comparisons in each. The stimuli order is balanced so that the stimuli that are from the same early-late combination and transition time are presented only once in a session. Further, an equal number of comparisons with stimuli at each of the transition time values used, is included in each session, i.e. there is 5 comparisons for 20 ms, for 60 ms, for 100 ms, and for 140 ms. With these restrictions, the order of the comparisons within each session was randomized.

6. RESULTS

Only preliminary results from the pilot experiment with one listener exist. The percentages of correct answers are shown in Figure 6, as the mean value of the correct answer percentage and plus/minus its standard deviation. It is suggested that the percentage of correct answer is lower when the transition time value is higher as expected.

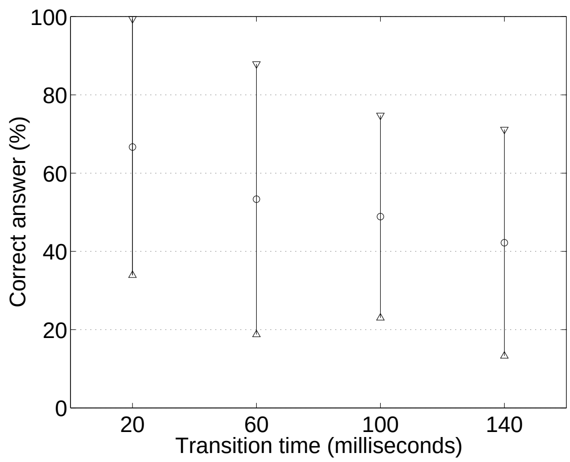

Figure 6: *Percentage of correct answers for the one listener from the pilot experiment.*

7. DISCUSSION

The authors abstain from any further discussion of the results, until more data exists.

8. REFERENCES

[1] L.W. Kinsler, A.R. Frey, A.B. Coppens, and J.V. Sanders, *Fundamental of Acoustics*, John Wiley and Sons, New York, 3rd edition, 1982.

[2] Durand R. Begault, "Perceptual effects of synthetic reverberation on three dimensional audio system," *J. Audio Eng. Soc.*, vol. 40, no. 11, pp. 895–903, Nov. 1992.

[3] John S. Bradley and Gilbert A. Soulodre, "The influence of late arriving energy on spatial impression," *J. Acoust. Soc. Am.*, vol. 97, no. 4, pp. 2263–2271, Apr. 1995.

[4] Graham Naylor and Jens Holger Rindel, "Predicting room acoustical behaviour with the ODEAN computer model," *Presented at the 124th Meeting of Acoustical Society of America*, New Orleans, USA, Nov. 1992, (Paper 3aAA3).

[5] J. Martin, D. Van Maercke, and J-P. Vian, "Binaural simulation of concert halls: A new approach for binaural reverberation process," *J. Acoust. Soc. Am.*, vol. 94, no. 6, pp. 3255–3264, Dec. 1993.

[6] Heinrich Kuttruff, *Room Acoustics*, Elsevier Science Publishers, 3 edition, 1991.

[7] Per Rubak and Lars G. Johansen, "Artificial reverberation based on a pseudo-random impulse response - part II," *Presented at the 106th Convention of the Audio Engineering Society*, München, Germany, May 1999, (preprint 4900).

[8] Schröeder M., "Natural sounding reverberation," *J. Audio Eng. Soc.*, vol. 36, no. 9, 1962.

[9] D. Griesinger, "Practical processors and programs for digital reverberation," in *Proc. 7th Audio Eng. Soc. Int. Conf.* Audio Eng. Society, 1989.

[10] William G. Gardner, *Applications of Digital Signal Processing to Audio and Acoustics*, chapter 3, pp. 85–131, Kluwer Academic, 1998.

[11] Jean-Marc Jot, Laurent Cervean, and Oliver Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," *Presented at the 103rd Convention of the Audio Engineering Society*, New York, USA, Sept. 1997, (preprint 4629).

[12] W. Reichardt and U. Lehmann, "Raumeidruck als oberbegriff von raumlichkeit und halligkeit, erlauterungen des raumeindrucksmasses," *Acoustica*, vol. 40, pp. 174–183, 1978.

[13] Flemming Christensen and Henrik Møller, "The design of VALDEMAR - an artificial head for binaural recording purposes," *Presented at the 109th Convention of the Audio Engineering Society*, Los Angles, California, USA, Sept. 2000, (preprint 5253).

[14] Søren Krarup Olesen, Jan Plogsties, Pauli Minnaar, Flemming Christensen, and Henrik Møller, "An improved mls measurement system for acquiring room impulse responses," *Proceedings of NORSIG 2000, IEEE Nordic Signal Processing Symposium*, Kolmården, Sweden, June 2000.