Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Equal reverberance matching of running musical stimuli having various reverberation times and SPLs

Doheon Lee, Densil Cabrera, William L. Martens

Faculty of Architecture, Design and Planning, The University of Sydney, NSW 2006, Australia

PACS: (1) 43.66.Lj, 43.55.Br

ABSTRACT

This paper examines effects of listening level and reverberation time on the reverberance of running musical stimuli. A listening test was conducted which tested an anechoic music stimulus convolved with synthetic RIRs having a range of listening levels and reverberation times: in the test, subjects adjusted the reverberance of a musical stimulus (by adjusting the decay rate of an impulse response convolved with dry music) to match that of reference stimuli. In this way, we constructed equal reverberance contours as a function of sound pressure level and reverberation time. The experiment results confirm that the listening level and reverberation time both have a significant effect on reverberance: increased listening level or reverberation time leads to greater reverberance. Loudness-based predictors of reverberance outperform the conventional reverberance predictors.

INTRODUCTION

Reverberation is one of the key factors considered in evaluating the acoustical conditions of auditora [1]. In the present study, we focus on the human perception of reverberation, which is known as '*reverberance*'. Although reverberation and reverberance may seem similar, there is a fundamental dinstintion in concept. The former is based on the objective physical characteristics of sound whereas the latter is based on the subjectively perceived characteristics of sound – and like other psychoacoustical relationships, there is unlikely to be a 1:1 correspondence between the two. Of course, there is a relationship between the two, and Sabine's seminal study of reverberation employed listening as a measurement technique to derive the concept of reverberationn time (T) [2]. However, as it is generally agreed that the reverberation time is not an ideal measure of reverberance, there have been many studies to refine the calculation of reverberation time or to develop new measures. By adapting the study by Haas [3], which showed the importance of early reflections in the human perception of sound, Atal *et al*. [4] proposed early decay time (EDT). EDT is similar to reverberation time except that it uses a shorter evaluation range, which is based on a linear regression line from the peak to 10 dB below the peak of the Schroeder reverse integration curve. ISO 3382-1 describes details of calculations and measurement methods of the two parameters as theses are widely used as objective reverberance predictors [5].

Soulodre and Bradley [6] invesitagated the extent to which the two conventional reverberance predictors (EDT and T) are correlated with the reverberance of a selected music stimulus. Their results show that EDT yields higher correlations to reverberance than the reverberation time (T). The highest correlation was observed for EDT averaged over all octave band values from 125 Hz to 4 kHz (r = 0.971) and the lowest correlation was found for the T averaged over all octave bands from 125 Hz to 4 kHz (r = 0.740). Although the study shows that the EDT strongly agrees with the reverberance, Lee and Cabrera [7, 8] noted that there is an issue in using EDT (or T) as a predictor of the reverberance, because although the sound pressure decay rate is independent of listening level, reverberance is likely to vary with listening level. The importance of listening level is supported by Hase *et al.* [9], who investigated the effect of sound pressure level and T on the reverberance of music and speech. According to their results, the two factors independently contribute to the reverberance of the tested music and speech stimuli. Moreover sound pressure level had the stronger effect on reverberance in their study (this tendancy is more significant for the tested music stimulus). The auditory system is not simple, and sound pressure level does not take account of many auditory features, such as auditory filter banks, auditory temporal integration, spectral masking, the functions relating auditory excitation to specific loudness and so forth [10]. By contrast, objective loudness models (such as Dynamic Loudness Model of Chalupper and Fastl [11] or Time-varying Loudness Model of Glasberg and Moore [12]) more accurately predict loudness variations over time because the models strive to account for these loudness complexities. It may be hypothesised that reverberance is closely related to the loudness decay over time, and so loudness decay may be plausibly used for predicting reverberance.

Loudness-based reveberberance predictors can be derived in a similar manner to the conventional T or EDT parameters, but using the loudness decay function obtained from a time-varying loudness model. This is helped by the fact that the loudness decay functions of room impulse responses are approximately exponential (at first). One question is what evaluation ranges should be used for the loudness-based reverberance predictors. According to Stevens [13], loudness is proportional to pressure raised to a power of 0.6 for tones of moderate frequency and moderate sound pressure level. This corresponds to the well-known rule-of-thumb that doubling or halving loudness corresponds to ±10 dB. Hence, the loudness-based EDT (namely EDT_N as the subscript '*N*' stands for loudness) can be calculated by measuring the time taken for a linear regresion line of a loudness decay function from the peak loudness to a half of the peak loudness, multiplied by six. Like EDT_N, the loudness-based T (namely T_N) can be calculated by measuring the time taken for a linear regression line of a loudness decay curve from 0.708 of the peak loudness and 0.178 of the peak loudness, multiplied by 3 (in analogy to T_{20}).

This idea was realized in the study of Lee and Cabrera [7] by examining the reverberance of music. The study investigated the relation between the reverberance of music and its listening level. The music stimuli were an anechoic music sample convolved with RIRs recorded in three auditoria. In the experiment, subject adjusted a decay rate of the RIRs before convolving with the anechoic music sample so as to match the reverberance of the music stimuli to a referenc music stimulus. The reference music stimulus was the same anechoic music sample convolved with a RIR. The subjects' responses from the experiment were averaged and converted into a loudness-based reverberance predictor (T_N), and compared with conventional EDT and T. The results show that the listening level has a significant postive effect on the reverberance and the loudness-based reverberance predictor outperform the conventional reverberance predictors, in part because of their sensitivity to listening level. Similar results are observed when impulsive stimuli (RIRs directly listened to) were used [8].

In the present study we investigated an effect of the listening level on the reverberance of music. Although one of the previous studies examined the reverberance of music, it was the overall reverberance of music. According to Morimoto and Asaoka [14], the reverberance of music is categorized into two parts; (1) running reverberance and (2) terminal reverberance [14]. The former refers to reverberance given while a stimulus is being played, and the latter refers to reverberance after a stimulus is stopped [14]. As there are few opportunities to hear stopped reverberance when audience is listening to music (except when there are large temporal gaps between notes) and the overall reverbeance of music was tested in the previous study, this study focused on the running reverberance. This is related to the concept behind the EDT, which also asseses the running reverberance by using a relatively short evaluation range. Therefore a listening experiment was conducted, and the results were converted into the conventional reverberance predictors and EDT_N. The details of the experiment are described in next section.

EXPERIMENT METHOD

The listening experiment consisted of two parts. In PART I, reference stimuli were an anechoic music sample convolved with synthetic RIRs having a T_{oct} of 2 s. The subscript 'oct' means an average of corresponding parameter values over the octave bands from 125 Hz to 4 kHz. The listening level of the reference stimuli ranged from a L_{Aeq} of 60 dBA to 80 dBA with 10 dB steps. In PART II, reference stimuli were the same anechoic music sample convolved with synthetic RIRs having various T_{oct}s of 1 s, 1.4 s, 2 s and 3 s. The listening level of the reference stimuli for PART II was fixed at a L_{Aeq} of 60 dBA. For both PART I and PART II, the comparison stimuli were the same anechoic music sample convolved with synthetic RIRs having a range of listening levels (60 dBA, 70 dBA and 80 dBA). The T_{oct}s of the comparison stimuli were adjusted by subjects in the experiment so that the reverberance of the comparison stimuli was equally matched with the reverberance of the reference stimuli. The total number of tested stimuli was eighteen pairs. The experiment setup for PART I and II are represented in Figure 1.

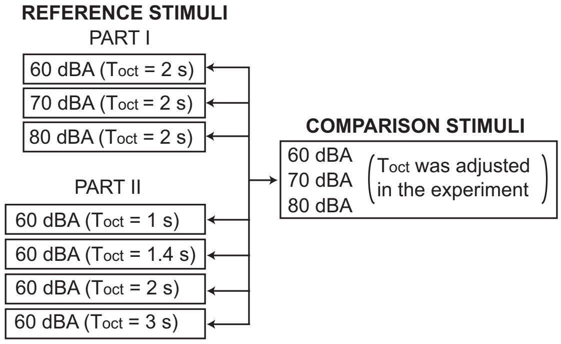

Figure 1. Experiment setups for PART I and II.

The listening experiment basically took the form of a magnitude-matching task. The reverberance of the comparison stimuli was adjusted by pressing 'More' or 'Less' buttons on MATLAB graphical user interface (GUI), which changed the exponential decay rate (i.e. T, as well as EDT) of synthetic RIRs before convolving with the anechoic music sample. An initial T_{oct} of the comparison stimuli was randomly chosen between a logarithmically distributed range from 0.5 s to 4.5 s, which corresponds to 57 steps. The stimuli were listened to via circumaural headphones (Sennheiser HD600) in an anechoic chamber, which has a background noise level below the threshold of hearing specifieid in ANSI S12.2 [15]. According to ISO 3382-1 [5], the just noticeable difference of reverberance corresponds to a 5% change of EDT_{mid} (the subscript 'mid' means an average of corresponding paramter values in the 500 Hz and 1 kHz octave bands). Therefore the pressing the 'More' or 'Less' button changed EDT (and T) of the synthetic RIRs by approximately 4%, which is slightly below one unit of JND of reverberance. Once a subject perceptually matched the reverberance of a comparison stimulus with that of a reference stimulus, they moved to the next pair by pressing 'Next' button on the GUI and repeated the process.

Since we are interested in the relationship between the listening level of a running music stimulus and its reverberance, the anechoic music sample was chosen to have a relatively constant sound pressure level over its playing time. The first 7.9 s of *Water Music* by Handel from Denon Test CD No.2 [16] was used for the anechoic music sample. According to the manufacturer of the CD [16], the music sample had been recorded in anechoic conditions meeting those specified in ISO 3745 [17]. In order to remove terminal reverberance from the stimuli, a very rapid decay was applied at the end of each stimulus (a linear decay of -60 dB over 0.05 s).

The synthetic RIRs were made from two filtered white noise signals. The lower frequency band spanned the 31.5 Hz to 4 kHz octave bands, and the upper band spanned the 8 kHz and 16 kHz octave bands. As RIRs from real auditoria normally have faster sound decay at higher frequencies than the lower frequencies, the sound decay of the two noise bands were separately adjusted by applying Equation 1.

$$p'(t) = p(t) \times \exp\left\{\frac{-\log_e(1000) \times t}{d}\right\} \qquad (1)$$

Here $p(t)$ is sound pressure of one of the white noise bands, t is time in seconds, d is a decay adjustment value and $p'(t)$ is sound pressure of the white noise band after a decay adjustment. In order to make the sound decay for the high frequencies twice that of the low frequencies, the d value for the high frequency band was always half that of the low band. Once the sound decay of the two white noise bands were appropriately adjusted, they were added in time domain and convolved with the anechoic music sample. After the convolution, the listening level of the comparison and the reference stimuli were adjusted according to the experiment calibration requirements.

Our aim was not to have highly realistic RIRs because RIRs from real auditoria are complex – with irregular early reflection sequences and frequency-dependent decay rates. However, we did emulate the gross energy structure of RIRs by drawing on the theory of Barron and Lee [18]. Their study found that energy distribution of RIRs is well-predicted from room volume, source-receiver distance and reverberation time. In order to apply the study to the synthetic RIRs, we set a medium-size virtual auditorium that has a volume of 15000 m^3 and a source-receiver position of 24 m with the initial time delay gap of 20 ms. As we used the music stimuli, the boundary between early and late reflections was chosen to be 80 ms. Figure 2 shows an example of a synthetic RIR with a gross energy distribution according to the Barron and Lee's study [18]. The T_{oct} of the example synthetic RIR is 2 s. We made some gain adjustments to these synthetic RIRs so that, when convolved with an anechoic music stimulus, the LAeq was 60, 70 or 80 dBA.

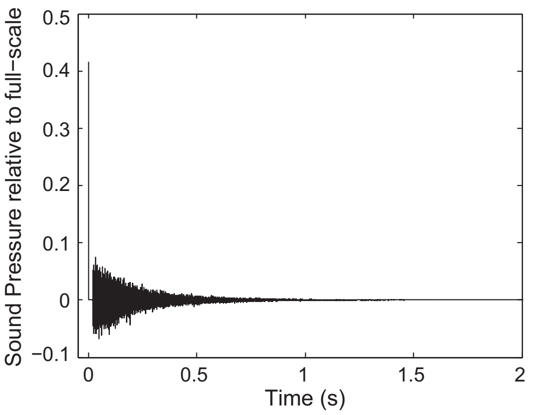

Figure 2. An example of synthetic RIR having a $T60_{oct}$ of 2 s

Fourteen subjects participated in the experiment. Twelve of them had an eduational background in acoustics including room acoustics. None of the fourteen subjects self-reported any hearing loss. Since this study emulates the Soloudre and Bradley's study [6], we used their definition of reverberance:

> The degree of perceived reverberation in a temporal sense. The blending of one sound into subsequent flowing sounds.

RESULTS

As seen in Figure 1, six of the eighteen pairs allow a subject to adjust the comparison to be physically identical to the reference stimulus. Hence the extent to which the subjects matched these stimuli is a straightforward method of assessing reliability of a subject's response (i.e., the ability of a subject to do the task). Figure 3 shows averaged mismatches of these stimuli for each subject. The averaged mismatches were converted into the JND of EDT as represented on the vertial axis of Figure 3. As seen in the figure, subjects 5, 9, 10 and 14 had mean errors greater than three times the JND of EDT (three times the JND of EDT corresponds to a modified EDT of between 1.7 s and 2.3 s for a reference EDT of 2 s). Hence these subjects were excluded from the further analyses.

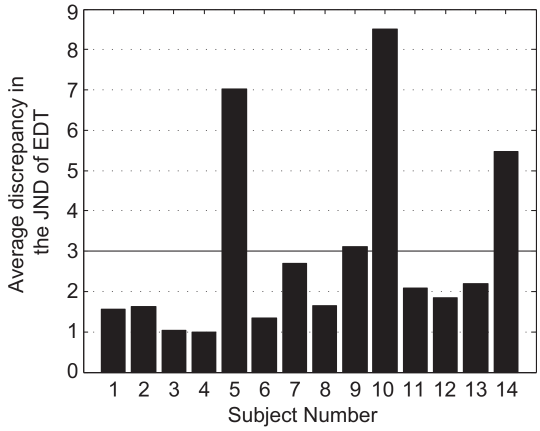

Figure 3. Averages of the unsigned JND of EDT discrepancy between EDTs of the reference stimuli and those of the comparison stimuli having the subjectively matched reverberance.

Table 1 shows the results of an ANalysis Of VAriance (ANOVA) executed on the subject responses for PART I. Using a confidence level of 95 %, values of *Prob>F* less than 0.05 mean that the corresponding variable has a siginificant effect on the subject responses. As seen in the table, the listening level of the comparison stimuli has a significant effect on the subject responses, while the listening level of the reference does not have a significant effect. Also there was not a significant interaction effect between the reference and the comparison stimuli levels.

Table 1. ANOVA for PART I

Variable	*Sum Sq*	*d.f*	*Mean Sq*	*F*	*Prob>F*
Ref Level	1.0962	2	0.5481	2.48	0.0904
Comp Level	1.5151	2	0.75755	3.42	0.0374
*Ref * Comp Level*	0.214	4	0.0535	0.24	0.9139
Error	17.9268	81	0.22132		
Total	20.752	89			

Table 2 shows the ANOVA results given the subjective responses collected in PART II. It also used a confindence level of 95 %. As seen in the table, both the T_{oct} of the reference stimuli and the listening level of the comparison stimuli significantly affect the subjective responses. However there is no interaction effect between these two variables.

Table 2. ANOVA for PART II

Variable	*Sum Sq*	*d.f*	*Mean Sq*	*F*	*Prob>F*
Ref T_{oct}	51.5986	3	17.1929	168.06	0
Comp Level	0.7593	2	0.3797	3.71	0.0276
*Ref T60 * Comp Level*	0.4125	6	0.0688	0.67	0.6724
Error	11.0483	108	0.1023		
Total	63.7987	119			

The subjects' responses were averaged, and RIRs possessing the average responses were generated for analysis. Analysis of these RIRs yielded conventional reverberance predictors (T_{oct}, T_{mid}, EDT_{oct} and EDT_{mid}) and the 'loudness EDT' (EDT_N). For EDT_N calculations, we used the Dynamic Loudness Model by Chalupper and Fastl [11], as implemented in *PsySound3* [19], to obtain the loudness decay curve. An issue was how to calibrate a level of the RIRs when these were input to the loudness model. As there is not a simple relation between the level of convolved stimuli and the level of RIRs, a L_{AFmax} of the RIRs was calibrated at the level of the convolved comparison stimuli. Figure 4 shows an equal reverberance contour constructed from the subject responses of

PART I, plotted as a function of the listening level of the comparison stimulus. The idea here was to provide a graphical indication of the levels of T_{oct} that correspond to equal reverberance percepts for the combination of listening levels for both the comparison and reference stimuli. Symbols represent T_{oct} derived from RIRs possessing the average responses for a given reference stimulus, and trends are shown by linear regression lines. The vertical lines that pass through each of the plotting symbols show error bars for a corresponding dataset with 5 %. In order to show the error bars clearly, the symbols and the error bars are slightly offset. This figure quantifies how much T_{oct} adjustment is required for each comparison stimulus, so as to match their reverberance to that of the reference stimulus (which had a reverberation time of 2 s). As seen in the figure, it is observed that a lower T_{oct} is required for a comparison stimulus that has a higher listening level than the reference stimulus, so as to match their reverberance equally.

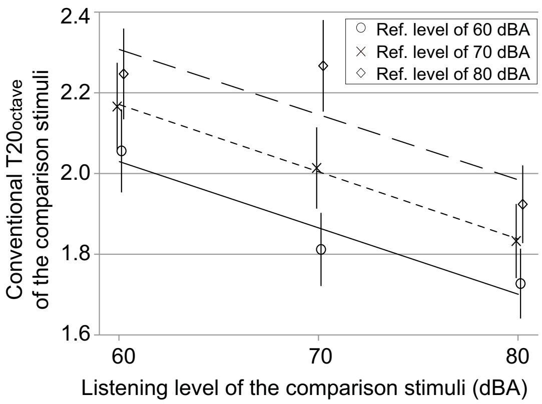

Figure 4. Equal reverberance contour derived from PART I, where the reference stimulus has a reverberation time of 2 s.

Figure 5 shows equal reverberance contours derived from the subject responses of PART II as a function of the listening level of the comparison stimuli. As seen in the figure, a lower T_{oct} is required for a comparison stimulus when its listening level is higher than the listening level of a reference stimulus, so as to have matched reverberance. This tendancy becomes stronger when the reference stimuli have a higher T_{oct} (and diminishes to insignificance when T_{oct} is 1 s).

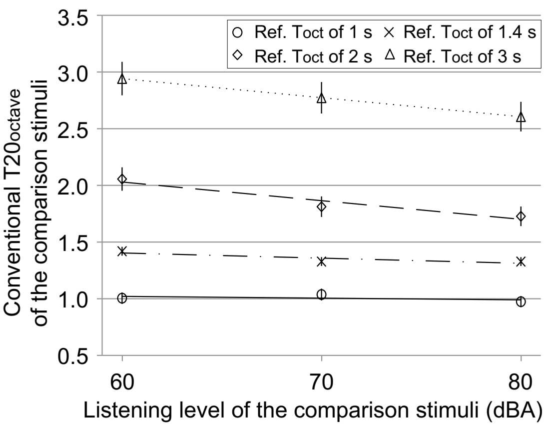

Figure 5. Equal reverberance contours derived from the results collected in PART II

Figure 6 compares the EDT_N with the conventional reverberance predictors derived from the subject responses of PART I. The vertical axis of the figure shows the coefficient of variation (which is the standard deviation divided by the mean of a data set) for a number of predictors derived from RIRs generated from the comparison stimuli that subjectively matched the reverberance of the standard stimulus presented at listening levels that varied over a 20 dB range. The horizontal axis is the level of the reference stimuli. Since the examined reverberance predictors always have positive values, the standard deviation is likely to yield a small number for a data set having lower values. The coefficient of variation takes account of the mean value of a data set, and so removes mean-related biases that would be found in the standard deviation. A smaller value of the coefficient of variation indicates a better prediction of the reverberance because the reverberance of each set of comparison stimuli was matched to that of a single reference stimulus. As seen in the figure, EDT_N outperforms the conventional reverberance predictors for all the listening levels tested in the experiment.

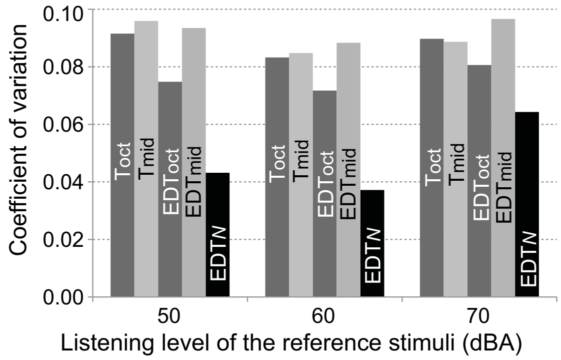

Figure 6. Coefficient of variation for the RIRs generated from mean subject responses in PART I as a function of listening levels of the reference stimuli

Figure 7 compares the EDT_N to the conventional reverberance predictors derived from the subject responses of PART II. As was found in PART I, the EDT_N values yielded the best match to reverberance. The match is even better than it was for the EDT_{oct} which gave the best match in the study by Soulodre and Bradley [6].

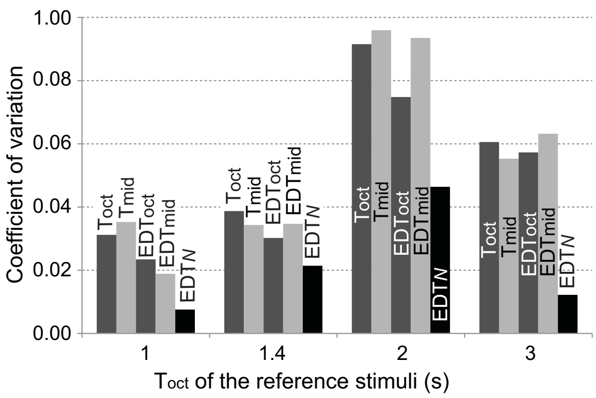

Figure 7. Coefficient of variation for the RIRs generated from mean subject responses in PART II as a function of T_{oct} of the reference stimuli

DISCUSSION

As seen in Table 1, the listening level of the reference stimulus does not significantly affect the subject responses, but this is likely to be due to the small number of the subjects. As this study is preliminary, more subjects will be tested in follow-up work.

As seen in the trend lines in Figure 4, the required T_{oct} to compensate the reverberance difference due to listening level variation is almost constant. For example, the T_{oct} for the situation where the comparision stimulus is 70 dBA and the reference stimulus is 80 dBA is similar to the T_{oct} for the situation where the comparison stimulus is 60 dBA and the reference stimulus is 70 dBA, and so on. Also the figure shows that T_{oct} is close to 2 s when the comparison and the reference stimulus have a same level of 70 dBA. Since T_{oct} of the reference stimuli was 2 s for PART I, this supports the notion that the subject responses are reliable.

The effectiveness of EDT_N in modelling reverberance is evident in Figures 6 and 7. Figure 8 shows the EDT_N for the synthetic RIRs having a L_{AFmax} from 60 dBA to 80 dBA and a T_{oct} of 1 s, 1.4 s, 2 s and 3 s. The symbols indicate the EDT_N derived from the synthetic RIR *without any reverberance adjustment*, and trends are shown as linear regression lines. The bars on the symbols extend to the EDT_N from the synthetic RIRs derived from the comparison stimuli having the reverberance adjustment. As seen in the figure, once the reverberance is subjectively matched, the EDT_N is also closely matched. It also appears that the greater reverberance (which is quantified in EDT_N) needs to be compensated for matching the reverberance when the T_{oct} of the synthetic RIRs increses. Hence this accounts for the phenomenon observed in Figure 5, which is a steeper slope is found for the trend lines of the reference stimuli having a higher T_{oct} values.

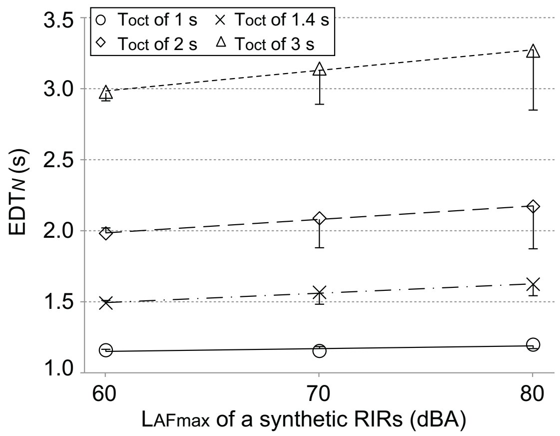

Figure 8. EDT_N derived from the synthetic RIRs *without the reverberance adjustment* (symbols) and from the synthetic RIRs possessing the averaged subject responses of PART II (bars)

CONCLUSION

This paper examined the relation between listening level of a running signal and its perceived reverberance. A loudness-based reverberance predictor, termed EDT_N, was found to outperform more conventional reverberance predictors (such as T and EDT). The experiment results also clearly showed that listening level has a significant effect on the reverberance of the selected running musical stimulus.

REFERENCES

1 L.L. Beranek, *Concert Halls and Opera Houses: Music, Acoustics, and Architecture* (Springer, New York, 2004)

2 W.C. Sabine, *Collected Papers on Acoustics* (Harvard University Press, Cambridge, 1923)

3 H. Haas, "The influence of a single echo on the audibility of speech" *J. Aud. Eng. Soc.* **20**, 146-159 (1951, translated 1972)

4 B.S. Atal, M.R. Schroeder and G.M. Sessler, "Subjective reverberation time and its relation to sound decay" In: Proceedings of the 5th Int. Cong. Acoust. (1965)

5 ISO 3382-1 Acoustics – measuremet of room acoustics parameters – Part 1: Performance spaces (2009)

6 G.A. Soulodre and J.S. Bradley, "Subjective evaluation of new room acoustic measures" *J. Acoust. Soc. Ame.* **98**, 294-301 (1995)

7 D. Lee and D. Cabrera, "Equal reverberance matching of music" In: Proceedings of Acoust. 2009 (2009)

8 D. Lee and D. Cabrera, "Effect of listening level and background noise on the subjective decay rate of room impulse responses: using time-varying loudness to model reverberance" *App. Acoust* doi:10.1016/j.apacoust.2010.04.005 (2010)

9 S. Hase, A. Takatsu, S. Sato, H. Sakai and Y. Ando, "Reverberance of an existing hall in relation to both subsequent reverberation time and SPL" *J. Sound. Vib.* **232**, 149-155 (2000)

10 E. Zwicker and H. Fastl, *Psychoacoustics: facts and models*, 2nd updated Edition, (Springer, Berlin, 1999)

11 J. Chalupper and H. Fastl, "Dynamic Loudness Model (DLM) for normal and hearing-impaired listeners" *Acoustica* **88**, 378-396 (2002)

12 B.R. Glasberg and B.C.J Moore, "A model of loudness applicable to time-varying sounds" *J. Aud. Eng. Soc.* **50**, 331-342 (2002)

13 S.S. Stevens, "The measurement of loudness" *J. Acoust. Soc. Am.* **27**, 815-829 (1955)

14 M. Morimoto and A. Asaoka, "Multi-dimentional analysis of *Reverberance*" In: Proceedings of the 18th Int. Cong. Acoust. (2004)

15 ANSI S12.2 *American national standard criteria for evaluating room noise* (1995)

16 Denon Professional Test CDs, CD-ROM, Nippon Columbia, Japan

17 ISO 3745 Acoustics - Determination of sound power levels of noise sources using sound pressure – Precision methods for anechoic and hemi-anechoic rooms (2003)

18 M. Barron and L.J. Lee, "Energy relations in concert auditoriums I" *J. Acoust. Soc. Ame.* **84**, 618-628 (1988)

19 D. Cabrera, S. Ferguson, F. Rizwi and E. Schubert, "'PsySound3': a program for the analysis of sound recordings" In: Proceedings of Acoustics'08 (2008)