Decay Envelope Manipulation of Room Impulse Responses: Techniques for Auralization and Sonification

Densil Cabrera, Doheon Lee, Manuj Yadav and William L. Martens

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

ABSTRACT

Room impulse responses (RIRs) are very commonly used to represent the acoustic response of rooms for the derivation of acoustical parameters and for auralization. This paper presents a set of signal processing techniques that can be used to enhance the usefulness of recorded RIRs for convolution-based room simulations (which could be classed as a type of auralization), which include using the noise floor to extend the decay, and manipulating the RIR to represent arbitrarily different yet plausible room conditions. The paper also considers how the manipulation of the decay slope can be used to make other features of RIRs more audible, which could have applications in RIR sonification.

INTRODUCTION

The room impulse response (RIR) has become a powerful and widely-used representation of room acoustic conditions for an acoustic signal transmitted between a source and receiver. Room acoustical parameters, such as reverberation time, strength factor, clarity index (International Organization for Standardization 2009) and speech transmission index (International Electrotechnical Commission 2003) may be derived from such RIRs. Binaural room acoustical parameters, including interaural cross-correlation coefficient (IACC), can be derived from binaural room impulse responses (BRIRs) (International Organization for Standardization 2009). Spatial parameters may also be derived from multichannel RIRs describing multiple directional responses (e.g. Gover *et al.* 2004). RIRs are also widely used for auralization by convolution with dry audio signals, either recordings or in real time (Vorländer 2008). In this paper we describe ways of manipulating RIRs, some of which could support applications involving auralization.

The second purpose for the manipulations of RIRs described in this paper is sonification. Analogous to visualisation, sonification is a process by which data are presented to be listened to, so that objective properties of the data can be discerned by ear (see Hermann 2008; Dombois and Eckel forthcoming). Previously we have proposed some approaches to RIR sonification, beyond the trivial case of listening to the RIR without modification (Cabrera & Ferguson 2007). These have included very simple manipulations such as time reversal (which increases the audibility of early reflections due to temporal masking asymmetry), time-stretching (also to increase the audibility of discrete reflections, either by simple resampling, or by techniques such as the phase vocoder) and autoconvolution (which increases the audibility of spectral features). In a related field – the sonification of head-related impulse responses – we developed a sonification technique that simultaneously enhances the audibility of spectral, temporal and spatial features (Cabrera and Martens 2011).

In this paper we focus on the decay envelope of RIRs, considering how the envelope can be modified and even removed. The emphasis of this paper is on techniques, and the techniques described have been used in a variety of projects conducted in our research group.

DECAY MANIPULATION

Extrapolating reverberant decay through the noise floor

Although there are many techniques for maximising signal-to-noise ratio in RIR measurements, measured RIRs inevitably have some level of background noise. If this noise is sufficiently low, such RIRs may be used for listening via convolution with dry signals. However, there are some situations where a noise floor can create audible artefacts. This issue arose in some listening tests that we have conducted, and rather than simply truncating or executing a broadband fade-out of the noise floor, we perform a multi-band extrapolation of the reverberation through the noise.

The point of operating on multiple spectral bands is that the decay rate and the signal-to-noise ratio are both likely to vary across the frequency range of measured RIRs. Usually we divide the spectrum into octave bands, except that the highest band is a high-pass filter, and the lowest band is a low-pass filter. For typical measurements, we find that eight bands (125 Hz – 16 kHz) are workable, but this could be extended to nine or ten bands if the low frequency signal to noise ratio is adequate. Zero phase filtering is used (via Matlab's *filtfilt*) so that synchrony between the bands is maintained. We use 6^{th} order Butterworth filters. Figure 1 gives an example of an octave-band filtered measured RIR, and Figure 2 shows the 4 kHz band in more detail, which has a noise floor more than 60 dB below the direct sound level.

There are various methods that could be used to identify the noise floor, and to identify the decay slope so that it can be applied to the noise floor. Our current approach is simple – which is to model the decay envelope as the sum of an exponential decay and a steady state noise. In order to find the decay envelope, we derive the squared octave band RIR, and smooth this envelope using a low-pass filter with a cut-off frequency of 4 Hz. Again we use zero phase filtering so as to avoid introducing time offsets, and the order of the filter depends on the octave band – lower octave bands have a higher order filter applied. Extracting the envelope (rather than operating on raw values) is helpful because we fit the function on a logarithmic scale (i.e., expressed in decibels), and using the envelope avoids very low sample values (which could be as low as -∞ dB). We identify the peak, and normalize it to 0 dB. To this decay envelope (starting from the peak), we fit a curve in the form of equation 1.

$$L(t) = 10\log_{10}\left(10^{at/10} + b\right) \qquad (1)$$

In this equation, t is time in seconds, and $L(t)$ is the fitted function. The parameter a controls the slope of the decay, and it equals the level decay after 1 second (notwithstanding the noise floor). Hence an initial estimate of reverberation time can be taken as $-60/a$. The parameter b specifies the noise floor – i.e., the noise level is $10\log_{10}(b)$ relative to the envelope peak. Note that the smoothing that is applied to the envelope smooths the initial peak, thereby bringing the noise floor somewhat higher relative to the peak. The middle chart of Figure 2 shows the decay envelope of an octave-band filtered measured RIR, together with the fitted curve $L(t)$.

As we have now separated the decay envelope parameter from the noise envelope parameter, we can now derive a gain compensation function from the ratio between a noise-free decay envelope and the modelled envelope. Equation 2 expresses this as an amplitude gain, $g(t)$ – i.e. a series that can multiply the octave band filtered RIR directly.

$$g(t) = \sqrt{\frac{10^{at/10}+b}{10^{at/10}}} \qquad (2)$$

It makes no sense to apply $g(t)$ to the entire octave band RIR because that would unduly smooth out the early decay that is unaffected by background noise. We apply it from the point of the envelope that is 10 dB above the noise floor. Figure 1 and the bottom chart of Figure 2 show the effect of applying this process to a measured RIR – and Figure 2 shows this process combined with reverberation time adjustment, as described in the next section.

Once this process is applied to each of the RIR bands, they can then be recombined by summation. Usually each band has a distinct reverberation time and signal to noise ratio, so that the transition between the true decay and the decaying noise occurs at a different point in time for each band (such as in Figure 1). Using this approach it is feasible to have a decay seamlessly extending well-beyond the limit of audibility, and indeed down to the lowest quantization level of a 16-bit or 24-bit waveform.

It might be recalled that a related approach to modelling RIRs was proposed by Xiang (1995) for the derivation of reverberation time. He applied a non-linear regression to the reverse-integrated squared RIR so that the effect of noise on the measurement could be removed (by including the effect of reverse-integrated steady state noise in the model). While our intention here is not reverberation time measurement, it is possible that the technique described here could be used to 'clean up' RIRs prior to reverse integration in the calculation of reverberation time of noise-affected RIRs.

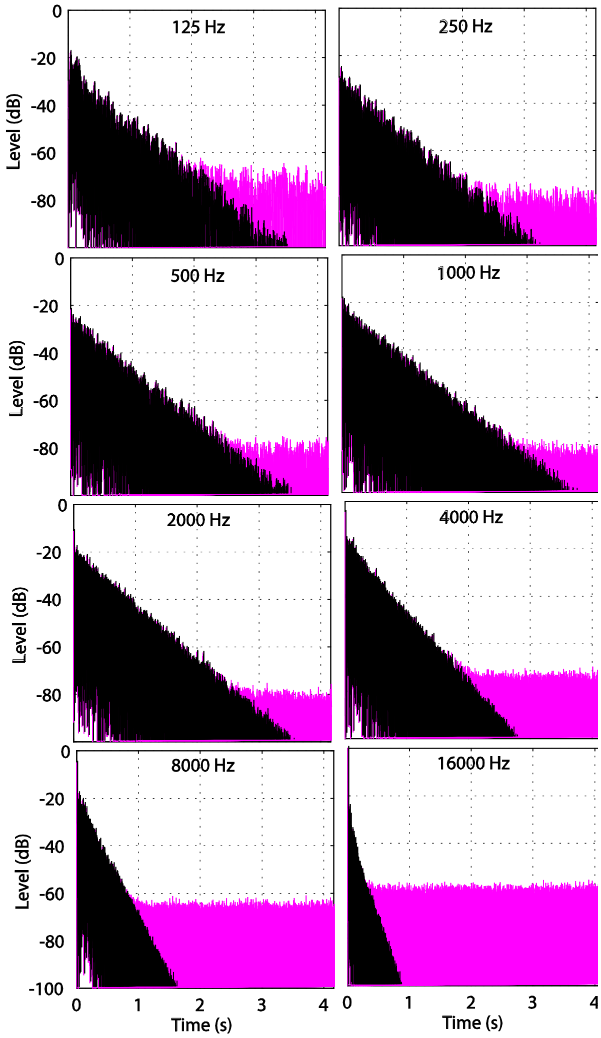

Figure 1. Example of treating a measured RIR for noise-floor removal in eight octave bands (except that the highest and lowest bands are high-pass and low-pass filters respectively). The original RIR is shown in pink (squared, and expressed in dB), and the modified RIR is shown in black for each spectrum band. The pink function is essentially identical to the black function until about 10 dB above the noise floor.

Although the assumptions underlying the fitted function may seem simplistic, in fact this procedure works well for typical recorded RIRs. It could be applied in narrower bands (for example, one-third octaves), but applying it in octave bands tends to provide excellent results. Curve-fitting issues can arise (i) when the signal to noise ratio is very low (for example, in octave bands outside the measurement range of the loudspeaker), (ii) when the reverberant decay deviates considerably from an exponential function (for example, in coupled rooms), and (iii) when the noise floor envelope is not constant. The first of these curve-fitting issues is avoided by choosing appropriate spectrum bands for processing (and using high quality measured RIRs). Should the second and third issues arise, they could be dealt with by a modified curve-fitting process. This could be done by introducing two additional parameters that control the decay rate and gain of a second decay slope, which has a longer reverberation time and reduced gain compared to the primary decay function. However, for automated processing, criteria might need to be developed to distinguish a true double-slope decay from a single-slope decay that has a decaying noise tail – since in the latter case the noise should be adjusted to match the single-slope decay, while in the former case the double-slope decay should be preserved (and extended). Another potential problem could occur with tonal background noise. In such a case it is better to substitute band-limited synthetic noise that follows the derived exponential decay function for the actual noise of the measured RIR (*cf.* Menzer and Faller 2010).

Adjusting reverberation time

The reverberation time of a RIR may be adjusted by multiplying the RIR by an exponential function. A positive exponent increases reverberation time, and a negative exponent reduces it. The exponential damping constant, δ of a RIR band is related to reverberation time, T, by equation 3.

$$\delta = \frac{\ln(10^6)}{2T} \approx \frac{6.91}{T} \tag{3}$$

Equation 4 expresses δ in relation to an ideal exponential reverberation decay envelope (neglecting fine features of the wave), where x represents envelope amplitude, t is time in seconds, and x_0 is the value of x at $t = 0$.

$$x^2(t) = x_0^2 e^{-2t\delta} \tag{4}$$

A new reverberation time is required with its envelope described by equation 5 (and a damping constant of δ_1; we will use δ_0 in subsequent equations to denote the original RIR's damping constant).

$$y^2(t) = x_0^2 e^{-2t\delta_1} \tag{5}$$

We need to find $a(t)$ such that $y(t)=x(t).a(t)$. Using equations 4 and 5, we set

$$x^2(t)e^{-2t\delta_1} = a^2(t)x^2(t)e^{-2t\delta_0} \tag{6}$$

so

$$a^2(t) = e^{-2t(\delta_1-\delta_0)}. \tag{7}$$

Hence an adjustment to reverberation time can be applied quite simply if the original band exponential damping constant, δ_0, and the desired exponential damping constant, δ_1, are known:

$$y(t) = x(t)e^{-t(\delta_1-\delta_0)} \tag{8}$$

In equation 8, $x(t)$ is the original RIR (or an octave band oe narrower band component of it), and $y(t)$ is the RIR with a modified reverberation time. The exponent is not multiplied by 2 because we are operating on pressure rather than pressure squared.

A subtlety of reverberation time measurement in this context is that the time spanned by the evaluation range of T20 or T30 depends on the reverberation time. For example, the T20 evaluation range is fixed as -5 dB to -25 dB (this is the range over which a linear regression is made on the reverse-integrated energy decay function, expressed in decibels, in order to determine the decay slope), but this corresponds to differing time periods as the decay slope is changed. Since RIR decays are not exactly exponential, this can yield a different reverberation time value to the one aimed for. Therefore we use an iterative process to refine the decay rate adjustment to yield a specified T20 value.

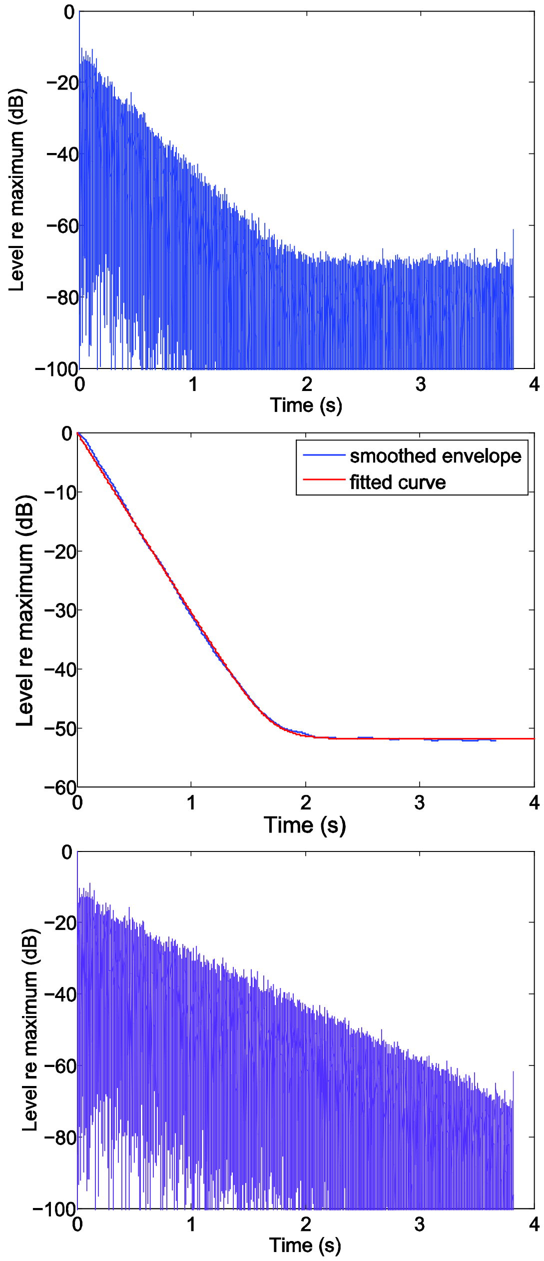

Figure 2. Example of treating a measured RIR for noise-floor removal and reverberation time change. The top chart shows the RIR filtered into one octave band (centred on 4 kHz) – values are squared and expressed in decibels for chart readability. The middle chart shows the smoothed envelope function of the octave band RIR, together with the fitted curve as described in the text. The bottom chart shows the modified octave band RIR, with the noise tail used as an extension of the decay, and the reverberation time changed from 1.9 s to 4.0 s.

We tend to apply decay rate adjustment in octave bands (using 6th order Butterworth zero-phase filters), so that any RIR can be transformed to one with the desired octave band reverberation time. Using octave bands derived in this way, the reverberation time of individual bands can be adjusted with very little interaction between bands, so that physically unrealistic reverberation spectra (with large contrasts between adjacent bands) can be made. However, finer (or, indeed, coarser) spectral resolution than that achieved by octave bands can be applied just as easily.

Extrapolating the reverberation decay through the noise floor (as described in the previous section) is especially important when reverberation time is being increased. Without this, the noise may grow in the tail of the modified RIR, leading to disturbing audible artefacts. Figure 2 gives an example of an octave band filtered RIR that has had its reverberation time increased from 1.9 s to 4.0 s, and such a change would not have been useful without treating the noise floor.

We have used these techniques in a series of listening tests of reverberance (e.g., Lee and Cabrera 2010; Lee *et al.* 2009 & 2010). In those experiments, participants adjusted the decay rate of one stimulus to match the decay rate of another, and so the implementation required reasonably fast processing (especially when the varied RIR needed to be convolved with a dry signal to prepare it for listening). This processing was implemented in Matlab. We have also used this processing in the preparation of recorded RIRs for a real-time binaural auralization system for one's own voice (the prototype system is described in Cabrera *et al.* 2009).

Removing the decay envelope

As a special case of reverberation time adjustment, a RIR can be given infinite reverberation time, meaning that the envelope is approximately steady state. There is more than one way to do this, but to follow on from the decay rate adjustment method outlined in the previous section; one approach is to multiply the RIR, $x(t)$, by an exponential function that *grows* with the same constant as the RIR's decay, δ_0. In equation 9, $y(t)$ is the RIR with decay removed.

$$y(t) = x(t)e^{(t\delta_0)} \qquad (9)$$

Another way of doing this might be to derive a heavily smoothed envelope function, and to divide the RIR by that function.

The purpose of transforming RIRs in this way may not be immediately obvious. One reason for doing it might be to create a generic format for storing the fine temporal data from RIRs (in multiple bands) which can be transformed into RIRs possessing a desired reverberation time spectrum with a minimum of calculation (in time-critical applications). Another, more interesting, reason for doing it could be for listening.

In listening to an RIR, decay is almost always its most obvious audible feature. Hence, removing decay from the RIR allows a listener to focus more on more subtle features, such as the evolution of the spectrum over time, the fine structure of early reflections, or the evolution of interaural short-term cross-correlation (in the case of binaural RIRs). For such listening purposes, the RIR can be processed unfiltered. An example of such processing is shown by Figure 3 as a spectrogram. This example was generated without noise floor treatment, so that the usefulness of such treatment is made more evident – hence a growing noise floor is seen from 1.4 s onwards in the treated RIR. Apart from this artefact, listening to this example makes the evolution of the reverberation spectrum much more obvious after treatment than before.

In the case of non-exponential reverberation decays (such as double-slope decays), it is not possible to achieve a steady envelope using this approach. In a double slope decay, if the first part of the slope is transformed to have an infinite reverberation time, then the latter part will grow (i.e., it will have a negative reverberation time). Such deviations from steady state may draw a listener's attention to an RIR's deviation from exponential decay.

A further use for RIRs without decay could be as part of an RIR manipulation process, such as the one described in the next section.

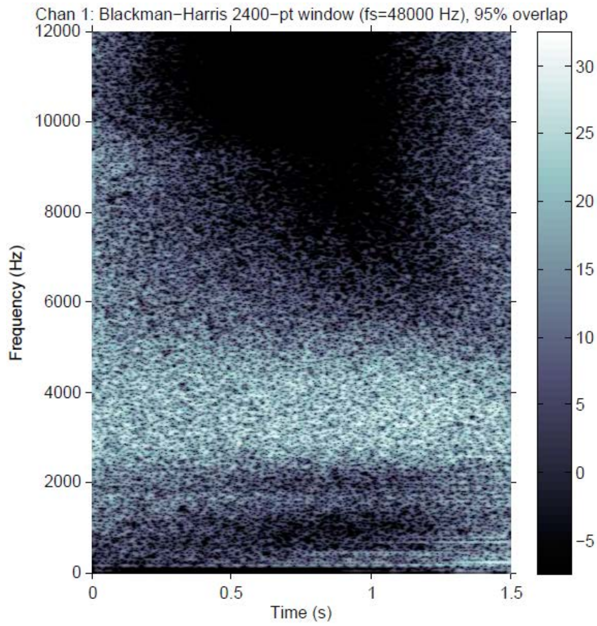

Figure 3. Spectrogram of an RIR that has been processed to have an infinite broadband reverberation time (albeit truncated to a period of 1.5 s).

Adjusting fine envelope contrast

The manipulations described so far are concerned with the large scale form of the decay envelope. Once a gross decay envelope is known, it is then possible to manipulate the extent to which the fine structure of the RIR deviates from that envelope. A simple approach to this is to remove the decay, then manipulate the non-decaying data, and finally to apply the desired decay rate (which could merely involve reinstating the original).

As indicated by Equation 10, the fine fluctuations in a decay-less RIR, $x(t)$, can be derived by taking the magnitude of its Hilbert transform, $|\mathrm{H}(x(t))|$. This envelope can then be normalised (divided by its maximum value) and raised to an exponent, n, so as to yield a compression/expansion function for the decay-less RIR. The product of this function and the decay-less RIR yields the result, $y(t)$. To expand the fine fluctuation range, n is greater than zero.

$$y(t) = x(t)\left[\frac{|\mathrm{H}(x(t))|}{\max|\mathrm{H}(x(t))|}\right]^n \qquad (10)$$

More complex approaches can be taken to gain mapping, as an alternative to using a simple exponent. For example, a threshold can be defined, above which the envelope is unchanged, and below which the envelope is expanded or compressed. In practice, applying a lowpass filter (with a cut-off frequency below 20 Hz) can be beneficial in smoothing the envelope prior to application as described above.

The purpose of manipulating RIRs in this way is to adjust the apparent 'echo density', although in fact the manipulation does not change the actual echo density (which could be achieved by time-stretching or compressing the RIR, for example by using a phase vocoder algorithm).

CONCLUSION

This paper has described a set of techniques that can be used to manipulate RIRs in simple and powerful ways. Uses of the techniques include the preparation of material for auralization, and to enhance RIR feature audibility for sonification. The decay slope is the most prominent feature of RIRs, and manipulating it can either be done because of its importance, or else to reduce its prominence so that other features can be better appreciated. The decay rate manipulations described in this paper are implemented in a set of Matlab functions, which are available from the authors.

REFERENCES

Cabrera, D & Ferguson, S 2007, 'Sonification of sound: tools for teaching acoustics and audio', in *Proceedings of the 13th international conference on auditory display*, Montreal, Canada

Cabrera, D & Martens, WL 2011, 'Sonifying head-related transfer functions, in Y Suzuki, D Brungart, Y Iwaya, K Iida, D Cabrera & H Kato (eds), *Principles and applications of spatial hearing*, World Scientific, New Jersey, pp. 473-83.

Cabrera, D, Sato, H, Martens, WL & Lee, D 2009, 'Binaural measurement and simulation of the room acoustical response from a person's mouth to their ears', *Acoustics Australia*, vol. 37, no. 3, pp. 98-103.

Dombois, F & Eckel, G forthcoming, 'Audification', in T Hermann, A Hunt & J Neuhoff (eds), *The sonification handbook*, Logos, Berlin, pages in press.

Gover, BN, Ryan, JG & Stinson, MR 2004, 'Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array', *Journal of the Acoustical Society of America*, vol. 116, no. 4, pp. 2138-2148.

Hermann, T 2008, 'Taxonomy and definitions for sonification and auditory display', in *Proceedings of the 14th international conference on auditory display*, Paris, France.

International Electrotechnical Commission 2003, *Sound system equipment – part 16: objective rating of speech intelligibility by speech transmission index*, IEC 60268-16 Ed. 3.0, International Electrotechnical Commission, Geneva.

International Organization for Standardization 2009, *Acoustics – measurement of room acoustic parameters – part 1 performance spaces*, ISO 3382-1:2009, International Organization for Standardization, Geneva.

Lee, D & Cabrera, D 2010, 'Effect of listening level and background noise on the subjective decay rate of room impulse responses: Using time-varying loudness to model reverberance', *Applied Acoustics*, vol. 71, no. 9, pp. 801-11.

Lee, D, Cabrera, D & Martens, WL 2009, 'Equal reverberance matching of music', in *proceedings of Acoustics 2009*, Adelaide, Australia.

Lee, D, Cabrera, D & Martens, WL 2010, 'Equal reverberance matching of running musical stimuli having various reverberation times and SPLs', in *proceedings of the international congress on acoustics*, Sydney, Australia.

Menzer, F & Faller, C 2010, 'Investigations on an early-reflection-free model for BRIRs', *Journal of the Audio Engineering Society*, vol. 58, no. 9, pp. 709-23.

Vorländer, M 2008, *Auralization*, Springer, New York.

Xiang, N. 1995, 'Evaluation of reverberation times using a nonlinear regression approach', *Journal of the Acoustical Society of America*, vol. 98, no. 4, pp. 2112-2121