On the visualization and modification of room impulse responses for sound design

Frank Melchior[1][2], Diemer de Vries[1]
[1] *TU Delft, Faculty of Applied Physics, The Netherlands*
[2] *Fraunhofer IDMT, Ilmenau, Germany*

Introduction

The use of psycho-acoustics in user interfaces for sound design can enhance the usability of sound design tools. To make use of such knowledge a psycho-acoustic layer in between the signal processing and the user interface is required. This layer makes use of perceptual models for visualization and interaction. This paper describes the extension of an interaction model for room impulse responses (RIR) based on energy decay curves (EDC) and energy decay relief (EDR) plots.

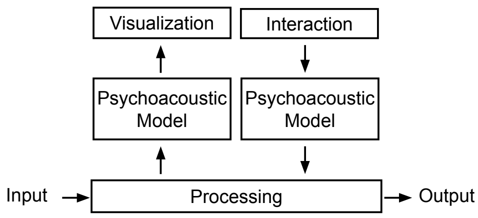

Figure 1: Block diagram of the integration of psycho-acoustic in sound design systems.

Inverse energy decay curve

The EDC of a RIR plotted with an logarithmic amplitude scale is highly correlated with the perception of the decay of a diffuse sound field. Detailed fluctuations as shown in a time plot of a RIR are not perceived in detail. The modification of the envelope of measured RIR is a well known tool in sound design. The drawback of such a modification is that the envelope defined by the user is not highly correlated to the perception. For this reason a direct modification of the energy decay curve is proposed for sound design application. After the definition of the desired EDC by the user an envelope can be generated to modify the impulse response in a way that the new curve is achieved. The EDC $s(t)$ of an impulse response $h_1(t)$ can be calculated using the Schroeder integral [1]:

$$s(t) = \int_{t}^{\infty} h^2(\tau) d\tau \qquad (1)$$

In case of a sampled room impulse response $h_1(n)$ with N samples the following expression is used:

$$s_1(m) = T \sum_{n=m}^{N} h_1^2(n) \qquad (2)$$

with $f_{sample} = \frac{1}{T}$. If a user modifies the energy decay curve we get a modified version $s_2(n)$ as shown in the upper part of figure 2. To calculate the required envelope directly form the two curves $s_1(n)$ and $s_2(n)$ their difference is required:

$$s_d(n) = s_1(n) - s_2(n) \qquad (3)$$

From the result the required envelope $e(n)$ is calculated with:

$$e(n) = \sqrt{1 - \frac{s_d(n) - s_d(n+1)}{h_1^2(n)T}} \qquad (4)$$

The generated envelop $e(n)$ using $s_1(n)$ and $s_2(n)$ is shown in the lower part of figure 2. Equation 4 represents the required envelop for the modification of the impulse response $h_1(n)$ to get the desired RIR $h_2(n)$ using:

$$h_2(n) = e(n) \cdot h_1(n) \qquad (5)$$

The block diagram in figure 3 summarizes the signal processing. The new layer introduced for the optimized interaction process can be found in the middle representing a simple "psycho-acoustic" model.

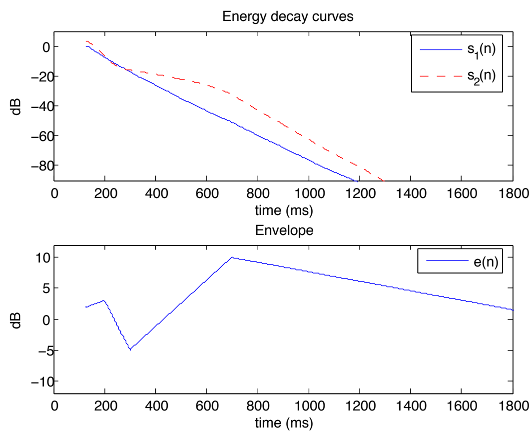

Figure 2: Original and modified energy decay curve with resulting envelop for impulse response processing.

Modified energy decay relief plot

The EDC and its inversion presented in the first part of the paper was originaly proposed to be used in the time domain. Jot et al. have extended the EDC to the energy decay relief (EDR) plot [2]. The visualisation can be interpreted as a frequency dependent EDC. From a signal processing point of view the use of short time Fourier transformation (STFT) is a appropriate solution to filter the RIR and calculate the EDR. In [3] a method for intuitive modification of EDR plots was proposed. The work

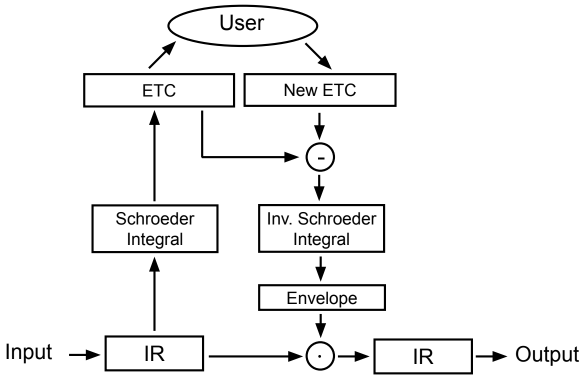

Figure 3: Block diagram for the modification of the energy decay curve (ETC).

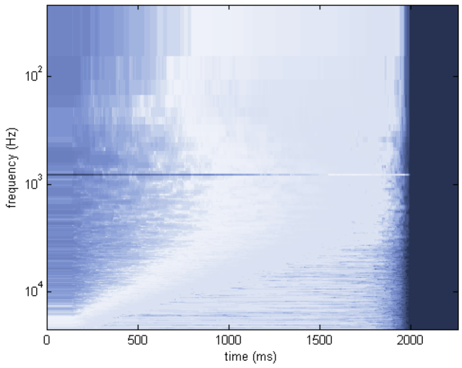

Figure 4: Short time Fourier spectrum representation of a EDR with one attenuated frequency bin.

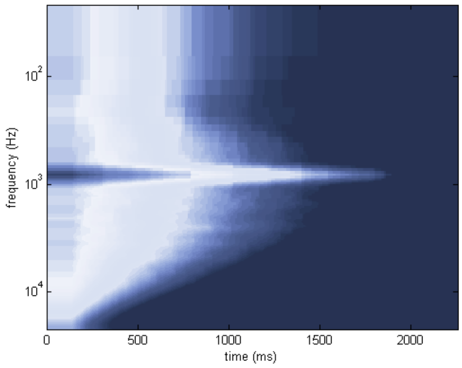

Figure 5: Internal spectrum representation of a EDR with one attenuated STFT frequency bin.

presented here is an extension of this proposal by using the internal spectrum of the auditory system described in [4]. To model the auditory filter Patterson proposed the following window in the frequency domain $W(f, f_c)$ for a given center frequency f_c:

$$W(f, f_c) = \left[1 + \frac{4|f - f_c|}{ERB(f_c)}\right] \cdot exp\left[-\frac{4|f - f_c|}{ERB(f_c)}\right] \quad (6)$$

In this equation $ERB(f_c)$ defines the Equivalent Rectangular Bandwidth at frequency f_c. Moore and Glasberg found the following empirical expression for $ERB(f_c)$:

$$ERB(f_c) = 6.23f_c^2 + 93.39f_c + 28.52) \cdot 10^{-3}(kHz) \quad (7)$$

For the graphical representation of the EDR plot a new spectrum corresponding to equation 6 is used:

$$|S_{int}(f_c, t)|^2 = \frac{\int_{-\infty}^{\infty} W(f, f_c)|S(f, t)|^2 df}{\int_{-\infty}^{\infty} W(f, f_c) df} \quad (8)$$

$S(f, t)$ are the time dependent spectra resulting from a STFT. The denominator of equation 8 results in a flat spectrum for white noise. This representation can be understood as a smoothing of the spectrum because single frequencies are smeared over a large frequency range. On the other hand detailed modifications of the impulse responses become visible depending on their auditory relevance. The smoothed EDR plot is calculated using:

$$\zeta(f_c, t) = \int_t^{\infty} |S_{int}(f_c, \tau)|^2 d\tau \quad (9)$$

Figure 4 presents a EDR based on a short time Fourier representation. One frequency bin is attenuated. Figure 5 is generated using the internal spectrum $\zeta(f_c, t)$.

For the modification of the plot through the user the inverse energy decay plot is used in each frequency band separately. The weighting of the frequencies around a certain frequency selected by a user is calculated using equation 6. The resulting envelopes are used to modify the magnitude spectrum of the underlaying STFT.

Summary

The principles of sound design based on the integration of psycho-acoustic knowledge into the user interface and processing chain has been presented in this paper. These principles are implemented for the modification of measured room impulse responses. Through the large knowledge of psycho-acoustics due to extensive research in this field several more principles can be adopted to optimize the sound design interaction process. Such principles have to be evaluated in further user studies.

The presented work is part of the project EDcine, supported by the IST 6th framework program of the European Commission (http://www.edcine.org).

References

[1] M.R. Schroeder.: New method for measuring reverberation time,J. Acoustic Soc. America 37, 1965

[2] J.-M. Jot, L. Cerveau, O. Warusfel.: Analysis and synthesis of room reverberation based on a statistical time-frequency model, 103rd AES Convention, 1997.

[3] F. Melchior, J. Langhammer, D. de Vries.: A new approach for direct interaction with graphical representations of room impulse responses for the use in wave field synthesis reproduction, 120th AES Convention, 2006

[4] W. de Bruijn.: Application of wave field synthesis in videoconferencing, PhD Thesis, TU Delft, 2004