HYBRID REVERBERATION PROCESSOR WITH PERCEPTUAL CONTROL

Thibaut Carpentier, Markus Noisternig, Olivier Warusfel
UMR STMS IRCAM-CNRS-UPMC,
1, place Igor Stravinsky
75004 Paris, France
{tcarpent, noisternig, warusfel}@ircam.fr

ABSTRACT

This paper presents a hybrid reverberation processor, i.e. a real-time audio signal processing unit that combines a convolution reverb for recreating the early reflections of a measured impulse response (IR) with a feedback delay network (FDN) for synthesizing the reverberation tail. The FDN is automatically adjusted so as to match the energy decay profile of the measured IR. Particular attention is given to the transition between the convolution section and the FDN in order to avoid audible artifacts. The proposed reverberation processor offers both computational efficiency and flexible perceptual control over the reverberation effect.

1. INTRODUCTION

During the last decades, two main approaches for digital artificial reverberation processing have been widely used in music and film production [1]: convolution reverbs and delay-network techniques. In this paper, we present and discuss a novel hybrid reverberation processor that combines both methods and overcomes some of the limitations of earlier approaches. A hybrid reverberation effect processor should have the following properties:

- the hybrid reverberation effect should be perceptually indistinguishable from a pure convolution reverb for a given room impulse response;
- the algorithm should fulfill the constraints for real-time audio signal processing;
- the algorithm should be computationally efficient and thus attractive for practical applications;
- the processing method should provide a flexible high-level control over the perceived room effect.

This paper is organized as follows: The remainder of Section 1 briefly discusses the motivation for this study and offers a summary of earlier works and current state of the art methods for reverberation effect processing. Section 2 details the technical aspects of the proposed method. In Section 3, we present the results of a case study. In Section 4, we extend the hybrid reverberator with a perceptual control paradigm. Section 5 discusses some of the limitations of the proposed method and outlines possible future improvements.

1.1. Convolution-based reverberators

The acoustic transfer path between an emitter and a receiver in a room is usually modeled as a linear time-invariant system, which is fully characterized by its impulse response (IR). With this linear model, the room reverberation can be reproduced by convolving an anechoic input signal with the respective room impulse response. This convolution-based "auralization" approach guarantees for an authentic and natural listening experience.

Due to the increase in available processing power and recent advances in the development of computationally efficient low latency algorithms for frequency domain filtering (such as, e.g., the block-partitioned FFT convolution [2] and frequency delay lines [3]), convolution-based reverberation processing became widely applied during the last few decades. However, the computational cost of this method depends on the length of the processed IR. This may become a problem when recreating the reverberation of large concert halls and opera houses, where the length of the IR is typically in the order of a few seconds.

A survey on available convolution-based reverberation rendering software and hardware devices shows that the control over the reverberation effect is, in general, limited to only a few low-level parameters. Typically, the early-to-reverb ratio can be modified by adjusting the gains of the respective time sections of the IR. Often the decay time can be varied too, i.e., either increased or reduced. This can be achieved, for example, by resampling the original IR or by applying an exponentially decaying gain curve to the late reverberation tail. More advanced IR transformations often yield artifacts that result in an unnatural or unpleasant sounding reverb. This clearly limits the range of possible IR transformations in current convolution-based reverberators.

1.2. Parametric reverberators (FDNs)

Jot and Chaigne [4] used feedback delay network (FDN) processing structures for digital reverberation rendering. FDN simulate the statistical properties of the late room reverberation in a computationally efficient way. They are scalable and allow for a continuous tuning of the time and frequency behavior of the room response.

A commonly reported drawback of FDN rendering is the lack of authenticity in the early part of the room response. This is typically linked to transient coloration effects or from insufficient echo and/or modal densities, as it takes some time to build up dense reflection patterns with feedback loop structures.

1.3. Motivation for developing hybrid reverberators

This work aims at developing a hybrid reverberator that combines both convolution processing for the early part of the IR and FDN for late reverberation rendering. The hybridization approach shows several advantages over full convolution processing. Early reflections (ER) typically arrive within less than $50 - 200$ ms. Applying convolution filtering to this part of the IR comes with a low computational cost, while it preserves the naturalness and spectral signature of the room response. The late reverberation decay, which may be several seconds for large rooms, can be accurately modeled with computationally efficient FDNs. The feedback loop structure offers flexible control over the rendering parameters and can be adapted to perceptually-motivated control methods (see Sec. 4 for further details). The two main challenges for the design of such hybrid processor are:

- the estimation of model parameters from the original (e.g., measured) IR for automatic tuning of the FDN;
- to guarantee smooth transitions between the two processing stages (i.e. at the transition between early reflections and reverberation tail) without perceptible artifacts.

1.4. Related works

The idea of combining FIR filter for early reflection modeling with a recursive topology for modeling late reverberation decays dates back to early works on digital artificial reverberation (see e.g. [5–7]), even though actual attempts at hybridization only appeared in the late 2000s.

Stewart [8, 9], for instance, proposed a hybrid reverberator using a 16-channel FDN for generating the late reverberation. This reverberator automatically estimates the FDN parameters from the energy decay relief (EDR). More precisely, the reverberation time (RT) is estimated in each frequency band. The initial spectrum of the FDN is, however, not taken into account. A Hann window assures smooth cross-fading between the concatenated sections (i.e. early reflections and late reverberation) and minimizes perceptible artifacts. Although Stewart et al.'s method is very similar to what is proposed in this paper (see Sec. 2), they only demonstrate that such a hybrid reverberator is viable. To the authors' knowledge, it has never been realized in practice. It should be further noted that this cross-fading approach is not well suited for real-time implementations. The rising edge of the Hann window is applied at the beginning of the late reverberation, which is not possible in real-time.

A similar approach is taken in [10], without providing detailed information on the crossfade in between the two sections. Here, the reverberation decay times are estimated in only two frequency bands and a 16-channel FDN is adjusted to match the original IR at the transition points. No additional spectral shaping is applied to the FDN.

Abel et al. [11] model a plate reverberator with a hybrid processing unit. This method (which is inspired from [8]) first estimates the spectral decay times and then applies them to the FDN. For equalizing the FDN a short FIR filter, which is obtained from a minimum-phase version of the impulse response, is applied to the transition region. The transition between the convolution section and the FDN is accomplished by means of a power-complementary crossfade.

Greenblatt et al. [12] further extended the methods presented in [8] and [11] by improving the window-based crossfade between the convolution and FDN sections. This method allows for any arbitrary window shape and length as it is subtracted from the convolution part.

In [13], Lee et al. generate the ER section using conventional convolution techniques and the late reverberation part with a so-called "switched convolution (SC)" technique. The SC processor consists of a recursive comb filter that is convolved with a short noise segment. The transition between the two processes uses the cross-fading technique developed in [11].

Other works mainly focus on the optimization of the different processes: Heise et al. [14] proposed an optimization strategy for matching the settings of two different audio processors. As a case study they tune an algorithmic reverb so that it mimics a convolution reverb processor. The optimization procedure evaluates the differences between the actual response and the target response on the basis of psychoacoustic features. As a principal measure the euclidian distance between MFCC vectors is applied.

A hybrid reverberation processor with a Moorer structure for the reverb tail is used by Primavera et al. [15–18]. It is based on an iterative optimization algorithm (see [19] for details) to determine the parameters of an IIR filter structure (i.e. delay line lengths, gains, and damping factors) that jointly minimize different cost functions. The cost functions are obtained by comparing the synthesized IR with the real IR in both the time and frequency domain.

Holm-Rasmussen et al. [20] apply linear predictive coding to fit a synthesized reverberation tail to a measured IR. For synthesis sparse FIR filters are used.

Several works investigate different time-frequency representations to estimate the model parameters that best approximate a given room impulse response (see e.g. [21–23] for further details). Most methods apply the short-time Fourier or wavelet transform; the parameters of the filter structures are estimated using the Prony or Steiglitz-McBride method.

2. PROPOSED METHOD

In this paper we propose a method that (automatically) tunes a FDN unit to best approximate the time-frequency response of the original IR. Schroeder (see e.g. [24, 25]) statistically modeled the late reverberation of a room as exponentially decaying Gaussian random process. It is shown in the following that, when applying Schroeder's statistical model, the FDN can be fitted with arbitrary accuracy to both the reverberation decay profile and the initial spectrum of the original IR.

2.1. IR truncation

A time-domain room impulse response can only be Gaussian when a sufficient number of reflections overlap, i.e. when the echo density in a room is sufficiently high enough. The stochastic model for late reverberation is thus only valid for frequencies higher than the Schroeder frequency [24] and times later than the mixing time (t_{mix}). The mixing time determines the transition between the ER and the reverberation tail and thus defines the cross-over point between the convolution section and the FDN section of a hybrid reverberator. Several estimators for the mixing time have been proposed in literature (see e.g. [26–32]), with varying results that, in general, strongly depend on the estimation parameters (e.g. the size of the analysis window). An objective comparison of the performance of these various estimators is beyond the scope of this paper and is postponed to future publications. (Note that Lindau et al. [33] presented a comparative study on the estimation of the perceptual mixing time). For the remainder of this article it is assumed that the mixing time is estimated with sufficient accuracy and that the IR can be modeled as decaying Gaussian random process for times $t \geq t_{mix}$.

2.2. Feedback delay network reverberator

The hybrid reverberation engine presented in this paper is based on IRCAM's parametric reverberation engine, which is part of the sound spatialization software "Spatialisateur" (Spat~). This FDN-based reverberator consists of a "lossless prototype" (i.e. a reverberator with infinite reverberation time that is based on lossless unitary feedback matrix structures) combined with absorptive IIR filters (see [34] for more details). With this processing structure one can achieve arbitrary time and modal densities, low tonal coloration, and independent control of the frequency envelope and decay characteristics [4]. The Spat~ reverberation processor can be controlled by a set of perceptual descriptors (see Sec. 4). These descriptors rely on a simplified model of the IR's time-frequency energy distribution that is reduced to four time segments and three frequency bands (cf. Fig. 1). The Spat~ model separates the IR into three sections: (a) "early" for the very first discrete echoes, (b) "cluster" for the late and more diffuse early reflections with a dense reflection pattern, and (c) "late reverb" for the late reverberation. The cluster is synthesized with multi-tap delay lines feeding a decorrelation unit; the late reverb is generated by a delay-network (that is fed by the output of the cluster section) with typically 4 to 16 feedback channels.

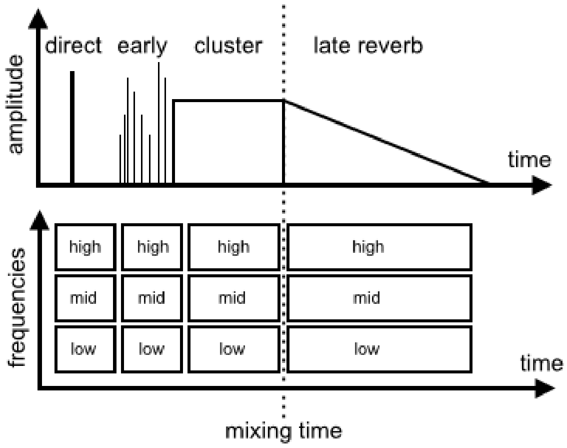

Figure 1: *Time-frequency IR model of the FDN-based reverberator Spat~: echogram (top) and time-frequency distributions (bottom).*

Hybridization requires to adjust the FDN model parameters (i.e. the reverberation profile and the initial frequency spectrum) to the time-frequency envelope of the original IR. This is achieved by analyzing the energy decay relief of the original IR.

2.3. Energy decay relief analysis

The energy decay relief (EDR) is the ensemble average of the time-frequency representation of the reverberation decay after the interruption of the excitation signal (see [35]). It represents the spectral energy density of the IR over time. The EDR is a generalization of Shroeder's energy decay curve (EDC), which allows for a time-frequency representation of the IR. It can be used to accurately estimate the model parameters of exponential reverberation decays. Given an impulse response $h(t)$, the EDR writes:

$$\mathrm{EDR}_h(t,f) = \left| \int_{\tau=t}^{\tau=\infty} h(\tau) e^{-j2\pi f \tau} d\tau \right|^2 . \quad (1)$$

Eq. (1) can be efficiently computed, e.g., through backward integration of the short-time Fourier spectrum of the impulse response. Following the procedure of [35], the reverberation time $\mathrm{RT}(f)$ can be estimated for any frequency f. Measured impulse responses are usually corrupted by measurement noise, which distorts the computed EDR and results in biased estimates of the decay times. In practice, the analysis of the EDR is restricted to a frequency-dependent time interval in which the hypothesis of exponential energy decay holds.

The absorptive filter g_i in the i^{th} feedback channel of the FDN is then chosen such that the logarithm of its magnitude response is proportional to the delay length and inversely proportional to the reverberation time. With reference to [4] and by neglecting the absorptive filter's phase response, the filter equation writes as

$$20 \cdot \log_{10} \left| g_i(e^{j2\pi f}) \right| = \frac{-60}{\mathrm{RT}(f)} \cdot \tau_i , \quad (2)$$

where τ_i is the delay length (in seconds) of the i^{th} inner channel. Spat~ implements the absorptive filter g_i as a three-band parametric filter with adjustable crossover frequencies. The estimated $\mathrm{RT}(f)$ is thus averaged and reduced to three frequency bands.

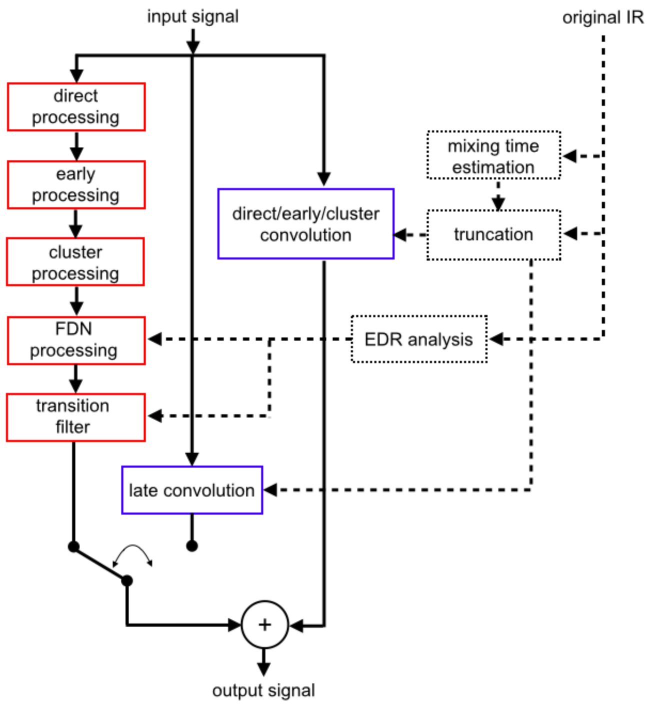

Figure 2: *Hybrid reverberator processing structure. Blocks with dashed-lines indicate offline processing. The "direct/early/cluster convolution" module performs the convolution of the IR truncated to the time interval* $[0 - t_{mix}]$.

2.4. Transition filter (spectral correction filter)

With reference to Jot et al. [4, 26], the EDR is not only characterized by the reverberation time $\mathrm{RT}(f)$ but also by the initial power spectrum $\mathrm{P}(f)$. In theory, the FDN's initial power spectrum is a zero-mean white Gaussian process that is independent of the decay characteristics. This assumption does not always hold in practice, due to approximations in the derivation of the FDN's correction filter [4] and the FDN-channel lowpass filters that simulate the air absorption of the reflection paths [36]. Note that the air absorption filters are not compensated by the FDN's correction filter.

An additional spectral correction is thus introduced to match the initial spectrum of the FDN with the EDR of the original IR at the mixing time:

$$\text{correction}(f) = \sqrt{\frac{\text{EDR}_h(t_{mix}, f)}{\text{EDR}_{\text{FDN}}(t_{mix}, f)}}. \tag{3}$$

A linear-phase filter is derived from the magnitude response of the spectral correction in Eq. (3) and then applied to the FDN (see Fig. 2). This spectral correction filter guarantees for a smooth and continuous time-frequency envelope of the hybrid reverberation processor at time $t = t_{mix}$.

2.5. Statistical aspects

The modal density D_m (i.e. the average number of modes per Hz) of a FDN with N feedback channels is related to the total length of the delay units by the following equation:

$$D_m = \sum_{k=1}^{N} \tau_k. \tag{4}$$

A crucial requirement for convincing artificial reverberation is to satisfy the assumptions of Schroeder's statistical model, i.e. to operate above the "Schroeder frequency". This condition corresponds to a modal overlap of at least 3:1 (see [35] for details), which is equivalent to

$$D_m \geq \text{RT}_0, \tag{5}$$

where RT_0 denotes the average reverberation time. Eqs. (4) and (5) are used to adjust the delay times of the inner loops of the FDN structure.

2.6. Current real-time implementation

The proposed hybrid reverberator is implemented in C++ and available as an external object (spat.hybrid$\sim$) for Max/MSP® as part of the Spat~ package. The external object first loads the IR and then performs the above-mentioned EDR analysis. This initial processing step is performed offline. Once all IR parameters are determined the external object enables the real-time processing of the input audio stream. To ease perceptive comparisons, one can switch between convolution and FDN modeling for the late reverberation tail (cf. Fig. 2). The real-time convolution is implemented as a zero-latency partitioned FFT algorithm adapted from [2]. The N uncorrelated output channels of the FDN can be either summed up to produce a mono output signal, or distributed over several loudspeakers creating a convincing spatial diffuse field out of a mono IR.

3. RESULTS

This section presents the results of a case study applying the above-mentioned algorithms to the IR of a large factory hall. In order to preserve a certain objectivity the IR was taken from a commercial library. The file is about 9 s long and the average decay time is $\text{RT}_0 \approx 4.5$ s. The mixing time is taken as $t_{mix} \approx 200$ ms. The FDN consists of $N = 8$ feedback channels and the crossover frequencies are set to 2.5 and 7 kHz.

Fig. 3 (bottom figure) demonstrates that the EDCs of the original and the hybrid IR are in good agreement; the upper figure compares the frequency-dependent reverberation time before and after applying the hybridization process. For frequencies higher than 1.5 kHz, the decay profile is in good agreement with that of the original IR. However, for lower frequencies an error of approximately $\pm 10\%$ can be observed. This error results from the use of 2^{nd}-order absorptive filters in the FDN loop, which determines the overall shape of the RT(f) curve. For the given example, the 2^{nd}-order shelving filters cannot approximate the original RT(f) curve with sufficient accuracy in the low and very high frequency bands.

The choice of 2^{nd}-order filters is motivated by the results of earlier studies on the perceptual characterization of the acoustic quality of concert halls, opera houses, and auditoria. Kahle [37] showed that controlling the reverberation time in three frequency bands covers the full range of perceptual attributes for a large set of room acoustic qualities. First informal listenings tests using the shown case study (among other examples) indirectly confirm these results. Despite the biased RT(f) depicted in Fig. 3, preliminary results indicate that listeners cannot distinguish sounds processed with the hybrid reverberator from those that have been convolved with the original IR. Anyway, more detailed listening experiments are needed to verify these early results.

The proposed hybrid reverberation processing model does not limit the number of absorption filter frequency bands. Higher-order parametric filters have been successfully implemented, but at the expense of a higher computational cost. The presented three-band filter model provides a good trade off between model accuracy and computational load.

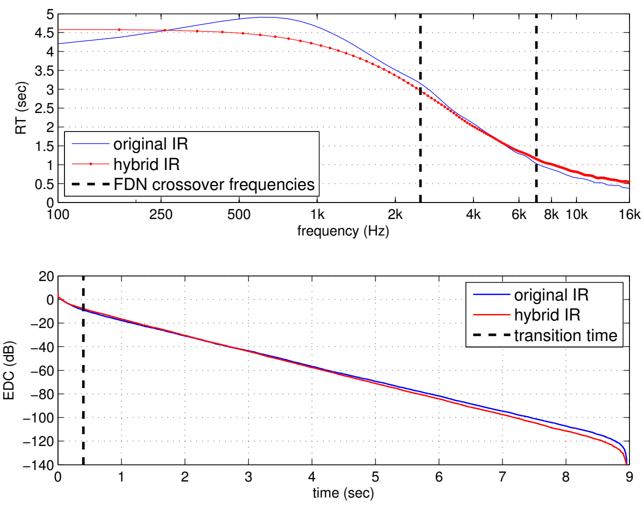

Figure 3: *Estimated reverberation time (top) and energy decay curve (bottom) for both the original and the hybrid IR.*

Fig. 4 (top figure) compares the EDR derived from the original IR with that from the FDN, both evaluated at the transition time, t_{mix}. The high frequency damping in the FDN spectrum (red dotted curve) results from the air absorption filters. Fig. 4 (bottom figure) depicts the magnitude response of the spectral correction filter that provides a smooth transition in between the convolution and the FDN. We perform a critical band smoothing before the magnitude response is transformed into a 256-taps FIR filter for real-time implementation. The actual length of the spectral correction filter is a tradeoff between the modeling accuracy and the computational efficiency.

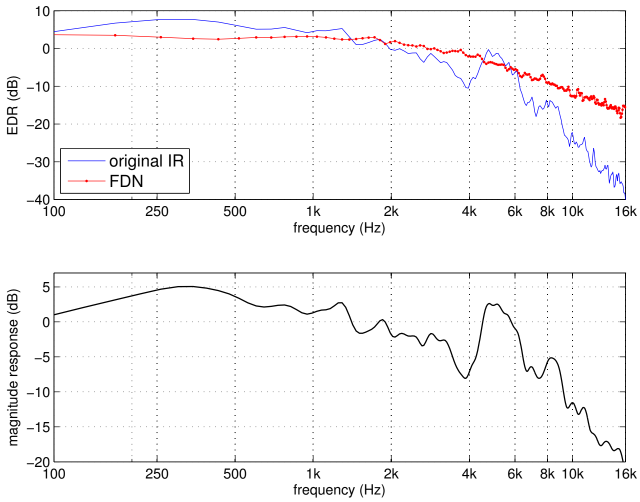

Figure 4: *EDR of the original IR and the FDN at the transition time (top) and the magnitude response of the transition filter (bottom).*

Fig. 5 illustrates the EDR of the original and the hybrid IR. Visual inspection confirms a good agreement between the EDRs. A closer look at the right figure shows that the reverberation time is slightly underestimated, i.e. the RT of the hybrid model is too short in a frequency range from 500 to 1500 Hz. The spectral correction filter shapes the FDN to match the spectrum of the original IR at the transition point. Without this correction filter, the hybrid reverberator would not reproduce the frequency boost around 5000 Hz. In this regard, the proposed processing method clearly outperforms standard FDN implementations.

Informal listening experiments confirm that the hybrid model generates perceptually indistinguishable results for various test signals (e.g. Dirac impulse, percussive sounds, male/female speech, music, etc.). As mentioned above, more detailed listening experiments are needed to verify these early results.

The computational advantage of the proposed hybrid reverberation processor over a pure convoler depends on many parameters, such as the length of the original IR (i.e. the reverberation time), the mixing time, the number of feedback channels in the FDN, the audio I/O latency of the processing environment, and so on. Therefore, it is difficult to draw general conclusions on the cpu load from comparisons with a pure convolution processor. With the different parameter settings that were tested in this case study, we gained about 35% of cpu load compared to an optimized real-time convolution algorithm.

4. PERCEPTUAL CONTROL

In the 1980s and 1990s, IRCAM has undertaken a series of room acoustic measurements and listening tests in different European concert halls. The aim was to establish a set of perceptual descriptors for the acoustic quality of concert halls (see e.g. [37–40]). Multidimensional data analysis (more specifically Individual Differences Scaling analysis, INDSCAL; see [41]), revealed a set of nine mutually independent perceptive descriptors for describing the room acoustic quality. It has been shown that these descriptors correlate well with some objective room acoustic criteria (see [38] for more details). In order to control the room effect along the relevant perceptual dimensions most of the proposed descriptors require both a temporal and spatial weighting; some of them do also require a spectral weighting in order to obtain satisfactory results. An in-depth discussion of the set of descriptors is beyond the scope of this paper. We only give one example to allow for a more general understanding of the perceptual control of the hybrid reverberator. For instance, the "DirE" descriptor refers to the energy of the "temporally extended" direct sound energy and controls the perceived presence of a sound source in a reverberant environment. It is computed from the temporally segmented impulse response as illustrated in Fig. 1. In the following, E_{R0} refers to the estimated energy of the direct sound ($0-20$ms), E_{R1} to the energy of the early reflections ($20-40$ms), E_{R2} to the energy of the cluster (i.e. the late reflections; $40-100$ms), and E_{R3} to the energy of the late reverberation tail (>100ms), respectively. DirE can be computed from these energy estimations as follows:

$$\text{DirE} = E_{R0} + E_{R1} + E_{R2,\text{excess}} + 0.18 \times E_{R2,\text{masked}} \tag{6}$$

with

$$\begin{aligned} E_{R2,\text{excess}} &= \max(0, E_{R2} - E_{R40}), \\ E_{R2,\text{masked}} &= \min(E_{R2}, E_{R40}), \\ E_{R40} &= E_R|_{[0,40\text{ms}]} = E_{R0} + E_{R1}. \end{aligned}$$

Jullien and Kahle [38–40] have, for instance, shown that the DirE parameter represents well Lochner and Burger's "energy ratio criterion" [42] for the intelligibility of speech. As a result one can control the perceived presence of sound source by controlling the gain of the different time sections of the impulse response. For more details on the perceptive descriptors, please refer to [37–40].

IRCAM's parametric FDN-based reverberator applies the perceptive descriptors in a similar way, and they have been proven useful in many music productions. If we now apply them to the hybrid reverberation process, we have to modify the signal processing structure given in Fig. 2 so that it represents the time-segmented structure given in Fig. 1. The convolution segment (i.e. for the time interval from $t = 0$ to $t = t_{mix}$) is split into three sub-segments corresponding to the sections "direct", "early", and "cluster". The output signals of these subsections are first time aligned with delay lines and then filtered with three-band parametric shelving filters, which are controlled by the perceptual model parameters. Fig. 6 depicts the extended processing model (for simplicity, the data analysis modules are not shown in this figure). When the direct/early/cluster/late filters are flat, the hybrid reverberation unit represents the original IR. When the user manipulates the perceptual factors, the parameters of the filters are updated accordingly. This allows to smoothly modulate the acoustical quality of the IR by navigating along the different perceptual axes. For instance the so-called "source presence" factor controls "DirE" and creates a convincing effect of proximity or remoteness of the sound source by simultaneously adjusting the direct/early/cluster/late levels (E_{R0}, E_{R1}, E_{R2}, E_{R3}) according to the structured model.

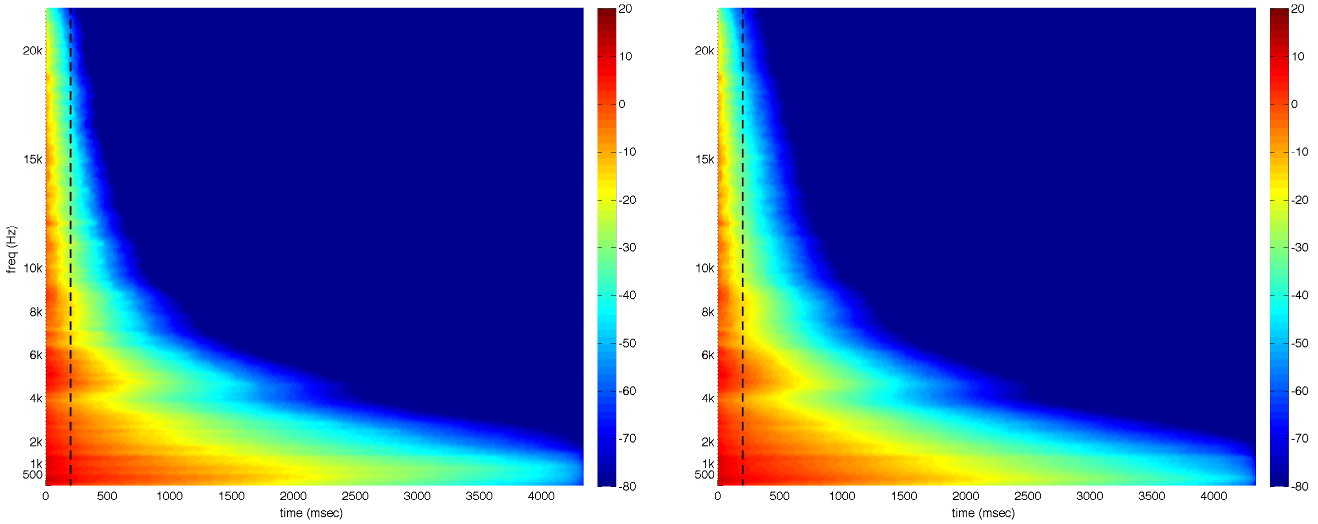

Figure 5: *EDR (in dB) of the original (left) and hybrid (right) IR. The dashed line represents the transition time,* $t = t_{mix}$.

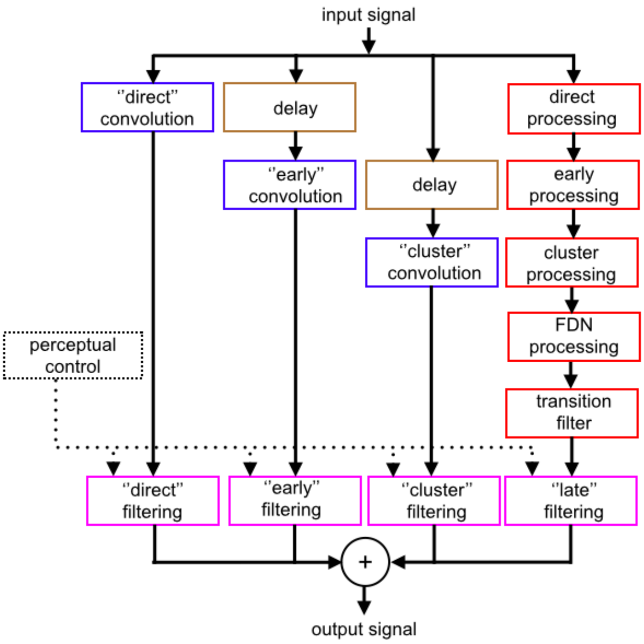

Figure 6: *Process chart of the hybrid reverberator with perceptual control. Blocks in blue correspond to convolution segments. Blocks in red correspond to parametric reverberation. Delay lines (brown blocks) ensure time-alignment of the convolution segments. Blocks in magenta correspond to three-band filters controlled by the perceptual model.*

5. CONCLUSIONS AND PERSPECTIVES

This paper considered both the theory and implementation of a hybrid reverberator that combines convolution processing for early reflections with feedback delay networks for late reverberation rendering. The proposed method first estimates the reverberation time and exponentially decaying envelope in different frequency bands from the original impulse response. These parameters are then used to control the FDN processing. Particular attention is paid to the smooth transition from convolution rendering to FDN processing; the power spectrum is matched at the transition point (given by the mixing time) in each frequency band.

An analysis of different room impulse responses (see also Sections 2 and 3) indicated that three-band shelving filters in each FDN channel may not always succeed to model the late reverberation with sufficient accuracy. The model accuracy strongly depends on the EDR profile of the original IR. Future work will focus on the analysis of the RT(f) curve in order to automatically determine the minimum number of required frequency bands (and corresponding crossover frequencies) to keep the modeling error below a given threshold. Increasing the number of FDN filter bands significantly increases the computational cost. However, with the rapid increase in available processing power real-time implementations may become feasible.

The modal density of a FDN should satisfy Schroeder's suggestions for natural sounding and high quality artificial reverberators (cf. Section 2). A useful extension of the hybrid processor would be estimating the modal density of the original IR to automatically adjust the FDN to these parameters.

The proposed method is based on the stochastic model of late reverberation and thus excludes, e.g., non-exponential decays, flutter echoes, and spring reverbs. Nonetheless it would be possible to extend the technique to IRs exhibiting a double-slope exponential decay; such decay profiles have gained interest in recent years and have been observed in concert halls such as, e.g., the Boston Symphony Hall. Both the EDR analysis and the FDN rendering can be extended to that purpose. Adapting the EDR analysis of [35] to multiple-slope exponential decays do not raise conceptual difficulties. The design of FDNs with multiple decay slopes is currently under investigation.

In this paper we focused on single channel impulse responses for a mono input signal and mono (or multichannel) output signal(s). A multichannel extension to directional room impulse responses (DRIRs) is currently under development. DRIRs are typically measured with spherical microphone arrays. Preliminary results of multichannel EDR analysis and DRIR denoising have been published in [43, 44]. The proposed methods perform a joint analysis of the EDR of all the microphone cells in order to preserve the spatial coherence between them. Hybrid convolution reverberators operating in the modal domain are used for higher-order Ambisonics rendering.

6. ACKNOWLEDGMENTS

This study was supported in part by the French ANR CONTINT project "Sample Orchestrator 2" (SOR2).

7. REFERENCES

[1] V. Välimäki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel, "Fifty years of artificial reverberation," *IEEE Audio, Speech, and Language Processing*, vol. 20, no. 5, pp. 1421 – 1448, July 2012.

[2] W. G. Gardner, "Efficient convolution without input-output delay," in *Proc. of the 97th AES Convention*, San Francisco, CA, Nov. 10-13, 1994.

[3] G. García, "Optimal filter partition for efficient convolution with short input/output delay," in *Proc. of the 113rd AES Convention*, Los Angeles, CA, Oct. 5-8, 2002.

[4] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *Proc. of the 90th AES Convention*, Paris, France, Feb. 19-22, 1991.

[5] M. R. Schroeder, "Digital simulation of sound transmission in reverberant spaces," *J. Acoust. Soc. Amer.*, vol. 45, no. 1, pp. 424 – 431, 1969.

[6] J. A. Moorer, "About this reverberation business," *Computer Music Journal*, vol. 3, no. 2, pp. 13 – 28, 1979.

[7] F. R. Moore, *Elements of Computer Music*, Prentice Hall, Englewood Cliffs, NJ, USA, 1990.

[8] R. Stewart and D. Murphy, "A hybrid artificial reverberation algorithm," in *Proc. 122nd AES Convention*, Vienna, Austria, May 5-8, 2007.

[9] R. Stewart, "Hybrid convolution and filterbank artificial reverberation algorithm using statistical analysis and synthesis," M.S. thesis, University of York, 2006.

[10] R. Stewart and M. Sandler, "Real-time panning convolution reverberation," in *Proc. of the 123rd AES Convention*, New York, NY, Oct. 5-8, 2007.

[11] J. S. Abel, D. P. Berners, and A. Greenblatt, "An emulation of the emt140 plate reverberator using a hybrid reverberator structure," in *Proc. of the 127th AES Convention*, New York, NY, Oct. 9-12, 2009.

[12] A. B. Greenblatt, J. S. Abel, and D. P. Berners, "A hybrid reverberation crossfading technique," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, Dallas, TX, Mar. 14-19, 2010, pp. 429 – 432.

[13] K. Sup Lee, N. J. Bryan, and J. S. Abel, "Approximating measured reverberation using a hybrid fixed/switched convolution structure," in *Proc. of the 13th Int. Conference on Digital Audio Effects (DAFx-10)*, Graz, Austria, Sept. 6-10, 2010.

[14] S. Heise, M. Hlatky, and J. Loviscach, "Automatic adjustment of off-the-shelf reverberation effects," in *Proc. of the 126th AES Convention*, Munich, May 7-10, 2009.

[15] A. Primavera, L. Palestini, S. Cecchi, F. Piazza, and M. Moschetti, "A hybrid approach for real-time room acoustic response simulation," in *Presented at the 128th AES Convention*, London, May 22-25, 2010.

[16] A. Primavera, S. Cecchi, L. Romoli, P. Peretti, and F. Piazza, "An advanced implementation of a digital artificial reveberator," in *Proc. of the 130th AES Convention*, London, May 13-16, 2011.

[17] A. Primavera, S. Cecchi, L. Romoli, P. Peretti, and F. Piazza, "Approximation of real impulse response using IIR structures," in *Proc. of the 19th European Signal Processing Conference*, Aug. 2011.

[18] A. Primavera, M. Gasparini, S. Cecchi, L. Romoli, and F. Piazza, "Hybrid Reverberation Algorithm: a Practical Approach," in *Proc. of the AIA-DAGA Conference*, Merano, Italy, Mar. 18-21, 2013.

[19] G. Constantini and A. Uncini, "Real-time room acoustic response simulation by an iir adaptive filter," *Electronics Letters*, vol. 39, pp. 330 – 332, 2003.

[20] B. Holm-Rasmussen, H.-M. Lehtonen, and V. Välimäki, "A new reverberator based on variable sparsity convolution," in *Proc. of the 16th Int. Conference on Digital Audio Effects (DAFx-13)*, Maynooth, 2013.

[21] J. C. Sarris and G. E. Cambourakis, "Time frequency analysis and parametric approximation of room impulse responses," in *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, 2003, vol. 6, pp. 297 – 300.

[22] M. Schonle, N. Fliege, and U. Zolzer, "Parametric approximation of room impulse responses based on wavelet decomposition," in *Applications of Signal Processing to Audio and Acoustics, 1993. Final Program and Paper Summaries., 1993 IEEE Workshop on*, 1993, pp. 68 – 71.

[23] M. Schoenle, N. Fliege, and U. Zoelzer, "Parametric approximation of room impulse responses by multirate systems," in *Acoustics, Speech, and Signal Processing, 1993. ICASSP-93., 1993 IEEE International Conference on*, 1993, vol. 1, pp. 153 – 156.

[24] M. R. Schroeder, "Eigenfrequenzstatistik und Anregungsstatistik in räumen," *Acustica*, , no. 4, pp. 456 – 468, 1954, English translation: M. R. Schroeder, "Statistical Parameters of the Frequency Response Curves of Large Rooms", J. Audio Eng. Society, vol. 35, no. 5, pp. 299 - 306, 1987.

[25] M. R. Schroeder and H. Kuttruff, "On frequency response curves in rooms. comparison of experimental, theoretical and monte-carlo results for the average frequency spacing between maxima," *J. Acoust. Soc. Am.*, , no. 34, pp. 76 – 80, 1962.

[26] J.-D. Polack, "Playing billiards in the concert hall: The mathematical foundations of geometrical room acoustics," *Applied Acoustics*, vol. 38, no. 2–4, pp. 235–244, 1993.

[27] J. S. Abel and P. Huang, "A simple, robust measure of reverberation echo density," in *Proc. of the 121st AES Convention*, San Francisco, CA, Oct. 5-8, 2006.

[28] T. Hidaka, Y. Yamada, and T. Nakagawa, "A new definition of boundary point between early reflections and late reverberation in room impulse responses," *J. Acoust. Soc. Am*, vol. 122, no. 1, pp. 326 – 332, July 2007.

[29] P. Huang and J. S. Abel, "Aspects of reverberation echo density," in *Proc. of the 123rd AES Convention*, New York, NY, Oct. 5-8, 2007.

[30] R. Stewart and M. Sandler, "Statistical measures of early reflections of room impulse responses," in *Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*, Bordeaux, France, Sept. 10-15, 2007.

[31] G. Defrance, L. Daudet, and J.-D. Polack, "Using matching pursuit for estimating mixing time within room impulse responses," *Acta Acustica united with Acustica*, vol. 95, no. 6, 2009.

[32] G. Defrance and J.-D. Polack, "Estimating the mixing time of concert halls using the eXtensible Fourier Transform," *Applied Acoustics*, vol. 71, no. 9, pp. 777–792, 2010.

[33] A. Lindau, L. Kosanke, and S. Weinzierl, "Perceptual evaluation of model- and signal-based predictors of the mixing time in binaural room impulse responses," *J. Audio Eng. Soc.*, vol. 60, no. 11, pp. 878 – 898, 2012.

[34] J.-M. Jot, "Real-time spatial processing of sounds for music, multimedia and interactive human-computer interfaces," *ACM Multimedia Systems Journal (Special issue on Audio and Multimedia)*, vol. 7, no. 1, pp. 55 – 69, 1997.

[35] J.-M. Jot, L. Cerveau, and O. Warusfel, "Analysis and synthesis of room reverberation based on a statistical time-frequency model," in *Proc. 103rd AES Convention*, New York, NY, Sept. 26-29, 1997.

[36] J. Huopaniemi, L. Savioja, and M. Karjalainen, "Modeling of reflections and air absorption in acoustical spaces – a digital filter design approach," in *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct 1997.

[37] E. Kahle, *Validation d'un modèle objectif de la perception de la qualité acoustique dans un ensemble de salles de concerts et d'opéras*, Ph.D. thesis, Laboratoire d'Acoustique de l'Université du Maine, Le Mans, 1995.

[38] J.-P. Jullien, "Structured model for the representation and the control of room acoustic quality," in *15th International Congress on Acoustics (ICA)*, Trondheim, Norway, 1995, pp. 517–520.

[39] E. Kahle and J.-P. Jullien, "Some new considerations on the subjective impression of reverberance and its correlation with objective criteria," in *Acoustical Society of America, 127th annual Meeting (Sabine Centennial Symposium)*, Cambridge, MA, USA, June 1994, pp. 239–242.

[40] E. Kahle and J.-P. Jullien, "Subjective listening tests in concert halls: Methodology and results," in *Proc. of the Int. Congr. on Acoustics*, Trondheim, Norway, June 1995, pp. 521 – 524.

[41] J. D. Carroll and J. J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of 'eckart-young' decomposition.," *Psychometrika*, vol. 35, pp. 283 – 319, 1970.

[42] J. P. A. Lochner and J. F. Burger, "The subjective masking of short time delayed echoes by the primary sounds and their contribution to the intelligibility of speech.," *Acustica*, vol. 8, pp. 1 – 10, 1958.

[43] T. Carpentier, T. Szpruch, M. Noisternig, and O. Warusfel, "Parametric control of convolution-based room simulators," in *Proc. of the Int. Symposium on Room Acoustics*, Toronto, Canada, June 9-11, 2013.

[44] M. Noisternig, T. Carpentier, T. Szpruch, and O. Warusfel, "Denoising of directional room impulse responses measured with spherical microphone arrays," in *Proc. of the 38th Annual Convention of Acoustics (DAGA)*, Oldenburg, Germany, March 10-13, 2014.