Reverberation Modeling Using Velvet Noise

Matti Karjalainen and Hanna Järveläinen

[1] *Helsinki University of Technology, Lab. of Acoustics and Audio Signal Processing, P.O.Box 3000, FI-02015 TKK, Finland*

Correspondence should be addressed to Matti Karjalainen (`matti.karjalainen@tkk.fi`)

ABSTRACT

Room reverberation consists of a multitude of reflections from surfaces and objects in a room. Particularly the late reverberation tail resembles noise with an exponential decay envelope. Artificial reverberation algorithms try to simulate this in a computationally efficient manner. Some proposed algorithms are based on the convolution with a sparse FIR filter corresponding to a randomized sparse sequence of unit impulses. In this paper we search for such sequences with minimal impulse density vs. maximal smoothness of the noise-like characteristics. Such noise is called here "velvet noise", because it can sound smoother than the Gaussian noise. The perceptual characteristics of velvet noise are described by results from listening experiments and auditory analysis. Reverb algorithms based on velvet noise are discussed and analyzed.

1. INTRODUCTION

The physical principles of room reverberation are known well and can be measured or modeled computationally fairly accurately. From a frequency domain point of view, room responses are composed of eigenmodes, i.e., decaying sinusoids with modal density N'_{mode} (modes per Hz) increasing as a function of frequency f by $N'_{\text{mode}}(f) = 4\pi V f^2/c^3$, where V is the volume of the room and c the velocity of sound [1]. Below a critical frequency the modes are increasingly sparse and may be perceived as separate resonant components if they are strong enough.

In the time domain the same phenomena can be understood as a series of reflections from room surfaces with increasing reflection density N'_{refl} as a function of time t by $N'_{\text{refl}}(t) = 4\pi c^3 t^2/V$. After early reflections the room impulse response is approaching an exponentially decaying noise signal with spectrum changing in time.

Generation of artificial reverberation [2] can be based on principles that simulate the frequency or time domain behavior or their mixture. Most reverb algorithms can be considered as special digital filters that utilize primarily the temporal decomposition to delayed impulse-like components. Classical examples thereof span from the Schroeder and Moorer reverbs to feedback delay networks (FDNs) [3–6], which can be computationally highly efficient but demanding to mach to given room responses. Modal decomposition in the frequency domain for reverberation modeling may bring design flexibility but with increased computational load, see e.g. [7]. The same is true with time-domain processing such as convolution directly by FIR filtering or through FFT.

In this paper we study the possibility of efficient reverberation modeling by sparse FIR filtering, optimized through auditory criteria instead of physical facts. The late reverberation composed of randomly positioned reflections can be roughly approximated by a sequence of impulses weighted by an exponential decay curve. While full FIR filtering requires a filter order of $N = f_s T$, where f_s is the sample rate and T is the length of the needed response, a sparse FIR can be much more efficient when filter taps can be positioned sparsely enough.

The impulse density required for perceptually good late reverberation has been studied for example by Schreiber [8]. He found that an impulse sequence of position-randomized Dirac impulses requires an impulse density that is roughly equal to the applied bandwidth, up to about 2 kHz bandwidth, in order to sound smooth without roughness. For wider bandwidths than 2 kHz it was found that the impulse density required remains in about 2000 spikes per second.

In a recent study Rubak and Johansen have developed the idea of randomized impulse train with exponentially decaying envelope into a practical reverberation algorithm [9, 10]. They ended up with impulse densities of 2000 to 4000 spikes per second for high-quality noise and reverberation. They propose a recursive filter structure built from a sparse FIR for pseudo-random processes and a low-pass filter in the feedback loop to control the frequency-dependent decay profile of late reverberation.

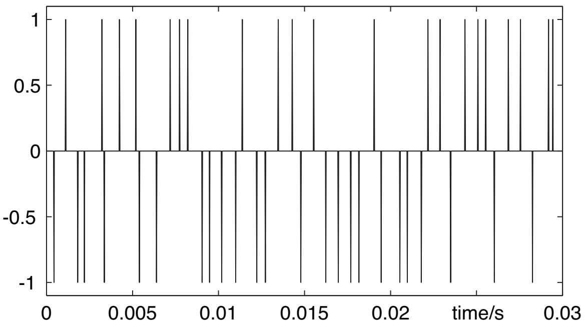

Fig. 1: A 30 ms sequence of velvet noise with impulse density of 1500 spikes per second.

The present paper follows similar guidelines by investigating further the perceptual requirements of pseudorandom impulse sequences and their applicability in reverberation modeling. It has been found that certain sparse impulse sequences are perceived as noise and late reverberation even smoother than those based on corresponding Gaussian noise. Therefore we call them "velvet noise". The reason to this perceptual behavior is the smooth auditory envelope of such noises. Velvet noise has somewhat similar properties as the so-called low-noise noise [11–13], although the generation mechanism is very different. For basic research literature on the perception of temporal properties of sounds, see [14, Chapter 5] and [15, Chapter 6].

Sections 2-4 of this paper investigate the basic principles and properties of velvet noise. Section 5 presents artificial reverberation modeling based on this principle, and discussion in Section 7 concludes the paper.

2. VELVET NOISE

In order to improve the spectral flatness and to remove the DC component, the Dirac sequence of [8] can be replaced by a similar sequence but with randomly varying sign, as was also proposed in [9]. With impulse density above 1500 spikes per second it has best smoothness, with even less roughness than with Gaussian white noise. In the subsections below we investigate the perceptual properties of such velvet noise.

2.1. Generation of velvet noise

The velvet noise $n(k)$ of this study is constructed as a sum of position- and sign-randomized unit impulses $u(k)$ as:

$$n(k) = \sum_{m=0}^{M} a(m)u(k - \mathrm{round}(\frac{T_\mathrm{d}}{T_\mathrm{s}}(m + \mathrm{rnd}(m)))) \quad (1)$$

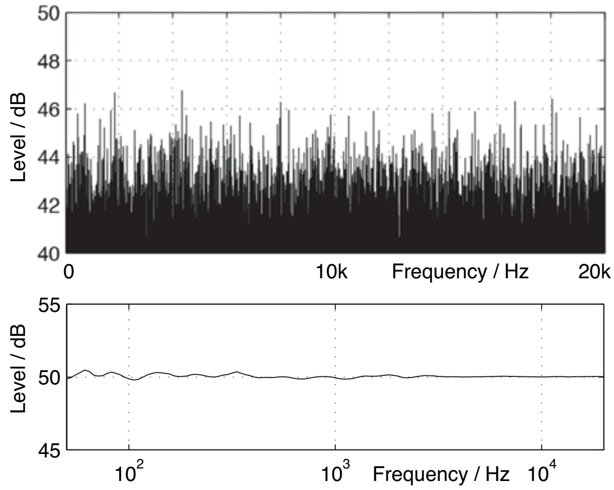

Fig. 2: Example spectrum of a 2 seconds sequence of velvet noise with impulse density of 1500 spikes per second: (top) linear frequency scale, (bottom) 1/3-octave smoothed spectrum. For longer sequences the ripple of smoothed spectrum at low frequencies is reduced.

where 'rnd(m)' is a random number of uniform distribution between $0 \ldots 1$, 'round' is rounding of a real number to integer, $a(m)$ is randomly varying sign value of amplitude $\{-1, +1\}$, T_d is the average time interval between impulses (the inverse of impulse density), and T_s is the sampling rate[1]. We have tried also with high sampling rates for improved time resolution of impulse positioning, but it was found that a typical audio rate, such as 44100 Hz, is high enough.

Figure 1 shows a 30 ms impulse sequence with average density of 1500 spikes per second. Figure 2 depicts the logarithmic spectrum of a 2 seconds long sequence and its 1/3-octave smoothed version. The spectrum is white and approaches flatness as the length of the sequence increases. The amplitude distribution is discrete, consisting of values $\{-1, 0, +1\}$ only.

3. LISTENING EXPERIMENTS

Perception of velvet noise was studied in a listening experiment in comparison to Gaussian white noise. The objective was to find out how its smoothness decreases and turns into roughness with decreasing average impulse density.

[1]This is similar to [9] but may differ in fine detail.

Table 1: Test conditions in the listening experiment.

Condition	Velvet noise	RASI noise
Broadband	x	x
Lowpass at 1.5 kHz	x	x
1 ERB, f_c = 1.5 kHz	x	
1 octave, f_c = 1.5	x	
3 octaves, f_c = 1.5 kHz	x	

Table 2: Smoothness grading scale.

Grade	Value (range)
Very rough	-3 (-3.5 ... -2.5)
Rough	-2 (-2.5 ... -1.5)
Slightly rough	-1 (-1.5 ... -0.5)
Equal	0 (-0.5 ... 0.5)
Slightly smooth	1 (0.5 ... 1.5)
Smooth	2 (1.5 ... 2.5)
Very smooth	3 (2.5 ... 3.5)

random-amplitude sparse impulse (RASI) noise, similar to the velvet noise but whose pulse amplitudes were normally distributed with mean value = 0 and standard deviation = 1. RMS values of the stimuli were then equalized for listening. Average pulse densities were 600, 800, 1000, 1500, 3000, 7000, and 20000 1/s. Signal duration was 2 seconds. Broadband (100 ... 20000 Hz), lowpass (100 ... 1500 Hz), and 1-ERB [16], 1-octave and 3-octave bandpass velvet noise, centered around 1.5 kHz, were tested. RASI noise was tested only in the broadband and lowpass conditions. The test conditions are summarized in Table 1. The reference signal was always Gaussian white noise.

The experiment was made in a silent room using the GuineaPig3 test software [17]. Signals were played back through headphones. Subjects controlled the playback of stimulus A and reference Ref freely through the GuineaPig graphical user interface. The task was to judge, how much rougher or smoother the noise stimulus sounded compared to the reference. A 7-step grading scale was used, as specified in Table 2. The grade was given by moving the grade bar with the mouse, and it was recorded with precision of one decimal. Each trial was repeated twice, which resulted in a total of 102 trials and average duration of 18 minutes. Playback order of the trials was randomized.

Nine subjects participated in the experiment. They were staff of the laboratory, had previous experience in psychoacoustic testing, and were between 23 and 32 years of age. Before the test run each subject was trained by playing a few examples of all test conditions. The subjects were adviced to concentrate on irregularities in the signal envelope.

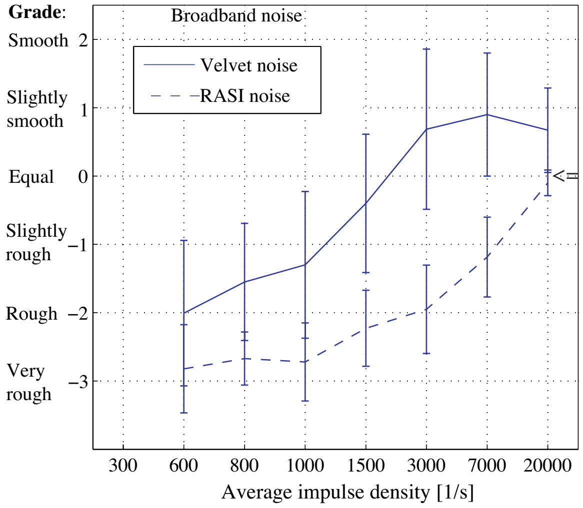

Fig. 3: Test results for broadband velvet and RASI noise. The error bars give one standard deviation up and down.

The results of two subjects were excluded from final analysis because of poor consistency. Correlation coefficients were calculated for each subject between the repeated trials, and only those subjects, whose mean correlation was above 0.5, were included in the analysis. The average correlation of the seven remaining subjects was 0.69.

Figure 3 shows the results for broadband velvet noise and RASI noise. At low pulse densities both sound rougher than Gaussian white noise. As the pulse density increases, velvet noise is perceived even smoother than Gaussian white noise, while RASI noise approaches the smoothness of Gaussian white noise monotonously. The difference between the grades for velvet and RASI noise is significant for all pulse densities at $\alpha = 0.05$ level (one-way ANOVA).

From Fig. 3 it can be concluded that for impulse densities higher than about 2000 per second the velvet noise is perceived smoother than the Gaussian white noise, thus motivating the term "velvet noise" as a specific form of "low-noise noise". At all frequencies it is perceived smoother than the RASI noise. Based also on subjective impression it can be concluded that velvet noise sequences are useful as a basis of late reverberation at least

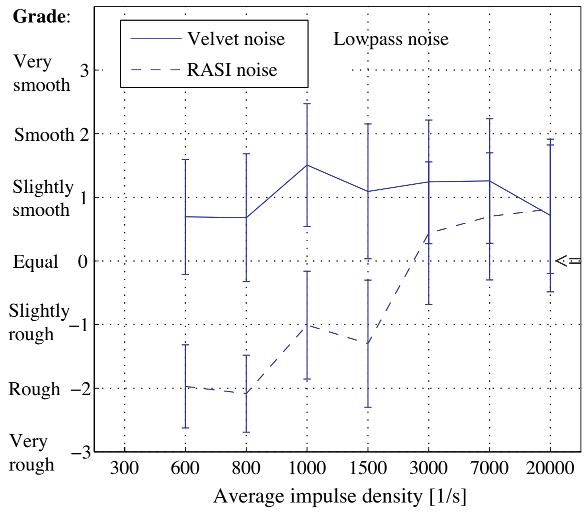

Fig. 4: Test results for lowpass velvet and RASI noise. The error bars give one standard deviation up and down.

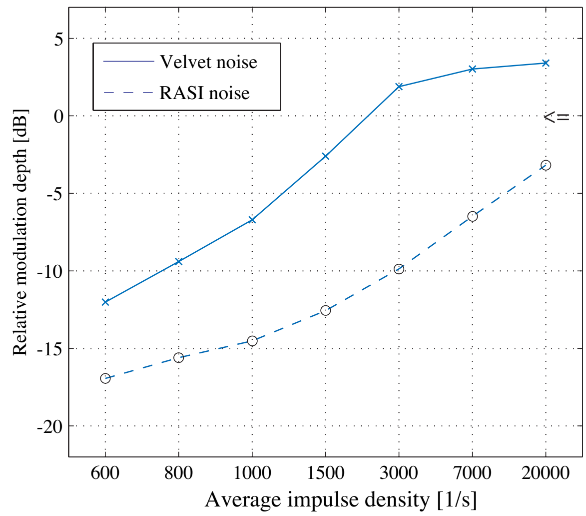

Fig. 5: Auditory simulation of the smoothness of noise signals as a function of the average impulse density for broadband velvet and RASI noise. The reference (0 dB level) is a Gaussian white noise. Compare with Fig. 3.

down to about 1500 impulses per second. Below 1000 imp/sec the roughness of the noise starts to sound disturbing and individual spikes increasingly stand out from the noise.

The results from a similar listening experiment with 1.5 kHz lowpass-filtered noises show somewhat similar behavior, see Fig. 4. However, velvet noise is perceived smoother than full-band Gaussian white noise at all pulse densities, while RASI noise only reaches the smoothness of broad-band Gaussian white noise (reference) at about 3000 impulses/second and beyond. It can be noticed also that the smoothness of velvet noise remains down to very low impulse rates, which is characteristic to narrow bandwidths in general. In a bandpass experiment it was noticed that for an octave band around 1.5 kHz the roughness increases around 800 impulses/second, for a 1/3-octave band at about 500 impulses/second, and for an ERB bandwidth at 300 impulses/second. From a reverberation point of view this means, when splitting the processing into subbands, that each subband can use lower impulse rate than a single wideband processing requires.

4. AUDITORY ANALYSIS

From studies with so-called low-noise noise [11] it is obvious that the flatness of temporal envelope of a signal, when analyzed in the auditory critical bands, makes the velvet noise perceived smoother than Gaussian white noise. In this section we analyze how a simple auditory simulation of different noises correlates with the results of the listening experiments above.

The envelope fluctuation due to the inherent modulation in a noise can be computed by (full-wave) rectifying the noise signal and by low-pass filtering to correspond to the temporal resolution of the auditory system. The modulation depth is obtained as the ratio of variance and mean value of the envelope signal, and this ratio is further expressed in decibels. Figure 5 shows the difference in modulation depth of velvet noise vs. Gaussian noise and RASI noise vs. Gaussian noise as functions of impulse density. The curves are plotted so that increasing modulation (roughness) yields lower dB-ratings and less modulation (smoothness) means higher dB-values. Lowpass filtering of the envelope was done with a 4th-order Butterworth filter having a cutoff frequency at 350 Hz.

When comparing the curves in Fig. 5 with the corresponding curves in Fig. 3, the similarity of shapes between them can be noticed. One scale step in roughness rating in Fig. 3 corresponds to about 5 dB of change in the modulation depth. Similarity between listening test results and simulation patterns was found also for the lowpass noises, although absolute level shifting of the sparse noises was needed in relation to the Gaussian reference noise.

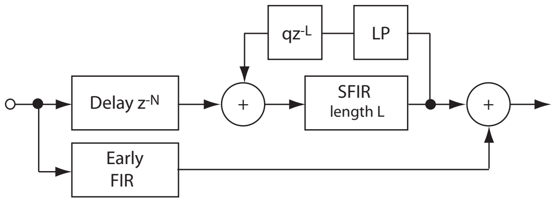

Fig. 6: Basic structure of the sparse FIR reverberator proposed in [9, 10] consisting of an early response part (left) and a recursive late reverberation part (right).

5. REVERBERATION MODELING

Artificial reverberation can be made by taking a suitable length sequence of velvet noise and temporally weighting it by an exponential decay envelope. According to the results of perceptual studies above, approximately 1500-2000 impulses are needed per second for high-quality full-band noise-like late reverberation. This means for example a $2 \cdot 1.5 \cdot 1500 = 4500$ tap sparse FIR (SFIR) for a 2-second T_{60} concert hall impulse response with a 90 dB dynamic range. While this is only about 3 % of a full-tap FIR filter, it remains still time-consuming to compute.

5.1. Recursive SFIR structures

A way to reduce computational load is to use a shorter SFIR with a feedback loop as proposed in [9, 10]. The early part of a room response can be implemented as a separate SFIR, to which the late recursive part is concatenated as shown in Fig. 6. In this section we will discuss methods to optimize the computation of recursive SFIR structures for high-quality late reverberation.

The rules of designing the recursive SFIR are as follows. The desired decay envelope $e(t)$ is obtained from reverberation time T_{60} as

$$e(t) = 10^{-3t/T_{60}} \tag{2}$$

which is used as a weighting function to the sparse tap coefficients of the SFIR. The feedback gain q in Fig. 6 (first assumed to be a real-valued constant) needs to be set to

$$q = 10^{-3T_{\text{loop}}/T_{60}} \tag{3}$$

in order to get a smooth continuation of the decay envelope due to the recursive structure.

The recursive SFIR loop delay cannot be shortened below about 0.2 seconds without perceiving disturbing periodicity in the reverberation [9, 10]. The periodicity appears in two forms: (a) temporally as jumps and clicks if T_{60} is frequency-dependent so that q in Eq. (3) is a digital filter and (b) as spectral coloration because the recursive SFIR is basically a comb filter.

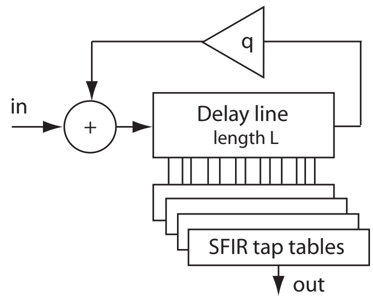

Fig. 7: Time-varying reverberation using tap parameter table switching and interpolation in a velvet noise SFIR.

With many reverb algorithms, time variance of filter parameters is applied to reduce perceivable periodicities. We now propose such a solution to the recursive SFIR reverberator. Figure 7 shows a time-varying recursive SFIR, where the time-variance is realized by switching the sparse FIR tap position and amplitude tables once for each loop-around period of the delay-line data. The recursive SFIR tap coefficients can be precomputed as independent randomized coefficients into so many tables that looping around the tables does not introduce perceivable periodicity.

A new problem appears, since due to the abrupt switching of filter coefficients there will be highly disturbing discontinuities in the filter output signal at switching points. This can be removed by interpolation between subsequent tap parameter tables. Because the sequences to be interpolated are incoherent noise, the interpolation must be done linearly in power, which means using square root law in amplitude. Figure 8 shows the interpolation functions of two subsequent tap amplitude tables.

5.2. Realization of frequency-dependent decay

The next question is how to make the decay rate frequency-dependent, as happens in real rooms, without artifacts in sound. There are basically two alternatives:

- to divide the audio range to subbands that are processed separately and finally to add the bands, or
- to use a lowpass filter in cascade with coefficient k in the feedback loop as shown in Fig. 6.

The feedback filter method has been used in [9, 10]. We found it quite problematic as such for short SFIR

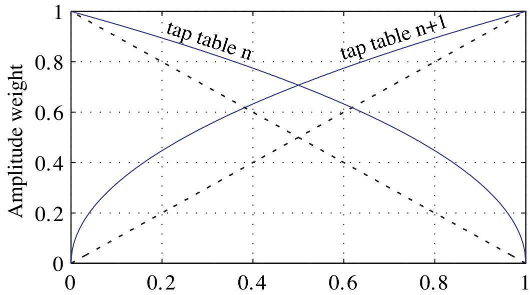

Fig. 8: Square-root law for interpolation of two subsequent SFIR tap-table amplitudes. Linear amplitude interpolation is plotted by dotted curves.

loops due to audible discontinuities (clicks and jumps in waveform). The time-varying tap-vector interpolation removes this quite efficiently, however. Therefore we can go down to 100 ms and even below, depending on the signal to be reverberated, without this problem appearing, but the spectral coloration remains, as will be discussed later below.

Another method to realize frequency-dependent T_{60} is to use subband processing. Octave or 1/3-octave bands can be applied, or just 1-2 separate bands for high frequencies, where the reverberation time is remarkably shorter than for low and mid frequencies. Subband processing naturally adds complexity and processing latency. Multirate processing is an attractive choice to reduce computational load, but at the expense of overall complexity. One advantage to be taken into account is that for narrower bands the impulse rate of the velvet noise can be reduced. For example for a 1.5 kHz lowpass-filtered signal it can be down to about 500-800 spikes per second.

6. MODELING EXAMPLE

To test the velvet noise reverberation method we decided to apply it to the simulation of a measured concert hall impulse response. Data was taken from [18] where high-quality impulse responses are available for a concert hall in Pori, Finland. The specific case chosen is the response 'S1_R3o.m', measured from a stage position to the floor quite far from the stage (see [18] for documentation).

Figure 9(a) plots the first 100 ms of the measured impulse response where the direct sound, a strong reflection around 17 ms, and a cluster of reflections starting from about 47 ms are clearly visible. The rest of the impulse response after about 70 ms consists of temporally diffuse late reverberation.

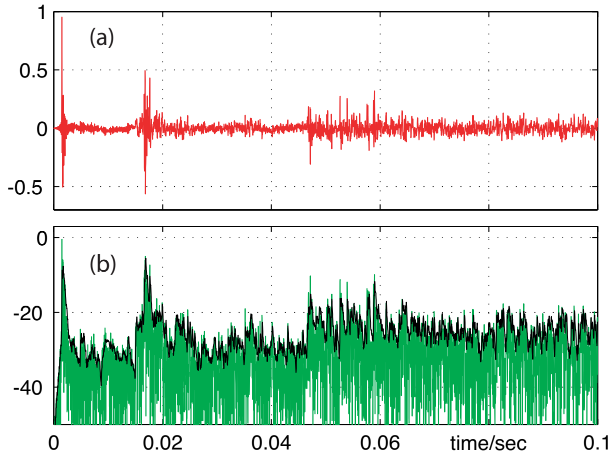

Fig. 9: First 100 ms of measured impulse response of concert hall: (a) time waveform, (b) dB-scaled envelope (green) and 1 kHz lowpass smoothed envelope (black).

6.1. Direct sound modeling

The measured direct sound is very short (1-2 ms) and somewhat colored spectrally, which might be more due to the measurement equipment than by physics of sound propagation. Accurate modeling of the direct sound is important from a perceptual viewpoint, particularly for binaural and multichannel reproduction. Therefore we applied AR-modeling (linear prediction) to it in order to get a good spectral match to the original one. Particularly warped linear prediction [19] and Kautz filter modeling [20–22] are efficient means to obtain very low order (2-8) models, yet perceptually good for the direct sound as well as for isolated reflections.

6.2. Early response modeling

In this particular case the early sound, starting after the direct sound and up to about 100 ms, was modeled as an SFIR, with impulse density of 1500 per second and temporally weighted by the smoothed amplitude envelope shown by black curve in Fig. 9(b). This automatically brings approximately correct temporal properties to the early response. Additionally, the response is spectrally shaped by a cascaded low-order IIR filter (order of 4 was enough in this case), obtained by AR modeling of this part of the response. This does not provide a frequency-dependent decay within the early response, which however is not very remarkable during the first 100 ms.

Table 3: Late reverberation time $T60$ of the hall for octave bands.

f/Hz	125	250	500	1k	2k	4k	8k
$T60$/s	2.7	2.5	2.4	2.3	2.1	1.7	1.1

A more important problem in this simple approach is that the randomness of impulses in fitting the temporal envelope makes timbral deviations from the target timbre. We simply repeated the design many times with different randomizations and selected the one that sounded closest to the original.

6.3. Late response modeling

The late part after 100 ms in the measured response was modeled using the recursive SFIR structure. The input to this block was taken from the same AR-modeled IIR filter as for the early response (see Fig. 10). This was done to guarantee maximal continuity after the end of the early part. The task was to design the late reverberation parameters so that the recursive SFIR used is computationally efficient, yet avoiding disturbing artifacts in sound.

The T_{60} reverberation time for the hall in octave bands is as shown in Table 3. Equation (3) was applied to this data to compute the attenuation factor as a function of frequency and to design a loop IIR filter (LP in Fig. 6) to approximate this behavior. A filter of order 4 was found good enough, designed by the function `yulewalk` in Matlab. The main remaining parameter to experiment with was the loop length of recursive SFIR filter in order to find out how short it can be made without artifacts. This will be discussed in the next subsection.

6.4. Perceptual evaluation

The concert hall reverberation model with different parameter values was evaluated by informal listening to typical audio signals convolved with the model. We listened to the impulse response as such and response to transient sounds in music signals, speech, and continuous sine waveforms.

The impulse response and transient sound responses behave typically well with a carefully calibrated model. The length of recursive SFIR loop can be shortened down to 50-100 ms before temporal repetition appears. Particularly for speech signals this is no problem; in fact, speech allows for the velvet noise impulse density to be lowered even down to 500-1000 spikes per second. Spectral coloration, however, starts to be a problem for short delay loops.

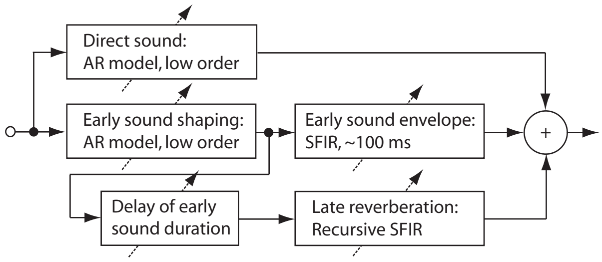

Fig. 10: Block diagram of the concert hall simulator implementation.

Continuous sine waves are the most problematic case for short recursive SFIR loops, which can be understood because the filter structure is a comb filter with regularly spaced dips and peaks in the frequency domain. To keep the ripple around ± 1 dB the length of the delay-line should not be below 1/3 of T_{60}, which means about 0.7 seconds for a concert hall.

If the signals to be reverberated do not contain long steady-state partials or harmonic complexes (for example in speech there seldom are steady-state parts longer than 150 ms), this problem does not appear easily. Therefore the design of the reverberation model can be optimized to the application at hand and signals to be processed in it.

6.5. Sound examples available in the Web

Sound examples processed with the concert hall model are available at `http://www.acoustics.hut.fi/demos/VelvetReverb/`. These include impulse responses as well as convolved transient, speech, and steady-state sounds in comparison for the original measured response and the modeled response.

7. DISCUSSION

The example presented above demonstrates that reverberation modeling based on velvet noise is a practically useful method for reverb design, although it is not as efficient computationally as feedback delay networks or other classical reverb algorithms. Among advantages of the presented approach are practically zero processing latency and easy parametrization of models.

Parametric controllability of the reverb algorithm is characterized in Fig. 10 by arrows. The amplitude of the direct sound is of course easy to control, and its spectral shape is also controllable by AR-filter coefficients, if desired. The initial delay between the direct sound and the first reflection as well as the level of the noise-like floor between them are controllable. Discrete reflections can be separated and controlled in amplitude, position, and spectral shape, if desired, although this part was modeled as a whole in the example above. For late reverberation, T_{60} and general spectral shape can be controlled, although changing the reverberation time profile requires computing a new temporal envelope according to Eq. (3) or new tap coefficient tables including this envelope.

The modeling example above was related to a concert hall impulse response. Different challenges are faced when modeling small rooms, although in general these cases are simpler and end up with more efficient computation due to short room reverberation. The initial gap between the direct sound and first reflections is short and the reflection density grows relatively high very rapidly. Reverberation times are typically shorter, and there can be flutter type of echoes, which need recursive filter structures. An essential difference is also due to the appearance of discrete modes at low frequencies below the critical (Schroeder) frequency. This might require using a filterbank to realize these modes separately.

Modeling of small rooms with velvet noise reverberation is out of the scope of this paper, and remains a task for a future study. Other topics for detailed future work are the design of reverberation for binaural and multichannel reproduction, as well as optimization and evaluation of realtime performance of the algorithms.

8. ACKNOWLEDGMENTS

The work of Hanna Järveläinen was based on the Academy of Finland project 105651 (MAPS).

9. REFERENCES

[1] A. D. Pierce, *Acoustics — Introduction to Its Physical Principles and Applications*. Woodbury, NY.: Acoust. Soc. Am., 1991, second printing.

[2] W. G. Gardner, *Reverberation Algorithms*, ch. 3 in "Applications of Digital Signal Processing to Audio and Acoustics" (ed. M. Kahrs and K. Brandenburg). Boston: Kluver Academic Publishers, 1998.

[3] M. R. Schroeder, "Natural sounding artificial reverberation," *J. Audio Eng. Soc.*, vol. 10, no. 3, pp. 219–223, 1962.

[4] M. R. Schroeder, "Digital simulation of sound transmission in reverberant spaces," *J. Acoust. Soc. Am.*, vol. 47, no. 2, pp. 424–431, 1970.

[5] J. A. Moorer, "About this reverberation business," *Computer Music Journal*, vol. 3, no. 2, pp. 3255–3264, 1979.

[6] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *AES 90th Convention,* Preprint 3030, (Paris, France), 1991.

[7] M. Karjalainen and H. Järveläinen, "More about this reverberation science: Perceptually good late reverberation," in *AES 111th Convention,* Preprint 5415, (New York, USA), 2001.

[8] L. Schreiber, "Was empfinden wir als gleichförmiges Raushen?," *Frequenz,* vol. 14, no. 12, pp. 399–403, 1960.

[9] P. Rubak and L. G. Johansen, "Artificial reverberation based on a pseudo-random impulse response," in *AES 104th Convention,* Preprint 4725, (Amsterdam), 1998.

[10] P. Rubak and L. G. Johansen, "Artificial reverberation based on a pseudo-random impulse response II," in *AES 106th Convention,* Preprint 4900, (Munich), 1999.

[11] J. Pumplin, "Low-noise noise," *J. Acoust. Soc. Am.*, vol. 78, no. 1, pp. 100–104, 1985.

[12] J. W. Hall III, J. H. Grose, and W. M. Harmann, "The masking level difference in low-noise noise," *J. Acoust. Soc. Am.*, vol. 103, no. 5, pp. 2573–2577, 1998.

[13] D. A. Eddins, "Monaural masking release in random-phase and low-noise noise," *J. Acoust. Soc. Am.*, vol. 109, no. 4, pp. 1538–1549, 2001.

[14] B. C. J. Moore, *Introduction to the Psychology of Hearing,* Fifth Edition. Academic Press, 2003.

[15] D. A. Eddins and D. M. Green, *Temporal Integration and Temporal Resolution*, ch. 6 in "Hearing" (ed. B. C. J. Moore), 2nd Edition. Academic Press, 1995.

[16] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing Research*, vol. 47, pp. 103–138, 1990.

[17] J. Hynninen, *A software-based system for listening tests*. MSc Thesis. Helsinki University of Technology, 2001.

[18] J. Merimaa, T. Peltonen, and T. Lokki, *Concert Hall Impulse Responses — Pori, Finland: Reference*. Helsinki University of Technology, 2005. Report and data available at: `http://www.acoustics.hut.fi/projects/poririrs/`.

[19] A. Härmä and et al., "Frequency-warped signal processing for audio applications," *J. Audio Eng. Soc.*, vol. 48, no. 11, pp. 1011–1031, 2000.

[20] T. Paatero and M. Karjalainen, "Kautz filters and generalized frequency resolution — Theory and audio applications," *J. Audio Eng. Soc.*, vol. 51, no. 1/2, pp. 27–44, 2003.

[21] T. Paatero, "Modeling of long and complex responses using kautz filters and time-domain partitions," in *Proc. 12th European Sig. Proc. Conf. (EUSIPCO 2004)*, (Vienna), 2004.

[22] M. Karjalainen and T. Paatero, "Equalization of loudspeaker and room responses using Kautz filters: Direct least squares design," *EURASIP J. Applied Signal Processing*, 2007, accepted for publication.