More about this reverberation science: Perceptually good late reverberation

Matti Karjalainen and Hanna Järveläinen

Helsinki University of Technology

Laboratory of Acoustics and Audio Signal Processing,
P.O. Box 3000, FIN-02015 HUT, Finland
matti.karjalainen@hut.fi

ABSTRACT

The perceptual aspects of reverberation are less well known than the acoustic principle itself and its DSP-based simulation in artificial reverberators. In this paper, a series of psychoacoustic experiments are reported, along with their interpretation using auditory modeling, in order to reveal the underlying principles of late reverberation perception. Motivated by the results, a simple technique for reverb design is proposed.

0   INTRODUCTION

Reverberation is found in acoustics whenever there is a rich set of reflections or resonant modes, which cause coloration of the time-frequency response of a sound source. In addition to rooms and halls, many musical instruments, such as string instruments with a body or resonant strings or the piano with its soundboard, exhibit reverberant behavior. Three-dimensional resonators create a modal structure where the mode density increases as a function of frequency. From a time domain viewpoint, a dense distribution of reflections creates a diffuse, exponentially decaying impulse response.

The modal density in a diffuse reverberant room is high. It is proportional to the volume and the square of frequency so that at 1 kHz for a medium sized living room it is about 10-30 modes/Hz and for a concert hall few thousand modes per Hertz. At low frequencies, below a critical frequency (often called the Schroeder frequency) the field is not diffuse and dense of modes so that individual modes have to be considered separately.

The question of how we perceive dense modal patterns has not been studied as carefully as the physical basis of reverberation. Developers of artificial reverberators [1, 2] know that a too low mode density makes the sound metallic. The impulse response of a reverberant system should not be too regular in time or frequency, or otherwise perceptually 'good' diffuse reverberation is not achieved.

Three research approaches meet in this study (a) physical knowledge and modeling of reverberation, (b) signal modeling and DSP as means of artificial reverberation and experimen-

tal tool, and (c) psychoacoustics and auditory modeling for understanding how we perceive the phenomena. For relevant literature, primarily for modeling and artificial production of reverberation, see [2, 3, 4, 1, 5, 6, 7, 8, 9, 10].

In this study we are interested in the perceptual requirements of good late reverberation, restricting the investigations to monaural aspects.

The following questions were asked:

- What is the just noticeable difference (JND) in reverberation decay rate?
- What is the modal density needed per critical band?
- Optimal frequency-domain spacing of modes?
- Irregularity of time-domain response needed in each critical band?
- Does frequency modulation of modes help reducing modal density? What kind of modulation is needed?
- Frequency resolution needed for decay rate vs. frequency?
- Frequency resolution needed for proper magnitude response?
- Can we explain the perceptual criteria using an auditory modeling approach?
- Can we find a simple late reverb design based on these principles?

In this paper we answer some of these questions based on listening experiments, auditory modeling, and experimentation on simple digital filter models for modal responses.

1 IDEALIZED LATE REVERBERATION

Statistically, for a regular room [11], the modal density N'_{mode} (modes/Hz) for high frequencies is

$$N'_{\text{mode}}(f) = 4\pi V f^2 / c^3 \qquad (1)$$

and the reflection density N'_{refl} (reflections/second) for late reverberation is

$$N'_{\text{refl}}(t) = 4\pi c^3 t^2 / V \qquad (2)$$

where c is sound velocity and V is the volume of the room. It can be easily evaluated from (1) and (2) that for medium-to-large spaces, mid-to-high frequencies, and late reverberation, both modal density and reflection density are high compared to the spectral and temporal resolution of human auditory perception. Although for resonators of musical instruments this may not be fully true, it is useful to generalize and idealize reverberant behavior for perceptual investigations in order to find requirements for good late reverberation.

For this purpose, let us idealize reverberant responses in the following ways:

1. For determining JND of decay time (or for T_{60}), an idealized reverberant impulse response is assumed to be exponentially decaying Gaussian noise.
2. For determining perceptually sufficient modal density, a decay time approaching infinity can be assumed as the extreme case. Using flat spectrum (by critical-band resolution or within frequency range of interest) helps experimental configuration. As an idealized reference case, Gaussian noise can be used since it is found that it represents good uncolored late reverberation.

For the first case, JND of decay time is tested below for typical values of reverberation (0.5, 1.0, and 2.0 seconds) both for the impulse response itself and for a speech signal convolved with such idealized reverberation. For the second case, perceptually sufficient modal density, several techniques to synthesize dense modal distribution are tested by subjective experimentation. For these noise-like signals the requirement is assumed to be, as far as spectral flatness is valid, the lack of perceivable regularities in temporal envelopes within each critical band. The use of an auditory model is proposed to test this. If the modal density is too low, tonality or metallic timbre is perceived due to short-term periodicity or quasi-periodicity in amplitude envelope in critical band(s) of interest.

In this study we investigate the perception of reverberation using idealized impulse response $h(t)$ of the form

$$h(t) = \sum_{i=1}^{N} A_i e^{-\tau_i t} \sin(\omega_i t + \phi_i) \qquad (3)$$

consisting of a set of decaying sinusoids, i.e., modal decays. With high enough number of such sinusoids with properly selected angular frequences $\omega_i = 2\pi f_i$, decay parameters τ_i, initial levels A_i, and initial phases ϕ_i, any (monaural) late reverberation can be approximated. Decay parameter τ can be computed from a given reverberation time T_{60} by

$$\tau = \frac{1}{T_{60}} \log(10^3) \approx \frac{6.91}{T_{60}} \qquad (4)$$

The impulse response of Eq. (3) can be realized by a discrete-time parallel filterbank of second-order filters

$$
\begin{aligned}
H(z) &= \sum_{i=0}^{M} A_i \frac{b_{0,i} + b_{1,i} z^{-1}}{1 + a_{1,i} z^{-1} + a_{2,i} z^{-2}} \\
a_{1,i} &= -2 r_i \cos(\Omega_i) \\
a_{2,i} &= r_i^2 \\
b_{0,i} &= \sin(\phi_i) \\
b_{1,i} &= r \cos(\phi_i) \sin(\Omega_i) - r \sin(\phi_i) \cos(\Omega_i) \\
\Omega_i &= 2\pi f_i / f_s \\
r_i &= e^{-\tau_i / f_s}
\end{aligned}
\qquad (5)
$$

where i is mode index, M is the number of second order sections. Other symbols are as defined for Eq. (3).

While in real physical reverberation systems the decay parameters τ_i have a wider distribution [3, 1], in artificial reverberation by (3) these parameters can be made equal (within a critical band or frequency range of interest) in order to minimize the number of modes needed [7]. This can be understood easily in the case of very long reverberation time, because the only way to have contribution from each mode over a long time period is to make decay rates equal. Also, making initial amplitude parameters A_i approximately equal (at least by critical-band resolution) maximizes the contribution of each modal component to obtain a low level of perceived modulation periodicity. Such "nonphysical" distribution of modal parameters helps reducing the required modal density of perceptually uncolored reverberation compared to required modal density produced physically [1].

Initial phases ϕ_i can be randomized between $0 \dots 2\pi$ radians. If they are not randomized, the onset part of impulse response from (3) will have perceivable artifacts.

2 ANALYSIS OF REVERBERATION THROUGH AUDITORY MODELING

To analyze perceptual properties of a steady-state noise-like signal with no transients or onsets, a relatively simple computational *auditory model*[1] is applicable. Figure 1 depicts a general block diagram that includes a filterbank[2] for frequency selectivity, half-wave rectification for envelope detection, a low-pass filter for simulating the synchrony loss in neural firings towards high frequencies, a block for firing rate adaptation after signal level changes, and a block for temporal integration/masking/modulation analysis.

A gammatone filterbank [12] is typically used to simulate the frequency selectivity of the auditory system. Neural synchrony cutoff can be simulated for example by a second order lowpass filter with 1–2 kHz cutoff frequency. The adaptation block can be omitted when steady-state sounds are analyzed.
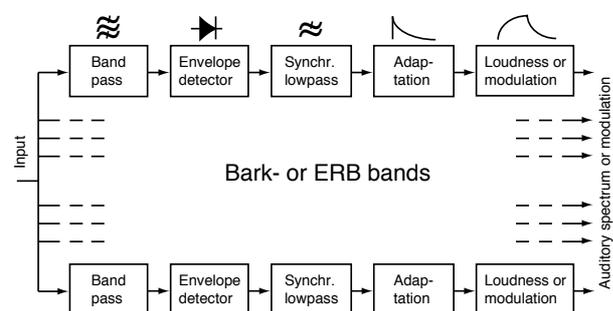

Fig. 1: Block diagram of an auditory filterbank model.

The last block of Fig. 1 is a (nonlinear) lowpass filter when loudness is computed, or a bandpass filter when modulation within critical bands is analyzed as is done in this study. Figure 2 illustrates how this filter should be sensitive to amplitude modulation at different modulation frequencies for a sinusoidal carrier of 1 kHz and two different levels, as well as for modulation of white noise. The increased sensitivity (lower detection threshold) for higher signal levels of modulated sine wave is due to increased spreading of excitation to neighboring critical bands (off-band listening) which does not take place in the case of modulated wideband noise.
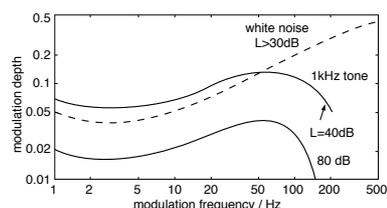

Fig. 2: JND of amplitude modulation as a function of modulation frequency for a 1 kHz tone, levels 40 and 80 dB, and for wideband noise (after [13]).

2.1 Modulation regularity in critical bands

For a stationary noise, the amplitude envelope of a critical band filter output shows no repeating periodicity, although the noise introduces low-level random self-modulation. For reduced modal density, the envelope starts to exhibit increasing (quasi)periodicities which reduce the perceived quality when used as idealized reverberation.

To illustrate the effect of reduced modal density, Figures 3a-d show four cases of envelope spectrum of a 160 Hz wide band (920–1080 Hz $\sim$ 1 Bark) of modal distribution around the center frequency of 1 kHz. Sinusoids of equal level and logarithmic frequency spacing are mixed together to simulate modes (with infinitely slow decay). In (a) there are only 5 partials, in (b) 20, in (c) 80, and in (d) 320 partials within this critical band. The curves illustrate modulation spectra, i.e., spectra of half-wave rectified and 1 kHz lowpass-filtered envelopes of the critical band for modulation frequencies 0–180 Hz (equal modulation sensitivity weighting was applied).
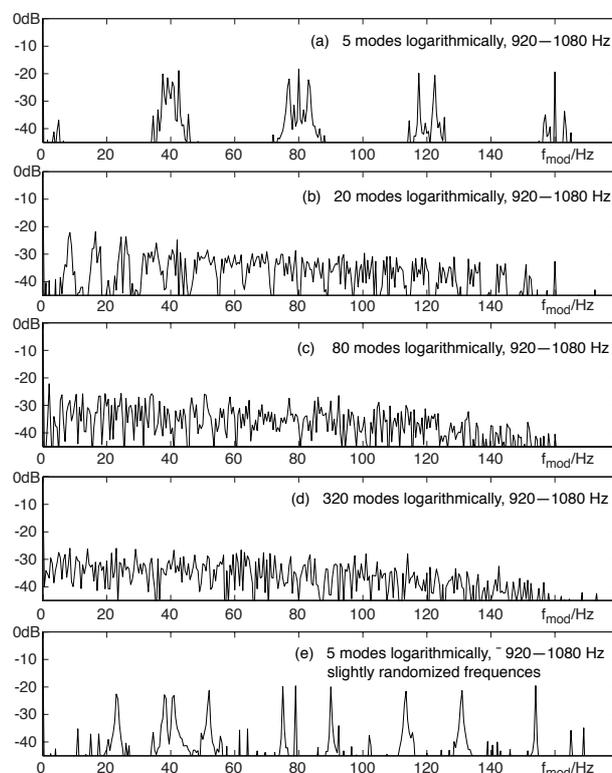

Fig. 3: Envelope modulation spectra within critical band 920–1080 Hz for different mode densities of logarithmically distributed modes: (a) 5 modes, (b) 20 modes, (c) 80 modes, and (d) 320 modes. Case (e) corresponds to (a) except that modal frequencies have been randomly displaced around their logarithmically distributed positions.

---

[1] The concept *auditory model* refers to any method of simulating the auditory system or its functionality. The terms *perceptual model* and *psychoacoustic model* are more often used in audio, contrasting to physiologically motivated modeling.

[2] Perceptual models are often based on applying Fourier transform first, frame by frame, and then using frequency domain processing for warping, spreading, masking, etc. In the present study we are interested in the temporal fine structure within each critical band, and thus a filterbank auditory model is necessary.

For low mode densities, strong metallic roughness as well as (quasi)periodicity is perceived due to modulation components of 15–100 Hz about 20 dB below average envelope level. When the modal density is increased, both roughness and periodicities are less pronounced. Case (c) with 80 modes sounds already quite random, and further increase in modal density makes the signal approach truly random bandpass noise.

According to Fig. 2, for modulation frequencies around 4 Hz, a modulation depth of about 0.04 (-28 dB in level) corresponds to JND threshold of modulation perception for broadband noise, and about 0.02 (-34 dB in level) for loud sine or narrowband signal. This corresponds relatively well to cases (c)–(d) in Fig. 3 and we might expect that the JND threshold of modal density to reach idealized noise-like perception could in this case be around 100.

Minor randomization of the modal frequencies helps making the resulting sound more random. Particularly the perceived periodicity can be reduced but the roughness at low modal densities is not decreased essentially. Figure 3e plots the envelope modulation spectrum for a 5 modes per critical band case, otherwise equivalent to case (a) but having slightly randomized positioning of mode frequencies.

3 LISTENING EXPERIMENTS AND RESULTS

Very little information on the perceptual features and requirements of reverberant systems is available, such as JND of reverberation time of required modal density [14]. Thus the main goal of this study was set to conduct listening experiments in order to estimate them. Such data is important for example in evaluating the adequacy of the design of real or artificial reverberation. Also, comparing such data to what an auditory model predicts could show how well we can explain and model the underlying phenomena.

The JNDs for reverberation time and the effects of reduced modal density were measured in two separate listening tests. Four subjects participated in both experiments. None of them reported any hearing defects. The tests were carried out in a listening room by headphone listening. The average sound level was over 80 dB, even though the listeners were allowed to set the master volume to a comfortable level. A rehearse cycle was played prior to the actual test to ensure consistent judgments.

3.1 JND of reverberation time

In the first experiment the just noticeable difference in reverberation time was investigated. An idealized, spectrally dense modal decay was generated based on Gaussian noise, its amplitude envelope being shaped with an exponentially decaying window. For simplicity, all frequencies had the same reverberation time up to 10 kHz which was the cutoff frequency of the synthetic impulse response. Although in real rooms the highest frequencies typically decay faster than mid frequencies, relatively constant reverberation time up to several kHz can be found in concert halls or listening test rooms. Thus the idealized reverberation of this study is not too far from reality.

Room impulse responses obtained in this way, with $T_{60} = 0.5$, 1.0, and 2.0 seconds, were taken as standard tones. T_{60} was then varied both up and down from the reference value by 4 steps of 5 % each. Two variations of excitation signal were used: an impulse and a speech signal. Combining these two variables resulted in six experimental cases altogether.

Thresholds were measured for detecting either an increase or a decrease in reverberation time, expressed as T_{60} required for 75 percent correct answers [15]. The experiment was a same vs. different pairwise comparison test, in which the subjects were required to distinguish between the standard tone with T_{60} fixed and stimuli whose reverberation time varied between 80 % and 120 % of that of the standard. A stimulus was present in half of the trials, and each such trial was judged four times. The rest of the trials included the standard tone twice. The playback order of the sound pairs was randomized within cases, and the standard was the first sound within a pair twice and the second one twice [16].

A reported difference was either a *hit* or a *false alarm*, depending on whether a stimulus was present in the trial. A measure of sensitivity, percentage correct $P(C)$, was derived for each condition from the proportion of hits ($p(H)$) and false alarms ($p(FA)$) as follows:

$$P(C) = \frac{p(\text{hit}) + (1 - p(\text{false alarm}))}{2} \qquad (6)$$

The values of the function given above range between 0.5, which corresponds to chance level, and 1.0, which corresponds to 100 % performance. The detection threshold was found by estimating the T_{60} required to reach the midpoint (i.e., the 75 % point) of the $P(C)$ function. If the threshold was not directly evident in the data, it was interpolated between the nearest higher and lower scores. The advantage of this measure is that it eliminates the effects of response bias that is caused by favoring either same or different [15].

Listening to impulse responses

Fig. 4 plots the results of the reverberation time JND test when the excitation was an impulse, i.e., the subjects listened to the pure impulse response. The two topmost figures show the median and the 25 % and 75 % quartiles of the judgments for the upper and lower thresholds, respectively. The mean results are summarized in the bottom figure and in Table 1 relative to each reference value of T_{60}.

The mean results vary between 5.0 % and 7.2 % for the upper JNDs and 5.0 % and 12.5 % for the lower JNDs. An analysis of variance (ANOVA) [17] was conducted to reveal the possible effect of T_{60}. The result was insignificant for both upper and lower thresholds ($p = 0.27$ and $p = 0.06$, respectively). Thus the data could be collapsed over T_{60}, resulting in an overall mean tolerance between 92 % and 106 % of the center value of T_{60}.

Listening to reverberated speech

To obtain the JNDs of reverberation time for a more realistic case, the impulse responses in the previous case were used to convolve with a short (0.8 seconds) sample of speech to obtain the reference and test signals. The experiment was otherwise arranged as the one with pure impulse responses.

The results are depicted in Fig. 5. The topmost figures show again the boxplots of the upper and lower individual JNDs. The mean thresholds for reverberated speech are seen in the bottom figure, ranging between 4.6 % and 6.7 % for the upper JNDs and between 3.3 % and 9.6 % for the lower JNDs. The variation was slightly higher for the lower thresholds, for which a significant effect of T_{60} was observed ($p = 0.015$). The lower JNDs for $T_{60} = 0.5$ s seem to be closer to the reference value than for $T_{60} = 1.0$ or 2.0 s. However, no systematic behavior of the thresholds as a function of T_{60} was found.

The threshold data for both impulse responses and reverberated speech were checked for a significant effect of the excitation signal. Such effect was found for the lower JNDs for T_{60} = 0.5 s ($p = 0.007$) and T_{60} = 2.0 s ($p = 0.043$). However, the differences are relatively small. Table 1 summarizes the mean results for both excitation types.

Table 1. Mean upper and lower relative JNDs for each center value of T_{60} in experiment 1.

Sample		T_{60}=0.5 s	T_{60}=1.0 s	T_{60}=2.0 s
Impulse	Upper	1.072	1.066	1.050
Impulse	Lower	0.927	0.875	0.950
Speech	Upper	1.060	1.067	1.046
Speech	Lower	0.967	0.904	0.913

A basic difficulty in this kind of experiment is that subjects cannot easily focus on differences in T_{60} but rather they listen to any perceivable differences. Because the signals are shaped from noise with inherent self-variation, differences can be perceived in highly attentive listening even between parametrically equivalent signals. Thus the extreme JND of below ±10 %, as found above, may be too strict for practical purposes.
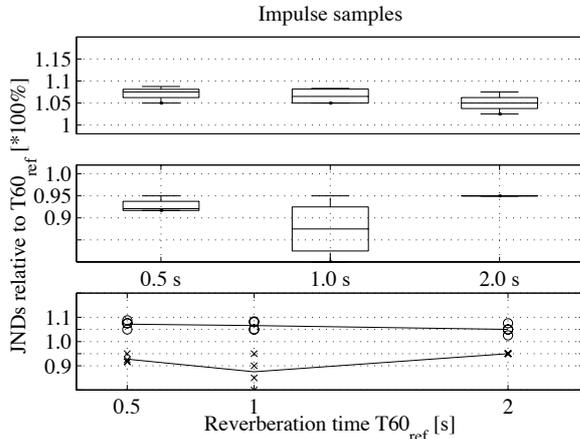
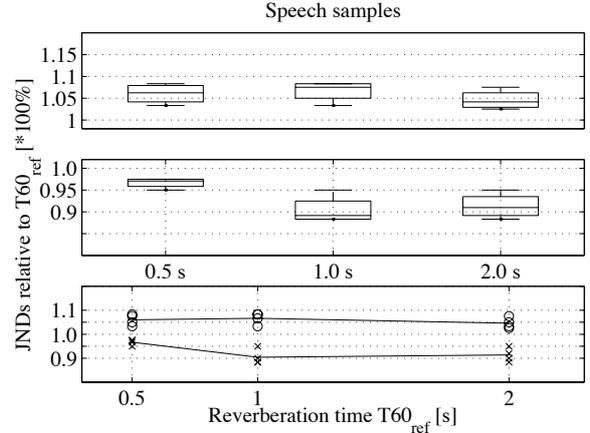Fig. 4: Boxplot of the relative upper JNDs (top) and lower JNDs (middle) of reverberation time for impulse responses. Bottom: Relative upper and lower JNDs as a function of T_{60}. (x) = individual lower thresholds, (o) = individual upper thresholds. Means connected with solid lines.

3.2 Minimum modal density of perceptually diffuse reverberation

As mentioned above, a low modal density of a reverberant response sounds rough, metallic, or periodically fluctuating. For high enough modal density the auditory perception of reverberation becomes spectrally dense and temporally diffuse, and further increase in modal density does not change the character of perception. This minimal modal density of perceptually diffuse reverberation was the second target of subjective experimentation in this study.

The first question before starting experimentation is the optimal distribution of modes. If modal frequencies are positioned too regularly, periodic fluctuation or other similar artifacts can be perceived. We approached this question by informal listening of reverberation based on some basic distributions.

Fig. 5: Boxplot of the relative upper JNDs (top) and lower JNDs (middle) of reverberation time for speech samples. Bottom: Relative upper and lower JNDs as a function of T_{60}. (x) = individual lower thresholds, (o) = individual upper thresholds. Means connected with solid lines.

The first idea was to try logarithmically uniform positioning of modal frequencies. For a single critical band (Bark band) of 920–1080 Hz, noise-like perceptually diffuse signal from Eq. (3) in a non-decaying case was achieved when approximately 80 modes were used. Randomized repositioning of the modal frequencies (about ±30–40 % of their spacing interval) helped in reducing the required modal density to about 30 for the same critical band.

In an informal broad-band (80 Hz – 10 kHz) experiment with logarithmically distributed modal frequencies and reverberation time of 1 second it was found that more than 1500 modes are needed to avoid artefacts. Otherwise the impulse response exhibits some metallic timbre due to insufficient modal density at high frequencies. When some arbitration of modal frequencies was applied as mentioned above, the required number of modes to avoid artifacts was found to be about 1200, and even with a lower number of modes the artifacts were not highly noticeable.

As another simple case of modal positioning, linear frequency distributions without and with minor repositioning were listened to informally. The nominal frequencies were linearly distributed between 80 Hz and 10 kHz. With low mode densities there is a clearly audible periodicity (flutter echo effect) in the reverberation. Without randomization the minimum number without artifacts was above 1500. Slight randomization did not reduce essentially the required number of modes but made flutter echo with lower mode densities clearly less pronounced than without randomization.

A known technique to design maximally non-periodic sequences is to apply prime numbers [18]. Using the principle here means to position modal frequencies proportionally to prime numbers. In an informal listening the resulting reverberation was found to be very similar to uniform distribution slightly randomized.

The next distribution that was experimented was uniformity on the Bark scale. Without frequency randomization it requires more than 1000 modes for the 80 Hz to 10 kHz range.

With randomly uniform repositioning of ± 0.4 Barks of the mode frequencies it yields better results than the methods above so that already 700 modes result in a useful reverberation quality, although a higher number of modes is needed for perfect randomness. (In formal listening tests reported below the requirements for perfectly diffuse reverberation were found slightly more strict.)

The reverberation filterbanks obtained by modal distributions discussed above are time-invariant, since a fixed set of modal frequencies is used even if the positions are randomized. A different strategy is to implement an invariant filterbank by modulating modal frequencies and thus enhancing the randomness of response. Such time-variance is often used in artificial reverberation. The type, amount, and frequency of modulation must be selected carefully to avoid noticeable degradations of sound quality with some critical signal types.

This strategy was applied to the parallel filterbank reverberation filter by modulating each mode frequency of a uniform Bark bank with an independent low-pass random signal. Because finding the best modal modulation strategy is a complex multidimensional optimization problem, no formal listening tests were conducted. Instead of that, it was found in preliminary and informal listening that the modal density can be reduced to about 10 per Bark, i.e., about 240 for full audio range, without deteriorating the reverb useless. The metallic timbre is reduced, but signal modulation artifacts emerge easily along with increased modulation depth that is needed for lowest modal densities.

3.3 Results of formal experiments

Based on the preliminary experiments described above, another set of formal listening experiments was carried out to find the minimum modal density for perceptually diffuse reverberation. A variable number of modes were uniformly distributed on Bark scale with static randomization of frequency. 5000 modes were used for the standard tone, a number that was found perceptually diffuse in preliminary listening. Five test tones were then generated with their modal density reduced to 1500, 1000, 710, 500, and 305. The experiment was run using both impulse and speech excitations and fixed reverberation times T_{60}= 0.5 s, 1.0 s, and 2.0 s. Combining the conditions produced six experimental cases.

The task was to grade the quality of each stimulus compared to the standard tone on a scale from 1.0 to 5.0. The subjects were particularly asked to observe the known effects of insufficient modal density: metallic timbre, roughness, or periodic fluctuations. The labels 'Very metallic', 'Rather metallic', 'Somewhat metallic', 'Very little metallic', and 'Not at all metallic' were associated to the numeric grades from 1 to 5, respectively.

The judgments were given by moving a continuous slider of the graphical UI and recorded by one decimal. The subjects could switch freely between the standard and the stimulus until they were ready to make a judgment. A rehearse cycle was used to make the test material familiar to the listeners and to encourage them to use the entire scale. In the actual test all stimuli, including a standard-standard pair per each experimental case, were judged three times.

The results of the modal density experiment are seen in Figures 6, 7, and 8 for reverberation times T_{60}= 0.5 s, 1.0 s, and 2.0 s, respectively. The boxplots of the top figures present the median grades given to each test tone for both impulse and

speech excitations as well as the 25 % and 75 % quartiles. The bottom figures show the mean results for impulse and speech samples separately as well as their common average.

For all reverberation times, 1500 modes were considered enough to produce as good quality as the standard (with 5000 modes). With decreasing modal density, the grades decreased from around 5 to less than 2. Tones with 1000 modes received an average grade 4 or higher in all cases. An effect of reverberation time was observed for both impulse responses ($p = 0.002$) and speech excitations ($p = 0.002$) but a significant effect of excitation was only observed for T_{60}= 0.5 s ($p = 0.031$), for which the speech samples received better judgments for lower mode densities.
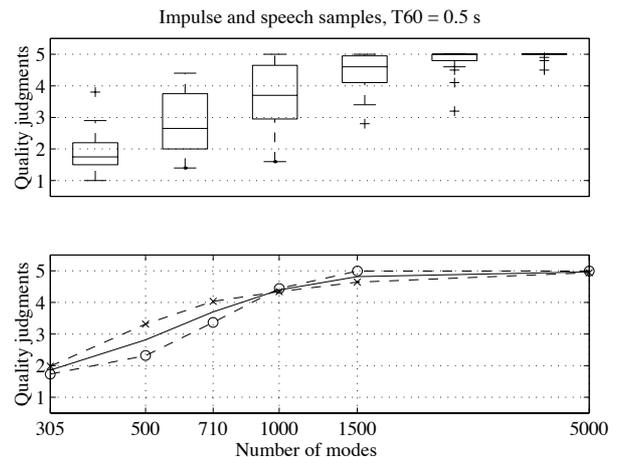

Fig. 6: Top: Boxplot of the results of the modal density experiment for T_{60}=0.5 s combined for impulse and speech samples. Bottom: Mean results for speech samples (–x–), impulse samples (–o–), and their common average (solid line).
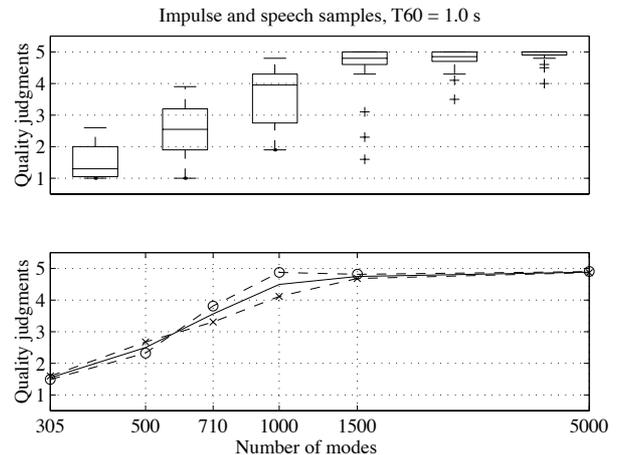

Fig. 7: Top: Boxplot of the results of the modal density experiment for T_{60}=1.0 s combined for impulse and speech samples. Bottom: Mean results for speech samples (–x–), impulse samples (–o–), and their common average (solid line).
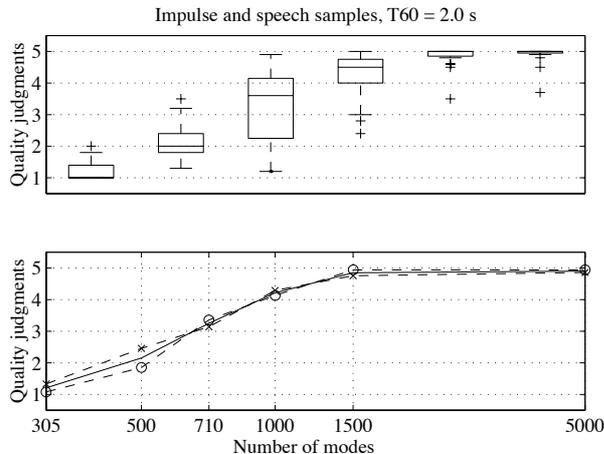
Fig. 8: Top: Boxplot of the results of the modal density experiment for T_{60}=2.0 s combined for impulse and speech samples. Bottom: Mean results for speech samples (–x–), impulse samples (–o–), and their common average (solid line).

4 DESIGN OF MODAL FILTER REVERBS

Based on the experiments described above, a simple approach to reverb design can be taken by implementing late reverberation as a parallel filter structure of Eq. (3) or as a filter derived from it, such as a direct form filter[3]. We call this approach the Modal Filter Reverb (MFR) approach. From the results of listening experiments it can be concluded that for a time-invariant case, filter orders 2000 to 3000 are needed for highest quality reverberation because each mode reserves a second order filter section. Already 700 modes (order 1400) may be useful depending on application. In time-varying cases, filter orders down to 500–700 are found useful at least for specific applications.

The approach can be easily applied to arbitrary reverb design, not only to late reverberation. In the case of the parallel structure, a cascaded filter (such as an FIR or any suitable digital filter) can be attached that simulates direct sound, early reflection, or any non-diffuse initial part of response. In principle it could be possible to apply some iterative pole-zero filter design techniques, such as in [19], to find a nearly optimal modeling of a target response by a single filter. Perceptual aspects should take an important role in such optimization because physical reality is not the goal of such modeling.

One of the advantages of such modal filter reverbs is that they can be designed easily to meet desired parametric behavior, i.e., desired modal density, reverb time, and magnitude response as function of frequency. The MFR structure is also well suited to specific colored reverberation such as musical instrument body modeling and other related audio effects, both in time-invariant and -variant versions.

A disadvantage of MFRs is their clearly higher computational cost compared to delay-based structures, such as feedback delay networks (FDNs) [7]. This is due to the fact that each

mode is realized separatately instead of generating a bunch of modes by each feedback delay loop.

The computational complexity of MFRs is, however, much lower than that of direct FIR-based convolution. For example for T_{60} of two seconds and sample rate of 44100 Hz an FIR of $1.5 \cdot 2.0 \cdot 44100 \approx 130000$ taps is needed for a 90 dB decay range. A modal IIR filter of order 1400 is two orders of magnitude more efficient to compute than full convolution. Time-variant versions of MFRs can be even more efficient. The relatively high computational cost of MFRs compared to simple delay feedback structures is becoming less a decisive factor in the future since according to Moore's law of computation power growth one MFR reverb channel will require only a fraction of available processing power on a typical processor in ten years or sooner.

5 DISCUSSION AND CONCLUSIONS

This study has concerned some problems related to perceptually high-quality late reverberation. Listening experiments have been carried out to find the just noticeable difference in reverberation time and minimum modal density requirement of spectrally and temporally diffuse reverberation. An auditory modeling approach is suggested for analyzing the lack of artifacts (coloration and periodicities) in reverberation.

The listening experiments show that for a proper modal parameter distribution a much lower modal density is enough for high-quality reverberation than is found or needed in real acoustic rooms and halls. While for physically produced reverberation a modal density of 3/Hz at 1 kHz, i.e., $3 \cdot 160 \approx 500$ modes for the Bark band around 1 kHz are needed [1], the 'homogenized' modal parameter distribution applied here needs only about $1500/24 \approx 60$ modes for the same Bark band.

The idea of using an auditory modeling approach to evaluate reverberation quality and lack of artifacts is proposed. More work will be needed to make this approach a practical tool in reverberation studies and reverb design. Such modeling can also yield better perceptual understanding of reverberation in general.

Finally, a straightforward approach to realize artificial reverberation based on parallel filter structures has been discussed. Although the method is computationally more expensive than reverbs using delay or feedback delay structures, its flexibility makes it useful at least in specific cases, such as experimentation on reverberation and special audio effects.

ACKNOWLEDGMENTS

REFERENCES

[1] W. Kuhl, "Notwendige Eigenfrequenzdichte zur Vermeidung der Klangfärbung von Nachhall," in *Proc. 6th International Congress on Acoustics*, (Tokyo, Japan), pp. E–69–72, 1968.

---

[3]This assumes that the transfer function can be expanded into a rational expression (ratio of polynomials) and implemented without numerical precision problems.

[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 Rverberation Business," *Computer Music Journal*, vol. 3, no. 2, pp. 3255–3264, 1979.

[6] D. Griesinger, "Practical Processors and Programs for Digital Reverberation," in *Proc. AES 7th Int. Conf.*, (Toronto, Canada), pp. 187–195, 1989.

[7] J.-M. Jot and A. Chaigne, "Digital Delay Networks for Designing Artificial Reverberators," in *Preprint 3030, AES 90th Convention*, (Paris, France), 1991.

[8] W. G. Gardner, "The Virtual Acoustic Room," Master's thesis, MIT, 1992.

[9] P. Rubak and L. G. Johansen, "Artificial Reverberation Based on a Pseudo-Random Impulse Response, part I," in *Preprint 4725, AES 104th Convention*, (Amsterdam), 1998 May.

[10] A. Czyzewski, "A Method of Artificial Reverberation Quality Testing," *J. Audio Eng. Soc.*, vol. 38, no. 3, pp. 129–141, 1990 March.

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

[12] B. C. J. Moore, R. W. Peters, and B. R. Glasberg, "Auditory Filter Shapes at Low Center Frequencies," *J. Acoust. Soc. Am.*, vol. 88, no. 1, pp. 132–140, 1990 July.

[13] E. Zwicker and H. Fastl, *Psychoacoustics — Facts and Models.* Berlin: Springer-Verlag, 1990.

[14] P. Huang, S. Serafin, and J. O. S. III, "Modeling High-Frequency Modes of Complex Resonators Using a Waveguide Mesh," in *Proc. Conf. Digital Audio Effects (DAFX-00)*, (Verona, Italy), 2000 Dec.

[15] W. A. Yost, *Fundamentals of Hearing – An Introduction.* New York: Academic Press, 3rd ed., 1994.

[16] J. P. Guilford, *Psychometric Methods.* New York: McGraw-Hill, 1956.

[17] R. S. Lehman, *Statistics and Research Design in the Behavioral Sciences.* Belmont, California: Wadsworth Publishing Company, 1991.

[18] M. R. Schroeder, *Number Theory in Science and Communication.* New York, NY: Springer Verlag, 1986.

[19] T. Paatero and M. Karjalainen, "Kautz Filters and Generalized Frequency Resolution — Theory and Audio Applications," in *AES 104th Convention*, (Amsterdam), 2001 May.