PERCEPTUAL RESONATORS FOR INTERACTIVE WORLDS

Dylan Menzies

Zenprobe Technologies
dylan@zenprobe.com

ABSTRACT

The simulation of sounds in an interactive world requires efficient and flexible algorithms, so that a potentially large number of different sounds can be generated simultaneously, in real-time. Diffuse resonators, such as wooden doors, pose a particular challenge because existing delay-feedback based methods of simulation are relatively expensive. A *perceptual resonator* is presented here as an alternative. In this the perceptually relevant characteristics of a resonator are simulated rather than all the low-level acoustics, gaining efficiency and control at the expense of some loss of detail, and strict linearity. The resonator also forms an interesting tool for perceptual investigation and a protoype for a new synthesis class.

1. INTRODUCTION

There has been a growing interest in the simulation of sounds in interactive worlds, using physical models that are stimulated according to dynamical interactions in the world. This can be traced back to the work of Hahn [1], in which impacts between graphically rendered objects were modelled with modal resonators. Modal resonators themselves have a long history predating digital computers. Van den Doel [2, 3] expanded this to continuous contact excitations, still using modal resonators. The continuous excitations depend on the dynamics of the points of contact between resonating objects. Modal resonators are useful for simulating a variety of objects but they are not suitable for objects with *diffuse* resonance. These form a broad and important class, and an audio-enabled interactive world simulator should be expected to deal with them competently. Other digital linear structures exist for handling diffuse resonances, but the computational expense is considerably greater per object for similar quality. This paper investigates the nature of the diffuse resonance, reviews techniques that have been previously been applied and then proposes a new resonator paradigm that combines physical modelling with perceptual modelling. Some examples are described, but no assessment is made in terms of listener tests or objective perceptual comparison. It is suggested that the effectiveness of the technique will be quite obvious to most listeners.

The words *filter, response, resonator* shall be used interchangeably to mean a causal function of the input signal, so that the output can depend in a completely general way on the entire history of input. For the broad class of naturally occurring complex sound generating mechanisms no more can be assumed than this. Appending the word *linear* or using "*impulse response*" indicates the familiar linear filter to which spectral analysis can be applied, and in most cases digital approximation.

2. DIFFUSE AND NON-DIFFUSE RESONANCE AND THEIR SYNTHESIS

2.1. Non-diffuse

Linear filters for acoustic world simulation can be broadly divided into diffuse and non-diffuse resonators. Non-diffuse resonators have well defined resonances that can be simulated using parallel banks of second order resonator sections, as shown in Figure 1.

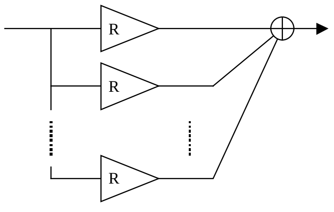

Figure 1: *Modal resonator structure*

The resonances of small solid objects are typically dominantly non-diffuse; cups, stones, coins.

2.2. Diffuse reverberation

Diffuse resonators have many resonances that are closely spaced, and hence the perception of clearly pitched resonance is lost and the impulse response is more noise like, rather than sinusoidal.

The characteristic physical quantity that determines the diffuseness of an object is the time, T, taken for sound to travel across its linear dimensions, since this provides a lower bound for resonant frequencies, 1/T. T depends on the size of the object and speed of sound in it. For large T, the onset of dense resonance starts at a lower frequency, sometimes called the *schroeder frequency* [11]. Also the duration of a typical real impulse from an impact becomes smaller relative to T, which reduces the blurring of the diffuse response as seen in the time domain.

The most obvious example of diffuse resonance is found in room reverberation. Natural concert hall reverberation has 1000s of modes per Hz, and of the order of 1000 effective modes are necessary in a good artificial reverberator [11]. If the acoustic tail of some reverberation is listened to in isolation it is perceived as coloured noise. This observation was confirmed by Karjalainen [11], who found the tail to be indistinguishable from noise. The most efficient methods of simulating reverberation employ delay-feedback structures and date back to the work of Schroeder [4]. More recently a class of delay-feedback structures called *feedback delay networks, FDNs*, originally proposed by Gerzon, have been used to simulate reverberation [5]. See Figure 2. These have the advantage that their spectral and decay profiles can be fitted to data using a design procedure.

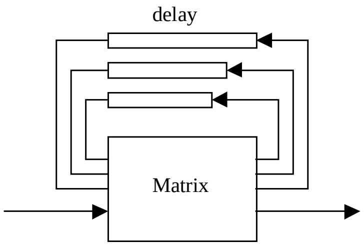

Figure 2: FDN structure

2.3. Diffuse object resonance

The sound generated when solid objects are struck may also have significant diffuseness. There appears to be little about this in the literature, so we shall rely on some simple observations and analogy with reverberation, which has been studied in more detail. See [11] for a recent study. An example of the connection with reverberation is provided very directly by ‘plate reverb’, an old method of generating reverberation using a large metal plate. A more familiar and important example of diffuseness from the viewpoint of world simulation is the resonance of a wooden door or table. The decay of the impulse response is short, so the diffuseness is not as apparent, although it is still perceptually important. In Figure 3 a dominant low frequency mode can be picked out with noise-like diffuse resonance superimposed. A spectrogram plot has been omitted because it is a poor method for showing up the noise component, which is low level and dispersed in the frequency domain. By listening to the decay section alone using an audio editor, the perception of the diffuseness is much clearer. This is because the perception of the complete impulse is dominated by the initial impact. The diffuseness decays exponentially like reverberation, and builds up over an initial onset period much as reverberation does but much faster. It can be expected that early-reflection peaks exist but it is very doubtful that these can play a significant perceptual role except for very large objects for which the timescale approaches that of reverberation.

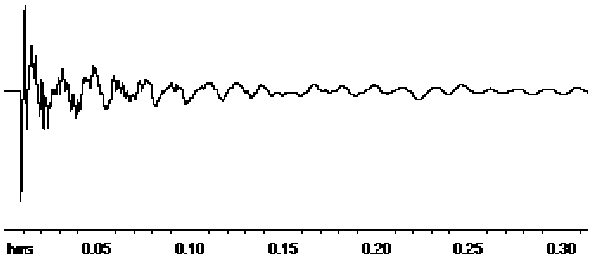

Figure 3: Recording of door being hit with a knuckle

Mixed in with the diffuse response are a few discernible low frequency modal resonances. Examples of simulating this kind of resonance are found in the music technology literature, again using delay-feedback structures: Digital waveguides can be used to explicitly solve the wave equation within a well defined geometrical shape, [6, 7, 8, 9]. See Figure 4.

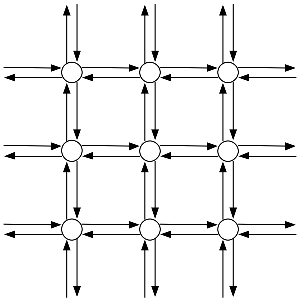

Figure 4: Section of a 2D waveguide lattice

This has been used for percussion instruments[1] [7,8]. Simpler and less geometrically determined waveguide structures can be devised in order to reduce cost. However, to create a sufficiently ‘smooth’ sounding diffuse resonance is still expensive relative to the cost of a modal resonator for a non-diffuse sound. It is also difficult to match a structure to given impulse or spectral data.

The *Data Driven* synthesis method described in [10] provides a very different approach, with direct input of control data rather than audio. The output is formed by mixing a basis of sound grains according to a bayesian predictor in lag-space. Diffuseness is achieved by mixing in more diffuse-sounding grains. This has been used to produce compelling if not hi-fidelity violin-like response. Nothing is assumed about the physics of sound generation. The following approach has something of the flavour of this, but with attention to perceptual relevance in the sound, and the physical process.

3. DIFFUSE PERCEPTUAL RESONATOR

3.1. Noise modes

We now consider the creation of a perceptual model for diffuse resonance based on the observations of Section 1. An important component in the sound is perceived as coloured noise, the part that is costly to simulate by physical modelling. The spectral profile of the noise is not static, however, and is determined by the changing excitation signal. This can be modelled approximately by summing a bank of dynamically controlled noise bands. Each band is modulated by a simple leaky-integrator, which is in turn fed by the level of the [input filtered by a matching band filter]. See Figure 5.

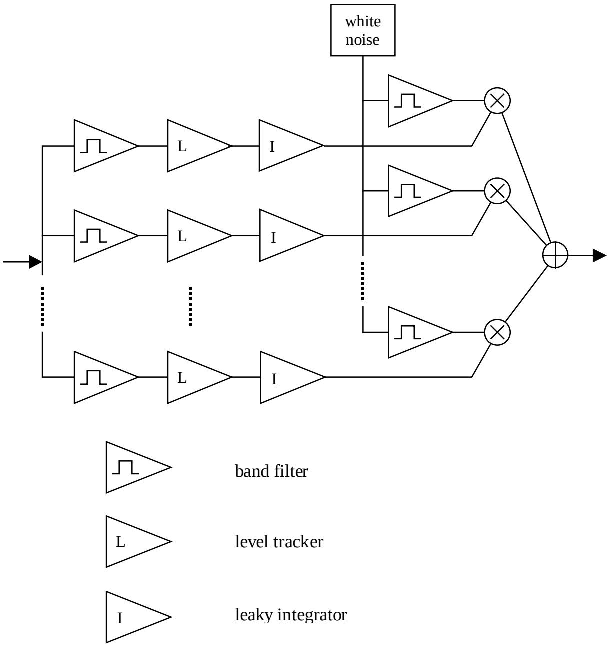

Figure 5: Noise band response generator

Crudely speaking, excitations with high frequency content will lead to high frequency diffuse output, and similarly for low frequency. The processing channel for each band shall be referred to as a *noise mode* by analogy with linear modes. The main feature required of the integrator is that its decay is exponential, with time constant matching the time constant of diffuse decay in the corresponding band. This is analogous to the time constant of reverberation. For instance to simulate the door knock of Figure 3, the time constant is of the order of 0.1 seconds. Such an integrator is implemented very simply as a first order filter with impulse response shown in Figure 6.

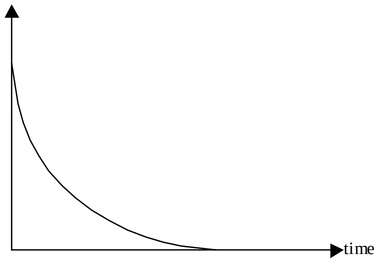

Figure 6: Simple first order leaky-integrator impulse response

An important refinement is to simulate the delay in the build up of maximum diffuseness, by softening the rise time of the impulse response. This is achieved simply by adding a second first order section in series with the

---

[1] Note the same approach can be taken for reverberation, but it is very costly since a large 3D lattice is required, and so it is only suitable for accurate offline acoustic simulation.

first, using a smaller time constant. The response is shown in Figure 7. No attempt is being made to recreate the early reflections since they are predicted to be unimportant by virtue of their very short time scale. In any case, early reflection simulation would destroy the simplicity and efficiency of this approach.

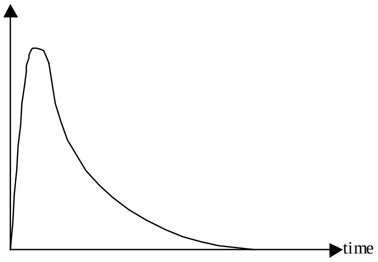

Figure 7: Modified integrator impulse response

The time constants of the level trackers need to be small compared with the integrator time bases. Just taking the absolute value of the input works well in practise. The collective state of the integrators contains the current spectral profile of diffuse energy within the modelled object. This structure is not linear, but the form of the spectral output has the correct response, because of the decorrelating effect of noise convolution.

3.2. Hybrid resonator

Generally, a diffuse resonance is accompanied by more distinct resonances in the lower frequencies. These can be accomodated by combining the noise band generator with a modal resonator, see Figure 8.

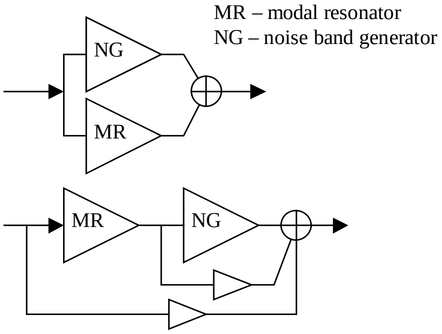

Figure 8: Hybrid resonator structures

Connecting in series, with direct and modal signals mixed in gives the most flexibility and the best results.

To create a working example, 3 modes and 3 noise modes were chosen to roughly match the decay times and frequency content of the sound in Figure 3. The model was then excited with a simple triangular pulse to simulate a knock, generating the response shown in Figure 9. An informal listener evaluation suggests that the example is close to the original and more importantly is virtually indistinguishable, *in character*, from wood resonance. Reducing the noise mode mix progressively to zero, removes the ‘woodiness’ to the point where it sounds synthetic. Figure 10 shows the knock-response of a resonator simulating an object such as a chair leg. The time constants are similar but the modes and noise modes are set at higher frequencies.

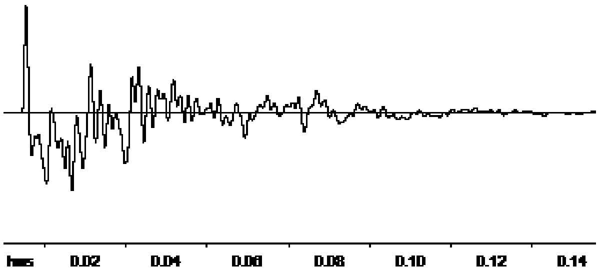

Figure 9: Artificial knock-response from a hybrid resonator

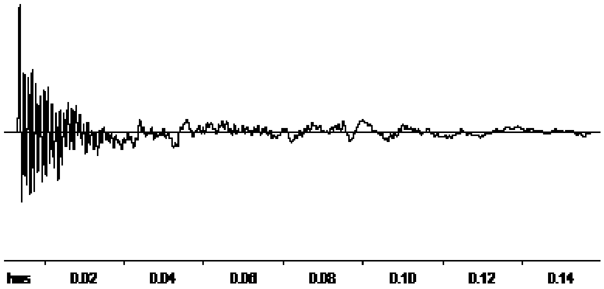

Figure 10: Hybrid resonator with high frequency response

Finally the perceptual resonator was incorporated into a physical world audio library, and substituted for modal resonators in existing demonstrations of physical audio in dynamic interactive worlds. For instance a man-hole cover spinning to the ground becomes a wooden wheel. The observed behaviour verifies the expected dynamic spectral variation according to the changing excitation signal, and fulfils the original goal.[2]

3.3. Evaluation

A variety of convincing wood-like responses can be generated with as few as 3 bands. The computational cost is about the same or less than a typical simple modal resonator, and compared with delay-feedback structures, virtually no memory is required. The level of diffuseness can be varied at will, providing an interesting confirmation of the importance of diffuseness to the perception of wood-like resonances.

4. CONCLUSION AND DISCUSSION

The perceptual resonator is analogous to perceptually based resynthesis using noise, transients and sines: A response function is constructed rather than a signal. As well as being realistic and cheap, the perceptual resonator is naturally parameterised, so that manual adjustments can easily be made to a sound for aesthetic or musical reasons. In the context of entertainment systems such as computer games, this process of fixing parameters is perhaps more useful than data-driven calibration, because the designer has a mental image of the desired response and may not wish to have to search through sample libraries to find an appropriate template. A few noise modes are sufficient to produce interesting behaviour. From the entertainment viewpoint this is more important than high fidelity. Unusual and exaggerated effects can be readily created without worries of stability.

Further work on calibration will be useful to investigate the perceptual basis of this technique. One possible route is to use algorithms that can separate a recording into noise and sine components, such as SMS [12]. The sine part can be analysed in a conventional way for modal content, while the noise part can be analysed for noise band decomposition and decay times.

In the context of the data driven synthesis method described in Section 2.3, a new hybrid is conceivable in which basis grains are supplemented with physically and perceptually motivated generating ‘atoms’. The data fitting procedure must fit these atoms to the desired response in an analogous way to the composition of a perceptually resynthesized signal consisting of noise, transients and harmonic signals.

Beyond linear resonators, there exist a great variety of non-linear responses that would be usefully simulated in an acoustic world. For this, the perceptual resonator approach provides an alternative to explicit physical modelling, permitting simpler and more intuitive calibration.

5. REFERENCES

[1] J. K. Hahn, J. Geigel, J.W. Lee, L. Gritz, T. Takala, and Mishra, S. "An Integrated Approach to Sound and Motion", *Journal of Visualization and Computer Animation*, 6(2):109-123.

[2] K. van den Doel, P.G. Kry and D.K. Pai. “FoleyAutomatic: Physically-based Sound Effects for Interactive *Graphics* Simulation and Animation”, *Computer ACM SIGGRAPH 01 Conference Proceedings*

[3] K. van den Doel, “Sound Synthesis for Virtual Reality and Computer Games”, *PhD thesis, University of British Columbia*, 1998

[4] M. R. Schroeder, “Natural-sounding artificial reverberation”. *I. Audio Eng Soc.* 10(3)

[5] D. Rochesso, and J. O. Smith, "Circulant and elliptic feedback delay networks for artificial reverberation". *IEEE trans. Speech & Audio* 5(1), 1997

[6] J. O. Smith. “Physical modeling using digital waveguides”, *Computer Music Journal*, 16(4):75-87, 1992

[7] J. O. Smith. “Physical Modeling Synthesis Update” *The Computer Music Journal*, vol. 20, no. 2, pp. 44-56, Summer 1996.

[8] S. A. Van Duyne and J. O. Smith, "Physical modeling with the 2-D digital waveguide mesh," in *Proc. 1993 Int. Computer Music Conf., Tokyo*, pp. 40--47, Computer Music Association, 1993.

[9] S. Van Duyne, J. Smith. “The 3D Tetrahedral Digital Waveguide Mesh with Musical Applications”, *Proceedings International Computer Music Conference*, Cuba, 2001.

[10] B. Schoner, C. Cooper, C. Douglas, “Data-driven Modeling and Synthesis of Acoustical Instruments”, *Proceedings International Computer Music Conference*, 1998

---

[2] Examples will become available at *www.zenprobe.com/pub*

[11] M. Karjalainen, H. Jarvelainen, “More About This Reverberation Science: Perceptually Good Late Reverberation”, *Preprint 5415, AES111 New York (2001) Convention*

[12] X. Serra, “Musical Sound Modeling With Sinusoids Plus Noise”, *Musical Signal Processing, Swets & Zeitlinger Publishers, 1997*