Proceedings of the 18th International Conference on Digital Audio Effects

DAFx15

Trondheim, Norway
November 30 - December 3, 2015

Proceedings published by:
Department of Music and Department of Electronics and Telecommunications
Norwegian University of Science and Technology
November 2015
Editors: Peter Svensson and Ulf Kristiansen

Credits:
Logo design DAFx-15: Amund Ulvestad
Cover design: Sigurd Saue and NTNU Grafisk senter
Cover photo: Endre Forbord

Web sites:
www.ntnu.edu/dafx15
www.dafx.de

ISSN 2413-6700 (Print)
ISSN 2413-6689 (Online)

Foreword

Welcome to DAFx-15 in Trondheim

It is with great pleasure that we welcome you to the 18th International Conference on Digital Audio Effects in Trondheim, Norway from Monday November 30 to Thursday December 3, 2015. For the first time the DAFx conference series revisits a host city, 16 years after Trondheim hosted the 2nd DAFx conference in 1999 at the Norwegian University of Science and Technology (NTNU), organized primarily by the Department of Electronics and Telecommunications. Since then the Department of Music has established an ambitious Music Technology group and the two departments are jointly hosting this year's edition of the DAFx conference. We hope that you will sense an inspiring mix of music and audio technology in the program that we present.

DAFx-15 offers four days of scientific program intermingled with concerts and social events. The bulk of the conference consists of 61 papers selected for presentation Tuesday through Thursday: 31 oral presentations and 30 poster presentations. We have organized the oral presentations in 11 sessions, covering topics such as virtual analog, sound synthesis, audio analysis, physical modelling, spatial audio, audio coding, speech applications and audio effects. The posters are deliberately not organized into topics to favor exchanges between scientists of the same field of research. There are 10 poster presentations each day, including two short announcement sessions in the auditorium.

We wanted the conference to reflect how advances in audio technology finds application in music performance and specifically asked for contributions related to real-time applications of digital audio processing. This theme have not made a strong impact on the contributed papers, but we have strengthened it through keynotes and tutorials. We have invited three distinguished keynote speakers to kick off each conference day with an inspiring talk. On Tuesday, the merited artist and researcher Marije Baalman will discuss digital audio as open source hardware and software in a performance context, and she will return as performer at the evening concert. On Wednesday, the brilliant young researcher Kurt Wegner will present status quo on wave digital audio effects at CCRMA in a keynote prepared with Julius Smith. Finally, Dr Franz Zotter from IEM, Graz, and chair of a prize-winning initiative on virtual acoustics within the German acoustical society, will present a keynote on Ambisonics audio effects on Thursday.

Three tutorials launches the conference on Monday. Peter Svensson, professor of electroacoustics at NTNU, starts with a tutorial on "Sound field modeling for virtual acoustics". Professors Øyvind Brandtsegg and Trond Engum will follow with a tutorial/musical demonstration of "Cross-adaptive effects and real-time creative use" together with invited musicians. The third tutorial will be given by Xavier Serra who will present the AudioCommons initiative that "aims at bringing Creative Commons audio content to the creative industries".

There will be several musical contributions during the conference, but the main concert event is held on Tuesday evening at Rockheim, the national museum of popular music. The social program also includes a welcome reception on Monday evening, right next to Nidaros cathedral, Norway's national sanctuary and it concludes with a conference dinner on Wednesday at the idyllic Ringve Museum, the national music museum.

We are proud to present to you the proceedings of DAFx-15. From this year on the publication is registered as an annual serial publication with an ISSN number. We hope this will make it even more attractive to contribute to future DAFx conferences. We have to thank all the great people that helped us make this year's conference happen, including the local organization team, all the reviewers from the DAFx-15 programme committee and the DAFx board members. Warm thanks go to the organizers of DAFx-13 and DAFx-14, who have been very helpful and responsive to our questions along the way. We would also like to extend our thanks to this years sponsors: Native Instruments GmbH, Audiokinetic Inc., Soundtoys Inc., Ableton AG and Aalberg Audio AS, and to our supportive institution, Norwegian University of Science and Technology (NTNU).

But most of all we would like to thank the DAFx community who makes these conferences possible by contributing great scientific work and enlightened discussions whenever we meet. We hope that DAFx-15 can play a similar role as earlier conferences in stimulating progress in the field of digital audio. When we presented our idea of DAFx-15 last year, we promised cold, darkness and snow. As we write these words, the first snow is falling heavily in the cold winter dark, so we will most likely keep our promise. Now it is all up to you, researchers, authors and participants at DAFx-15 to make this a memorable conference!

Welcome to Trondheim!

The DAFx-15 conference committee

Sigurd Saue (Music Technology): General Chair

Jan Tro (Electronics and Telecommunications): Vice Chair, Social program

Peter Svensson (Electronics and Telecommunications): Program Chair

Ulf Kristiansen (Electronics and Telecommunications): Program Co-chair

Øyvind Brandtsegg (Music Technology): Concerts Coordinator

Tim Cato Netland (Electronics and Telecommunications): Technical Coordinator

Conference Committees

DAFx Board

Daniel Arfib (CNRS-LMA, Marseille, France)
Nicola Bernardini (Conservatorio di Musica 'Cesare Pollini", Padova, Italy)
Francisco Javier Casajús (ETSI Telecomunicación - Universidad Politécnica de Madrid, Spain)
Laurent Daudet (LAM / IJLRA, Université Pierre et Marie Curie (Paris VI), France)
Philippe Depalle (McGill University, Montreal, Canada)
Giovanni De Poli (CSC, University of Padova, Italy)
Myriam Desainte-Catherine (SCRIME, Université Bordeaux 1, France)
Markus Erne (Scopein Research, Aarau, Switzerland)
Gianpaolo Evangelista (University of Music and Performing Arts Vienna)
Emmanuel Favreau (Institut National de l'Audiovisuel - GRM, Paris, France)
Simon Godsill (University of Cambridge, UK)
Robert Höldrich (IEM, Univ. of Music and Performing Arts, Graz, Austria)
Pierre Hanna (Université Bordeaux 1, France)
Jean-Marc Jot (DTS, CA, USA)
Victor Lazzarini (National University of Ireland, Maynooth, Ireland)
Sylvain Marchand (L3i, University of La Rochelle, France)
Damian Murphy (University of York, UK)
Søren Nielsen (SoundFocus, Arhus, Denmark)
Markus Noisternig (IRCAM, France)
Luis Ortiz Berenguer (EUIT Telecomunicación - Universidad Politécnica de Madrid, Spain)
Geoffroy Peeters (IRCAM, France)
Rudolf Rabenstein (University Erlangen-Nuremberg, Erlangen, Germany)
Davide Rocchesso (IUAV University of Venice, Department of Art and Industrial Design, Italy)
Jøran Rudi (NoTAM, Oslo, Norway)
Mark Sandler (Queen Mary University of London, UK)
Augusto Sarti (DEI - Politecnico di Milano, Italy)
Lauri Savioja (Aalto University, Espoo, Finland)
Xavier Serra (Universitat Pompeu Fabra, Barcelona, Spain)
Julius O. Smith (CCRMA, Stanford University, CA, USA)
Alois Sontacchi (IEM, Univ. of Music and Performing Arts, Graz, Austria)
Marco Tagliasacchi (Politecnico di Milano, Como, Italy)
Todor Todoroff (ARTeM, Bruxelles, Belgium)
Jan Tro (Norwegian University of Science and Technology, Trondheim, Norway)
Vesa Välimäki (Aalto University, Espoo, Finland)
Udo Zölzer (Helmut-Schmidt University, Hamburg, Germany)

Organizing committee

Sigurd Saue
Jan Tro
Øyvind Brandtsegg
Ulf Kristiansen
Peter Svensson
Tim Cato Netland

Program Committee, Reviewers

Trevor Agus
Jens Ahrens
Roland Badeau
Søren Bech
Stefan Bilbao
Øyvind Brandtsegg
Jean Bresson
Michael Bürger
Andres Cabrera
Christophe D'Alessandro
Bertrand David
Kristjan Dempwolf
Philippe Depalle
Christian Dittmar
Ross Dunkel
Dan Ellis
Gianpaolo Evangelista
Sebastian Ewert
Bernhard Feiten
Federico Fontana
Volker Gnann
Pierre Hanna
Christian Hofmann
Craig Jin
Jari Kleimola
Malte Kob
Sebastian Kraft
Ulf Kristiansen
Wolfgang Kropp
Tapio Lokki
Simon Lui
Tom Lysaght
Jaromir Macak
Piotr Majdak
Sylvain Marchand
Rémi Mignot
Damian Murphy
Thibaud Necciari
Søren Nielsen
Jyri Pakarinen
Jouni Paulus
Geoffroy Peeters
Peter Pocta
Stephan Preihs
Ville Pulkki
Rudolf Rabenstein
Tor Ramstad
Josh Reiss
Mark Sandler
Augusto Sarti
Sigurd Saue
Gerald Schuller
Xavier Serra
Jan Skoglund
Julius Smith
Clifford So
Alex Southern
Nicolas Sturmel
Fabian-Robert Stöter
Peter Svensson
Sakari Tervo
Jan Tro
Tony Tew
Joseph Timoney
Toon van Waterschoot
Maarten van Walstijn
Christophe Vergez
Tuomas Virtanen
Vesa Välimäki
Carl Haakon Waadeland
Rory Walsh
Yi-Hsuan Yang
Udo Zoelzer
Franz Zotter

Table of Contents

Oral session 2: Physical Modeling

Oral session 3: Audio Effects

Posters

Oral session 4: Audio and Music Analysis

Keynote 2

Oral session 5: Audio Coding and Implementation

Posters

Oral session 6: Spatial Audio and Auralization

Oral session 7: Virtual Analog

Posters

Oral session 8: Speech Applications

Keynote 3

Oral session 9: Perceptually Based Applications

Posters

Oral session 10: Virtual Analog Approaches

Oral session 11: Physical Modelling

Posters

REVERBERATION STILL IN BUSINESS: THICKENING AND PROPAGATING MICRO-TEXTURES IN PHYSICS-BASED SOUND MODELING

Davide Rocchesso
roc@iuav.it

Stefano Baldan
stefanobaldan@iuav.it

Stefano Delle Monache
sdellemonache@iuav.it

Iuav University of Venice, Venice, Italy

ABSTRACT

Artificial reverberation is usually introduced, as a digital audio effect, to give a sense of enclosing architectural space. In this paper we argue about the effectiveness and usefulness of diffusive reverberators in physically-inspired sound synthesis. Examples are given for the synthesis of textural sounds, as they emerge from solid mechanical interactions, as well as from aerodynamic and liquid phenomena.

1. INTRODUCTION

Artificial reverberation has always been part of the core business of digital audio effects [1, 2]. Its main purpose is that of giving ambience to dry sounds, mimicking propagation, absorption, and diffusion phenomena, as they are found in three-dimensional enclosures, at the architectural scale.

Ideally, artificial reverberators are linear time-invariant systems whose impulse response looks and sounds like a decaying noise. In the response of a real room, the early pulses correspond to the early reflections coming from the walls, and the density of pulses rapidly increases in time as a result of multiple reflections and scattering processes. It is often assumed that late reverberation is ideally represented as an exponentially-decaying Gaussian process [3, 4]. Essentially, a good reverb creates a kaleidoscopic and seemingly random multiplication of incoming pulses.

Feedback delay networks [4, 5] (FDN) are often used as the central constituent of reverberators, because they are efficient and their stability can be accurately controlled. FDNs can be parameterized according to reference room geometries [6, 7] or to recorded impulse responses [8], but they are also quite usable as instrument resonators or time-varying modulation effects [9].

In the Ball-within-the-Box model [6] the normal modes of a rectangular room are related to geometrical directions of standing waves, and diffusion is treated as a continuous transfer of energy between harmonic modal series, by means of a single scattering object represented by the feedback matrix of a FDN.

In this paper we propose the use of reverberation units, namely FDNs, as constituents of physics-based sound synthesis models, whenever the goal is that of thickening the distribution of elementary events occurring in mechanical interactions, or to give account of scattering and propagation phenomena.

In fact, reverberation phenomena do not occur only in air at the architectural scale. As it is obviously deduced from the historical success of spring and plate reverb units, vibration in solids can have a clear reverberation character [10].

The textural character of many everyday sounds is indeed determined by dense repetitions of basic acoustic events, which can be assimilated to reverberation in a wide sense. The key for simulating reverberation phenomena is to achieve a high event density, or echo density in reverberation terms. Abel and Huang [11] proposed a robust measure, called normalized echo density (NED), that can be used to characterize reverberant responses. Such measure can reveal the buildup of echoes at various levels of diffusion for a reverberation system. They also showed that NED is a good predictor of texture perception, regardless of the bandwidth of each single echo (or event) [12].

Section 2 recalls the structure of a FDN and illustrates the realization considered in this paper. Section 3 explains how reverberation is used in the context of solid interaction synthesis, namely to differentiate between scraping and rubbing. Section 4 shows how reverberation is used for the simulation of some aerodynamic phenomena. Section 5 points to uses of diffuse reverb for the synthesis of massive liquid sounds.

2. THE CORE COMPONENT

A FDN is described by the following equations:

$$\begin{aligned} y(n) &= \sum_{i=1}^{N} c_i s_i(n) + d x(n) \\ s_i(n+m_i) &= \sum_{j=1}^{N} a_{i,j} s_j(n) + b_i x(n) \end{aligned} \quad (1)$$

where $s_i(n)$, $1 \leq i \leq N$, are the outputs of a set of N delay lines at discrete time n, and x and y are respectively the input and output signal. $a_{i,j}$, b_i, c_i and d are real numbers, acting as weighting and recombining coefficients.

The diffusive behavior of FDN reverberators is determined by the feedback matrix $\mathbf{A} = [a_{i,j}]_{N\times N}$. To ensure that the diffusion process preserves energy, such matrix should be lossless [5]. To speedup convergence towards a Gaussian distribution of echoes, all coefficients of $\mathbf{A}$ should have the same magnitude [4]. A third requirement, especially for large matrices, is efficiency, i.e., the possibility to have sub-quadratic complexity for matrix-vector multiplies. Some lossless, maximally-diffusive, and efficient matrices, such as Hadamard matrices, have been proposed in the literature [4, 5].

In this paper we consider a realization having the three properties of energy preservation, equal-magnitude coefficients, and efficiency, which is based on a circulant feedback matrix defined from a Galois sequence [13]. Circulant matrices afford matrix-vector multiplies in $O(N \log N)$ time by means of FFT or, alternatively, they admit a particularly simple implementation of such multiplies, whose parallelization is straightforward.

The sample-by-sample computation of the reverberator based on a 15×15 circulant matrix [13] can be organized as follows:

```
double SDTReverb_dsp(SDTReverb *x, double in) {
  double a, b, c, d, *s, out;
  int i;

  out = 0.0;

  for (i = 0; i < 15; i++) {
    s = &x->v[i];
    b = s[1] + s[2] + s[3] + s[5] +
        s[6] + s[9] + s[11];
    c = s[0] + s[4] + s[7] + s[8] + s[10] +
        s[12] + s[13] + s[14];
    a = 0.25 * (b - c);
    d = SDTDelay_dsp(x->delays[i], in + a);
    x->v[i] = x->g[i] *
              SDTOnePole_dsp(x->filters[i], d);
    out += x->v[i];
  }
  memcpy(&x->v[15], x->v, 14 * sizeof(double));
  return out / 15.0;
}
```

The main artifice for such a compact code is the juxtaposition of two copies of the outputs of the delay lines (`memcpy` operation), which allows the exploitation of the circulant structure as 15 independent iterations of a `for` loop, which could be efficiently parallelized. The proposed implementation is actually included as a Cycling'74 Max external in the Sound Design Toolkit (SDT) [14], a collection of physics-based sound models for the aural rendering of basic acoustic phenomena, such as contacts between solids, liquid and aerodynamic interactions[1]. `SDTReverb_dsp()` uses the SDT implementation of delay lines and one-pole IIR filters for frequency-dependent damping.

This maximally diffusive yet efficient FDN [13] takes six arguments as input parameters: the size of the room along the x, y and z axes (l_x, l_y, and l_z), a geometric randomness coefficient (between 0 and 1), the global reverberation time and the reverberation time at 1 kHz.

Virtual room dimensions are used to compute the lengths of the delay lines, which play a key role in the system response as they represent the fundamental periods at which the virtual environment resonates. In our implementation, delay times are computed to simulate the bouncing period of stationary plane waves in a rectangular room [6].

Each delay time is the reciprocal of

[1]The Sound Design Toolkit and its source code are available at `https://github.com/skat-vg/sdt`

$$f = \frac{c}{2}\left[\left(\frac{n_x}{l_x}\right)^2 + \left(\frac{n_y}{l_y}\right)^2 + \left(\frac{n_z}{l_z}\right)^2\right]^{\frac{1}{2}}, \tag{2}$$

where c is the speed of sound in the medium, and the triplets $[n_x, n_y, n_z]$ belong to the set

1,0,0	2,1,0	1,1,0	1,2,0	0,1,0
0,2,1	0,1,1	0,1,2	0,0,1	1,0,2
1,0,1	1,1,1	1,2,1	2,1,1	2,0,1.

Perfect rectangular enclosures provide a good reference model to distribute echoes in time. However, a subtle artifact called sweeping echo can arise from rigid geometrical specifications [15]. That is why, in our realization, a randomness coefficient is used to add some irregularity to the delay times, thus simulating slight deviations from a strict rectangular reverberation box.

Delay lines are implemented as follows [16]: A single circular buffer is swept by two allpass interpolated readers, updated and crossfaded at a 16 sample alternation rate. This arrangement allows to handle fractional as well as continuously varying delay times, preserving intonation accuracy and avoiding audible glitches in the feedback loop.

As the FDN matrix is lossless, to achieve a finite reverberation time the recirculating signal is multiplied by an attenuation coefficient g_i before entering the delay lines, yielding an exponential decay. To achieve a -60 dB attenuation ($\frac{1}{1000}$ of the initial amplitude) on a delay line of period τ at a given reverberation time T, the attenuation coefficient is computed as follows:

$$g_i = 10^{\frac{-3\tau}{T}}. \tag{3}$$

Frequency-dependent attenuation is achieved by applying a simple one-pole lowpass filter at the output of each delay line, implemented by the difference equation

$$y(n) = (1 + a)x(n) - ay(n - 1). \tag{4}$$

The cutoff frequencies are computed in order to obtain a filter attenuation of 60 dB at 1 kHz after a given time T'. Remembering to take into account the frequency independent attenuation coefficient, using a sampling period t_s the filter response at 1 kHz must be

$$g_\omega = \frac{10^{\frac{-3\tau}{T'}}}{g_i} = \frac{|1 + a|}{\sqrt{a^2 + 2a\cos(2000\pi t_s) + 1}}. \tag{5}$$

Solving this quadratic equation gives two possible values of a, but only one solution preserves the stability of the lowpass filter ($|a| <= 1$). That value is therefore the desired feedback coefficient.

2.1. Echo buildup time

The FDN multiplies elementary acoustic events by recirculating them through a set of delay lines. The density of "echoes" increases in time, with a speed that depends on delay line lengths which, in turn, depend on the size of the virtual resonator that the FDN is modeling. To quantify and represent the rapidity of echo buildup we use the NED measure [11], which is based on a sliding window and counts the number of impulse response taps which lie outside the standard deviation of the windowed samples, normalized to give 1 for a Gaussian distribution of values. In our implementation of NED we used a 20 ms Hanning window, gradually shrinking to 10 ms at the beginning of the impulse response. Figure 1 shows the buildup of the normalized echo density for a small (0.1 m), a medium (1.0 m), and a large (10.0 m) box. While for small and medium size boxes the buildup of a Gaussian process is practically instantaneous, it is clear how that takes over a hundred milliseconds in the case of boxes at architectural scale. This kind of diffuse reverb inevitably adds spaciousness and depth to any texture it will be applied to. Conversely, a resonator about 1 m in size starts promptly with a high echo density, and stabilizes around a NED value of 1 after about 10 ms. A small box also starts with a high echo density, but the tail of its response tends to ring, and that explains the dip in the dotted line of figure 1.

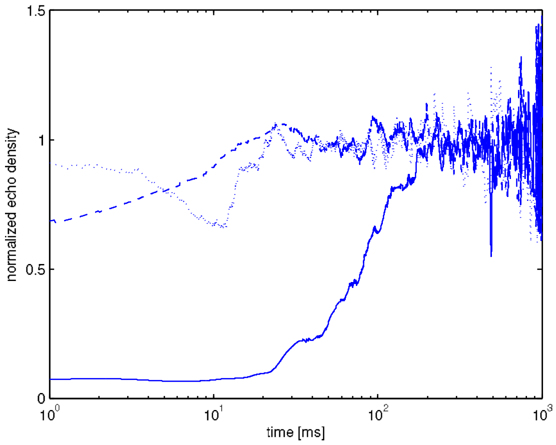

Figure 1: *Normalized echo density buildup for three maximally diffusive FDNs, modeling cubes of size* 0.1 m *(dotted),* 1.0 m *(dashed), and* 10.0 m *(solid), with maximal value for the randomness parameter.*

3. USE 1: SCRAPING/RUBBING

To produce sliding noises, Van Den Doel [17] proposed to bandpass filter a fractal noise, at a central frequency proportional to the velocity between the two surfaces in contact. The rugosity of the surface could be controlled by means of the fractal dimension. Similarly, Conan et al. proposed lowpass filtering a noise, with a biquad cutoff frequency proportional to velocity [18]. Since they showed that the density of impacts is the main discriminant for the perceptual distinction between rubbing and scraping, they proposed using a pulse-based noise generator: Pulses are generated by a Bernoulli process, and their temporal density can be controlled. The amplitude of pulses has a uniform distribution.

To realize a scraping/rubbing sound generator we take a different approach: We consider the rugosity of a surface, imported as an audio file, and we explicitly model an object sliding on it. The thickness of such object determines the number of impact points, the larger the object the higher the number of contacts. However, to increase the temporal density of micro-events further it is necessary to thicken the pattern of impacts, and a maximally-diffusive reverberator is the key component to increase the impact density.

The scraping model belongs to the set of physics-based algorithms available in the SDT. In the scraping model an audio signal is interpreted as surface profile in order to modulate the collisions between a point-mass exciter and a resonating object [19]. The sound model is implemented as Cycling'74 Max patch, shown in Figure 2. An arbitrary audio signal is acquired in order to provide a roughness profile to drive the sound synthesis (e.g., a sawtooth waveform at 50 Hz is displayed in the figure). The signal buffer length is 1000 ms, ideally corresponding to a 1000 mm-long surface. The "sliding parameters" layer is used to interpret the stored surface profile and drive the impact model accordingly. The vertical penetration of the probe sets the threshold level of the roughness profile above which the signal is detected, while the probe width parameter sets the size of the sliding window on the roughness profile (in mm, large = rubber, small = sharp object). The virtual probe is advanced every Δt ms by a distance $\Delta x = v\Delta t$, where v is the sliding velocity in m/s. The "velocity profile" box allows to draw velocity trajectories, that is describing the temporal unfolding of specific gestures, as in sawing, filing, scratching, and so forth. Additional parameters are Δt in ms and the diameter of a single contact area in cm. The sound quality of the single impact is described in the two boxes at the bottom of the GUI displayed in figure 2 (i.e., stiffness, contact shape, energy dissipation affecting the occurrence of bouncing phenomena, and modes of resonance), in order to characterize the scraping/rubbing on the surface profile with auditory impressions of different materials (e.g., metal, wood, plastic, glass, etc.)[2].

Figure 3 shows the spectrograms of the sounds produced by 1-second sliding gestures at constant speed, with a sharp (left, top) or with a wide (right, top) probe. The supporting surface is a sawtooth wave similar to the one shown in figure 2. The wide probe, as compared to the sharp probe, hits the surface asperities at many more points, thus producing a denser distribution of elementary impact noises. This multiplication of acoustic events is similar to early echoes in room reverberation. To increase the event density further we introduce the FDN reverberator, whose effects as a small or large square box (with maximum randomness), and as a damped or reflective enclosure, are also illustrated in figure 3. It is clear that both the enlargement of the probe and the introduction of diffusive reverberation increase the density of micro-impacts, and this adds a degree of freedom for the sound designer. Similarly, in classic reverb design one must decide how to split memory and operations between early reflections (simulated by a FIR structure) and diffuse reverberation (FDN) [20].

Beside increasing the temporal thickening of the sound texture, the introduction of a diffusive reverberator affects other timbral aspects. For this purpose, we conditioned the reverb parameters in order to reduce the occurrence of evident timbral effects, trying to maximize the multiplication of micro-impacts. Based on some informal listening tests, we associated the probe width ($1.0 - 100.0$ mm) with the global reverb time in the range of $0.2 - 3.0$ s, while keeping the reverb time at 1 kHz shorter (a 0.4 factor of the global reverb time). Different values of FDN damping give different impressions of surface material, being it metallic for low damping, or wood-like for high damping. The velocity ($0.0 - 1.0$ m/s) is inversely associated to the room size, in the range of $1.3 - 0.3$ m, thus producing the compression of the room size for high velocity. While poorly justified by the physics, we

[2]Audio examples of various gestures and materials are available at `https://soundcloud.com/skat-vg/sets/dafx2015-audio-examples`

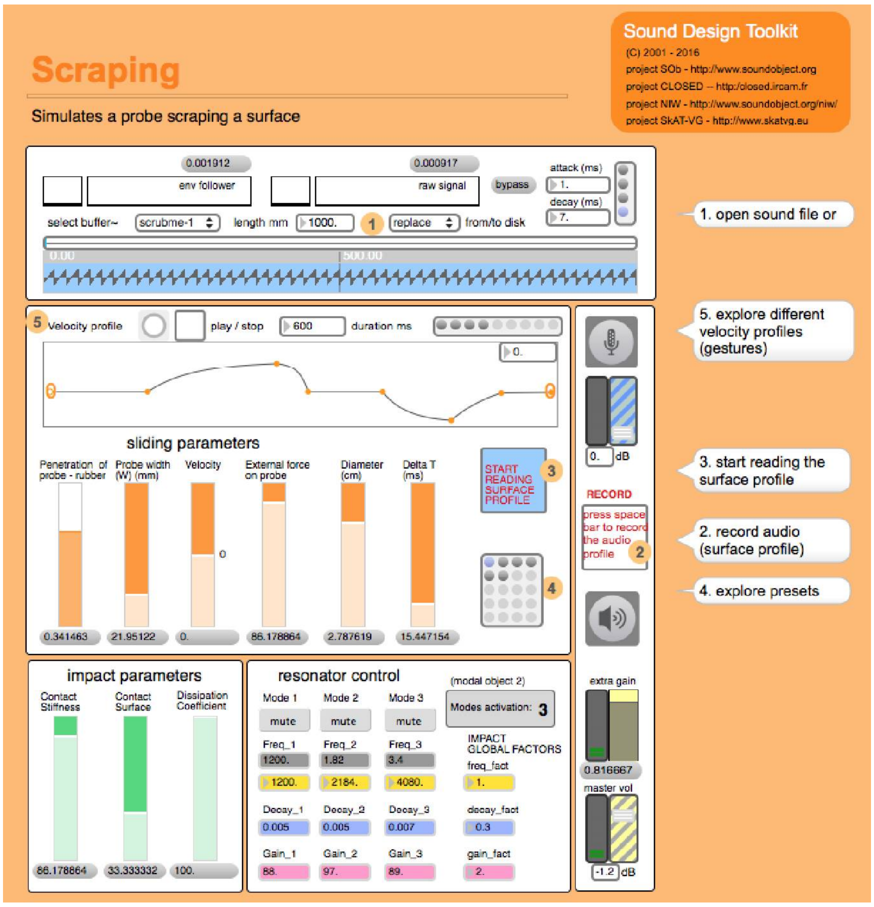

Figure 2: *Cycling'74 Max GUI of the scraping model.*

found this control strategy suitable to emphasize the typical glissando effect which occurs in quick sliding gestures and in back and forth motions.

Playing with the three size parameters of the FDN produces timbral effects, which can give the impression of extension and distance of the surface. For instance a texture of sparse, sharp micro-impacts with a room size of 14 m, a global time reverb of 14 s, and a damping at 5.6 s can easily convey the impression of a heavy rain on roofing iron sheets (see footnote 2).

4. USE 2: EXPLOSIONS

Powerful explosions, as well as objects traveling at supersonic speed such as rifle bullets or cracking whip tails, create shock waves, namely a sudden peak in pressure followed by a negative expansion tail. The algorithms in the Sound Design Toolkit simulate this event with a Friedlander waveform [21], which approximates the pressure change caused by an exploding point source emitting a spherical shock wave. The air then flows back to restore atmospheric pressure, generating a blast wind. This air flow is rendered by bandpass filtered white noise, modulated in amplitude by the Friedlander waveform as the wind intensity follows more or less the profile of the initial shock wave.

Real world explosions, however, are almost never perfectly impulsive. When happening in air, the initial shockwave is likely to generate some chaotic turbulence as it propagates. The blast can also get reflected by the ground or other obstacles, generating Mach stems and other kinds of interference. If the explosion transfers part of its energy to a solid object, such as the ground or the body of a rifle, vibrations propagate also through the solid. If the material is of non-uniform density, (e.g. rocks, gravel or soil), the wave is subject to different propagation speeds, reflections and refractions. Some explosions, then, cannot be approximated by a point source emitting spherical waves: Lightning bolts, for example, generate a simultaneous cylindrical shockwave across their length, and different wave sections interact as they meet because of the bolt tortuosity. These interactions create complex temporal patterns and have a direct effect on the resulting sound [22, 23].

The explosion model implemented in the Sound Design Toolkit exploits our maximally diffusive FDN to simulate scattering, diffusion, interferences and other kinds of interaction caused by the phenomena described above, adding complexity to the initial blast wave and improving the realism of the acoustic result. Figure 4 displays the block diagram of the whole explosion synthesis model.

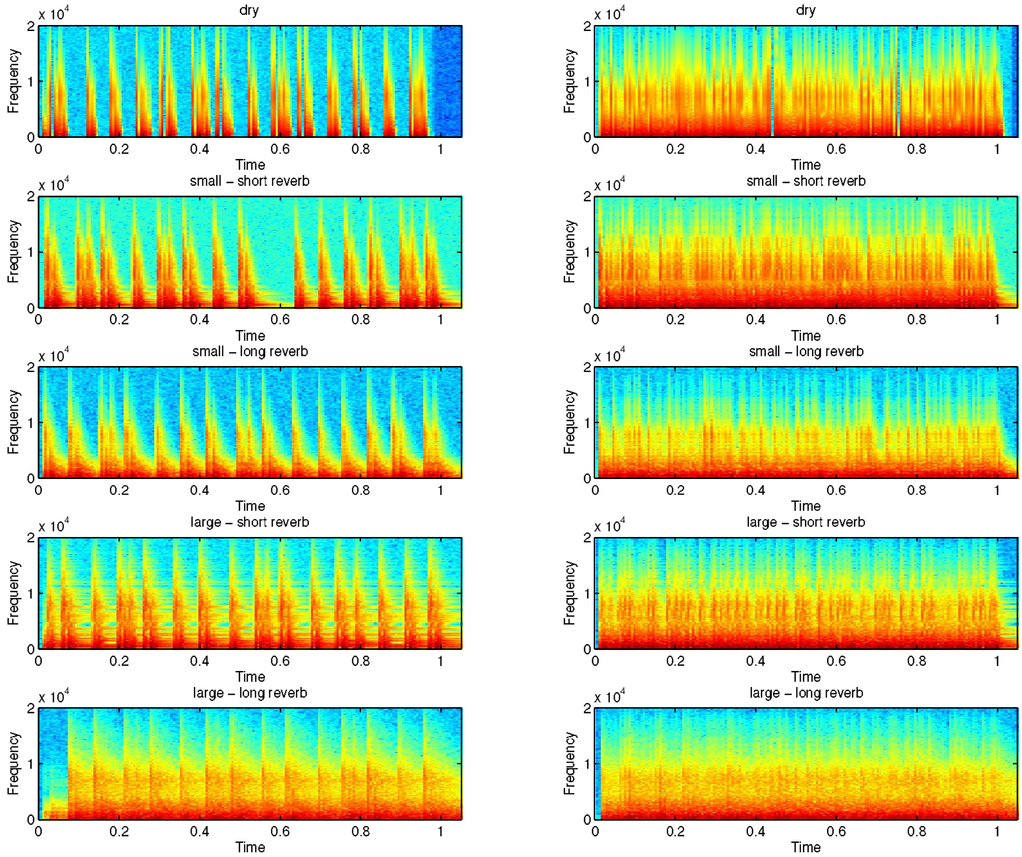

Figure 3: *Spectrograms resulting from the exploration at constant speed of a surface sawtooth texture, with a sharp* (0.002 m, *left) or wide* (0.020 m, *right) probe. Multiplication of impacts is obtained by means of a maximally-diffusive reverberator corresponding to a small* (0.1 m) *or large* (1.0 m) *resonator, with short* (0.2 s, *highly damped) or long* (3.0 s, *slightly damped) reverberation tail.*

The Friedlander waveform is initially scattered by the artificial reverb, and subsequently processed by a lowpass filter. The cutoff frequency of the filter is inversely proportional to the square root of the listening point distance, to simulate frequency dependent attenuation caused by the impedance of the medium in which acoustic waves propagate. The resulting waveform as it is represents the explosion shock wave, but it is also applied as an amplitude envelope to a band limited noise generator to simulate the blast wind. These two outputs are finally sent into two independent delay lines, to model a different propagation velocity for each component.

The reverberation algorithm is applied just after the Friedlander wave generator, with all of its parameters bound to a single free variable, namely the duration of the audible diffusive tail. By empirical trial and error, we have found that acoustically convincing results can be achieved adjusting the length of the delay lines so that an incoming shockwave is recirculated about a hundred times on average before becoming inaudible. For example, a diffusive tail of 1 second would require delay lengths to vary around an average value of 10 milliseconds. Moreover, higher frequencies are slightly damped by setting the reverberation time at 1 kHz to 90% of the whole reverberation time.

5. USE 3: SPRAYS AND SPLASHES

Liquids are involved in a great number of different and very complex sound events, such as dripping, splashing, burbling, pouring, fizzling and so on. However, the main source of these sounds is not the liquid in itself but rather its population of bubbles, trapped by its movement or generated by cavitation phenomena. The pressure of the surrounding liquid mass applies energy on the gas, transforming the bubble in a pulsating oscillator and quickly converting it to a spherical form. Spherical bubbles have a very well known

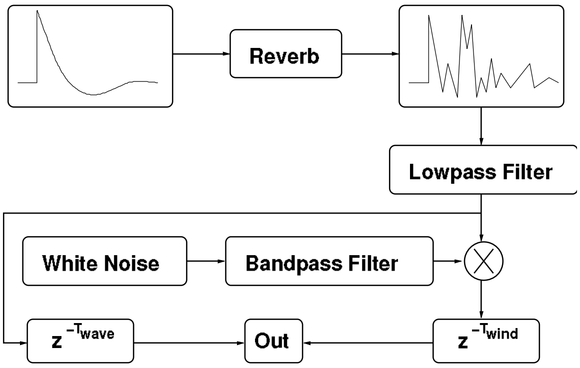

Figure 4: *Block diagram of the explosion sound synthesis model as implemented in the Sound Design Toolkit. The reverberator is cascaded to the Friedlander waveform generator to simulate scattering and diffusive phenomena in the wave propagation.*

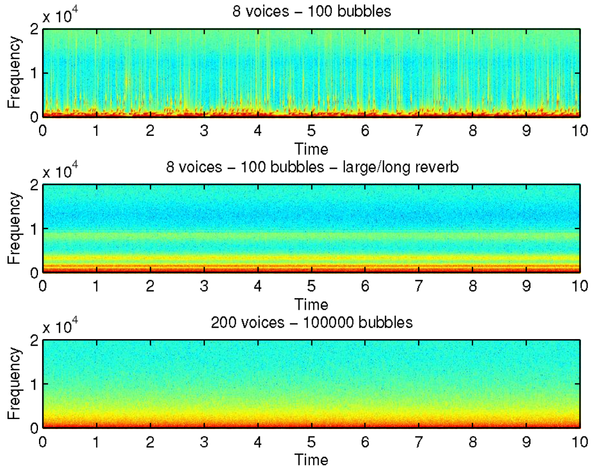

Figure 5: *Spectrograms of the sounds produced by a relatively small number of bubbles, without (top) and with (middle) reverb. The spectrogram at the bottom is obtained by a massive generation of bubbles.*

acoustic behavior [24], allowing to devise a physically informed approach for the synthesis of liquid sounds. A single bubble can be modeled by a simple exponentially decaying sinusoidal oscillator, whose amplitude and frequency envelopes are strictly dependent on the bubble radius and depth. The Sound Design Toolkit renders liquid sounds through additive synthesis, by means of a polyphonic sinusoidal oscillator bank driven by a stochastic model for bubble generation [25].

One of the main drawbacks of this method is that it requires a high degree of polyphony to produce convincing simulations, especially for very dense textures like rainstorms or roaring waterfalls, or bursting events like water splash crowns. If too many bubbles are generated compared to the available number of voices in the oscillator bank, amplitude envelopes get updated so often that the modulated sinusoids become almost stationary, resulting in a ringing rather than bubbling sound. On the other side, increasing the number of voices leads to a greater computational load. Artificial reverberation can be used as a cheaper alternative to generate dense sound events, replicating and diffusing the sound of each single bubble instead of generating more bubbles through actual polyphony.

To demonstrate the use of our FDN reverberation algorithm for this particular purpose, a very dense waterfall sound simulation was produced. The fluid flow synthesis model of the Sound Design Toolkit was configured with the highest level of polyphony achievable in real time on the target machine: 100000 bubbles per second distributed across 200 simultaneous voices (figure 5, bottom). Polyphony and event density was then reduced to 1000 bubbles per second on 32 voices and 100 bubbles per second on 8 voices (figure 5, top), generating sparser textures more similar to a small stream or to a gentle dripping (see footnote 2). Finally, artificial reverberation was applied to the latter sounds, using long decay times (10 seconds or more) and large room sizes (30 meters or more) (figure 5, middle). The use of reverb on the sparse sound events allowed to recover the dense character of the waterfall, although with a noticeable difference in timbre. This is visible in figure 5, where it is clear that reverberation preserves the characteristic formants present in the parsimonious bubble generation.

6. CONCLUSION

Earth, water, air and fire: The sonic textures found in everyday soundscapes can still be largely attributed to the four classic elements. Aristotle related the elements to the two sensible axes "hot – cold" (attributes extensively used by media theorist Marshall McLuhan) and "dry – wet". It may be a coincidence, but the adjectives dry and wet have a very precise meaning in audio technology and practice. They refer to the presence or addition, in a sound signal, of reverberation.

Reverberation is multiplication of audible events and processes, thickening of textures, or texturization of acoustic elements. Such multiplication occurs in rooms at the architectural scale, but it could also occur in physical interactions between solids, in gasses, or in liquids.

In this article we argue about the importance of a good, efficient, and versatile diffusive reverberation model in the toolkit of a sound designer. We found this component to be important to convert scraping to rubbing, to generate convincing thunders and explosions, and to get showers out of single drops. These are all applications that stretch the use and interpretation of artificial reverberation beyond a purely spatial boundary, to embrace sound textures at large.

7. ACKNOWLEDGMENT

The authors are pursuing this research as part of the project SkAT-VG and acknowledge the financial support of the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 618067.

8. REFERENCES

[1] James A Moorer, “About this reverberation business,” *Computer music journal*, vol. 3, no. 2, pp. 13–28, june 1979.

[2] Ville Pulkki, Tapio Lokki, and Davide Rocchesso, *Spatial Effects*, pp. 139–183, John Wiley & Sons, Ltd, 2011, in Digital Audio Effects, Udo Zölzer, ed.

[3] Matti Karjalainen and Hanna Jarvelainen, “More about this reverberation science: Perceptually good late reverberation,” in *Audio Engineering Society Convention 111*, Nov 2001.

[4] Jean-Marc Jot, “Efficient models for reverberation and distance rendering in computer music and virtual audio reality,” in *Proc. 1997 International Computer Music Conference*, Thessaloniki, Greece, 1997, pp. 236–243.

[5] Davide Rocchesso and Julius O. Smith, “Circulant and elliptic feedback delay networks for artificial reverberation,” *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 1, pp. 51–63, Jan 1997.

[6] Davide Rocchesso, “The Ball within the Box: a sound-processing metaphor,” *Computer Music Journal*, vol. 19, no. 4, pp. 47–57, 1995.

[7] Enzo De Sena, Hüseyin Haciihabiboglu, Zoran Cvetkovic, and Julius O. Smith III, “Efficient synthesis of room acoustics via scattering delay networks,” *Audio, Speech, and Language Processing, IEEE/ACM Transactions on*, vol. 23, no. 9, pp. 1478–1492, Sept 2015.

[8] Jean-Marc Jot, “An analysis/synthesis approach to real-time artificial reverberation,” in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Mar 1992, vol. 2, pp. 221–224 vol.2.

[9] Vesa Norilo, “Exploring the vectored time variant comb filter,” in *Proc. Int. Conf. on Digital Audio Effects*, Erlangen, Germany, 2014.

[10] Cynthia Bruyns Maxwell, “Real-time reverb simulation for arbitrary object shapes,” *Proc. Int. Conf. on Digital Audio Effects, Bordeaux, France*, 2007.

[11] Jonathan S Abel and Patty Huang, “A simple, robust measure of reverberation echo density,” in *Audio Engineering Society Convention 121*, 2006.

[12] Patty Huang, Jonathan S Abel, Hiroko Terasawa, and Jonathan Berger, “Reverberation echo density psychoacoustics,” in *Audio Engineering Society Convention 125*, 2008.

[13] Davide Rocchesso, “Maximally diffusive yet efficient feedback delay networks for artificial reverberation,” *IEEE Signal Processing Letters*, vol. 4, no. 9, pp. 252–255, Sept 1997.

[14] Stefano Delle Monache, Pietro Polotti, and Davide Rocchesso, “A toolkit for explorations in sonic interaction design,” in *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound*, Piteå, Sweden, 2010, pp. 1:1–1:7.

[15] E. De Sena, N. Antonello, M. Moonen, and T. van Waterschoot, “On the modeling of rectangular geometries in room acoustic simulations,” *IEEE/ACM Trans. on Audio, Speech, and Language Processing*, vol. 23, no. 4, pp. 774–786, April 2015.

[16] Scott A van Duyne, David A Jaffe, Gregory Pat Scandalis, and Timothy S Stilson, “A lossless, click-free, pithcbendable delay line loop interpolation scheme,” in *Proc. Int. Computer Music Conference*, Thessaloniki, Greece, 1997, pp. 252–255.

[17] Kees van den Doel, Paul G. Kry, and Dinesh K. Pai, “Foleyautomatic: Physically-based sound effects for interactive simulation and animation,” in *Proc. Conf. on Computer Graphics and Interactive Techniques*, New York, NY, USA, 2001, SIGGRAPH, pp. 537–544, ACM.

[18] Simon Conan, Etienne Thoret, Mitsuko Aramaki, Olivier Derrien, Charles Gondre, Richard Kronland-Martinet, and Solvi Ystad, “Navigating in a space of synthesized interaction-sounds: Rubbing, scratching and rolling sounds,” in *Proc. Int. Conference on Digital Audio Effects*, Maynooth, Ireland, 2013, pp. 202–209.

[19] Stefano Papetti, Federico Avanzini, and Davide Rocchesso, “Numerical methods for a nonlinear impact model: a comparative study with closed-form corrections,” *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 19, no. 7, pp. 2146–2158, 2011.

[20] Davide Rocchesso and William Putnam, “A Numerical Investigation of the Representation of Room Transfer Functions for Artificial Reverberation,” in *Proc. XI Colloquium Mus. Inform.*, Bologna, Italy, Nov. 1995, AIMI, pp. 149–152.

[21] Oleg Mazarak, Claude Martins, and John Amanatides, “Animating exploding objects,” in *Proc. Graphics Interface*, 1999, pp. 211–218.

[22] Andy Farnell, *Designing sound*, Mit Press, Cambridge, MA, 2010.

[23] R.D. Hill, “Channel heating in return-stroke lightning,” *Journal of Geophysical Research*, vol. 76, no. 3, pp. 637–645, 1971.

[24] M Minnaert, “On musical air-bubbles and the sounds of running water,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 16, no. 104, pp. 235–248, 1933.

[25] Kees van den Doel, “Physically based models for liquid sounds,” *ACM Trans. on Applied Perception*, vol. 2, no. 4, pp. 534–546, 2005.