PebbleBox and CrumbleBag: Tactile Interfaces for Granular Synthesis

Sile O'Modhrain
Media Lab Europe
Sugar House Lane
Dublin 8, Ireland
sile@media.mit.edu

Georg Essl
Media Lab Europe
Sugar House Lane
Dublin 8, Ireland
georg@mle.media.mit.edu

ABSTRACT

The PebbleBox and the CrumbleBag are examples of a granular interaction paradigm, in which the manipulation of physical grains of arbitrary material becomes the basis for interacting with granular sound synthesis models. The sounds made by the grains as they are manipulated are analysed, and parameters such as grain rate, grain amplitude and grain density are extracted. These parameters are then used to control the granulation of arbitrary sound samples in real time. In this way, a direct link is made between the haptic sensation of interacting with grains and the control of granular sounds.

Keywords

Musical instrument, granular synthesis, haptic

1. INTRODUCTION

Interaction with objects in the world around us is a richly multisensory experience. Casting a pebble into a pond, we both see the ripples resulting from the disturbance of the water's surface and hear the impact of the stone on the water as a disturbance of the air. If we are close enough and the stone is big enough, we might also get wet. Furthermore, the interaction of stone and water makes certain information explicit - the size of the splash is correlated with both the size of the stone and the force with which it was thrown, and the sound it makes provides information about the depth of the water. Thus the physical laws that govern the behaviour of stones falling into water give rise to an event which is perceived via many sensory channels which each encode, in their different ways the complexity of the event. The perceptual system therefore has a number of representations of the event upon which to draw. In this paper, we suggest that it is possible to build a methodology for sound control upon commonalities between the behaviour of physical objects and that of sound objects which share many of their physical properties. In particular, we focus on the technique of granulation, presenting two instances of expressive instruments for live control of granular synthesis.

Granular synthesis has long been an important and widely used compositional technique in Computer Music. Its literature is too extensive to be sufficiently reviewed here — we refer the reader to a recent comprehensive exposition by Curtis Road [21].

Granular synthesis has also been used in live computer music performance including novel interfaces for expressive control of granulated sound. For example, in "The Lobster Quadrille" [24], Dan Trueman used his sensor-augmented violin bow, (the RBow [25]), to play granular models. Additionally a number of controllers related to granular synthesis have been proposed. These include Timothy Opie's Fish [15, 16] Gadd and Fels' MetaMUSE [7], Perry Cook's PhISM and FoleyMat controllers [2, 3] and the MIDI keyboard and laptop based Creatovox by Roads [21]. Cook also proposed a granular approach to Gait synthesis [3] which is also related to other footware controllers [18]. While all of these controllers drive granular synthesis, and have some haptic feel to them, they usually do not retain the haptic component of the granular interaction itself. For example, Cook's PhISM shakers retain the form factor and weight of an acoustic shaker, but the moving particles (pebbles or the like) are removed and replaced by rigidly anchored electronics. Hence the performer does not feel the particle interaction - they feel the coarse haptic experience but not the fine detail. This also holds for Gadd and Fels' MetaMUSE [7] and the RBow [25]. In the case of the Opie and Road's controllers, the control gesture is abstracted from the interaction and neither level is captured directly.

Our interest here is in retaining the haptic features that are relevant for the parametric control of the sound synthesis algorithm. To the best of our knowledge, this goal has not been explicitly stated elsewhere in the literature. While musical devices that have implicit haptic components have been explored elsewhere e.g. the Musical Playpen and Musical Candy Bowl of Weinberg and coworkers [28, 27] which employed spatially distributed accelerometers, these were not used for tight musical coupling or control of event-based granular synthesis.

2. DESIGN GOALS

The overarching goal of our work on haptic controllers for computer-based musical instruments is to uncover instances of

coupling, however loose, between the haptic and auditory senses and to build on these couplings to develop new paradigms for

instrument control. The examples presented here represent a sub-set of such controllers, those based on interactions that are mediated by physical objects,the properties they embody and the manipulation strategies they invoke.

For more details on experimental investigations into the importance of haptic feedback for musical performance see [14]).

Since the current goal was to build a controller that couples the feel and sound of granular events, it was impor-

tant to incorporate into the interface the manipulation of elements that could objectively or subjectively give rise to granular sounds. Two different interaction paradigms were developed, playing with a hand-full of pebbles and crushing a bag of brittle material. Both are somewhat complex environmental events, whose temporal patterns give rise to important perceptual cues [12, 26].

Therefore, there is a need to sense these temporal events . This poses a number of problems. Firstly, given the nature of the sounds of interest, the events are likely to be spatially distributed. Moreover, the sound-producing mechanism may be internal to the objects interacted with crinkling paper,or may be a result of their destruction (for example crushing cornflakes.)

Finally, while the coupling between temporal events as they are perceived by both the haptic and auditory system should be relatively tight, we are interested in leaving other parameters such as dynamics and timbre open for exploration by the performer.

3. CONTROLLER DESIGN

At the heart of our design of haptic controllers for granulated sounds is a recognition that there exist a class of sounds which are produced by our actions on objects in the world. Thus dragging, dropping, scraping and crushing give rise to to correlated touch and sound events [22]. As noted earlier, such events also bear many signatures of other physical characteristics of the materials and actions involved. However, it is possible to imagine a further class of events where the feel of an object and the sound it produces are less strongly correlated - for example, when playing with pebbles in ones hand, the haptic sensation one feels is that of the pebbles against the hand, while the sound of the interaction stems from the colliding of pebbles within the hand. This loose correlation between feel and sound is appropriate for this experience and in its looseness provides an opportunity to decouple the haptic experience from the sound source. This is the opportunity we build on in our granular synthesis controllers. The first example, the Pebble Box, is based on the manipulation of objects in the environment - the manipulation of pebbles in a tray. The second, the Crumble Bag, is based on the manipulation of an ensemble of grains contained in a malleable skin.

3.1 Object Interaction: PebbleBox

The PebbleBox is designed to allow for direct manipulation and ecological behavior of objects in a relatively unconstrained way. See Figure 1.

The design consists of a foam-padded table with an inlaid actively powered microphone (see Figure 2). The purpose of the foam is to eliminate the possibility of objects colliding with the container and to damp the sounds of objects dropped or rolled inside the box. However, interactions and disturbances are still picked up by the embedded microphone. Additionally, the microphone picks up interactions in a limited range above the device, i.e. the interaction of objects held in the hands just above the box.

Typical sounds are the collision of objects with the foam padding and collisions between objects.

Haptic feedback is a result of the direct manipulation of the objects in the PebbleBox. The flexibility of this approach allows for the manipulation of any collection of small objects — we have experimented with polished stones, ball
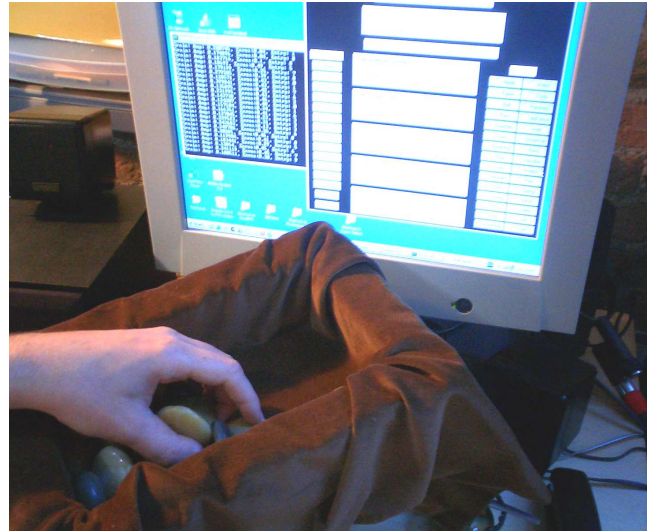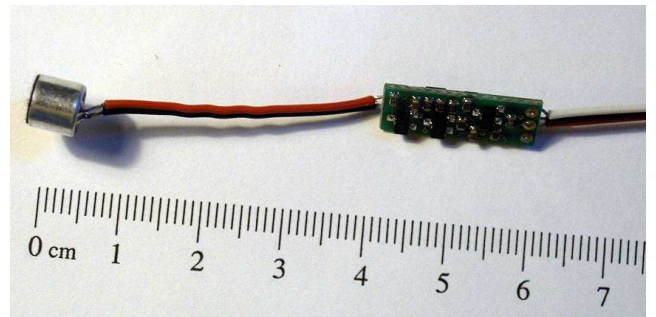

Figure 1: The PebbleBox.

Figure 2: Microphone used for both devices.

Figure 3: The CrumbleBag with cereal, coral and Styrofoam fillings.
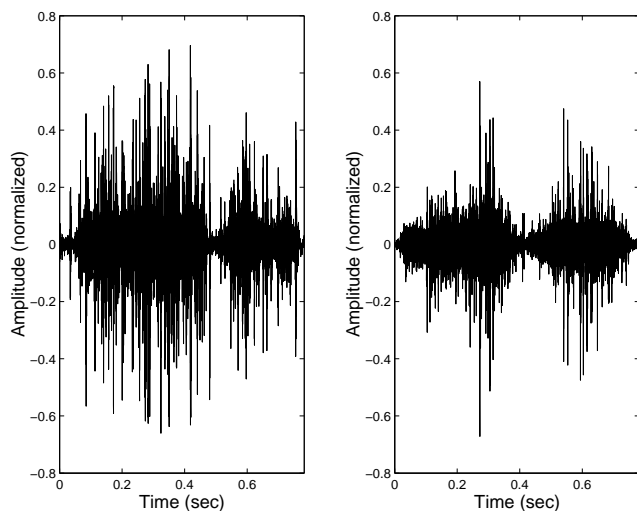

Figure 4: Grabbing of the CrumbleBag with Styrofoam filling in a plastic bag (left) and a cloth bag (right).

bearings and crumbling paper. Each material suggests its own gestures: grabbing, dropping, tapping, crumbling, shuffling, rolling and so forth.

3.2 Grabbing Action: CrumbleBag

The *CrumbleBag* is a flexible bag made of neoprene, into which different granular materials can be placed. The concept, it is derived from the sand-bags used by traditional Foley artists. Through the use of grabbing gestures, the artist articulates foot-steps and the material used in the bag defines the property of the material that is being stepped on (for example corn flakes for leaves and cornstarch for snow [4]).

So far, we have experimented with corn-flakes and ground coral (in plastic and cloth lining bags), Styrofoam beads, and a metallic chain as filling examples, each yielding a very different set of dynamic control parameters (See Figure 3.)

Figure 4 compares the effect of plastic bag versus cloth bag. The recorded instance is one grabbing event. The cloth sound is more muffled whereas the noise created by the plastic adds somewhat to the plastic bag sound, notice
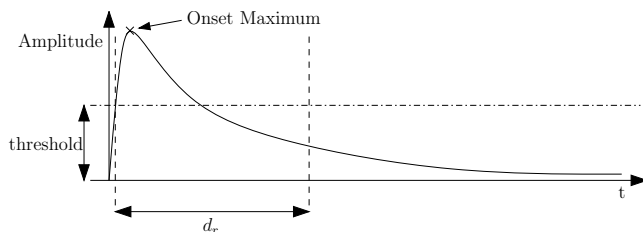

Figure 5: Threshold based grainification scheme. The curve displays an amplitude envelope of an event. d_r is the retrigger delay, preventing detection of new onsets.

however the overall similarity of grain envelope and temporal progression.

Haptic components of the interaction can still be felt through the bag. For example the breaking of cereal or the shifting of coral sand will be felt and the material resistance is maintained.

4. AUDIO-DRIVEN GRANULAR SYNTHESIS

Live audio based sound manipulation is a known concept. It has for instance been used by Jehan, Machover and coworkers[10, 11], though in their case the relationship between audio event and haptic action was not explicitly retained, as the audio was driven by an ensemble mix of traditional acoustical musical instruments as opposed to single instrument granular events.

Granular processing is usually related to what Lippe called "Granular Sampling" [13] but can also be Wavelet inspired processing [21, see for a review]. Neither of these processing paradigms adequately captures the properties we require for intimate interactive control and hence we draw from music, speech and sound retrieval literature for ideas to arrive at practical real-time "granular analysis" algorithms that allow for the grain-level control, that we are looking for.

4.1 Grainification Process

To use the raw audio signal as a driver for granular synthesis, the signal stream needs to be analyzed for granular events. This procedure is somewhat different from granular sampling and we will call it *grainification*. It does, however, relate to event detection as described by Puckette [20].

The parameters that we considered desirable were event detection in the temporal range of perception ($> .1s$), an amplitude measure of a granular event and a measure of spectral content.

The procedure is constrained by the real-time nature of the design goal. Firstly, we are bound by causality and hence any consideration for oncoming data translates into delay. Also the amount of processing is bound by the playback buffer length, which in turn translates into delay.

Given these constraints we employ the following procedure in the current prototype: A very basic onset and retrigger prevention algorithm which also includes a moving short-time average zero-crossing average. The onsets are detected by thresholding followed by a local maximum detection. We do not employ averaging for envelope as we assume that the events have impulsive onsets, hence the first gradient should be expected to lead to a strongest maximum. This
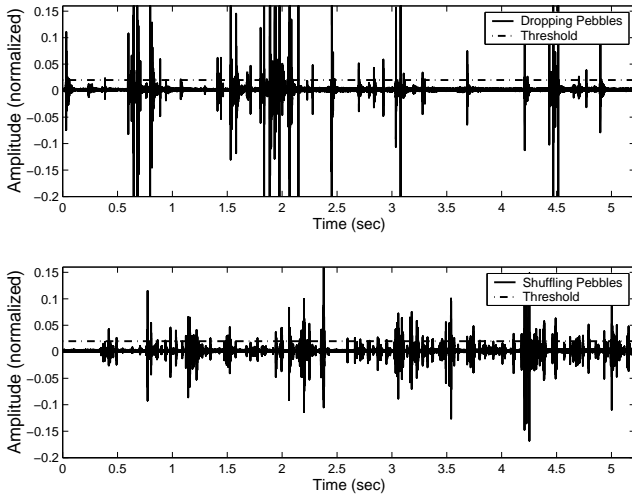
Figure 6: Thresholding of dropping pebbles (top) and pebbles shuffled in one hand (bottom).

amplitude is then used as an immediate measure of grain strength. After a grain event is detected, the detection for further granular events is postponed until a certain time has expired (a so-called retriggering delay d_r). The purpose is to detect only events that lie in the temporal range of perception ($t > 0.05 - 0.1s$ or alternatively $f < 10 - 20Hz$). The second purpose of this procedure is to avoid spurious retriggering by the decaying oscillation of the detected grain , waiting until the grain has decayed below the detection threshold. Hence the inherent signal assumptions are rapid onset, decaying envelope events, where the decay is of the order of the retriggering delay d_r or faster. For this reason this procedure would not be meaningful for the class of sustained sounds which would be inherently less suited to the type of temporal pattern that we are trying to extract. The relationship of thresholding and retrigger delay to a grain amplitude envelope can be seen in Figure 5. The final measure we employ is moving average zero-crossing count. Over a short-time moving window, the number of zero crossings are calculated. This value is used as a spectral measure. The number of zero-crossing is bound from below by the lowest present frequency in a signal [5] and has a correspondence overall with the dominant spectral content of a signal (i.e. the spectral centroid) [19, 17].

We found that despite these assumptions and the simplicity of implementation of this procedure reliable grain detection and believable control is achieved and hence more advanced methods were not concerned. Figure 6 shows two audio signals as detected, including the threshold. The first signal shows pebbles being dropped into the PebbleBox and the second displays a handful of pebbles being shuffled in the player's hand above the pebble box. As can be seen the dropping are more distinctly temporally separated events, whereas the shuffling creates a denser pattern. As can be seen the impulsive assumption of the signal as well observed, and grains are well-separated from background noise.

The real-time implementation is based on STK's real-time audio duplexing. We found an input and output buffer size of 128 to work without clicks or missed buffers. This buffer size, at $22050Hz$ corresponds to a basic delay of $11.6ms$.

Typically grain estimation windows of 100 samples were used leading to a total delay of around $16.1ms$. Performance measures are taken on a $1.6GHz$ Pentium 4 PC running Windows XP with 256 MB ram and a SoundMAX Integrated Digital Audio device by Analog Devices.

5. EXAMPLES AND APPLICATIONS

To test the controller in a real application, the extracted data needs to be mapped to sound generation mechanisms. This is the *mapping problem*, which has seen both theoretical and experimental advances [8; 9; 23, for example].

In principle the sensed data can be mapped arbitrarily. Here we consider the application of our controller design to two types of granular synthesis. The first is based on recorded dictionaries of environmental sounds. The second uses parametric physically informed models developed by Perry Cook [1, 2, 3].

5.1 Recorded Environmental Sound Grains

We implemented a prototype grain dictionary based on recordings of natural sounds. 30 grains were explored using between one and 12 recordings of comparable events. More recordings were used when similar interactions led to different sonic experiences, as for example water splashing or the buckling of a can, or where the detail of the interaction is hard to control and hence leads to variation as in the case of walking, or the shuffling of coins.

The grains are played back based on the granular parameters in the grainification process. The onset time triggers a variable playback event with the playback amplitude defined by the grain onset amplitude. The playback rate, as a measure of the grains overall frequency, was varied with the average zero crossing at the instance of onset. In the absence of the last procedure, the sound is repetitive and multiple entries in the dictionary of similar grain instances are necessary. Three grains are found to be still too likely to have consecutive instances of equal sound events, whereas this was improved with 8 grains. In the presence of variable frequency the monotonous appearance of the sound disappears even for only one recorded grain. In the case of multiple grain recordings for one grain event in the dictionary, a particular instance is chosen at random. The relationship between recorded collision sounds and final sound using a Hammer grain using the PebbleBox can be seen in Figure 7.

5.2 Physically Informed Parametric Models

In order to explore parametric models, we used Perry Cook's shaker based granular synthesis as implemented in his STK software [3] (see the left button row in Figure 8).

Here the mapping of grain onset time and amplitude relates to time and amount of energy infused into the physically inspired model. The zero-crossing average is mapped to the center resonance frequency of the models. These models have inherent stochastic variability. Also some do respond more immediately to energy infusion than others. This does affect the perception of playability, and in general a strong correlation of energy infusion to granular events is desirable. For details on the parametric model synthesis we refer the reader to [1, 2, 3].
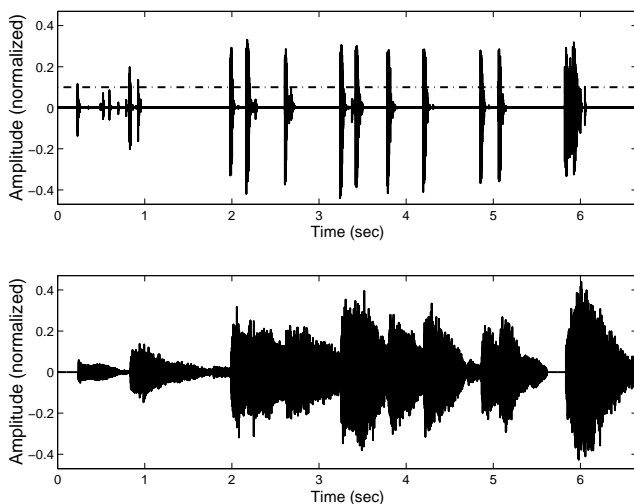
Figure 7: Recorded signal of the PebbleBox (top) and granulated response using a Hammer grain (bottom) of the complete granulation process.
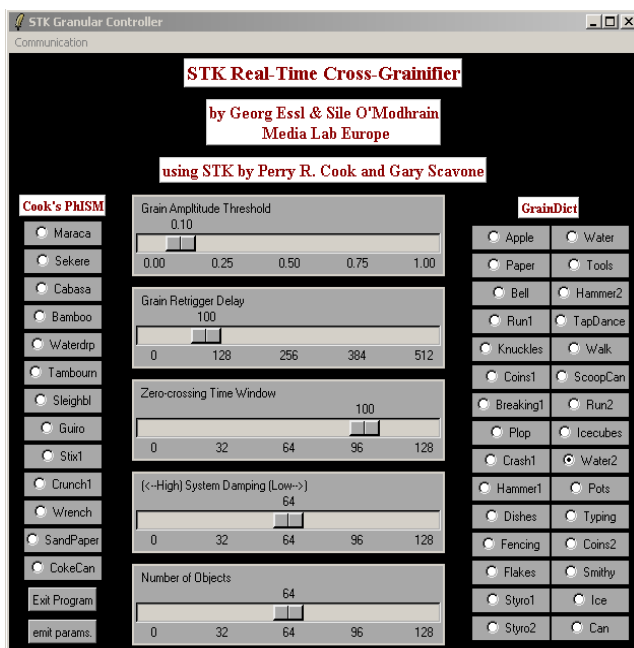

Figure 8: The interface of the Grainification and Synthesis Application GUI implemented in STK.

6. CONCLUSION

The advantage of the proposed design is its simplicity and low cost, and the flexibility which supports the exchange of interaction materials for varied haptic performance experiences and spectral control over the granular synthesis process. Drawbacks include the possibility of environmental noise that could interfere with the performance. In practice, however, we found that ongoing background conversation and extraneous sounds in a large shared office space did not affect the performance of the device unless the speaker were in the immediate vicinity of the device. Another drawback of the current design is the lack of spatial information, a drawback that might be overcome by multiple channel microphone recordings at various positions inside the PebbleBox or the CrumbleBag. Finally only basic granular features are currently extracted and additional degrees of freedom may be desirable.

To our knowledge, this is the first controller for granular synthesis which maintains the individual mapping of fine-structure temporal behavior of the order of $10 - 100ms$ of granular event to haptic interactions, while not having the haptic interaction abstracted from the expected sounding mechanism. In the particular implementation of onset grains, the mapping is flexible, but remains intuitive for sounds that have comparable temporal patterns to the recorded sounds. The haptic feel of the instrument is modular and can be adjusted by exchanging interacting objects, in the case of the PebbleBox, and by varying filling materials, in the case of the CrumbleBag. In this way a variety of environmentally based sounds, like dropping or shuffling of objects can be performed. Also imitated gestures such as walking can be controlled by enacting the characteristic time patterns of sounds. Furthermore, the opportunity exists to create new abstract sounds by imposing unconventional temporal patterns, which don't mimic the behaviour of these environmental sounds.

In summary, we have presented an environmentally based haptic controller which maintains the feel of physical granular processes and allows for real-time performance of granular synthesis methods.

7. REFERENCES

[1] P. R. Cook. Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal*, 21(3):38–49, 1997.

[2] P. R. Cook. Toward Physically-Informed Parametric Synthesis of Sound Effects. In *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA-99)*, pages 1–5, New Paltz, NY, October 17-20 1999.

[3] P. R. Cook. *Real Sound Synthesis for Interactive Applications*, chapter Chapter 13, pages 149–168. A K Peters, Ltd., 2002.

[4] S. Ebersole. CSU-Pueblo Television Production Handbook. Retrieved online on January 24, 2004 at `http://faculty.colostate-pueblo.edu/samuel.ebersole/handbook/sfx.html`.

[5] A. Eremenko and D. Novikov. Oscillation of Fourier Integrals with a spectral gap. Retrieved online on January 20, 2004 at `http://www.math.purdue.edu/~eremenko/dvi/novik1011.pdf`, May 30 2003.

[6] Federation of American Scientists. Tempest. Retrieved online on January 26, 2004 at `http://www.fas.org/irp/program/security/tempest.htm`.

[7] A. Gadd and S. Fels. MetaMuse: A Novel Control Metaphor for Granular Synthesis. In *Proceedings ACM Conference on Computer Human Interaction*. SigCHI, ACM, 2002.

[8] A. Hunt, M. M. Wanderley, and M. Paradis. The importance of parameter mapping in electronic instrument design. In *Proceedings of the 2002 Conference on New Instruments for Musical Expression (NIME-02)*, pages 149–154, Dublin, Ireland, May 24-26 2002.

[9] A. Hunt, M. M. Wanderley, and K. R. Towards a Model for Instrumental Mapping in Expert Musical Interaction. In *Proceedings of the International Computer Music Conference (ICMC-00)*, pages 209–212, Berlin, Germany, August 27-September 1 2000.

[10] T. Jehan, T. Machover, and M. Fabio. Sparkler: An audio-driven interactive live computer performance for symphony orchestra. In *Proceedings of the International Computer Music Conference*, Göteborg, Sweden, September 16-21 2002.

[11] T. Jehan and B. Schoner. An audio-driven, spectral analysis-based, perceptual synthesis engine. In *Proceedings of the 110th Convention of the Audio Engineering Society*, Amsterdam, Netherlands, 2001. Audio Engineering Society.

[12] D. Keller and J. Berger. Everyday sounds: synthesis parameters and perceptual correlates. In *Proceedings of the VIII Brazilian Symposium of Computer Music*, Fortaleza, Brazila, 2001.

[13] C. Lippe. Real-time granular sampling using the ircam signal processing workstation. *Contemporary Music Review*, 10:149–155, 1994.

[14] S. O'Modhrain. *Playing by Feel: Incorporating Haptic Feedback into Computer-Based Musical Instruments*. PhD thesis, Stanford University, Palo Alto, California, November 2000.

[15] T. Opie. Granular synthesis: Experiments in live performance. In *Proceedings of the Australasian Computer Music Conference 2002*, pages 97–102, Fitzroy, Australia, 2002.

[16] T. Opie. Creation of a real-time granular synthesis instrument for live performance. Master's thesis, QUT, Brisbane, Australia, 2003.

[17] C. Panagiotakis and G. Tziritas. A Speech/Music Discriminator Based on RMS and Zero-Crossings. Retrieved online on January 20, 2004 at `http://www.csd.uch.gr/~tziritas/papers/sound_seg_class_00.pdf`, 2004.

[18] J. Paradiso. Electronic Music Interfaces: New Ways to Play. *IEEE Spectrum Magazine*, 34(12):18–30, December 1997. Available online at `http://www.spectrum.ieee.org/select/1297/muse.html`.

[19] G. Peeters and X. Rodet. Automatically selecting signal descriptors for Sound Classification. In *Proceedings of the International Computer Music Conference*, pages 455–458, Göteborg, Sweden, September 16-21 2002.

[20] M. S. Puckette. Theory and Techniques of Electronic Music. Retrieved online on January 22, 2004 at `http://crca.ucsd.edu/~msp/techniques/latest/book.pdf`, December 8 2003.

[21] C. Roads. *Microsound*. MIT Press, Cambridge, Massachusetts, 2001.

[22] D. Rocchesso and F. Fontana, editors. *The Sounding Object*. PHASAR Srl, Florence, Italy, 2003. Available online at `http://www.soundobject.org/SObBook/SObBook_JUL03.pdf`.

[23] J. B. Rovan, M. M. Wanderley, S. Dubnov, and P. Depalle. Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. In *Proceedings of Kansei - The Technology of Emotion Workshop*, Genova, Italy, October 3-4 1997. Available online at `http://www.ircam.fr/equipes/analyse-synthese/wanderle/Gestes/Externe/kansei_final.pdf`.

[24] D. Trueman. The lobster quadrille. available online at `http://www.xs4all.nl/~mwais/confusioninfo.html`, 1999.

[25] D. Trueman and P. R. Cook. BoSSA: The Deconstructed Violin Reconstructed. In *Proceedings of the International Computer Music Conference*, pages 232–239, Beijing, China, October 22-27 1999.

[26] W. H. Warren and R. R. Verbrugge. Auditory Perception of Breaking and Bouncing Events: A Case Study in Ecological Acoustics. *Journal of Experimental Psychology*, 10(5):704–712, 1984.

[27] G. Weinberg. Expressive Digital Musical Instruments For Children. Master's thesis, MIT Media Laboratory, Cambridge, MA, 1999.

[28] G. Weinberg. The Musical Playpen: An Immersive Digital Musical Instrument. *Personal Technologies*, 3:132–136, 1999.