Auditory feedback through continuous control of crumpling sound synthesis

Roberto Bresin
Department of Speech, Music and Hearing
KTH, Stockholm, Sweden
roberto@kth.se

Stefano Delle Monache, Federico Fontana, Stefano Papetti, Pietro Polotti
Dipartimento di Informatica
University of Verona, Italy
{SURNAME}@sci.univr.it

Yon Visell
Centre for Intelligent Machines and CIRMMT
McGill University, Montreal, Canada
yon@cim.mcgill.ca

ABSTRACT
A realtime model for the synthesis of crumpling sounds is presented. By capturing the statistics of short sonic transients which give rise to crackling noise, it allows for a consistent description of a broad spectrum of audible physical processes which emerge in several everyday interaction contexts. The model drives a nonlinear impactor that sonifies every transient, and it can be parameterized depending on the physical attributes of the crumpling material. Three different scenarios are described, respectively simulating the foot interaction with aggregate ground materials, augmenting a dining scenario, and affecting the emotional content of a footstep sequence. Taken altogether, they emphasize the potential generalizability of the model to situations in which a precise control of auditory feedback can significantly increase the enactivity and ecological validity of an interface.

Author Keywords
Interaction design, auditory feedback, walking interfaces, crumpling sound.

ACM Classification Keywords
H.5.5 Information Interfaces and Presentation: Sound and Music Computing—*Modeling*

INTRODUCTION
It is a common everyday experience to hear sounds that, despite conveying information about a well-defined event, are formed by a sequence of short sonic transients. Despite a strong integration of such sounds at the cognitive level, they are clearly perceived to be composed of shorter, semantically indistinguishable units which, taken together, concur to define the event. It is the ecological experience of the physical world that enables listeners to attach individual labels to the composite sounds. We will refer to them as *crumpling sounds*.

Crumpling sounds are generated by a broad spectrum of physical phenomena, many of those manifesting themselves in everyday contexts: a fire burning, a cob of corn kernels exposed to the heat of a flame, a crunched piece of paper, a can crushing, a moka that spills the coffee out of the kettle. All these phenomena belie a common temporal process originating with the transition toward a minimum-energy configuration of an ensemble of microscopic systems, by way of a set of dynamic events whose energy and transition time depend on the individual characteristics of each system and the amount of power it absorbs while changing configuration. Such a process cannot be appreciated at the microscopic level by our auditory system. Conversely, listeners can discriminate well if a fire is extinguishing or revitalizing, if popcorn is ready or needs more heating to cook well, and if a piece of paper or a can have been crunched gently or, instead by means of a strong action. In this sense, crumpling sounds are informative of the macroscopic state of a composite system, furthermore they disclose dynamic aspects which depend on the macroscopic properties of the system, as well as on the overall energy that is being dissipated. Interestingly, physics provides an exact formalization of *particle systems*, which share similarities with systems producing crumpling sounds concerning their damping and acoustic transmission properties [18].

The idea of sonifying sequences of short transients to provide information about the event causing them initially resulted in feed-forward paradigms [12]. The audification of earthquakes and underground explosions through the auditory display of seismic signals was probably the first attempt to infer qualitative information about physical phenomena by means of sound [8]. Such signals produce crumpling sounds once they are transposed onto an audible time scale [5]. More recently, the synthesis of such sounds has taken on a role in the realm of enactive interfaces, thanks to the development of physical model-based algorithms for the parametric synthesis of so-called "particle" sounds. Such techniques have been used to control and inform the generation in real time of hand clapping and walking sounds [4, 17]. On the side of artistic production, the use of such sounds dates back to granular synthesis [11] and has inspired the development of physically-informed virtual musical instruments, such as maracas [3].

The research presented here addresses the problem of designing a real time parametric sound generation model that is applicable to diverse multimodal interaction scenarios where crumpling can augment the informational value of the auditory feedback. The generality of the approach is made possible by two assumptions that strengthening practicality of their use for rendering the perceived acoustical properties of a wide range of ecological as well as non-ecological [5] phenomena: *i*) the intensity E of the short transients that

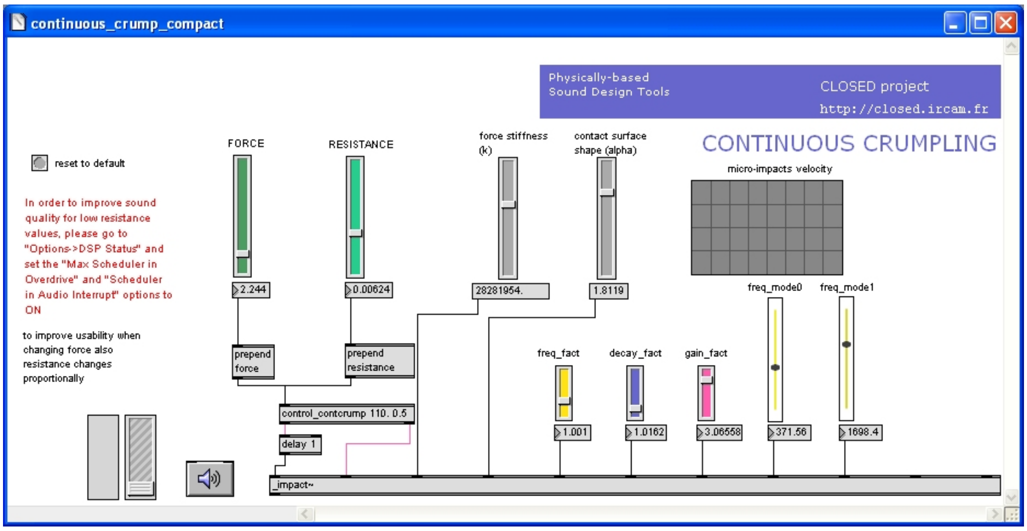

Figure 1. A patch for Max/MSP from the Sound Design Toolbox, implementing the crumpling model.

are produced by phenomena giving rise to crackling noise follows the known stochastic *power law*, whose probability $p(E) = E^{\gamma}$ characterizes different types of noise depending on the value taken by γ (for instance, in the case of crumpling paper it is $-1.6 < \gamma < -1.3$) [20]; *ii*) the temporal sequences of transients can be modeled by stationary *Poisson processes* provided their arriving times are assumed not to depend on each other [16]. Besides their mathematical foundation, whose understanding requires some facility with physics and probability theory, both assumptions are quite general and translate into modeling constraints that can be easily determined. Indeed, power laws as well as Poisson distributions with independent arrival times are involved in many fundamental processes in physics [19]. With some speculation, the generality of such processes allows, to think of crumpling in phenomenological terms just as with other ecological sounds, whose perception is inherently anchored to the type of sound source they evoke [9].

Even if the two assumptions above determine the dynamics of a process, the sound generation model under design is still left with a crucial degree of freedom, given by the specification of transient characteristics. This specification in fact provides the model with information on the physical attributes of the system in which the process takes place. In this paper we will show that by adopting a general form of *nonlinear impactor* to prototype the transient, the model can accommodate different interaction contexts via a proper tuning of the impacting parameters [1].

After a short introduction to the model, we present three example applications involving high level control of crumpling sounds for the simulation of different everyday sonic experiences. While a rigorous experimental validation of the interaction contexts presented in this work has still to be made, the judgments that have been obtained from users during informal testing of the ecological validity, acceptability, and engagement of such contexts are promising. In this sense, the prototype applications described in what follows address a key passage toward a design process for multimodal interfaces displaying crumpling sounds, and they foreshadow for such sounds a significant role as components of closed-loop interaction paradigms providing contextual auditory feedback through their continuous control.

CONTINUOUS CONTROL OF CRUMPLING SOUNDS

This research stems from previous work [7], in which a preliminary model for the synthesis of crumpling sounds was proposed, tailored to the rendering of can crushing and walking, or running, events. It relied on a fairly sophisticate control strategy for the distribution of an initial amount of (potential) energy across the temporal process, and operated *i*) at the macroscopic level, by determining the arrival time and intensity of the crumpling process respectively using power law and Poisson statistics, and well as *ii*) at the microscopic level, by driving the physical parameters of mass, elasticity and dissipation of both the impacting and the impacted bodies [1].

That model did not permit to control the evolution of an event continuously. The limitations of such a control strategy become clear as soon as we consider closed-loop interactions, whose feedback can not only be triggered by an event when it starts, but conversely must be informed by every instantaneous change concerning the evolution of the event. Recalling the examples mentioned above, consider the continuously changing force of a hand that crushes a sheet of paper, or a walking task that takes place on floors comprising several types of ground materials in such a way that different surfaces interact with the foot during a single step. Motivated by such ideas, we have designed a more flexible sound synthesis model, that still relies on the same macroscopic rules (i.e. transients ruled by power law plus Poisson arrival times) and microscopic prototype (i.e. nonlinear impact), meanwhile being continuously controllable.

In the new model, synthesis of crumpling sounds is driven by two parameters: *force*, which sets the instantaneous power that enters the overall system, and *resistance*, which determines how resistant the microscopic systems are, on average, to the applied force before their state changes; The higher the resistance, the smaller the average number of crumpling events and, consequently, the stronger their average intensity in front of an exerted force. Force and resistance are respectively mapped onto the power law γ parameter and the *density* of the stochastic temporal process via simple algebraic relationships. Concerning the physical attributes of the system, parameters of *center frequency*, *decay time*, *stiffness*, and *contacting surface* of the impacting bodies are exposed to control. These have been demonstrated to play a significant perceptual role in the auditory discrimination of colliding bodies [10].

Based on these design specifications, two *external* objects have been included in the Sound Design Tools (SDT) library,[1] synthesizing respectively the crumpling events and the interaction between colliding bodies. These externals have been realized in C++ upon the flext library[2], to ensure cross-compilation under Windows and MacOS and compatibility with the Max/MSP[3] real time environment as well as the PureData open platform[4]. The example patch, shown in Figure 1, refers to the `control_contcrump` external, which computes the intensity and arrival time of every transient, and the `_impact~` abstraction, which contains the external computing the collisions. Furthermore the patch exposes the controls that drive the sound synthesis.

CRUMPLING UNDERFOOT: THE ECOTILE PROJECT

A floor component called the *EcoTile* has been created with the aim of supplying interactive audio-haptic simulations of ecological ground materials in walking, based on the crumpling model. Walking is an enactive activity which is continuously negotiated in direct contact with its surroundings, dependent on the mutual morphology and material properties of both. Informative ground materials for walkers have long played a role in the design of urban environments. Passive haptic indicators are commonly used to signify important locations including stairways, crosswalks or subway platforms, and outdoor paths are designed to be readily distinguished by their material properties. However, the possibility of conveying these kinds of information through active devices embedded in real spaces in which people walk remains basically unknown. Devices that have been developed for interactive simulation in walking have been typically confined to laboratories or other closed environments.

The EcoTile project aims at providing dynamic control over the ecological information that is available in walking, allowing to interactively shape the perceived auditory and haptic material qualities of the ground. The project aims to realize an architectural interface for enaction in walking, as opposed to a wearable interface, such as a shoe. One advantage is that an interactive tile can be utilized for the design of diverse spaces, without requiring users to don special equipment to experience it.

Several research groups have developed instrumented floors for the sensing of walking, dancing, or running movements. For an overview, see chapter 2 of the book by Wanderley and Miranda [14]. The PhOLIEMat developed by Cook [4] is somewhat unique as a floor-based device for the control of synthesized walking sounds via the feet (evoking the work of the Foley artist in film). To our knowledge no project has yet addressed the design of a floor for interactive simulation of surface material properties in walking, whether through the auditory or haptic channel.

Device Design and Characterization

The prototype EcoTile (Figure 2) was designed to explore the capability of an actuated but otherwise rigid interactive platform to interactively simulate and convey the experience of walking on various materials. This aim was partly motivated by the notion that the auditory information alone is capable of conveying considerable information about the properties of surfaces that are walked upon. As a result, an even modestly successful haptic stimulus that is correlated with such an auditory signal may be enough for a convincing percept.

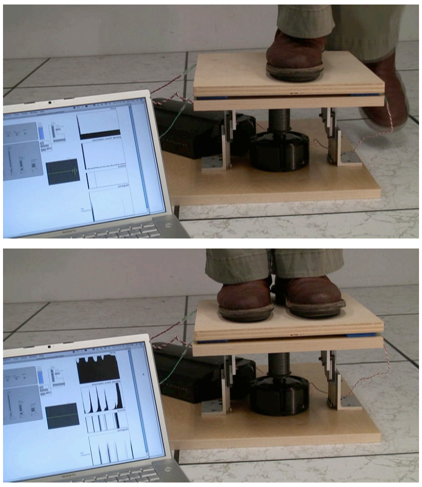

Figure 2. The interactive EcoTile with subject walking on it (top and bottom), next to the software component running on a laptop computer. The display of the latter shows the force, event density, and a sonogram of the resulting audio/haptic signal.

In the prototype, the user walks on or otherwise interacts via the feet with the tile, which interactively delivers a signal designed to mimic the sensation provided by a given surface type. The stimulus consists of a vibration transmitted to the foot and simultaneously an acoustic signal transmitted to the ear, but originating near the platform. For this prototype, a snowy surface was selected as phenomenologically interesting and as a material that might be readily identified by the user through interaction[5]. The device is pictured in the

[1]Sound Design Tools. Publicly available at for testing and integration in sound design projects [2].

[2]`http://grrrr.org/ext/flext/`

[3]`www.cycling74.com`

[4]`www.puredata.org`

[5]Audiovisual documentation can be viewed at `http://cim.mcgill.ca/~yon/HS-Video`; naturally, it does not capture

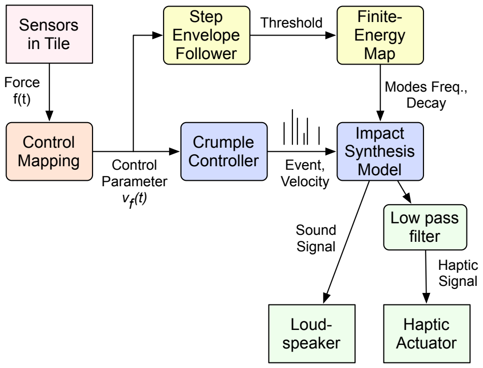

Figure 3. Diagram of the footstep instrument, with force mapping (orange), SDT crumpling model (blue), and finite energy map (yellow).

images of Figure 2. The physical tile is a 34 cm by 34 cm vertically actuated platform. It uses a D-Box motion control actuator and four Interlink force sensing resistors, with accompanying electronics. The simulation is run on a single computer in the software environment Max/MSP. Sound is reproduced by a powered loudspeaker located at the base of the tile.

Physically Based Auditory and Haptic Synthesis

Interactive auditory and haptic stimuli in EcoTile are generated continuously in response to users' footsteps. They are synthesized in real time by a simulation of crumpling snow underfoot, in a way that is driven by the force data from the footstep. An overview of the approach is presented in Figure 3. The method utilizes the crumpling and impact model components of the SDT described above. In order to simulate the decay of the initial energy provided by the footstep, the modal frequencies and decay times were controlled based on the finite energy model described in [7], with the energy level controlled by a simple two-state footstep envelope follower. Parameters of the synthesis model were tuned by hand to approximate walking on moderately well-packed snow.

Control Parameter Mapping

The net force signal $f(t)$ provided by the EcoTile sensors was not utilized directly, but a control parameter v_f was mapped onto the crumpling model's density parameter, via:

$$v_f = \begin{cases} |df/dt| & \text{if} \quad df/dt < 0 \\ 0 & \text{else} \end{cases} \tag{1}$$

This choice represents a practical simplification in which crumpling occurs during walking only as weight is being transferred from the foot onto the tile, ignoring sensations produced as weight merely shifts from one part of the sole to another. The variation in force on the tile generated by weight shifts nonetheless contributes a qualitatively appropriate crumpling, due to the fact that there is some modulation of the overall force on the tile as shifting occurs.

EcoTile: Qualitative Assessment and Future Work

the level of haptic interactivity.

In the prototype, auditory and haptic stimuli are derived from the same instance of the crumpling model. Due to actuator limitations, the haptic signal lacks frequency components above 100 Hz, whereas the human haptic channel is sensitive to those up to 1000s of Hz. Based on feedback from users, the auditory signal with the limited haptic signal appeared sufficient to convey the impression of snow. Ten users were asked to explore interaction with the device by stepping onto it, in stride or in isolation, with one or both feet. Several expressed surprise at the level of veridicality of the experience of the virtual snow. We felt this could be attributed to the synchronization between feedback channels, to the realism of the sound signal, and to the unexpected nature of interacting with virtual snow in a laboratory.

Clearly a more systematic assessment is needed. While the device was designed to aid the fusion of haptic and auditory signals into a single multisensory percept, the level of psychophysical consideration that has gone into the design has so far been limited. Nonetheless, the degree of fusion achieved seems remarkable. As noted above, the synthesis parameters were tuned by hand to approximate well-packed snow. This method is unsatisfactory to the extent that it requires an intuitive exploration of the large parameter space.

A larger and more modular floor is being designed and built, based on a lower cost, wider bandwidth vibrotactile actuator. In tandem with the design revision, a set of experiments is being undertaken by colleagues in psychology, to assess the salience of each modality of the features of the corresponding signals in contributing to the perception of material properties in walking. In addition, synchronized force, acoustic, and vibrotactile recordings are being collected from walking on diverse surfaces. This data will be used to refine the control parameter mapping used with the synthesis model, and may allow the tile application to feed information back to the synthesis algorithm design.

CRUMPLING SOUNDS IN A DINING SCENARIO

In the context of the EU *CLOSED* project, a particular interest is presently devoted to sonic interaction scenarios related to the kitchen and dining [2]. These scenarios have an obvious strong semantic value and provide convenient setting in which to evaluate functional, aesthetic and emotional aspects of sound in the spirit of the CLOSED project. Recently a specific workbench has been developed and tested, namely a dining table named Gamelunch [15]. The *Gamelunch* (the name is the fusion of the Balinese "Gamelan" orchestra and the word "lunch") is the place where a set of experiences have been realized and will be tested, aiming at evaluating the role of the auditory aspects in the dining context (Figure 4). No formal and rigorous evaluation experiment has been performed yet.

Nevertheless, the creation of a virtual (continuous) sonic feedback, augmenting actions such as cutting a water melon or stirring a soup, established the intuitive relevance of the auditory feedback at the phenomenological level. Some ecological parameters were uncovered, while a coherence between the action and the sonic feedback was maintained.

Figure 4. The Gamelunch.

As an example, the action of stirring a bowl of soup can provide a coherent feedback in the sense that the sound follows the dynamics of the movement of the spoon in the liquid. At the same time, the sound can be designed to convey the notion of a solid sandy material instead of a liquid one, while maintaining a correspondence between the embodied action and a coherent virtual sonic response, even as the auditory feedback contradicts the haptic and visual feedback. From such a contradiction the unitary multisensory perception of the action is broken and the auditory feedback is instead isolated and placed in perspective. This seems an effective way to induce the user to think about and realize the importance of the sonic aspects in interaction with everyday artefacts.

In this spirit, several such sonic interaction examples were implemented. In one case, friction models for rubbing and squeaking sounds were grafted onto cutting and lifting actions in the interaction with fork and knife, respectively. The continuous sound feedback clearly affected both the sensation of the consistency of the food (a watermelon) and the sensation of the weight of the cutlery [15].

The simple (non-continuous) crumpling model has appeared to be extremely effective for a variety of cases involving powder-like sounds (the sandy-soup) or tearing sounds (the sticky fruit). However, the limitations noted above concerning the original, discretely-triggered version of the crumpling sound are apparent. The new continuous crumpling model offers a more powerful tool and new opportunities to expand and test interactive cases in which the auditory component is more stringent in the sense of a closed-loop and continuous feedback interaction.

In another case, a scenario surrounding the dressing of a salad was implemented. A bowl with a flexion sensor hidden among the leaves is sonified by means of a continuous crumpling sound activated by the action of the two forks mixing the salad. The system also records the quantity of action performed up to a certain time and varies the characteristics of the sound, which becomes gradually more fluid relative to the dry granularity of the initial timbre. The auditory feedback gives an instantaneous sense of the intensity of mixing action. At a longer time-scale, it provides a sense of the effectiveness of the mixing action and of the distribution of the dressing in the salad thanks to the change in timbre of the sound.

AFFECTIVE FOOTSTEP SOUNDS

In this section we present the main design ideas of a recent study in which a model for the synthesis of natural footstep sounds was designed [6], and preliminarily assessed.

In an earlier model of natural walking and running footstep sounds [7], the pace of footsteps was controlled by tempo curves which were derived from studies in music performance, since strong similarities between walking/running and music performance have been found in prior research. A listening test for the validation of that model highlighted the way in which footstep sequences that were generated using expressive tempo curves, derived from music performance, were perceived as more natural by listeners compared to sequences having a constant pace. Using this study as starting point, and in the context of the Affective Diary project [13], we have developed a model of footstep sounds for simulating the presence of people walking in a virtual environment. The design choice was that the footstep sounds should communicate the gender, age, weight, and emotional intention of the virtual walker.

We used the crumpling/impact sound model that is part of the SDT library, tuned to simulate different materials of the ground. Gravel, dirt, soft wood, snow and grass were used for the crumpling sound model; rubber, glass, steel and wood for the impact sound model. The timing in footstep sequences was controlled by using a footstep model developed after measurements of real walkers, who were asked to walk with emotional intentions (happiness, sadness, fear and anger), as well as with their natural (neutral) pace. In interactive listening tests, subjects could adjust pace and material to determine the gender of a virtual walker.

Preliminary results show that subjects associated both different pace and material to the two genders (Figure 5). Female walkers were identified by faster pace (the time interval between to footsteps was about 0.5 s for females and 0.8 s for males), higher resonant frequency for impact sounds (glass and steel sounds for female; rubber and wood sounds for males) and for crumpling sounds (mainly gravel and snow sounds for females; dirt and soft wood sounds for males).

It was also tested how subjects would change the emotion of footstep sound sequences. Subjects could control the sound on a two-dimensional activity-valence space in which pace characteristics (regularity and timing) were changed dynamically. Results are promising despite the existence of some confusion between angry and happy footstep sounds. This confusion should be overcome by improving the continuous control over the real-time change of the acoustical characteristics of the ground, thus allowing for a gradually changing perception of both the gender and emotional intention of a virtual walker. This idea is under test for presentation at the workshop.

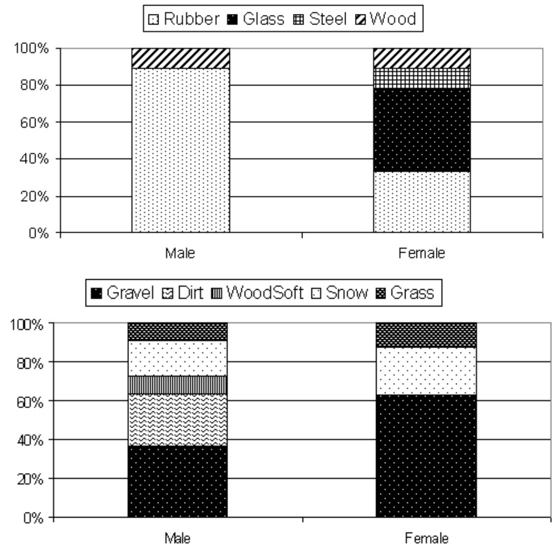

Figure 5. Subjects' percentage choices of different ground materials in association to walker's gender. The upper figure shows the choices for impact sound models. The lower figure shows subjects' preferences for different tunings of the crumpling sound model.

CONCLUSIONS

Controlling the rich spectrum of sounds that the crumpling model continuously changes in front of parameter variations is the most difficult task that must be faced by this research. At the same time, further explorations are needed using many impactor synthesizers simultaneously, generating more sophisticated sound attributes from the materials involved in the crumpling simulation. These two qualities must be addressed in tandem, in order to perform experiments whose outcome may be more definitive than the incomplete and preliminary results reported here. The research to date reveals that crumpling sounds have the potential to become effective conveyors of enactive knowledge in interactive contexts simulating or augmenting everyday experience.

REFERENCES

1. F. Avanzini and D. Rocchesso. Modeling collision sounds: Non-linear contact force. In *Proc. Conf. on Digital Audio Effects (DAFX-01)*, pages 61–66, Limerick, Ireland, December 2001.
2. CLOSED (Closing the Loop of Sound Evaluation and Design). http://closed.ircam.fr.
3. P. Cook. Physically inspired sonic modeling (phism): Synthesis of percussive sounds. *Computer Music Journal*, 21(3), 1997.
4. P. R. Cook. Modeling Bill's gait: Analysis and parametric synthesis of walking sounds. In *Proc. Audio Engineering Society 22 Conference on Virtual, Synthetic and Entertainment Audio*, Espoo, Finland, July 2002. AES.
5. Crackling noise. Web published at http://simscience.org/crackling/index.html.
6. A. DeWitt and R. Bresin. Sound design for affective interaction. In *Proc. Affective computing and intelligent interaction ACII 2007*, pages 523–533, Lisbon, Portugal, Sep. 12-14 2007.
7. F. Fontana and R. Bresin. Physics-based sound synthesis and control: crushing, walking and running by crumpling sounds. In *Proc. Colloquium on Musical Informatics*, pages 109–114, Florence, Italy, May 2003.
8. G. E. Frantti and L. A. Leverault. Auditory discrimination of seismic signals from earthquakes and explosions. *Bull. seism. Soc. Am.*, 55:1–25, 1965.
9. B. Giordano. Everyday listening: an annotated bibliography. In D. Rocchesso and F. Fontana, editors, *The Sounding Object*, pages 1–16. Edizioni Mondo Estremo, Florence, Italy, 2003.
10. B. L. Giordano and S. McAdams. Material identification of real impact sounds: effects of size variation in steel, glass, wood and plexiglass plates. *J. of the Acoustical Society of America*, 119(2):1171–1181, 2006.
11. Granular synthesis. Web published at http://en.wikipedia.org/wiki/Granular_synthesis.
12. G. Kramer. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Addison-Wesley, Reading, MA, 1994.
13. M. Lindström, A. Støahl, K. Höök, P. Sundström, J. Laaksolahti, M. Combetto, A. Taylor, and R. Bresin. *Affective diary: designing for bodily expressiveness and self-reflection*. ACM Press, 2006. Work in progress paper.
14. E. R. Miranda and M. M. Wanderley. *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*. A-R Editions, Middleton, WI, 2006.
15. S. Delle Monache, P. Polotti, S. Papetti, and D. Rocchesso. Gamelunch: A physics-based sonic dining table. In *Proc. Int. Computer Music Conf.*, volume 2, pages 41–44, Copenhagen, Denmark, August 2007.
16. A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, 2nd edition, 1984.
17. L. Peltola, C. Erkut, P. Cook, and V. Valimaki. Synthesis of hand clapping sounds. *IEEE Trans. on Audio, Speech and Language Processing*, 15(3):1021–1029, 2007.
18. M. Saeki. Analytical study of multi-particle damping. *J. Sound and Vibration*, 281:1133–1144, 2005.
19. M. R. Schroeder. *Fractal, Chaos, Power Laws: Minutes from an Infinite Paradise*. W.H. Freeman & Company, New York, NY, 1991.
20. J. P. Sethna, K. A. Dahmen, and C. R. Myers. Crackling noise. *Nature*, (410):242–250, March 2001.