Simon Conan, Etienne Thoret, Mitsuko Aramaki, Olivier Derrien, Charles Gondre, Sølvi Ystad, and Richard Kronland-Martinet
Laboratoire de Mécanique et d'Acoustique
CNRS, UPR 7051
31 chemin Joseph Aiguier
13402 Marseille Cedex 20, France
{conan, thoret, aramaki, derrien, gondre, ystad, kronland}@lma.cnrs-mrs.fr

An Intuitive Synthesizer of Continuous-Interaction Sounds: Rubbing, Scratching, and Rolling

Abstract: In this article, we propose a control strategy for synthesized continuous-interaction sounds. The framework of our research is based on the *action–object* paradigm that describes the sound as the result of an action on an object and that presumes the existence of sound invariants (i.e., perceptually relevant signal morphologies that carry information about the action's or the object's attributes). Auditory cues are investigated here for the evocations of rubbing, scratching, and rolling interactions. A generic sound-synthesis model that simulates these interactions is detailed. We then suggest an intuitive control strategy that enables users to navigate continuously from one interaction to another in an "action space," thereby offering the possibility to simulate morphed interactions—for instance, ones that morph between rubbing and rolling.

Synthesis of everyday sounds is still a challenge, especially the control of sound-synthesis processes. Indeed, it is of interest to intuitively control sounds obtained with a synthesis model, that is, to be able to create sounds that carry or evoke specific information. To achieve this, we need to offer users the possibility to create and control sounds from semantic descriptions of sound events or from gestures. Intuitive sound-synthesis control provides interesting alternatives to indexed sound databases in domains such as the development of video games (Lloyd, Raghuvanshi, and Govindaraju 2011; Böttcher 2013), and is of great interest for sound design (Farnell 2010), sonification (Dubus and Bresin 2013), and virtual and augmented reality—for instance, for motor rehabilitation (Danna et al. 2013; Rodger, Young, and Craig 2013).

In previous studies, intuitive control of sounds based on acoustic descriptors or features has been proposed through so-called feature synthesis (Hoffman and Cook 2006). Other authors either have suggested that sounds can be directly generated from semantic descriptions of timbre (Gounaropoulos and Johnson 2006; Le Groux and Verschure 2008), or have evoked motion of the sound source (Merer et al. 2013). Aramaki et al. (2006, 2009a) proposed the use of semantic labels describing the perceived material, size, and shape of the object producing a sound to intuitively control an impact-sound synthesizer. In particular, that control strategy allowed the user to apply a continuous control to the perceived material to simulate continuous transitions (i.e., morphing) from one material to another (e.g., from glass to metal, through a continuum of ambiguous materials).

This article is devoted to the synthesis and control of continuous-interaction sounds. By continuous interaction we mean any kind of friction phenomena (Akay 2002) or rolling. We look at a subset of continuous interactions—i.e., rubbing, scratching, and rolling sounds. Similarly to the continuous control space of perceived material offered by Aramaki and co-workers (2009a), we would like to present a control space for actions that enables continuous transitions, for instance, from rubbing to rolling. From a synthesis point of view, we have a generic model that allows for such continuous sound transformations. We have achieved this by investigating previous synthesis models: some based on physical modeling or physically informed considerations (Gaver 1993; Hermes 1998; van den Doel, Kry, and Pai 2001; Rath and Rocchesso 2004; Stoelinga and Chaigne 2007) and others on analysis-synthesis schemes (Lagrange, Scavone, and Depalle 2010; Lee, Depalle, and Scavone 2010). Such a generic tool is of interest for sound design and for fundamental studies in sound perception (Aramaki et al. 2009b; Micoulaud-Franchi et al. 2011).

Computer Music Journal, 38:4, pp. 24–37, Winter 2014
doi:10.1162/COMJ_a_00266

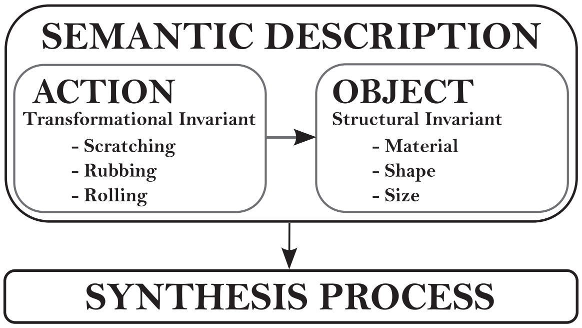


Figure 1. The action–object framework for the synthesis and control of sounds.

In the next section, we describe the *action–object* paradigm within which we developed our synthesizer. We then examine sound morphologies that convey the evocation of rubbing, scratching, and rolling interactions, and describe how these interactions can be reproduced by synthesis. Further, a control strategy of the proposed synthesis model is presented. In the last section, we provide some general conclusions and outline future research.

The Action–Object Paradigm

Inspired by William Gaver (1993), who proposed an independent synthesis of actions and objects, we developed a conceptual description of sounds through an action–object paradigm. This concept consists of considering the sounds as resulting from an action on a resonant object—e.g., "plucking a metal string" or "hitting a wooden plate." This approach suggests the existence of specific acoustic patterns in the perceived signal, enabling the auditory identification of objects, on the one hand, and actions, on the other. For instance, a vibrating string produces a particular spectral content that enables the listener to recognize it, whether it is bowed (e.g., violin), plucked (e.g., guitar), or struck (e.g., piano). Similarly, it is possible to recognize a bottle by the sound it produces, whether it bounces or breaks (Warren and Verbrugge 1984), or a bouncing or rolling cylinder, and the material it is made of (Lemaitre and Heller 2012).

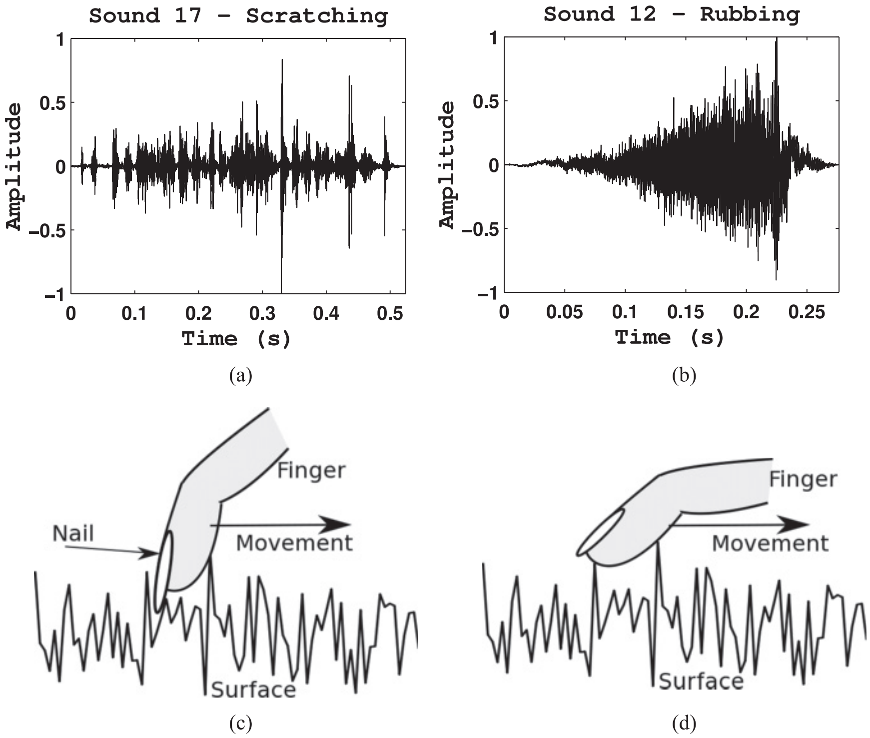
The psychological theory underlying this paradigm is known as the ecological approach of perception, first introduced by James Gibson for visual perception (Gibson 1966, 1979; for a more accessible introduction to Gibson's theory, refer to Michaels and Carello 1981). He proposed that the perception of visual events is constrained by our interactions with the surrounding world, and, more precisely, that the recognition of the properties of a visual event is provided by invariant structures contained in the sensory flow. Concerning the auditory recognition of acoustic events, this theory was first exploited by Warren and Verbrugge (1984) and later formalized by McAdams and Bigand (1993). It supposes the existence of *invariant structures* that carry the necessary information for the recognition of sound events. The so-called invariants are split in two categories: *structural invariants*, which enable recognition of physical properties of a sounding object (its material, shape, etc.) and *transformational invariants* describing the type of change or the action on the object (breaking, rolling, etc.).

Some studies have already identified such acoustic invariants. For instance, it has been shown that impact sounds contain sufficient information to perceptually discriminate the material (Wildes and Richards 1988) or the size (Lakatos, McAdams, and Caussé 1997; Carello, Anderson, and Kunkler-Peck 1998) of the sound-producing, impacted objects. In particular, it was shown that the perception of material is mainly related to frequency-dependent damping of spectral components (Klatzky, Pai, and Krotkov 2000; Tucker and Brown 2002; Giordano and McAdams 2006) and to roughness (Aramaki et al. 2009b). A study by Warren and Verbrugge (1984) revealed that, from the rhythm of a series of impact sounds, it is possible to predict if a glass will break or bounce. More recently, Thoret and colleagues (2014) highlighted that, by listening to friction sounds produced when someone is drawing, subjects were able to recognize (to a certain extent) the shape that was drawn and that the relevant information was conveyed by the velocity profile of the writer's gesture. The general action–object framework, on which the synthesis model presented in this article is based, is illustrated in Figure 1.

Invariants Related to Continuous Solid Interactions

In this section, acoustic invariants related to rubbing, scratching, and rolling interactions

Figure 2. Recorded sounds 100 percent associated with scratching (a) and with rubbing (b), as well as schematic representations of a nail scratching a surface (c) and a finger rubbing a surface (d). In the schematic representations, the y*-axis represents the detailed surface height (greatly exaggerated for clarity); the* x*-axis represents the direction along which the finger is dragged across the surface (also corresponding approximately to time, as the finger is dragged from left to right).*



(a) (b)

(c) (d)

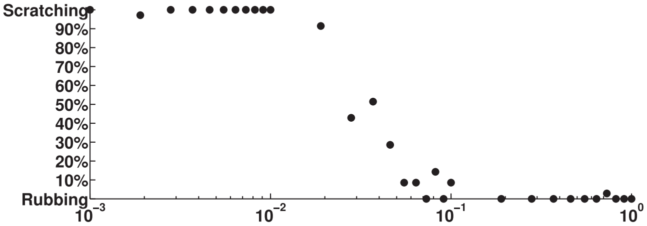
are described. Technical details related to the implementation of both the object and the action parts will be provided later.

Rubbing and Scratching Interactions

To our knowledge, the auditory ability to distinguish rubbing sounds from scratching sounds has not previously been formally investigated. To ascertain that it is possible to distinguish a sound that evokes rubbing from one that evokes scratching and to reveal the signal properties responsible for this ability, a series of perceptual experiments was conducted by Conan and colleagues (2012; the related paper and experimental material are available online at www.lma.cnrs-mrs.fr/∼kronland/RubbingScratching). We summarize the main results here. From the first experiment, qualitative analyses were made of recorded sounds for which all subjects agreed that the sound evoked one of the interactions, scratching or rubbing. These analyses implied that rubbing sounds resulted from a higher temporal density of impact events than scratching sounds. A sound that was associated with scratching by all subjects is plotted in Figure 2a, and a sound that was associated with rubbing by all subjects is in Figure 2b. The differences in temporal density of impacts can be explained as follows. Sounds produced when scratching a surface, for instance, with a nail, are due to the interaction between surface irregularities and the nail and therefore can be considered to be the result of successive impacts. Scratching a surface can be considered as scanning deeper into the surface than rubbing, which implies that each surface irregularity is encountered one after another and more intensely than in the case of rubbing. In rubbing, several surface irregularities are encountered simultaneously and less intensely, implying a more noisy sound, due to the higher density of impacts (see Figures 2c and 2d). The perceived surface irregularities therefore seem to be closely linked to the type of interaction.

These observations led us to set up a second experiment to validate the previous hypothesis. The experiment investigated how impact density, as a

Figure 3. Results of the experiment with synthesized sounds. The x*-axis represents the temporal density of impacts (log scale, increasing from left to right), and the* y*-axis represents the percentage of association to the scratching category for each sound.*



relevant acoustic cue, could serve to distinguish rubbing from scratching sounds. The stimuli used for the experiment were synthetic friction sounds generated with different impact densities. The synthesis model was based on the pioneering work of Gaver (1993), relying on phenomenological considerations, and later improved by van den Doel, Kry, and Pai (2001). It consists of simulating the interaction force as a result of successive "microimpacts" of a plectrum on the irregularities of a surface. The successive impacts are modeled by low-pass filtered noise with a cutoff frequency related to the plectrum velocity, and the roughness of the surface is defined by the nature of the noise. Thirty synthetic sounds, representing a continuous transition from low- to high-density impact series, were generated. These sounds were then presented randomly to 35 subjects who were asked to associate each sound with one of the two interaction categories, rubbing or scratching. The results of the perceptual experiment are displayed in Figure 3 (details provided in Conan et al. 2012) and confirm our hypothesis that low impact densities are associated with scratching and high impact densities with rubbing, and there is a less clear perception of the type of interaction for moderate densities (at intermediary positions of the continuous transition).

In summary, these experiments allowed us to conclude that one invariant (possibly among others) that contributes to the discrimination between these two interactions is the temporal density of the impacts contained in the interaction force. That is, the higher the impact density in the signal, the higher the probability for the sound to evoke rubbing. Similarly, lower impact density in the signal tends to evoke scratching. Therefore, from a synthesis point of view, rubbing and scratching interactions can be simulated by controlling the characteristics of the impact series.

Rolling Interaction

The rolling interaction has been fully investigated by Conan and co-workers (2014). The main points are summarized in this section.

Similarly to rubbing and scratching interactions, we address here the determination of acoustic morphologies related to the auditory perception of rolling: Which signal information is responsible for the recognition of a rolling object? To answer this question, we investigated a physics-based model of rolling sounds. In the literature, most authors consider that the physics of a rolling ball is similar to the physics of a bouncing ball. The model of a bouncing ball generally takes into account a nonlinear sphere–plane interaction that relates the force f applied to the sphere to the penetration x and the penetration velocity $\dot{x}$ of the sphere into the surface (following the model of Hunt and Crossley 1975):

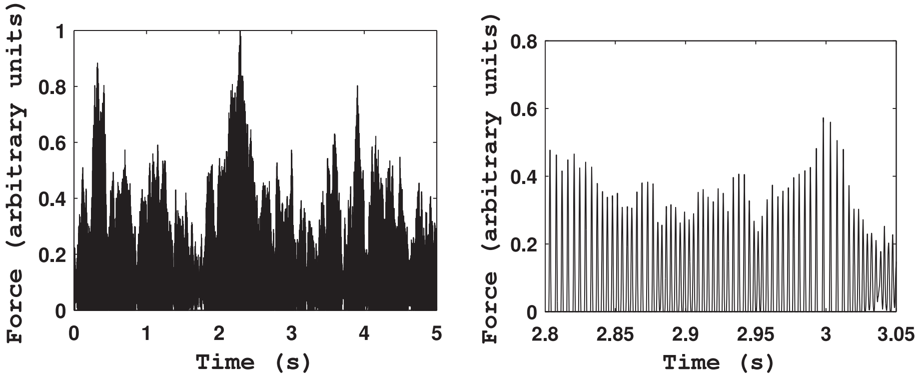
$$f(x,\dot{x}) = \begin{cases} kx^{\alpha} + \lambda x^{\alpha}\dot{x}, & x > 0 \\ 0, & x \leq 0, \end{cases} \tag{1}$$

where k is the stiffness and λ the damping weight of the force. The parameter α takes into account the local geometry around the contact surface ($\alpha = 3/2$ according to Hertz's theory of contact mechanics, cf. Johnson 1987). By taking into account the effect of the gravity on the ball, this model simulates the behavior of a bouncing ball (see, for instance, Falcon et al. 1998; Avanzini and Rocchesso 2001).

To adapt this model to the simulation of a rolling ball (Rath and Rocchesso 2005; Stoelinga and Chaigne 2007) or a rolling wheel (Nordborg 2002), consider that the rolling object moves along an irregular surface and the height of the irregularities is added as a perturbation to the penetration term x in Equation 1. The rolling interaction can be considered as a ball that bounces on surface irregularities with a randomly changing height.

We simulated the rolling of a ball over an irregular surface using the fourth-order Runge-Kutta method

Figure 4. Simulated rolling interaction force (a) and a magnified (zoomed in) excerpt (b). The force parameters are $\alpha = 3/2$, $k = 10^7 N/m^{3/2}$, $\lambda = 10^7 N\text{-}sec/m^{5/2}$. The ball has a mass of 5g and a velocity of 0.5m/sec. The surface is assumed to be fractal with $\beta = 1.2$ and a maximum amplitude of $10^{-9} m$.



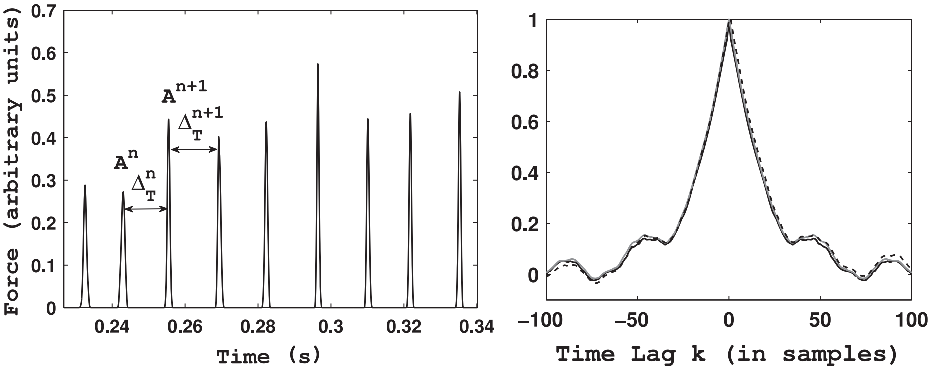
(a) (b)

(as described by Papetti, Avanzini, and Rocchesso 2011, where numerical issues of Equation 1 are studied). Note that we do not consider here the vibration of the surface itself, as do Rath and Rocchesso (2005). The surface, therefore, is modeled by noise with a specific spectrum, adjusted according to tribological observations (Ben Abdelounis et al. 2010) and phenomenological considerations (van den Doel, Kry, and Pai 2001; Rocchesso and Fontana 2003). In practice, the spectrum is characterized by $S(\omega) \propto 1/\omega^{\beta}$ where β enables the control of the perceived roughness. The amplitude of $S(\omega)$ is normalized to provide a maximum asperity of $10^{-9} m$. Also called fractal noise, such a spectrum accurately models most physical surfaces. An example of a simulated nonlinear interaction force f with such a surface is plotted in Figure 4.

Based on results from informal listening tests, such an interaction force f was found to convey sufficient information to evoke a rolling object (and was often perceived as a small, hard marble ball). Moreover, it has been shown that this model produces rolling sounds spontaneously recognized as such by naive listeners (Rath 2004). From a signal point of view, this force can be considered as a series of impacts (see Figure 4b). This assumption has already been exploited by Dik Hermes (1998) for the purpose of sound synthesis, and by Lagrange and colleagues (2010) in an analysis-synthesis context. In particular, if we consider the amplitude of the impacts and the time interval between two successive impacts as time series, the autocorrelation and the cross-correlation between the previously simulated time series obtained from the interaction force f can be computed. As shown in Figure 5b, these time series have a strong autocorrelation. That is, successive impacts have strong mutual dependencies. They are also strongly cross-correlated. These observations are coherent with the physics of a bouncing ball (recall that the rolling model is derived from a bouncing model), since within a bounce event, successive impacts are mutually related. Hence, a first conclusion that can be made is that the temporal structure of impact series associated with a rolling interaction seems to follow a specific pattern.

Another important characteristic of the rolling interaction force f is the dependency of the contact time on the impact velocity, which is related to the amplitude of the impact. This dependency has been studied by several authors (Chaigne and Doutaut 1997; Avanzini and Rocchesso 2001) and seems to be an important auditory cue that is responsible for evoking the rolling interaction. Indeed, to informally test this assumption, we first detected all the impact durations and amplitudes in the simulated force plotted in Figure 4. We then created two modified versions of this force by replacing the impact windows by a raised cosine window (van den Doel, Kry, and Pai 2001) that fits the original impact shapes well, one with a duration depending on the impact amplitude, the other with a fixed duration. It was always found that the version with varying, amplitude-dependent impact durations clearly produced the most realistic evocations

Figure 5. Notations for the impacts amplitudes A *and time interval* Δ_T *series (a). Autocorrelation of* Δ_T *(black) and* A *(gray) time series, cross-correlation between* A *and* Δ_T *time series (dashed black) (b). The amplitudes of the correlations functions are normalized.*



(a)

(b)

of the rolling interaction, confirming previous findings.

Finally, some authors have suggested that a rolling object should never be considered perfectly spherical or perfectly homogeneous, so they have applied a sinusoidal modulation to the interaction force (Hermes 1998; Rath and Rocchesso 2005). From a perceptual point of view, the asymmetry of rolling objects is likely to contribute to the sensation of velocity. This auditory cue has been shown to strongly influence the perception of rolling objects' size and speed (Houben 2002). Hence, we considered amplitude modulation as a significant sound effect for evoking the sound of rolling.

Summary of Invariants Related to Continuous Interactions

The previous considerations allowed us to conclude that the interaction forces associated with rubbing, scratching, and rolling interactions can all be represented as impact series. Based on results from listening tests, it was shown that the temporal density of impacts conveys the information needed to recognize rubbing or scratching sounds. As opposed to rubbing and scratching interactions, the intrinsic structure of the impact series (correlations and statistics of durations and amplitudes) seems to be an indispensable signal morphology for evoking the rolling interaction. In the next section, we will describe a generic model to simulate rubbing, scratching, and rolling sounds based on the reproduction of the characteristics of these impact series.

Implementation of the Generic Sound-Synthesis Model

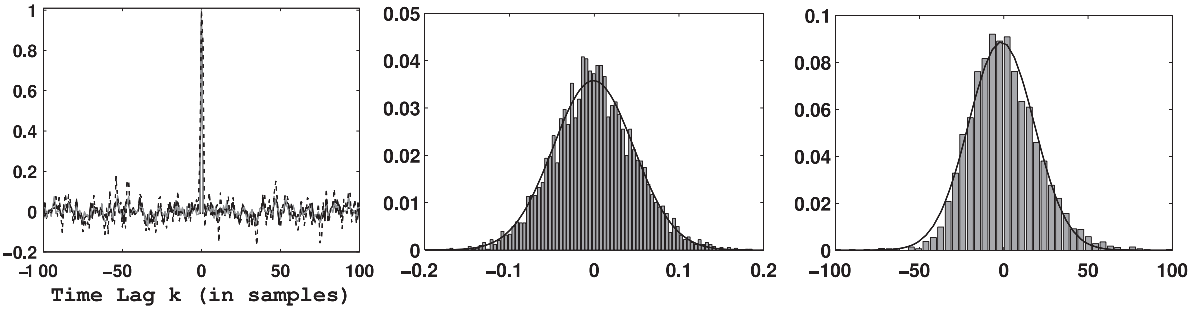
As described in the previous section, the interaction force carries relevant perceptual information related to rubbing, scratching, and rolling interactions. For these three types of interaction, the forces are series of impact events with specific relations. Such a signal can be formally described as:

$$f(t) = \sum_n A^n \, \phi^n(t - T^n), \qquad (2)$$

where A^n and T^n are, respectively, the amplitude and time-position of each impact, and ϕ^n represents the "impact pattern," i.e., the shape of the n^{th} impact. Because the absolute time position of the impact is not a relevant parameter, we consider the time interval between successive impacts, defined as $\Delta_T^n = T^{n+1} - T^n$, in the remainder of this article (cf. Figure 5).

The aim of this section is to describe a synthesis process that is generic enough to simulate rubbing, scratching, and rolling interactions and, further, to design an intuitive control allowing a continuous navigation between these interactions. According to the action–object paradigm, the interaction forces constitute the action part of the general action–object framework, in which actions and

Figure 6. Autocorrelation of ${}^{c}\tilde{\Delta}_{T}$ (black) and ${}^{c}\tilde{A}$ (gray) time series, cross-correlation between ${}^{c}\tilde{A}$ and ${}^{c}\tilde{\Delta}_{T}$ time series (dashed black) are displayed in (a). The amplitudes of the correlation functions are normalized. Probability density of ${}^{c}\tilde{A}$ (b) and ${}^{c}\tilde{\Delta}_{T}$ (c). The gray bars are the measures and the black line is the Gaussian fit.



(a) (b) (c)

objects can be simulated independently and freely associated. The properties of the resonant object (such as perceived material or shape) are included in the object part (as proposed by Gaver 1993) based on the synthesizer developed by Aramaki et al. (2006, 2009a).

In practice, the implementation process that associates the action with the object is based on a source-filter model, and the resulting sound is obtained by convolving the force (source) with the impulse response of a resonant object (filter bank). It is important to note that such a process based on a source-filter model does not limit the synthesis possibilities to the simulation of linear interaction phenomena. Indeed, nonlinear effects that are perceptually relevant can be taken into account in the source part. This has, for instance, been successfully accomplished for piano tone synthesis (Bensa, Jensen, and Kronland-Martinet 2004) and for synthesis of flute sounds (Ystad and Voinier 2001), but also for synthesis of nonlinear friction sounds (such as squeaking sounds, see Thoret et al. 2013). The implementation of the action and object parts will be described in the following sections.

Action Part Implementation

As previously presented, the force f that conveys the perceptual information about the interaction type (rolling, scratching, or rubbing) can be modeled as an impact series (Equation 2). The specific behaviors of the amplitude A^{n} and time interval Δ_{T}^{n} series seem to be an important perceptual cue associated with these interactions. We will experimentally characterize the time series ${}^{c}A = A - \mu_{A}$ or ${}^{c}\Delta_{T} = \Delta_{T} - \mu_{\Delta_{T}}$ (i.e., the centered versions of A and Δ_{T}) in order to, first, delineate a synthesis scheme for rolling sounds. Then we will describe how this model can be extended to simulations of rubbing and scratching sounds.

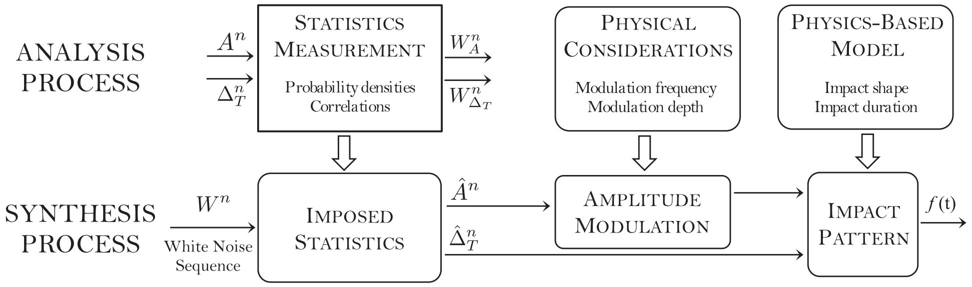
As pointed out in the section on the rolling invariants, A and Δ_{T} are strongly autocorrelated (as are ${}^{c}A$ or ${}^{c}\Delta_{T}$). We consider these two time series as autoregressive moving average processes. To characterize their behaviors, these series are "whitened" (as in Ninness, Wills, and Gibson 2005), and we experimentally noted that the whitening filters need have no more than one pole and one zero. Let X be one of the two processes ${}^{c}A$ or ${}^{c}\Delta_{T}$; we can then write in the z domain:

$$X(z) \approx H_{X}(z)\tilde{X}(z), \quad H_{X}(z) = \frac{1 + \sum_{i=1}^{p} b_{i} z^{-i}}{1 + \sum_{i=1}^{q} a_{i} z^{-i}}, \tag{3}$$

where $\tilde{X}$ is the whitened version of X. As $\tilde{X}$ is white, we can properly estimate its probability density function and transform $\tilde{X}$ into W_{X}, which follows a uniform law, thanks to the inverse-transform sampling method (i.e., $W_{X} = F_{X}(\tilde{X})$, where F_{X} is the cumulative distribution function of the random variable $\tilde{X}$).

Figure 6a displays the autocorrelation of ${}^{c}\tilde{\Delta}_{T}$ (black line) and ${}^{c}\tilde{A}$ (gray line), the whitened version of ${}^{c}\Delta_{T}$ and ${}^{c}A$. The A and Δ_{T} time series were obtained from the force f computed with the physics-based model. As experimentally observed, this figure shows that a one-pole, one-zero whitening filter is well suited to whiten the time series. The probability densities of ${}^{c}\tilde{A}$ and ${}^{c}\tilde{\Delta}_{T}$ are plotted in Figure 6b and Figure 6c,

Figure 7. Schematic representation of the analysis–synthesis process.



respectively, showing that for the rolling force, they can be modeled as Gaussian processes. The analysis process of the impact series is schematically described in Figure 7.

In Figure 6a, $C_{(W_A^n, W_{\Delta_T}^n)}(k)$ the cross-correlation between the whitened process W_A and W_{Δ_T} is displayed. Given the long-term autocorrelation of ${}^cA$ and of ${}^c\Delta_T$, we assume that $W_{\Delta_T}^{n-1}$ has a poor influence on W_A^n (and, similarly, that W_A^{n-1} has a poor influence on $W_{\Delta_T}^n$), and that $C_{(W_A^n, W_{\Delta_T}^n)}(k)$ is proportional to $\delta(k)$, the unit impulse that equals 1 if $k = 0$ and equals 0 elsewhere. This led to the proposition of the synthesis scheme presented in Figure 7 where, according to the previous observations on $C_{(W_A^n, W_{\Delta_T}^n)}(k)$, we start from a single white and uniform process W to synthesize both $\hat{A}$ and $\hat{\Delta}_T$ (the hat designates for "estimated values," and no hat designates "measured values").

A sinusoidal modulation of the amplitude time series is further included in the synthesis process:

$$s(t) = [1 + m \sin(2\pi \nu_m t)], \tag{4}$$

with $m \in [0, 1]$ the modulation depth, controlling the perceived asymmetry of the rolling ball. The modulation frequency follows:

$$\nu_m \propto \frac{v}{R} \tag{5}$$

where v is the transversal velocity and R the ball radius (i.e., the modulation frequency increases with the rolling object's speed and decreases with its size.) As previously explained, this is done to emphasize the perceived velocity and size, as proposed in several studies (Hermes 1998; Houben 2002; Rath and Rocchesso 2005).

Finally, to simulate the whole force, we consider the "impact pattern" ϕ^n (cf. Equation 2) defined as a raised cosine function (van den Doel et al. 2001) :

$$\phi^n(t) = \begin{cases} \frac{1}{2}\left[1 + \cos\left(\frac{2\pi t}{t_0^n}\right)\right], & t \in \left[-\frac{t_0^n}{2}, \frac{t_0^n}{2}\right], \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where t_0^n is the duration of the nth impact. As previously discussed, the impact duration varies with the impact amplitude, and we suggest the following relation:

$$t_0^n = \zeta \cdot (A^n)^{-\theta}, \tag{7}$$

where ζ is a constant that depends on the mass of the ball and on the stiffness k (see Equation 1), A^n is the n^{th} impact amplitude, and θ is a positive value that controls the strength of the dependence between the impact amplitude and the duration. This mapping was chosen based on several physical studies of impact sounds (Chaigne and Doutaut 1997; Avanzini and Rocchesso 2001).

Although this model was derived to synthesize rolling sounds, it is also well suited to reproduce rubbing and scratching sounds. As previously highlighted, the latter interaction forces are satisfactorily reproduced with stochastic processes such as white noise. Hence, by ignoring correlations between impacts, which in the present model corresponds to setting the (a_1, b_1) coefficients of the filters $H_A(.)$ and $H_{\Delta_T}(.)$ (Equation 3) to zero. In this case, the amplitude series A follows a Gaussian process and Δ_T follows an exponential distribution.

Object Part Implementation

The synthesis of impacted resonant objects was previously addressed by using an additive synthesis process (Aramaki and Kronland-Martinet 2006; Aramaki et al. 2006, 2009b). In this case, the frequencies of the oscillators corresponded to the eigenfrequencies of the resonant object. Noise could be added at the output of the oscillators to simulate the stochastic part of the impact. Then, the summed signal of noise and sinusoids was filtered into different frequency bands according to the Bark scale (Zwicker and Fastl 1990). A different time-varying envelope was applied to each band to take into account the frequency dependency of the damping. This synthesis process is effective for simulating single impacts, but not suitable when increasingly complex interactions are to be combined with the object. For instance, to synthesize bouncing objects, the signal of noise plus sinusoids has to be triggered precisely, as do the time-varying gains of the filter bank. This becomes computationally too expensive for more complex and continuous interactions such as rubbing, scratching, or rolling.

To overcome this limitation, a subtractive synthesis process is used to implement the object part. A resonant filter bank, whose frequencies and damping coefficients are tuned according to the eigenmodes and material properties of the resonant object, is used to simulate the object's impulse response. Each resonant filter's impulse response is an exponentially decaying sinusoid (Mathews and Smith 2003). With this approach, complex and continuous interactions can be fed into the resonant filter bank in a computationally efficient manner. A similar synthesis algorithm has already been proposed by van den Doel and Pai (2003).

Intuitive Sound-Synthesis Control

In this section, we detail our control strategy for continuous navigation between rubbing, scratching, and rolling interactions. The intuitive control of the perceived material of the object part was suggested by Aramaki et al. (2011), and the interface allows the user to morph between different material categories (i.e., wood, metal, or glass) by moving a cursor in a two-dimensional space. The material categories are represented by anchors (prototypes) in the navigation space and correspond to specific sets of parameters. Thus, by interpolating spectral and damping parameters between the anchors according to the cursor's displacement, the synthesis process enables one to morph between materials and to create in-between (or hybrid) materials. Note that high-level control of the resonant object's shape is also available and allows the user to specify whether the object is one-dimensional (string-bar), two-dimensional (plate-membrane), or three-dimensional (shell).

Similarly to this material space for the control of objects, we present a navigation space for an intuitive control of the interaction type based on the definition of a "prototype" for each interaction. We will present here the navigation strategy between these prototypes, which are defined from the low-level parameters of the synthesis model as:

Impact model—two parameters control the impact duration t_0^n (Equation 7) in the chosen impact model (Equation 6): ζ and θ.

Probability density—the probability density is sampled as a set of discrete values, which are used to derive the cumulative distribution function F_X (the cumulative sum of the probability density), giving ${}^{c}\hat{\tilde{A}}$ and ${}^{c}\hat{\tilde{\Delta}}_T$ series. Two sets are defined, for the ${}^{c}\hat{\tilde{A}}$ and ${}^{c}\hat{\tilde{\Delta}}_T$ series respectively, and are written as $\mathbb{P}_A$ and $\mathbb{P}_{\Delta_T}$.

Whitening filters—as previously pointed out, one pole and one zero are sufficient for these filters. Each filter is described by a set of two coefficients (a_1, b_1), and each set is noted $\mathcal{C}_A$ and $\mathcal{C}_{\Delta_T}$, for $H_A(.)$ and $H_{\Delta_T}(.)$, respectively.

Offset coefficients—these correspond to the centered values μ_A and μ_{Δ_T} for $\hat{A}$ and $\hat{\Delta}_T$, respectively.

Amplitude modulation—the parameters are the modulation depth m and the frequency modulation ν_m (depending on the size and speed of the rolling object).

The global set of parameters is called: $\mathfrak{P} = \{\zeta, \theta, \mathbb{P}_A, \mathbb{P}_{\Delta_T}, \mathcal{C}_A, \mathcal{C}_{\Delta_T}, \mu_A, \mu_{\Delta_T}\}$.

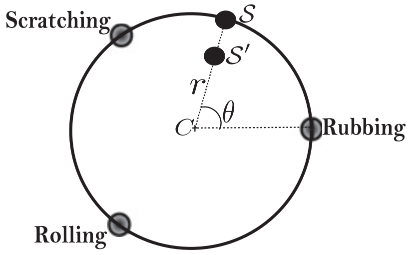


Figure 8. Schematic control space of the interaction sound synthesizer.

Prototypes for Interaction Sounds

In this section, we will offer a set of parameters that can be used to calibrate "prototypes" for the three interaction sounds. Regarding the rubbing and scratching interactions, the precise parameter values have been documented in another paper (Conan et al. 2012). Regarding the rolling interaction, we have documented the parameter values, as well as the rolling controls such as size, velocity, and surface roughness, in a recent article published in *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (Conan et al. 2014).

Rubbing Prototype

As shown by Conan et al. (2012) and noted in the section on sound invariants, small time intervals between impacts are perceptually associated with rubbing sounds (a maximum impact density implies a source signal that is white noise). For the rubbing sound prototype, we set the parameters to obtain a Gaussian white noise: the coefficients of $\mathcal{C}_A$ and $\mathcal{C}_{\Delta_T}$ were equal to zero, $\mu_A = \mu_{\Delta_T} = 0$, $\theta = 0$ (no dependency between impact amplitude and duration), $\mathbb{P}_{\Delta_T}$ was defined to get one impact at each sample, and $\mathbb{P}_A$ to follow a Gaussian distribution. The parameter ζ is set to get an impact duration of one sample and is used to control the perceived size of the object that rubs (longer impact durations, corresponding to low-pass filtering, evoke bigger objects). This set of parameters is called $\mathfrak{P}_{\text{rub}}$.

Scratching Prototype

The set of parameters for the scratching sound prototype is the same as for the rubbing prototype except for the Δ_T^n time series. As described previously, the scratching is associated with a low impact density (i.e., to high Δ_T^n values), and we propose that the Δ_T series follows an exponential distribution (defined empirically by Conan et al. 2012). This parameter set is called $\mathfrak{P}_{\text{scratch}}$.

Rolling Prototype

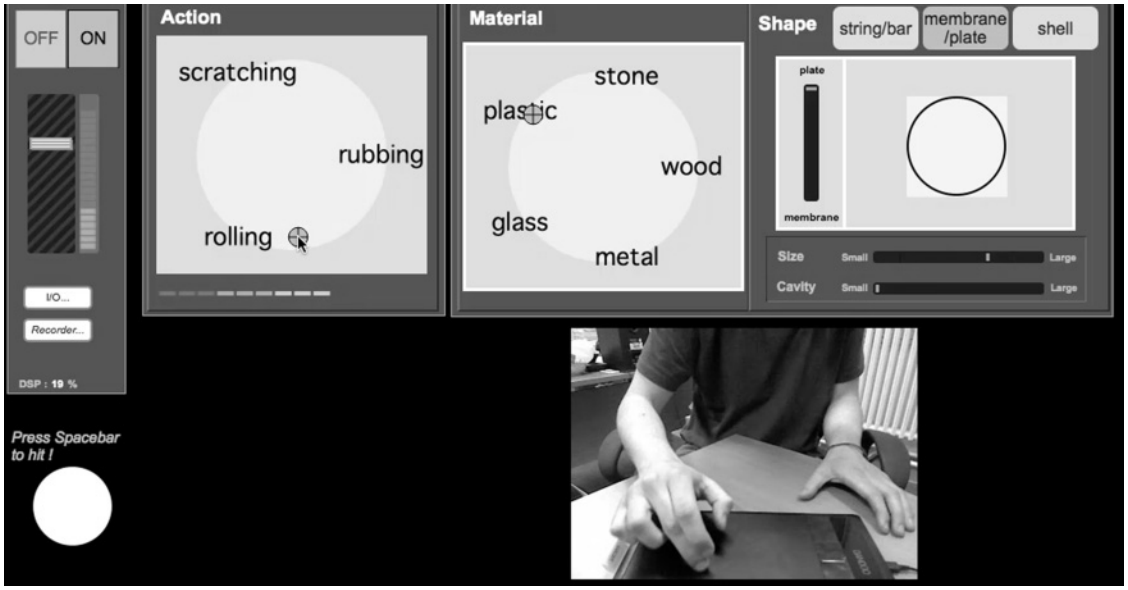
We set $\mathbb{P}_A$ and $\mathbb{P}_{\Delta_T}$ as Gaussian distributions. The values of the filter coefficients $\mathcal{C}_A$ and $\mathcal{C}_{\Delta_T}$, which are nonzero here, are based on the experimental analysis of the rolling force previously simulated. θ is set to 0.29, based on numerical simulations of the impact model (Equation 1). From a perceptual point of view, as ζ controls the contact duration, it is related to the size of the rolling ball. The filter coefficients $\mathcal{C}_A$ and $\mathcal{C}_{\Delta_T}$, as well as the probability densities $\mathbb{P}_A$ and $\mathbb{P}_{\Delta_T}$, are linked to surface parameters such as roughness. The perceived asymmetry lies within the interval $[0, 1]$. The modulation frequency is set to $\nu_m = 3V/S$, with $S \in]0, 1]$ the perceived ball size and $V \in [0, 1]$ the perceived ball velocity. This parameter set is called $\mathfrak{P}_{\text{roll}}$.

Navigation Strategy

The control strategy that we designed is inspired by the one described by Aramaki and colleagues (2011) to control the perceived material in an impact sound synthesizer. The three sound prototypes of perceived material (wood, metal, and glass, for which synthesis parameters were determined based on behavioral and electrophysiological experiments) are placed on the border of a disk that represents the so-called material space. The user can thereby navigate continuously between wood, metal, and glass by moving a cursor in this space, and the synthesis parameters are interpolated according to the distance to the three prototypes. A similar control space dedicated to interaction sounds is schematized in Figure 8.

The sound prototypes for rubbing, scratching, and rolling are placed as anchors on the circumference of a unit disk, at angles of 0, $2\pi/3$, and $4\pi/3$, respectively. Along this circumference, a sound $\mathcal{S}$, characterized by its angle θ, is defined by the

Figure 9. Synthesizer interface. One can see the user controlling the synthesizer by means of a graphical tablet that captures gesture velocity.



parameters $\mathfrak{P}_{\mathcal{S}}(\theta)$ as follows:

$$\mathfrak{P}_{\mathcal{S}}(\theta) = T(\theta)\,\mathfrak{P}_{\text{rub}} + T\left(\theta - \frac{2\pi}{3}\right)\mathfrak{P}_{\text{scratch}} + T\left(\theta - \frac{4\pi}{3}\right)\mathfrak{P}_{\text{roll}}. \tag{8}$$

The function $T(\theta)$ is defined as follows:

$$T(\theta) = \begin{cases} -\dfrac{3}{2\pi}\theta + 1, & \theta \in \left[0; \dfrac{2\pi}{3}\right[\\ 0, & \theta \in \left[\dfrac{2\pi}{3}; \dfrac{4\pi}{3}\right[, \\ \dfrac{3}{2\pi}\theta - 2, & \theta \in \left[\dfrac{4\pi}{3}; 2\pi\right[\end{cases} \tag{9}$$

Inside the disk, a sound $\mathcal{S}'$, characterized by both its angle θ and radius r, is defined by the parameters $\mathfrak{P}_{\mathcal{S}'}$:

$$\mathfrak{P}_{\mathcal{S}'}(\theta, r) = (1 - r)\mathfrak{P}_C + r\mathfrak{P}_{\mathcal{S}}(\theta), \tag{10}$$

where

$$\mathfrak{P}_C = \frac{1}{3}\left(\mathfrak{P}_{\text{rub}} + \mathfrak{P}_{\text{scratch}} + \mathfrak{P}_{\text{roll}}\right), \tag{11}$$

and $\mathfrak{P}_{\mathcal{S}}(\theta)$ is as defined in Equation 8.

In addition to the navigation in this "action space," the gesture is taken into account in the control strategy. Indeed, for such continuous interactions, the underlying gesture is a fundamental attribute that can be conveyed in the dynamics of the sound (Merer et al. 2013; Thoret et al. 2014). Following the synthesis process discussed by van den Doel, Kry, and Pai (2001), the resulting interaction force is low-pass filtered with a cutoff frequency that is directly related to the relative transversal velocity between the object that interacts (hand, plectrum, etc.) and the surface. When associated with a biological law, a specific calibration of the velocity profile enables the evocation of a human gesture (Thoret et al. 2014). The synthesizer interface is displayed in Figure 9. A demonstration video showing intuitive navigation in the action space and a gestural control on a graphic tablet is available online at www.lma.cnrs-mrs.fr/~kronland/CMJ2014 and is also at www.mitpressjournals.org/doi/suppl/10.1162/COMJ_a_00266/.

Conclusions

In this article, we described specific signal morphologies that are related to the auditory perception of rubbing, scratching, and rolling interactions. Phenomenological considerations, physical modeling, and qualitative signal analysis were investigated, and we concluded that the interaction forces conveyed the relevant perceptual information about the type of interaction. A generic synthesis model aiming at reproducing these interaction forces (characterized by particular statistics of impact series) was designed. Then an intuitive control space that enables continuous transitions between these interactions was described.

Further studies will be done to expand this control space to other interactions such as nonlinear friction (squeaking, squealing, etc.; also see Thoret et al. 2013; Avanzini, Serafin, and Rocchesso 2005). The synthesis parameters associated with rubbing and scratching could be refined from the analysis of a large set of recorded sounds by using methods such as those discussed by Lagrange, Scavone, and Depalle (2010). The influence of the physical velocity profiles (rolling marble in a bowl, sliding object on an inclined plate, etc.) on the perceived interaction can also be studied, as already done by Thoret and co-workers (2014) on the evocation of human gestures by using specific velocity profiles. Such synthesis tools and morphing capabilities are of interest for motor rehabilitation purposes (Danna et al. 2013).

More interestingly, the proposed action–object framework is suitable for the creation of *sound metaphors*. This means that by freely combining objects with actions, unheard-of action–object combinations could be synthesized. With further experiments on new sound textures, the salience of the sound morphologies related to the evocations of the actions highlighted in this study can be accurately examined. For instance, is it possible to modify a given sound texture by means of the highlighted invariants, so that this texture evokes a rolling interaction? The determination of such transformational methods is useful in various domains related to musical as well as sonification applications. In particular, direct applications can be found in the current “MétaSon” project, in which specific sound-design issues are raised by the automobile industry. Because an increasing number of electric cars making very little noise will be used in the future, they should be equipped with sounds that clearly evoke a rolling object and that are recognizable as potentially dangerous objects for pedestrians. The model offered, together with the intuitive control strategy, constitutes a relevant tool for such investigations.

Acknowledgments

This work was funded by the French National Research Agency (ANR) under the MétaSon: Métaphores Sonores (Sound Metaphors) project (ANR-10-CORD-0003) in the Contenus Numériques et Interactions (CONTINT) 2010 framework (information online at metason.cnrs-mrs.fr/home.html). The authors Conan, Thoret, Aramaki, Gondre, Ystad, and Kronland-Martinet are with the Laboratoire de Mécanique et d’Acoustique (LMA), Centre National de la Recherche Scientifique (CNRS), Unité Propre de Recherche (UPR) 7051, Aix-Marseille Université, and École de Centrale Marseille. Derrien is with the Université de Toulon, LMA-CNRS-UPR 7051.

References

Akay, A. 2002. “Acoustics of Friction.” *Journal of the Acoustical Society of America* 111(4):1525–1548.

Aramaki, M., and R. Kronland-Martinet. 2006. “Analysis-Synthesis of Impact Sounds by Real-Time Dynamic Filtering.” *IEEE Transactions on Audio, Speech, and Language Processing* 14(2):695–705.

Aramaki, M., et al. 2006. “A Percussive Sound Synthesizer Based on Physical and Perceptual Attributes.” *Computer Music Journal* 30(2):32–41.

Aramaki, M., et al. 2009a. “Thinking the Sounds: An Intuitive Control of an Impact Sound Synthesizer.” In *Proceedings of the International Conference on Auditory Display*, pp. 119–124.

Aramaki, M., et al. 2009b. “Timbre Perception of Sounds from Impacted Materials: Behavioral, Electrophysiological and Acoustic Approaches.” In S. Ystad, R. Kronland-Martinet, and K. Jensen, eds. *Computer Music Modeling and Retrieval: Genesis of Meaning in Sound and Music*. Berlin: Springer, pp. 1–17.

Aramaki, M., et al. 2011. “Controlling the Perceived Material in an Impact Sound Synthesizer.” *IEEE*

Transactions on Audio, Speech, and Language Processing 19(2):301–314.

Avanzini, F., and D. Rocchesso. 2001. "Modeling Collision Sounds: Non-Linear Contact Force." In *Proceedings of the COST-G6 Conference on Digital Audio Effects*, pp. 61–66.

Avanzini, F., S. Serafin, and D. Rocchesso. 2005. "Interactive Simulation of Rigid Body Interaction with Friction-Induced Sound Generation." *IEEE Transactions on Speech and Audio Processing* 13(5):1073–1081.

Ben Abdelounis, H., et al. 2010. "An Experimental Study on Roughness Noise of Dry Rough Flat Surfaces." *Wear* 268(1):335–345.

Bensa, J., K. Jensen, and R. Kronland-Martinet. 2004. "A Hybrid Resynthesis Model for Hammer-String Interaction of Piano Tones." *EURASIP Journal on Applied Signal Processing* 2004(7):1021–1035.

Böttcher, N. 2013. "Current Problems and Future Possibilities of Procedural Audio in Computer Games." *Journal of Gaming and Virtual Worlds* 5(3):215–234.

Carello, C., K. Anderson, and A. Kunkler-Peck. 1998. "Perception of Object Length by Sound." *Psychological Science* 9(3):211–214.

Chaigne, A., and V. Doutaut. 1997. "Numerical Simulations of Xylophones, I: Time-Domain Modeling of the Vibrating Bars." *Journal of the Acoustical Society of America* 101(1):539–557.

Conan, S., et al. 2012. "Perceptual Differences between Sounds Produced by Different Continuous Interactions." In *Proceedings of the Acoustics 2012 Nantes Conference*, pp. 581–586.

Conan, S., et al. 2014. "A Synthesis Model with Intuitive Control Capabilities for Rolling Sounds." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22(8):1260–1273.

Danna, J., et al. 2013. "Handwriting Movement Sonification for the Rehabilitation of Dysgraphia." In *Proceedings of the International Symposium on Computer Music Multidisciplinary Research*, pp. 200–208.

Dubus, G., and R. Bresin. 2013. "A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities." *PLOS ONE* 8(12):e82491.

Falcon, E., et al. 1998. "Behavior of One Inelastic Ball Bouncing Repeatedly Off the Ground." *The European Physical Journal B: Condensed Matter and Complex Systems* 3(1):45–57.

Farnell, A. 2010. *Designing Sound*. Cambridge, Massachusetts: MIT Press.

Gaver, W. 1993. "How Do We Hear in the World? Explorations in Ecological Acoustics." *Ecological Psychology* 5(4):285–313.

Gibson, J. J. 1966. *The Senses Considered as Perceptual Systems.* Boston: Houghton Mifflin.

Gibson, J. J. 1979. *The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin.

Giordano, B., and S. McAdams. 2006. "Material Identification of Real Impact Sounds: Effects of Size Variation in Steel, Glass, Wood, and Plexiglass Plates." *Journal of the Acoustical Society of America* 119(2):1171–1181.

Gounaropoulos, A., and C. Johnson. 2006. "Synthesising Timbres and Timbre-Changes from Adjectives/Adverbs." In F. Rothlauf, et al., eds. *Applications of Evolutionary Computing*. Berlin: Springer, pp. 664–675.

Hermes, D. 1998. "Synthesis of the Sounds Produced by Rolling Balls." Internal IPO report no. 1226, IPO, Center for User-System Interaction, Eindhoven, The Netherlands.

Hoffman, M., and P. Cook. 2006. "Feature-Based Synthesis: Mapping Acoustic and Perceptual Features onto Synthesis Parameters." In *Proceedings of the International Computer Music Conference*, pp. 536–539.

Houben, M. 2002. "The Sound of Rolling Objects, Perception of Size and Speed." PhD dissertation, Technische Universiteteit, Eindhoven.

Hunt, K., and F. Crossley. 1975. "Coefficient of Restitution Interpreted as Damping in Vibroimpact." *Journal of Applied Mechanics* 42(2):440–445.

Johnson, K. L. 1987. *Contact Mechanics*. Cambridge: Cambridge University Press.

Klatzky, R., D. Pai, and E. Krotkov. 2000. "Perception of Material from Contact Sounds." *Presence: Teleoperators and Virtual Environments* 9(4):399–410.

Lagrange, M., G. Scavone, and P. Depalle. 2010. "Analysis/Synthesis of Sounds Generated by Sustained Contact between Rigid Objects." *IEEE Transactions on Audio, Speech, and Language Processing* 18(3):509–518.

Lakatos, S., S. McAdams, and R. Caussé. 1997. "The Representation of Auditory Source Characteristics: Simple Geometric Form." *Attention, Perception, and Psychophysics* 59(8):1180–1190.

Le Groux, S., and P. Verschure. 2008. "Perceptsynth: Mapping Perceptual Musical Features to Sound Synthesis Parameters." In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 125–128.

Lee, J., P. Depalle, and G. Scavone. 2010. "Analysis/Synthesis of Rolling Sounds Using a Source-Filter Approach." In *Proceedings of the International Conference on Digital Audio Effects*, pp. 483–487.

Lemaitre, G., and L. Heller. 2012. "Auditory Perception of Material Is Fragile While Action Is Strikingly Robust." *Journal of the Acoustical Society of America* 131(2):1337–1348.

Lloyd, D. B., N. Raghuvanshi, and N. K. Govindaraju. 2011. "Sound Synthesis for Impact Sounds in Video Games." In *Proceedings of the Symposium on Interactive 3D Graphics and Games*, pp. 55–62.

Mathews, M., and J. Smith. 2003. "Methods for Synthesizing Very High Q Parametrically Well Behaved Two Pole Filters." In *Proceedings of the Stockholm Musical Acoustics Conference*, pp. 405–408.

McAdams, S., and E. Bigand. 1993. *Thinking in Sound: The Cognitive Psychology of Human Audition*. Oxford: Oxford University Press.

Merer, A., et al. 2013. "Perceptual Characterization of Motion Evoked by Sounds for Synthesis Control Purposes." *ACM Transactions on Applied Perception* 10(1):1–24.

Michaels, C., and C. Carello. 1981. *Direct Perception*. Englewood Cliffs, New Jersey: Prentice-Hall.

Micoulaud-Franchi, J., et al. 2011. "Categorization and Timbre Perception of Environmental Sounds in Schizophrenia." *Psychiatry Research* 189(1):149–152.

Ninness, B., A. Wills, and S. Gibson. 2005. "The University of Newcastle Identification Toolbox (UNIT)." In *Proceedings of the World Congress of the International Federation of Automatic Control*. Available online at ww.sigpromu.org/reports/EE05010.pdf. Accessed June 2014.

Nordborg, A. 2002. "Wheel/Rail Noise Generation Due to Nonlinear Effects and Parametric Excitation." *Journal of the Acoustical Society of America* 111(4):1772–1781.

Papetti, S., F. Avanzini, and D. Rocchesso. 2011. "Numerical Methods for a Nonlinear Impact Model: A Comparative Study with Closed-Form Corrections." *IEEE Transactions on Audio, Speech, and Language Processing* 19(7):2146–2158.

Rath, M. 2004. "Interactive Realtime Sound Models for Human–Computer Interaction: A Sound Design Concept and Applications." PhD dissertation, University of Verona, Department of Computer Science.

Rath, M., and D. Rocchesso. 2004. "Informative Sonic Feedback for Continuous Human–Machine Interaction: Controlling a Sound Model of a Rolling Ball." *IEEE Multimedia Special on Interactive Sonification* 12(2):60–69.

Rath, M., and D. Rocchesso. 2005. "Continuous Sonic Feedback from a Rolling Ball." *IEEE Multimedia* 12(2):60–69.

Rocchesso, D., and F. Fontana. 2003. *The Sounding Object*. Florence: Mondo Estremo.

Rodger, M. W., W. R. Young, and C. M. Craig. 2013. "Synthesis of Walking Sounds for Alleviating Gait Disturbances in Parkinson's Disease." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 22(3):543–548.

Stoelinga, C., and A. Chaigne. 2007. "Time-Domain Modeling and Simulation of Rolling Objects." *Acta Acustica United with Acustica* 93(2):290–304.

Thoret, E., et al. 2013. "Controlling a Non Linear Friction Model for Evocative Sound Synthesis Applications." In *Proceedings of the International Conference on Digital Audio Effects*, pp. 224–230.

Thoret, E., et al. 2014. "From Sound to Shape: Auditory Perception of Drawing Movements." *Journal of Experimental Psychology: Human Perception and Performance* 40(3):983–994.

Tucker, S., and G. Brown. 2002. "Investigating the Perception of the Size, Shape and Material of Damped and Free Vibrating Plates." Technical Report CS-02-10. University of Sheffield, Department of Computer Science.

van den Doel, K., P. Kry, and D. Pai. 2001. "Foleyautomatic: Physically-Based Sound Effects for Interactive Simulation and Animation." In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 537–544.

van den Doel, K., and D. K. Pai. 2003. "Modal Synthesis for Vibrating Objects." In K. Greenebaum and R. Barzel, eds. *Audio Anecdotes III: Tools, Tips, and Techniques for Digital Audio*. Wellesley, Massachusetts: A K Peters, pp. 99–120.

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

Wildes, R. P., and W. A. Richards. 1988. "Recovering Material Properties from Sound." In W. A. Richards, ed. *Natural Computation*. Cambridge, Massachusetts: MIT Press, pp. 356–363.

Ystad, S., and T. Voinier. 2001. "A Virtually Real Flute." *Computer Music Journal* 25(2):13–24.

Zwicker, E., and H. Fastl. 1990. *Psychoacoustics: Facts and Models.* Berlin: Springer.