NAVIGATING IN A SPACE OF SYNTHESIZED INTERACTION-SOUNDS: RUBBING, SCRATCHING AND ROLLING SOUNDS

S. Conan, E. Thoret, M. Aramaki, O. Derrien, C. Gondre, R. Kronland-Martinet, S. Ystad

Laboratoire de Mécanique et d’Acoustique
Aix-Marseille Univ.
Centrale Marseille
Marseille, France
conan@lma.cnrs-mrs.fr

ABSTRACT

In this paper, we investigate a control strategy of synthesized interaction-sounds. The framework of our research is based on the action/object paradigm that considers that sounds result from an action on an object. This paradigm presumes that there exists some sound invariants, i.e. perceptually relevant signal morphologies that carry information about the action or the object. Some of these auditory cues are considered for rubbing, scratching and rolling interactions. A generic sound synthesis model, allowing the production of these three types of interaction together with a control strategy of this model are detailed. The proposed control strategy allows the users to navigate continuously in an “action space”, and to morph between interactions, e.g. from rubbing to rolling.

1. INTRODUCTION

Synthesis of everyday sounds is still a challenge and especially the control of sound synthesis processes. Indeed, it is of interest to control intuitively the sounds obtained with a synthesis model, i.e. to be able to create sounds that carry a specific evocation, information. To achieve this, we need to offer the users the possibility to create sounds from semantic descriptors of the sound events or from gestures. Previous studies by Hoffman and Cook [1] have approached this problem by proposing intuitive control of sounds through acoustic descriptors or features (so-called feature synthesis) and an intuitive synthesizer of impact sounds has been developed by Aramaki et al. [2] to control the perceived material, size and shape of impacted objects.

This paper is devoted to the synthesis of continuous-interaction sounds. By continuous interaction we mean any kind of friction [3] or rolling. In this paper, we look at a subset of continuous interaction sounds: rubbing, scratching and rolling. Synthesis models for such sounds have already been proposed in previous studies, some based on physical modeling or physically informed considerations [4, 5, 6, 7, 8], others on analysis-synthesis schemes [9, 10]. The synthesizer proposed in [2] allows the user to control the perceived impacted material sounds and to morph continuously from one material to another (e.g. from glass to metal trouch a continuum of ambiguous materials). We would like to control the synthesis of interaction sounds in the same way, (e.g. being able to synthesize a rubbing sound and slowly transform it into a rolling one). Such a tool could be useful for sound design and at a more fundamental level be used to study sound perception as done by Aramaki et al. [11] and Micoulaud et al. [12].

We adopted the {action/object} paradigm in order to propose such a control strategy of interaction-sounds synthesis. This paradigm is based on the assumption that a sound results from an action on an object, e.g. “plucking a metal string” or “hitting a wood plate”, as proposed in [4]. The underlying hypothesis is the existence of sound invariants, i.e. sound signal morphologies that carry specific meanings about the action or the object. For instance, a string produces a particular spectrum that allows the listener to recognize it, even if it is bowed (violin), plucked (guitar) or hit (piano). Likewise, it is possible to recognize that a cylinder bounces even if it is made of glass, wood or metal [13]. This is inspired from the ecological theory of visual perception as proposed by Gibson [14] (for a more accessible introduction to Gibson’s theory, refer to [15]), and formalized as an explanation of sound sources recognition by McAdams [16].

Some studies have already identified sound invariants. For instance, to a certain extent, it has been shown that impact sounds contain sufficient information to discriminate the materials of impacted objects [17]. Other authors showed that the perceived material is mainly related to the damping of spectral components [18, 19, 20] and to sound roughness [11]. A study by Warren and Verbrugge revealed that from the rhythm of a series of impact sounds, it is possible to predict if a glass will break or bounce [21].

In this paper, we will focus on sound invariants related to the actions of rubbing, scratching and rolling to propose an intuitive control of these actions. For this purpose we will first identify some invariants related to the auditory perception of these interactions. Then, we will describe how the concerned interactions can be reproduced by synthesis. Further, a control strategy of the proposed synthesis model is presented. Then in the last section we give some general conclusion and propose future research.

2. INVARIANTS

In this section, we examine sound signal morphologies that are related to the perception of rubbing, scratching and rolling. The results on rubbing and scratching sounds are based on preliminary studies by Conan et al. [22], which are briefly recalled below [23].
2.1. Rubbing and Scratching Interactions

To our knowledge, the auditory ability to distinguish rubbing sounds from scratching sounds has not been investigated yet. A listening test with recorded sounds was therefore set up to identify signal properties that enable to distinguish between rubbing and scratching from the auditory point of view.

- **Participants.** 14 participants took part in the experiment (4 women, 10 men, mean age=30.64 years; SD=12.05).

- **Stimuli.** Twenty monophonic recordings of friction sounds on ten different surfaces (2 different sounds on each surface, one by acting with the fingertips, the other with the nails) were recorded with a Roland R05 recorder at 44.1 kHz sampling rate.

- **Apparatus.** The listening test interface was designed using MAX/MSP and sounds were presented through Sennheiser HD-650 headphones.

- **Procedure.** Subjects were placed at a desk in front of a computer screen in a quiet room. They were informed that they were to classify twenty sounds in two categories, rubbing and scratching. Before the session started, the twenty stimuli were played once. Then, the subjects had to evaluate the evoked action for each sound by classifying each sound in one of the two categories “rub” or “scratch” using a drag and drop graphical interface. They could listen to each stimulus as often as they desired. No time constraints were imposed and sounds were placed in a random position on the graphical interface across subjects.

Results are presented in figure 1. Three sounds were 100% associated to scratching (number 11, 16, 17) and six sounds were 100% associated to rubbing (number 3, 4, 7, 9, 12, 20).

Qualitative signal analysis led us to suppose that, if considering that friction sounds result from successive impacts of a plectrum on an irregular surface, rubbing sounds result from a higher temporal density of impacts than scratching sounds (i.e. impacts occur more frequently in rubbing sounds than in scratching sounds). A sound signal that is 100% associated to scratching and of a sound that is 100% associated to rubbing are plotted in figure 2 (left and right parts respectively). The differences can be explained as follows. On one hand, scratching a surface results from a deeper scan of the surface, which implies encountering each surface’s irregularity one after another and more intensely. On the other hand, rubbing a surface results from a soft contact between the finger and several surface irregularities, implying a more noisy and stationary sound because of the higher density of impacts (see figure 3).

In summary, rubbing and scratching interactions can be simulated by series of impacts on a surface. One possible invariant (among others), that contributes to the discrimination between these two interactions is the temporal density of the impacts, i.e. the more (respectively the less) impacts occur in the sound signal, the more the sound is related to rubbing (respectively to scratching).

To validate the previous hypothesis made from the qualitative observations, an experiment with controlled synthetic stimuli was set up. The stimuli were synthetic friction sounds generated with different impact densities to test the hypothesis. The synthesis model was based on the pioneering work of Gaver [4], improved by Van den Doel et al. [6] and consisted in simulating the sound as a result of successive micro-impacts of a plectrum on the asperities of a surface. The successive impacts were modeled by a low pass filtered noise with a cutoff frequency related to the gesture velocity while the roughness of the surface was controlled by the nature of the noise. By generating low to high density impact series, continuous transitions between scratching and rubbing were then generated. The synthesized stimuli were presented randomly to the subjects who were to associate each stimulus to one of the two interaction categories (rubbing or scratching). The results confirmed the empirical hypothesis: small impact densities are associated to scratching and high impact densities to rubbing (more details in [22]).

In summary, rubbing and scratching interactions can be simulated by series of impacts on a surface. One possible invariant (among others), that contributes to the discrimination between these two interactions is the temporal density of the impacts, i.e. the more (respectively the less) impacts occur in the sound signal, the more the sound is related to rubbing (respectively to scratching).

2.2. Rolling Interaction

The aim of the previous section was to highlight the perceptual information which is relevant to perceptually differentiate rubbing interactions from scratching, i.e. the acoustical morphology which characterized these interactions. In this section we address the same question concerning the auditory perception of rolling: what is the signal information which allows us to recognize a rolling ball? To answer this question, we investigate a physical model of rolling.

In the literature, most authors consider that the physics of a rolling ball is close to the physics of a bouncing ball. The models

---

**Figure 1:** Results of the experiment with recorded sounds. On the X-axis, the number of the sound. The Y-axis represents the percentage of association to scratching for each sound.

**Figure 2:** Left: A recorded sound 100% associated to scratching. Right: A recorded sound 100% associated to rubbing.

**Figure 3:** Left: A nail which is scratching a surface. Right: A finger which is rubbing a surface (asperities which are not “seen” by the finger are circled). For clarity, dimensions of the asperities are exaggerated on the scheme.
Simulated rolling force (a zoom on the right figure). The force parameters are $\alpha = \gamma = 3/2$, $k = 10^3 \text{N/m}^{3/2}$, $\lambda = 10^7 \text{N s/m}^{3/2}$. The ball has a mass of 5g and a velocity of 0.5m/s. The surface is assumed to be fractal with $\beta = 1.2$ and a maximum amplitude of $10^{-9} \text{m}$.

Procedures generally take 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:

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

(1)

where $k$ is the stiffness and $\lambda$ damping weight of the force. By taking into account the effect of the gravity on the ball, this model can simulate the behavior of a bouncing ball (see for instance [23] and [24] for an application to bouncing sound synthesis).

Studies that addressed the simulation of a rolling ball [7, 8] or a rolling wheel [25] consider that the rolling object moves along an irregular surface. The height of this irregular surface is added as a perturbation to the penetration term $x$ in equation (1). It can be seen as a bouncing ball whose height changes randomly.

We simulated the rolling of a ball over an irregular surface (we do not consider here the vibration of the surface as in [21]) thanks to a Fourth-Order Runge-Kutta scheme (as described in [26], where numerical issues of the model [4] are studied, instead of the K-method [22]). An example of a simulated nonlinear interaction force $f$ is plotted in figure 4 where the used parameters are listed.

In the numerical model, the surface is modeled by a noise with a specific spectrum adjusted according to tribological observations [30] and phenomenological considerations [6, 31]. In practice, the spectrum is $S(\omega) \propto 1/\omega^3$. Also called fractal noise, such a spectrum models accurately most surfaces. $\beta$ controlled the perceived roughness, and the amplitude was normalized to provide a maximum of $10^{-3} \text{m}$.

From informal listening tests, it has been concluded that the force $f$ evokes a rolling object such as a small hard marble. One can note that this force can be considered as a series of impacts, see figure 4. This assumption has already been exploited by Hermes [5] for sound synthesis purposes and by Lagrange et al. [9] in an analysis-synthesis context. The first observation is that the temporal structure of the impact series seems to follow a specific pattern (see figure 4 left). Let us consider $(A, \Delta \tau)$ as time series, with $A^n$ and $\Delta \tau^n$ respectively the amplitude of the $n^{th}$ impact and the time interval between the $n^{th}$ and the $(n+1)^{th}$ impact (see figure 5 left). The Fourier spectrum of $\Delta \tau$ time series is plotted in figure 5 right. One can note that this time series has a strong auto-correlation, i.e. that successive impacts have strong mutual dependencies. The amplitude series $A$ exhibits the same behavior. $A$ and $\Delta \tau$ are also strongly cross-correlated as can be seen in figure 6. These observations are coherent with the physics of a bouncing ball (recalling that the rolling model is derived from a bouncing model) which relates the impacts to each other.

Another important aspect is the fact that the contact time of the impact depends on the impact velocity, related to the amplitude of the impact in $f$. This dependency has been studied by several authors [32, 24] and seems to be an important auditory cue that is responsible for the rolling interaction evocation. Indeed, to informally test this assumption, we first detected all the impact time locations and amplitudes in the simulated force in figure 4. We then created two modified versions of this force signal by replacing the impact windows by a raised cosine window [33] that fits well the original simulated impacts, one with a duration depending on the impact amplitude, the other with a fixed duration. For different fixed durations, it was always found that the version with varying impact durations depending on the impact amplitudes clearly produced the most realistic evolutions of the rolling interaction, which confirms previous findings.

2.3. Invariants Related to Continuous Interactions

As described in this section, rubbing, scratching and rolling interactions all can be represented as impacts series. Thanks to listening tests, it was shown that the temporal distribution of successive impacts convey the information of rubbing or scratching sounds. For the rolling interaction, the temporal structure of the impacts (their correlations and statistics) may also convey the rolling information. As opposed to rubbing and scratching interactions, a specific relation between the impact durations and amplitudes seems to be an important signal morphology that conveys information about the rolling interaction. These assumptions need to be formally verified but this is out of the scope of this paper. In the next section, we will propose a model to simulate these impact series.

---

Footnote 1: When $\alpha = \gamma$, the model is the Hunt and Crossley one [28, 29].
3. IMPLEMENTATION OF A GENERIC MODEL OF CONTINUOUS INTERACTIONS

As described in the previous section, the force applied to a surface carries the information about the type of interaction. For rubbing, scratching and rolling, the forces are series of impacts, with specific relations between each other. The aim of this section is to propose a synthesis process which is generic enough to simulate different continuous interactions and navigate continuously between them. In order to cohere with the \{action/object\} paradigm proposed in the section\[1\] the force previously described is modeled as the source part of a source-filter model. In fact, by convolving the force with an impulse response of a resonant surface, the resulting sound conveys the information on both the action and the object. The object properties such as material and shape are included into the filter part of the source-filter model (as proposed by Gaver \[4\]). The implemented filters are described in \[34\]. The control strategy is similar to the one proposed by Aramaki et al. \[2\], which enables to navigate continuously between the perceived materials. Aspects linked to the perceived object are not presented here.

3.1. Modeling the Source Signal

As presented in the section\[2\] the force that conveys the information about the interaction type (rolling, scratching or rubbing) can be considered as impacts series. The specific behavior of the time series $A^n$ and $\Delta^n_T$, respectively representing the amplitude of the $n^{th}$ impact and the time interval between the $n^{th}$ and the $(n+1)^{th}$ impact, seems to be an important cue to discriminate these interactions. We will experimentally characterize the series $^A A$ or $^\Delta A_T$, centered version of $A$ and $\Delta_T$, in order to propose a synthesis scheme.

As pointed out in \[2\] $A$ and $\Delta_T$ can be strongly autocorrelated (so are $^A A$ or $^\Delta A_T$). We consider these two time series as autoregressive moving average (ARMA) processes. To characterize their behavior, these series are whitened (the method that we use is described in \[15\]), and we experimentally noted that no more than 2 poles - 2 zeros whitening filters are needed. So let $X$ be one of the two processes $^A A$ or $^\Delta A_T$, we can write in the $z$ domain:

$$X(z) \approx H_X(z) \bar{X}(z), \quad H_X(z) = \frac{B(z)}{A(z)} \tag{2}$$

where $B(z) = 1 + \sum_{i=1}^{\infty} b_i z^{-i}$, $A(z) = \sum_{i=0}^{\infty} a_i z^{-i}$ and $\bar{X}$ is the whitened version of $X$. Then, as $\bar{X}$ is white, we can model its probability density function, and transform $\bar{X}$ into $W$ which follows a uniform law, thanks to the inverse transform sampling method (i.e. $W = F_X(\bar{X})$, where $F_X(\cdot)$ is the cumulative distribution function of $\bar{X}$). The analysis scheme is presented in figure 7.

Results of the analysis on $\Delta_T$ time series (from the force $f$ computed in the section\[2\]) are presented on top and middle part of figure 8 (results on $A$ series are similar and therefore not displayed). The top figure displays the autocorrelation of $^\Delta A_T$, whitened version of $^\Delta A_T$, showing that a 2 poles - 2 zeros whitening filter is well suited to whiten the time series. The probability density of $^\Delta A_T$ is plotted in the middle figure, showing that for roll force, $^\Delta A_T$ and $^A A$ can be modeled as gaussian processes. In the bottom figure, $C(W_A, W_{\Delta_T})(k)$ the cross correlation between $W_A$ and $W_{\Delta_T}$ is displayed.

![Figure 7](image7.png)

**Figure 7:** Analysis scheme of $A^n$ and $\Delta^n_T$ series.

![Figure 8](image8.png)

**Figure 8:** Top: $^\Delta A_T$ autocorrelation. Middle: probability density of $^\Delta A_T$ (grey bars are the measure and the black line is the gaussian fit). Bottom: cross correlation between $W_A$ and $W_{\Delta_T}$.
In this section, we will detail our control strategy for continuous navigation between rubbing, scratching and rolling. First, the prototypes of each interaction will be given. Then, we will present the navigation strategy between these prototypes.

Let us first recall the low level parameters of the model:

- **Impact Model.** Two parameters control the impact duration (eq. 5) in the chosen impact model (eq. 4): \( \mu \) and \( \nu \).
- **Probability Density.** The probability density is defined as a list of discrete values, which are used to derive the cumulative distribution function \( F_X \) (cumulative sum of the probability density), leading to obtain \( \hat{\Delta}A \) and \( \hat{\Delta}_T \) series (see figure 4). These lists, respectively for \( \hat{\Delta}A \) and \( \hat{\Delta}_T \) series will be noted \( P_\Delta^A \) and \( P_\Delta^T \).
- **Filters.** As pointed out in 3, no more than 2 poles and 2 zeros are needed for the filters. Then, each filter will be described by a set of 5 coefficients \((a_0, a_1, a_2, b_1, b_2)\), each set will be noted \( C_A \) and \( C_{\Delta_T} \), respectively for \( H_A(.) \) and \( H_{\Delta_T}(.) \).
- **Offset coefficients.** \( \mu_A \) and \( \mu_{\Delta_T} \) respectively for \( A \) and \( \Delta_T \).

The global set of parameters will be called:

\[
\Psi = \{ K, \nu, P_A, P_{\Delta_{\Delta_T}}, C_A, C_{\Delta_T}, \mu_A, \mu_{\Delta_T} \}.
\]

4.1. Interaction Sounds Prototypes

4.1.1. Rubbing Prototype

As showed in 22 and recalled in section 2.1 small intervals between impacts are associated to rubbing sounds (a maximum impact density implies a source signal which is a white noise). For the rubbing prototype, we then tune the parameters to obtain a gaussian white noise: setting the coefficients of \( C_A \) and \( C_{\Delta_T} \) to zero for \( i \geq 1 \), \( \mu_A = \mu_{\Delta_T} = 0 \), \( \nu = 0 \) and \( K = 1 \) (to get an impact duration of 1 sample), \( P_\Delta^A \) to get 1 impact at each sample and \( P_\Delta^T \) following a gaussian distribution. This set of parameters will be called \( \Psi_{\text{rub}} \).

Following the model proposed by Van den Doel et al. 6, the generated source signal is lowpass filtered with a cutoff frequency which is directly related to the relative transversal velocity between the object that rubs (hand, plectrum...) and the rubbed surface. This enabled the simulation of the sound producing gesture when associated to a biological gesture law 36.

4.1.2. Scratching Prototype

The scratching prototype is associated to a low impact density, i.e. to high \( \Delta_T^{\nu} \) values and, as proposed empirically in 22, the \( \Delta_T \) series follows an exponential distribution. The rest of the parameters set is the same as in the rubbing prototype. Again, the velocity between the interacting objects controls the cutoff frequency of a lowpass filter. The parameter set will be called \( \Psi_{\text{scratch}} \).
On the circumference of the circle, a sound (wood, metal and glass). Eecters are weighting of the three prototypes of perceived material space to get the desired impact sound, and the synthesis parameters, are placed on the border of a circle which represent the so-determined thanks to behavioral and electrophysiological experi-
ters. Figure 10: Schematic control space of the interaction sound syn-
thesizer.

4.1.3. Rolling Prototype

In section 4.1.2 we developed our synthesis model based on an experimental analysis of a simulated rolling force. We saw that \( P_A \) and \( P_{\Delta r} \) are gaussian distributions. The filter coefficients \( C_A \) and \( C_{\Delta r} \) are here non zero, and we take \( \nu = 1/5 \). As the whole rolling model necessitates a more precise perceptual calibration that requires an individual study, we will only give qualitative remarks on the parameters’ influence in this article. Presently, the complete perceptual mapping of the rolling model is not fully established and is left for further studies, so for now the synthesis parameters are controlled by the following observations. As \( K \) controls the contact time, it is related to the size of the rolling ball. The mean value of \( P_{\Delta r} \) is related to the perceived velocity. The filters coefficients \( C_A \) and \( C_{\Delta r} \) are linked to surface parameters such as roughness. As in the scratching and rubbing prototypes, the resulting source signal is also lowpass filtered according to the ball’s transversal velocity. The parameter set will be called \( \Psi_{\text{roll}} \).

4.2. Navigation Strategy

The control strategy that we adopt is based on the one proposed by Aramaki et al. to control the perceived material in an impact sound synthesizer \cite{43}. In this paper, the authors proposed to allow the users to navigate continuously in a material space that comprises wood, metal and glass. The three prototypes of perceived material (wood, metal and glass), which synthesis parameters were determined thanks to behavioral and electrophysiological experiments, are placed on the border of a circle which represent the so-called material space. Hereby, the user can move a cursor in this space to get the desired impact sound, and the synthesis parameters are weighting of the three prototypes of perceived material (wood, metal and glass).

A similar control space of interaction is schematized in figure 10. On the circumference of the circle, a sound \( S \), characterized by its angle \( \theta \) in the control space, is generated with the parameters \( \Psi_S (\theta) \) as follows:

\[
\Psi_S (\theta) = T(\theta) \Psi_{\text{rub}} + T \left( \theta - \frac{2\pi}{3} \right) \Psi_{\text{scratch}} + T \left( \theta - \frac{4\pi}{3} \right) \Psi_{\text{roll}}
\]

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

\[
T(\theta) = \begin{cases} 
-\frac{1}{\pi} \theta + 1 & \text{, } \theta \in [0; \frac{2\pi}{3}] \\
0 & \text{, } \theta \in [\frac{2\pi}{3}; \frac{4\pi}{3}] \\
\frac{1}{\pi} \theta - 2 & \text{, } \theta \in [\frac{4\pi}{3}; 2\pi] 
\end{cases}
\]

Inside the circle, a sound \( S' \), characterized by its angle \( \theta \) and radius \( r \), is generated with the parameters \( \Psi_{S'} \):

\[
\Psi_{S'} = (1 - r) \Psi_C + r \Psi_S (\theta)
\]

where

\[
\Psi_C = \frac{1}{3} (\Psi_{\text{rub}} + \Psi_{\text{scratch}} + \Psi_{\text{roll}})
\]

and \( \Psi_S (\theta) \) defined in equation 6. Even if this control space is not perceptually validated and calibrated, it yields good results (see the demonstration video on the webpage associated to this paper: \url{http://www.lma.cnrs-mrs.fr/~kronland/InteractionSpace}).

5. CONCLUSIONS

In this paper, we described specific signal morphologies that are related to the auditory perception of rubbing, scratching and rolling interactions. Both phenomenological considerations and physical modeling were investigated as well as qualitative signal analysis. A generic synthesis model of these three interactions that enables continuous transitions between the categories has been proposed. Although a precise perceptual calibration is still to be carried out, the control is already convincing.

Further studies will be done to expand this control space to other interactions such as nonlinear friction (squeaking, squealing...) \cite{38}. The rubbing and scratching models are also quite simple and could be improved by an analysis method as proposed by Lagrange et al. \cite{9}. Finally, the influence of the physical velocity profiles (rolling marble in a bowl, sliding object on an inclined plate...) is to be studied, as already done by \cite{16} with the velocity of human gestures.

6. REFERENCES


