Perceptual differences between sounds produced by different continuous interactions

S. Conan, M. Aramaki, R. Kronland-Martinet, E. Thoret and S. Ystad

Laboratoire de Mécanique et d'Acoustique, 31, Chemin Joseph Aiguier - 13402 Marseille
Cedex 20
conan@lma.cnrs-mrs.fr

The presented study is part of a general framework consisting in designing an intuitive control strategy of a generic synthesis model simulating continuous interaction sounds such as scratching, sliding, rolling or rubbing. For that purpose, we need to identify perceptually relevant signal properties mediating the recognition of such sound categories. Some studies tend to suggest the existence of acoustic features related to velocity or periodic modulations but, to our knowledge, the auditory distinction between these interactions is still not well-known and no formal investigations have been conducted so far. This study aimed at presenting a perceptual evaluation method, with brings to the light the differences between two continuous friction sounds, rubbing and scratching. By a perceptual experiment with recorded sounds, we noted that listeners unanimously classified some sounds in one category or the other. Then, an analysis of the recorded signals let us hypothesize on a feature that may be responsible of this distinction. This hypothesis on a characteristic morphology was tested by synthesizing sounds and using it in a second perceptual experiment. Results support that this typical morphology is responsible for the evocation of rubbing or scratching, thereby useful for the design of intuitive control of the generic synthesis model.

1 Introduction

Sound synthesis of realistic everyday sounds has been studied for plenty of years, and almost everyday sounds are nowadays synthesizable with different classes of models such as physical or signal models. Nevertheless they do not provide easy controls for non expert users. Therefore, an important issue is to give easier controls based on semantic descriptions. For instance, different studies focused on impact sound synthesis and how to control these models without controlling all the complex sets of parameters of the physical model. In [1, 2, 3], Aramaki et al. developed a physically based model of impact sounds and set up perceptual experiments which allow controlling the generation of impact sounds by describing the perceived material (e.g. wood, metal, glass), the shapes of the objects (e.g. plate, bar, string) and the size of the object which is impacted.

In this study we will focus on continuous interactions, and particularly rubbing ("to rub" = french word "frotter") and scratching ("to scratch" = french word "gratter"). Sound synthesis of such friction sounds have been studied by Gaver [4] and Van den Doel et al. [5]. They proposed a physically based model which enables to synthesize realistic friction sounds. The control parameters of this model do not provide a way to morph between rubbing and scratching. The present study aims at investigating these two interactions with a phenomenological approach of the underlying physical phenomenon. The goal is to propose, in a synthesizer, an intuitive control of the interaction perceived through the sound in a continuous way.

Ecological acoustic theories [6, 7, 8] have highlighted the relevance of different perceptual attributes related to an action from the auditory point of view. Perceptual invariants were defined to explain the perceptual process of such sonic events. According to their definition, the perception of an action is linked to the perception of a *transformational* invariant, i.e. sound morphologies which are linked to the action. In [9], Warren et. al. highlighted that the perception of breaking or bouncing events was linked to the specific rhythms of the sequences of impacts involved by such sound events. Here the differences between rubbing and scratching are studied with the same behavioral approach. Scratching and rubbing correspond to a sustained contact between a plectrum and a surface. The auditory differences seem due to the type of interaction between the plectrum and the surface. In this study, we propose to investigate these perceptual differences by searching a transformational invariant linked to each type of friction.

First, we will investigate listeners abilities to distinguish these two interactions from recorded sounds with a forced categorization experiment. In a second time, signal differences will be investigated with respect to the underlying physical phenomenon. Then a sound synthesis model with an associated control strategy enabling to generate rubbing and scratching sounds will be proposed. In a fifth section, a perceptual morphing based on the variation of the roughness of the surface will investigate the possibility to classify the two interactions according to the variation of only one control parameter. At last, the perspectives opened by this study will be discussed.

2 Perceptual categorization of recorded sounds

To highlight the possibility to distinguish rubbing from scratching sounds from the auditory point of view, a listening test with recorded sounds is set up[1].

2.1 Method

Participants 14 partipants took part in the experiment (4 women, 10 men, mean aged=30.64 years ; SD=12.05).

Stimuli Twenty monophonic recordings of a person who is rubbing and scratching on different surfaces were recorded with a Roland R05 recorder at 44.1 kHz sampling rate. We made the hypothesis that rubbing an object is like a "distant scanning" of the object's surface, therefore more related to scan the surface with fingertips, and that scratching is a deeper scanning of this object and so it is related to scan the object's surface with the nails (see figure 4, used for the description of the synthesis model in 4.2). To investigate this hypothesis, two recordings were done for each surface, one obtained by interacting on the surface with the fingertips and one another with the nails.

Apparatus The listening test interface was designed using MAX/MSP[2] and sounds were presented through Sennheiser HD-650 headphones.

2.2 Procedure

Subjects were placed at a desk in front of a computer screen in a quiet room. They were informed that they were to
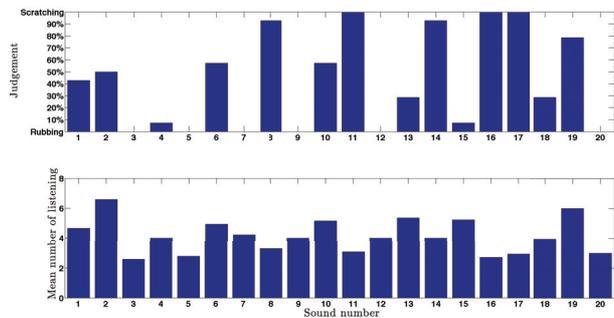

Figure 1: Results of the experiment with recorded sounds. On the X-axis, the number of the sound. Top : Judgement, the Y-axis represents the percentage of association to scratching for each sound. Bottom : Mean number of times each stimulus has been played.

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" in a drag and drop graphical interface. They could listen to each stimulus as many times as they want. No time constraints were imposed and sounds were placed in a random position on the graphical interface across subjects.

2.3 Data Analysis & Results

For each subject, the selected category for each sound is collected. The number of times each sound has been played is also collected. The "scratch" category was arbitrarily associated with the value 1, and the "rub" category with the value 0. For each sound, the values were averaged across subjects and a percentage of association to the rubbed and scratched category was associated.

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). A one-way ANOVA revealed no significant difference in the number of times sound has been played with $F(19, 260) = 0.96$, $p = 0.51$.

2.4 Discussion

Two sets of sounds can be determined, either sounds were associated quasi-exclusively to one category, or sounds led to a more ambiguous categorization with a lower percentage of association in one category.

The high percentage of association obtained for several sounds in each category allowed us to conclude that the rubbing and scratching interactions are distinct interactions from the perceptual point of view and that they can be distinguished. The ambiguity observed for some sounds supports the idea that the perception of these two interactions is not categorial and that some sounds could be assessed in a non-consensual way.

3 Signal analysis

In this section, we will investigate signal properties related to the perceptual differences between the two different
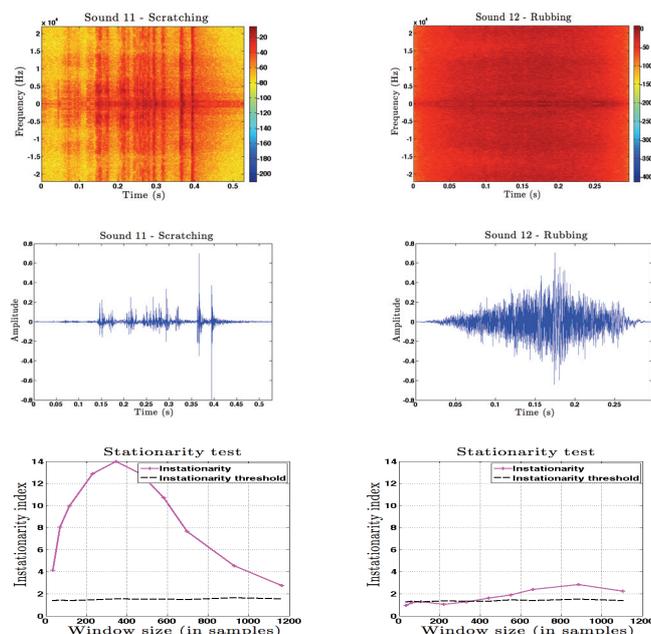

Figure 2: Left column : Sound 11, associated at 100% to scratching. Right column : Sound 12, associated at 100% to rubbing. Top row : Time frequency representation. Middle row : temporal representation. Bottom row : stationarity test on a portion of the signal. The dashed black line represents the calculated stationarity threshold and the magenta line the index of nonstationarity.
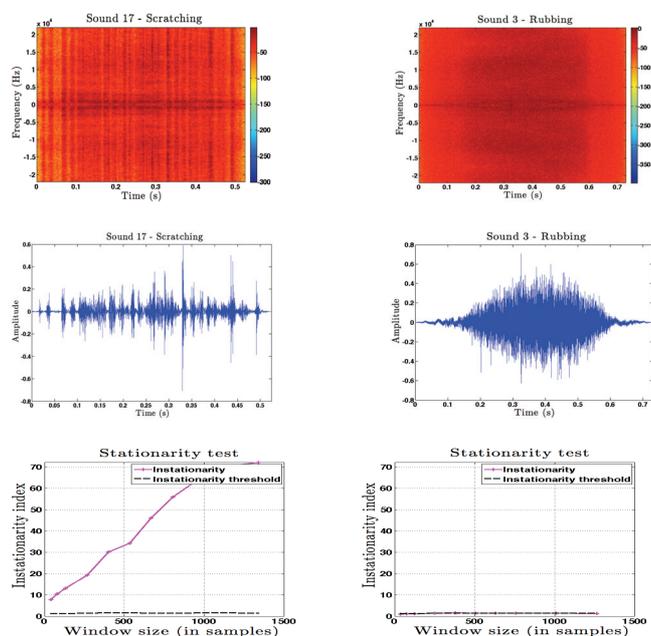

Figure 3: Left column : Sound 17, associated at 100% to scratching. Right column : Sound 3, associated at 100% to rubbing. Top row : Time frequency representation. Middle row : temporal representation. Bottom row : stationarity test on a portion of the signal. The dashed black line represents the calculated stationarity threshold and the magenta line the index of nonstationarity.

411

categories investigated by the experiment presented in the previous section. In a qualitative study, we will focus on two sounds 100% associated to rubbing (sounds 3 and 12) and two sounds 100% associated to scratching (sounds 11 and 17) in order to highlight features specific of rubbing and scratching, see figures 2 and 3 (Sounds were grouped according to the excited surface : sounds 11 and 12 were recordings on a corrugated paper while sounds 17 and 3 were recordings respectively on sandpaper and on a synthetic sofa cover).

The first observation is that sounds which are associated to rubbing seem to be more constant in the time frequency domain. Contrariwise, we can discern more distinct energy peaks in the sounds associated to scratching. Our first assumption was that sounds that are associated to scratching are less stationary than sounds which are associated to rubbing. To test this stationarity, we used the method proposed by Xiao et al. [10]. This method quantifies the stationarity of the signal at different scales (i.e. for different window sizes) by comparing local spectra to the mean spectrum. In practice, for a given window size, the statistical distribution is obtained by considering the spectrum with random phases (to obtain stationarized "surrogates"). This process is repeated for different window sizes and it allows testing the null hypothesis and quantifying the nonstationarity for different scales of observation. As the global envelope of the sound (that we believe associated to the velocity and pressure of the action) strongly influences this test, we performed this test on a small part of each sound, in which the influence of the global amplitude of the sound can be neglected (the duration of each part is at least 50ms, but for some sounds not so much because of their short duration). These tests allows us to note that sounds associated to scratching are less stationary than sounds associated to rubbing. For example, sounds 11 and 12 were recorded on the same surface (corrugated paper), respectively by interacting with the nails and with the fingertips. Although both sounds 11 and 12 are nonstationary (see bottom row of figure 2), the sound associated to scratching is clearly more nonstationary than the rubbing sound.

From this short qualitative analysis, we hypothesize that the perception of scratching an object is due to sparse impacts and the perception of rubbing an object is due to a denser distribution of impulses. These considerations gave us cues to build a synthesis model of rubbing and scratching an object. This model will be presented in the next section.

4 Friction Sound Synthesis

The analysis of recorded friction sounds led us to hypothesize that the perception of the nature of interactions rubbing or scratching interactions is linked to a density of impacts on the surface. To control the effect of impact's density on the perception of a friction sound, we implemented a friction sound synthesis model. This model, firstly proposed by Gaver in [4] and improved by Van den Doel in [5], provides a suitable tool to investigate the friction sounds and their perception. The parameters can be controlled independently from each other. It is therefore possible to generate sounds of different impact's densities.
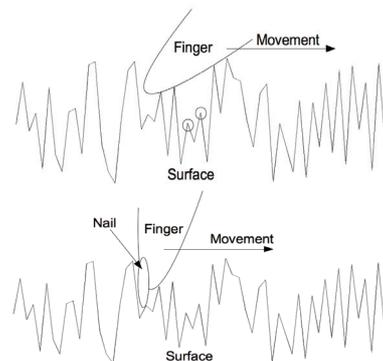

Figure 4: Top : A finger which is rubbing a surface (asperities which are not "seen" by the finger are circled). Bottom : A nail which is scratching a surface.

4.1 Physically Based Model

This model considers that friction sounds are the result of the successive micro-impacts of a plectrum on the asperities of a surface. The velocity of the plectrum directly controls the velocity of occurences of the successive impacts. Otherwise the pressure controls the intensity of each impact. The profile of the surface is modeled by a noise where the heights of the asperities are linked to the roughness of the surface. In the present study the pressure is assumed to be constant. Other controls on the material or the shape of the object are available, but not described here.

In practice, the successive impacts are modeled by a noise low pass filtered with a cutoff frequency linked to the velocity of the gesture. The nature of the noise, which is described in the following section, is controlled by the density of impacts. This control parameter modifies the perception of the interaction (rubbing or scratching).

4.2 Impact's Density Control

As exposed previously, there are noticeable differences between the sound produced when rubbing or scratching it an object. We hypothesized that a major difference is due to the temporal density of impacts : an object which is rubbed would contain a lot of impacts while a scratched one less. When considering someone rubbing an object with his fingertips (see figure 4), the contact between the two interacting surfaces is not very intense, as the fingers don't "reach" each microscopic asperity, and don't interact with one asperity after one another but several at the same time. It could be understood as a "distant scanning" of the surface which results in a constant contact between the two interacting surfaces. When considering scratching a surface with the nails, the contact is more intense and is more like a "deep scanning" of the surface as the nails tend to reach the macroscopic asperities one by one. Hence the sound produced by scratching seems more like a succession of impacts whereas the sound produced by rubbing is more noisy and constant.

In the friction sound synthesis model, these differences could be modeled as a series of impulses with different amplitudes and which are more or less spaced in time : each sample of the impulse series is the result of a Bernoulli process, $0 = no\ impulse$ or $1 = impulse$ with a probability of an impact equal to ρ. The amplitude of each impulse is ran-
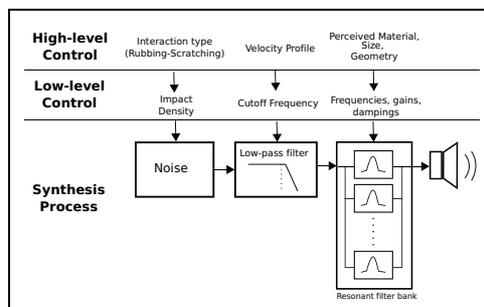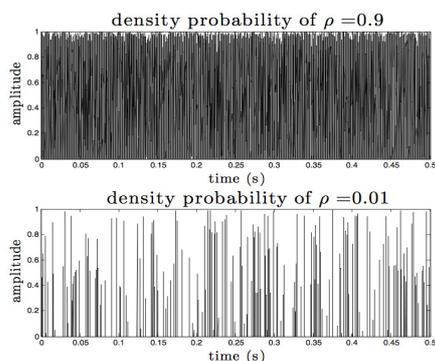
Figure 5: Physically Based Model of Friction

Figure 6: Top : A quasi-constant interaction (high ρ probability of impact), which represents rubbing interaction. Bottom : A more sparse interaction, which represents scratching interaction.

domly affected according to a uniform law. The figure 5 sums up the general scheme of this synthesis model and the control possibilities. Two interaction patterns with different ρ values are represented in figure 6.

5 Perceptual categorization of synthesized sounds

The following experiment has been designed to evaluate the influence of the density parameter ρ (see 4.2) on the perception of the two interactions (rubbing and scraping). A listening test with a 2-AFC procedure was conducted. It supports the hypothesis that the distinction between rubbing and scratching is based on the impact density. These results also highlight the ambiguous perception of this kind of interaction for a range of impact density values[3].

5.1 Method

Participants Thirty five participants (9 women, 26 men, mean aged=30.11 years ; SD=12.01) took place in the experiment. 6 of them also participated in the first experiment. They were all naive about this experiment.

Stimuli Thirty one sounds were synthesized using previous synthesis model with different values of density parameter ρ. We chose $\rho \in [0.001, 1]$, logarithmically spaced. The velocity profile (figure 7) used to control the model was recorded
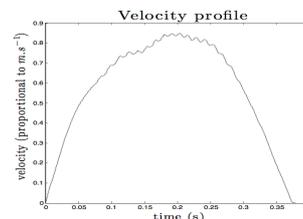

Figure 7: Velocity profile used to generate the stimuli.

using a graphic tablet. The graphic tablet used to record the velocity profile is a Wacom Intuos 3 which records the position of a specific pen (Wacom Grip Pen) at 200 Hz with a spacial resolution of 5.10^{-3} mm. The velocity could then be computed and resampled at the audio rate. To synthesize the vibrating surface, an impulse response of a hard structure (stone) generated by an impact sound synthesizer [2] was used.

Apparatus The listening test interface was designed using Max/MSP and the subjects used Sennheiser HD-650 headphones. Participants took part of the experiment in a quiet office.

5.2 Procedure

The subjects were informed that they were to hear thirty one stimuli and that they would have to judge whether the sounds evoked rubbing or scraping. Before the judgement task, they listened to two distinct stimuli (one with a density probability $\rho = 0.0073$, simulating scratching according to our hypothesis and the other with $\rho = 0.91$, simulating rubbing) which were presented in a random order. The aim of presenting these two examples was to show to the listener what kind of different sounds he would hear, but he was not informed about the kind of interaction associated with each sound.

For the judgement task, the subjects could listen to each stimulus maximum two times in order to determine whether it evoked rubbing or scraping. The 31 stimuli were presented in random order.

5.3 Data Analysis & Results

Results are presented in figure 8. There is a clear perceptual distinction between scratching and rubbing at the extremities of the continuum, and the association of poor impact density with scratching and high impact density with rubbing validate our hypothesis. The perception between the two interactions is not clearly categorical. This less clear area at the intermediate positions on the continuum highlight the ambiguous perception of this kind of interaction for approximately $\rho \in [0.01, 0.09]$. This ambiguous range of density values is also supported by the mean number times the stimuli were played which increases in this area (see figure 8, bottom).

6 Conclusion

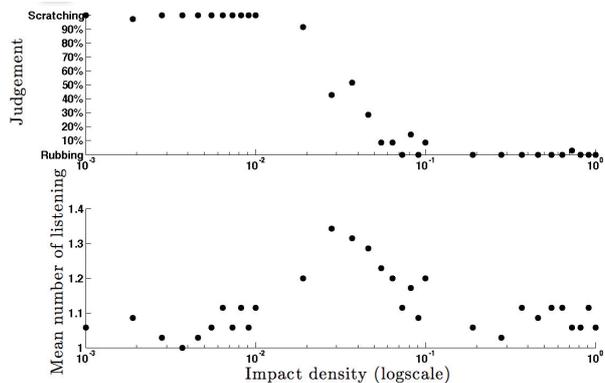In this study, a behavioral approach was used to understand the perception of continuous interactions, especially

---

[3]All stimuli used in this experiment and the interface test are available on http://www.lma.cnrs-mrs.fr/~kronland/RubbingScraping

Figure 8: Results of the experiment. On the X-axis, the impact density ρ is presented on a logscale, increasing from left to right. Top : Judgement, the percentage presented on the Y-axis represents the percentage of association to scratchingfor each sound. Bottom : The Y-axis represents the mean number times each stimulus was played.

the distinction between rubbing and scratching a surface. A listening test was conducted on 20 recorded sounds and results support the fact that listeners could distinguish sounds that rub from sounds that scratch. A qualitative analysis on signal differences gave us cues and let us formulate an hypothesis on the existance of a transformational invariant that may be responsible for the evocation of rubbing or scratching. This suspected invariant, the impact density, was implemented in an existent continuous interaction model [5]. In a second experiment, a sound continuum from rubbing to scratching was generated with this modified continuous interaction model, and were further tested in a perceptual experiment. Results clearly support our hypothesis on the transformational invariant "impact density" which is partly responsible of the rubbing-scratching categorization, with an ambiguity for a particular band of density values. As our model permits a continuous control of the impacts density, a continuous high-level control, with a morphing between the two categories "to rub" and "to scratch" was included and calibrated in the previous synthesis model.

Although further studies are required to improve the synthesis model, this study shows that a simple transformational invariant such as the impact density can convey information on the nature of the continuous interaction, here the difference between rubbing and scratching an object.

The stationarity test seems to be a good descriptor to characterize different friction sounds, although further studies are required to take into account the global amplitude profile of the sounds due to pressure and velocity. Such a descriptor would lead to new possibilities in term of synthesis. Indeed, a composer or a sound designer may want to synthesize a sound which scratches or rubs like a recorded audio file he likes. By computing this descriptor, he could automatically get the synthesis parameters and then control and modify the synthesis process to obtain a sound with the same behaviour on another sound texture for example.

7   Acknowledgments

References

[1] Aramaki, M. and Kronland-Martinet, R. Analysis-synthesis of impact sounds by real-time dynamic filtering. Audio, Speech, and Language Processing, IEEE Transactions on, 14, 695–705 (2006)

[2] Aramaki, M. and Gondre, C. and Kronland-Martinet, R. and Voinier, T. and Ystad, S. Thinking the sounds: an intuitive control of an impact sound synthesizer. Proceedings of the 15th International Conference on Auditory Display, Copenhagen, Denmark May 18 - 22 (2009)

[3] M. Aramaki, M. Besson, R. Kronland-Martinet, and S. Ystad. Controlling the Perceived Material in an Impact Sound Synthesizer. IEEE Transaction On Speech and Audio Processing, 19(2) :301–314, February 2011.

[4] Gaver, W.W. How do we hear in the world? Explorations in ecological acoustics. Journal of Ecological psychology, 5, 4, 285–313 (1993)

[5] Van Den Doel, K. and Kry, P.G. and Pai, D.K. FoleyAutomatic: physically-based sound effects for interactive simulation and animation. Proceedings of the 28th annual conference on Computer graphics and interactive techniques, ACM, 537–544 (2001)

[6] McAdams, S.E. and Bigand, E.E. Thinking in sound: The cognitive psychology of human audition. Oxford Science Publication. Chapter 6 (1993)

[7] Gaver, W.W. What in the world do we hear?: An ecological approach to auditory event perception. Ecological psychology, 5, 1, 1–29 (1993)

[8] Michaels, C.F. and Carello, C. Direct perception. Prentice-Hall Englewood Cliffs, NJ (1981)

[9] Warren, W.H. and Verbrugge, R.R. 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 (1984)

[10] Xiao, J. and Borgnat, P. and Flandrin, P. Testing stationarity with time-frequency surrogates (2007)