Rubbing a Physics Based Synthesis Model: From Mouse Control to Frictional Haptic Feedback

Published
DOI:10.5281/zenodo.6797501

RUBBING A PHYSICS BASED SYNTHESIS MODEL: FROM MOUSE CONTROL TO FRICTIONAL HAPTIC FEEDBACK

Marius George Onofrei
Dept. of Design & Media Technology
Aalborg University
Copenhagen, Denmark
monofr11@student.aau.dk

Federico Fontana
Human Computer Interaction Lab
University of Udine
Udine, Italy
federico.fontana@uniud.it

Stefania Serafin
Multisensory Experience Lab
Aalborg University
Copenhagen, Denmark
sts@create.aau.dk

ABSTRACT

This paper investigates three kinds of interactions for a friction based virtual music instrument. The sound synthesis model consists of a bank of mass-spring-dampers individually excited via rubbing. A nonlinear static friction model capable of reproducing the characteristic stick-slip phenomenon observed in frictional interaction is employed, allowing for dynamic variation of the sliding friction.

The different controls developed allow for gradually increasing the interplay between performer and instrument. The key excitation parameters, e.g., the rubbing velocity and the rubbing normal force are controlled using three different interfaces: a standard mouse, a Sensel Morph, and a 3D Systems Touch X. The Sensel Morph is a touchpad with pressure sensitivity, allowing for a natural exertion of the normal force; the 3D Systems Touch X is a haptic device that renders both resistance to the applied normal force, as well as the stick-slip motion resulting from the friction interaction.

A preliminary user study aiming to compare the experience of performing with the different interfaces was carried out. The results indicate that the haptic feedback provides a more intuitive and enjoyable experience. However, extra features do not necessarily improve the user interaction, as the results suggest a preference for the mouse over the Sensel.

1. INTRODUCTION

Virtual musical instruments can be defined as computer programs that generate digital audio, typically for music. The interaction with such software applications is, in essence, an interaction with the computer itself. The use of the mouse and/or keyboard as control interfaces to such instruments is a straightforward choice. Other forms of controllers designed specifically for electronic music instruments (including audio applications) exist, with the most popular one being the MIDI keyboard, which makes use of a standardized serial protocol, which allows for communication between controllers and synthesizers.

Using such interfaces, however, can introduce a disconnection between the gestural actions and the sound produced by the virtual instrument. In fact, Cadoz proposes to preserve a consistent link between traditional instrumental gestures and their sound [1]. This disconnect, present in most virtual instruments, is particularly evident when considering physics based sound synthesis, as pointed out by O'Modhrain, who highlights the example of using a piano keyboard to control a physical model of a trumpet [2]. Nuances in the timber and dynamics of the sound being played are lost in a simple note on/note off system. Extra layers of controls, such as knobs controlling an ADSR envelope, or the mod wheel could be used to add these different nuances, but at the cost of extra complexity and with loss of an intuitive musical link between action and perception.

Controllers that more accurately mimic the excitation mechanisms of the instruments being simulated could provide a more satisfying performance experience. Virtual reality (VR) environments prove to be a fruitful ground for such explorations. As an example, Willemsen et al. made use of the haptic device Phantom Omni, to control a model of the tromba marina - a bowed string instrument [3]. Fontana et al. made use of the VR environment Unity 3D, together with the same haptic device to build a simulation of a plucked guitar [4].

A focus on intuitive musical gestures is also a critical part of the new interfaces for musical expression (NIME) community, where for example a very popular iPhone audio app, Smule's Ocarina, was first introduced [5]. Part of the appeal of this app, which simulates a wind instrument, is the use of breath-control as excitation mechanism, adding a layer of realism to the interaction.

One could argue that the realistic replication of all aspects contributing to the interaction with a real instrument is redundant. Replicating the excitation mechanism in the digital domain, however, allows to push the synthetic model to physical limits and even beyond. Consider as an example the work of Huynh who in [6] explores the effects on perception when exciting resonators in unnatural ways: "bowing plates" or "blowing strings". Even when aiming to model an instrument one-to-one, the virtual domain allows for unconventional interactions, as Onofrei et al. showed in the case of a friction drum model where the position of the friction excitation on the drum membrane could be changed in real-time, something not possible with a real instrument counterpart [7].

This paper explores the question of how to control a physics-based instrument model in an intuitive way, with the aim of providing a satisfying user experience. A simple audio application consisting of a bank of mass-spring-damper elements individually excited by means of rubbing was built. This simple and somewhat abstract physical interaction was chosen as to investigate three control strategies which were then developed with the aim of progressively affording musical ”rubbing” gestures, as well as providing the corresponding haptic feedback. A preliminary user study was then carried out evaluating the different setups. An important note is that the friction feedback used in our system results from the actual simulation of the friction interaction; the haptic friction force is the same as the force which excites the sound synthesis model. This is something that was not attempted in previous similar research, e.g. in both [3] and [8] the friction haptic feedback resulted from a different model than the one used for auditory feedback.

In Section 2 we describe the physical model behind the sound synthesis and the numerical method used to solve the resulting system of equations. Section 3 deals with the implementation of the model: it presents results of our off-line prototype with the aim of validating the nonlinear static friction model, then introduces the real-time application by describing its various features and controls. Finally the three user interfaces and their respective control strategies are presented. An evaluation of the resulting setups and its outcome is presented next in Section 4 and final concluding remarks are given in Section 5.

2. SOUND SYNTHESIS

The sound synthesis presented in this paper is based on the mass-spring-damper resonator model excited via frictional interaction. This resonator is the mechanical equivalent of a damped harmonic oscillator, which produces decaying sinusoids, and is of central importance to musical sound synthesis, both in abstract sound synthesis algorithms such as additive, subtractive or FM synthesis, as well as in physical modelling synthesis [9].

2.1 Physical Model

The model consists of a mass m [kg] connected to a rigid support through a parallel system of a linear spring, of spring constant k [N/m], and linear damper, of damping coefficient c [kg/s], as illustrated in Fig. 1. This mass is then excited via a rigid and weightless stick, which is pressed onto it with normal force N [N] and moved with a rubbing velocity v_r [m/s]. Its displacement relative to rest position at time t [s] is denoted $u(t)$ [m] and can be described by the following 2$^{\text{nd}}$ order ordinary differential equation (ODE):

$$m\partial_t^2 u + c\partial_t u + ku = -F_{fr}, \tag{1}$$

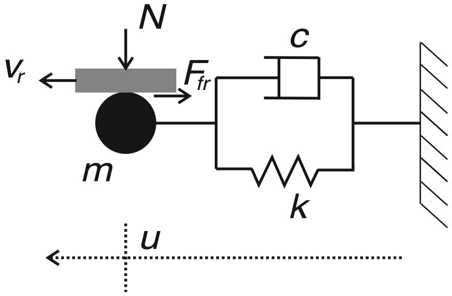


Figure 1: Mass-spring-damper system excited via friction.

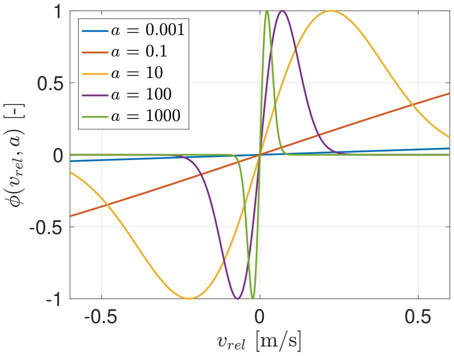


Figure 2: Shape of nonlinear friction characteristic given in Equation (3) for various values of a [s^2/m^2].

where F_{fr} [N] is the nonlinear frictional force resulting at the contact between the stick and the mass, modelled as:

$$F_{fr} = N\phi(v_{rel}, a). \tag{2}$$

This equation scales the applied normal force N with a nonlinear function $\phi(v_{rel}, a)$ which depends on the relative velocity between the stick and the mass, v_{rel} [m] and a friction parameter a [s^2/m^2] that gives the shape of the nonlinear function, and essentially controls how ”sticky” the interaction is. The nonlinear function is defined in Equation (3), while Equation (4) gives the relative velocity as the difference between the time derivative of the mass’s displacement and the rubbing velocity of the stick:

$$\phi(v_{rel}, a) = \sqrt{2a}v_{rel}e^{-av_{rel}^2+0.5}, \tag{3}$$

$$v_{rel} = \partial_t u - v_r. \tag{4}$$

Fig. 2 shows the nonlinear characteristic $\phi(v_{rel}, a)$ for various values of a. This friction model was introduced by Bilbao in [9] and it interpolates between an approximation of a viscous friction model (for small values of a) and a Stribeck model (for large values of a), where the discontinuity around zero relative velocity is smoothed out by the exponential function, making the friction characteristic continuous and differentiable, thus easier to compute numerically. Due to its stability and small number of parameters, which makes it more predictable and easy to work with, it has been fairly used for real-time audio simulations of bowed strings, e.g. in [10, 11].

For the case of the undamped system, i.e. the damping coefficient $c = 0$, it is well known that the natural resonance frequency of the system, f_0, is given by:

$$f_0 = \frac{1}{2\pi}\sqrt{k/m}. \tag{5}$$

Introducing damping will slightly skew this value by a factor determined by the critical damping factor, ζ, i.e., a ratio of the damping c over the amount of damping required to reach critical damping. The resonance frequency of the damped system, $f_{0,d}$ is given by:

$$\begin{aligned} f_{0,d} &= f_0\sqrt{1-\zeta^2}, \quad \text{with:} \\ \zeta &= \frac{c}{2m\sqrt{k/m}}. \end{aligned} \tag{6}$$

As can be deduced from the equations, the effect of damping on the resonance frequency becomes significant for values close to the critical damping, which is not the case for the current application.

With respect to the sound synthesis, the audio output is given by the displacement of the mass.

2.2 Solving the System

A numerical model can be figured out using the finite difference time domain (FDTD) method. Since the model is lumped, i.e. its parameters are not distributed in space, only the time domain needs to be discretized. The continuous time is sampled in discrete time steps indexed by $n \in \mathbb{N}$ as such: $t = nT$, where $T = 1/f_\text{S}$ [s] is the temporal step, f_S [Hz] being the sampling frequency.

With this discretization in place, the continuous function describing the displacement of the mass, $u(t)$, can be approximated by a grid function u^n, over the previously introduced temporal grid. Such discretizations are similarly carried out for the other functions: $F_{fr}(t) \approx F_{fr}^n$, $N(t) \approx N^n$, $v_r(t) \approx v_r^n$, $v_{rel}(t) \approx v_{rel}^n$ and $a(t) \approx a^n$. Moreover, the 1st and 2nd order time derivatives from Equation (1) can be approximated using the following operators:

$$\partial_t u \approx \delta_{t\cdot}u^n = \frac{1}{2T}\left(u^{n+1} - u^{n-1}\right), \tag{7a}$$

$$\partial_t^2 u \approx \delta_{tt}u^n = \frac{1}{T^2}\left(u^{n+1} - 2u^n + u^{n-1}\right). \tag{7b}$$

These approximations replace the continuous variables presented in the previous subsection, so that Equation (1) is discretized as:

$$m\delta_{tt}u^n + c\delta_{t\cdot}u^n + ku^n = -F_{fr}^n = -N^n\phi\left(v_{rel}^n, a^n\right). \tag{8}$$

From this, an update equation can be calculated for u^{n+1}:

$$u^{n+1} = u^n\frac{4m - 2kT^2}{2m + cT} - u^{n-1}\frac{2m - cT}{2m + cT} - F_{fr}^n\frac{2T^2}{2m + cT} \tag{9}$$

However, the resulting friction force at the current time step, F_{fr}^n, depends on the displacement of the mass at the following step, u^{n+1}. This dependence results from the insertion of Equation (7a) into Equation (4) and subsequently Equation (2). This issue can be addressed by making use of the following identity:

$$\delta_{tt}u^n = \frac{2}{T}\left(\delta_{t\cdot}u^n - \delta_{t-}u^n\right), \tag{10}$$

with $\delta_{t-}u^n = \frac{1}{T}\left(u^n - u^{n-1}\right)$ being a backward time-difference approximation to $\partial_t u$.

Substituting $\delta_{tt}u^n$ as such in Equation (8) and by making use of the definition of the relative velocity in the discrete domain:

$$v_{rel}^n = \delta_{t\cdot}u^n - v_r^n, \tag{11}$$

one can figure out the following equation:

$$\begin{aligned} f\left(v_{rel}^n\right) &= \frac{N^n}{m}\phi\left(v_{rel}^n, a^n\right) + v_{rel}^n\left(\frac{c}{m} + \frac{2}{T}\right) + b^n = 0, \\ b^n &= v_r^n\left(\frac{c}{m} + \frac{2}{T}\right) + u^n\left(\frac{k}{m} - \frac{2}{T^2}\right) + u^{n-1}\frac{2}{T^2}. \end{aligned} \tag{12}$$

This can be solved for v_{rel}^n using e.g. a Newton-Raphson iterative scheme, limited to a maximum of 99 iterations in order to avoid audio drop-outs. This value is then used to calculate F_{fr}^n, which allows for the solution to be propagated in time by computing u^{n+1} from Equation (9).

2.2.1 Stability

The stability of the finite difference scheme can be studied by means of an energy analysis of the discrete system, based on the work of [9] and [12]. Equation (8) is multiplied with $\delta_{t\cdot}u^n$, resulting in:

$$\begin{aligned} &m(\delta_{t\cdot}u^n)(\delta_{tt}u^n) + c(\delta_{t\cdot}u^n)^2 + \\ &k(\delta_{t\cdot}u^n)u^n = -(\delta_{t\cdot}u^n)N^n\phi\left(v_{rel}^n, a^n\right), \end{aligned} \tag{13}$$

Introducing $\delta_{t+}u^n = \frac{1}{T}\left(u^{n+1} - u^n\right)$ as the forward time-difference approximation to $\partial_t u$, the following identities hold:

$$(\delta_{t\cdot}u^n)(\delta_{tt}u^n) = \delta_{t+}\left(\frac{1}{2}(\delta_{t-}u^n)^2\right), \tag{14a}$$

$$(\delta_{t\cdot}u^n)u^n = \delta_{t+}\left(\frac{1}{2}(u^n u^n - 1)\right). \tag{14b}$$

Using these together with Equation (11), Equation (13) can be rewritten as:

$$\begin{aligned} &\delta_{t+}\left(\frac{m}{2}(\delta_{t-}u^n)^2\right) + \delta_{t+}\left(\frac{k}{2}(u^n u^n - 1)\right) = \\ &- c(\delta_{t\cdot}u^n)^2 - N^n v_{rel}^n\phi(v_{rel}^n) - N^n v_r^n\phi(v_{rel}^n). \end{aligned} \tag{15}$$

These terms have each an energetic interpretation. The first term in the left side of the equation is the discrete rate of change (δ_{t+}) of the kinetic energy, $\mathfrak{t}$, and the second is the rate of change in potential energy, $\mathfrak{v}$. Their sum gives the rate of change of the total energy, $\mathfrak{h}$.

$$\delta_{t+}\mathfrak{h}^n = \delta_{t+}\mathfrak{t}^n + \delta_{t+}\mathfrak{v}^n, \tag{16a}$$

$$\mathfrak{t}^n = \frac{m}{2}(\delta_{t-}u^n)^2, \tag{16b}$$

$$\mathfrak{v}^n = \frac{k}{2}(u^n u^{n-1}). \tag{16c}$$

Looking at the right side of Equation (15), the first term can be viewed as a loss in the system due to damping, $\mathfrak{q}$, while the next two terms represent the power dissipated and supplied by the excitation via the stick, $\mathfrak{p}_d$ and $\mathfrak{p}_s$:

$$\mathfrak{q}^n = -c(\delta_{t\cdot}u^n)^2, \tag{17a}$$

$$\mathfrak{p}_d^n = -N^n v_{rel}^n \phi(v_{rel}^n), \tag{17b}$$

$$\mathfrak{p}_s^n = -N^n v_r^n \phi(v_{rel}^n). \tag{17c}$$

It is evident that $\mathfrak{q}^n \leq 0$, as the damping coefficient $c \geq 0$. Similarly, $\mathfrak{p}_d^n \leq 0$ since the normal rubbing force N^n cannot be negative and, following the definition of the friction characteristic $\phi(v_{rel}^n, a^n)$ given in Equation (3), sign(v_{rel}^n) = sign($\phi(v_{rel}^n, a^n)$). This leaves the supplied power, $\mathfrak{p}_s^n$, which has indeterminate sign. However, an upper bound can be found for it, as $|\phi(v_{rel}^n, a^n)| \leq 1$. It follows that $\mathfrak{p}_s^n \leq |N^n v_r^n|$, meaning that the rate of change of total energy, $\mathfrak{h}^n$, is bounded by the following inequality:

$$\delta_{t+}\mathfrak{h}^n \leq |N^n v_r^n|\,, \tag{18}$$

with the external excitation inputs N^n and v_r^n being finite and thus, assuring that the energy in the system will not explode.

Additionally, for guaranteeing stability of the numerical scheme, it must be ensured that the total numerical energy, $\mathfrak{h}^n$, cannot be negative. This is done by starting from Equation (16) and expanding the finite difference operators, then rearranging the equation in a quadratic form:

$$\left(\frac{2T^2}{m}\right)\mathfrak{h}^n = (u^n)^2 + (u^{n-1})^2 + 2\left(\frac{kT^2}{2m} - 1\right)u^n u^{n-1}. \tag{19}$$

Based on this formulation the following condition must hold in order for $\mathfrak{h}^n$ to be positive definite:

$$\left|\frac{kT^2}{2m} - 1\right| < 1, \tag{20}$$

from which the final stability condition on the temporal step T results:

$$T < 2\sqrt{\frac{m}{k}}. \tag{21}$$

3. IMPLEMENTATION

The aim was to implement a real-time polyphonic audio application, consisting of a bank of mass-spring-dampers tuned to various musical frequencies. Additionally it was desired to have dynamic control of the damping as well as the friction parameter a, which controls how ”sticky” the interaction is.

3.1 Prototype Model Results

A prototype implementation of the sound synthesis model described in the previous section was first carried out in Matlab, whose offline setting allowed for a more in-depth analysis of the results. A particular focus was on whether the friction model produces results in line with the analytical results. Fig. 3 shows the displacement of a mass during a rubbing simulation of 0.1 seconds, meant to highlight the stick-slip behavior of the frictional interaction. The relative velocity between the stick and the mass periodically hovers around zero, meaning the two elements are stuck to each other until an abrupt slip, during which this velocity increases drastically. This results in a sawtooth waveform for the displacement of the mass and in similar sudden jumps for the resulting frictional force.

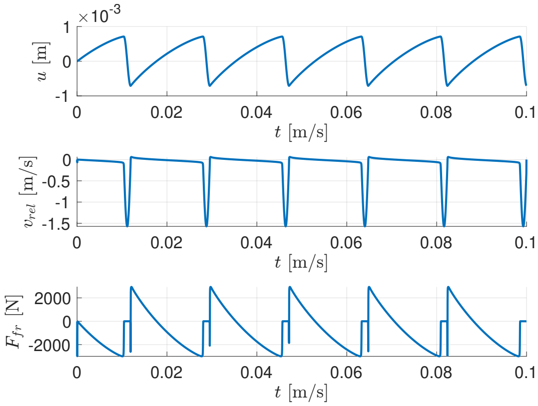


Figure 3: Results of a rubbing simulation with $m = 1$ [kg], $k = 4.2992\text{E}6$ [N/m], $c = 0$ [kg/s], $a = 100$ [s^2/m^2], $N = 3000$ [N], $v_r = 0.1$ [m/s], $f_\text{S} = 44100$ [Hz].

3.2 Real-Time Application

The real-time application was written in C++ using the JUCE framework. A demo video is available at [13] and Fig. 4 shows a snapshot of the application during use. Eleven mass elements are placed in the center of the application window, appearing as circles filled with different red tones. The darker the tone the lower the natural frequency. They are tuned starting from a resonance frequency of 55 Hz up to 330 Hz with a constant interval of 27.5 Hz. The simulation runs at a sampling frequency $f_\text{S} = 44100$ [Hz] and a corresponding temporal step $T = 0.0227$ [ms], ensuring numerical stability based on the condition given in Equation (21). The oscillation of these masses, i.e. the displacement obtained from the numerical simulation described in Section 2.2 is used as the audio output signal. This same displacement is updated in the application window at a rate of 24 frames per second, thus providing visual feedback to the user. Perhaps a better alternative for the sound output is the velocity of the masses, which is more proportional to the radiated sound of a physical instrument and also has more high-frequency detail. Two applications with the different sound output choice are available at [14]. These can be controlled using the mouse or a laptop’s touch pad, as described in more detail in Section 3.3.

The model of the stick used to excite the masses is portrayed as a long rectangle with varying opacity. When the stick is not in contact with the masses, it is displayed as grey; when in contact, its color changes to orange. Opacity increases proportionally with the amount of normal force with which the stick is pressed onto the masses, N. How this normal force is controlled is discussed in detail in the

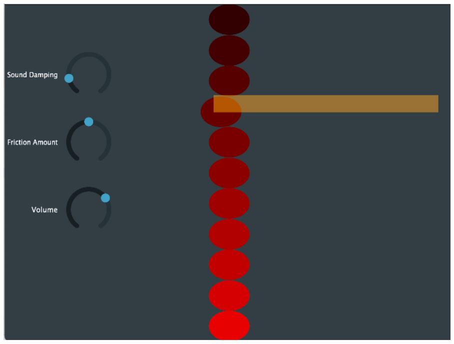


Figure 4: Snapshot of the audio application window. The stick (yellow rectangle) is exciting one of the masses (red circles), which is displaced from its rest position.

Figure 5: 3D Systems Touch X haptic device.

next subsection. When the stick is in contact with one mass, the velocity of its longitudinal movement is mapped to the rubbing velocity v_r. These are the two excitation parameters, as was illustrated in the sketch of the physical model given in Fig. 1. Of course, the stick can only rub one mass at a time, actually the one over which it is placed.

When the stick is lifted, the masses will continue to oscillate with an exponential decay that can be varied via the software knob labeled ”Sound Damping”, ranging from zero damping, i.e. constant sustained sound, to an amount large enough to remove all the release of the sound.

The second knob in the audio application, named ”Friction Amount”, controls the friction parameter a whose effect on the model was discussed in Section 2.1. When it is all the way to the left, the friction interaction will be completely smooth and the stick will have no effect on the masses and slide over them. When it is turned all the way to the right, the interaction becomes ideally adhesive and the masses will not slide, conversely they will be dragged by the stick. In this way the stick can be used to pluck the masses, thus creating a sound with a more sharp attack.

The linear range of values obtained from the knobs was mapped to the physical model parameters using a heuristically determined exponential mapping that was found to give a reasonable perception of their effect. Another such heuristic approach was applied by scaling the displacement of each mass before the output to the audio application. This was due to the fact that mass-spring-dampers tuned to lower frequencies produce larger amplitudes than oscillators resonating at higher frequencies, holding an identical excitation. This effect is due to the smaller stiffness of the former.

The last knob, ”Volume”, controls the global sound volume and is mapped over a logarithmic scale going from -30 dB to 3 dB.

In order to avoid audio clicks and with the intent of having a good dynamic range of the sound, a simple limiter was added, based on the design given in [15]. Finally as an extra safety measure, a DC filter, was also included.

3.3 Control Strategies

Three control strategies were developed for the app with an aim to mimic with increasing accuracy the gestural action of rubbing a mass using a stick.

3.3.1 Mouse

First off, the interaction with the app was based on the universal control device for virtual musical instruments, i.e., the mouse. Incidentally, the action of ”rubbing” is somewhat inherent in the dragging of the mouse, so the two gestures are in fact rather comparable.

The position of the mouse is mapped to the position of the center of the stick in the application window. When the mouse is clicked, the stick enters in contact with one mass, which can then be dragged longitudinally. The rubbing normal force, N, can be modified using the mouse wheel. Of course if one does not have a mouse at hand, the trackpad of a laptop can be used instead. In the case of a MacBook trackpad, the equivalent of tilting the mouse wheel is sliding two fingers up & down over it.

3.3.2 Sensel Morph

The second control strategy was implemented using a device named Sensel Morph, a tablet-sized pressure sensitive touchpad which is especially fast and sensitive.

Behind this control, was the idea to include the pressing action as part of the instrumental gesture, an action which can be naturally mapped to the rubbing normal force. With this in place, a more dynamic control of this excitation parameter is available for the user.

As opposed to the mouse control, where a binary mapping for putting the stick in contact with the masses is available, in the case of the Sensel this mapping cannot be explicitly set on and off. It was decided to map the position of the stick to the point where a finger first touches the device. The touch by a second finger signals that the stick must be put in contact with a mass, with normal force dependent on its pressure.

3.3.3 3D Systems Touch X

For the final control strategy, the purpose was not only to simulate the instrumental gesture even further but also include realistic haptic feedback based directly on the physical model, thus reinforcing the multisensory interaction with the audio application. This control could be achieved by using the 3D Systems Touch X haptic device, which is illustrated in Fig. 5 together with a reference coordinate system, used in the following descriptions. It is a 6-degree of freedom pen-shaped robotic arm, equipped with a series of motors which can provide 3-degree of freedom force feedback.

The position of the tip of the pen in the x-z plane was mapped to the location of the stick in the application window. The movement of the arm in the z direction was bounded by introducing two very stiff haptic x-y planes along the z axis, creating the feeling of hitting a wall. Another such haptic plane was modelled parallel to x-z at elevation y = 0. However this plane is elastic, providing force feedback proportional to the penetration along the y direction, corresponding to negative y coordinate values. When $y > 0$ conversely no feedback is felt, with this situation being mapped to the case when the stick is hovering above the masses. When the pen tip is at y = 0, contact with the masses occurs. The normal rubbing force, N, is then mapped to the force feedback felt in the y direction, coming from the elastic plane. Lastly, the movement velocity of the pen tip along the x direction is mapped to the rubbing velocity v_r.

In order to increase the realism of ”rubbing”, the friction force resulting from the physical model simulation is mapped to the force feedback in the x direction. In this manner, users can actually feel the stick-slip interaction between the stick and the masses.

4. EVALUATION

A user study was carried out aiming to evaluate the overall experience of using the virtual instrument with the different control strategies. The evaluation followed the guidelines given by Barbosa et al. in [16], with a focus on clearly stating the goal, methods and criteria used. As for the audio output signal, it was given by the displacement of the masses, since the velocity option was not developed at that time.

The goal was to compare the control strategies and investigate whether mimicking the instrumental gesture benefits the experience. Furthermore, there was a desire to collect feedback from users in order to improve the audio application and the interaction. This was achieved by means of both quantitative and qualitative methods, namely by: (1) a questionnaire composed of 7 statements, each related to a single attribute/criteria, which could be rated on a 7-point Likert scale, from ”Strongly Disagree” to ”Strongly Agree”. The statements and their associated attributes are given in Table 1. They are formulated in such a way that responses on the upper part of the scale (agreement) indicate positive opinions with respect to the attributes. (2) A qualitative interview was carried out after the subjects completed the questionnaire. Here, the focus was on investigating whether there was an understanding of what the different controls of the application did and how they affected the sound, i.e. turning the different knobs or rubbing the masses with different velocities and pressure. Lastly, their feedback regarding possible improvements to the system was noted.

Attribute	Statement
1. *intuitiveness*	The interaction with the objects was intuitive.
2. *playability*	I obtained sounds as by my intentions.
3. *enjoyment*	The interaction was fun.
4. *expressiveness*	I obtained enough types of sounds.
5. *difficulty*	It was easy to obtain the same sound twice.
6. *realism*	The interaction was realistic.
7. *precision*	The interaction was precise.

Table 1: The questionnaire items with corresponding anchors of the 7-point (1-7) Likert scale (Strongly Disagree - Strongly Agree).

4.1 Participants

A total of 14 participants took part in the user study, mostly comprising of students and staff at University of Udine in Italy. All but 2 persons had experience with playing music instruments, albeit not in any professional manner. Only one person had considerable experience using virtual instruments.

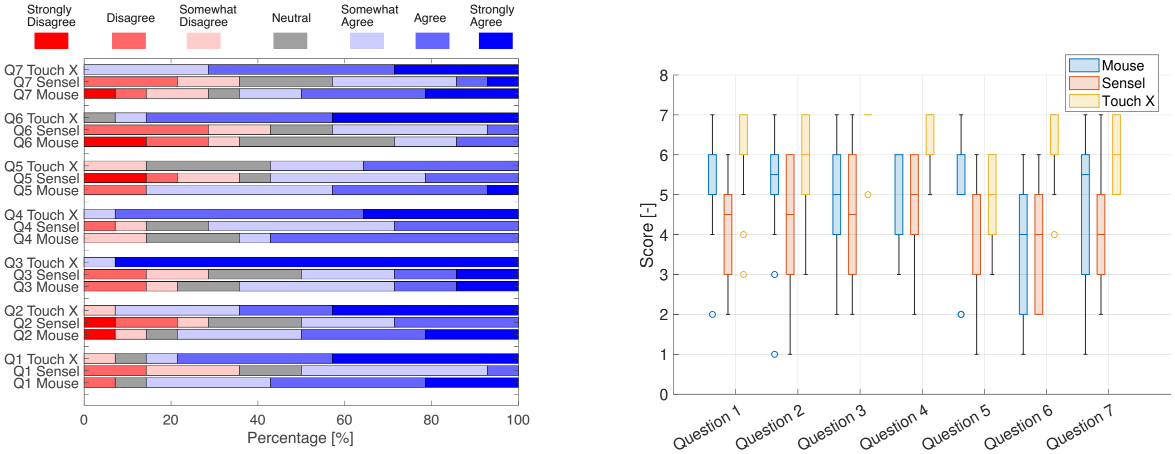
4.2 Procedure and Task

Each session started with a verbal introduction to the participants about the audio-haptic application and the physical model behind it. Showing the sketch of the system illustrated in Fig. 1 to the participants was found to be helpful in this regard.

Three instances of the app were opened, overall fitting on a large monitor. Each was respectively controlled by a device among those presented in Section 3.3. A brief tutorial on how to use the devices was carried out, during which the auditory feedback was turned off. This avoided users to experience how they could produce sounds of different quality instead of letting them explore the interface by themselves. The visualisation of the oscillation of the masses, however, was left on, to clearly illustrate the app’s response to control. Users were not instructed about the haptic feedback from the Touch X, as it was desired they discover it themselves.

After the briefing, the audio was turned on via a pair of Genelec 8020 active loudspeakers, by means of an RME Babyface PRO sound interface. Subjects were then free to use any of the three apps and their respective controllers in whichever order they prefer including the possibility to move across them more than once. Before they started, they were encouraged to experiment with the different knobs and try to get a feeling of what such controls did. Every participant’s activity was observed and notes were taken. Here, using loudspeakers as opposed to headphones was useful, as the experimenter could, among other things, identify different sonic preferences of the subjects or infer pressures applied to the stick via the auditory feedback.

Once they experimented with all three controllers, they were requested to fill out the questionnaire. For each ques-



(a) (b)

Figure 6: (a) Visualisation of the quantitative data illustrated as stacked bars. The three control devices: Mouse, Sensel and Touch X are grouped together for each of the 7 questions: Q1 to Q7. (b) Boxplot visualising the results related to the 7 questionnaire statements for the three control devices.

tion, they would provide a score for each controller, i.e., they went through the questions only once. This strategy highlighted the comparison between the devices, as the scores were more relative to each other. Once the questionnaire was filled, participants were interviewed about their experience.

Every session lasted between 10 and 30 minutes, depending on how long the users chose to use the apps, as there was no fixed time limit for that.

4.3 Results

The data from the questionnaire are presented in this section, followed by the qualitative findings from the interview.

4.3.1 Quantitative Data

An overview of the data can be provided by a stacked bar plot, as illustrated in Fig. 6a. This plot gives an idea of the relationships among such data and allows for an informal comparison between the devices. It can be seen that most responses for the control strategy using the Touch X haptic device fall in the positive Likert scale range (”Agree” to ”Strongly Agree”). Results are more mixed in the case of the two other controllers.

Furthermore, the Likert scale data can be treated as ordinal and a good indication on how the results are spread out can be achieved by box plot visualisation, as shown in Fig. 6b. This gives an assessment of the results in terms of central tendencies (median, mode), interquantile ranges and minimum/maximum ratings. Here again, one can see that the Touch X consistently scores better than the two other devices. When it comes to *enjoyment*, the users almost unanimously ”Strongly Agree” that the interaction was fun, with only a single score of ”Agree”. A similar positive distribution of scores is found with respect to *expressiveness*, *realism* and *precision*, with each of these attributes receiving answers solely in the positive range. The lowest scores for the Touch X were found for the *difficulty* statement, which is not surprising considering that none of the participants had used such a device before and a learning curve for its use is to be expected.

A somewhat unexpected result is that the mouse was generally preferred over the Sensel. It seems that the benefit for participants to rely on an intuitive dynamic pressure control is not enough, at least at first exposure, to overcome the familiarity they already have with the mouse and, at least to some extent, with trackbars aboard laptops. In this sense, the dragging gesture afforded by a mouse may have found an ideal match with the experimental task. Consequently, subjects found the mouse to be more intuitive, and easier to play with. One additional reason for this preference over the Sensel could be that using one finger to move the stick and two fingers to rub the masses could be difficult to learn, in spite of the availability nowadays of two-finger commands in several trackbars e.g. for scrolling windows.

As a matter of fact, the mouse scored best with respect to the *difficulty* attribute, suggesting that it was the easiest to use of the three devices. Again, this score testifies its ubiquity in computer setups.

4.3.2 Qualitative Data

This section summaries the insights obtained by observing participants during the task, and from the interviews they gave after filling out the questionnaires.

First off it is worth mentioning that all participants could easily explain and identify the effect of the ”Sound Damping” knob, with answers such as: ”[it] controls the release of the sound” or ”how long it [the mass] continues to vibrate”. When it comes to the ”Friction Amount” knob users were less certain, but did have some correct intuitions regarding its mechanism. For instance one said that it controls ”how strong you have to press to make a sound”, while another mentioned that it ”changes the perceptions of hardness or viscosity. . . like stirring a soup”. Indeed, with low friction values one needs to press harder to excite the masses, but with high values, corresponding to sticker interactions, this is not the case. Here, rubbing does indeed feel more viscous (one of the definitions of the word being "having a thick or sticky consistency"). Users mentioned that using the Touch X haptic device did help them better understand this parameter. After further explanation users could easily identify ranges for the knob where they could feel the stick-slip feedback. One reason why perhaps they were not so aware of this mechanism at first try, is that most of them focused on rubbing the masses in an effort to produce musical sounds, which is achieved at lower rubbing pressures, where the magnitude of the friction force feedback is consequently smaller. Another trend in the interviews was that there were plenty of references to the sounds of friction, with one user comparing the sound directly to that of bowed instruments like the violin, while another saying it "sounds like when rubbing a wet glass".

Some users spent a considerable amount of time with the applications and were noticeably experimenting with the devices and the sounds, while others were quite more reserved. In fact one of the persons with no experience with any musical instrument had very slow and unnatural movements, which resulted in unmusical sounds, seemingly reflecting a difficult interaction. This is a statement to resemblance of physics-based sound synthesis with real-world instruments, in the sense that one needs to learn the virtual instrument in order to fully enjoy the experience.

5. CONCLUSION

In this paper, a friction based physical modelling sound synthesis application was introduced, for which three different control strategies were developed. With these, the aim was to differently replicate the gestural action of rubbing in terms of control. We assumed that preserving the link between the excitation gesture of the physical system and the resulting sound would enhance the quality of the experience. This expectation was evaluated via a preliminary user study which showed a clear preference for controlling the app via the 3D Systems Touch X haptic device. The device not only mimicked the instrumental gesture, but also provided haptic feedback in terms of pressure response and the friction resulting from the physical model. This result is another indication of the importance of implementing intuitive interfaces for virtual instruments, and how haptic feedback plays a key role in this process.

6. REFERENCES

[1] C. Cadoz, "Instrumental gestures and musical composition," in *ICMC*, 1988.

[2] M. S. O'Modhrain, "Playing by feel: incorporating haptic feedback into computer-based musical instruments," Ph.D. dissertation, Stanford University, 2001.

[3] S. Willemsen, R. Paisa, and S. Serafin, "Resurrecting the tromba marina: A bowed virtual reality instrument using haptic feedback and accurate physical modelling," in *Proceedings of the 17th Sound and Music Computing Conference*, 2020.

[4] F. Fontana, R. Paisa, R. Ranon, and S. Serafin, "Multisensory plucked instrument modeling in unity3d: From keytar to accurate string prototyping," *Applied Sciences*, vol. 10, p. 1452, 2020.

[5] G. Wang, "Designing smule's ocarina : The iphone's magic flute," in *Proceedings of the 2009 International Conference on New Interfaces for Musical Expression*, 01 2009.

[6] E. Y. Huynh, "Bowed plates and blown strings: Odd combinations of excitation methods and resonance structures impact perception," Master's thesis, Department of Music Research, McGill University, Canada, 2019.

[7] M. G. Onofrei, S. Willemsen, and S. Serafin, "Real-time implementation of a friction drum inspired instrument using finite difference schemes," in *Proceedings of the 24th International Conference on Digital Audio Effects (DAFx20in21)*, 2021.

[8] S. Sinclair, G. Scavone, and M. Wanderley, "Audio-haptic interaction with the digital waveguide bowed string," in *Proceedings of the 2009 International Computer Music Conference, ICMC 2009*, 2012.

[9] S. Bilbao, *Numerical Sound Synthesis*. John Wiley and Sons, Ltd, 2009.

[10] S. Willemsen, S. Serafin, S. Bilbao, and M. Ducceschi, "Real-time implementation of a physical model of the tromba marina," in *Proceedings of the 17th Sound and Music Computing Conference*, 2020.

[11] P. Christensen, D. Overholt, and S. Serafin, "The daïs: A haptically enabled nime for controlling physical modeling sound synthesis algorithms," in *Proceedings of the 2020 International Conference on New Interfaces for Musical Expression*, 2020.

[12] S. Willemsen, "The emulated ensemble: Real-time simulation of musical instruments using finite-difference time-domain methods," Ph.D. dissertation, Aalborg University, 2021.

[13] M. G. Onofrei, "Audio-haptic frictional excitation of a mass-spring-damper physical model," https://www.youtube.com/watch?v=raLtqhdOYOM, Youtube.

[14] M.G.Onofrei, "Bowed mass-spring-damper apps," https://www.dropbox.com/sh/xclkp9ukz9qryn2/AAAk5XXyn0iU5E1Po9LU5x1Ma?dl=0, Dropbox.

[15] P. Dutilleux and U. Zölzer, "Chapter 5 - nonlinear processing," in *DAFX - Digital Audio Effects*. John Wiley & Sons, 2002.

[16] J. Barbosa, J. Malloch, M. Wanderlay, and S. Huot, "What does "evaluation" mean for the nime community?" in *Proceedings of the 2015 International Conference on New Interfaces for Musical Expression*, 2015.