

Higher Level Control Objects for Mass-Interaction Physical Modeling

Matthew Skarha
Music Technology
McGill University, Montreal, QC
matthew.skarha@mail.mcgill.ca

1 Introduction

Walking down the street with my headphones on, I observe that I am listening to both continuous-time signals (my footsteps, the cars passing by) and digitally-processed discrete-time signals (lofi house mix on SoundCloud through my headphones). However, my brain more-or-less perceives them as both continuous-time, in a way blurring the lines between the analog and digital worlds. At the center of this blurring is, of course, signal processing. Wielding this idea has numerous applications, particularly for digital sound synthesis for music creation. If we can understand how to model mechanical and acoustical systems that are governed by physical laws using discrete-time and discrete-space mathematical formalisms, we can simulate these systems with a control that might not be feasible in the analog realm. This is then the task of physical modeling sound synthesis.

2 Physical Modeling for Sound Synthesis

Within the realm of computer music, a number of formalisms for physical modeling have been proposed over the past forty years. A few are:

- finite difference schemes
- lumped models
- digital waveguides
- modal synthesis
- finite element modeling
- state-space techniques

Certainly each of these formalisms have their own pros and cons and will be suitable or unsuitable depending on the system. Often combinations of techniques can be used to achieve a more precise model. In 2003, Castagne and Cadoz proposed 10 general criteria for evaluating physical modeling techniques oriented to music creation [1]:

1. How efficient is the algorithm?
2. How faithful are the synthesized sounds?
3. How diverse are the categories of instruments that can be modelled?
4. Is the scheme exclusively dedicated to sound synthesis or more general?
5. How robust is sound plausability?
6. How modular is the technique?
7. How intuitive and effective is the associated mental model?
8. How deep is the modeling process enabled by the scheme?
9. Do generation algorithms exist?
10. Is there a friendly musician-oriented environment for using the scheme?

While it is not the task of this project to compare and contrast modeling schemes, I think these are good criteria to keep in mind when considering this project.

3 Mass-Interaction Physical Modeling

3.1 Background

Mass-Interaction Physical Modeling is one of the oldest techniques for digital sound synthesis via physical modeling. It was pioneered beginning in 1978 by the CORDIS-ANIMA system developed at ACROE in Grenoble, France [2]. The basic idea is to represent physical systems in the form of lumped networks composed of two main components: mass objects and interaction objects. Mass objects represent material points in a given space with some inertial behavior while interaction objects represent a specific type of coupling between mass objects (i.e. visco-elastic, collision, non-linear). By doing this, we avoid the need to explicitly define a mathematical model such as a partial differential equation system with boundary conditions (as is needed in finite difference methods).

Within the past few years, there has been somewhat of a renewed interest in mass-interaction physical modeling, driven by the release of a Max/MSP package `mi-gen` [3] (github) and a Faust [4] (github) library. These tools have certainly improved the accessibility of MI-modeling thanks to the inclusion of a model scriptor that translates a high level model into lower level code.

3.2 Basics

In order to model the physics of masses and interactions in discrete-time, we will apply a second order central difference scheme to Newton's second law. For a point mass, we have

$$f = ma = m \frac{d^2x}{dt^2}$$

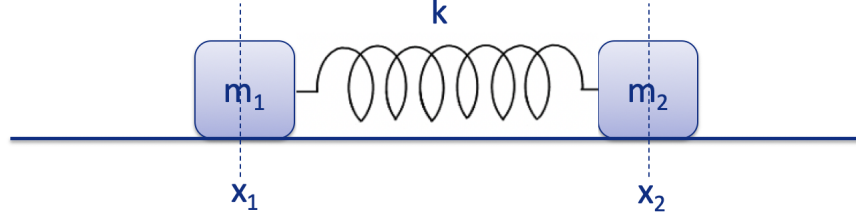


Figure 1: A visco-elastic spring connecting two masses, m_1 and m_2 at positions x_1 and x_2 , respectively.

where f is the force applied to the mass, m is its inertia, a is its acceleration, and x is its position.

Using a second order central difference scheme with sampling period ΔT , we have

$$f(t) = m \frac{x(t + \Delta T) - 2x(t) + x(t - \Delta T)}{\Delta T^2}.$$

To discretize, we simply convert the sample period into a sample and collect the continuous-time parameters into a discrete-time inertial constant $M = \frac{m}{\Delta T^2}$.

$$f[n] = M(x[n + 1] - 2x[n] + x[n - 1]).$$

Rearranging terms, we can achieve a difference equation that describes the discrete-time position of a mass as a result of the external forces being applied:

$$x[n + 1] = 2x[n] - x[n - 1] + \frac{f[n]}{M}.$$

A very common object in mass-interaction modeling is a visco-elastic spring connecting two masses. Consider a spring with stiffness parameter k and damping parameter z . Hooke's Law tells us that the force needed to extend or compress a spring by some distance x scales linearly with respect to that distance. In the case of a spring connecting masses as shown in Figure 1, we have

$$f_{s,1 \rightarrow 2} = -k(x_2 - x_1)$$

where x_1 and x_2 are the positions of the two masses m_1 and m_2 , respectively. In discrete-time, this is

$$f_{s,1 \rightarrow 2}[n] = -K(x_2[n] - x_1[n])$$

where $K = k$ is the discrete-time stiffness parameter.

If we approximate damping to the first-order, we have the following relationship describing the force from m_1 onto m_2 as a result of the damper:

$$f_{d,1 \rightarrow 2} = -z \frac{d(x_2 - x_1)}{dt}.$$

Using a first-order backward difference scheme to discretize the derivative, we have

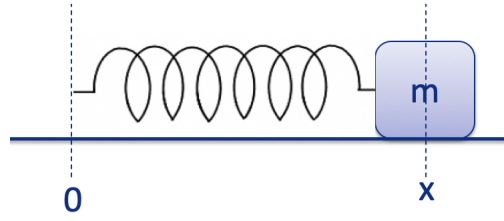


Figure 2: A visco-elastic spring connecting a mass and a fixed point.

$$f_{d,1 \rightarrow 2} = -z \frac{(x_2(t) - x_1(t)) - (x_2(t - \Delta T) - x_1(t - \Delta T))}{\Delta T},$$

where ΔT is the sampling period. In discrete-time, we have

$$f_{d,1 \rightarrow 2}[n] = -Z[(x_2[n] - x_2[n - 1]) - (x_1[n] - x_1[n - 1])].$$

Thus, we have computed the discrete-time external forces for a spring connecting two masses. More common, however, is the world-reknown linear harmonic damped oscillator (mass on a spring) where one end of the spring is now fixed to a point, as shown in Figure 2. Writing the sum of the forces on mass m as a result of the spring and the damper, we have

$$f_{tot,m}[n] = f_{s,m}[n] + f_{d,m}[n]$$

$$f_{tot,m}[n] = -Kx[n] - Zx[n] + Zx[n - 1].$$

We can now substitute this expression into the discrete-time Newton's second law from above and rearrange terms to achieve a difference equation that describes the motion of a harmonic oscillator.

$$x[n + 1] = \left(2 - \frac{K + Z}{M}\right)x[n] + \left(\frac{Z}{M} - 1\right)x[n - 1]$$

Notice that the position of the mass at time $n + 1$ simply depends on a linear combination of the position of the mass at time n and $n - 1$, where the coefficients are simply terms in the stiffness parameter K , the damping parameter Z , and the inertial parameter M . In practice, these parameters are normalized to 1. Although they are functions of the sample rate, they do provide a direct view of stability conditions for the system. In order to stay in an oscillatory regime, we must have $4M > K + 2Z$. A derivation of stability conditions for the linear harmonic oscillator can be found in the appendix of [3].

3.3 Building Topologies

Based on this formalism, modeling with MI involves building a geometrical model by positioning and connecting masses together through interaction objects, and by specifying the parameters and initial conditions for each one. Figure 3 shows a basic physical model of a triangle (the percussion

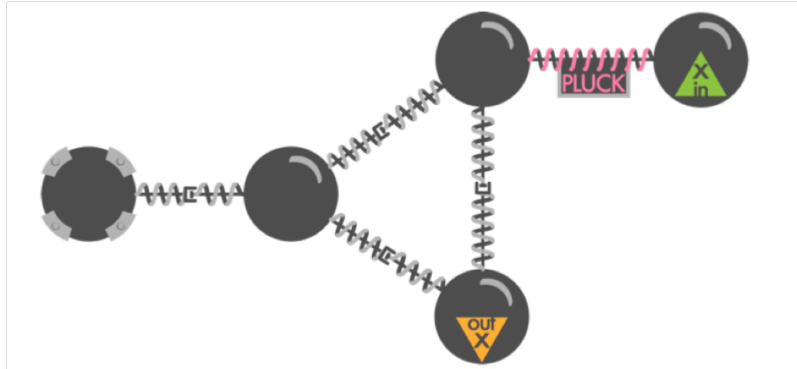


Figure 3: *Physical model of a triangle. Taken from [4].*

instrument). Three masses are joined by dampened springs with one of the masses fixed to a point. The system is then excited by a pluck at the input module and the output is taken from one of the masses. Because sound is essentially just air being pushed, we can listen to the motion of a given mass as the output of our system.

One main advantage arises when comparing MI modeling with other physical modeling paradigms: It introduces the ability to build and excite virtual mechanical constructions that are not bounded by realism. We know from more traditional musical acoustics research that many nonlinear behaviors exist in acoustic musical instruments which contribute to their unique and interesting timbres. It is often the task of other physical modeling techniques to model these nonlinearities for a given system with the goal of achieving a more realistic sound synthesis. With MI modeling, we are "atomically" building structures grounded in Newton's second law, meaning nonlinear behavior will arise naturally if a system is excited properly and maintains numerical stability. If we can model and excite the right structure, theoretically we can discover new interesting behaviors and sounds for use in an artistic process.

One main disadvantage, however, is the task at hand. It is difficult to develop an intuition of how to model structures to achieve a certain timbre or behavior, especially when there are three parameters to adjust for every object. The goal of this project, then, is to simplify this process by introducing higher level (timbral) control objects to build MI geometries.

4 Towards Higher Level Control of MI Systems

This project attempts to introduce higher level control objects for mass-interaction physical models. This was done by first assessing what a set optimal control parameters would look like in this context. Next, a series of experiments were conducted to build models for the relationships between existing parameters and optimal parameters. And finally, a Max patch was developed with these controls as input to determine if it is a more intuitive approach to MI modeling.

4.1 Optimal Control Parameters

If an artist wants to use mass-interaction modeling as a sound synthesis paradigm in their creative process, what parameters would they be interested in controlling? Considering the victories of previous synthesizers, I contend pitch and decay rate would be suitable as a start.

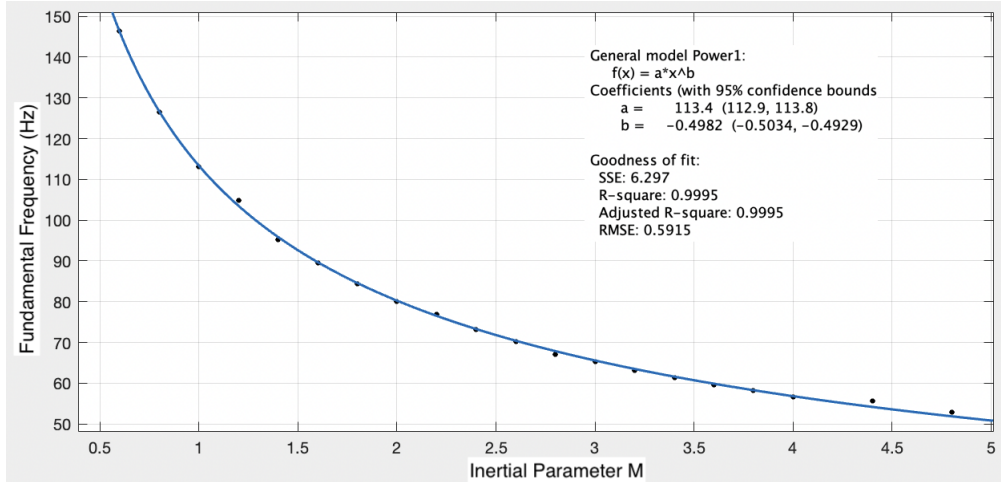


Figure 4: *Inertial parameter M vs. fundamental frequency. A least-squares fit indicates a first-order power law relationship.*

As for timbral control, we can use spectral descriptors to build a control object. Spectral descriptors are a set of perceptually-relevant algorithms that can be computed on any audio signal to learn information about that signal. Some examples are spectral centroid, decrease, kurtosis, and spread. Sound events are analyzed in terms of various input representations including the short-term Fourier transform, harmonic sinusoidal components, and an auditory model based on the equivalent rectangular bandwidth concept. A number of audio descriptors are then derived from each of these representations to capture temporal, spectral, spectrotemporal, and energetic properties of the sound events [5].

There are often correlations between spectral descriptors and perceptual attributes. For example, spectral centroid is known to be correlated with brightness. I contend that we can exploit this idea to obtain higher level control over mass-interaction physical models.

4.2 Experiments

To determine the relationship between the inertial parameter and the pitch, an experiment was carried out by varying the inertial parameter M and estimating the fundamental frequency, as shown in Figure 4. A 50 mass string was used as the test structure. The data was fit to a first-order power law of the form $f(x) = ax^b$ and a 0.9995 R-square value was achieved. Thus, to control the inertial parameter with pitch, an inverse function $f^{-1}(x)$ is taken and solved for $f(x)$:

$$M = af_0^b$$

$$f_0 = \sqrt[b]{M/a},$$

where M is the discrete-time inertial parameter, f_0 is the fundamental frequency in Hz, $a = 113.4$, and $b = -0.4982$. It is important to note that while the first order power law relationship holds across different structures, the coefficients of the model do not and can therefore not be generalizable. This drastically increases the difficulty for pitch control via inertial parameter modulation and it would likely be easier to use known pitch shifting techniques.

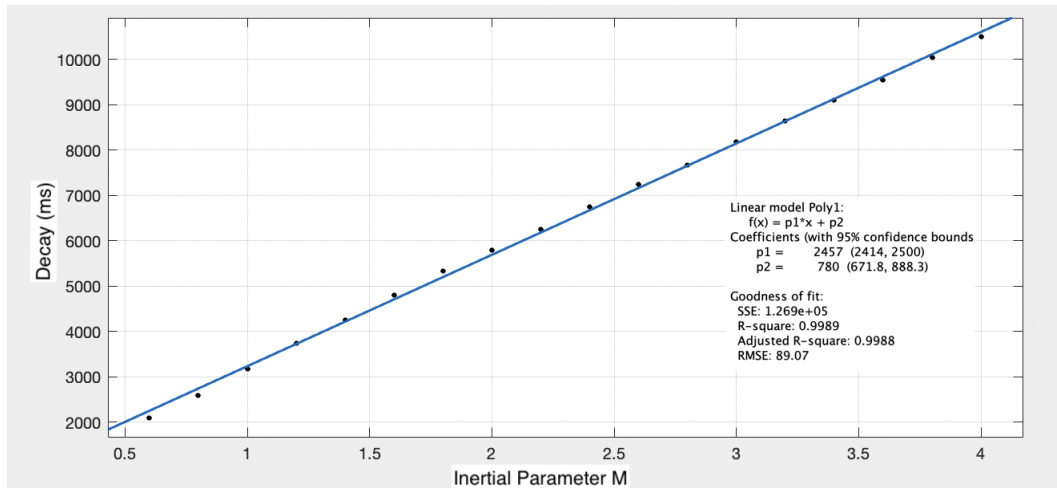


Figure 5: Inertial parameter M vs. decay time (ms). A least-squares fit indicates a linear relationship.

Next, an experiment was carried out to determine the relationship between the inertial parameter and the decay time, as shown in Figure 5. This indicated a linear relationship between the two variables and a control object was constructed accordingly. Again, while the fit coefficients are not generalizable for any given structure, the linear relationship holds.

To determine what a timbral control object would look like, a number of experiments were carried out using spectral descriptors to try to determine the relationship between MI geometry and timbre. First, several spectral descriptors were measured against the number of masses in an ideal string to determine the timbral effect of masses. Some of the results are shown in Figure 6. Adding masses to a string therefore decreases centroid, increases spectral decrease, increases kurtosis, and decreases slope.

Then, six mid-level macro objects were constructed using the Mass Interaction Model Scripser (MIMS) provided in the mi-gen Max/MSP toolbox: string (ideal string fixed at both ends), stiff string (a string accounting for 2nd order stiffness, fixed at both ends), a chain (string but not fixed at end), mesh (2d string with open boundary conditions), closedMesh (entirely fixed boundary conditions), and cornerMesh (fixed points at each of the four corners). Then, the parameters were tuned so that all objects had the same fundamental frequency and damping (it doesn't make sense to compare timbres of objects of different pitches/decays). The objects were measured against spectral descriptors (centroid, decrease, kurtosis, spread, skewness, rolloff) to try to determine their effect on the timbre. Some of the results are shown in Figure 7. The main takeaway from these experiments was that not fixing points has the effect of boosting low frequencies (as in the case of the chain).

4.3 Max Patch

A Max patch was constructed with a novel "timbral control" object based on the results of the experiments in the previous section. The object is based off of Max's *patrstorage* object while allows for linear interpolation between presets. Six timbral "modes" are contained as presets in this object, between which the user can interpolate. The modes were determined by varying the

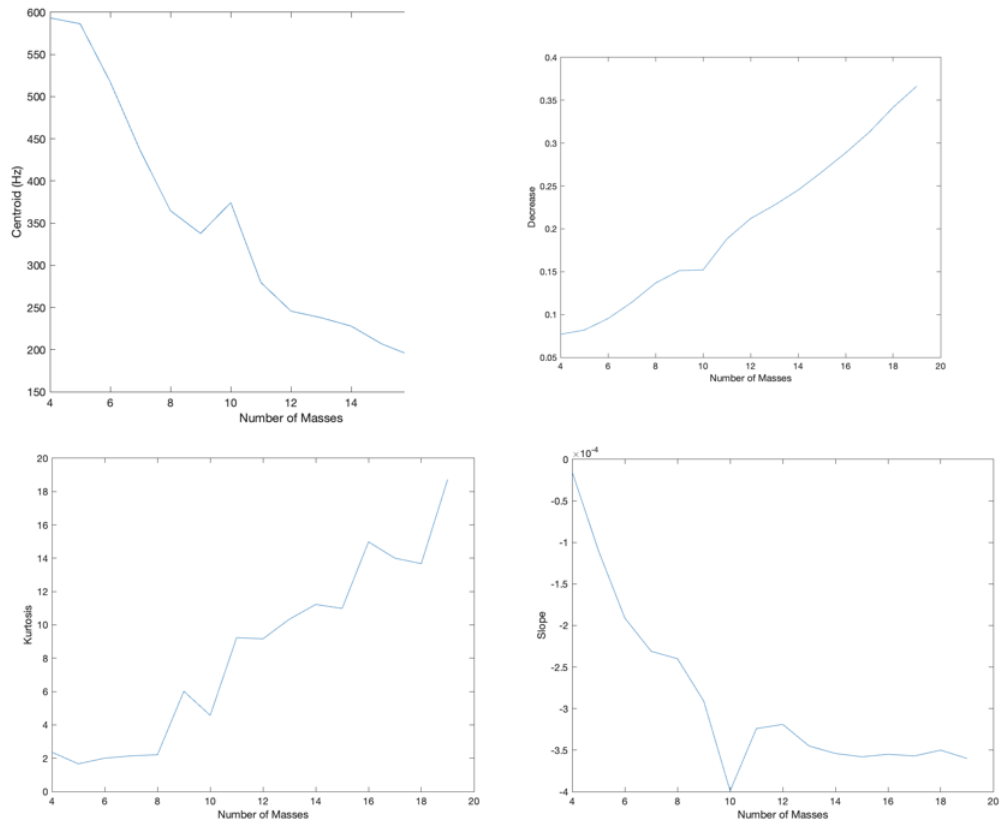


Figure 6: Some results of experiment testing timbral effects of number of masses on an ideal string.

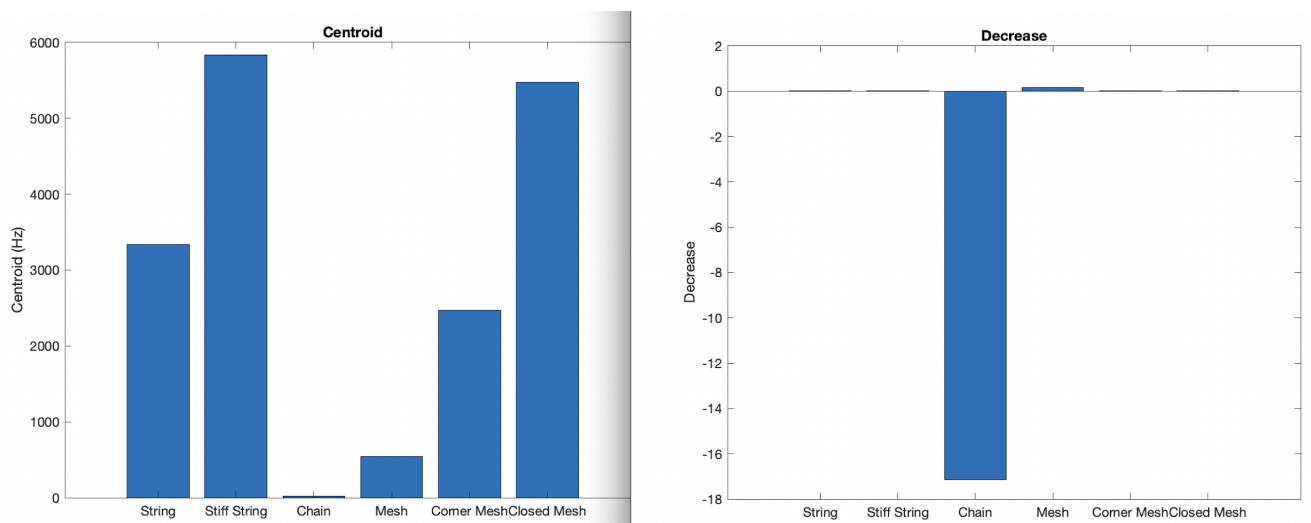


Figure 7: Some results of experiment testing timbral effects of various mid-level structures.

gain of each of the six macro objects described earlier while holding the other five constant and determining their effect on six spectral descriptors (centroid, kurtosis, decrease, skewness, rolloff). If a spectral descriptor quantity reached a maximum as the gains were varied, it was determined to be a timbral "mode" and was added to the object.

5 Conclusions

Mass-Interaction Physical Modeling is a physical modeling sound synthesis technique performed by representing physical systems in the form of lumped networks composed of two main components: mass objects and interaction objects. Modeling with MI involves building a geometrical model by positioning and connecting masses together through interaction objects, and by specifying the parameters and initial conditions for each one. However, it is difficult to achieve intuition over how to build physical models since the objects are so low-level. This project sought to create higher level timbral control objects by experimenting with spectral descriptors across various topologies.

While I have made some progress towards higher level control of MI models, there is still far more to be done. The task ended up being far more difficult than I expected and would require a much deeper understanding in order to implement such an object in a synthesizer. Nonetheless, I believe MI modeling has endless potential as a physical modeling technique and I look forward to continue studying it in the future.

6 References

- [2] N. Castagne, C. Cadoz, "10 Criteria for evaluating physical modelling schemes for music creation", Proceeding of the Digital Audio Effects Conference DAFX03, London, UK, 2003
- [1] Cadoz, C., Luciani, A., Florens, J. L. (1993). *CORDIS-ANIMA: A Modeling and Simulation System for Sound and Image Synthesis: The General Formalism*. *Computer Music Journal*, 17(1), 19-29. doi:10.2307/3680567
- [2] Castagné, N., Cadoz, C. (2002). *GENESIS : a Friendly Musician-Oriented Environment for Mass-Interaction Physical Modeling*. *International Computer Music Conference Proceedings*.
- [3] Leonard, J., Villeneuve, J. (2019, 2019-05-28). *mi-gen : An Efficient and Accessible Mass-Interaction Sound Synthesis Toolbox*. Paper presented at the SMC 2019 - 16th Sound Music Computing Conference, Malaga, Spain.
- [4] Leonard, J., Villeneuve, J., Michon, R., Orlarey, Y. (2019). *Formalizing Mass-Interaction Physical Modeling in Faust*.
- [5] Peeters G, Giordano BL, Susini P, Misdariis N, McAdams S. The Timbre Toolbox: extracting audio descriptors from musical signals. *J Acoust Soc Am*. 2011 Nov;130(5):2902-16. doi: 10.1121/1.3642604. PMID: 22087919.
- [6] Kontogeorgakopoulos, A., Claude, C. (2007). *Cordis Anima Physical Modeling and Simulation System Analysis*.

[7] Leonard, J., Cadoz, C. (2015). Physical Modelling Concepts for a Collection of Multisensory Virtual Musical Instruments. Paper presented at the Proceedings of the international conference on New Interfaces for Musical Expression, Baton Rouge, Louisiana, USA.