SOUND CONTROLLED MUSICAL INSTRUMENTS BASED ON PHYSICAL MODELS

Andrew Johnston | Benjamin Marks | Linda Candy
Creativity & Cognition Studios | ELISION Ensemble | Creativity & Cognition Studios
Faculty of IT, University of Technology Sydney | 36 Longueval St., Moorooka, Queensland, Australia | Faculty of IT, University of Technology Sydney
aj@it.uts.edu.au | ben@optusnet.com.au | linda@lindacandy.com

ABSTRACT
This paper describes three simple virtual musical instruments that use physical models to map between live sound and computer generated audio and video. The intention is that this approach will provide musicians with an intuitively understandable environment that facilitates musical expression and exploration. Musicians live sound. exerts ‘forces’ on simple mass-spring physical models which move around in response and produce sound. Preliminary findings from a study of musicians’ experiences using the software indicate that musicians find the software easy to understand and interact with and are drawn to software with more complex interaction - even though this complexity can reduce the feeling of direct control.

1. INTRODUCTION
In this paper we describe some of our recent work which makes use of virtual physical models to mediate between live, acoustic sounds and computer-generated sounds. We are interested in exploring the use of physical models in software which facilitates musical exploration by musicians playing acoustic instruments. That is, we are interested in using the physical models as ‘interfaces for musical expression’ - as opposed to using them to synthesise sound directly.

Our approach to investigating the use of physical models in this way has been ‘art-driven’. We have collaboratively created a two-part composed work for solo trombone and software based on virtual physical models. This collaboration is between two people, a composer/trombonist (the ‘composer’) and a technologist/trombonist (the ‘technologist’). The composer has a masters degree in composition and performance, specialising on the trombone. He also completed a masters degree in computing and currently works as a lecturer in a faculty of Information Technology.

There are, of course, a large number of applications designed to facilitate musical expression, so in order to briefly clarify the nature of our work we state here that:

- The software responds in real-time.
- The musician can exert ‘force’ on a virtual physical model by playing into a microphone.
- A representation of the physical model is shown on-screen during performance and is visible to both performer and audience.
- As the physical model moves it causes sound to be produced. That is, the audience hears computer-generated sounds (mediated by the physical model) as well as acoustic sounds.
- The musician only interacts with the software via the microphone. There are no additional buttons, mice or other controllers.

Because this software responds only to the sound produced by the musician, and not to input from any other device, we have come to think of it as an extension of the instrument. It may also be seen as a musical instrument in its own right which, unlike traditional musical instruments, is controlled by sound rather than more direct physical interaction. Throughout this paper, we will refer to the software as ‘virtual instruments’.

2. USE OF PHYSICAL MODELS
The use of physical models in computer music is a very active research area. A particular focus is on their use in audio synthesis to both produce realistic sounds and explore new sounds that have a basis in familiar sounds. In our work however, we are interested in using physical models to provide a link between performer and computer-generated audio-visuals. The work of Momeni and Henry [6] [5] has been a key influence here.

We began to use mass-spring models very early on in the project. The term ‘models’ in this sense does not refer to actual physical objects that exist in the real-world,
but rather to the building of ‘virtual models’ made up of masses and links that move around the computer display. In our work, this model can be thought of as a kind of interactive virtual sculpture that is controlled by sound. The sculpture is built by positioning various objects in virtual space and linking them together with virtual ‘springs’ of various length, rigidity, etc. Because this sculpture obeys the laws of Newtonian physics, it responds in ways that appear natural when forces are applied to it. In our case, the forces are mapped to characteristics of the music that is played. So, for example, if the loudness of the input sound is mapped to the quantity of force exerted on the sculpture then playing a loud note will cause a large amount of force to be applied to the model and, depending on its structure, it may bounce around the screen, change shape and so on. In our work, these movements also cause the computer to output sounds.

To put it simply, the musician’s live sound exerts force on parts of the physical model/virtual sculpture which causes it to move in physically plausible ways. Figure 1 shows a high-level view of how this works. Note that while it does not necessarily have to be the case, in our work the visual output is a direct representation of the physical model itself. The intention is that the musician has a feeling of direct control over the ‘virtual sculpture’ with their playing and that the audience can readily perceive this.

We took the decision to use physical models in this way for a number of reasons. The most important in terms of providing an engaging experience for musicians and audiences appear to be:

- The musician feels in control of the visual and sonic behaviour of the computer;
- Because the models respond in a way that is consistent with our experience of everyday physical objects, their behaviour is intuitively understandable;
- Because the movements of the model which produce sound are based on realistic physical motions, we feel the resulting sounds have an ‘organic’ quality. (We note however that the use of physical models does not guarantee this however, as some musicians who used our software felt that not all of the sounds produced by the virtual instruments had this quality.)

### 3. THREE VIRTUAL INSTRUMENTS

In this section we will describe each of the three virtual instruments we have developed. The first two were developed for the two part performance work *Partial Reflections* for solo trombone and virtual instruments, a work we have presented in a concert setting with the soloist playing onstage beside a large screen showing the software’s visual output. Each part is self-contained and is based upon a single physical model. We briefly outline each of the physical models and the mappings between the live sounds, the forces exerted on the virtual model and resulting computer-generated sounds. In addition we describe a third virtual instrument, based on the first two, in which the interaction has been simplified.

Technically, the software is implemented using the open source programming environment Pure Data (pd) (http://pure-data.info). The Pure Data libraries Graphical Environment for Multimedia (GEM) (http://gem.iem.at/) and Physical Modeling for Pure Data (pmpd) [5] are used for 3D graphics and physical modeling respectively.

#### 3.1. Partial Reflections Part I (PR1)

Following Perry Cook’s principle of “Instant music, subtlety later” [1] our initial exploratory steps make use of extremely simple physical models. In fact the physical model we used for PR1 is an almost unchanged example model provided with pmpd. Twelve masses are linked together to form a string, fixed at one end (figure 2).

The musician exerts force on the model by playing into a microphone connected to the computer. The pd object *fiddle* [7] is used to analyse the input signal and extract various musical parameters such as pitch, volume and strength and frequency of partials.

The mapping is simple. The pitch of detected notes selects the target mass and the volume determines the amount of force. Each mass is associated with a particular pitch-class. So, if the trombonist plays an A with a frequency of 440Hz then the mass second from the top will have force exerted on it. The amount of force will be proportional to the volume of the note. In response, the mass will move around and, because it is linked to other masses in the string, this will cause all the other masses to move in response.
The on-screen spheres which represents the masses glow brightly when they are receiving force as a result of the musician’s playing. When not receiving force their glow gradually fades.

They effect is that the musician can ‘push’ masses in the virtual model around by varying the pitch and volume of their playing.

In addition to responding to live sound, the virtual model also causes sounds to be produced using simple additive synthesis. As previously mentioned, each mass is associated with a particular pitch class. Of course, by definition the term ‘pitch-class’ refers to all notes with the same name and not to a note with a specific frequency. Obviously the musician may play each pitch-class in several octaves. When the software detects an A for example it makes a note of its frequency and associates that frequency with the A mass. It will then play a pure tone (sine wave) with that frequency at a volume proportional to the current velocity of that particular sphere. This means that while any sphere on the screen is moving at a speed greater than an adjustable threshold value, the system is generating sounds. Conceptually, one could say that the masses ‘store’ the frequencies of the notes played by the musician. Note that each mass will only store one frequency at a time, so if the musician plays another A with a different frequency (perhaps in another octave or simply a slightly flatter or sharper note) then the frequency of the new A will replace that of the old.

In order that the generated sounds are more interesting, the partials of the player’s sound are treated the same way. That is, they are ‘stored’ by the masses. We use fiddle’ to identify the two strongest partials in the sound and these are associated with the appropriate mass. eg. If the sounded A has partials with a pitch-class of E and G, then the E mass and the G mass will store the frequencies of those partials and play them back when they move. The effect is difficult to describe, but in a way is a kind of simple resynthesis of the musician’s live sound, mediated by the physical model.

3.2. Partial Reflections Part II (PR2)

The physical model at the core of this movement is again very simple (figure 4). The model is made up of twelve masses, each one associated with one of the twelve pitch-classes of the equal-tempered scale. Each of these masses is linked to a fixed central point. Initially, the spheres spin very rapidly around this point very close to the centre. When the musician plays, each note causes the associated sphere to be pushed in an anti-clockwise direction which makes the sphere accelerate and thus spin out further from the central point.

The software records the first 100ms of each note (the attack) and this recording is associated with the sphere of the appropriate pitch class. Each time the sphere completes a half turn around the central point, the software plays back the recorded sound linked to that sphere with one additional modification: the higher the orbit, the slower the playback. The effect of slowing the playback is to lower the pitch of the played-back note. So, if the sphere has a very high orbit (because it has had a lot of force exerted on it) then the note that plays back every half rotation will be pitched quite low. An example may help to clarify

![Figure 3. Visual display of PR1 during performance.](image)

![Figure 4. The mass-spring model for the Partial Reflections Part II (PR2). The masses orbit about a fixed central point.](image)
this behaviour. When the software starts the spheres are spinning rapidly around a central point at a very low altitude. If the performer plays a Bb several things happen:

1. The Bb sphere has force exerted on it in proportion to the volume of the Bb attack;
2. The first 100ms of the note are recorded and associated with the Bb sphere;
3. In response to the force, the Bb sphere is pushed out into a higher orbit;
4. Every half turn, the 100ms of recorded Bb is played back, but with pitch shifted down by an amount proportional to the distance of the sphere from the central point;
5. When the performer stops playing Bbs the Bb sphere gradually spins back to the central point and as it does so the audio playback gradually returns to its original pitch.

3.3. Charmed Circle (CC)

The virtual instruments for Partial Reflections were designed to be used in performances of music specifically composed for them. As such they had some characteristics which made them less generally usable as musical instruments. For example, because the physical model for Part I was a string fixed at one end, playing notes that were

associated with the masses at the non-fixed end caused a lot more movement in the model than playing notes associated with masses at the fixed end. This suited the music that was written for that instrument, but we were interested to see what would happen if we made the instrument a little more generic. In addition, a number of musicians who had seen and played with Partial Reflections had suggested this, and indicated that a simpler, more generic instrument might also be useful in teaching.

We therefore created a simplified version of the virtual instrument from PR1, Charmed Circle (CC), altering the interaction in several ways. Firstly, the three musical features extracted from the live audio were reduced to two: pitch and volume. Secondly, to ensure that all musical pitches had equal influence on the model, the physical structure was redesigned to arrange the masses in a circle (figure 7). Finally, the links between the masses were removed so that forces exerted on one mass would have no impact on the others.

The underlying physical model is slightly more complicated than the screenshot might indicate. Each of the spheres which are visible on the screen actually sit at the midpoint of a long string, fixed at both ends. The 'far end' of all the strings is a single fixed point in the distance, behind the spheres. In the middle distance are the visible spheres, floating in the middle of the string. In the 'foreground' (ie. closest to the viewer) are 12 separate fixed points arranged in a circle that anchor the other end of the strings. Neither the fixed points nor the string are visible, but their existence may be discerned by observing the behaviour of the visible spheres.

When the musician plays a note, the software determines which pitch is being sounded. Once again, each sphere is linked to a particular pitch-class so that whenever a particular note is played, the sphere associated with that pitch-class has force exerted on it. The force is always exerted in an outward direction, so the sphere is effectively pushed outwards by the musician’s note. The force is proportional to the volume of the note, so loud notes have a greater impact on the string of spheres than soft notes.

When the musician stops playing, the sphere that was
The simple additive synthesis technique used in PR1 was also used in CC. That is, for each pitch class the frequency of the sounded note is ‘stored’ by the associated mass. A sine wave is played at that frequency at a volume that is proportional to the velocity of the mass. This means that the pitch of note played on the acoustic instrument sets the frequency of the sound produced by the vibrating string.

4. DESIGN CRITERIA AND USERS’ EXPERIENCES

In the previous sections we have outlined the characteristics of three virtual instruments. As we designed them we informally developed a set of principles or design criteria that guided our work. The criteria we discuss here are broad and relate mainly to the interaction design. There were of course a number of other criteria - that the software be cross-platform and make use of open source tools where possible for example - but these do not directly impact on the experience of musicians while they are using the software.

The criteria that emerged included:

1. The instruments should respond in a way that seems natural. In the prototypes we’ve tried to do this by basing the instruments on virtual physical models that respond in physically plausible ways.

2. Instrument response should be consistent. i.e. Two perceptually identical notes should appear to have the same effect on the virtual instrument.

3. Instruments should be conceptually simple but allow skilled musicians to create complex effects;

4. The instruments should have a sense of ‘character’ - be interesting, engaging and motivate the musician.

5. The musician should feel in control of the instrument, but it should retain the ability to surprise. That is, the musician may be stimulated to try something new, discover something about their technique or gain insight into some aspect of their music making.

6. The instrument should encourage a playful, exploratory approach, especially in new users. They should encourage the musician to consider questions such as ‘What does it do if I play...?’ and ‘How can I make it...?’

7. The relationship between live sound, the behaviour of the virtual instrument and the resulting sounds should be apparent to observers (eg. audience members).

A number of these criteria align with well-known criteria in this domain such as those proposed by Wessel and Wright [11] and it seems possible that at least some of them are generally applicable. In order to validate the criteria to some extent and help us understand their impact on users we are now investigating musicians experiences with the virtual instruments. We have therefore obtained feedback from expert musicians on whether the design criteria listed above have been met and, more broadly, what impact using the virtual instruments had on their music making, composition and teaching.

4.1. Methodology

‘Traditional’ Human Computer Interaction approaches have focused on measuring user performance when carrying out various well-defined tasks such as navigating a website or entering figures into a spreadsheet. Software designed to facilitate musical expression presents a problem in this context as it is difficult to formulate tasks to assign to users which are measurable but also meaningful[10]. If the aim was to produce a general-purpose musical instrument for performing music in a well-established tradition, then this would be simpler. Tasks such as playing a scale, trilling, etc. could be assigned and measurements made to ascertain how successfully users were able to execute them. The benefit of this approach is that it would be possible to objectively compare two different virtual musical instruments in terms of their playability.

However, where the instrument is intended to create new and unusual sounds - to explore new languages of composition and performance - this approach is problematic. Part of the rationale for creating these instruments was that they disrupt habitual ways of thinking about music so that musicians are stimulated to try new ways of playing and composing. Measuring how effective they are at facilitating performance of current styles of music might be interesting, but it would not necessarily help us learn more about designing to encourage divergent thinking.

Our approach was to give the musicians freedom to use the software in any way they wished - to make music with it in order to explore its potential. We used the concurrent think-aloud technique [3] in order to gain insight into their experience. That is, we asked the musicians to ‘think aloud’ as they interacted with the software. This does present some problems as our participants have to date been brass and woodwind players, who are obviously unable to speak and play their instrument at the same time. As we did not wish to interrupt the flow of performance, we did not ask musicians to interrupt their music-making to make comments. Instead we simply asked them to verbally report what they were thinking and perceiving as frequently as they were able during their time using the software. This meant that they played for some time,
commented on what was happening, played some more, made further comments and so on.

A disadvantage of this approach was that having to continually stop to comment on what they were doing may have prevented the musicians from becoming fully immersed in the music. That is, they may not have attained the ‘flow state’ which seems to be very important in creative work [2]. However, there is significant evidence that the think-aloud procedure has minimal impact on the cognitive processes of research subjects engaged in problem-solving tasks [3]. It does not seem unreasonable to assume that this extends to creative tasks also.

The instruction to participants was:

“We would like you to report what you perceive or think while you are interacting with the software. We would like to get as complete a report of what is going through your mind as possible, so please don’t worry if what you say is inconsistent or incomplete - just report what is going through your mind at the time. Don’t feel that you must break your concentration to make a report. Please try to give reports at a time that feels natural and appropriate to you.”

Following the familiarisation process, the musicians were asked to prepare and perform a short piece using the software. The instruction was:

“We would like you to prepare and perform a piece of music with the software. You can use manuscript paper if you wish to jot down any ideas but please don’t feel that you have to. There are no constraints on the style or duration of the piece; we’re just interested in how you use the software and your experiences in so doing. Please remember that we are evaluating the software not your performance.

“If you would like any aspect of the software adjusted let me know and I will do what I can to accommodate your request.”

After interacting with the software, we conducted a semi-structured interview with the musicians to explore interesting issues that had arisen during the session, and to elicit further comments on their experience. The questions were:

1. Tell me about the piece of music you wrote.
   (a) Why did it have these characteristics?
2. Do you have some comments about how easy or hard it was to write for or perform with the software?
3. Do you have some comments about the sound produced by the software?
4. Do you have some comments about the visual display?
5. While you were interacting with the work, did you become aware of any particular characteristics of your playing?
6. Did you play differently today than you normally would?
   (a) In what ways?
   (b) Can you say why?
7. Do you have any suggestions or proposals for how we might improve or extend this software?
8. Can you think of any uses for this software?
9. Is there anything I should have asked you?
10. Do you have some comments or questions about what we’ve done today? Is there anything we should have done differently?

5. SOME PRELIMINARY FINDINGS

We have only recently commenced our evaluations of the software. So far, six professional musicians (each with over 10 years professional experience) have participated in our study. We deliberately chose musicians who had an interest in contemporary music - especially those who also compose and/or improvise. Each session took approximately two hours and was video recorded, transcribed and analysed using techniques from grounded theory [4] [8]. The software Transana (http://www.transana.org) was used to facilitate this process. In addition, an observer was present and took notes at all sessions to provide an additional perspective.

This investigation of the users’ experience is ongoing. As such, the observations we report here are very preliminary but give an indication of some of the issues that are emerging. In this section we present some of the comments made by musicians during evaluations along with brief notes on our findings to date.

Q: “Do you have any comments about how easy or hard it was to use?” A: “Insanely easy. Very straight forward.”

The musicians that have participated to date found the software easy to use and have no difficulty in understanding the mapping of their sound to physical forces.

“What I like about it is the sensitivity; the ability to change the sound via my sound. What I’m finding challenging - I’m not saying it’s wrong - but what I’m finding challenging is if I have an effect that I like, to be able to guarantee that same effect on demand.”

“I’m always coming from a western harmonic tradition which has coloured my view of the whole thing, so I want to be able to set up the harmonies. When you have the sliding pitch [in PR2] it takes that away from you.”

Perhaps because all participants to date have been predominantly orchestral musicians, they felt that to use the software in a performance they would have to be in control. Mostly they did not feel comfortable if the software responded inconsistently. For these musicians, the issue of control - of knowing that they can “guarantee” a particular musical effect in a performance situation is very important.
Q: “Ok, so you mentioned about the more complicated harmonic response. Was there anything else about it that made you think of it more as a person?” A: “...It’s not linear. Maybe I’m not using the right terms, but it [PR1] swelled at times, so it gives you a feeling of conversation. Whereas the other one [PR2] felt specifically like a direct response to what I just played, where this feels more like a conversation.”

On the other hand, musicians liked the response to be rich and not necessarily linear or predictable. For example, a bug in the software for PR2 caused the playback sound to drop in pitch suddenly when the masses span out a long way from the central point. This was noticed by one of the musicians and the (minor) bug was removed. However, after playing with the bug-free version for a short time, the musician asked for the bug to be put back. He liked the unusual, surprising effect and was happy to have it in the software, even though it wasn’t consistent with the behaviour of the virtual instrument as a whole. For him, though it was critical that he understood exactly what conditions triggered the effect - he remained in control.

For several musicians, the interaction with the slower moving virtual instruments (PR1 and CC) seemed to be similar to interacting with a human musician in some ways. Partly this seemed to be because the sonic response was considered more complex than the samples that PR2 played back. In addition PR1’s slower, less mechanical movements seemed to appeal more ‘natural’ and life-like to the musicians.

It’s interesting that the musicians thought that the more complex virtual instruments (PR1 and CC) responded more consistently than Charmed Circle, even though all instruments used the same pitch tracking technique. Perhaps because this instrument is simpler and has a more direct mapping between live sound and behaviour, any inconsistencies in pitch recognition were more obvious to the user.

“What I like about the chord clusters is, if you’re thinking of...so like I’m writing a piece for trombone and organ and you’ve got a huge chord cluster. Well you’ve only got X amount of fingers. But with this [PR2] you’ve got the ability to really stack up the clusters.”

“I like it [CC], but it’s static and always comes back to the one point. I’d like spheres to explode or something like that when it was really loud. I’m interested in the one where they move in 3D. So, it’s good but I think it’s a little too limited for use in performance.”

In general, the musicians did feel that the virtual instruments facilitated complex musical effects, but there was not unanimous agreement on this. PR1 and PR2 were generally favoured because the visual and sonic feedback from these instruments was felt to be more complex and engaging. CC, with its simplified, more generic style was found to be more limiting. This is interesting because PR1 and PR2’s complexity comes at the cost of reduced controllability to some extent. CC offers far greater direct control over individual notes - more like a traditional instrument - but musicians felt it did not add enough to what they played into it. Comments on the audio indicate that more nuanced control over the timbre of the sounds that CC produces is an area to pursue. Scanned synthesis [9] is an interesting possibility here.

Q: “Did you find you were distracted by the visual aspect?” A: “I think at first it’s a lot to take on board. So you play something, you see it happen and you think ‘ok that’s good, what if I do this? Will it move further? Will it do this?’ I know there was one note I was hitting really loudly and it was going right to the top of the screen and I thought, ‘That’s going to bang into the C [mass].’ Then I realised [it didn’t work that way]. So there was an interesting visual aspect. Once I got more used to it I think that tonal things would predominate. At the moment it’s like a new toy.”

One musician commented that the desire to create interesting visual and musical effects simultaneously could be problematic, resulting in musical phrases being produced simply because they achieved a certain visual appeal. The quote above illustrates this problem. In general however, the musicians seemed to accept that the tight coupling between visual and sonic effects was simply part of the constraints that they needed to work within - part of the musical ‘problem’ that set the boundaries for their music-making. Of course because we want our software to encourage musical exploration in many ways the musicians use of the software in this way is actually quite desirable - see design criteria 6 (above).

6. CONCLUSION

In this paper we have presented three simple virtual musical instruments based on physical models. Design criteria that emerged from the process have been presented along with some preliminary observations from an ongoing investigation into musicians’ experiences using the instruments.

There is a lot of work still to do. Our hope is that by carefully analysing the data we have collected during the qualitative evaluations of this software we can improve our understanding of the relationships between the design criteria that emerged and the users’ experiences with the resulting software. Of particular interest are how the specific manifestations of these criteria influence the ability of musicians to express themselves musically. When we get the balance right, it seems that musicians deeply engage with the virtual instrument and interact with it as if it were a human player and not simply a digital ‘effect’ (such as delay or harmonisation). How specific design criteria influence this feeling of engagement is a critical question that we are very motivated to explore further.

7. ACKNOWLEDGEMENTS

We would like to express our gratitude to the developers of Pure Data, Graphical Environment for Multimedia, Physical Modeling for Pure Data and Transana for providing the tools to make this work possible. Thankyou to the musicians for being so generous with their time and providing
us with such insightful feedback.

8. REFERENCES


