INCORPORATING HAPTIC FEEDBACK INTO INTERFACES FOR MUSIC APPLICATIONS

M. SILE O'MODHRAIN

Center for Computer Research in Music and Acoustics (CCRMA), Stanford University

CHRIS CHAFE

Center for Computer Research in Music and Acoustics (CCRMA), Stanford University

ABSTRACT

Though musicians rely primarily on their sense of hearing to monitor and adjust the sound being produced by their instrument, there exists a second path through which valuable information about the instrument's behavior can be observed - namely the feedback received via the haptic senses, the senses of touch and kinesthesia. The present study tested the hypothesis that leveraging off the musician's unconscious use of combined auditory and haptic cues by adding haptic feedback to computer-based musical instruments would improve the "playability" of these instruments.

Twenty experienced musicians played a series of short melodies on an unfamiliar computer-based musical instrument. The instrument's controller was capable of producing force feedback, making it possible to systematically vary the instrument's "feel." Results indicate that the presence of haptic feedback improved playing accuracy by approximately 23 percent. Moreover, haptic feedback which was correlated with changes in the musical parameter being controlled produced greater playing accuracy than uncorrelated haptic feedback.

KEYWORDS: music, performance, synthesis, haptics, force-feedback, theremin

INTRODUCTION

When a musician plays an instrument, they perform certain actions with the expectation of achieving a certain result - a musical performance. As they play, they monitor the behavior of their instrument and, if the sound is not quite what they expect, they will adjust their actions to change it. In other words, the musician has effectively become part of a control loop - they constantly monitor the output from their instrument and subtly adjust bow force, breath pressure, or whatever control parameter is appropriate. A violinist, for example, uses their sensitivity to force and vibration to control bow velocity. A trombone player can "feel" where the resonant modes of their instrument are by an increase in vibrations fed back to their lips via the mouthpiece [1].

Sophisticated sound synthesis techniques such as "physical modeling" [2] provide composers and performers with the opportunity to change any aspect of their instrument, often in real time. Potentially, a player can alter the size, shape and even the material composition of an instrument as they play. The challenge presented by such flexibility is how to provide the performer with access to appropriate control parameters. The solution which we propose is to leverage off the musician's existing sensitivity to the relationship between their instrument's "feel" and its sound. By adding haptic feedback to sound synthesis and manipulation tools, we will be able to couple a whole repertoire of physical interactions with instrument control parameters.

BACKGROUND

The enhancement of computer-based instrument interfaces with haptic feedback dates back to the late 1970's, when Claude Cadoz and his colleagues built an experimental device that relayed forces, mapped to some aspect of a sound synthesis model, back to a user's hand by means of a motorized joystick. The development of their subsequent series of experimental devices, which they call "Gestural Force-feedback Transducers", is predicated on the realization that the physical instrument-performer relationship is bidirectional, i.e. the performer both transmits information to their instrument in the form of gestures, and receives information back from the instrument in the form of tactile/kinesthetic feedback. Further, they suggest that, in many cases, the instrumental gesture is the best way to communicate appropriate sound control parameters. Their system therefore exploits a modular design in which the gestural controller (with its appropriate force-feedback) can be associated with any sound synthesis model [3].

In a similar study, Gillespie [4] modeled the kinematics of a grand piano's action and built a one-octave force-feedback piano keyboard to display forces derived from this model to the player's hand. Whereas Cadoz's system derives the forces relayed to the player from aspects of the sound synthesis model being controlled, Gillespie derived his force-feedback directly from the mechanics of the instrument's action itself.

In a further study [5] the present authors were able to demonstrate that the relationship between haptic and auditory cues need not be based on the behavior of a known musical instrument. With an eye toward incorporating haptic feedback into interfaces for sound editing applications, we recorded two different performances of the opening bars of a Beethoven piano sonata and derived a parameter that described the increase and decrease of musical "tension", as projected by two experienced pianists. We then mapped these changes in musical tension to changes in stiffness of a virtual wall so that, as the music progressed, it was possible to "feel" the expressive contour imposed by the player.

Though these studies have incorporated haptic feedback into instruments, none have provided empirical evidence for the utility of haptics in music applications. The experiment presented here was carried out to provide some measures for assessing the importance of haptic feedback in controlling computer-based instruments such as those discussed above.

EXPERIMENT

One useful way to think of the relationship between a performer and their instrument is to model the performer/instrument interaction as a simple feedback controller [6].

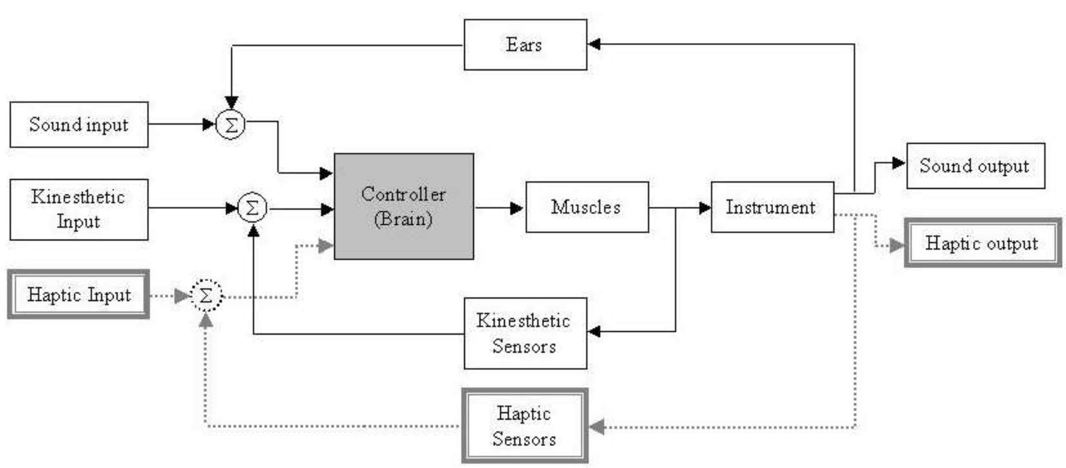

Figure 1. Musician as feedback controller.

The principal inputs to the system - the sound (auditory), feel (haptic), and layout (kinesthetic) of a given instrument - are fed via their associated sense organs to the controller, the brain. Based on these inputs and on knowledge of musical goals, the controller issues commands to the appropriate motor organs to modify the behavior of the instrument. The results of these actions are fed back to the controller via auditory, haptic and kinesthetic sensors. With this model in mind, our study tested the hypothesis that adding force-feedback to interfaces for computer-based musical instruments would improve the accuracy with which they could be played. The instrument we chose as our model was the theremin, an early electronic instrument which uses electric field sensing to gauge the position of the player's hands in space. The right and left hands control the pitch and amplitude respectively of a continuously sounding tone. There is no mechanical coupling between instrument and player.

In theremin performance, information about the state of the instrument is limited to auditory and kinesthetic feedback alone (as indicated by the dashed and grayed-out portions in Figure 1). The instrument is considered to be extremely difficult to master (only a handful of players have ever achieved a level of skill comparable with that of a concert performer.) By happy accident, we discovered that coupling the player's hand to the instrument's antenna via a simple elastic band made it much easier to control. This lead us to hypothesize that the increases and decreases in tension in the elastic band were providing additional feedback that was somehow making it easier to judge the amplitude of changes in the control parameter.

In order to discover what kind of haptic feedback would be appropriate for the control of a continuously varying parameter - in this case pitch - we built a virtual theremin in software and coupled it to a haptic display. We then measured the accuracy with which a player could play a melody in conditions where changes in force mapped directly to changes in pitch against conditions where the force-feedback was constant.

Method

Participants. Twenty members of the Stanford Symphony Orchestra participated in this study. They ranged in age from 18 to 28 with a mean age of 22; 11 participants were female; 20 were predominantly right-handed; and all were experienced musicians with an average of 13 years of musical training. In a pre-test questionnaire, 16 participants said they spent between 10 and 20 hours practicing or performing each week, while the remaining 4 practiced for more than 20 hours. All received $20 gift certificates for participating in the study.

Apparatus. Our experimental apparatus consisted of a haptic display device, the "Moose," and a PC with a software MIDI synthesizer (see [7] for a description of the haptic display hardware.) The various force conditions were generated in software in real time. Sample melodies were "played back" via MIDI, while the pseudo-theremin's sound was produced by the PC's internal speaker (since this produced the most realistic theremin sound. It was also spatially separated from the MIDI synth sound output, making it easier for participants to separate the two audio sources.) A computer-controlled metronome, synchronized with the tempo of the current melody, ensured rhythmic accuracy.

Stimuli. The experiment employed a 6 by 12 within-subjects factorial design, with repeated measures on each condition. The factors were feedback condition (six levels, Table 1) and melody (12 levels.)

The melodies used were opening phrases of melodies taken from the Themefinder database, maintained by the Center for Computer-Assisted Research in the Humanities (CCARH), at Stanford (see `http://www.themefinder.org`) The melodies chosen were diatonic, ranging in length from 9 to 16 notes with no rests or directly repeated pitches. All contained both a rising and falling perfect fifth (A perfect fifth being the interval between the first and fifth notes of a major or minor scale.) This was to enable a comparison of overshoot (both rising and falling) across all melodies and all conditions (see Results - Overshoot section below.)

Both force conditions and melodies were presented randomly, each force condition appearing twice and each of 12 melodies (selected at random from a pool of 18) appearing only once.

Table I. Six force-feedback conditions.

1) Viscous damping	changes in force depend on velocity	$F = Bv$
2) Constant positive force	in the direction of pitch change	$F = 0.2$
3) Constant negative force	opposed to the direction of pitch change	$F = -0.2$
4) Positive spring	force increases as pitch increases	$F = Kx$
5) Negative spring	force decreases as pitch increases	$F = -Kx$
6) No force feedback	(like the original theremin)	$F = 0$

F is force in Newtons, B is the damping coefficient in N/M/Sec, v is velocity, K is spring stiffness, a constant, x is position.

Procedure. Each participant completed 12 experimental trials. Before beginning the first trial, participants completed a short training period (approx. 10 minutes) in which they were given a practice melody and asked to play this melody in all six force conditions. When the participant was comfortable playing the practice melody in time and in tune and had "felt" each force condition, they proceeded to the first experimental trial. The format of all experimental trials was the same: The participant was given a score of the melody to be played and then listened to it once. Thereafter, they were free to listen to the melody again, play along with it or practice on their own until they were ready to "perform" the melody. The experimenter then recorded one performance of the melody and proceeded to the next trial (each trial lasted approx. 3 minutes). Data were recorded to disk for later analysis (data was non-audio, consisting of pitch frequency, force, position and time measurements sampled at 1 KHz).

Results

Accuracy. To test the hypothesis that adding force feedback would improve the accuracy with which participants could perform simple melodic phrases, we calculated an RMS error for each performance by comparing it against a computer-generated template for that melody. If the type of force feedback made it easier for the player to control the instrument, then the error for melodies played under that force condition should be correspondingly lower. Our results, summarized in Figure 2 below, indicate that participants were least accurate when no force feedback was provided (mean RMS error = 45.16) and most accurate when changes in force mapped directly to changes in the parameter being controlled (mean RMS error = 34.79, a decrease in RMS error of 23%.)

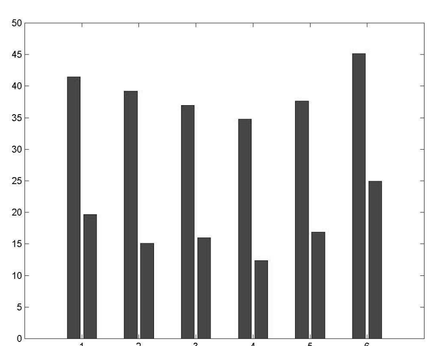

Figure 2. Mean and STD for accuracy by force conditions 1 - 6 of Table 1 above.

Evidently, one reason the standard deviation of the mean RMS error is so large is that there is a large variation in the difficulty of the melodies (with the normalized RMS error on each melody ranging from 1.46 to 0.55 where the average difficulty of all melodies is about 1.0).

Overshoot. An alternative way to evaluate the effect of the six force-feedback conditions on the controllability of our instrument is to look at the time taken to move from one pitch to the next and the amount by which the player overshot the target. In particular, the player/instrument relation can be modeled as a linear second order system [8]. A desired pitch change of a fifth will cause the actual instrument response to have an associated rise time and overshoot. In the same way, a desired step change for a linear second-order system will result in a response with some overshoot and an associated rise time. Table II below shows the characteristics of the step response for a change in pitch of a perfect fifth (a movement amplitude of 0.25 of an inch) measured for each force condition.

Table II. Overshoot data. Percentage overshoot and movement time for rising and falling perfect fifths.

Force	Rising % OS	Rising t (msec)	Falling % OS	Falling t (msec)
damped	15.2 ($\zeta = 0.5$)	333.9	11.7	294.4
pos const	11.2	245.2	13.2	263.9
neg const	14.0	249.3	13.9	237.8
pos spring	12.3	304.3	11.0	256.9
neg spring	9.9 ($\zeta = 0.6$)	296.9	11.5	263.5
no force	11.0	269.6	11.9	256.8

The poorest response was for the viscous damper rising fifth, while the best response was for the negative spring rising fifth (overshoot = 15 vs. 10 %, time delay = 334 vs. 297 msec). For a linear second-order system the overshoot to a step response can be expressed as a damping ratio. Using equation 1 below the damping ratio associated with the 15 percent overshoot of the viscous force-feedback condition turns out to be approximately 0.5, while that associated with the 10 percent overshoot in the negative spring condition is 0.6.

$$\frac{d^2x}{dt^2} + 2\zeta\,\omega_n\,\frac{dx}{dt} + \omega_n^2 = u(t) \qquad \text{(equation 1)}$$

where t is time, x is the dependent variable (instrument pitch), ω_n is the systems natural frequency in radians, and $u(t)$ is the step input.

DISCUSSION

Playing Accuracy

Our measure of playing accuracy, the RMS error resulting from a comparison of a given "performance" of a melody against a computer-generated template of that melody with "perfect" timing and pitch, is somewhat stringent because in real life no musician would play this strictly. However, as all experienced musicians know, it is impossible to infuse a performance with expressive nuances of timing and pitch if a strict model of the music does not exist even if only in the mind of the performer. Having instructed our performers to play as "accurately" as possible, we can therefore assume that they are doing their best to reproduce this internal stringent model.

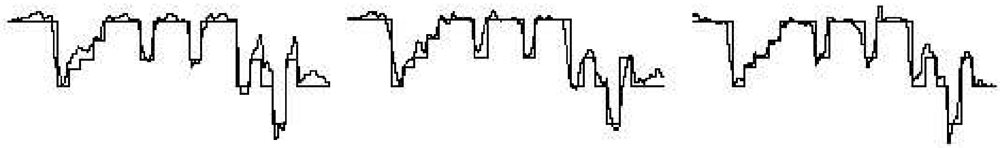

Figure 3. Three theremin performances of the same melody by different subjects under different force conditions (1, 3 and 4, above). Measured pitch curves are shown inscribed over a template computer performance (consisting of instantaneous pitch changes between perfectly tuned pitches).

As predicted, the mean RMS errors for the force conditions we tested indicate that the presence of haptic feedback provides cues that improve the "playability" of our instrument. Why should this be so? One explanation might be that the presence of force-feedback, regardless of its nature, is providing additional cues (contact forces, etc.) which make it easier for the player to control the amplitude of their movements. If this is the case, then the five test conditions where some form of haptic feedback was available should fair better than the non-force condition. As can be seen in figure 4, this is indeed the case. Whereas performances in the non-feedback condition resulted in a mean RMS error of 45 Hz, the mean RMS error of all other conditions combined was 38 Hz. However, if damping of limb oscillation was the only factor responsible for improvement in performance, the five force-feedback conditions tested would fair equally well. The fact that there are small but consistent differences in errors across conditions suggests that something other than damping of limb oscillation is assisting players.

The force-feedback condition which produced the most consistent and smallest RMS errors was the positive spring. Here, it will be recalled, change in pitch was correlated with change in force so that as pitch rose, force increased. The most probable explanation for the greater accuracy here is the effect of stimulus congruence, i.e. better performance with congruent than with incongruent stimulus combinations [9]. Perhaps the human mechanism for pitch production, which requires more physical effort to produce higher notes, predisposes us to associate greater effort with higher pitch. Secondly, in real-world musical instruments with continuous pitch, the logarithmic nature of pitch results in a non-monotonic distribution of "notes" along the instrument's length. For stringed instruments, for example, playing a rising chromatic scale on one string would result in smaller and smaller inter-note distances as pitch increased. Thus the non-monotonic nature of pitch may somehow conjoin, albeit subconsciously, with non-monotonic force-feedback (and help counter the present theremin's linear pitch arrangement).

Overshoot

As MacKenzie [10] points out musical tasks belong to a class of interactions that are constrained primarily by time (as opposed to pointing tasks which are dominated by spatial constraints.) For such tasks, a move proceeds as accurately as possible and terminates at a specified time. In the case of the theremin, the overshoot and time delay associated with the approach to a target pitch accurately describe the player's ability to control this parameter. Describing the player/instrument interaction in terms of the step response of a linear second-order system yields yields one value, the damping ratio, which encapsulates the player's ability to operate within the temporal constraints of a musical task given any force condition. Though the model adequately describes the controllability of the instrument, it is worth noting some interesting nuances that arise within individual force-feedback conditions. Firstly, the results for rising and falling leaps are not symmetrical under all conditions. Even in the no force condition, players consistently under-estimated rising fifths and overshot falling fifths. Again, this may relate to vocal pitch production. It is well known by choral conductors that singers will undershoot large upward leaps, where more physical effort is required to achieve pitch accuracy, and will overshoot downward leaps because they simply relax too much.

Are such differences, idiomatic of particular force conditions, significant? A recent study by Huron, et al. [11] showed that the dynamics associated with playing a particular instrument over a number of years shape the player's internal representation of musical articulation, in general. As Huron points out, the player may or may not learn to compensate for the behavior of the instrument or, as is more probably the case, the dynamics idiomatic of the instrument's action are incorporated into the "sound" associated with performances on that instrument.

SUMMARY AND FUTURE WORK

The present study has provided empirical evidence supporting the hypothesis that adding haptic feedback to interfaces for computer-based musical instruments improves the player's ability to control these instruments. Moreover, it has been shown that those force conditions where haptic feedback was congruent with auditory feedback resulted in better performance than conditions where auditory and haptic feedback were not correlated. The results, obtained within the context of a musical performance, indicate that modeling the player's ability to control a given parameter as the step response of a linear second-order system adequately captures the temporal constraints of a musical task. However, most musical instruments (real or simulated) require the skilled player to control many parameters simultaneous (pitch, amplitude, timbre, etc.) Understanding the player/instrument interaction in terms of a control system therefore requires us to take this additional complexity into account. We are currently working to learn more about the interaction of multiple control loops, each with its associated haptic, auditory and kinesthetic feedback, and each with its own temporal constraints. Armed with these tools, we will be able to press toward our ultimate goal - namely to extend the study to control of parameters for a more complex and realistic player/instrument interaction.

ACKNOWLEDGMENTS

The authors wish to thank Craig Sapp for his help in preparing the musical stimuli for this study and Herb Rauch for his help in developing mathematical and statistical models.

REFERENCES

1. Chafe, C. "Tactile Audio Feedback." *Proc. of Intl. Computer Music Conf.* Tokyo (1993), 76-79.

2. Smith, J.O. "Physical Modeling using Digital Waveguides." *Computer Music J.* 16(4) (1992), 74-91.

3. Cadoz, C. et al., A Modular Feedback Keyboard Design. *Computer Music J.* 14(2) (1990), 47-51.

4. Gillespie, R.B. "Haptic Display of Systems with Changing Kinematic Constraints: The Virtual Piano Action." Ph.D. Thesis, Stanford University, 1996.

5. Chafe, C. and O'Modhrain, M. S. "Musical Muscle Memory and the Haptic Display of Performance Nuance." *Proc. of Intl. Computer Music Conf.* Hong Kong, (1996), 427-431.

6. Gillespie. R.B. "Haptics." *Music, Cognition, and Computerized Sound: An Introduction To Psychoacoustics,* Edited by Perry Cook. MIT Press, Cambridge, Mass. (1999) 229-245.

7. Gillespie, R.B. and O'Modhrain, M. S. "The Moose: A Haptic User Interface for Blind Persons." *Proc. of the 3rd Annual WWW6 Conf.* Santa Clara, California, (1997). Also available as STAN-M-95, CCRMA Tech. Rep. Music Dept., Stanford University.

8. Ramos, C. "How Models Help Us to Understand the Interactions Between Posture and Voluntary Movements." *IEEE Eng. in Medicine and Biology Magazine,* 11(4) (1992), 91-92.

9. Ben-Artzi, E. and Marks, L.E. "Visual and Auditory Interaction in Speeded Classification: Role of Stimulus Difference." *Perception & Psychophysics*, 57(8) (Nov, 1995), 1151-1162.

10. MacKenzie, I.S. "Fitts' Law as a Research And Design Tool in Human-Computer Interaction." *Human-Computer Interactions,* 7(1) (1992), 91-139.

11. Huron, D. and Berec, J. "Musical Organization and Instrumental Idioms: A Case Study of the B-flat Valve Trumpet." *Computing and Musicology,* (2000) in press.