Chapter 4

Theremin and Variations

4.1 Introduction

One useful way to think of the relationship between a performer and an instrument is to model the performer-instrument interaction as a simple feedback controller. (See Gillespie, 1999.) A block diagram of the controller is shown in Figure 4.1.

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, the experiments reported here tested the hypothesis that adding force feedback to interfaces of computer-based musical instruments would improve the accuracy with which they could be played. The instrument we chose as our model was the Theremin (shown in Figure 4.2 being played by its inventor.) The Theremin is an early electronic instrument that 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.

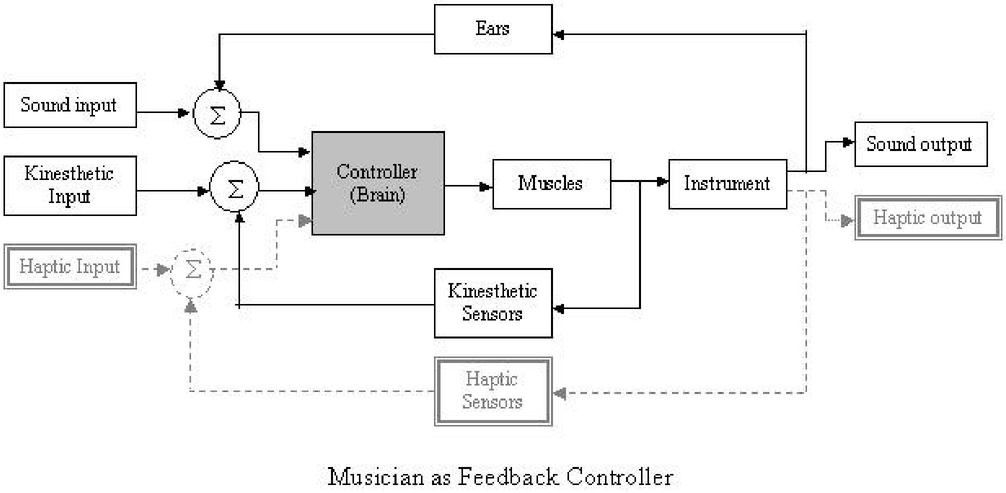
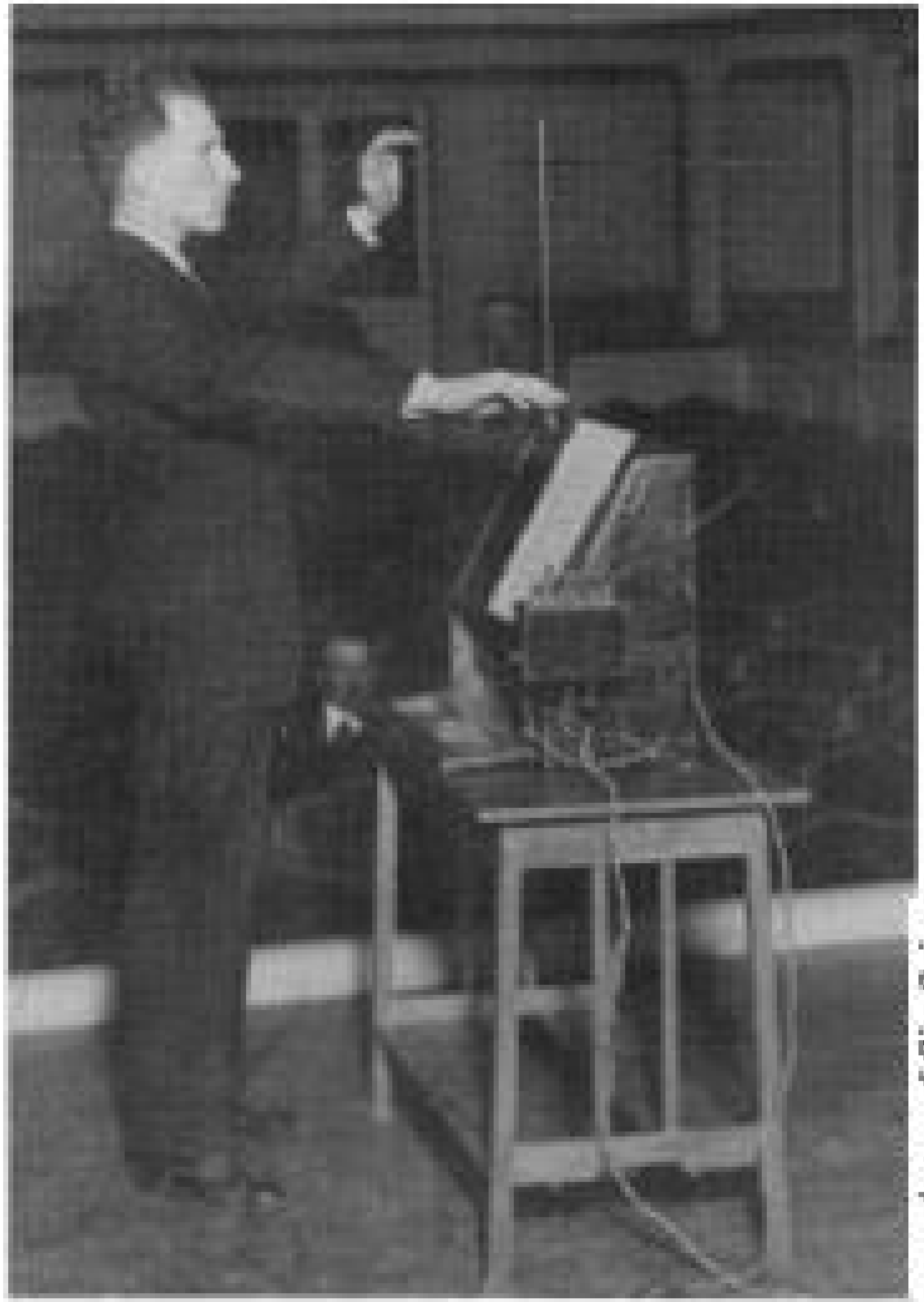


Fig. 4.1. Musician as feedback controller.

4.2 Experiment I: Variations on a Theremin

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 4.1). By coupling the player's hand to the instrument's antenna via a simple elastic band, we discovered that the instrument became much easier to control. This led 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 parameter being controlled. 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 under different conditions of force feedback. We compared relative accuracy when changes in force were correlated to changes in pitch to conditions where force and pitch were uncorrelated.



Courtesy of Big Briar

Fig. 4.2. Leon Theremin playing his invention.

4.2.1 Experiment

Participants

Twenty-one 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; 15 were predominantly right-handed; and all were experienced musicians with an average of 13 years of musical training. In a pre-test questionnaire (see appendix A), 16 participants said they spent between 10 and 20 hours practicing or performing each week, while the remaining 5 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. The Moose is a prototype haptic display developed at the Center for Computer Research in Music and Acoustics (CCRMA) by Brent Gillespie O'Modhrain and Gillespie (1996).

The various force feedback conditions were generated via software in real time. The virtual Theremin's sound was produced by the PC's internal speaker. A computer-controlled metronome, synchronized with the tempo of the current template melody, ensured rhythmic accuracy. The stimulus melodies were played back through MIDI so that their timbre was distinct from that of the virtual Theremin.

Stimuli

The melodies used for this study were opening phrases of melodies taken from the Themefinder database, maintained by the Center for Computer-Assisted Research in the Humanities (CCARH), at Stanford University (see Appendix B). The melodies chosen were diatonic (major or minor), ranging in length from 9 to 16 notes with an average of 12 notes. Melodies contained no rests or directly repeated pitches. All

Fig. 4.3. Prototype haptic display device: the Moose. Two linear voice-coil motors are coupled to the white puck in the center of the Moose's workspace by two perpendicularly-oriented double flexures. The puck's position is tracked by two linear encoders and the whole is interfaced to the host computer via a digital I/O card.

contained both a rising and falling perfect fifth. (A perfect fifth is the interval between the first and fifth notes of a major or minor scale.) This was to enable a comparison of movement overshoot for both rising and falling intervals across all melodies and all conditions (see Results - Overshoot, Section 4.2.2).

Each participant completed 12 trials. For each trial, one of six different force conditions was selected at random (see Table 4.1).

Figure 4.4 illustrates the relationship between pitch and force for each of the six force conditions listed in Table 4.1).

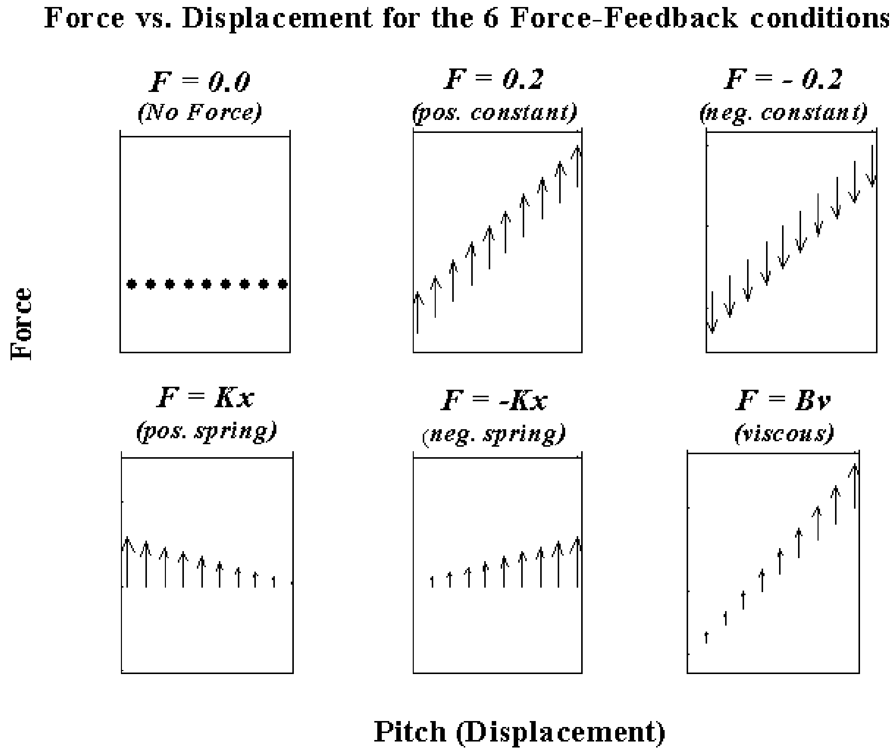


Fig. 4.4. Relationship between pitch and force for the six force feedback conditions.

For each participant, each force condition occurred twice and each of 12 melodies (selected at random from the pool of 18 in Appendix B) appeared only once.

Procedure

Each participant completed 12 experimental trials. Before beginning the first trial, participants completed a short training period (approximately 10 minutes) in which they were given a practice melody and asked to play this melody in all six force conditions. When participants were comfortable playing the practice melody in time and in tune and had experienced each force condition, they proceeded to the first experimental trial. The format of all experimental trials was the same: the subject was given a score of the melody to be played and then heard the melody once. Thereafter, the participant was free to do one of three things: (i) listen to the melody

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, and x is position

Table 4.1. Six force-feedback conditions

again, (ii) play along with it or (iii) practice alone until he or she was ready to "perform" the melody. At this point, the experimenter recorded one performance of the melody and proceeded to the next trial (each trial lasted approx. 3 minutes). The data was recorded to disk for later analysis. Recorded data was non-audio, consisting of pitch frequency, force, position and time measurements sampled at 1 kHz.

4.2.2 Results

In preliminary inspection of the data showed that both the best and worst performers were significantly different from the remaining participants. We therefore did not include their data in subsequent analyses because we judged them to be significant outliers.

Playing Accuracy

In order to estimate playing accuracy, we scored each performance against a computer-generated template for the same melody. Each score represented the average RMS error of pitch in Hertz across the entire melody for that performance.

We then averaged scores over all participants for melodies played in each force condition to test whether the force condition had a significant effect on performance. If a given form of force feedback made it easier for the player to control the instrument, then the score for melodies played under that force condition should be correspondingly lower. Our results, summarized in Figure 4.5, indicate performance tended to be better in all conditions where force feedback was present (conditions 1-5) than in the no-feedback condition (condition 6). Participants were least accurate in force condition 6, i.e. when no force feedback was provided (mean score = 45.16) and most accurate in force condition 4, when changes in force mapped directly to changes in the parameter being controlled (mean score = 34.79, a decrease of 23%).

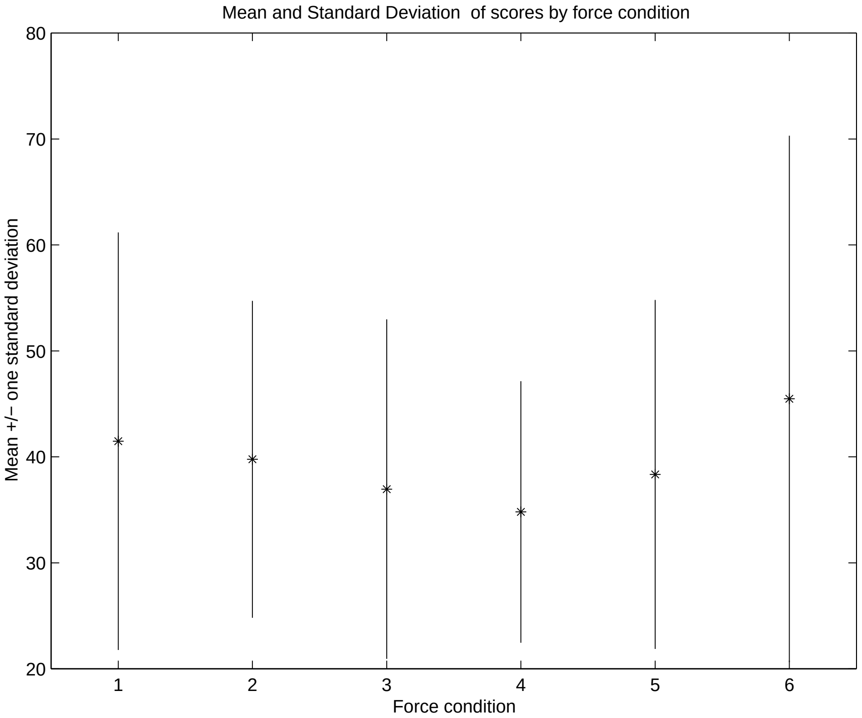


Fig. 4.5. Histogram showing mean score and standard deviation for each force condition.

The standard deviations of the average scores for the six force conditions are relatively large. Some melodies were significantly more difficult than others and some subjects were more skilled than others. The normalized score for each melody ranged from 1.58 to 0.57 (where the normalized difficulty of all melodies is about 1.0). Also, the normalized score for each subject ranged from 1.16 to 0.79 (where the normalized difficulty for all subjects is about 1.0). Both of these factors contributed to the variability in scores within each force condition. Figure 4.6 shows the distribution of normalized scores by melody and subject.

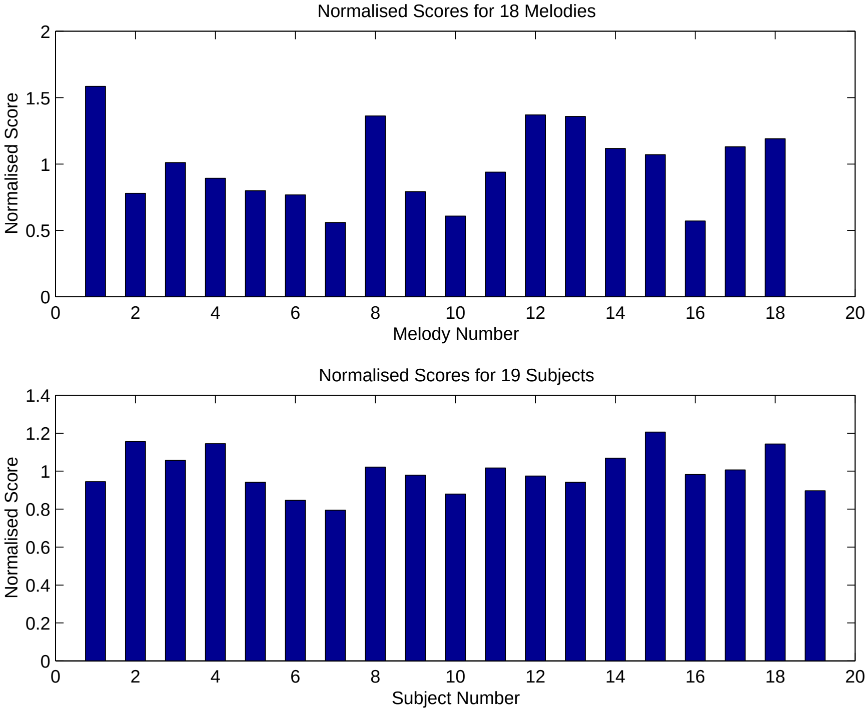


Fig. 4.6. Distribution of normalized scores by melody and subject.

We used a standard least squares regression to remove the effects of differences in melody difficulty and subject skill. The standard deviation of the residual scores was reduced by 33% and their variance by 55%.

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 points in melodies where players had to make large leaps and see how accurately they could reach a target pitch. Our approach was to model the player/instrument as a linear second-order system

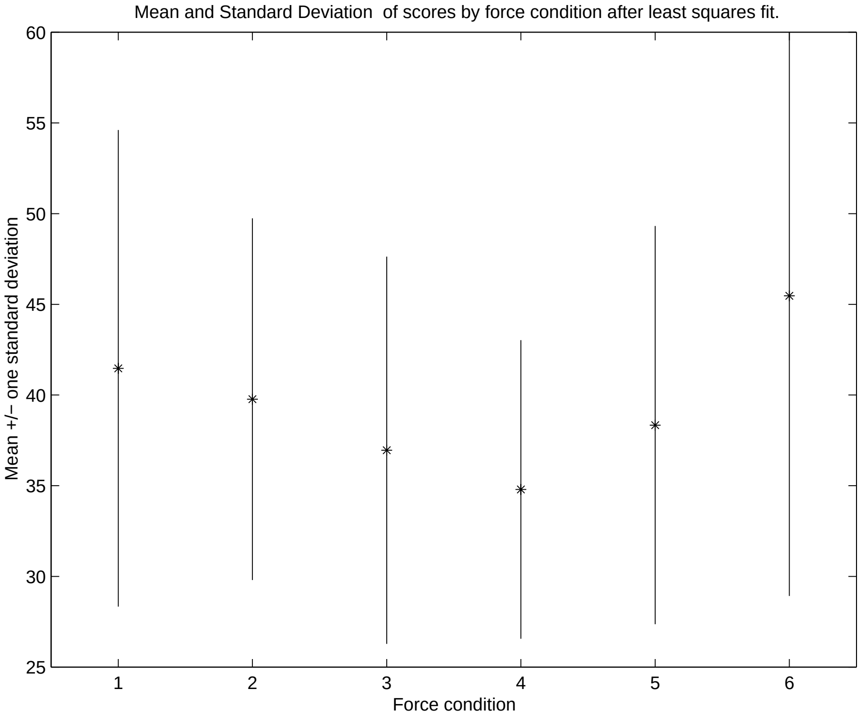


Fig. 4.7. Mean and standard deviation of melody scores by force condition after least squares fit.

(Ramos, 1992). A desired pitch change of a fifth will cause the actual instrument response to have some overshoot and some rise time. In the same way, a desired step change for a linear second-order system will result in a response with some overshoot and some rise time.

In Table 4.2 we show the characteristics of the step response for a change in pitch of a perfect fifth (a movement amplitude of 0.25 inches) measured for each force condition. Percent overshoot is defined as the movement overshoot or undershoot as a percent of the amplitude of the movement required.

The poorest response in terms of overshoot was for force condition 1, the viscous damper rising fifth, while the best response was for force condition 5, the negative spring rising fifth (overshoot = approx. 15% vs. 10%, rise time = 334 vs. 297 msec). For a linear second-order system the overshoot to a step response can be expressed as fractional damping, ζ.

Force	Rising %OS	Rising t (msec)	Falling %OS	Falling t(msec)
damped (1)	15.2 ($\zeta = 0.5$)	333.9	11.7	294.4
pos const (2)	11.2	245.2	13.2	263.9
neg const (3)	14.0	249.3	13.9	237.8
pos spring (4)	12.3	304.3	11.0	256.9
neg spring (5)	9.9 ($\zeta = 0.6$)	296.9	11.5	263.5
no force (6)	11.0	269.6	11.9	256.8

Table 4.2. Overshoot data. Percentage overshoot (OS) and movement time for rising and falling perfect fifths.

$$\frac{d^2x}{dt^2} + 2\zeta\omega_n\frac{dx}{dt} + \omega_n^2 = u(t) \tag{4.1}$$

where t is time, x is the dependent variable (instrument pitch), ω_n is the system's natural frequency in radians, and $u(t)$ is the unit step input.

Using equation 4.1, the fractional damping associated with the 15% overshoot of the viscous force feedback condition turns out to be approximately 0.5, while that associated with the 10% overshoot in the negative spring condition is 0.6.

4.2.3 Discussion

Playing Accuracy

Our measure of performance accuracy was the least-squares distance in Hertz from a given performance to a computer-generated template of that melody. The least-squares solution, which is less sensitive to small deviations from the template than to large ones, proved to be a useful tool for this task being tolerant of small musical deviations but intolerant to gross errors.

Figure 4.8 shows 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).

Fig. 4.8. Three Theremin performances of the same melody by different subjects under different force conditions.

If the hypothesis is borne out and the presence of haptic feedback is providing additional cues that improve the playability of the virtual Theremin, then the five test conditions where some form of haptic feedback was available should fare better than the control condition where no feedback was available. As can be seen in Figure 4.7, this is indeed the case. Whereas performances in the no-feedback condition (force condition 6) resulted in a mean score of 45, the mean score of all other conditions combined was 38.

The force feedback condition that produced the smallest scores was the positive spring (force condition 4). Here, it will be recalled, change in pitch was correlated with change in force so that as pitch rose, force increased. Viscous damping (force condition 1), on the other hand, was the least successful of all force conditions. The condition that fared worst was the condition closest to the real Theremin, namely the no-force condition. Thus the presence of even inappropriate force feedback was apparently better than no force feedback at all.

However, overall, the differences in feedback condition were not significant. Therefore we concluded that changing force feedback condition had little effect on performance.

Fractional Damping

As (MacKenzie, 1992) has pointed 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 rise time and overshoot associated with the approach to a target pitch together accurately describe the player's ability to control movement duration and movement amplitude. Modeling a player's performance as the step response of a linear second order system yields a single quantity, fractional damping, that encapsulates both rise time and overshoot behavior. Fractional damping can therefore be used as a measure of a player's ability to operate within the temporal constraints of a musical task given any of our force conditions. 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. In the positive spring condition, for instance, players underestimate rising leaps and overshoot falling leaps (see Figure 4.9).

Such dynamic behavior, idiomatic of a particular force condition, can therefore influence the characteristic sound of even the most basic computer-based musical instrument. A recent study by Huron and Berec (1995) showed that the dynamics associated with playing a particular (traditional) instrument over a number of years shape the player's internal representation of musical articulation in general. As Huron and Berec 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. If these force-dependent differences in performance persist over longer periods of exposure to the instrument, they suggest that the ability to associate haptic feedback with the production of sound in computer-based instruments could potentially make available a whole repertoire of physically-based behaviors for composers and performers to explore.

Temporal Invariants in Movement

In the analysis of our performance data, one clear pattern of movement from note to note began to emerge. Each player developed a characteristic movement trajectory

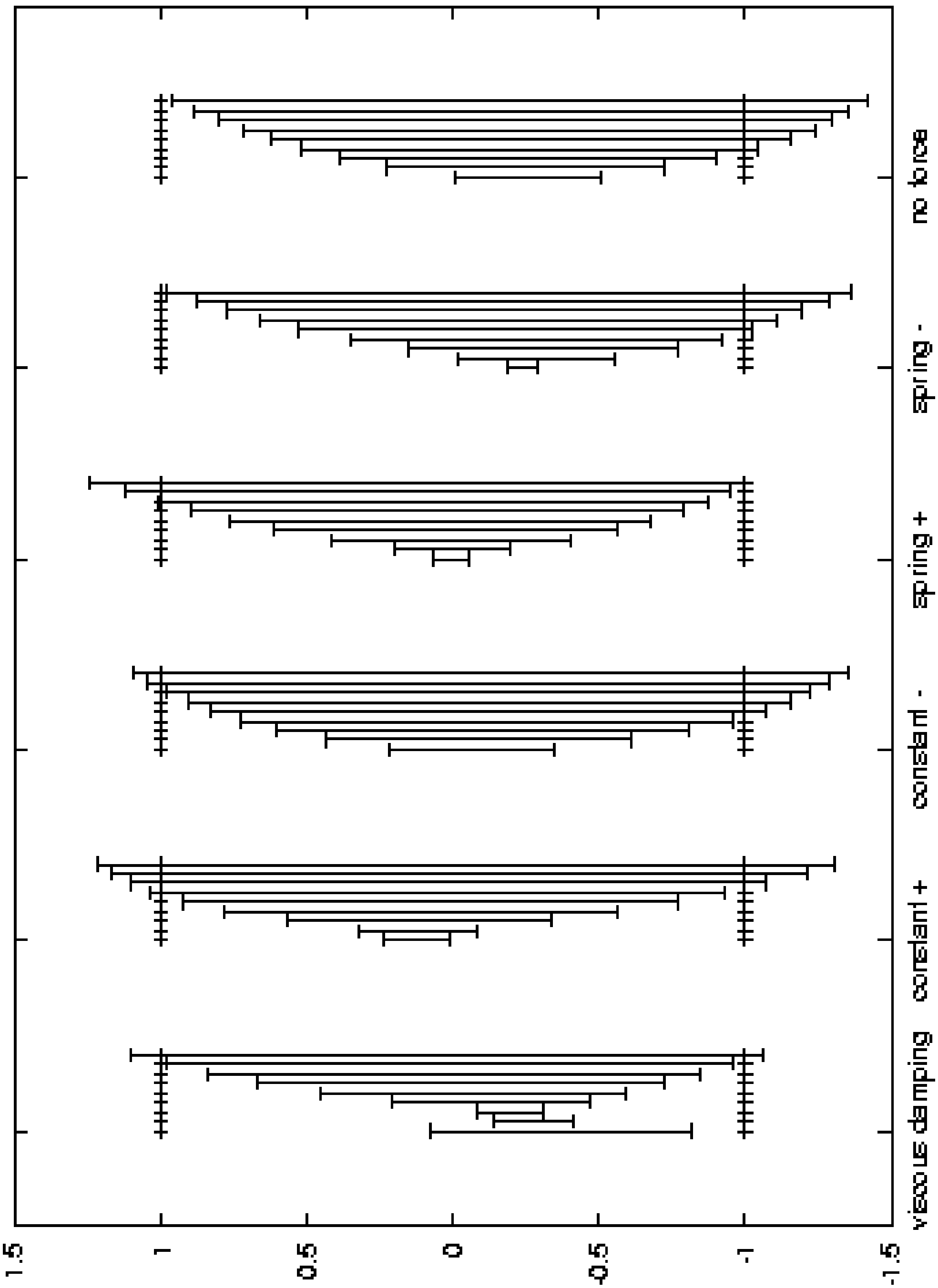


Fig. 4.9. Overshoot plot for each of the six force feedback conditions.

whose duration was dependent not on the lengths of notes before and after it, nor on the tempo of the melody, but on the amplitude of the movement itself. Figure 4.10 is a plot of the movement trajectories for a performance of the opening measures of the main theme from Tchaikovsky's Swan Lake. Note the differences in trajectories to and from the highest pitch, B, for leaps of different sizes and between notes of different durations. Here, each quarter note beat lasts 1 second (i.e. MM = 60).

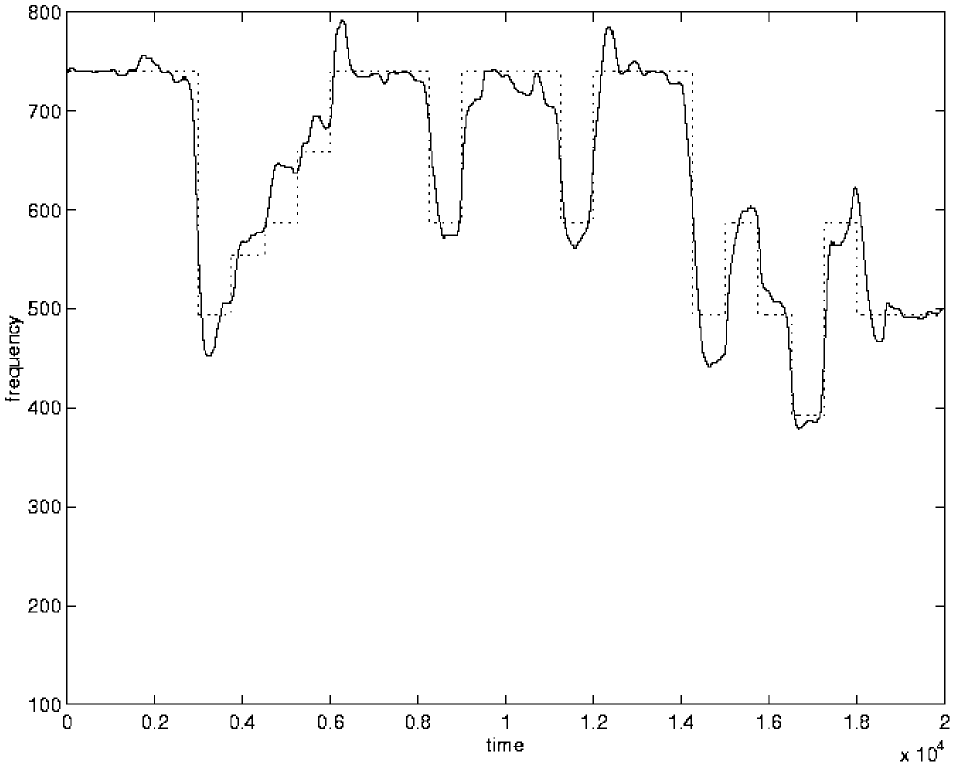


Fig. 4.10. Theme from "Swan Lake".

Why is it the case that movement velocity is constant regardless of musical timing constraints? The answer most likely lies in the structure of the underlying motor program that controls the movement itself. One question that has occupied motor behavior researchers for many years is whether skilled movements are a product of stored programs that control precise timing and amplitude of individual movements, or whether these programs are more general and control relative movement timing and force parameters (see Chapter 3). In arguing for the latter, (Schmidt, 1976) has proposed a generalized theory for motor programs in which some elements of the program are fixed while others are varied in accordance with the goal of the movement at any given time. In support of this argument, Schmidt cites a study by (Lashley, 1942) in which subjects were asked to write the words "motor equivalence" using different end-effectors (non-dominant hand, pencil in teeth, etc.) Lashley discovered that some elements of the writer's individual style persisted in all examples, regardless of the end-effector used. The resulting handwriting is shown in Figure 4.11 taken from (Lashley, 1942).

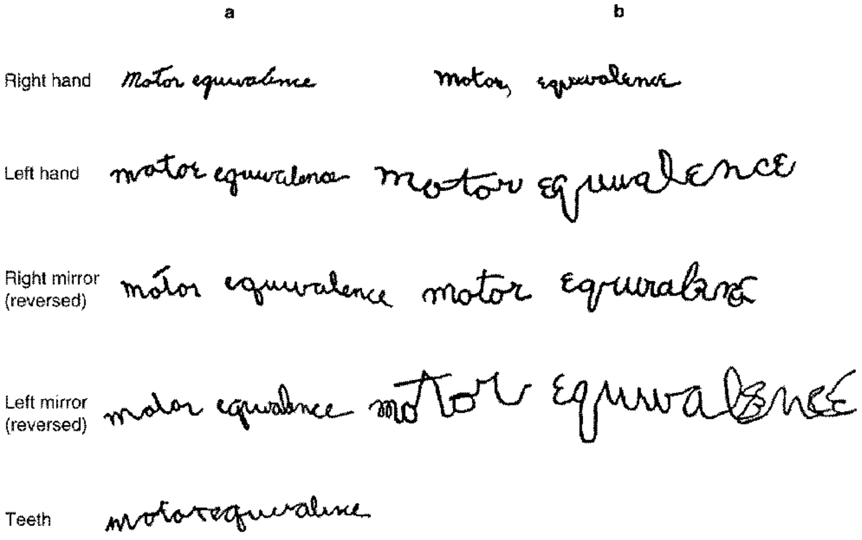


Fig. 4.11. Lashley's handwriting example. Columns A and B represent samples from two different subjects.

Schmidt theorized that those features which are invariant, that in some way are fundamental to these written words, are structured in the motor program while those aspects of the movement that are relatively superficial (speed, effector used) are then parameters of the program. The constant duration of individual movements, independent of movement amplitude, which can be seen in Figure 4.10 might therefore be considered to be further evidence for the existence of invariant components in human motor programs, but now within the context of a task that has a timing structure imposed by an external entity, music. (See Chapter 3 for further discussion of motor control in music.) An interesting question then arises: Is that quality of musical performance which is characteristic of an individual performer partly a function of how programs for individual movements are represented in motor memory? If so, then clearly feedback which influences the building of the mental representation of an instrument's response is part of a player's model of the dynamics of that instrument.

4.2.4 Summary of Experiment I

In summary, the results of this experiment lead to the following conclusions:

1. The existence of force feedback in a computer-based instrument marginally improves performance of a simple musical task.
2. That the specific nature of the force feedback provided seems to have little effect on the size of this performance improvement.
3. The observation that movement duration was a product not of musical time constraints, but of movement constraints, provides further evidence to support Schmidt's theory of generalized motor programs, now within the context of music.

4.3 Experiment II: Adaptation

One question which arises from Experiment I is whether the presence of force feedback is providing feedback which is of long-term use in performance or whether players might treat it as an artifact of the interface which they will overcome given sufficient practice.

As Shadmehr *et al.* (1995) has shown, humans excel in their ability to adapt rapidly to the variable dynamics of their arm as their hand interacts with the environment. Given sufficient practice (approximately 700 discrete reaching movements) in a novel force field, Bhushan and Shadmehr demonstrated that subjects had internalized its dynamics; when force feedback was suddenly removed, trajectories of hand movements showed clearly that subjects were expecting to encounter certain forces at particular points and had "planned ahead" to account for these disturbances. Based on these findings, we conjectured that given sufficient practice, players would "learn" the mapping between the haptic and auditory response of the virtual Theremin and learn to compensate for any feedback which was impeding their performance. We hypothesized that, if players were given long enough to play in one force condition, they would adapt to it and the differences between the spring and viscous damper feedback conditions would be greatly reduced or disappear altogether.

4.3.1 Experiment

Results of Experiment I showed that different force conditions had differing effects on the tragectories of melodies performed and that these effects changed for individual performers as force condition was varied. Experiment II was designed to see whether any improvements in performance due to the presence of force feedback would be amplified or attenuated if players were exposed to a single force condition for an extended period of time.

Participants

15 volunteers, none of whom had taken part in experiment I, were recruited from the faculty and student population at Stanford's Center for Computer Research in Music and Acoustics (CCRMA.) As with experiment I, all were experienced musicians and none had any known hearing or motor impairments.

Apparatus

The apparatus was the same as that used in experiment I.

Stimuli

The principal difference between experiments I and II was in experimental design. In experiment I all participants played 12 melodies, two melodies in each of six force conditions. In this experiment, participants were randomly assigned to one of three experimental groups — positive spring feedback, viscous damper or no force feedback. All participants played all 18 melodies from experiment I. Melodies were presented in random order with one exception: Melody 18 was used as the training melody and was presented again at the very end of the experiment. Finally, a 19th trial was added in which, without participants' prior knowledge, all force feedback was removed and players recorded one more performance of the training melody.

Procedure

As with experiment I, participants were given an initial period of time to get used to the device. A performance of the control melody (melody 18) was then recorded and represented base-level performance in the force feedback condition to which the participant was assigned. Participants then completed 18 further trials. For each trial, a first attempt at the melody was recorded; a maximum of four practice trials were then permitted before a final recording of the melody was made.

4.3.2 Results

Convergence of Force Conditions

In order to estimate "playing accuracy," we again scored each performance against a computer-generated template for the appropriate melody. Each score represented the RMS error in Hertz across the whole melody for that performance. For each melody, two recordings were made and the lower of the two scores was kept as the score for that trial.

We first measured the performance of the control melody when it was played at the start of the experiment and compared it with performance of the same melody recorded at the end of the experiment just before force feedback was removed. A comparison of the mean scores for first and last melodies between the three force conditions revealed no significant difference in terms of improvement in performance. We therefore concluded that players in all three groups had learned to compensate for the dynamics of the particular force condition to which they had been asigned and it was having no long term effect on their performance.

But how quickly did players adapt to each force condition? To gauge the slope of adaptation over the duration of the experiment for each force condition, we divided each participant's 18 trials into 6 groups of 3 trials. We averaged each group of three trials for each subject individually, to obtain a graph for performance over time and plotted the mean scores for each group of trials by force condition to see if there was a difference in the rate of adaptation between conditions. Figure 4.12 shows means scores over time for all three force conditions.

From the plot in Figure 4.12 it can be observed that curves for both force feedback conditions (spring and damper) are very different from those for the no-feedback condition. The curves for the two force conditions drop steeply at their start and rise to a peak in trials 13-15. This peak represents a point approximately 3/4 through the hour-long experiment where participants began to tire. (The knowledge that there were only 18 melodies may have caused an improvement in performance for the last few trials.)

4.3.3 Discussion

As with experiment I, the principal finding in this study was that the presence of force feedback, regardless of its specific nature, produced some improvement in performance. The plots in Figure 4.12 show that performance in spring and damper conditions improved significantly after the first three trials.

In terms of the time line of the experiment, this period of fairly marked improvement

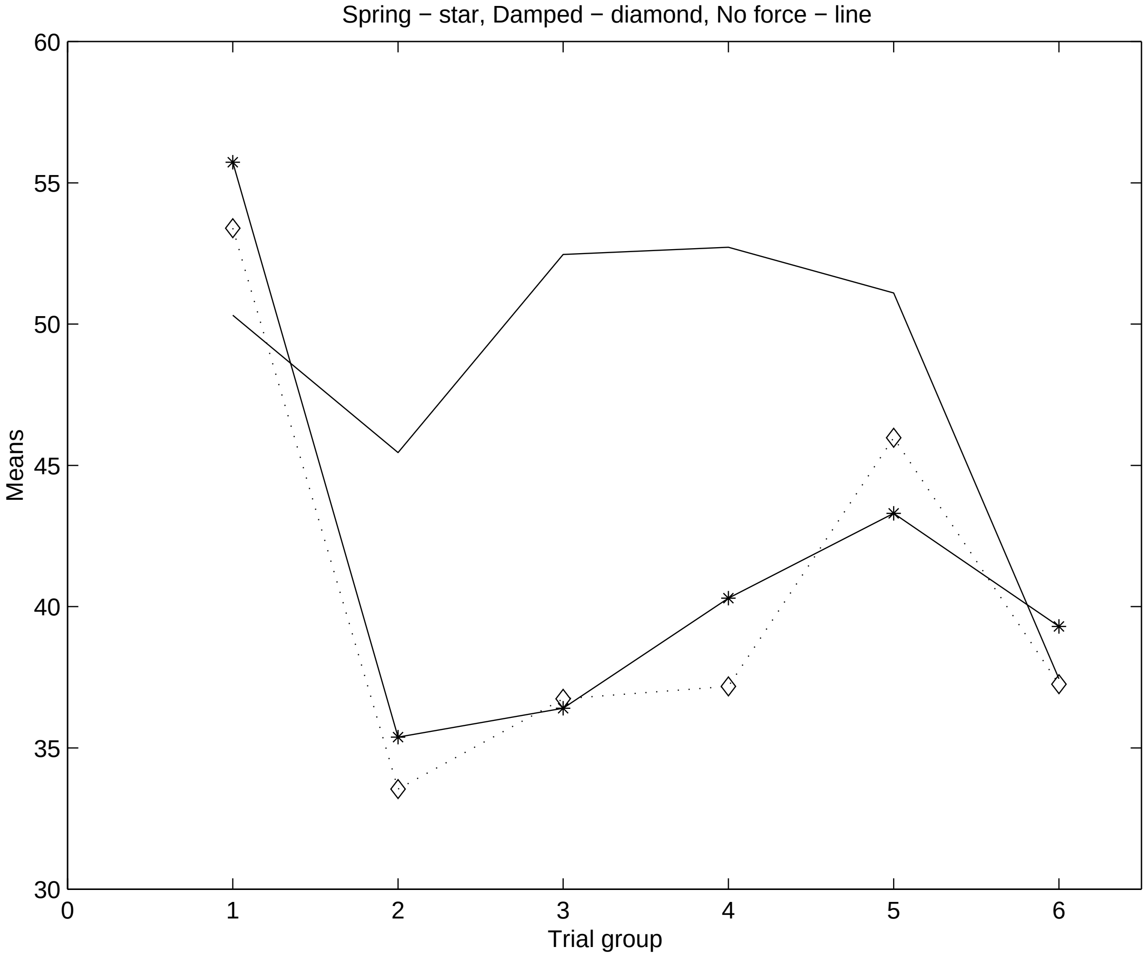


Fig. 4.12. Mean scores for each group of trials by force condition. Diamonds indicate damped force, "*" indicates spring and "-" represents the no force condition.

represents a period about 10 minutes into the 1-hour experiment. By this point, participants had figured out the demands of the task and had settled into the routine of each trial. In terms of the stages of motor acquisition defined by Fitts (1964), this would appear to correspond to the start of the second stage where strategies for achieving the task goal have been selected and are gradually being refined.

In determining the overall duration of the study, we were guided by Shadmehr's observation that it took approximately 750 discrete movements in a novel force field to cause participants to anticipate the dynamics of the force field in making simple reaching movements (Shadmehr *et al.*, 1995). By this reasoning, we calculated that participants would need to play eighteen melodies containing approximately twelve inter-note movements each, and would need to practice each melody four times (for a total of 860 individual movements) to approach Shadmer's estimate. However, we underestimated the extra cognitive load that would be imposed by performing simple melodies as opposed to making reaching movements. Participants tired sooner than our informal pilot study had indicated they might. It is likely that shortening the experiment would have produced a more marked difference between performances of the control melody at the start and end of the experiment reflecting the difference in scores for no-feedback and feedback conditions observed in trials seven to twelve (see Figure 4.12.)

Lastly, to be certain we had forced players to internalize the dynamics of the force condition in which they played, we removed all force feedback at the very end of the experiment and recorded an encore performance of the training melody.

Figure 4.13 is a plot showing the template for the training melody (the solid line), the final performance in the assigned force condition (the solid thin line) and the encore performance once all force feedback had been removed (the dotted line). The force condition assigned to this player was the viscous damper. It can be clearly seen that, when force feedback was removed, large leaps were over-estimated while small intervals remained more or less the same.

The results of this study lead us to conclude that, like Shadmehr, we had caused players to build an internal model of the dynamics of the instrument, the inverse of which became apparent once force feedback was removed.

4.4 Summary and Conclusions

The present studies have provided evidence to support the hypothesis that adding haptic feedback to computer-based musical instruments can improve a player's ability to control these instruments. However, they have also shown that the specific nature of the force feedback has little effect on performance. Given practice, it seems that

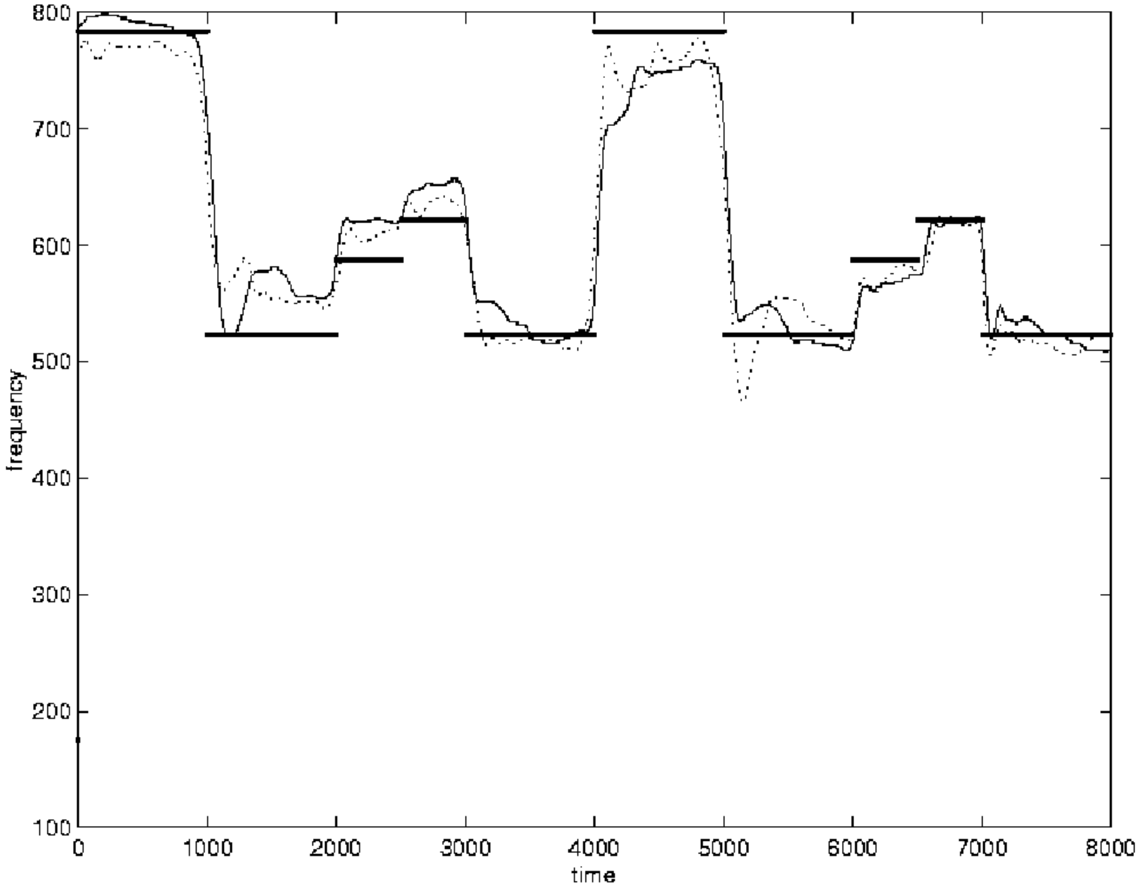


Fig. 4.13. Overlay of training melody with actual performances with feedback (solid) and no feedback (dashed).

the presence of force feedback continues to produce small performance gains over a condition where no force feedback is available.

These results, obtained within the context of a musical performance on a virtual Theremin, indicate that the presence of force feedback can accelerate the initial phase of modeling the behavior of an instrument.

However, most musical instruments require a player to control many parameters simultaneously (pitch, amplitude, timbre, etc.). Moreover, the mapping between the player's actions and the instrument's response to a given performance gesture is seldom linear. In our next sequence of experiments (Chapter 5) we created and tested a more complex virtual instrument, a bowed string.