Perception of Sound Renderings via Vibrotactile Feedback

Ricardo Pedrosa[1] Karon E. MacLean[1]

Department of Computer Science, The University of British Columbia, Vancouver, Canada

ABSTRACT

The vibration behavior of acoustic music instruments, as perceived through touch, is known to play an important role in the interaction between performers and music instruments. This research explores the relevance and utility of including a tactile simulation of this behavior in computer music interfaces, in a snapshot taken during early learning, and including tradeoffs of using this channel for other information. Vibrotactile (VT) renderings of different types of sounds were presented to subjects via a handle fitted with vibration actuators. Subjects (1) evaluated consistency of the sound with the VT rendering, (2) identified from a "lineup" the VT rendering that did *not* match a sound, and (3) provided data on how inclusion of VT cues in a rhythm-tapping task affected perception quality of the VT rendering. For the last, a distractor task allowed us to measure both the usability of the cues and the degradation in perception of the VT rendering. All renderings were either played directly through vibration transducers, or first altered by modifying the sound's frequency content. VT playback was delivered at intensities similar to those experienced in traditional acoustic music instruments.

Subjects indicated that specific alterations to the original, direct VT rendering were consistent with the source sound more often than the original was, and were able to differentiate between correct and incorrect renderings of a source sound. However, the latter ability was masked when subjects were provided with and able to effectively utilize extra VT cues added to the VT feedback to improve their performance, suggesting that the VT cues were of greater utility than the VT mimicking. We discuss the relevance of these findings on the design of computer music interfaces.

KEYWORDS: vibrotactile feedback, perception, music computer interfaces.

1 MOTIVATION AND RELATED WORK

The interaction between performer and musical instrument is complex, involving multiple sensory as well as motor channels, among which the haptic channel (used here to encompass proprioceptive and tactile) is crucial. Any music instrument responds to a performer's action by producing a sound. For traditional acoustic instruments, performer actions directly modify the behavior of a vibrating element. The diversity of mechanical interactions used in acoustic music instruments (hitting, blowing, plucking, strumming, bowing, stepping, etc.), and of musical instruments themselves share one common factor: the instrument body's amplification of the sound emanating from the active element causes the body itself to vibrate. A portion of those vibrations reach the performer; others are dissipated in interaction with the instrument, or attenuated by the skin's limited vibration perception range and the signal's low vibratory intensity [17].

Nevertheless, the vibrotactile (VT) feedback as sensed by the performer plays an important role in supporting performance and enhancing the interaction [4], for example during performance for stringed instruments. Percussive instruments, on the other hand are known to rely more on proprioceptive feedback [1].

In computer based music instruments, especially those whose interaction method does not mimic a particular traditional instrument, haptic feedback available to the performer is not mechanically coupled to sound generation. Any dependence of touch feedback on the instrument's state as a sound generator must be explicitly designed and displayed. Marshall and Wanderley described several devices that can be used to provide VT feedback in digital music instruments and presented various instruments that provide the feeling of an acoustic instrument by integrating sound production into the instrument, generating VT feedback which is directly related to the sound produced [10].

Prior work has explored the incorporation of haptic feedback into computer music interfaces, from translation of traditional instruments to virtual implementation [3][6][10] to new interfaces and sound-generation techniques [2][5]. However, most of these studies aim to improve performance by enhancing controllability, either in traditional acoustic [7] or new [15] instruments.

How important is it to reproduce the VT "feel" of traditional acoustic instruments (i.e. a feel consistent with the sound generated) in new computer music interfaces? This channel could also be used for other performance-enhancing information unavailable in traditional acoustic instruments [15], and thus there may be a tradeoff. Perceptual saturation or overload is likely to degrade the interaction experience.

In order to answer this, we first ask how to best design VT feedback to improve the instrument-playing experience: by simulating vibrations as they naturally occur in an acoustic instrument; or through usable inclusion of performance-enhancing cues, as part of the definition of that instrument's interface?

This study is part of a larger work in which we are assessing the inclusion of a broader range of haptic feedback in a gesture-based computer music interfaces [12][13]. Here, we address two specific research questions:

1. Regarding reproducing the vibratory behavior of traditional music instruments ("**VT mimicking**"):
    a. Is there a perceivable difference between a VT signal containing the full spectrum of a sound and one filtered to the frequency bandwidth of human vibratory perception?
    b. Is tactile recognition of a sound enhanced through further modifying (e.g. spectral shaping) the original sound?
    c. Does the *type* of sound (vocal, percussive, multi-instrument) influence preference of VT rendering method?
2. Regarding inclusion of extra "**VT cues**" (for rhythm, dynamics, multi-instrument coordination, etc.):
    a. When attention is focused on a "distractor" task reliant on VT cues, does sensitivity to a change in the VT signal mimicking the sound diminish?

---

[1] {rpedrosa,maclean}@cs.ubc.ca

IEEE World Haptics Conference 2011
21-24 June, Istanbul, Turkey
978-1-4577-0298-3/11/$26.00 ©2011 Crown

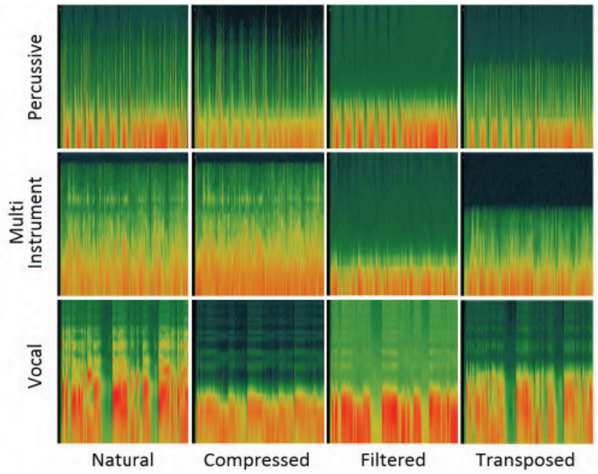


Figure 1. Spectrograms of some of the signals used in the study.

We designed a 3-part study to explore these questions. We describe the signals used, overview the study's overall structure, detail each sub-study then present and discuss their results.

2 VIBROTACTILE SIGNALS AND MUSICAL EXCERPTS

2.1 VT Rendering – "Mimicking"

Signal Frequency Content: We used two types of VT signals (Figure 1). ***"Natural"*** renderings directly reproduced the sound through the VT transducers. "***Altered***" renderings were spectrally modified from the original audio signal, through one of:

a. ***Filtering***: Original signal passed through a low-pass filter with a cut-off frequency of 1 kHz (perceptual limit of vibrations for the nonglabrous skin [17]).
b. ***Transposing*** (frequency scaling): Sound spectrum re-allocated to the skin's perceived range, maintaining the original relationship between frequency components.
c. ***Compression*** (as known in audio processing): Signal dynamic range is reduced by narrowing the difference between high and low audio levels. Here, we sought to reduce attack and release times of signal components, forcing transducers to react more rapidly and thereby transmit more energy and increase perceived loudness of original signal. Unlike filtering, this method increases a signal's high frequency components.

Signal Amplitude: A major risk of vibration rendering is numbness when the renderings are too strong, especially for extended durations. Vibration amplitudes measured in stringed instruments have been observed up to 20dB above the perceivable skin range [1]. We normalized vibrations to within this value.

To optimize amplitude level, in a pilot test we rendered a vocal music excerpt (the most difficult to perceive of those we used) through the vibration display (Section 3.1). The rendering's amplitude was reduced trial by trial and 5 subjects reported when they stopped feeling the vibrations, for three repetitions. The average value of their responses was used as the lower limit for perception with this display, a conservative approach. A value 20dB greater was used to normalize amplitude for the rest of the music excerpts renderings used here.

Incorrect Rendering*:* For some study tasks, we produced "incorrect" renderings of sound excerpts by mixing (superposing) the relevant excerpt with a portion of another excerpt of the same type, e.g. vocal or percussive. The location in the mix was also chosen such that the VT representation of these elements was noticeable. In the particular case of the final task, the additional elements were mixed into the source track with a gain of 6dB above the average level of the sound used as a base.

2.2 Performance Task and VT Cues

One of the tasks required adding timing cues to the VT signal. Cues were created by modulating a 250Hz sine wave with a pulse 60ms wide. The resulting short burst was triggered during the track, to help subjects tap in synchrony with a rhythm.

We note that our study's aims do not include characterizing VT cues or their impact on performance, e.g. relative to traditional cues used in teaching musical skills. Instead, we use the cues and their performance metrics as a workload instrument to study how perception of VT mimicking signals is degraded by distraction, as described further in Sections 3.1 and 3.4.

2.3 Sound Sources: Musical Excerpts

The VT renderings described above were sourced from nine auditory music excerpts: three each of only percussive instruments, only voice, and various instruments and voice. This selection covers as much as possible of the possible auditory variety in dynamics and frequency available from a music instrument. The excerpts ranged in length from 13 to 19 seconds, to accommodate different sound types. Vocal tracks contained voices of males, females, children and choral ensembles; percussion tracks contained various solo instruments, from hand drums to drumsets; multi instrument tracks contained samples from music genres of classical, Latin, rock and fusion.

3 METHODS

3.1 Study Overview

Our study had two sections. The first tested how different VT renderings of the same sound were perceived, and included two tasks; specifically, (Task 1) how accurately a particular vibration rendering represents a sound and (Task 2) how each VT rendering method supports the identification of a vibration signal that does not match a sound. The second section was designed to test the effect of additional cues on perception of a "primary" vibration rendering the sound (Task 3). The effect of the cues on performance accuracy was also measured, to understand workload distribution between the primary task (here, perception of VT mimicking signal) and the "distractor" task (for our present purpose, tapping out a rhythm, in some cases aided by VT cues).

For the two tasks in the first section we used all nine musical excerpts as described in 2.1. Through pilots we found that despite optimization of amplitude between tracks, some types of music were perceived as stronger than others (i.e. vocal excerpts felt softer than percussive tracks). This prompted us to experimentally block presentation of sounds and corresponding VT renderings by sound type. Specifically, for each excerpt subjects simultaneously heard the sound and felt a vibration signal rendered from the sound, once for each of the four renderings methods described

Figure 2. The handle designed to contain the VT transducers. From left to right: Handle opened to show the actuators, lateral view, rear view and grip used for last task in the study.

above. Order of sound type and order of rendering within sound type were randomized between subjects.

Task 3 was to test whether including VT cues supporting a particular music task diminished perceivability of VT mimicking; and also if those cues effectively aided performance at a primary music-related task. We used the primary music-related task as a distraction by informing the subject we were testing performance accuracy instead of how the VT was perceived.

3.2 Apparatus

The vibration display was a handle designed to be attached to a Sensable Technologies Phantom Premium and containing 2 vibration actuators (Tactaid [16]; Figure 2). The Tactaid is a high Q resonant type of vibrator with a peak response around 250 Hz, well matched to the human detection threshold curve which is essentially a bandpass filter around 200-300 Hz. While not used here, the Phantom is employed in other studies in this series for which we require a consistent protocol. Subjects were instructed to use the same grip as would be natural if it were attached to the Phantom. The VT rendering is generated in a computer and sent via audio port to the handle actuators. A MIDI interface (Korg padKontrol [9]) was used to record the user tapping performance during the last task of the study. The software components controlling the reproduction of the sound samples, the generation of the VT renderings and the recording of the subject's performance were developed using Pure Data [14].

3.3 Participants

Eighteen participants, recruited from the university campus population through ads posted in public locations (11 female and 7 male, five left handed, aged 19 to 39) took part in the study. Five participants were skilled music players, 9 reported ability to play some music instrument without particular skill, and 6 reported no music training or ability to play music instruments. From the participants with music skills, only one self-reported ability to play a traditional drumset; the rest played melodic instruments. Two participants (both male) had a previous exposure to the rhythms used in the second section of the study. Participants were compensated with $10 for 1 hour of time.

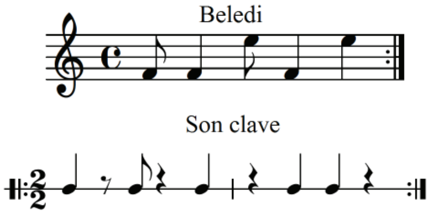
3.4 Procedures and Metrics

Task 1: Consistency between VT Renderings and Sound

We asked subjects to rate consistency of the sound they were hearing and the VT rendering they simultaneously felt using a Likert scale (1: *not consistent* to 10: *highly consistent*). After all four renderings for an excerpt were presented, subjects were asked to choose which one rendering they liked most before continuing to the next excerpt. This process was repeated for every excerpt, which were blocked on sound type. Subjects were asked to rate with a "0" those VT signals they couldn't properly feel.

Task 2: Identification of the Wrong VT Rendering

For each VT rendering, we asked subjects to identify which single rendering out of 4 presented for that sound had "elements that are not present in the sound". The sound/vibration presentation was the same as in the previous task, including blocking scheme: subjects heard a sound and felt a vibration and were asked to respond as above. For each rendering felt, subjects were then asked to report how confident they were in their selection of the wrong rendering on a Likert scale (1: *not confident* to 5: *very confident*). Subjects were again asked to signal with a "0" any particular rendering they could not feel and no other evaluation was requested in this case.


Beledi

Son clave

Figure 3. Western musical notations of rhythms used in Task 3.

Task 3: Effect of VT Cues on Ability to Identify VT Mimicking Nuances

Subjects were asked to perform 6 trials consisting of tapping with the handle on a drum MIDI interface (physical pad) following percussion rhythms, under various VT feedback conditions. They did this twice, for an easier and a harder rhythm. Subjects could hear the acoustic sound of the handle hitting the drum pads, but no percussion sounds were synthesized.

Pilot testing revealed that the percussive sounds were well represented by their *natural* VT renderings, so we used the natural rendering for the VT Mimicking signal.

Subjects were informed that we wanted to determine how some cues in the VT feedback improve their performance at tapping out different rhythms. However, unlike Task 2, subjects were *not* informed that in some cases the VT Mimicking signal was modified (Section 2.1).

During this task, subjects were asked to perform the following six trials in sequence for each rhythm:

1st Trial: An 8 bar audio track with a soloist playing a rhythm was presented. Rendering reproduced sound always. Subjects were asked to tap along the soloist following the rhythm.

2nd Trial: The rhythm was heard only during the first and last two bars of an 8 bar track, with 4 bars of silence in between. VT rendering reproduced sound when present. Subjects were asked to tap through silence following the rhythm heard initially and continue playing to end of track.

3rd Trial: The rhythm was heard as in previous trial. VT rendering reproduced sound when present and included a strong pulse every time subject was supposed to tap, as a rhythm cue. This cue was available during the silence gap. Subjects were asked to tap with the rhythm using the cues as guidance if needed.

4th Trial: An audio track with an ensemble playing an excerpt of a piece was presented. VT rendering reproduced the sound in the audio track. One musician in the ensemble played the rhythm. Subjects were asked to play along with this musician. Subjects repeated the performance of this track twice (Trial 4a and 4b).

5th Trial: The audio presentation is the same as the previous task. The VT signal included a cue for the rhythm the subject should play, and was also modified by including elements not present in the original sound. After the track ended, subjects were asked if they felt that the VT mimicking was inconsistent with the sound they were hearing.

6th Trial: The same audio track and VT as in Trial 5 was presented. Subjects were asked to do the same as before (tap following his part in the ensemble) but also to pay attention to the VT mimicking in order to identify the presence or not of elements extraneous to the sound.

As an introduction and training for this task, subjects were first presented with the sound of a metronome running a 4/4 beat at 120 bpm, and executed the first three trials.

Subjects then were presented with two rhythms of increasing complexity (Figure 3), and asked to carry out Trials 1-6 in the order above. First they faced a Middle Eastern drum rhythm known as *Beledi* or *Masmoudi Saghir*, commonly used in oriental dancing. This rhythm contains two different sounds or pitches, but subjects were asked to ignore the difference in the sound and to

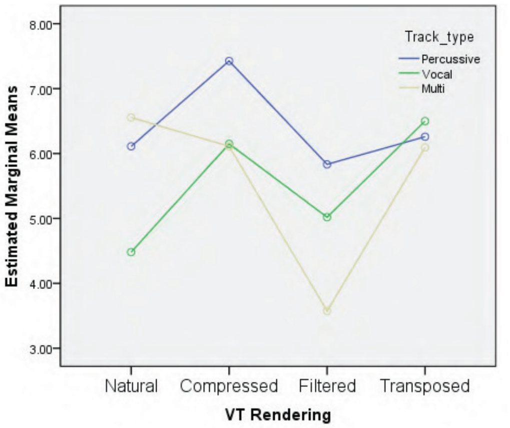


Figure 4. Estimated means for consistency ratings of VT rendering type and sound. A high value indicates consistency (Task 1).

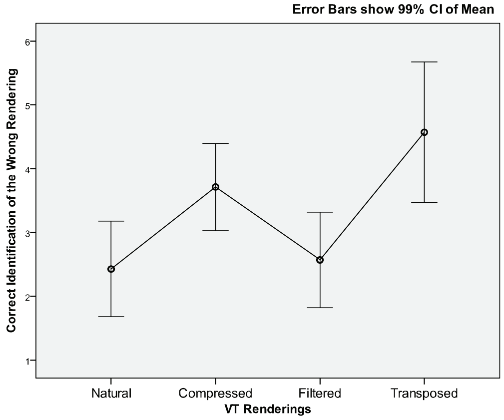


Figure 5. Estimated means for correct identification of wrong VT rendering (Task 2)

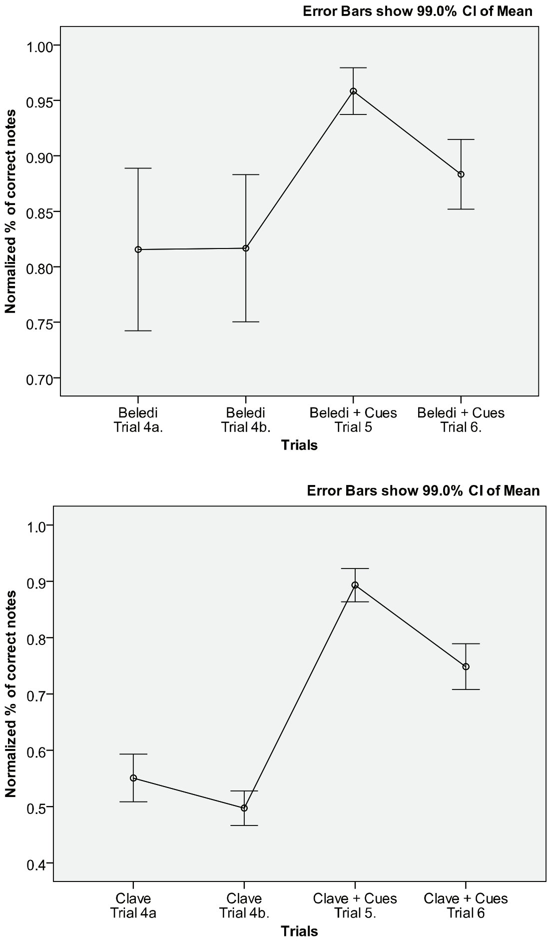


Figure 6. Estimated means of tapping performance for (top) Middle Eastern and (bottom) Cuban rhythms (Task 3)

concentrate on timing. Then, subjects were presented with a more challenging Afro-Cuban rhythm known as the *Son Clave*, difficult to follow especially for those not from a Latin culture.

Data from Trials 1, 2 and 3 were not analyzed, but used as further training for each rhythm. By Trial 4, subjects had heard and played the rhythm 3 times. Trials 4 to 6 are considerably more complex than the previous three. Subjects had to identify the rhythm in the ensemble and play along with it. Sometimes the sound from the instrument playing the rhythm is barely audible, so the subject had to rely on the rhythmic structure he or she has already played three times to maintain synchronization.

We are interested in early stages of learning, since a user must go through this for any new music interface; this protocol, through use of medium to hard rhythms which are unknown to most of the subjects, remained within this learning zone: six trials (each lasting <=20s) are far less than what is required to achieve a perfect performance for this largely novice population

For each subject, the inclusion of extraneous elements to the rhythm in Trials 5 and 6 only happened once, either in the Middle Eastern or in the Afro-Cuban performance. Nine subjects were presented with an incorrect rendering of the Beledi rhythm and nine with an incorrect rendering of the Clave rhythm. Tapping performance was recorded in every trial, and subjects were asked to identify presence/absence of an incorrect VT Mimicking signal.

The wrong renderings in Task 3 were relatively more salient than in Task 2 (Section 2.1). As the cues for rhythm were also superimposed on the VT Mimicking signal, we wanted to ensure that any other added element was noticeable as well, and only neglected if not necessary for fulfilment of the task.

Subject performances were scored against a computer-generated template of the rhythm. Subjects performances where quantized to eighth notes, the shortest note used in the reference rhythms; i.e. a note was correct if it occurred at the same time as one in the pattern. This eliminates small timing nuances that could be considered as fitting within the target rhythmic structure. Notes executed on time were counted and the percentage of correct notes was calculated with respect to the total number of note occurrences per track. Percentage values were normalized to a value between 0 and 1, where 1 corresponds to all notes correct.

4 Results

Task 1: Consistency between VT Rendering and Sound

We analyzed subjects' rating of VT rendering and sound consistency using a two-way repeated-measures ANOVA, with independent variables of sound and rendering type. Mauchly's test indicated that the sphericity had been violated for the interaction effect between the music and rendering type ($\chi^2(20)=49.418, p<.001$); therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon=0.512$).

Rendering method had a significant effect on reported consistency of sound and vibration ($F(3,51)=22.84, p<.001$) and there was a significant interaction effect between the music type and the rendering method ($F(3.072,52.23)=9.90, p<.001$) (Figure

Table 1: Identification of the wrong VT rendering after performing a rhythm along an ensemble (Task 3).

Rhythm	Number of participants recognizing the Wrong VT rendering	
	After 5th Trial	After 6th Trial
Beledi	2 (22.2%)	8 (88.9%)
Clave	1 (11.1%)	6 (66.7%)

4). There was no significant difference in the preference of one rendering over the other.

Contrast analysis revealed that the *transposed* rendering had significantly higher consistency with the sound when compared to the *natural* (*F(1,17)=12.39, p=.003*) and *filtered* renderings (*F(1,17)=28.32, p<.001*). Regarding interactions, for percussive and multi instruments music excerpts, significant differences were found between the perception of the *compressed* and *transposed* renderings (*F(1,17)=20.94, p<.001*) and between the *filtered* and *transposed* renderings (*F(1,17)=9.016, p=.008.* For vocal and multi instruments music excerpts, a significant consistency difference exists in *natural* and *transposed* renderings (*F(1,17)=53.35, p<.001*).

Task 2: Identification of the Wrong VT rendering

We used a one-way ANOVA to compare the means of correctly identified "wrong" renderings for each VT rendering. Due to experiment size constraints, only 1 rendering per audio track (9 throughout the Task) was rendered incorrectly (Section 2.1).

There was a significant effect of rendering method on the identification of the VT renderings that did not match the sound (*F(3,24)=20.1, p<.001*) (Figure 5). Planned contrasts revealed that the *transposed* rendering significantly improved detection of the wrong rendering over the *natural* (*t(24)=4.564, p<.001*) and *filtered* renderings (*t(24)=6.261, p<.001*).

Task 3: Effect of VT Cues on Ability to Identify VT Mimicking Nuances

We applied a one-way ANOVA to compare the means of the percentage of correct notes (percussion hits) played in time for each one of the rhythms independently, for Trials 4, 5 and 6, since trials 1 to 3 were considered training.

Inclusion of VT cues significantly improved the tapping execution of both the Middle Eastern ("easier") and Cuban ("harder") rhythms (*F(3,51)=36.245, p<.001* and *F(3,51)=400.537, p<.001*); Figure 6. Two-tailed t-tests with corrections for multiple comparisons also show that for Trial 5 (cues presented) tapping performance was significantly better than in Trial 6 (cues presented and subjects also assessing whether VT signal contained elements extraneous to the sound for both rhythms (Middle Eastern: *t(17)= 10.694, p<.001* and Cuban: *t(17)= 9.723, p<.001*).

5 DISCUSSION

Improving VT Consistency through Signal Reshaping:

Spectral reshaping can improve the perceptual consistency of a vibrotactile rendering to its source sound over direct reproduction, as demonstrated by the results of Task 1 and 2.

In Figure 4, we see that for percussive and vocal excerpts, the perceived sound consistency of the VT rendering is similar for the *natural* and *filtered* renderings, while the multi-instrument excerpts dive for the *filtered* case. The latter may be explained by spectral content of the original signal (Figure 1): vocal and percussive sounds have reduced frequency content relative to multi-instrument sounds, and thus these excerpts retained more of their content when filtered.

Something very interesting happened with the *compressed* rendering. This method's primary characteristic is to decrease the dynamic range of a sound signal by emphasizing the quieter sounds. The main consequence of this process is the inclusion of high frequency components (>20kHz), falling outside both the auditory and tactile perception range. Compared with the transposition method, no drastic (if any) spectral modification occurred in the tactile perception range.

However, the trend in Figure 4 shows greatest impact for VT Mimicking using this rendering method (the multi-instrument signal consistency actually drops slightly from the *natural* rendering). This might be attributable to perception of the energy transmitted by the vibrations. While human skin perception is limited to vibrations of around 1kHz, humans are able to notice when a higher frequency signal is suddenly turned off or on; this is thought to be due to the skin acting as a low pass filter [8]. Compressing the signal as we did here decreases the attack times of the sound components. In other words, slowly rising levels are sped up, increasing the energy carried by the signal. Thus a compressed audio signal is *heard* more loudly (energy range is compressed into the audible range), but without significant changes to the perceived frequency content. Conversely, the skin's LPF quality allows it to detect those highly sped-up components. While not characteristic of the source signal, these components appear synchronized in time with sound events, and this might explain why our *compressed* VT signals were perceived as more consistent with the sound than a *natural* (for lower-frequency original content) or *filtered* signal.

Transposing the signal into the skin's perceptual range produced the most consistent results across the full range of sounds we examined. It was also the only one to significantly improve probability of detecting a discrepancy between the sound and the VT Mimicking (Figure 5). Both results validate transposition as the preferred method to use if we were to design a display to convey a broad range of sounds through vibrations.

Improving Rhythm Following Performance

Task 3 results show that tapping performance improved when additional cues were added, for both rhythms used (Figure 6). The first two data points came from repetition of reference Trial 4 (playing the rhythm in an ensemble with no extra cues). During these trials, subjects consistently gave an average (relative to range seen for all Task 3 trials) performance of the Beledi rhythm and a rather poor one of the Son Clave. These levels confirmed our premise that for this population, these two rhythms would have medium to hard difficulty levels, with the Afro-Cuban rhythm the more difficult to learn.

Tapping performance significantly increased when timing cues for note occurrence were introduced (Trial 5). This indicates that the cues did provide assistance and were correctly interpreted. After finishing this trial, subjects were asked if they had detected VT Mimicking that was no longer consistent with the sound. Table 1 show the number of participants (out of 9 for each rhythm) that correctly identified the wrong rendering when it occurred, after finishing the 5th and 6th trials. Two subjects recognized an inconsistency with the Beledi rendering, and one for the Clave after the 5th Trial. This implies that these changes went unnoticed when subjects were focused trying to perform as best as they could; despite the slight amplification of the modified signal component over the base and unmodified signals.

During Trial 6, subjects were asked to attend to the VT rendering while playing the rhythm to determine if the VT Mimicking was right or not, but told that the main task was still to play the rhythm in time. Figure 6 and Table 1 together show that Trial 6 tapping performance significantly decreased but far more subjects were able to identify occurrence of a wrong rendering, indicating that attending to the VT rendering was indeed distracting for the tapping task. This also shows that the cues inserted in the VT feedback did not completely mask irregularities added to the VT Mimicking, as they were still noticeable by at least some subjects. Also, the average performances were not 100% perfect, a sign that subjects were still learning the rhythms.

Some subjects reported that the presence of the cues were helpful not only for playing but also for understanding the rhythm. One subject reported that for the clave rhythm, when the cues were present: *"I could hear the rest of the players building the piece around the rhythm I was playing"*.

6 CONCLUSIONS

In this study we used a set of cues that effectively improved performance. Unlike the auditory cues typically used for this purpose (metronomes, clapping, etc), the VT cues did not obscure the musical audio signal that subjects were asked to follow.

This study took place during the early learning stage, and its principal goal was to assess perception of vibrotactile feedback accompanying sounds and assessment of its consistency to the sounds. Performance cues were used as a tool to study this question; and while it is of interest that the cues improved performance compared to no cues, our objective was not the design or evaluation of those cues. Further testing is necessary to assess different ways of presenting these cues and to understand the role of performing-enhancing cues in defining an interface as a music instrument.

The results of the first two tasks (Figures 4-5) demonstrate that subjects are able to differentiate between a correct and an incorrect VT rendering of a particular sound and that there are better ways to render the sound through a VT signal than simply reproducing it through transducers.

Results from the third task show that in this early stage of learning, the vibration renderings which conveyed a particular status of the instrument (as told by the instrument's own sound-generated vibrations) pass unnoticed when extra cues for important performance parameters are introduced and effectively used by the performer. Trial 5 improves sharply over both runs of the 4th Trial, and meanwhile ability to recognize wrong VT mimickings drops.

With acoustic music instruments, a similar process occurs. During early learning stages, novice players tend to put little effort into choosing a specific instrument; given an adequate model, they are too busy developing motor and playing skills to attend to instrument nuances such as how it vibrates. This situation changes as performers gain experience, and the movements and instrument manipulations necessary to generate a particular sound become second nature. At this point, musicians start spending more time at music stores trying to find a particular instrument that simply *"feels right"*. In this sense we can see the knowledge about the instrument behavior as felt through VT feedback *"growing"* as the musician spends more time with the instrument.

For computer-based music interfaces these results are particularly relevant, as this space can be used to provide meaningful cues to accelerate and improve the learning process without overloading the perception of the instrument itself. Such cues will be extraneous for skilled musicians.

For those music interfaces that radically depart from traditional acoustic instruments, these results can lead to new guidelines. For instance, we can hypothesize that cues properly designed as a learning aid and displayed in synchrony with some music parameters could become a part of the perception of the interface as a music instrument. The existence of this "signature" must be examined through longitudinal studies that track extensive use of such interfaces.

Acknowledgments. We acknowledge advice from Drs. Hamel, Kingstone and van de Panne as well as UBC's SPIN Research Group. The long sessions of discussions with Dr. Hamel were particularly instrumental to our study design. Funded by Canada's NSERC Discovery Grant Program; studies carried out under UBC Ethics approval H01-80470.

REFERENCES

[1] Askenfelt, A., Jansson, E.V. On vibration sensation and finger touch in stringed instrument playing. Music Perception. Spring 1992, Vol. 9, No. 3. pp. 311-350.

[2] Bongers, B. The use of active tactile and force feedback in timbre controlling electronic instruments. In Proceedings of the 1994 International Computer Music Conference, pp. 171-174, 1994.

[3] Cadoz, C. Lisowski, L. Florens, JL. A Modular Feedback Keyboard Design. Computer Music Journal, Vol. 14, No. 2, New Performance Interfaces 2. (Summer, 1990), pp.47-51.

[4] Chafe, C. Tactile Audio Feedback Proc. Intl. Computer Music Conf. Tokyo, Sept. 1993

[5] Essl, G., O'Modhrain, S. Scrubber: an interface for friction-induced sounds. Proceedings of the 2005 conference on New Interfaces for Musical Expression. Vancouver, Canada. pp. 70 – 75. 2005.

[6] Gillespie, B. The Virtual Piano Action: Design and Implementation Proceedings of the International Computer Music Conference, Aahus, Denmark, Sept 12--17, 1994. pp. 167--170.

[7] Grosshauser, T., Hermann, T. Augmented Haptics – An Interactive Feedback System for Musicians. In Alinsoy, Jekosh and Brewster (eds.): HAID 2009, LNCS 5270, pp. 100-108, 2009.

[8] Klatzky, R.L.,Lederman, S. Touch. In Hanbook of Psychology: Experimental Psychology. Healy, Proctor and Healy (Eds.), Wiley, PP. 147-168, 2003.

[9] Korg padKontrol: MIDI Studio/Drum Controller. Product specification page. (http://www.korg.com/product.aspx?&pd=414).

[10] Marshall, M.T., Wanderley, M.M. Vibrotactile feedback in digital musical instruments In NIME '06 Proceedings of the 2006 conference on New interfaces for musical expression

[11] Nichols, C. The vBow: Development of a Virtual Violin Bow Haptic Human- Computer Interface. Proceedings of the 2002 Conference on New Interfaces for Musical Expression (NIME-02), pages 29-32, Dublin, Ireland, May 24-26 2002.

[12] Pedrosa, R., MacLean, K. E. (2009). Evaluation of 3D Haptic Target Rendering to Support Timing in Music Tasks. in Proceeding of The International Conference on New Interfaces for Musical Expression (NIME '09), Pittsburgh, Pennsylvania, USA, pages 19-24, 2009.

[13] Pedrosa, R., MacLean, K. Perceptually Informed Roles for Haptic Feedback in Expressive Music Controllers. In A. Pirhonen and S. Brewster (Eds.): HAID 2008, LNCS 5270, pp. 21–29, 2008.

[14] Pure Data: Real-time graphical programming environment for audio, video and graphical processing. http://puredata.info/

[15] Rovan, J. Hayward, V. Typology of Tactile Sounds and their Synthesis in Gesture-Driven Computer Music Performance. In Trends in Gestural Control of Music. Wanderley, M., Battier, M. (eds). Editions IRCAM, Paris, 2000.

[16] Tactaid and tactilator website (www.tactaid .com)

[17] Verillo, R.T. Vibration sensation in humans. Music Perception. Spring 1992, Vol. 9, No. 3. pp. 281-302