Tactile Teacher: Sensing Finger Tapping in Piano Playing

Chih-Pin Hsiao [1,3], Richard Li [1,2], Xinyan Yan [1,2], Ellen Yi-Luen Do [1,2,3,4]

[1] ACME Lab, GVU Center Georgia Institute of Technology
[2] College of Computing Georgia Institute of Technology
[3] College of Architecture Georgia Institute of Technology
[4] CUTE Center National University of Singapore

{chsiao9, lichard49, voidpointer, ellendo}@gatech.edu

ABSTRACT

In a piano lesson, a student often imitates the teacher's playing in terms of speed, dynamics, and fingering. While this learning model leverages one's visual and even audial perception for emulation, it still lacks an important component of piano playing – the tactile sensation. We seek to convey the tactile sensations of the teacher's keystrokes and then signal the student's corresponding fingers. We implemented an instrumented fingerless glove called Tactile Teacher to detect finger taps on hard surfaces. Since finger taps generate acoustic signals and cause vibrations, we embedded three vibration sensors on the glove and use machine learning algorithms to analyze the data from the sensors. After a brief training procedure, this prototype can accurately identify single finger tap in a very good performance at above 89% accuracy, and two finger taps resulted in accuracy around 85%.

Author Keywords

Sensing; Piano Learning; Tactile Sensations; Wearable Computing

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Human Factors; Design; Measurement.

INTRODUCTION

In a piano lesson, the teacher instructs the student by demonstrating the physical playing of the piece. The student observes by watching the teacher's hands, following along in the score, and listening to the minute details of the sounds produced by the teacher. Clark et al. describes these common representations when conveying motor skills as the common ground [3] in which novices (e.g. students) usually benefit from different sensory inputs produced by masters (e.g. teachers). For instance, in a real piano lesson, a teacher would tap his or her fingers on the palm or on the back of the hands of the students so they would experience the sensation of the finger taps from the teacher. This real life example inspired us to create Tactile Teacher to capture the teacher's fingering, and enhance the learning experience for the students. The result is an implementation that adds an additional medium – touch sensation – for assisting students in learning piano.

In this paper, we start by discussing the importance of sharing common representations for learning motor skills. Based on the experiences of other researchers, we explore the possible designs and implementations of Tactile Teacher, a device that can detect finger taps without covering or hindering the fingers. As shown in Figure 1-right, the fingerless glove employs three piezo vibration sensors mounted on the back of the user's hands. With a micro-controller and a machine-learning program on the host computer, we obtained accurate results when tapping one finger (89%) and two fingers simultaneously (85%). We conclude the paper with discussions about the potentials of this implementation in addition to enhancing learning in a piano lesson.

TEI '15, January 15–19, 2015, Stanford, California, USA.

http://dx.doi.org/10.1145/2677199.2680554

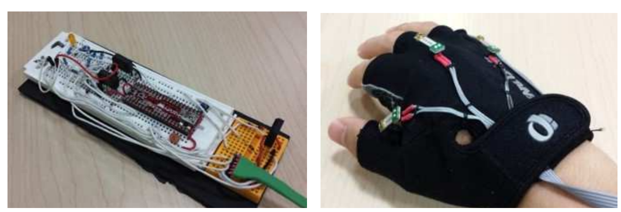

Figure 1. Tactile Teacher Prototype with a microcontroller (left) directly connected to an instrumented glove (right)

RELATED WORK

Sharing Tactile Sensations

Sharing common representations is crucial in learning, especially for motor skills [3]. Sharing haptic feedback for communication and learning is particularly interesting because touch sensation is invisible and difficult to convey through verbal descriptions. Thus, researchers have created prototypes that employ preset haptic stimuli to enhance the learning procedures of various motor skills, such as handwriting [14], memorizing force sequences [10], and remembering finger patterns [6,13].

Chellali et al. argues that real-time haptic communication will dramatically improve the learning curve for motor skills [2]. Typically, the haptic communication channel needs to be coupled with verbal or visual channels to maximize the effects [1]. It would be beneficial to engage in a multi-modal experience to enhance the learning of motor skills, especially with the experience of haptic communication. However, very few projects (e.g. [11]) discuss the types and means of sharing tactile representations in real-time learning scenarios, especially for those applications that require immediate responses.

Haptic Feedback for Guidance and Learning

Several researchers employ haptic or tactile guidance for providing additional information and improving the usability when using multi-touch devices, e.g. [8,15]. However, our task of assisting the piano learning process also requires the sensing of speed, dynamics, and fingerings on the originating hand and then rendering these sensations on the receiving hand. Mobile Music Touch (MMT) [6] employs various vibration motors to convey the tactile sensations for piano teaching purposes. The MMT glove addressed the rendering side of our task, and demonstrated the passive learning effects. However, the data used for vibration output is hardcoded in the program and lacks the spontaneity as well as the artistic aspects of timing and the force of hitting the piano keys that are demonstrated by the pianist (e.g., a teacher or a virtuoso). Therefore, in addition to building a piano glove that renders the finger tapping for students, we designed a sensing glove to capture the piano fingering from the teachers.

Designs of Always Available Sensing Devices

Instead of instrumenting a piano with sensors, the design decision was made to provide an always-available input device for capturing finger taps that can be applied to any regular piano or hard surfaces. Various approaches can be taken to achieve this goal. For instance, computer vision could help detect the finger movements and the location of the fingers [4]. This approach would require intensive computation and is error prone without rigorous configurations. An alternative method is to wear sensors on hands or arms for detecting motion. We have seen many different efforts in this approach. Sensing muscle movements by using EMG is one feasible method for detecting finger movements as an always-available input device [12]. However, EMG sensors cannot accurately recognize the timing of finger tapping on hard surfaces, which is crucial to piano playing. The other approach is to wear an array of vibration sensors on arms for detecting bio-acoustic signals on the skin [5]. In this implementation, users need to train the system every time before use since the locations of the sensors can vary when wearing the armband. In addition, this implementation is not capable of sensing multiple finger taps simultaneously. Therefore, we implemented a lightweight sensing glove with a fun and engaging way to train the system to distinguish fingers and recognize the strength of the finger taps.

HARDWARE DESIGN AND SENSING

Glove Design

Designing the Tactile Teacher glove was a challenging task. Pianists would not appreciate wearing a glove that would affect dexterity. Thus, the traditional pressure-sensing glove that employs piezo pressure sensors on the fingertips is not appropriate. We also needed to avoid additional materials on the fingers and employ as few sensors as possible. We initially employed only two piezo vibration sensors (MiniSense 100, Measurement Specialties, Inc.) on the back of the hand. Table 1 Column 1 shows that this configuration is capable of classifying the signals from one-finger taps (93%), but with very bad results when testing two-finger taps (< 50% recognition rate on average). To improve the results of recognizing two-finger taps, we added a third sensor. We experimented with many different sensor placements with these three sensors. Placing the sensors on the wrist (Figure 3B) would remove the burden of wearing a glove, which would be ideal for the pianist. However, putting the sensors on the back of the hand - closer to the fingers - returned better results. In the following sections, we describe the sensor placements and the training of the machine learning classifier.

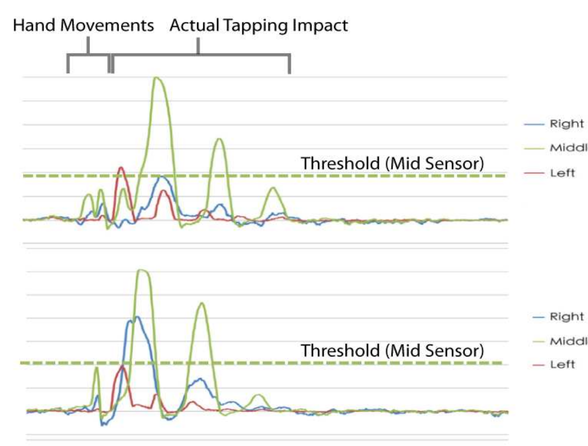

Figure 2. Top: signals of index finger tapping. Bottom: tapping of index-ring finger combination.

The Waveform from Different Finger Combinations

Figure 2 shows sample data received by the host computer with the sensor placements in Figure 3-D. The blue, green and red lines represent signals received from the right, middle and left sensors, respectively. These signals exhibit characteristics that would impact the machine learning results. For instance, in the upper chart of Figure 2 (one-finger tap), the tallest spike in the green line appears before the tallest spike in the blue line while the lower chart (two-finger tap) depicts how the tallest spike in the blue line happens only slightly before the tallest spike in the green line. This relationship is due to the smaller distance from the right sensor to the tapping finger (ring finger) in a double-finger tap than a single-finger tap – the vibration needs to travel from the index finger to the right sensor. Likewise, in a ring-finger tapping event, the tallest spike in the blue line would appear before the tallest spike in the green line, in reverse order of the spikes shown in the upper chart of Figure 2.

Hardware and Data Processing

As shown in Figure 1, we employ a Maple Mini board (from LeafLabs) [7] to sample three ADC channels at 6.5KHz from piezo vibration sensors and send the data to a computer for further data processing, before running a machine-learning program for classification. This data processing includes exponential smoothing and filtering. When the smoothed data from one of the three channels exceeds a triggering threshold, the subsequent samples from all of the channels are considered as a tap event. Likewise, if the samples fall under the threshold, the signal is considered to have exited the tapping session. As shown in Figure 2, each tap comes with the tallest spike and few shorter spikes, each of which lasts for about 12-17ms (58-83Hz). To ensure that the samples influenced by a tap event are filtered, the key parameters, such as the threshold, smoothing factor, and number of samples around the exceeding time, are tuned.

After smoothing and filtering, 174 features can be extracted from each three-sensor-input event for machine learning purposes. These features include the three averages and square averages of each channel, as well as the three ratios and square ratios between the 2-combinations of the three channels, totaling to twelve features. The system also computes a 256-point Fast-Fourier Transform (FFT) for transforming input from the time domain to the frequency domain. Because the frequency of the tap waveform is roughly between 50 to 100 Hz, only the lower nine bins (23 Hz each) are used for computing the normalized decibels; FFT is also applied to the log10 of the sensor data, totaling 162 features from FFT. Finally, peak time difference and peak value ratios are calculated for each 2-combinations of the three sensors' waveform, giving six features. The 174 features are then classified using the Support Vector Machine (SVM) package from Weka [9]. The latency of computing the data for each tap event, from receiving the data on the host computer to finishing the computation, is about 128ms on an Intel® Core™ i5 processor.

Training Procedure

Since people have different sized hands, the system requires an initial training of the machine-learning model by playing the well-known tune "Ode to Joy". This procedure not only makes the training session more enjoyable than simply repeatedly tapping on the same key multiple times, but it also gives the machine learning algorithm a chance to observe the natural variations of finger movement that may only occur when the player is tapping the key organically. During the training session, the program records all of the data from the three vibration sensors to build the model described above. Since the program knows the finger taps required for the tune, it can map the set of features to each finger as it progresses.

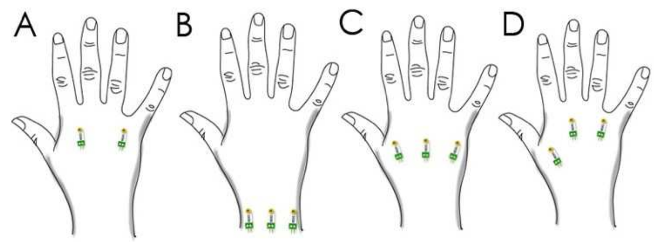

Figure 3. Different sensor placements and orientations

Table 1. Sensor placements versus tap accuracy percentage.

Finger	1	2	3	4	5	Single Avg.	1-3	2-4	3-5	Total Avg.
Pl.A	100	100	90	85	90	93	70	20	60	76.875
Pl.B	100	100	50	75	45	74	55	60	70	69.400
Pl.C	75	55	50	35	55	54	80	65	70	60.600
Pl.D	95	100	90	70	90	89	100	75	85	88.125

Results from Different Sensor Configurations

In order to find the best sensor placements, we have conducted the experiments with the machine learning classifier described above. The experiment results were obtained by the following procedure: after the model-training procedure, 20 taps for each finger and finger combination were performed, and each tap's correctness was verified. Sensor locations were changed between trials for experimenting results from different sensor placements. We picked four example placements as shown in Figure 3, and reported the results in Table 1, where finger numbers 1 through 5 correspond to the thumb, pointer, middle, index and pinky fingers, respectively.

As predicted, the best result is derived from sensors placed on the back of the hand in a radial fashion. (Figure 3-D). In our observations, the orientations of the sensors are also important. Specifically, the placements of the sensors aligned toward the spreading directions of the fingers obtained maximal sensibility. The optimal placement is to align one sensor in the same direction of the middle finger, and the other two sensors placed facing towards the in-between directions of the rest of the fingers (i.e., between thumb and index finger, between ring finger and pinky).

CONCLUSION

We are excited that Tactile Teacher can capture teacher's finger tapping with high accuracy. We have demonstrated the prototype to a piano teacher in a demo session who responded with positive feedback and showed enthusiasm about the prototype and its idea. While she would refuse to wear a complete glove with fingers covered, she believes she could wear our fingerless glove in piano playing and teaching that could benefit novice piano learners.

In the future, we would explore capturing more of the embodied experience of piano playing for a more engaging kinesthetic learning. For instance, a pianist may use body movements to increase the dynamics of the music, and achieve complex, multi-limb movements. Eventually we may also use additional sensors worn on the arms or other body parts for detecting and conveying these supplemental features. These sensors may also provide data from which we can more intelligently filter the vibration sensor signals, rather than the currently rudimentary approach of setting a threshold through experimental data.

In addition to complementing piano lessons, we believe this prototype can lead us to many other applications that were not possible before, especially those that require fine dexterity and a high level of motor skills, such as learning other (musical) instruments that also require finger movements. This prototype can also be extended and applied to the rehabilitation procedure, including regaining the capability of using hands of a post-stroke patient.

SUMMARY

In this paper, we have described an implementation of a light weight fingerless glove to recognize the finger taps in piano playing with only three vibration sensors placed on the back of the hand. These sensors, passed into a machine learning classifier, provide sufficient results of both keystrokes by a single finger (89% accuracy) and by two fingers tapping (85% accuracy) simultaneously. These results are encouraging and give us the confidence to test the prototype in a real world setting with a real pianist giving a lesson. We conclude the paper with the positive feedback from a pianist and the many rich possibilities derived from this glove configuration.

ACAKNOWLEDGEMENT

We thank our collaborators in the Contextual Computing Group, especially Prof. Thad Starner and Scott Gilliland, for feedback and supports on the initial developments of the Tactile Teacher, members of ACME Lab for thoughtful insights, and the support from the National Research Foundation, Prime Minister's Office, Singapore.

REFERENCES

1. Chellali, A., Dumas, C., and Milleville-Pennel, I. Influences of haptic communication on a shared manual task. *Interacting with Computers 23*, 4 (2011), 317–328.
2. Chellali, A.M., Dumas, C., Milleville-Pennel, I., and others. WYFIWIF: A haptic communication paradigm for collaborative motor skills learning. *Proceedings of the Web Virtual Reality and Three-Dimensional Worlds 2010*, (2010), 301–308.
3. Clark, H.H. and Brennan, S.E. Grounding in communication. *Perspectives on socially shared cognition 13*, 1991 (1991), 127–149.
4. Erol, A., Bebis, G., Nicolescu, M., Boyle, R.D., and Twombly, X. A Review on Vision-Based Full DOF Hand Motion Estimation. *Computer Vision and Pattern Recognition - Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*, 75–82.
5. Harrison, C., Tan, D., and Morris, D. Skinput: appropriating the body as an input surface. *In Proc. SIGCHI '10*, ACM (2010), 453–462.
6. Huang, K., Starner, T., Do, E., et al. Mobile music touch: mobile tactile stimulation for passive learning. *In Proc. SIGCHI '10*, ACM (2010), 791–800.
7. LeafLabs. The Maple Mini. http://leaflabs.com/.
8. Ledo, D., Nacenta, M.A., Marquardt, N., Boring, S., and Greenberg, S. The HapticTouch Toolkit: Enabling Exploration of Haptic Interactions. *In Proc. TEI '12*, ACM (2012), 115–122.
9. Machine Learing Group at the University of Waikato. Weka. http://www.cs.waikato.ac.nz/ml/weka/.
10. Morris, D., Tan, H.Z., Barbagli, F., Chang, T., and Salisbury, K. Haptic feedback enhances force skill learning. *In Proc. WHC '07*, (2007), 21–26.
11. Sallnäs, E.-L., Rassmus-Gröhn, K., and Sjöström, C. Supporting Presence in Collaborative Environments by Haptic Force Feedback. *ACM Trans. Computer-Human Interaction 7*, 4 (2000), 461–476.
12. Saponas, T.S., Tan, D.S., Morris, D., Balakrishnan, R., Turner, J., and Landay, J.A. Enabling always-available input with muscle-computer interfaces. *In Proc. UIST '09*, ACM (2009), 167–176.
13. Seim, C.E., Quigley, D., and Starner, T.E. Passive Haptic Learning of Typing Skills Facilitated by Wearable Computers. *In Proc. CHI '14 Extended Abstracts on Human Factors in Computing Systems*, ACM (2014), 2203–2208.
14. Yoshikawa, T. and Henmi, K. Human skill transfer using haptic virtual reality technology. *experimental Robotics VI 250*, (2000), 351–360.
15. Zimmermann, S., Rümelin, S., and Butz, A. I Feel It in My Fingers: Haptic Guidance on Touch Surfaces. *In Proc. TEI '13*, ACM (2013), 9–12.