World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:4, No:7, 2010

Portable Virtual Piano Design

Yu-Xiang Zhao, Chien-Hsing Chou, Mu-Chun Su, and Yi-Zeng Hsieh

***Abstract*—**The purpose of this study is to design a portable virtual piano. By utilizing optical fiber gloves and the virtual piano software designed by this study, the user can play the piano anywhere at any time. This virtual piano consists of three major parts: finger tapping identification, hand movement and positioning identification, and MIDI software sound effect simulation. To play the virtual piano, the user wears optical fiber gloves and simulates piano key tapping motions. The finger bending information detected by the optical fiber gloves can tell when piano key tapping motions are made. Images captured by a video camera are analyzed, hand locations and moving directions are positioned, and the corresponding scales are found. The system integrates finger tapping identification with information about hand placement in relation to corresponding piano key positions, and generates MIDI piano sound effects based on this data. This experiment shows that the proposed method achieves an accuracy rate of 95% for determining when a piano key is tapped.

***Keywords*—**virtual piano, portable, identification, optical fiber gloves.

I. INTRODUCTION

THE piano is an instrument with a long history, and a lovely timbre enjoyed widely. The piano, however, is an instrument with many limitations; its expensive price, large size, and inconvenience to move. To solve the spatial and movement problems associated with traditional pianos, this study presents a portable virtual piano system. This system used finger bending information detected by optical fiber gloves to distinguish the tapping motions of piano keys. We designed an image tracking method which provided hand position and movement information using video images and linking them with corresponding scales. The finger tapping identification results and corresponding scales were then integrated to play MIDI sound effects.

In recent years, data gloves have been widely promoted and applied to many research fields [1], such as human-computer interaction [2]-[5], gesture recognition [6]-[13], virtual reality [13]-[16], and medical aid and rehabilitation [14]-[16]. The most common data glove application is gesture recognition. Gesture recognition is extremely critical for sign language recognition systems. Currently, data gloves are classified into three types:

Yu-Xiang Zhao is with Ta Hwa Institute of Technology, Department of Computer Science and Information Engineering, Taiwan (e-mail: yxzhao@thit.edu.tw).

Chien-Hsing Chou is with Tamkang University, Department of Electrical Engineering, Taiwan (e-mail: chchou@mail.tku.edu.tw).

Mu-Chun Su and Yi-Zeng Hsieh are with National Central University, Department of Computer Science and Information Engineering, Taiwan (e-mails: muchun@csie.ncu.edu.tw, eizon@csie.ncu.edu.tw).

(1) Optical fiber data gloves [4], [6]-[8]
(2) Resistive data gloves [2]-[3], [8]-[11]
(3) Mechanical data gloves [5], [12], [15]-[16]

Generally speaking, the attrition rate of the optical fiber gloves is the lowest among the three types of gloves, whereas the attrition rate of the resistive gloves is the highest. Mechanical gloves have a higher accuracy rate for gesture identification; however, they can be too heavy for some applications. Some data gloves are already sold commercially, such as Power Glove, Cyber Glove, and 5DT Glove [17]-[18]. Prices are higher for those products with better accuracy and stability. For the purposes of this paper, we used Data Glove 5 virtual reality gloves produced by the 5DT company, as Fig. 1 shows. The optical fiber sensors in each finger of the glove extract information on the degree of finger bending. These gloves meet the needs of this paper because they provide accurate information, are comfortable, and easy to wear.

Fig. 1 Optical fiber gloves

II. SYSTEM DESIGN

As Fig. 2 shows, this system framework includes three units: finger tapping identification, hand movement and positioning identification, and system integration and MIDI sound effect simulation. The following are introductions of the three major units.

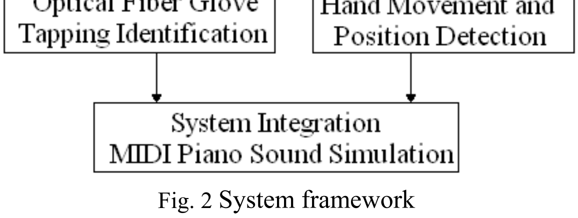


Fig. 2 System framework

A. Finger Tapping Identification

Based on a model of typical finger behavior (shown in Fig. 3), a fast and accurate method are proposed to determine when fingers made key tapping motions. Finger bending information, amplitude changes, and duration of finger bending are extracted using optical fiber gloves. This information forms a basis for calculation and identification of finger tapping motions.

Fig. 3 Finger tapping behavioral model

Different glove users have inconsistent relaxed states for their fingers. Finger tapping speeds and amplitudes vary as well. Therefore, before using the system, users were asked to initialize calibration. The calibration required every finger to repeat a tapping motion three times and recorded the measured values to adjust the parameters in the tapping identification algorithm. The proposed tapping identification algorithm included four steps:

Step 1. Bending Information Extraction of Optical Fiber Gloves

The optical gloves used in this study can read finger bending information up to 200 times per second; values range between the whole numbers 0 and 4,095 (12 bits). Figure 4(a) shows the bending information read by a single glove finger. When a finger bends downwards, the measured value becomes smaller. When the bending motion is restored, the measured value becomes larger.

Step 2. Moving Average Operation

We used moving average to smooth the data curve which got from the optical gloves. This approach is similar to the effect of a low-pass filter. Fig. 4(b) shows the results after smoothing. Only wide-range bending changes were preserved. In this paper, five-points moving average is applied in this step. To filter out minor, unintended finger movement, we the average value of five measured values, and the average values to is then used to replace the measured values. This approach is similar to the effect of a low-pass filter. Fig. 4(b) shows the results after smoothing. Only wide-range bending changes were preserved.

Step 3. First-Order Difference Operation

In first-order difference operations, the data obtained after the five-point moving average was used to reveal the relative changes in finger bending. The results are shown in Fig. 4(c). This method filters baseband sections of signals so that we can focus on relative changes in finger bending.

Step 4. Finger Tapping Identification

Based on the signals obtained after step 3, we observed that when a finger was tapping, the signals would first increase, then decrease dramatically, and finally increase. Based on these features, we first set two thresholds (T1 and T2) to detect the change of the signals. Fig. 4(d) shows the thresholds setting. If the detected signals start to become larger and have positive values, which are beyond threshold T1, then the tapping motion has started. If the detected signals then start to become larger and have negative values, which are beyond the threshold value T2, the tapping motion has ended. The occurrence of and length of time spent tapping can be determined through the identification of relative motions. In this identification method, the later change is called the restored motion, which is the finished motion after finger tapping. After the system detects a tapping motion, if the restored motions are not detected within the default maximum tapping time of the system, then the system automatically finishes recognizing the tapping motion, so as to prevent continuing piano sounds from causing sound confusion.

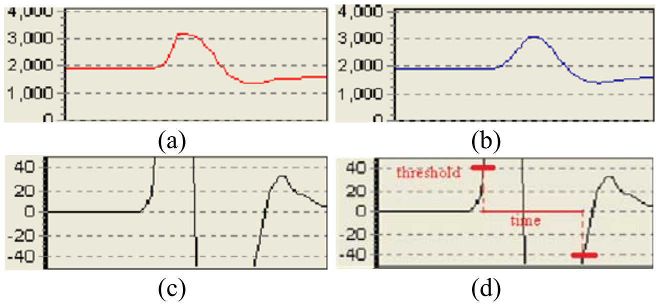


Fig. 4 Finger tapping identification process: (a) Original data; (b) Moving average operation, (c) First-order difference operation (d) Threshold setting and finger tapping identification

In the tapping identification method above, choosing the appropriate threshold values, T1 and T2, is critical for tapping identification calculation accuracy. The T1 and T2 threshold settings should be set uniquely for each finger. Different users also tap at different strengths, therefore, setting T1 and T2 thresholds can be very difficult. To solve the problem of setting threshold values, this paper uses the genetic algorithm as a basis and makes online adjustments according to the calibration information provided by the user's initialization so that corresponding threshold values can be automatically generated for different users.

We first asked 5 users to participate in training data creation. Each finger of every user carried out a tapping motion 10 times and the bending signal generated by each finger tapping was recorded. The training data containing 250 records of tapping motions are collected. Afterwards, we used the genetic algorithm to search for appropriate threshold values. Fifty chromosomes were randomly generated; each chromosome contained 10 sets of T1 and T2 parameters. When the 10 sets of T1 and T2 parameters were used in the tapping identification calculations, the finger tapping identification was used to set the fitness function. After 100 iterations, we retained 5 sets of chromosomes that had the highest accuracy rates as references of original parameters for different users when they made adjustments online.

For actual online use, each user is asked to perform a tapping motion three times for each finger to further fine-tune the thresholds. Since there was little time for optimal parameter search, we used the 5 most accurate sets of chromosomes as a basis and randomly fine-tuned to generate 20 new chromosomes. Three iterations of the genetic algorithm utilized the 30 tapping motion records provided by the users online. The T1 and T2 parameter collections of chromosomes with the highest accuracy rates were then used as the threshold values for the system's actual tests of users.

B. Identification of hand movement and positioning

To identify hand movement and positioning, we proposed an image-based tracking and positioning method. As shown in Fig. 5, the digital video camera was set up directly in front of both hands. Using digital images obtained from the video camera, piano keys corresponding to each finger could be determined. The following are introductions to the three major steps of the tracking and positioning calculation method used in this study:

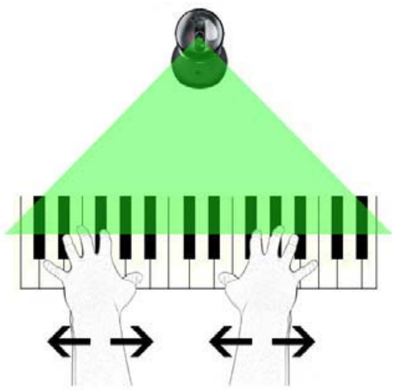


Fig. 5 Identification of hand movement and positioning

Step 1. Searching for the colored ball and calculating the center point

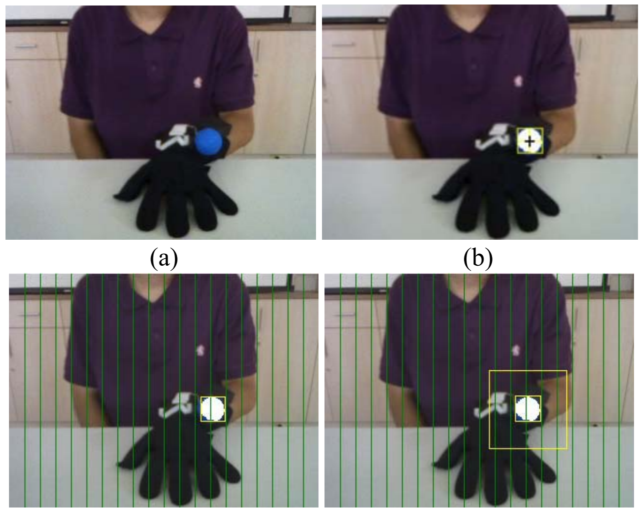
As shown in Fig. 6(a), a blue ball is attached to one of the gloves. Based on the color of the ball, the exact position of the colored ball was isolated from the entire image. Then, the center position of the colored ball was calculated for further tracking, as indicated in Fig. 6(b). In addition, users could select appropriate ball colors based on the environment and the background to increase searching accuracy.

Step 2. Dividing images into blocks

We vertically divided captured video images into a fixed number of blocks. Each block corresponded to a piano key of a specific scale, as Fig. 6 (c) indicates. When the colored ball's center point was located in a specific block, we could determine the piano scale that corresponded to the fingers.

Step 3. Region tracking and positioning

Due to the need for real-time tracking of the colored ball's position, we extended the search region from the current center point of the colored ball to 2.5 times the size of the block, as Fig. 6 (d) shows. Using the local search method can reduce the required operations for tracking the colored ball. If the colored ball could not be found in the new search region, we returned to Step 1 and searched the entire image. Otherwise, we repeatedly carried out Steps 2 and 3.



(a) (b) (c) (d)

Fig. 6 Hand tracking and positioning: (a) The colored ball attachment; (b) Colored ball center point calculation; (c) The image divided into blocks; (d) Regional tracking and positioning

C. System integration and simulation of MIDI sound effects

In this unit, the finger tapping identification information is integrated with hand position. If we identified tapping motions, then based on hand position, MIDI sound effects simulated the virtual piano keys being played. Fig. 7 shows the MIDI simulated piano interface applied in the integration system. The piano sounds were set using the standard MIDI, with 127 instrument sounds to choose from.

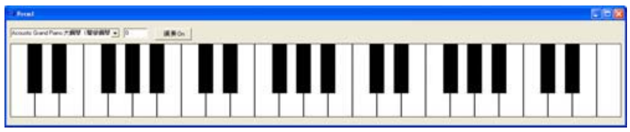


Fig. 7 Illustration of MIDI simulated piano software

III. Experiment Results

We tested for finger tapping identification using optical fiber gloves, hand movement and positioning detection, and the ability to play music using this program. To test finger tapping identification, the system randomly selected a finger and asked the user to tap with the finger. Each test required 50 taps with different fingers. A total of 7 users were tested. A total of 350 records were collected. Table I shows the results. The average accuracy rate is 94.9%.

To test hand movement and position detection, we set the resolution of the video camera at 320 240 and evenly divided the picture into 20 blocks. The search radius was set to be from the center of the colored ball outward to 2.5 times the block width (60 pixels). We had to consider the relative positions of both hands as well as the video camera so that piano key positions corresponding to fingers could match the video camera images. Therefore, the relative positions of both hands and the digital video camera were adjusted before undertaking the experiment. During testing, the system arbitrarily selected a block and asked the user to move inside the block. Each test required 50 movements to different blocks. A total of 7 users were tested. One hundred and forty tests were recorded. Table II shows the results. The average moving time was 1.4 seconds.

In testing playing music on the virtual piano, we designed a piece of music (Fur Elise) for playing. The music required the user to press 139 piano keys in different positions. Table III shows the seven users' total required time for successfully playing the entire piece of music. On average, users needed 115 seconds to play the entire piece of music and 0.83 seconds to play a note correctly.

TABLE I
TAPPING IDENTIFICATION OF OPTICAL FIBER GLOVES

user / test	1	2	3	4	5	6	7	Average	Average Accuracy
Correct	336	319	341	331	338	326	334	332.1	94.9%
Incorrect	14	31	9	19	12	24	16	17.9	

TABLE II
HAND MOVEMENT DETECTION AND POSITIONING EXPERIMENT

user / test	1	2	3	4	5	6	7	Average
Time (sec.)	66.9	73.9	64	71.2	72.0	73.4	69.4	70.1

TABLE III
MUSIC PLAYING EXPERIMENT

user / test	1	2	3	4	5	6	7	Average
Time (sec.)	106	126	104	118	125	113	114	115

IV. DISCUSSION AND CONCLUSION

The primary goal of this study is to design a portable virtual piano. The system has two major parts: the finger bending identification of optical fiber gloves and the identification of hand movement and positioning. In the calculation of finger tapping identification, we used the genetic algorithm as a basis and made appropriate threshold value adjustments. The experiment results revealed a satisfactory accuracy rate.

For hand tracking and positioning calculations, if the picture resolution of the video camera is higher, the required calculations would be increased. In considering the need of real-time tracking, then, researchers must select appropriate picture resolution to carry out the experiment. In addition, during object tracking, search range settings seriously affect the outcome. The wider the search range, the greater amount of calculation is required.

These experimental results indicate that on average, hand movement requires 1.4 seconds. The reason for this is that the required virtual piano keys being tapped are not often exactly where they are expected to be. When the system is in actual performance, hand movements are not always required, leading to better performance outcomes.

ACKNOWLEDGMENT

This work was supported by the National Science Council, Taiwan, R.O.C., under the Grant NSC 98-2218-E-233-002.

REFERENCES

[1] D. J. Sturman and D. Zeltzer, "A Survey of Glove- based Input", *IEEE Computer Graphics and Applications*, vol.14, no.1 pp. 30-39, 1994.
[2] G. C. Sun, Self-balancing Control and Manipulation of a Glove Puppet Robot on a Two-wheel Mobile Platform, Master's thesis, Institute of Electrical Control Engineering, National Chiao Tung University, 2009.
[3] J. S. Hu, J. J. Wang, and G. Q. Sun, "The Glove Puppet Robot: X-puppet," *IEEE/RSJ 2008 International Conference on Intelligent Robots and Systems (IROS '08)*, Nice, France, 2008.
[4] M. C. Su, C. H. Hung, C. K. Yang, H. T. Yang, Z. J. Huang, Y. X. Zhao, and D. Y. Huang, 2005, "Design of Optical Fiber Glove and Its Application," in *2005 13th National Conference on Fuzzy Theory and Applications*, Kaohsing, Taiwan.
[5] J. M. Hollerbach, "Anthropomorphic Robot and Human Interactions", in *Proc. of 1st. Intl. Symposium on Humanoid Robots*, 1996.
[6] S. S. Fels and G. E. Hinton, "Glove-Talk: a neural network interface between a Data-Glove and a speech synthesizer", *IEEE Trans. On Neural Networks*, Vol. 1, pp. 2-8, 1993.
[7] M. C. Su, Y. X. Zhao, C. H. Hung, and M. Chunag, "A Fuzzy Rule-Based Approach to Hand Shape Recognition Systems," in *The Joint Conference on AI, Fuzzy System, and Grey System*, Taiwan, Dec. 2003.
[8] Y. X. Zhao, A Portable Speaking-Aid System for Taiwanese Sign Language, Master's thesis, Department of Electrical Engineering, Tamkang University, 2002.
[9] C. H. Hung, A Trajectory-based Approach to Gesture Recognition, Master's thesis, Department of Computer Science and Information, Engineering, National Central University, 2006.
[10] M. C. Su, Y. X. Zhao, H. Huang, and H. F. Chen, "A Fuzzy Rule-Based Approach to Recognizing 3D Arm Movements," *IEEE Trans. on Neural Systems and Rehabilitation Engineering*, vol.9, no. 2, pp. 191-201, June 2001.
[11] M. C. Su, C. H. Hung, and Y. X. Zhao, "A SOMART System for Gesture Recognition," *WSEAS Transactions on Computers*, vol. 5, pp. 2764-2771, 2006.
[12] V. Tartter and K. Knowlton, "Perception of sign language from an array of 27 moving spots," *Nature*, vol. 2, pp. 676-678, 1981.
[13] J. Weissmann and R. Salomon, "Gesture recognition for virtual reality applications using data gloves and neural networks," *Proceeding of the International Joint Conference on Neural Networks*, Vol.3, pp.2043-2046, 1999.
[14] G. Burdea, S. Deshpande, B. Liu, N. Langrana, and D. Gomez, "A Virtual Reality-Base System for Hand Diagnosis and Rehabilitation," *IEEE Transaction on Rehabilitation Engineering*, vol. 6, no. 2, pp. 229-240, 1997.
[15] V. Popescu, G. Burdea, and M. Bouzit, "Virtual Reality Simulation for a Haptic Glove," *IEEE Computer Animation Proceedings*, Geneva, Switzerland, pp. 195-200, 1999.
[16] V. Popescu, and G. Burdea, "A Virtual-Reality-Base Telerehabilitation System with Force Feedback," *IEEE Transactions on Information Technology in Biomedicine*, vol. 4, Issue: 1, pp. 45-51, March 2000.
[17] Power Glove, http://en.wikipedia.org/wiki/Power_Glove.
[18] 5DT Data Glove & CyberGlove, http://www.vrlogic.com/html/datagloves.html.
[19] S. Vance, L. Migachyov, and W. Fu, I.Hajjar, "Fingerless glove for Interacting with data processing system", United States Patent, Patten No: 6,304,840, Oct.16, 2001.