Using the Kinect to Detect Potentially Harmful Hand Postures in Pianists

Mengyuan Li, *IEEE Student Member*, Paola Savvidou, Bradley Willis and Marjorie Skubic, *IEEE Senior Member*

Abstract— **Pianists who practice hours per day may have a risk for developing playing-related musculoskeletal injuries if they do not play with proper hand alignment. In order to detect the harmful, misaligned hand postures (such as wrist flexion and extension, knuckle collapse, and ulnar and radial deviation) and analyze the injury risk, a motion capture system was developed using the Microsoft Kinect depth camera. Data were captured on professional pianists and student pianists from the School of Music, University of Missouri. Features extracted from the 3D point clouds reconstructed from the Kinect depth images are used for hand posture evaluation. Results are included for different hand postures of participating pianists.**

I. Introduction

Almost every pianist or piano teacher has experienced pain or knows someone who has experienced pain from playing the piano [1]. Prevalence rates for Playing Related Musculoskeletal Disorders (PRMDs) in pianists vary widely from 26% to 93%. [2] As many as 12% of professional classical musicians with PRMD have been reported to give up their profession permanently [3]. Thus, it would be very helpful if an automated assessment was available to detect a high risk of injury so that corrections could be made early.

There are many causes of playing-related injury, such as not enough warm-up time, practicing too long, not enough rest, awkward postures, and static muscular activity. Some related work has been done. Mohamed et al. found a statistically significant difference in hand temperatures between pianists with and without pain [4]. Other studies are based on captured motion. Rosety-Rodriguez et al. analyzed the active range of motion of wrist joints of young players affected by Repetitive Strain Injury (RSI); their results showed that pianists who could perform both maximal flexion and extension were rarely affected by RSI [5]. Neninger et al. studied the use of wearable devices and markers for motion capture to investigate hand injuries among musicians [6]. One drawback of this approach is that the wearable devices can limit the movement of the player or may cause unnatural movement. Also, marker-based motion capture systems are typically expensive and may limit access.

The Microsoft Kinect device uses infrared (IR) light to generate depth images which can give a 3D representation after reconstruction. This provides a cost-effective sensing device to replace the expensive 3D motion capture equipment. In addition, it offers a significant advantage in being non-wearable which has no effect on the playing movements. In related work, Hadjakos used the Kinect for motion capture of pianists [7] and compared the performance of the Kinect with 2D marker tracking. Leap Motion is another IR motion sensor device we explored. It focuses on the hand and fingers which is ideal for our application. Unfortunately, due to its limited-sized and fixed upside-down pyramid-shaped working space, it made it difficult to detect the hands on keyboard and give coverage of the entire piano keyboard.

The work presented in this paper uses the Microsoft Kinect. A hand posture feature extraction strategy is proposed, along with a landmark detection algorithm to automatically capture key points of the hand. Data have been collected with professional pianists and both undergraduate and graduate-level student pianists. Features extracted from raw depth images provide a representation of the different hand postures and are proposed for hand posture clustering and misaligned posture detection.

II. Metrology

A. System Setting

One Microsoft Kinect was set up above the piano using a camera stand, facing down to the piano keyboard, see Fig. 1. The data were collected using the near mode of the Kinect with a working range of 500mm to 3000mm at a frame rate of 30fps. An example of our raw data is shown in Fig. 2.

Figure 1. Kinect-Piano setting

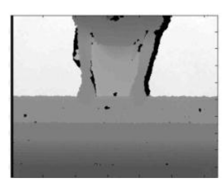

Figure 2. Raw depth image

B. System Overview

In our system, the raw depth images are captured using Microsoft Kinect. The foreground (player) is then extracted

M. Li and M. Skubic are with the Electrical and Computer Engineering Department, University of Missouri, Columbia, MO 65211 (emails: mlq3f@mail.missouri.edu, skubicm@missouri.edu).

P. Savvidou is with School of Music, University of Missouri, Columbia, MO 65211 (email: savvidoup@missouri.edu).

B. Willis is with Missouri Orthopaedic Institute, Columbia, MO 65212 (email: WillisBW@health.missouri.edu).

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

using a background subtraction algorithm. The hand samples are then segmented using a robust approach, utilizing a two-stage calculation (initial bounding box calculation and morphology optimization) and temporal smoothing along frame sequence. This then produced the following landmark positions in 3D: (1) fingertip of the longest finger, (2) hand center, (3) wrist center, and (4) a point in the mid forearm. Features are extracted from the reconstructed 3D information of the hand sample and landmark positions. See Fig. 3.

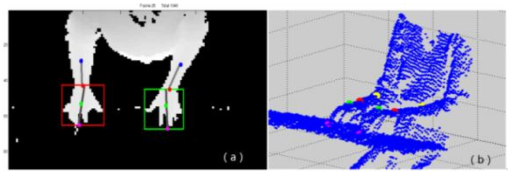

Figure 3. (a) Segmented left and right hands in the red and green bounding boxes and landmarks detected. The pink, green, red and blue dots represent fingertip of the longest finger, hand center, wrist center and a point in the mid forearm, respectively. (b) Reconstructed 3D point cloud. The pink, green, red and yellow dots represent fingertip of the longest finger, hand center, wrist center and a point in the mid forearm respectively.

C. Feature Extraction

Side view

For hand postures with problems of collapsed knuckles, wrist extension and wrist flexion, the side view outline of the hand gives a good representation of the misalignment. See Fig. 4. The following features are extracted from the side view, i.e., the vertical plane: (1) the *CAR* feature (hand **C**enter height to hand **A**rch height **R**atio) and (2) the *WAV* feature (**W**rist **A**ngle – **V**ertical).

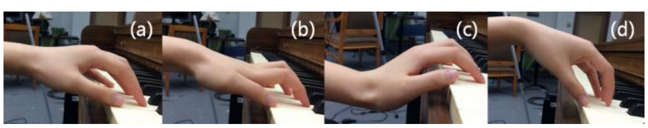

Figure 4. Side view of the different hand postures. (a) neutral; (b) collapsed knuckles; (c) wrist extension; (d) wrist flexion.

CAR feature. This feature is proposed to detect collapsed knuckles. After the hand is segmented, 3D reconstruction is done on the hand pixels, using the keyboard plane as the zero-height plane. We use the highest point in each row of the segmented hand sample as the side view hand outline height, as shown in Fig. 5(a). Then the height is calculated for each point along the side view outline to the line defined by the wrist and fingertip, to extract the side view hand shape, as shown in Fig. 5(b, c). The hand center height $H(c)$ (blue dots) and the arch peak height $H(p)$ (red dots) are then used in the ratio, as shown below.

$$CAR = H(c) \,/\, H(p) \qquad (1)$$

The *CAR* value can range from a negative value to 1. From our experiments, the negative value can be as small as -5; however, a negative *CAR* value carries much less information than a *CAR* value in the range of 0 to 1. As shown in Fig. 5(c), if *CAR* is negative, the collapse is severe and there is much less difference between a -1 *CAR* and a -4 *CAR* than between a 0.8 *CAR* and 0.2 *CAR*. In other words, when the *CAR* drops from 0.8 to 0.2, it can be a change from neutral to strong collapse while a *CAR* drop from -1 to -4 is a change from a very strong collapse to a very, very strong collapse. To address this problem, we normalize *CAR* into a range of 0 to 1 with a normalized information distribution as shown below.

$$CAR' = e^{CAR-1} \qquad (2)$$

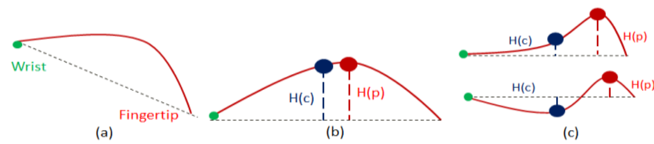

Figure 5. (a) Hand side view; (b) Hand arch shape side view of neutral hand posture; (c) Hand arch shape side view of collapsed knuckles hand posture. The green dots represent the wrist, the blue dots represent the hand center, and red dots represent the hand arch peak.

For a neutral hand posture, CAR' is close to 1, while for a collapsed knuckles hand posture, CAR' is close to 0. The CAR' feature gives a representation of the degree of collapse in the knuckles.

WAV feature. The vertical wrist angle gives a good representation of the wrist extension and flexion degrees. The angle between the forearm extension line and the hand extension line in the vertical plane is calculated as the vertical wrist angle, i.e., the *WAV* feature. Flexion yields positive *WAV* values, while extension yields negative *WAV* values, as shown in Fig. 6.

The forearm extension line is defined by the landmarks on the mid forearm and the wrist center. The hand extension line is defined by the wrist center and the hand center. Note that the landmarks sit on the surface of the arm and hand.

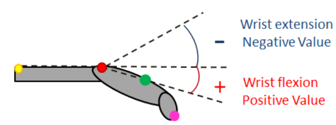

Figure 6. Vertical wrist angle; yellow, red, green and pink dots represent point in mid forearm, wrist center, hand center and fingertip.

Top view

The top view provides an indication of ulnar and radial deviation misalignments. See Fig. 7. This is captured by the *WAH* feature (**W**rist **A**ngle – **H**orizontal).

WAH feature. This feature is computed similarly to the *WAV* feature except the angle is computed in the horizontal plane. Ulnar deviation yields positive values, while radial deviation yields negative values. The 3D coordinates of the mid forearm center, wrist center and the fingertip of the longest finger are used to calculate the angle, as shown in Fig. 7.

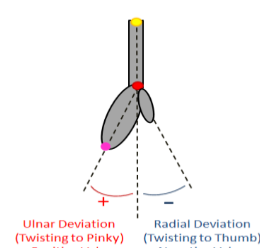

Figure 7. Horizontal wrist angle; yellow, red and pink dots represent point in mid forearm, wrist center and fingertip, respectively.

III. Data collection

An IRB-approved study was conducted to collect data from 2 professional pianists (1 male and 1 female) and 10 student pianists (2 males and 8 females) in the School of Music, University of Missouri. Three different types of data were collected. Participants were first asked to hold static hand postures for 10 seconds each. Secondly, participants were asked to slowly move through the different positions for 4 cycles, as outlined below. For example, participants moved continuously from a neutral position to the collapsed knuckles position and back to the neutral position for 4 cycles. Thirdly, participants were asked to play five music selections in their normal hand postures. These pieces were selected to elicit a range of hand postures, including misaligned hand positions, e.g., from students who may use poor hand postures.

- Static samples *(Hold the hands for 10 seconds)*
    - Neutral Position
    - Collapsed Knuckles
    - Wrist Extension
    - Wrist Flexion
    - Ulnar Deviation (Deviation toward Pinky)
    - Radial Deviation (Deviation toward Thumb)
- Continuous samples *(Move back and forth slowly for 4 cycles)*
    - Collapsed knuckles to neutral position
    - Wrist extension to flexion through neutral position
    - Ulnar to radial deviation through neutral position
- Selected pieces *(Play each three times at a given speed)*
    - Chord progression in C major
    - D major Scale (two octaves)
    - D major Arpeggio (two octaves)
    - Chorale "Rejoice, O My Soul", Op. 68, no. 4 (R. Schumann)
    - Little Prelude No.2 in C major, BWV 939 (J.S. Bach)

IV. Experimental Results

In this paper, 5 different misaligned hand postures in total (collapse in knuckles, wrist extension, wrist flexion, ulnar deviation and radial deviation) are discussed. In this section, we present results of the proposed features for determining misaligned hand postures from a neutral position, using the data collection described in Section III. Note that all of the results in this paper are for the right hand, due to space limitations.

We first present an example of the continuous, slow movement through different hand postures. Fig. 8 shows the movements in the proposed feature space. The red traces represent neutral position samples. The green traces represent the back and forth movement from neutral to collapsed knuckles. The blue traces show movement from wrist extension to flexion through the neutral position. Finally, the orange traces show the movement from ulnar deviation (toward pinky) to radial deviation (toward thumb).

In Fig. 8, we can clearly see the trace of the movements in the feature space. A control variate method was used here to allow movement along one axis only. For example, in the movement between neutral and collapse, we control the wrist posture such that no flexion, extension or twisting is demonstrated. As shown in Fig. 8, the green trace moves along the CAR' axis, roughly perpendicular to the WAV-WAH plane. Similarly, the blue trace shows a back and forth movement along the WAV axis from extreme wrist extension to extreme flexion passing through the neutral. The orange trace shows a change along the WAH axis.

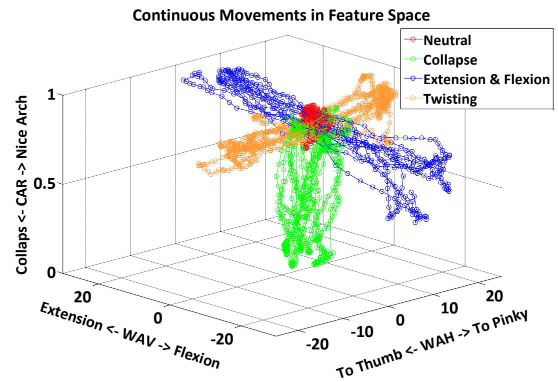

Figure 8. Continuous Movements in Feature Space. The three axes are WAV in degree, WAH in degree and CAR'. The red shows the neutral hand posture while the green trace represents back and forth movement from neutral to collapse. The blue trace represents back and forth movement from wrist extension to flexion passing through the neutral; flexion has a positive WAV value. The orange trace represents back and forth movement between ulnar deviation (toward pinky) and radial deviation (toward thumb) passing through the neutral; ulnar deviation has a positive WAH value.

The data of these continuous movements are processed to get the extreme values of each feature for the 5 misaligned postures: CAR' for collapse, WAV for wrist extension and flexion and WAH for ulnar and radial deviation. Fig. 9 shows the result of the neutral and extreme values of each feature for the 12 subjects. The red circles represent the neutral positions, and the blue and green markers represent the extremes. From Fig. 9, we can see the difference in neutral hand postures of different subjects and also the active range of motion of the hands of different subjects. We found that most pianists share a neutral CAR' around 0.85, a neutral WAH angle around 0 degrees, and a neutral WAH angle around 10 degrees. This indicates that 10 degrees of ulnar deviation is quite normal and neutral for most of the pianists, especially for the situation when they need to play in the middle of the keyboard (just in front of them) with two hands close to each other.

Data from the music selections show the range of hand postures used by each subject as the piece is performed. We present examples of these in Fig. 10-12 for subject 1, who is a professional pianist. The figures show the histogram of each feature for 3 of the music selections. The gray bars show the histogram. The red line shows the neutral position. The blue lines represent the radial deviation extreme for WAH, wrist extension extreme for WAV, and collapse extreme for CAR'. The green lines represent the ulnar deviation extreme for WAH and wrist flexion extreme for WAV.

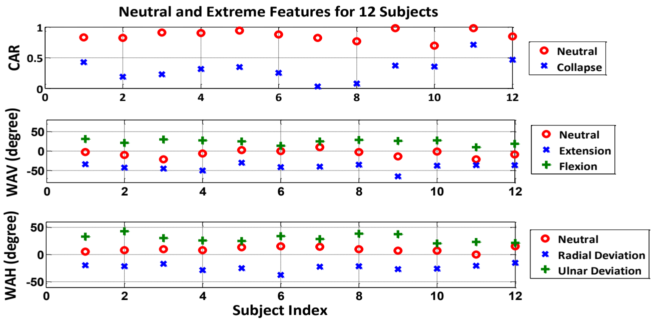

Figure 9. Feature values for Neutral and Extremes of 12 subjects. The x-axes show the subject index and the y-axes show the feature values. The red circles represent the neutral feature, while the green crosses and blue crosses show the extreme features.

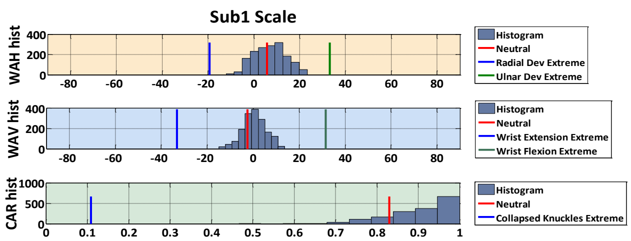

Figure 10. Histograms of each feature for the D major Scale (two octaves). The gray bars represent the feature histograms. The red line shows the neutral position while the blue and green lines represent the extremes.

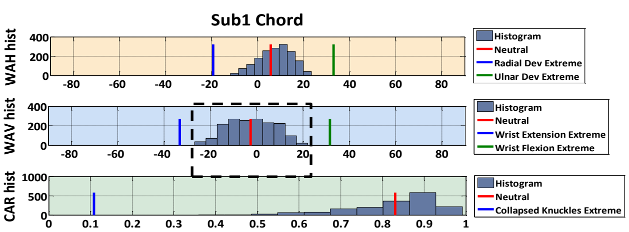

Figure 11. Histograms of each feature for the Chord progression in C major. The gray bars represent the feature histograms. The red line shows the neutral position while the blue and green lines represent the extremes. The black box highlights a slightly wider range of wrist extension and flexion in chord than in other pieces.

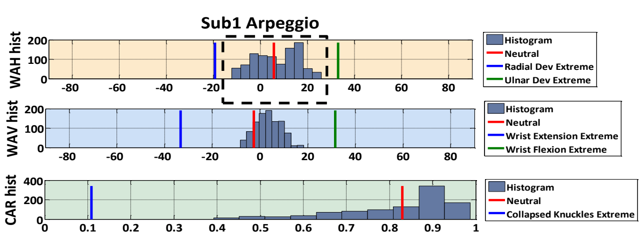

Figure 12. Histograms of each feature of D major Arpeggio (two octaves). The gray bars represent the feature histograms. The red line shows the neutral position while the blue lines represent the extreme. The black box highlights a slightly wider range of ulnar and radial deviation in arpeggio than in other pieces.

From these figures, we observe that in pieces with notes concentrated in the middle part of the keyboard or moving back and forth from left to right such as scale, chord and arpeggio, the wrist has certain degrees of deviation, wrist extension and flexion around the neutral. Most of the samples sit near the neutral position with slight ulnar deviation, which is confirmed to be normal with both piano professor and our result shown in Fig. 9. For chord (Fig. 11), we see a wider range of wrist extension and flexion in feature WAV comparing to the scale and arpeggio due to the more vertical wrist movements in chord. For the arpeggio (Fig. 12), we see a wider range of ulnar and radial deviation and fewer samples in neutral for feature WAH comparing to the scale and chord due to the more hand shifts in the arpeggio. Overall, the histograms show a good hand posture for subject 1, which is expected for this professional pianist.

V. Conclusion

Our experiments show that the Kinect provides good sensing capabilities for capturing the hand postures of pianists. The setting is non-obtrusive and has no effect on the player's movement. It is portable and inexpensive which makes it accessible to many pianists. The results show that the WAH, WAV and CAR' features extracted from the reconstructed 3D point cloud and landmarks extracted from the raw Kinect depth images can provide a good representation of the hand posture. They are promising for a screening assessment to detect misaligned hand postures, including postures that show a combination of misalignment types. It is expected that some misaligned hand postures will occur during performance; however, to avoid injury, it is important that pianists move back to a neutral position instead of holding a misaligned posture.

For future work, we will collect more data of the selected pieces of music. This will include subjects playing in their natural postures of neutral and misaligned positions, as well as subjects with and without a history of injury. We will investigate clustering and correlation of our features with simultaneous measurement of standard static goniometry measurements. Also, a piano professor will label the samples to build up a dataset for future classification studies.

Acknowledgment

The research is funded by the Mizzou Advantage program. The authors are grateful for the support of the professors and students who participated in the study.

References

[1] Mark, Thomas Carson. *What every pianist needs to know about the body: A manual for players of keyboard instruments: piano, organ, digital keyboard, harpsichord, clavichord*. GIA Publications, 2003.

[2] Bragge, Peter, Andrea Bialocerkowski, and Joan McMeeken. "A systematic review of prevalence and risk factors associated with playing-related musculoskeletal disorders in pianists." *Occupational Medicine* 56.1 (2006): 28-38.

[3] Parry, C. B. "Prevention of musicians' hand problems." *Hand clinics* 19.2 (2003): 317-324.

[4] Mohamed, Safaa, Monique Frize, and Gilles Comeau. "Assessment of piano-related injuries using infrared imaging." *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*. IEEE, 2011.

[5] Rosety-Rodriguez, M., Ordóñez, F. J., Farias, J., Rosety, M., Carrasco, et al. "The influence of the active range of movement of pianists' wrists on repetitive strain injury." *European Journal of Anatomy* 7.2 (2014): 75-77.

[6] Neninger, C. R., Sun, Y., Lee, SH., Chodil, J. "A Complete Motion and Music Capture System to Study Hand Injuries among Musicians." *Emergency Management & Robotics for Hazardous Environments*: 1-11.

[7] Hadjakos, Aristotelis. "Pianist Motion Capture with the Kinect Depth Camera."*Proceedings of the International Conference on Sound and Music Computing, Copenhagen, Denmark*. 2012.