Computer Analysis of the Indirect Piano Touch: Analysis Methods and Results

Aristotelis Hadjakos	Erwin Aitenbichler	Bernhard Wetz	Max Mühlhäuser
TU Darmstadt	*TU Darmstadt*	*HfMDK Frankfurt*	*TU Darmstadt*
telis@tk.informatik.tu-darmstadt.de	*erwin@tk.informatik.tu-darmstadt.de*	*b.wetz@inspiration-b.de*	*max@tk.informatik.tu-darmstadt.de*

Abstract

Indirect touches - touches that originate from above the key - play an important role in piano technique. Analysis methods are presented and applied in a study with piano students performing different touches in slow motion. Colored markers that were attached to the players' fingers were tracked and the angles in the joints were determined. Methods to judge the regularity of the measured movements are introduced and applied to the obtained dataset. Further, phenomena that we found in the motion graphs are discussed.

1. Introduction

Analyzing the technique of a piano player, two types of touch are distinguished [4]:

- Direct touch and
- Indirect touch.

A direct touch begins with the finger in contact with the key. At the starting point of a direct touch, the finger is at rest. The finger then continuously accelerates the key. On the contrary, the indirect touch begins with the finger above the key. When the finger hits the key, it has already attained a considerable speed. This is a key difference to direct touch with implications on the sound because of noise being generated when the finger hits the key. In this paper we examine the indirect touch.

For normal touch, the finger is flexed in the knuckle (1st joint). The 2nd and 3rd joints contribute to the finger's motion by flexion or extension [4]. Henceforth, we will call a touch with flexion of the 2nd and 3rd joint a flexion-touch and a touch with extension of the 2nd and 3rd joint an extension-touch.

The remaining paper is organized as follows. In section 2 we discuss related work. Next, we describe the design goals and the approach of our touch analysis software. In section 4, we describe the study, which was made with piano students. Typical results of the analysis are shown in section 5. We present formal methods for the analysis (section 6) and apply them on the user study (section 7). Conclusion and future work sections follow.

2. Related work

Music via Motion (MvM) [9] is a framework that allows mappings from physical movement to multimedia events. MvM uses video tracking and other sensor technology for the acquisition of movement data.

The Conductor's Jacket [8] is used to gather data from conductors. It consists of various sensors integrated into clothing. The Conductor's Jacket measures physical motion and physiological activity. It has EMG-sensors, sensors for breathing, body temperature, galvanic skin reaction, and heart rate. Additionally, there are position and orientation sensors attached at various points.

Schoonderwaldt et al. [10] developed a bow tracking system that consists of a combination of optical tracking and acceleration sensors. The system uses EyesWeb [2] as a framework for tracking colored markers on the bow.

The Hyperbow [11] is a commercial carbon fiber bow with attached sensors. The tracking of the bow position is done by oscillators attached to the ends of the bow and an antenna at the frog of the instrument. Flexion sensors measure the force that is applied to the bow when pressing it against a string.

A commercial visual tracking system (Selspot) was used by Dahl [3] to track movement trajectories of drummers performing a rhythmical pattern. Individual movement habits were found, which the players kept consistently in all playing conditions.

The Piano Pedagogy Research Lab of the University of Ottawa uses sensor technology to analyze piano playing with the computer [1]. The researchers create tools for teaching, research, and for prevention of piano playing related health problems [7].

To measure the behavior of grand piano actions, Goebl et al. [5] attached acceleration sensors to selected keys and corresponding hammers of the examined pianos. Direct and indirect touches were executed on the prepared instruments and resulted in different behaviors of the piano action. In direct touches the motion of the hammer starts immediately with the motion of the key. In indirect touches, the hammer motion starts several milliseconds after key motion. In contrast to Goebl's work we directly examine the finger motions of the players performing indirect touches.

3. Analysis software

3.1. Design goals

The design goals for the analysis software were:
Touch Analysis: The software should capture the finger motion and create a graphical representation.
Automation: The process should be automatic and require as little human intervention as possible.
Low Cost: The acquisition hardware should not be expensive. A system based on off-the-shelf components would also be affordable for hobbyists.
Extensibility: The software should support extending the analysis. Two areas are relevant: (1) the analysis of motions originating from the wrist or elbow and (2) the refinement of the analysis by using better and more expensive acquisition hardware.

3.2. Approach

The finger motion of the piano player is recorded as a video. To enable tracking, colored markers are attached to the player's finger. The video is analyzed in three steps. First, the positions of the markers are tracked. This is done by *MotionAnalyzer*, which was developed for this purpose. The second step is to compute the angles between phalanges of the finger with *AngleExtractor*, a command line program. Third, a graphical representation is generated. This is done with *ToDat* that generates a file that is used as input for Gnuplot.

3.3. MotionAnalyzer

MotionAnalyzer (see figure 1) is used for tracking colored markers attached to the player's finger. *MotionAnalyzer* displays a still image from the video. The user first defines the reference colors by clicking

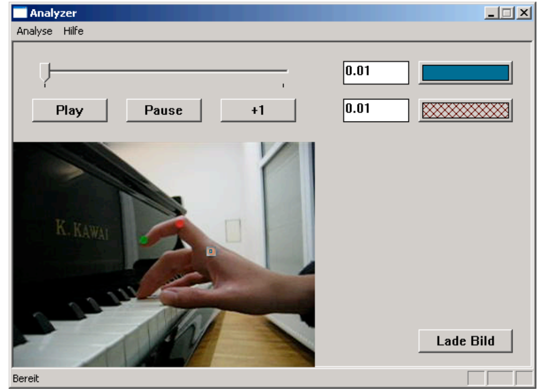

Figure 1. MotionAnalyzer

on the markers to be tracked in the image. The program will then search for these markers and extract their coordinates. The user can also adjust the sensibility of the recognition for each color individually.

In each still image, which is extracted from the video, *MotionAnalyzer* searches for the marker colors. It computes the Euclidian distance of each pixel to the defined reference colors taking into account the RGB components of the color. If the Euclidian distance falls below a threshold, which the user defines for each color individually, the pixel is grouped to the reference color. The median of the x- and y-coordinates of the so collected pixel is than the recognized position. *MotionAnalyzer* was implemented using C++ and DirectShow on a Windows XP platform.

3.4. AngleExtractor

AngleExtractor computes the angles of the finger joints from the data provided by *MotionAnalyzer*. *AngleExtractor* computes the angle between two line segments that are formed by the positions of three successive markings. The order of the markings is given by the order in which they were defined in *MotionAnalyzer*. An additional angle is computed at the position of the first marking: It is the angle between the line segment of the first and second marking and an imaginary horizontal line intersecting the first marking. See figure 2 for an example of the computed angles.

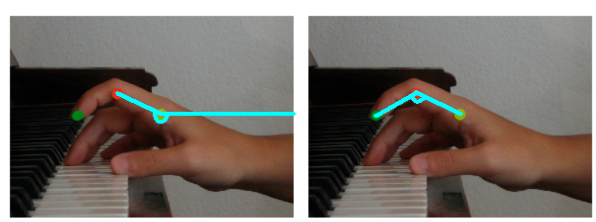

Figure 2. Measured angles

4. Experiment

A user study with five piano students was conducted at the HfMDK Frankfurt. The students were recorded playing indirect flexion- and extension-touches. They played the touches with the index and little fingers of both hands resulting in 40 samples of data. A Canon Ixus 30 digital camera was used, which provided a video in MJPEG format with a resolution of 320x240 pixels at a temporal resolution of 60 fps. The students were asked to slow down the motion artificially because the camera does not have enough temporal resolution to capture the motion at original speed. A metronome was used to make sure that the students executed the motions with a defined rhythm and to hint on the velocity of the motion. Markers were attached to the knuckle (1^{st} joint), the 2^{nd} joint, and the fingertip. The two resulting angles were measured (see figure 2). Because hand and arm are at rest, there is no need to add markers behind the 1^{st} joint.

5. Observation

The figures generated from the angle measurements visualize two angles as a function of time. The upper graph represents the angle in the 1^{st} joint; the lower graph represents the angle in the 2^{nd} joint. Figure 3 shows a series of flexion-touches. It begins with the finger resting on the key. After a short time, the player lifts the finger and the angles in the 1^{st} and 2^{nd} joints increase. After reaching the maximum height, the angles in the joints decrease again until the finger hits the key.

Figure 4 shows a series of extension-touches. It starts with the finger resting on the key. After a short time, the player lifts the finger. While the angle in the 1^{st} joint increases, the angle in the 2^{nd} joint decreases. After the finger reaches the maximum height, the angles in the 1^{st} joint decrease while the angles in the 2^{nd} joint increase.

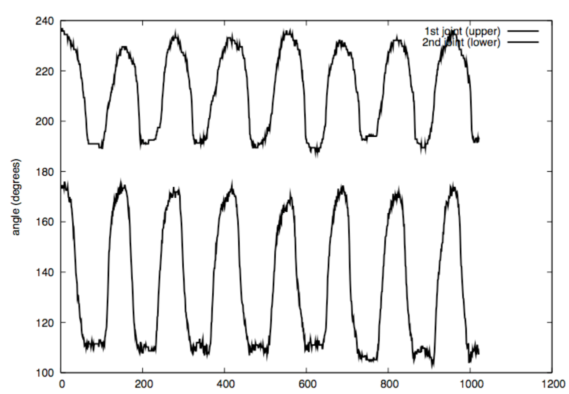

Figure 3. Flexion-touch

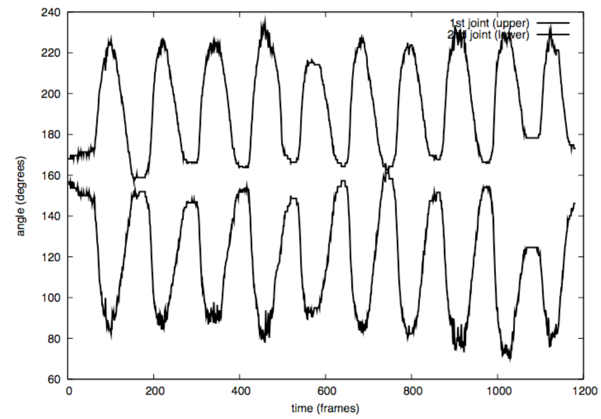

Figure 4. Extension-touch

6. Analysis Methods

6.1. Levels and ways

The gathered data was analyzed with formal methods. Before these methods can be applied, the motion curves have to be segmented. The following segments are distinguished (see figure 5):

- Preparation way,
- Preparation level,
- Hit way, and
- Hit level.

Preparation level and hit level are phases of relative stability. There is only little motion and the present motions extinguish each other. The hit level occurs when the finger hits the key. The preparation level occurs when the finger reaches its end position above the key. The preparation way is the transition from hit level to preparation level. The hit way is the transition from preparation level to hit level.

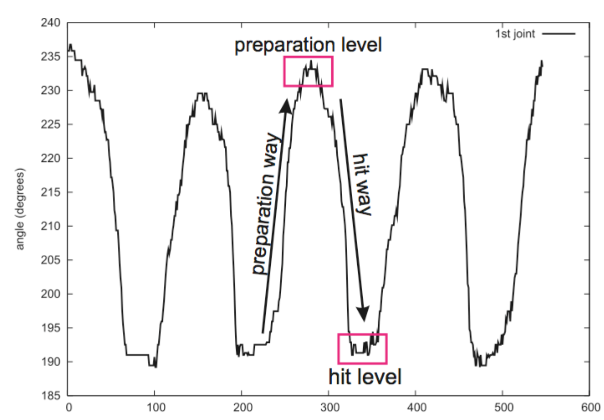

Figure 5. Segments

6.2. Properties of levels and ways

The height of a preparation level is the maximum angle achieved in the preparation level. The height of a hit level is the minimum angle achieved in the hit level. The above definitions of height apply for the 1st joint executing either flexion-touch or extension-touch and for the 2nd joint executing flexion-touch.

For the 2nd joint executing an extension-touch the preparation and the hit levels are vertically flipped because the hit level occurs when the 2nd joint is fully extended. For this case, the height of a preparation level is the minimum angle achieved in the preparation level and the height of the hit level is the maximum angle achieved in the hit level.

The length of the hit way is the difference between the heights of the connected levels.

6.3. Equality and translation measures

Equality and translation measure, which are defined in this section, are aids for judging the regularity of a series of touches.

The equality measure (E) is defined as the fraction of the shortest hit way to the longest hit way of a series of touches (see figure 6). For the translation measure (T) the heights of four levels of a series of touches have to be considered:

- Lowest hit level,
- Highest hit level,
- Lowest preparation level, and
- Highest preparation level.

The translation measure is the fraction between the minimum distance between a preparation and hit level and the maximum distance between a preparation and hit level (see figure 7).

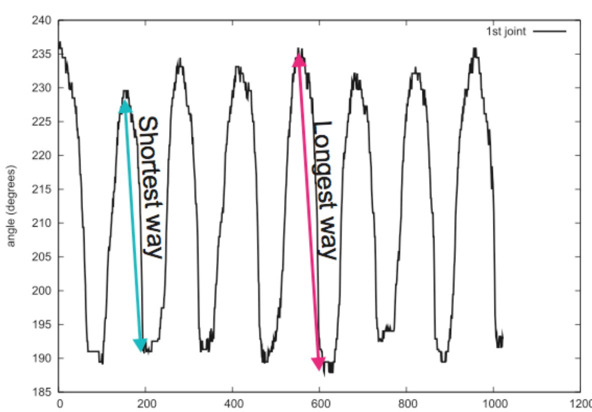

Figure 6. Equality measure

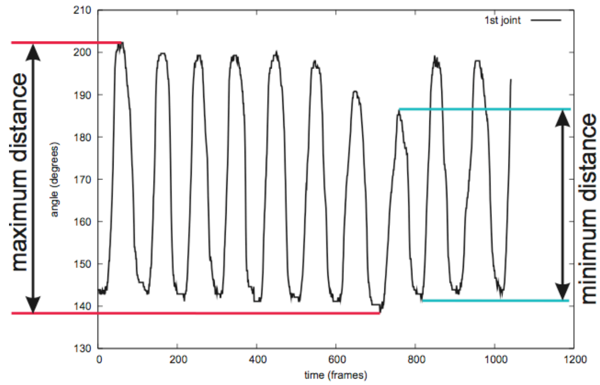

Figure 7. Translation measure

If the motion in a joint reaches similar preparation level and hit level heights in a series of touches, the equality measure of that joint's movement will have a value close to 1. However, an offset the preparation and hit level height of the same amount is not detected by the equality measure. It is, however, detected by the translation measure. The combination of equality (E values) and translation (T values) measure has implications for the analyzed motion as can be seen in table 1.

Table 1. Implications of E and T values on the analyzed motion

	E big	E small
T big	Regular motion	Impossible, $T \leq E$
T small	Translation	Irregular, possibly also translated

To support the analysis process, the *ET* program was developed. It calculates the E and T values given the preparation and hit levels. The user provides this information by marking the preparation and hit levels with bounding boxes in the GUI of the *ET* program.

7. Analysis

7.1. Quantitative analysis

In our dataset of 40 samples, the motion of the 1st joint tended to be more regular than the motion of the 2nd joint in both, flexion-touches and extension-touches. In 80% of the cases the E value of the 1st joint was higher than the E value of the 2nd joint. In 87.5 % of the cases, the T value of the 1st joint was higher than the T value of the 2nd joint.

Extension-touches and flexion-touches of each finger were compared, giving 20 pairs to be considered. In 80 % of the pairs the T value of the 2^{nd} joint was higher when executing flexion-touches. However, only in 60 % of the pairs the E value of the 2^{nd} joint was higher when executing flexion-touches. Overall, in the combination of E and T values, it seems that the flexion-touch motion could be more regular.

A list of the measured E and T values and more details about the analysis process can be found in [6].

7.2. Empirical evaluation

In some graphs of the 2^{nd} joint executing a flexion-touch phenomena could be seen in our dataset: enter-drop, leave-drop, early movement and complete irregularity.

An enter-drop occurs if the angles drop below the hit level before returning to the hit level again (see figure 8). This phenomenon is called enter-drop because it occurs when the hit level is entered.

A leave-drop occurs if the angles drop below the hit level when the preparation should begin (see figure 8). This phenomenon is called leave-drop because it occurs when the hit level is left.

A graph with distinct enter-drops and leave-drops can be seen in figure 9.

Early movement of the 2^{nd} joint occurs if the beginning of the movement of the 2^{nd} joint precedes the movement of the 1^{st} by a substantial amount (see figure 9).

Some graphs of flexion-touches of our dataset show strong enter-drops and early movements of the 2^{nd} joint, e.g. the graph in figure 9. The movement can be described as follows:

1. The finger is in preparation level. The 2^{nd} joint is fully stretched.
2. The finger is flexed in the 2^{nd} joint (early movement) while the first joint stays in rest.
3. The flexing of the finger in the 1^{st} joint begins.
4. While the finger is still considerably far from the key, the 2^{nd} joint has already reached the minimum angle.
5. The finger is stretched in the 2^{nd} joint and flexed in the first joint until the finger reaches the key.

The described motion is not a correct execution of a flexing-touch. A beginning flexion-touch is aborted in favor of an extension-touch.

Some graphs of the 2^{nd} joint were so irregular that the preparation and hit levels could not be identified (see figure 10).

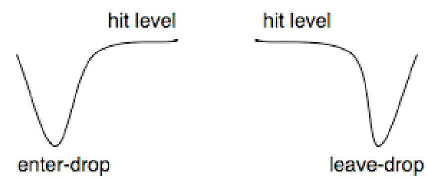

Figure 8. Enter-drop and leave drop

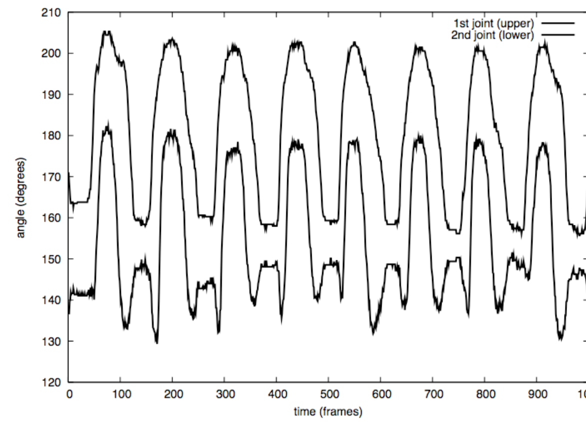

Figure 9. Enter-drops, leave-drops, and early movement

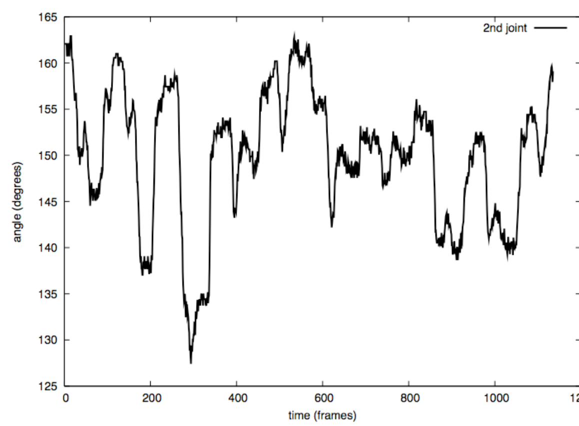

Figure 10. Complete irregularity

8. Conclusion

By tracking visual markers attached to players' fingers we calculated the angles in the joints and visualized them. For analyzing the graphs we introduced the E and T values that can be used for estimating the regularity of the movements. For the calculation of the E and T values, the graphs were segmented to preparation level, hit level, preparation way and hit way. Properties of these segments were defined.

The methods introduced in this paper help to interpret motion graphs and give a judgment about the regularity of the motion. Although our study can be expanded, for example by using a high frame rate camera, important tools for the analysis of the indirect piano touch were introduced and can serve as a basis for further research.

9. Future work

Our approach can be extended towards online generation of the graphical representations. If the graphs were generated in real-time they could serve as direct visual feedback about the regularity of the touches and could be used in a pedagogical setting.

If we could distinguish flexion-touches and extension-touches automatically and in real-time, this could be used to implement a special electronic piano. The flexion- and extension-touches would have different timbres. This piano could be useful for learning and teaching the different touches and as an instrument with an additional expressive parameter.

10. References

[1] Birkett, S., G. Comeau, M. Frize, C. Herry, R. Know, and P. Payeur, Challenging 300 years of piano teaching practices with 21st century technology: New teaching and research tools in piano pedagogy, Round table conducted at the Canadian University Music Society (CUMS) Conference, London, Ontario, 2005

[2] Camurri, A., P. Coletta, A. Massari, B. Mazzarino, M. Peri, M. Richetti, A. Ricci, and G. Volpe, Toward real-time multimodal processing: EyesWeb 4.0, AISB 2004 Convention: Motion, Emotion and Cognition, Leeds, 2004

[3] Dahl, S., Playing the accent – comparing striking velocity and timing in an ostinato rhythm performed by four drummers, Acta Acustica United with Acustica vol. 90, 2004

[4] Gat, J., The technique of piano playing, 2nd edition, Collet's Holding, London, 1965

[5] Goebel, W., R. Bresin, A. Galembo, Touch and temporal begavior of grand piano actions, Technical Report, Austrian Institute for Artificial Intelligence, Vienna, 2005

[6] Hadjakos, A., Computeranalyse des Leggieroanschlags, Diploma Thesis, HfMDK Frankfurt, 2007

[7] Herry, C., L., M. Frize, R. A. Goubran, and G. Comeau, Evolution of the surface temperature of pianists' arm muscles using infrared thermography. Paper presented at the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), Shanghai, 2005

[8] Marrin, T. and R. Picard, The Conductor's Jacket: a device for recording expressive musical gestures, Proceedings of the International Computer Music Conference (ICMC 1998), Ann Arbor, USA, 1998

[9] Ng, K. C., Music via motion: transdomain mapping of motion and sound for interactive performances, Proceedings of the IEEE vol. 92(4), 2004

[10] Schoonderwaldt, E., N. Rasamimanana, and F. Bevilacqua, Combining accelerometer and video camera: Reconstruction of bow velocity profiles, Proceedings of the International Conference on New Interfaces for Musical Expression (NIME06), Paris, 2006

[11] Young, D., The Hyperbow controller: real-time dynamics measurement of violin performance, Proceedings of the International Conference on New Interfaces for Musical Expression (NIME02), Dublin, 2002