SYSSOMO: A Pedagogical Tool for Analyzing Movement Variants Between Different Pianists

Aristotelis Hadjakos Erwin Aitenbichler Max Mühlhäuser
Telecooperation Group, Department of Computer Science, TU Darmstadt, Germany
E-mail: {telis, erwin, max} @tk.informatik.tu-darmstadt.de

Abstract

The visualization of arm movement can provide valuable additional information for piano teaching. However, movement visualizations are often difficult to understand for musicians. SYSSOMO uses score-following to synchronize two performances of the same piece in order to overlay motion data, MIDI, and video. Using this technique, the user can easily compare the differences between the performances of a teacher and a student.

To capture arm movement, sensors have to be unobtrusive for good user acceptance, and lightweight to minimize independent movement of the sensor. We have developed and built sensing hardware, called MotionNet, to capture arm movements. MotionNet is composed of a set of sensor units and a host unit, which communicate via CAN. This architecture helps us to reduce size and weight of the sensor units.

1. Introduction

The movement of the arm plays an important role in piano technique. The arm has to bring the finger, which executes the next touch, in an optimal position to strike the key. Arm movement can also be used to execute a touch while the finger remains slightly fixated to transduce the movement to the key.

We were able to identify piano playing variants that are distinguishable on gyroscope and accelerometer sensors attached at different positions of the player‘s arm. E.g., it is possible to distinguish variants of tremolo playing. We have published these result in [7].

In our experience, graphs of movement data are often difficult to understand for musicians. It is our approach to overlay two performances, for example the performance of a teacher and the performance of a student, so that users can easily see the differences.

The SYSSOMO system (SYnchronization through Score-following of SOmatic MOtion data) visualizes data captured by gyroscopes and accelerometers, which are attached to the user‘s arm. SYSSOMO has the ability to synchronize two performances and to overlay the motion data visualization. To support different learning scenarios, including distance learning, SYSSOMO provides further features: additional modalities, i.e., audio, MIDI, piano roll, and video, efficient storage, and annotation of performances.

The rest of this paper is organized as follows. In section 2 we discuss related work. The design of MotionNet is presented in section 3. In section 4, the synchronization mechanism of SYSSOMO and the additional features are discussed. A usage example is presented in section 5. Finally, the paper is concluded in section 6.

2. Related Work

2.1 Piano Teaching Systems

A variety of piano teaching systems have been proposed, including the famous Piano Tutor [4], the fingering suggestion and animation system by Lin et al. [9], the scale evaluation system by Akinaga et al. [1], the exercise generation and feedback system by Kitamura and Miura [8], the analysis and exercise generation system by Mukai et al. [12], the piano duo support system by Oshima et al. [14], and our recent sonification system [6]. SYSSOMO is a piano teaching system that uses performance visualization to provide an additional information source for piano tuition.

2.2 Performance Visualization Systems

Performance visualization systems process input data that is captured during the performance and visualize certain aspects of the performance.

Performance visualization systems need not to model “correct play” and are therefore not limited to the musical learning scenarios that the developer has foreseen. Because of the high level of detail, which is often provided by performance visualization systems, the interpretation of the visualizations can be difficult and require an advanced user. However, performance visualization systems can provide information that is otherwise unaccessible and can therefore be a valuable additional information source for instrument tuition.

2.2.1 Motion Data-based Performance Visualization Systems

Montes et al. [10] visualized surface EMG signals from relevant muscles to teach thumb touches. The EMG signals show different muscular activation patterns when comparing professional pianists with non trained persons.

Riley uses a system to capture MIDI, video, and surface EMG of finger muscles for teaching the piano [15, 16]. The system visualizes MIDI data as piano roll to visualize musical issues, video for frame-to-frame motion analysis, and surface EMG signals to identify unnecessary tension.

Mora et al. [11] developed a system that overlays a 3D mesh of a suggested posture on a video of the student's performance. The student can visualize the differences and adopt the suggested posture. To generate the 3D mesh, the posture of a professional pianist was recorded using motion capturing.

The 3D Augmented Mirror (AMIR) system [13] is a performance visualization system for string instruments. Markers that are attached to the instrument, the bow, and the player‘s body are tracked with a VICON. The AMIR provides several visualizations including a 3D rendering of the performance.

The visualization systems of Bouenard et al. [3] is a performance visualization system for timpanists. Markers that are attached to the drumstick and the player‘s body are tracked with a VICON. The system shows an animation of a 3D virtual character and provides several visualizations of kinematic and dynamic cues. Kinematic cues include plots of position, velocity, and acceleration as well as combinations of position-velocity and velocity-acceleration.

In contrast to other motion data based performance visualization systems, SYSSOMO provides the ability to synchronize two performances so that they can be compared easily.

2.2.2 MIDI-based Piano Performance Visualization Systems

The pianoFORTE [18] system uses MIDI input and visualizes tempo, articulation, and dynamics of the performance. The visualization is based on a music score, which is annotated to reflect dynamics and articulation of the performance. Tempo is visualized by a speed-o-meter. Missed and wrong notes are marked in the score.

The MIDIATOR [17] uses MIDI data to visualize note timings and note volumes. Note timings are shown in a piano roll representation and allow the analysis of rhythm, articulation, and tempo. It is possible to superimpose two performances. However, the two performances have to be played in exactly the same tempo.

The practice tool for pianists by Goebl and Widmer [5] generates visual feedback from MIDI input in real-time. The practice tool finds reoccurring patterns by autocorrelation and visualizes them. By placing successive patterns, represented over each other in a piano roll fashion, timing deviations can be seen. Additionally, the practice tool provides several other visualizations including a visualization of automatically extracted beats, a chord timing visualization and a piano roll overview.

SYSSOMO uses score-following to automatically synchronize two performances and can therefore, differently than the mentioned MIDI-based performance visualization systems, superimpose performances with tempo variations.

3. Sensors

We designed and built a custom sensing platform for SYSSOMO. Our sensors capture arm movements by inertial measurement, based on gyroscopes and accelerometers. MotionNet was developed according to the following requirements. We wanted our sensor system to be:

Unobtrusive: As the sensors have to be worn during piano performance, the sensors have to be as unobtrusive as possible to be accepted by users.

Lightweight: Since sensors can never be fixed perfectly to the arm, a low weight is essential. Otherwise the mass inertia of sensors would lead to uncontrolled, independent movements of the sensor itself, which has a negative impact on the accuracy of measurements.

Real-time capable: To support sonification of movements and live remote performance and teaching, the communication channel between the sensors and the host must have a high bandwidth and a low latency.

Configurable: The user should be able to balance wearing comfort and level of detail of the captured motion data. It should be possible to add or remove sensors easily.

Low cost: A large part of the intended target users, e.g., small musical schools or pianists with no professional ambition (yet) should be able to afford the system as well.

Because we could not find any system on the market that meets our requirements, we started building our own hardware. To date, wireless sensor platforms (e.g., Crossbow [19]) are in general not suitable for our application. Wireless communication bandwidth is too little, media access control (MAC) only scales up to a few nodes, and the message delay is significantly higher compared to wired sensor networks. In addition, such sensors must be powered by a battery that adds weight. Off-the-shelf inertial sensors (e.g., InertiaCube [20], Xsens [21]) are very expensive, i.e., about ten times more expensive than a sensor of our system. The InertiaCube is not based on a bus, which raises the question of how to connect many sensors to a host. The closest system is the Xsens MTx Xbus system. It would have been usable for our application, but was out of our budget. When sampling at 100 Hz, the Xsens sensors compare to our system as follows:

Feature / System	Xsens MTx	MotionNet
Max. number of sensors	10	80
Weight	30g	10g
Magnetic field sensor	yes	no
Power consumption	360mW	170mW
Cost (approx.)	EUR 1.750,--	EUR 150,--

3.1 Hardware

The hardware of the MotionNet system consists of an arbitrary number of *sensor units* and a *host unit*. All units are interconnected via a CAN bus (ISO 11898). Because CAN uses up to 29 bits for addressing, the maximum number of sensors is only determined by the total bandwidth of the bus (1 Mbit/s) and physical deployment considerations.

3.1.1 Host Unit

The *host unit* is used to connect a computer to the sensors. It powers the bus and all sensors and bridges from the CAN bus to RS232 (up to 921.6 kbps).

When the system is powered up, the host broadcasts a discovery message. Each sensor unit receiving this message sends a reply back to the host containing its identifier. This allows the host to create a list of sensors present on the bus.

During measurement, the host unit broadcasts sensor trigger messages at a rate of 100Hz. This allows to synchronize all sensors. The host unit then waits until all discovered sensor units have replied with their sensor values and sends the collected result to the computer via the serial interface.

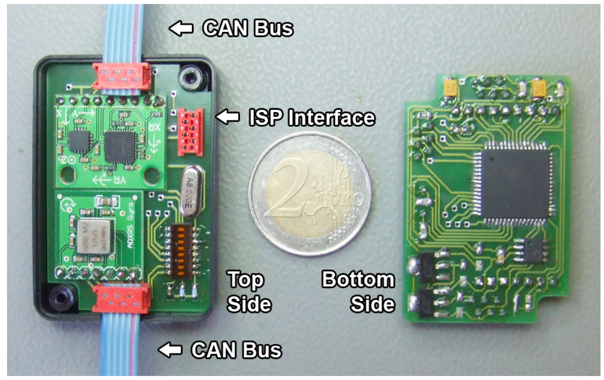

Figure 2: Sensor unit

3.1.2 Sensor Unit

Each sensor unit (Figure 1) has a unique identifier, which unambiguously describes the position where it has to be worn by the user. This identifier is configured with hardware DIP switches.

The central component of the sensor unit is an Atmel AT90CAN128 microcontroller. 3D acceleration is measured with an ADXL330 sensor and 3D rotation is measured with a combination of two gyroscope sensor chips: an IDG300 for the *x*- and *y*-axes and an ADXRS300 for the *z*-axis.

The sensor unit therefore provides six measurements (6-DOF) in total (Figure 2).

3.2 Mounting

Possible positions for the sensor units are the back of the hand, the wrist, and the upper arm. The user can choose which combination of sensor positions suits her best. We used a combination of Velcro fasteners and rubber bands to fixate the sensors to the arm.

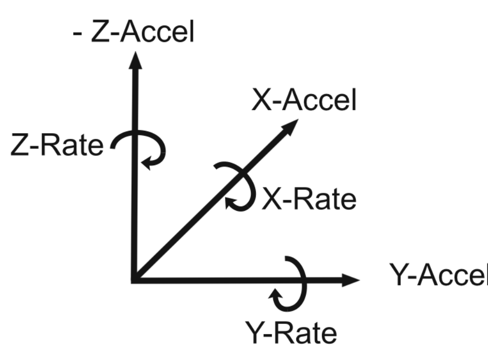

Figure 1: Measured axes

4. SYSSOMO

The following scenarios should be supported by SYSSOMO:

Single user: The user wants to analyze her playing movements with SYSSOMO. She might want to compare an old recorded performance with a new performance. Alternatively, she might want to compare her performance with the performance of her teacher.

Two users: Two users want to compare their performances and study the differences of their playing movements.

Distance learning: The student records motion data and sends it to her teacher over the Internet. The teacher compares the recording of the student to a reference recording, which the teacher has generated beforehand.

The following requirements can be derived from the described usage scenarios:

Visualization: Motion data should be displayed.

Comparison: The system should support the users to compare different performances.

Persistence: It should be possible to make recordings of the motion data of a performance and to load recordings at a later time.

Multimedia: The system should also support audio, video, and MIDI data.

Encoding: To allow transmission over the Internet, the size of the data to be transmitted has to be kept small.

Annotation: The system should allow to annotate motion data.

4.1 Visualization

SYSSOMO has four sources of input:

- Gyroscope and accelerometer data from the MotionNet
- MIDI data (we used a Kawai K-15 ATX MIDI-enabled upright piano)
- Video data (we used a Genius Slim 1322AF webcam)
- Audio data (we used an external microphone)

SYSSOMO provides the following visualizations (Figure 3):

- Motion visualization of the gyroscope and accelerometer data
- Piano roll of the MIDI data
- Video

The motion visualization shows the graphs of the gyroscope and accelerometer data of a sensor node. The zero-line is marked along with +/-50°/sec markings for the gyroscope and +/-1g for the accelerometer signal. The user can display several motion visualizations simultaneously for the different sensor units attached to the arm. For the interpretation of the graphs, it can be important to know when a note was played. Therefore, we included dots that represent the played notes at the bottom of the motion visualization. These dots also help the user to switch between the motion visualization and the piano roll.

4.2 Comparison and Persistence

SYSSOMO allows the user to save recordings of the performance. This includes the motion data, MIDI, audio, and video data. When a saved performance is opened, the user can control the visualization playback of the motion data, MIDI, audio, and video data with a control similar to a tape deck control (play, pause, fast forward, fast backward). The user can also hear the audio or the MIDI recording.

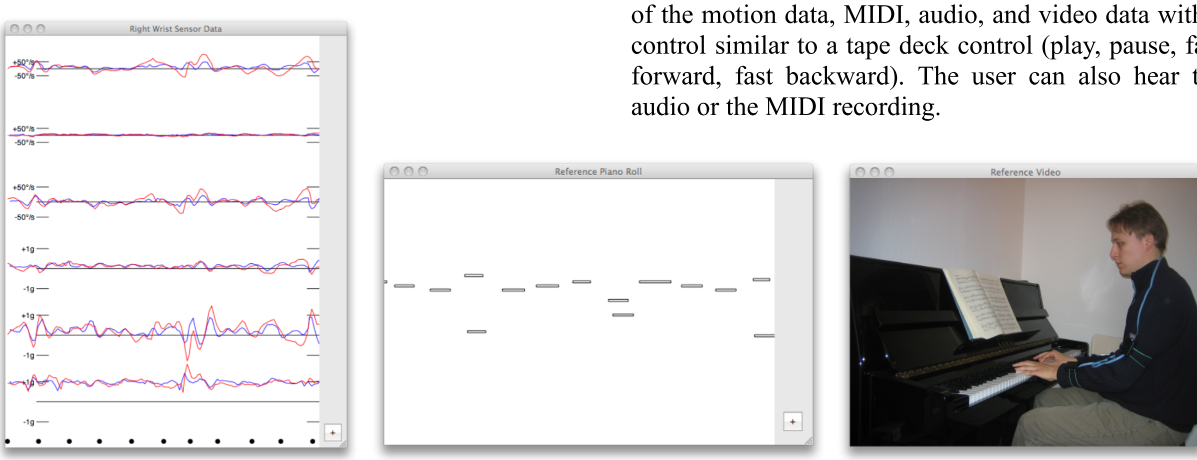

Figure 3: Motion data visualization, piano roll, and video

SYSSOMO provides the ability to align two performances of the same piece, facilitating their comparison. SYSSOMO distinguishes between the **reference performance** and the **dependent performance**: While the reference performance stays unchanged, the dependent performance is modified so that reference and dependent performance are synchronized.
To modify the dependent performance, SYSSOMO needs to compute a time map, which translates from dependent performance time to reference performance time. To this purpose, SYSSOMO aligns the MIDI data of the performances. SYSSOMO synchronizes the performances using Bloch's and Dannenberg's polyphonic score follower [2]. The algorithm matches MIDI events of the dependent performance to MIDI events of the reference performance. The timestamps of the matched MIDI events provide synchronization points between dependent and reference performance. The time map is completed by using linear interpolation between the synchronization points. Future developments could include audio-based score-following to support acoustic pianos without MIDI interface. By using the time map, SYSSOMO can synchronize motion data, MIDI, and video of the two performances. Future developments could include audio time-stretching of the dependent performance.

When a user plays the same passage at very different speeds, the playing movements usually have different forms, e.g., the size of the playing movements usually reduces with increased tempo. Furthermore, the time scaling of movement data introduces artifacts that are noticeable at very different tempi. The users should therefore normally use SYSSOMO to synchronize performances with comparable tempos.

4.3 Multimedia and Encoding

To enable transmission over the Internet, the size of the stored recording has to be kept relatively small. The data rate of the captured motion data is small compared to the amount of data for audio and video capturing so it did not seem necessary to compress the motion data. For video, the user can choose between H.264 and MPEG4 codecs with different resolutions. For audio, the user can choose between the high-fidelity ALAC codec and the AAC codec, which is available for different bitrates.

4.4 Annotation

The user can annotate the motion data visualization and the piano roll. An annotation can relate to a single point in time or a time interval. An annotation (Figure

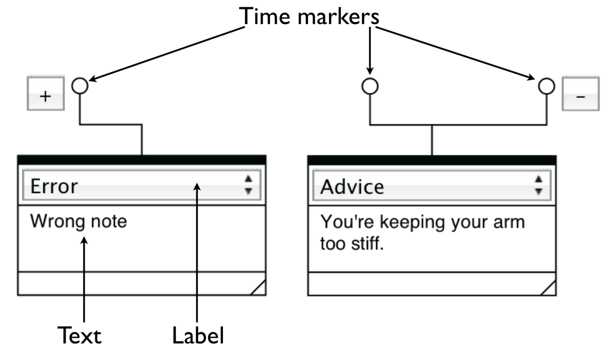

Figure 5: Annotations

4) consists of a label, like “Advice“ or “Question“, a text, and one or two time markers. A time interval is represented by two time markers. If the annotation relates to one point in time, a single time marker is used.

5. Usage Example

In the following, we want to demonstrate the interaction with SYSSOMO in a distance learning scenario.

The student wears the motion sensors and captures a performance with SYSSOMO. Using SYSSOMO, she plays back the performance: She hears the music, sees the motion data visualization, the piano roll, and the video. She uses the annotation tools to mark questions and any mistakes that she is aware of. She then sends the annotated performance to her teacher.

The teacher loads the student‘s performance as the reference performance. For comparison, the teacher loads her own performance of the same piece as dependent performance. Hearing the music, the teacher identifies technical problems of the student and uses the motion data visualization to study differences in the movement (see figure 5 for an example). The teacher then uses annotation to give advice. Finally, when the teacher is done annotating the performance, she sends back the annotated performance to the student.

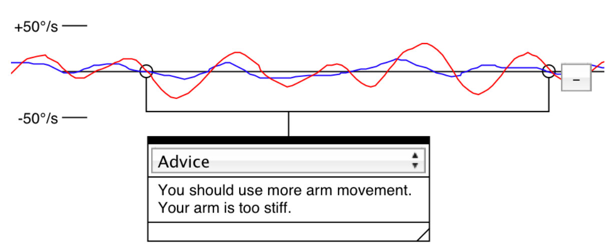

Figure 4: The signal shows the rotation of the arm. The teacher (red) uses more arm movement than the student (blue).

6. Conclusion

The visualization of arm movement can provide valuable additional information for piano teaching. However, movement visualizations are often difficult to understand for musicians. Therefore, SYSSOMO overlays two performances of the same piece so that the user can more easily identify differences of the arm movement. As we have shown previously [7], playing variants are visible on gyroscope and accelerometer signals of sensors that are attached to the pianist‘s arm.

To support different pedagogical scenarios, including distance learning over the Internet, further features were included in SYSSOMO: additional modalities (audio, MIDI, piano roll, video), efficient storage, and annotation of performances.

MotionNet provides SYSSOMO with rate and acceleration measurements of the arm movement. Separating functionality between sensor units and a host unit helps us to reduce size and weight of the sensor units, which is important for user acceptance and to minimize independent movement of the sensor units.

7. References

[1] S. Akinaga, M. Miura, N. Emura, and M. Yanagida. Toward realizing automatic evaluation of playing scales on the piano. In International Conference on Music Perception and Cognition (ICMPC), 2006.

[2] J. J. Bloch and R. B. Dannenberg. Real time computer accompaniment of keyboard performances. In ICMC 1985, 1985.

[3] A. Bouenard, S. Gibet, and M. M. Wanderley. Enhancing the visualization of percussion gestures by virtual character animation. In NIME 2008 Proceedings, 2008.

[4] R. B. Dannenberg, M. Sanchez, A. Joseph, R. Joseph, R. Saul, and P. Capell. Results from the piano tutor project. In Proceedings of the Fourth Biennial Arts and Technology Symposium, pages 143–150, 1993.

[5] W. Goebl and G. Widmer. Unobstrusive practice tools for pianists. In Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC9), 2006.

[6] A. Hadjakos, E. Aitenbichler, and M. Mühlhäuser. The elbow piano: Sonification of piano playing movements. In NIME 2008 Proceedings, 2008.

[7] A. Hadjakos, E. Aitenbichler, and M. Mühlhäuser. Potential use of inertial measurement sensors for piano teaching systems: Motion analysis of piano playing patterns. In Proceedings of the 4th i-Maestro Workshop on Technology-Enhanced Music Education, 2008.

[8] T. Kitamura and M. Miura. Constructing a support system for self-learning playing the piano at the beginning stage. In International Conference on Music Perception and Cognition, 2006.

[9] C.-C. Lin and D. S.-M. Liu. An intelligent virtual piano tutor. In Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications, pages 353–356, 2006.

[10] R. Montes, M. Bedmar, and M. S. Martin. EMG biofeedback of the abductor pollicis brevis in piano performance. Biofeedback and Self-Regulation, 18(2), 1993.

[11] J. Mora, W.-S. Lee, G. Comeau, S. Shirmohammadi, and A. E. Saddik. Assisted piano pedagogy through 3d visualization of piano playing. In HAVE 2006 – IEEE Intl. Workshop on Haptic Audio Visual Environments and their Application, pages 157–160, 2006.

[12] M. Mukai, N. Emura, M. Miura, and M. Yanagida. Generation of suitable phrases for basic training to overcome weak points in playing the piano. In Proc. of International Congress on Acoustics, 2007.

[13] K. Ng, T. Weyde, O. Larkin, K. Neubarth, T. Koerselman, and B. Ong. 3d augmented mirror: A multimodal interface for string instrument learning and teaching with gesture support. In ICMI ’07: Proc. of the 9th Intl. Conference on Multimodal interfaces, 2007.

[14] C. Oshima, K. Nishimoto, and N. Hagita. A piano duo support system for parents to lead children to practice musical performances. ACM Trans. Multimedia Comput. Commun. Appl., 3(2), 2007.

[15] K. Riley. New views of piano playing through motion analysis and midi technology. Journal of Technology in Music Learning, 2007.

[16] K. Riley, E. E. Coons, and D. Marcarian. The use of multimodal feedback in retraining complex technical skills of piano performance. Medical Problems of Performing Artists, 20(2), 2005.

[17] S. Shirmohammadi, A. Khanafar, and G. Comeau. Midiator: A tool for analyzing students’ piano performance. Revue de recherche en éducation musicale, 24:35-48, 2006.

[18] S. W. Smoliar, J. A. Waterworth, and P. R. Kellock. pianoforte: A system for piano education beyond notation literacy. In MULTIMEDIA ’95: Proceedings of the Third ACM International Conference on Multimedia, 1995.

[19] Crossbow Technology: Wireless Sensor Networks, http://www.xbow.com/, last visited: 04.07.08

[20] INITION: InertiaCube, http://www.inition.com/, last visited: 04.07.08

[21] Xsens Technologies: Xbus Kit, http://www.xsens.com/, last visited: 04.07.08