ANS EPRRSD - 13th Robotics & remote Systems for Hazardous Environments • 11th Emergency Preparedness & Response
Knoxville, TN, August 7-10, 2011, on CD-ROM, American Nuclear Society, LaGrange Park, IL (2011)

A COMPLETE MOTION AND MUSIC CAPTURE SYSTEM TO STUDY HAND INJURIES AMONG MUSICIANS

Neninger, C.R., Sun, Y.
Department of Computer Science and Engineering
University of South Florida
4202 E. Fowler Ave.
Tampa FL 33620
USA
cneninge@mail.usf.edu; yusun@cse.usf.edu

Lee, SH., Chodil, J.
School of Music
University of South Florida
4202 E. Fowler Ave.
Tampa FL 33620
USA
slee@usf.edu; jchodil@mail.usf.edu

ABSTRACT

This paper presents a complete motion and music capture system to study hand injuries among musicians and eventually to provide experimental results for researchers to better understand the hand and arm motions that cause and reflect musical performance skills,and the hand and arm injuries with great detail. The music performance is recorded from a hybrid YAMAHA N3 AvantGrand piano that not only produces the sound of an acoustic concert grand piano and has the same touch sensitivity and feel to the pianist, but also it has an electronic Musical Instrument Digital Interface (MIDI) that allows the musical performance to be recorded with computer through a USB cable. We integrated a motion capturing system composing of a 5DT DataGlove and a Natural Point passive marker-based motion capture system to capture the fine control of finger and arm motion of the musicians without any occlusion. The location of the markers are selected for easy computation of the joint angles of the shoulder, elbow, and the wrist. The 5DT DataGlove captures 14 degree-of-freedom (DOF) including two DOF for the thumb and three DOF for each of the rest of the fingers. Computer software is developed to capture totally 29 DOF data along with the MIDI data at 1000 Hz. A graphical rendering system is also developed to illustrate the motion.

Key Words: Motion Capture system. Musical performance. Biomechanical measurement.

1 INTRODUCTION

Motion capture systems provide an excellent way to model human motion. Modeling helps to extract and compare the components that conform the motion, making it easier to compare among subjects. The subject of motion modeling has become extremely useful in the medical field and rehabilitation aspects of medicine. Modeling helps investigators and subjects to understand what components represent most of the motion and how these components differ among healthy and injured subjects.

Musicians' play can in some cases result in irremediable injury to hand motion. In these cases, injury cause the professional career of musicians to halt, in many occasions indefinitely. The application of motion modeling to musicians' play is a very recent analysis method to study hand and arm motion. This paper presents a data collection setup to allow fine motion modeling in the future.

2 LITERATURE REVIEW

2.1 On Motion Capture Integration Systems

The motion capture of a musician's hand motion during performance is not a particularly new field. Numerous studies have been conducted, with similar abstract approaches to the one presented here. Goebl and Palmer reported a study involving motions of pianist in the search of the effect of digits feedback in piano performance on the accuracy of timing [1]. The study exemplified two musical sequences to emphasize the wanted motion. The motion capturing (MoCap) system in place was a passive marker system with 25 infrared markers glue onto the musician's right hand. The music was recorded using the MIDI output of an electronic instrument. Headphones were used to isolate the subject and reduce external noises that might affect the performance. Palmer, *et al* did another similar study that concerned the motion of clarinet's players [2]. The motion of the musician was captured using an active marker MoCap system, meaning the markers have unique identifiers to help distinguish their position and avoid the overlapping with other markers that may jeopardize their true positions. The system was then synchronized with the sound emitted from the instrument rather than the electronic output, such as the MIDI. Other studies lack the sound synchronization as well [3]. Engel, *et al.*, concluded a study involving musician's performance, but took exclusively the data from the electronic instrument (MIDI), and was concerned with the expected correlation of finger-note pairs, from which the motion was explained [4]. This study used the resultant MIDI output to deduce whether anticipatory sensations played a role in the timing or motion of the pianist.

The motion of the hand in the previous studies lack the proper level of precision and accuracy, due to the systems in place not being adept to capture finger-sized movements. Further, their sampling speeds were too slow to record accurately a continuous motion. Some studies that dealt with the capture of hand motion, especially during grasping, used DataGloves in one or both hands, both for its easy setup and accurate data acquisition [5][6]. In some cases ,the use of a DataGlove limited the natural movement of the subjects, but to a minimal degree.

2.2 On Marker Placement

Marker placement was a subject of extensive research to ensure the use of efficient methods, specifically for the purpose of rigid bodies. Kirk defines marker placement for grouping into rigid bodies in [7]. The study shows the location of the markers as superficial extensions to the bone structure underneath. Multiple markers offer better representation of the bone structure. Their position also plays a role in later bone identification. Triangulation of the markers position creates a distinct form that is visibly classifiable from each other, benefiting the clustering of the various rigid bodies [8]. This paper also explains how to extract rigid bodies based on skeleton pieces and recreate the entire skeleton from this information. Additional information on marker placement was obtained from the OptiTrack Documentation Manual, which explain the weaknesses of the particular system in use and placement for increased efficiency [9][10][11].

3 EXPERIMENTAL SETUP

We have integrated a MoCap system that is composed of three main devices: a hybrid piano with MIDI output; a motion capture system with infrared cameras and passive markers; and a right-handed DataGlove with 14 motion sensors. They are controlled by a single PC. A console program, once started, initializes all the devices and waits for each individual device to be ready before continuing. Once all the devices are ready, it polls all the devices, and consequently records the individual state onto different data files. The program was written in C++, within the Microsoft Visual Studio IDE. The utilization of such IDE provided an easier integration of all the library components necessary to interact with all devices.. The overall design of the program is described at the following high-level, abstract figure.

The main program basically created two additional threads to deal with each device independently. However, since they belong to a similar parent process, all child thread share basic resources, such as the time instance, meaning the system time calls for all threads is synchronized. The program takes advantage of this timestamp to synchronize the data, acquired independently for each device into a single time period. The resources and concepts for each device are explained subsequently.

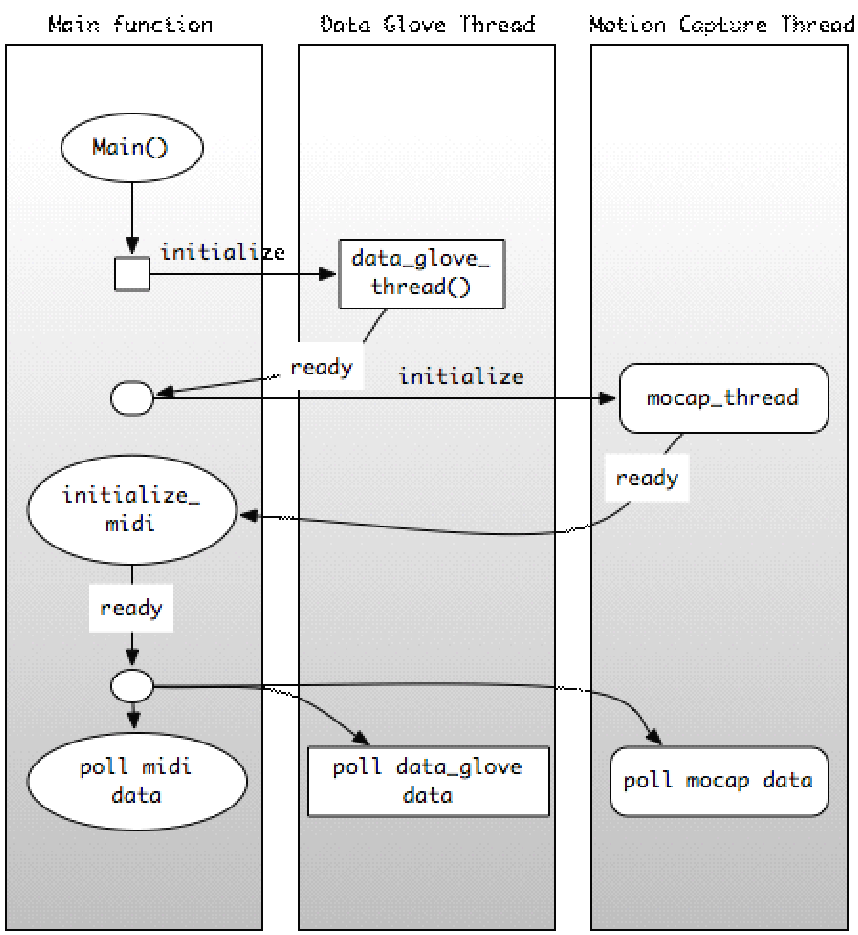

Fig. 1. Abstract design of the main.cpp source file, that polls data simultaneously from all devices, using the system's timestamp as a synchronization mechanism.

3.1 The AvantGrand Hybrid Piano and PortMIDI library

The YAMAHA AvantGrand, N3 version, is a state-of-the-art instrument. It is unique for its combination of classical components, such as the traditional key-hammer mechanism; and electronic components, which generates both the actual sound notes and the Musical Instrument Digital Interface (MIDI) data. The MIDI is a digitized computer language specifically designed to described musical notes based on a series of basic components such the note name, key-press and key-release times, and the speed of key-press, which are similar in various musical instruments. The MIDI data helps to obtain an electronic signature of the piano performance, and to guide the other devices to start or stop by polling data.

Within the midi_initialize and midi_poll functions in Figure 1, the program directly gets access to the PC's serial port that is connected to the piano and checks the port's buffer for any recent events emitted from the piano. The actual interface and system's calls are implemented through a cross-platform library, dedicated to MIDI I/O, called PortMIDI. PortMIDI is developed by PortMedia, an NSF-funded organization that specializes in providing an interface to MIDI-capable instrument using several computer languages.

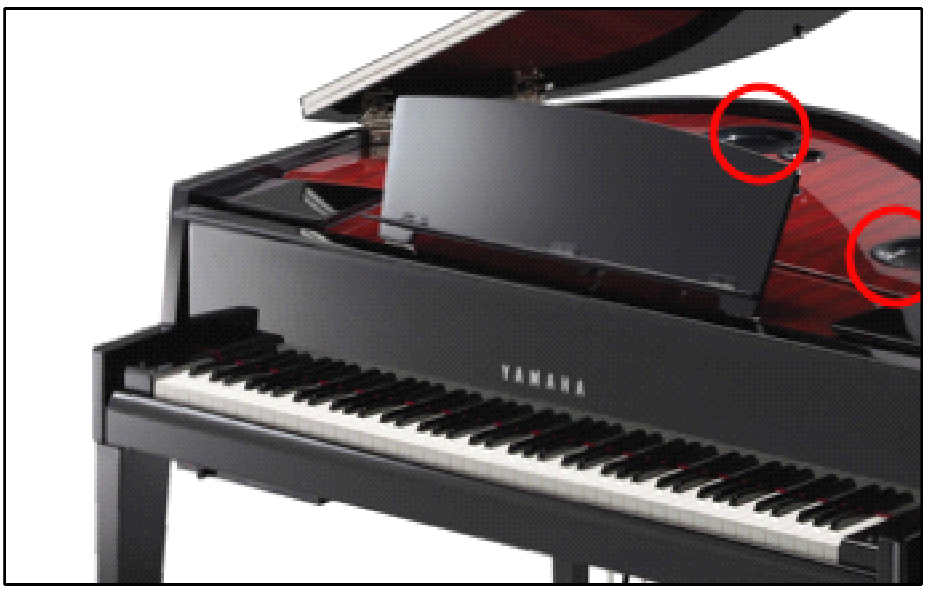

Fig. 2. YAMAHA AvantGrand, N3 version, hybrid piano, with electronic speakers highlighted

3.2 The 5DT DataGlove 14 Ultra and 5DT SDK

The motion of the subject's hand is captured with a 5DT DataGlove. The 5DT 14DOF DataGlove is a glove with open fingertips. Fiber optics run from the base of the glove to the every tip and between adajent fingers. The following shows a picture of the DataGlove placed on the hand of one of the subjects.

Fig. 3. The 5DT DataGlove 14 Ultra. Notice the encircled passive markers of the MoCap systems used to caspture the glove spacial location and orientation.

As noted by the glove's product name, it captures 14 of the 22 DOFs of the human hand [12]. Specifically, the glove's fiber optics capture the joint angles of the proximal interphalangeal articulations for each finger and the thumb, resulting in 5 DOFs; next, the metacarpophalangeal joints of each finger and the thumb are also captured, including the side-to-side motion of each finger, which is independently capture by an additional 4 sensors placed between the 4 fingers and the thumbs. The 5 metacarpophalangeal sensors add 5 DOFs to the model, which with the 4 inter-finger sensors complete the 14 DOFs model. The following table shows the joint captured by each sensor, while the figure shows the schematic view of the joints captured by the glove.

Table I. Sensor-Joint capture relationship

Sensor #	Joint	Sensor #	Joint
1	Thumb PIP	8	Middle MP
2	Thumb MP	9	Middle-Ring IMP
3	Thumb-Index IMP	10	Ring PIP
4	Index PIP	11	Ring MP
5	Index MP	12	Ring-Pinky IMP
6	Middle-Index IMP	13	Pinky PIP
7	Middle PIP	14	Pinky MP

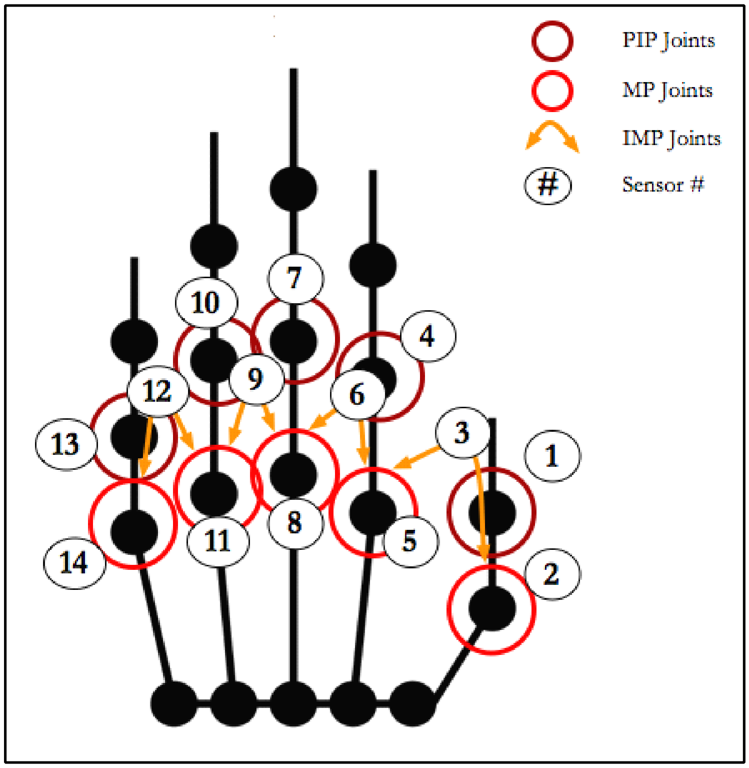

Fig. 4. Schematic showing the joints capture by the DataGlove: 5 Proximal Interphalangeal, 5 Metacarpophalangeal, and 4 Inter-finger Metacarpophalangeal, for a total of 14 DOFs.

The glove's manufacturer provides a DataGlove Manager Simulator, and a Software Development Kit (SDK) as an interface to get access to the glove's data in real time. Our program calls some of the SDK functions to first check for the connection of the glove to the system, and then initialize the glove, calibrate the glove input data, and store the data.

The glove calibration is done through a set of function calls, specifically fdSetCalibrationAll() or fdResetCalibrationAll() [13]. The raw values obtained from the glove are used to update the maximum and minimum normalized values, which are then used to normalize the entire data set. The normalized output is given by,

$$O = R - R_{min} / R_{max} - R_{min} \cdot V_{max} \qquad \text{for } R = \text{raw value}, \ V = \text{normalized value}$$

The values are then saved into an ASCII text file in double precision mode and normalized to [0,1] range. Although the data transfer frequency of the glove stands usually at 75 Hz, the polling rate set by the console program is given by 1 KHz. Each poll event is timestamped with the current system's clock, to provide an easier synchronization during the data's analysis with the other two devices.

3.3 The NaturalPoint OptiTrack Motion Capture System and NaturalPoint TrackingTools API

The DataGlove is integrated into a larger scale NaturalPoint OptiTrack MoCap system. The MoCap systems consists of six OptiTrack infrared cameras, model FLEXV100-R2, that capture the spatial position of a series of passive markers. As shown in Figure 5, the markers are positioned onto the subjects right shoulder, right arm and right hand, in a way they align with the underlying bone structure. Three markers are placed on the shoulder, aligned with the scapula bone; two are placed in the upper arm, aligned with the humerus bone; three markers are placed in a triangular fashion on the forearm, aligned with the ulna; another three are placed at the dorsal part of the hand, on top of the metacarpals. These three sets of markers are used to create a set of rigid bodies within the MoCap system interface, which allows not only the translational position, but also the relative orientation of each of the 4 sets. Altogether, the system provides 13 DOFs: three DOFs of the shoulder muscles, yaw, pitch and roll; one DOFs of the elbow, pitch; three DOFs of the wrist, also giving yaw, pitch and roll; and finally six DOFs related to the torso's movement, 3 translational and 3 rotational DOFs. The attachement of the markers to the subject are shown in Figure 5.

(a) (b) (c)

Figure 5. Passive makers position on the subject body. The 4 encircled groups of marker represent the 4 rigid bodies of the MoCap system, and they relate to the 4 main parts of the body which motion is being captured: the shoulder (c), the upper arm (b), the forearm (b) and the hand (a).

Overall, the integration of the MoCap and the DataGlove account for 13 DOFs and 14 DOFs respectively, bringing the total of DOFs capture by the system to 27 DOFs.

Figure 6. Cameras position with respect to the hybrid piano. Notice that only a portion of the cameras are shown, as some stand behind the position of the above photograph.

4 RESULTS

The collected data from our integrated system are independently stored in tabulated, ASCII text files, with each column defining a certain parameter or feature, and each row describing an event or observation.

4.1 MIDI Data Collection

MIDI data describes a number of parameters, some more particularly important than others. For the purpose of this study, the parameters collected concern only standard key events, and the variables attributed to a single event. MIDI events are divided as classes of events based on their *status byte*, which differentiates between key events, pedal events, and the like. The variables we are interested in are key events, namely *status byte* = 144. Normally, a MIDI outputs an event for every single touch in the piano. The key-on and key-off events are distinct from each other, with the main difference that key-off has 0 value for its velocity.

Table II. MIDI Data Parameters of Interest

Name	Description
Event #	The number of the event in the current execution
Time	The starting time for the event, in milliseconds
Status Type	The type of events is originated from the piano: 144 standard key note, 160 polymorphic aftertouch, 13X standard pedal hold
Note Name	The name of the note in the standard musical scale
Note #	The decimal number attributed to the note
Velocity	A 7-bit velocity scale value, where higher number denotes higher velocity
Duration	Duration of the key event (key-on event time - key-off event time), in milliseconds

For further analysis of the data and a graphical visualization of the data, a MATLAB program charts the duration and the velocity of each key event in the text file on a 2D coordinate system, with x-axis denoting time in seconds and y-axis the note name.

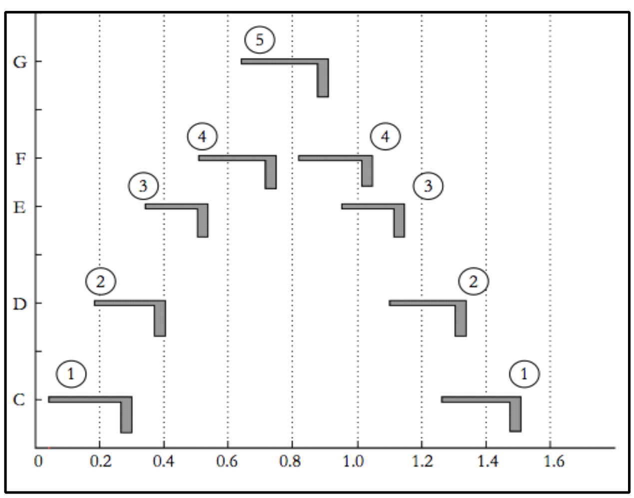

Figure 7. Graphical representation of the MIDI data for each key event: x-axis represent time in seconds and y-axis represent note name; each horizontal bar denoted duration, while each vertical bar is proportional to the velocity of the key event.

4.2 DataGlove Data Collection

The number of features described by the DataGlove are 14, as explained before. These relate to the angle values of the principal hand joints, after calibration and are normalized to 1. The data is once again stored in tabulated, ASCII text files, the first column denotes the time stamp in milliseconds, and the subsequent columns denote the features in the order they appear in Table I. A hand skeleton model was developed to simulate the motion of the hand based on the output of the DataGlove. The resultant model is fairly accurate in reflecting the full motion of the fingers, however, there are some mistakes in the range values for some of the joints, which extend beyond the range of 0 to 90 degrees.

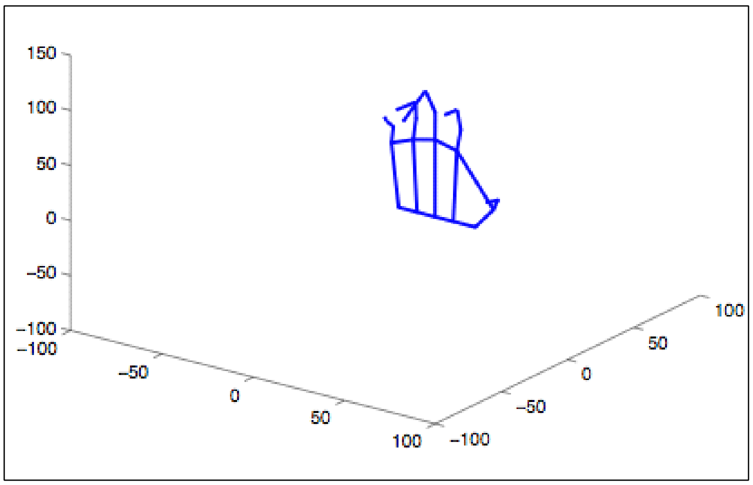

Figure 8. Using a fix hand model in 3D coordinates, the DataGlove angles are used to control the flexion of the joints, using pre-defined angle range for each joint: 90 degrees for all fingers MP, and PIP; 30 degrees for fingers IMP.

5 CONCLUSION AND FUTURE WORK

The experimental design for this study provides an unprecedented capture of 27 DOFs, for the entire right arm and hand during piano playing. The integration of this three devices, a MIDI-output capable piano, a 14-sensor DataGlove, and a passive marker motion capture system, describes the synchronization motion at several points of the performance, which makes the comparison between subjects an extremely comprehensive process.

The use of this integrated system has the purpose of conducting a through analysis of the motion of pianist. Part of the future analysis will include the evaluation of the error rate that the system outputs in comparison with optimal standards and actual measurements. Other extensions to the project will include an optimization of the hand model, with emphasis on improving its realistic display of DataGlove data.

6 ACKNOWLEDGMENTS

This study is made possible by a generous Neuroscience Collaborative Grant at the University of South Florida (USF), which was awarded by a competition in Motion Disorder category. The

study is ongoing with ten international artist-pianists, two dozen undergraduate and graduate pianists at the USF and other universities, and injured pianists who are referred to us by colleagues and physicians. We thank the participating pianists at all levels.

7 REFERENCES

1. Goebl, W., Palmer, C. "Tactile feedback and timing accuracy in piano performance." *Exp Bran Res*, 186:471-479. (2008)
2. Palmer, C., Carter, C., Koopmans, E., Loehr, J. "Movement, planning, and music: Motion coordinates of skilled performance." *The inaugural International Conference on Music Communication Science.* (2007)
3. Ferrario, V.F., Macri, C., Biffi, E., Pollice, P., Sforza, C. "Three-Dimensional Analysis of Hand and Finger Movements during Piano Playing." *Medical Problems of Performing Artists*, 22:18–23. (2007)
4. Engel, K.C., Flanders, M., Soechting, J.F., "Anticipatory and sequential motor control in piano playing." *Exp Brain Res,* 113:189-199. (1997)
5. Lin, J., Wu, Y., Huang, T. "Modeling the Constraints of Human Hand Motion." *Beckman Institute University of Illinois at Urbana-Champaign Urbana, IL* 61801. (2000)
6. Matsuo, K., Murakami, K., Hasegawa, T., Tahara, K., Kurazume, Ryo. "Measurement of static constraints imposed by a human hand on a grasped object." *Graduate School of Information Science and Electrical Engineering, Kyushu University.* (2009)
7. Kirk, A., O'Brien, J.F., Forsyth, D.A., "Skeletal Parameter Estimation from Optical Motion Capture Data." *University of California, Berkeley.*
8. Kiser, J. "Motion Capture Technology: The Basics." *C-Motion Inc.* (2010).
9. NaturalPoint. "Optitrak Manual: Hand Marker Placement." http://www.naturalpoint.com/optitrack/support/manuals/arena/HandMarkers.htm (2010).
10. NaturalPoint. "Optitrak Manual: Upper Arm Marker Placement." http://www.naturalpoint.com/optitrack/support/manuals/arena/UArmMarkers.htm (2010).
11. NaturalPoint. "Optitrak Manual: Upper Back Marker Placement." http://www.naturalpoint.com/optitrack/support/manuals/arena/UBackMarkers.htm (2010)
12. Soechting, J.F., Flanders, M. "Flexibility and Repeatability of Finger Movements During Typing: Analysis of Multiple Degrees of Freedom." *Journal of Computational Neuroscience*, 4:29–46. (1997)
13. Fifth Dimension Technologies. "5DT Data Glove Ultra Series: User's Manual." www.5DT.com. (2011)