Tai Chi: Training for Physical Tasks in Virtual Environments

Category: Research

Abstract

We present a wireless virtual reality system and a full-body training application built with the system. Our primary contribution is the creation of a virtual reality system that tracks the full-body in a working volume of 4 meters by 5 meters by 2.3 meters high to produce an animated representation of the user with 42 degrees of freedom. This, combined with a wireless audio/video broadcast, belt-worn electronics weighing under 3 pounds, and a lightweight head-mounted display, provide a wide area, untethered virtual environment system that allows exploration of new application areas. Our secondary contribution is our attempt, and failure, to show that user interface techniques made possible by such a system can improve training for a full-body motor task. We tested several immersive techniques, such as providing multiple copies of a teacher's body positioned around the student and allowing the student to superimpose his body directly on top of a teacher's body. Surprisingly, none of these techniques were significantly better than mimicking how Tai Chi is traditionally taught, where we provided one virtual teacher directly in front of the student.

CR Categories: I.3.7 [Three-Dimensional Graphics and Realism]: Animation—Virtual reality

Keywords: virtual environments, motion capture, training for physical tasks

1 Introduction

With recent advancements in wireless technology and motion tracking, virtual environments have the potential to deliver on some of the early promises of effective training environments and compelling entertainment experiences. In this paper, we describe and evaluate a system for training for a physical task, Tai Chi. The system is wireless, freeing the student from the encumbrance of trailing wires. The basic approach is to place a student of Tai Chi in a virtual environment with a virtual instructor, similar to how students currently learn Tai Chi in the real world by mimicking the motions of a teacher. Our overall hypothesis is that virtual environments allow us to take advantage of how physical activities, such as Tai Chi, are currently learned, while enhancing the learning environment in ways not possible in the real world. The student's motion is recorded with an optical motion capture system that allows real-time capture of whole body motion. Captured motion is used to animate representations of the student and teacher and displayed via a wireless head-mounted display (HMD), allowing students to have a natural representation of their body in the virtual environment and easily understood feedback on their movement patterns.

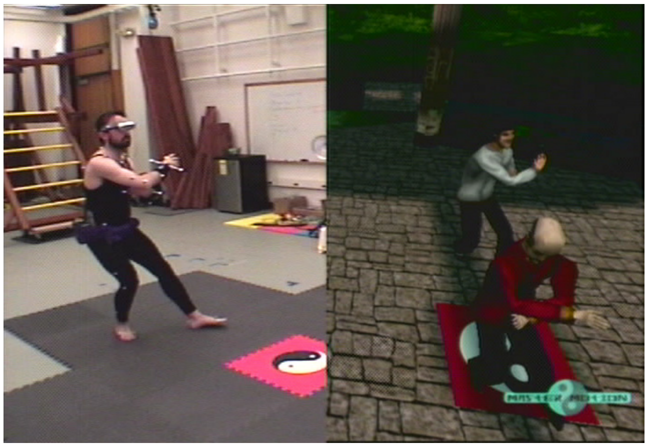


Figure 1: **Left: A student in the Tai Chi trainer. Right: The virtual world. Student is in white, teacher is in red**

We view this system as a first step towards effective and more general systems for training for physical tasks. Physical training is an economically important application with a significant impact in both health and safety. Tai Chi is itself important because there are demonstrated positive effects in preventing falls in the elderly and other health improvements [Wilson and Datta 2001].

We chose Tai Chi because it is composed of slow motions, making on-line feedback and correction possible. Tai Chi is a challenging training application because the sequence of movements, the *form*, is complicated and the standards for performance are exacting. Novices struggle to memorize the sequence of movements but more experienced students spend 20 years or more refining their movements. Because the emphasis is on balance and the shape of the body during slow movements, students are able to adjust their motion based on feedback during the sequence. Teachers in conventional Tai Chi classes will instruct a student to "drop the left shoulder" or "touch the foot only lightly to the ground" as the student is performing a motion.

Figure 1 shows a user wearing the wireless virtual reality equipment (a belt and an HMD) and the virtual environment with the real-time data applied to a student model. In the virtual environment, the student is rendered with a white shirt; and the master is shown wearing red.

Real-time, full body capture of the student's motion allowed us to experiment with different forms of visual feedback. We implemented several user interaction techniques, such as providing multiple simultaneous copies of the teacher surrounding the student, and allowing the student to superimpose his body directly inside the teacher's body. While the former could be done via projection or CAVE-like systems, the latter technique is not readily implemented without the use of an HMD. We compared various interfaces objectively by computing the error in limb position between student and teacher, and subjectively via post surveys of our subjects. Surprisingly, both the objective measures and subjective surveys show that none of the virtual environment enabling techniques did better than merely mimicking a real world situation, with one teacher in front of the student.

---

2 Background

Wireless virtual environments with full body capture have only become possible in the past few years. To our knowledge, ours is the first wireless, full body virtual environment system usable for a broad variety of applications. The only other wireless system of which we are aware was a highly specialized dismounted soldier training system that required the participant to wear 25 pounds of equipment [Molnar et al. 1997], severely limiting its applicability. The STRICOM system was constructed for combat training and used a General Reality CyberEye HMD and Premier Wireless for the transmittal of a bi-ocular NTSC video image. The motion of the soldier and rifle were captured with a Datacube-based optical system.

The technology for wireless whole body capture in real-time has been available for a number of years, first with magnetic sensors and more recently with commercial optical motion capture systems. Researchers have explored how to make magnetic systems less burdensome for the user by using fewer sensors and supplementing the data with inverse kinematics for the limbs whose motion is not directly measured [Badler et al. 1993; Semwal et al. 1998; Molet et al. 1996]. Technologies for capturing some of the motion of the user via video have also been developed [Maes et al. 1997].

With the system described in this paper, our goal has been to explore the use of virtual environments for training. A number of virtual environments have been developed in which the goal is to teach the student decision-making strategy for such situations as hostage negotiations, earthquakes, parachute descent and fire fighting [Ruisseau et al. 2000; Tate et al. 1997]. Fewer virtual environments have been built that attempt to train the user for physical tasks. Davis and Blumberg built a virtual aerobics trainer in which the user's actions were captured and recognized and he or she was instructed to "get moving" or given feedback such as "good job!" [Davis and Bobick 1998]. The user's silhouette was captured via infrared light and matched to a set of pre-recorded templates. They did not assess the effectiveness of the system as a trainer. Becker and Pentland built a Tai Chi trainer [Becker and Pentland 1996] in a modified version of the Alive system that provided better depth information. Their system used a hidden Markov Model to interpret the user's gestures and achieved a greater than 90% recognition rate on a small set of 18 gestures. When gestures are recognized but do not receive a high score, the system provides feedback by playing back the segment of the motion in which the student's motion differed most from that of the teacher.

Yang performed some preliminary user studies with his virtual environment called "Just Follow Me" [Yang 1999]. He used a *Ghost* metaphor to show the motion of a teacher to a subject. This form of visual feedback is similar to the layout that we called superimposition although the rendering styles differed. Subjects executed slow and fast hand motions in one of two test conditions: while seeing the teacher's motion displayed in a first person view or while viewing the teacher's motion on a third person video display. Although he did not quantitatively assess the results, qualitative assessment indicated that the virtual environment produced more accurate trajectories.

3 Training Environment

Our training environment uses the Vicon real-time optical motion capture system from Oxford Metrics. The recorded motion is filtered to reduce noise and recorded for off-line analysis. The students wear a wireless HMD through which they see a rendered virtual environment containing animated representations of both the student and the teacher. The head-mounted display is extremely light (3.4 oz), which is important for a task involving body posture.

Figure 2: **The virtual environment for Tai Chi**

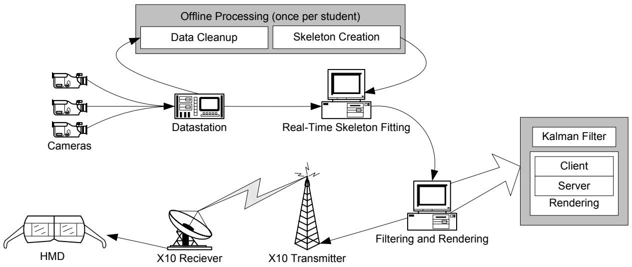


Figure 3: **Data pipeline and system diagram**

3.1 Virtual World

The virtual environment places the student in a pavilion on an island in the middle of a calm lake (Figure 2). On the horizon the sun is setting behind the mountains. The environment was designed to have a quiet and peaceful appearance as those are the circumstances in which Tai Chi should be practiced.

In the physical laboratory, floor mats were used demarcate the capture region for the optical motion capture system. To provide further continuity between the physical and virtual worlds, the virtual world has a Yin Yang symbol on the floor in the same place that the symbol appears in the real world painted on a floor mat. These physical and visual cues allowed the student to easily position him or herself in the appropriate part of the laboratory and virtual environment.

The student wears a Spandex unitard with reflective markers applied to the clothing and skin. The student also wears a fabric hip belt that holds the batteries and the receiver for the HMD. These weighed 3 pounds in total. We used a marker set with 41 14 mm markers that is an adaptation of a standard biomechanical marker set. Markers were placed at every major joint and along every major bone, and additional markers provided redundant information for critical body segments (head, hands, and feet). When subjects begin the experiment, they are asked to perform a calibration motion in which they move each joint through its full range of motion. This system uses this information to compute a 17 bone skeleton with bones of the correct length for that student.

3.2 Real-time Optical Motion Capture

Twelve Vicon cameras were placed around the laboratory to provide good coverage of a 4 m×5 m capture region that is 2.3 m high. The lens of each camera is surrounded by infrared emitters that reflect IR light off of the retroreflective tape on the markers. The cameras capture 1024×1024 bitmap images at 60 fps. The images show the markers as white dots on a black background. These locations on the image plane of each camera are converted to a set of points in $\mathbf{R}^3$ by the Oxford Metrics Vicon 512 Datastation hardware.

The locations of the markers are sent to the Oxford Metrics Tarsus real-time server which fits the student's skeleton to the three-dimensional marker locations. The fitting process runs at between 30-60 fps on a 1 GHz Pentium III. The variability in frame rate appears to be caused by occasional spurious "ghost markers" and large changes in the data. The real-time server provides our filtering process with world-space position and orientation vectors for each of the 17 bones.

3.3 Filtering

After the skeleton has been fit to the real-time data, the position and orientation of each limb is filtered to reduce noise. Because the fitting process occurs in real time and must be done very quickly, two types of errors occur. The first is a gross error due to marker mislabeling which causes bones to "flip" out of position. The other error is a general jittering of the skeleton that results from inaccuracy of the fitting process and marker position reconstruction. Flipping errors, which usually occur during high frequency motion, are rare in our system since Tai Chi motions tend to be slow. Jittering, however, is noticeable regardless of the speed of the student's motion and needs to be corrected.

The mean residual error of markers in the fitting process was between 3.9 mm and 7.2 mm for our students, which is enough to cause distracting movement in the virtual environment. In particular when viewing the environment in first-person with the camera rigidly attached to head, those small errors become large changes in the view.

In our analysis of the real-time data, we found that our system could be considered a linear stochastic system and the process/measurement error in out system approximately followed a Gaussian distribution, thus we implemented a Discrete Kalman filter [Welch and Bishop 2001; Murphy 1999]. Sul and his colleagues also used a Kalman filter to post-process motion capture data to smooth the motion and maintain the kinematic constraints of a hierarchical model [Sul et al. 1998]. We, however, chose to ignore the bone length constraints during the filtering process. This still gave us acceptable visual results and avoided any extra lag due to the processing time needed to satisfy the constraints.

The Discrete Kalman filter requires the input signal to be in a linear space. This constraint was not a problem for our position vectors (which are vectors in $\mathbf{R}^3$), but it was a problem for our orientations (quaternions representing points in $\mathbf{S}^3$). In order to apply this filter to the orientation data, we linearized the orientation space using a local linearization technique [Lee and Shin 2002]. Once our orientations were in a linear space ($\mathbf{R}^3$), we applied the Discrete Kalman filter in the same way we did with the position vector. After filtering, the orientation vector was converted back to a unit quaternion.

Because the virtual camera was rigidly mounted to the student's head, jitter that may have been unnoticeable on other parts of the body was magnified and resulted in an unacceptably shaky first-person view. To correct this, we applied a more aggressive set of filter parameters to the head than to the rest of the body. Thus making the first-person point of view stable while not sacrificing the agility of the rest of the body.

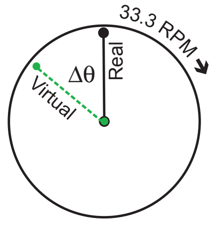


Figure 4: **Latency measurement technique (top-down view of turntable)**

3.4 Rendering

The rendering was performed by the Lithtech 3.1 game engine using a client/server architecture. The server acted as an intermediary, broadcasting commands to control the layout of students and teachers in the environment and the playback position and speed. The server also recorded data for both the virtual teacher and student for off-line analysis. The client took the data from the filtering application, mapped it onto a character, and rendered the virtual world with the teacher and student models. The client, server, and the filtering process all resided on a 733 MHz Pentium II with a GeForce3 graphics accelerator. The rendering speed averaged 17 fps, and was always at least 14 fps.

After the scene is rendered, a scan converter converted the signal from VGA to NTSC. It was then sent to an off-the-shelf X10 wireless video transmitter that broadcast the video from a location 3 m above the student's head in the center of the capture area. This signal was received by an X10 video receiver on the student's belt, and then sent to the Olympus Eye-Trek glasses we used for our HMD. The rendering was bi-ocular, with the same image being sent to each eye via separate LCD screens. The Eye-Trek had a 37.5° horizontal and 21.7° vertical field of view, each screen was 640×360 pixels, and it weighed 3.4 oz. Because a lightweight HMD was a requirement for this application, we had to compromise on spatial resolution of the display itself. Therefore, our decision to degrade VGA to NTSC in order to broadcast it implied little or no further loss of resolution. An important detail in rendering the image for the student was to block out other light sources, so that the virtual environment was the only thing the subject saw. Rather than using a cowling on the HMD, we found that turning off the lights in the laboratory was just as effective and far more comfortable for the subjects.

3.5 Latency

The latency of the motion capture and rendering process was 200 ms. To measure our end-to-end system latency, we put markers on a record player revolving at $33\frac{1}{3}$ RPM to capture the motion of the turntable and then added representations of those markers to our existing Tai Chi environment (Figure 4). We then projected our display onto the physical turntable as it rotated. The angle difference $\Delta\theta$ between the projected and real markers is proportional to latency such that $latency = k\Delta\theta$ where $k = 60s/(33\frac{1}{3} * 2\pi) \approx 0.286$. Our measured $\Delta\theta$ was .70 radians, giving us an end-to-end latency of 200 ms.

A latency of 200ms is high for an interactive virtual environment, but the effect on the student's performance was somewhat mitigated by the slow nature of movements in Tai Chi. In addition, the students tended to lag behind the teacher by an average of 610 ms as described in the Results section.

4 Assessment

In order to evaluate the effectiveness of our system for learning Tai Chi, we tested 40 volunteers to analyze how well they learned the motions of a previously recorded teacher, given different user interfaces involving the teacher and student in the virtual space.

4.1 Design Considerations

There is a large potential design space for how to represent the student and teacher in a full body motion training task. We made the constraining assumptions that the student would always see a first-person view, that the teacher and student would be represented in a fairly realistic way (both would remain humanoid, be of the correct proportions, etc.), and that the feedback that the students received would be entirely visual and based on body positions alone.

Given these assumptions, the variables we identified were:

- Number of copies of the teacher
- Number of copies of the student
- Orientation and placement of the teachers and students
- Rendering styles of the teacher/student

Our first decision was to fix the orientation of the teacher and student so that they always faced that same direction. This decision was based both on early pilot studies and the fact that traditional Tai Chi classes are taught this way. After a number of pilot experiments, we identified five representative points in the design space, pictured in Figure 5. These conditions are:

- **One on One**: One teacher with one student
- **Four Teachers**: Four teachers surrounding one student
- **Side By Side**: Four teachers standing next to four students with one student in the center.
- **Superimposition 1**: Five normally rendered students with a red wireframe teacher superimposed on them.
- **Superimposition 2**: Five wireframe and transparent green students with a red stick figure teacher superimposed inside them.

4.2 Experiment Design

The majority of our 40 volunteers were college students with no previous formal motion training. Each subject experienced four of the five layouts. In each layout, the subjects were asked to match the teacher as they performed twelve repetitions of the same motion after once watching the motion from any viewpoint the student chose. Each layout could be associated with one of four different motions. The order in which subjects experienced the layouts was randomized to minimize learning effect, however, the first layout of each test session was always paired with motion 3. If there had been a learning effect from one motion to the next, this would have allowed us to discard motions 1, 2 and 4, and only analyze data for motion 3.

All four motion animations were segments of the same Tai Chi form performed by a Tai Chi teacher and captured in an offline motion capture session. Each motion was approximately twenty seconds long and had distinguishing characteristics. Motion 1 featured simple hand movements and a 90 degree turn to the right; motion 2 had little hand motion, a 180 degree turn of the feet, and some turning of the upper body; motion 3 featured slow hand movements, large movement of the feet, and some turning the thorax; and motion 4 featured swift hand movement and turning of the thorax and hips but little movement of the feet. Based on participants' error measurements, we can say that all motions were of approximately equal difficulty except for motion 1 which was significantly easier (Figure 6(c)).

4.3 Results

4.3.1 Error Measurement

In order to measure how closely the student mimicked the teacher's motion we use 12 bones for comparison: the upper and lower arms, hands, upper and lower legs, and feet. Each bone is defined using the skeletons hierarchy with a parent end and a child end. For example the upper arm has a parent end corresponding to the shoulder and a child end corresponding to the elbow. We measure the error for each bone as the difference between teacher and student child end positions, once the parent end positions and bone lengths have been normalized. Given student bone end positions S_p and S_c and teacher bone end positions T_p and T_c the error E for that bone is:

$$E_{limb} = \left\| \frac{S_c - S_p}{\|S_c - S_p\|} - \frac{T_c - T_p}{\|T_c - T_p\|} \right\|$$

We then average the errors over all limbs to get the total error for one frame of data: $\overline{E}_{frame} = \sum_{i=1}^{i=12} E_i/12$. To obtain the total error for one repetition of the motion, we sum over the n frames in that repetition: $\overline{E}_{rep} = \sum_{i=1}^{n} \overline{E}_i/n$.

Finally, though there are 12 repetitions for each motion, we only consider the last four that the students performed. Before the 9th trial, subjects were reminded that there were four repetitions remaining, and encouraged to do their best on the remaining trials. Averaging over these four gives a reasonable indication of how well the student learned the form. So the total error, for one student, given one layout and one motion, is:$\overline{E}_{trial} = \sum_{i=9}^{i=12} \overline{E}_i/4$

4.3.2 Static Pose Matching

We gave students unlimited time to match the beginning pose of the teacher. Once they were satisfied with this pose, they told us they were ready to begin. This experiment allowed us to see how well students were able to match a static pose for each of they layouts they saw.

Figure 6(a) shows the error in matching a static pose, keeping layout constant and averaging over all students, all motions and all repetitions of the motion. We can say with statistical confidence that Superimposition 2 gets smaller error values than One on One or Four Teachers.

4.3.3 Data Analysis Details

The error measurement for full motions is slightly more complicated. There are two major sources of error which are common for all students who participated in the experiment: time lag and yaw offset. If the student is not oriented in precisely the same way as the teacher, the error in yaw will add an additional error to all bones in the student's body. Even if the student performs the form perfectly, if there is a few degree offset in the orientation the student will have a large error. To minimize this, we consider the student's starting pose and find the yaw offset that minimizes this static pose error. We found that when considering all possible offset angles, from -30° to 30° there was one clear point where error was minimized (Figure 7(a) shows a typical case).

In addition to the yaw shift, as the students performed the motion they all tended to follow behind the master to some degree. Because they were still doing the motion correctly, just not at the same time

(a) One on One

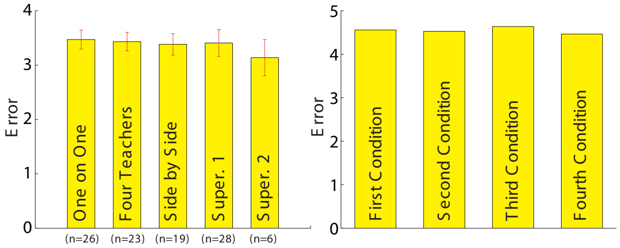
(b) Four Teachers

(c) Side by Side

(d) Superimposition 1

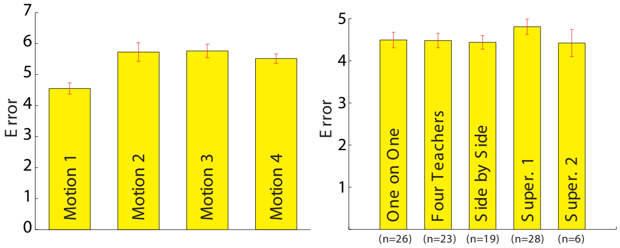
(e) Superimposition 2

Figure 5: **The five conditions**



(a) Static Pose Matching

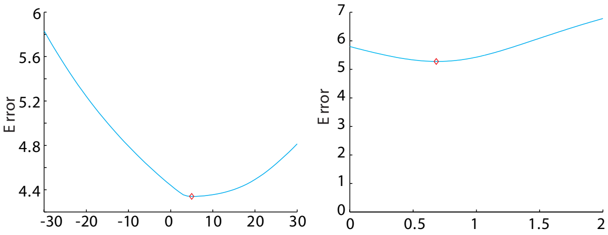
(b) Learning Effects



(c) Motion Difficulty

(d) Layout Comparison

Figure 6: **Error analysis results**



(a) Yaw Shift (degrees)

(b) Time Shift (s)

Figure 7: **Change in error under different amounts of yaw and time shifting**

as the master, we searched for a time shift that would minimize the error. For each repetition of the motion, we considered all possible time shifts from 0 to 120 frames (0 to 2 seconds). For each value, we compared the shifted student data with the original teacher data to get the error value. Again we found that there was one clear shift value which minimized the error (Figure 7(b) shows a typical case). The average time shift was approximately 610 ms.

We also needed to make sure there was no learning or fatigue effect. Figure 6(b) shows that when we averaged all conditions performed first, second, third, and fourth, the order of performance had no effect on error. Before we could average results from different motions we needed to compensate for different difficulties in the motions. Figure 6(c) shows that on average, motion 1 was significantly easier than the other motions. For our final analysis, we scaled the error values to normalize this difference in difficulties.

Having dealt with these issues, we can now compare all of the

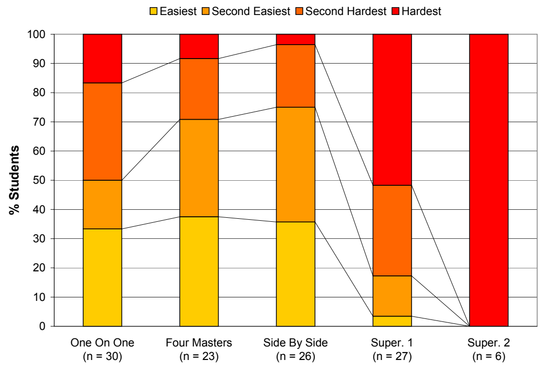


Figure 8: **Survey Results: Relative difficulty of each motion**

layouts with each other. Figure 6(d) shows the overall result. The only statistically significant conclusion we can draw is that Superimposition 1 is worse than the others by a small amount. One on One, which most closely mimics a traditional teaching environment, is as good as the other techniques.

4.3.4 Survey Results

After each experiment ended, we gave a survey to the subjects. We asked them to rank the four layouts they experienced, from easiest to hardest. As Figure 8 shows, the two Superimposition trials were considered to be significantly harder than the other trials. In fact, all of the people who tried Superimposition 2 thought it was the most difficult. Surprisingly, although participants considered Superimposition 2 very difficult compared to the other layouts, their error was approximately the same as the non-superimposed layouts.

5 Conclusions

Our primary contribution is the engineering of a wireless virtual reality system that tracks 41 points on the human body in a working volume of 4 meters by 5 meters by 2.3 meters high. Our system has a positional accuracy of roughly 5mm, a latency of 200ms, and maintains a rendering frame rate of 17fps using a commodity PC. Using a 3.4 ounce sunglasses-style HMD and a 3 pound waist pack, ours is the first wireless virtual environment system useful for a large variety of applications.

To our surprise, none of the layouts of students and teachers had a substantial effect on learning a motor control task involving full body motion. This result contradicts the initial intuition of virtual reality researchers we have discussed this project with, who generally assumed superimposition would make the task much easier. If other studies come to similar conclusions for other motor control tasks, the virtual environment research community may need to more carefully examine which tasks are best suited for immersive training, because full body motion motor tasks are often cited as example applications.

We have several theories about why our virtual environment training did not improve performance. In our experiments, students were expected to perform motor movements while simultaneously watching the movements of the teacher's avatar. Other studies of simpler movements have indicated that this simultaneous activity may interfere with learning [Pomplun and Mataric 2000; Schmidt and Wulf 1997]. Pomplun and Mataric compared the performance of subjects making arm movements when the subjects were instructed to rehearse while watching a video presentation of the motion and when they were instructed to just watch the video before performing the task. The subjects who did not rehearse performed substantially better. A number of our subjects sat out some of the trials in the middle of each session and just watched the performance of the teacher, perhaps unconsciously recognizing that watching and rehearsing simultaneously was not an effective learning strategy. With the exception of the superimposition layouts, our rendering styles were chosen to represent the student and teacher in as realistic a fashion as possible given the limited polygon count possible in a virtual environment. However, experimental evidence indicates that subjects tend to track the hands when watching or imitating a teacher performing arm movements [Mataric and Pomplun 1998]. If these results generalize to Tai Chi, a rendering style that emphasized the position and orientation of the hands might be more effective.

We evaluated the motion of the students during trials 9-12 in the virtual environment. A better evaluation technique would have been to remove the visual stimulus of the images of the teacher and student and assess the student's performance on an independent performance [Schmidt and Wulf 1997]. In addition to assessing retention in the absence of feedback, this assessment technique would have allowed us to compare the effectiveness of more standard visual presentations such as a video of a teacher projected on the wall of the laboratory and perhaps understand whether the small field of view and low resolution of the HMD was affecting the student's ability to learn in the environment.

We can not conclusively rule out that our virtual environment was at fault; one can always worry that if only our fidelity had been better, the latency had been lower, or we had just been clever enough to try another rendering style or user interface, the results would be different. However, our system is substantially better in both tracking area and spatial resolution than most, tracks the position of the entire body well, and addresses what has been identified as one of the biggest remaining problems: the umbilical cord [Usoh et al. 1999]. Although our latency is admittedly large, we chose a task where latency should not be important, and also measured our errors with a technique that would account for the students' latencies in responding to the teacher. Therefore, we can at least conclude that if virtual environments are to be helpful for this task, they must either be at higher quality that we have engineered, or utilize better user interaction techniques than we devised, or both.

References

BADLER, N. I., HOLLICK, M., AND GRANIERI, J. 1993. Real-time control of a virtual human using minimal sensors. *Presence 2*, 82–86.

BECKER, D., AND PENTLAND, A. 1996. Using a virtual environment to teach cancer patients to t'ai chi. In *Proceedings of International Conference on Automatic Face and Gesture Recognition*.

DAVIS, J. W., AND BOBICK, A. F. 1998. Virtual PAT: A virtual personal aerobics trainer. In *Workshop on Perceptual User Interfaces*.

LEE, J., AND SHIN, S. Y. 2002. General construction of time-domain filters for orientation data. To appear in IEEE Transactions on Visualization and Computer Graphics, available at http://wings.graphics.cs.cmu.edu/jehee/Papers/TVCG2001.pdf.

MAES, P., DARRELL, T., BLUMBERG, B., AND PENTLAND, A. 1997. The ALIVE system: Wireless, full-body interaction with autonomous agents. *Multimedia Systems 5*, 2, 105–112.

MATARIC, M. J., AND POMPLUN, M. 1998. Fixation behavior in observation and imitation of human movement. *Cognitive Brain Research 7*, 2, 191–202.

MOLET, T., BOULIC, R., AND THALMANN, D. 1996. A real-time anatomical converter for human motion capture. In *EGCAS '96: Seventh International Workshop on Computer Animation and Simulation*, Eurographics.

MOLNAR, J., ANSCHUETZ, R., JONES, C., AND KELLY, P. 1997. Immersion of a live individual combatant into a virtual battlespace. In *The 14th Annual AESS/IEEE Dayton Section Symposium*, 27–34.

MURPHY, K., 1999. Kalman filter toolbox for Matlab. Available at http://www.cs.berkeley.edu/ murphyk/Bayes/kalman.html.

POMPLUN, M., AND MATARIC, M. J. 2000. Evaluation metrics and results of human arm movement imitation. In *Proceedings, First IEEE-RAS International Conference on Humanoid Robotics (Humanoids-2000)*.

RUISSEAU, J. Y., TODESCHINI, L., AND GORZERINO, P. 2000. Learning and training simulator virtual reality development for parachute descent. In *Proceedings of the IEA 2000/HFES 2000 Congress*, 527–529.

SCHMIDT, R. A., AND WULF, G. 1997. Continuous concurrent feedback degrades skill learning: Implications for training and simulation. *Human Factors 39*, 4, 509–525.

SEMWAL, S., HIGHTOWER, R., AND STANSFIELD, S. 1998. Mapping algorithms for real-time control of an avatar using eight sensors. *Presence 7*, 1, 1–21.

SUL, C., JUNG, S., , AND WOHN, K. 1998. Synthesis of human motion using Kalman filter. In *CAPTECH*, 100–112.

TATE, D. L., SIBERT, L., AND KING, L. T. 1997. Virtual environments for shipboard firefighting training. In *Proceedings of the 1997 Virtual Reality Annual International Symposium*.

USOH, M., ARTHUR, K., WHITTON, M. C., BASTOS, R., STEED, A., SLATER, M., AND BROOKS, F. P. 1999. Walking > walking-in-place > flying, in virtual environments. In *Siggraph 1999, Computer Graphics Proceedings*, Addison Wesley Longman, Los Angeles, A. Rockwood, Ed., Annual Conference Series, ACM Siggraph, 359–364.

WELCH, G., AND BISHOP, G. 2001. An introduction to the Kalman filter. In *ACM SIGGRAPH 2001 Course Notes*.

WILSON, C. J., AND DATTA, S. K. 2001. Tai Chi for the prevention of fractures in a nursing home population: An economic analysis. *Journal of Clinical Outcomes Management 8*, 3.

YANG, U. 1999. Just Follow Me: An immersive vr-based motion training system. In *International Conference on Virtual Systems and Multimedia*.

Tai Chi: Training for Physical Tasks in Virtual Environments

Philo Tan Chua Rebecca Crivella Bo Daly Ning Hu Russ Schaaf
Todd Camill Jessica Hodgins Randy Pausch
Entertainment Technology Center, CMU

Category: Research
Format: print

Contact: Jessica Hodgins
CMU, Robotics Institute, 5000 Forbes Ave, Pittsburgh, PA 15213-3890
phone: (412) 268-6795
fax: (412) 268-6436
email: jkh@cs.cmu.edu

Estimated # of pages: 7

Keywords: virtual environments, motion capture, training for physical tasks

We present a wireless virtual reality system and a full-body training application built with the system. Our primary contribution is the creation of a virtual reality system that tracks the full-body in a working volume of 4 meters by 5 meters by 2.3 meters high to produce an animated representation of the user with 42 degrees of freedom. This, combined with a wireless audio/video broadcast, belt-worn electronics weighing under 3 pounds, and a lightweight head-mounted display, provide a wide area, untethered virtual environment system that allows exploration of new application areas. Our secondary contribution is our attempt, and failure, to show that user interface techniques made possible by such a system can improve training for a full-body motor task. We tested several immersive techniques, such as providing multiple copies of a teacher's body positioned around the student and allowing the student to superimpose his body directly on top of a teacher's body. Surprisingly, none of these techniques were significantly better than mimicking how Tai Chi is traditionally taught, where we provided one virtual teacher directly in front of the student.