Experiments in Human Motion

Ekaterina Mirochnik
eim@cs.cmu.edu

Abstract

The application of motion capture technology is quite useful and adaptable for gathering information about the way things move naturally. In this paper I will discuss two research projects that utilized a motion capture system. The first, “Interference”, studied how a human being responds to visual stimuli that contradict a suggested action. The second, “General Captures”, explored and developed an adaptable method to share captured motion data for use in personal projects or research. Motion capture does have its limitations; this paper will discuss several notable ones encountered throughout the course of these projects.

Motion Capture Overview

The Carnegie Mellon University Motion Capture Lab utilizes a VICON camera array with 12 cameras. Each camera is either attached to a ceiling rack mount or standing on a tripod, depending on the desired distance from the subject and size/layout of the markers. Each marker is a plastic sphere wrapper in reflective tape and attached to the subject with double-sided tape.

The first step in capturing a subject’s motion is calibrating the system to properly locate each camera. This is accomplished via triangulation by waving a wand with three markers positioned at set distances from one another. The floor plane also needs to be located by placing a flat frame with markers on it. Now the system is ready to capture.

The markers are placed on the subject in a manner consistent with the type of capture desired. For a full body capture the person puts on a full-body spandex suit and 41 markers are placed to construct a skeleton. For a less involved capture, e.g. of an isolated body part, fewer markers are used.

The motion data is captured using VICON IQ software on a workstation interfaced with the camera controllers. In order for IQ to create a subject-specific skeleton, Range of Motion and Motorcycle Pose captures should be taken. The Range of Motion requires the subject to rotate and flex each joint sequentially. The Motorcycle Pose is a stationary pose where all joints are bent.

The experimental motion capture data is then recorded in real time using IQ. After this data is captured it needs to be cleaned. The cleaning process removes ghost markers and fills in gaps caused by subject occlusion. The cleaned data may be exported to the public-domain binary C3D file format[i], which expresses the raw marker position data. To use the cleaned data as a basis for 3D animation in a commercial package, e.g. Alias Maya, one may use VICON BodyBuilder software to export it to Acclaim ASF/AMC formats, which express the skeletal motion metadata[ii].

Interference

This project studied whether human participants associate their own motion with life-like renders more so than with simplistic forms. An actor was motion-captured and videotaped moving his hand vertically and horizontally to the beat of a metronome. His motion was then applied to three different 3D models and rendered into video animations. The four resulting videos were: a simple small sphere; an ellipsoid humanistic skeleton; a textured human model; and the original video (*Figure 1*).

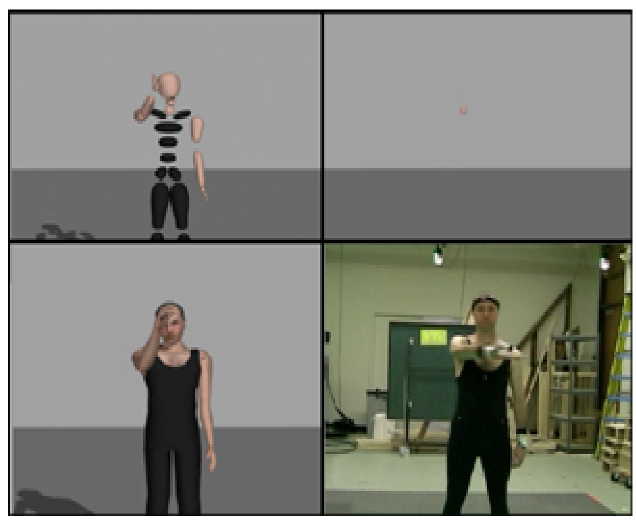

Figure 1: *Samples from videos (Interference I)*

Each type of video has two orientations corresponding the direction of the actor’s hand motion: vertical or horizontal. Each orientation has two different textual prefaces: one that says “Move Horizontally” and one that says “Move Vertically”. Thus, the text may correspond to the action in video they preface, or they may contradict it. In total, there are sixteen videos resulting from the combination of the above variables.

All sixteen videos were shown to 64 human experimental participants. Sixteen distinct orderings of the videos were employed, with four subjects seeing each ordering. Each participant was prepared by placing a reflective marker on their right hand. They were then told that they will see a series of videos prefaced by textual directions that tell the participant to move their hand vertically or horizontally. The participant was told to follow these directions, regardless of whether or not they correspond to the following video.

Three practice runs were performed in which the participant saw two horizontal / vertical instructions with corresponding video motion then one horizontal instruction with contradictory (vertical) video motion. The experimenter verified that the participant was following the textual directions and not simply mimicking the video motion projected before them.

Following the preparation practice runs, the experiment began and 16 videos were shown with 10-second breaks between each one to avoid participant fatigue. For each video, the participant's motion was captured as a separate take labeled with the video type, orientation, and whether or not the textual directions were contradictory. The motion was then cleaned; the resulting data was processed through a statistical analysis package to quantify smoothness or hesitation in motion.

The results demonstrated that the original video was hardest to follow, followed by the textured human model, then the ellipsoid skeleton, then the simple sphere. This difficulty ordering held for all videos, whether prefaced by corresponding or contradictory textual directions. It seems apparent that the more complex videos, with extraneous visual data on the screen, proved distracting to experimental participants and impeded their ability to smoothly mimic the original rate of motion.

One problem encountered was that the participants did not know whether to mirror the video or to move with it; e.g. whether the rendered model's left hand should correspond to the participant's left hand.

Following the first round of experimentation, a second round was performed with modifications designed to confirm our preliminary findings and to avoid several of the common problems encountered early on. A new set of videos was produced, all of them computer-rendered. The first was the original textured model holding a bright orange ball in the hand he was moving; the second had this model faded to be 15% of its opacity while still holding an opaque orange ball; the third features only the orange ball moving; and the final is the model with no ball (*Figure 2*).

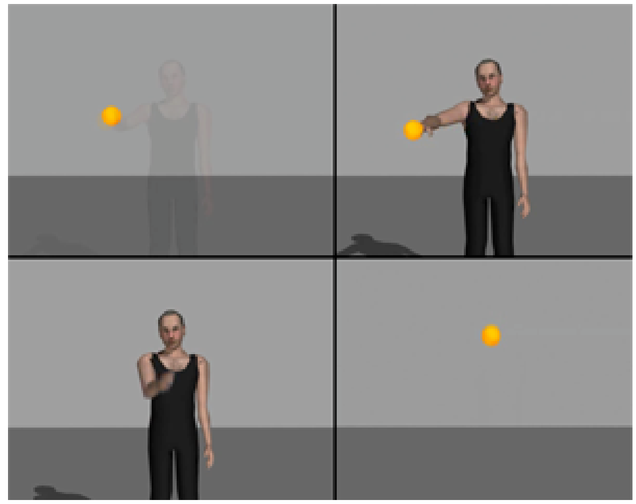

Figure 2: *Videos samples from Interference II*

For this phase there were three markers placed on the participant: two on the right and one on the left. This way both hands could be tracked and the skeleton fit could tell the difference between the right and left hand. The participant performed sixteen trials using their right hand and sixteen using their left hand.

General Subjects

The CMU Motion Capture Lab maintains an online database accessible to the general public[iii], allowing anyone to use captured and cleaned motion data for research projects. In order to keep this database full of high-quality sample motions we constantly capture average participants along with those having special talents. Through advertisements to the campus community we secure the participation of a variety of individuals.

When a participant arrives we explain to them what they will be doing and give them an overview of how the motion capture system works. The participant is then asked to put on a black spandex suit, which the motion capture technician covers with forty-one strategically placed markers (*Figure 3*). The participant is then guided through "Range of Motion" and "Motorcycle Pose" captures as described above.

For those with special talents, we ask them to show us specific moves and stances of motion that are unique to their talent. For instance, a person that does karate is asked to do some individual kicks and punches along with some fighting combinations.

When someone without a motion talent comes in we ask them to perform a list of tasks that vary from walking, running, and climbing steps to folding

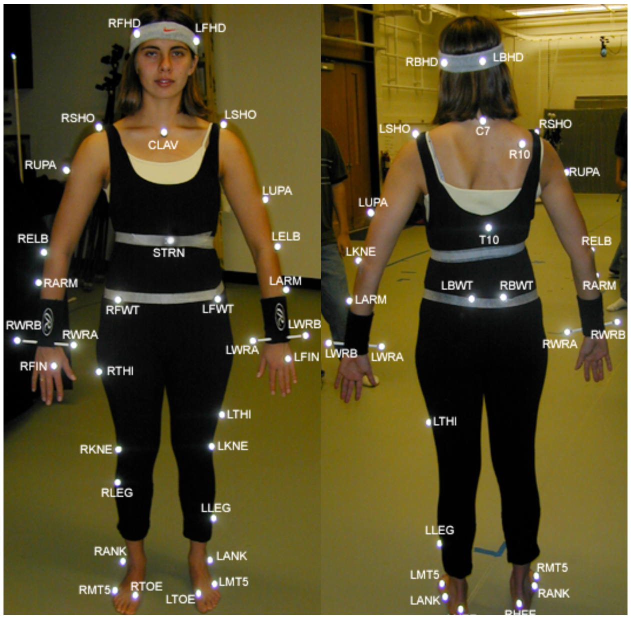

Figure 3: *Marker placement for full body capture*

laundry, or showing emotionally-charged actions like angrily throwing things or gleefully prancing.

Limitations

Although motion capture is an extremely useful technique, allowing us to obtain realistic information about the way that people move, it is not without its limitations.

Due to the constraint of the space in which the cameras are positioned, motions that are extensive like pole vaulting or long periods of uninterrupted running are very hard to capture as a whole. They can be captured in parts and stitched together with great effort.

Depending on the motion to be captured, different marker sets are employed. Therefore it is impractical to capture both a large-scale and a small-scale motion simultaneously. For instance, a human sneeze involves full-body motion as well as detailed facial motion. To capture both, two similar sneezes must be captured, using the appropriate-scale marker set each time. To use the small markers effectively, the cameras need to be positioned closer to the subject and thus the possible range of motion is constricted. The two separate captures must be manually pieced together.

Conclusion

The application of motion capture technology can produce extremely valuable results. The 3D animation industry has exploited motion capture techniques to quickly create animations with realistic nuances that a traditional animator might be unable to produce.

Not just a commercially-driven technology, motion capture provides researchers with an excellent tool to analyze real-world motion. It can express complex or qualitative motions as a quantitative collection of marker trajectories, suitable for traditional mathematical analyses.

The CMU Motion Capture Lab is pioneering many techniques to effectively capture and analyze motion data; this work will certainly help to drive motion capture technology to higher levels of maturity and commercial/research value.

References

[i] http://www.c3d.org/
[ii] http://www.motionrealityinc.com/software/asfamcspec.html
[iii] http://mocap.cs.cmu.edu/