AUTOMATIC MOTION-CAPTURE DATA CORRECTION

on 24-form Taiji Routine

# Summary

Mathematically speaking this is a problem of associating 39 human body joint names (labels) to a set of noisy, incomplete, or redundant marker points in the 4-dimensional spatiotemporal space. Our goal is to start with an initial set of label-marker associations generated by a commercial motion capture software package (VICON), and improve the label-marker associations to achieve better, whole-body, human Taiji movement trajectories of all joints. The evaluation methods include: visual evaluation and quantitative evaluation of trajectory smoothness and cross-subject/time comparison.

# Problem Description

Ideally, the VICON motion capture system detects all markers on a person, computes their 3D locations, labels each with which positional marker it is (e.g. Left Knee) and tracks them over time to form trajectories. However, several things can go wrong, due to occlusion, subject movement beyond the motion capture volume (visibility), or ambiguity due to close proximity of markers. In this case there are several post-processing (cleaning) steps that can be performed, some of which can be done automatically as scripts by the VICON software, and some of which need to be done (currently) by hand.

There are 6 conditional states that the motion capture data can be in after the data is initialized and the auto-labeling model is fit to the data. The auto-labeling tool is not 100% accurate and some of the problems that can arise are described below.

1. Correctly Labeled Marker: A marker is detected by the tracking hardware, the auto-fit correctly labeled the marker, and the marker has no capture noise/errors. This state happens with 97-99% of the data collected for a typical motion capture sequence. No additional processing is required.
2. Unlabeled Marker: A marker is correctly detected, but is not labeled. In this case, the 3D position of the marker position is present in the data, but the label for whatever reason was not associated with the marker. The 3D position is consistent through time but may have temporal gaps of 1 or more frames. These markers are often correctly labeled before and after the unlabeled sections, making it easy to identify.
3. Mislabeled Marker: A marker is detected but labeled incorrectly. In this case, the marker has been labeled but the label is for the wrong location. This could be caused by the labels of two nearby points being swapped resulting in two markers being incorrectly labeled (e.g. Left Knee label in the Left Tibia marker, Left Inner Wrist label on the Left Outer Wrist marker). This occurs more frequently with markers that are symmetrically associated or have spatial proximately over a long time period. This issue is identifiable as the markers are often correctly labeled before and after the period of mislabeling.
4. False Marker: A marker that is in the dataset but does not associate with any of the 39 labels that define a subject. This are incorrectly detected markers in 3D space that are the result of false positives from the tracking hardware. This can happen due to noise, reflections off of shiny objects, or influence from markers/reflections outside the capture volume.
5. Missing Markers: A marker is not detected, and thus missing from the data. In some cases, it is possible to automatically recover the missing 3D marker position. For example, some rigid body parts have four markers on them. If only one is missing but the other three are correctly detected, it is possible to compute the relative position and motion of the fourth from the visible three.  Unfortunately, four points per rigid body part is the exception rather than the rule. There are existing techniques for estimating the location of missing markers using 3D spline fitting over small temporal gaps and markers that are associated with the missing marker (replacing the heel maker using the ankle or toe). This becomes very difficult when missing markers have large spatial separation or no markers in the part segment (entire hand or foot is missing).
6. Noisy Marker: A marker that is mostly in the correct location and properly labeled but through visual inspection have significant noise or tracking errors. The issues fall into 2 major subsets; small spatial variation at high frequency (1-50mm, >10Hz) and large spatial variation at low frequency (50mm-1m, <2Hz). Most of these issues result from data quality issue as the result of hardware limitations or marker tracking at the limits of the tracking volume.

# Current State of the Art

The problem of cleaning motion capture data automatically has yet to be solved, particularly when they are relatively long (>1sec) gaps of missing data, caused by extended occlusions or when a person’s body parts (segments) are near or outside the boundary of the capture space. It should also be noted that the resolution of one issue may not transition a marker in to the “correct” category. This means that an unlabeled marker that is then correctly labeled does not mean that it does also have noisy marker characteristics or sub-sections of missing marker characteristics.

To put the scale of the data issues in to perspective, consider the volume of data that is required to be 100% correct to be considered clean. While a 5 minute sequence has a total of ~1.2 Million markers and most of the collected data is correct (~98%) from automatic labeling, the remaining marker issues represent ~23,000 corrections that require resolution. On average, the time to correct these remaining markers requires ~6 hours of manual changes to complete the cleaning process. However, the issues can be resolved in significant groups by labeling a marker for 100 or more frames at once.

An interesting data-driven or learning approach is presented in Baumann et.al. “Data-driven completion of motion capture data”. The related work section of that paper would also be good to look at for identifying previous approaches to automated cleaning (as would be papers since 2011 that refer back to this one). It is important to check which publication since 2011 has cited this paper, and any other newer paper on this type of work.

# References

Baumann, Kruger, Zinke, Weber; “Data-driven completion of motion capture data” Workshop on virtual reality interaction and physical simulation, Dec. 2011, Institute of Computer Science, Bonn University, Lyon, France, Eurographics Association. <http://cg.cs.uni-bonn.de/aigaion2root/attachments/2011_MotionCleaning.pdf>