Skeletal Parameter Estimation from Optical Motion Capture Data

Adam Kirk James F. O'Brien David A. Forsyth
University of California, Berkeley

Overview

In this sketch we present an algorithm for automatically estimating a subject's skeletal structure from optical motion capture data without using any *a priori* skeletal model. Our algorithm consists of a series of four steps that cluster markers into groups approximating rigid bodies, determine the topological connectivity between those groups, locate the positions of the connecting joints, and project those joint positions onto a rigid skeleton. These steps make use of a combination of spectral clustering and nonlinear optimization. Because it does not depend on prior rotation estimates, our algorithm can work reliably even when only one or two markers are attached to each body part, and our results do not suffer from error introduced by inaccurate rotation estimates. Furthermore, for applications where skeletal rotations are required, the skeleton computed by our algorithm actually provides an accurate and reliable means for computing them. We have tested an implementation of this algorithm with both passive and active motion capture data and found it to work well. Its computed skeletal estimates closely match measured values, and the algorithm behaves robustly even in the presence of noise, marker occlusion, and other errors typical of motion capture data.

Methods

The first step of our method is to cluster markers into groups that represent rigid bodies. In an ideal rigid body, the points on the body do not move with respect to each other through time, i.e. the standard deviation in distance between points on the body is zero. Therefore, to determine rigid bodies our method clusters based on the standard deviation in distance between all pairs of markers. Using all frames to compute the standard deviation in distance between two markers can be expensive and sensitive to errors induced by noisy frames, so the method calculates this quantity only over a jittered uniform sampling of frames. Markers in the sampled set of frames are segmented using spectral clustering. To correct for errors due to sampling, the algorithm employs a random sample consensus (RANSAC) based procedure. Rather than using one sampling of frames, the method uses several different samplings, each of which produce a possible clustering of markers. The optimal clustering has the smallest average standard deviation between markers in a group over all samplings.

Our method determines skeleton topology by testing for possible joints between all rigid body pairs. For each pair, our method attempts to solve a nonlinear optimization problem for a point that remains fixed with respect to the markers on those bodies. More specifically, this point minimizes the average variance in distance between itself and each marker. Unfortunately, the trivial solution to this optimization is a point infinitely far away. To prevent the procedure from finding the trivial solution, a small cost proportional to the average distance between markers and the point is added to the function. Optimization residuals provide a metric for the likelihood that a pair of rigid bodies should be connected. The optimal skeleton minimizes the sum of these residuals. By treating rigid body groups as nodes in a fully connected graph, and using the residuals as edge weights in that graph, the optimal skeleton is found by computing the minimum spanning tree. At this stage, the algorithm is only trying to determine if rigid bodies are connected. As such, the method speeds computation by subsampling frames. This approach is similar to [O'Brien et al. 2000]. However, they used magnetic motion capture data, which includes rotation information and allows the problem to be expressed in linear form.

Once the topology of the skeleton is identified, the method proceeds with finding joint positions. For this, the method repeats the minimization procedure in the previous stage, however all frames are used. Additionally, the length term in the optimization is dropped. This term is eliminated because the correct joints are known, so the method no longer needs to protect against joints far from the skeleton.

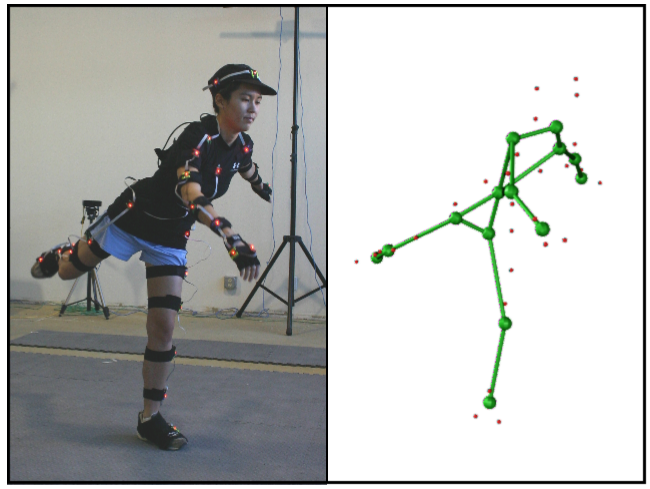

Figure 1: Automatic skeletal reconstruction for human subject

The final step of our algorithm is to project the joint and marker positions onto a rigid body skeleton. This procedure is done one rigid body at a time. For each rigid body, the only frames used are the ones in which all the markers on that body appear. Within this set, a single frame is picked at random to serve as the target. For every other frame, the method computes the affine transformation that best transforms the marker and joint positions to equal those of the target frame. If all the frames are lined up using their respective transformations, several small clouds of points appear, representing each of the markers and joints connected to the rigid body. For each cloud of points the average position is used as the true offset of the marker or joint.

If the rotation at each joint is desired, it can be found using inverse kinematics (IK). Since the rigid body skeleton and the offset of each marker from a bone is known, IK is used to find the rotations that minimize the distance from the marker positions on a bone and the input data. Estimating joint rotations using skeleton information is considerably better than just using marker data. Knowledge about relative joint position removes ambiguity in the estimation procedure, which results in more accurate rotation estimates, even if there is only one marker on a body segment.

Results

We tested this method on human data gathered from a PhaseSpace active motion capture system and a Vicon passive motion capture system. When using calibration motion capture sequences as input we were able to accurately reconstruct visually plausible human skeletons, as seen in Figure 1. To provide further validation, we constructed a three-link chain of aluminum rods connected by universal joints. We motion captured this mechanical device using a PhaseSpace active marker system and were able to reconstruct the length of the middle rod within less than a centimeter, which is within the accuracy limits of the motion capture system.

References

O'BRIEN, J. F., BODENHEIMER, R. E., BROSTOW, G. J., AND HODGINS, J. K. 2000. Automatic joint parameter estimation from magnetic motion capture data. In *Proceedings of Graphics Interface 2000*.