Optimal Segmentation and Understanding of Motion Capture Data

Xiang Huang, M.A.Sc Candidate
Department of Electrical and Computer Engineering
McMaster University

Supervisor: Dr. Xiaolin Wu

7 Apr, 2005

Outline

- Introduction
- Optimal Segmentation of Motion Capture Data
  - Mathematical Formulation of Optimal Segmentation
  - Algorithms for Optimal Segmentation
  - Fitting Motion Capture Data to Dynamic Model
- Understanding Motion Capture Data
  - Motion Primitives and Motion Sequence
  - Segmenting Motion Sequence into Meaningful Motion Episodes
- Selected References

Introduction

- Motion capture systems attach markers to 3D objects (Fig. 1) and track the markers'positions as the objects move.

- Motion capture data can be considered four-dimensional: at each time instance a frame of 3D coordinates of the makers are recorded.

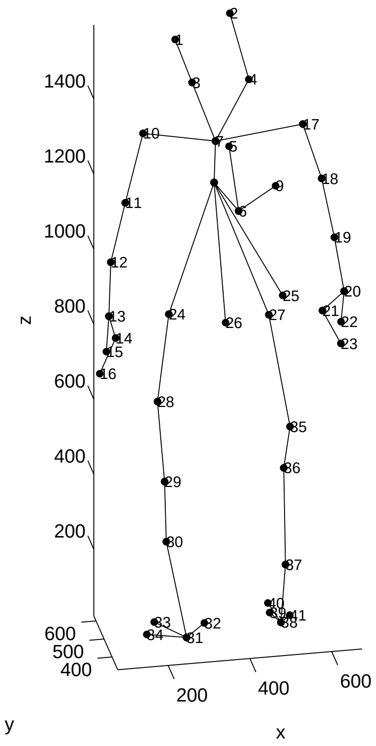

Figure 1: Human body with 41 Markers attached on it.

Introduction

- Like in most other fields, abstraction, summarization and modeling of raw observation data is of fundamental value for motion capture research and applications. Temporal segmentation is a natural process of extracting semantic elements and identifying syntactic structures of the actions in an acquired 3D data sequence.

- The key design is to fit a parametric dynamic model to the input motion sequence, and implicitly extract the segments at the time instances when the discontinuities of model parameter values occur.

- The parameters of the dynamic model has clear mechanical and geometrical meanings. Therefore, the motion segments generated by this model provide key constructs in scene analysis and motion synthesis, hence help understanding the motion capture data.

Mathematical Formulation of Optimal Segmentation

- In Fig. 2, the dotted line is the original trajectory (two dimensional) of 150 time samples. It is segmented in to 14 segments. The solid line is the approximately reconstructed trajectory by linear interpolation.

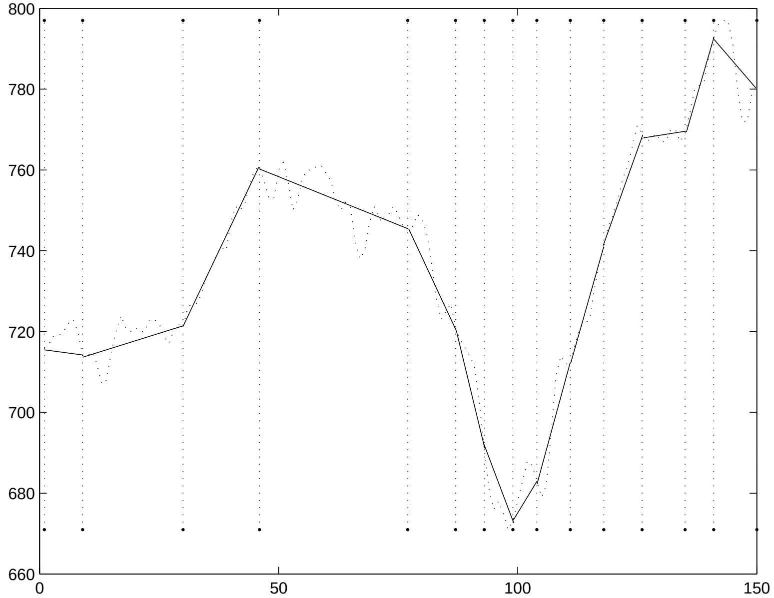

Figure 2: Linear Segmentation.

Mathematical Formulation of Optimal Segmentation

- Assume there are N time samples and denote the $3D$ position of marker $\mathbf{p}_i$ at time sample n as $\mathbf{p}_i^{(n)}$. Marker $\mathbf{p}_i$'s original trajectory is $\mathbf{P}_i = (\mathbf{p}_i^{(1)}, \mathbf{p}_i^{(2)}, \cdots, \mathbf{p}_i^{(N)})$,

- Let $\mathcal{S}_i = (i_1, i_2, \cdots, i_{K_i})$ as the segmentation's key points, $1 = i_1 < i_2 < \cdots < i_{K_i} = N$. Samples of $\mathbf{p}_i$ between any two adjacent key points will be fit to a dynamic model and can be approximately reconstructed by that dynamic model later.

Mathematical Formulation of Optimal Segmentation

- $$e_i(a,b] = \|\hat{\mathbf{P}}_i(a,b], \mathbf{P}_i(a,b]\| \tag{1}$$

  is the approximation error between the original trajectory and the reconstructed trajectory using dynamic model from time step $a+1$ to time step b.

- Markers belong to the same rigid part of human body should be segmented identically in time. The approximation error of all the markers on rigid body U_m is

  $$D_m(\mathcal{S}_m) = \sum_{\mathbf{p}_i \in U_m} \sum_{k=1}^{K_m} e_i(m_{k-1}, m_k]. \tag{2}$$

- Optimal motion segmentation for rigid body U_m is to minimize $D_m(\mathcal{S}_m)$ over all possible partitions of N frames.

Algorithms for Optimal Segmentation

- The optimal segmentation problem can be casted into two variants of a discrete optimization problem and solve it by a graph theoretical approach.
- First variant is to minimize the distortion while the segmentation number is given.
- Second variant is to minimize the segmentation number while satisfying prespecified error bounds.
- It could be seen that the algorithms to be developed here are general, independent of specific dynamic models for data fitting.

Algorithms for Optimal Segmentation

- In the first variant, construct a complete direct acyclic graph $G = \langle V, E\rangle$ (Fig. 3). The vertex set V contains $N + 1$ nodes labelled $0, 1, 2, \cdots, N$, node n corresponding to time stamp n. The edge set E consists of $N(N + 1)/2$ directed edges: edge (a, b) from node a to node b exists if and only if $a < b$.

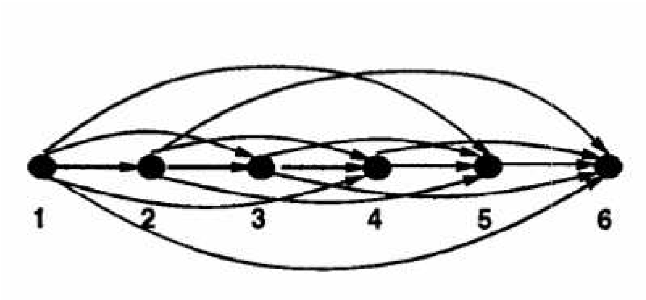

Figure 3: A complete directed acyclic graph (DAG) of 6 Nodes

Algorithms for Optimal Segmentation

- Assign to each edge $(a, b) \in E$ the weight

$$w(a, b) = \sum_{\mathbf{p}_i \in U_m} e_i(a, b]. \tag{3}$$

Then the optimal K_m segmentation is determined by the the K_m-link shortest path form node 0 to N. This problem can be solved in $O(K_m N^2)$ time by dynamic programming.

Algorithms for Optimal Segmentation

- The second variant of minimizing K_m while meeting the error bound $D_m \leq \mathbb{D}_m$ can be casted into a Lagrangian optimization problem of minimizing $K_m + \lambda D_m$ with a binary search on the value of λ.

- We adopt the same DAG of previous case, but change the edge weight to

$$w(a, b) = \sum_{\mathbf{p}_i \in U_m} (1 + \lambda e_i(a, b]). \tag{4}$$

Minimizing $K_m + \lambda D_m$ for a given λ can be solved by computing the conventional shortest path from node 0 to N, which can be solved in $O(N^2)$ time by dynamic programming.

Algorithms for Optimal Segmentation

- In the second variant, if the minimax approximation criterion is used, then $D_m(\mathcal{S}_m)$ is of the L_∞ norm:

$$D_m = \max_{\mathbf{p}_i \in U_m, 1 \le k \le K_m} e_i(m_{k-1}, m_k]. \quad (5)$$

- Although the objective function is no longer additive, the problem remains solvable by a shortest path algorithm. We prune all the edges $(a, b) \in E$ from the DAG G if there exists a marker $\mathbf{p}_i \in U_m$ such that $e_i(a, b] > \mathbb{D}_m$. Then we assign the unit weight 1 to all the survived edges. Now it is immediate that minimizing K_m under the constraint $D_{\infty,m} \le \mathbb{D}_m$ is equivalent to finding the shortest path from node 0 to node N in the edge-pruned graph.

Fitting Motion Capture Data to Dynamic Model

- The role of motion segmentation is to provide a sequence of semantically meaningful and well-defined motion alphabets to be interpreted by later stage. This view leads us to the dynamic model based on classic Newton kinematics.

- All marker movements are specified in relative motions about a joint to which they are connected. If marker $\mathbf{p}_i$ is on an ideal rigid body, then it moves on a sphere that is centered at the joint $\mathbf{o}$ and has the radius $d(\mathbf{p}_i, \mathbf{o})$ (Fig. 4).

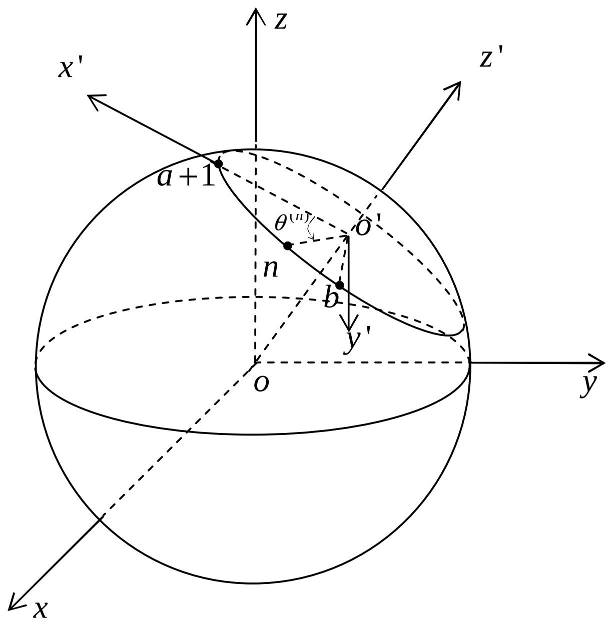

Figure 4: Marker move on the surface of a sphere.

Fitting Motion Capture Data to Dynamic Model

- In a segment $(a, b]$, assume marker $\mathbf{p}_i$ has a constant angular acceleration τ, then it leaves a circular trajectory $\widehat{ab}$ on the sphere. The motion can be modeled by Newton kinematics equation:

$$\theta^{(n)} = \hat{\theta}^{(a)} + \omega^{(a)}(n-a) + \frac{1}{2}\tau(n-a)^2. \tag{6}$$

- The dynamic model with constant angular velocity is simple yet reasonable as human movement contains a lot of accelerating and decelerating processes (Fig. 5).

Fitting Motion Capture Data to Dynamic Model

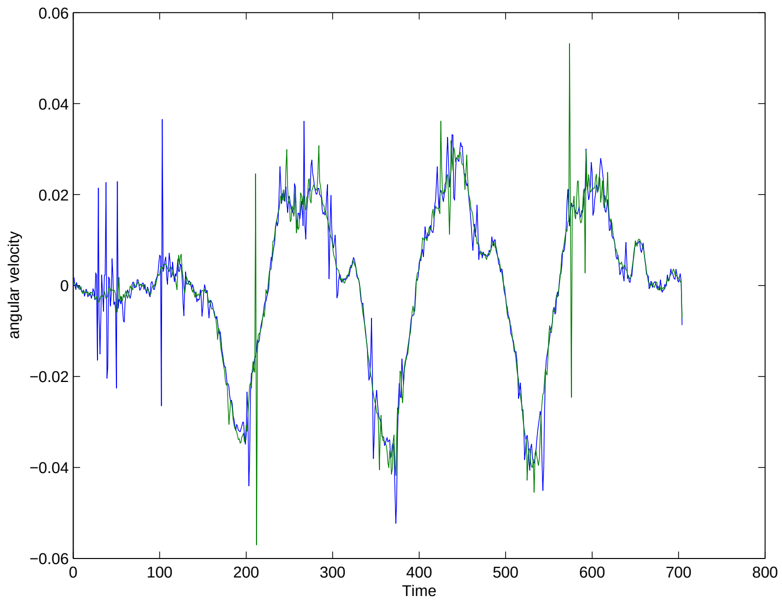

Figure 5: Angular Velocity of two markers belong to the same rigid body.

- We first find a plane fit to sample points $\mathbf{p}_i^{(a+1)}, \mathbf{p}_i^{(a+2)}, \cdots, \mathbf{p}_i^{(b)}$ in least square sense, as these samples are not exactly in the same plane because of noise and model mismatch.

Fitting Motion Capture Data to Dynamic Model

- Then as markers on the same rigid body have the same angular velocity, we can solve a least square fitting problem to get the parameters ${\theta_i}^{(a)}$, $\omega^{(a)}$ and τ.

$$\min_{\hat{\theta}_i^{(a)},\omega^{(a)},\tau} \sum_{\mathbf{p}_i \in U} \sum_{a<n\leq b} \left[\theta_i^{(n)} - \left(\theta_i^{(a)} + \omega^{(a)}(n-a) + \frac{1}{2}\tau(n-a)^2\right)\right]^2 \qquad (7)$$

- With the estimated parameters, the $\hat{\theta}_i^{(n)}$ in time n can be reconstructed via equation (6). Transforming $\hat{\theta}_i^{(n)}$, $n \in (a, b]$, back to the original coordinate system yields $\hat{\mathbf{p}}_i^{(n)}$, and hence the approximation error $e_i(a, b]$ of the model can be calculated in the motion segmentation process.

Motion Primitives and Motion Sequence

- Introducing a quantization step, the parameter vector in each segment can be represented by a symbol named motion primitive, and the total K_m segments can be represented by a sequence of length K_m, which is called motion sequence.

- We could apply sequence analysis algorithms on motion sequence to help understanding human movements.

- Human movement can be called body language. In motion primitive sequence, a motion primitive is analogical to an English character, and a higher level structured subsequence named motion episodes is analogical to English vocabulary.

Segmenting Motion Sequence into Meaningful Motion Episodes

- Segmenting a sequence into episodes is just like finding the word boundaries after removing all the spaces and punctuation from a text. E.g. ”eachpersonassignedasbelowshouldsubmitanindividualpostertobepresented” could be segment to ”each person assigned as below should submit an individual poster to be presented”.

- Episode is a subsequence not only occurs frequently but also meaningful. E.g., though in English both ”the” and ”th” occurs frequently, ”the” is an episode while ”th” is not.

- Cohen [1] used ”boundary entropy” and ”frequency” rules and segment characters into words with some accuracy.

References

[1] Paul R. Cohen and Niall M. Adams. An algorithm for segmenting categorical time series into meaningful episodes. In *IDA*, pages 198–207, 2001.

[2] Yan Li, Tianshu Wang, and Heung-Yeung Shum. Motion texture: a two-level statistical model for character motion synthesis. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 465–472, New York, NY, USA, 2002. ACM Press.