Motion Data Compression

Lecture 18

Taku Komura
Motivation

• Motion capture data is not so small
• 120 frames per second, 41 markers, 10 seconds -> 10MB
• So want to compress the data
Compression techniques

• Compression is used for image, video and sound files
  
  – Temporal coherence (sound, video)
    • Subsequent sound/images are similar
  – Spatial coherence (video, images)
    • Adjacent pixels have similar color

  – For human motion, we can use both temporal / spatial coherence
Approach to compress human motion data

• Principal component analysis (PCA)
  – Segment and do PCA within each segment
  – Cluster motions and then do PCA within the clusters

• Wavelet transform
  – Wavelet transform the trajectories of the joint angles individually
  – Keep higher frequencies for DOF that are more important
PCA

- PCA is a dimensionality reduction technique
- Retains characteristics of a dataset that contribute most to its variance
PCA-based compression

• Although the dimensionality of the data is high, usually the distribution of the data is biased

• In PCA, we find the principal components of the data

• The original data can be composed by linearly combining the principal components
Eigenvectors

• Say there are many samples like
  - \{x_1,x_2,x_3,\ldots, x_n\}
  - By using PCA, we can compute the eigenvectors
    \{y_1,y_2,y_3,\ldots, y_n\}
  - Usually, you only need the first few to represent the original data

\[ x_i = \sum_j c_j y_j + \mu_x \]

\[ \mu_x \] is the average of \( x_i \)
PCA-based approach

- There are many postures that are similar in motion data
- Arbitrary postures can be composed by linearly combining similar postures
- Make use of such a feature
PCA-based compression of motion segments

• In motion segments, the pose vectors lie near a much lower dimensionality
• PCA works very effectively
• Keep the first $k$ eigenvectors such that the residual variance covered by the discarded eigenvectors is less than a preset threshold
Segment-based approach

- We can first segment the motions when the behavior changes
  - Walking to running
  - Walking to stop and mowing grass
- Segment when the new motion part cannot be represented well by the principal components in the last $K$ frames
Experiments - Parameters

• PCA error tolerance for *non-normalization* markers
  – 6 Foot markers: 10mm/marker
  – Rest markers: 30 mm/marker
Experiments – Distortion Measures

• Distortion rate (%) for a whole sequence

\[ d = 100 \frac{\| A - \tilde{A} \|}{\| A - E(A) \|} \]

– where \( A \) is a \( 3m \times N \) data matrix containing the original motion sequence collected from \( m \) markers over \( N \) frames.

– \( \tilde{A} \) is the reconstructed result of the same motion sequence after decompression.

– \( E(A) \) is an average matrix in which each column consists of the average marker positions for all the frames.
Experimental Results

Breakdance

<table>
<thead>
<tr>
<th>Sequence</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>#segments</td>
<td>6</td>
<td>9</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>#principle components</td>
<td>4.6</td>
<td>10.9</td>
<td>3.1</td>
<td>4.2</td>
</tr>
<tr>
<td>#control points</td>
<td>117</td>
<td>211</td>
<td>235</td>
<td>221</td>
</tr>
<tr>
<td>Compression ratio</td>
<td>1:55</td>
<td>1:18</td>
<td>1:62</td>
<td>1:56</td>
</tr>
<tr>
<td>Distortion rate</td>
<td>5.1</td>
<td>7.1</td>
<td>5.1</td>
<td>5.4</td>
</tr>
<tr>
<td>Compression time (ms/frame)</td>
<td>1.3</td>
<td>1.4</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>Decompression time (ms/frame)</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Application to other parts of the body

- PCA is also effective for reducing the facial motion data.
- Many muscles move concurrently.
Wavelets

- Divides continuous signals into different frequency components
- Wavelet basis can be localized in space
  - In other compression methods such as Fourier transform, the basis function affects the whole curve
  - In wavelets, it can affect only a local area
Fourier Basis vs. Wavelet Basis

- Wavelet basis

- Fourier basis
Wavelet Transform

- Represent the given curve by the sum of wavelet bases
  - $j$ is the highest resolution
- We can linearly compute the lower resolution coefficients

\[
x = \sum_{j} \Phi^j c^j
\]

\[
c^{j-1} = A^{j} c^j
\]

\[
d^{j-1} = B^{j} c^j
\]

\[
c^j = P^{j} c^{j-1} + Q^{j} d^{j-1}
\]
Wavelet Transform
Standard wavelet compression

• Wavelet transform 62 signals

• Yield vector $\mathbf{w}_i$ ($1 \leq i \leq 62$) counting how many coefficients are kept for each signal
Vector $w_i$
Optimized coefficient selection

- $w_i$: initial number of coefficients for DOF $i$
- Based on how much the DOF affects the movements at the tip
  - Head: small
  - Thigh: large
Optimized coefficient selection

- Build vector $\mathbf{m}_i$ that favors some signals more than others
- Saving more coefficients for those important ones
- Fixed choice for $\mathbf{m}_i$? Bad!
  - Depends on complexity of signals
  - Depends on the poses
Start with \( \mathbf{m}_i = \mathbf{w}_i \)
Randomly select $i$
reduce $m_i$
Find optimal $j$

Increase $m_j$

Repeat…
Inverse kinematics correction

• Problem: Noticeable sliding feet

• Add positional channels for the feet, use IK

• Again do the wavelet transform for the trajectory of the feet
Comparison

• Joint + temporal coherence
  – Cannot access individual signals
  – 30:1 up to 35:1 [Arikan 06]

• Temporal coherence alone
  – 35:1 (this work)
  – Access to subset of joint data
  – Low computational requirements
Readings

• Philippe Beaudoin et al. “Adapting Wavelet Compression to Human Motion Capture Clips”, GI 2007
• Arikan, O., “Compression in Motion Capture Databases”, SIGGRAPH 2007
• Liu et al. “Segment-based Human Motion Compression”, SCA07