

Lecture

Information Retrieval for Music and Motion

Meinard Müller and Andreas Baak

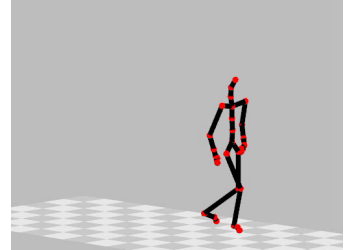
Summer Term 2008

Motion Retrieval



Motion Capture Data

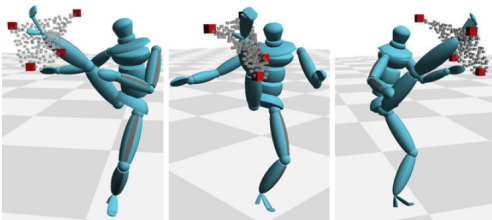
- Digital 3D representations of motions
- Computer animation
- Sports
- Gait analysis



2

Motion Capture Data

Application: Motion Morphing

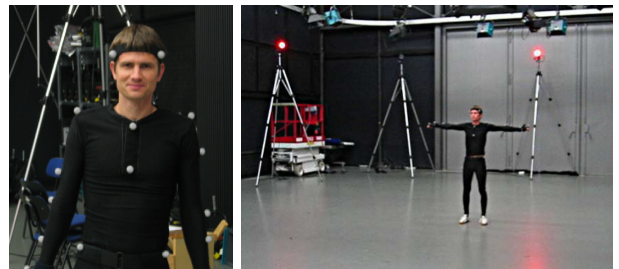


From Kovari/Gleicher (SIGGRAPH 2004)

3

Motion Capture Data

Optical System



4

Motion Capture Data

Mechanical and magnetic systems



<http://www.metamotion.com/gypsy/gypsy-motion-capture-system.htm>

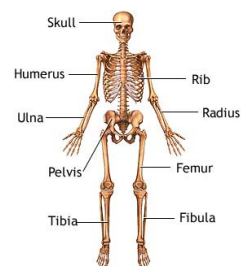


http://vrlab.epfl.ch/research/MC_motion_capture.html

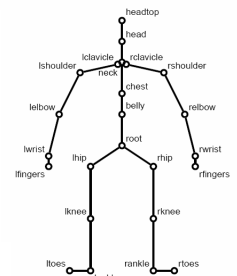
5

Motion Capture Data

Skeletal kinematic chain



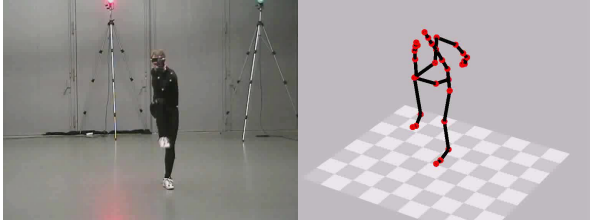
<http://apps.uwhealth.org/health/adam/graphics/images/en/9065.jpg>



6

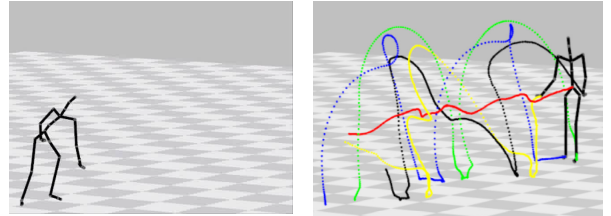
Motion Capture Data

Conversion: Marker → Skeleton



7

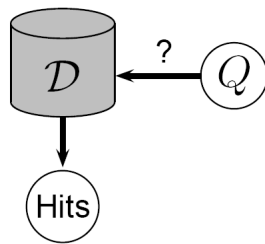
Motion Capture Data



8

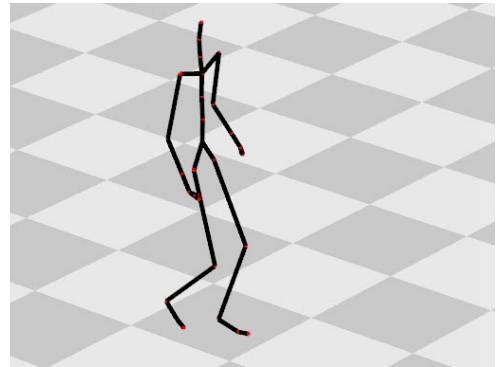
Motion Retrieval

- \mathcal{D} = MoCap database
- Q = query motion clip
- **Goal:** find all motion clips in \mathcal{D} similar to Q



9

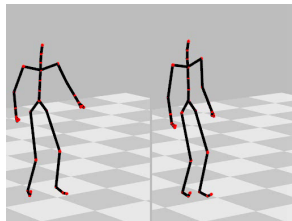
Motion Retrieval



10

Motion Similarity

- **Numerical** similarity vs. **logical** similarity
- Logically related motions may exhibit significant **spatio-temporal** variations

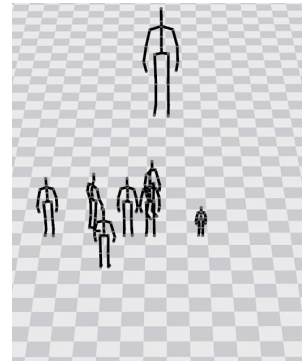


11

Motion Similarity

Global Transforms

- Translation
- Spatial scaling
- Rotation
- Reflection
- Temporal Scaling

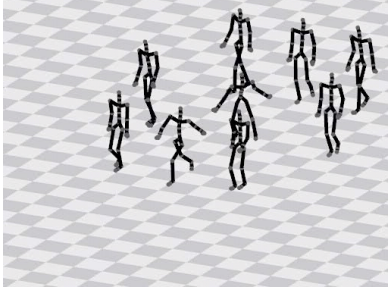


12

Motion Similarity

Motion Styles

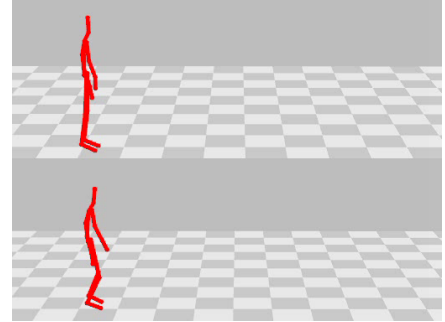
- Cheerful walking
- Furious walking
- Limping
- Tiptoeing
- Marching



13

Motion Similarity

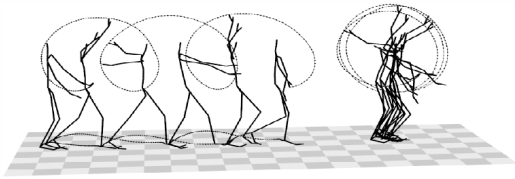
Spatio-Temporal Deformations



14

Motion Similarity

Partial Similarity

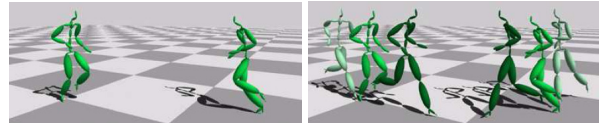


15

Local Similarity Measure

Point cloud (Kovar & Gleicher)

$$c^{3D}(D(n), D(m)) := \min_{\theta, x, z} \left(\sum_{i=1}^K w_i \|p_i - T_{\theta, x, z}(p'_i)\|^2 \right)$$

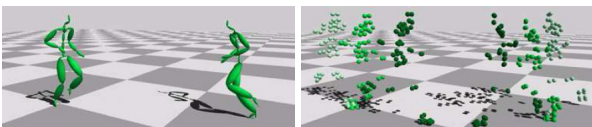


16

Local Similarity Measure

Point cloud (Kovar & Gleicher)

$$c^{3D}(D(n), D(m)) := \min_{\theta, x, z} \left(\sum_{i=1}^K w_i \|p_i - T_{\theta, x, z}(p'_i)\|^2 \right)$$

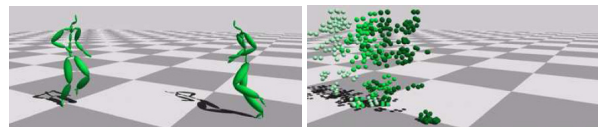


17

Local Similarity Measure

Point cloud (Kovar & Gleicher)

$$c^{3D}(D(n), D(m)) := \min_{\theta, x, z} \left(\sum_{i=1}^K w_i \|p_i - T_{\theta, x, z}(p'_i)\|^2 \right)$$



18

Local Similarity Measure

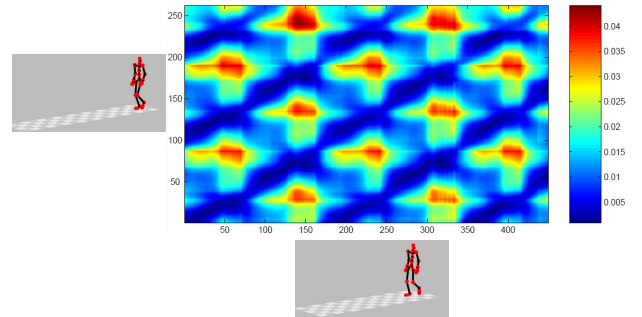
Quaternions

$$c^{\text{Quat}} : \mathcal{J} \times \mathcal{J} \rightarrow [0, 1]$$

$$c^{\text{Quat}}(j, j') := \sum_{b \in B} w_b \cdot \frac{2}{\pi} \cdot \arccos |\langle q_b | q'_b \rangle|$$

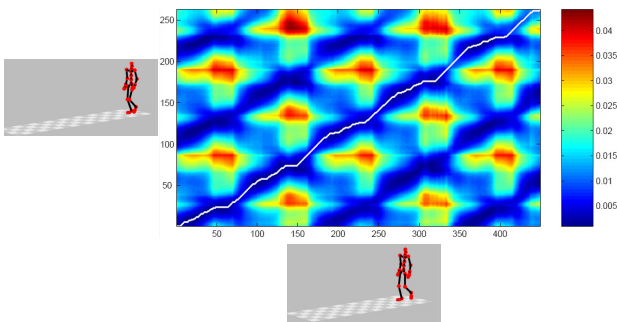
19

Dynamic Time Warping (DTW)



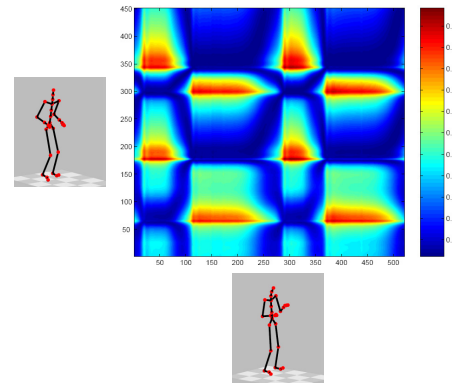
20

Dynamic Time Warping (DTW)



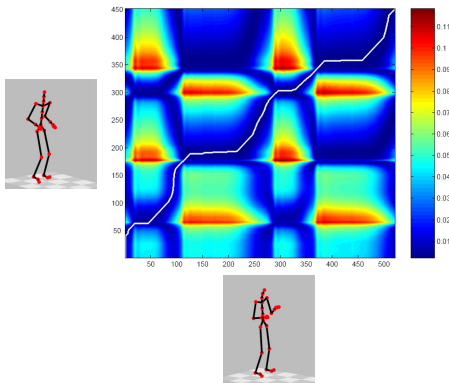
21

Dynamic Time Warping (DTW)



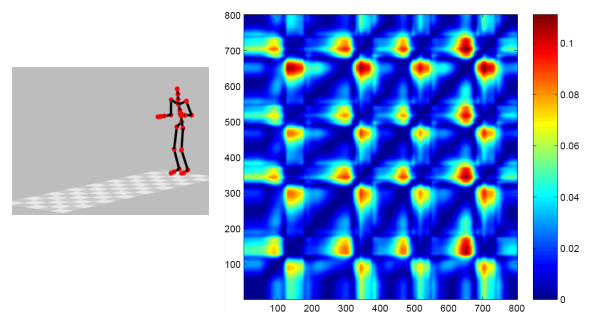
22

Dynamic Time Warping (DTW)



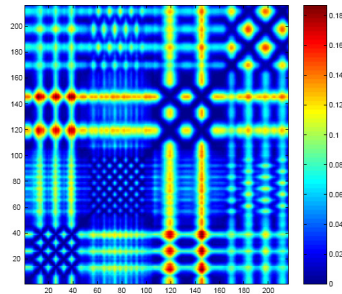
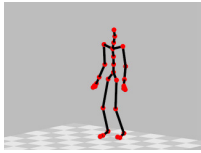
23

Self-Similarity Matrix



24

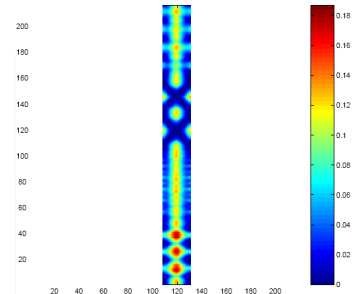
Self-Similarity Matrix



- Given: motion database (one single document)
- Compute: self-similarity matrix

25

Self-Similarity Matrix



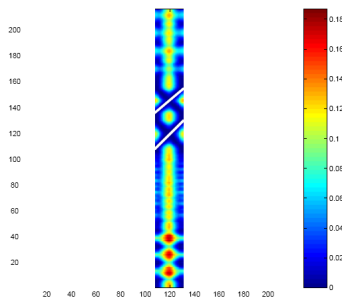
Query



- Query: segment of motion database
- Consider similarity matrix over query

26

Self-Similarity Matrix



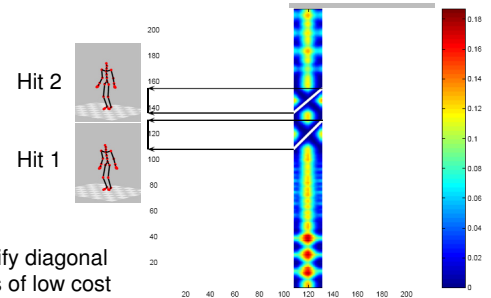
Query



- Identify diagonal paths of low cost

27

Self-Similarity Matrix



Query



Hit 2



Hit 1

- Identify diagonal paths of low cost
- Project paths onto vertical axis

28

Some Drawbacks

- DTW-based techniques computationally expensive
 - ↪ do not scale to large databases
- Rely on numerical features
 - ↪ hard to identify logically related motions
- No user-specified "center of attention,"
 - ↪ incorporation of a-priori knowledge not possible

29

Other Recent Approaches

- Wu et al. (IPPR 2003):
 - identify candidates for start and end frames
 - use DTW to compute actual distance from query
- Keogh et al. (VLDB 2004):
 - identify motion clips differing by global scaling
- Forbes/Fiume (SCA 2005):
 - PCA-based local features
 - substring DTW for matching

30

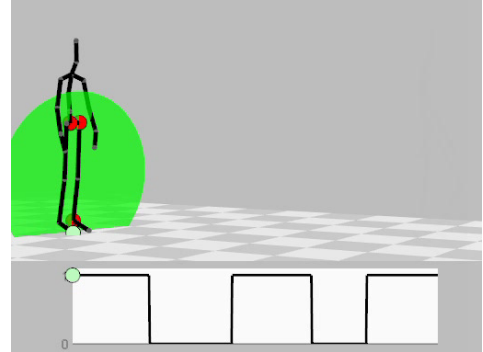
Our Approaches

- Introduction of relational features
 - ↪ accounting for spatial deformations
- Introduction of adaptive temporal segmentation
 - ↪ accounting for temporal deformations
- Usage of linear time/space indexing techniques
 - ↪ scalable to large databases

Müller/Röder/Clausen (SIGGRAPH 2005)

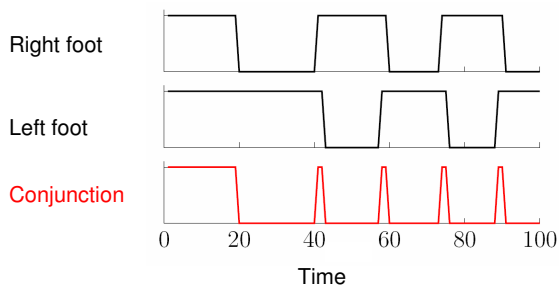
31

Relational Features



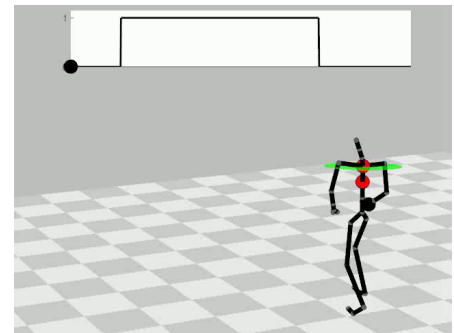
32

Relational Features



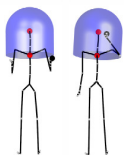
33

Relational Features



34

Relational Features



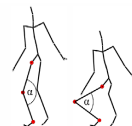
Left hand touching head?



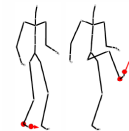
Both hands touching?

35

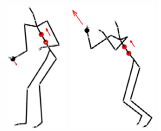
Relational Features



Right knee bent?



Right foot fast?

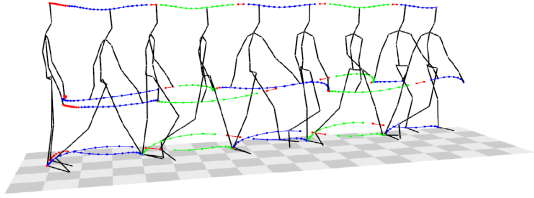


Right hand moving upwards?

36

Relational Features

Temporal Segmentation:



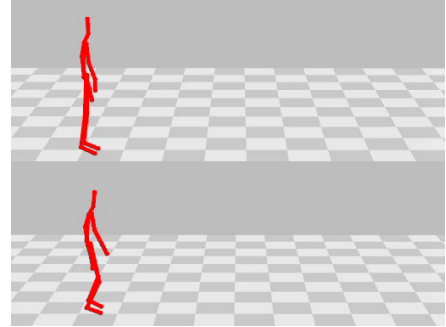
Induced feature sequence:

$$\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$$

37

Relational Features

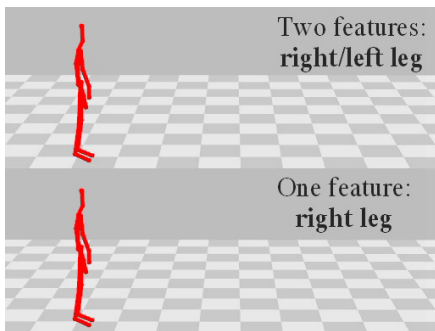
Spatio-temporal invariance



38

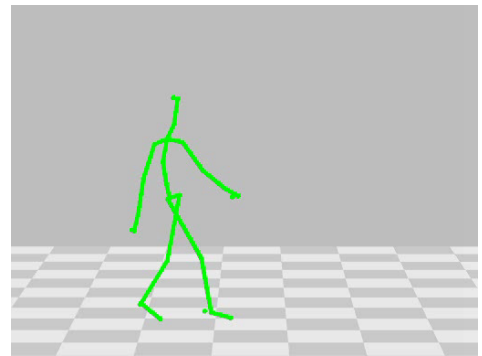
Relational Features

Feature Adaptivity



39

Motion Retrieval



40

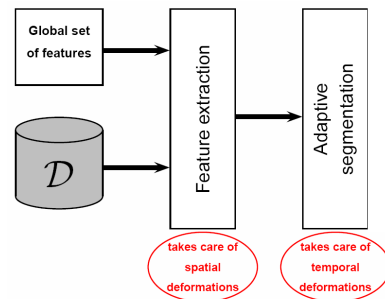
Motion Retrieval

Global set
of features



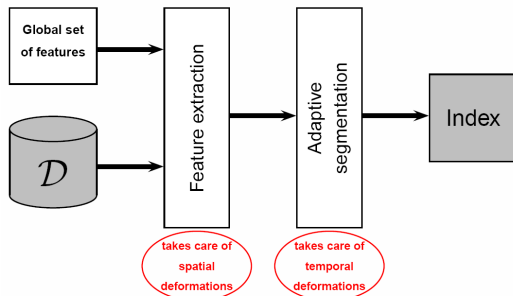
41

Motion Retrieval



42

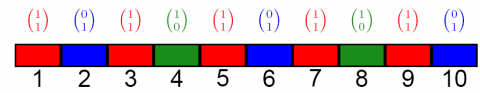
Motion Retrieval



43

Motion Retrieval

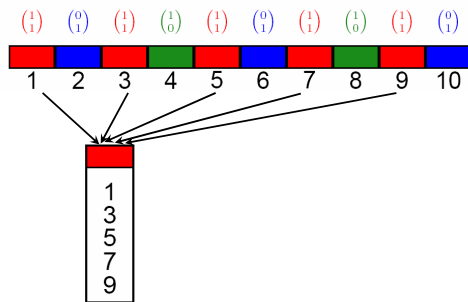
Indexing with inverted lists



44

Motion Retrieval

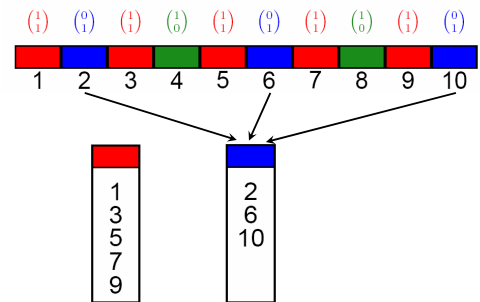
Indexing with inverted lists



45

Motion Retrieval

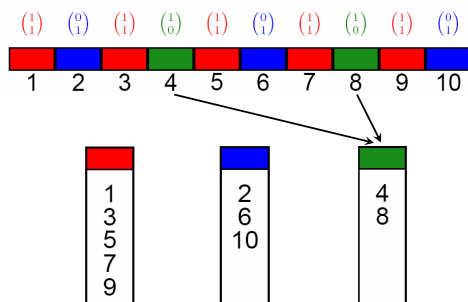
Indexing with inverted lists



46

Motion Retrieval

Indexing with inverted lists



47

Motion Retrieval

Preprocessing (Index)

- 3 hours of Mocap data
- 31 (manually designed) boolean features

Database	Index
1,200,000 frames	230,000 segments
370 MB	7.54 MB

- Index construction: 376 seconds
- Index time and index size **linear** in #(segments)
- Index is **query independent**

48

Motion Retrieval

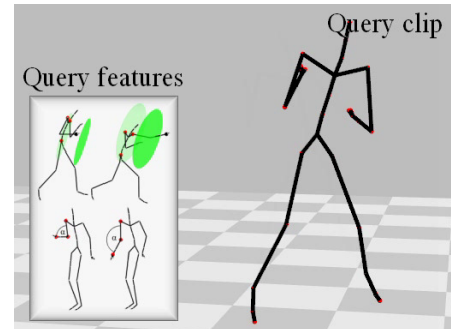
Query and retrieval stage

- Query motion clip
- Optional selection of preferences
 - feature selection
 - degree of fault tolerance
 - ranking strategy
- Automatic conversion of query into feature sequence
- Retrieving hits based on inverted lists
- Typical query response times: 10-300 ms

49

Motion Retrieval

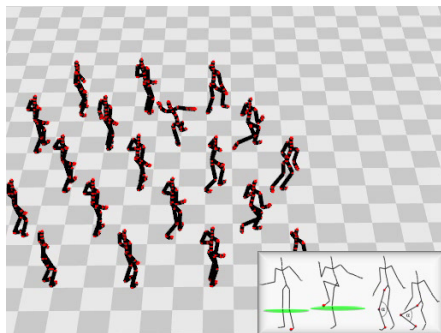
Results: Punch



50

Motion Retrieval

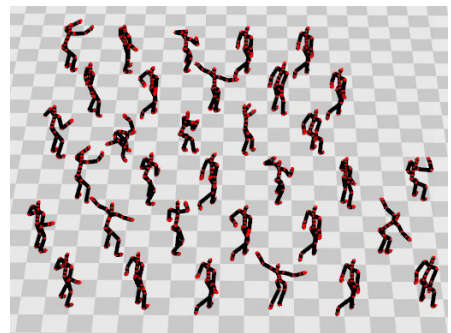
Results: Kick



51

Motion Retrieval

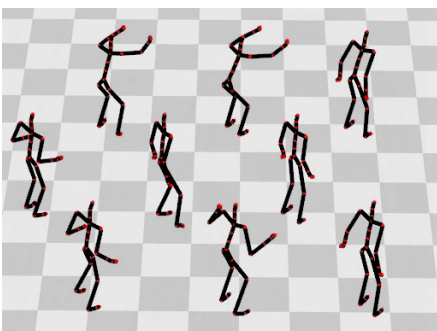
Results: Squat (unranked)



52

Motion Retrieval

Results: Squat (top 9 ranked)



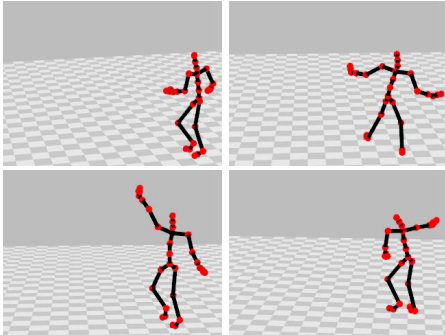
53

Strengths and Weaknesses

	Strength	Weakness
Retrieval	Efficiency	Rigid False positives/negatives Ranking?
Feature Design	Clear semantics	Ad-hoc Automation?
Feature Selection	A-priori knowledge	Critical Automation

54

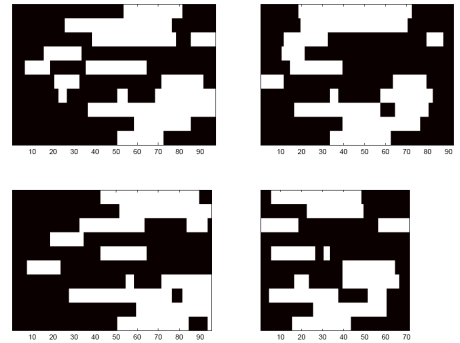
Motion Templates



Müller/Röder (SCA 2006)

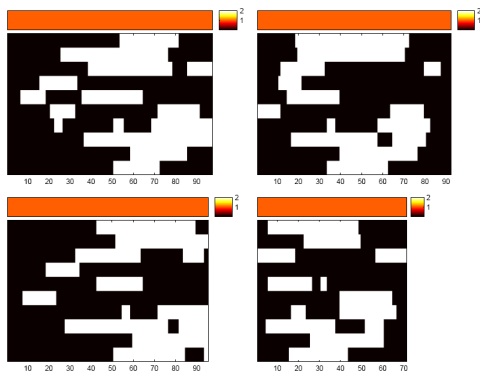
55

Motion Templates



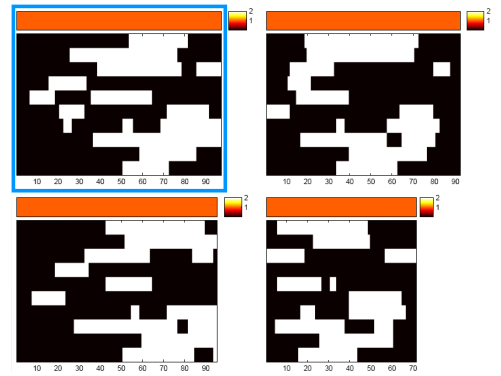
56

Motion Templates



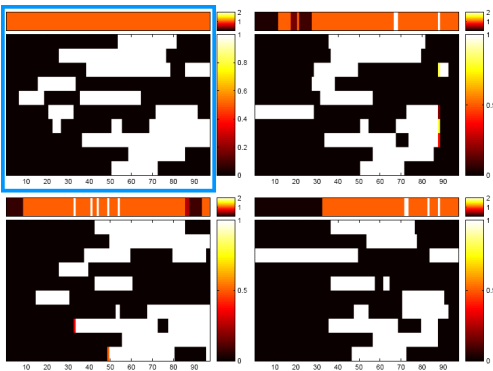
57

Motion Templates



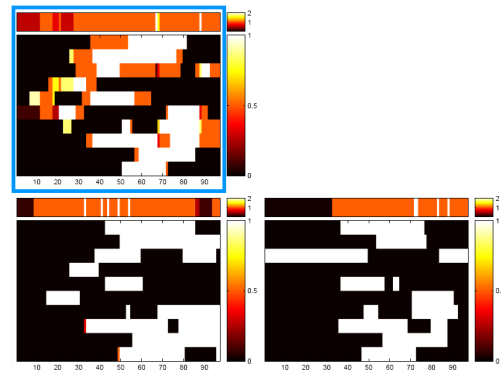
58

Motion Templates



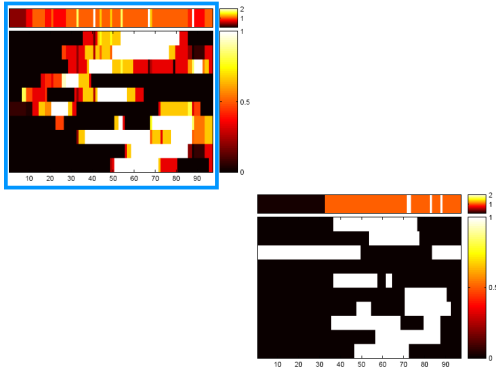
59

Motion Templates



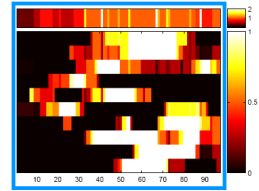
60

Motion Templates



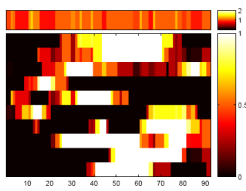
61

Motion Templates



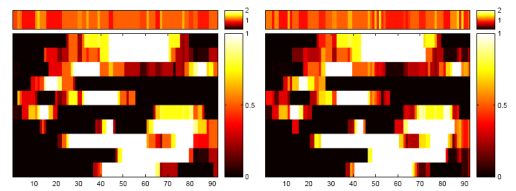
62

Motion Templates



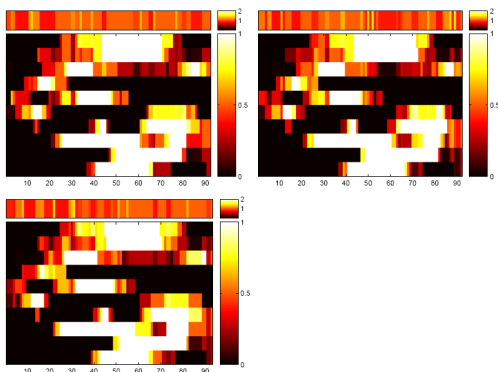
63

Motion Templates



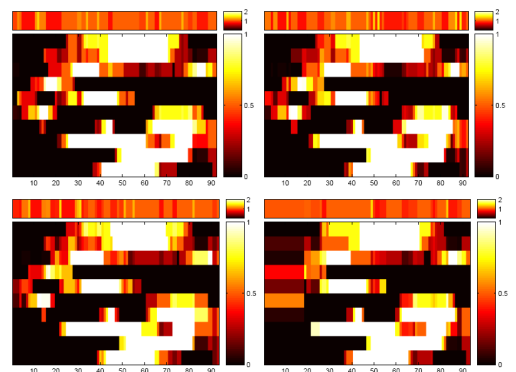
64

Motion Templates



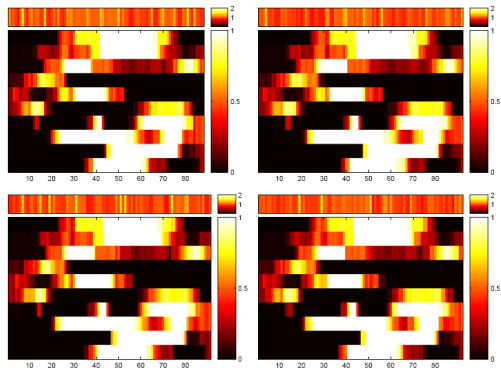
65

Motion Templates



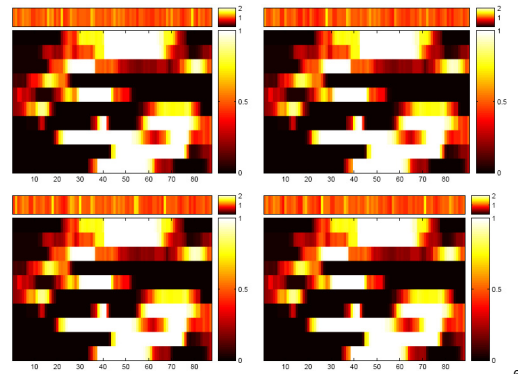
66

Motion Templates



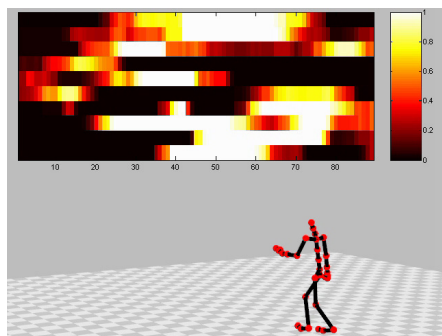
67

Motion Templates



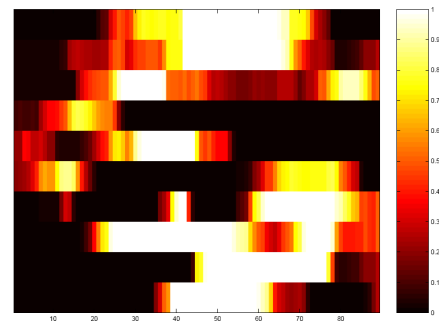
68

Motion Templates



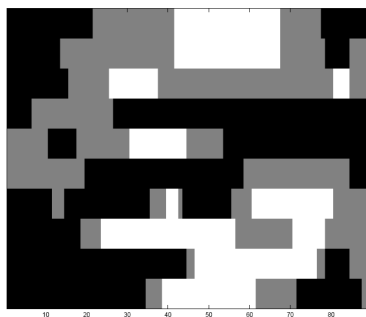
69

Motion Templates



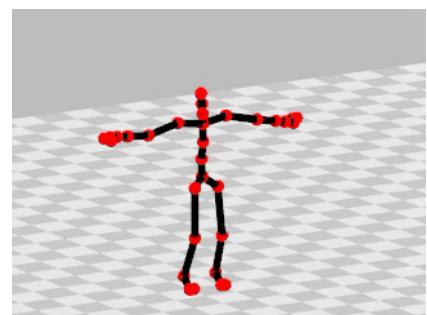
70

Motion Templates



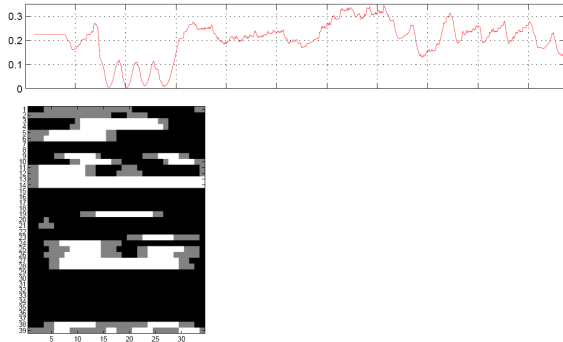
71

MT-based Motion Retrieval



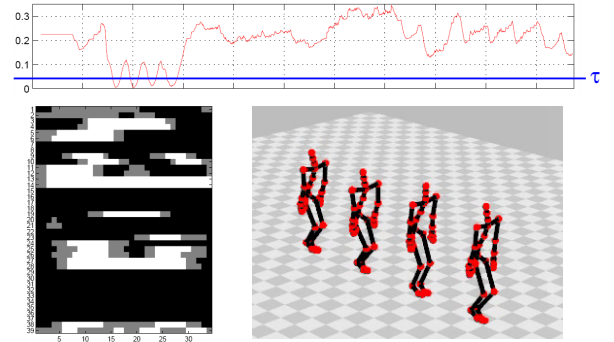
72

MT-based Motion Retrieval: Jumping Jack



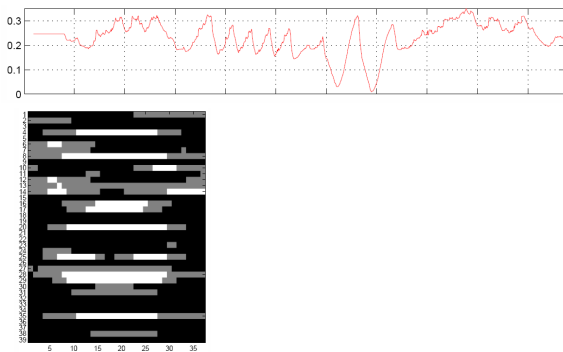
73

MT-based Motion Retrieval: Jumping Jack



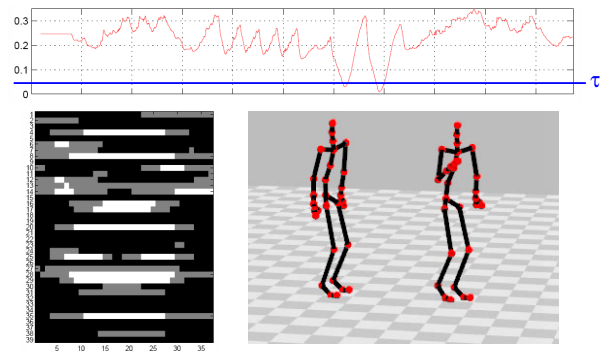
74

MT-based Motion Retrieval: Elbow-To-Knee



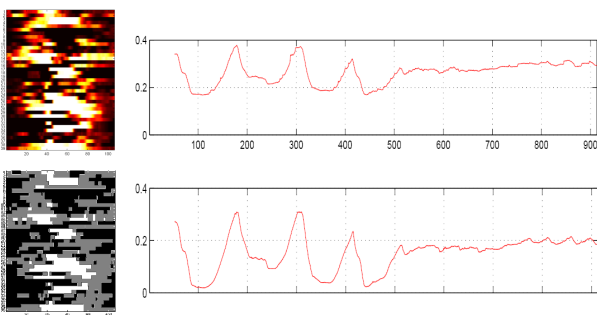
75

MT-based Motion Retrieval: Elbow-To-Knee



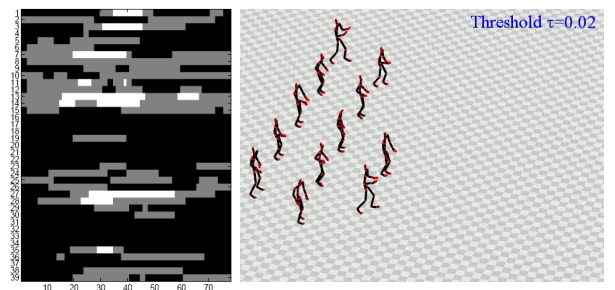
76

MT-based Motion Retrieval: Cartwheel



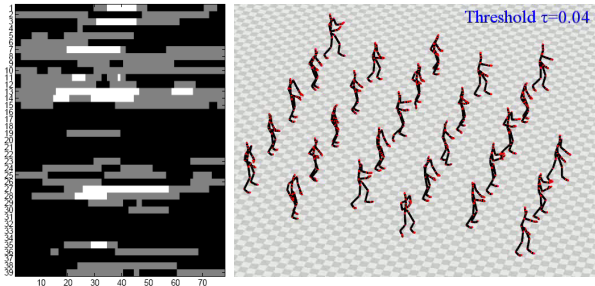
77

MT-based Motion Retrieval: Throw



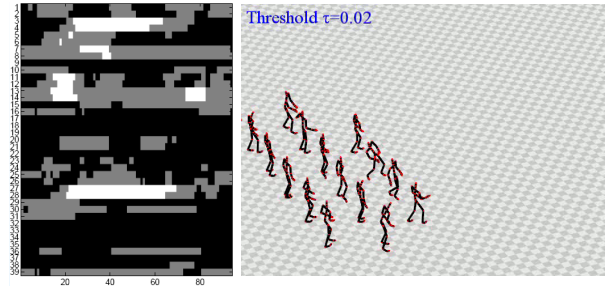
78

MT-based Motion Retrieval: Throw



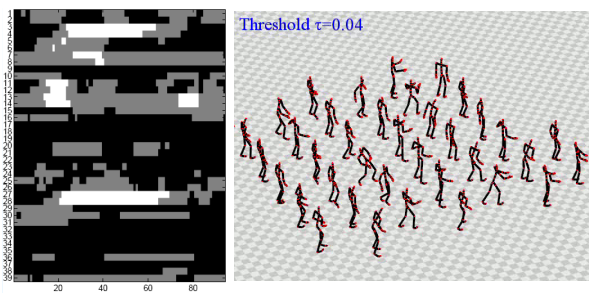
79

MT-based Motion Retrieval: Basketball



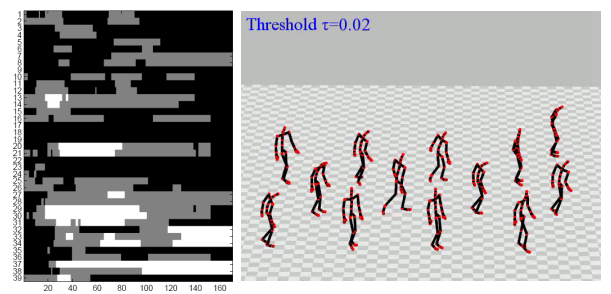
80

MT-based Motion Retrieval: Basketball



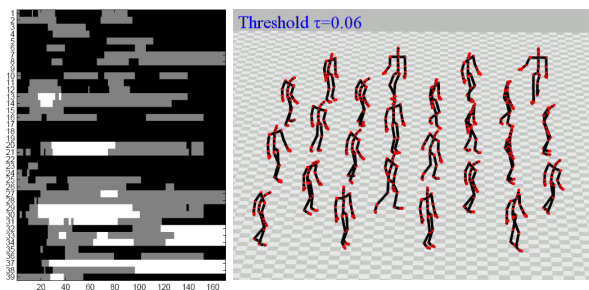
81

MT-based Motion Retrieval: Lie Down Floor



82

MT-based Motion Retrieval: Lie Down Floor



83

Problems and Future Work

- **Efficiency:** MT-based matching is linear in database size
- **Hit quality:** MT-based matching has problems with short motions with few characteristic aspects
- **Current work:** Combine MT-based matching with aspects of exact matching:
 - “Hard constraints” such as keyframes
 - Index-based preselection

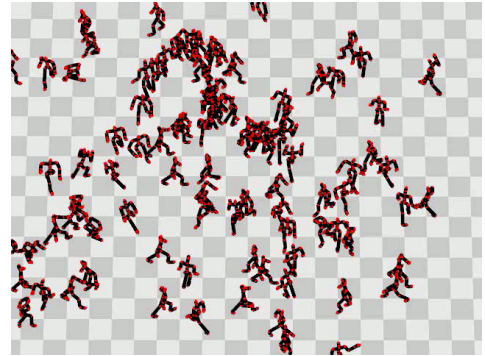
84

Conclusions

- Automated data organization
- Handling object deformations
- Handling multimodality
- Synchronization (alignment)
- Efficiency

85

Conclusions



86