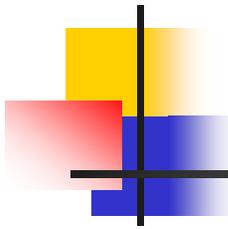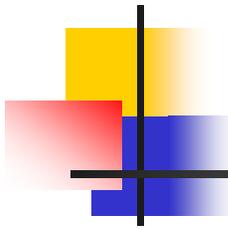# Efficient Content-Based Retrieval of Motion Capture Data

Meinard Muller, Tido Roder, Micheal Clausen

University of Bonn

SIGGRAPH 2005

Problem

Identify and extract those motion clips from a large motion database that are in some sense "*logically similar*" to a query motion (and user specification)

Crucial Notion - Similarity

- Two motion clips "logically similar" if the are variation (spatial or temporal) of same action
- Logical Similarity ~ Spatio-temporal invariance
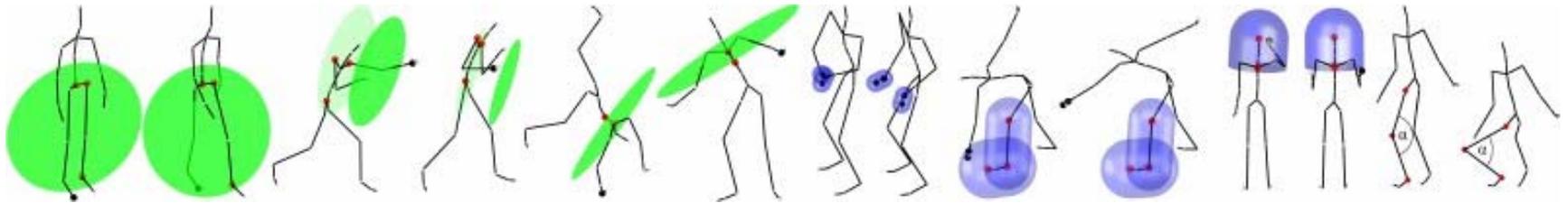- Logical Similarity != Numerical Similarity
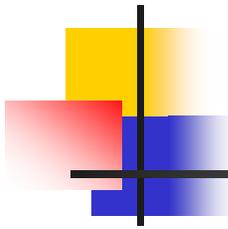
Approach

- Geometric Features
  - Describe motion at high semantic level
  - Robust against spatial variations
- Adaptive Segmentation
  - Robust against temporal variations
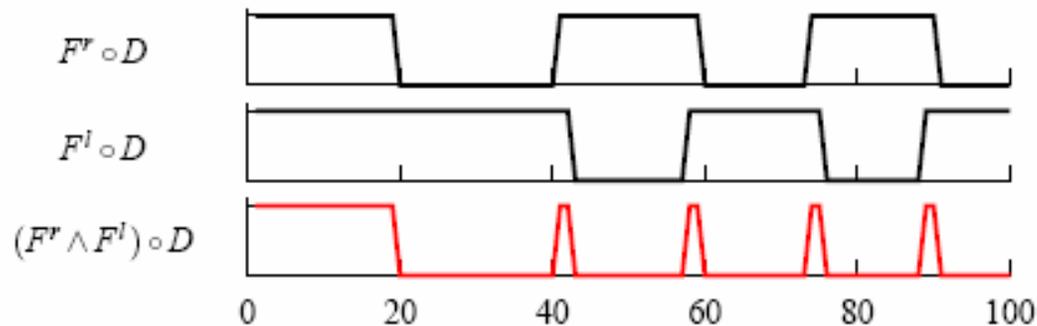 - Transforms motion into sequence of geometric configurations

Geometric Features

- Boolean Functions
- Describe relation between body parts

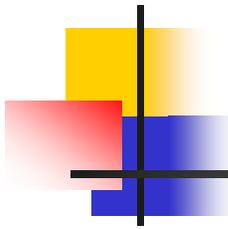# Geometric Features

- Concatenate to from feature vector F(P)

$F^r \circ D$

$F^l \circ D$

$(F^r \wedge F^l) \circ D$

0   20   40   60   80   100

- Invariant to Euclidean motion, scaling, spatial variations
- Coarse description of motion

Geometric Features

- ## Total of 31 features

Feature set	Features and corresponding abbreviations
F_ℓ	right/left foot in front (F_ℓ^1/F_ℓ^2), right/left foot raised (F_ℓ^3/F_ℓ^4), right/left foot fast (F_ℓ^5/F_ℓ^6), right/left knee bent (F_ℓ^7/F_ℓ^8), right/left leg sideways (F_ℓ^9/F_ℓ^{10}), legs crossed (F_ℓ^{11})
F_u	right/left hand in front (F_u^1/F_u^2), right/left hand raised (F_u^3/F_u^4), right/left arm sideways (F_u^5/F_u^6), right/left elbow bent (F_u^7/F_u^8), right/left hand fast (F_u^9/F_u^{10}), arms crossed (F_u^{11}), hands touching (F_u^{12})
F_m	right/left hand touching any leg (F_m^1/F_m^2), right/left hand touching head or neck (F_m^3/F_m^4), right/left hand touching hip area (F_m^5/F_m^6), torso bent (F_m^7), root fast (F_m^8)

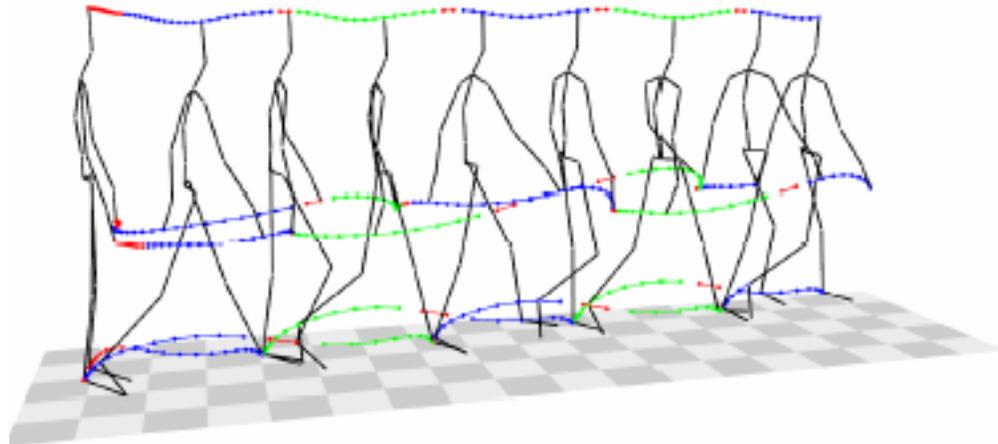- ## Divide post-dependant space into octant
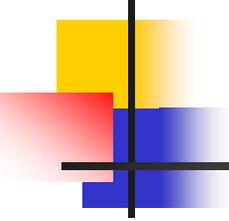
Geometric Features-Critique

- How many to select
  - Over or under specification
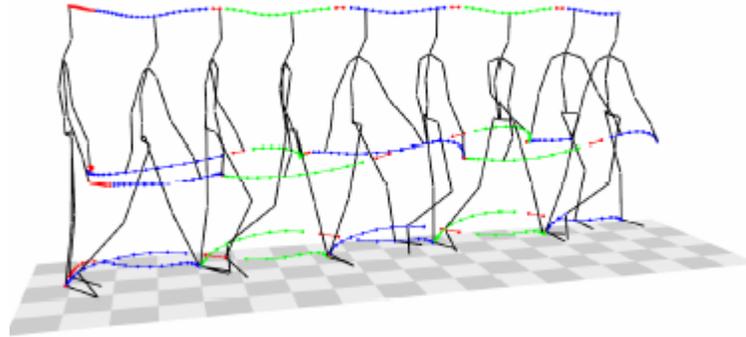- How to select
- Broad range of motions

Adaptive Segmentation

- Two poses F-equivalent if $F(P_1)=F(P_2)$
- F-segment is a sequence of consecutive F-equivalent poses
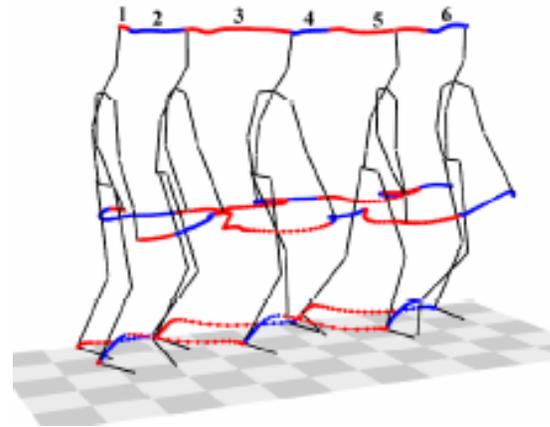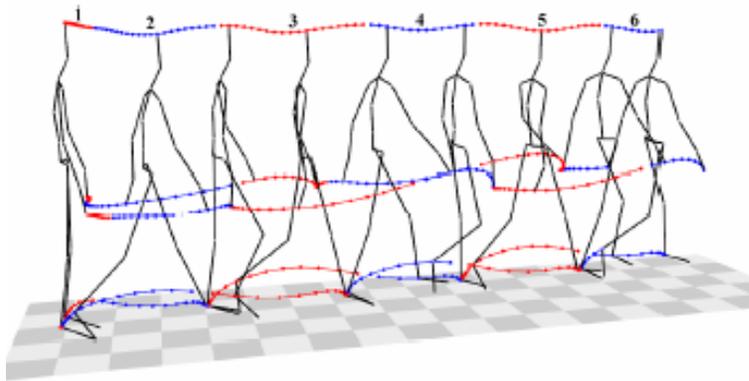
Adaptive Segmentation

- Each segment correspond to unique feature vector

- Segments induce feature vector sequence F[D]

$$F^2[D_{\text{walk}}] = \left(\binom{1}{1}, \binom{0}{1}, \binom{1}{1}, \binom{1}{0}, \binom{1}{1}, \binom{0}{1}, \binom{1}{1}, \binom{1}{0}, \binom{1}{1}, \binom{0}{1}\right).$$
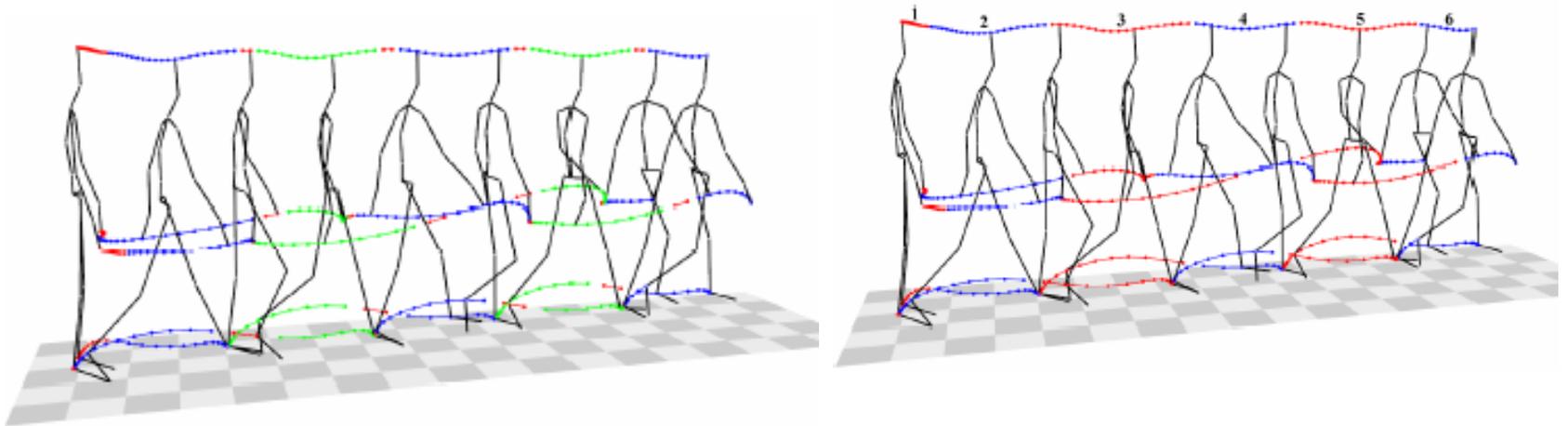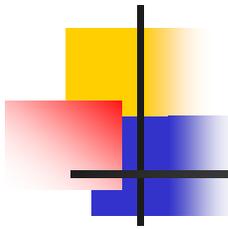
Adaptive Segmentation

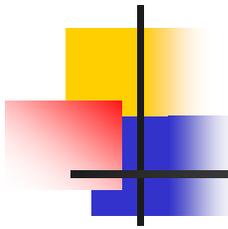- Time invariance

Adaptive Segmentation

- Adapts to selected features
  - Short F = long segments = coarser description

Adaptive Segmentation

- Compare at "*segment level*" rather than "*frame level*"
  - Use F-feature sequence F[D]
  - Computational saving
  - Spatio-temporal invariance incorporated

Indexing and Retrieval

- Fix feature set

- Construct query-independent index structure of F – inverted list

$$L\left(\binom{1}{1}\right) = \{1,3,5,7,9\}, \ L\left(\binom{0}{1}\right) = \{2,6,10\}, \ L\left(\binom{1}{0}\right) = \{4,8\}, \ L\left(\binom{0}{0}\right) = \emptyset.$$

$$F[\mathcal{Q}] = \left(\binom{1}{0},\binom{1}{1},\binom{0}{1}\right)$$

$$
\begin{aligned}
H_{\mathcal{D}}(F[\mathcal{Q}]) &= L\left(\binom{1}{0}\right) \cap \left(L\left(\binom{1}{1}\right) - 1\right) \cap \left(L\left(\binom{0}{1}\right) - 2\right) \\
&= \{4,8\} \cap \{0,2,4,6,8\} \cap \{0,4,8\} \\
&= \{4,8\}.
\end{aligned}
$$
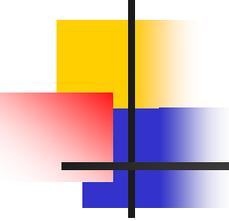
Indexing and Retrieval

- Large memory requirement for indexing
  - Solution:
    - Break feature set, Fu, Fl, Fm
    - Query individual indexes
    - Post process with additional merging/intersection

- Exact match at feature and segment level while still allowing lot of variability at frame level
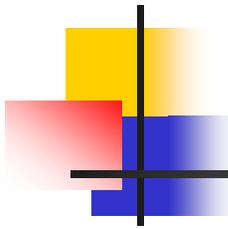
Results - Indexing

- O(n) in number of frames
- Index size proportional to # segments
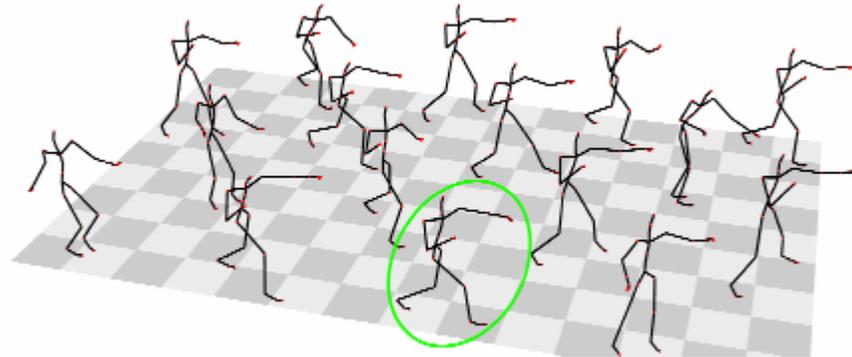- #segments = 3 to 5% of # framesc

Index	f	2^f	#(lists)	#(frames)	#(segs)	MB	$\frac{bytes}{seg}$	t_r	t_f	t_i	Σt
I_ℓ^{60}	11	2048	409	425,294	21,108	0.72	35.8	26	10	6	42
I_ℓ^{180}	11	2048	550	1,288,846	41,587	1.41	35.5	71	26	13	110
I_u^{60}	12	4096	642	425,294	53,036	1.71	33.8	26	13	10	49
I_u^{180}	12	4096	877	1,288,846	135,742	4.33	33.4	71	33	25	129
I_m^{60}	8	256	55	425,294	19,067	0.60	33.0	26	20	3	49
I_m^{180}	8	256	75	1,288,846	55,526	1.80	34.0	71	54	12	137

Results - Retrieval

- 10,000 random trials

Type, #(segs)	1–9 hits			10–99 hits			$\geq$ 100 hits		
	μ_h	σ_h	μ_t (ms)	μ_h	σ_h	μ_t (ms)	μ_h	σ_h	μ_t (ms)
exact, $|Q| = 5$	3.0	2.4	16	44	28	20	649	567	144
exact, $|Q| = 10$	1.7	1.6	17	34	22	26	239	147	71
exact, $|Q| = 20$	1.1	0.6	19	32	26	36	130	5	52

Results