Motion synthesis

- Goals:
  - generate human motions that “look human” and “do what you want”
  - Synthesis
    - with control; with interaction
  - Evaluation
    - what “looks human?”
- Features
  - Motion composes across the body and across time
    - so the number of available motions is huge
  - Multiple constraints on the appearance of motion
    - physics;
    - motor control system;
    - internal motion goals;
    - nearby objects;

Key problems

- What makes a motion look human?
  - can we tell good motions from bad?
- How do we describe human activities?
  - with what vocabulary? at what time scales?
- How do nearby objects affect our description
  - interactions and context

Motion synthesis difficulties

- People are good at spotting poor motion
  - and it sometimes matters
- Motions can be very fast and very detailed
  - high accelerations, contacts create major issues
- Authoring is mysterious
  - how does one specify constraints on activity usefully?
- Complexity
  - interactions with objects, etc. create a need for families of motion
  - motion composes in nasty ways
  - motions should interact with objects, users, etc.
- Control
  - character should be manageable
  - have some capability to cope on its own

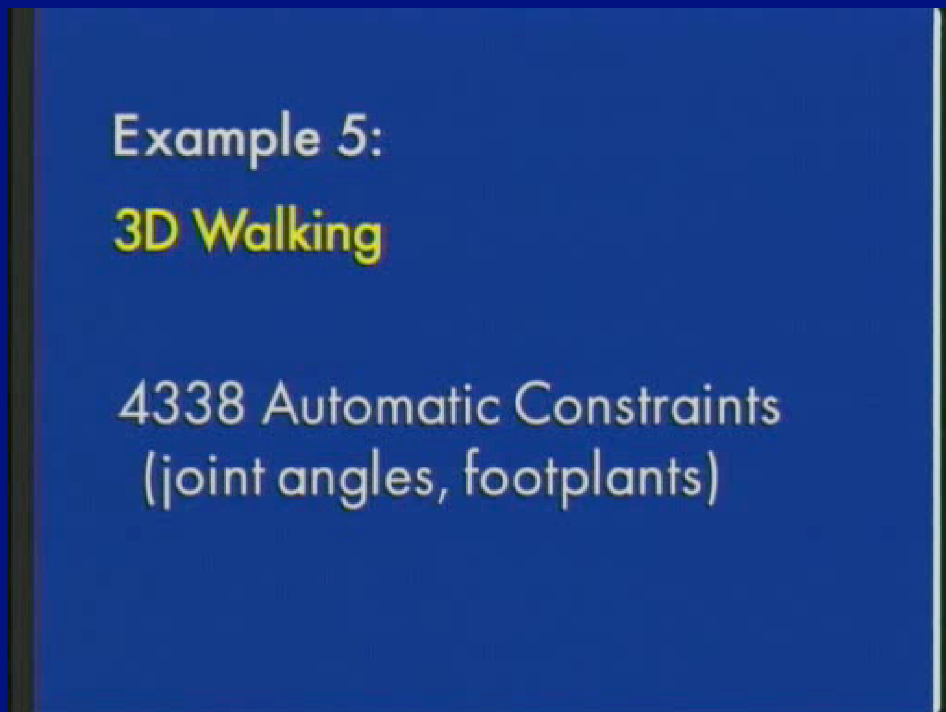
Motion synthesis, cont

- Motion composes across the body and across time
  - so the number of available motions is huge
- Multiple constraints on the appearance of motion
  - physics;
  - motor control system;
  - internal motion goals;
  - nearby objects;

Motion synthesis

- Methods
  - By animator
  - By kinematic control
    - profound difficulties with ambiguity
  - By combining observations
    - old tradition of move trees; also (Kovar et al 02, Lee et al 02, Arikan+Forsyth 02, Arikan et al 03,Gleicher et al 03)
  - By physical models
    - old tradition; (Witkin+Kass, 88; Witkin+Popovic 99; Funge et al 88; Fang+Pollard 03, 04)
  - By biomechanical models
    - old tradition; Liu+Popovic 02; Abe et al 04; Wu+Popovic 03; Liu+Popovic 02)
  - By statistical models
    - old tradition (e.g. Ramsey+Silverman 97); Li et al 02; Safanova et al 04; Mataric et al 99; Mataric 00; Jenkins+Mataric 04;

Variational and Physical Methods

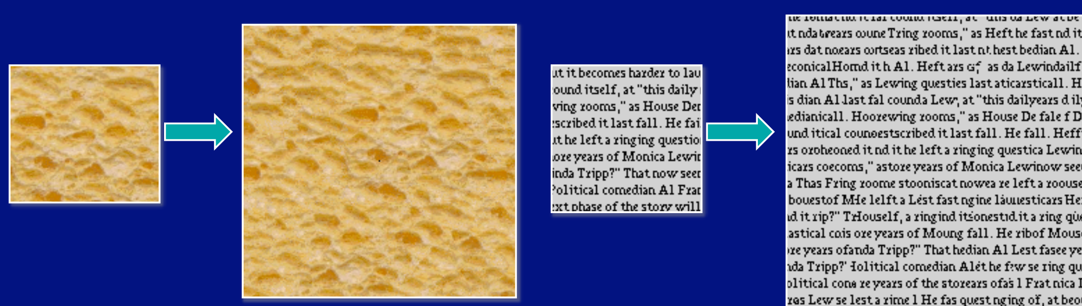


Data-driven motion synthesis

- Analogies
  - Text synthesis (Shannon)

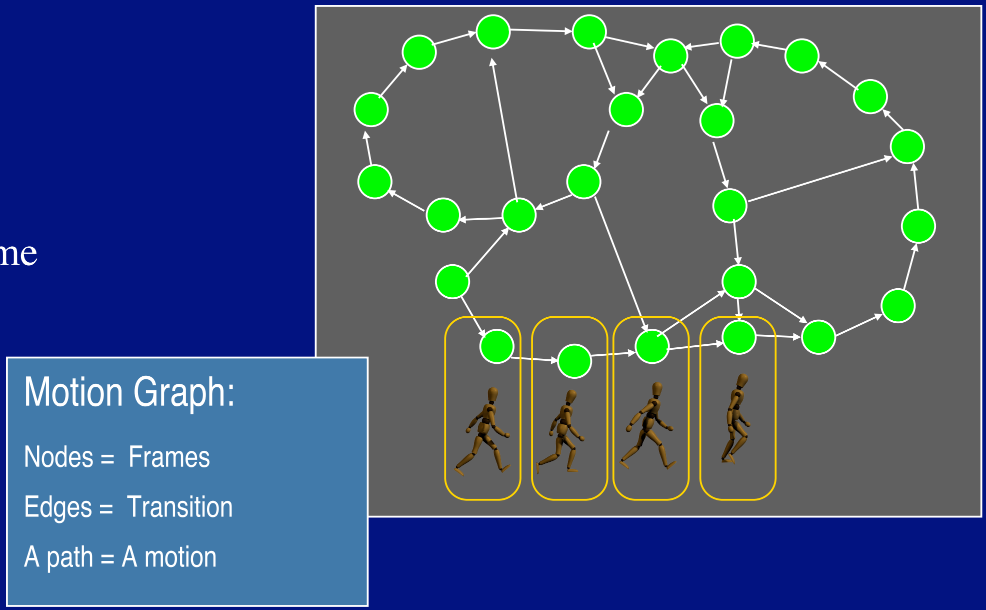
    "It means that in speaking with you, I am aware of how I think this is one of those questions that exposes a contradiction in our cultural cognitive disconnect the concept of authenticity exposes is, I believe, that we have inner and outer selves, and that the inner self is our real self. I personally find those ideas more misleading than helpful."

  - Texture synthesis (Efros+Leung `99; many others since)



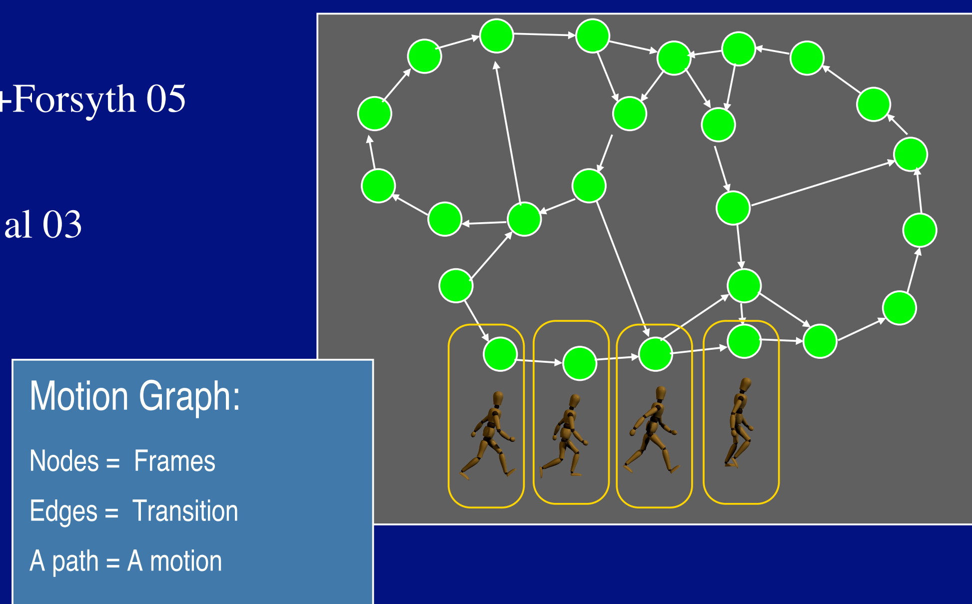
Motion graph

- Take measured frames of motion as nodes
  - from motion capture, given us by our friends
- Directed edge from frame to any that could succeed it
  - decide by dynamical similarity criterion
  - see also (Kovar et al 02; Lee et al 02)
- A path is a motion
- Search with constraints
  - root position+orientation
  - length of motion
  - occupy a frame at specified time
  - limb close to a point



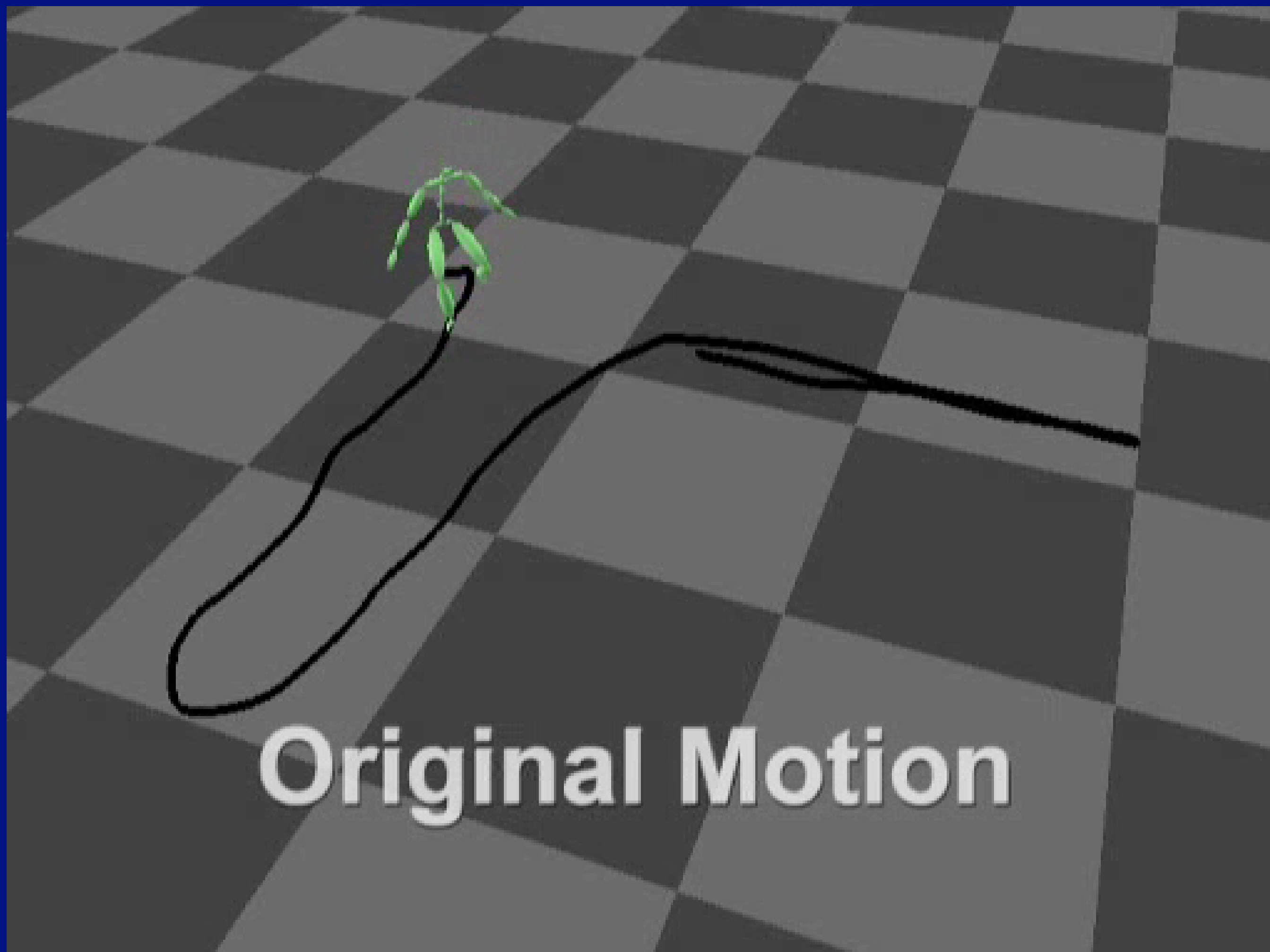
Search in a motion graph

- Local
  - Kovar et al 02
- With some horizon
  - Lee et al 02; Ikemoto, Arikan+Forsyth 05
- Whole path
  - Arikan+Forsyth 02; Arikan et al 03

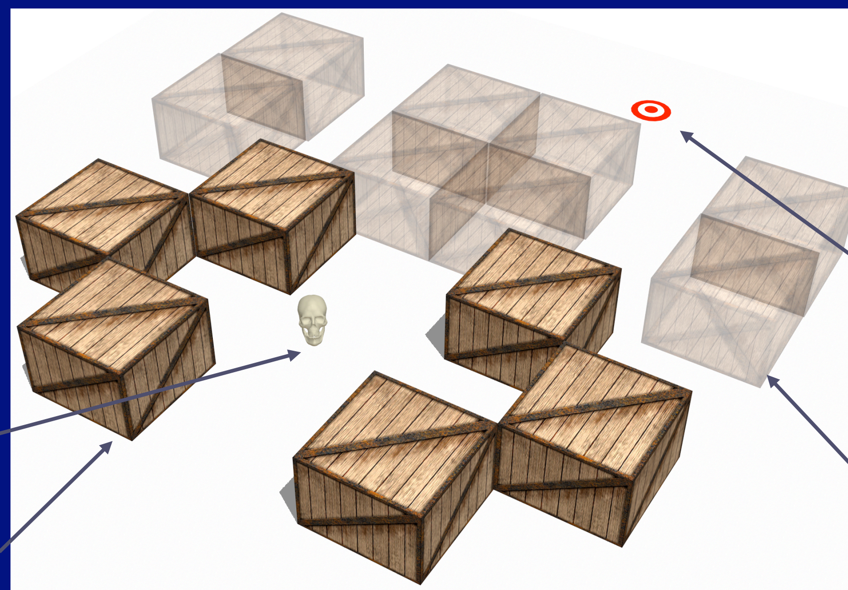


Local Search methods

- Choose the next edge (Kovar, Gleicher, Pighin 02)
  - ensure that one can't get stuck locally
  - but can't guarantee a goal is available on longer scale

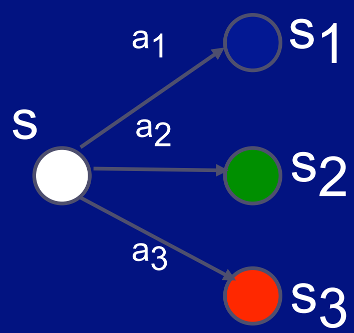

Original Motion

On-line control of motion synthesis



Agent travels along *motion graph*.
When he reaches a decision point, he must choose which branch to take so he can best meet his objectives.

Ikemoto+Arikan+Forsyth 05

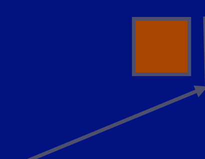


Value of state *s* obtained by comparing to a set of example states,
encoded using following weighted terms

Local geometry

Visible enemies

Distance to next waypoint
on global path plan

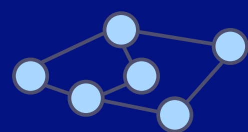


Ikemoto+Forsyth+Arikan 05

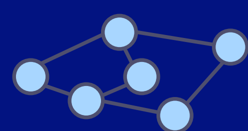
Reinforcement learning

Sample control parameters (*w*) for a random state (*s*)

w_1 → Fix for the motion graph → Generate motion → Reward 1

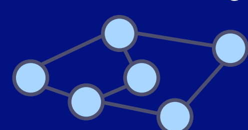


w_2 → Fix for the motion graph → Generate motion → Reward 2



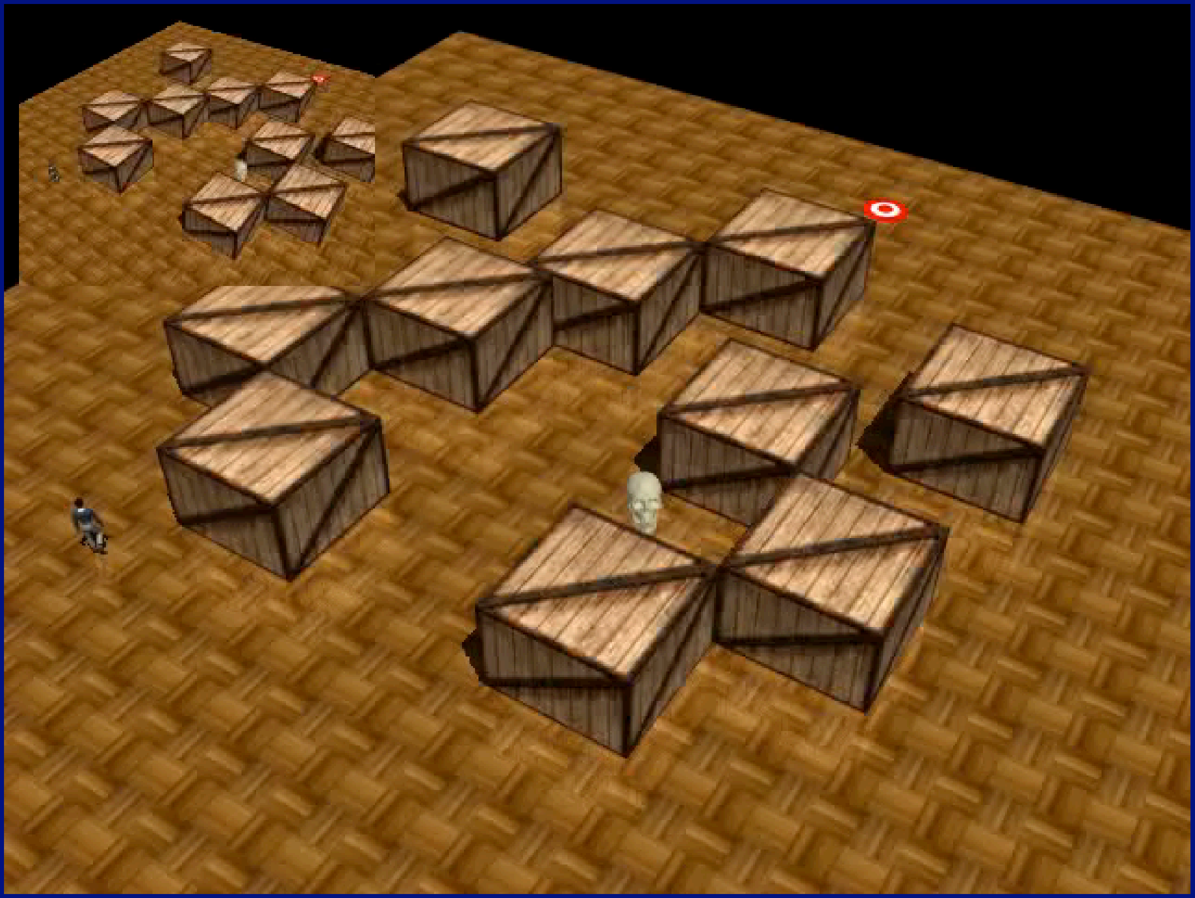
⋮

w_{3000} → Fix for the motion graph → Generate motion → Reward 3000



Fix control parameters for state *s* to be the *w* that yielded maximum reward

Ikemoto+Arikan+Forsyth 05

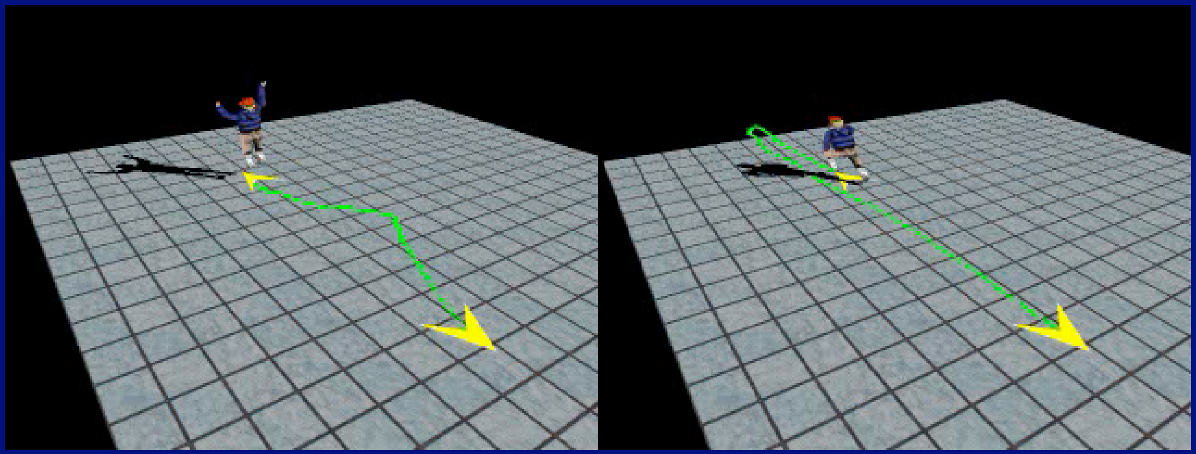


Ikemoto+Arikan+Forsyth 05

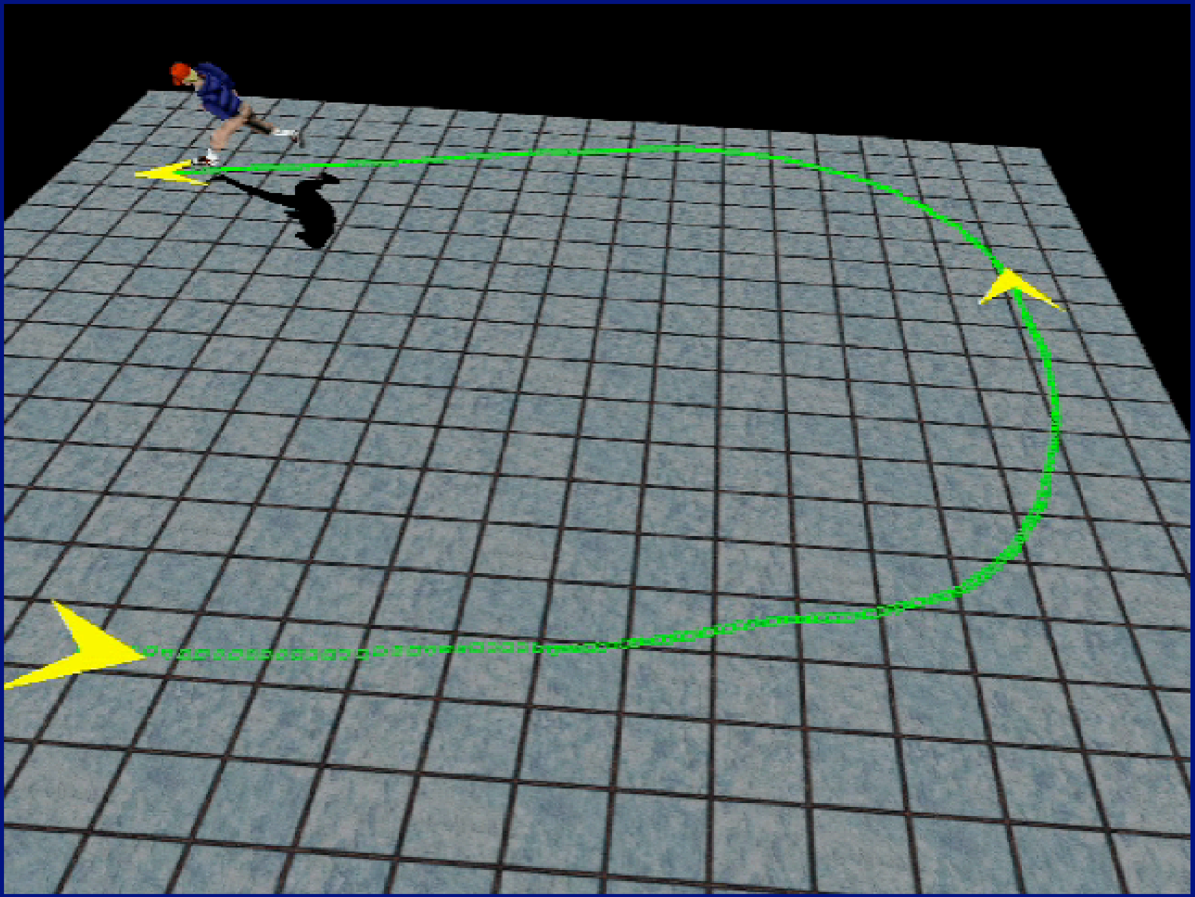
Ikemoto+Arikan+Forsyth 05

Characteristic properties of motion

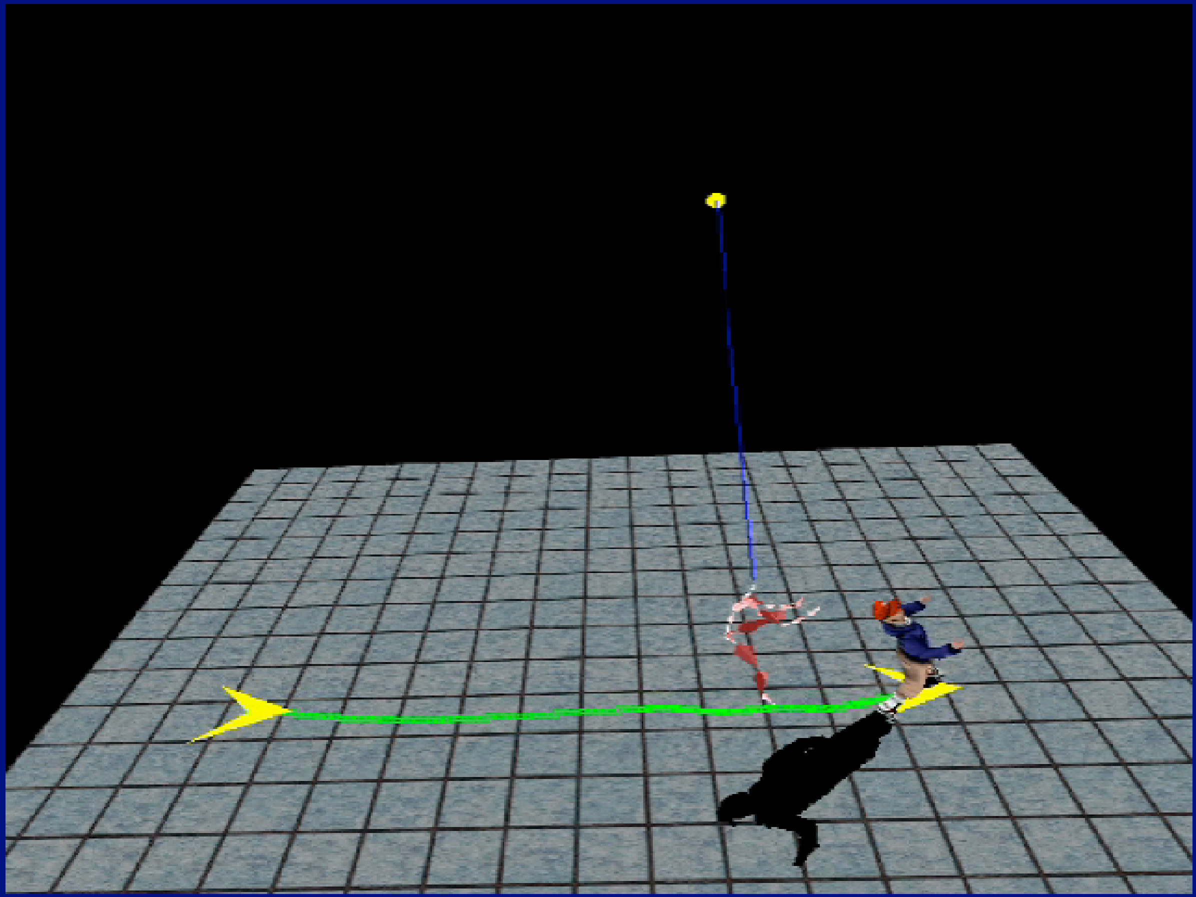
- Characteristic features
  - most demands are radically underconstrained
  - motion is simultaneously
    - hugely ambiguous
    - “low entropy”
- Suggests using “summaries”

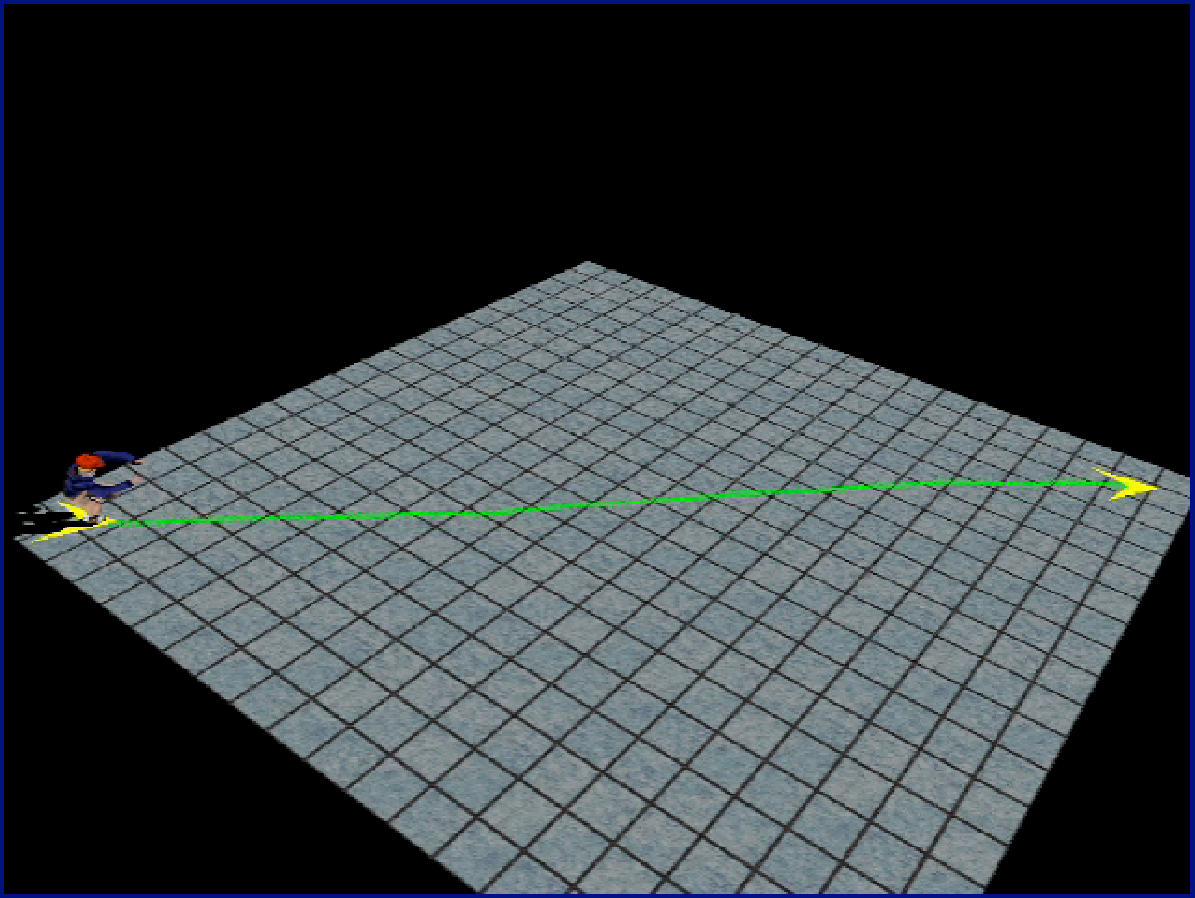


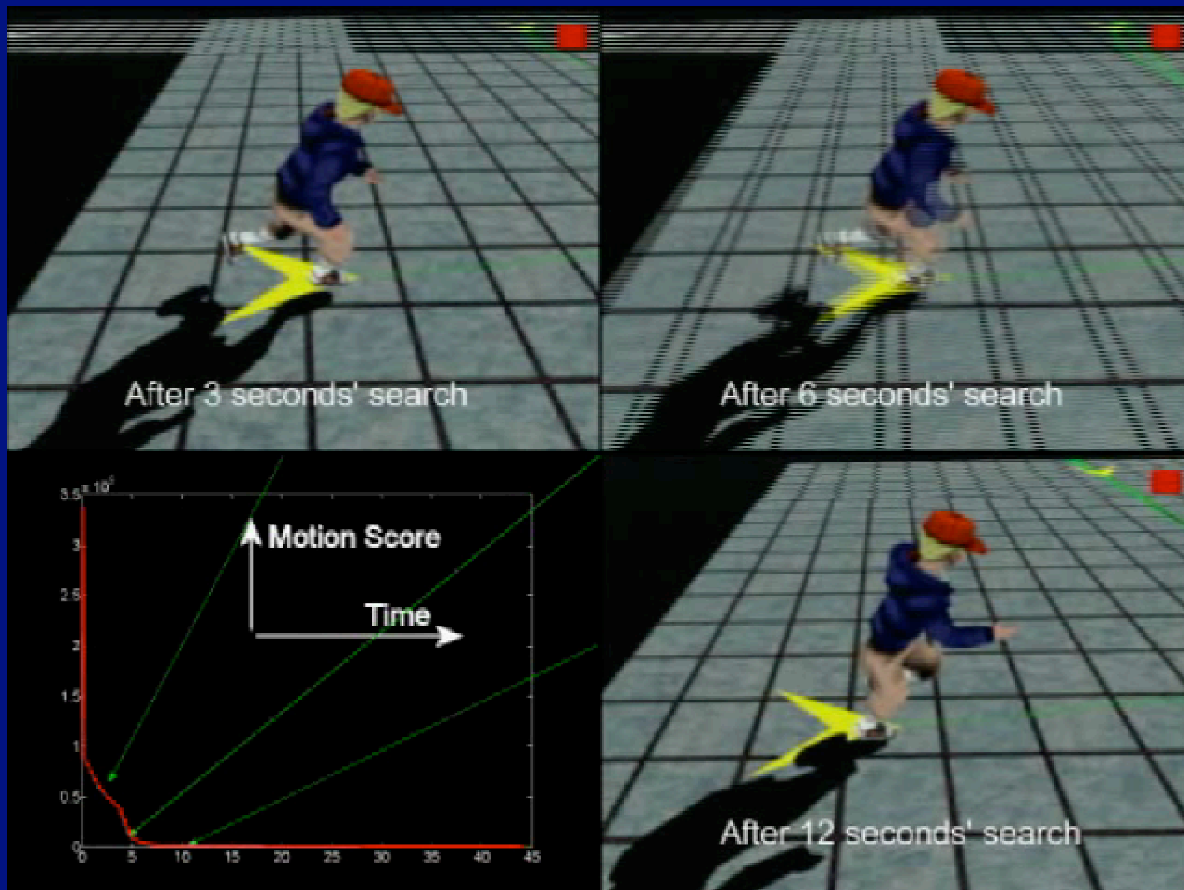
Arikan+Forsyth 02

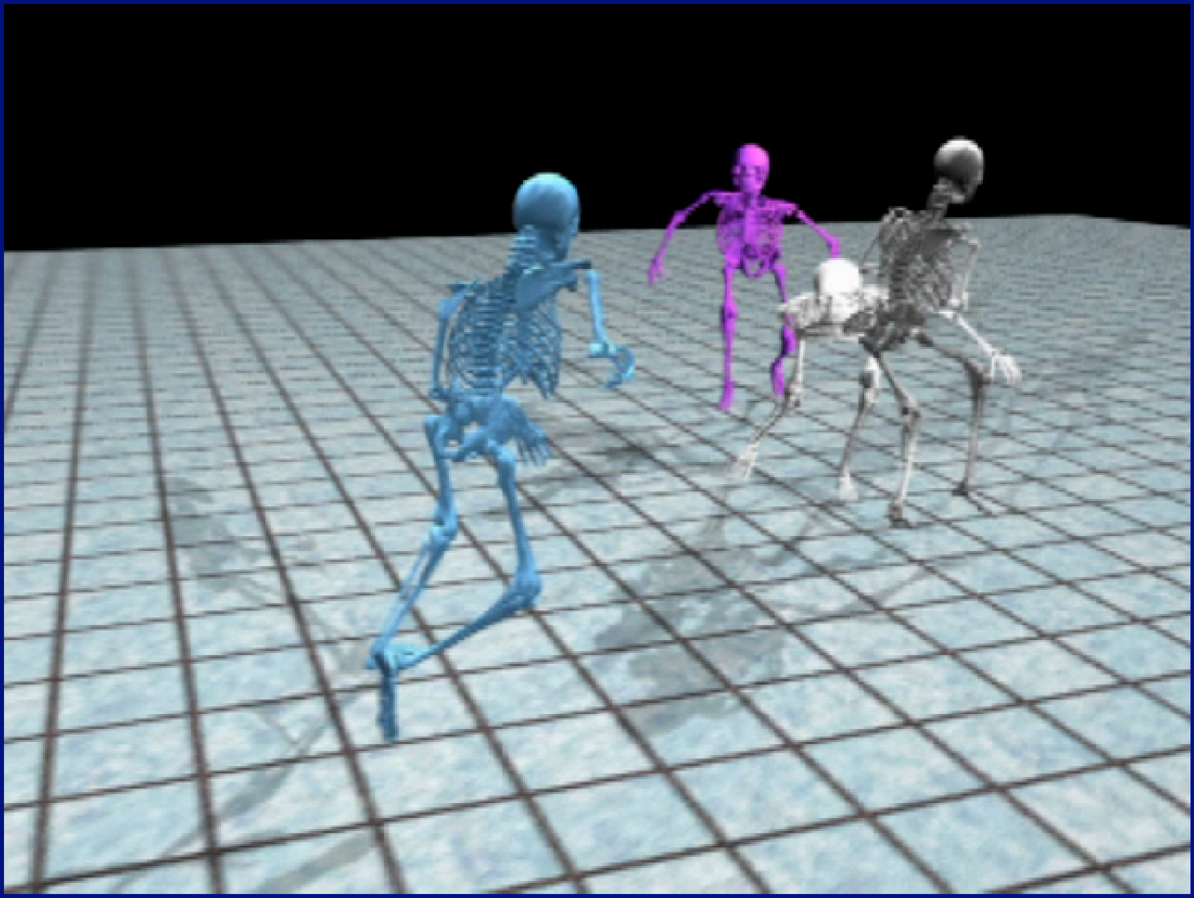


Arikan+Forsyth 02

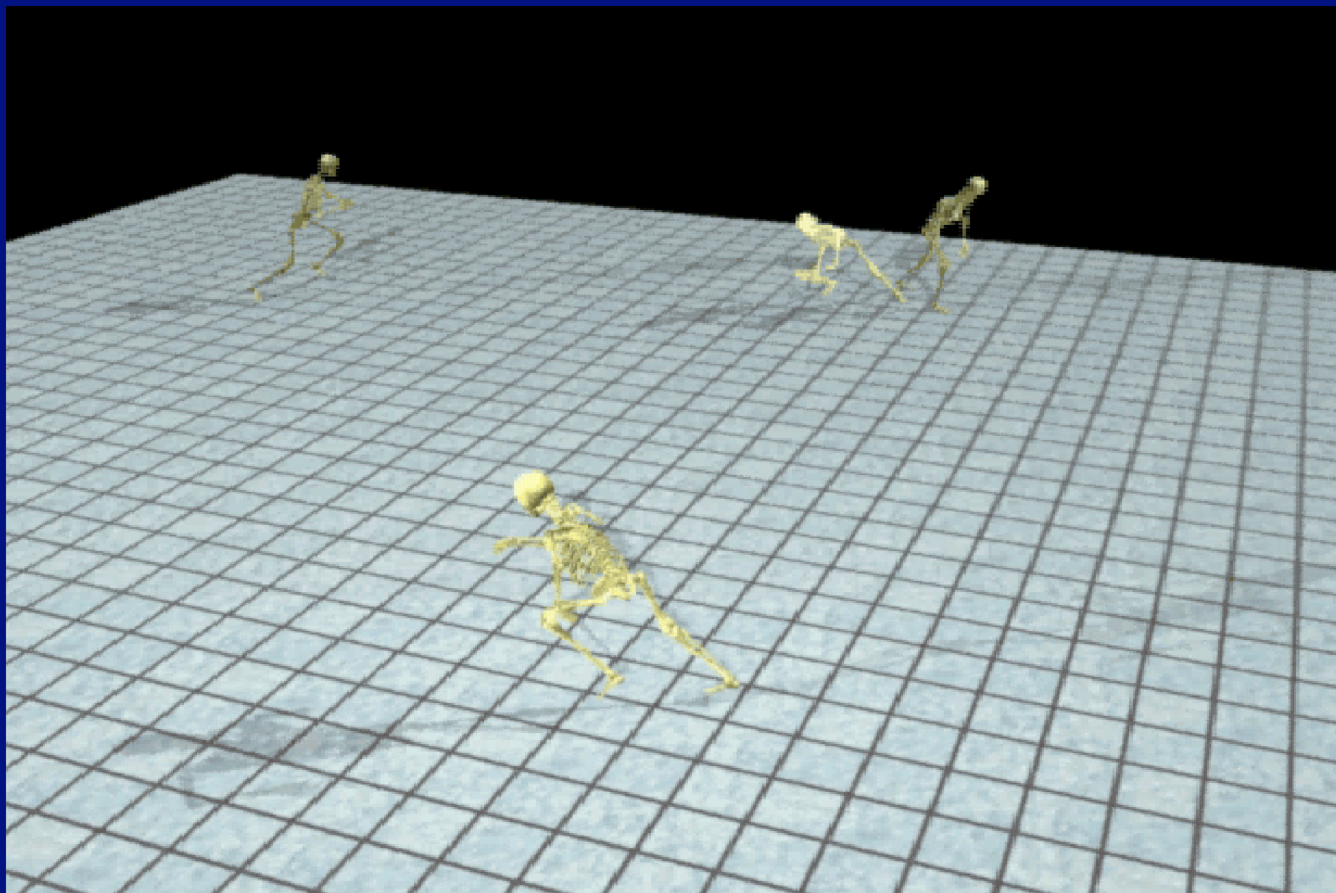




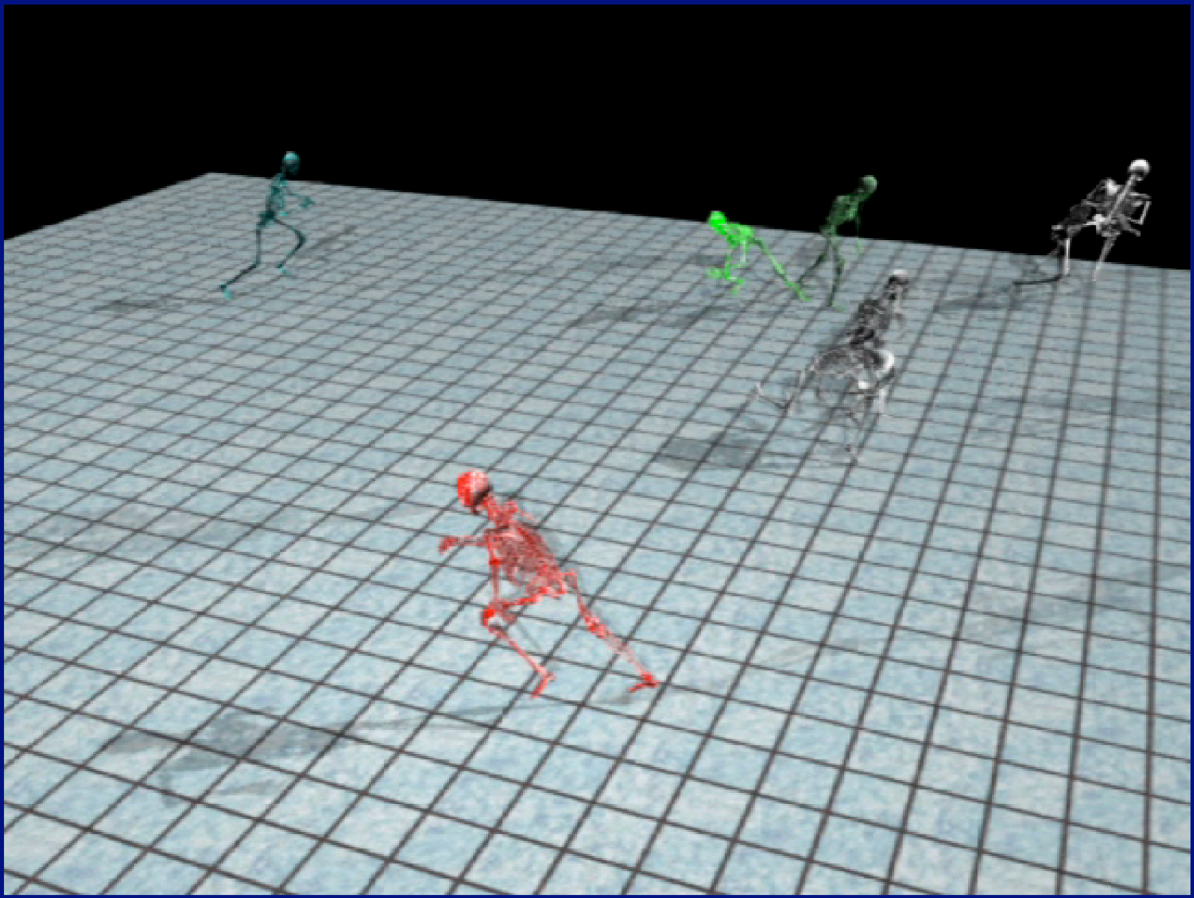

After 3 seconds' search
After 6 seconds' search
Motion Score
Time
After 12 seconds' search



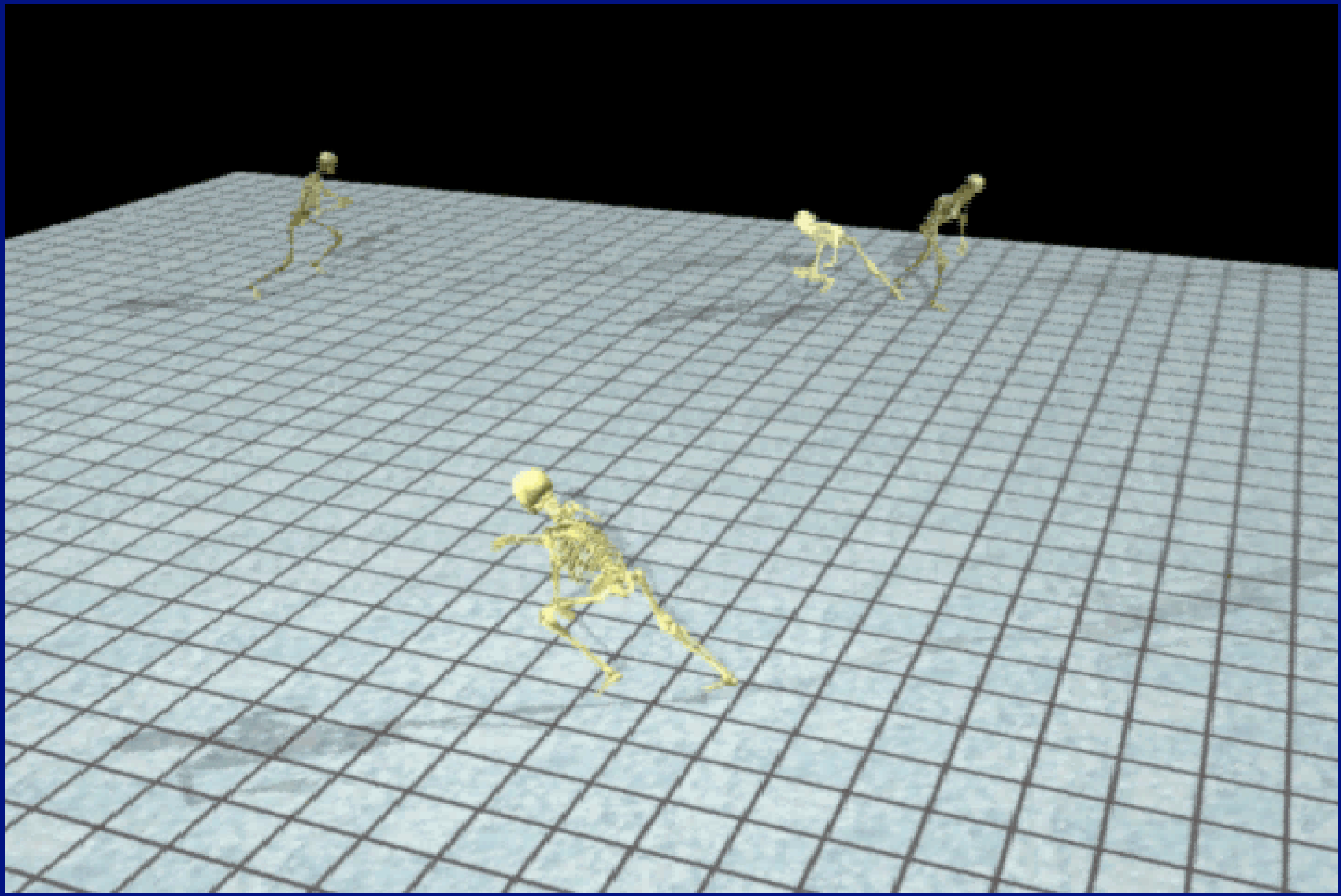
Arikan+Forsyth 02



Arikan+Forsyth 02



Arikan+Forsyth 02



Arikan+Forsyth 02

Limitations

- Can’t synthesize motions one hasn’t seen
  - but see later
- Long term structure of motion is strange
  - running backwards, etc.
- No on-the-fly control of motion or interaction
  - but see later
- Require more detailed control of “type” of motion
  - can deal with this

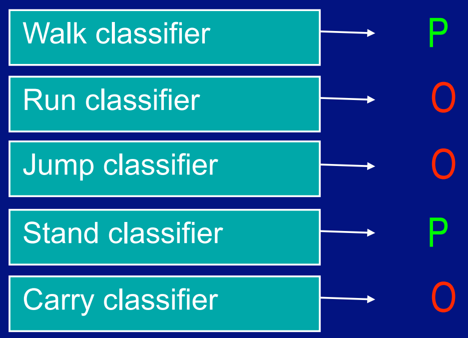
Synthesis with off-line control

- Annotate motions
  - using a classifier and on-line learning
  - efficient human-in-the loop training
- Produce a sequence that meets annotation demands
  - a form of dynamic programming

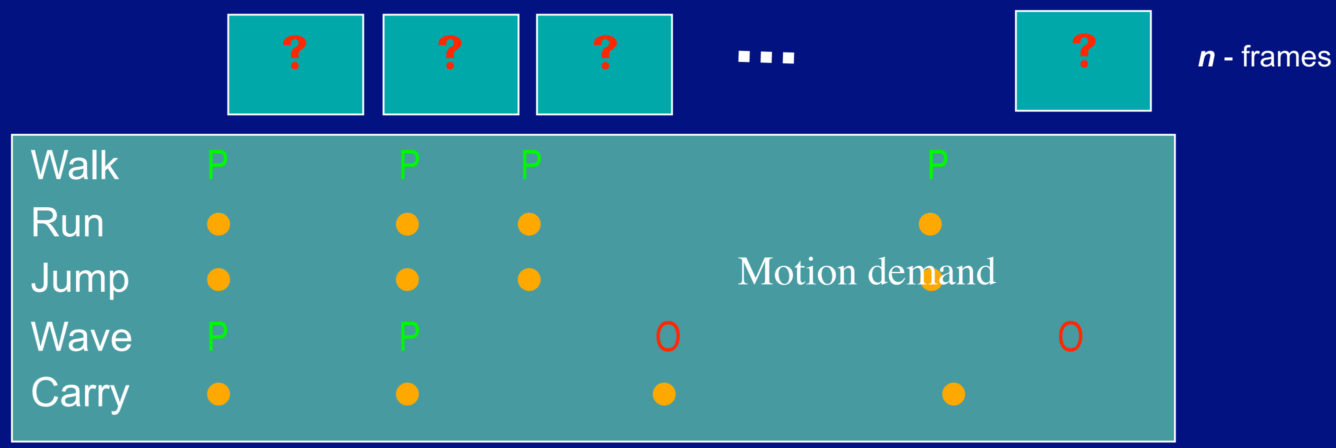
Annotation - desirable features

- Composability
  - run and wave;
- Comprehensive but not canonical vocabulary
  - because we don't know a canonical vocabulary
- Speed and efficiency
  - because we don't know a canonical vocab.

- Can do this with one classifier per vocabulary item
  - use an SVM applied to joint angles
  - form of on-line learning with human in the loop
  - works startlingly well (in practice 13 bits)



Arikan+Forsyth+O'Brien 03



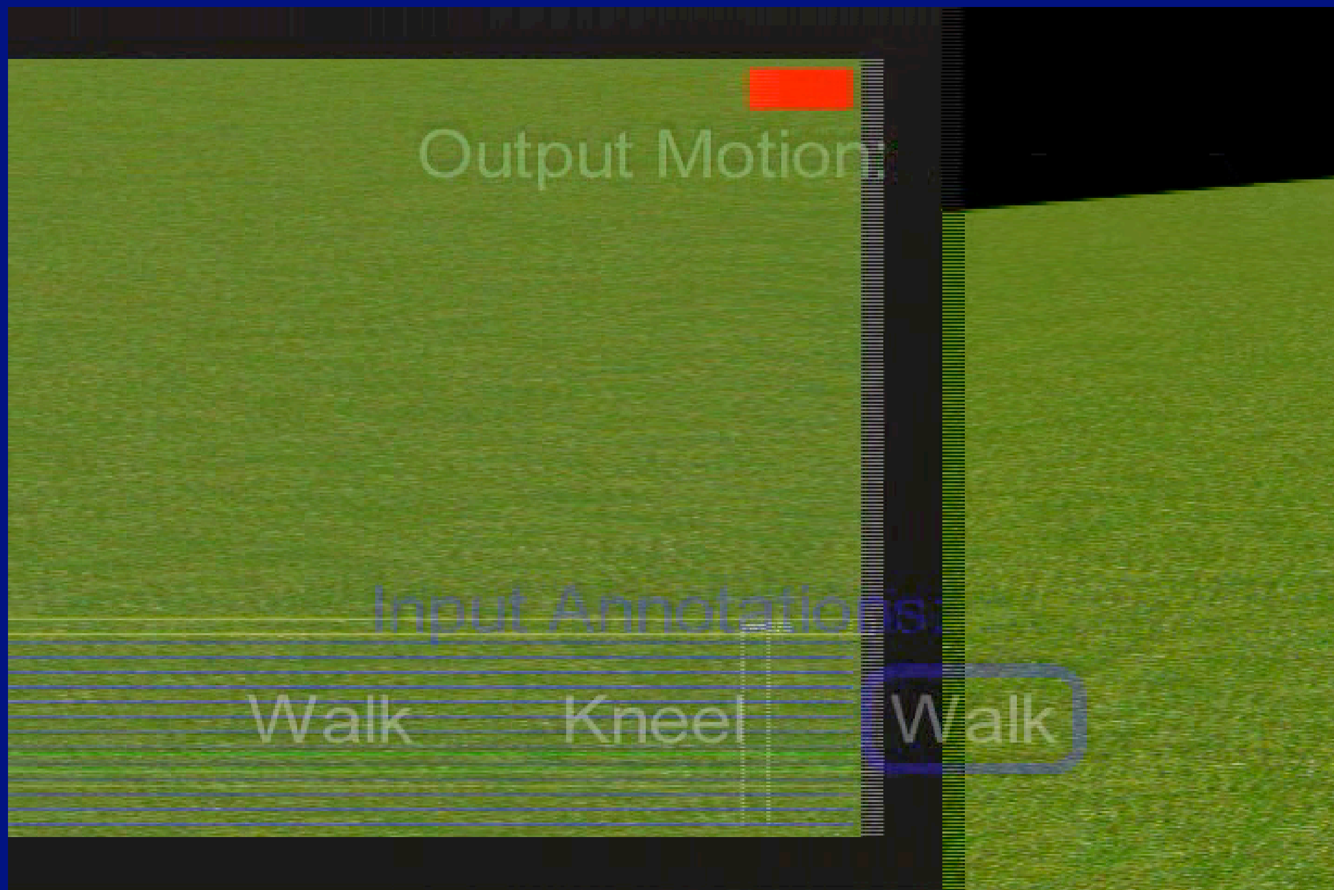
Synthesis by dynamic programming

Arikan+Forsyth+O'Brien 03

Dynamic programming practicalities

- Scale
  - Too many frames to synthesize
  - Too many frames in motion graph
- Obtain good summary path, refine
  - Form long blocks of motion, cluster
  - DP on stratified sample
    - split blocks on “best” path
    - find similar subblocks
      - DP on this lot
        - etc. to 1-frame blocks

Arikan+Forsyth+O’Brien 03



Arikan+Forsyth+O'Brien 03

Arikan+Forsyth+O'Brien 03

Still open

- Local control of synthesis
  - Long term structure of motion is strange
    - running backwards, etc.
  - essential for interaction

- Departing from data?
  - Can’t synthesize motions one hasn’t seen
  - essential for interaction

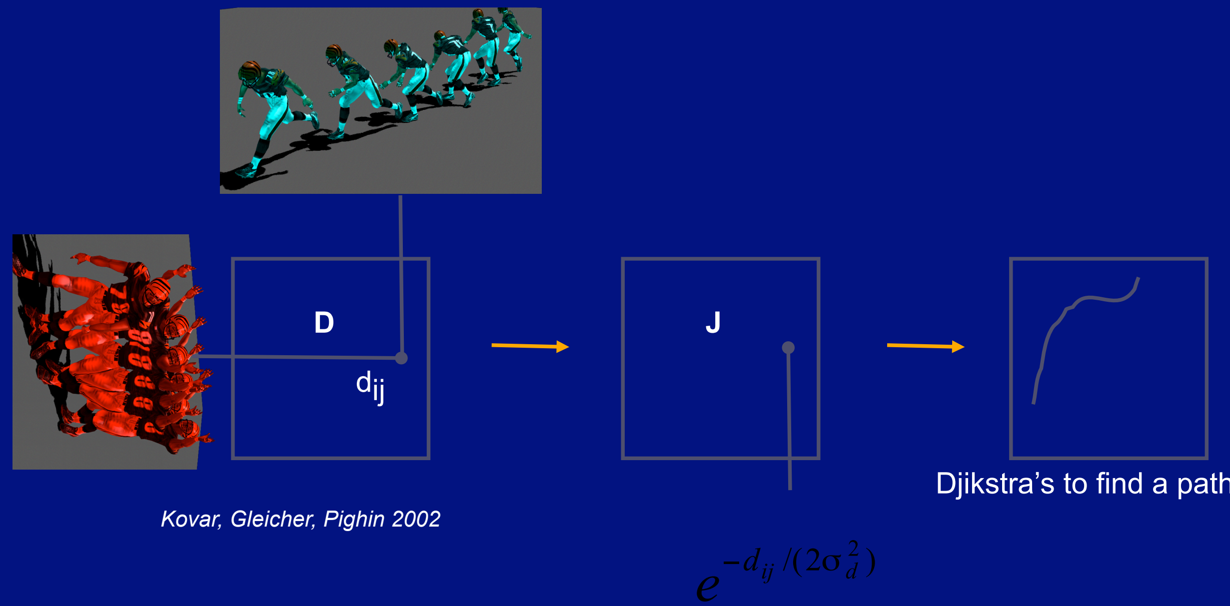
Transplantation

- Motions clearly have a compositional character
  - Why not cut limbs off some motions and attach to others?
    - we get some bad motions
  - build a classifier to tell good from bad
    - avoid foot slide by leaving lower body alone

Ikemoto+Forsyth 04

Loop {

Randomly pick a synthesis rule

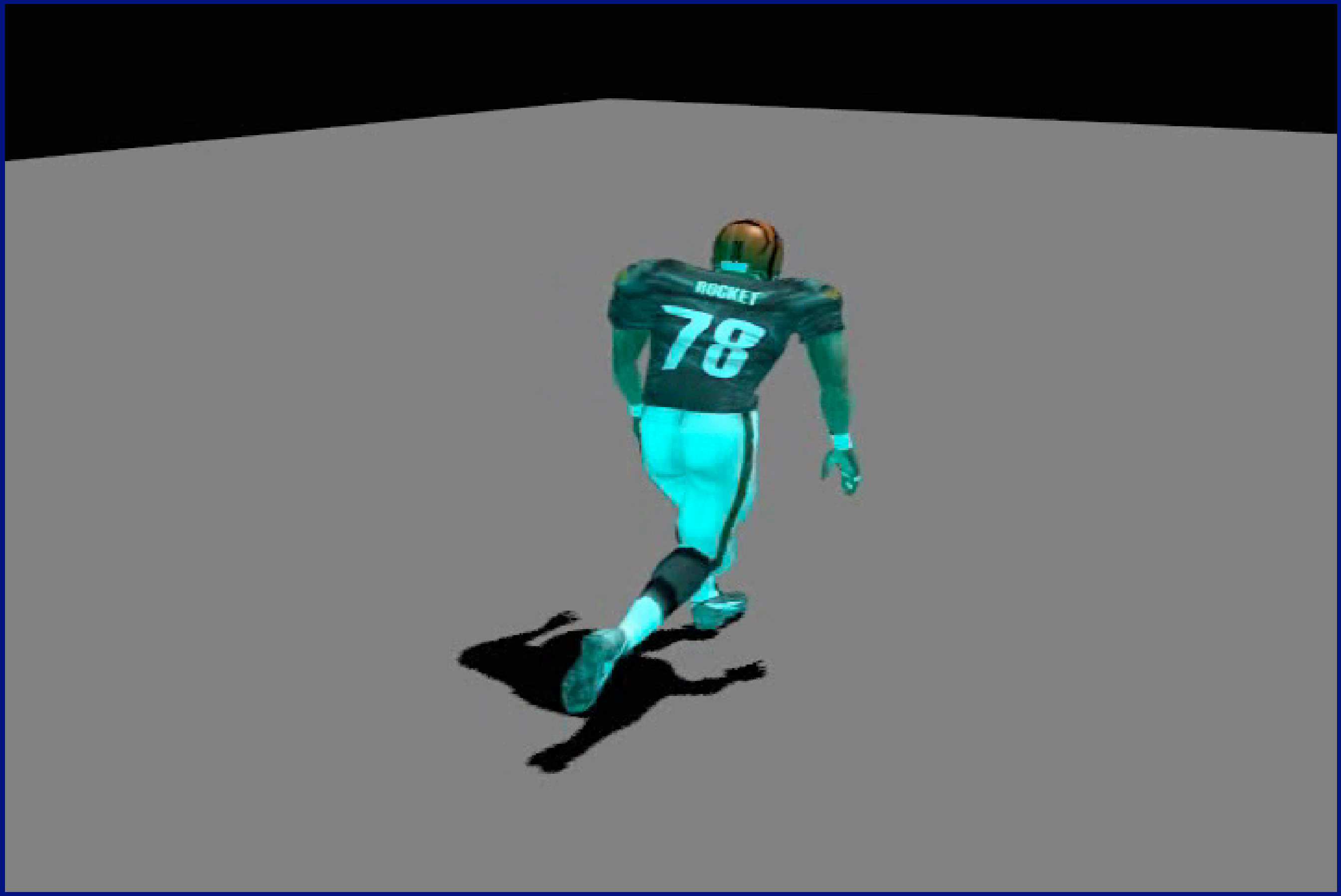


If successful, output candidate motions

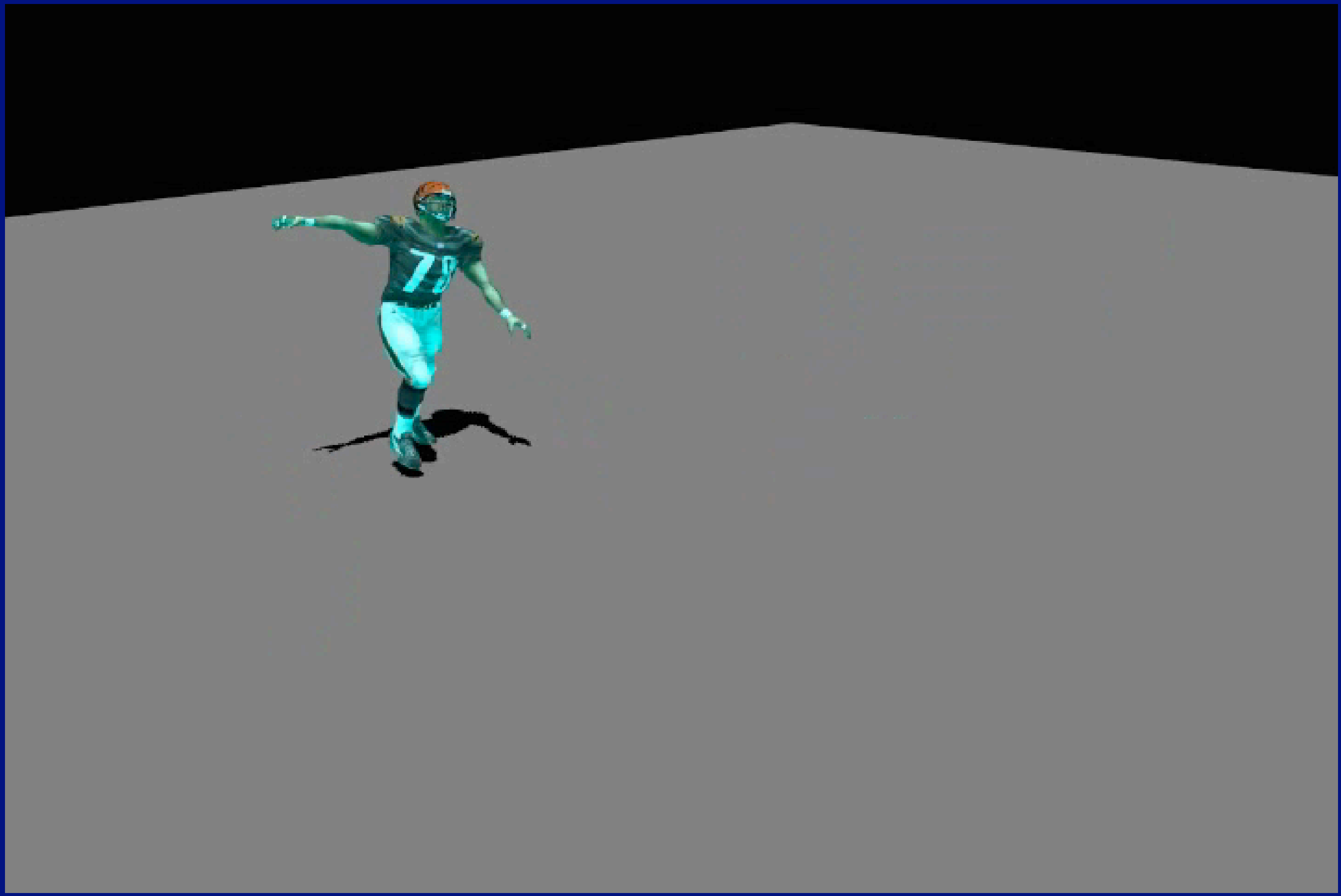
}

Ikemoto+Forsyth 04

Ikemoto+Forsyth 04



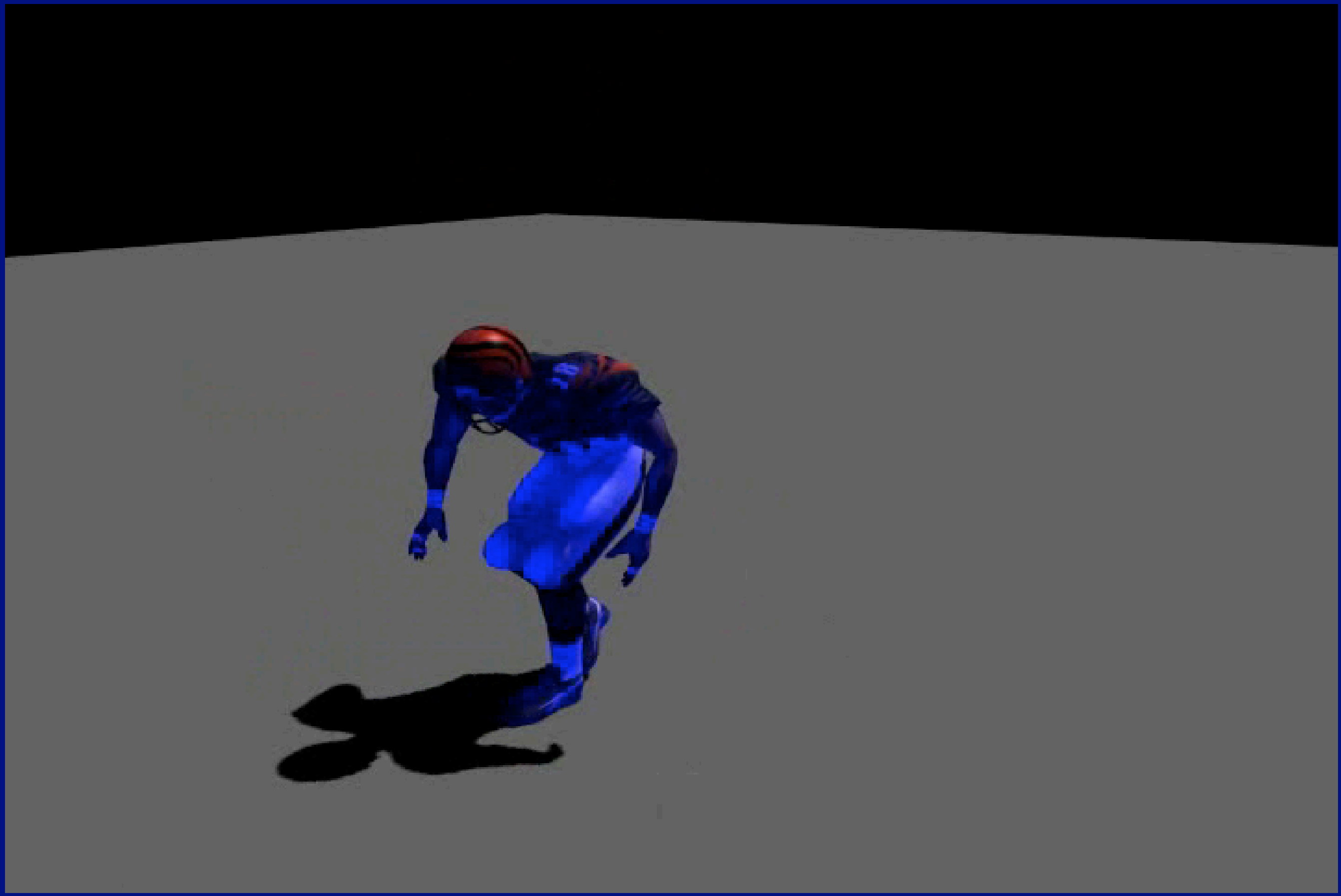
Ikemoto+Forsyth 04



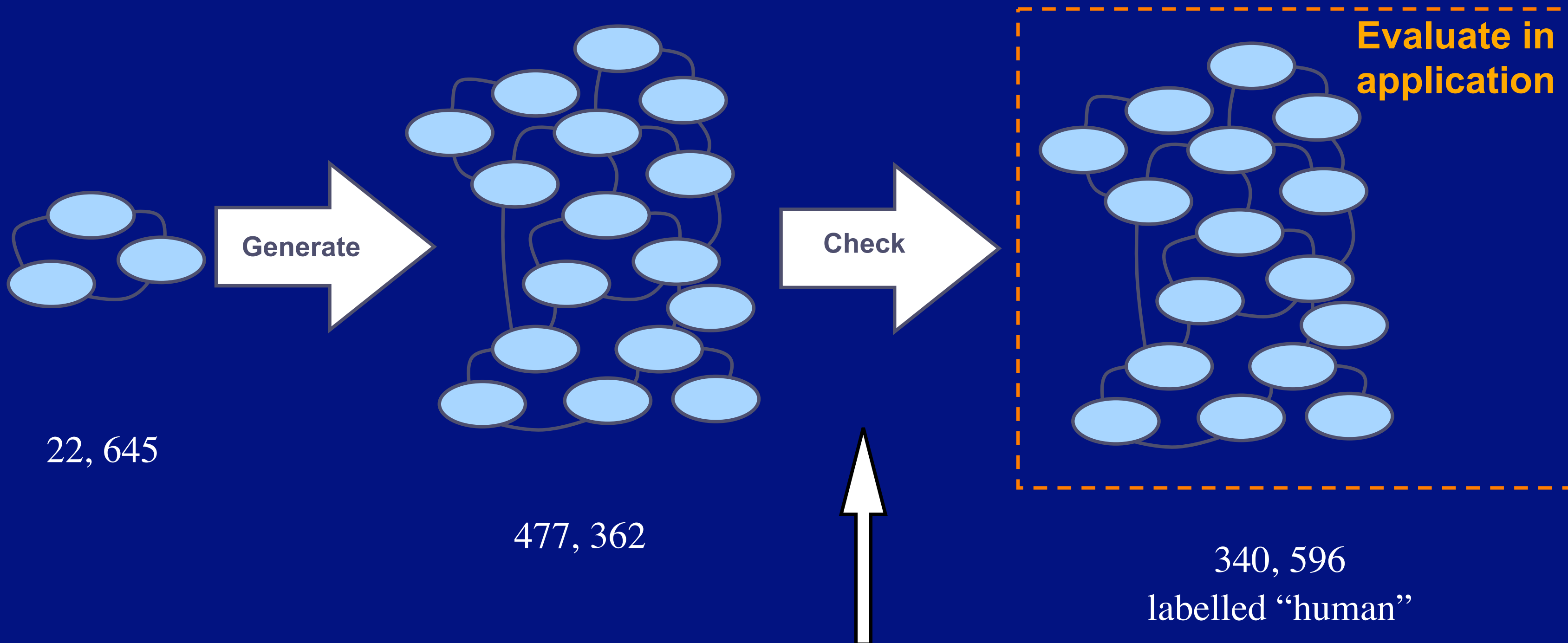
Ikemoto+Forsyth 04

Ikemoto+Forsyth 04

Ikemoto+Forsyth 04



Ikemoto+Forsyth 04



Classifier’s total error rate 13%

false positive rate 12%

But what does this mean in practice?

Ikemoto+Forsyth 04

Evaluate

Unreal Tournament 2004

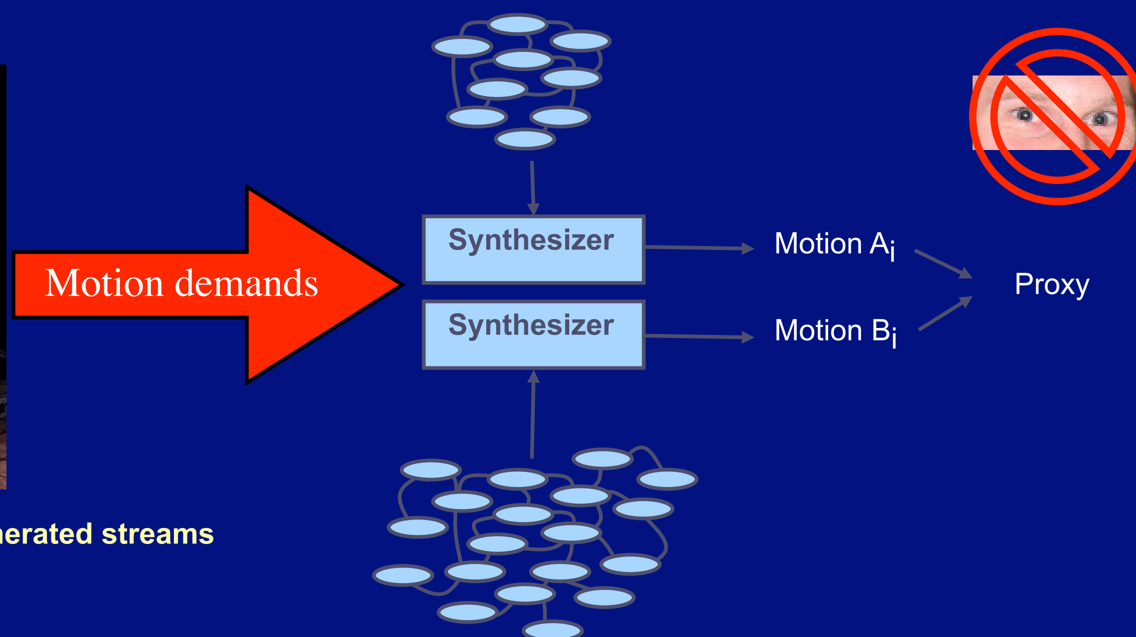
- Position
- Velocity
- Rotation
- Running/Falling

+ 40 hand-generated streams

- Stand
- Walk
- Run

- Jump

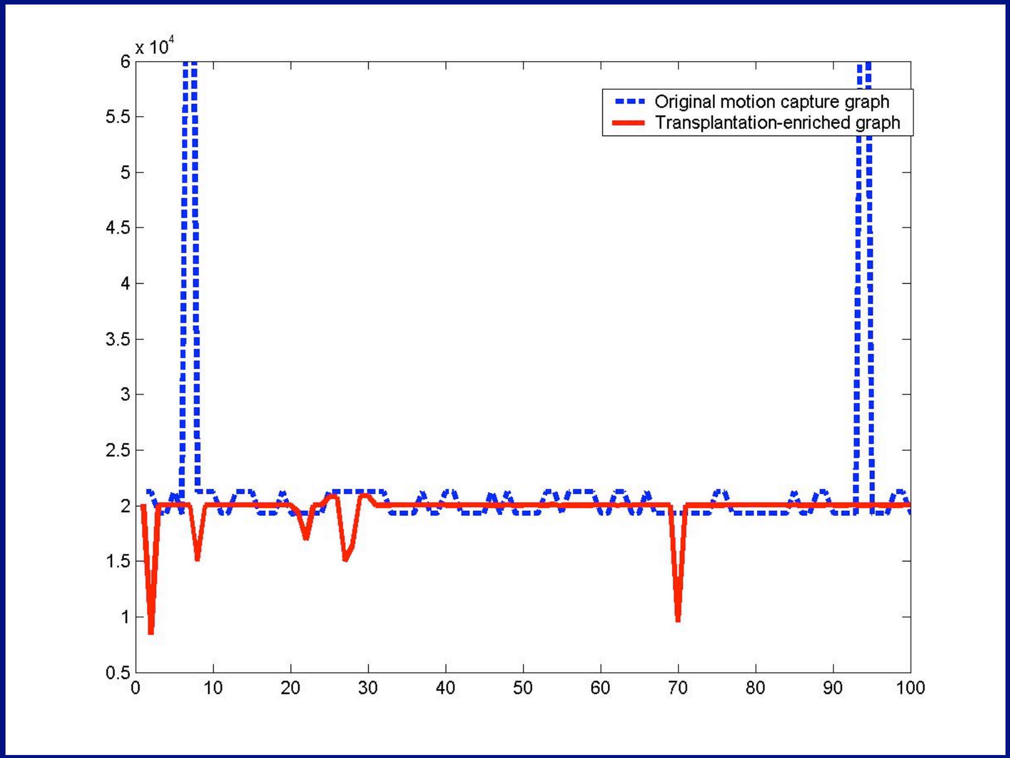
Original motion graph



Enriched graph

Ikemoto+Forsyth 04

Is the enriched graph better?



Ikemoto+Forsyth 04

Pushes and shoves

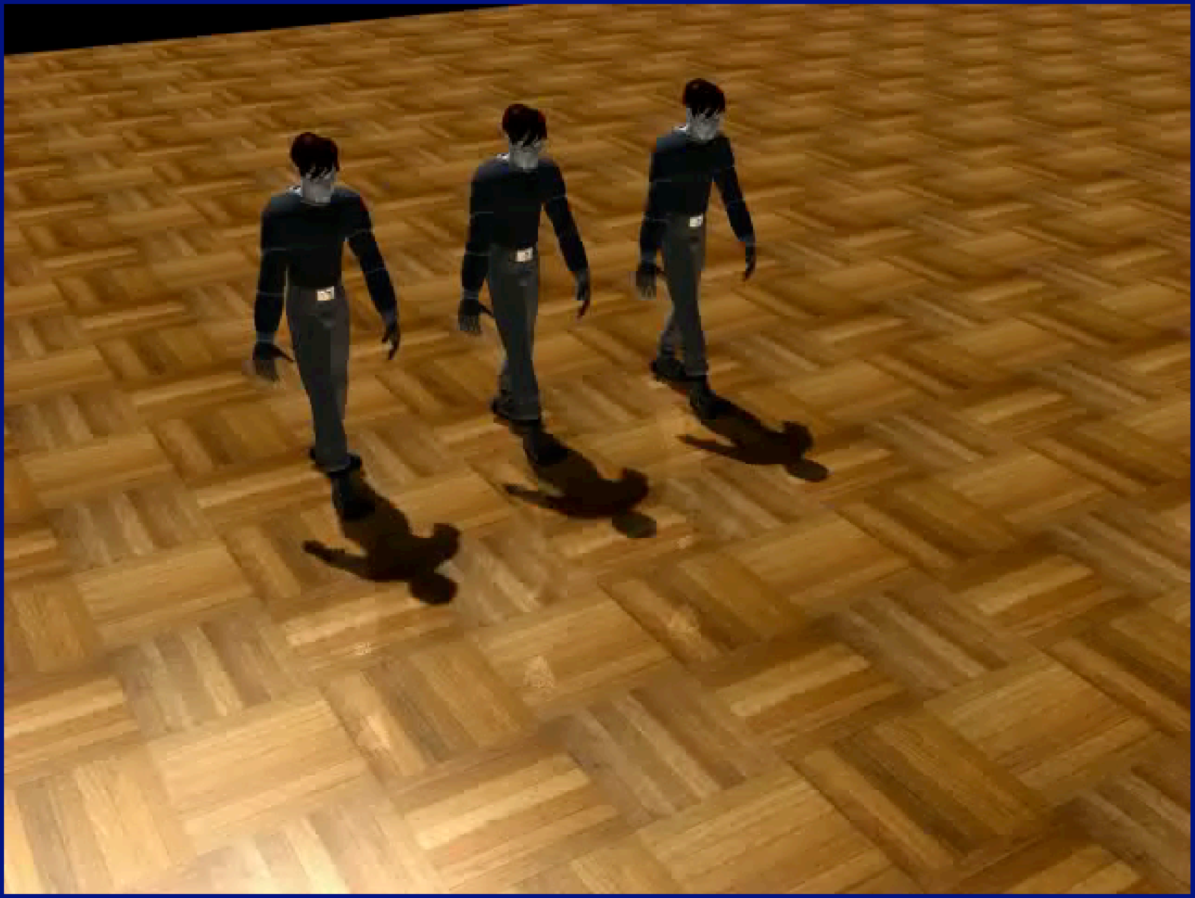
- Natural interaction --- push, pull, hit, shoot, etc
  - apply an impulse of given strength, direction
  - reaction time precludes much CNS involvement
    - Physics should be important
- Can’t serve impulse with observed data
  - too much data required unless you can guarantee limited impulses
- Strategy
  - deform each of many data items to serve given impulse
  - blend each to motion sequence
  - build regression model of motion quality to choose which to use

Arikan+Forsyth 05

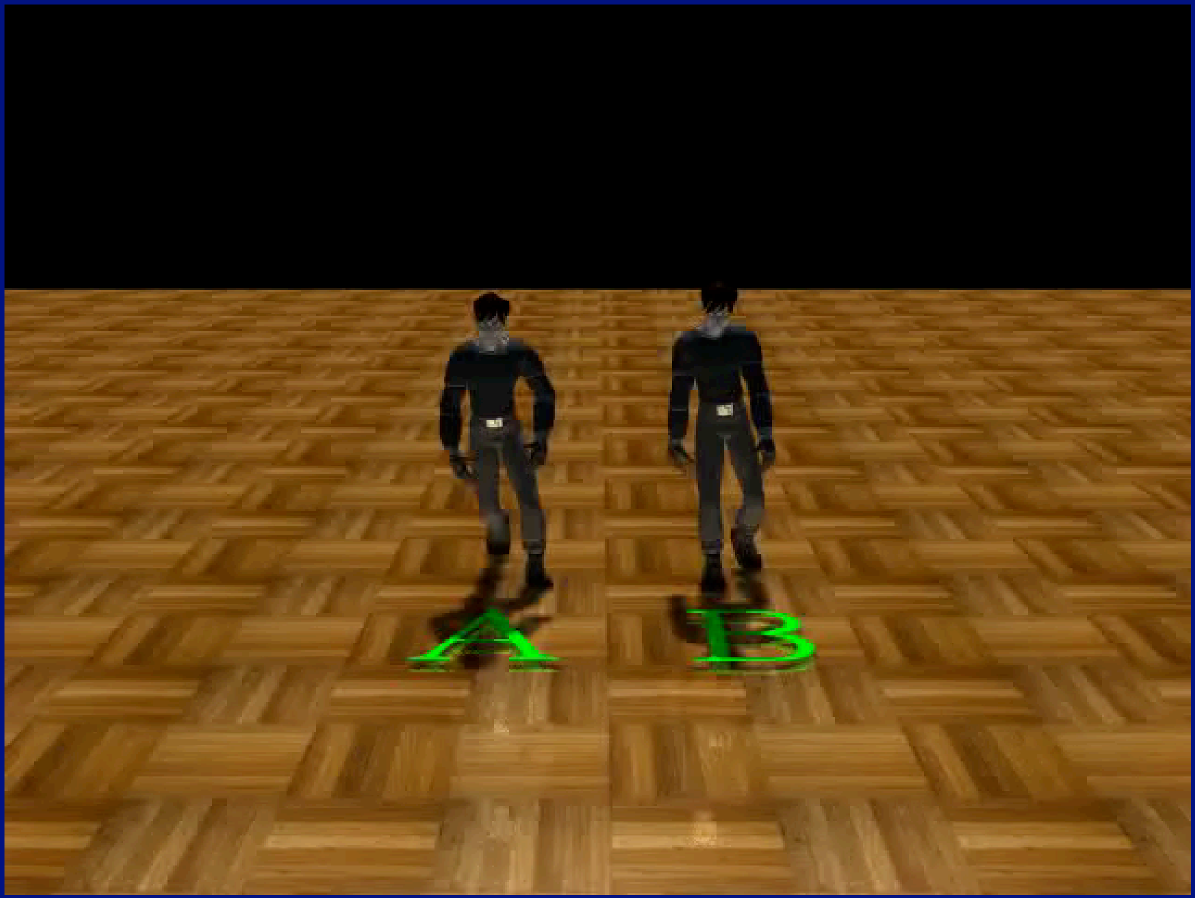
Arikan+Forsyth 05

Arikan+Forsyth 05

Arikan+Forsyth 05

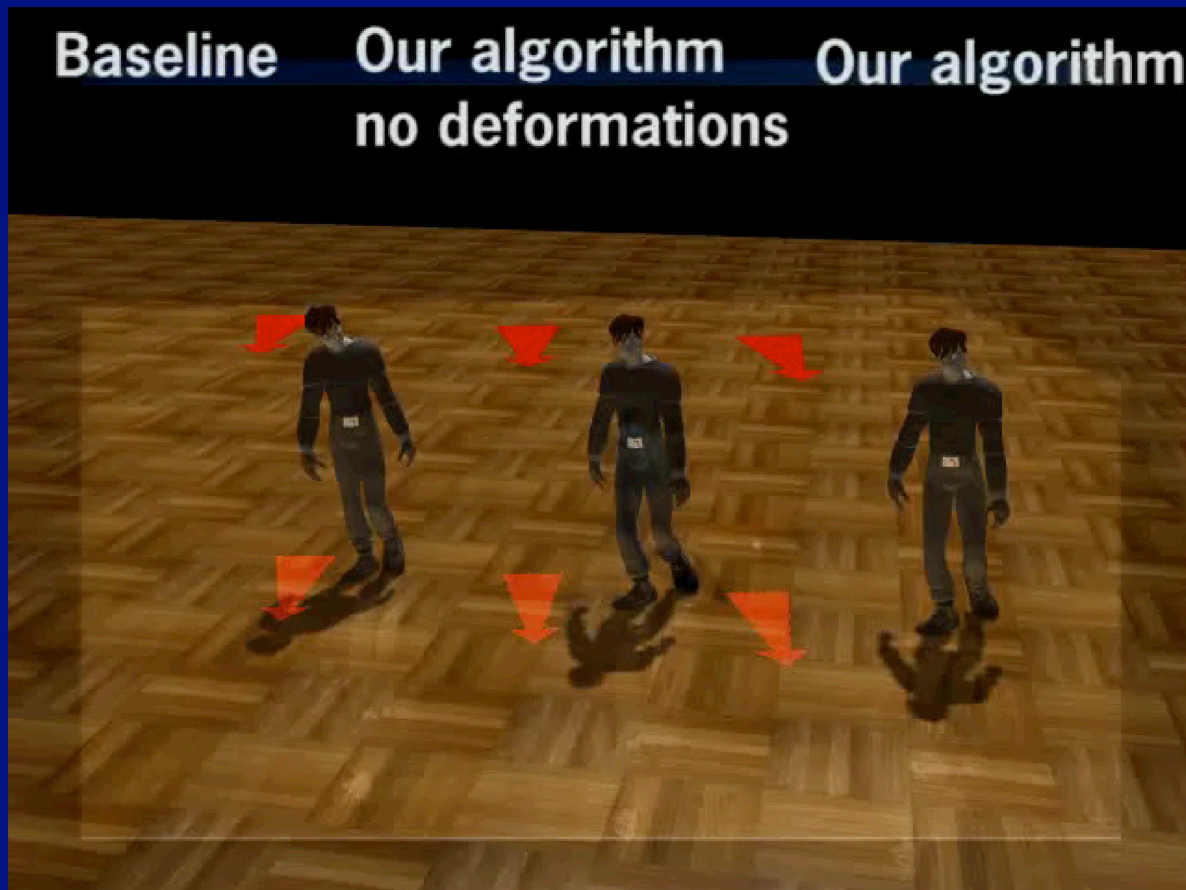


Arikan+Forsyth 05



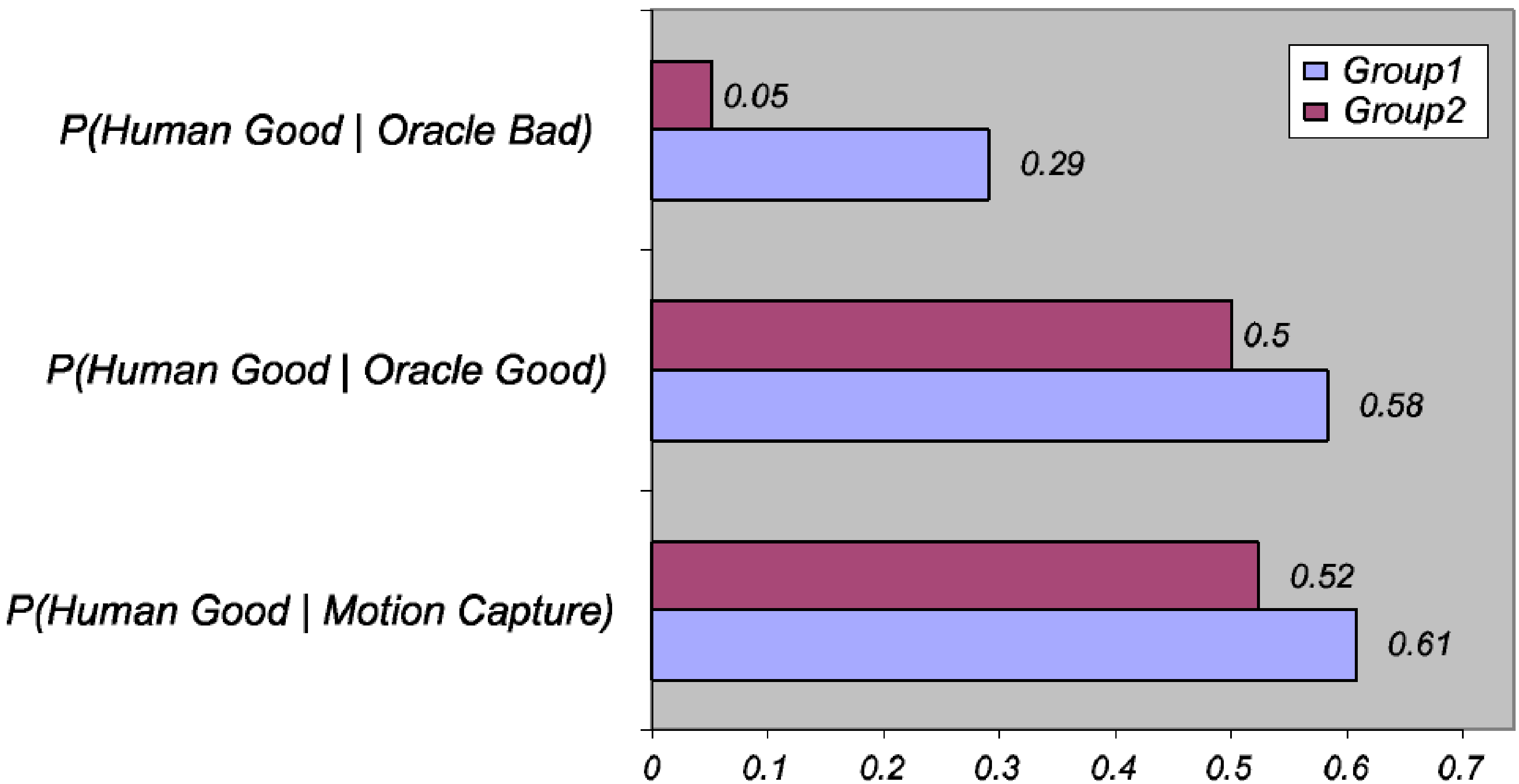
Arikan+Forsyth 05

Arikan+Forsyth 05



Arikan+Forsyth 05

How good are the motions?

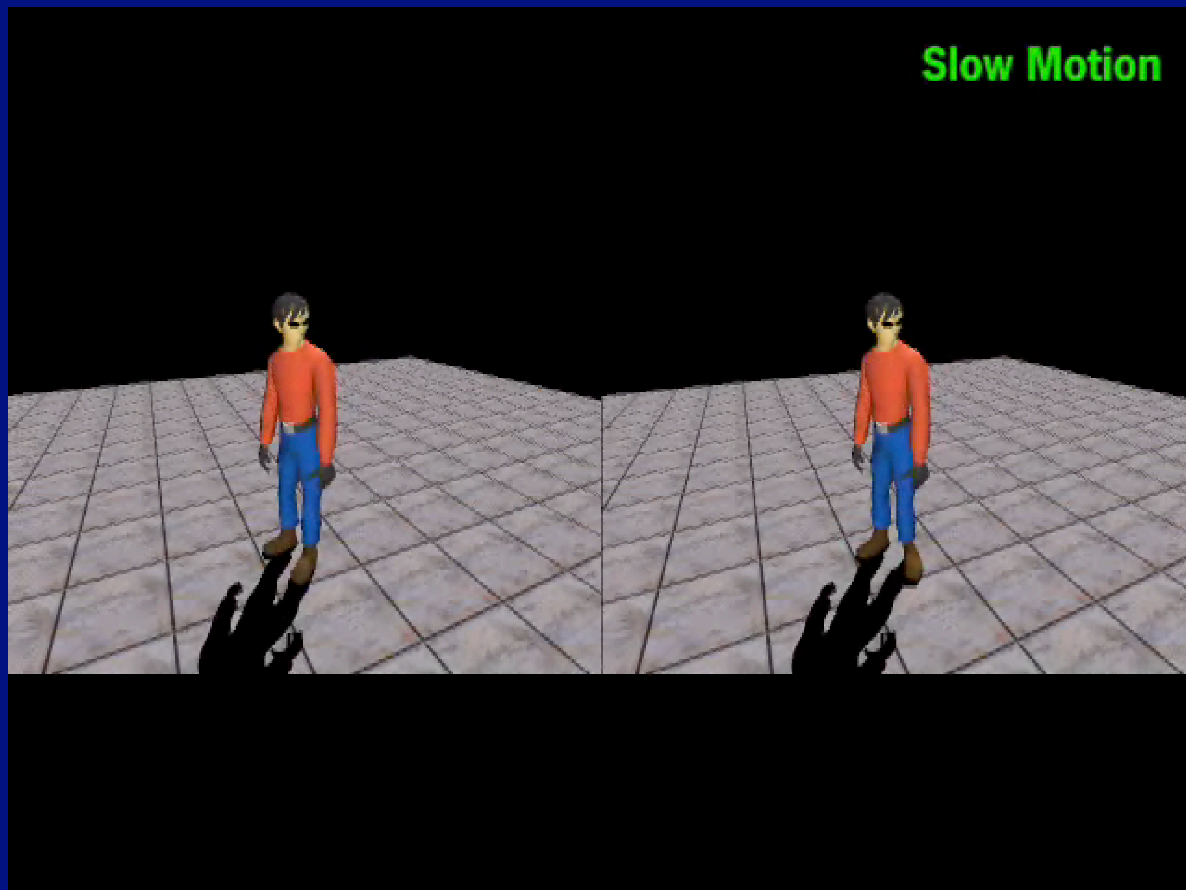


Arikan+Forsyth 05

Building Oracles

- Classifier (Ikemoto+Forsyth 04)
- Regression (Arikan+Forsyth 05)
- Ensemble of HMM’s (Ren et al 05)
- Nearest Neighbour (Ikemoto et al, in review)


Slow Motion


Slow Motion