Motion Capture Databases
Discussion of Open Problems

Jessica Hodgins
Carnegie Mellon University

mocap.cs.cmu.edu

6 hours of motion capture data

asf/amc format
Collected with the Helen Hayes Marker set
(41 markers)
Corresponding video
42% locomotion: 5% jumping, 3% running, and 33% walking
29% common scenarios: cleaning, waiting, gestures
16% physical activities: basketball, boxing, dance, exercise, golf, martial arts
7% interacting with the environment: rough terrain, playground equipment
6% two subjects interacting

Unique performances: pratfalls from a clown, break dancing, and several modern dance performances

Access

Freely available on the web via keyword search or download of the complete database.
What was it used for?

As of May 2005, over 40 technical papers published in graphics, vision, activity recognition, biomechanics, databases using mocap.cs.cmu.edu (with no affiliation to CMU). Most common area was activity recognition.

In May 2005, averaging 2000 accesses/month by unique IPs (after an attempt to cull for web crawlers).

What is missing?

Almost everything there is a “performance” and doesn’t necessarily reflect “natural” human activities.
Few behaviors involve interaction (with people or with objects other than the floor).
Hand and face motion was not recorded.
Degrees of freedom in back/shoulder were inadequate.
Video not calibrated or carefully synchronized.
Methods for access were very simple.
Kitchen Capture

Database of kitchen activities recorded using many different sensors:

- Motion capture with improved skeleton, hopefully hands
- Cameras (in room and on body)
- Microphones (in room and on body)
- Accelerometers
- Occasional use of other sensors: ground forces, object interaction forces, EMG, eyetracker

Long Captures: full meals cooked and consumed

Kitchen
Why the Kitchen?

Complex manipulation and mobility tasks
  • Requires an understanding of object manipulation and interaction
  • Record forces and hand motion
Long capture of natural activities
  • Mitigate effect of laboratory capture

Why the Kitchen?

Food preparation, consumption, and nutrition central to assistive technology applications
  • Aging in place
  • Obesity
  • Accidents in the home
  • Job coaching
What data are we recording?

Training data: labels/ground truth
   Hand-labeled semi-automatically
Tests: unlabelled data

Record long captures—full meals (cooking and consumption)
Record more modalities than will likely be necessary and
record at higher quality to provide a gold standard
Synchronize data
   Tightly within a modality
   As tightly as possible across modalities
      Via button switch light flashes/noise once/session

Sensors And Data

Mocap

Multiple Camera Video
Sensors And Data

accelerometer

microphones
Technical Challenges

Synchronization
Long recording sessions
Occlusions

Are we recording the right modalities with sufficient resolution in time and space?
Your input is welcome: jkh@cs.cmu.edu

Open Problems

Dimensionality reduction
  Data representation
  Poses or motion segments
Data compression
Segmentation
  Fine or coarse grain
Classification and clustering
Retrieval
  Full state
  Partial state (performance animation)
Generative models
  Filling of gaps due to occlusions (single marker or higher level)
  Generation of human motion

http://www.cs.cmu.edu/~christos/TALKS/SIGGRAPH-07-tutorial/