# NTT Visit:
# *Image Database Retrieval*
# *Variable Viewpoint Reality*

Professor Paul Viola

Collaborators: Professor Eric Grimson,
Jeremy De Bonet, John Winn, Owen Ozier,
Chris Stauffer, John Fisher, Kinh Tieu,
Dan Snow, Tom Rikert, Lily Lee,
Raquel Romano, Janey Hshieh, Mike Ross,
Nick Matsakis, Jeff Norris, Todd Atkins
Mark Pipes



## *Overview of Visit*

- Morning: Image Database Retrieval
  - Gatekeeper: Face detection and recognition
  - Complex Feature Image Database Retrieval (Tieu)
  - Flexible Template Retrieval (Yu)
- Interlude
  - Video/Audio Source Separation (Fisher)
  - Mathematical Expression Recognition (Matsakis)
- Lunch
- Visit Prof. Brooks lab

## *Overview of Visit - 2*

- Afternoon: Variable Viewpoint Reality
  - Real-time 3D reconstruction of people (Snow)
  - Automatic camera calibration (Snow + Lee)
  - Tracking of articulate human models (Lee + Winn)
  - Modeling of human dynamics (Viola + Fisher)



## *Show 3D Soccer software…*

## *VVR: Motivating Scenario*



- Construct a system that will allow each/every user to observe any viewpoint of a sporting event.
- Provide high level commentary/statistics
  - Analyze plays

*Paul Viola* *MIT AI Lab*

## *For example …*

Computed using a single view…
some steps by hand

*Paul Viola* *MIT AI Lab*

## *VVR Spectator Environment*

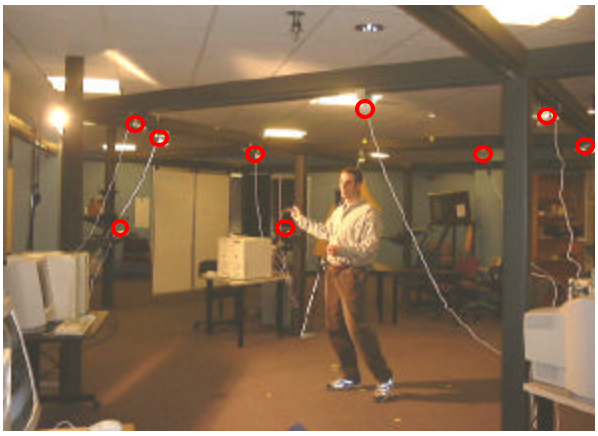
- Build an exciting, fun, high-profile system
  - Sports: Soccer, Hockey, Tennis, Basketball
  - Drama, Dance, Ballet
- Leverage MIT technology in:
  - Vision/Video Analysis
    - Tracking, Calibration, Action Recognition
    - Image/Video Databases
  - Graphics
- Build a system that provides data available nowhere else…
  - Record/Study Human movements and actions
  - Motion Capture / Motion Generation



## *Window of Opportunity*

- 20-50 cameras in a stadium
  - Soon there will be many more
- US HDTV is digital
  - Flexible, very high bandwidth digital transmissions
- Future Televisions will be Computers
  - Plenty of extra computation available
  - 3D Graphics hardware will be integrated
- Economics of sports
  - Dollar investments by broadcasters is huge (Billions)
- Computation is getting cheaper

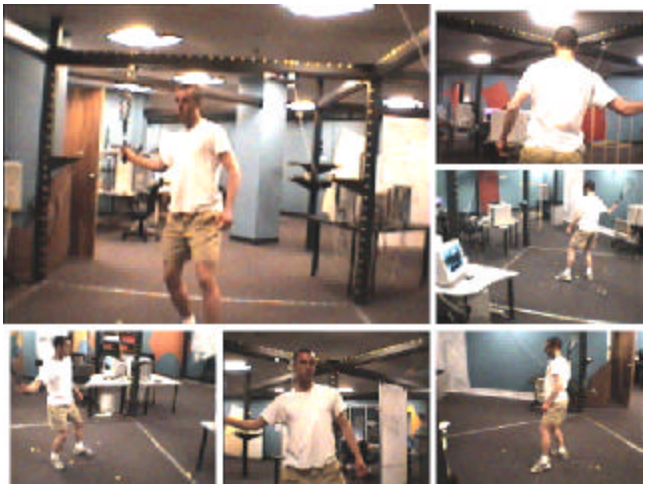
## *ViewCube:*
## *Reconstructing action & movement*



- Twelve cameras, computers, digitizers
- Parallel software for real-time processing

*Paul Viola* *MIT AI Lab*

## *The View from ViewCube*

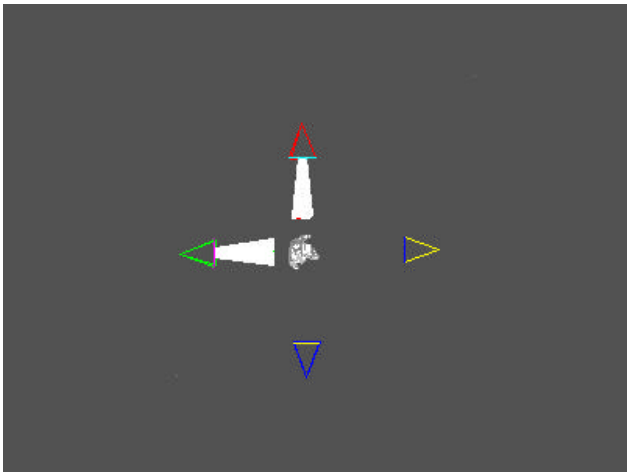


Multi-camera Movie

*Paul Viola* *MIT AI Lab*

## *Projecting Silhouettes to form 3D Models*

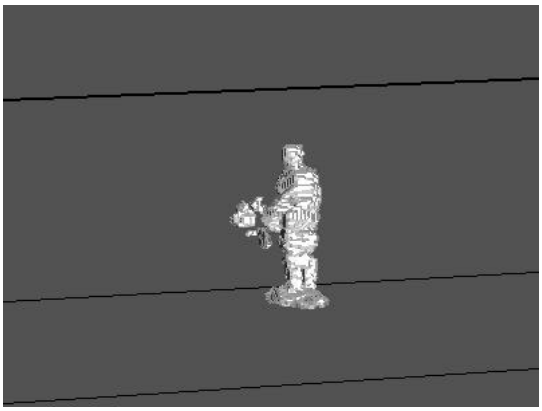
Real-time 3D Reconstruction is computed by intersecting silhouettes



3D Reconstruction Movie

*Paul Viola* *MIT AI Lab*

## *First 3D reconstructions ...*

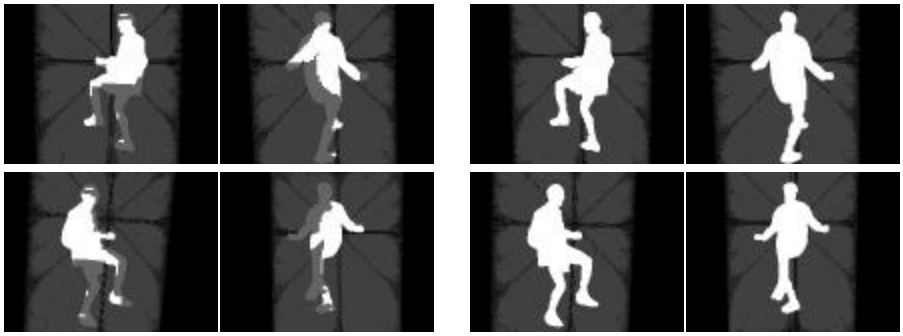


3D Movement Reconstruction Movie

*Paul Viola* *MIT AI Lab*

## *Dynamic Calibration*

**Calibration is a critical an expensive part 3D reconstruction**
**Our approach is dynamically self-calibrating**

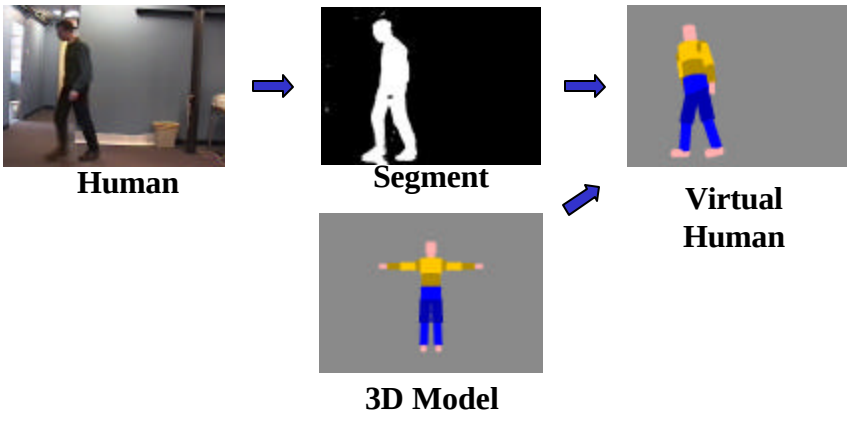


**Initial Calibration**
**- poor reconstruction**

**Automatically Improved**
**Calibration -> better results**

*Paul Viola*

## *Finding an articulate human body*



**Human**

**Segment**

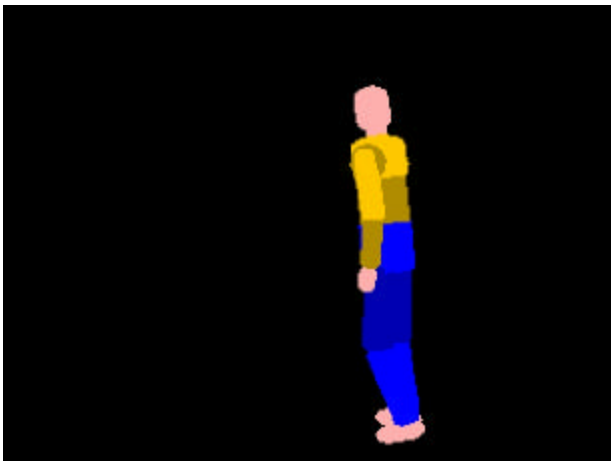
**Virtual Human**

**3D Model**

*Paul Viola*

*MIT AI Lab*

## *Automatically generated result:*



Body Tracking Movie



## *Analyzing Human Motion*

- Key Difficulty: Complex Time Trajectories
  Complex Inter-dependencies
- State of the Art: Hidden Markov Models
  - Have a discrete number of states
- Our Approach: Multi-scale statistical models

## *Human Motion is extremely complex:*

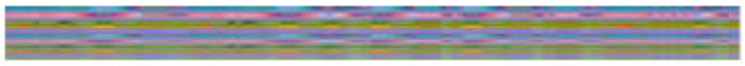

*Human Motion*
*is temporal texture*
*or*
*a temporal signature*

Motion Capture Movie

*Paul Viola* *MIT AI Lab*

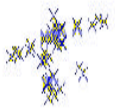
## *Texture Analysis and Synthesis can be applied to Movement Data*

**Synthesis**



**The texture synthesis tools can be applied to movement**

**The texture recognition tools can be applied to movement**



Motion Capture Movie

*Paul Viola*