Real-Time Video for Computer Entertainment

Richard Marks
Sony Computer Entertainment America
Research & Development Division

Sony Computer Entertainment America (SCEA)

- **R & D: 15 people in Foster City, CA**
- **Mission: Catalyze new ideas for computer entertainment**
- **Focus: Software for PlayStation2**
  - Advanced rendering
  - Intelligent characters
  - Physical simulation
  - Digital interfaces

PlayStation2, not PC

- **Very diverse end-user**
- **Platform is constant**
  - iLINK (IEEE1394) and 2 USB ports
  - Known compute capability (much)
- **Unique architecture**
  - Highly parallel
  - Micro-programmable
  - Data-centric

My Research Goals

- **Explore new capabilities generated by real-time video input to PS2**
  - Investigate alternative user-input mechanisms for PS2 (besides joystick)
  - Understand limitations of low-cost video input
- **Create new video-based entertainment models**

Real-time Video Research

- **Natural Interfaces**
  - SIGGRAPH 2000 sketch, emerging technology
  - Game Developer Conference 2001
- **Enhanced Reality**
  - ACM1
  - SIGGRAPH 2001

Natural Interfaces

- Intuitive
- Simple
- Enabling
- Enjoyable

⇒ Video-based interfaces, with and without props

Specification

- **Real-time**
  - 30 frames/second
  - Less than 3 frames total latency
- **Inexpensive**
  - Camera cost-of-goods target <$10
- **Robust**
  - Graceful failure/error recovery

Current Setup

- **USB webcam (<$50 retail)**
  - 30 Hz YUV411 video
  - 320x240 compressed,160x120 uncompressed
- **Video processing performed by core**
  - Decompression (bit-stream decode, IDCT)
  - Low-level image filters (smooth, threshold, etc.)
  - Segmentation, matching, tracking
- **Demo**

Medieval Chamber

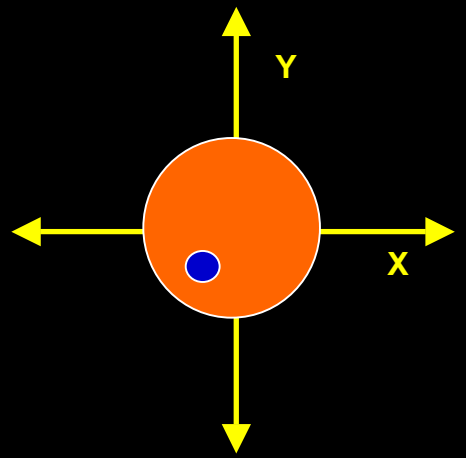
- **Multiple color-based tracking approaches**
  - Richard Marks
- **Advanced rendering including shadows, transparency, reflections, etc.**
  - Gabor Nagy
- **Physical simulation/collisions**
  - Eric Larsen

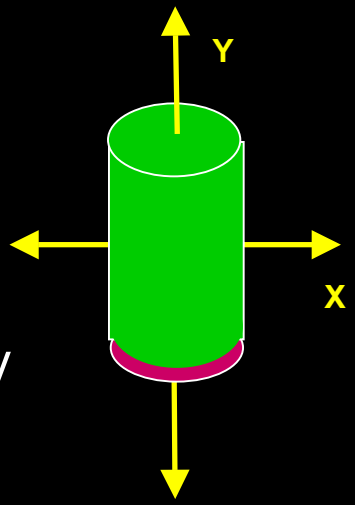
Medieval Chamber

- **Known camera, objects**
  - Spheres and cylinders have special projection properties
- **Tracking steps**
  - Color segmentation
  - Centroid, moment calculation
  - Windowed centroid, moment calculation
  - Situational probabilistic ambiguity resolution
  - Kalman filtering

Medieval Chamber

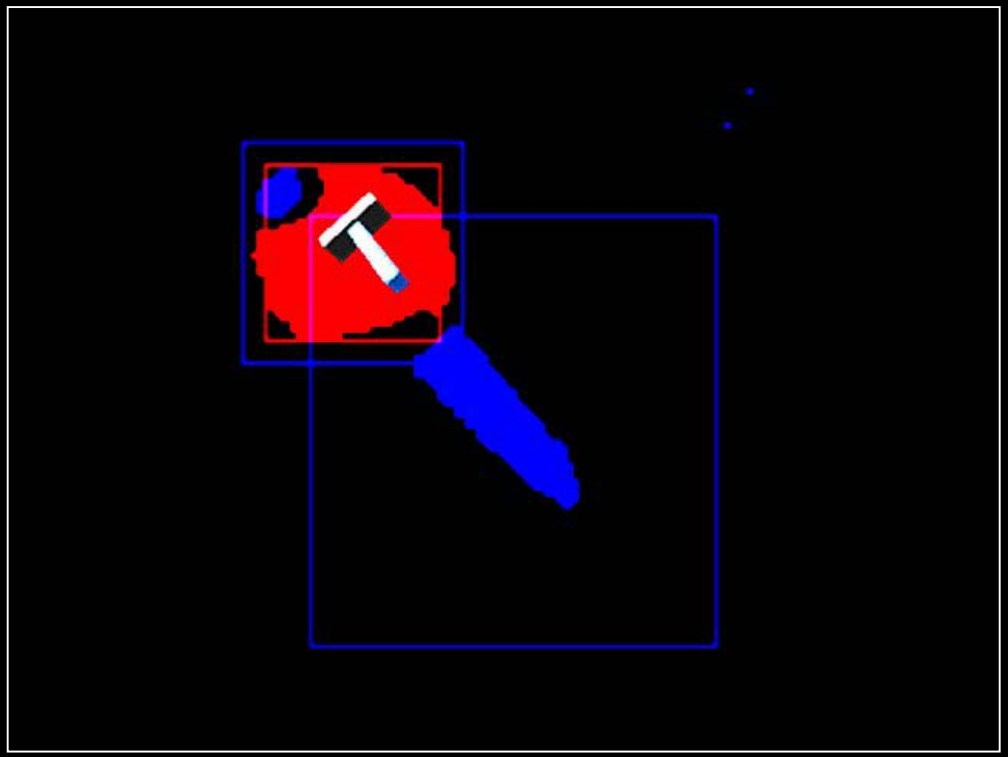
- Sphere
  - *x,y* from centroid,
  - *z* from principal moment
  - **R**x, **R**y from dot centroid (given *x,y*)
- Cylinder
  - *x,y* from centroid
  - **R**z from angle of principal moment
    - » Marker used to resolve ambiguity
  - *z* from secondary moment
  - Body **R**x from principal moment (given *z*)
    - » Foreshortening used to resolve ambiguity
  - Body **R**y from helical stripe





Medieval Chamber

- Combination of sphere and cylinder provides most robust tracking

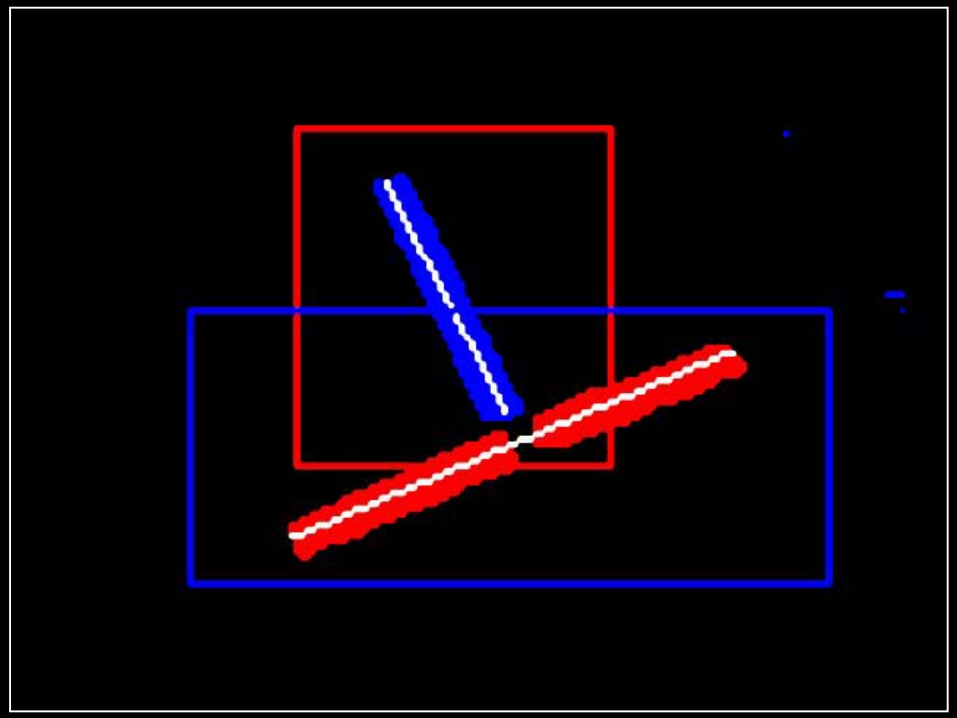


Marionette

- **Alternative form of character control**
- **Traditional marionette**
  - Darwin the Wizard, created by Daniel Oates
- **Virtual marionette**
  - 3D model by Care Michaud

Marionette

- Color segmentation
- Line fitting to find T shape
- T shape analysis to recover puppet parameters

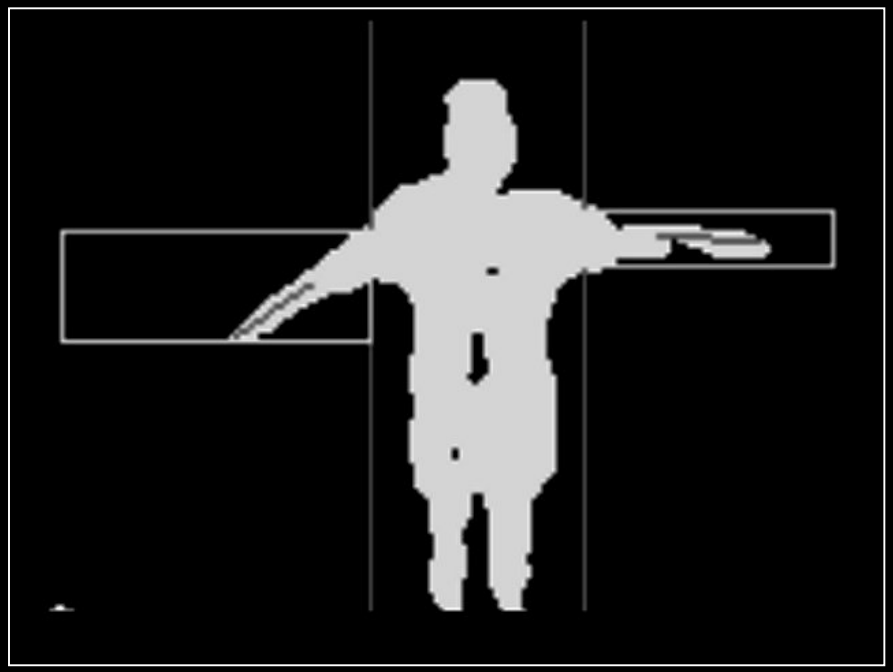


Fly

- **Flight simulation**
- **Procedural landscape**
  - Tyler Daniel
- **Relative arm angles determine bank angle**
- **Average arm angles determine attack angle**
- **Arm motion increases airspeed**
- **Shadow wings mimic arms**

Fly

- Centroid/ moment determines body extent
- Principal axis angles of outer regions correspond to arm angles



Conclusions

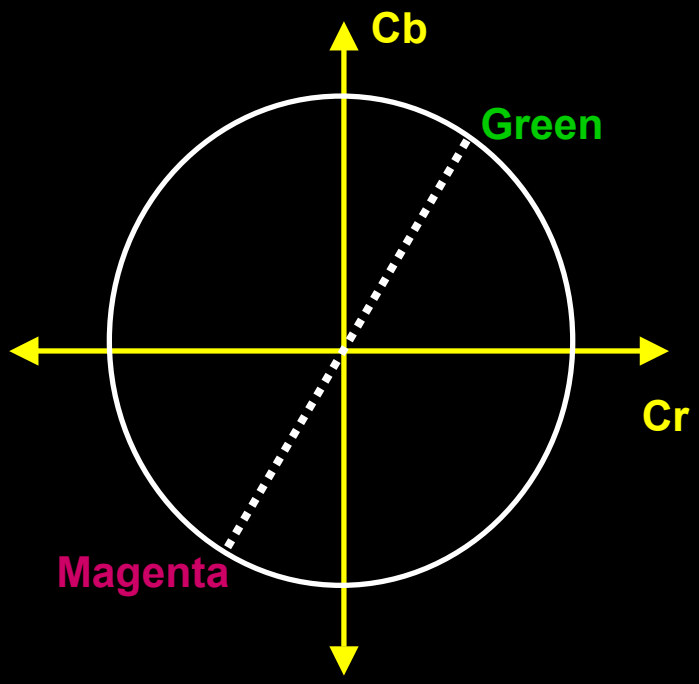
- Known props can be easily tracked and still be natural
- Precision more important than accuracy
- Secondary motion can enhance perceived response (and hide error)
- Display perspective important
- Make signal proportional to action
- Area-based measurement more robust and precise than lineal measures

Issues

- **Large variance in target environments**
  - Mirrors
  - Different colored lighting
  - Arbitrary background
- **Color processing sensitive to lighting color**
- **Inexpensive cameras poor in low light**
- **Some latency is unavoidable**
- **Low resolution video accentuates FOV vs. precision tradeoff**

Color Transitions

- Project (Cr, Cb) for each pixel onto a line
- Similar to barcodes, but selectable
- Maximal separation produces best results
- Robust to lighting variation
- Patents pending



Future Work

- Attach interfaces to *real* games
- Investigate interfaces for creative content generation
- Enhanced Reality research

Enhanced Reality: What is it?

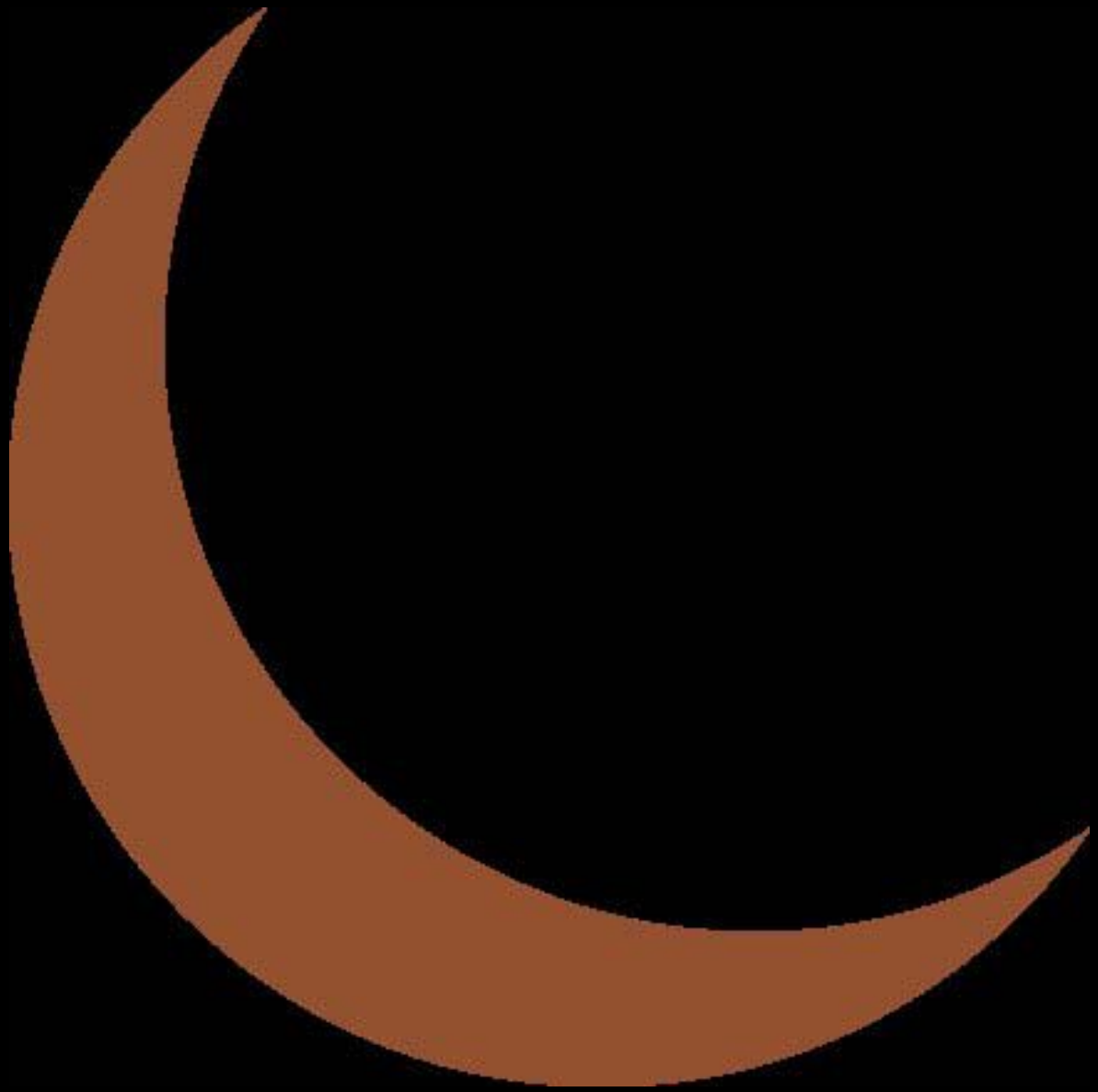
- Starting with live audio/video, enhance it by adding/modifying the content
- Not virtual reality!
- Similar to Augmented Reality (a la MIT Media Lab), but with entertainment focus
- Consists of enhancements to both user and environment

Virtual Pet

- Demo (witch)
- Video quality directly important to user experience

Virtual interaction

- Demo (butterflies)
- Rendering issues
- Holes
- Latency



Outline

- Background
- Video Input for PS2
- Natural Interfaces
- Conclusions
- Future Work

Planet Explorer

- **3D viewing, navigation**
- **Earth rendering**
  - Greg Corson
- **Rotating the ball rotates the earth**
- **Proximity of ball to camera adjusts zoom**

NEW ORLEANS

SIGGRAPH 2000

Planet Explorer

- Color segmentation/centroid to find ball
- Principal moment to adjust zoom
- Motion-estimation to measure rotations

Hand Puppet

- Very simple form of character control
- 3D models by Guy Burdick

Hand Puppet

- Color segmentation to locate hands
- Split screen for left/right hands
- Centroid, moments and principal moment orientation
- Split along principal moment, principal moment orientations provide mouth angle


NEW ORLEANS
2000
SIGGRAPH

My Background

- MIT, B.S. 1990 in Avionics
- Stanford, Ph.D. 1995 in underwater robotics
- Teleos Research, 10 person computer vision startup
- Autodesk, makers of AutoCAD, worked on photo-to-model research
- Digital Video Art, graphics consulting

Sony Computer Entertainment

- Subsidiary of Sony
- Responsible for PlayStation-related products
- SCEI (Tokyo)
- SCEE (London)
- SCEA (Foster City, CA--near SanFran)

PlayStation2