
NEW ORLEANS
2000
SIGGRAPH

Natural Interfaces via Real-Time Video

Richard Marks
Sony Computer Entertainment America
Research & Development Division

Natural Interfaces

- Intuitive
- Simple
- Enabling
- Enjoyable

⇒ Video-based interfaces, with and without props

Summary

- **Purpose**
  - Describe work in progress
  - Share practical knowledge we have obtained
- **Outline**
  - Background
  - Demonstrations
  - Conclusions

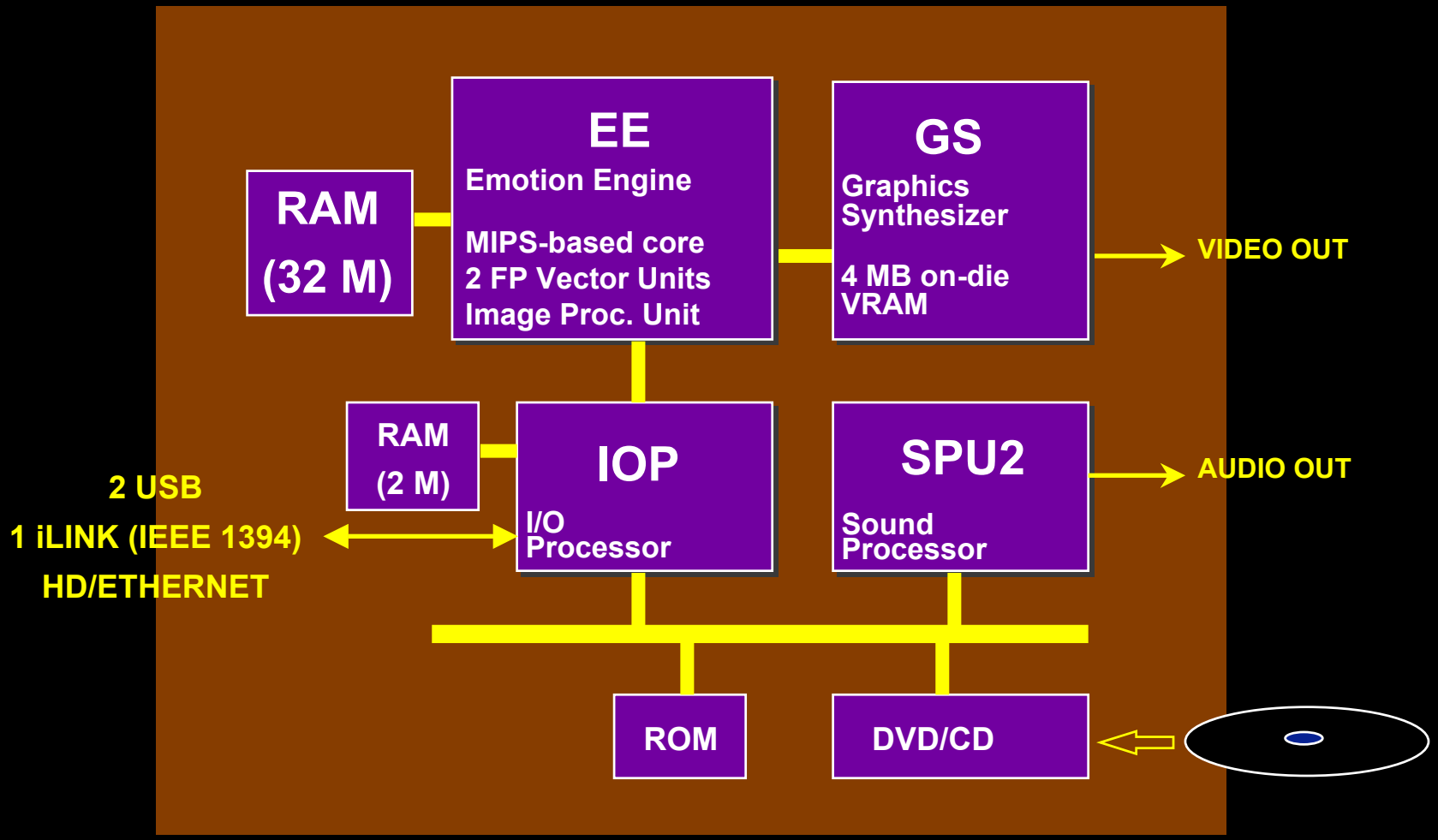
Sony Computer Entertainment America (SCEA)

- **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 a PC

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

PlayStation2



Current Setup

- **USB webcam (<$50)**
  - 30 Hz YUV420 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

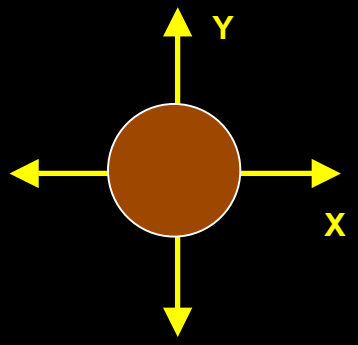
- **Multiple color-based tracking approaches**
  - Richard Marks
- **Advanced rendering including shadows, transparency, reflections, etc.**
  - Gabor Nagy
- **Physical simulation/collisions**
  - Eric Larsen

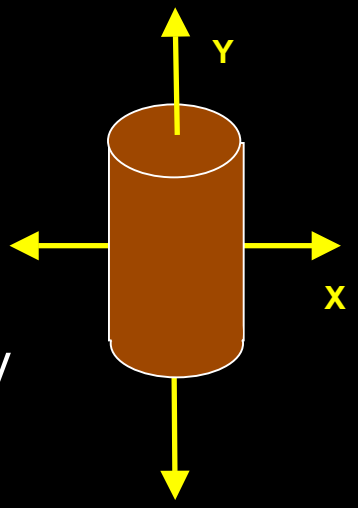
Medieval Chamber

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

Medieval Chamber

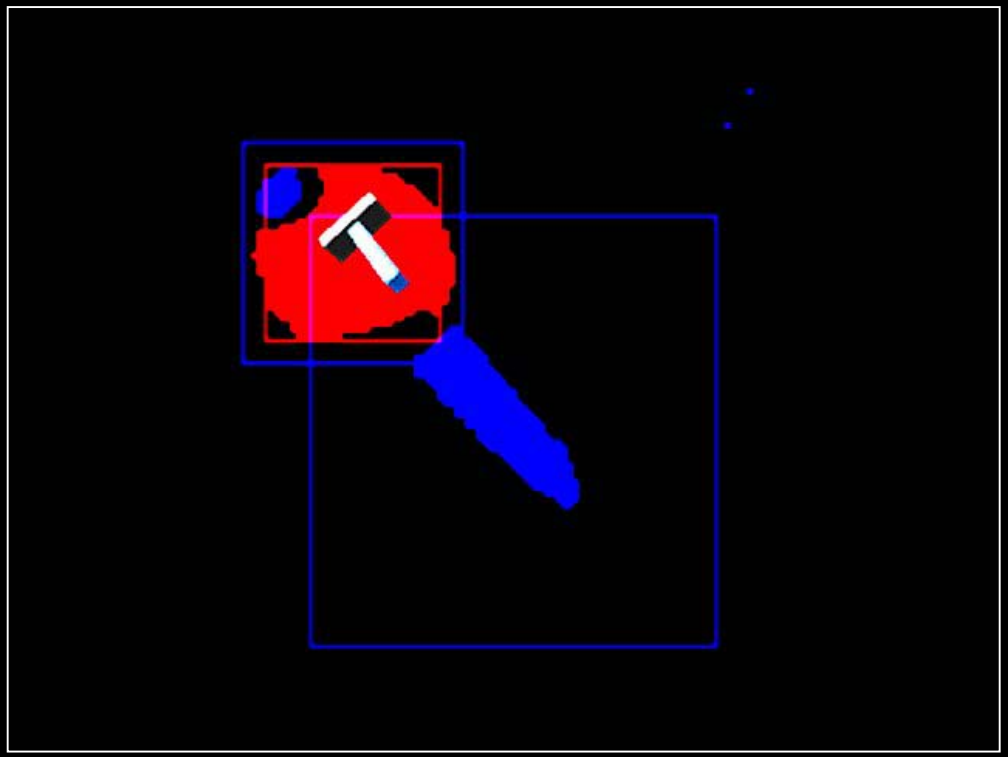
- Sphere
  - *x,y* from centroid,
  - *z* from principal moment
  - **R**x, **R**y from dot centroid (given *x,y*)
- Cylinder
  - *x,y* from centroid
  - **R**z from angle of principal moment
    - » Marker used to resolve ambiguity
  - *z* from secondary moment
  - Body **R**x from principal moment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





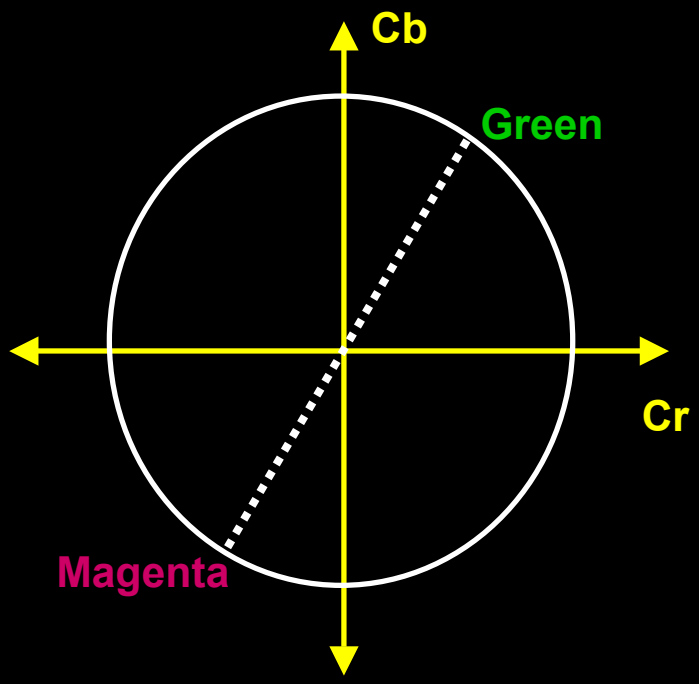
Medieval Chamber

- Combination of sphere and cylinder provides most robust tracking



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

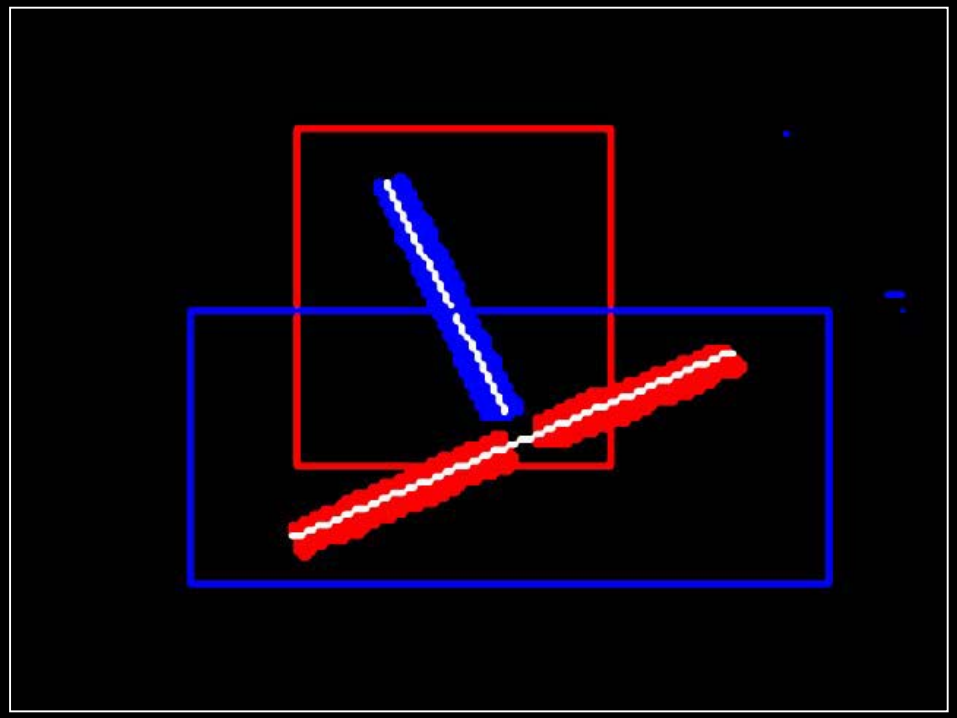


Marionette

- **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



Planet Explorer

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

Planet Explorer

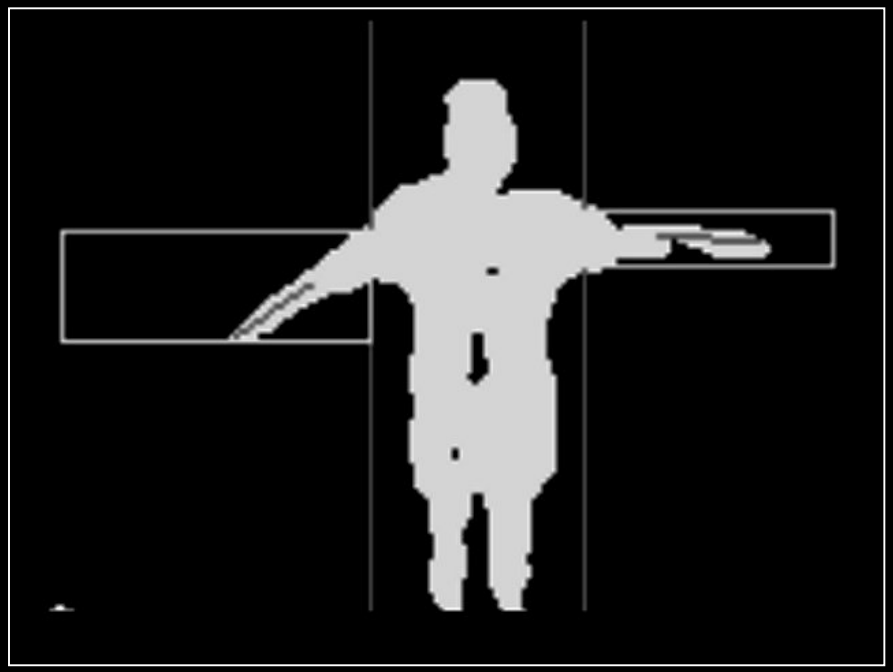
- Color segmentation/centroid to find ball
- Principal moment to adjust zoom
- Motion-estimation to measure rotations

Fly

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

Fly

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



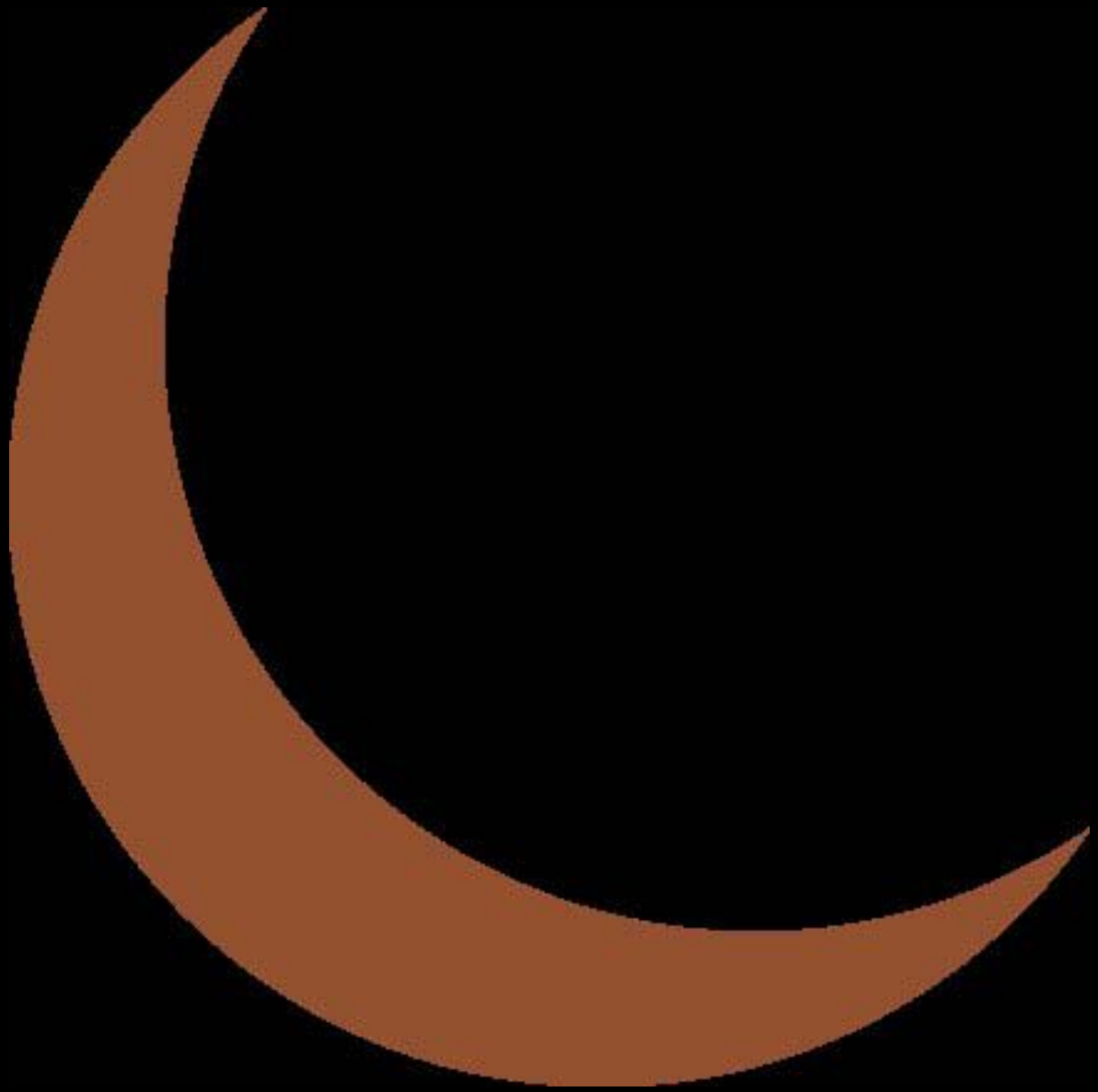
Conclusions

- Known props can provide more information and still be natural
- Area-based measurement more robust and precise than lineal measures
- Precision more important than accuracy
- Color sensitive to lighting, but color transitions are not

Conclusions

- Make signal proportional to action
- Secondary motion can enhance perceived response (and hide error)
- Display perspective important

$\Rightarrow$ Natural interfaces are viable via real-time video



NEW ORLEANS

SIGGRAPH 2000

Specification

- **Fun**
- **Intuitive**
- **Enabling**
- **Real-time**
  - 30 frames/second
  - Less than 3 frames total latency
- **Robust**
  - Graceful failure/error recovery