3D User Interfaces for Games and Virtual Reality

Lecture #3: Video Game Motion Controllers
Spring 2023
Joseph J. LaViola Jr.

3D Spatial Input Hardware – The Past

These Devices cost thousands of Dollars!!

Polhemus Patriot
Intersense IS-900
3rd Tech Hi Ball
3D Spatial Input Hardware – Today

Nintendo Wiimote
Microsoft Kinect
HTC Vive Controllers

These Devices cost hundreds of Dollars!!

Lecture Outline

- Discuss video game motion controller hardware characteristics
  - Nintendo Wiimote
  - Microsoft Kinect
  - PlayStation Move
- Case Studies
The Wiimote Device

- Wiimote features
  - uses Bluetooth for communication
  - senses acceleration along 3 axes
  - optical sensor for pointing (uses sensor bar)
  - provides audio and rumble feedback
  - standard buttons and trigger
  - uses 2 AA batteries
- Supports two handed interaction
  - can use 2 Wiimotes simultaneously
- Easily expandable
Wiimote Attachments

- Nunchuk
- Steering Wheel
- Zapper
- Wii Helm
- Boxing Gloves
- Sports Pack
- Fishing Reel

The Wiimote – Coordinates

Wiimote Coordinates

Spring 2023
The Wiimote – Optical Data

- Data from optical sensor
  - uses sensor bar
    - 10 LED lights (5 of each side)
    - accurate up to 5 meters
  - triangulation to determine depth
    - distance between two points on image sensor (variable)
    - distance between LEDs on sensor bar (fixed)
  - roll (with respect to ground) angle can be calculated from angle of two image sensor points

- Advantages
  - provides a pointing tool
  - gives approximate depth

- Disadvantages
  - line of sight, infrared light problems
  - only constrained rotation understanding

The Wiimote – Motion Data

- Data from 3-axis accelerometer
  - senses instantaneous acceleration on device (i.e., force) along each axis
  - arbitrary units (+/- 3g)
  - always sensing gravity
    - at rest acceleration is g (upward)
    - freefall acceleration is 0
  - finding position and orientation
    - at rest – roll and pitch can be calculated easily
    - in motion – math gets more complex
    - error accumulation causes problems
    - often not needed – gestures sufficient

- Advantages
  - easily detect course motions
  - mimic many natural actions

- Disadvantages
  - ambiguity issues
  - player cheating
  - not precise (not a 6 DOF tracker)
The Wii Motion Plus

- Current Wiimote device
  - gives user a lot of useful data
  - not perfect
    - ambiguities
    - poor range
    - constrained input
- Wii Motion Plus
  - moving toward better device
  - finer control
  - uses dual axis “tuning fork” angular rate gyroscope
  - true linear motion and orientation

Visualizing Wiimote Data

- Important to see data to understand device
Microsoft Kinect

- Kinect features
  - RGB camera
  - depth sensors
  - multi-array mic
  - motorized tilt
  - connects via USB
- Supports controllerless interface
- Full body tracking

Kinect – Hardware Details

- RGB Camera
  - 640 x 480 resolution at 30Hz
- Depth Sensor
  - complimentary metal-oxide semiconductor (CMOS) sensor (30 Hz)
  - infrared laser projector
  - 850mm to 4000mm distance range
- Multi-array mic
  - set of four microphones
  - multi-channel echo cancellation
  - sound position tracing
- Motorized tilt
  - 27° up or down
**Kinect – Extracting 3D Depth**

- Infrared laser projector emits known dot pattern
- CMOS sensor reads depth of all pixels
  - 2D array of active pixel sensors
    - photo detector
    - active amplifier
- Finds location of dots
- Computes depth information using stereo triangulation
  - normally needs two cameras
  - laser projector acts as second camera
- Depth image generation

**Kinect – Skeleton Tracking**

- Combines depth information with human body kinematics
  - 20 joint positions
- Object recognition approach
  - per pixel classification
  - decision forests (GPU)
  - millions of training samples
- See Shotton et al. (CVPR 2011)
Kinect 2

- RGB Camera
  - HD resolution
- Depth Sensor
  - time of flight
- microphone array
- ToF – illuminate it with a beam of pulsed light and calculate time it takes for the light to be detected on an imaging device

Kinect 2 – Other Differences

- Greater accuracy
  - three times the fidelity over Kinect
- Can track without visible light using an active IR sensor
- Has a 60% wider field of view
  - detect a user up to 3 feet from the sensor compared to six feet for the Kinect
  - track up to 6 skeletons at once
- Detect a player’s heart rate and facial expressions,
- Position and orientation of 25 individual joints (including thumbs),
- Weight put on each limb and speed of player movements
PlayStation Move

- Consists of
  - Playstation Eye
  - 1 to 4 Motion controllers
- Features
  - combines camera tracking with motion sensing
  - 6 DOF tracking (position and orientation)
  - several buttons on front of device
  - analog T button on back of device
  - vibration feedback
  - wireless

PlayStation Move – Hardware

- PlayStation Eye
  - 640 x 480 (60Hz)
  - 320 x 240 (120Hz)
  - microphone array
- Move Controller
  - 3 axis accelerometer
  - 3 axis angular rate gyro
  - magnetometer (helps to calibrate and correct for drift)
  - 44mm diameter sphere with RGB LED
    - used for position recovery
    - invariant to rotation
    - own light source
    - color ensures visual uniqueness

www.hardwaresphere.com
PlayStation Move – 6 DOF Tracking

Image Analysis
- find sphere in image
- segmentation
  - label every pixel being tracked
  - saturated colors more robust
- pose recovery
  - convert 2D image to 3D pose
  - robust for certain shapes (e.g., sphere)
- fit model to sphere projection
  - size and location used as starting point
  - 2D perspective projection of sphere is ellipse
  - given focal length and size of sphere, 3D position possible directly from 2D ellipse parameters

Sensor Fusion
- combines results from image analysis with inertial sensors (Unscented Kalman Filter)
- contributions
  - camera – absolute 3D position
  - accelerometer
    - pitch and roll angles (when controller is stationary)
    - controller acceleration (when orientation is known)
    - reduce noise in 3D position and determine linear velocity
  - gyroscope
    - angular velocity to 3D rotation
    - angular acceleration

Sensor Fusion (Unscented Kalman Filter)
- Invariant with:
  - $s_i = [x_i, y_i, z_i]$
  - $P_{k} = [x^2_i, y^2_i, z^2_i]$
  - $Q_{i} = [x^2_i + y^2_i + z^2_i + K^2_i]$

For $i = 1, \ldots, n$, calculate sigma points:
- $x_{i, 0} = s_i$
- $x_{i, 1} = s_i + \sqrt{\sigma^2_i} \cdot \mathbf{P}_{k_{i}}$
- $x_{i, 2} = s_i - \sqrt{\sigma^2_i} \cdot \mathbf{P}_{k_{i}}$

Time update:
- $s_{k+1} = F_{k+i} \cdot s_{k+1}$
- $P_{k+1} = \sum_{i=1}^{2n} \mathbf{x}_{i, 1, k+1} \cdot \mathbf{x}_{i, 1, k+1}'$
- $X_{0, k+1} = \mathbf{H}_{k+i} \cdot x_{i, 0, k+1}$
- $K_{i} = \sum_{i=1}^{2n} \mathbf{x}_{i, 0, k+1} \cdot \mathbf{x}_{i, 1, k+1}$

Measurement update equations:
- $P_{m+1} = \sum_{i=1}^{2n} \mathbf{P}_{i, 2, k+1} \cdot \mathbf{P}_{i, 2, k+1}'$
- $P_{m+1} = \sum_{i=1}^{2n} \mathbf{x}_{i, 0, k+1} \cdot \mathbf{x}_{i, 0, k+1}'$
- $K_{m+1} = P_{m+1} \cdot \mathbf{H}_{k+i} \cdot \mathbf{P}_{i, 2, k+1}$
- $s_{k+1} = s_{k+1} + K_{m+1} \cdot \mathbf{y}_{k+i}$

where, $s^2 = [x^2_i, y^2_i, z^2_i]$, $K^2 = [x^2_i + y^2_i + z^2_i + K^2_i]$. Appropriate scaling parameter $L$. Invariance of unscented error. $P_{m}$, measurement error. $P_{k}$, measurement noise error. $K_{m}$, weights as calculated in Eq. 15.
Case Studies

One Man Band

Bott et al., 2009
Real Dance

Charbonneau et al., 2009  Charbonneau et al., 2010  Charbonneau et al., 2011

Football

Williamson et al., 2010  Kinect Football by Andrew Devine
RealEdge – FPS

Williamson et al., 2011

Robots

Pfeil et al., 2013
Conclusions – Which to Choose?

- **Wiimote**
  - Positives
    - buttons
    - something to hold in hand
  - Negatives
    - not true 6 DOF
    - challenging to program
    - reasonable accuracy
    - no company support

- **Microsoft Kinect**
  - Positives
    - full body tracking
      - joint position
      - joint orientation (Kinect 2)
    - multimodal input
    - good SDK and support
  - Negatives
    - no buttons (temporal segmentation problem)
    - more data to process
    - not really designed with physical props in mind
    - latency issues (gesture recognition)
Conclusions – Which to Choose?

- PlayStation Move

  **Positives**
  - accurate and fast 6 DOF tracking
  - buttons
  - multimodal input
  - good SDK and support

  **Negatives**
  - requires PS3 (positive as well)
  - does not track full body (more restrictive)

Next Class

- Unity 3D Bootcamp
- Readings
  - Siggraph 2010, 2011 course notes on 3D UI and Video Game Hardware