Electrons, Photons, Phonons, Wave, Bits, and Industrial Design: Microsoft Kinect Sensor

Hot Chips 23

Dawson Yee
Scott McEldowney

Natural User Interface Hardware
Microsoft Corporation
Topics

1. User Experience Goals
2. Design Considerations
3. Product Requirements
4. Design Tactics
5. System Overview
   - Major components & IC’s
   - Depth
   - Audio
   - Thermal
   - Tilt
   - Mechanical structure
6. Other Considerations
   - Robustness
   - Test and Validation
   - Manufacturing and Supply chain
User Experience Goals

• Development of first large scale Natural User Interface System
  – Gesture, video and audio
  – State of the art
    • 3D Imaging
    • Array Microphone (Beam forming & Echo cancellation)

• New requirements
  – Play space range and field of view
  – Ambient light
    • Low lighting for video
    • High lighting for depth dynamic range
  – All room, clothing, etc.
  – Background noise

• Reliable and Affordable
Design Considerations

- Approachable (technology is hidden)
- Fits in with user environment
- Placement of device
- Ease of setup
- Discovery & ease of use
- Self recovery/diagnostics
- Graceful degradation
- Error notification
- Works with all Xbox 360 consoles
- Extensible for future applications and uses
Product Requirements

- Regulatory (EMC, RFI, RoHS,...)
- Reliability over time & cycles
- Operating temperature
- Hot spots
- Humidity
- Shipping stresses (vibration/shake/shock)
- User abuse
  - Drops
  - Dust
Design Tactics

• Don’t know:
  – Future application requirements
  – User base (new segment)
  – User expectations (about performance)

• Do know:
  – Physics
  – Basic function
  – Schedule
  – Cost target
  – Manufacturing and supply chain constraints

• Approach:
  – Work by design (versus work by test)
  – Understand material limits
  – Understand technical limits
  – Control what you know
  – Design margin /conservative design
System Overview

Near Infra-Red Illuminator (projector) and Depth Sensor

RGB Sensor
System Overview:
Overall Assembly & Major Components
Major Components & IC’s

- **USB 2.0 interface (Hub)**
- **Data**
  - Depth
  - RGB
  - Audio out
  - Audio in
  - Tilt

![Diagram](image-url)
Depth: Overall Depth Sensor Design

• An infrared projector combined with a monochrome CMOS sensor allows Kinect to see the room in 3-D

• Structured Light
  • Illumination source
  • Pattern generation
  • Detector/sensor

• IR Sensor
  – High sensor responsivity reduces power consumption
  – Large FOV with low distortion and high MTF lens system
  – Narrow band pass filter reduces interference from ambient incandescent lighting

• Infrared Projector
  – Near IR Laser Diode
  – Laser diode Considerations
    • Temperature control (hold to fraction of degree C)
    • Over operating temp range within boot time
    • Mode Hoping
    • Feedback from other optics
    • Slow ramp
    • Immune to transients
    • Over power/current → catastrophic optical damage (COD)
Depth: Radiometric Design

• Sensor Power Budget
  – Ambient light (Incandescent & Halogen lights, Sun) - SNR
  – Quantum efficiency of sensor (responsiveness in amps generated /optical watt of power)
  – Near vs. Far (dynamic range of sensor) – 1/R^2
  – Corner vs. Center (optics) – cos^4(q)
  – Minimum object size (resolution, illumination) (FOV/pixels)
  – Reflectivity of objects
  – Contrast of Imaging System (MTF)

• Narrow band pass optical filter
  – Blocking undesired ambient light
  – Passing illumination source wavelength
  – Incident angles – wavelength/transmission shift with angle
Optical Modules
Depth: Depth Error

- Calibration will ensure uniform accurate depth leaving the factory
- Must guarantee Uniform Across FOV, Temp, Time, Shake/Shock/Drop
- Mechanical alignment
  - Structured light principle is to measure illumination shift to sub-pixel levels so...
  - Sub-pixel shifts are important – tolerances measured in microns
  - Micron type deflections change depth
    - Drop
    - Temp cycles
    - Shipping
    - Metals and plastics
    - Change in lenses
    - Unintended stresses
Audio

- **Speech** commands (speech recognition)
- **Game chat** (directed full duplex with playback)
- **Video Conferencing** (further back, wider field, full duplex)
- Wideband audio 16kss @24bits
- 4 element beam forming input audio
  - Response matched to dB as built
  - Need to AEC each channel BEFORE beam forming
- Synchronization of Console 5.1 audio output with 4 microphone streams

- Latency
- Received FAN noise
  - e.g. quiet talker @ 3m -> equivalent at sensor
Tilt

- Determine the play space (see the floor)
- Tall & shorter players
- Motor
- Accelerometer
- Speed tilt to target
- Accuracy
- Power draw (peak vs. RMS)
- Wear /thermal effects
- Reliability
- Acoustic noise
Mechanical Structure

- Industrial Design
- Surface fit and finish
- Manufacturing & Assembly (DfX)
- Physical acoustics (microphone, fan)
- Thermal
  - Component operation and reliability
  - Optical elements, Depth error, Case temperature
  - Fan acoustics, vibration & air turbulence (speech, chat, video conferencing)
- Optical alignment
- Shake/shock/shipping/storage
- Impact
- Weight
Other Considerations: Robustness

- Watchdog timers
- Shutdown events
- Recovery from (unintentional) thermal overload
- Error event logging
- Power delivery
  - 12v and 5v for Xbox 360s
  - 5v from Xbox 360 + auxiliary 12v supply
  - 3m extender cable (voltage & power budget)
  - Peak loading (i.e. motor moves)
  - Transient immunity
- Firmware updates
- Ongoing reliability testing
Other Considerations: Test and Validation

- New category, new technology, new methods
- Stress validation (electrical/mechanical/optical/acoustic)
- USB stress (throughput/lost packets)
- Mechanical changes over temperature (optical, acoustic, cooling, tilt)
- Thermal stress & thermal capacity
- Stable operation time from cold or hot start
Other Considerations: Manufacturing and Supply chain

- Unknown new market - skepticism
- “Telecom quality and reliability at consumer price points”
- New suppliers – different industries
- Stretched supply chains (cross-applications)
- Blind & Buried VIA’s
- New assembly processes (molding, stamping, casting, coatings, glues (Tg & modulus, UV, outgassing)
Acknowledgement