Gesture Recognition: A Disruptive Technology?

Erik DeVito, MBA, CSEP–Acq
INCOSE Chesapeake Chapter
January 2012
Outline

- What is Gesture Recognition?
- Historical Highlights
- Technology Overview
- Industry Landscape
- Spotlight on Kinect
- What is Disruptive Technology?
- Is Gesture Recognition a Disruptive Technology?
- Implications for Systems Engineering?
Gesture Recognition is a technology that achieves dynamic human–system interactions that do not require physical, touch, or contact based input mechanisms.

- Types of Gesture Recognition:
  - Hands
  - Full Body
1919– The “Theremin” is an electronic musical instrument that does not require touch but accepts user gestures to create music (Billingshurst, 2011).

Image Source: io9.com
1970s–1980s Videoplace used “real time image processing of live video” to enable user interaction with a responsive display (Billingshurst, 2011).
1993– ALIVE system, used computer and video overlays that enabled user and “virtual agents” to interact in same display space (Maes, Darrell, Blumberg, & Pentland, 1995).
First Commercial Success

2003– The EyeToy, A 2D camera peripheral for the Sony Playstation 2 that translated “body movements into game controls” (Marks, 2010).
Technology Overview
2D Gesture Recognition

- Typically software based—Not requiring special cameras or sensors
  - Requires 3 steps
    - Segmentation
      - Separates pixels around the hands or body from background scene
    - Feature detection
      - Statistical—based on center of equivalent rectangle
      - Contour—based on the outline of users hands or body
    - Gesture recognition
      - Directional (Left, Right, Up, Down)
      - Action Based (Point, Reach, Grab, and Select)
      - Custom libraries

(Du & Li, 2000)
3D Gesture Recognition

- Uses embedded or separate device based cameras, projectors, and sensors to sense depth and motion
  - Microwave
    - Low-power, 9–15 GHz, approx. 2–3cm wavelength
    - Gaze, Eye and Facial Gesture tracking
  - Light Waves
    - Multi-Camera Triangulation
    - Time of Flight (CW or Modulated IR)
    - Structured Light (Grid or Pseudo Random IR)
    - Laser (Object Tracking and Scanning)
  - Ultrasonic Waves
    - Basic low resolution/Single Pass
    - Uses shifts in Doppler frequency to recognize movement

(Castaneda & Navab, 2011; Kalgaonkar & Raj, 2009)
Time of Flight

- Measures depth by recognizing return differences in infrared (IR) light that has been reflected off the user
  - Continuous Wave
    - Measures phase shift of return wave
  - Pulse Modulated
    - Measures absolute pulse return time

Image: Jeff Ludvick

(Castaneda & Navab, 2011)
Structured Light

Projects IR light into a “known pattern” and internal triangulation enables the device to recognize user’s place within scene
- Produces a 3D map of the reflected IR returns and re-calibrates it to projected pattern
- Depth and movement are obtained by continuously remapping the current scene and computing the shift from the previous scene
- Can use a combination of Grid and/or Pseudo Random “Speckles”

(Castaneda & Navab, 2011)
This video was shot with night vision and shows the Microsoft Kinect projecting IR Structured Light into a room. Shows combination of Grid and Pseudo Random “Speckles”

http://www.youtube.com/watch?v=6CFoOFZ6ifc
Industry Landscape
## Companies and Focus Areas

<table>
<thead>
<tr>
<th>Companies</th>
<th>Chip Development</th>
<th>Software/Middleware Development</th>
<th>Hardware Device Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMD</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>ASUS</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ceva</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognivue Corp</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Eedoo</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Extreme Reality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eyeSight</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>GestureTek</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H2i Technologies</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Intel</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>JDSU</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Mesa</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Microsoft</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Movea</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Oblong Industries</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Omek Interactive</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Panasonic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMD Technologies</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Pointgrab</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>PrimeSense</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Qualcomm</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Reactrix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samsung</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>SideKick</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SoftKinetic</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Texas Instruments</td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Gesture Recognition Devices

Images: Asus, PMD Tech, H2ii Inc., Lenovo Corp., SoftKinetic, Panasonic Corp
Spotlight on Kinect
The Device

- A Structured Light Gesture Recognition device
- Uses PrimeSense PS 1080 image processor connected to an IR light and 2 CMOS sensors to facilitate object acquisition and recognition
- Contains 2 microphones and connects to the Xbox 360 via USB 2.0

(Image: Microsoft)

(Microsoft Kinect Teardown, 2010)
The Architecture

[Diagram showing components and connections related to depth and color with labels for audio sources and a USB connection.]
The Story

- Released on Nov 4th 2010 for $149.99
- On the same day, Adafruit Industries announces $2,000 reward to the first person who can deliver a fully workable Open Source Driver
- Later that day after an article on CNET.com Microsoft responds with significant disdain and threatens that they’re going to “work closely with law enforcement and product safety groups to keep Kinect tamper-resistant”
- Adafruit raises the prize to $3,000 stating the company’s response as the reason for the increase

(The Open Kinect Project, 2010, Terdiman, 2010)
On Nov 10th 2010 Adafruit Industries announces a winner and releases open source drivers for Kinect.

On Dec 20th 2010 Microsoft denies that the Kinect was hacked but instead asserts that unencrypted USB was intentional.

Jan 2011, Kinect breaks world record for “Fastest-Selling Consumer Electronics Device” with 8m units in 60d (Announced Later).

Hacker/Project sites for the Kinect sprout up and gain popularity with videos and software downloads posted in abundance.

(Adafruit Industries, 2010; Bishop, 2010; Kinect Confirmed, 2011)
The Story

- On Feb 21st 2011 Microsoft announces Kinect for Windows Software Development Kit will be released in Spring 2011—Former employee and key Kinect developer confesses (on his blog) that he put up the $3,000 reward.

- Kinect for Windows beta SDK released in June 2011
  - Enabled non-commercial development of applications for Kinect.

- Company creates “official” project pages, blog, and “Kinect Website” to highlight what new and innovative projects people are creating with the Kinect.

(Lee, 2011)
Kinect sales for Xbox 360 steamed forward with robust holiday 2011 sales...At the time some estimated 11–14m sold

Kinect Accelerator–Startup incubator program announced–10 companies with innovative Kinect solutions/products/business plans will be selected on Jan 25th 2012 for $20k funding with mentoring by Microsoft developers and managers

Company announced @ CES 2012 that “Kinect is coming to Windows” on February 1st and that sales to date have reached 18m units

(Kinect Accelerator Program, 2011; Use the Power of Kinect, 2012)
Application Areas

- Health Care
- Education
- Mobile Phones
- Automotive
- Robotics/Computer Vision
- Home Media/Appliance Interaction
- Personal Computer Interaction
- Interactive Vending, Signage and Displays
- Security/Identity Management
- Motion Capture for graphics and film effects
- Video Games

Image: Jeff Ludvick
Video # 2

This video shows a person using a delta robot to grasp metal balls using only gestures and body movements. The robot is actuated by three servo motors which are controlled by an AVR microcontroller.

http://www.youtube.com/watch?v=-Ub8LKPkhos&feature=player_embedded
This video shows a person controlling a remote controlled quad copter.

http://youtu.be/A52FqfOi0Ek
This video shows the OmniTouch Natural User Interface created by Chris Harrison @ Carnegie Melon. It uses a modified Microsoft Kinect and Pico Projector to create a unique projection based user interface.

http://youtu.be/Pz17IbjOFn8
What is Disruptive Technology?
Definition

Have “very different package of attributes… mainstream customers historically value… often perform far worse along one or two dimensions important to those customers… mainstream customers are unwilling to use a disruptive product in applications they know and understand… disruptive technologies tend to be used and valued only in new markets or new applications; in fact, they generally make possible the emergence of new markets”

(Bower & Christensen, 1996)
The following were disruptive because they enabled people to do things that they couldn’t previously do with existing technologies:

◦ Telephone
◦ Transistor Radio
◦ Photocopier
◦ Personal Computers
◦ Digital Photography

Each had significant trade-offs, as the disruption occurred they were overcome

(Christensen, Anthony, & Scott, 2003)
Is Gesture Recognition a Disruptive Technology?
1. Signals of Change

2. Competitive Battles

3. Strategic Choices

(Christensen, Anthony, & Scott, 2003)
## Signals of Change

<table>
<thead>
<tr>
<th>Description</th>
<th>Disruptive Technology Theory</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are there signs that someone is capitalizing on opportunities for change?</td>
<td>Serving nonconsumers by introducing a disruption that was previously not convenient, too expensive or more technologically advanced for them to acquire.</td>
<td>As a new market innovator Microsoft has introduced a packaged that is convenient, inexpensive, and does not require the user to have special skills to use it.</td>
</tr>
</tbody>
</table>

(Christensen, Anthony, & Scott, 2003)
<table>
<thead>
<tr>
<th>Description</th>
<th>Disruptive Technology Theory</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the likely result of head to head battles?</td>
<td>The new market entrant has unique and/or abundant resources and is applying them in a way that causes incumbents to fail.</td>
<td>Microsoft is flush with resources. One of the top software firms in the world. This is no start up. Incumbents are scrambling for investments and partnership with other industry powerhouses.</td>
</tr>
</tbody>
</table>

(Christensen, Anthony, & Scott, 2003)
### Strategic Choices

<table>
<thead>
<tr>
<th>Description</th>
<th>Disruptive Technology Theory</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are firms making decisions that increase or decrease their ultimate chances for success?</td>
<td>Successful disruptors harness the forces of disruption by selecting the right value network and marketplace to make their stand.</td>
<td>Microsoft used the Xbox 360, the most popular video gaming device on the market as their launch pad for Gesture Recognition. Additionally, they have adopted a previously successful community development approach to help this technology further gain a foothold. Incumbents have been forced into &quot;also ran&quot; position within the marketplace.</td>
</tr>
</tbody>
</table>

(Christensen, Anthony, & Scott, 2003)
Is Gesture Recognition a Disruptive Technology?

- Yes!!!
  - It is highly likely that Gesture Recognition is in fact a Disruptive Technology.
  - The Kinect could act as the vehicle for this disruption. This is inline with both the definition of and theory behind Disruptive Technology.
  - Microsoft satisfies the 3 Step Process that helps determine if an industry is ripe for disruption.

(Christensen, Anthony, & Scott, 2003)
Implications for Systems Engineering?
“Exponential growth is deceptive. It starts out almost imperceptibly and then explodes with unexpected fury—unexpected, that is, if one does not take care to follow its trajectory”

(Kurzweil, 2006)
Affected SE Technical Processes

- Requirements Definition & Analysis
  - Understand thresholds for Gesture Recognition performance as it relates to stakeholder requirements or expectations.
    - Ground Truth vs. Marketing...I want to be able to use an Xbox Kinect to...!
    - Determine parameters will be used to define performance?

- Architectural Design
  - Identify architectural implications of adding Gesture Recognition into the system architecture.
    - Are Gesture Recognition interface requirements creating an issue for the system’s affordability/ROI?
    - Does Gesture Recognition play well with others—Interoperability Issues?

- Implementation & Integration
  - Gesture Recognition may evolve quickly so a balance must be maintained between selecting system element solutions too early or too late in the process.
    - Too early and you negate enhanced capabilities and opportunities for cost savings
    - Too late and you risk adopting technology that has advanced past your architectural interfaces. REWORK!
In general, international standards and best practices do not exist for Gesture Recognition. This poses some challenge for Systems Engineers. Bright spots do exist:

- OpenNi is an industry-led organization to promote compatibility and interoperability of Natural Interaction (NI) devices applications and middleware.
- Point Cloud Library (PCL) is a cross platform set of algorithms for creating data structures that represent a collection of multi-dimensional points to recreate or represent 3D objects.
Video # 5

- This video created in 2009 by Microsoft to showcase their vision for the future of productivity. Gesture Recognition tech is shown along with other technologies and provides compelling view of how things might evolve in the workplace.

- [http://youtu.be/t5X2PxtvMsU](http://youtu.be/t5X2PxtvMsU)
This video is from www.theverge.com and the shows host Josh Topolsky visiting with Microsoft Advanced Research to get a tour of their “Building 99” and see new innovations with the Xbox Kinect.

http://youtu.be/4DS_zGm_K84
Questions?
References

References


