Learning Objectives

- Understand what types of gestures exist
- Survey gesture technologies
- Understand approaches to developing gestures
- Survey research projects developing gestures
“I believe we will look back on 2010 as the year we expanded beyond the mouse and keyboard and started incorporating more natural forms of interaction such as touch, speech, gestures, handwriting and vision – what computer scientists call the ‘NUI’ or ‘Natural User Interfaces’.”

Steve Ballmer
(then-)CEO Microsoft
Ease versus Expressiveness

- **Keyboard** – expressive but not easy?

- **Mouse** – easy but not expressive?

- ‘**NUI**’ – easy and expressive?
# Mouse Gestures

## Navigation

<table>
<thead>
<tr>
<th>Gesture</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>History Forward</td>
<td>Close</td>
</tr>
<tr>
<td>History Backward</td>
<td>Up a Directory</td>
</tr>
<tr>
<td>Reload</td>
<td>Scroll Up (200px)</td>
</tr>
<tr>
<td>Forced Reload</td>
<td>Scroll Down (200px)</td>
</tr>
<tr>
<td>Homepage</td>
<td>Tabbed Browsing</td>
</tr>
<tr>
<td>New Document</td>
<td>Duplicate Tab</td>
</tr>
<tr>
<td>Duplicate Window</td>
<td>Next Tab</td>
</tr>
<tr>
<td>Minimize Window</td>
<td>Previous Tab</td>
</tr>
<tr>
<td>Maximize/Restore</td>
<td>New Tab</td>
</tr>
</tbody>
</table>

## Image Functions

<table>
<thead>
<tr>
<th>Gesture</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>Double Size</td>
</tr>
<tr>
<td>Open</td>
<td>Half Size</td>
</tr>
<tr>
<td>Hide Image</td>
<td></td>
</tr>
</tbody>
</table>

## Links Functions

<table>
<thead>
<tr>
<th>Gesture</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link in new window</td>
<td></td>
</tr>
<tr>
<td>Link in new tab</td>
<td></td>
</tr>
<tr>
<td>Horizontal Stack</td>
<td></td>
</tr>
<tr>
<td>Open every link dragged over (window)</td>
<td></td>
</tr>
<tr>
<td>Open every link dragged over (tab)</td>
<td></td>
</tr>
</tbody>
</table>

## Miscellaneous

<table>
<thead>
<tr>
<th>Gesture</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>View Source</td>
<td>Start over Image</td>
</tr>
<tr>
<td>View Cookies</td>
<td></td>
</tr>
<tr>
<td>View &lt;META&gt; Info</td>
<td></td>
</tr>
<tr>
<td>Add Bookmark</td>
<td></td>
</tr>
</tbody>
</table>

---

© C Anslow @ VUW
Gesture Limitations?

- Little relationship between operations
- Commands must be sequential
- Hard to refer to previous operations
- Arbitrary order of operations

– Marsh et al. (Study at US Naval Research Lab)
Types of Gestures?

- Cassell’s 1999 classification:
  - Conscious or spontaneous
  - Interactional or propositional

- Hummels and Stapers' classification:
  - Static versus dynamic
## Types of Gestures?

<table>
<thead>
<tr>
<th>Conscious / Witting</th>
<th>Unconscious / Spontaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emblematic</td>
<td>Iconic</td>
</tr>
<tr>
<td>Propositional</td>
<td>Metaphoric</td>
</tr>
<tr>
<td></td>
<td>Deictics</td>
</tr>
<tr>
<td></td>
<td>Beat</td>
</tr>
</tbody>
</table>

Justine Cassell 1999, MIT Media Lab
Types of Gestures?

**Gesture Styles**
- Deictic
- Gesticulation
- Manipulation
- Semaphores
- Sign Language

**Technologies**
- Perceptual
  - Vision
  - Audio
  - Remote sensing
- Non-Perceptual
  - Data gloves
  - Pens
  - Mouse / Joystick
  - Touch surfaces
  - Tangible devices
Exercise #1

- Imagine a technology that can accurately map both of your hands and all fingers in any pose and continuously.

- You are looking to create a mid-air gesture set for alarms. You’d at least want to support:
  - Opening/closing the alarm app
  - Turning on/off a preset alarm time
  - Turning off a ringing alarm
  - Pausing a ringing alarm
  - Adding an alarm time
Design of Touch Gestures

- Present result of a command, request gesture.
- Determine agreement (or lack thereof).
- Microsoft Surface, 24in x 18in, 1024x768

<table>
<thead>
<tr>
<th>REFERENTS</th>
<th>Mean</th>
<th>SD</th>
<th>REFERENTS</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Move a little</td>
<td>1.00</td>
<td>0.00</td>
<td>15. Previous</td>
<td>3.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2. Move a lot</td>
<td>1.00</td>
<td>0.00</td>
<td>16. Next</td>
<td>3.00</td>
<td>0.00</td>
</tr>
<tr>
<td>3. Select single</td>
<td>1.00</td>
<td>0.00</td>
<td>17. Insert</td>
<td>3.33</td>
<td>0.58</td>
</tr>
<tr>
<td>4. Rotate</td>
<td>1.33</td>
<td>0.58</td>
<td>18. Maximize</td>
<td>3.33</td>
<td>0.58</td>
</tr>
<tr>
<td>5. Shrink</td>
<td>1.33</td>
<td>0.58</td>
<td>19. Paste</td>
<td>3.33</td>
<td>1.15</td>
</tr>
<tr>
<td>6. Delete</td>
<td>1.33</td>
<td>0.58</td>
<td>20. Minimize</td>
<td>3.67</td>
<td>0.58</td>
</tr>
<tr>
<td>7. Enlarge</td>
<td>1.33</td>
<td>0.58</td>
<td>21. Cut</td>
<td>3.67</td>
<td>0.58</td>
</tr>
<tr>
<td>8. Pan</td>
<td>1.67</td>
<td>0.58</td>
<td>22. Accept</td>
<td>4.00</td>
<td>1.00</td>
</tr>
<tr>
<td>9. Close</td>
<td>2.00</td>
<td>0.00</td>
<td>23. Reject</td>
<td>4.00</td>
<td>1.00</td>
</tr>
<tr>
<td>10. Zoom in</td>
<td>2.00</td>
<td>0.00</td>
<td>24. Menu access</td>
<td>4.33</td>
<td>0.58</td>
</tr>
<tr>
<td>11. Zoom out</td>
<td>2.00</td>
<td>0.00</td>
<td>25. Help</td>
<td>4.33</td>
<td>0.58</td>
</tr>
<tr>
<td>12. Select group</td>
<td>2.33</td>
<td>0.58</td>
<td>26. Task switch</td>
<td>4.67</td>
<td>0.58</td>
</tr>
<tr>
<td>13. Open</td>
<td>2.33</td>
<td>0.58</td>
<td>27. Undo</td>
<td>5.00</td>
<td>0.00</td>
</tr>
<tr>
<td>14. Duplicate</td>
<td>2.67</td>
<td>1.53</td>
<td><strong>MEAN</strong></td>
<td>2.70</td>
<td>0.47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Form</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>static pose</td>
<td>Hand pose is held in one location.</td>
</tr>
<tr>
<td>dynamic pose</td>
<td>Hand pose changes in one location.</td>
</tr>
<tr>
<td>static pose and path</td>
<td>Hand pose is held as hand moves.</td>
</tr>
<tr>
<td>dynamic pose and path</td>
<td>Hand pose changes as hand moves.</td>
</tr>
<tr>
<td>one-point touch</td>
<td>Static pose with one finger.</td>
</tr>
<tr>
<td>one-point path</td>
<td>Static pose &amp; path with one finger.</td>
</tr>
<tr>
<td>Nature</td>
<td>Description</td>
</tr>
<tr>
<td>symbolic</td>
<td>Gesture visually depicts a symbol.</td>
</tr>
<tr>
<td>physical</td>
<td>Gesture acts physically on objects.</td>
</tr>
<tr>
<td>metaphorical</td>
<td>Gesture indicates a metaphor.</td>
</tr>
<tr>
<td>abstract</td>
<td>Gesture-referent mapping is arbitrary.</td>
</tr>
<tr>
<td>Binding</td>
<td>Description</td>
</tr>
<tr>
<td>object-centric</td>
<td>Location defined w.r.t. object features.</td>
</tr>
<tr>
<td>world-dependent</td>
<td>Location defined w.r.t. world features.</td>
</tr>
<tr>
<td>world-independent</td>
<td>Location can ignore world features.</td>
</tr>
<tr>
<td>mixed dependencies</td>
<td>World-independent plus another.</td>
</tr>
<tr>
<td>Flow</td>
<td>Description</td>
</tr>
<tr>
<td>discrete</td>
<td>Response occurs after the user acts.</td>
</tr>
<tr>
<td>continuous</td>
<td>Response occurs while the user acts.</td>
</tr>
</tbody>
</table>

Examples
Effects of the Display Angle and Physical Size on Large Touch Displays in the Work Place

VALCRI - http://valcri.org
https://www.youtube.com/watch?v=itTBWMKeyRI
ACM ISS 2017
Digital Collaborative Agile Cardwalls

XP 2017 / CMIS 2016

https://www.youtube.com/watch?v=fzCnjnpRiTI

© C Anslow @ VUW
Collaborative Business Process Modeling in Multi-surface Environments

CMIS 2016
https://www.youtube.com/watch?v=GFjj5sa3GvE
Properties of Touchlessness

• Proxemic

• Transfer of Matter

• Momentum & Pressure

• Constraints on Movement

• Haptic Feedback
Haptic

Design:
• Uses sense of touch.
• Can be combined into virtual reality.
• User receives feedback via tactile sensation.

Issues:
• Sense of touch very complex
  – Different nerves have different functions
  – Pressure
  – Temperature
  – Muscle
  – Hair movement
• Texture of surface particularly difficult.
Novint's “Falcon”

https://www.youtube.com/watch?v=gjAxGVH1JOM
PHANTOM

https://www.youtube.com/watch?v=REA97hRX0WQ
Ultra Haptics

https://www.youtube.com/watch?v=GDra4lJmJN0
Nintendo Wii

- Originally 3 axes of acceleration
  - Wii MotionPlus gyroscopes added 3 axes of orientational change
- Sensor bar has 2 groups of IR LEDs
- Wiimote has camera that detects the two LED groups
- Detects acceleration, not positional change
  - Can lead to “waggle” “cheating”
  - Small movements interpreted as full movements
http://johnnylee.net/projects/wii/
PlayStation Move

- 3-axis accelerometer and 3-axis gyroscope
- RGB LED on top of motion controller
  - Used by Eye to determine controller position
  - Generates own light, reducing scene light problem
  - Light can be adjusted to suit scene
  - Eye captures the 2D ellipse parameters of the sphere – calculate
- 1-1 mapping of controller motion to virtual object?
3D Cameras and Gestures

- 2D cameras can't handle arbitrary scenes
- 3D cameras incorporate Z-sensing
  - Powerful segmentation feature
  - Provides 3D spatial information
- Possible implementation:
  - Active infrared illumination and camera
- Possible alternative:
  - Pulsed light emitter with active shuttering of camera
Kinect - 2010
Skeletal Models
Leap Motion 2012

https://www.youtube.com/watch?v=_d6KuiuteIA
3D Gestures Classification and Literature Review


Video/Vision Costs/Benefits

- User adaptability and feedback
  - Fixed palette of gestures, or adapt during use

- Learnability
  - Learning rate and memorability for users
  - Performance versus gesture set size

- Accuracy
  - Hand within the camera view, moving, occluded, shape, colour
  - Gestures designed for camera accuracy versus intuitive
Video/Vision Costs/Benefits

- **Low mental load**
  - Recalling gestures and trajectory adds to load
  - Short and natural gestures

- **Intuitiveness**
  - Clear cognitive association with task
  - No complex shapes and unnatural combinations
  - Tied to cultural background, no commonality

- **Comfort**
  - Avoid requiring intense muscle tension - “Gorilla Arm” syndrome
  - No awkward repetitive gestures
Video/Vision Costs/Benefits

- **Lexicon size and multi-hand systems**
  - Single hand, dual hand, arm, body, ?
  - Recogniser for small # of gestures is best

- **Come as you are**
  - How encumbered can the user be, and in what environment
  - Impact on setup time
  - IR for near, far, ultrasonic, other imagers?

- **Reconfigurability**
  - Not providing a huge palette to start, but modifying the palette
  - Modify palette for different tasks
Video/Vision Costs/Benefits

- **Interaction space**
  - Where can gestures be recognised (virtual interaction envelope)
  - Stereo vision is best, but possible for mobile platforms?

- **Gesture spotting and the immersion syndrome**
  - Unintended movement recognition (Midas Touch problem)
  - Cue selection approaches (voice, buttons, gestures, ?)

- **Ubiquity and wearability**
  - Small sensors and cameras placed where on the body?
Minority Report - 2002

Oblong Technologies - GSpeak

https://www.youtube.com/watch?v=fe0fHTHEL9w
https://www.ted.com/talks/john_underkoffler_drive_3d_data_with_a_gesture
Postgraduate research at the University of Auckland
Kinect to Architecture

IVCNZ 2011, Auckland, NZ
20 August 2020

Presented by
Robert Amor
Department of Computer Science
in association with
Leroy D’Souza, Isuru Pathirana and Dermott McMeel
What is an appropriate gesture?

- Natural to user
  - Non fatiguing
  - Full body due to technology

- Iterated over a range of arm, leg, body postures
## Final gesture set

<table>
<thead>
<tr>
<th>Gesture</th>
<th>Functionality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point forward</td>
<td>Move forward</td>
</tr>
<tr>
<td>Step forward</td>
<td>Move forward</td>
</tr>
<tr>
<td>Step backward</td>
<td>Move backward</td>
</tr>
<tr>
<td>Point up</td>
<td>Move to an upper storey</td>
</tr>
<tr>
<td>Point down</td>
<td>Move to a lower storey</td>
</tr>
<tr>
<td>Point left/right</td>
<td>Pan camera left/right</td>
</tr>
<tr>
<td>“Hold-up” pose</td>
<td>Calibrate user to system</td>
</tr>
<tr>
<td>Push hand forward</td>
<td>Select</td>
</tr>
<tr>
<td>Cross arms</td>
<td>Return to starting position</td>
</tr>
<tr>
<td>Cross arms (hold for 2 seconds)</td>
<td>Return to model selection</td>
</tr>
</tbody>
</table>
Gestures adapted for each user

- Naïve approach used fixed distances to trigger a gesture
- Testing identified poor performance
- Adapted to individual users

$0.2 \times \text{User Height}$
Avatar in the scene?
Ease of use and accuracy

**Gestures**

- Cross
- Point to pan camera
- Back
- Down
- Up
- Forward (Leg)
- Forward (Hand)
- Point to browse models
- Push

<table>
<thead>
<tr>
<th>Gesture</th>
<th>Easy to use</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Point to pan camera</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Back</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Down</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Up</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Forward (Leg)</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Forward (Hand)</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Point to browse models</td>
<td>80%</td>
<td>80%</td>
</tr>
<tr>
<td>Push</td>
<td>80%</td>
<td>80%</td>
</tr>
</tbody>
</table>

**Percentage**
Responsiveness and memorability

<table>
<thead>
<tr>
<th>Gestures</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross</td>
<td></td>
</tr>
<tr>
<td>Point to pan camera</td>
<td></td>
</tr>
<tr>
<td>Back</td>
<td></td>
</tr>
<tr>
<td>Down</td>
<td></td>
</tr>
<tr>
<td>Up</td>
<td></td>
</tr>
<tr>
<td>Forward (Leg)</td>
<td></td>
</tr>
<tr>
<td>Forward (Hand)</td>
<td></td>
</tr>
<tr>
<td>Point to browse models</td>
<td></td>
</tr>
<tr>
<td>Push</td>
<td></td>
</tr>
</tbody>
</table>

- Responsive
- Memorability
Fatigue

- Cross
- Point to pan camera
- Back
- Down
- Up
- Forward (Leg)
- Forward (Hand)
- Point to browse models
- Push

**Gestures**

**Percentage**

**Level of Fatigue**
User Elicited Hand Gestures for VR-based Navigation of Architectural Designs

Karim Cissé, Aprajit Gandhi, Danielle Lottridge, Robert Amor

School of Computer Science
University of Auckland, New Zealand

14 August 2020
Navigating Complex 3D Environments
Leap Motion and Oculus Rift setup
**Gesture Development Process**

1. **Interview Professionals** (6)
   - Primary and alternate gestures

2. **Elicited Gestures** (64)
   - Popularity; Physical demand; Recognisability; Intuitiveness; Fit

3. **Filter Gestures**
   - Selected Gestures (8)

4. **Evaluation Study** (12)
Mid-Air Hand Gesture Set

(a) Move forward;
(b) Move forward double speed;
(c) Move backwards;
(d) Move up a floor;
(e) Move down a floor;
(f) Rotate right;
(g) Rotate left;
(h) Main menu.
Gesture Usability Ratings

- Intuitiveness
- Memorability
- Comfort

Gestures:
- Forward
- Backward
- Up a floor
- Down a floor
- Rotate left
- Rotate right
- Forward (2X speed)
- Main menu

Average rating:

- 5.00
- 4.50
- 4.00
- 3.50
- 3.00
- 2.50
- 2.00
- 1.50
- 1.00
- 0.50
- 0.00
Mean Self-Reported Task Load, Raw-TLX
Conclusions

- Design professionals identified domain and task specific gestures
- Developed a coherent mid-air gesture set
- Accurately recognised by Leap Motion controller
- Participants rated them as memorable, intuitive and physically non-demanding
- Identified need to reassess the rotate gesture
- Identified need for more gestures
- Identified need for social interaction features
Summary

- Many categories of gestures across touch and touchless
- Increasing sophistication and accuracy of devices
- Conflicting range of criteria to be considered in gesture design
  - Consider domain-user proposed gestures