UNIVERSITY OF CRETE COMPUTER SCIENCE DEPARTMENT

COURSE CS-564 (OPTIONAL)

ADVANCED TOPICS

IN HUMAN – COMPUTER INTERACTION

Course Convenor: Constantine Stephanidis

Recent and emerging interaction techniques: Touch and Gesture Interaction

Touch Interaction



Introduction



- Hands have an impressive ability to perform different kinds of manipulation tasks:
 - from working with miniature details
 - to lifting heavy objects
- The excellent performance of the hand in natural contexts is based on the efficiency of the interaction between sensors and muscles in the hand and the environment



Touch Sensing



- The foundation for all touch interaction is touch sensing
 - i.e., technologies that capture human touch and gestures
- This includes sensing touch using:
 - cameras or arrays of optical elements
 - laser rangefinders
 - resistance and pressure sensors
 - acoustics
 - etc.
- Touchscreens are built on top of these technologies



Touchscreens (1/2)

- Touchscreens are displays where the position of contact with the screen can be detected
 - Interaction is achieved either through pointed physical objects (e.g., pens) or fingers
- A touchscreen behaves as a two dimensional planar smart skin
 - Wherever it is touched, some virtual object can be activated
- The interaction paradigm is easily customizable to accommodate a range of applications and users



Touchscreens (2/2)

- They share some characteristics that make them ideal for many workplaces and public spaces:
 - Easy to learn
 - Durable
 - Require no additional work space
 - No moving parts
- Today they are used routinely in many applications and many devices such as:
 - point of sales terminals
 - mobile devices
 - games consoles
 - vehicle navigation systems
 - information kiosks



Multi-touch Technology

- Multi-touch technology allows multiple users to operate simultaneously
- Various technologies have been introduced to develop multi-touch capable devices
- These techniques can be classified into two categories
 - Computer Vision-Based
 - Purely Vision-Based
 - Everywhere Display [Pinhanez '03]
 - Vision- and Optical- Based
 - Microsoft Surface

- Sensor-Based
 - iPhone
 - SmartSkin [Rekimoto '02]
 - Diamond Touch [Dietz & Leigh '01]



Purely Vision-Based Systems

- Due to the decreasing cost and improved performance of computers, computer vision technology has been greatly improved
 - This enables the processing of real-time, and high-speed video signals
 - and is sufficient to meet the real-time interaction and human-computer interaction requirements
- Touches and their positions are identified through image processing techniques
- This type of multi-touch systems are highly portable
 - They can be used on any flat surface without the need for a dedicated display device
- However, the flexibility of pure vision systems comes at the cost of precision



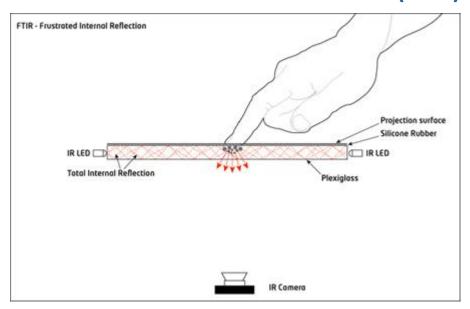
Vision- and Optical- Based Systems (1/2)

- Devices based on computer vision and optical Multi-touch technology:
 - have good scalability
 - have a relatively low cost
 - but they have a larger volume
- There are two kinds of computer vision and optical-based Multitouch systems:
 - Frustrated Total Internal Reflection (FTIR)
 - Diffused Illumination (DI)



Vision- and Optical- Based Systems (2/2)

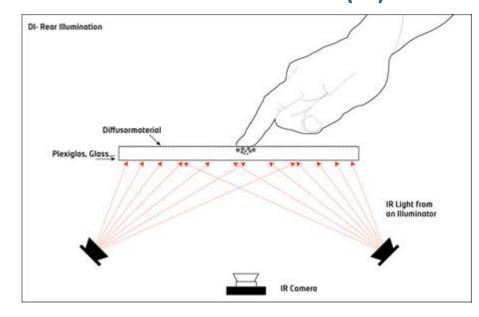
Frustrated Total Internal Reflection (FTIR)



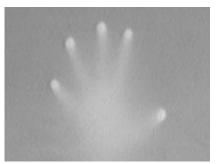
Beams of LED totally reflect within an acrylic medium and when occluded by a finger the scattered light reveals its position



Diffused Illumination (DI)



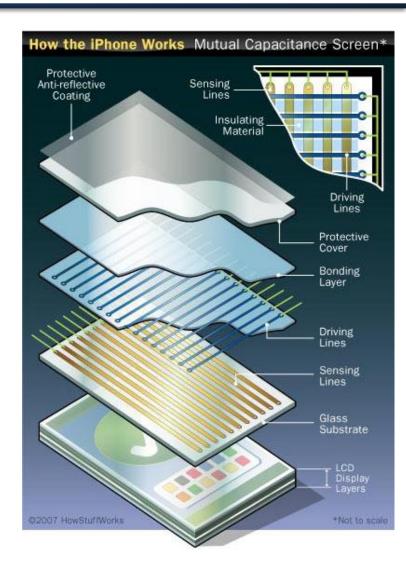
Infrared light shines a surface, a finger reflects more light upon touch thus revealing its position





Sensor-Based Systems

- Many Multi-Touch Devices based on sensor technology can simultaneously detect multiple touch points
- Unlike some of the computer-vision-based systems, sensor based systems are almost impossible to build from off-the-shelf components
- The cost is prohibitively high, and the environment temperature and humidity will affect the system performance
- However, because the sensor can be integrated in the surface, it can be used for mobile phones, PDAs and other small-screen handheld devices





The Key Technology of Multi-touch

- Multi-touch technology can be simply divided into two parts:
 - Hardware: serves to complete the information collection
 - Software: serves to complete the analysis of information which are finally converted into specific user commands
- Multi-touch technology should include the following components:
 - Multi-touch Hardware Platform
 - Each platform may have its own advantages and disadvantages
 - Accuracy of Selection
 - Precision has great significance on how to accurately track and locate contacts to achieve the freedom of gesture interaction
 - Identification Technology
 - Existing Multi-touch technology detects the contact without carrying information of users
 - The technology that can now identify the user's identity is Diamond Touch technology (which can identify up to four users)
 - Bimanual Interactive Technology
 - A more natural interaction process
 - Two-handed interaction techniques increase parallelism by using both hands of the user

Tangible and Haptic Interaction



Overview

- When touch is described as a skin sense, it is often considered as a passive receiver of stimulation from the environment
- In opposition to this view, a hand can be regarded as a perceptual system which relies on exploration to collect information, and active touch is in many contexts considered to be superior to passive touch
- To stress the importance of activity, the sense is often called "active touch" or "haptics"
 - The use of the latter two terms is not consistent in the literature
- Sometimes haptics provides more advanced information than vision, for instance concerning the weight of objects and the texture and hardness of surfaces



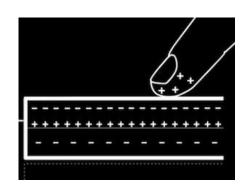
Haptic Interaction

- Most devices today are still in the stone age of haptics they can intelligently vibrate to communicate different things to the user—but that's about it
- While effective, this basic system is very one dimensional, in that the entire phone vibrates instead of just the key that you're pressing
- The next generation of haptics aims to make the tactile experience much more nuanced and useful, both on devices and in the air above them



Disney Research Tactile Display

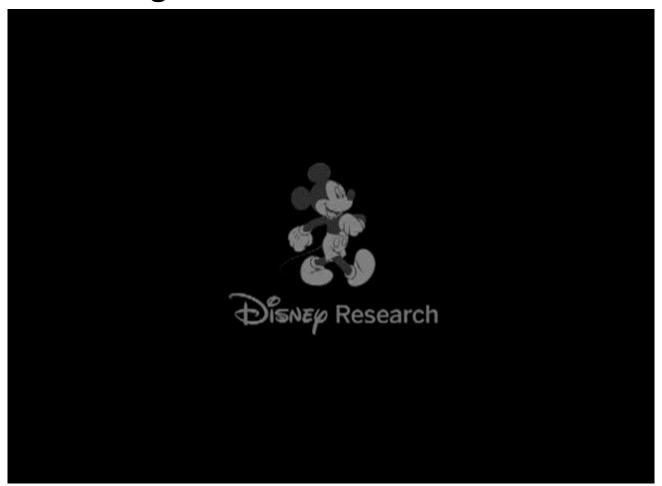
- Disney Research presented an algorithm that is able to translate 3D information in an image or video directly into tactile sensations on a special haptic display
- The display itself stays perfectly smooth, and instead modulates the friction at your fingertips to trick you into feeling like there's texture under them
- The display creates the illusion of friction using another Disney Research technology from several years ago called TeslaTouch, which uses oscillating electric charges to dynamically adjust the friction between your finger and the touch panel





Disney research videos (1/2)

Tactile rendering of 3D features on Touch Surfaces





Disney research videos (2/2)

Tesla Touch Technology

TeslaTouch A Tactile Texture Display



UltraHaptics

- UltraHaptics, from the University of Bristol's Interaction and Graphics research group, does away with tactile displays entirely, and brings touch interaction into the air
- A transducer array projects carefully calculated waves of ultrasonic sound into the air, which you can't see, hear, or feel
- At certain points, however, the waves come into focus and intensify substantially, displacing the air at those points and creating a pressure difference that you can feel
- The system can create multiple pressure points in different locations at the same time, and can even endow individual points with distinct tactile properties



UltraHaptics video

UltraHaptics:

multi-point mid-air haptic feedback for touch surfaces



Tangible Interfaces

- A Tangible User Interface (TUI) is a user interface that augments the real physical world by connecting digital information to everyday physical objects and environments
 - They allow users to quite literally grasp data with their hands
 - They often function as both input and output devices
- Tangible Interfaces draw upon the human urge to be active and creative with one's hands
- They can be realized by attaching micro sensors and actuators to physical objects



Application Domains

- Dominant application areas for TUIs seem to be:
 - Learning
 - Support of planning and problem solving
 - Programming and simulation tools
 - Support of information visualization and exploration
 - Entertainment
 - Play
 - Performance and music
 - Social communication



CapStones and ZebraWidgets

 Chan et al. [2012] use stackable gaming pieces and tangible widgets with moving parts to interact with an underlying Touchscreen



Enabling a game of tangible checkers, the touchscreen distinguishes a stack of two CapStones from the single CapStones surrounding it



- (a) A single Zebra Dial allows users to adjust the brightness of the underlying image.
- (b) Placing a second Zebra Dial on top, allows adjusting brightness and contrast



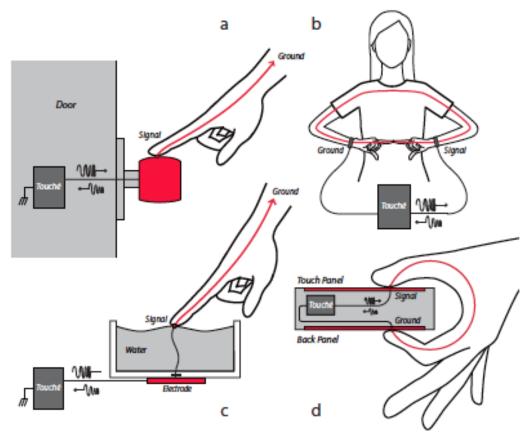
The Touché Approach (1/3)

- Touché proposes a novel technique that can:
 - detect touch events
 - recognize complex configurations of the human hands and body
- Uses the conductive properties of the human body to detect:
 - how the user is touching the object
 - how the user is connected to the ground
 - the current configuration of the human body and individual body properties
- As a user touches different objects or parts of their body, signals flow through slightly different paths and denote various messages



The Touché Approach (2/3)

- The user interacts with an object that is attached to a Touché sensor board via a single wire
- If the object itself is conductive, the wire can be attached directly to it
- Otherwise, a single electrode has to be embedded into the object and the wire attached to this electrode



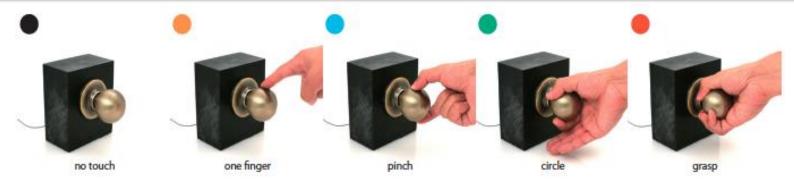


The Touché Approach (3/3)

- Example Touché Applications
 - Making everyday objects touch gesture sensitive
 - Sensing human bimanual hand gestures
 - Sensing human body configuration (e.g., pose)
 - Enhancing traditional touch interfaces
 - Sensing interaction with unusual materials (e.g., liquids)



Touché: Making objects touch and grasp sensitive



- If analogue or digital objects can be made aware of how they are being touched, held or manipulated, they could configure themselves in meaningful and productive ways
- Touché offers a lightweight, non-invasive sensing approach that makes it very easy to add touch and gesture sensitivity to everyday objects
- Example A smart doorknob that can sense how a user is touching it could have many useful features
 - closing a door with a tight grasp could lock it
 - closing it with a pinch might set a user's away message, e.g., "back in five minutes"
 - a sequence of grasps could constitute a "grasp password" that would allow an authorized user to unlock the door



Touché: Body Configuration Sensing

- Touché can be used to sense the configuration of the entire human body
- Example 1 a door could sense:
 - if a person is simply standing next to it
 - if they have raised their arm to knock on it
 - If they are pushing the door
 - or if they are leaning against it
- **Example 2** a chair or a table could sense the posture of a seated person:
 - reclined or leaning forward
 - arms on the armrests or not
 - one or two arms operating on the surface





Touché: Enhancing Touchscreen Interaction

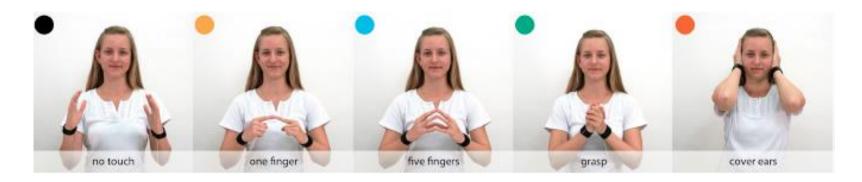
- Touché brings new and rich interaction dimensions to conventional touch surfaces by enhancing touch with sensed hand posture
- Example Touché could sense the configuration of:
 - fingers holding a device
 - if they are closed into a fist or held open
 - whether a single finger is touching, all five fingers, or the entire palm
 - the part of the hand touching the device
 - fingertips or knuckles





Touché: On-Body Gesture Sensing

- Unlike inanimate physical objects, the human body is highly variable and uncontrolled, making it a particularly challenging "input device"
- Touché takes advantage of the conductive properties of the body and uses the skin as a touch sensitive surface while being minimally invasive
 - Electrodes can be worn like a wristwatch





Touché: Sensing Gestures in Liquids

- By interacting with water, we do not mean using touch screens under water, but touching the water itself
- Touché can distinguish between a user touching the water's surface and dipping their finger into it
 - by placing an electrode on the bottom of the water vessel





Touché: Video

Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects

http://www.youtube.com/watch?v=uanM3YGflVw

Touché:

Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects

Munehiko Sato, Ivan Poupyrev, Chris Harrison

CHI 2012 Paper Video Figure



Gestures



Definition

"A gesture is a motion of the body that contains information. Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed."

[Kurtenbach and Hulteen (1990)]

"Any physical movement that a digital system can sense and respond to without the aid of a traditional pointing device such as a mouse or stylus. A wave, a head nod, a touch, a toe tap, and even a raised eyebrow can be a gesture."

[Dan Saffer – Designing Gestural Interfaces (2009)]



The role of gesture in communication

- Gesture plays an important role in pre-linguistic communication for babies
- As well as aids cognition and fully linguistic communication for adults
 - People use gesture to conceptually plan speech production
 - and to communicate thoughts that are not easily verbalized
- Gesture also plays a helpful role for the speaker
 - Gesturing has been shown to lighten cognitive load for both adults and children
- Systems that constraint gestural abilities (e.g,. having your hands stuck on a keyboard) are likely to hinder the user's thinking and communication



Gesture Classification

Several studies have dealt with the classification of gestures, examining the role
of gestures on their own, and along with speech as a representational format or
as a means for displaying thoughts not conveyed in speech

| Gesture | Description | Example |
|--------------------------|--|--|
| Beat | Small baton like movements | A short beat marking an important point of a conversation |
| Cohesive | Used to tie together thematically related but temporally separated portions of discourse | Hand gestures during a politician's speech, e.g., when highlighting a series of points |
| Deictic | Aspects of the discourse are spatialized or located in the physical space in front of the narrator | Put that there |
| Emblematic / Symbolic | Gestures that have a specific single meaning within each culture | The American V-for-victory |
| Iconic / pantomimic | Some feature of the action or event being described is depicted | Would you like to go [fishing] tomorrow?, accompanied by a gesture of throwing the fishing line to the water |
| Metaphoric | The represented concept has no physical form | The meeting went [on and on], accompanied by a rolling motion of the hand |



Human Based Gesture Vocabulary (1/2)

- Gesture research shows that there is no such thing as a universal gesture vocabulary
 - so a good gesture vocabulary may only match one specific application and user group
- The gesture vocabulary must be tailored for the specific task and contain gestures:
 - Easy to perform and remember
 - Intuitive
 - Metaphorically and iconically logical towards functionality
 - Ergonomic; not physically stressing when used often
- The system should be able to recognize gestures unambiguously



Human Based Gesture Vocabulary (2/2)

- Biomechanics and ergonomics tell about constraints in postures and the usage of the gestures, such as avoiding to stay in static positions, and moving joints too far from their neutral positions
- There are constraining relationships between individual joints within fingers and between neighboring fingers
 - These constraints are different in people
- The ergonomics show that it is important to make the recognition algorithms tolerant to de-stressing movements, which allows the user to avoid staying fixed in e.g. a static "residue" or "pointing" gesture
- Tolerance for deviations in gestures is desirable when implementing gesture interfaces
 - Because of varying hand shapes and posture performance

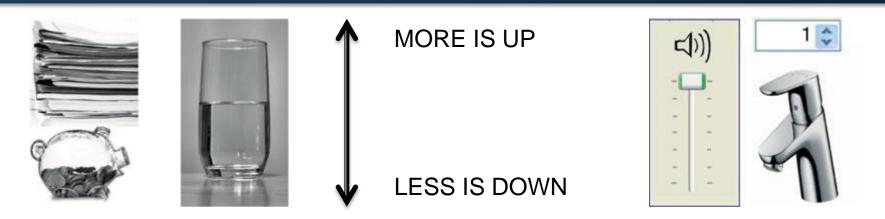


Image Schemas

- Image schemas are knowledge about the world that is expected of a typical user
- Everyone develops image schemas starting from their earliest interactions with the world as babies
- They are developed based on experiences with the physical world and they allow people to apply general concepts to multiple scenarios
 - Up-Down us an excellent example of an image schema
 - Up-Down is intuitively applied to many concepts like quantity or volume



Primary Metaphors



- Primary metaphors are the relationships between the image schema and the abstract concepts
- When designing natural gestures, primary metaphors should be leveraged whenever possible to express abstract concepts within the application
 - the primary metaphor MORE IS UP LESS IS DOWN can be found in a vertical sliding gesture for controlling the volume of speakers
 - when we walk along a path, waypoints in front of us will be reached at a later time from now. This
 pervasive experience grounds the metaphor THE FUTURE IS IN FRONT THE PAST IS BEHIND
 - familiarity tends to co-occur with physical closeness, forming the metaphor FAMILIAR IS NEAR –
 UNFAMILIAR IS FAR



Procedure to Finding Gestures

- Nielsen et al (2004), presented some important issues in choosing the set of gestures for an interface from a user-centred view such as the learning rate, ergonomics, and intuition
- A procedure was proposed for selecting appropriate gestures:

A. Find the Functions

Find the functions needed by the application and which the gestures will have to communicate

B. User Tests – Collect Gestures from User Domain

- The goal is to find the gestures that represent the functions found in step A
- This is done through experiments with people by taking them through scenarios under camera surveillance where they communicate the above functions

C. Analysis – Extract Gesture Vocabulary

 The video recorded data is evaluated to extract the gestures that the testees used in their interaction

D. Test – Benchmark the Chosen Gesture Vocabulary

Test the resulting gesture vocabulary



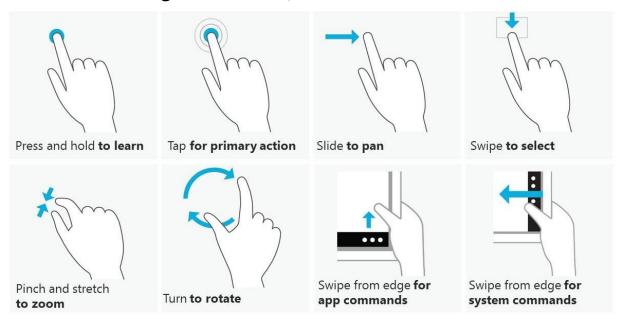
Gesture Interfaces Categories

- Gestures on a touch display or pad
 - as in the iPad or an ATM
 - also known as TUI (touch user interfaces)
- 2. Gestures through a controller
 - such as the Wiimote
- 3. Gestures identified by a vision-based system
 - such as the Kinect
 - also known as NUI (natural interaction interfaces)



Touchscreen gestures

- Interfaces manipulated by gestures performed by the finger or hand on a touchscreen or touchpad
 - Most common class of gesture interfaces
 - Point of Sales systems (POS) and kiosks are common public applications
- Multi-touch systems
 - Introduced to the mass market by Apple through MacBook's MagicPad and made popular through the iPhone/iPod Touch mobile devices







iOS 7 Gesture Interaction Guidelines

Avoid associating different actions with the standard gestures

 Unless your app is a game, redefining the meaning of a standard gesture may confuse people and make your app harder to use

Avoid creating custom gestures that invoke the same actions as the standard gestures

 People are used to the behavior of the standard gestures, and they don't appreciate being expected to learn different ways to do the same thing

Use complex gestures as shortcuts to expedite a task, not as the only way to perform it

As much as possible, always give users a simple, straightforward way to perform an action, even if it means an
extra tap or two. Simple gestures let users focus on the experience and the content, not the interaction

In general, avoid defining new gestures unless your app is a game

- In games and other immersive apps, custom gestures can be a fun part of the experience. But in apps that help people do things that are important to them, it's best to use standard gestures because people don't have to make an effort to discover them or remember them
- Always bear in mind that nonstandard gestures aren't discoverable and should rarely, if ever, be the only way to perform an action



Vision-based gesture interfaces

- Interfaces that are operated through gestures captured by one or more cameras
- Examples include the Gestix system (see figure) or Sign Language Interfaces



Gestix Gestural Interface

- Such interfaces may often be a subset of Perceptual Interfaces
 - Multimodal Interfaces that track one or more of the following: speech, expressions, body, head or eye movements, gestures



Example: the Gestix Gestural Interface

 Gestix is a sterile gesture interface for users, such as doctors/surgeons, to browse medical images in a dynamic medical environment



- A vision-based gesture capture system interprets user's gestures in real-time to navigate through and manipulate an image and data visualization environment
- Dynamic navigation gestures are translated to commands based on their relative positions on the screen
- A state machine switches from navigation gestures to others such as zoom and rotate



Gestix video

<u>Innovative Technology: Hand Gesture system</u> available at: http://www.youtube.com/watch?v=gSyjFLBsIHg





Applications

Apart from the devices mentioned:

- Sign Language Interfaces
- Remote device operation through gestures
 - The Clapper
 - a sound activated electrical switch
 - first mass market gestural interface
 - TV zapping through hand motions
 - PowerPoint Presentations
 - Public restrooms (faucets, toilet paper dispensers, lights)
- Gaming (Wii, Arcades)



Why use them (1/3)

More natural interactions

- Human beings are physical creatures; they like to interact directly with objects
- Interactive gestures allow users to interact naturally with digital objects in a physical way, like they do with physical objects

Less cumbersome or visible hardware

- With many gestural systems, the usual hardware of a keyboard and a mouse isn't necessary: a touchscreen or other sensors allow users to perform actions without this hardware
- This benefit allows for gestural interfaces to be put in places where a traditional computer configuration would be impractical or out of place, such as in retail stores, museums, airports, and other public spaces



Why use them (2/3)

- Eyes-free
 - Performed without having to watch the screen closely
- Silent interaction
 - Less obtrusive than e.g., voice input
- More flexibility
 - As opposed to fixed, physical buttons, a touchscreen, like all digital displays, can change at will, allowing for many different configurations depending on functionality requirements
 - Thus, a very small screen (such as those on most consumer electronics devices or appliances) can change buttons as needed
 - This can have usability issues, but the ability to have many controls in a small space can be a huge asset for designers
 - With non-touchscreen gestures, the sky is the limit, space-wise. One small sensor, which can be nearly invisible, can detect enough input to control the system. No physical controls or even a screen are required



Why use them (3/3)

More nuance

- Keyboards, mice, trackballs, styli, and other input devices, although excellent for many situations, are simply not as able to convey as much subtlety as the human body
- A raised eyebrow, a wagging finger, or crossed arms can deliver a wealth of meaning in addition to controlling a tool. Gestural systems have not begun to completely tap the wide emotional palette of humans that they can, and likely will, eventually exploit

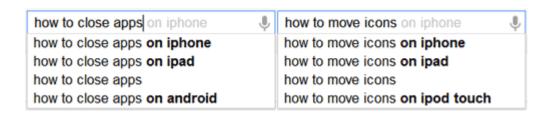
More fun

- You can design a game in which users press a button and an on-screen avatar swings a tennis racket
- But it is simply more entertaining—for both players and observers to mimic swinging a tennis racket physically and see the action mirrored on-screen
- Gestural systems encourage play and exploration of a system by providing a more hands-on (sometimes literally) experience



Challenges: Discoverability

- Similarly to command-line interfaces, gesture interfaces do not offer clues as to what the gestures actually are (memorization vs. GUI recognition)
 - GUIs offer verbs as commands in buttons ("delete", "apply"). Gestures
 are the verbs themselves
- Some gestures are becoming mainstream and are fairly obvious to discover, such as
 - pinch and expand for zoom, flick through lists, two-finger rotation, etc.
- Others are impossible to discover alone



Frequent searches on Google reveal the problem



Challenges: No common vocabulary

- Not all gesture interfaces use the same gestures for the same actions
 - One reason is patents, which is still a grey area
 - Technology is young and there is no standard way of doing things
 - The problem of discoverability means developers seek new approaches in the quest for an optimal, natural solution
- This is not an area for product differentiation. This is an area that requires standardization, predictability, and consistency with user expectations.



Challenges: Vision-based interfaces

- The challenges are similar to those of other vision-based interfaces (such as eye-tracking or head-pose)
 - Poor lighting conditions
 - Camera specifications of depth and angle
 - Demanding processing
 - Identifying similar gestures
 - Accuracy



Reasons NOT to have a Gestural Interface (1/3)

Heavy data input

 Although some users adapt to touchscreen keyboards easily, a keyboard is decidedly faster for most people to use when they are entering text or numbers

Inappropriate for context

- The environment can be non-conducive to a gestural interface in any number of situations, either due to privacy reasons or simply to avoid embarrassing the system's users
- Designers need to take into account the probable environment of use and determine what, if any, kind of gesture will work in that environment



Reasons NOT to have a Gestural Interface (2/3)

Reliance on the physical

- Gestural interfaces can be more physically demanding than a keyboard/screen
- The broader and more physical the gesture is (such as a kick), the more likely that some people won't be able to perform the gesture due to age, infirmity, or simply environmental conditions
 - E.g. pressing touchscreen buttons in winter gloves is difficult
- The inverse is also true: the subtler and smaller the movement, the less likely everyone will be able to perform it
- The keyboard on the iPhone, for instance, is entirely too small and delicate to be used by anyone whose fingers are large or otherwise not nimble



Reasons NOT to have a Gestural Interface (3/3)

Reliance on the visual

- Many gestural interfaces use visual feedback alone to indicate that an action has taken place (such as a button being pressed)
- In addition, most touchscreens and many gestural systems in general rely entirely on visual displays with little to no haptic affordances or feedback
- There is often no physical feeling that a button has been pressed, for instance. If your users are visually impaired (as most adults over a certain age are) a gestural interface may not be appropriate

This problem may actually have a solution soon



References

- Chang, R., Wang, F., & You, P. (2010, April). A Survey on the Development of Multi-touch Technology. In Wearable Computing Systems (APWCS), 2010 Asia-Pacific Conference on (pp. 363-366). IEEE. (Available through the course website)
- Hurtienne, J., Stößel, C., Sturm, C., Maus, A., Rötting, M., Langdon, P., & Clarkson, J. (2010). Physical gestures for abstract concepts: Inclusive design with primary metaphors. Interacting with Computers, 22(6), 475-484. (Available through the course website)
- Sato, M., Poupyrev, I., & Harrison, C. (2012, May). Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 483-492). ACM. (Available through the course website)



Additional References

- Chan, L., Müller, S., Roudaut, A., & Baudisch, P. (2012, May). CapStones and ZebraWidgets: sensing stacks of building blocks, dials and sliders on capacitive touch screens. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 2189-2192). ACM. Dietz, P., & Leigh, D. (2001, November).
- DiamondTouch: a multi-user touch technology. In Proceedings of the 14th annual ACM symposium on User interface software and technology (pp. 219-226). ACM.
- Pinhanez, C., Kjeldsen, R., Tang, L., Levas, A., Podlaseck, M., Sukaviriya, N., & Pingali, G. (2003, November). Creating touch-screens anywhere with interactive projected displays. In Proceedings of the eleventh ACM international conference on Multimedia (pp. 460-461). ACM.
- Rekimoto, J. (2002, April). SmartSkin: an infrastructure for freehand manipulation on interactive surfaces. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 113-120). ACM.



Questions?