Input / Output
Announcements

You hopefully submitted your first commentaries

Due at 5pm night before lecture

You will receive section and discussant assignments by early next week.

Sections start next Wed Apr 10

Email cs347@cs.stanford.edu with questions or requests
Last time

**Ubiquitous computing:** a vision in which computers “vanish into the background” rather than focus our attention on a single box

**Tangible computing:** a subset of ubiquitous computing in which all data, interaction, and representation are encoded physically

Dourish [2004]: “Tangible computing is of interest precisely because it is not purely physical. It is a physical realization of a symbolic reality.”
Ubiquitous computing requires that we provide nontraditional levers for interaction

What ought those look like? How ought they to work? And why?
Today: two big ideas

1. Input sensing architectures
2. Output approaches, and their tradeoffs
   Displays, vs.
   Augmented Reality, vs.
   Virtual Reality
...And one course theme

As an interdisciplinary field, HCI bridges perspectives and methods from multiple fields.

Today, we will see one example of how an idea iterates through and across these perspectives.
Input and sensing
Bolt. “Put-that-there”: voice and gesture at the graphics interface. SIGGRAPH '80.
Put That There

Contribution: combined gesture and voice input

In a closed world

With a toy goal

Using simple manipulation operations

Using a laser attached to the wrist

In many ways, our goal since 1980 has been to relax those assumptions
DigitalDesk

Contribution: fluid boundaries between digital and physical objects

- In a constrained space
- On a small set of tasks
- With predefined behaviors

Again, we work to relax these assumptions
Today, suppose you wanted to...
On-body input

[Harrison et al. 2010]

Appropriate our own skin as a widely-available input surface
Environmental audio input

[Laput et al. 2018]

Detect activities in the local environment and adapt

Evolved into Apple’s handwashing feature:
Fine-grained gesture input

[Li et al. 2016]

Provide expressive inputs beyond the limits of the display screen
Fine-grained gesture input

[Provided by Lien et al. 2016]

Provide expressive inputs beyond the limits of the display screen.
Fine-grained gesture input

[Lien et al. 2016]

Now integrated into a number of Google products
Muscle input
[Saponas et al. 2009]

Recognize hand gestures with on-body instrumentation
...and be an air guitar hero
Or many others...

Hand gesture detection with an instrumented glove [Glauser 2019]

Activity detection on the phone or watch [Consolvo et al. 2008]

Detect body pose or posture without instrumentation or cameras [Cohn et al 2012]

Recognize exercises in a gym [Khurana et al. 2018]
How would you do it?
Step 1: Sensor input  [Harrison et al. 2010]

Example: a series of highly tuned vibration sensors, each tuned to different resonant frequencies
Step 1: Sensor input [Laput et al. 2018]

Example: standard laptop, phone, or watch microphone
Step 1: Sensor input

Example: radar (radio wave reflections) scatters a wide beam and captures responses from many different parts of your hand as the waves reflect off of it.

[Lien et al. 2016]
Step 1: Sensor input

Example: attach EMG sensors in a band around your arm: forearm electromyography
Step 1: Sensor input

Audio
Video
Accelerometers
Vibration sensors
EM waves of many flavors (from radio to infrared to wifi)
Environmental EM waves
Anything else you can think of?
Step 2: Featurization

Transform the sensor input into a form that is maximally informative for a machine learning algorithm. The exact transformation depends on the sensor and application.

Example: audio feature engineering on recorded sound clips: Fourier transforms fed into spectrograms

[Laput et al. 2018]
Step 2: Featurization

Transform the sensor input into a form that is maximally informative for a machine learning algorithm. The exact transformation depends on the sensor and application.

Example: treat vibration data as audio data and derive similar features (FFTs, average amplitude of each sensor, amplitude ratios between pairs of sensors)

[Harrison et al. 2010]
Step 2: Featurization

Transform the sensor input into a form that is maximally informative for a machine learning algorithm. The exact transformation depends on the sensor and application.

Example: locations of RF scattering centers detected, temporal transformations of scatter center motion

[Lien et al. 2016]
Step 2: Featurization

Transform the sensor input into a form that is maximally informative for a machine learning algorithm. The exact transformation depends on the sensor and application.

Example: split the EMG signal into short segments and treat each one as a sample, calculate root mean square (RMS) ratios between channels, frequency energy, phase coherence ratios between channels

[Saponas et al. 2009]
Step 3: Train a classifier

Option 1: Impress your friends and maximize performance, use a deep learning architecture

"the bidirectional LSTM classifier to produce the best accuracy owing to the superiority of Recurrent Neural Networks (RNNs) like LSTMs and Bi-LSTMs in modeling long-range temporal dependencies and. Moreover, since"

[Bhalla, Goel, and Khurana 2021]
Step 3: Train a classifier

Option 2: Operate under power or speed constraints—like you’re launching a mobile product—and use something simple that works.

There are a number of powerful classification algorithms that can be used for temporal gesture recognition, such as

[...]

These algorithms are computationally expensive and are not suited for real time operation on low-power embedded platforms at high frame rates and small memory footprint. By benchmarking and comparing various algorithms we converged on a Random Forest classifier.

[Lien et al. 2016]
Architecture overview

(1) Collect raw sensor data
(2) Featurize the raw sensor data
(3) Train a classifier

Example from EM-Sense: [Laput et al. 2015]
Challenge: Midas touch

Unlike traditional interfaces, sensing-based interfaces do not ask for explicit intent.

As a result, they can trigger as false positives: activate when you don’t want them.

(Consider the woes of your friends who are named Alexa)

This is known as the Midas touch problem.

What to do?
Output
Displays: a limiting factor for ubicomp

How can a ubiquitous computing (ubicomp) system communicate back to people?

What if the person is nowhere near a display, and the information is situated out in the world?
Three approaches

There is no single solution here. Instead, we have developed a series of different explorations and technologies.

(1) Exotic displays
(2) Augmented reality
(3) Virtual reality
Exotic displays

Goal: develop ubiquitous computing technologies that extend the availability and affordances of existing displays
On-body displays [Mueller et al. 2020]

Worn displays [Lo et al. 2016]

Embedded in textiles [Daguin 2021]

Thin, electro-tactile feedback (35µm thick) [Withana et al. 2018]
Embedded in objects and the environment

Example: 3D printed display components

[Willis et al. 2012]
Embedded in objects and the environment

Example: mobile swarm robotics

[Le Goc et al. 2016]
Embedded in objects and the environment

Example: drones
[Yamada et al. 2017]
Embedded in objects and the environment

Example: drones

[Yamada et al. 2017]

(Though a remaining issue...)
(2) Augmented reality

AR: lay virtual information out into the physical world

These technologies often must sense the layout of the space around them, then project the digital information into that space
Projector-based AR

Mount projectors into the space to add digital augmentation

Shoulder-mounted projector turns surfaces into interfaces

[Harrison, Benko, and Wilson 2011]
Projector-based AR

Mount projectors into the space to add digital augmentation

[Wilson et al. 2012]
Head-mounted displays

(Technically, much use of the HoloLens is mixed reality, where the physical world and the digital world can interact with each other.)
Head-mounted displays

Enabling digital proxies of remote objects or participants

[Orts-Escolano et al. 2016]
(3) Virtual reality

VR: Head-worn display that occludes the surrounding world and instead embeds you in a fully digital world

“Such a display could literally be the Wonderland into which Alice walked”

[Sutherland 1965]
Psychological impacts of VR

High levels of embodiment mean that our virtual self-representations impact our behaviors [Yee and Bailensen 2007]
Tradeoffs

Exotic displays: require specialized hardware and devices, so difficult to generalize and power, but very flexible design space

AR: can extend the local environment with new behaviors and objects, but seams can show between the physical and digital

VR: fully immersive, but completely removes you from the physical environment, and the illusion breaks when you run into your couch

There is no perfect approach—select the modality that matches your needs.
Interdisciplinary perspectives in HCI
So far, we’ve mostly discussed *engineering*-style contributions to HCI. These involve envisioning and creating new technological approaches to human-computer interaction.

But, as we proceed through the course, we will encounter contributions that come from several other perspectives, including:

- Psychology & Social sciences
- Design
- Theory
Before today: “HCI is design process-iterated product”

After today

- An algorithm paper can be HCI
- A design paper can be HCI
- A qualitative paper can be HCI
- A critical theory paper can be HCI
- An EE/ME paper can be HCI
- A field experiment can be HCI

...
HCI interdisciplinarity

HCI succeeds by bringing together these perspectives.

Take each paper you read on its own disciplinary terms: e.g., what is expected of an evaluation of an engineering contribution is different than what is expected of a study in a social science contribution.

You will likely encounter ideas pursued from several of these perspectives simultaneously across papers. Let’s take an example in today’s context…
Can’t Touch This [Hammer 1990]

A perennial problem in AR and VR is haptic feedback: the digital environment looks realistic, but it has no physical substance, so you cannot actually touch it.

A traditional approach is to create active or passive haptic feedback:
Can’t Touch This

[Hammer 1990]

A perennial problem in AR and VR is haptic feedback: the digital environment looks realistic, but it has no physical substance, so you cannot actually touch it.

A traditional approach is to create active or passive haptic feedback:

[Choi et al. 2017]

[Tao et al. 21]

[Massie and Salisbury 1994]
Haptic illusions

Could we instead convince our tactile systems that we are feeling things that aren’t exactly physical reality?

Would this allow us to produce a wider variety of haptic sensations even with restricted hardware?
The Visual Dominance Effect [Rock and Victor 1964]

When touch conflicts with vision, our brains resolve the conflict in favor of vision.

Experiment: place a shape on a table, and ask participants to view it and feel it from behind plastic that distorts how the shape looks. Then, ask people to draw what they think the shape actually is.

Result: people drew distorted shapes that matched what they saw and not what they felt, often without being aware of the conflict at all.
Haptic retargeting
[Azmandian et al. 2016]

By warping the rendered version of your body or the world in VR, you pick up the same cube three times while thinking it’s three different cubes.
Crafting the Impossible

[Abtahi et al. 2022]

VR embodiments need not have any real-world equivalents

Scale up the size of the user’s arm

Scale up the user’s avatar when moving long distances

Argument: rather than making reality-based VR, aim to create beyond-real interactions in order to improve the experience
Or, just fake it. [Cheng et al. 2014]

Use people instead?
Summary

Input sensing architectures: sensors, features, ML
Output approaches, and their tradeoffs
  Displays, vs.
  Augmented Reality, vs.
  Virtual Reality

HCI interdisciplinarity: ideas are pursued from multiple perspectives
A reminder on commentaries

**Do:** engage with the core contributions —

**Step 1:** What is the point that this paper is trying to make?

**Step 2:** How effectively does it convince you of that argument? How could the argument be even more persuasive, on its own terms?

**Step 3:** What are the implications of the argument? **What future frontier projects might be inspired by this work?** What follow-up project would you work on?

**Don’t:** nitpick low-level details, harp on already-acknowledged limitations / future work, bring expectations from other HCI paper genres (“needs a user study!”), spend too much time summarizing, levy judgment (“I like this!”) without digging into why or implications
Refining the 3-steps

Do: engage with the core contributions —

**Step 1 (Reflection):** State the main point but then *reflect on why* the ideas in the reading made sense from the authors' perspectives.

**Step 2 (Synthesis):** How does the idea relate to your experiences today. How effectively does it convince you of that argument? *How could the argument be even more persuasive, on its own terms?*

**Step 3 (Future work):** What are the implications of the argument? Given the ideas presented in the paper, what would you want to work on, or how would you modify those ideas?


Daguin, Clara. Ambient Experiments. 2021


Schwarz, Julia. HoloLens 2 Demo. https://www.youtube.com/watch?v=ulHPPtPBgHk&t=1s


References


