

Input / Output

CS 347

Michael Bernstein

Announcements

You (hopefully) submitted your first commentaries

Due at noon before lecture

If you joined late and didn't do the readings—it will decrement your first week tutorial grade, but we drop the lowest tutorial grade, so don't fret

Tutorials start this week: watch your email for assignment

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?

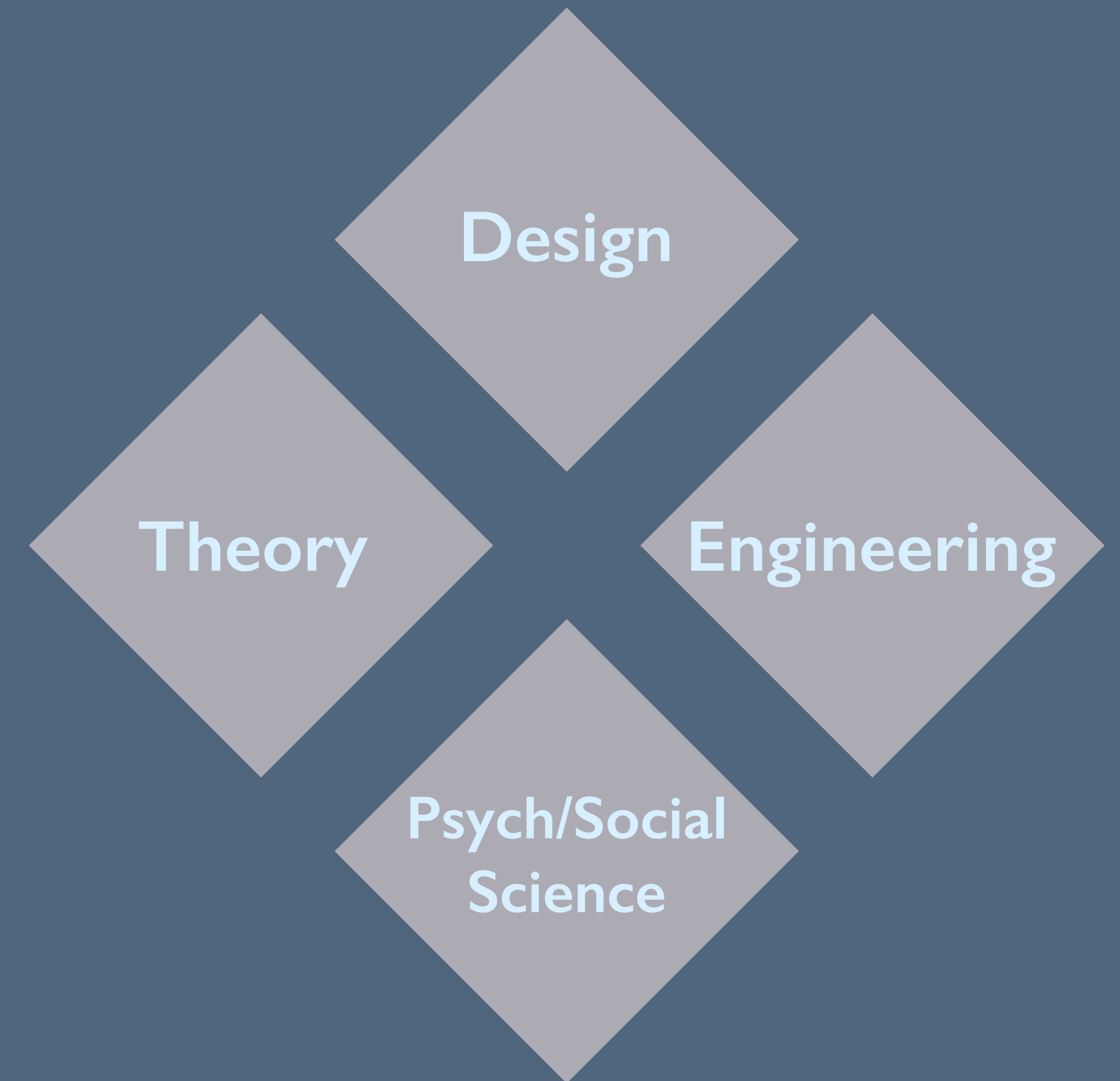
Today: two big ideas

1. **Input** sensing architectures
2. **Output** approaches, and their tradeoffs
 - Nontraditional 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

- Using highly constrained language to avoid ambiguity

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





Wellner. Interacting with paper on the DigitalDesk. CACM '93.

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

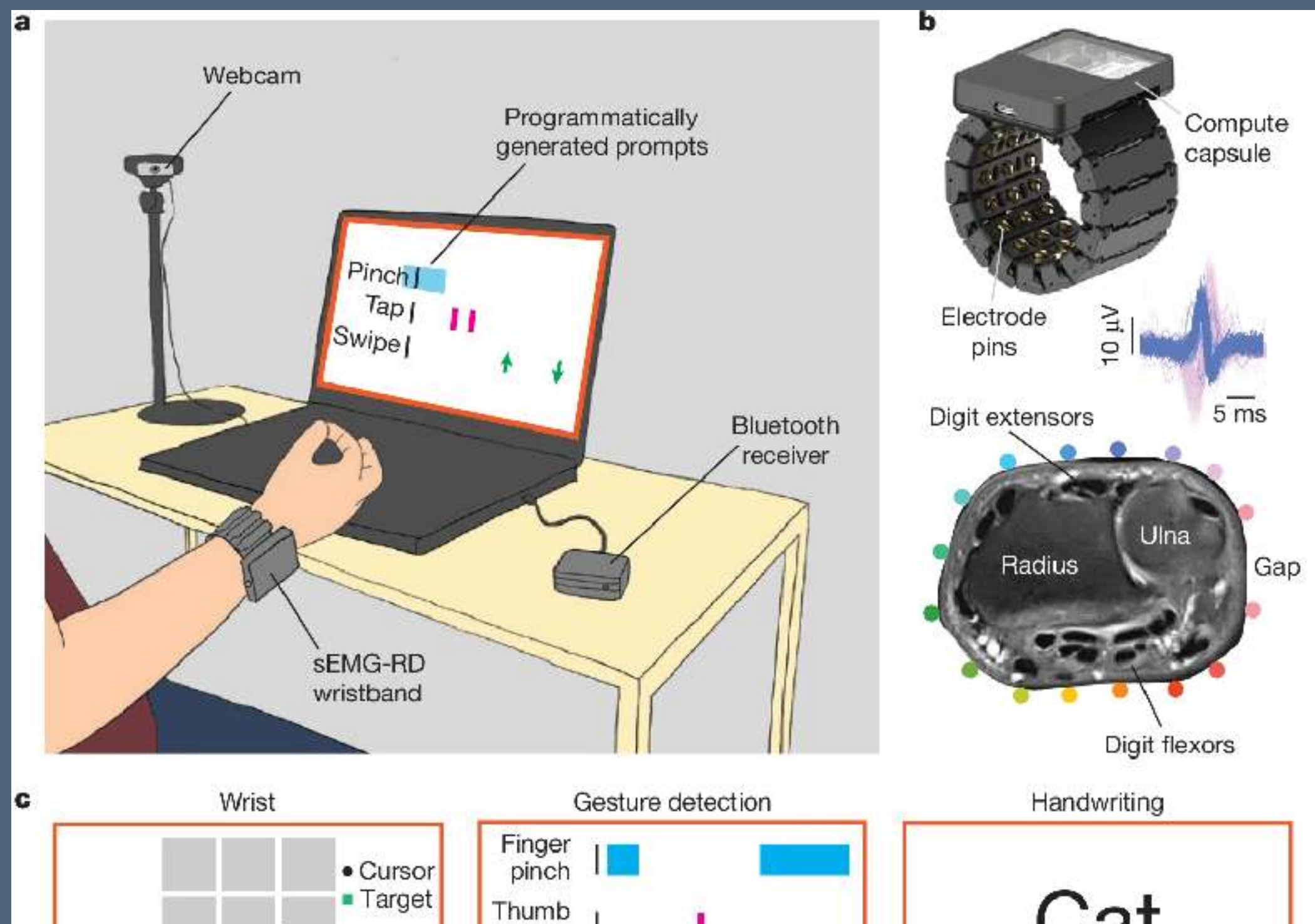
Today, suppose you wanted
to...

Muscle input

[Kaifosh et al. 2025]

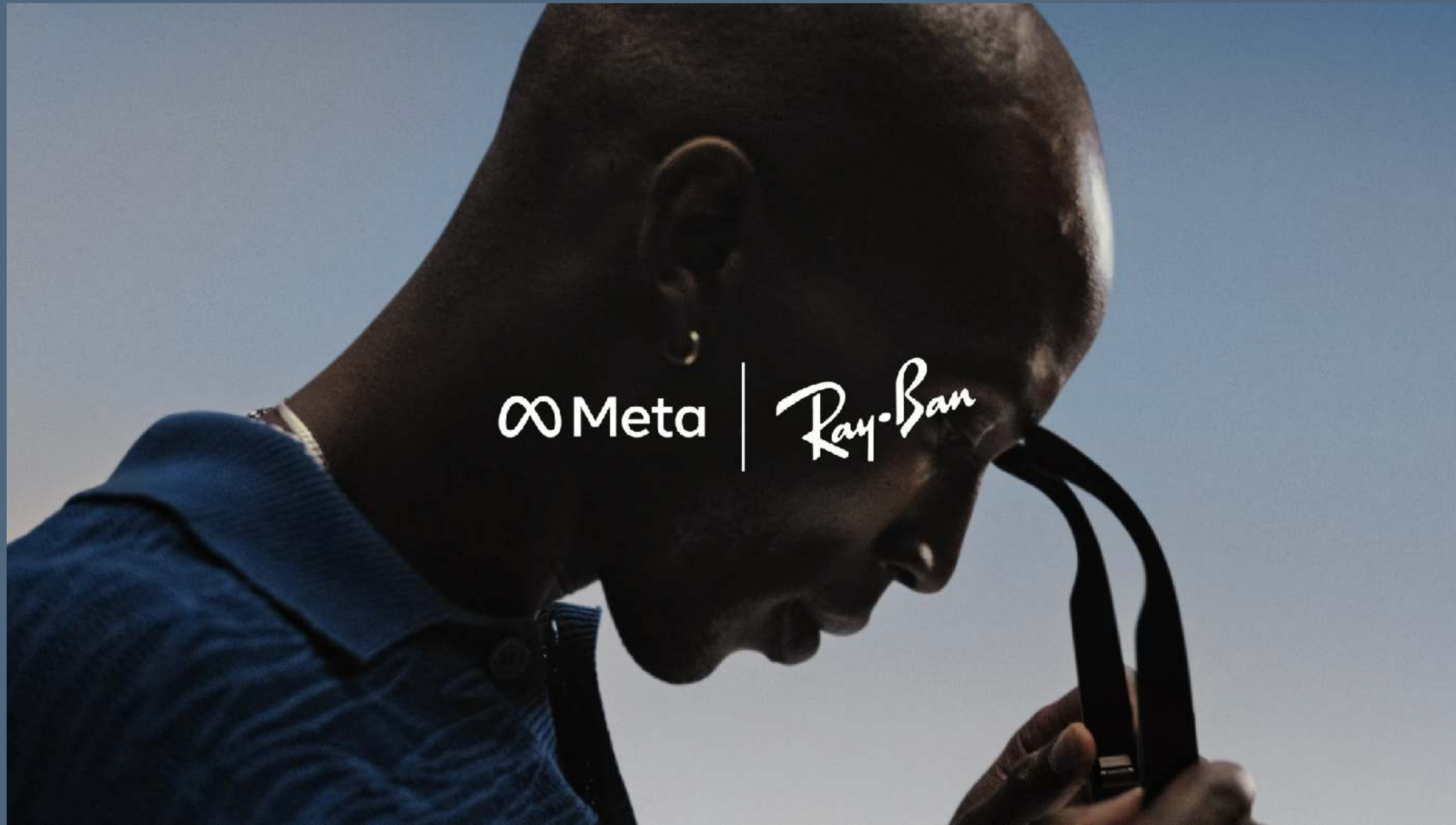
A generic non-invasive neuromotor interface for human-computer interaction.

Meta Neural Band.
(Same team productized)



Muscle input

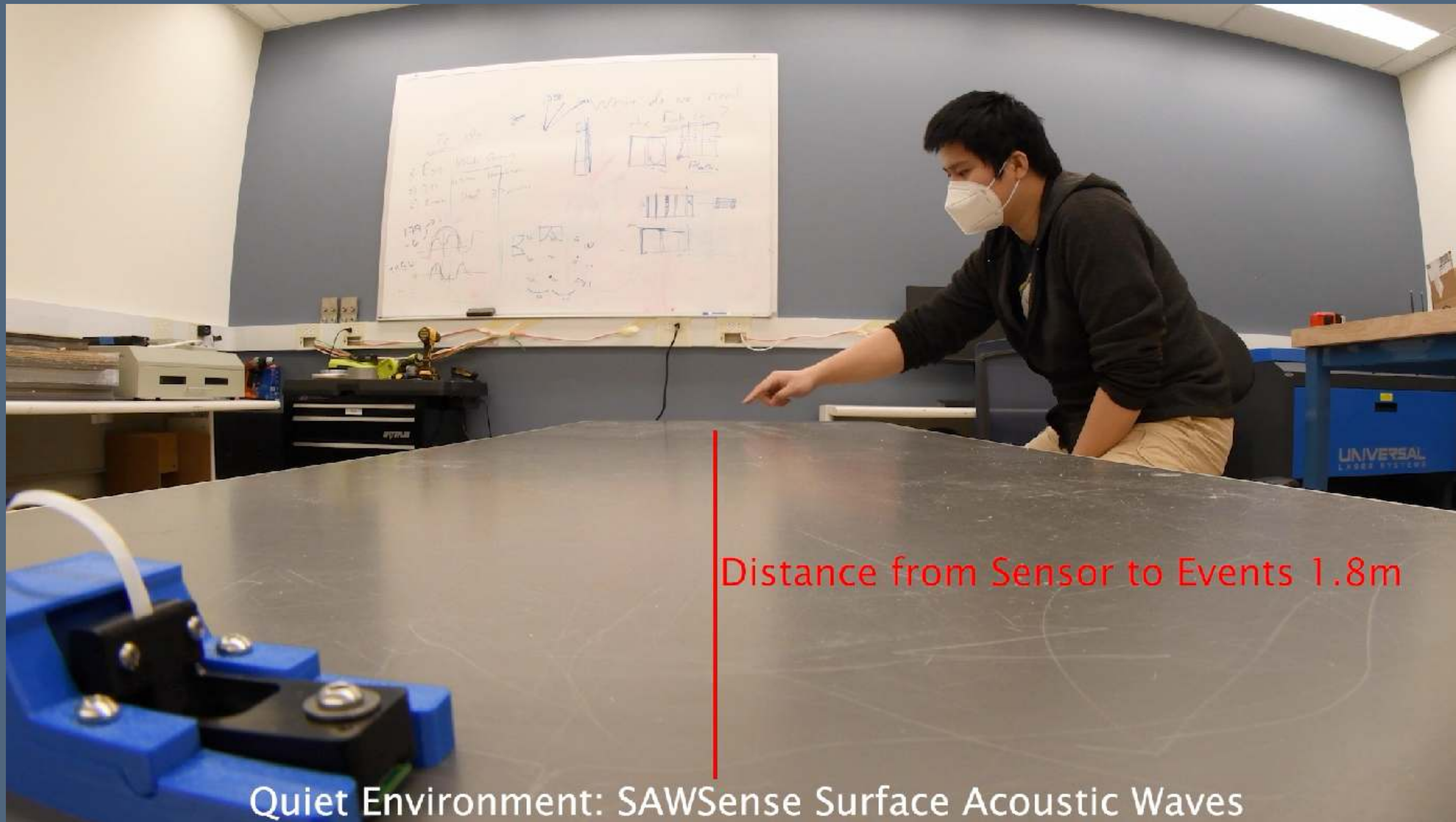
[Kaifosh et al. 2025.]



Gesture
recognition from
sensors on a
wrist-worn
device

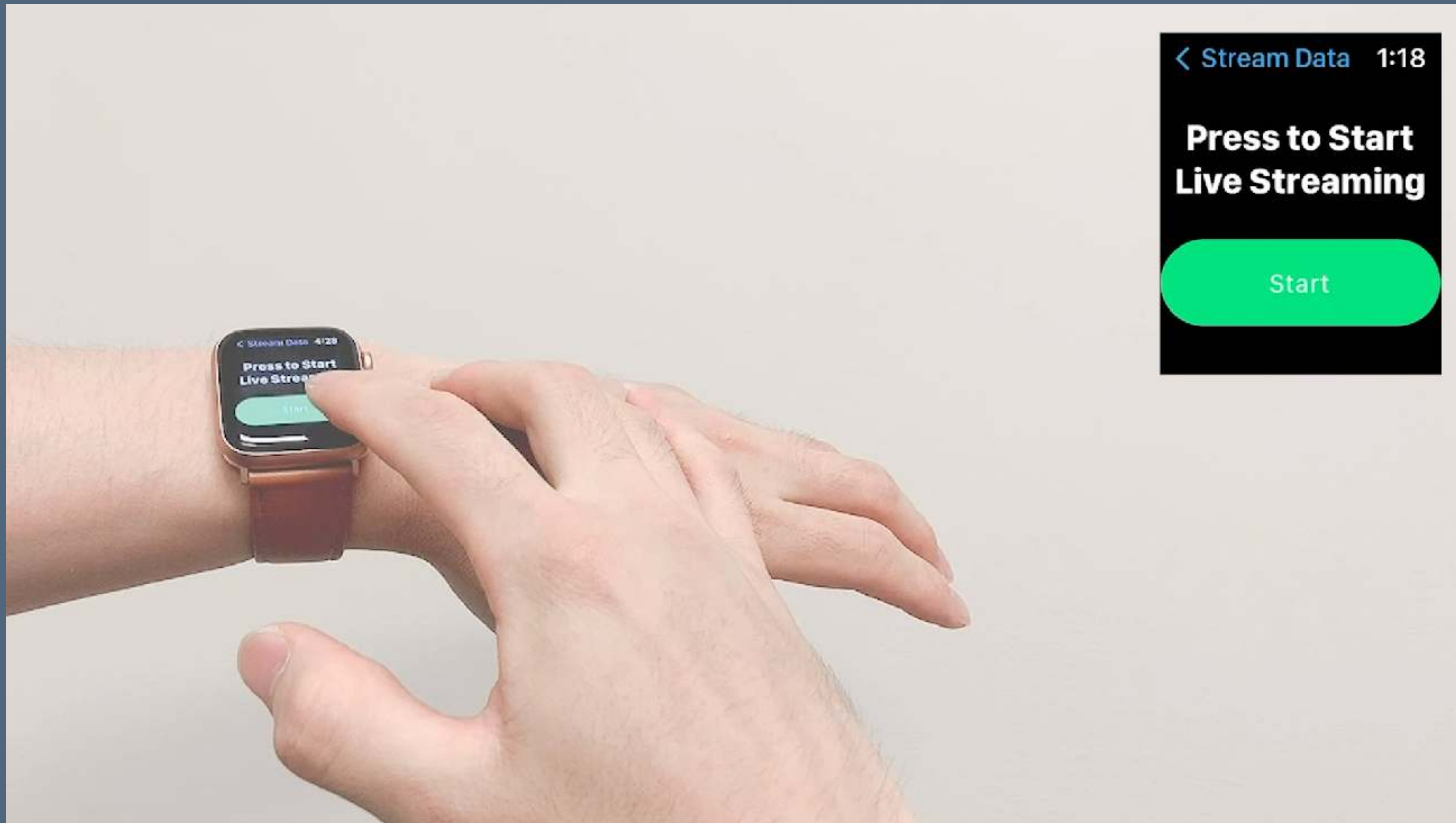
Environmental input [Iravantchi et al. 2023]

Detecting acoustic waves traveling along the surface of tables, furniture, 3D printed devices, fabric, etc.



Fine-grained gesture input

[Xu et al. 2022]



Provide
expressive
inputs
beyond the
limits of the
display screen

Or many others...

Hand gesture detection with an instrumented glove [Glauser 2019]

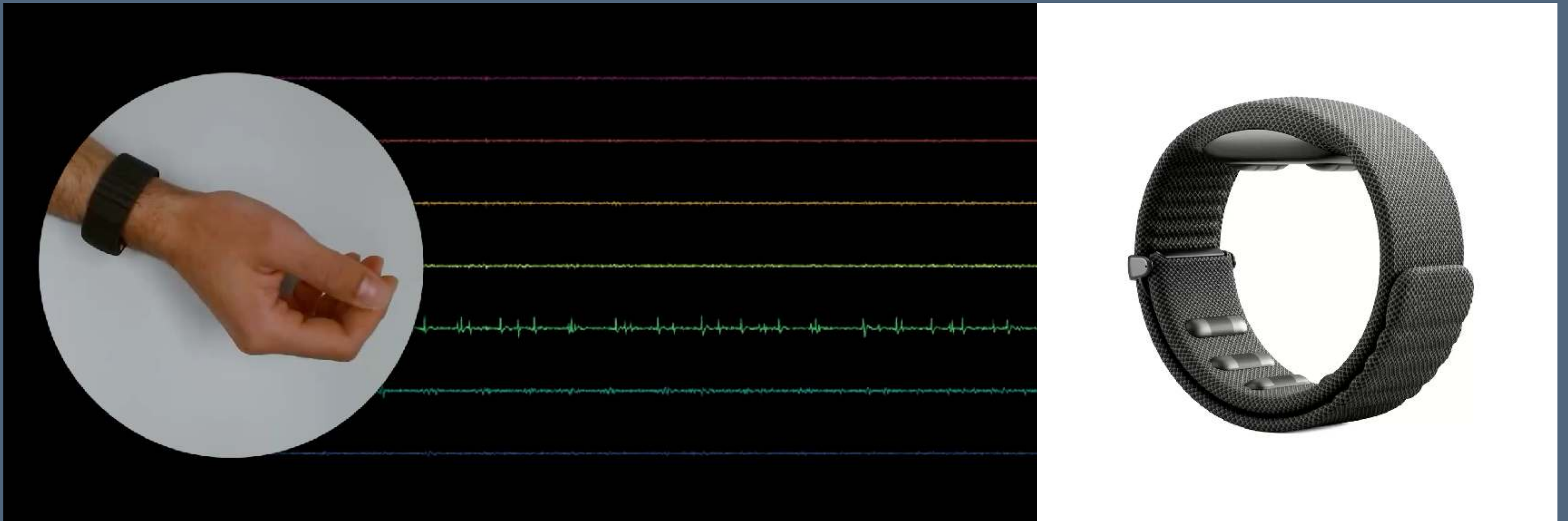
Activity detection on the phone or watch [Consolvo et al. 2008]

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

Recognize exercises in a gym [Khurana et al. 2018]

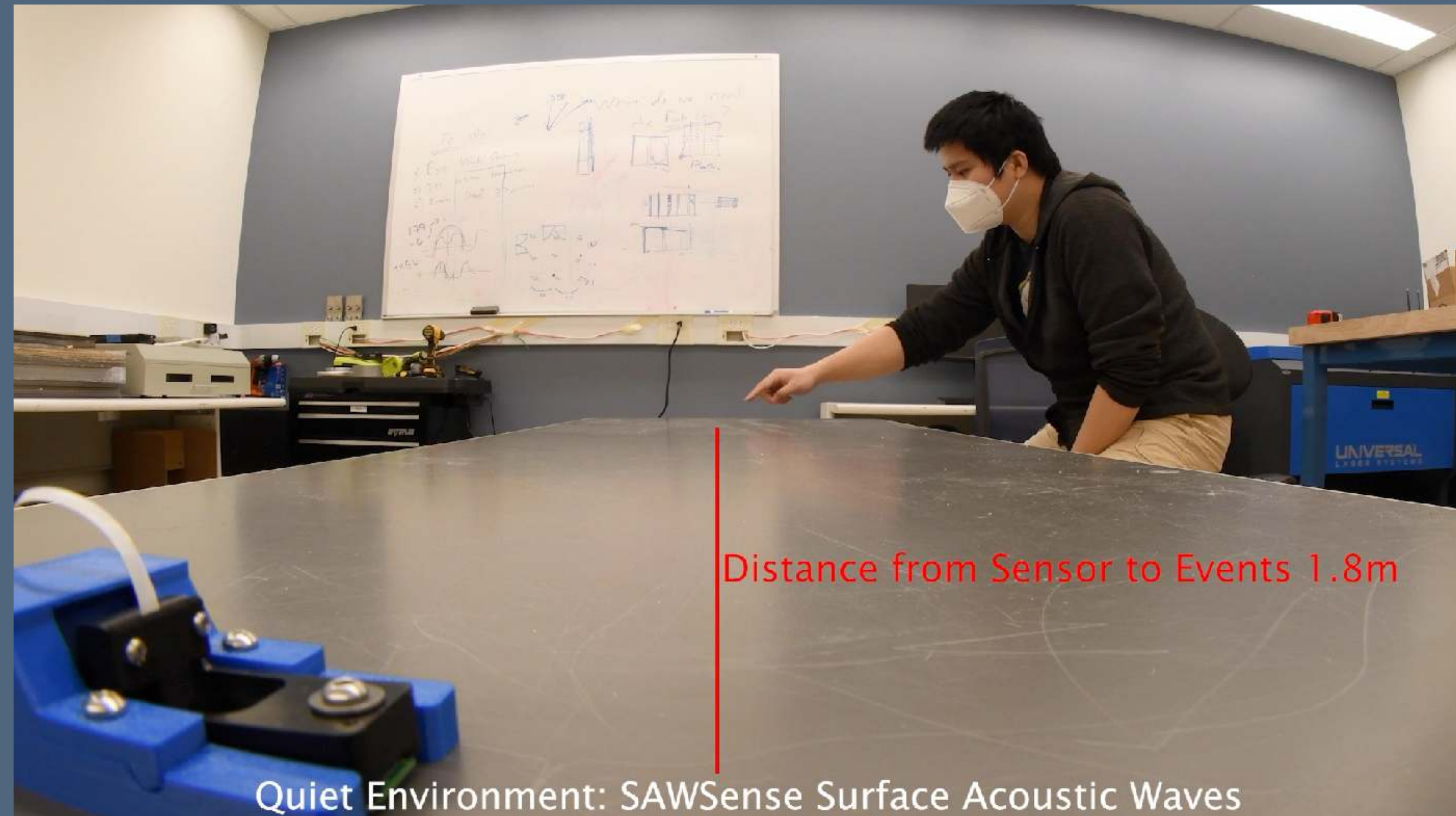
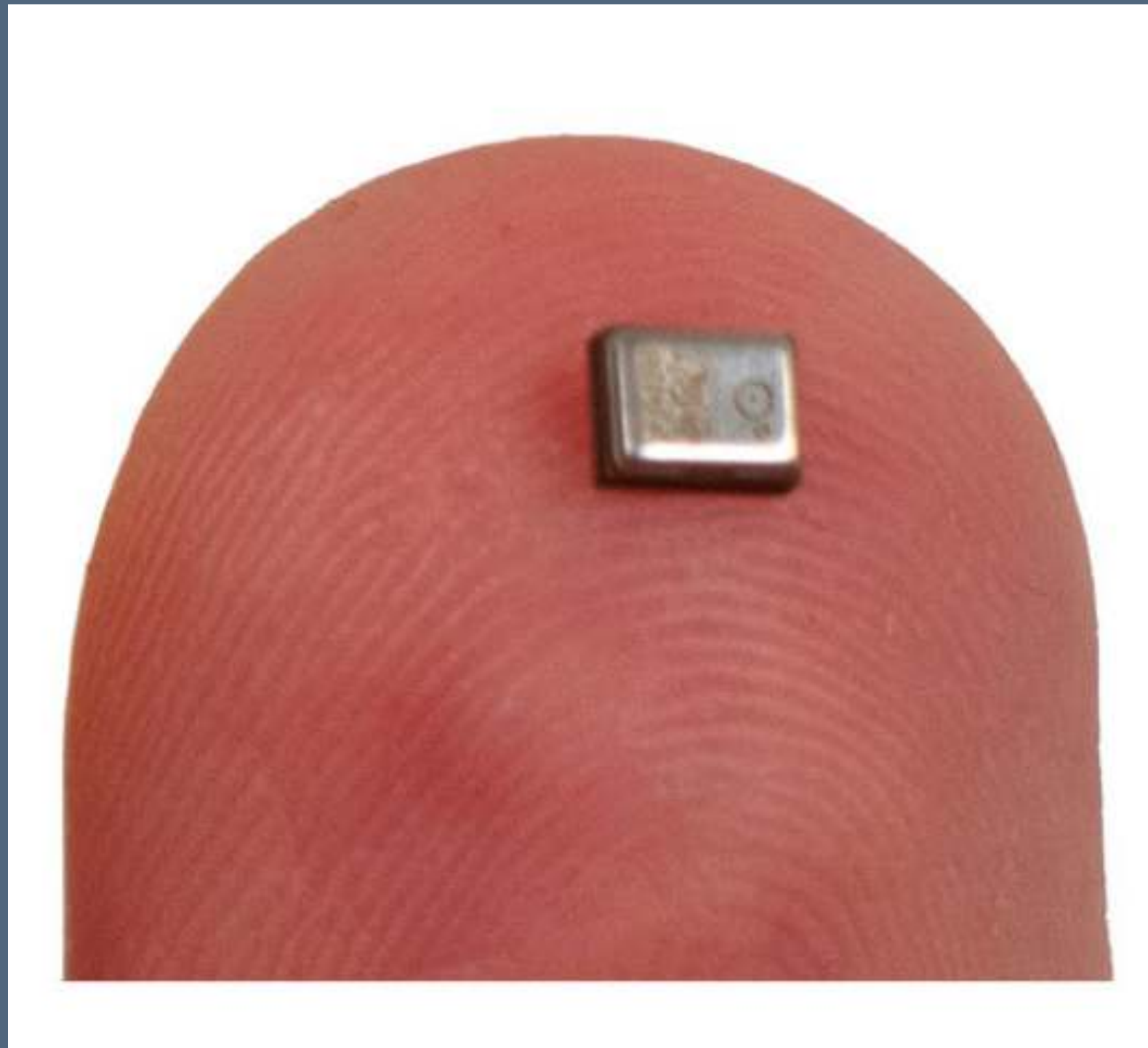
How would you do it?

Step 1: Sensor input [Kaifosh et al. 2025]



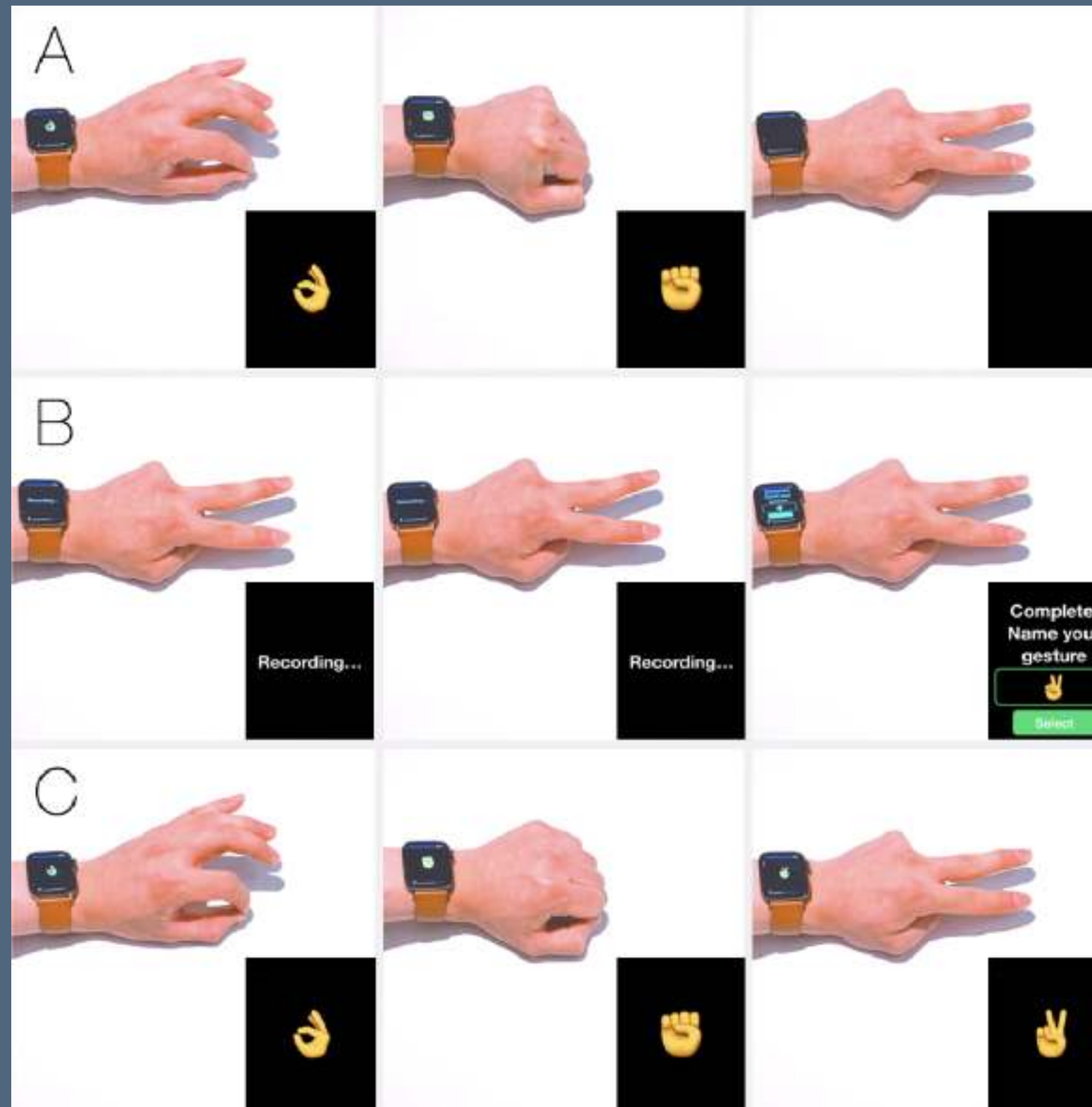
Read electrical signals from muscles on the wrist
This technology is called surface electromyography

Step I: Sensor input [Iravantchi et al. 2023]



Utilize the same sensor technology that earbuds use to hear you through bone conduction in a noisy room

Step 1: Sensor input [Xu et al. 2022]



Directly sample the accelerometer (how quickly is your speed changing?) and gyroscope (how fast are you rotating?) in watch

Step 1: Sensor input

Audio

Video

Accelerometers

Vibration sensors

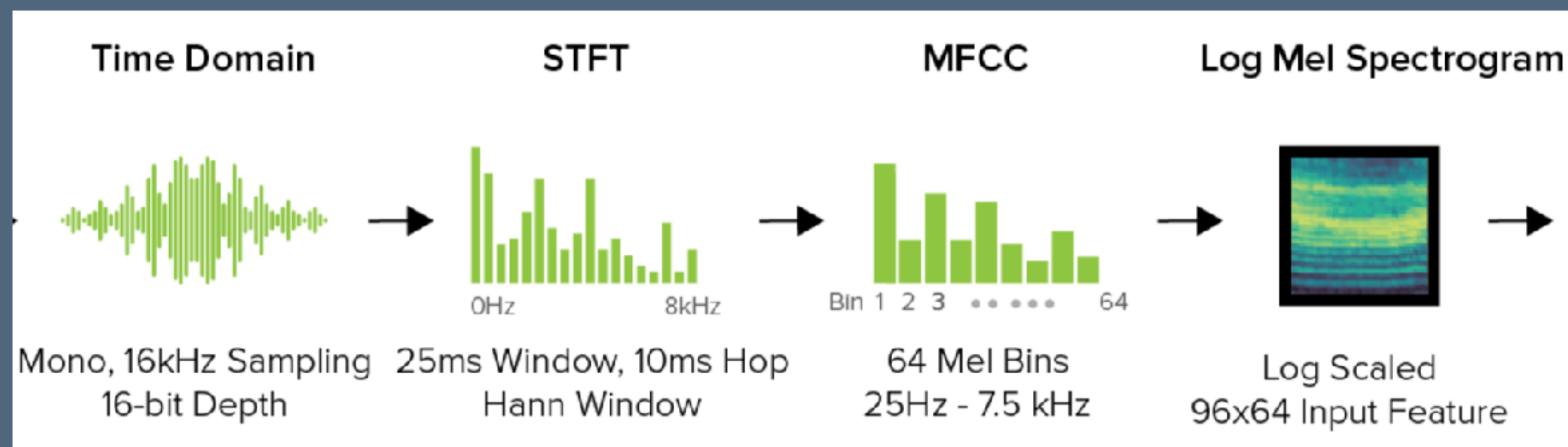
EM waves of many flavors (from radio to infrared to wifi)

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.

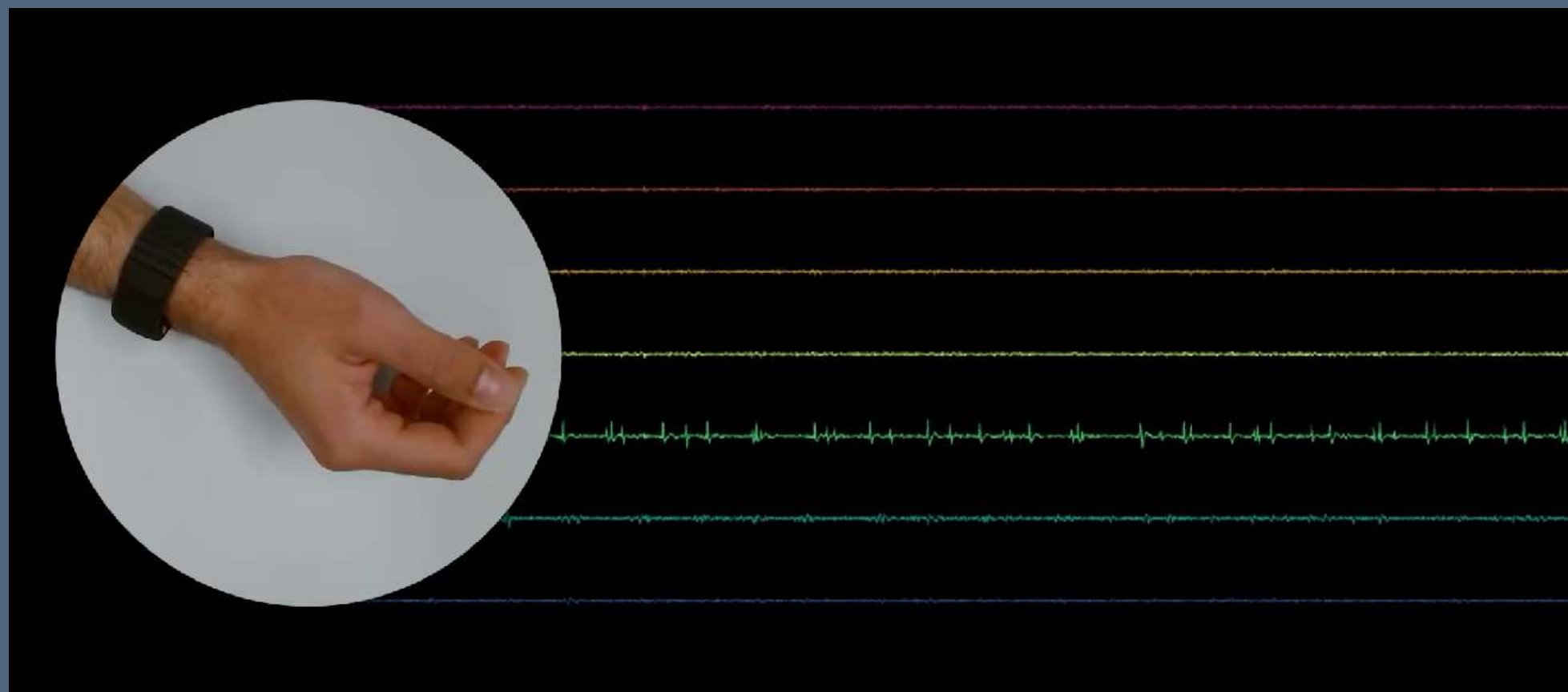


[Laput et al. 2018]

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

Step 2: Featurization

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

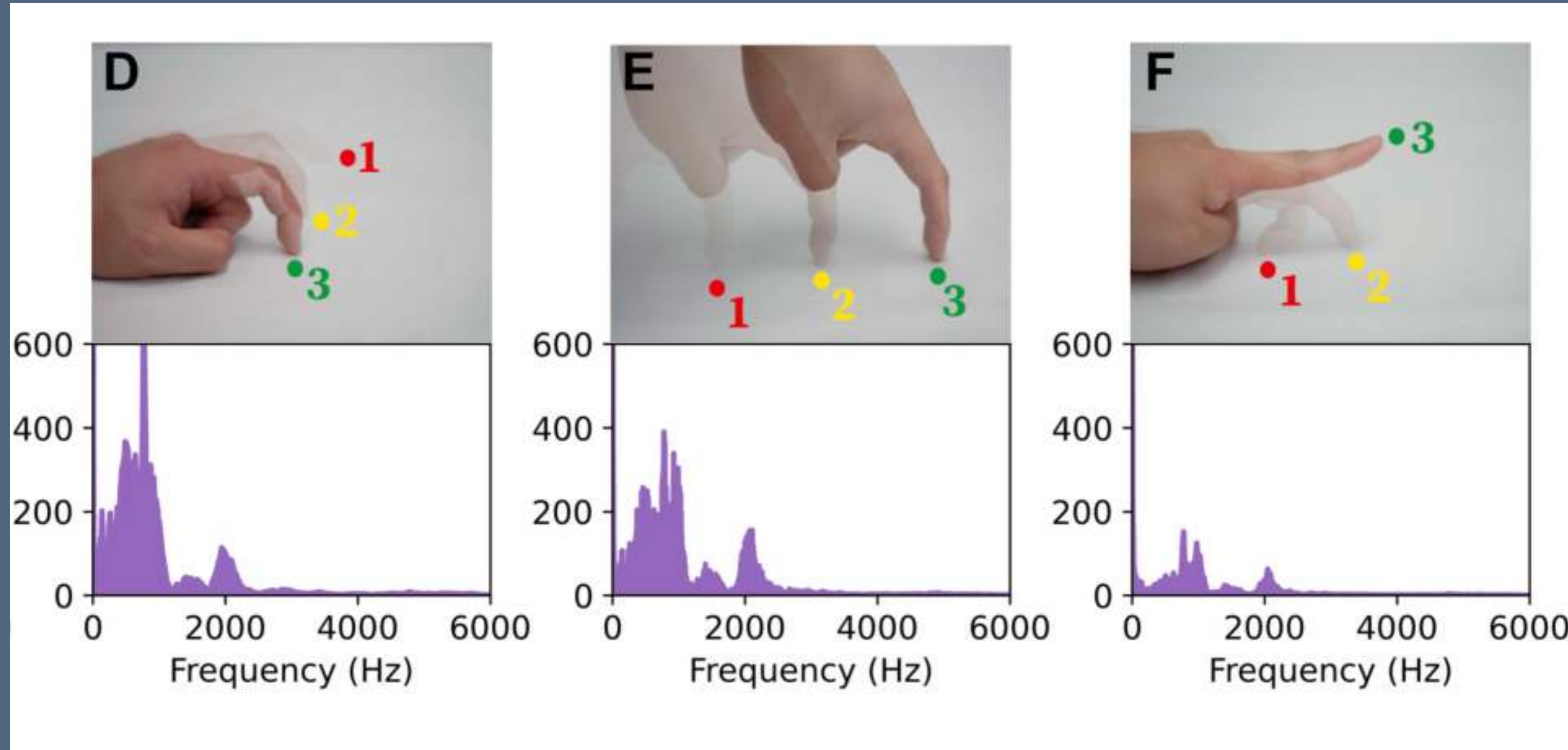


Example: derive multivariate power frequency (MPF) features, which are machine learning features often used in multi-dimensional time series data

[Kaifosh et al. 2025]

Step 2: Featurization

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

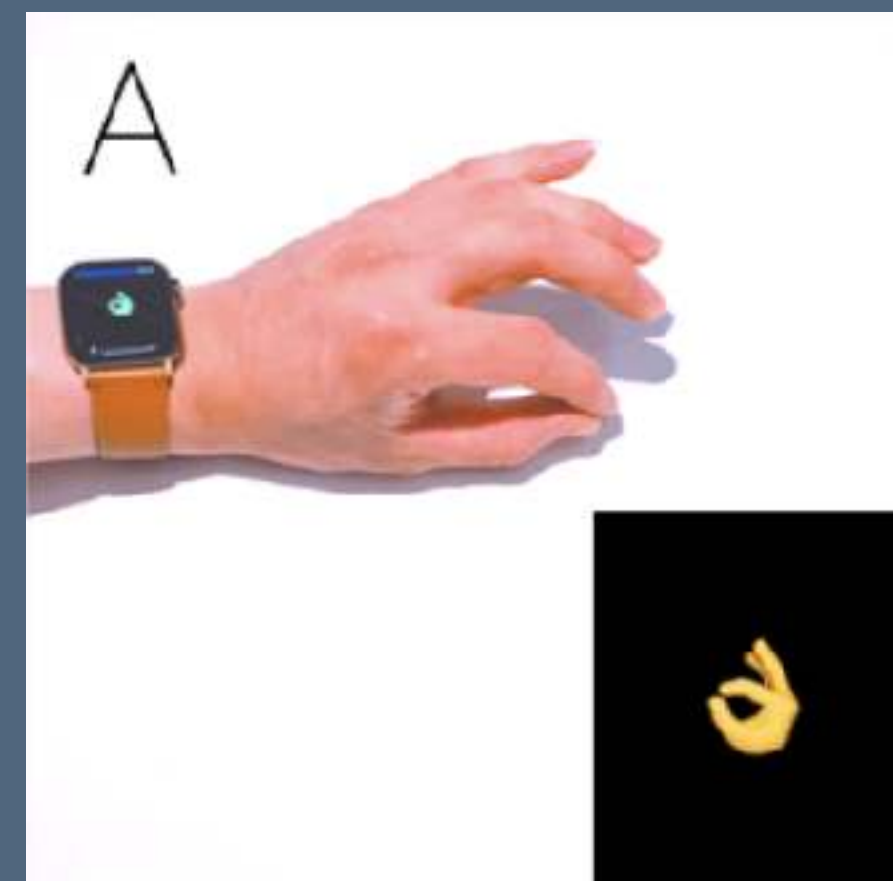
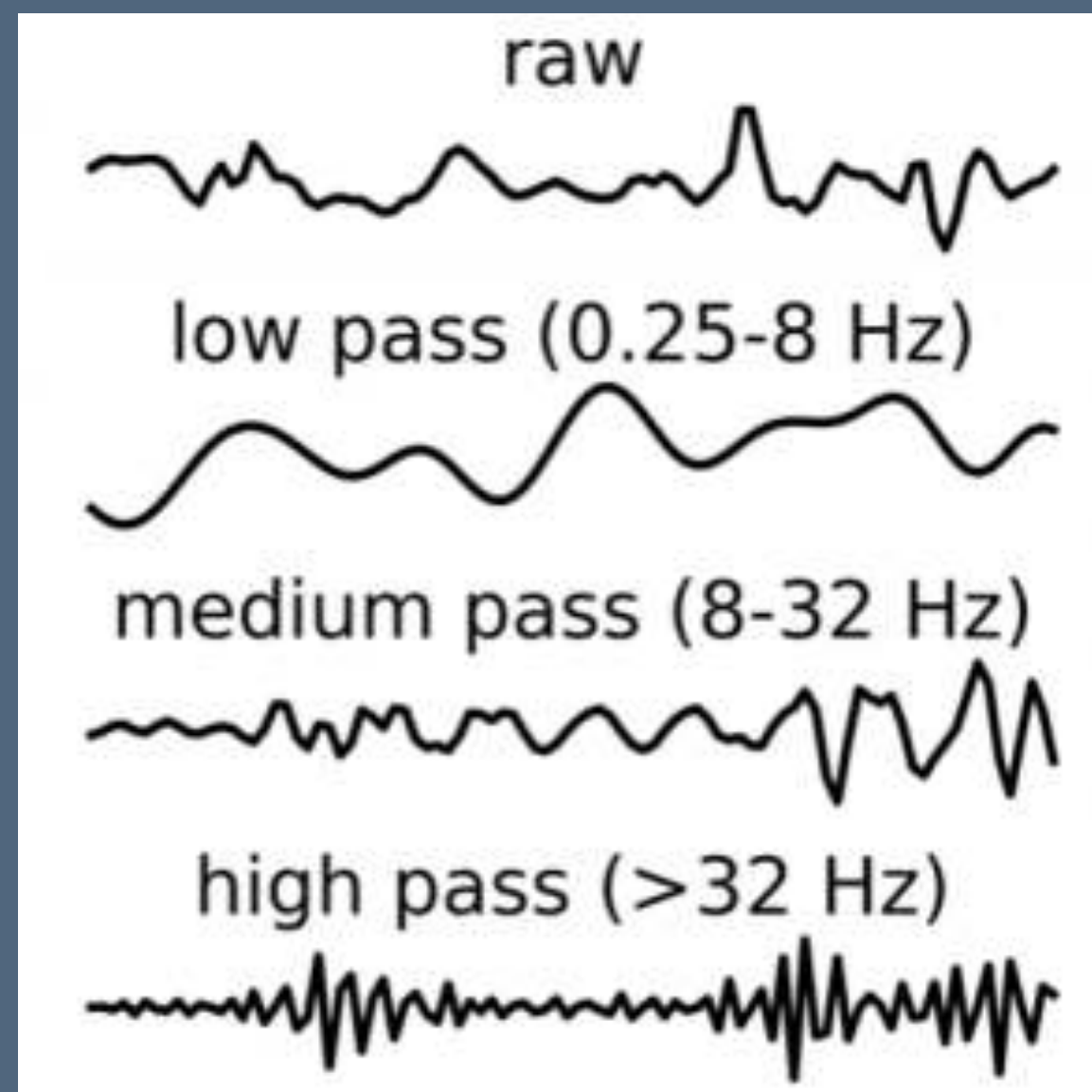


Example: FFT on sensor data, then Mel-Frequency Cepstral Coefficients (MFCC) to filter down to 128 features

[Iravantchi et al. 2023]

Step 2: Featurization

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



Example: take the raw accelerometer and gyroscope data, run low/medium/high bandpass filters on each one, and use those as efficient inputs

[Xu et al. 2022]

Step 3: Train a classifier

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

“ We adopt the concept of EfficientNet [79] to balance the number of trainable parameters and model performance. Specifically, for each input channel, we employed two inverted residual blocks [71] (*i.e.*, MBConv in Figure 2) to process the signals. We then concatenate the output of the six channels, and add one more separable convolution layer [15] to capture concatenated information with ”

[Xu et al. 2022]

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

“

for each gesture. We then explore the variety of SAWs generated by household appliances and cooking-related events and evaluate SAWSense's performance in activity recognition tasks in the kitchen. For both sets of evaluations, we present the results using a linear SVM, a Random Forest, and a 6-layer Multi-Layer Perceptron, representing three “tiers” of machine learning complexity. For consistency, all parameters except for the number of Random Forest estimators remain SciKit-Learn default. The number of estimators

”

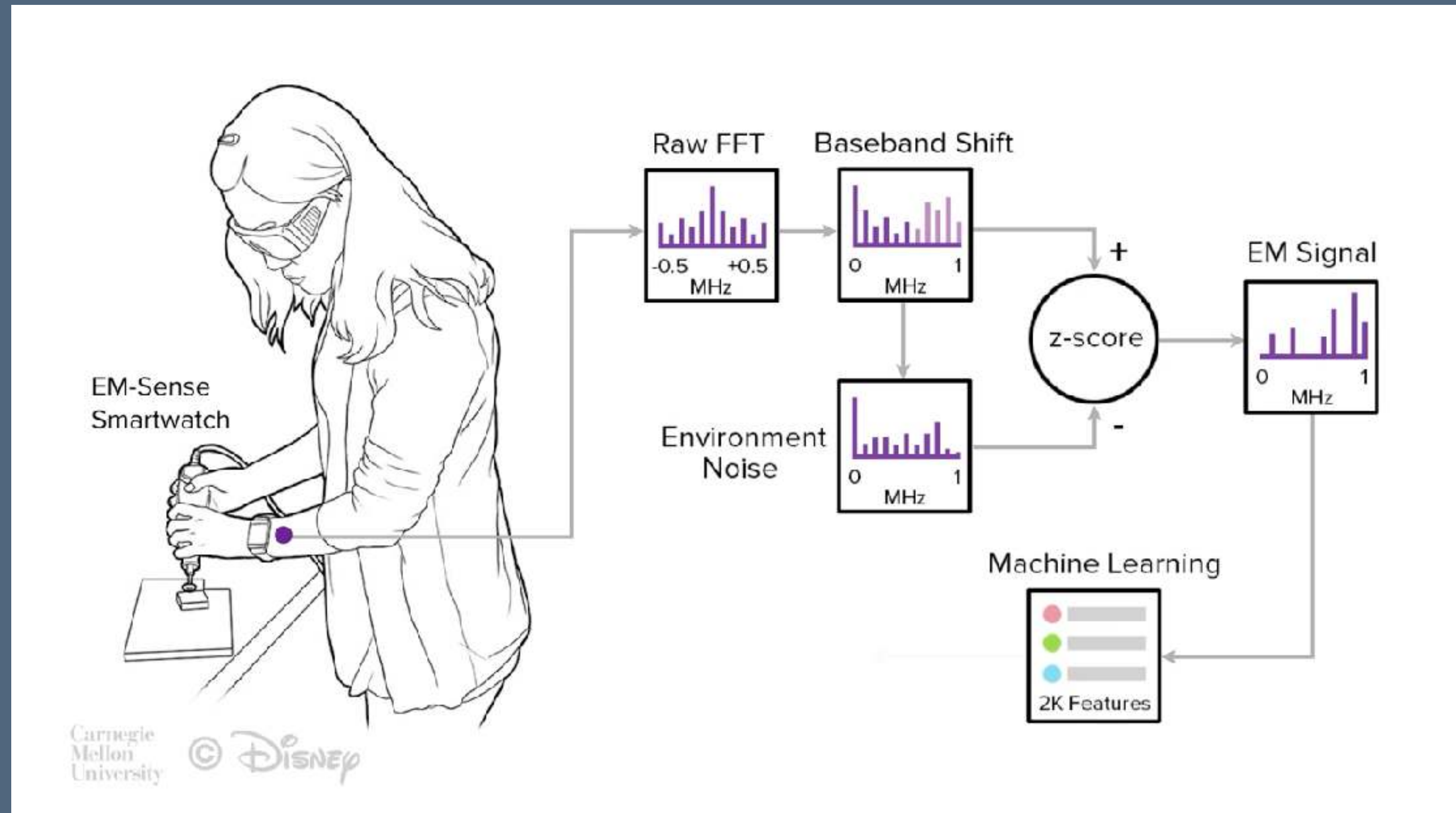
[Iravantchi et al. 2023]

Architecture overview

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

Features focus the learning algorithm on aspects of the data relevant to the task

Can be difficult to choose features



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

But aren't hand-tuned features losing favor?

End-to-end machine learning

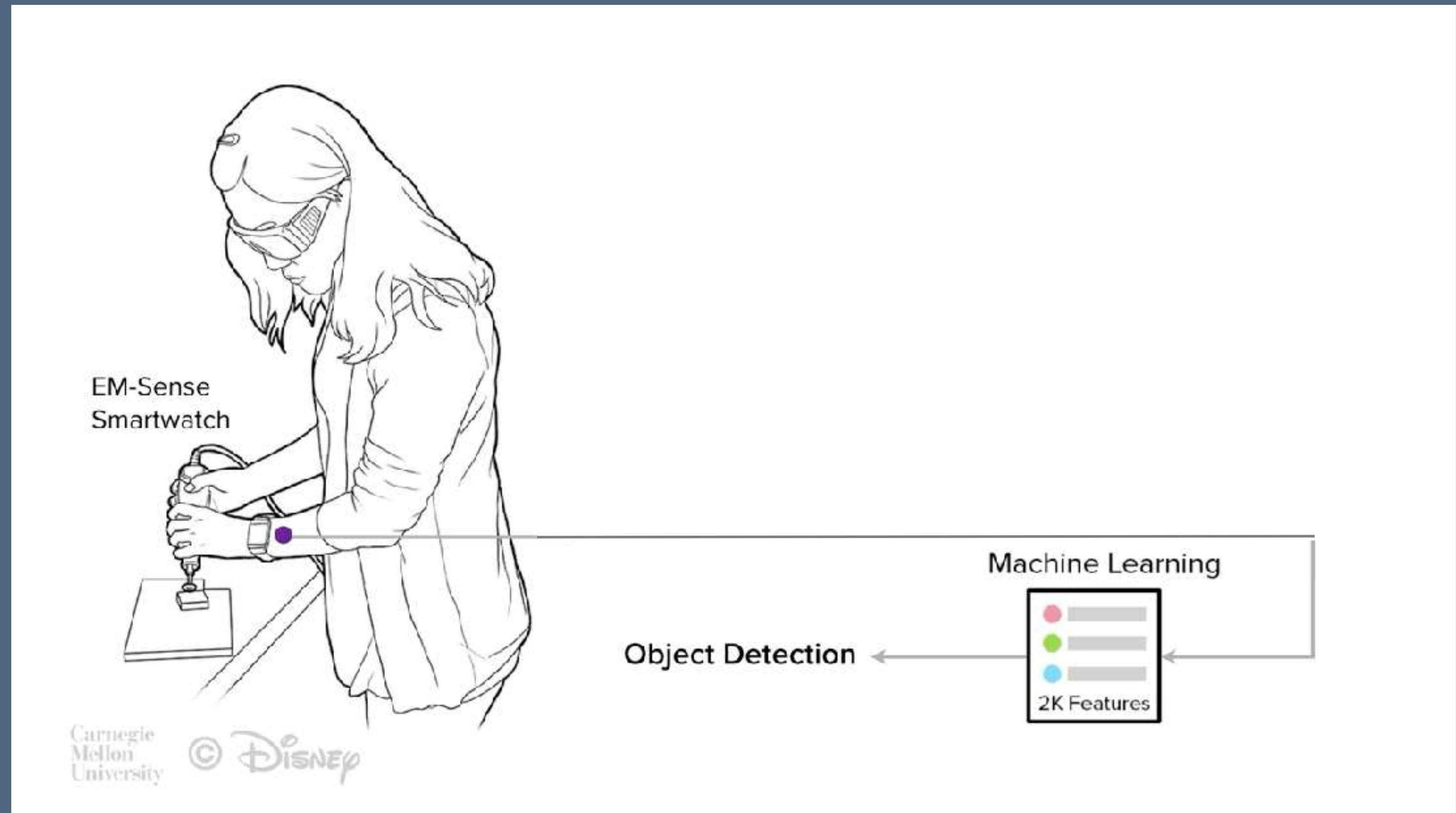
(1) Collect raw sensor data

~~(2) Featurize the raw sensor data~~

(3) Train a classifier

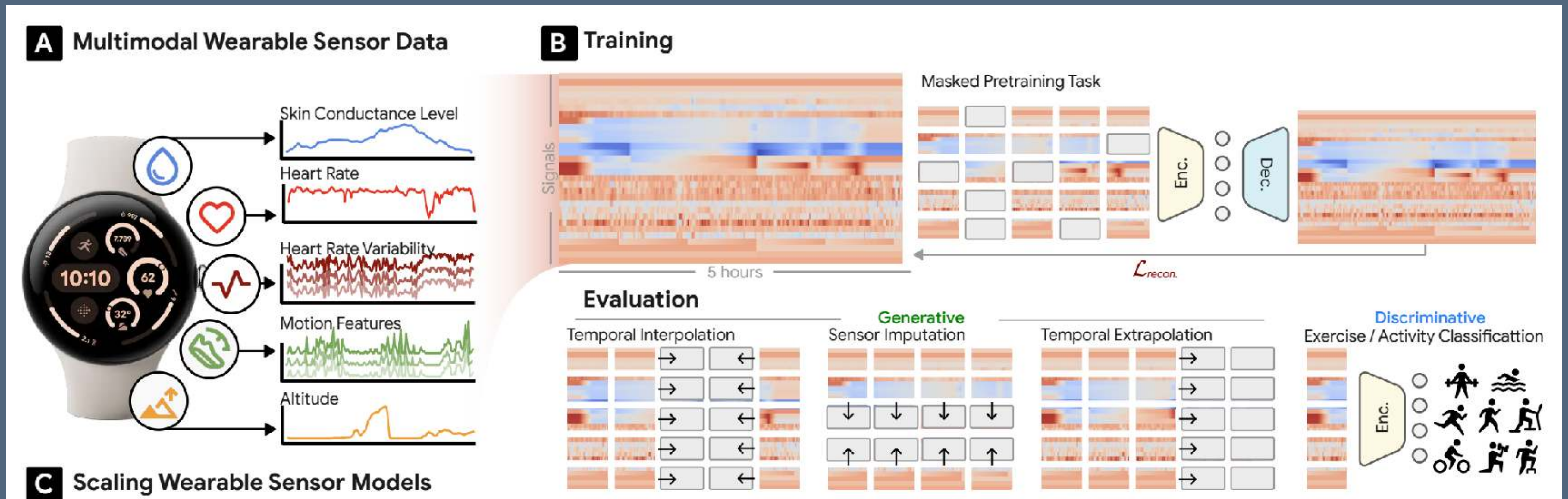
End-to-end is easier for developer; no need to design features

But can require far more data



Sensor foundation model

[Narayanswamy et al. 2025]



Trained on FitBit data from 40M hours of data from 165k people
Can now serve as a training backbone for many downstream tasks

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) Nontraditional displays
- (2) Augmented reality
- (3) Virtual reality

(I) Nontraditional displays

Goal: develop ubiquitous computing technologies that extend the availability and affordances of existing displays

Non-pixel display media



Example: hydraulic
fluid actuation

[Yu et al. 2025]

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]

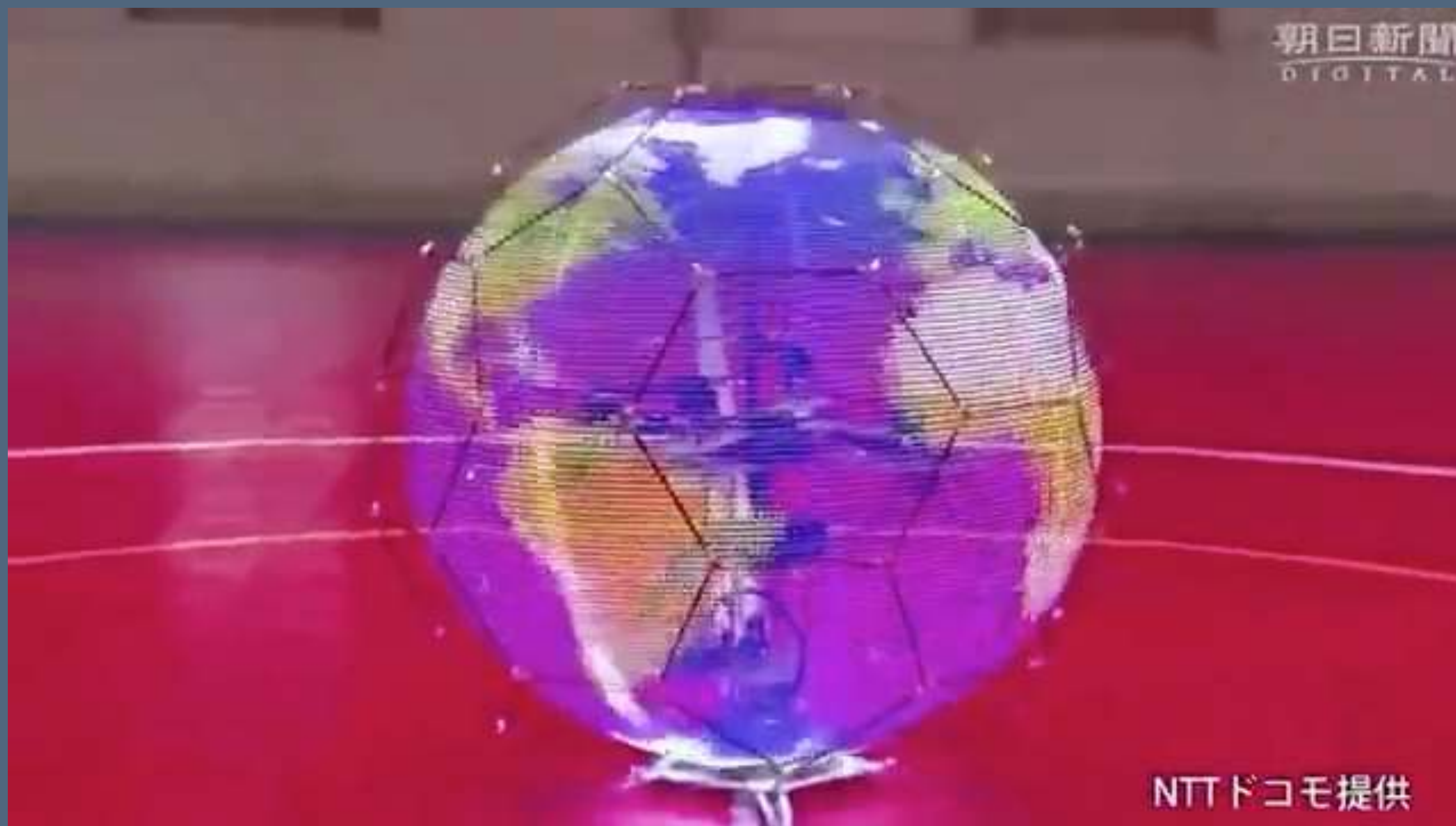
Embedded in objects and the environment



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

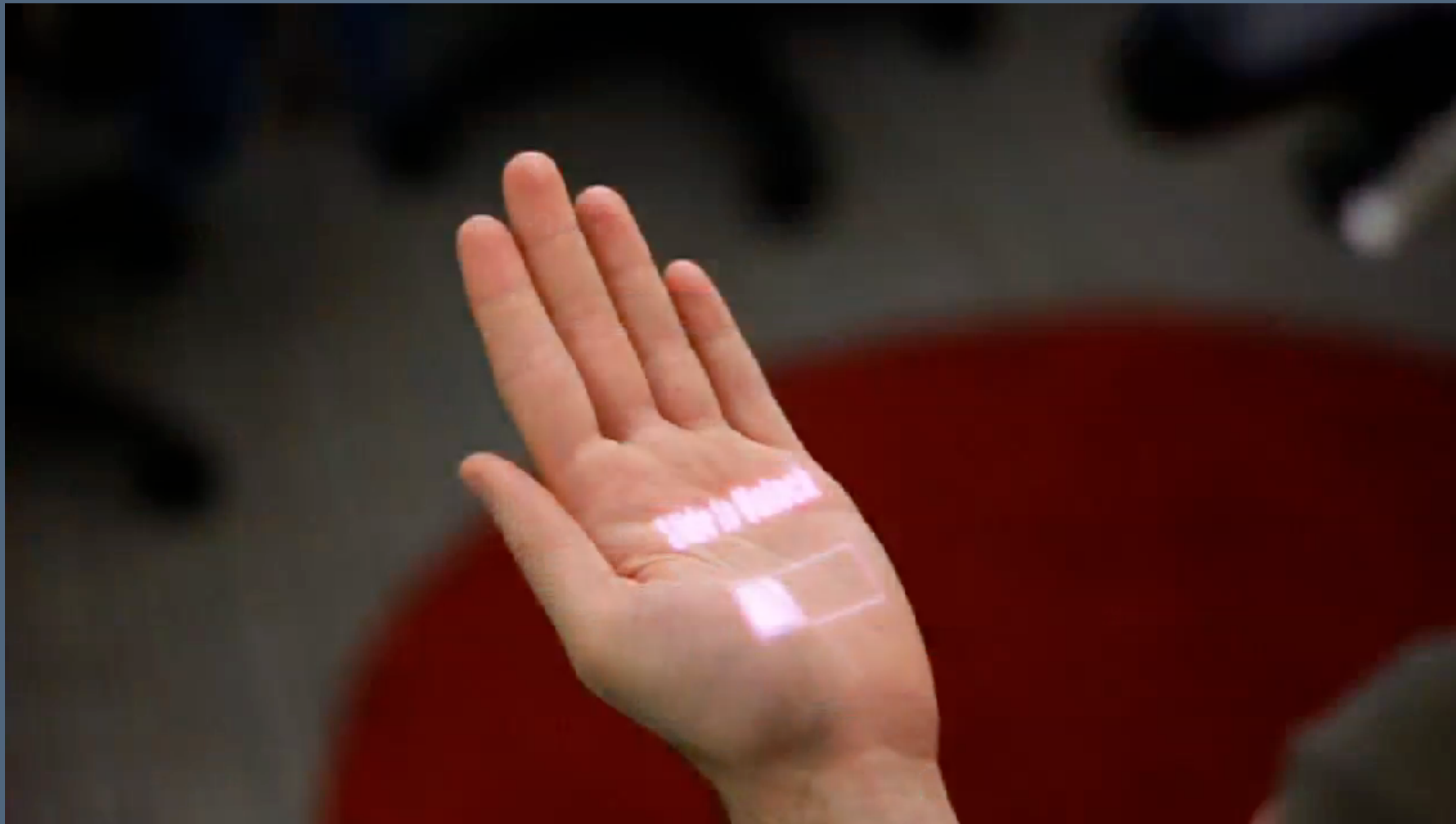
These technologies often must sense the layout of the space around them, then project the digital information into that space



[Image from NYT]

Projector-based AR

Mount projectors into the space to add digital augmentation



Body-mounted projector
turns surfaces into
interfaces

[Harrison, Benko, and Wilson 2011]

Projector-based AR

Mount projectors into the space to add digital augmentation

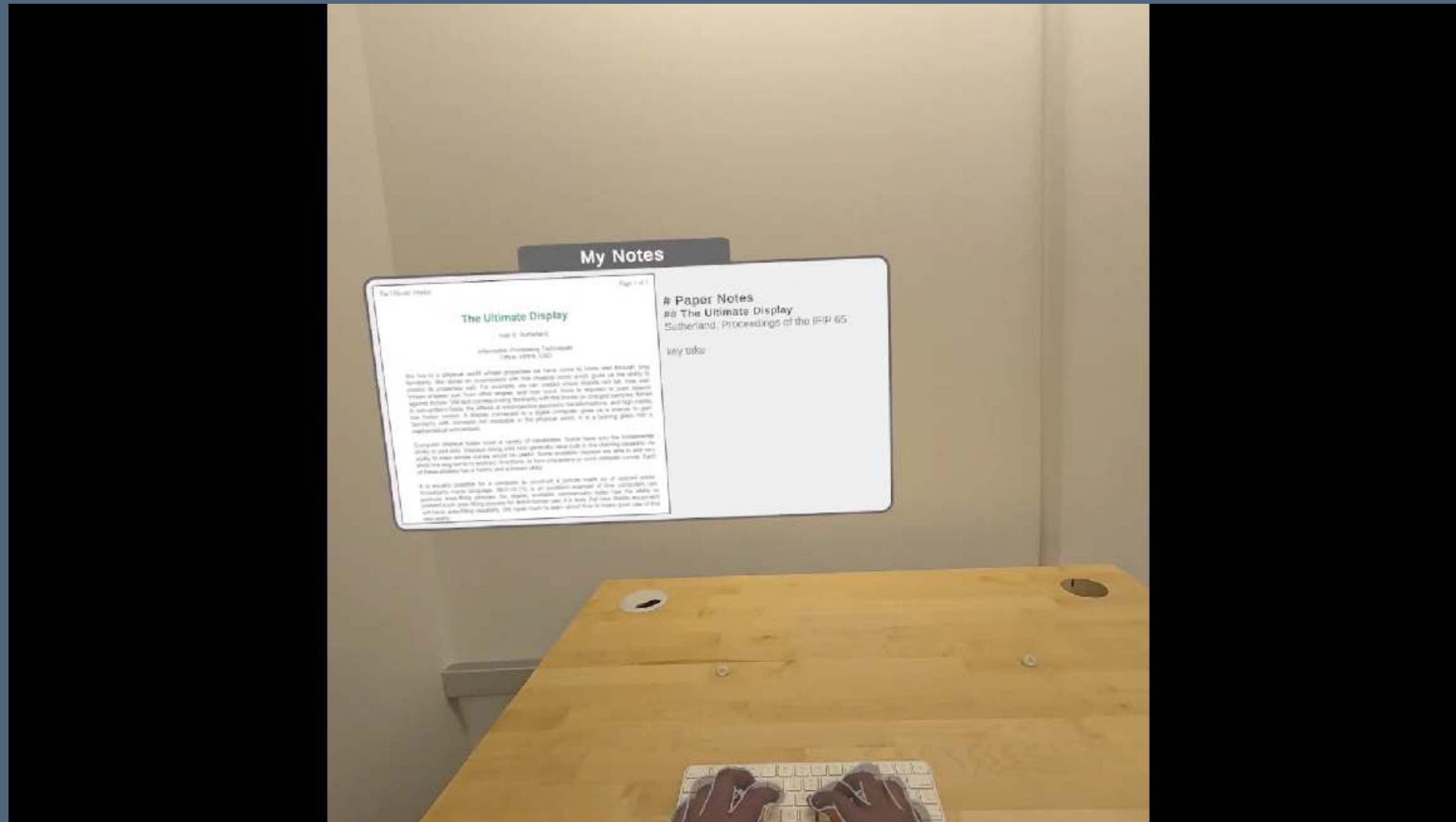


Body-mounted projector
turns surfaces into
interfaces

[Humane 2025 — RIP]

Passthrough display

Place digital elements into the physical world



In this case, automatically edit the scene to remove physical content that breaks the illusion [Kari and Abtahi 2025]

Remote People

Capture and transmit a wide range of 3D audiovisual cues



[Google Beam, Lawrence et al. 2021]

(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 Bailenson 2007]



Tradeoffs

Nontraditional 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

HCI interdisciplinarity

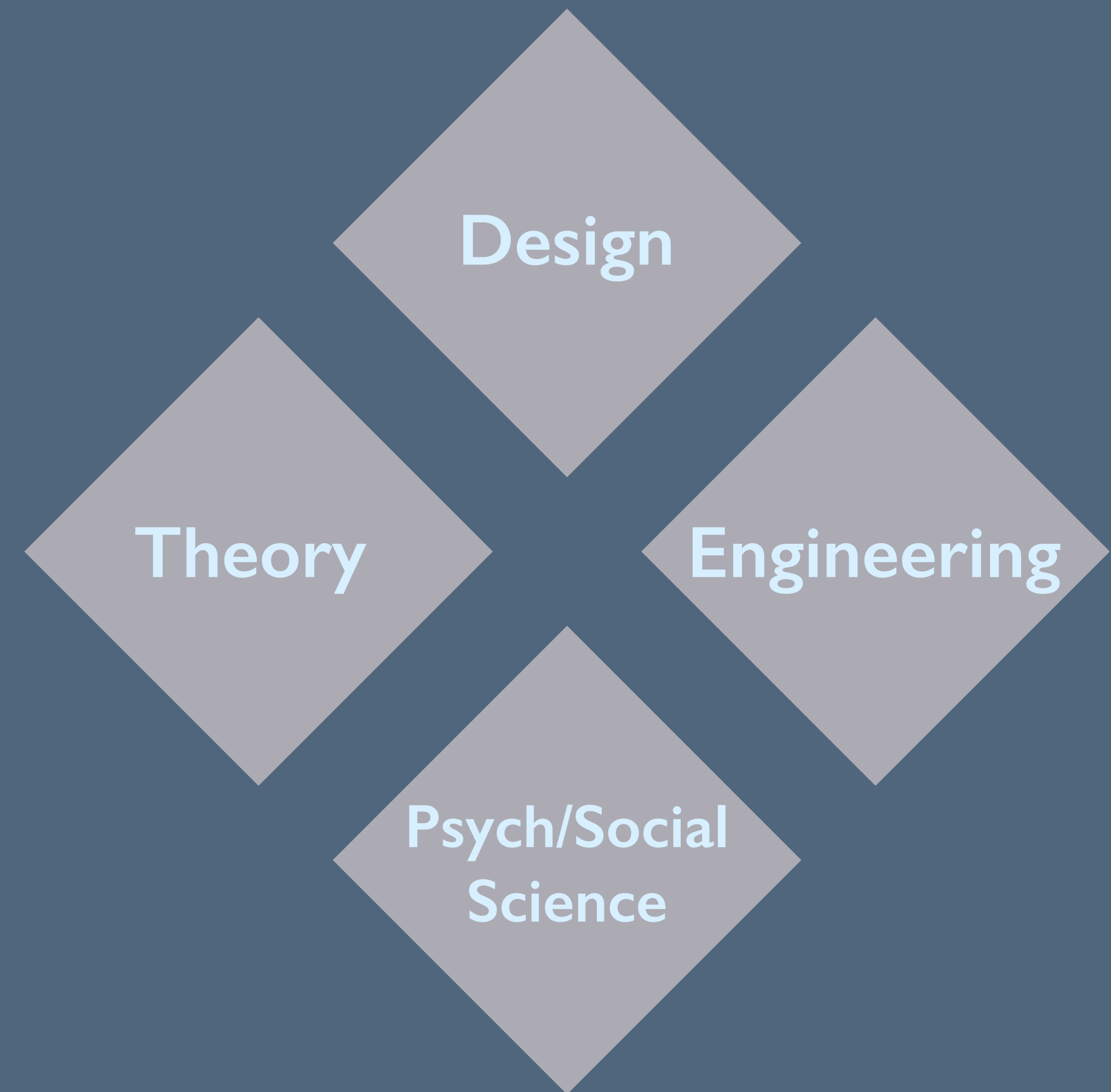
So far, we've mostly discussed **engineering**-style contributions to HCI. These involve envisioning and creating new technological approaches to human-computer interaction.

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

Psychology & Social Sciences

Design

Theory



HCI interdisciplinarity

Before today: “HCI is design process-iterated product”

After today

An algorithm paper can be HCI

A design paper can be HCI

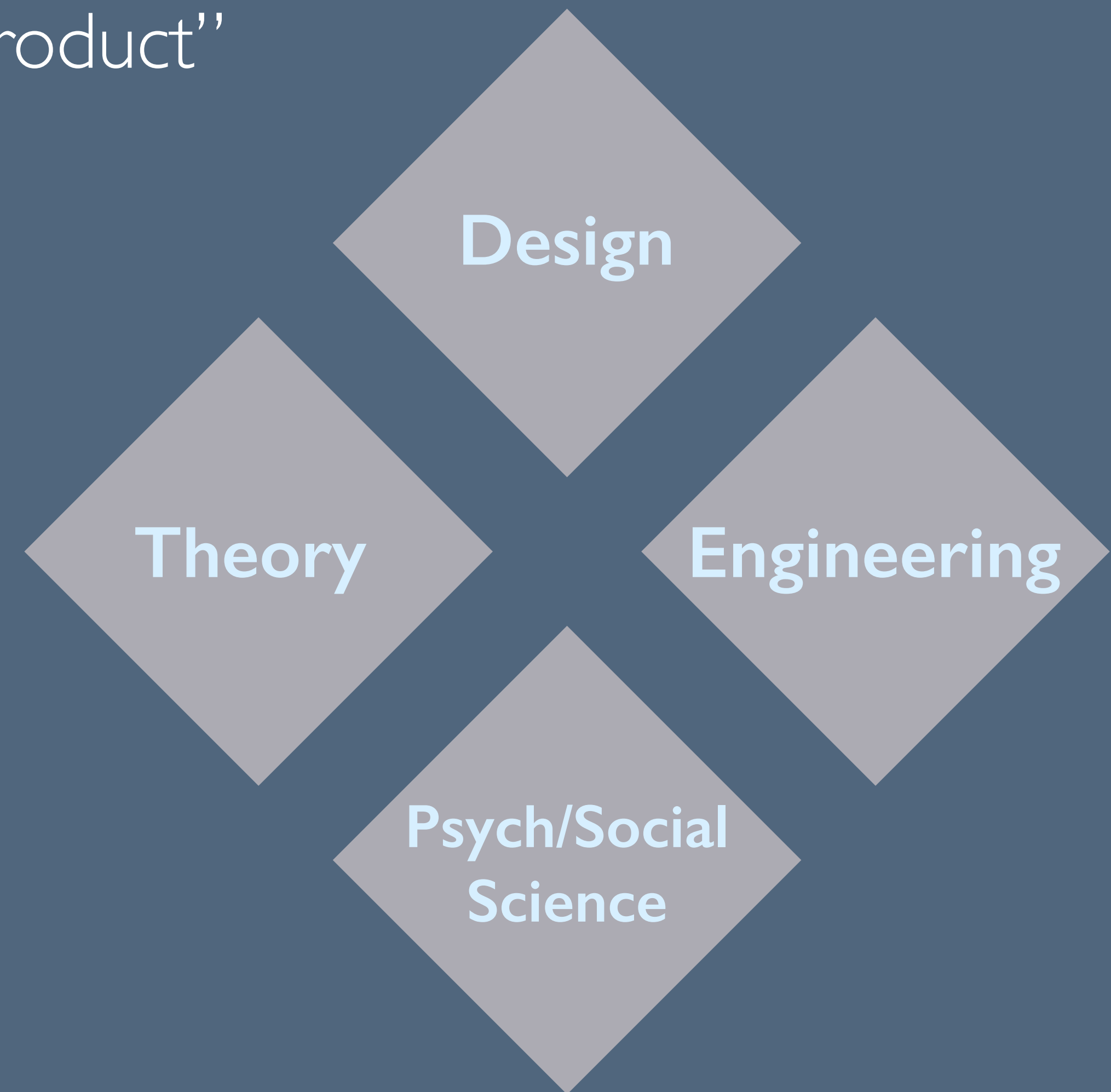
A qualitative paper can be HCI

A critical theory paper can be HCI

An EE/ME paper can be HCI

A field experiment can be HCI

...

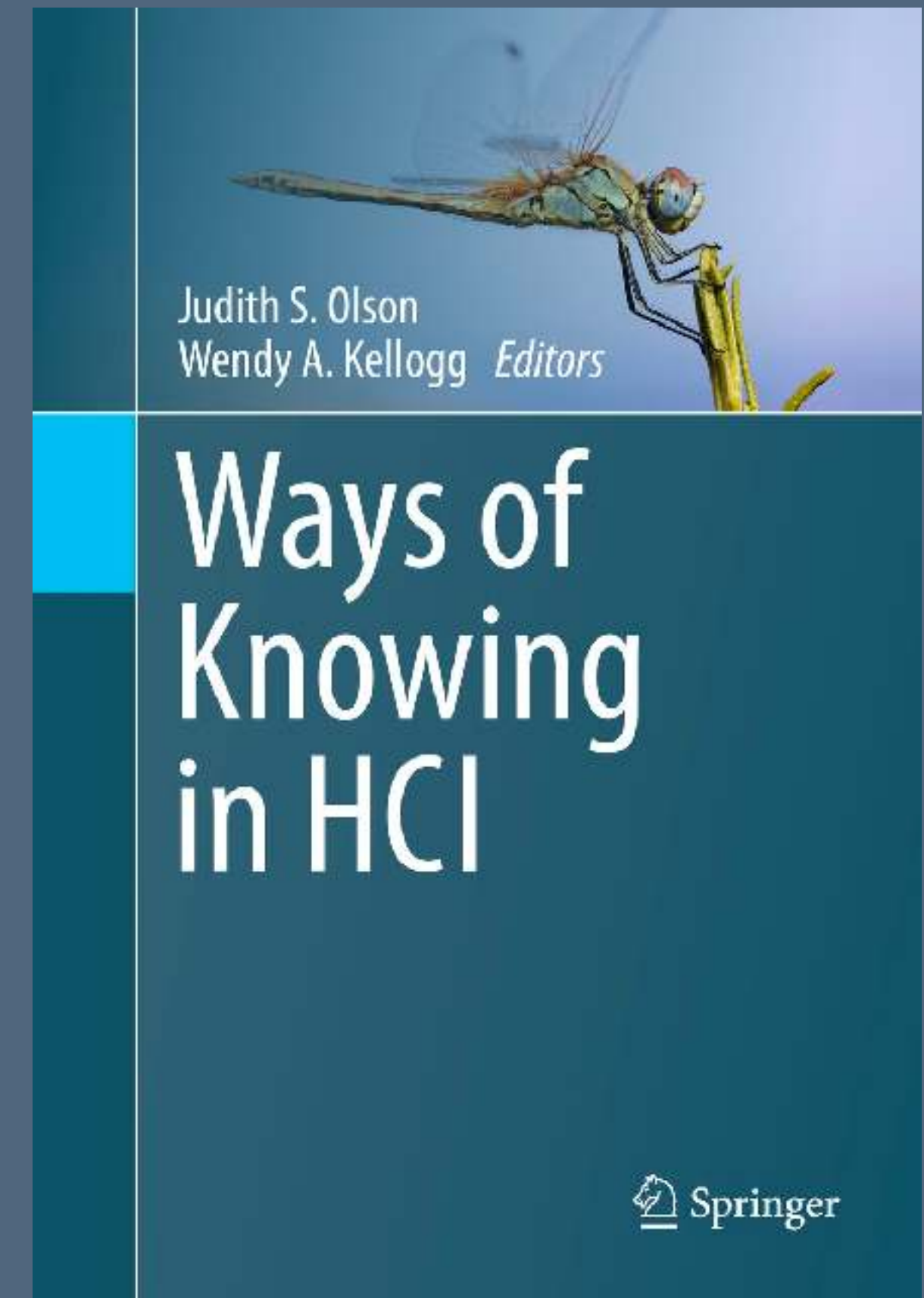


HCI interdisciplinarity

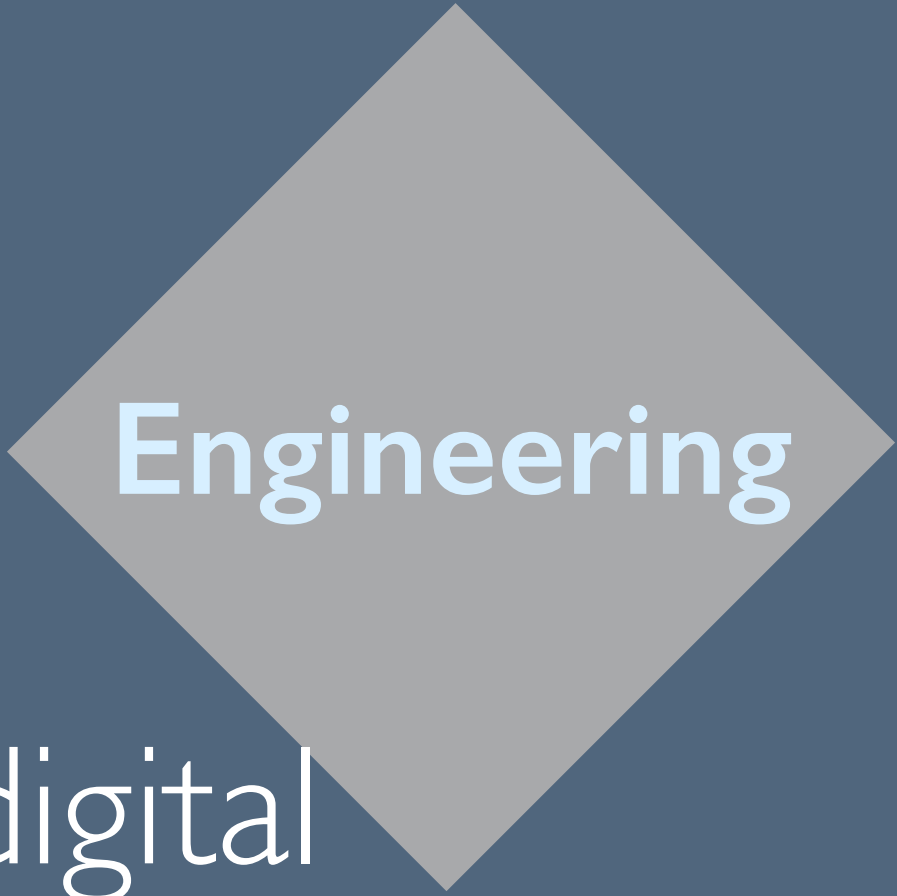
HCI succeeds by bringing together these perspectives.

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]



Engineering

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]

Engineering

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]

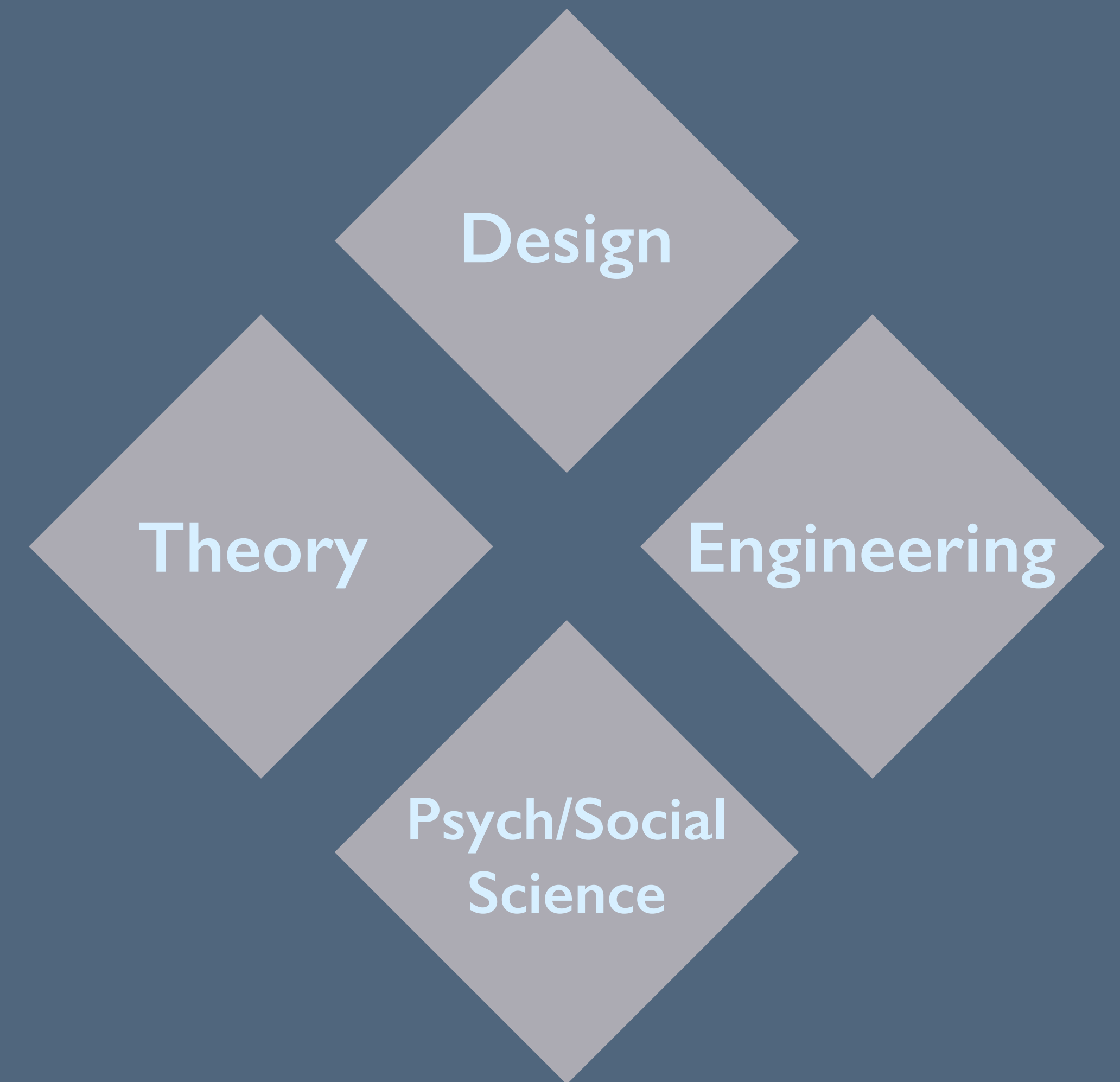


[Massie and Salisbury 1994]

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]

Psych/Social
Science

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

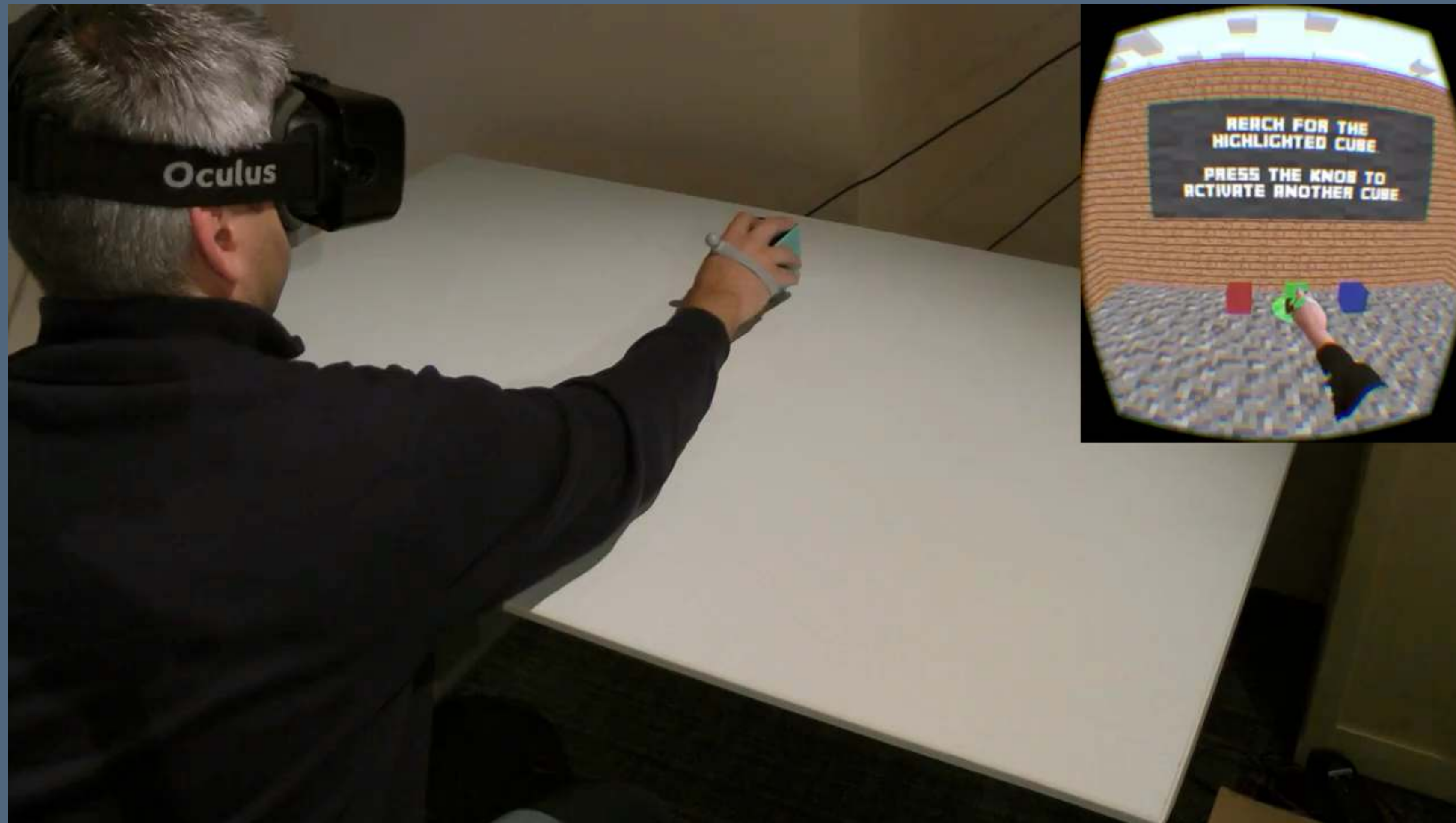
Experiment: place a shape on a table, and ask participants to view it and feel it from behind plastic that distorts how the shape looks. Then, ask people to draw what they think the shape actually is.

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]

Engineering



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

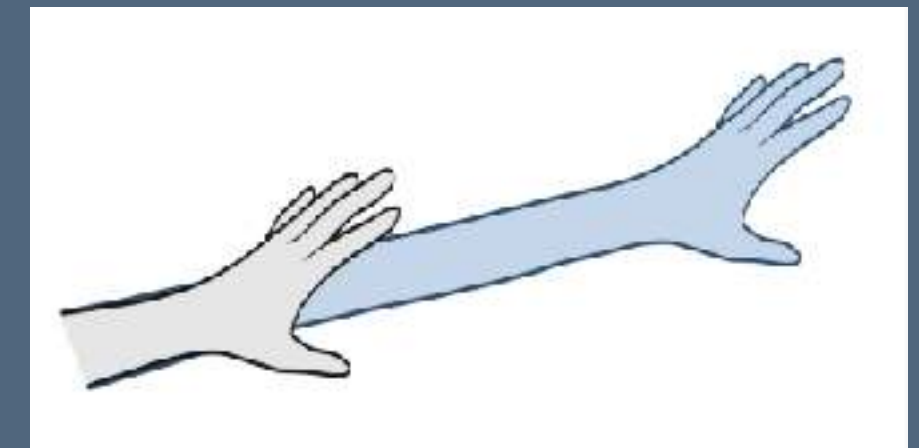
[Abtahi et al. 2022]

Theory

VR embodiments need not have any real-world equivalents

Scale up the size of the user's arm

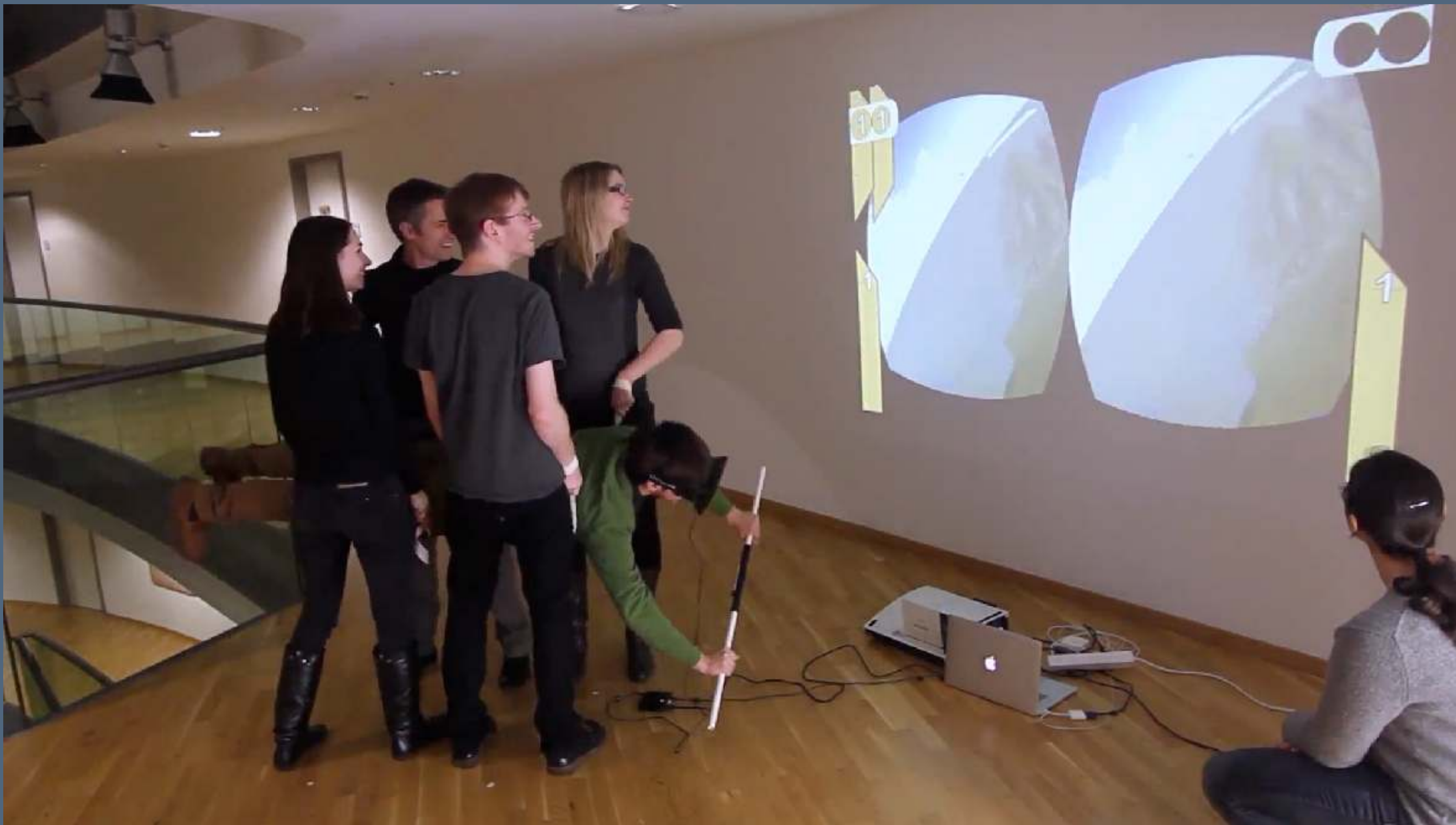
Scale up the user's avatar when moving long distances



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]

Design



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

References

Abtahi, Parastoo, et al. "Beyond being real: A sensorimotor control perspective on interactions in virtual reality." Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022.

Azmandian, Mahdi, et al. "Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences." Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016.

Bolt, Richard A. "'Put-that-there': Voice and gesture at the graphics interface." Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques. 1980.

Cheng, Lung-Pan, Patrick Lühne, Pedro Lopes, Christoph Sterz, and Patrick Baudisch. "Haptic Turk: A motion platform based on people." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014.

Choi, Inrak, et al. "Grabity: A wearable haptic interface for simulating weight and grasping in virtual reality." Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. 2017.

Cohn, Gabe, et al. "Humantenna: Using the body as an antenna for real-time whole-body interaction." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012.

Consolvo, Sunny, et al. "Activity sensing in the wild: A field trial of UbiFit Garden." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008.

Dourish, Paul. Where the Action Is: The Foundations of Embodied Interaction. MIT Press, 2004.

Glauser, Oliver, et al. "A stretch-sensing soft glove for interactive hand pose estimation." ACM SIGGRAPH 2019 Emerging Technologies. 2019.

Hammer, M.C. U Can't Touch This. 1990. Capitol Records.

References

- Harrison, Chris, Hrvoje Benko, and Andrew D. Wilson. "OmniTouch: Wearable multitouch interaction everywhere." Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. 2011.
- Humane. Ai Pin. 2024. Discontinued February 2025.
- Irvantchi, Yasha, Yi Zhao, Kenrick Kin, and Alanson P. Sample. "SAWSense: Using surface acoustic waves for surface-bound event recognition." Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. 2023.
- Kaifosh, Patrick, Thomas R. Reardon, and CTRL-labs at Reality Labs. "A generic non-invasive neuromotor interface for human-computer interaction." Nature 645 (2025): 702-711.
- Kari, Mohamed, and Parastoo Abtahi. "Reality Promises: Virtual-physical decoupling illusions in mixed reality via invisible mobile robots." Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology. 2025.
- Khurana, Rushil, et al. "GymCam: Detecting, recognizing and tracking simultaneous exercises in unconstrained scenes." Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2.4 (2018): 1-17.
- Laput, Gierad, et al. "EM-Sense: Touch recognition of uninstrumented, electrical and electromechanical objects." Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology. 2015.
- Laput, Gierad, et al. "Ubioustics: Plug-and-play acoustic activity recognition." Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. 2018.
- Lawrence, Jason, et al. "Project Starline: A high-fidelity telepresence system." ACM Transactions on Graphics 40.6 (2021): 1-16.
- Le Goc, Mathieu, et al. "Zooids: Building blocks for swarm user interfaces." Proceedings of the 29th Annual Symposium on User Interface Software and Technology. 2016.

References

Massie, Thomas H., and J. Kenneth Salisbury. "The PHANTOM haptic interface: A device for probing virtual objects." Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. Vol. 55. No. 1. 1994.

Meta. Meta Ray-Ban Display and Meta Neural Band. 2025.

Narayanswamy, Girish, et al. "Scaling wearable foundation models." International Conference on Learning Representations (ICLR). 2025.

Olson, Judith S., and Wendy A. Kellogg, eds. Ways of Knowing in HCI. Springer, 2014.

OpenAI. Introducing GPT-6 Astra. 2026. https://www.youtube.com/watch?v=IQNsdr-Qx_I

Rock, Irvin, and Jack Victor. "Vision and touch: An experimentally created conflict between the two senses." Science 143.3606 (1964): 594-596.

Sutherland, Ivan. "The ultimate display." Proceedings of the IFIP Congress. 1965.

Tao, Yujie, Shan-Yuan Teng, and Pedro Lopes. "Altering perceived softness of real rigid objects by restricting fingerpad deformation." Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology. 2021.

Wellner, Pierre. "Interacting with paper on the DigitalDesk." Communications of the ACM 36.7 (1993): 87-96.

Willis, Karl, et al. "Printed optics: 3D printing of embedded optical elements for interactive devices." Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology. 2012.

Xu, Xuhai, et al. "Enabling hand gesture customization on wrist-worn devices." Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022.

Yamada, Wataru, et al. "iSphere: Self-luminous spherical drone display." Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology. 2017.

Yee, Nick, and Jeremy Bailenson. "The Proteus effect: The effect of transformed self-representation on behavior." Human Communication Research 33.3 (2007): 271-290.

Yu, Tianyu, et al. "MorphingSkin: A skin-like platform that integrates multimodal hydraulic actuators based on flexible electroosmotic pumps." Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology. 2025.