HandSight: A Touch-Based Wearable System to Increase Information Accessibility for People with Visual Impairments

Lee Stearns

CS PhD Candidate | lstearns@umd.edu

Examining Committee:

Dr. Jon Froehlich
Chair / Advisor

Dr. Rama Chellappa
Co-Advisor

Dr. Leah Findlater
Member

Dr. Ramani Duraiswami
Member

Dr. Gregg Vanderheiden
Dean's Representative

MAKEABILITY LAB | inclusive design lab | DUB | HCIL Human Computer Interaction Laboratory | COMPUTER SCIENCE UNIVERSITY OF MARYLAND | UNIVERSITY of WASHINGTON | UNIVERSITY OF MARYLAND

Activities of Daily Living

Getting dressed

Preparing food

Wayfinding

Shopping

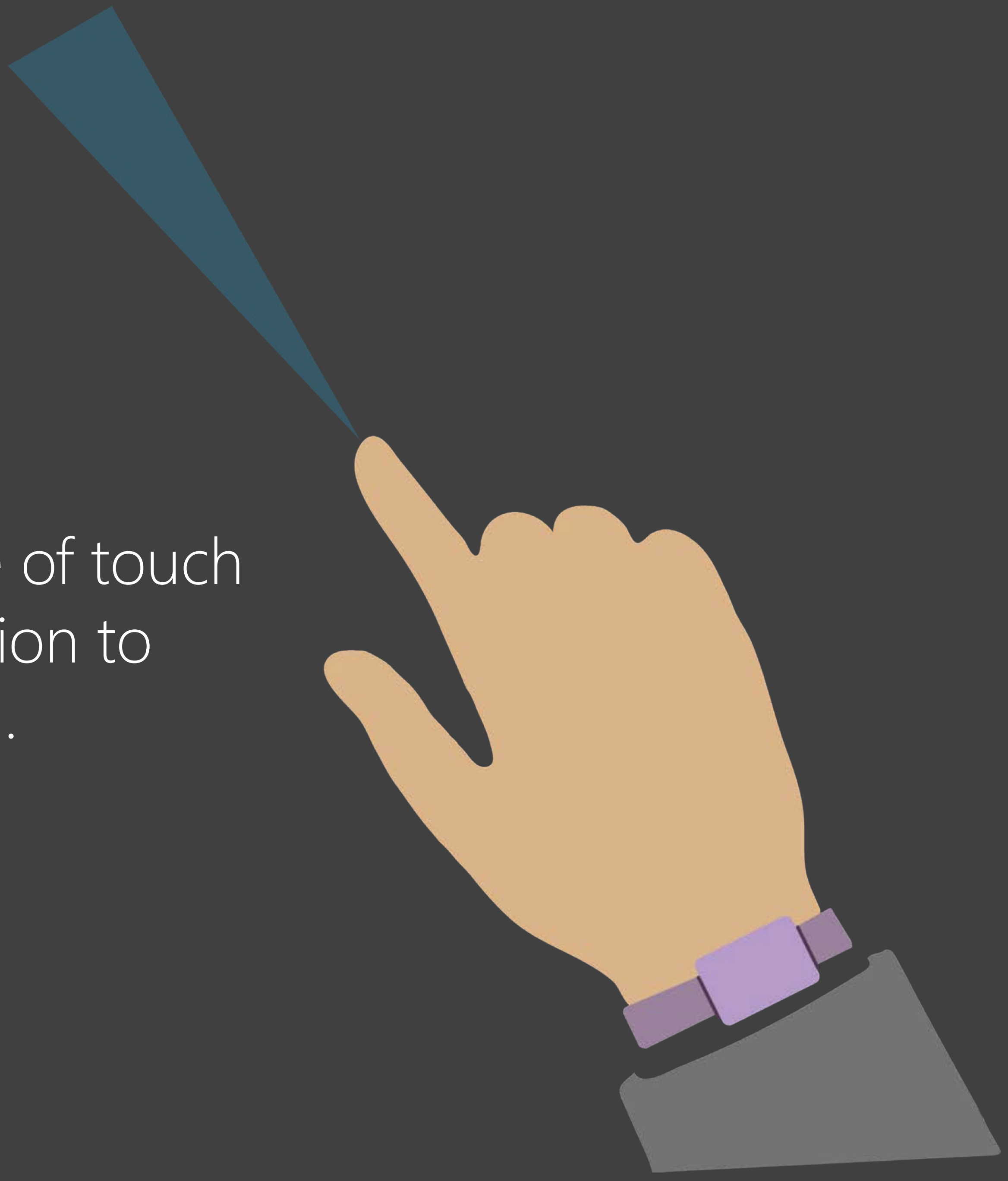
Previous research has used mobile cameras and computer vision for **at-a-distance tasks**

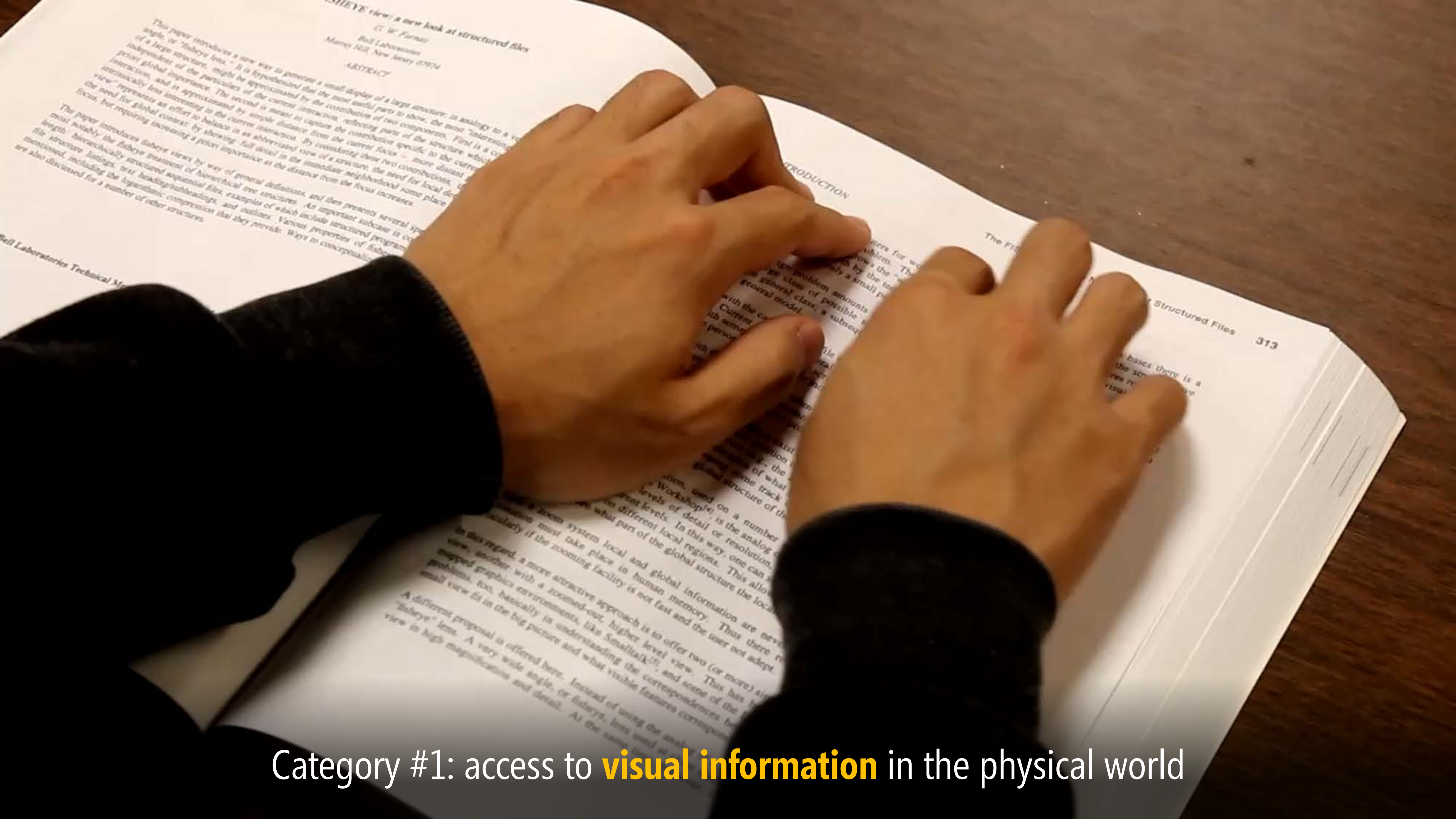
but few have attempted to provide access to **visual information** through **touch**

For people with visual impairments, touch is a **primary means** of obtaining information about the physical world

Research Goal

Augment visually impaired users' sense of touch with interactive, real-time computer vision to improve the accessibility of information.





Category #1: access to **visual information** in the physical world


Category #2: access to **digital information** by controlling computers or mobile devices

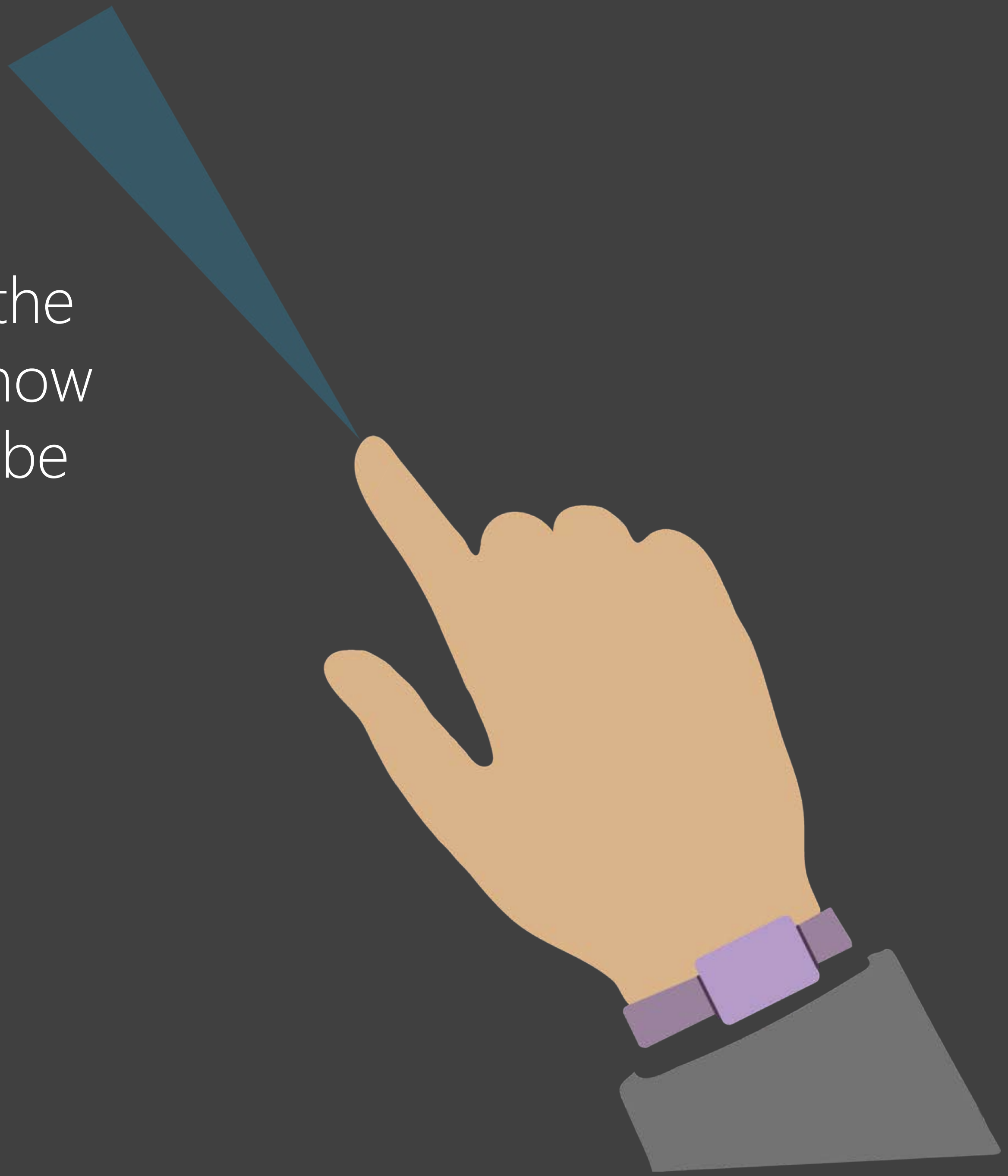
Open Questions

What is the best method to recognize the content that the user is touching, and how should information about that content be conveyed to the user?

HandSight

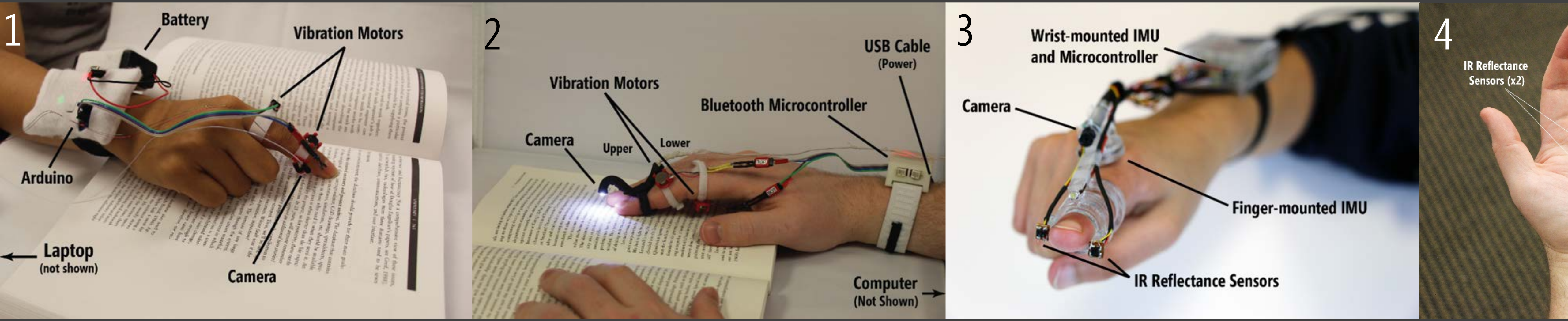
Key aspects:

1. Physical hardware
2. Real-time algorithms
3. Interactive interface
4. Usability evaluations



Summary of Contributions

Development and iterative refinement of **HandSight**, a novel wearable system to assist visually impaired users.

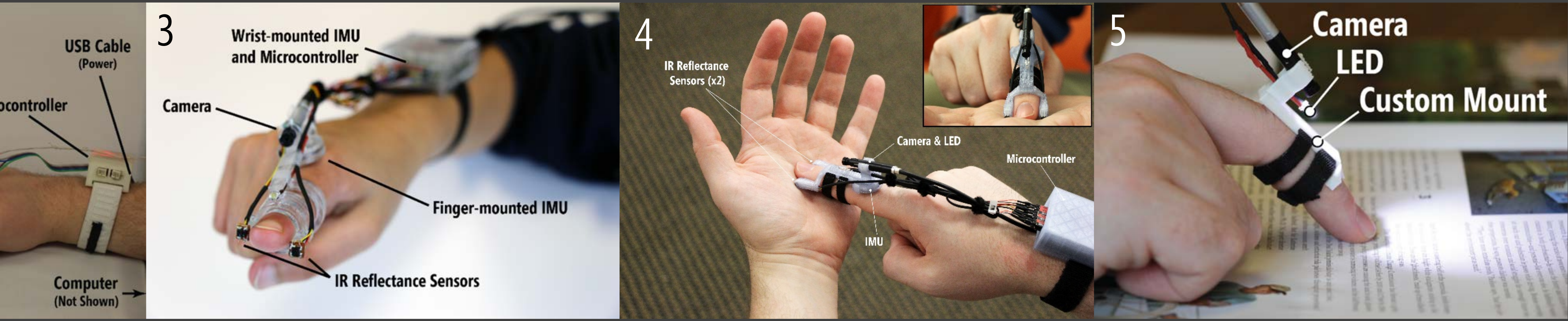


Reading and exploring printed documents

Controlling mobile devic
through on-body inpu

Summary of Contributions

Development and iterative refinement of **HandSight**, a novel wearable system to assist visually impaired users.



Controlling mobile devices
through on-body input

AR Magnification &
Identifying Colors/Patterns

Summary of Contributions

Development and iterative refinement of **HandSight**, a novel wearable system to assist visually impaired users.

Findings from **user evaluations** across a diverse set of tasks, which demonstrate advantages and disadvantages of our **finger-worn approach**.

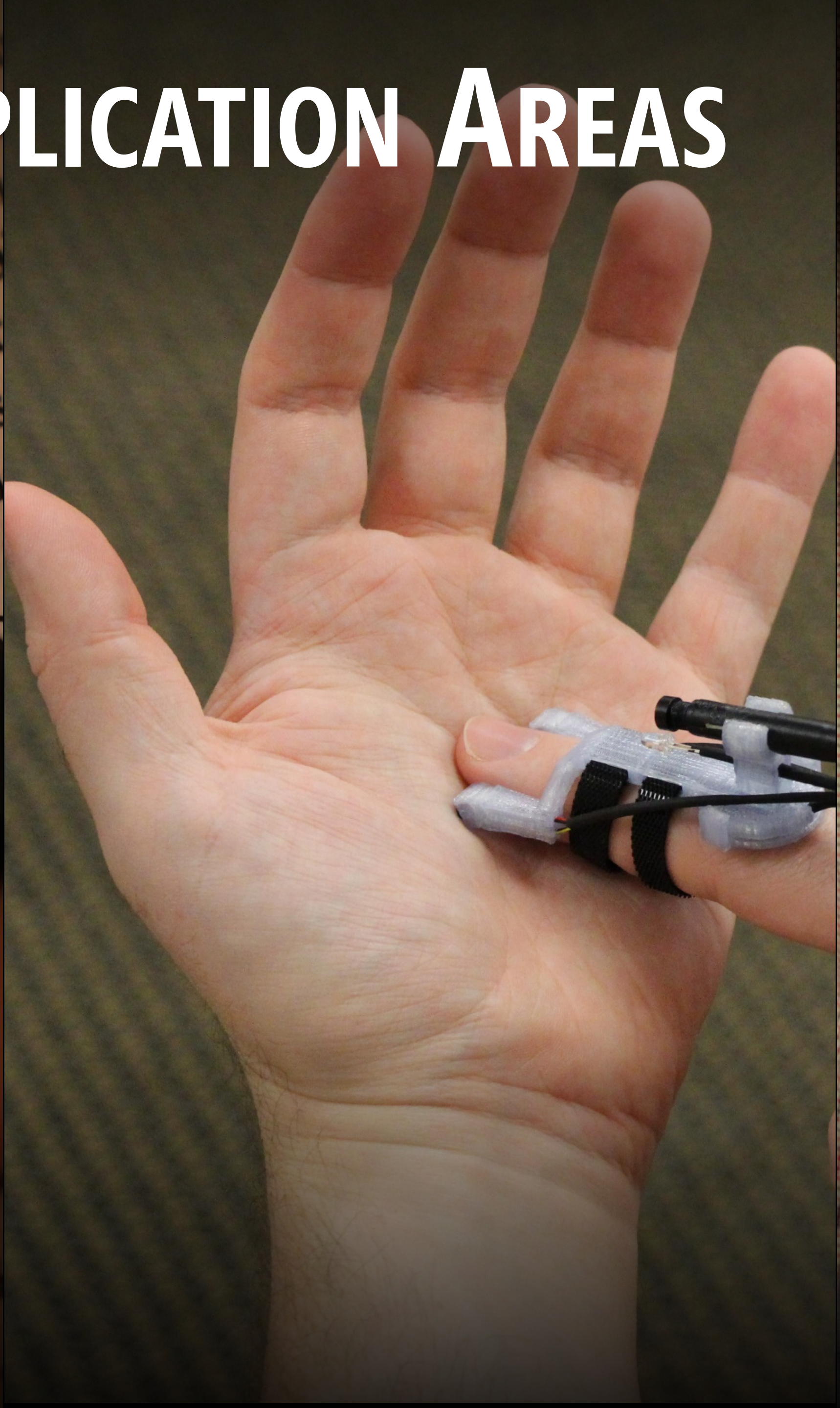
Summary of Contributions

Development and iterative refinement of **HandSight**, a novel wearable system to assist visually impaired users.

Findings from **user evaluations** across a diverse set of tasks, which demonstrate advantages and disadvantages of our **finger-worn approach**.

Design Implications for **future wearable assistive systems** and for the fields of accessibility, computer vision, AR and VR, and human-computer interaction.

HandSight Application Areas

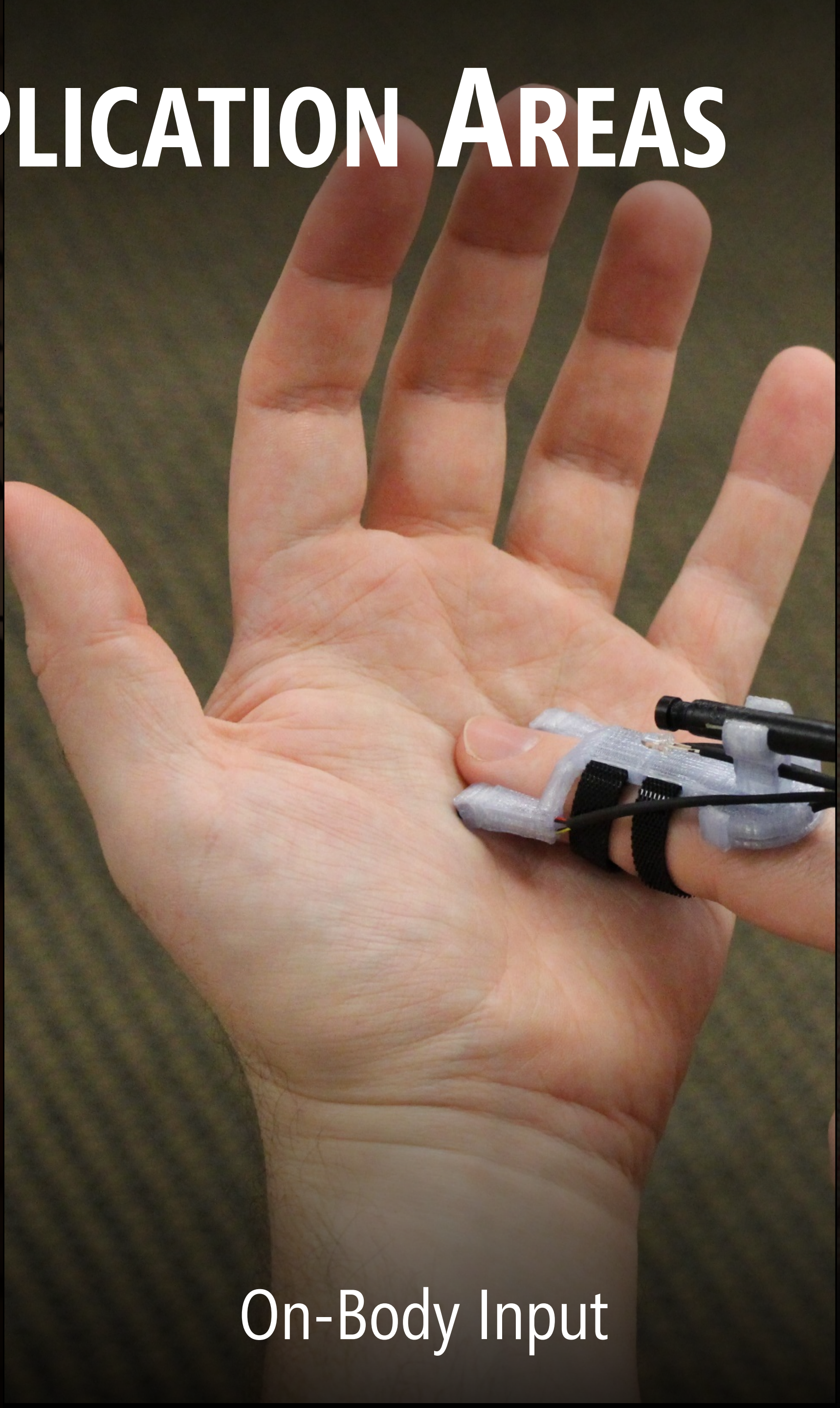


HandSight Application Areas

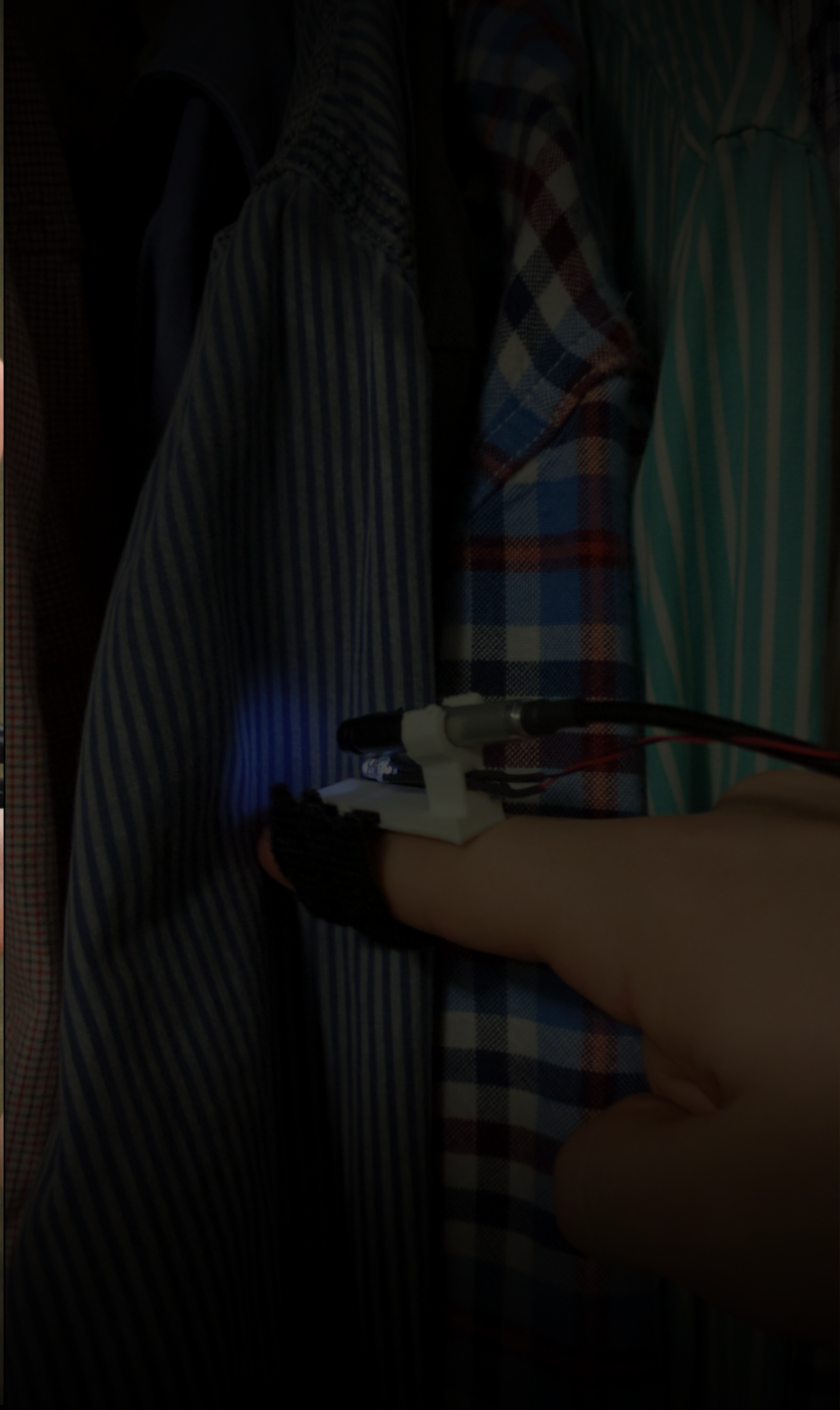
Reading/Exploring Text

HandSight Application Areas

Reading/Exploring Text



On-Body Input



HandSight Application Areas

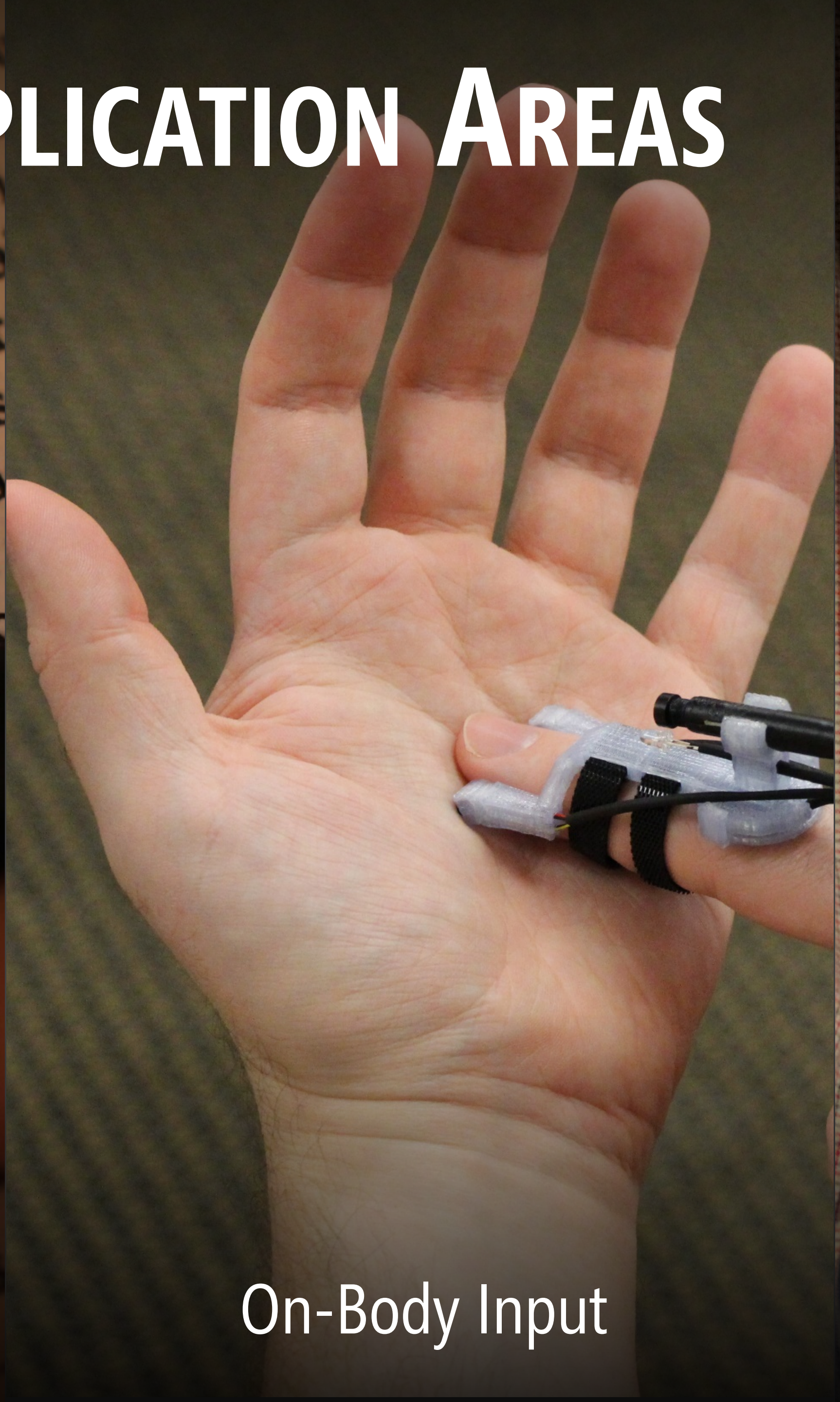
Reading/Exploring Text

On-Body Input

Clothing Colors/Patterns

HandSight Application Areas

Reading/Exploring Text

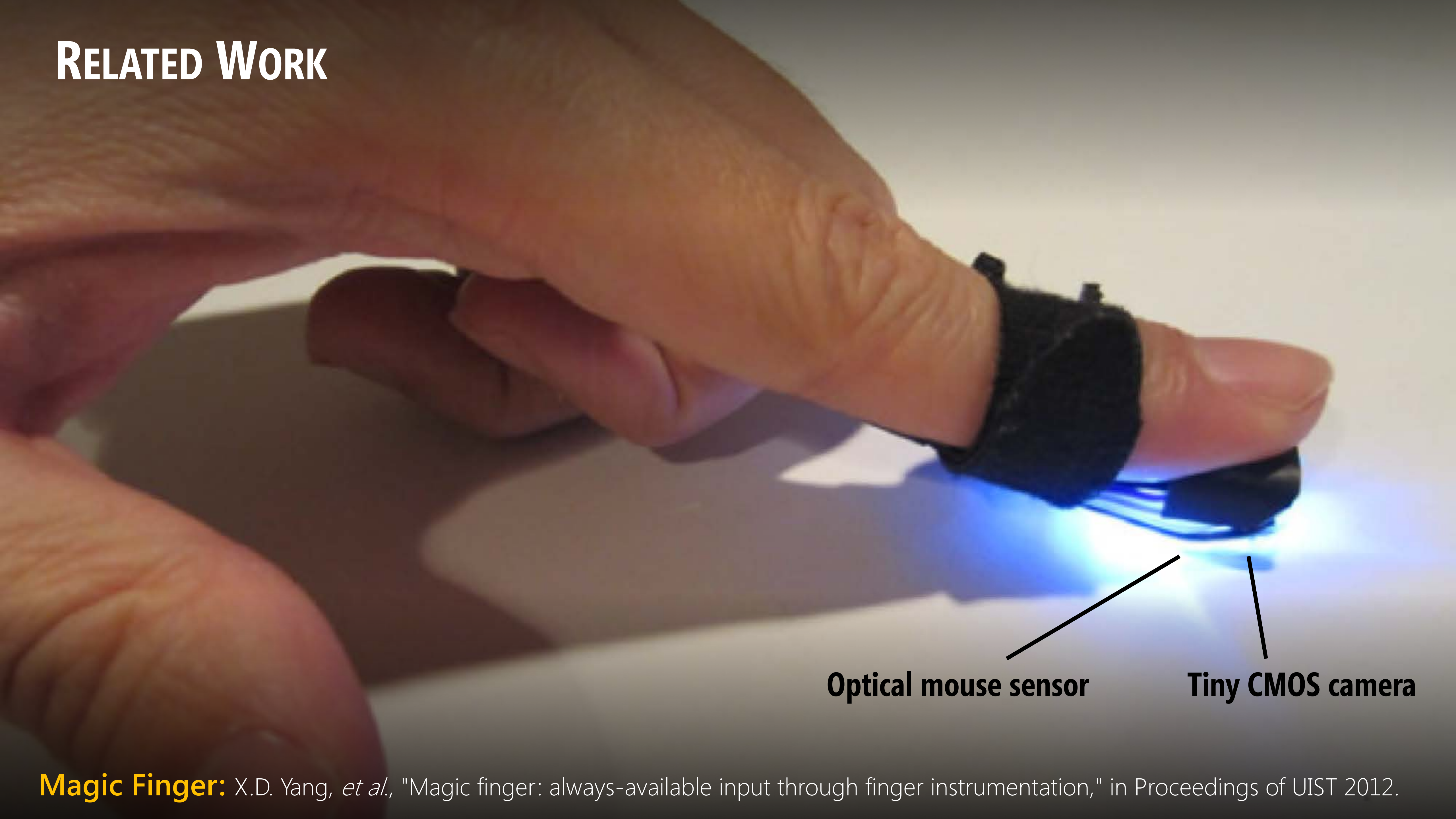


On-Body Input

Clothing Colors/Patterns

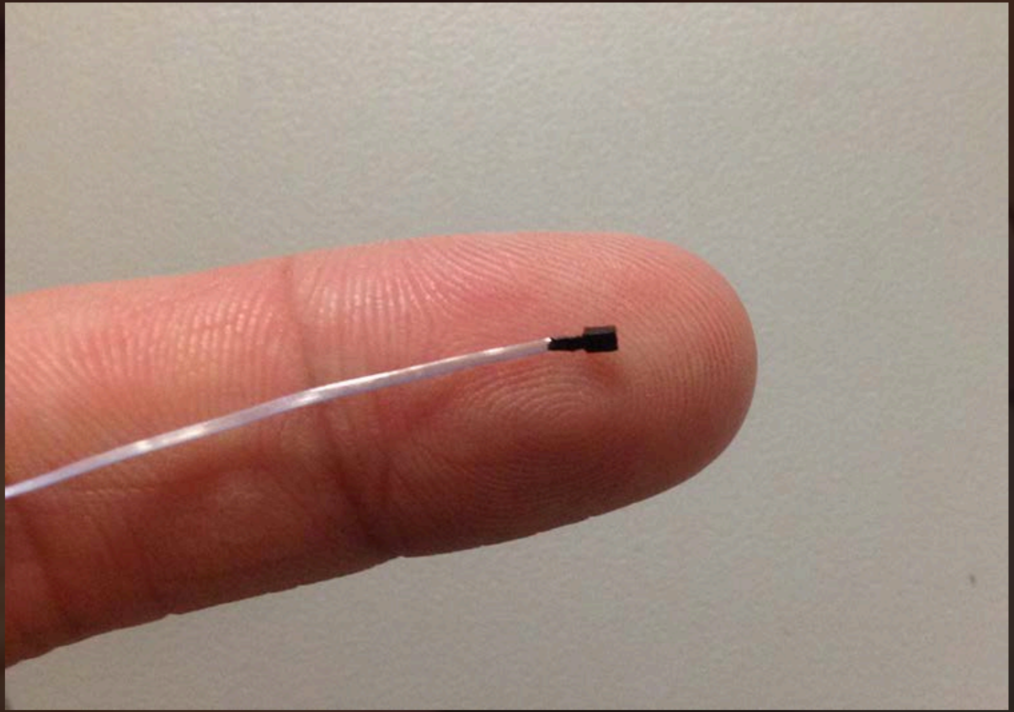
Related Work

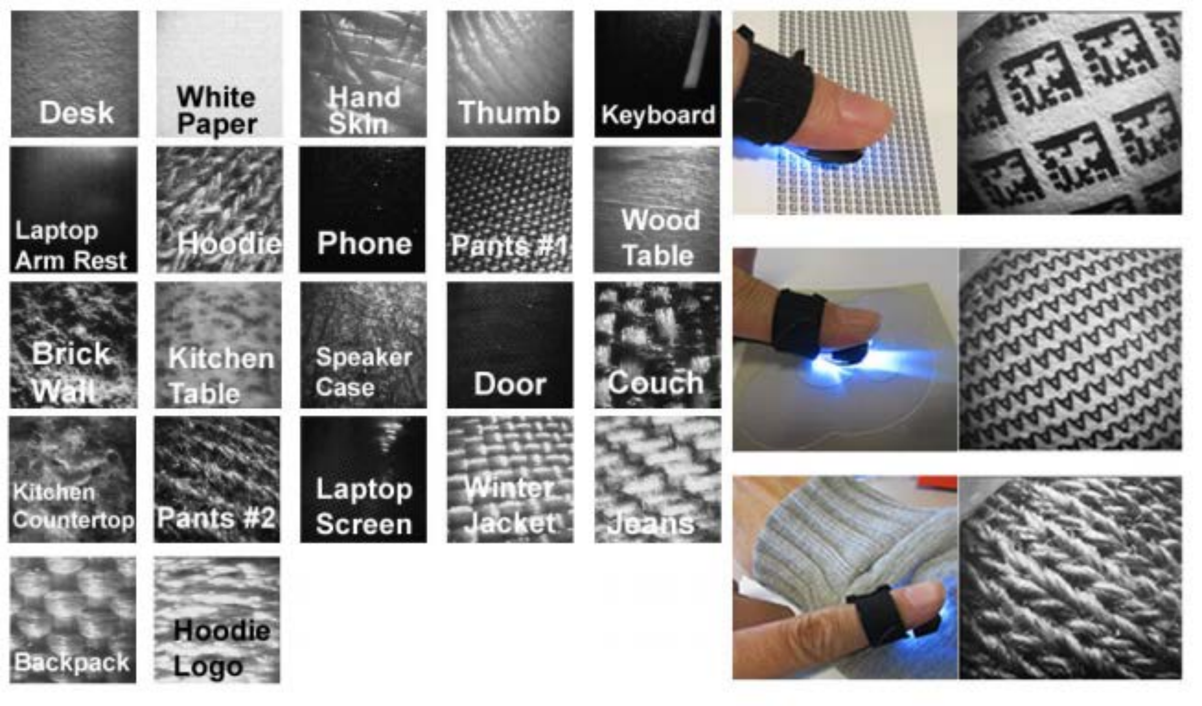
RELATED WORK



Magic Finger: X.D. Yang, *et al.*, "Magic finger: always-available input through finger instrumentation," in Proceedings of UIST 2012.

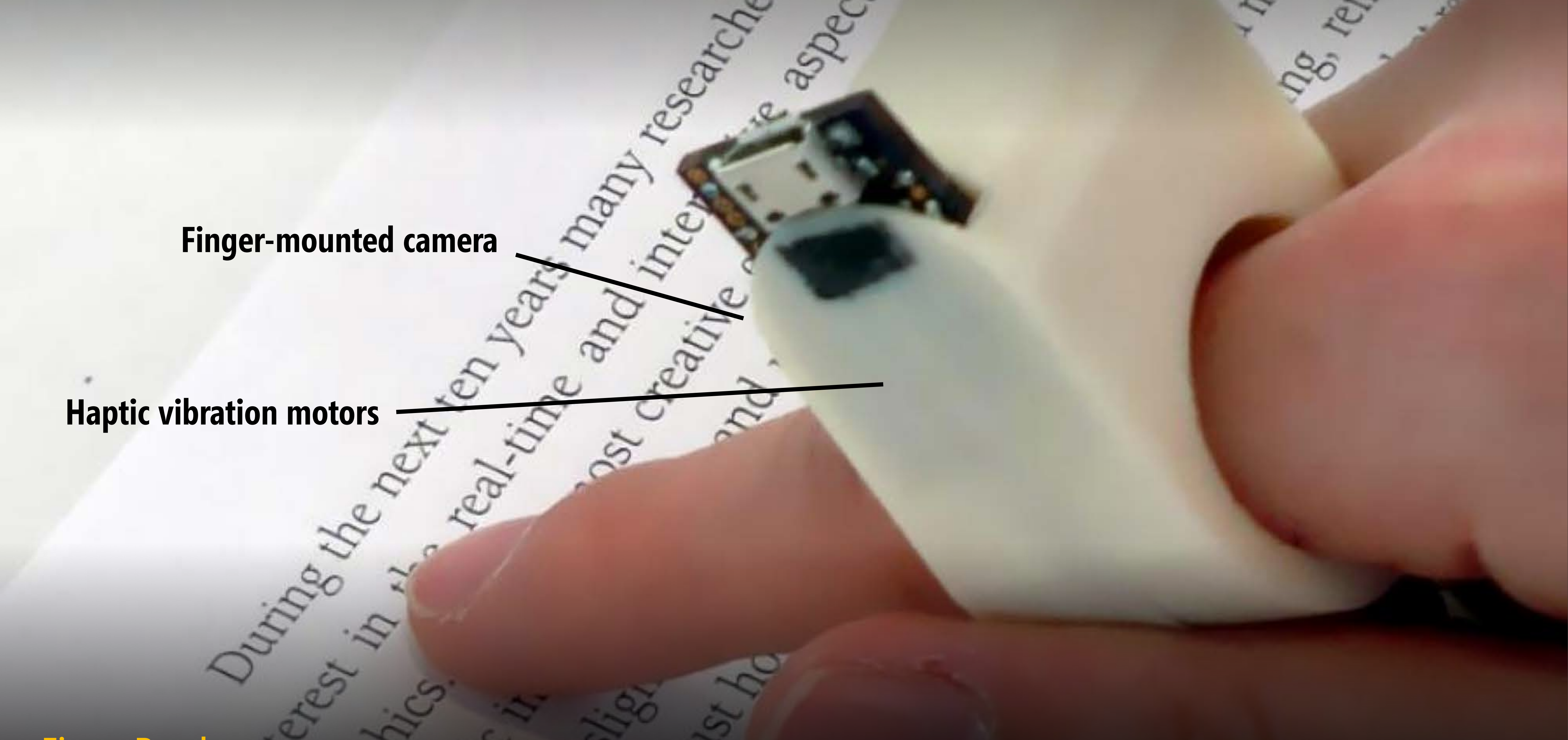
Related Work





Magic Finger: X.D. Yang, *et al.*, "Magic finger: always-available input through finger instrumentation," in Proceedings of UIST 2012.

Related Work



FingerReader: R. Shilkrot, *et al.*, "FingerReader: a Wearable Device to Explore Printed Text on the Go," in Proceedings of CHI 2015.

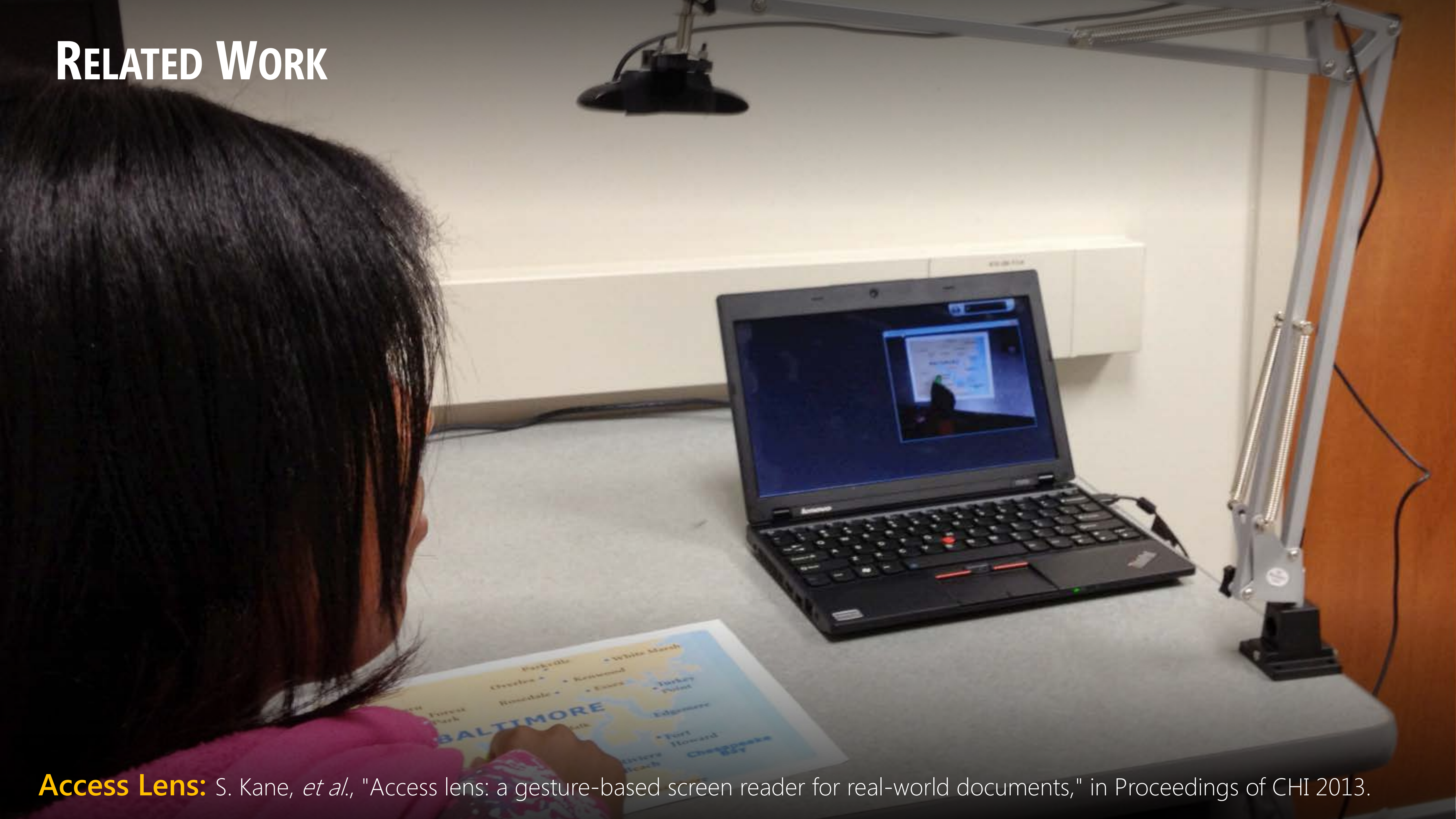
Related Work

FingerReader: R. Shilkrot, *et al.*, "FingerReader: a Wearable Device to Explore Printed Text on the Go," in Proceedings of CHI 2015.

Related Work

OrCam MyEye: http://www.orcam.com

Related Work

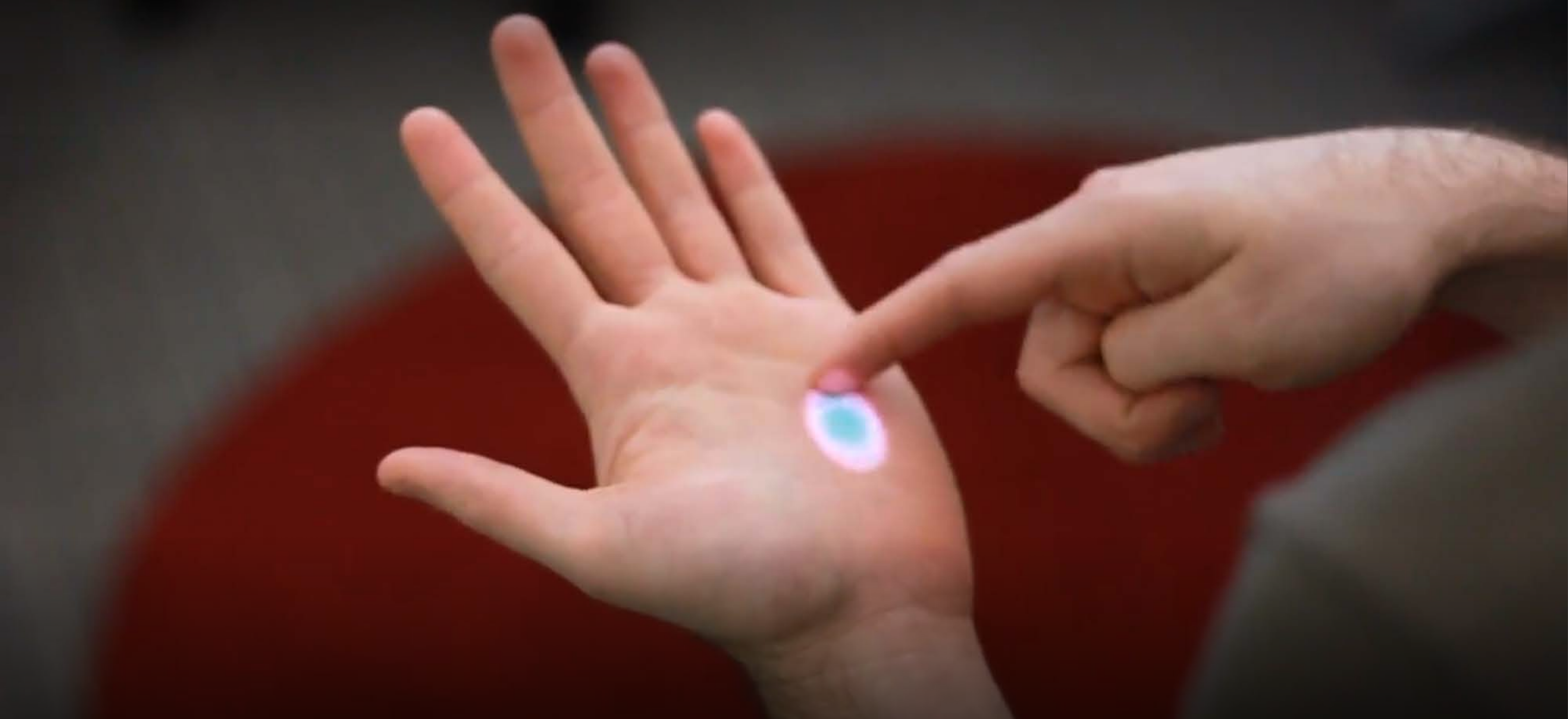


Access Lens: S. Kane, *et al.*, "Access lens: a gesture-based screen reader for real-world documents," in Proceedings of CHI 2013.

Related Work

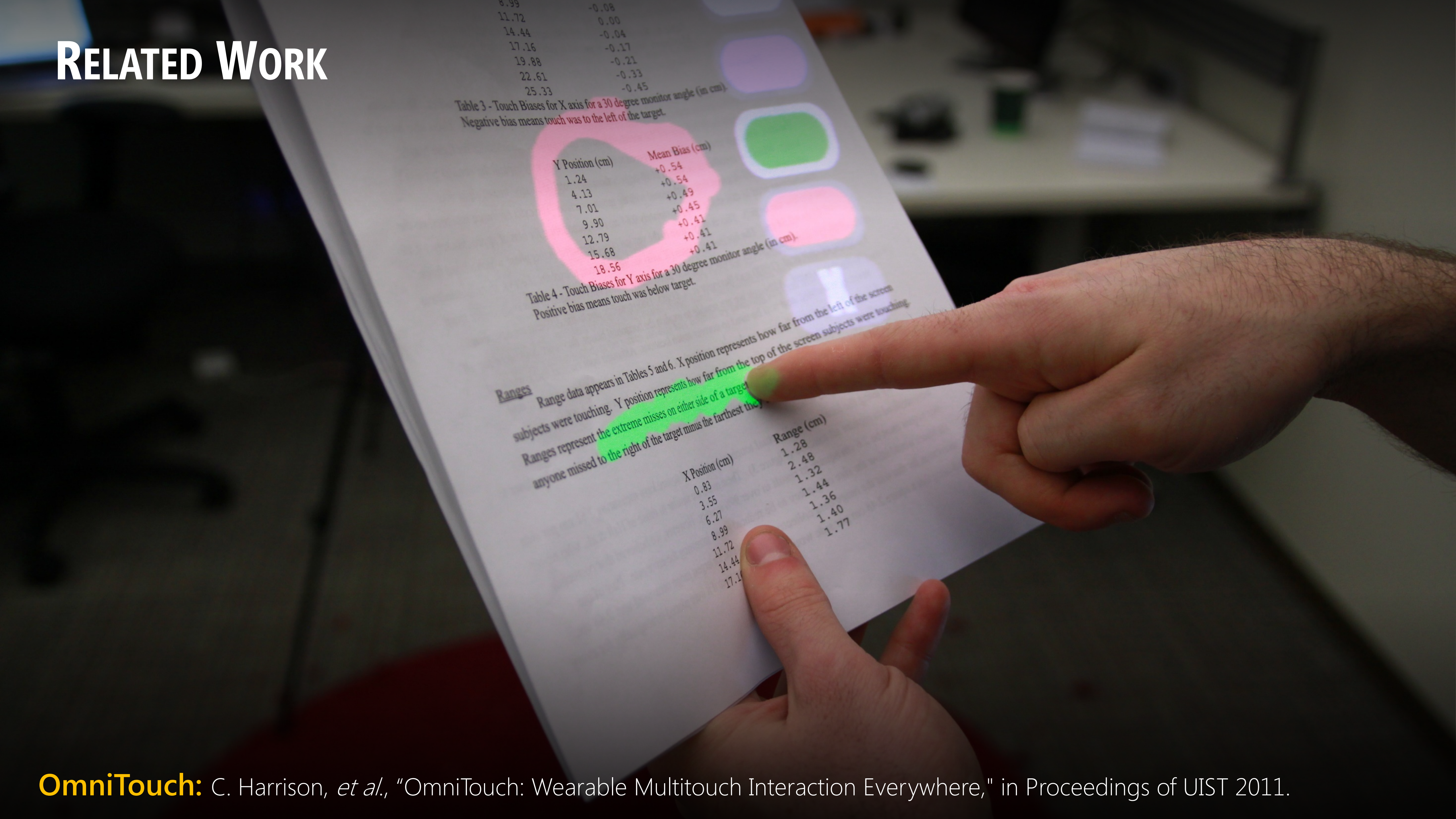
Access Lens: S. Kane, *et al.*, "Access lens: a gesture-based screen reader for real-world documents," in Proceedings of CHI 2013.

Related Work



OmniTouch: C. Harrison, *et al.*, "OmniTouch: Wearable Multitouch Interaction Everywhere," in Proceedings of UIST 2011.

Related Work

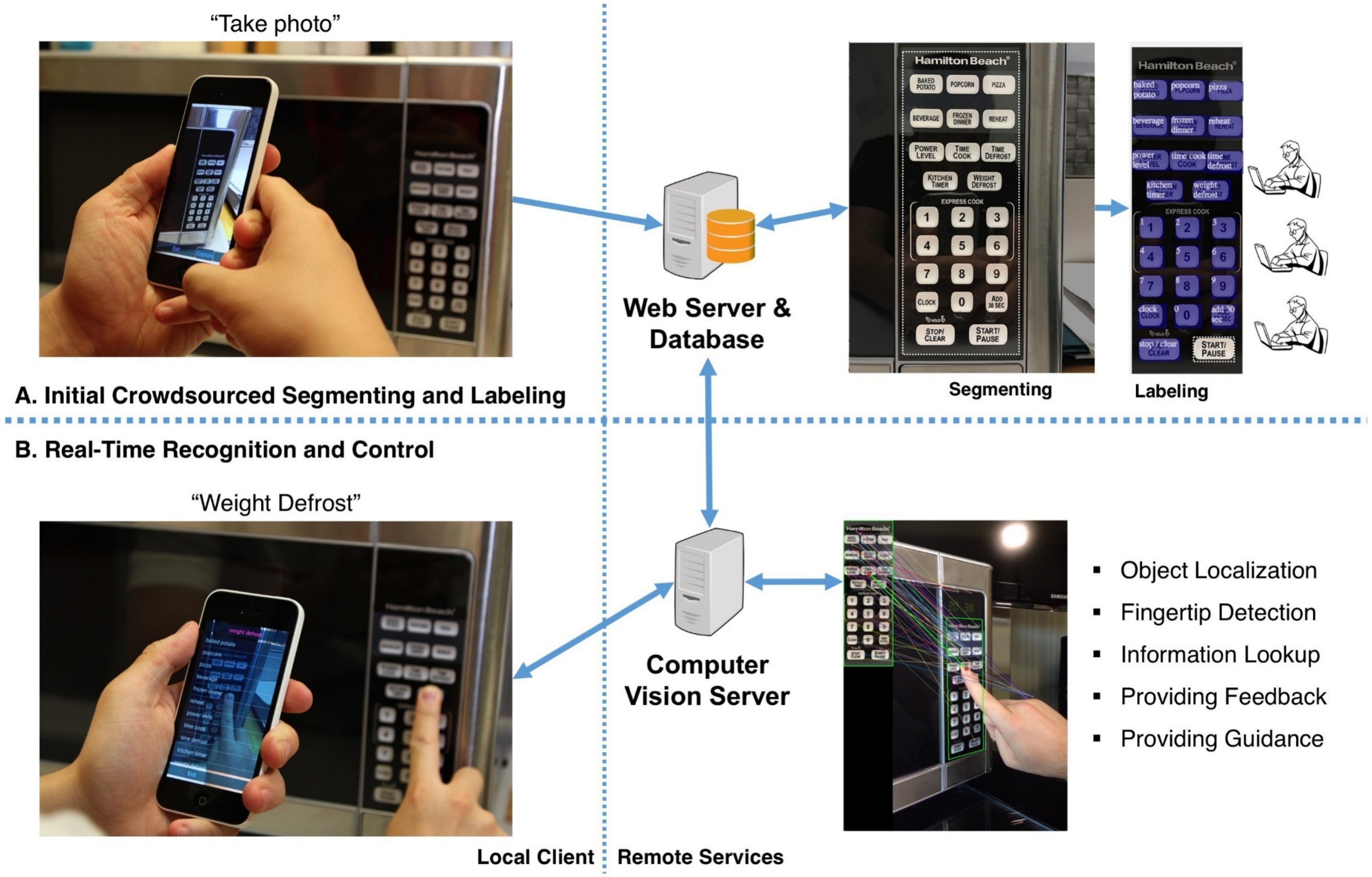


OmniTouch: C. Harrison, *et al.*, "OmniTouch: Wearable Multitouch Interaction Everywhere," in Proceedings of UIST 2011.

Related Work

VizLens: A. Guo, *et al.*, "VizLens: A Robust and Interactive Screen Reader for Interfaces in the Real World," in Proceedings of UIST 2016.

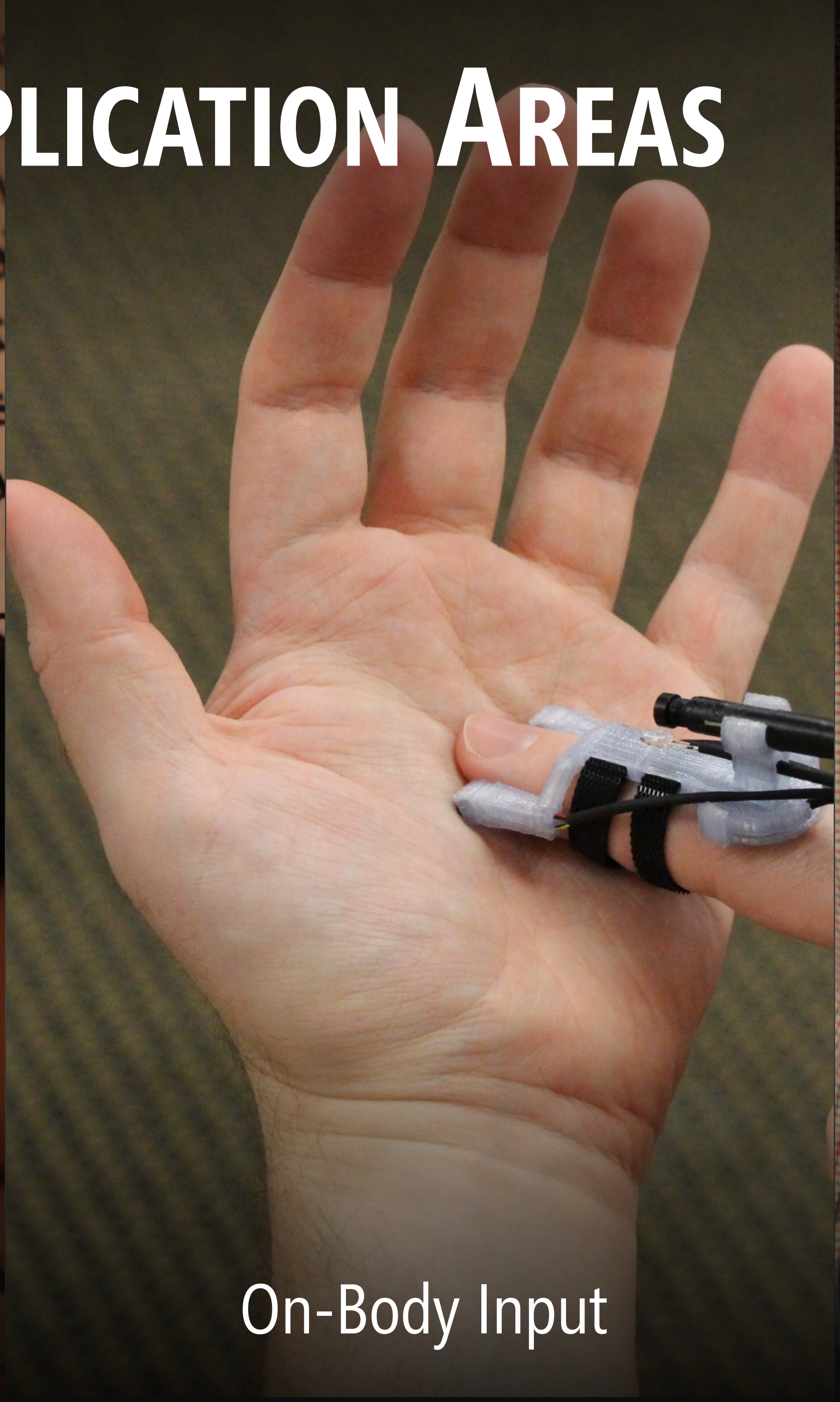
RELATED WORK



VizLens: A. Guo, *et al.*, "VizLens: A Robust and Interactive Screen Reader for Interfaces in the Real World," in Proceedings of UIST 2016.

HandSight Application Areas

Reading/Exploring Text



On-Body Input

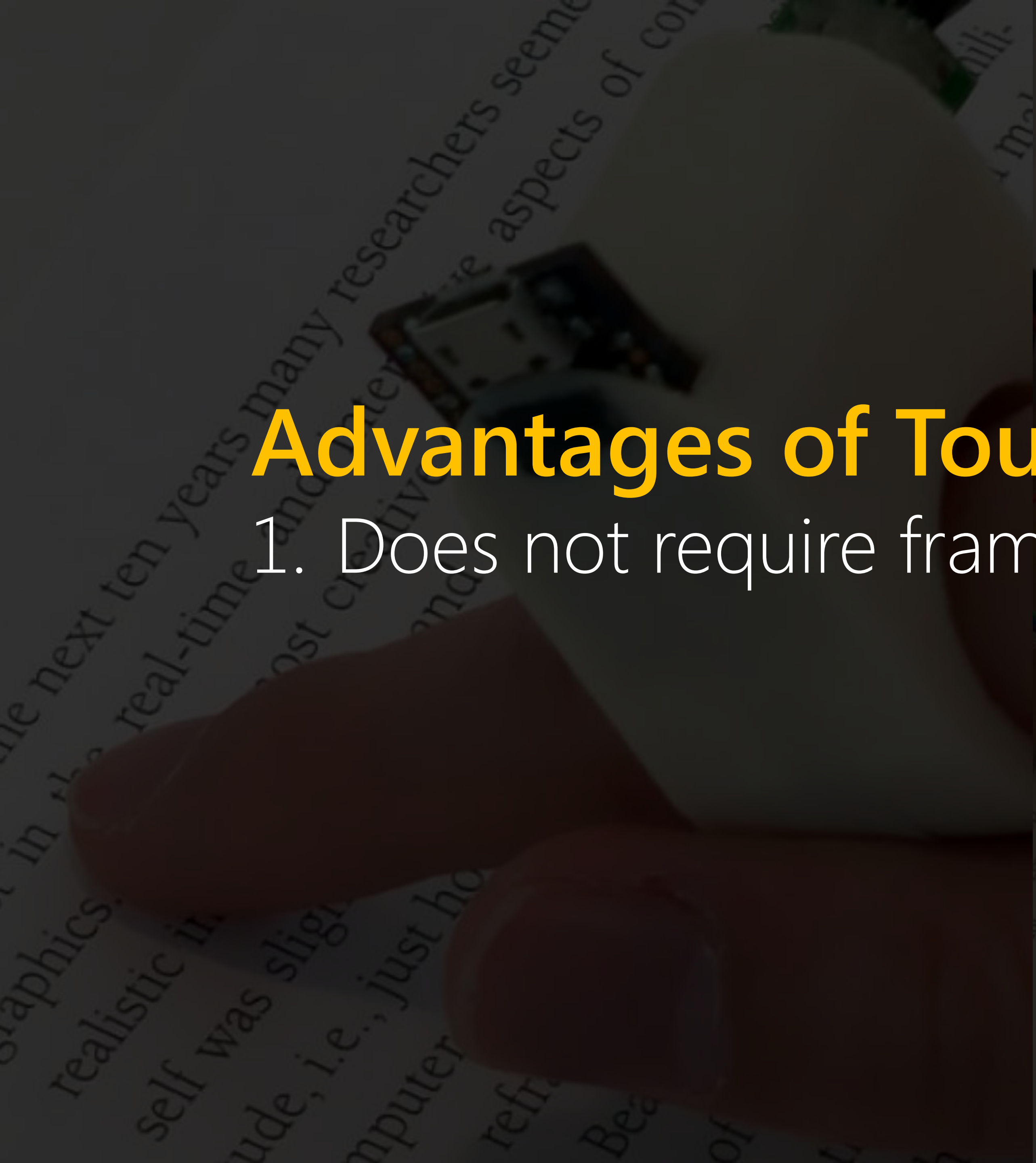
Clothing Colors/Patterns

HandSight Application Areas

Reading/Exploring Text

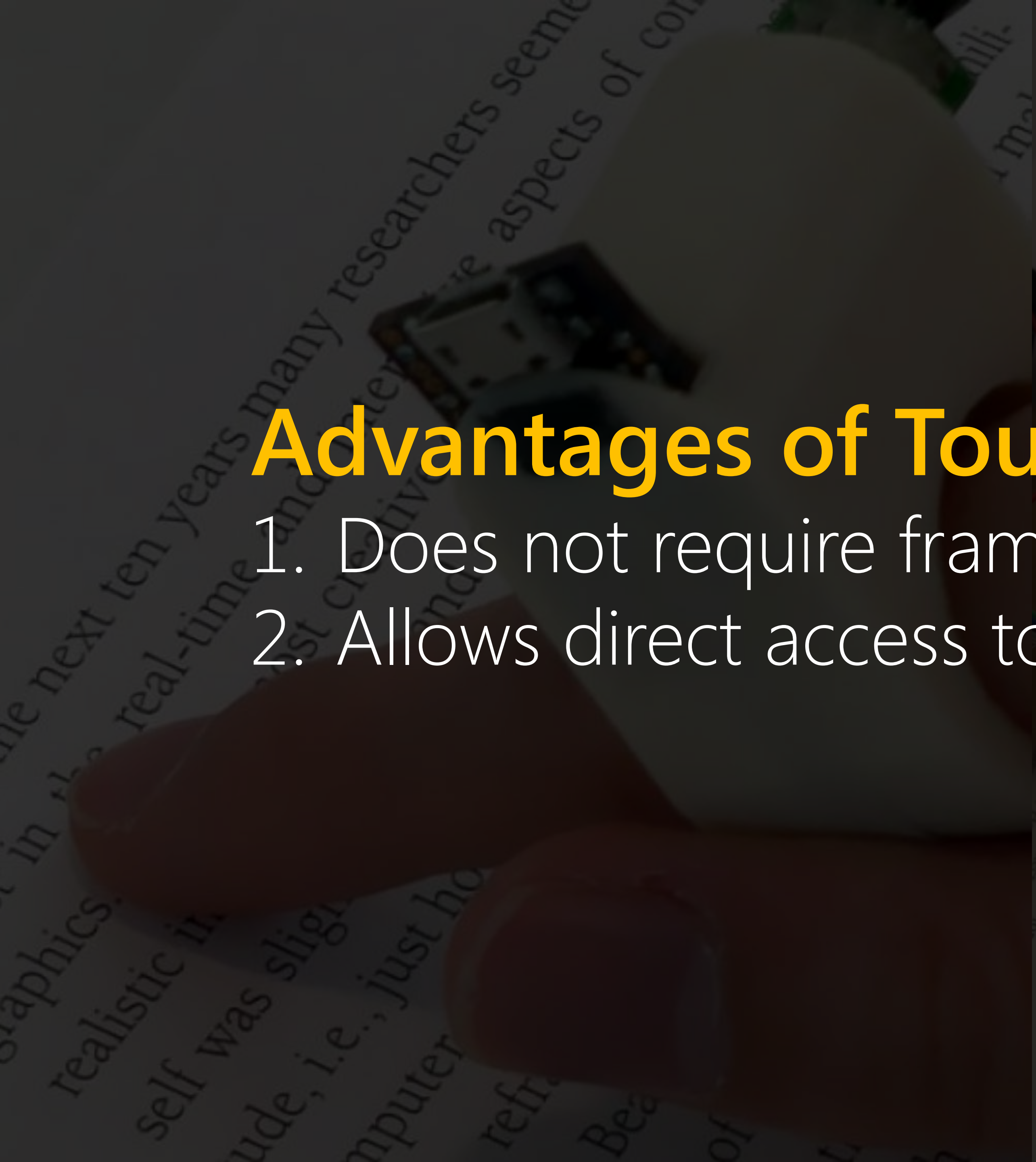
On-Body Input

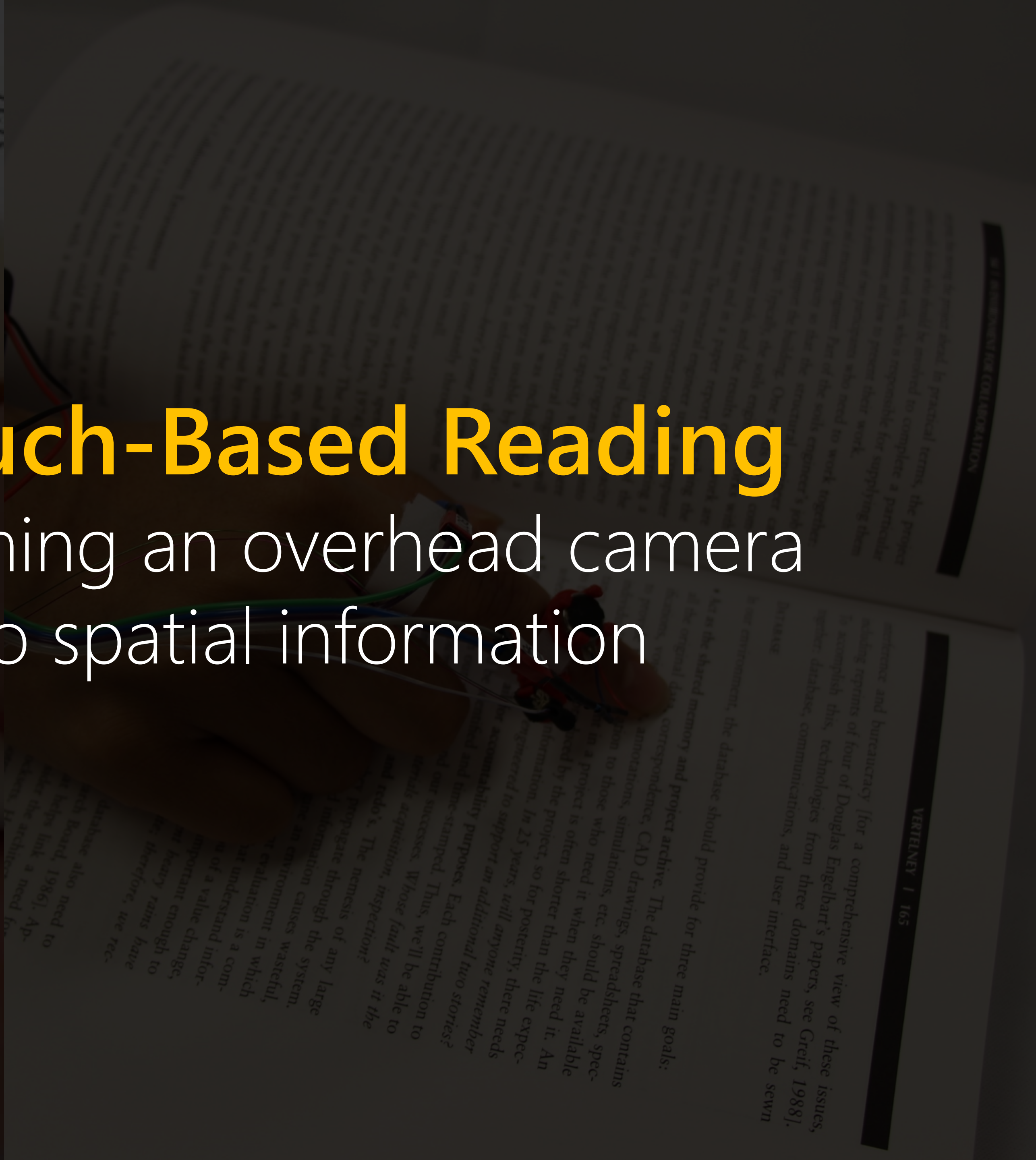
Clothing Colors/Patterns




Advantages of Touch-Based Reading

1. Does not require framing an overhead camera





Advantages of Touch-Based Reading

1. Does not require framing an overhead camera
2. Allows direct access to spatial information

Advantages of Touch-Based Reading

1. Does not require framing an overhead camera
2. Allows direct access to spatial information
3. Provides better control over pace and rereading

Advantages of Touch-Based Reading

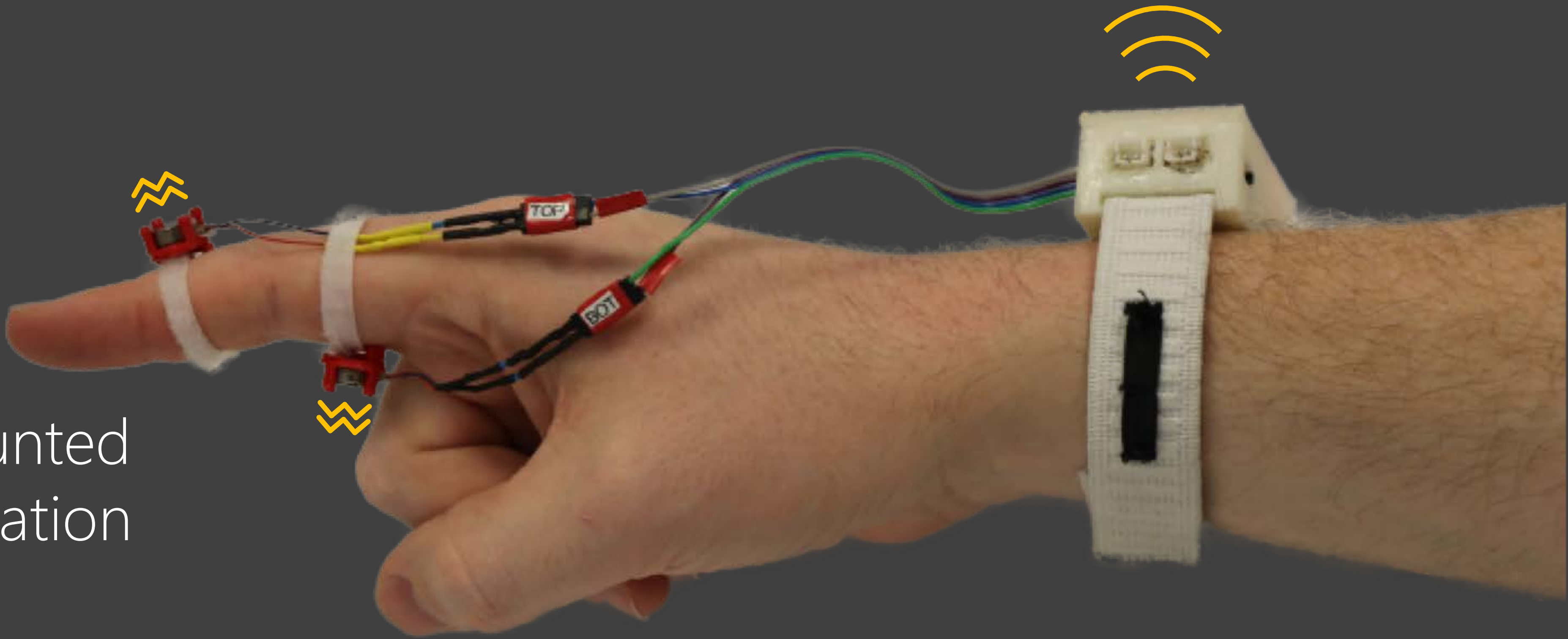
1. Does not require framing an overhead camera
2. Allows direct access to spatial information
3. Provides better control over pace and rereading

New Challenges

1. How to precisely trace a line of text?
2. How to support physical navigation?

Comparing Two Types of Directional Finger Guidance

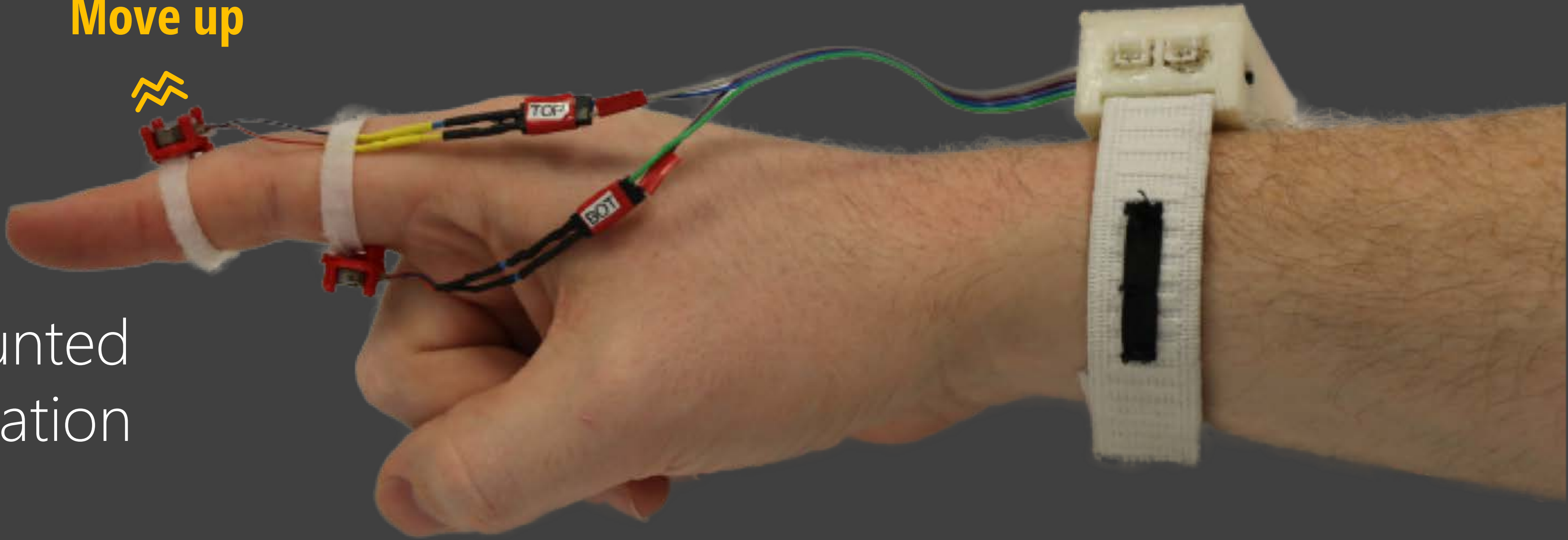
2. **Audio** via built-in or external speakers



1. Finger-mounted **haptic** vibration

Comparing Two Types of Directional Finger Guidance

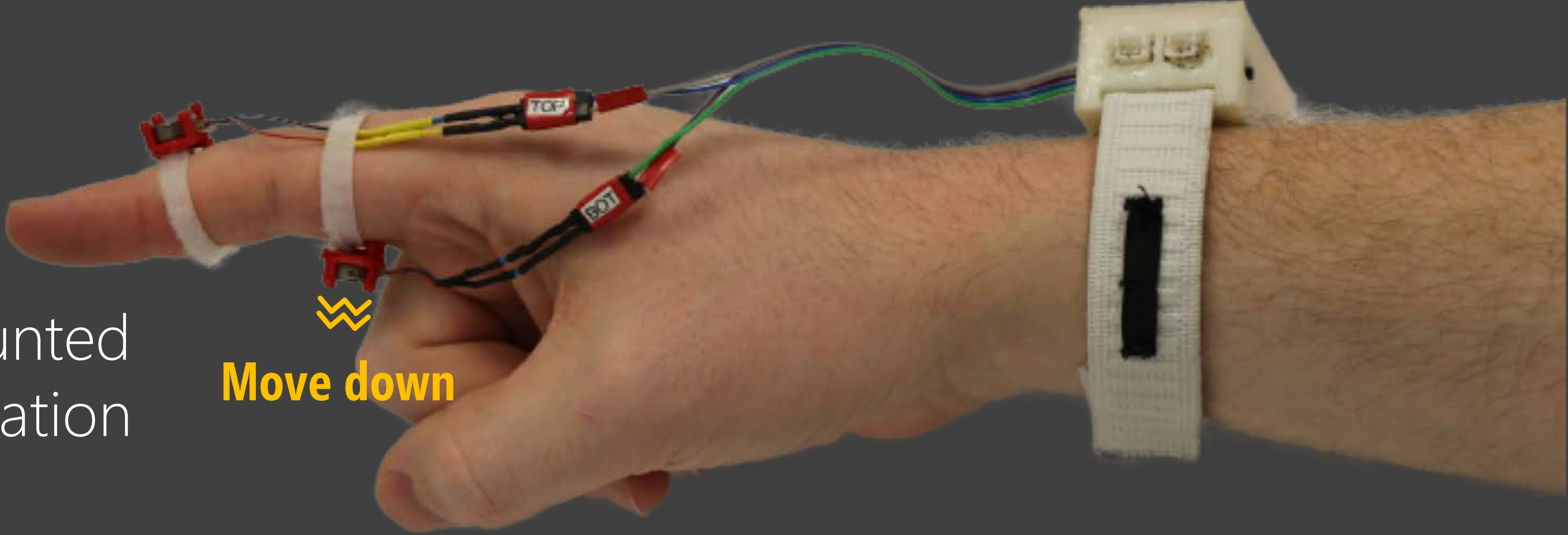
Move up



1. Finger-mounted **haptic** vibration

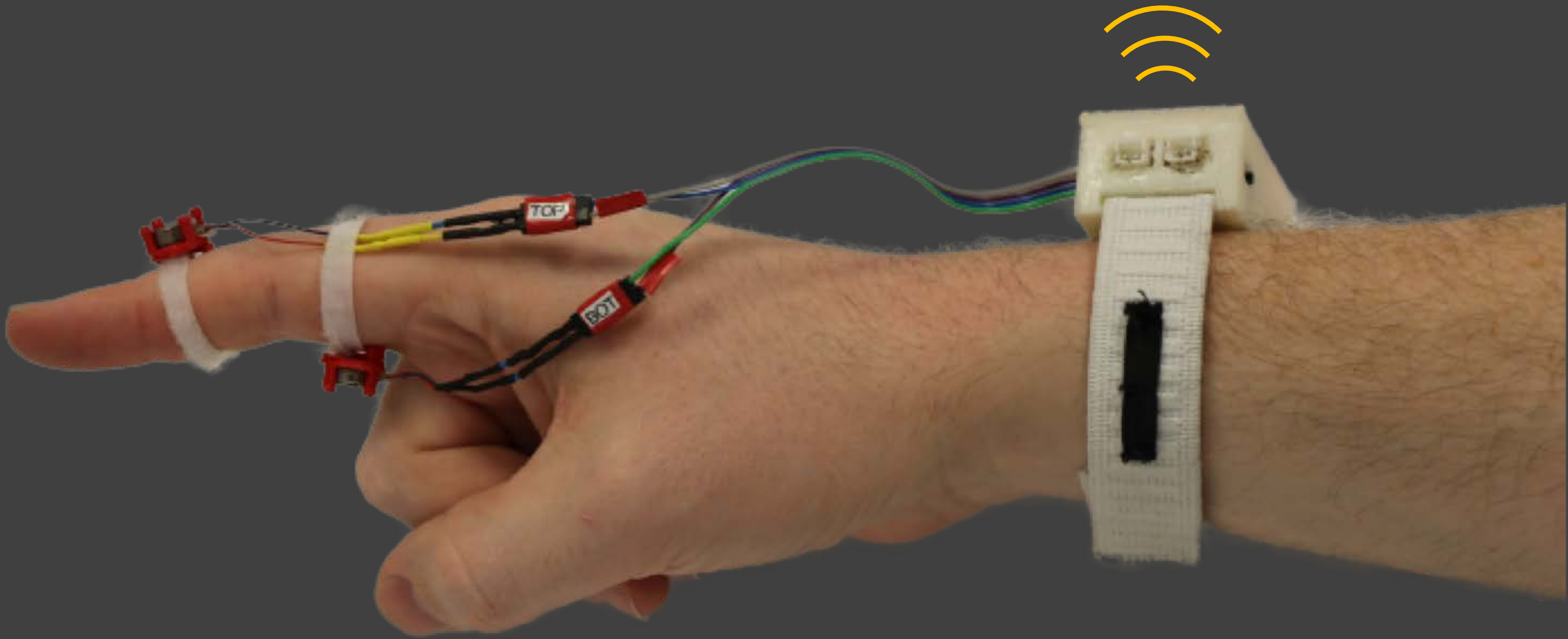
Comparing Two Types of Directional Finger Guidance

1. Finger-mounted **haptic** vibration



Comparing Two Types of Directional Finger Guidance

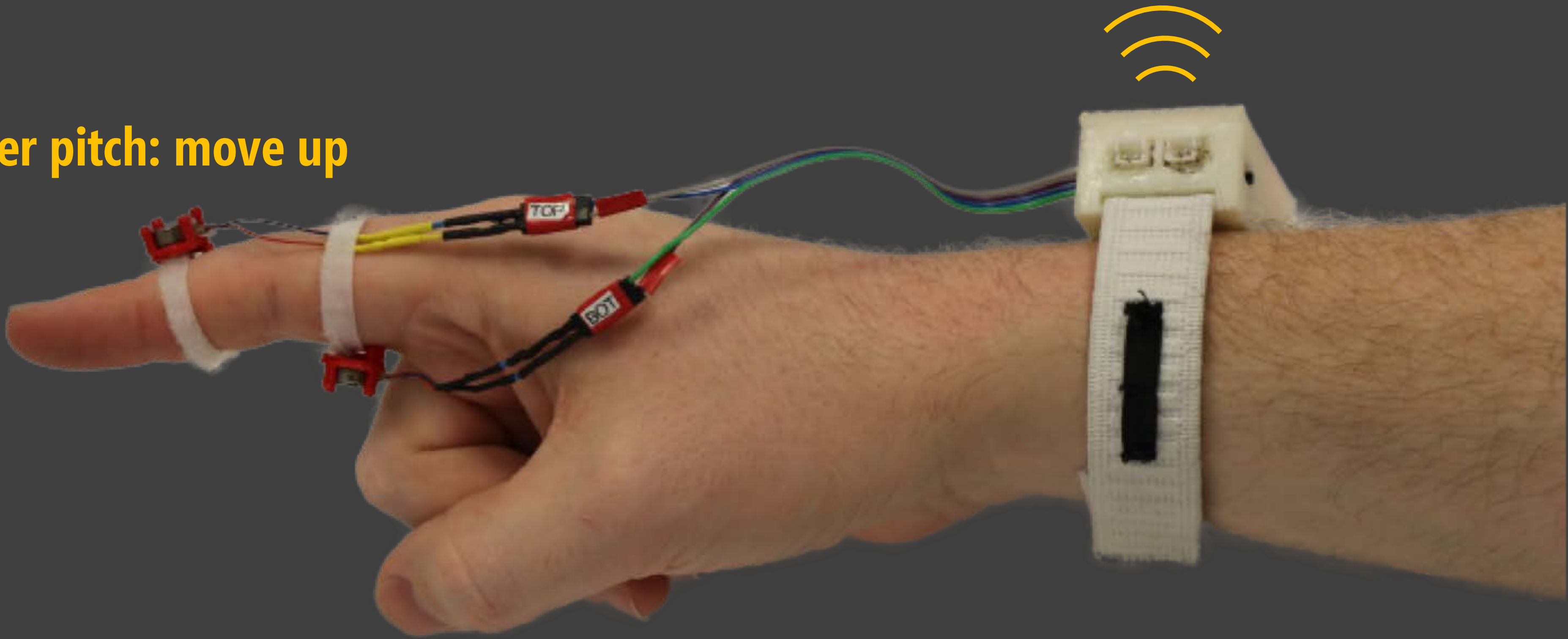
2. **Audio** via built-in or external speakers



Comparing Two Types of Directional Finger Guidance

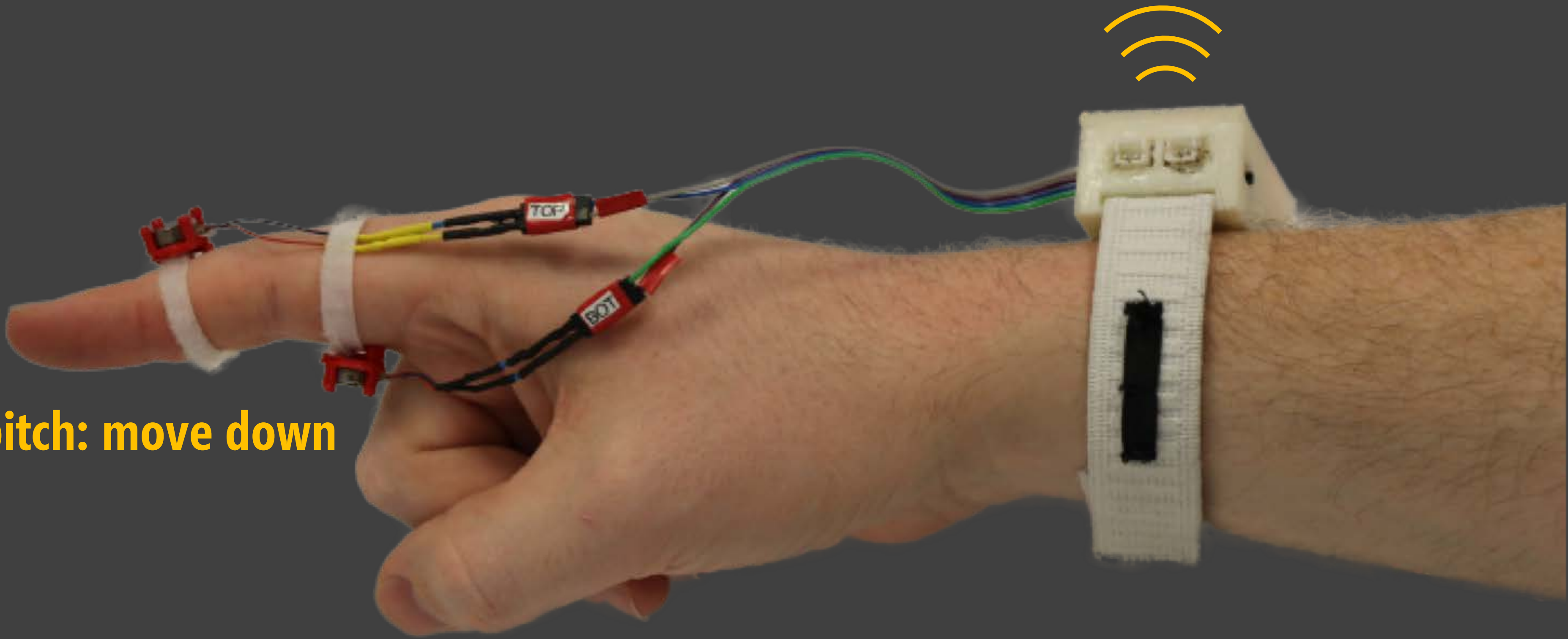
2. **Audio** via built-in or external speakers

Higher pitch: move up



Comparing Two Types of Directional Finger Guidance

2. **Audio** via built-in or external speakers

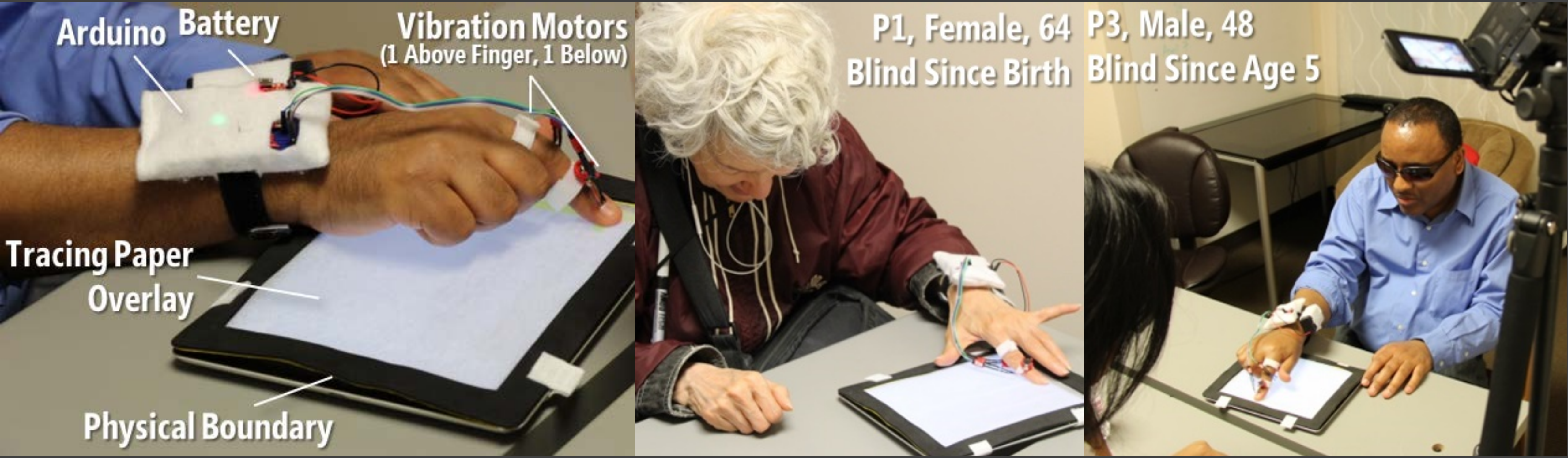


Reading Pilot Study
Methods

4 Participants

3 female, 1 male, ages 43-64

3 totally blind, 1 severe low vision



Lee Stearns, Ruofei Du, Uran Oh, Yumeng Wang, Leah Findlater, Rama Chellappa, and Jon E. Froehlich, **"The Design and Preliminary Evaluation of a Finger Mounted Camera and Feedback System to Enable Reading of Printed Text For the Blind,"** in *Proceedings of ECCV 2014 (ACVR Workshop)*.

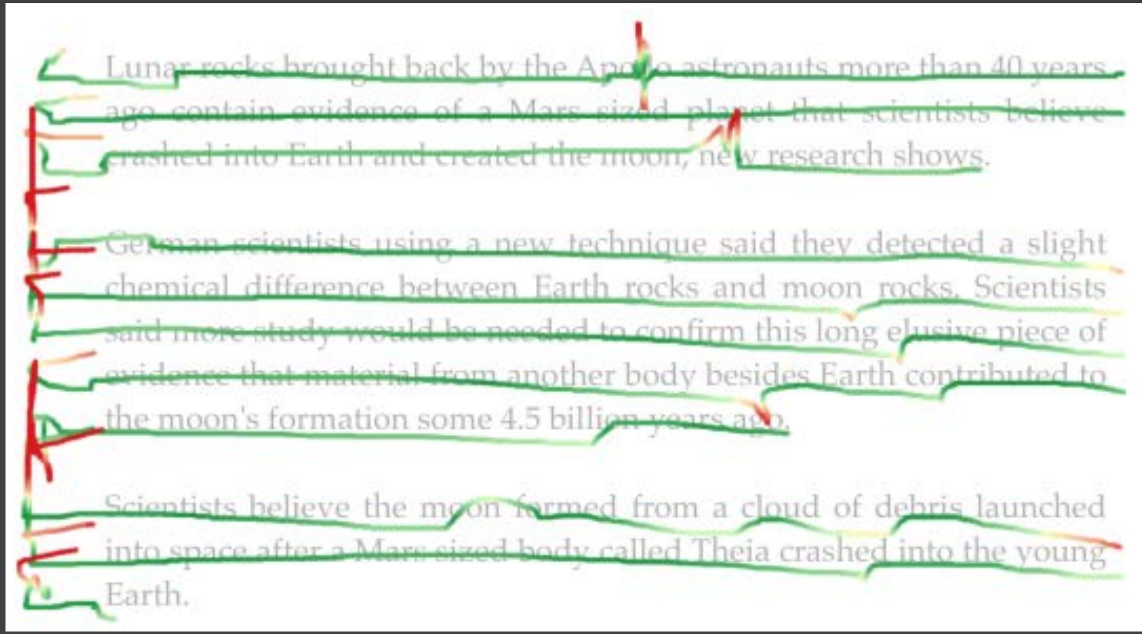
READING PILOT STUDY FINDINGS

4 Participants

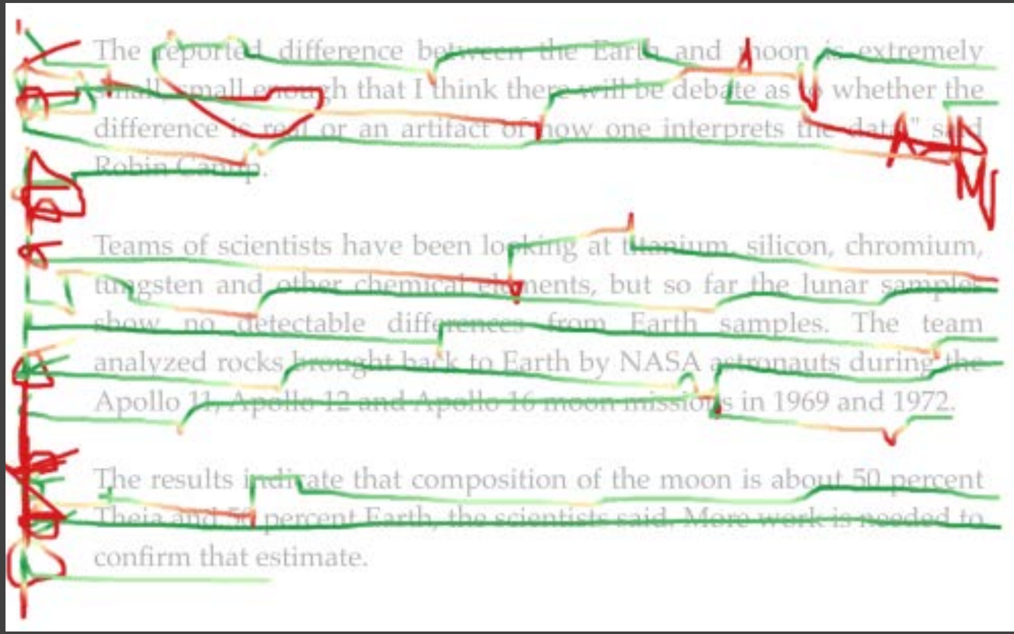
3 female, 1 male, ages 43-64

3 totally blind, 1 severe low vision

audio-only



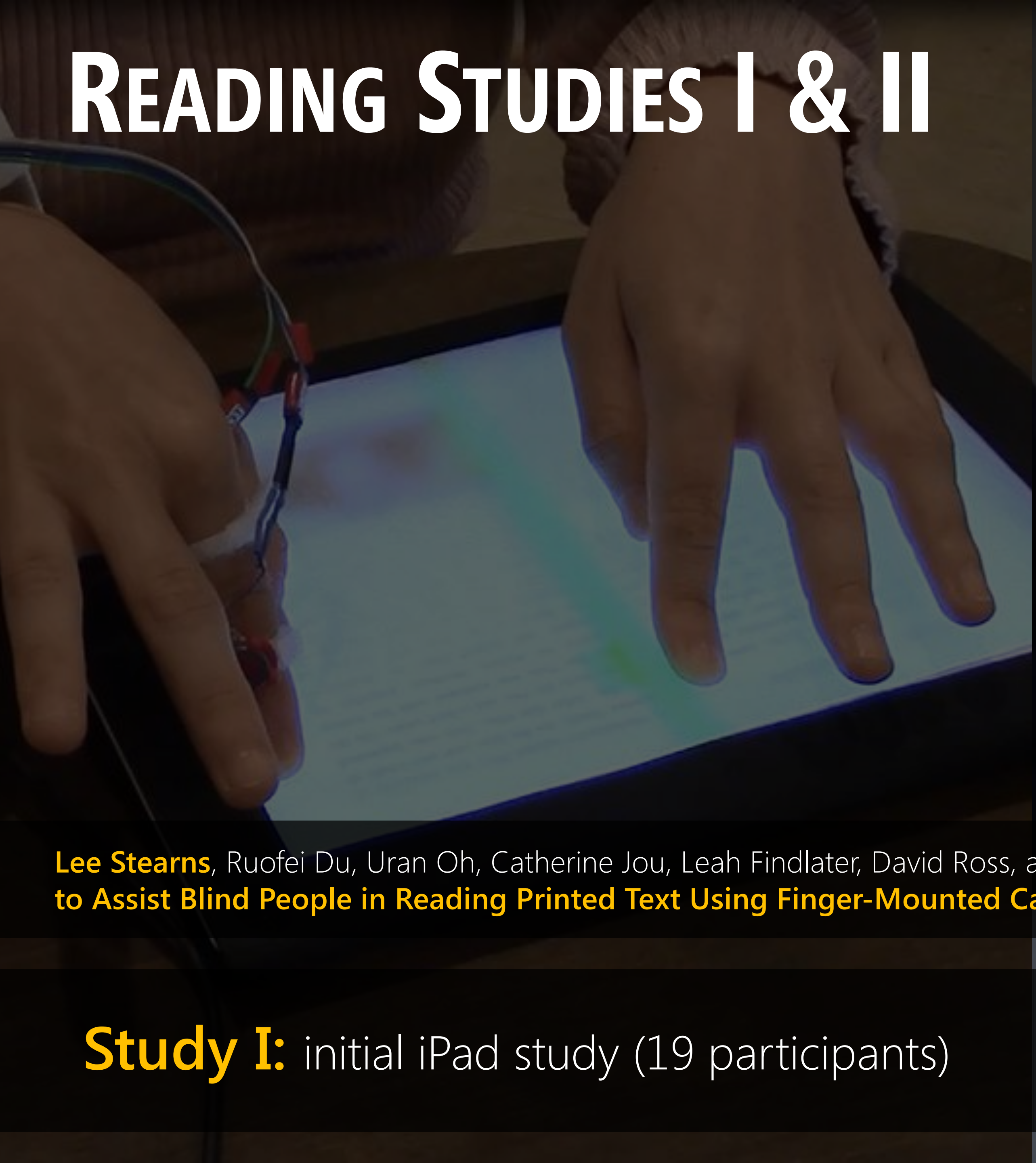
haptic-only

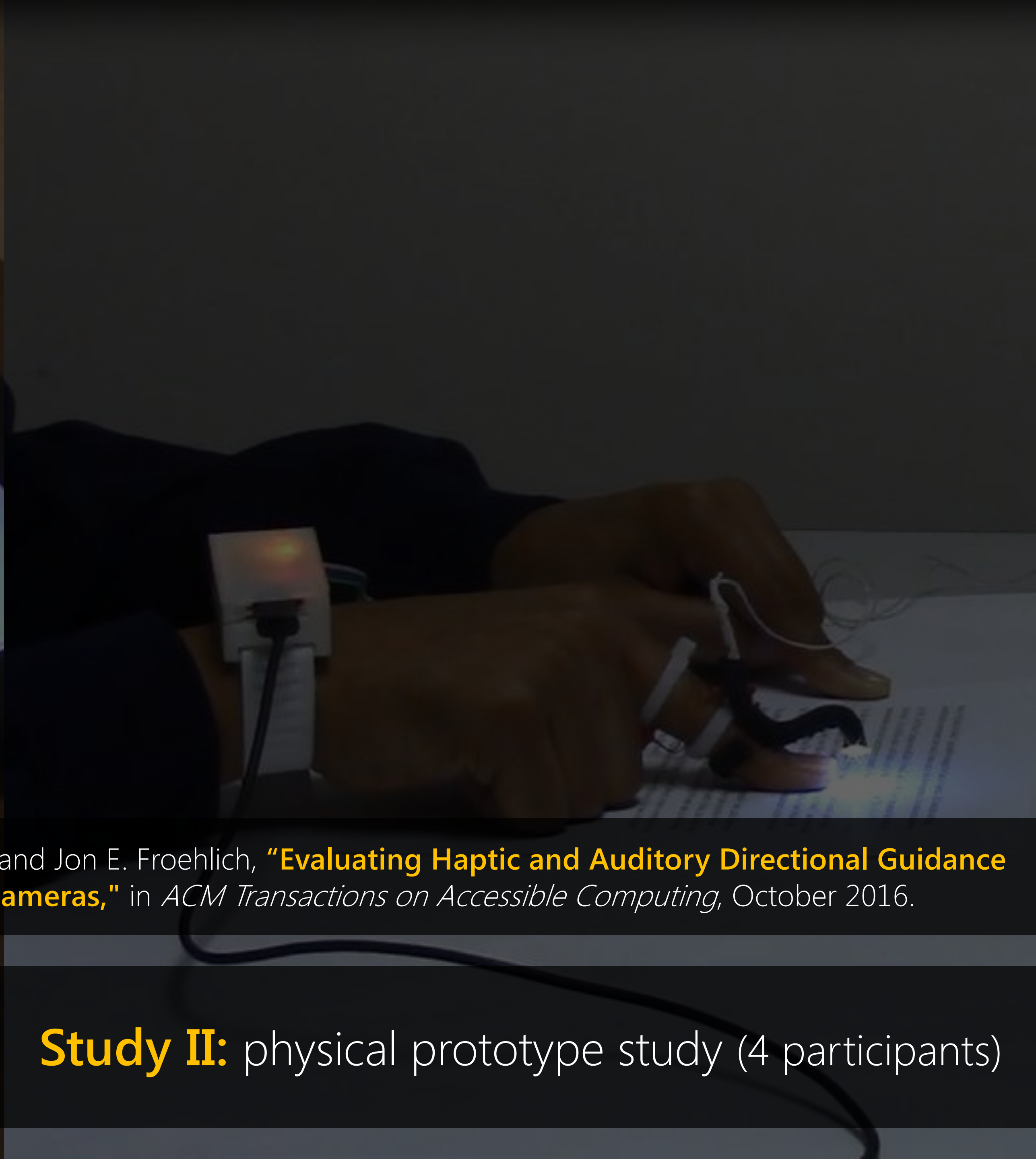


Example finger traces—**Red lines** mark drift off of the line

Audio was more accurate, about twice as fast, and preferred by 3 out of 4 participants

Reading Studies I & II

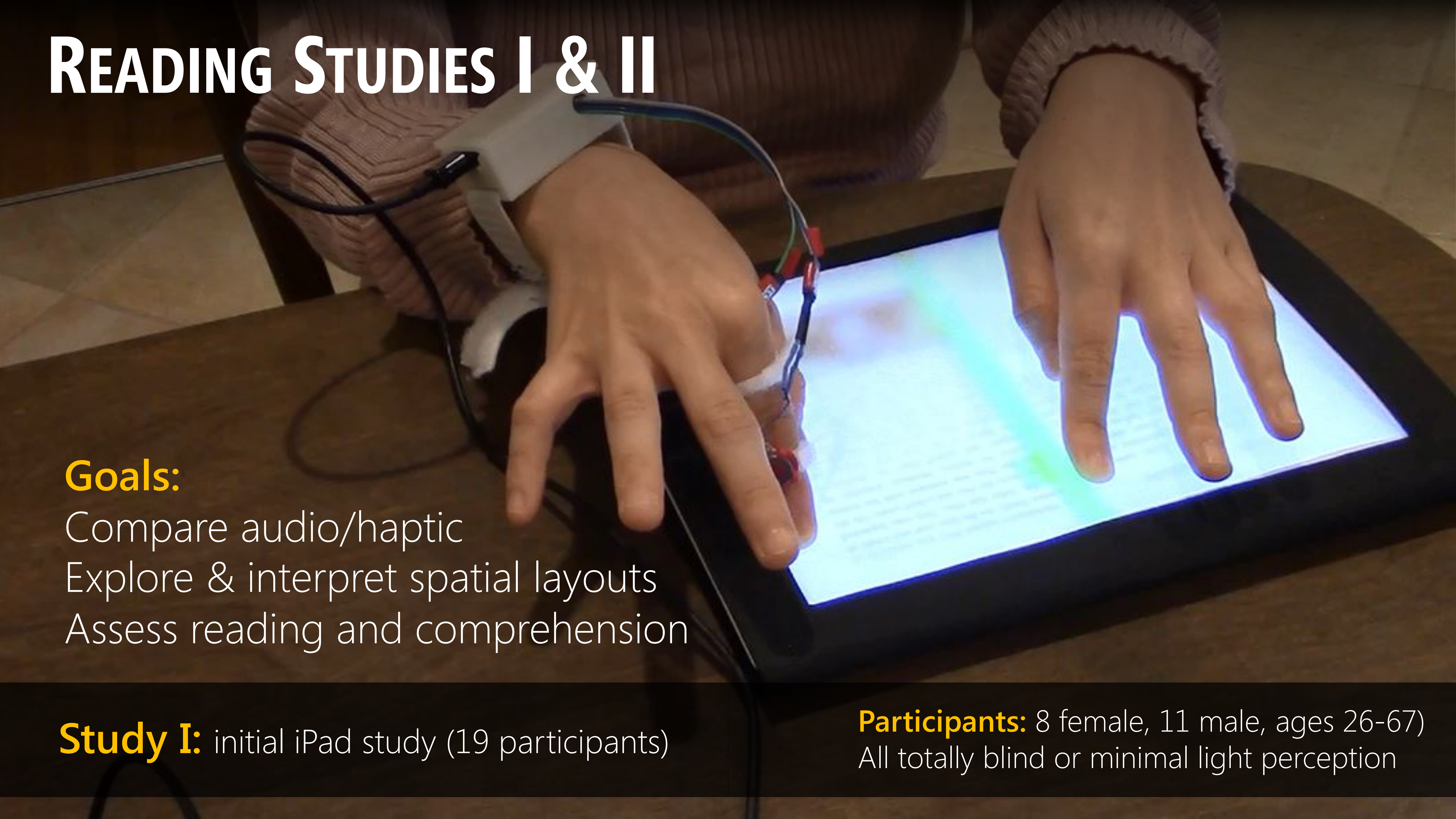




Lee Stearns, Ruofei Du, Uran Oh, Catherine Jou, Leah Findlater, David Ross, and Jon E. Froehlich, **"Evaluating Haptic and Auditory Directional Guidance to Assist Blind People in Reading Printed Text Using Finger-Mounted Cameras,"** in *ACM Transactions on Accessible Computing*, October 2016.

Study I: initial iPad study (19 participants)

Study II: physical prototype study (4 participants)



Reading Studies I & II

Goals:

Compare audio/haptic
Explore & interpret spatial layouts
Assess reading and comprehension

Study I: initial iPad study (19 participants)

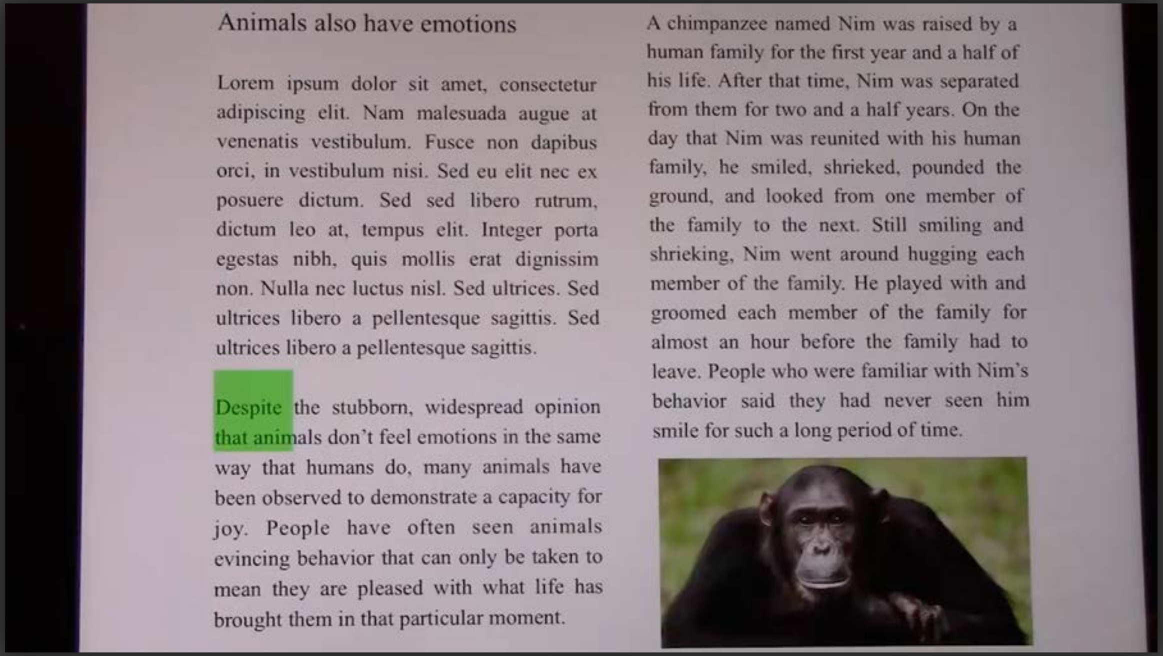
Participants: 8 female, 11 male, ages 26-67)
All totally blind or minimal light perception

Reading Study I

Used an iPad to focus on **user experience**, gather **finger trace** data

System Design

Exploration and Reading Modes

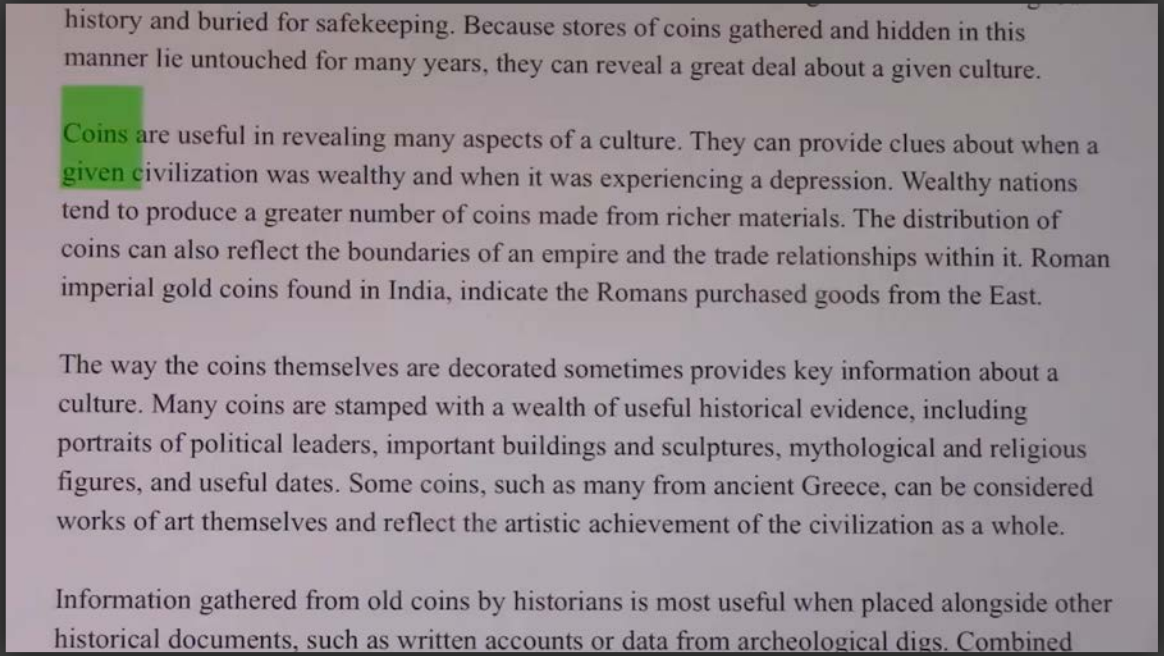

Animals also have emotions

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam malesuada augue at venenatis vestibulum. Fusce non dapibus orci, in vestibulum nisi. Sed eu elit nec ex posuere dictum. Sed sed libero rutrum, dictum leo at, tempus elit. Integer porta egestas nibh, quis mollis erat dignissim non. Nulla nec luctus nisl. Sed ultrices. Sed ultrices libero a pellentesque sagittis. Sed ultrices libero a pellentesque sagittis.

Despite the stubborn, widespread opinion that animals don't feel emotions in the same way that humans do, many animals have been observed to demonstrate a capacity for joy. People have often seen animals evincing behavior that can only be taken to mean they are pleased with what life has brought them in that particular moment.

A chimpanzee named Nim was raised by a human family for the first year and a half of his life. After that time, Nim was separated from them for two and a half years. On the day that Nim was reunited with his human family, he smiled, shrieked, pounded the ground, and looked from one member of the family to the next. Still smiling and shrieking, Nim went around hugging each member of the family. He played with and groomed each member of the family for almost an hour before the family had to leave. People who were familiar with Nim's behavior said they had never seen him smile for such a long period of time.

Exploration Mode


history and buried for safekeeping. Because stores of coins gathered and hidden in this manner lie untouched for many years, they can reveal a great deal about a given culture.

Coins are useful in revealing many aspects of a culture. They can provide clues about when a given civilization was wealthy and when it was experiencing a depression. Wealthy nations tend to produce a greater number of coins made from richer materials. The distribution of coins can also reflect the boundaries of an empire and the trade relationships within it. Roman imperial gold coins found in India, indicate the Romans purchased goods from the East.

The way the coins themselves are decorated sometimes provides key information about a culture. Many coins are stamped with a wealth of useful historical evidence, including portraits of political leaders, important buildings and sculptures, mythological and religious figures, and useful dates. Some coins, such as many from ancient Greece, can be considered works of art themselves and reflect the artistic achievement of the civilization as a whole.

Information gathered from old coins by historians is most useful when placed alongside other historical documents, such as written accounts or data from archeological digs. Combined

Reading Mode

System Design

Exploration Mode

Continuous audio feedback to identify content beneath finger

- **Flute sound**: text
- **Cello sound**: picture
- **Silence**: empty space

Animals also have emotions

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam malesuada augue at venenatis vestibulum. Fusce non dapibus orci, in vestibulum nisi. Sed eu elit nec ex posuere dictum. Sed sed libero rutrum, dictum leo at, tempus elit. Integer porta egestas nibh, quis mollis erat dignissim non. Nulla nec luctus nisl. Sed ultrices. Sed ultrices libero a pellentesque sagittis. Sed ultrices libero a pellentesque sagittis.

Despite the stubborn, widespread opinion that animals don't feel emotions in the same way that humans do, many animals have been observed to demonstrate a capacity for joy. People have often seen animals evincing behavior that can only be taken to mean they are pleased with what life has brought them in that particular moment.

Flute sound: text

Silence: empty space

A chimpanzee named Nim was raised by a human family for the first year and a half of his life. After that time, Nim was separated from them for two and a half years. On the day that Nim was reunited with his human family, he smiled, shrieked, pounded the ground, and looked from one member of the family to the next. Still smiling and shrieking, Nim went around hugging each member of the family. He played with and groomed each member of the family for almost an hour before the family had to leave. People who were familiar with Nim's behavior said they had never seen him smile for such a long period of time.

Cello sound: picture

System Design

Reading Mode

Right index finger to read, **left** to anchor start of line

Two directional guidance conditions: **audio** or **haptic**

- Used to stay on the line or find the start of the next line
- **Audio:** pitch of continuous audio
- **Haptic:** strength and position of vibration

Additional **audio cues** (same for both conditions)

- Start/end of line or paragraph
- Synthesized speech

history and buried for safekeeping. Be
this
manner lie untouched for many years,
culture.

Above the line: downward guidance
(low pitch or lower vibration motor)

Coins are useful in revealing many aspects of a culture. They can provide clues about when a given civilization was wealthy and when it was experiencing a depression. Wealthy nations
tend
e from richer m
coins
pire and the tra
imperial gold coins found in India, indicate the Romans purchased goods from the East.

Below the line: upward guidance
(high pitch or upper vibration motor)

Start/end of line or paragraph
(short but distinctive audio cues)

The way the coins themselves are decorated sometimes provides key information about a culture. Many coins are stamped with a wealth of useful historical evidence, including portraits of political leaders, important buildings and sculptures, mythological and religious figures, and useful dates. Some coins, such as many from ancient Greece, can be considered works of art themselves and reflect the artistic achievement of the civilization as a whole.

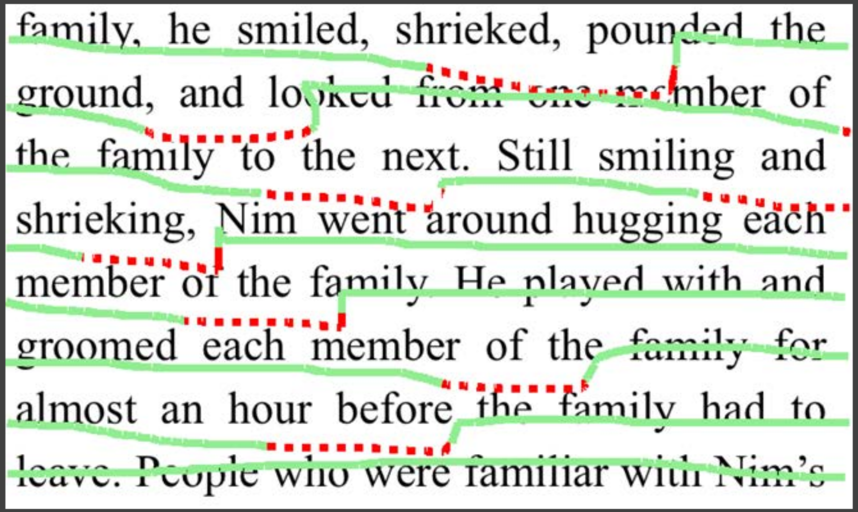
Information gathered from old coins by historians is most useful when placed alongside other historical documents, such as written accounts or data from archeological digs. Combined

Study I

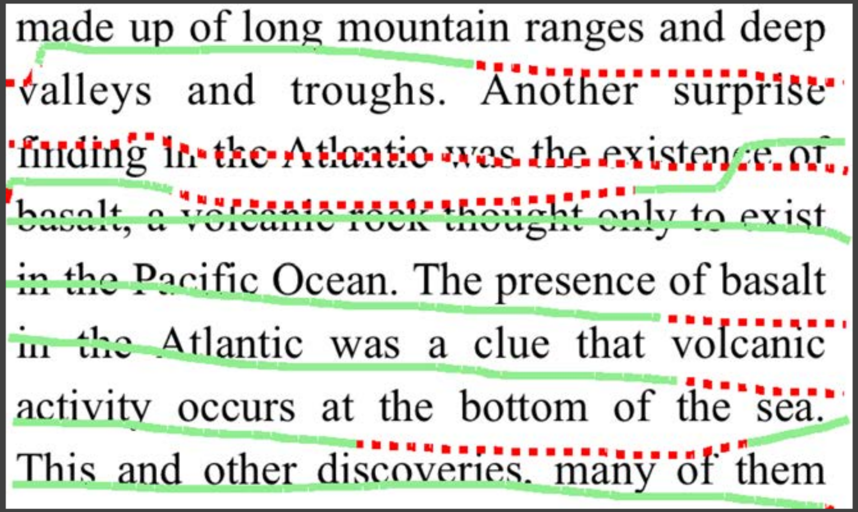
Findings

Haptic vs. Audio: Quantitative Performance (n=19)

audio



haptic

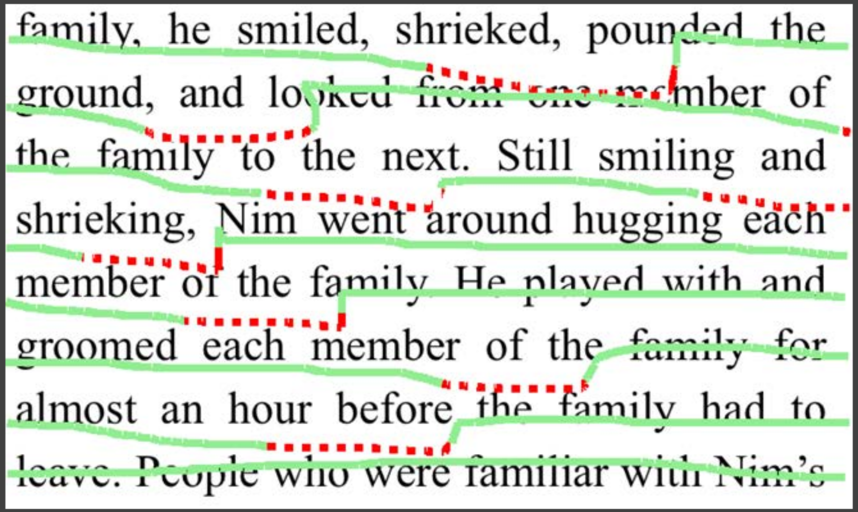


Example finger traces—**Dashed red lines** mark drift off of the line

STUDY I
FINDINGS

Haptic vs. Audio: Quantitative Performance (n=19)

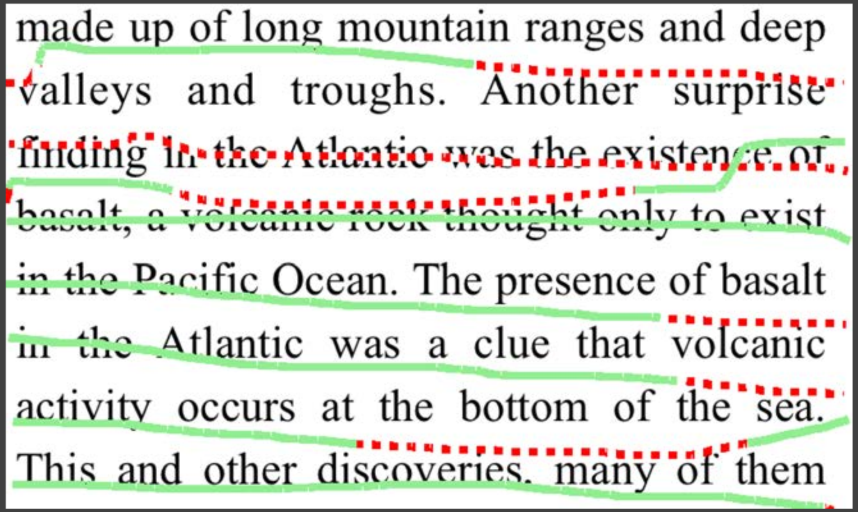
audio



0.21cm

Absolute error from line center

haptic



0.28cm

Absolute error from line center

Z_{19}=-2.374, *p*=.018, *r*=.54

Study I

Findings

Overall Reading Experience

Pros	
Low learning curve	
Flexible	
Direct control over speed	

STUDY I

FINDINGS

Overall Reading Experience

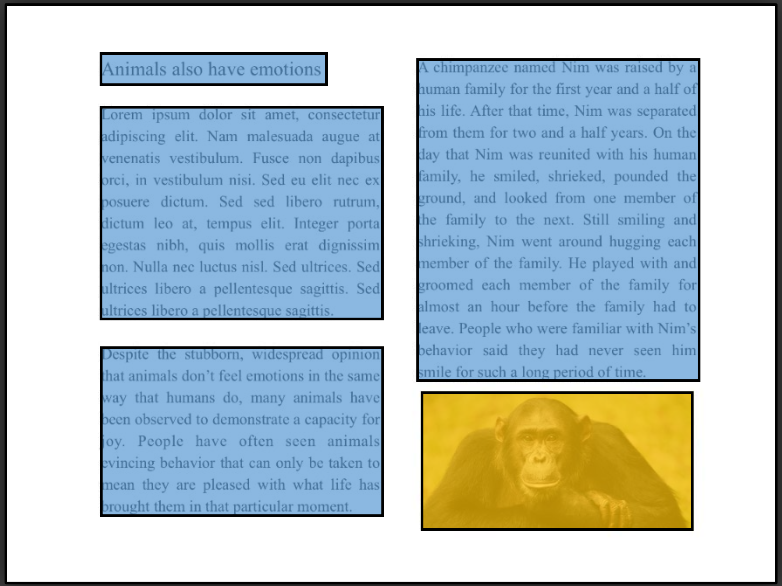
Pros	Cons
Low learning curve	Hard to use for reading
Flexible	High cognitive load may affect comprehension
Direct control over speed	

STUDY I

FINDINGS

Exploration Mode

Participants appreciated direct access to spatial information, and nearly all able to locate images and count the number of columns.

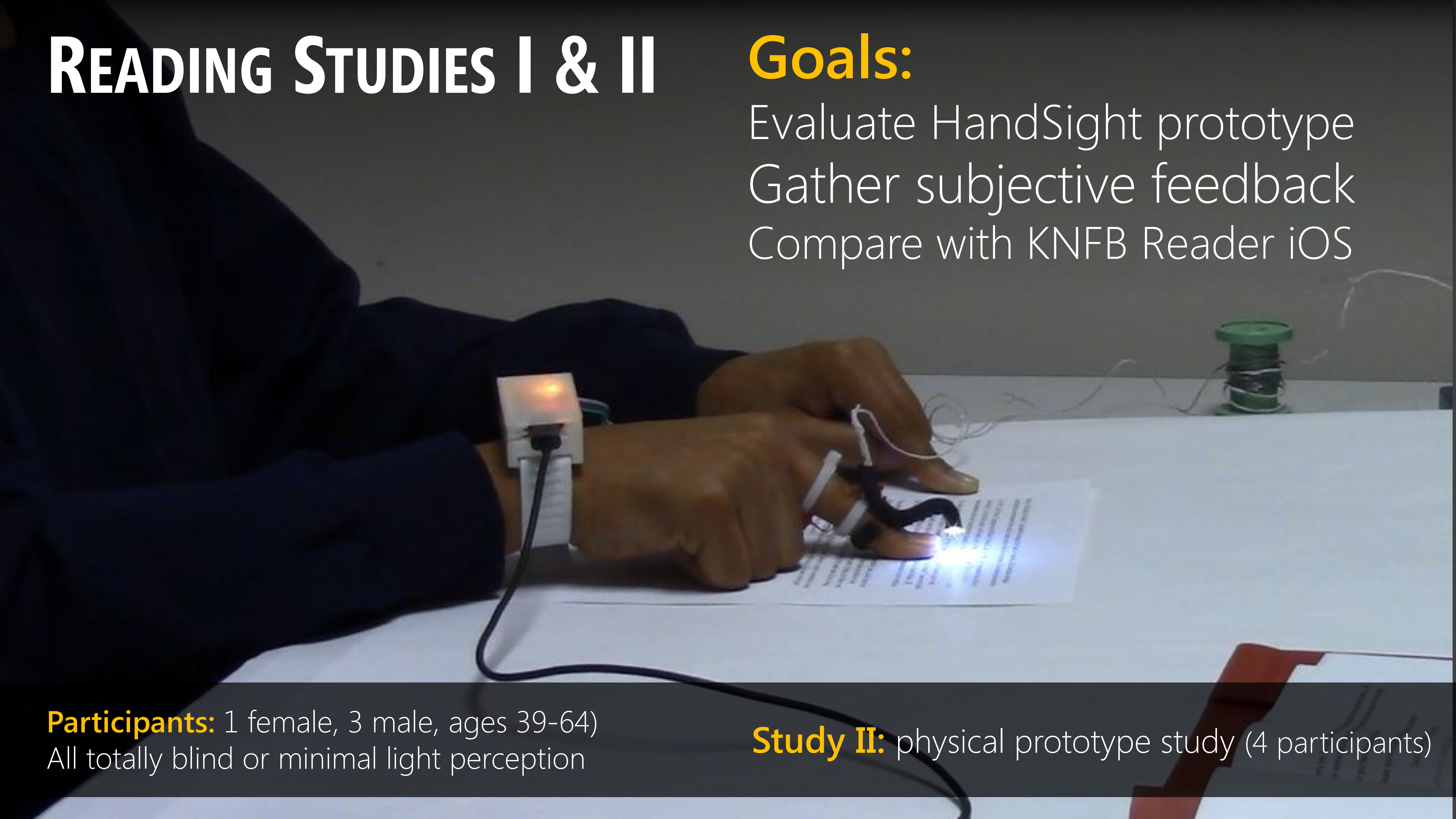


READING STUDIES I & II

Study I: initial iPad study (19 participants)

Study II: physical prototype study (4 participants)

Reading Studies I & II



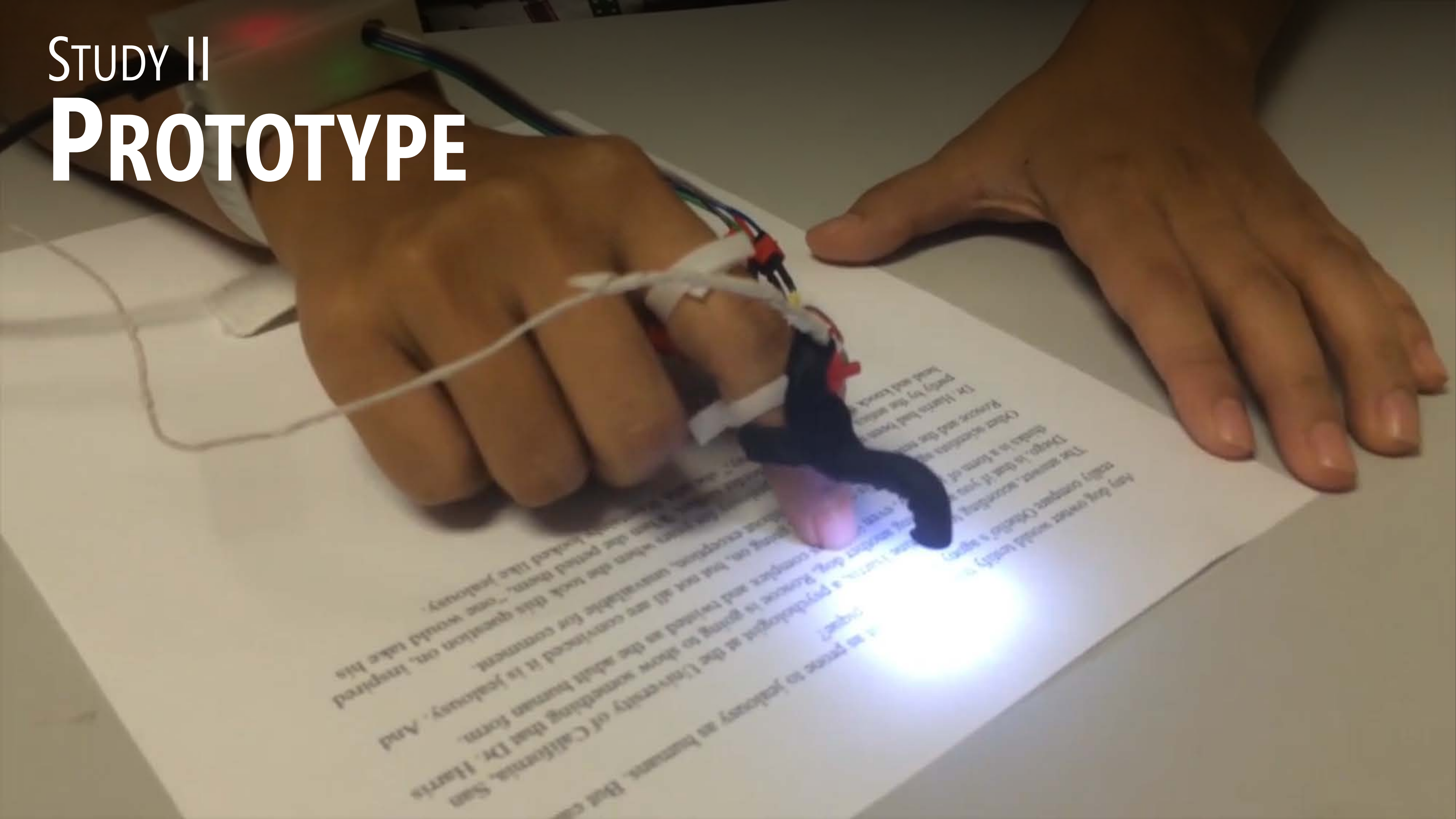
Goals:

Evaluate HandSight prototype
Gather subjective feedback
Compare with KNFB Reader iOS

Participants: 1 female, 3 male, ages 39-64)
All totally blind or minimal light perception

Study II: physical prototype study (4 participants)

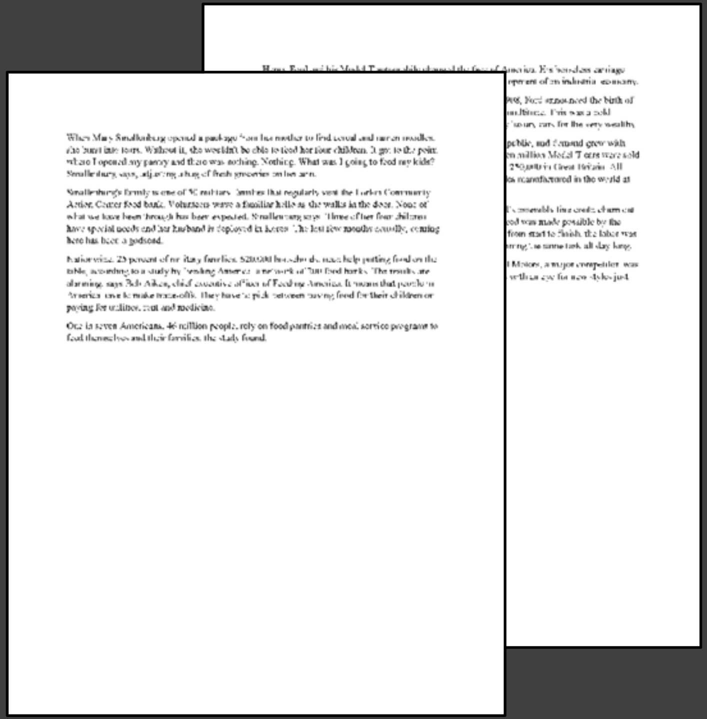
Study II
Prototype



Study II

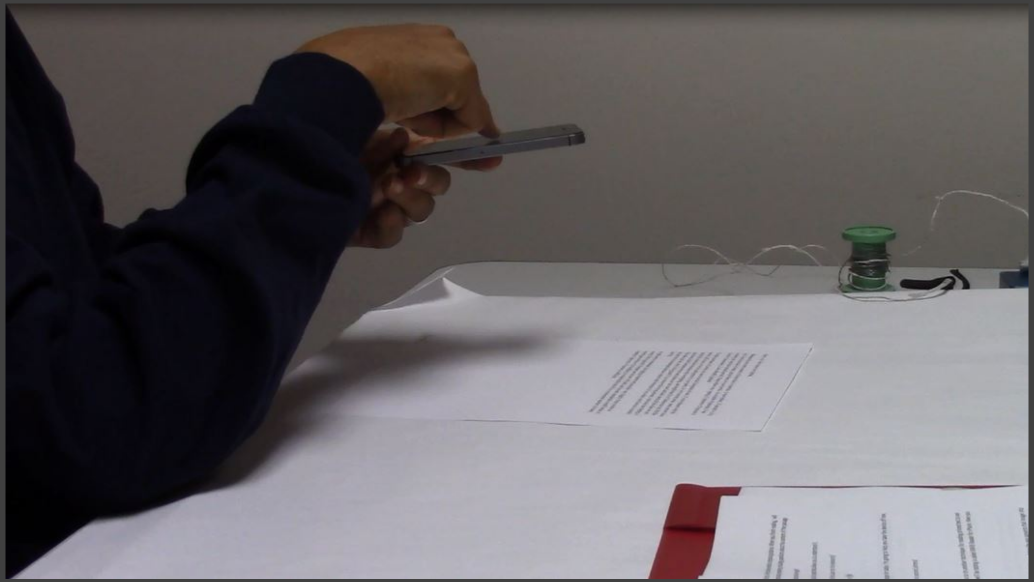
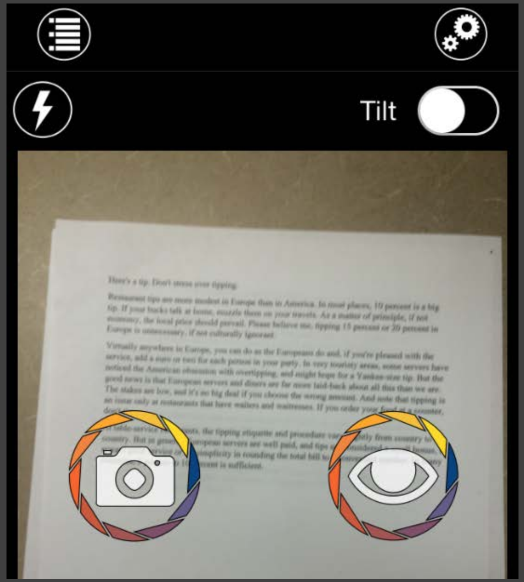
Method: HandSight

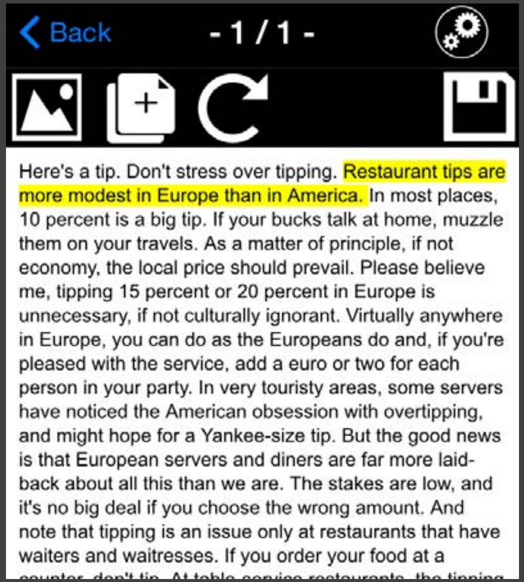
Participants used **preferred guidance** from Study I to explore and read physical documents



Study II
Method: KNFB Reader iOS

Photographed and read physical documents



IMPLICATIONS

ADVANTAGES AND DISADVANTAGES OF A FINGER-BASED READING APPROACH

Pros

Implications

Advantages and Disadvantages of a Finger-Based Reading Approach

Pros	
Spatial layout information	

Implications

Advantages and Disadvantages of a Finger-Based Reading Approach

Pros	
Spatial layout information	
Direct control over reading	

Implications

Advantages and Disadvantages of a Finger-Based Reading Approach

Pros	
Spatial layout information	
Direct control over reading	
Reduced camera framing issues	

Implications

Advantages and Disadvantages of a Finger-Based Reading Approach

Pros	
Spatial layout information	
Direct control over reading	
Reduced camera framing issues	
Efficient text detection and recognition	

Implications

Advantages and Disadvantages of a Finger-Based Reading Approach

Pros	
Spatial layout information	
Direct control over reading	
Reduced camera framing issues	
Efficient text detection and recognition	
* We observed these in our studies	

Implications

Advantages and Disadvantages of a Finger-Based Reading Approach

Pros	Cons
Spatial layout information	Slower, requires increased concentration and physical dexterity
Direct control over reading	
Reduced camera framing issues	
Efficient text detection and recognition	

* We observed these in our studies

Implications

Advantages and Disadvantages of a Finger-Based Reading Approach

Pros	Cons
Spatial layout information	Slower, requires increased concentration and physical dexterity
Direct control over reading	* Consistent with Shilkrot *et al.* 2014, 2015
Reduced camera framing issues	
Efficient text detection and recognition	
* We observed these in our studies	

Implications

Advantages and Disadvantages of a Finger-Based Reading Approach

Pros	Cons
Spatial layout information	Slower, requires increased concentration and physical dexterity * Consistent with Shilkrot *et al.* 2014, 2015
Direct control over reading	Importance of spatial layout information is unclear
Reduced camera framing issues	
Efficient text detection and recognition * We observed these in our studies	

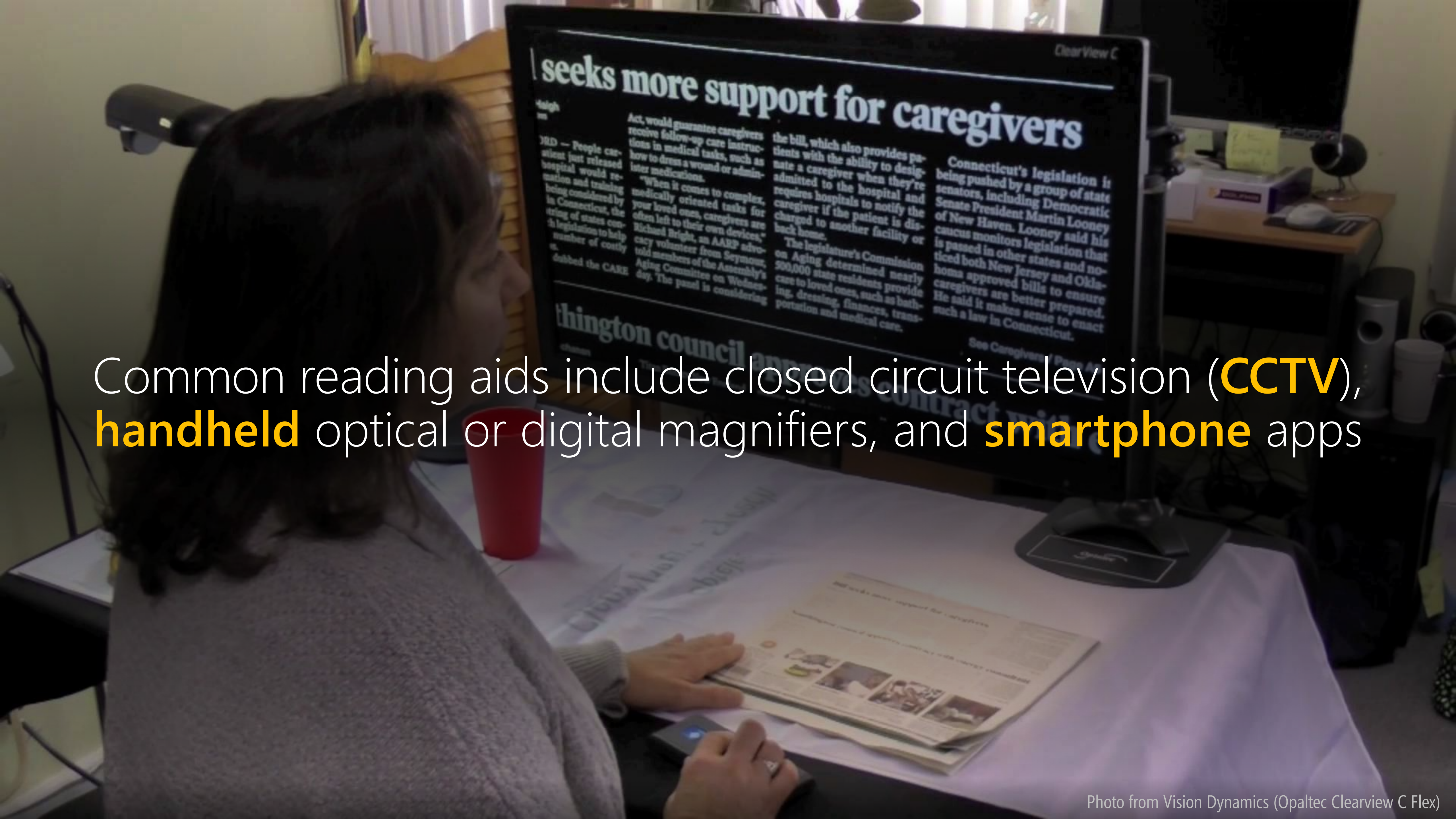
Contributions

- Implementation and systematic evaluation of **haptic and auditory cues** for directional finger guidance
- Identification of **tradeoffs** of both the finger guidance methods and touch-based reading in general in terms of speed, accuracy, and user preference
- Proof-of-concept **realtime system** for reading and exploring printed documents via touch

Extension for Low Vision Users

Lee Stearns, Victor De Souza, Jessica Yin, Leah Findlater, and Jon E. Froehlich, **"Augmented Reality Magnification for Low Vision Users with the Microsoft HoloLens and a Finger-Worn Camera,"** in *Proceedings of ASSETS 2017.*

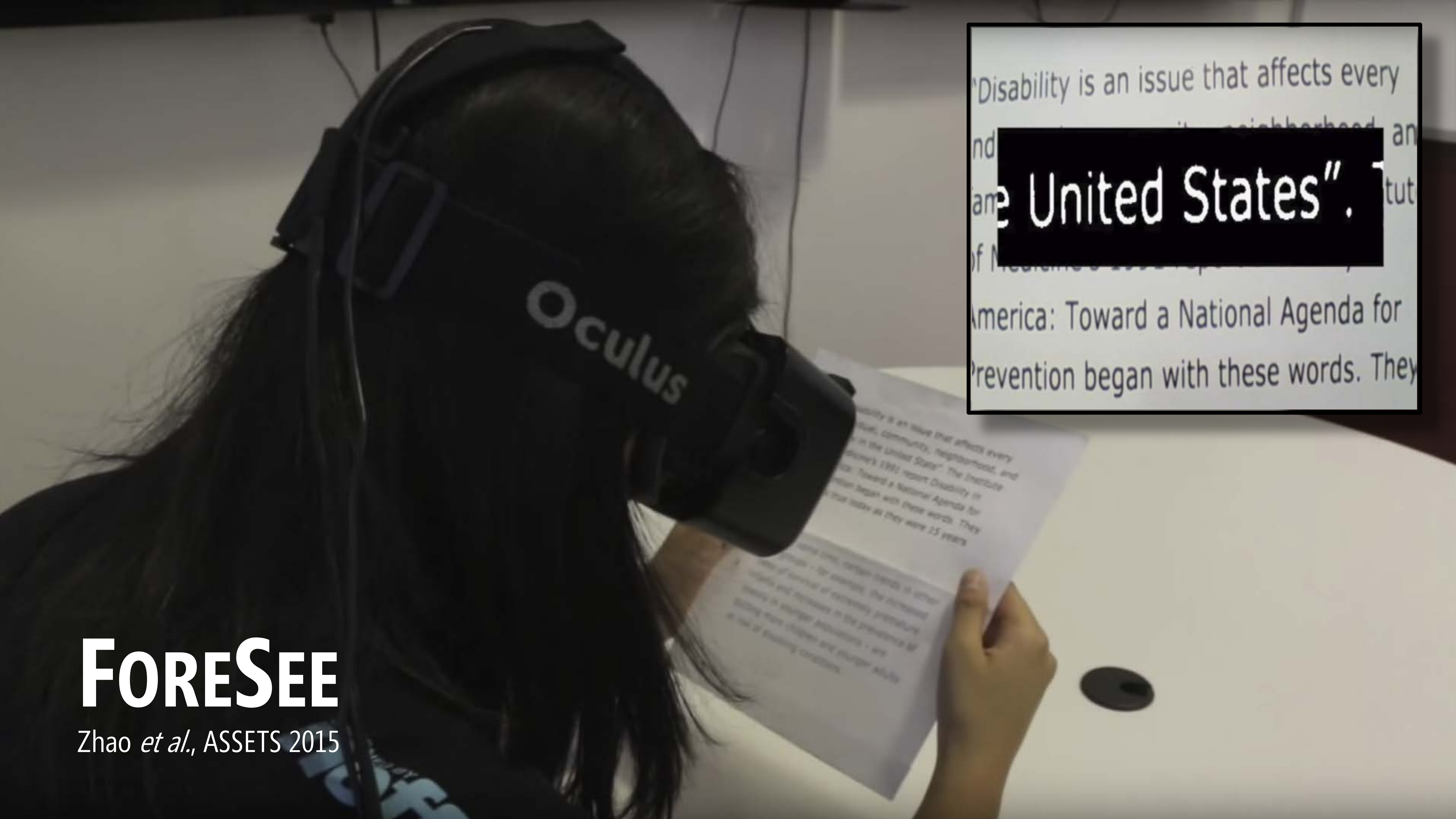
Lee Stearns, Leah Findlater, and Jon E. Froehlich, **"Design of an Augmented Reality Magnification Aid for Low Vision Users,"** in *Proceedings of ASSETS 2018 (To Appear).*



Common reading aids include closed circuit television (**CCTV**), **handheld** optical or digital magnifiers, and **smartphone** apps

Photo from Vision Dynamics (Opaltec Clearview C Flex)

Photo from Artefact Group (eSight 3)



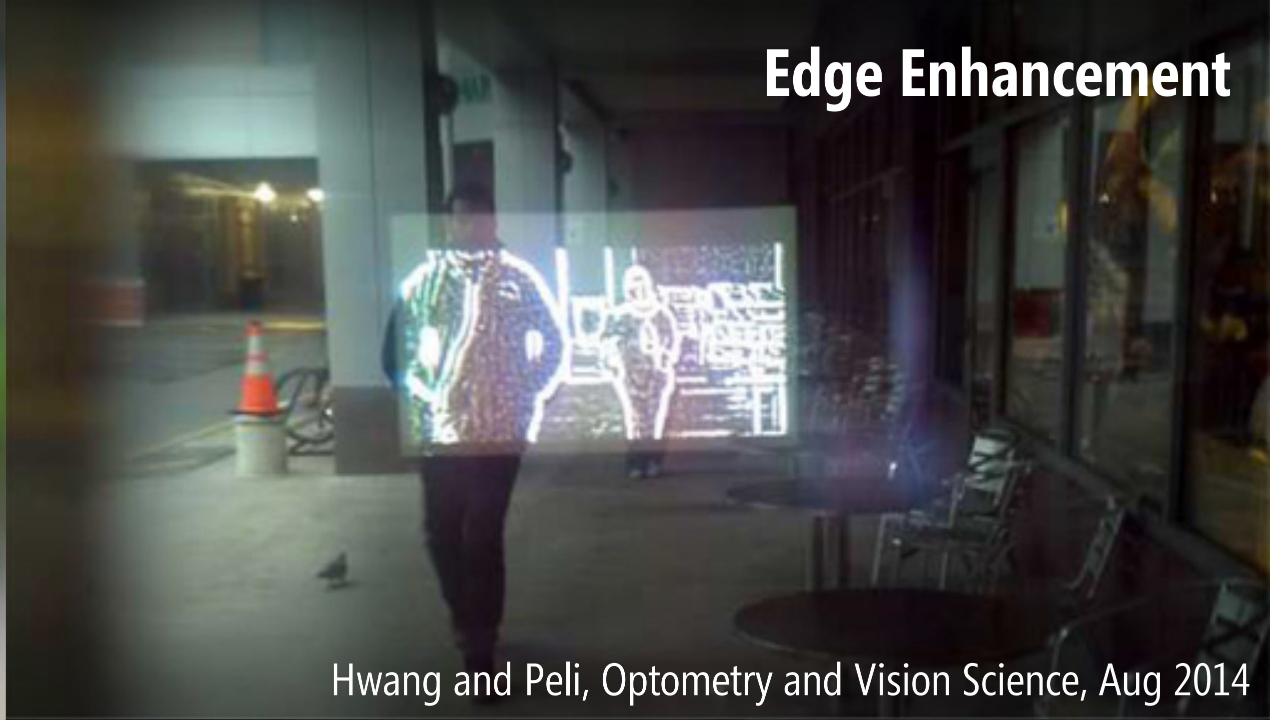
ForeSee

Zhao *et al.*, ASSETS 2015

Google Glass

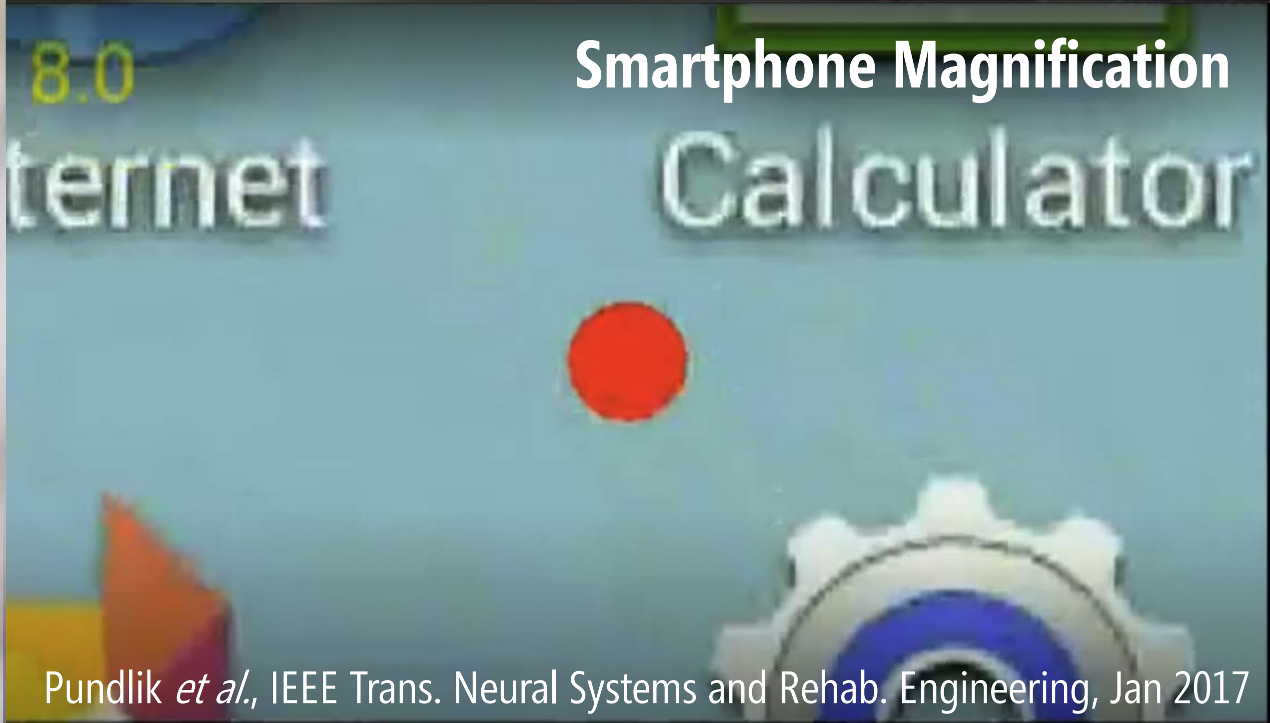
Photo from Wikimedia Commons

Edge Enhancement



Hwang and Peli, Optometry and Vision Science, Aug 2014

Smartphone Magnification



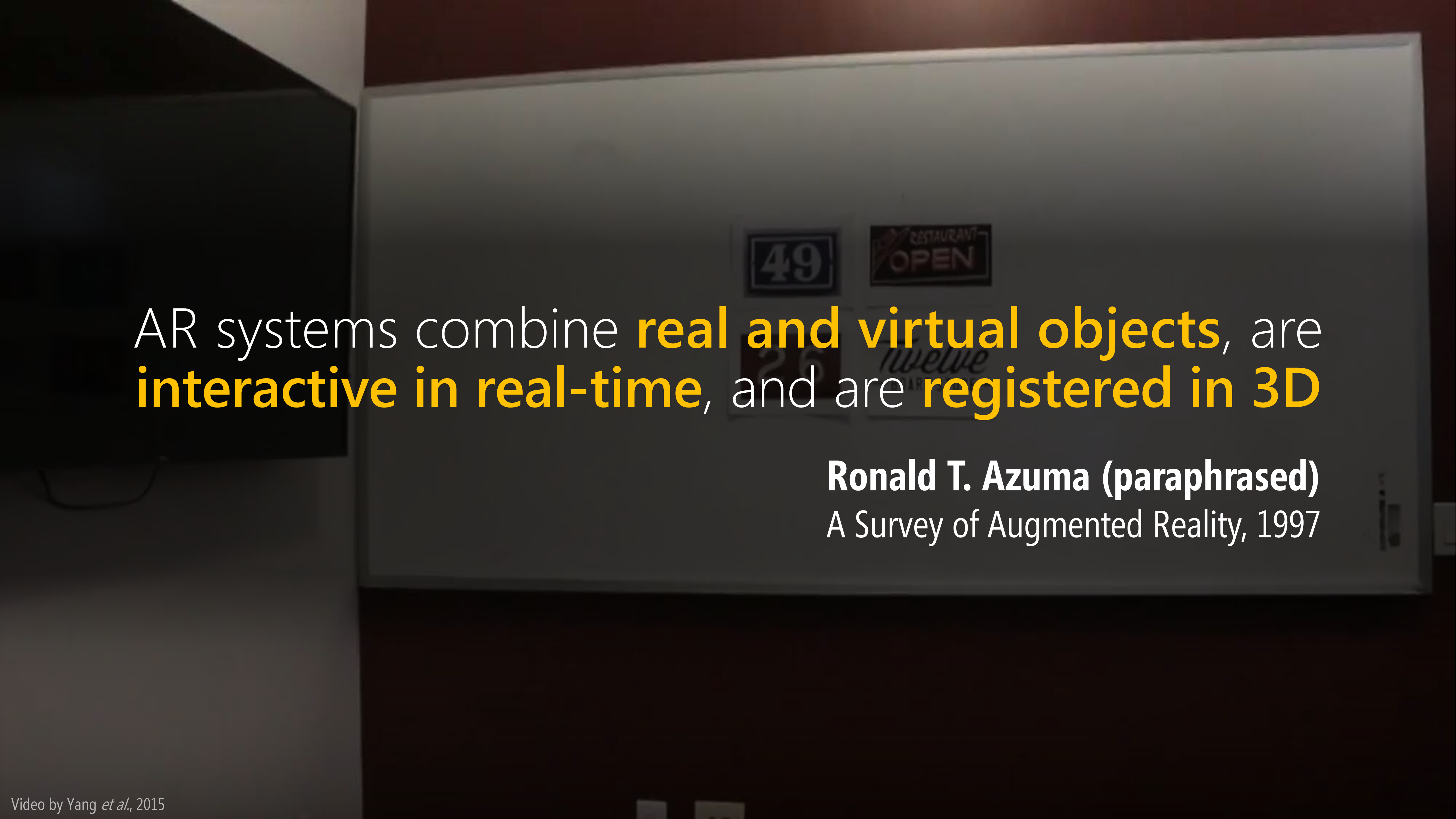
Pundlik *et al.*, IEEE Trans. Neural Systems and Rehab. Engineering, Jan 2017

Commercial Head-Worn Vision Enhancement Systems

eSight

NuEyes

IrisVision



AR systems combine **real and virtual objects**, are **interactive in real-time**, and are **registered in 3D**

Ronald T. Azuma (paraphrased)
A Survey of Augmented Reality, 1997

Video by Yang *et al.*, 2015

Microsoft HoloLens

Photo from Wikimedia Commons

Our Approach

DESIGN SPACE
GOALS

Augment rather than **replace** existing vision capabilities

Design Space
Goals

Augment rather than **replace** existing vision capabilities

Leverage **augmented reality** and persistent 3D content

Design Space
Goals

Augment rather than **replace** existing vision capabilities

Leverage **augmented reality** and persistent 3D content

Prioritize **customization** and **flexibility**

Initial Investigation: HoloLens
Design

Built-in camera to capture images

Two display modes:
Fixed 2D & Fixed 3D

Voice Commands to select mode

Image Enhancements:
Binary threshold & Invert colors

Initial Investigation: HoloLens
Observations

Camera resolution too low

Turning head to look at desired content was uncomfortable

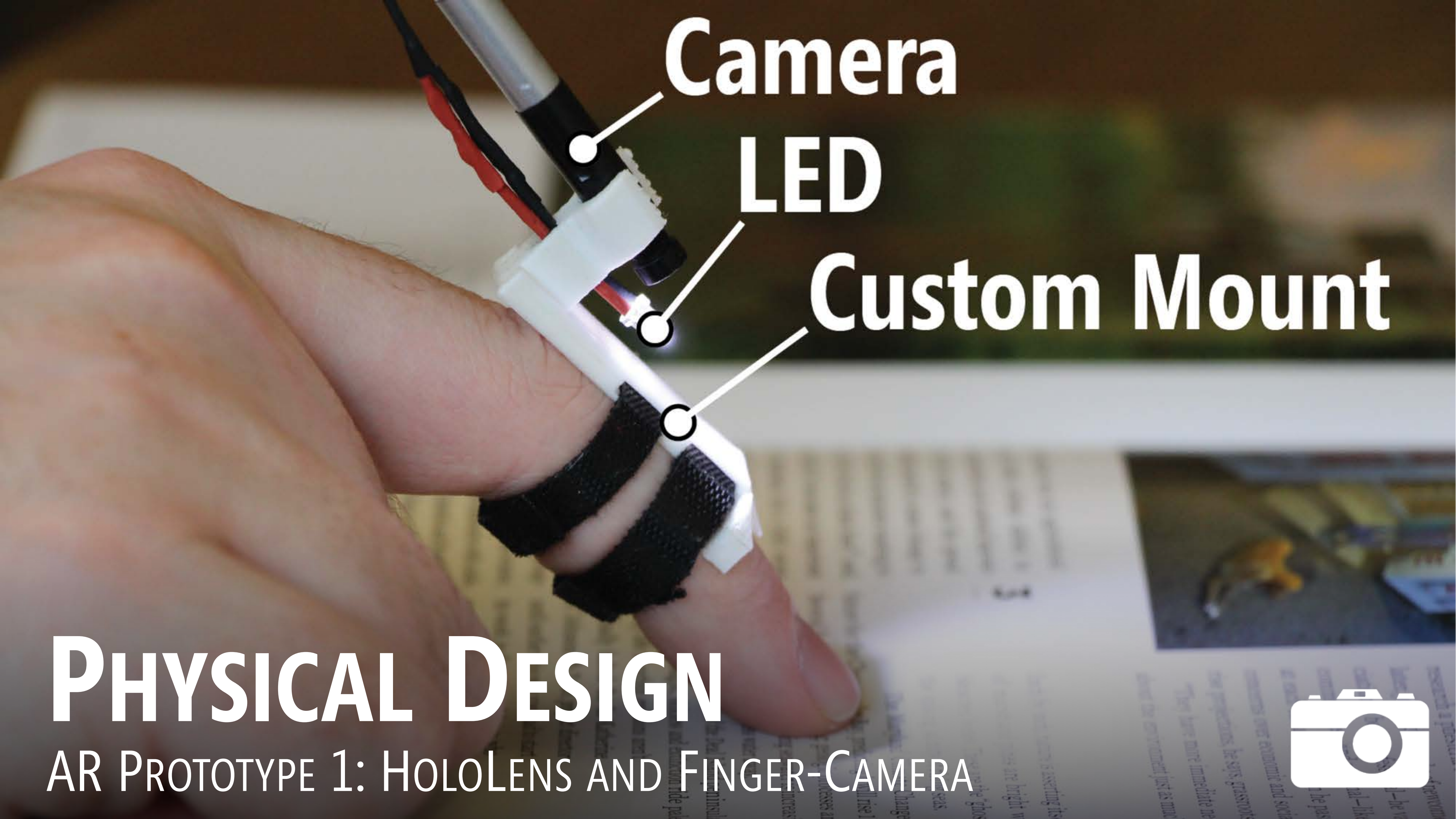
Voice commands cumbersome, imprecise, limited customization

AR Prototype 1

HoloLens and Finger-Camera

Physical Design

AR Prototype 1: HoloLens and Finger-Camera


Camera
LED
Custom Mount
PHYSICAL DESIGN
AR PROTOTYPE 1: HOLOLENS AND FINGER-CAMERA

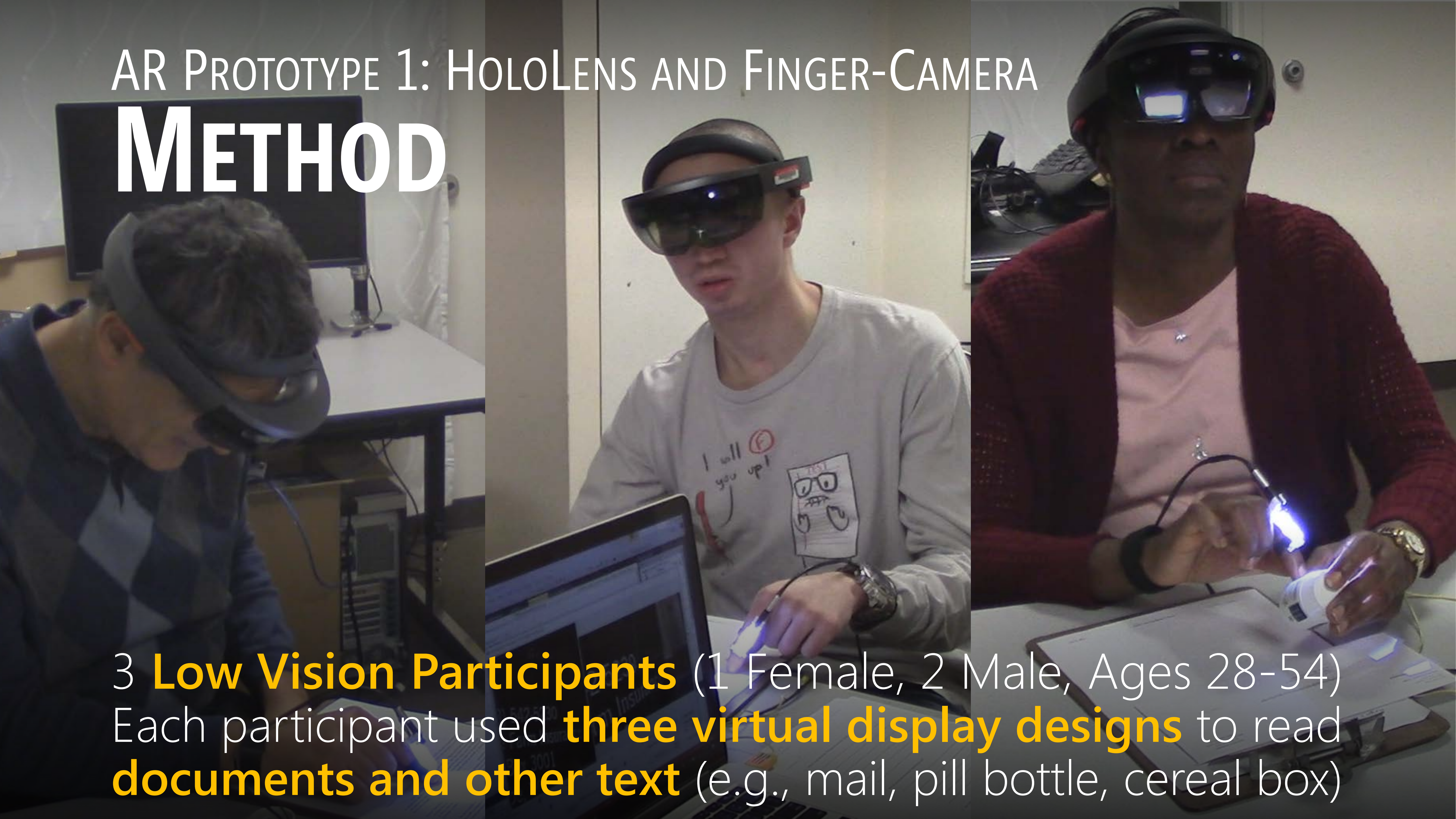

Virtual Display Design 1: Fixed 2D
Acts as a heads-up display, stays in the user's view at all times
nt just as much.
is asserting itself on
s are bright white
se are the "ghost fo
of rising seas.
Aa Bb
VIRTUAL DISPLAYS
AR PROTOTYPE 1: HOLOLENS AND FINGER-CAMERA

User Interactions

AR Prototype 1: HoloLens and Finger-Camera

AR Prototype 1: HoloLens and Finger-Camera

Method



3 **Low Vision Participants** (1 Female, 2 Male, Ages 28-54)

Each participant used **three virtual display designs** to read **documents and other text** (e.g., mail, pill bottle, cereal box)

AR Prototype 1: HoloLens and Finger-Camera

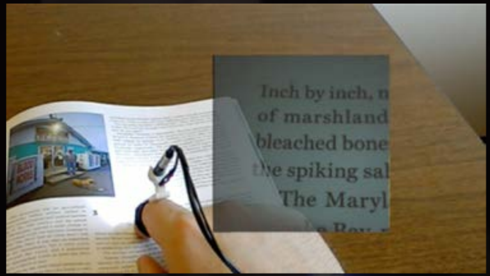
Method

They provided **feedback and suggestions** on their likes, dislikes, design preferences, ideas for improvements or new features

AR PROTOTYPE 1: HOLOLENS AND FINGER-CAMERA

FINDINGS

Virtual Display Designs



Fixed 3D (Vertical or Horizontal)

Reading experience similar to using a CCTV or handheld magnifier.

Finger Tracking

Can help to quickly search a document.

Fixed 2D

Always visible, required least concentration.

AR Prototype 1: HoloLens and Finger-Camera
Findings

Finger-Worn Camera

[+] Flexible, allows hands-free use

[–] Requires moving finger to read

[–] Small field of view (~3-4 lines)

AR Prototype 1: HoloLens and Finger-Camera
Findings

HoloLens Display

- [–] Low contrast due to transparency
- [–] Narrow view, center of vision

AR Prototype 1: HoloLens and Finger-Camera

Findings

User Input

- [–] Midair gestures difficult to use
- [–] Unable to make quick adjustments

AR Prototype 2

HoloLens and Smartphone

Physical Design

AR Prototype 2: HoloLens and Smartphone

AR Prototype 2: HoloLens and Smartphone

Physical Design

Virtual Displays

AR Prototype 2: HoloLens and Smartphone


Each design included several options for customization, including the position, size, and contrast/colors
User Interactions
AR Prototype 2: HoloLens and Smartphone

AR Prototype 2: HoloLens and Smartphone

Method

6 **Low Vision Participants** (3 Female, 3 Male, Ages 28-68)
Each participant used **three virtual display designs** to read **documents and other text** (e.g., mail, pill bottle, cereal box)

AR Prototype 2: HoloLens and Smartphone

Findings

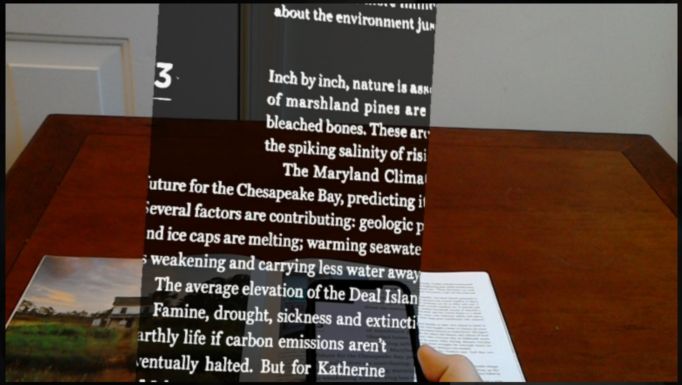
Participants were **more successful and positive** about their experience using this version of our system.

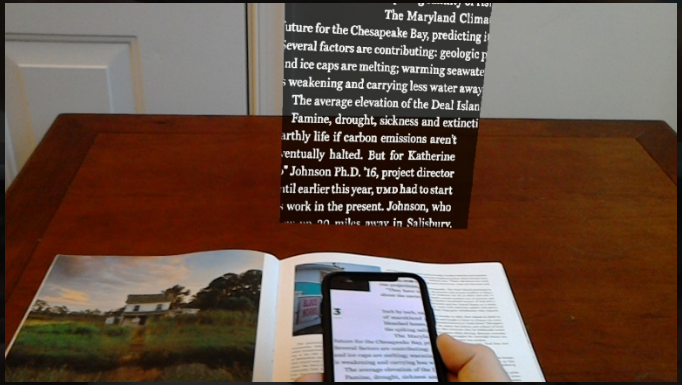
They were better able to experience the **AR aspects of our approach**, which most participants found promising.

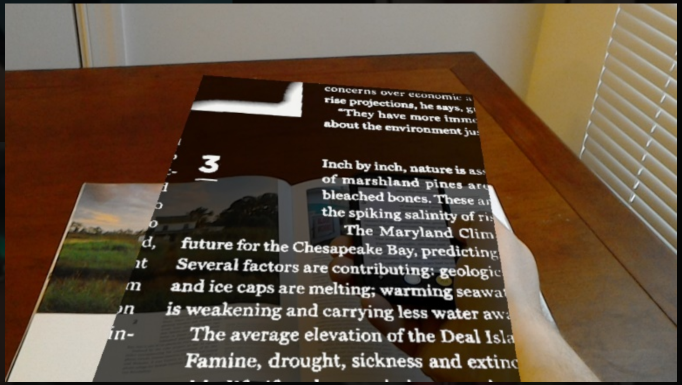
AR Prototype 2: HoloLens and Smartphone

Findings

Virtual Display Designs







Attached to Headset

Easier to focus on the text

Potentially distracting

Attached to World

Natural reading experience

Easier to multitask

Attached to Phone

Versatile

Intuitive interactions

AR Prototype 2: HoloLens and Smartphone
Findings

Smartphone

[+] Better camera

[+] More usable interactions

[–] No longer hands-free

[–] Too heavy for extended use

AR Prototype 2: HoloLens and Smartphone
Findings

HoloLens

Issues with contrast, field of view, and physical size and weight still present.

Participants with central vision loss struggled to use the system.

CONCLUSIONS

Strengths and Weaknesses of 3D AR for Magnification

[+] Enables new interactions not possible with other approaches

[+] Good for multitasking

Conclusions

Strengths and Weaknesses of 3D AR for Magnification

[+] Enables new interactions not possible with other approaches

[+] Good for multitasking

[–] May require more effort to use than fixed 2D display

CONTRIBUTIONS

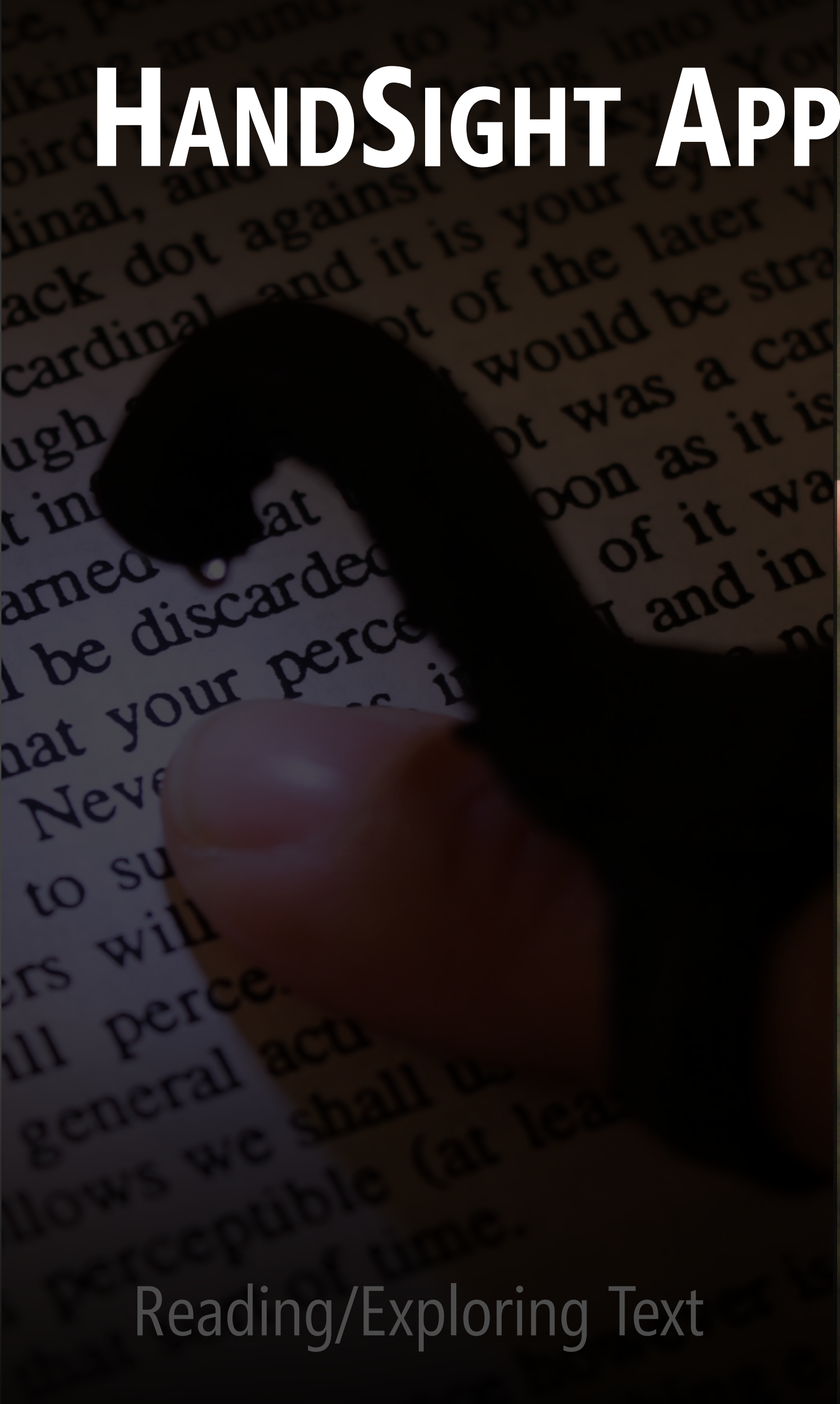
Design space exploration: **AR magnification** & **enhancement**

Implementation and evaluation of **proof-of-concept designs** with low vision users

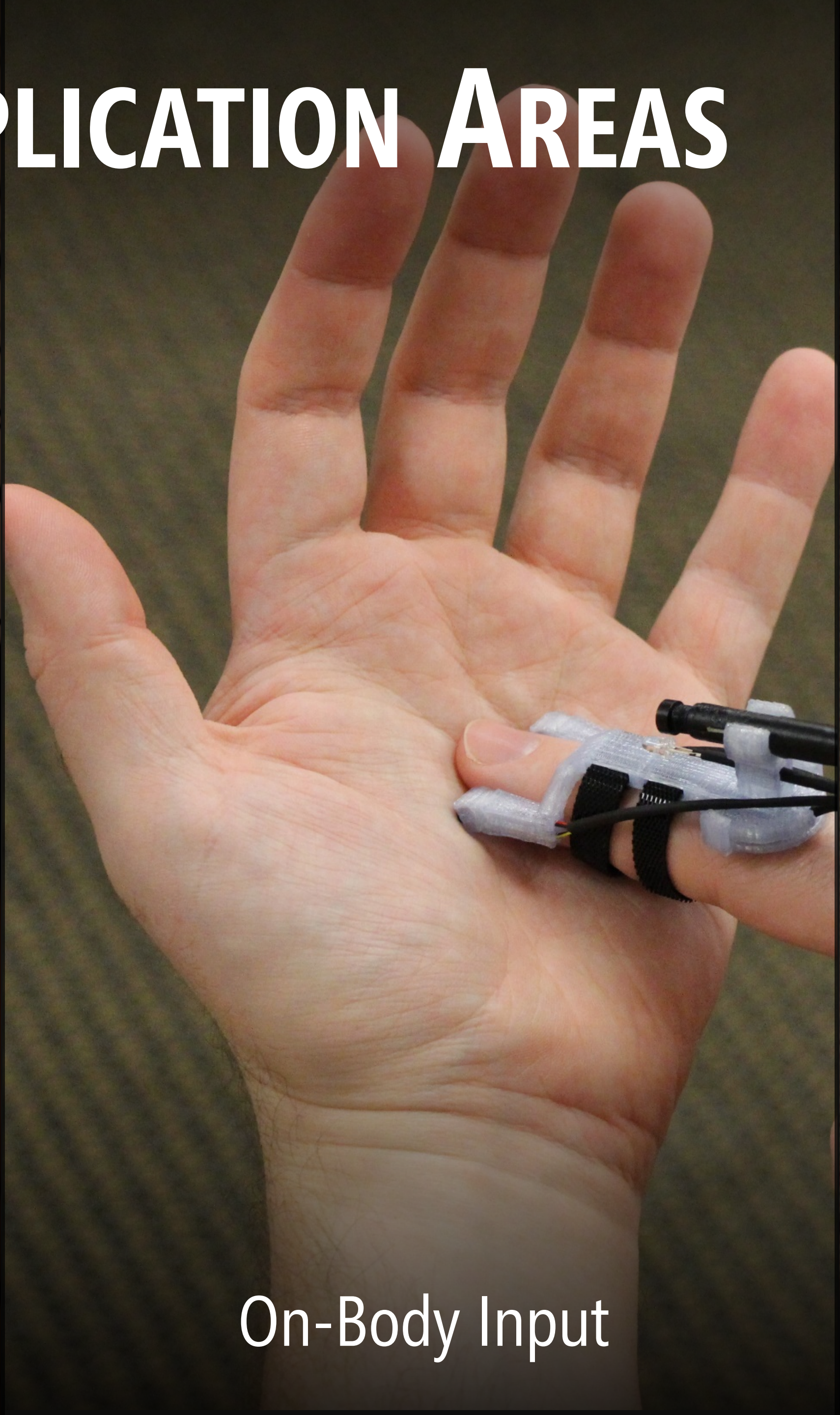
Proposed **design recommendations** for future AR vision enhancement aids

HandSight Application Areas

HandSight Application Areas

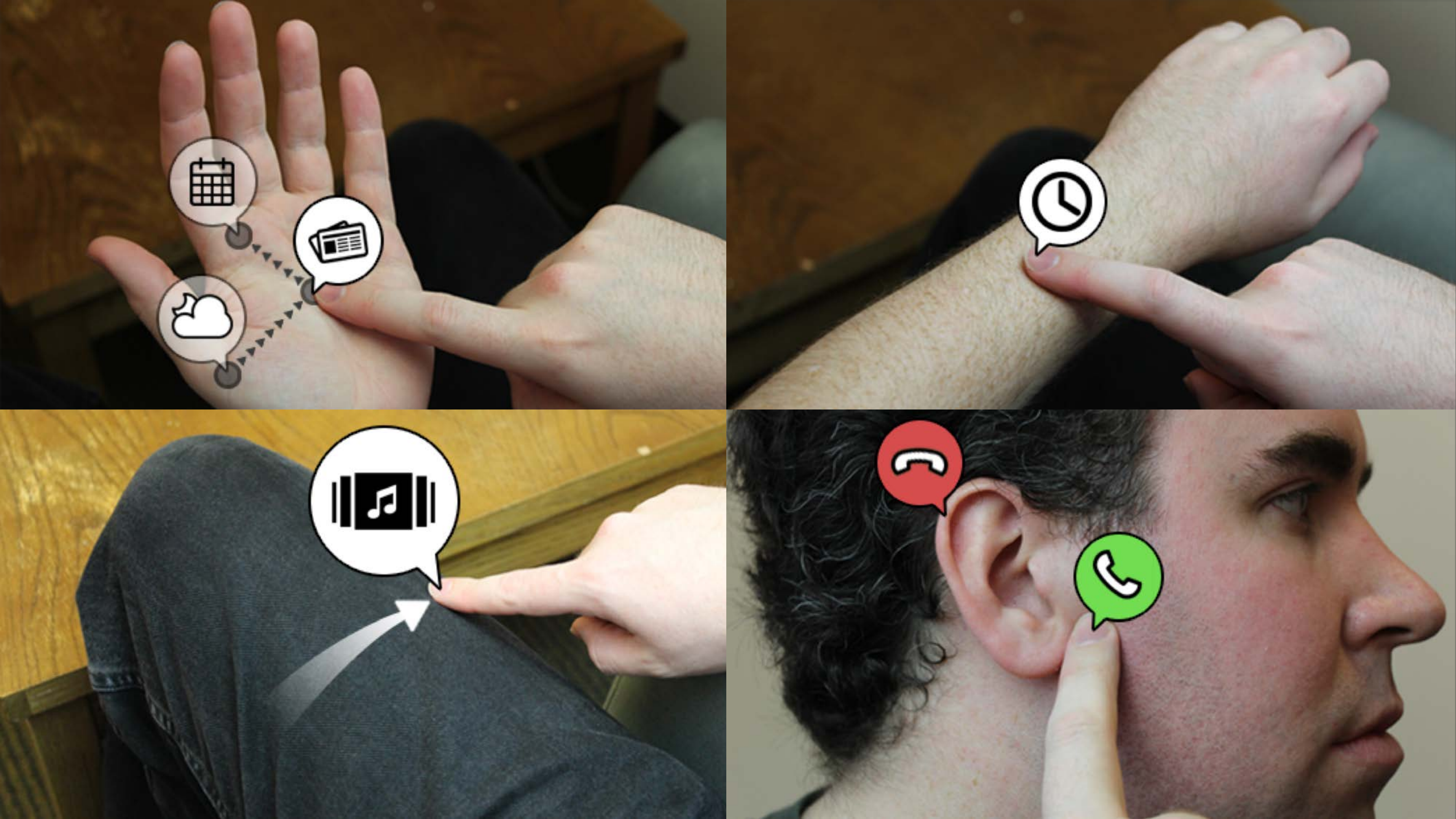


Reading/Exploring Text



On-Body Input

Clothing Colors/Patterns



On-Body Input
Using Finger-Worn Sensors

Advantages:

1. More easily **scalable** to other body locations and surfaces
2. Larger **input vocabulary**: touch position × relative gestures
3. Simplified **sensing and processing** due to sensor positioning

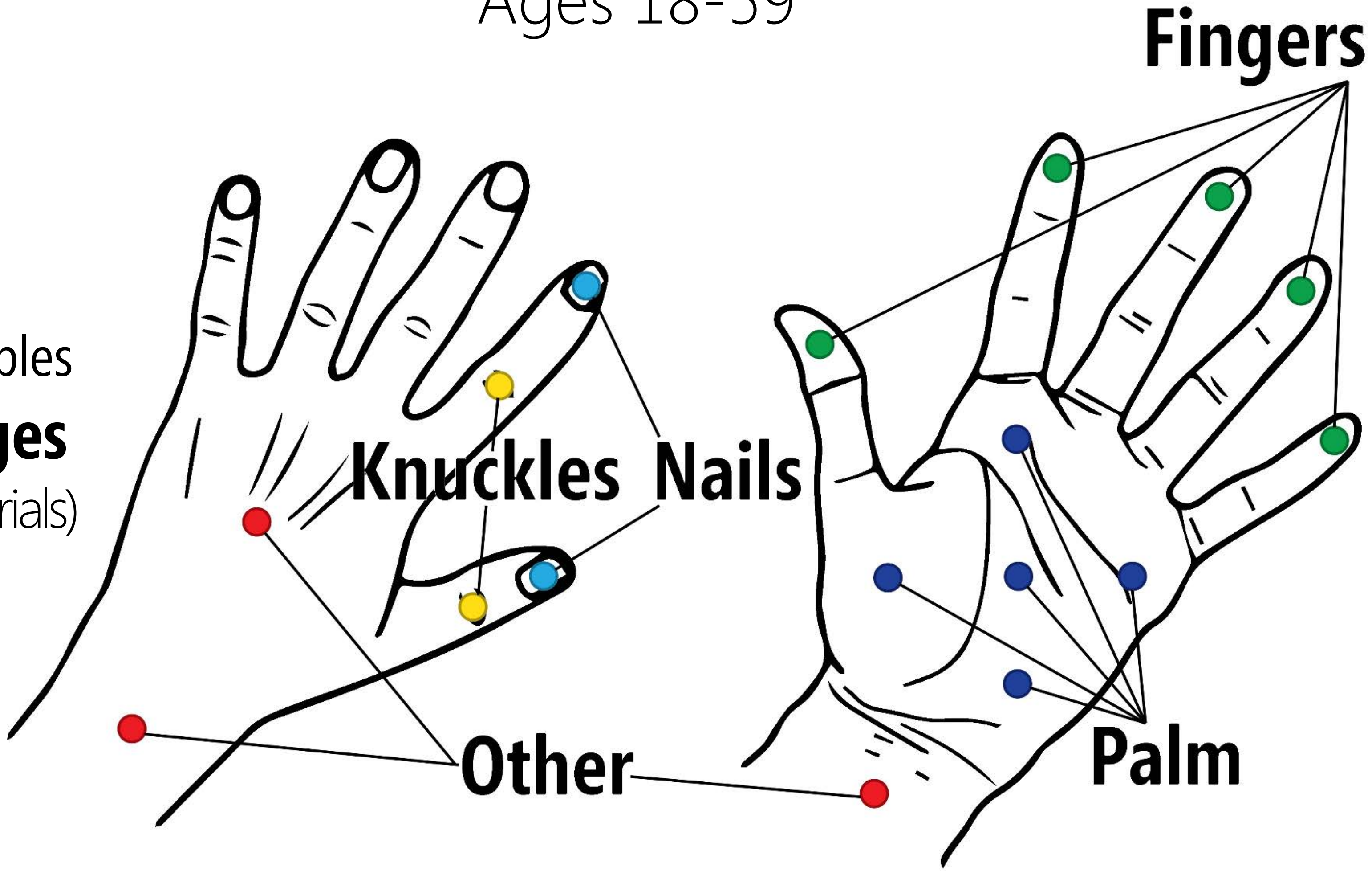
ON-BODY STUDY I
DATASET

30 Sighted Participants
23 Female, 7 Male
Ages 18-59

30 participants × 17 locations × 20 samples

Total dataset size: 10,198 images

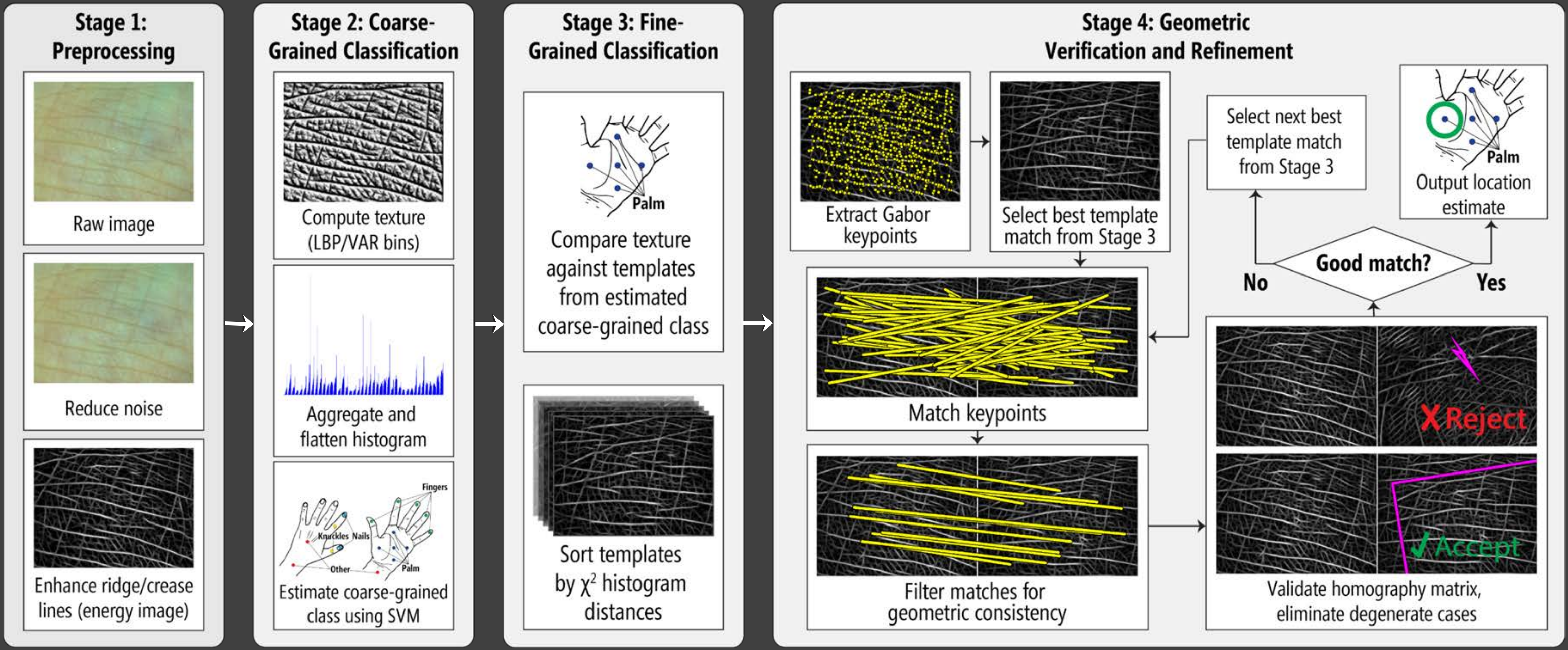
(one participant accidentally skipped two trials)



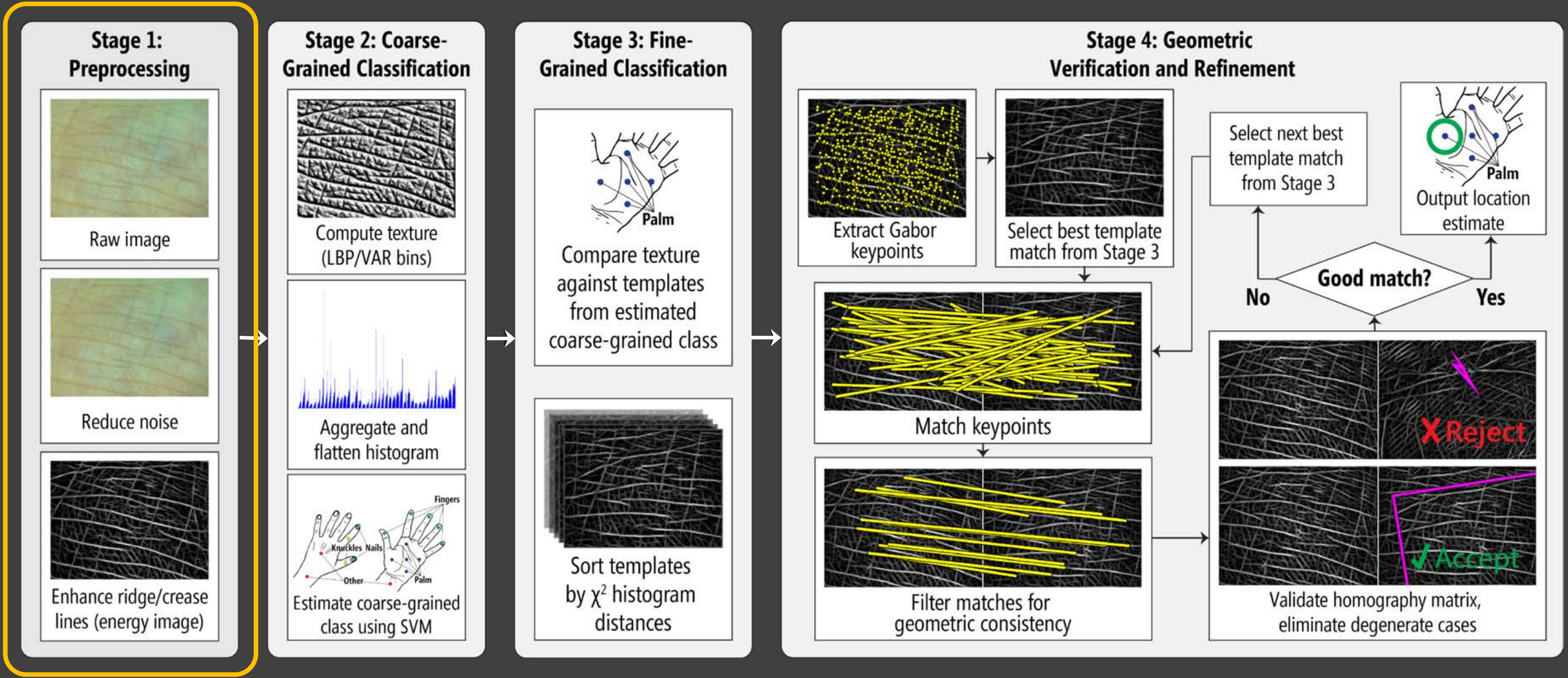
Lee Stearns, Uran Oh, Bridget J. Cheng, Leah Findlater, David Ross, Rama Chellappa, and Jon E. Froehlich, **"Localization of Skin Features on the Hand and Wrist from Small Image Patches,"** in *Proceedings of ICPR 2016*.

On-Body Study I

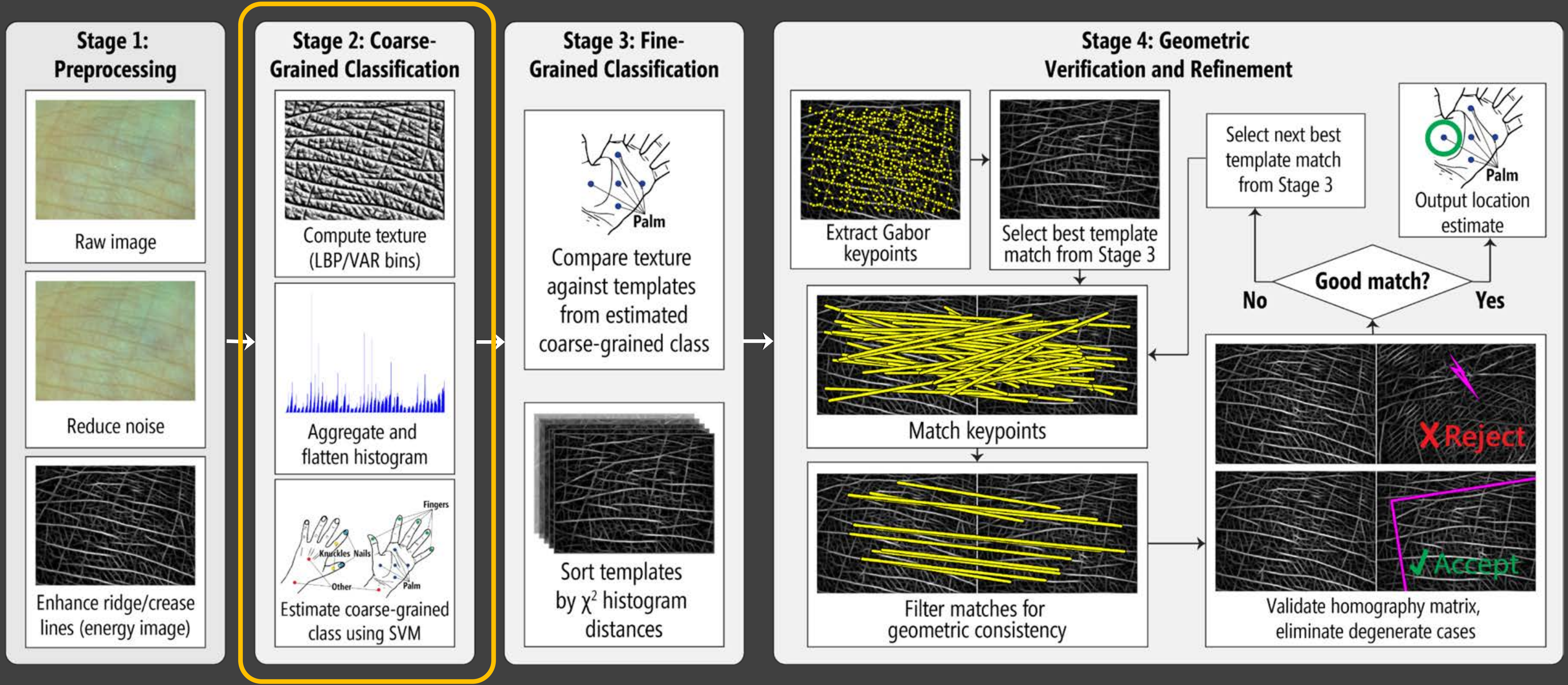
Algorithms



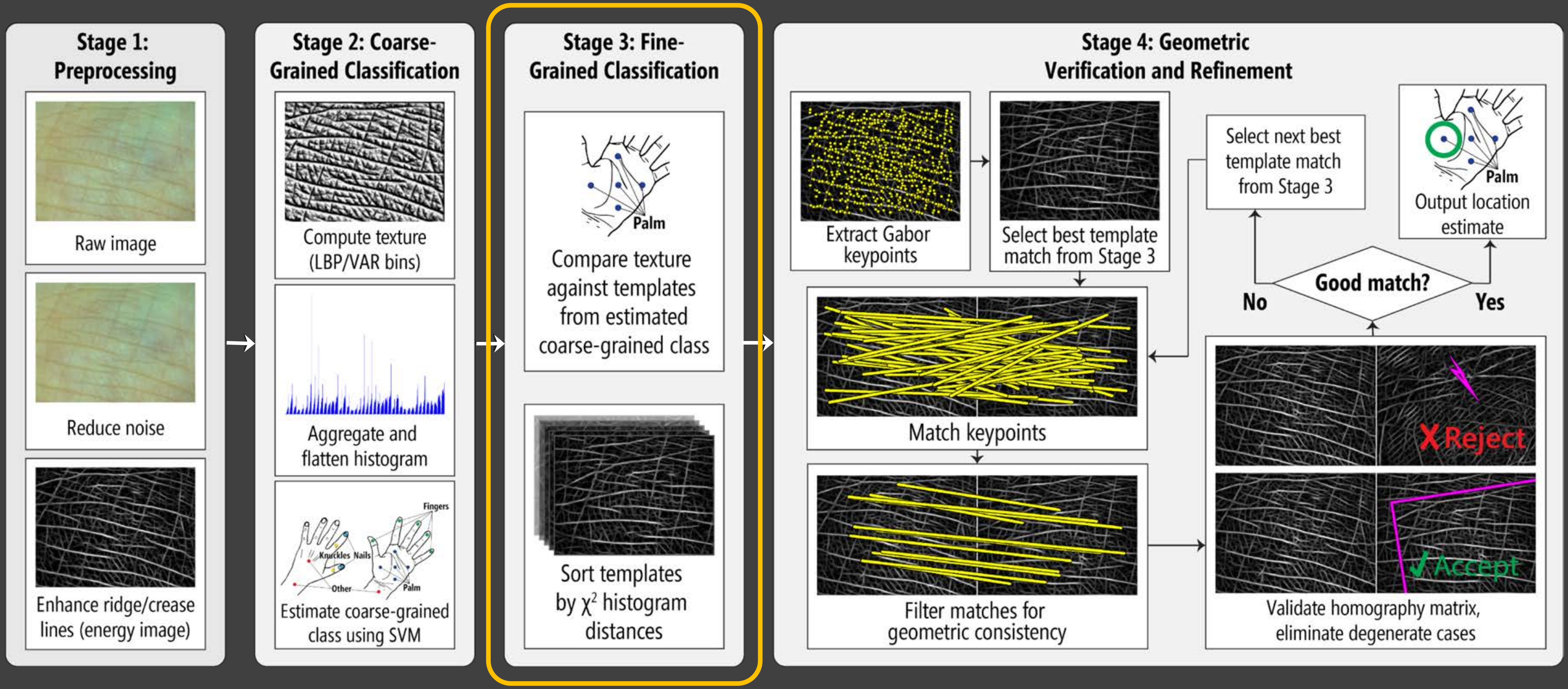
On-Body Study I
Algorithms



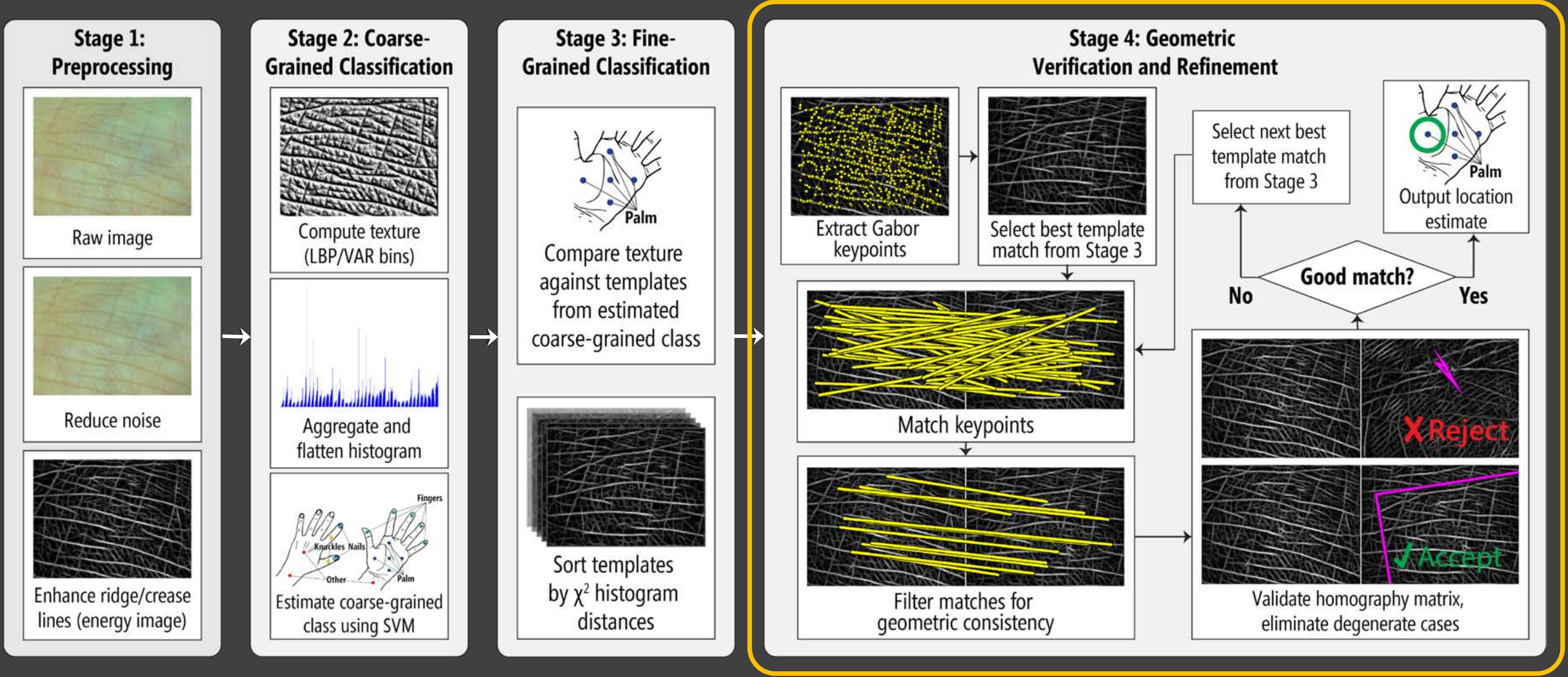
On-Body Study I
Algorithms



On-Body Study I
Algorithms



On-Body Study I
Algorithms



On-Body Study I

Experiments and Results

Within-Person Classification Experiment (n=30)

Coarse-Grained Localization (Stage 2)

(5 classes: ● finger, ● palm, ● knuckle, ...)

accuracy: **99.1%** (*SD*=0.9%)

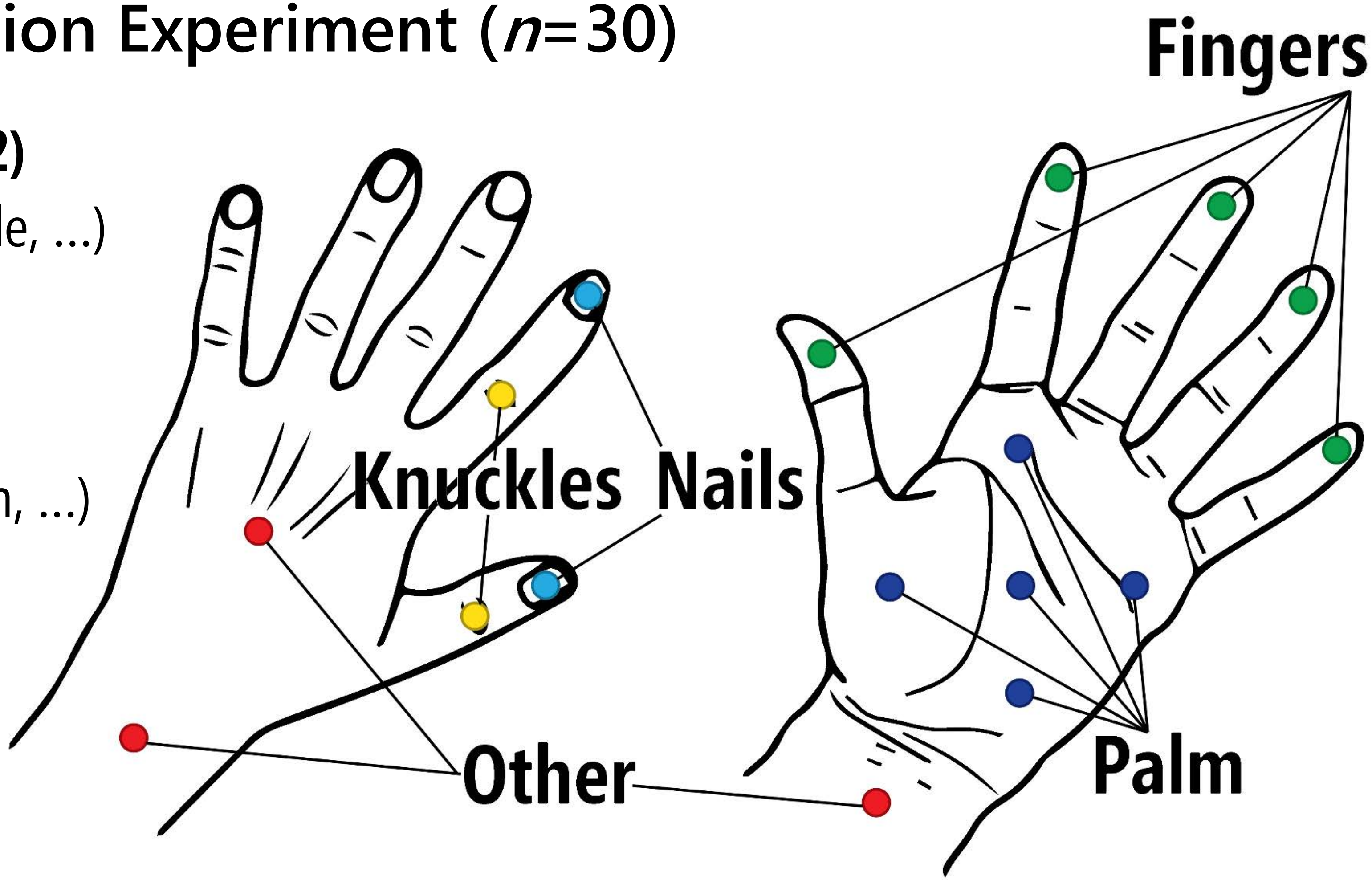
Fine-Grained Localization (Stage 3)

(17 classes: ● palm up, ● palm down, ...)

accuracy: **88.0%** (*SD*=4.5%)

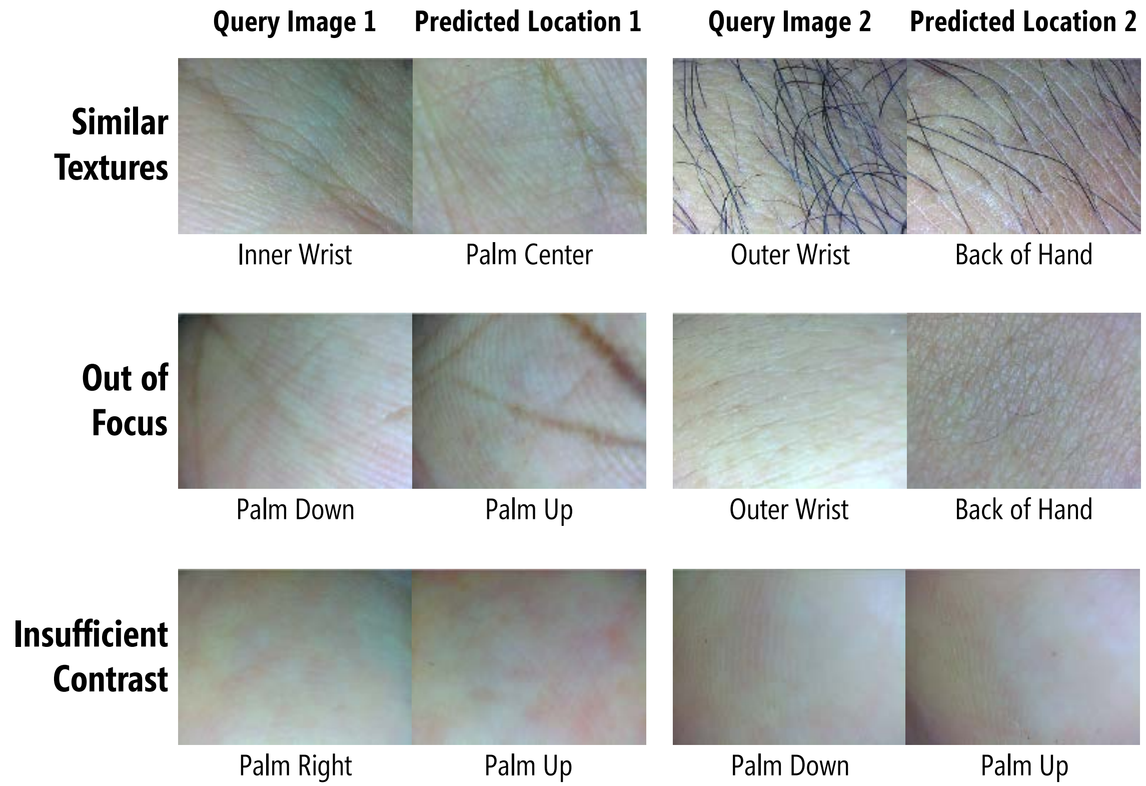
Fine-Grained Localization (Stage 4)

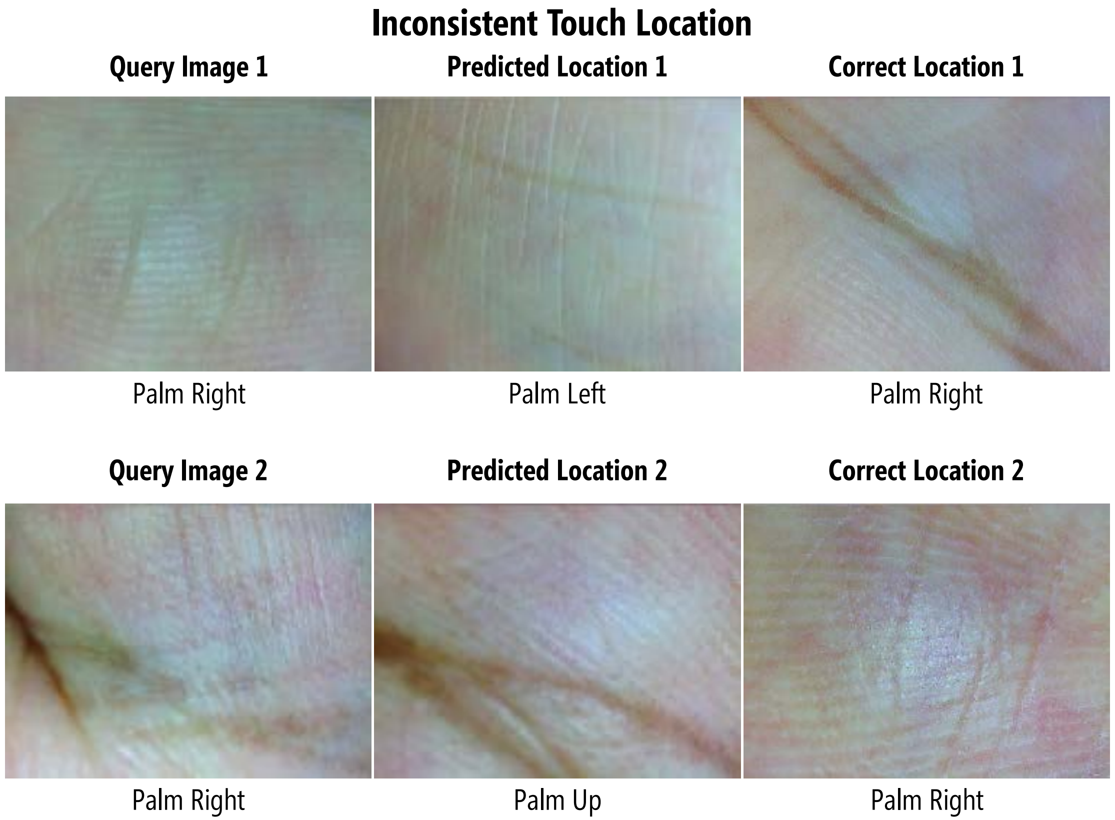
accuracy: **96.4%** (*SD*=2.3%)



On-Body Study I
Experiments and Results

Within-Person Classification Experiment (n=30)

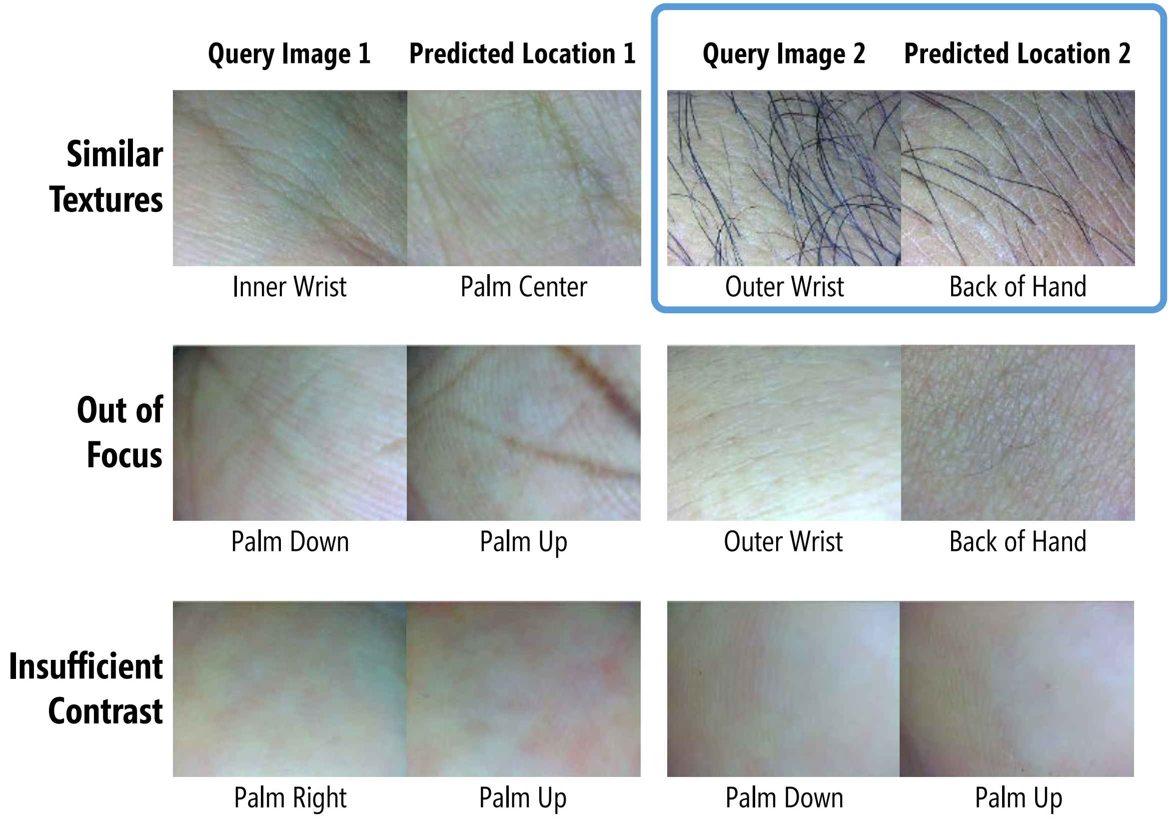


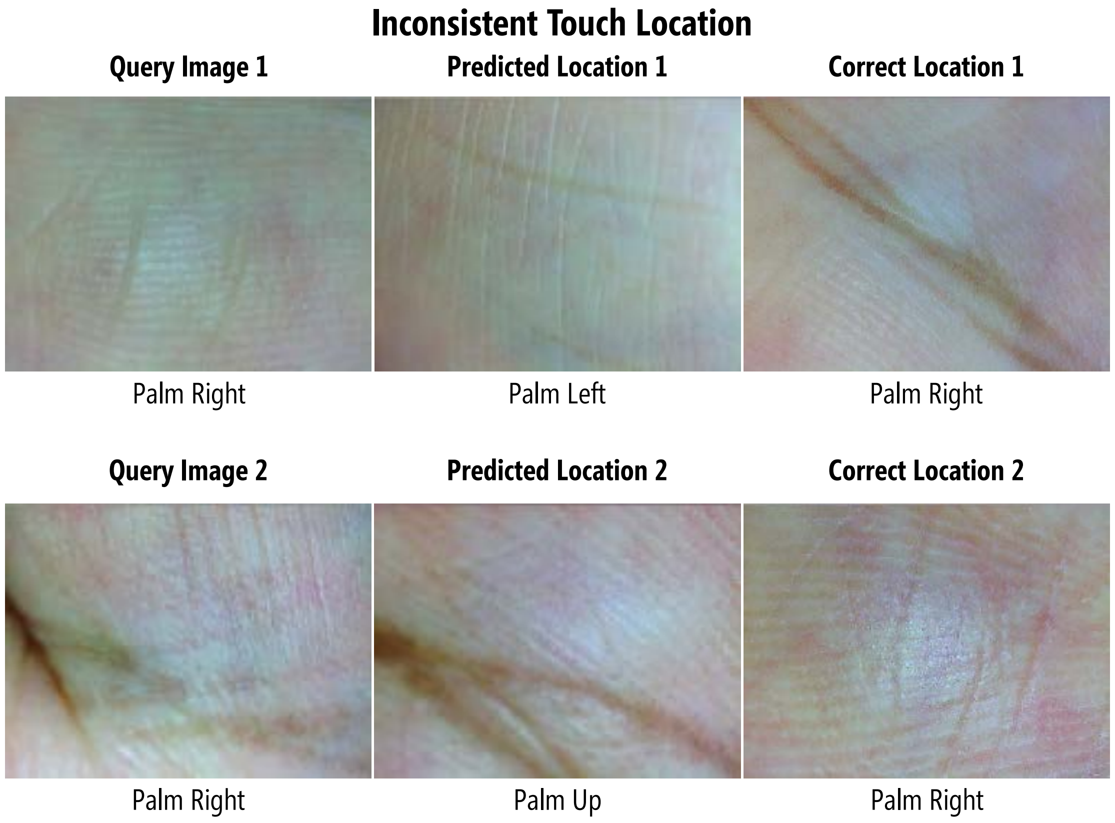


On-Body Study I

Experiments and Results

<u>Within-Person</u> Classification Experiment (*n*=30)

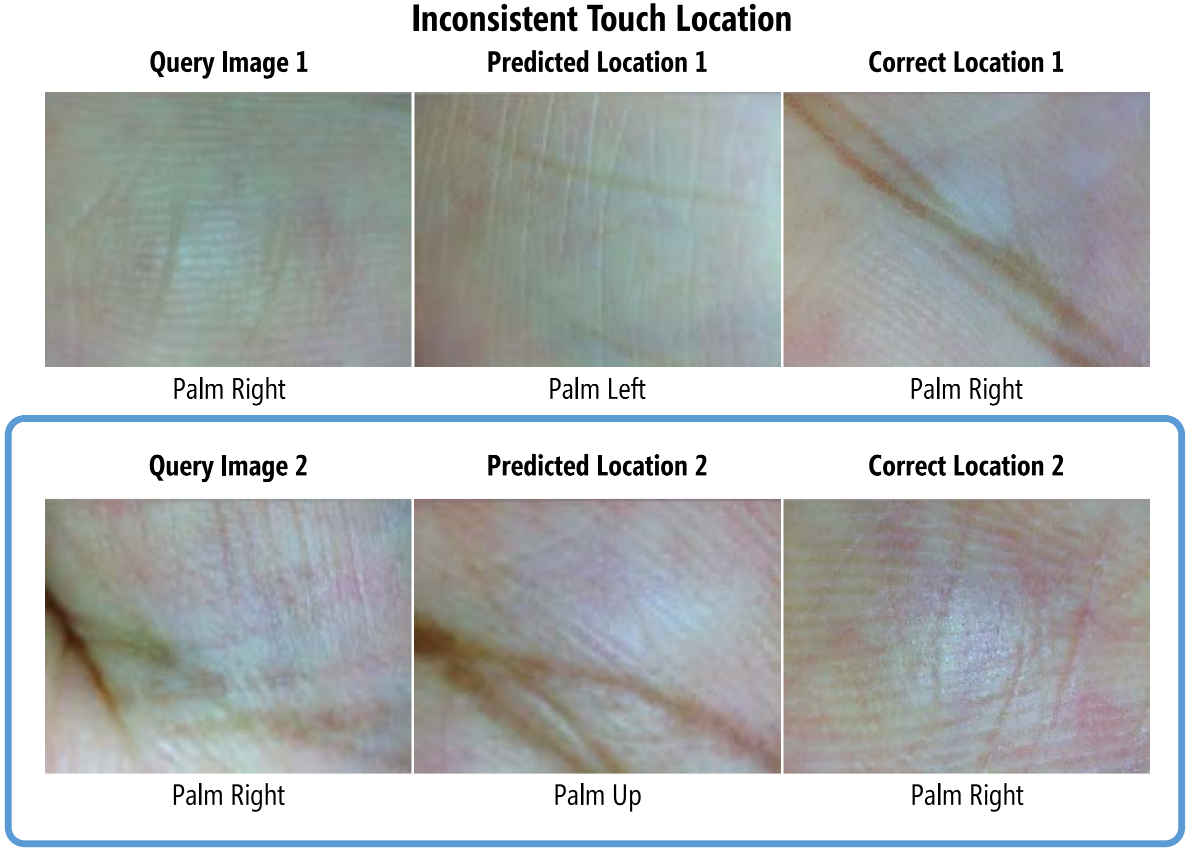




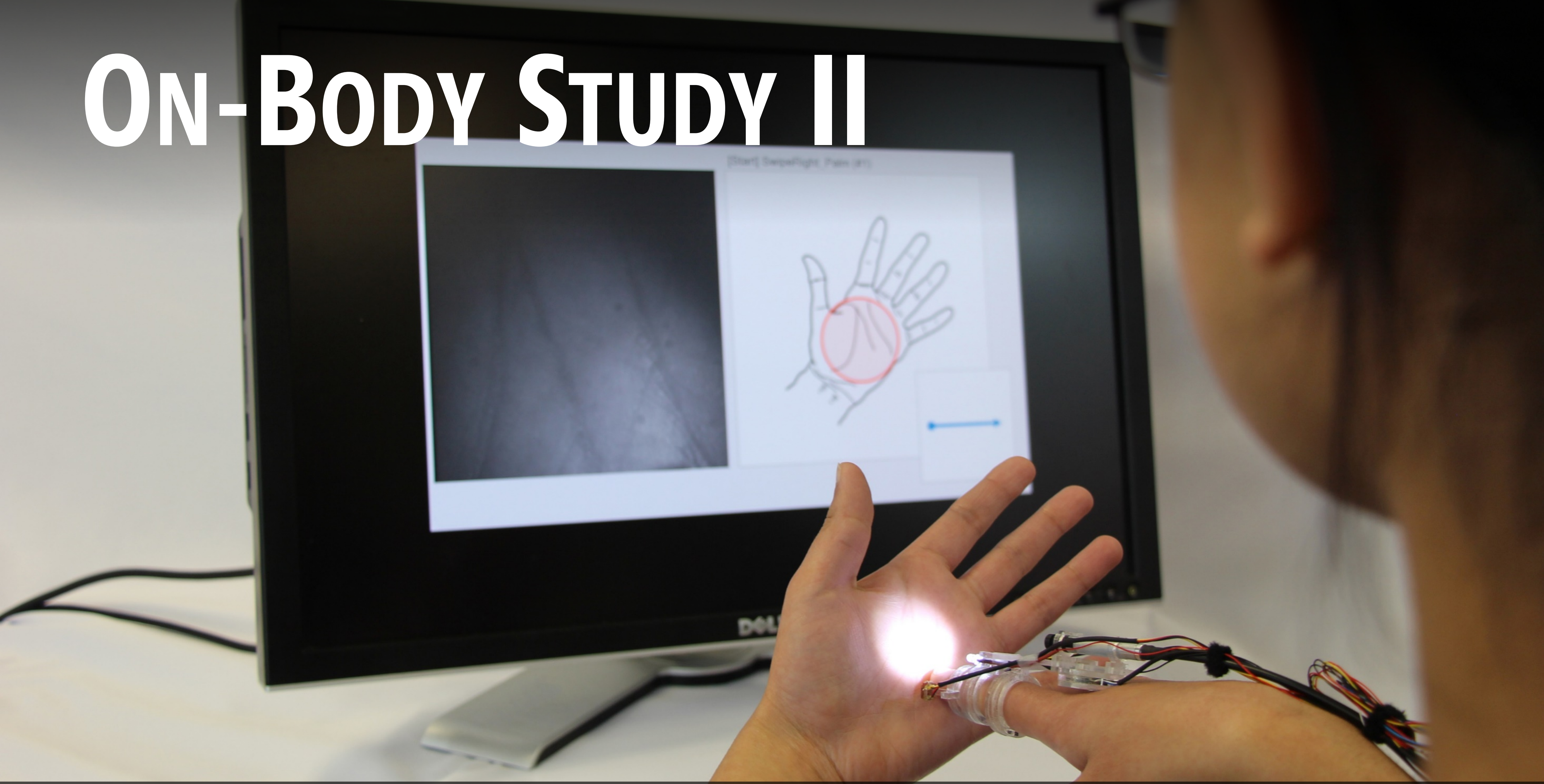
On-Body Study I

Experiments and Results

<u>Within-Person</u> Classification Experiment (*n*=30)

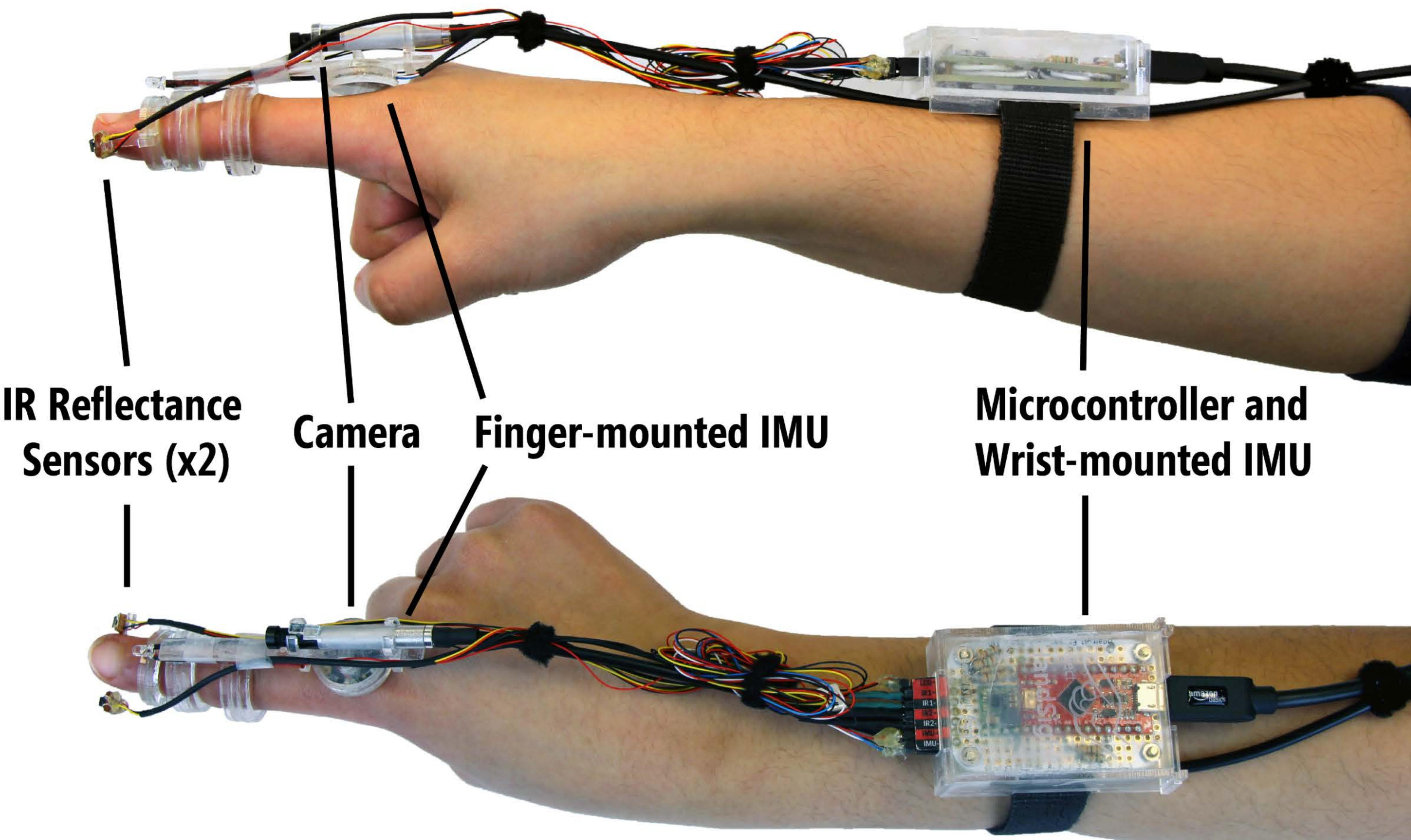


On-Body Study II



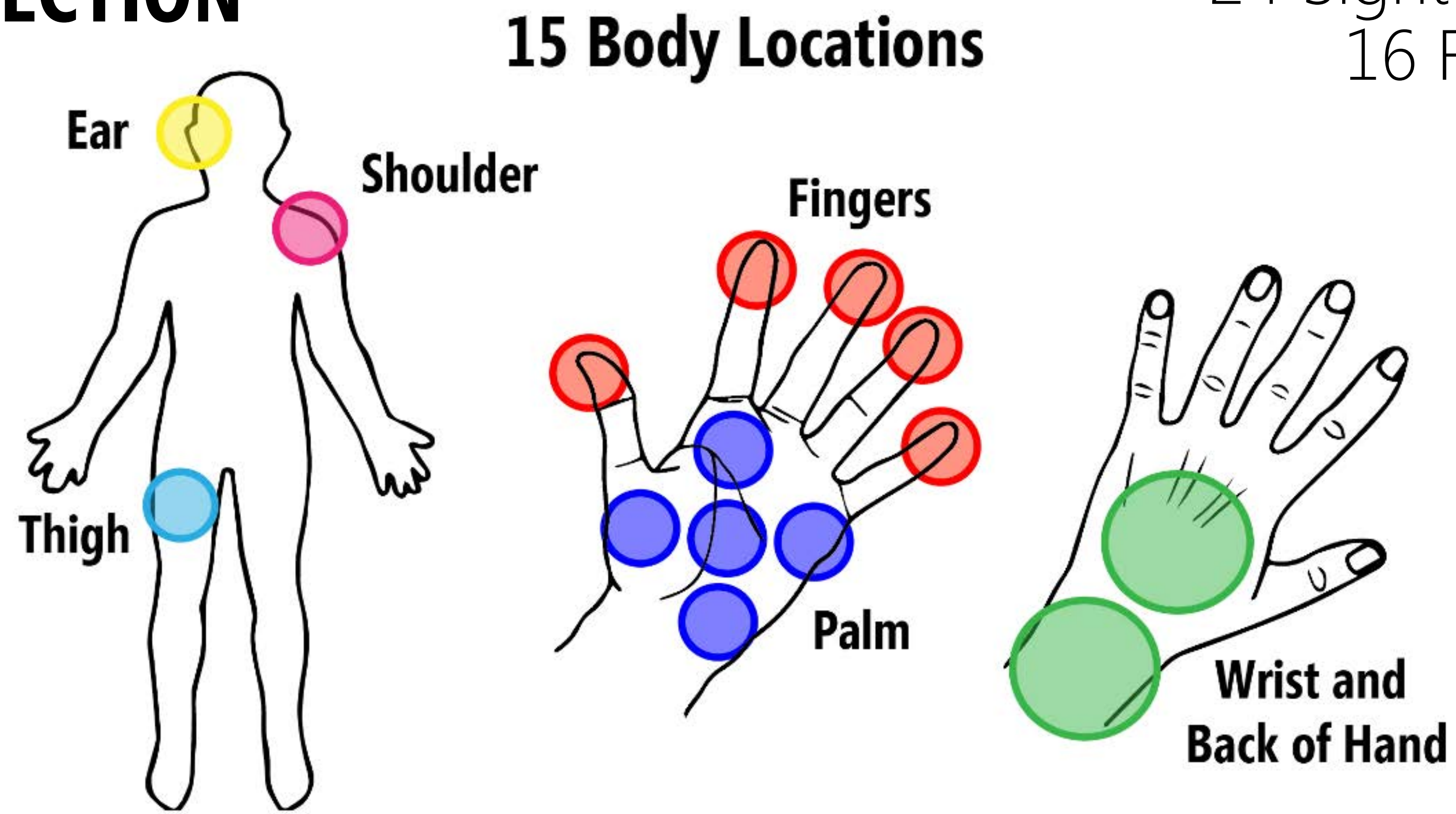
Lee Stearns, Uran Oh, Leah Findlater, Jon E. Froehlich, **"TouchCam: Realtime Recognition of Location-Specific On-Body Gestures to Support Users with Visual Impairments,"** in *Interactive, Wearable and Ubiquitous Technologies, December 2018.*

Offline Prototype



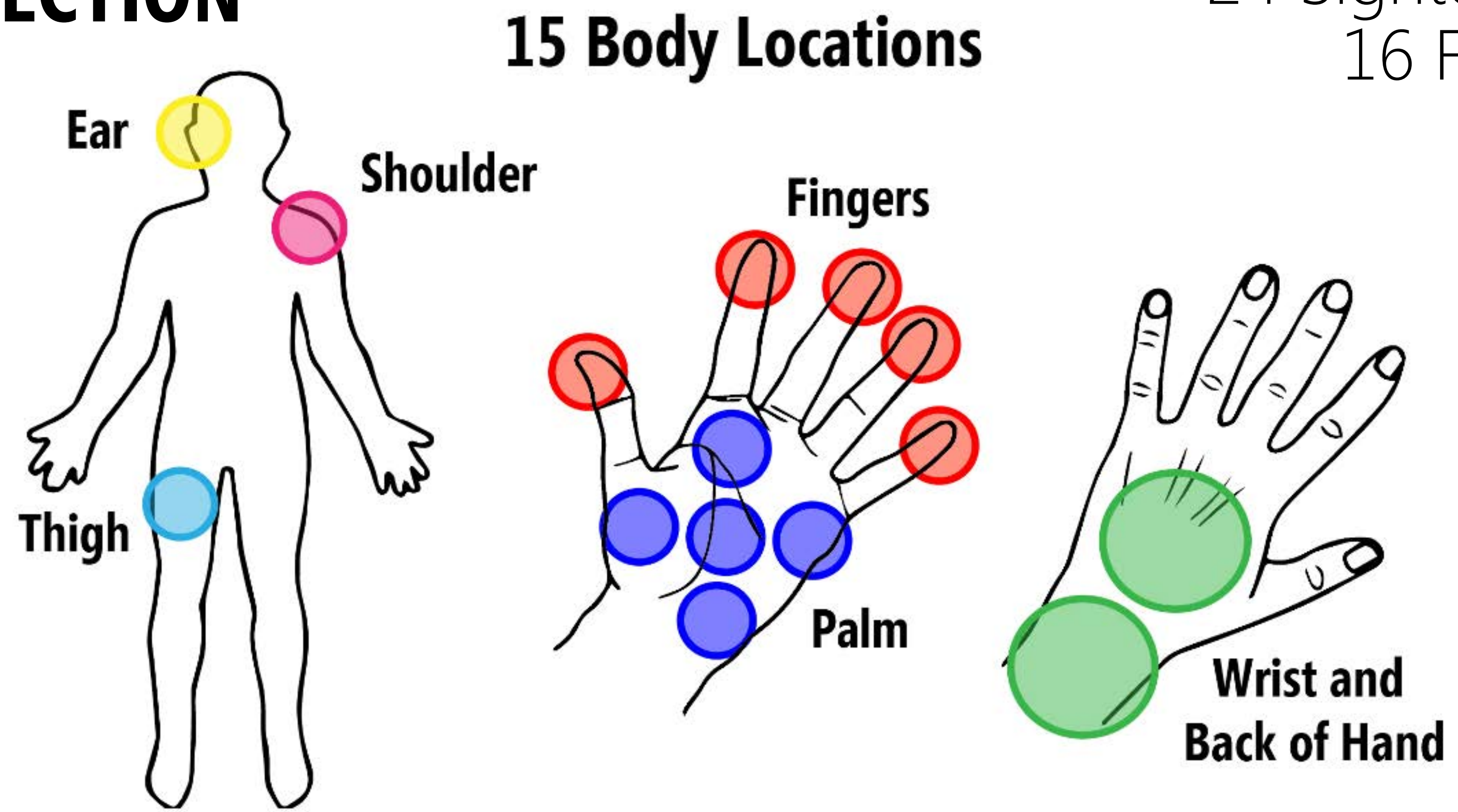
Data Collection

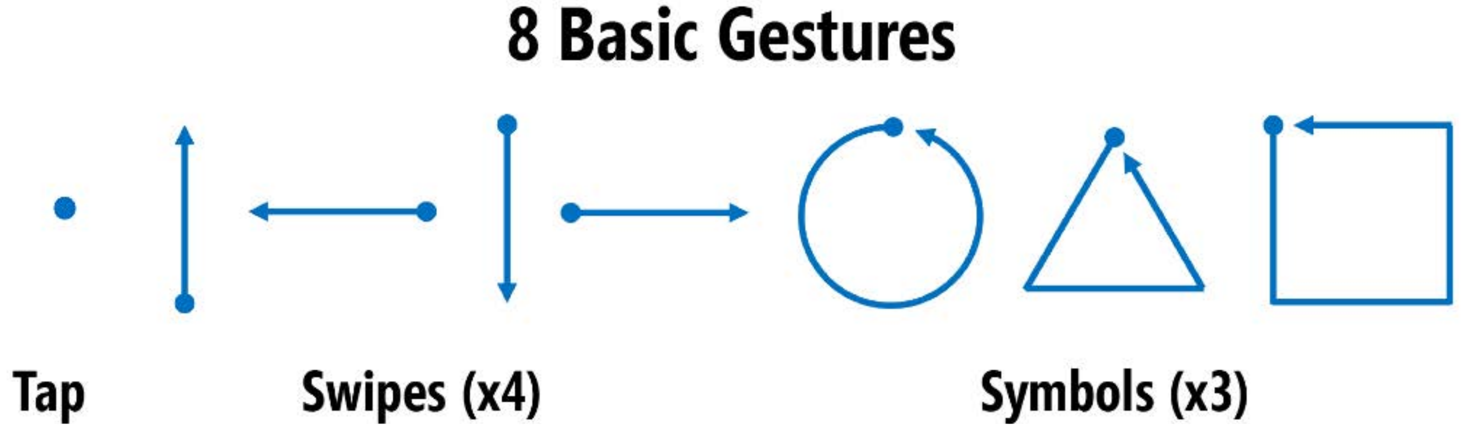
24 Sighted Participants
16 Female, 8 Male
Ages 19-51



Data Collection

24 Sighted Participants
16 Female, 8 Male
Ages 19-51

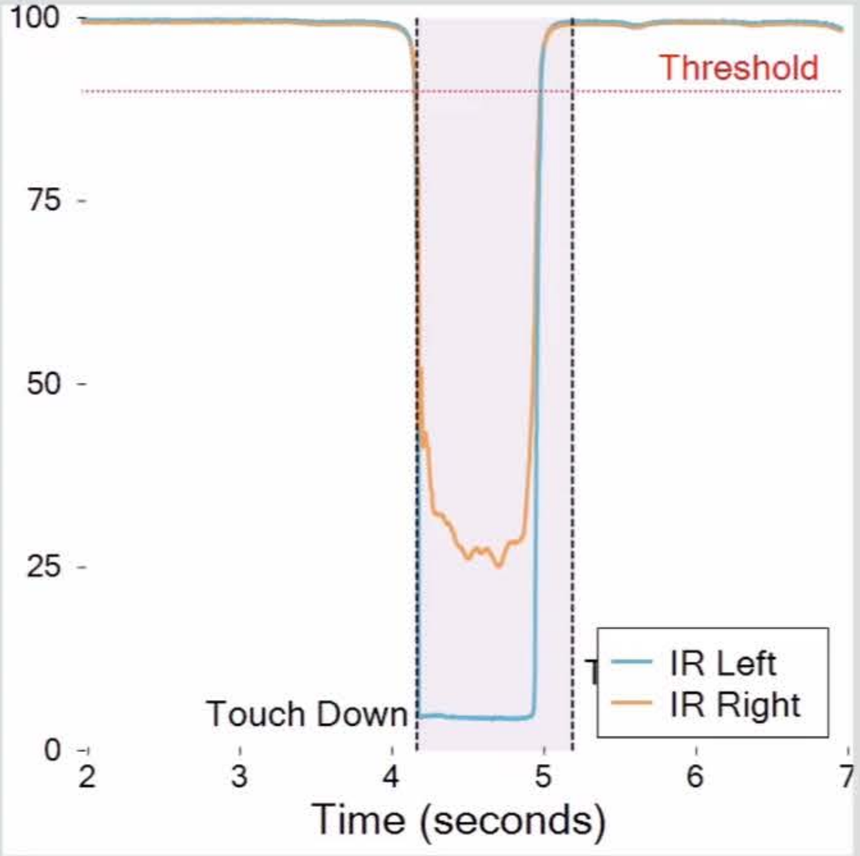




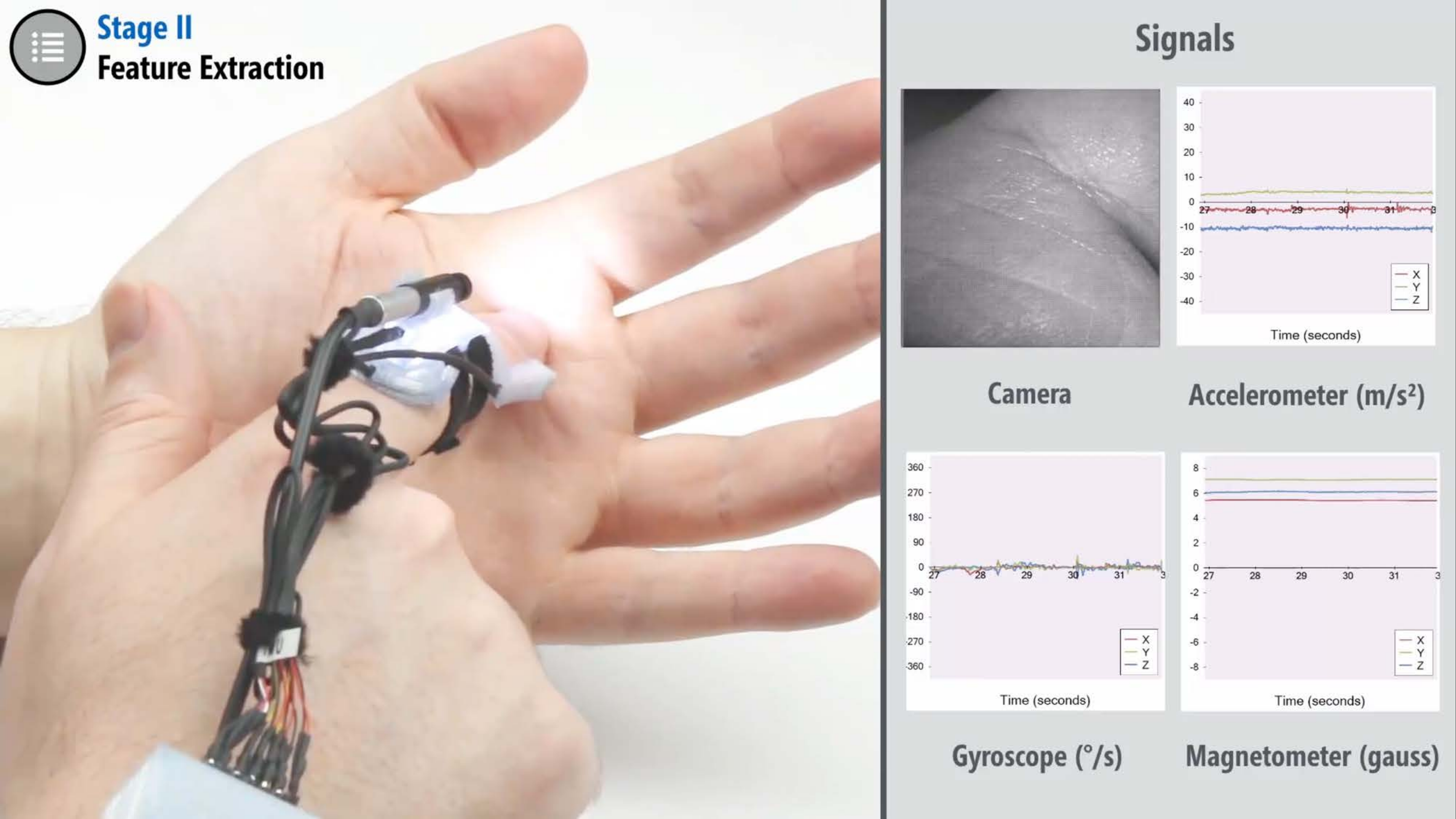
Stage I

Touch Segmentation

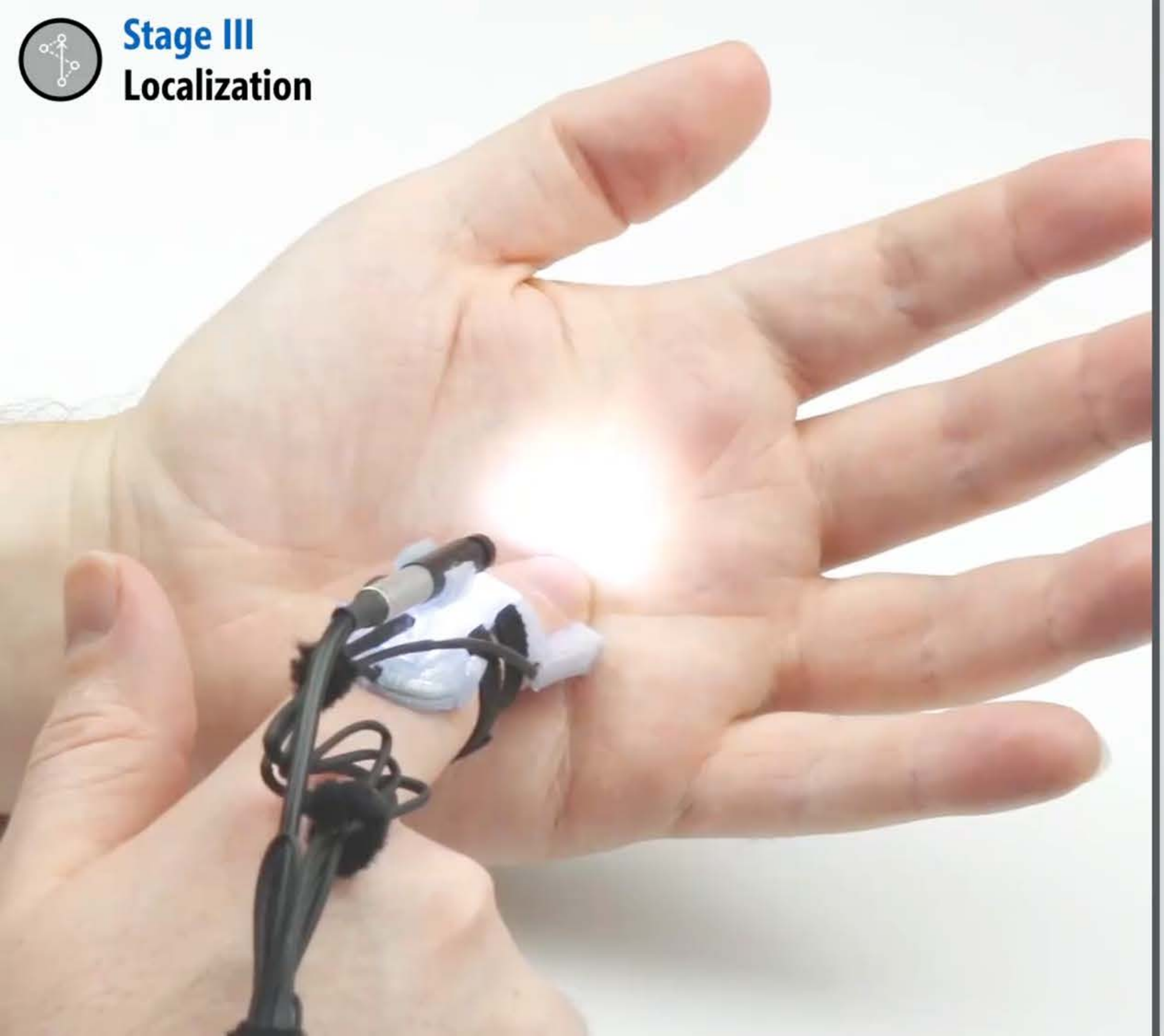
Signals



IR(% Max)

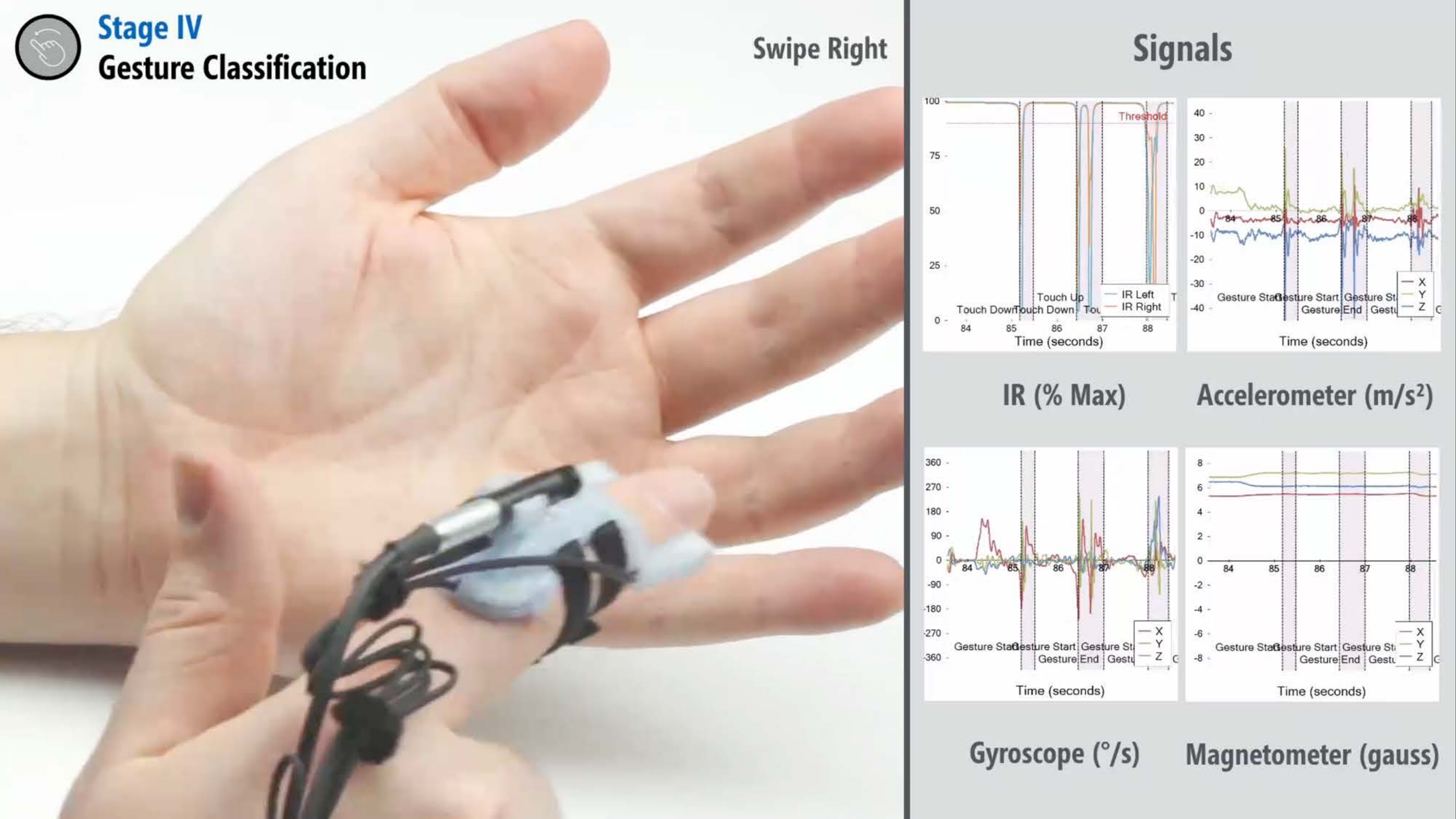

Stage II
Feature Extraction
Signals
40
30
20
10
0
-10
-20
-30
-40
27
28
29
30
31
X
Y
Z
Time (seconds)
Camera
Accelerometer (m/s²)
360
270
180
90
0
-90
-180
-270
-360
Time (seconds)
8
6
4
2
0
-2
-4
-6
-8
Time (seconds)
Gyroscope (°/s)
Magnetometer (gauss)

Stage III
Localization



Signals

Camera


Stage IV
Gesture Classification
Swipe Right
Signals
Threshold
Touch Up
Touch Down
IR Left
IR Right
Time (seconds)
Gesture Start
Gesture End
X
Y
Z
IR (% Max)
Accelerometer (m/s²)
Gyroscope (°/s)
Magnetometer (gauss)

On-Body Study I

Experiments and Results

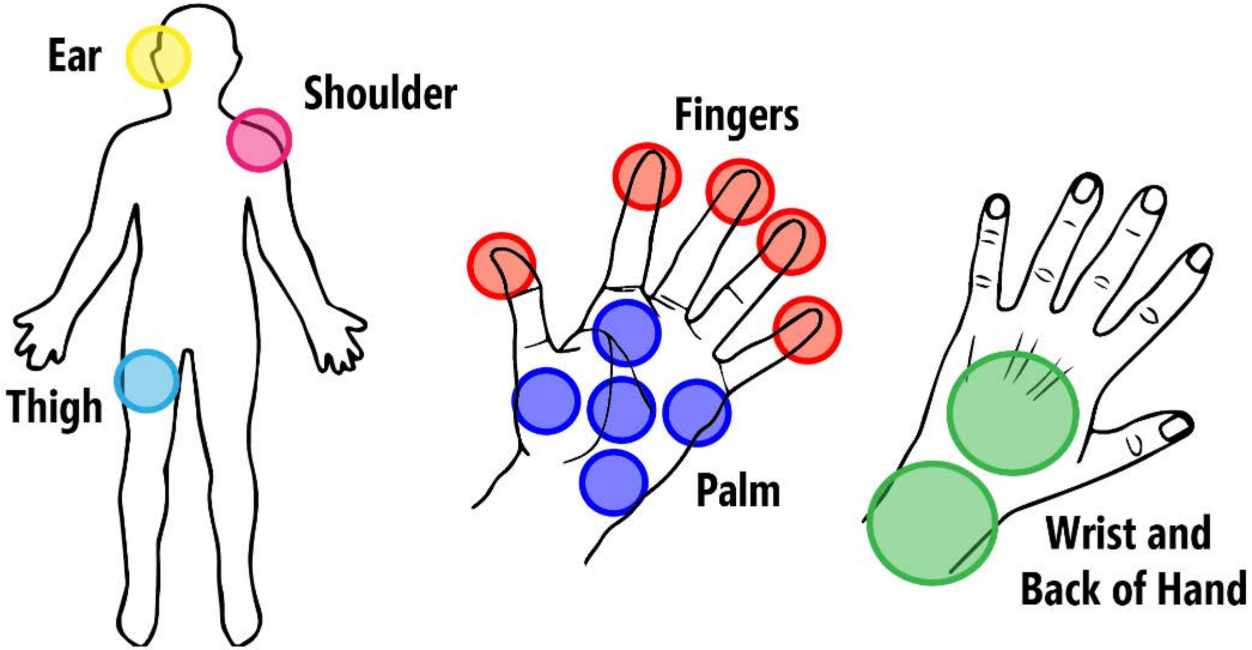
Within-Person Classification Experiment (*n*=24)

Coarse-Grained Localization

(6 classes: finger, palm, thigh, ...)
accuracy: **98.0%** (*SD*=2.3%)

Fine-Grained Localization

(15 classes: palm up, palm down, ...)
accuracy: **88.7%** (*SD*=7.0%)



ON-BODY STUDY I

EXPERIMENTS AND RESULTS

Within-Person Classification Experiment (*n*=24)

Coarse-Grained Localization

(6 classes: finger, palm, thigh, ...)

accuracy: **98.0%** (*SD*=2.3%)

Fine-Grained Localization

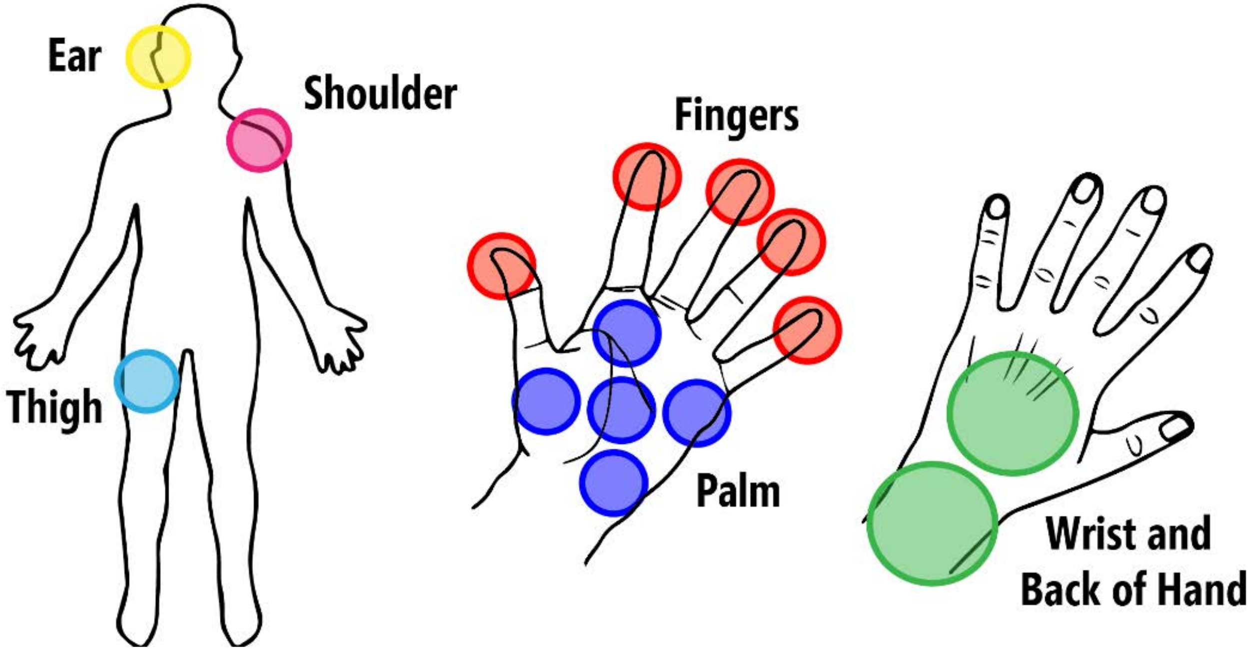
(15 classes: palm up, palm down, ...)

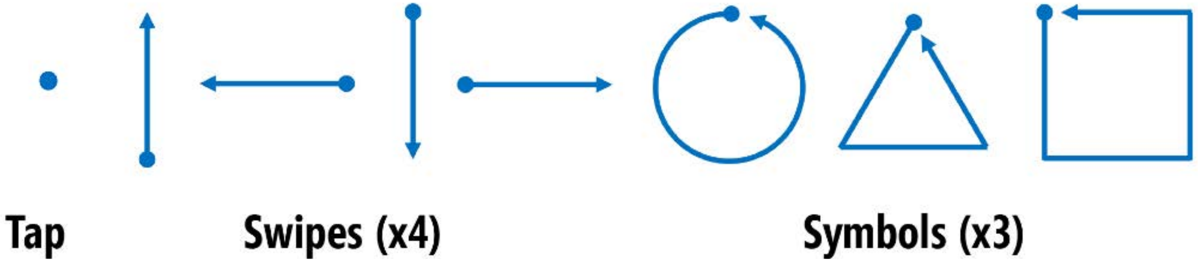
accuracy: **88.7%** (*SD*=7.0%)

Location-Specific Gesture Recognition

(24 classes: 3 locations × 8 gestures)

accuracy: **95.7%** (*SD*=3.2%)





ON-BODY STUDY III

12 Participants:
7 female, 5 male
Ages 29-65
All blind or low vision

Uran Oh, **Lee Stearns**, Alisha Pradhan, Jon E. Froehlich, Leah Findlater, **"Investigating Microinteractions for People with Visual Impairments and the Potential Role of On-Body Interaction,"** in *Proceedings of ASSETS 2017.*

Lee Stearns, Uran Oh, Leah Findlater, Jon E. Froehlich, **"TouchCam: Realtime Recognition of Location-Specific On-Body Gestures to Support Users with Visual Impairments,"** in *Interactive, Wearable and Ubiquitous Technologies, December 2018.*

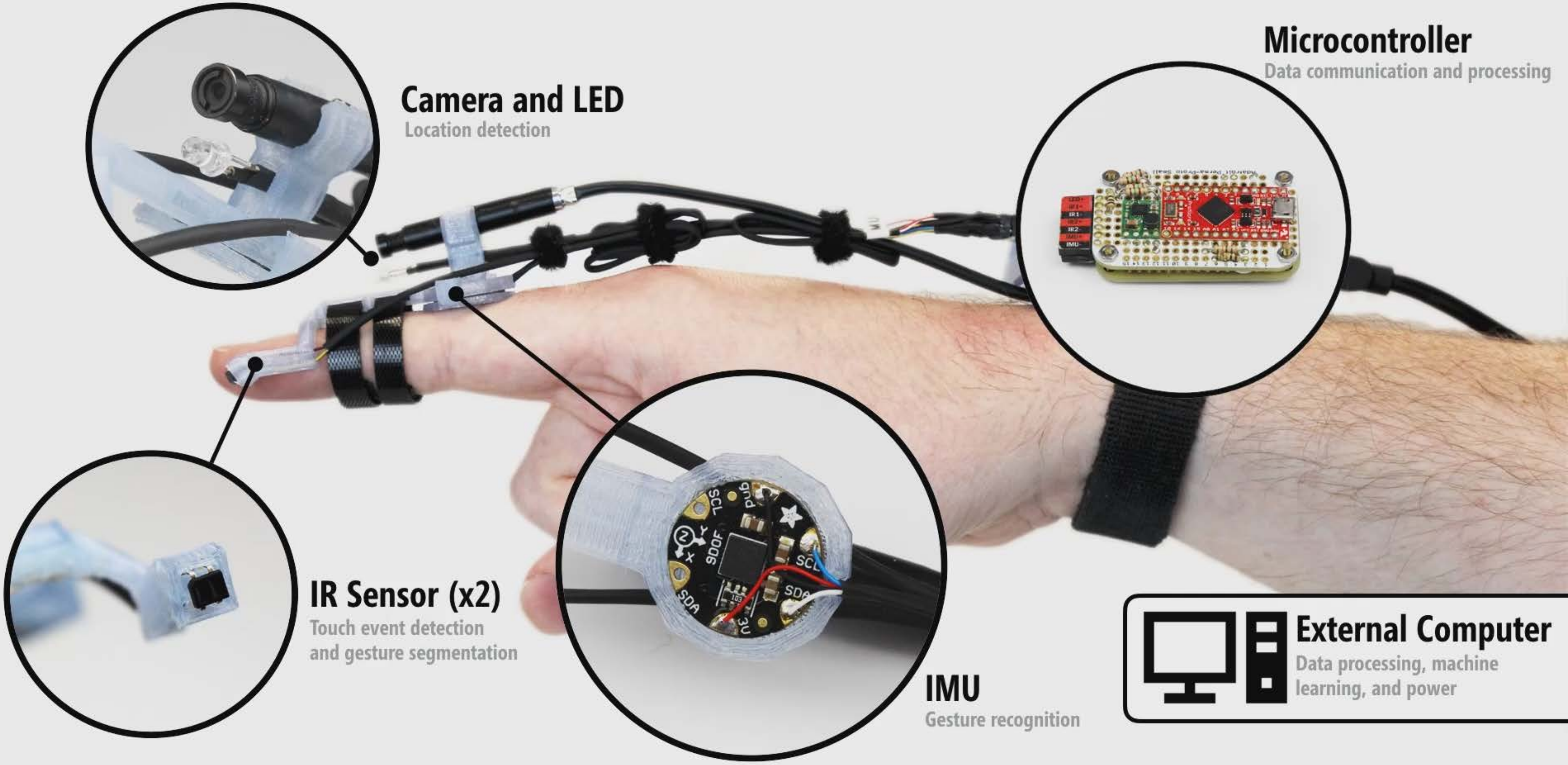
On-Body Study III

Algorithms

Real-time processing (~60fps)

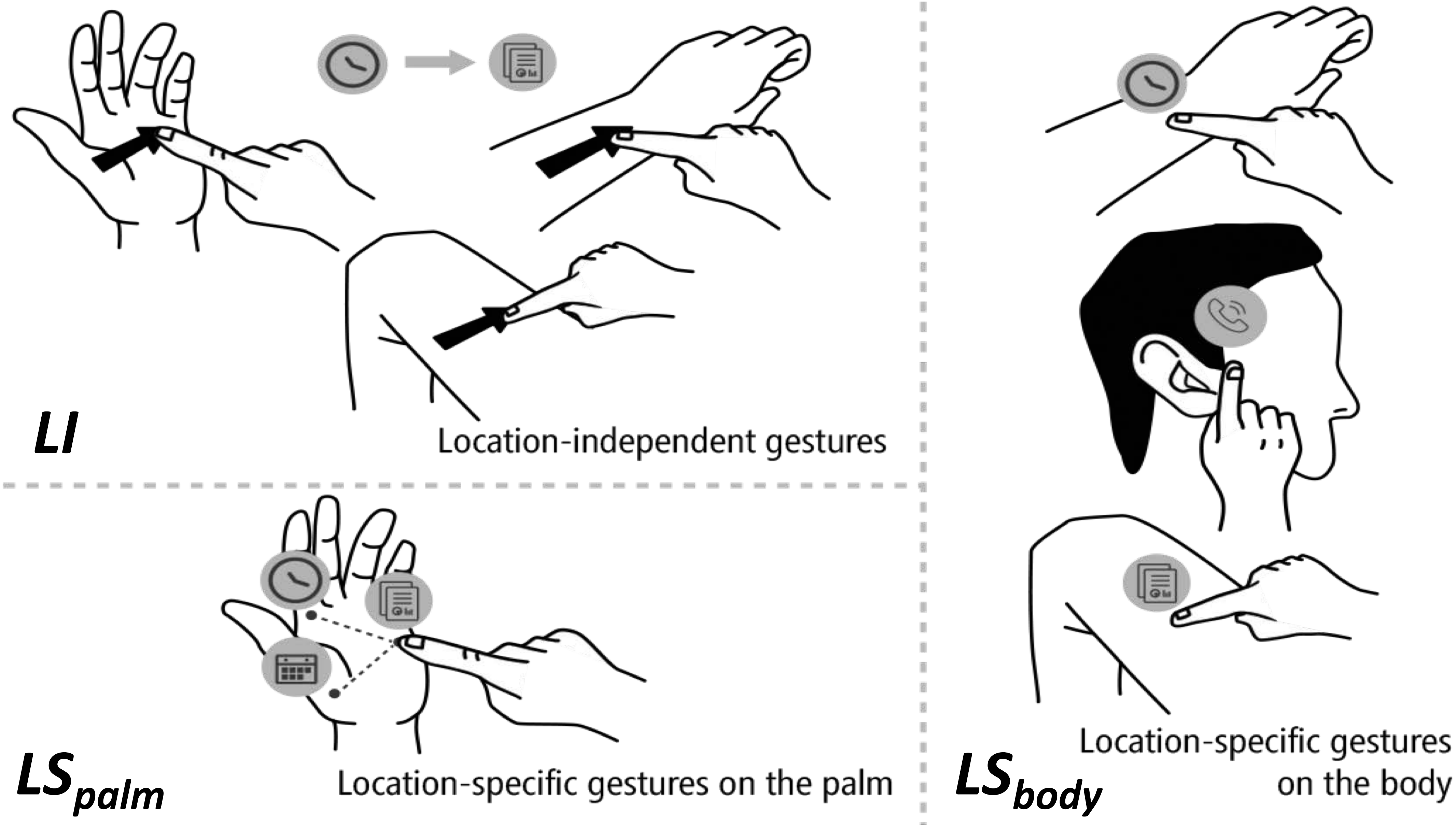
- Removed geometric verification stage (required 1-2s ***per image***)
- Combine predictions across 20 video frames (~300ms)
- Increase number of texture features per image from **1792** to **15,552**
- Reduced number of fine-grained locations (removed 5 fingertip classes)

Realtime Prototype



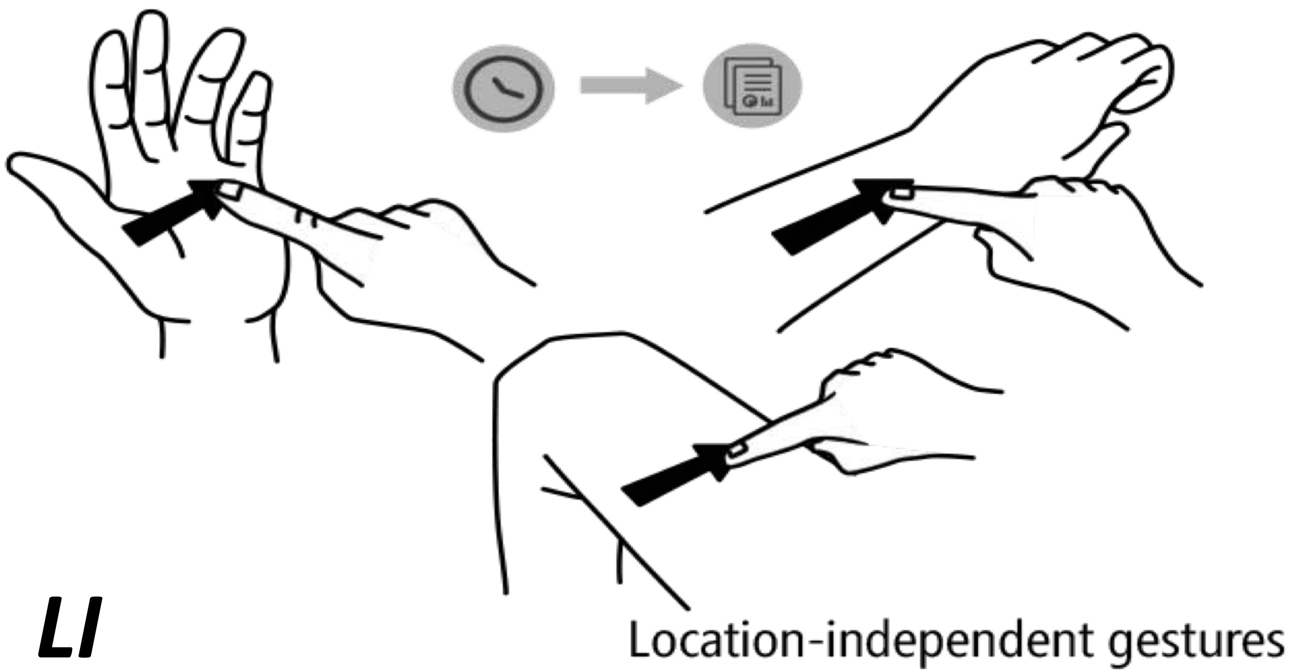
INTERFACE DESIGNS

Five applications:
Clock, Daily Summary, Notifications, Health and Fitness, Voice Input



Interface Designs

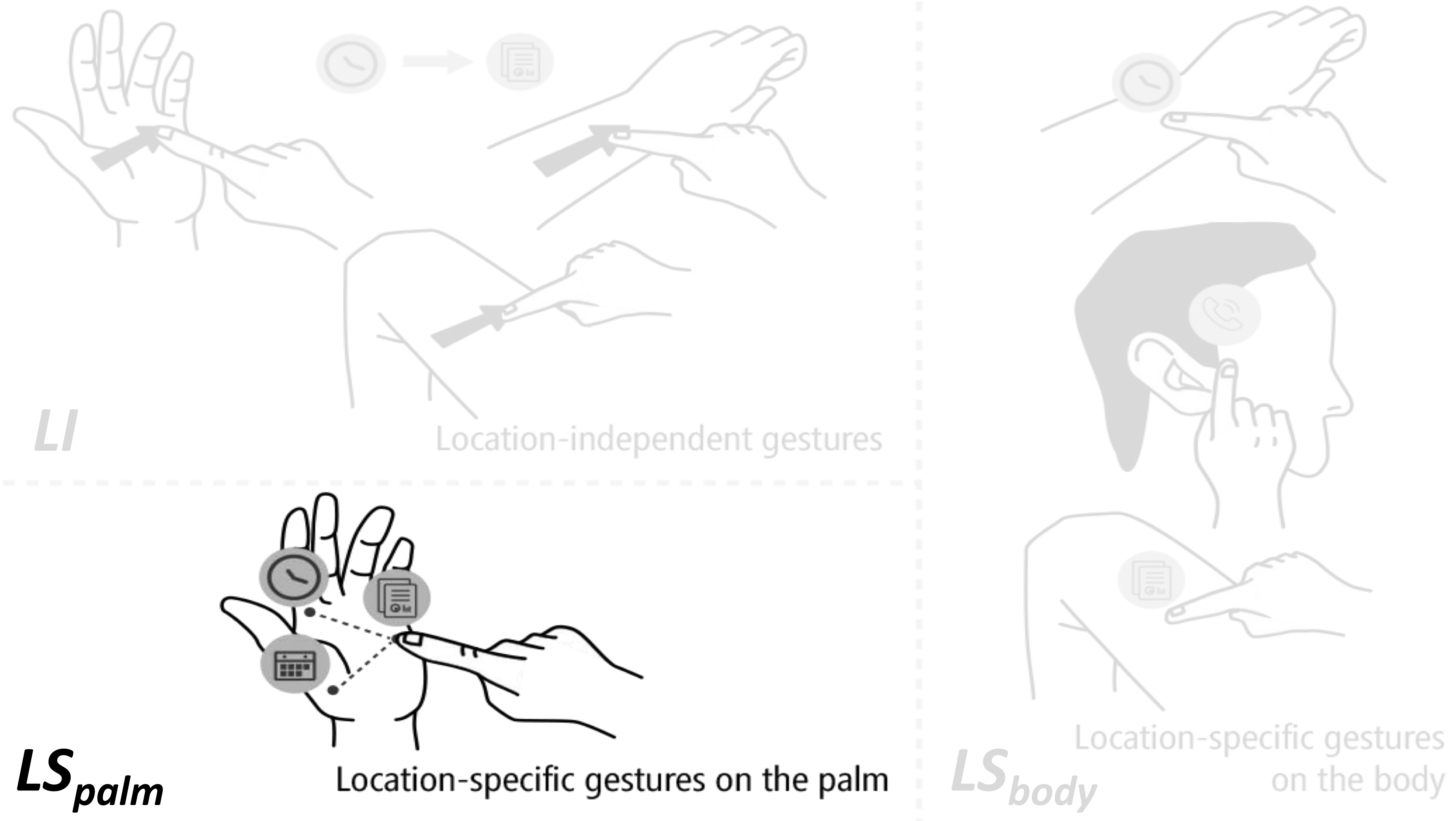
Five applications:
Clock, Daily Summary, Notifications, Health and Fitness, Voice Input




Health & Activities
Notifications
Daily Summary
Interaction Design 1: Location-Independent Gestures
Swipe left and right anywhere on the body to sequentially navigate a list of applications.

INTERFACE DESIGNS

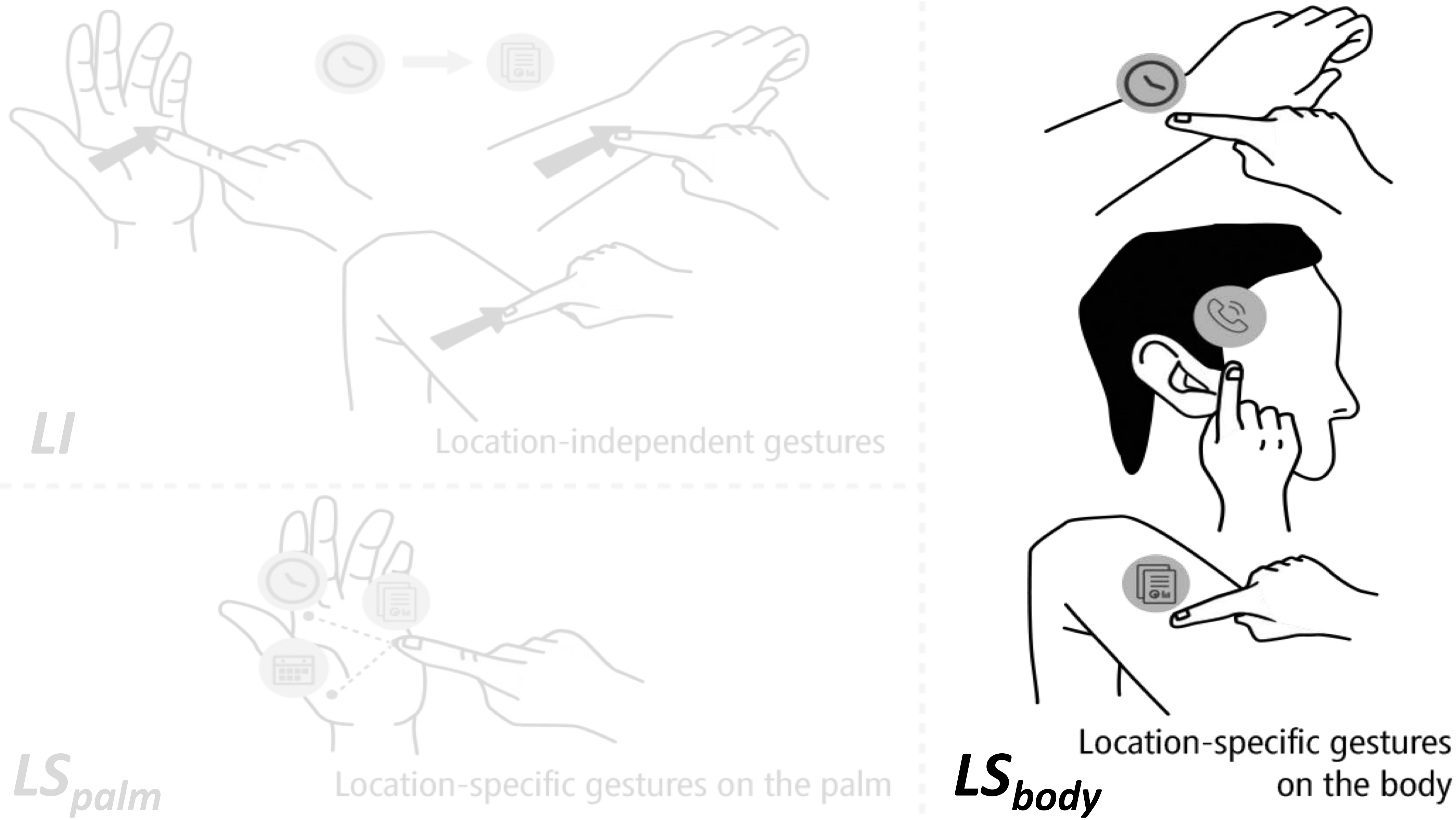
Five applications:
Clock, Daily Summary, Notifications, Health and Fitness, Voice Input

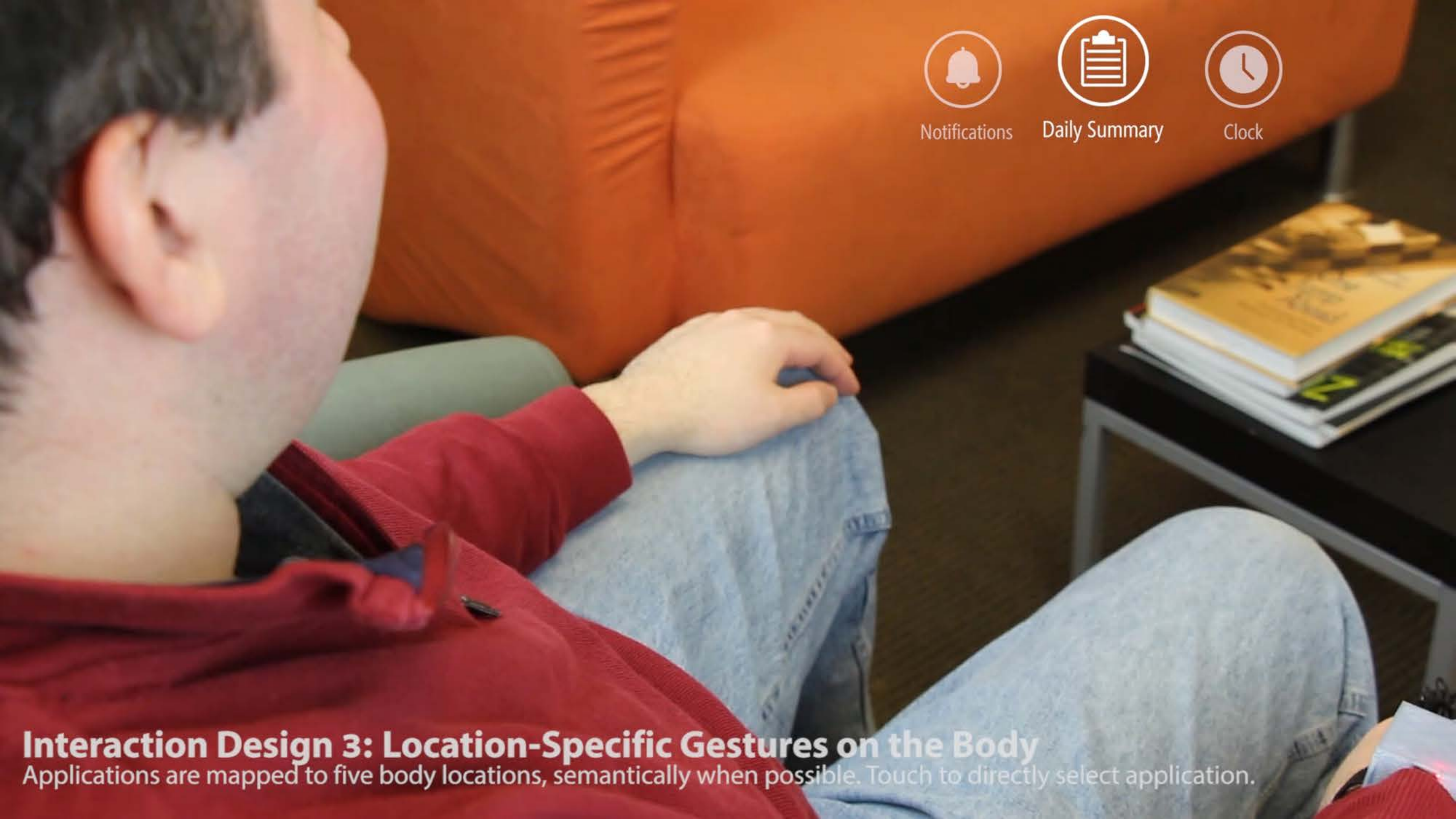



Voice Input
Health & Activities
Notifications
Interaction Design 2: Location-Specific Gestures on the Palm
Applications are mapped to five palm locations, select by directly touching a location. Supports "touch and explore".

INTERFACE DESIGNS

Five applications:
Clock, Daily Summary, Notifications, Health and Fitness, Voice Input

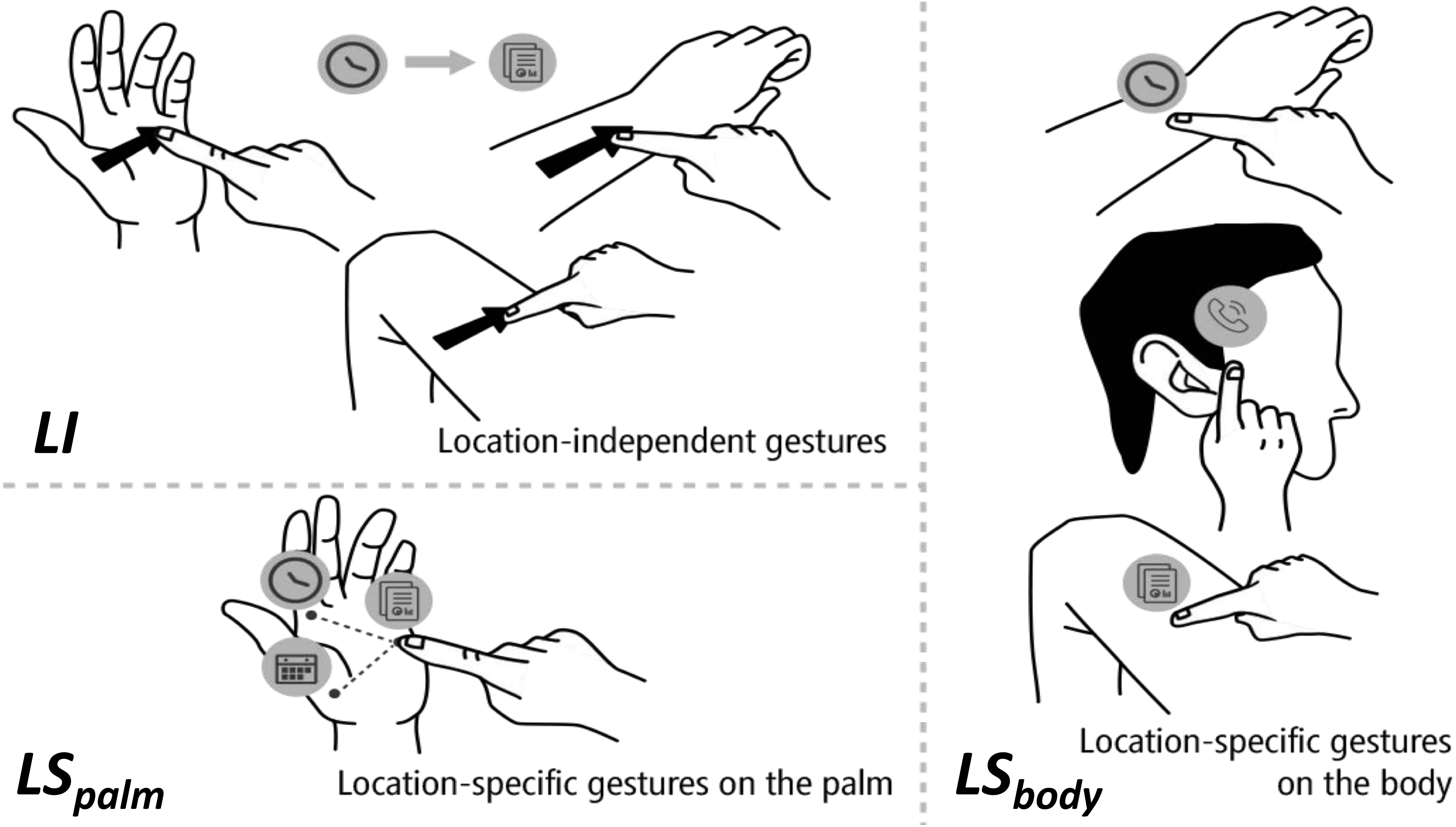




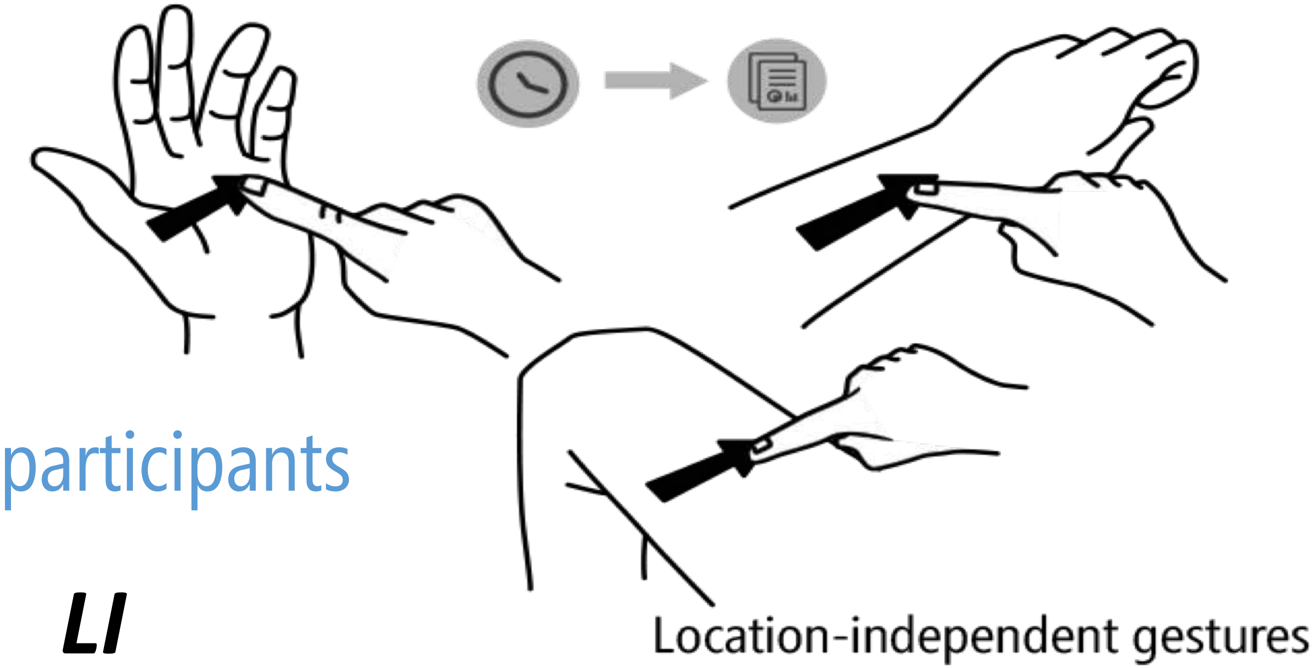
Interaction Design 3: Location-Specific Gestures on the Body

Applications are mapped to five body locations, semantically when possible. Touch to directly select application.

FINDINGS (12 Visually Impaired Participants)



FINDINGS (12 Visually Impaired Participants)



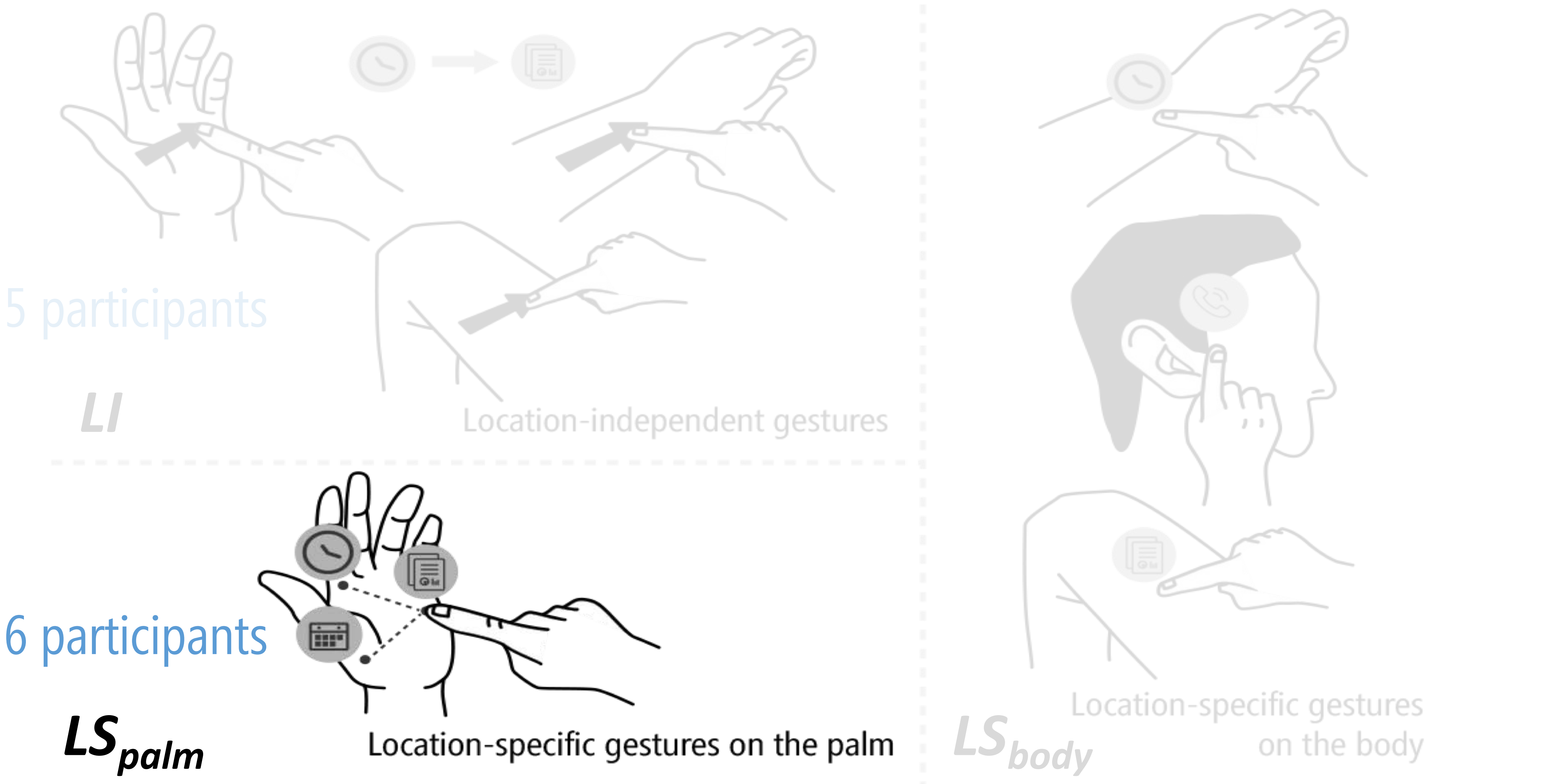
5 participants

LI

LS_{palm}

LS_{body}

FINDINGS (12 Visually Impaired Participants)



FINDINGS (12 Visually Impaired Participants)

5 participants

LI

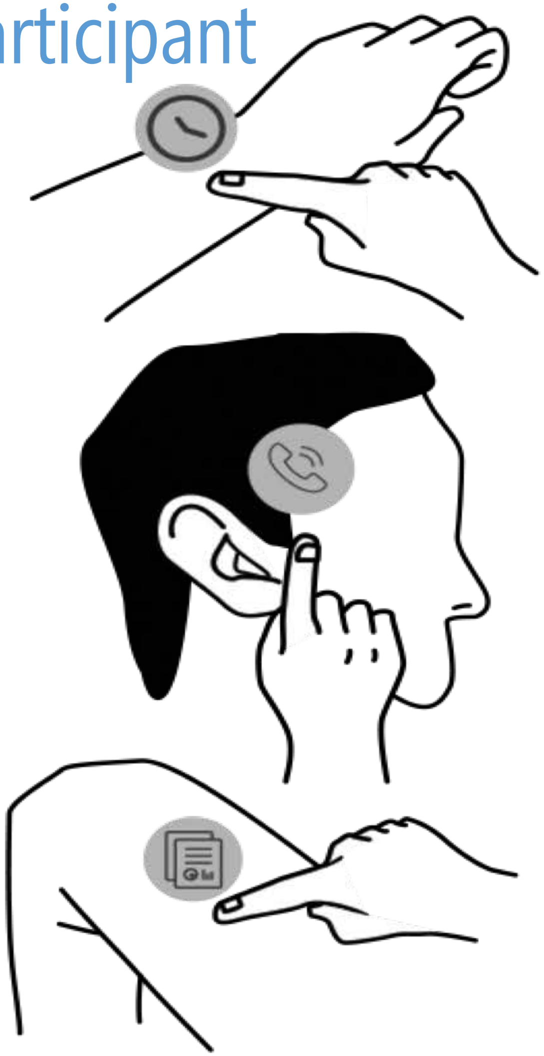
Location-independent gestures

6 participants

LS$_{palm}$

Location-specific gestures on the palm

1 participant



LS$_{body}$

Location-specific gestures on the body

Contributions

On-body input method using finger-worn sensors

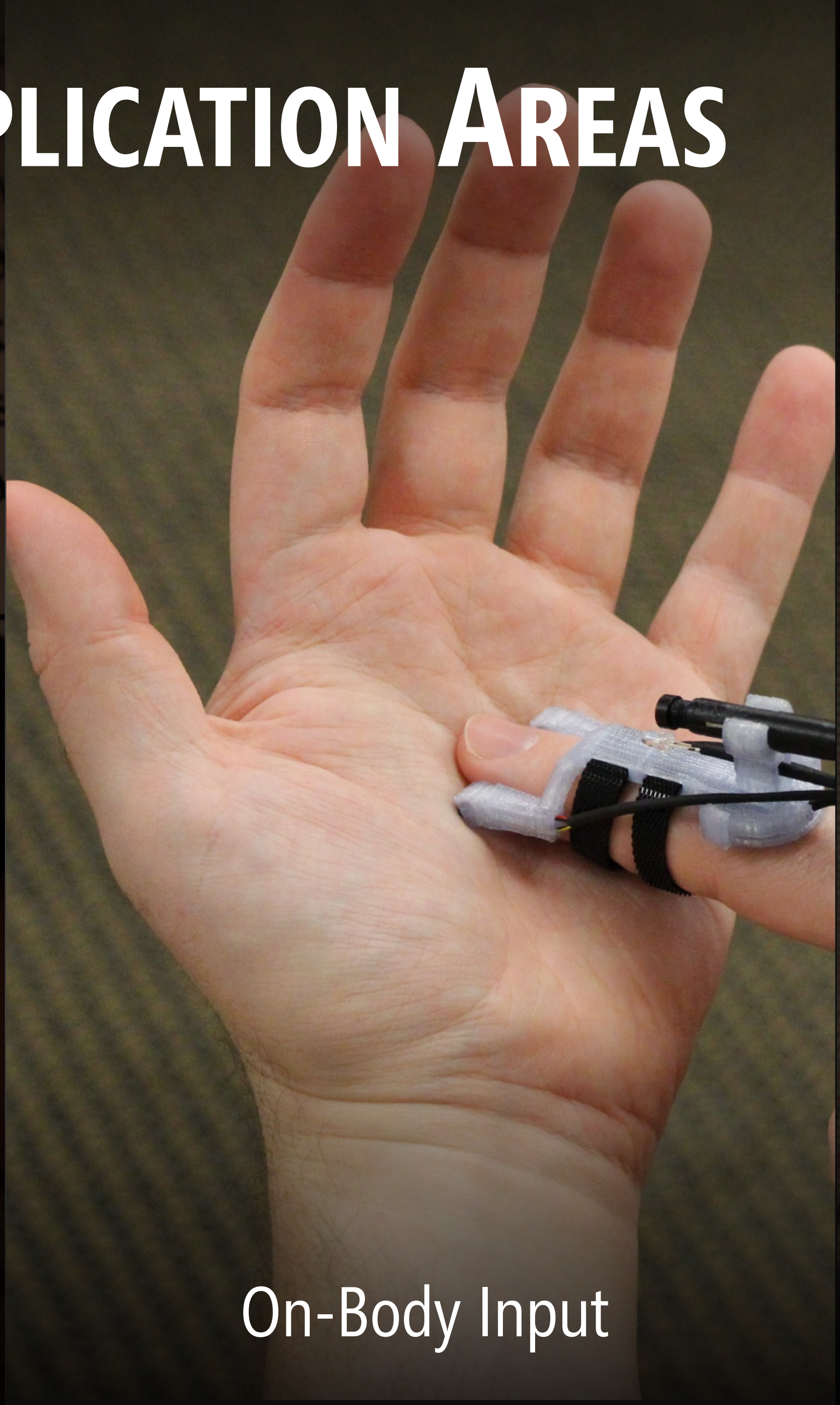
- Mitigates camera framing issues
- New types of contextual, **location-specific gestures**

Our findings demonstrate feasibility, with **high accuracy** and **realtime performance**

We identified **tradeoffs** that will impact the design of future **on-body interfaces** (*e.g.*, accuracy, usability)

HandSight Application Areas

Reading/Exploring Text



On-Body Input

HandSight Application Areas

Reading/Exploring Text

On-Body Input

Clothing Colors/Patterns

Identifying Clothing Colors and Patterns

Alexander Medeiros, **Lee Stearns**, Leah Findlater, Chuan Chen, Jon E. Froehlich, **“Recognizing Clothing Colors and Textures using a Finger-Mounted Camera: An Initial Investigation,"** in *ASSETS 2017 (Poster Track).*

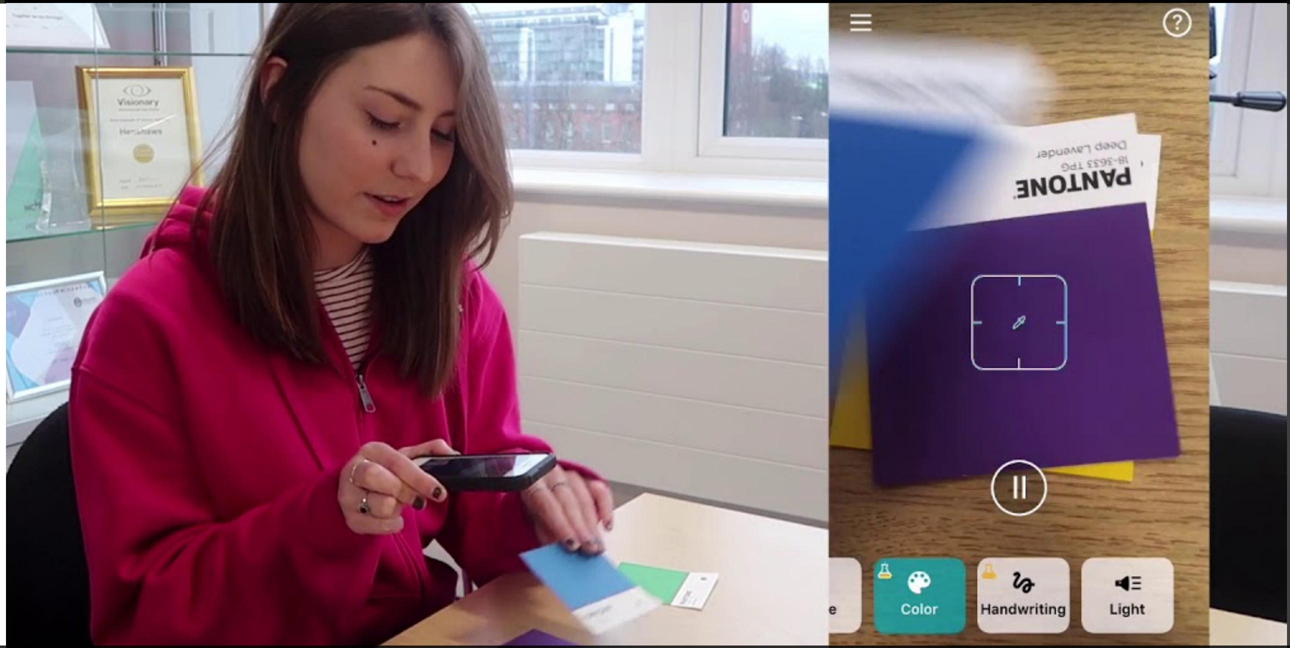
Lee Stearns, Leah Findlater, Jon E. Froehlich, **“Applying Transfer Learning to Recognize Clothing Patterns Using a Finger-Mounted Camera,"** in *ASSETS 2018 (Poster Track, To Appear).*

Commercial Color Recognizers

Brytec Color Teller

LedScope Android App

Colorino Color Identifier



Microsoft SeeingAI App

Commercial Color Recognizers

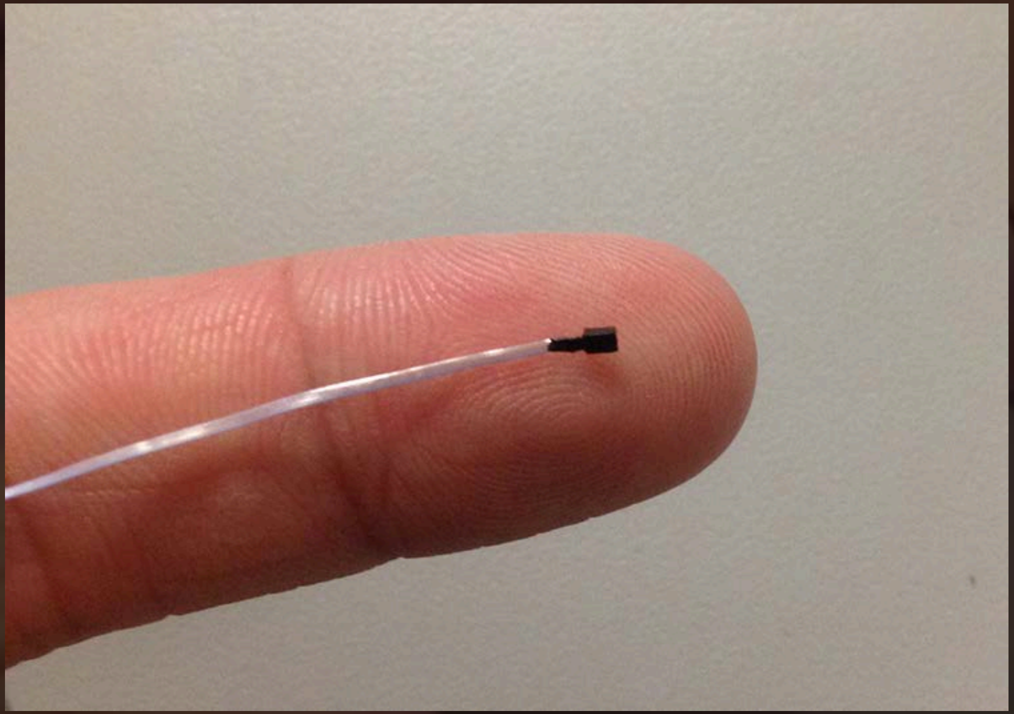
Limitations:

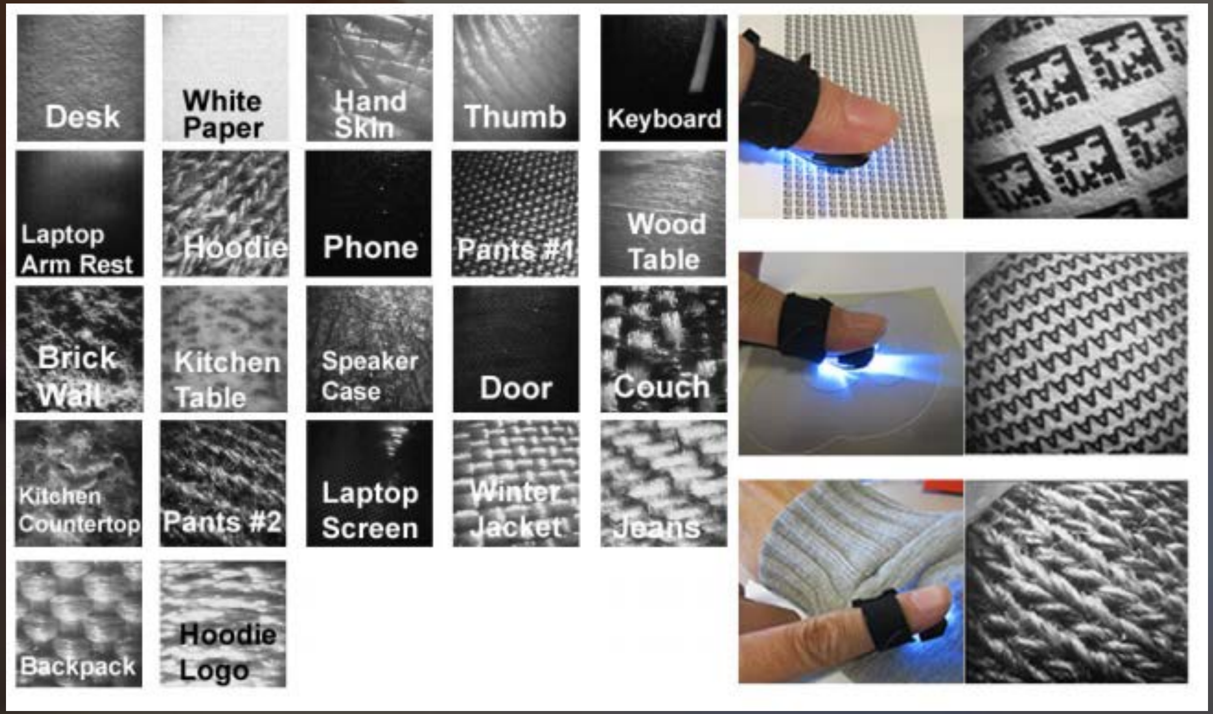
- Cannot recognize **patterns**, only **color**
- Do not allow users to quickly inspect **multiple locations**
- Accuracy affected by **lighting and distance**

Other Related Work

Access Lens: S. Kane, *et al.*, "Access lens: a gesture-based screen reader for real-world documents," in Proceedings of CHI 2013.

Other Related Work





Magic Finger: X.D. Yang, *et al.*, "Magic finger: always-available input through finger instrumentation," in Proceedings of UIST 2012.

OTHER RELATED WORK

X. Yang, *et al.*, "**Assistive Clothing Pattern Recognition for Visually Impaired People**," in *IEEE Trans. on Human-Machine Systems*, 2014.

Identifying Clothing
Colors and Patterns

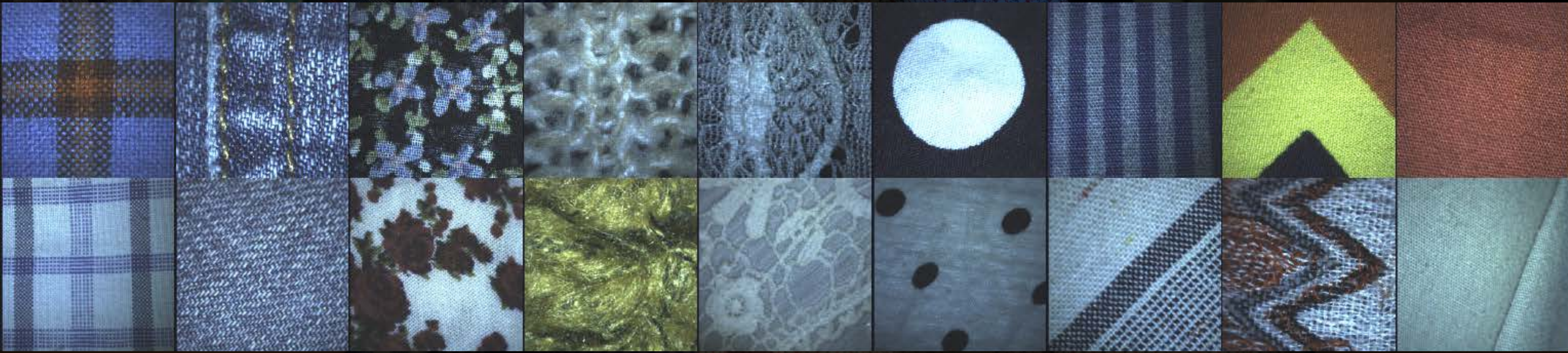
Advantages of HandSight:

- Can recognize visual patterns as well as color
- Constrains distance and lighting for consistent results
- Allows for interactive exploration of colors and textures

HandSight Clothing Pattern
Dataset

Collected **520 images** across **29 articles of clothing** and **9 categories of pattern**

Controlled and varied the **distance** (5cm *vs.* 12cm), **rotation** (0° *vs.* 45°), **perspective** of the camera (90° *vs.* 45°), and the **tension** of the fabric (taut *vs.* hanging naturally)



Checkered | Denim | Floral | Knitted | Lacelike | Polka-dotted | Striped | Zigzagged | None

Pattern Classification

Initial Exploration

Visual Texture Recognition Algorithms

* Adapted from Cimpoi, *et al.*, 2012

1. Deep convolutional activation features (DeCAF)

   Adapt a pretrained object classifier from the **ImageNet Large Scale Visual Recognition Challenge**, removing last two layers used for classification and using the outputs as a raw feature vector (**4096 DeCAF features**).

Pattern Classification

Initial Exploration

Visual Texture Recognition Algorithms

* Adapted from Cimpoi, *et al.*, 2012

1. Deep convolutional activation features (DeCAF)

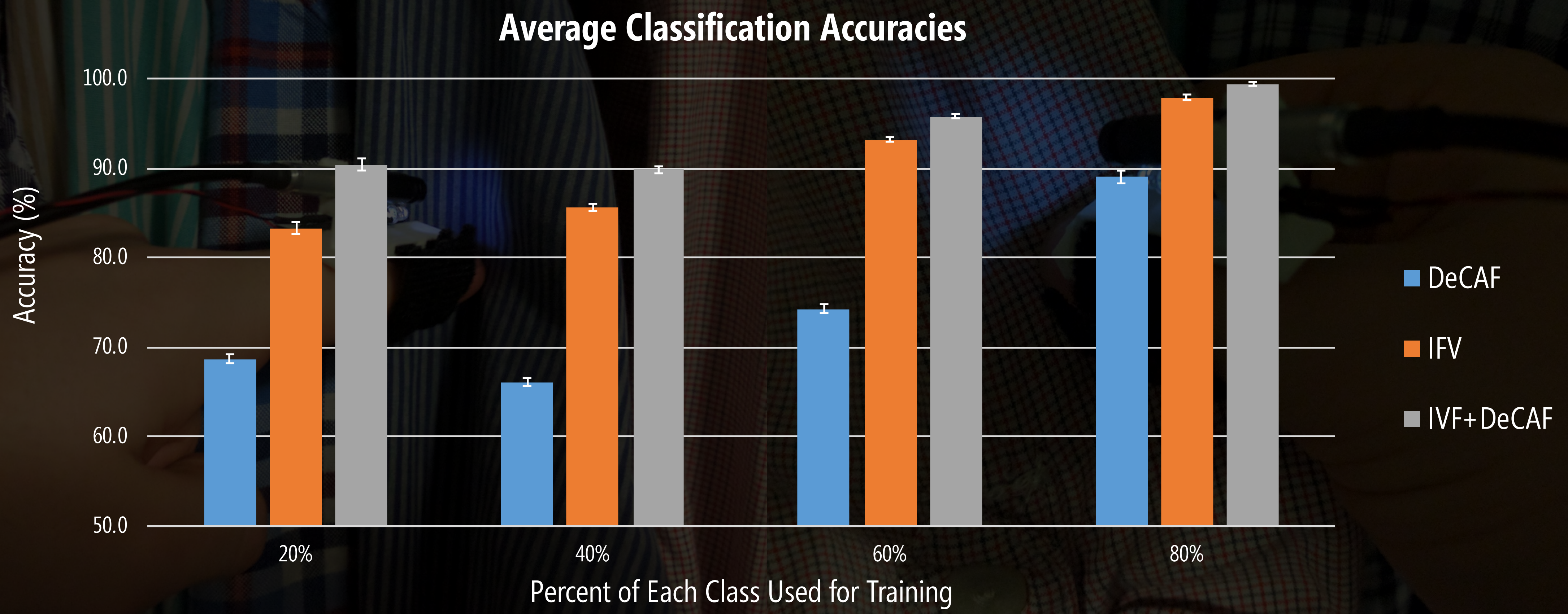
   Adapt a pretrained object classifier from the **ImageNet Large Scale Visual Recognition Challenge**, removing last two layers used for classification and using the outputs as a raw feature vector (**4096 DeCAF features**).

2. Dense SIFT features combined in an Improved Fisher Vector (IFV)

   Extract scale invariant features densely at multiple scales, then combine using the **Improved Fisher Vector** formulation. Results in a complementary set of features that captures important texture information (**40,960 IFV features**)

Pattern Classification

Initial Exploration



Pattern Classification

Initial Exploration

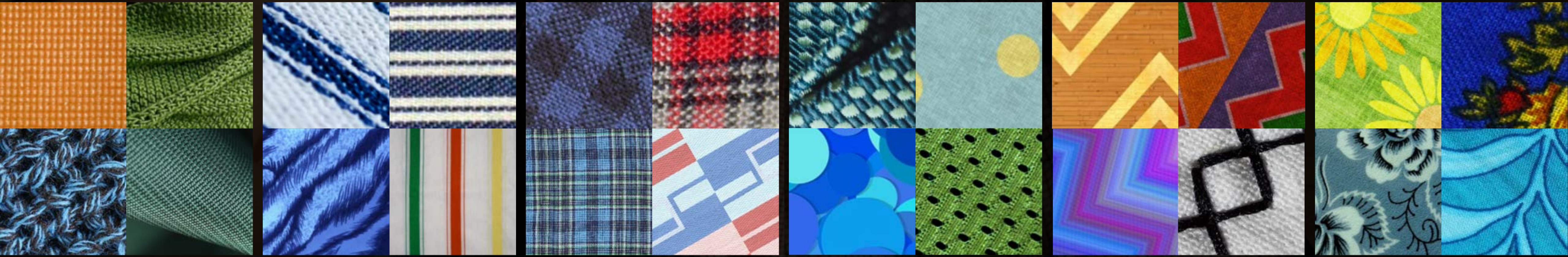
Highly controlled dataset—**risks overfitting**, **limits robustness**

Training process **not easily scalable**

Online Clothing Pattern Dataset

Built a larger and more varied dataset of images downloaded from **Google Images**

Focused only on **fabric patterns** that cannot easily be **distinguished by touch**

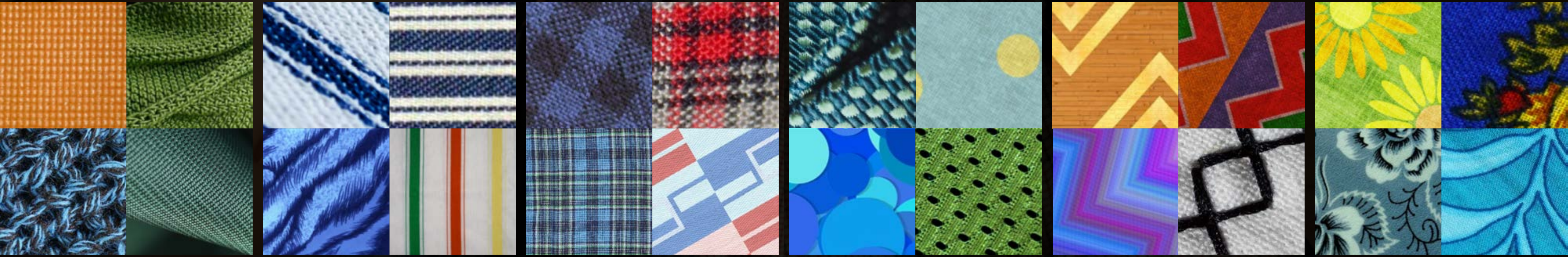


Solid Striped Checkered Dotted Zigzag Floral

Available online: https://github.com/lstearns86/clothing-pattern-dataset/

Online Clothing Pattern Dataset

- Built a larger and more varied dataset of images downloaded from **Google Images**
- Focused only on **fabric patterns** that cannot easily be **distinguished by touch**
- Downloaded **top 1000 images** for each category (*e.g.*, **“striped fabric”**)
- After removing erroneous results, duplicates, and cropping logos/backgrounds, contained **317–584 images per class (2764 total)**



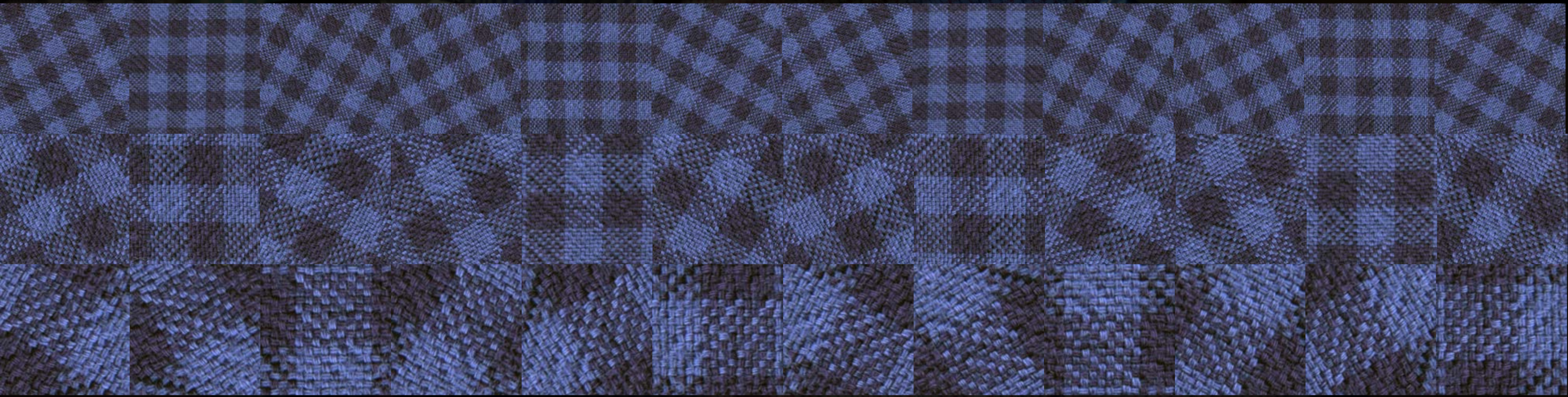
Solid | Striped | Checkered | Dotted | Zigzag | Floral

Available online: https://github.com/lstearns86/clothing-pattern-dataset/

Online Clothing Pattern Dataset

Synthetic variations: **rotation** (30° increments), **scales** (1–4, depending on resolution)

Final dataset: 8,232–17,304 samples per class, 77,052 total



Available online: https://github.com/lstearns86/clothing-pattern-dataset/

* Simplified visualization, actually used ResNet-101 architecture

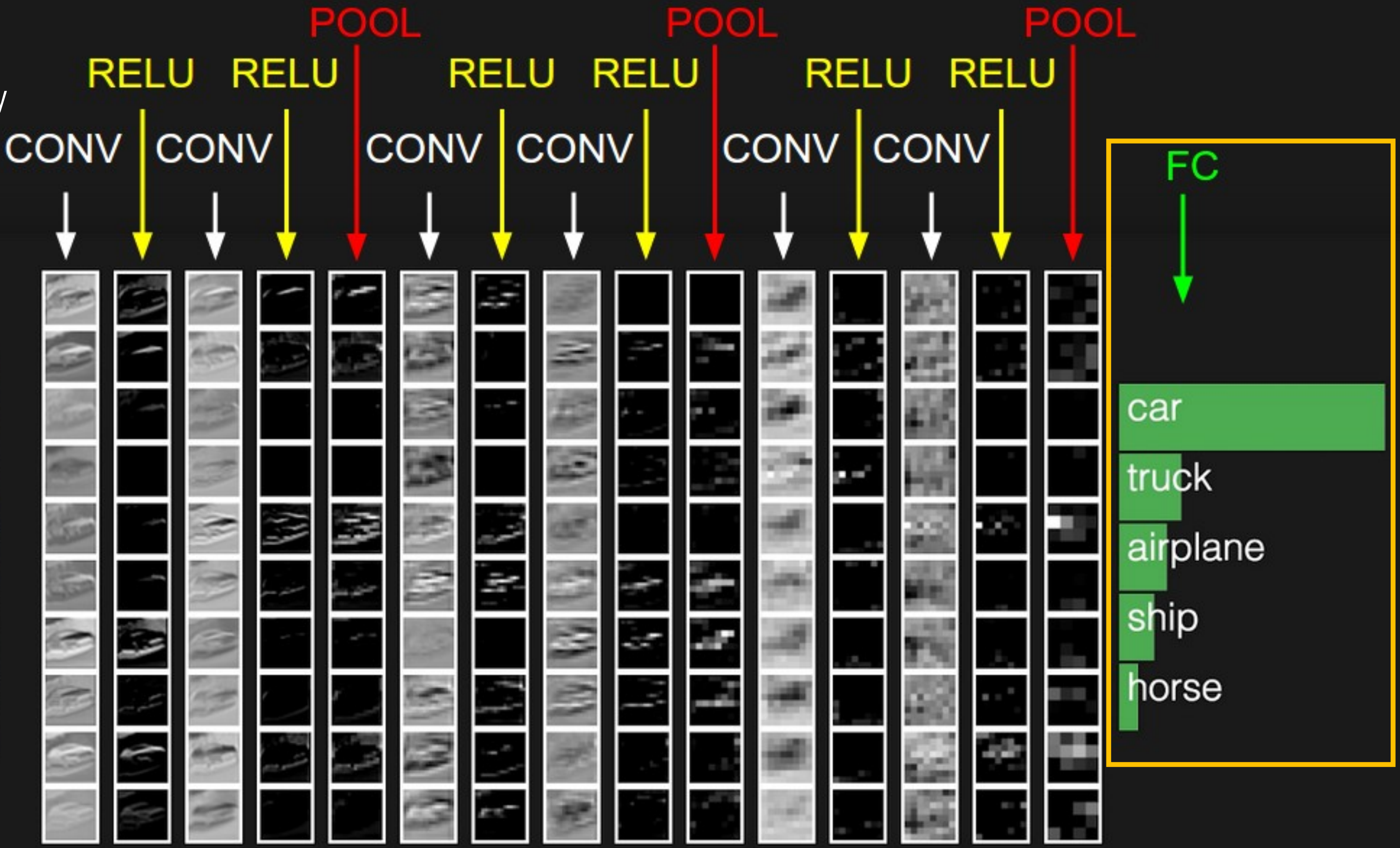


Image from Andrej Karpathy's *"Convolutional Neural Networks for Visual Recognition"* course, http://cs231n.github.io/convolutional-networks/

Pattern Classification

An End-to-End Deep Learning Approach

Randomly sampled 6400 images per class for **training**, and 1600 for **testing**

Validation: classification accuracy on the test set was **91.7%**

Accuracy on HandSight dataset (400 images) was **72.8%**.

PATTERN CLASSIFICATION

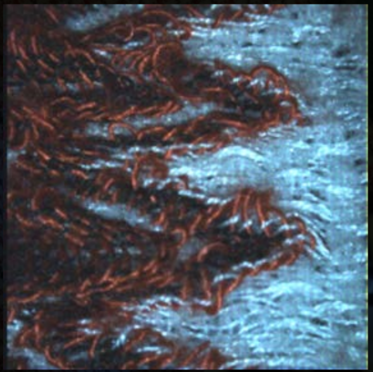
AN END-TO-END DEEP LEARNING APPROACH

Randomly sampled 6400 images per class for **training**, and 1600 for **testing**

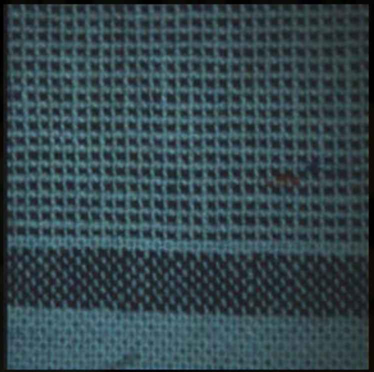
Validation: classification accuracy on the test set was **91.7%**

Accuracy on HandSight dataset (400 images) was **72.8%**.

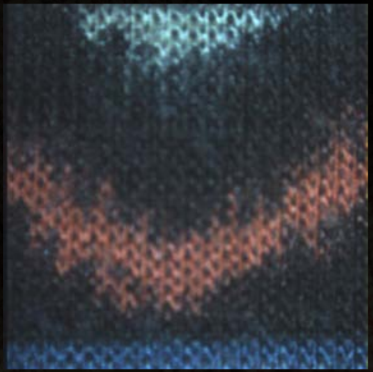
Fine-tuning the classifier with ~half of the HandSight images (N=36 per class) increases accuracy to **96.5%**



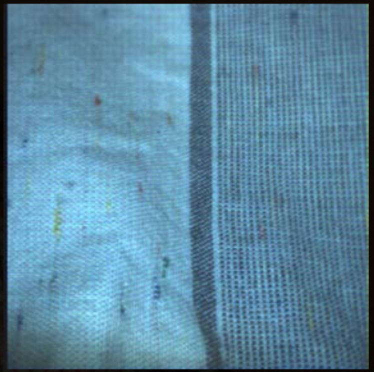
Zigzag → Floral



Striped → Checkered



Zigzag → Checkered



Striped → Solid

Ongoing Work

Identifying Clothing Colors

Identify **multiple colors** in a single image

Two methods: **K-means clustering** and **superpixel segmentation**

Original k=2 k=4

Original Superpixels Color clusters

Ongoing Work

Identifying Clothing Colors

What is the best way to **convey** this information to **users**?

User-configurable level of detail:

- Color names
- Number of colors
- Frequency of feedback

Contributions

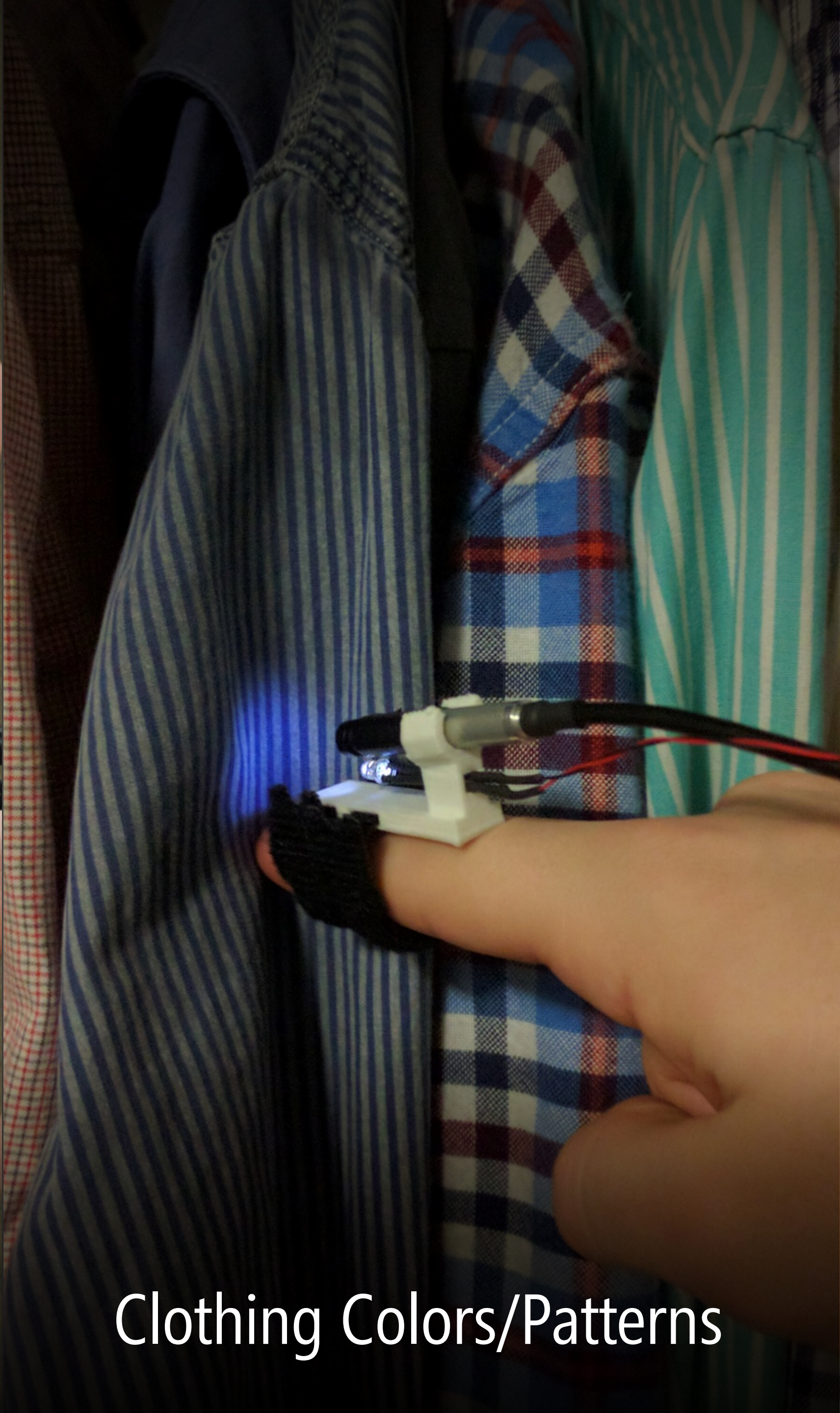
- Two datasets of **fabric pattern images**
    - **529 images** collected systematically using HandSight.
    - **77,052 images** assembled from online sources and augmented synthetically (rotations, scaling, and cropping).
- **Preliminary algorithmic results** demonstrating the feasibility of recognizing clothing patterns with a finger-worn camera.

Conclusions

And Future Research Directions

Reading/Exploring Text

On-Body Input



Clothing Colors/Patterns

Conclusions

And Future Research Directions

Alternative or Supplementary Camera Locations

Conclusions
And Future Research Directions

Camera on the User's Finger or Wrist	Camera on the User's Upper Body
Enables interactions anywhere the user can touch	Wider field of view, more contextual information
Reduced issues with occlusion	Easier to localize and track hand/finger position
Better resolution, more detail at the touch location	Can use larger, higher-fidelity hardware
Simplified processing of content beneath finger	Less likely to interfere with motion or touch sensitivity
Easier to recognize relative gestures	

Alternative or Supplementary Camera Locations

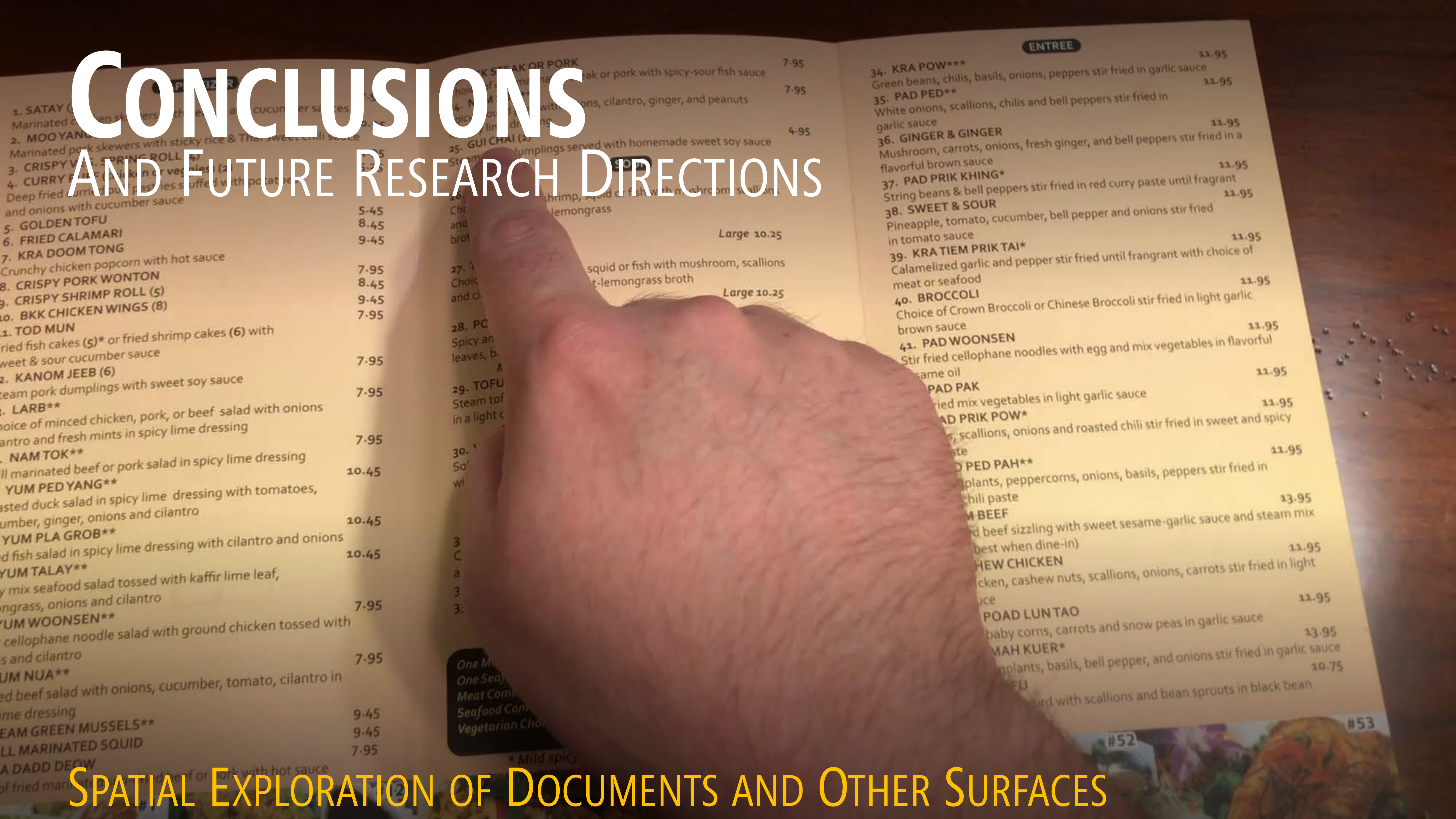
Conclusions

And Future Research Directions

Spatial Exploration of Documents and Other Surfaces

Conclusions

And Future Research Directions



Spatial Exploration of Documents and Other Surfaces

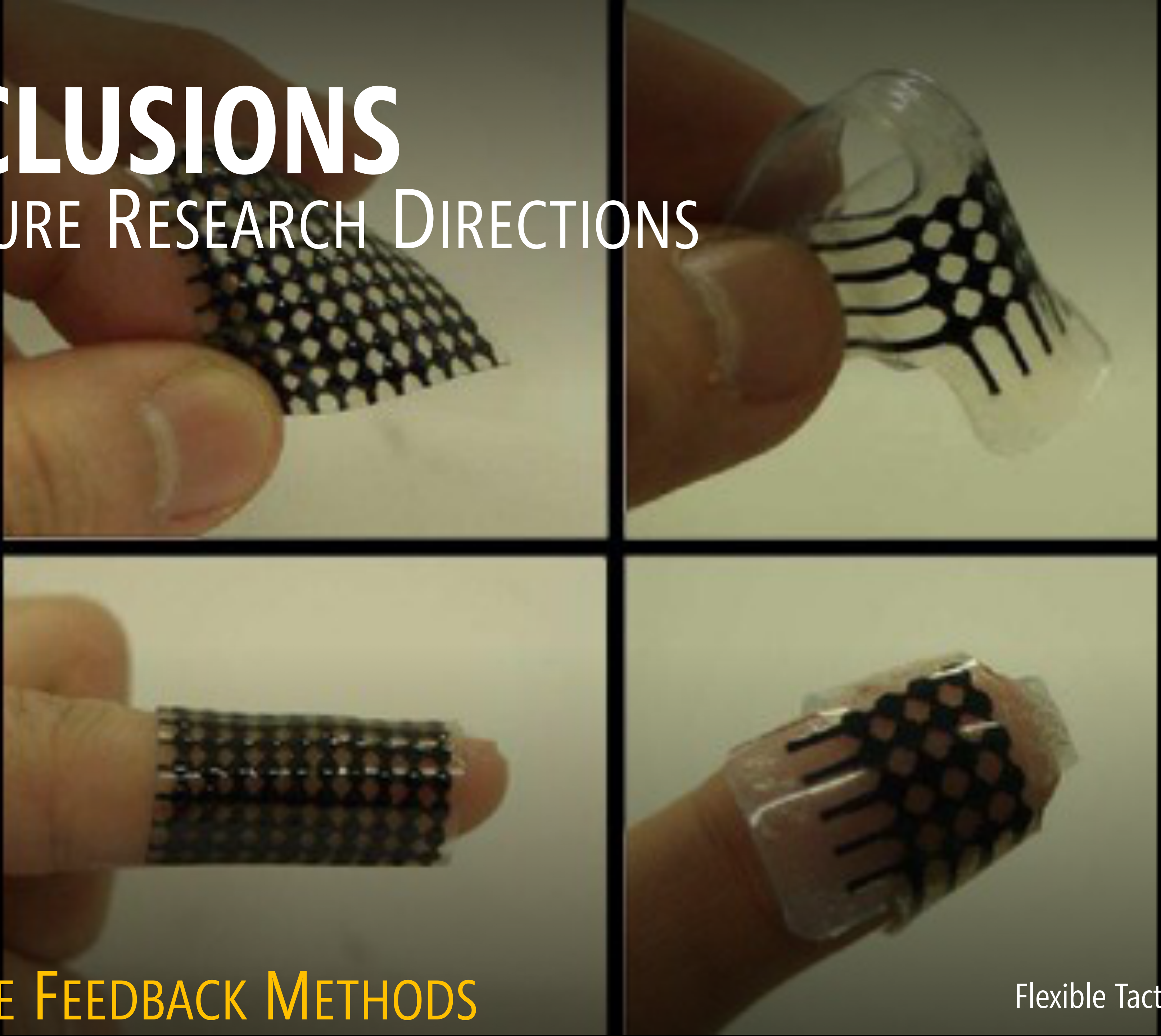

Conclusions
And Future Research Directions
Spatial Exploration of Documents and Other Surfaces

Conclusions

And Future Research Directions

TeslaTouch, Xu *et al.* 2011

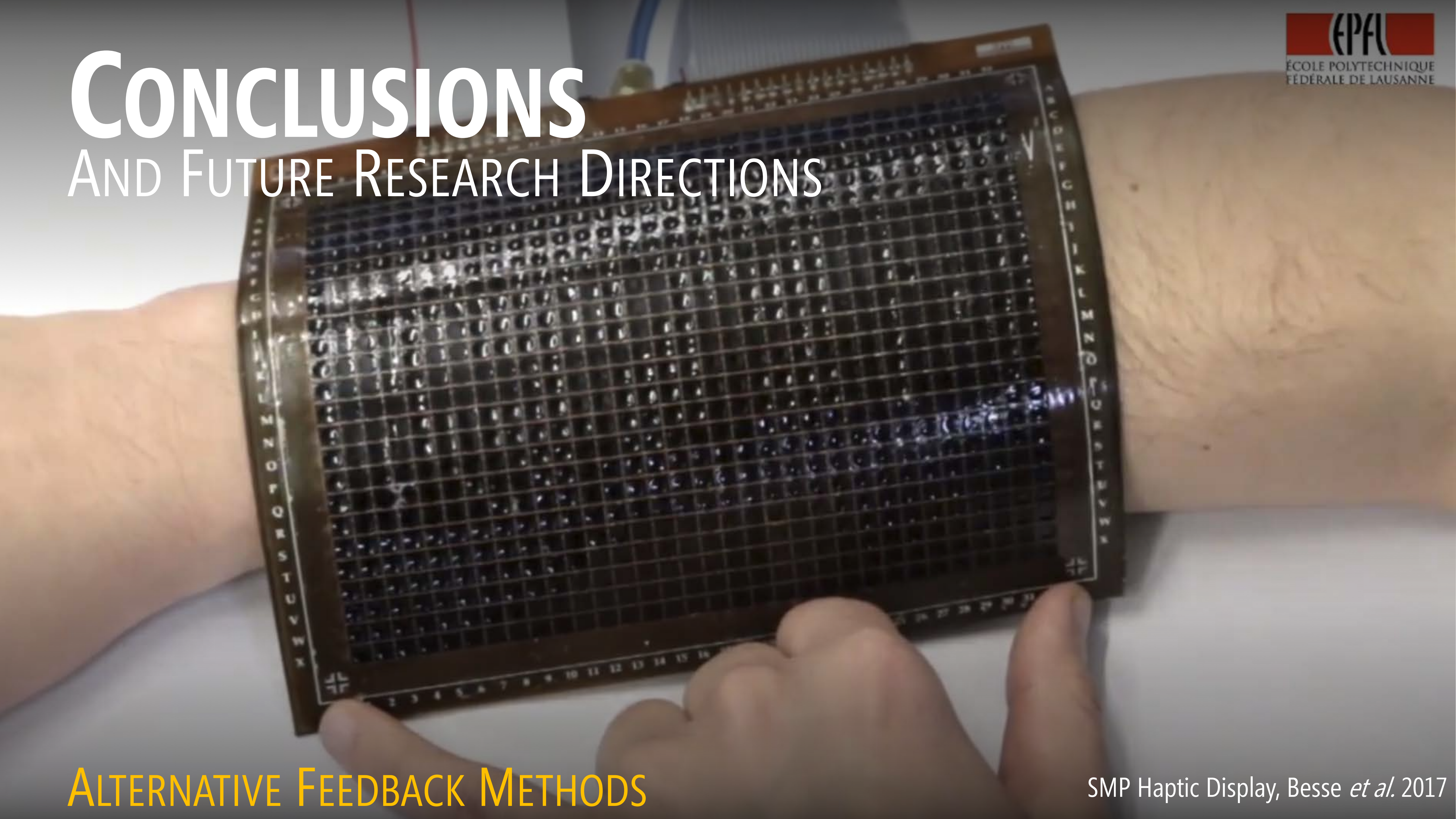
Conclusions
And Future Research Directions



Flexible Tactile Display, Choi *et al.* 2004

Alternative Feedback Methods

Conclusions
And Future Research Directions



ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Alternative Feedback Methods

SMP Haptic Display, Besse *et al.* 2017

Conclusions

And Future Research Directions

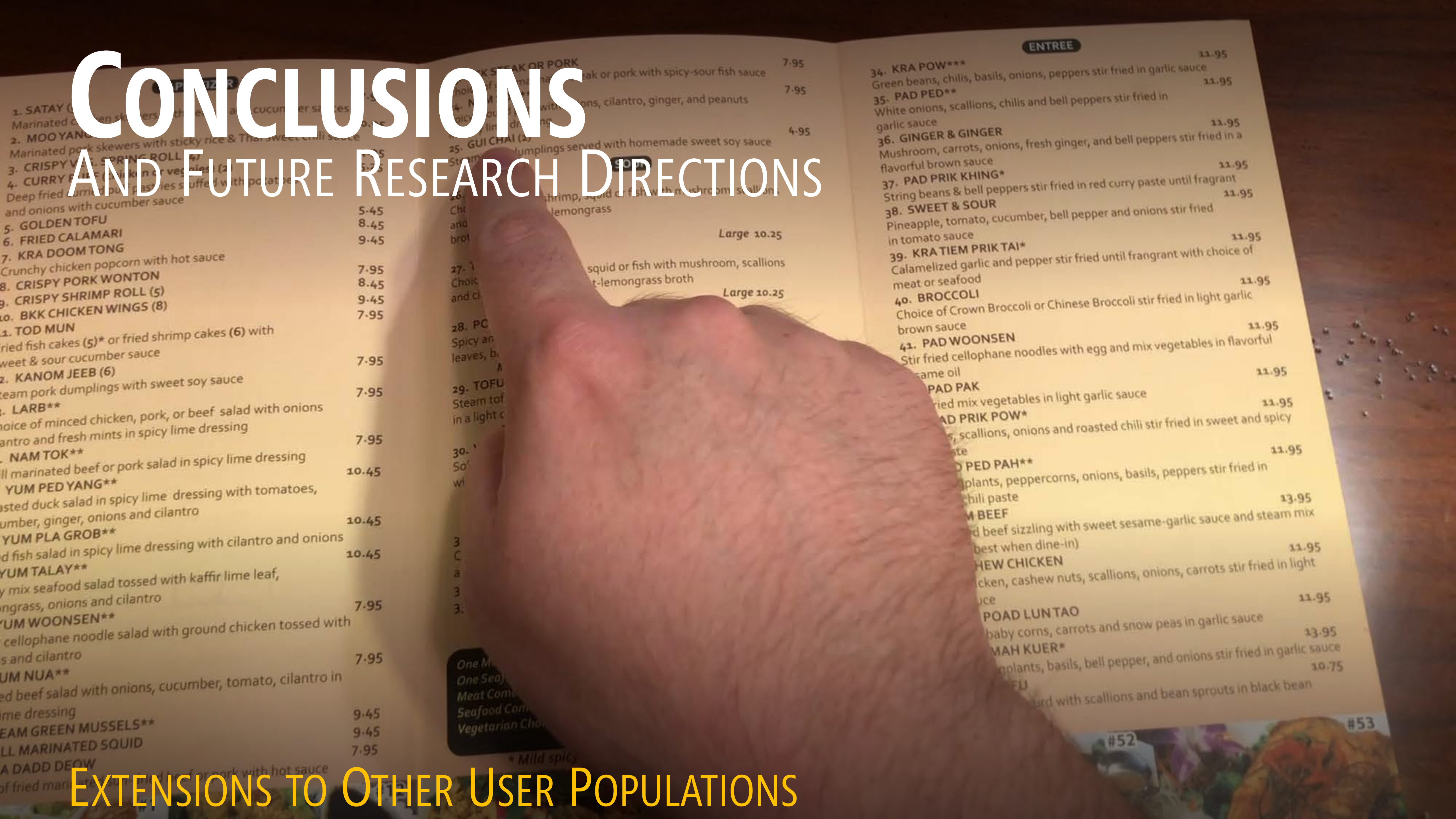
Extensions to Other User Populations

CONCLUSIONS
AND FUTURE RESEARCH DIRECTIONS

EXTENSIONS TO OTHER USER POPULATIONS

Conclusions

And Future Research Directions



Extensions to Other User Populations

Conclusions

And Future Research Directions

Extensions to Other User Populations

Acknowledgements

Professors
- Jon Froehlich, UMD / UW
- Rama Chellappa, UMD
- Leah Findlater, UMD / UW

Collaborator
- David Ross, Atlanta VA

Funding Source
- Office of the Assistant Secretary of Defense for Health Affairs

Graduate Students
- Uran Oh, UMD / CMU
- Jonggi Hong, UMD
- Ruofei Du, UMD
- Liang He, UMD / UW
- Alisha Pradhan, UMD
- Darren Smith, UMD

Undergraduate / High School Students
- Victor DeSouza, UMD
- Meena Sengottuvelu, UMD
- Jessica Yin, Poolesville HS
- Eric Lancaster, UMD
- Bridget Cheng, Cornell
- Ji Hyuk Bae, Poolesville HS
- Alexander Medeiros, UMD
- Chuan Chen, Poolesville HS
- Harry Vancao, UMD
- Catherine Jou, UMD
- Tony Cheng, UMD
- Yumeng Wang, UMD

MAKEABILITY LAB · inclusive design lab · DUB · HCIL Human Computer Interaction Laboratory · COMPUTER SCIENCE UNIVERSITY OF MARYLAND · UNIVERSITY of WASHINGTON · UNIVERSITY OF MARYLAND