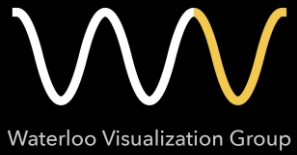
Exploring Unimanual Around Ear Off-Device Gestures for Earables

Shaikh Shawon Arefin Shimon,

Ali Neshati, Junwei Sun,

Qiang Xu, Jian Zhao



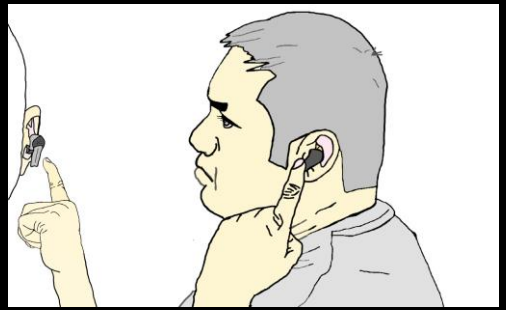
Earables: Ear-mounted Wearable Devices

Usually paired with external device

- (e.g., Laptop, Mobile, External Display)

Most popular: In-ear Earables

- Small Form factor
- Subtility in Public space



In-Ear Earable (Earbud)
(e.g., AirPod, Galaxy Buds)

Over-Ear Earable (Headphone) [2]

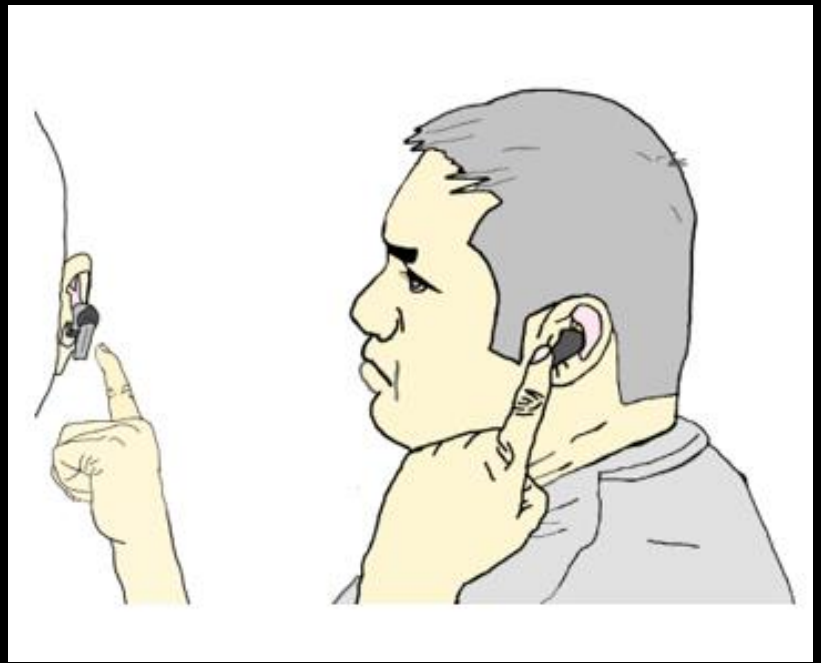
Behind-Ear Earable
(Hearing Aid) [1]

Lobule-Mounted Earable
(Smart Ear-Stud Jewelries)

In-Ear Earbles: Limited Physical Input Space

- Space for Tap & Vertical Swipe only
  - Restricts providing complex input to paired devices

- Solution? – Off-device Earable Inputs
  - Head Motion
  - Full Body Motion
  - Facial Expression
  - Silent Commands
  - **Manual gestures**



Preferred Off-device Earable Gestures

- Unimanual (One-handed) hand-to-face gestures most preferred.
  - Spontaneous & Unconscious
  - Upto 23 hand-to-face contacts/Hr.
- Subtle unimanual gestures more socially acceptable than large uni/bi-manual gestures.

Viable Input Space For Manual Gestures?

- Above neck for both midair and onskin spaces
- Midair space:
  - above and around outer ear and head
- Onskin space:
  - around outer ear
  - Touching outer ear can dislocate device

Midair space: Above neck and around face and ears

Onskin space: Physical area on neck, face and head except outer ear.

Preferred Off-device Unimanual Gesture Shapes?

Tap

Swipe
Up / Down / Front / Back

Pinch
In/Out

- Suggested by prior earable elicitation studies
- Considered subtle by end-users

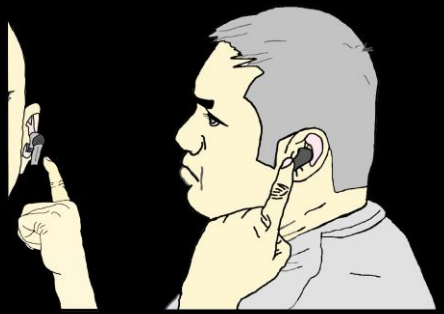
Using Gesture Regions to Expand Inputs

Segment interaction space to create multiple gesture regions

Expanding input set by Gesture Shape X Location

Small Number of Shapes = Higher gesture Memorability

Location-based gesture reusability unexplored for viable space



Physical Interaction with In-ear Earable

Tap on Nose | Tap on Cheek | Tap behind Ear

Off-Device Earable Input Using Location-based Onskin Taps

Research Questions

- **RQ1:** Whether and how does gesture performance vary between in mid-air and on-skin space?

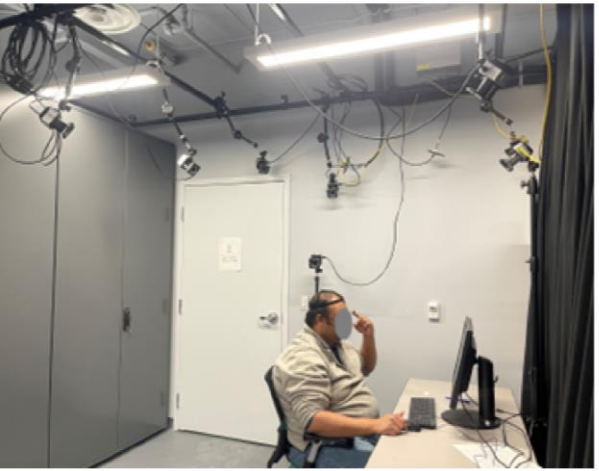
- **RQ2:** Whether and how does gesture performance vary when the number of segments increases in a chosen interaction space?

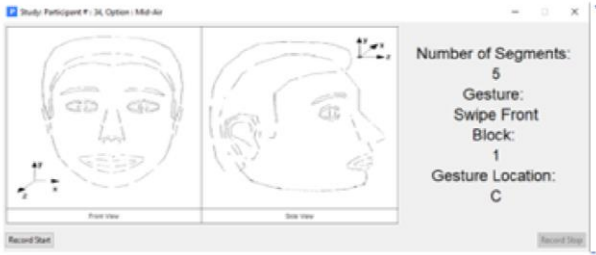
- **RQ3:** Whether and how does end-user consensus exist on most and least preferred regions across a fixed number of gesture regions in a particular interaction space?

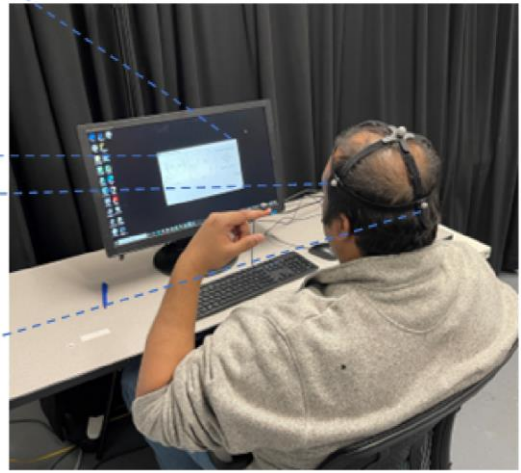
Independent and Dependent Variables

- **Independent Variables:**
  - **IV1** - Interaction spaces
    - Midair
    - Onskin
  - **IV2** - Gesture Shapes:
    - Swipe Up
    - Swipe Down
    - Swipe Front
    - Swipe Back
    - Pinch In
    - Pinch Out
    - Tap
  - **IV3** –Gesture Region Density:
    - 3 regions
    - 5 Regions
    - 7 Regions

- **Dependent Variables:**
  - **DV1** – Gesture Time (ms)
  - **DV2** – Gesture Path Length (mm)
  - **DV3** – Gesture Accuracy (%)
- **Other Assessment Metrics:**
  - NasaTLX Workload Assessment
  - Relative Gesture Ranking
  - Questionnaire and Discussion







Apparatus

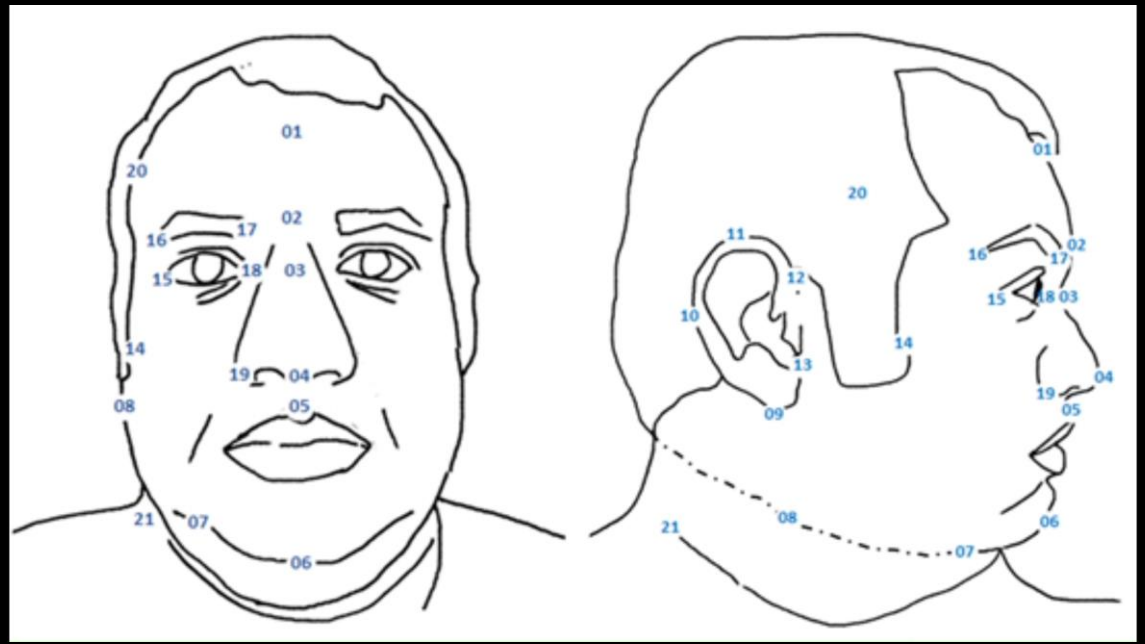
- 8 Vicon Cameras, Vicon Marker on Finger and Head
- Desktop application for showing gesture task and recording Vicon gesture motion data, Vicon tracker 3.9 running on background
- All motion start from white tape on the table, followed by delimiter gesture
- Metric measurement starts when fingers reaches 30cm from head
- Non-dominant hand for gesture and dominant hand to control apparatus (using keyboard and mouse)

Study Design

- Within-subject Experiment - 11 Male, 7 Female (Total 18) participants (M = 26.37, SD = 5.520)
- 2 counterbalanced half studies (corresponding to each interaction space): 75 ~ 90 minutes each.
- 42 Experimental Conditions:
  2 interaction spaces × 3 region density patterns × 7 Gestures
  - Counterbalanced region density patterns within each half study
  - Regions labeled A-G and boundary marked on a page with front and side view of head
- 2 Trials / Region, 420 Samples / Participants, 7560 Gesture motions analyzed

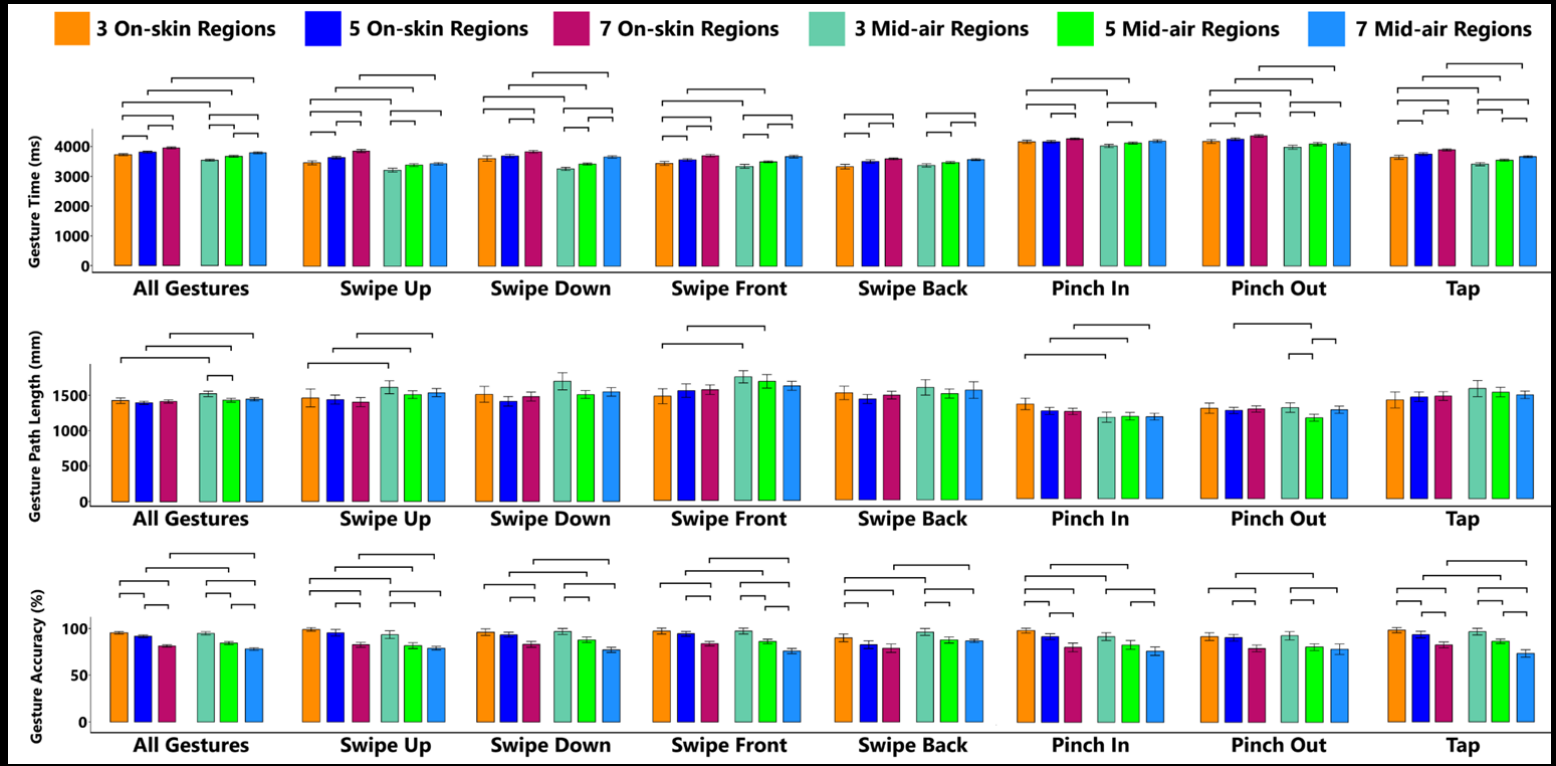
Data Collection and Analysis

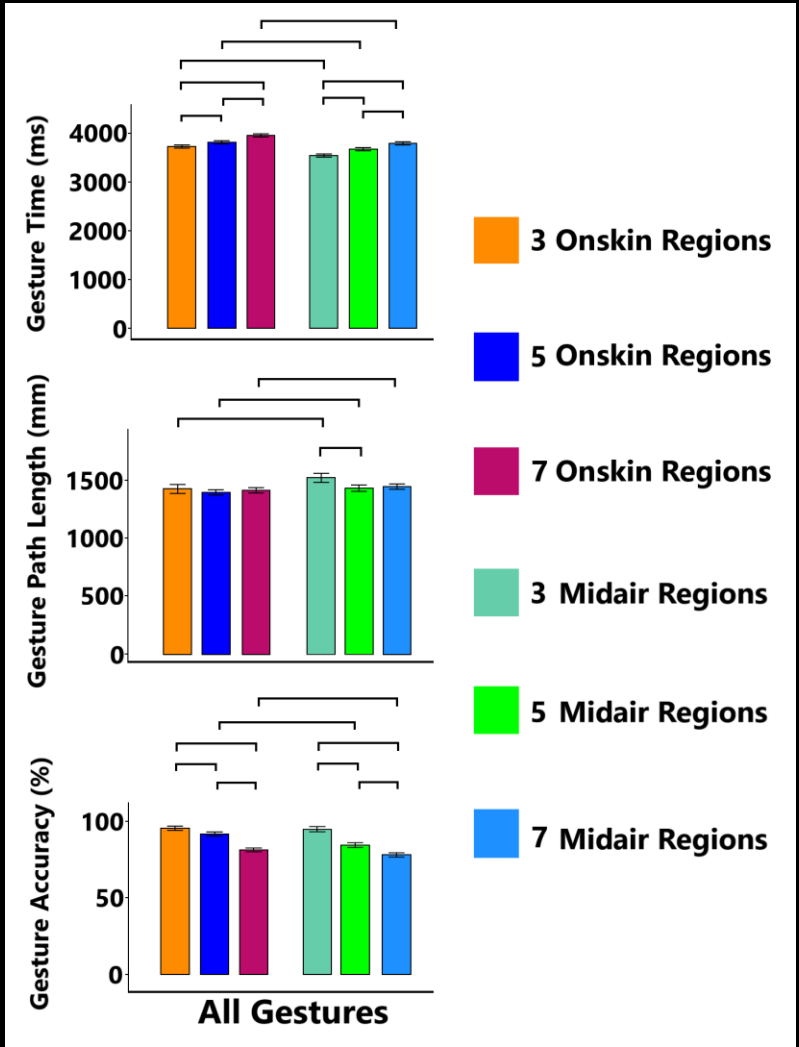
- Built adjustable 3D model of face for individual participants by recording face data points
- Transferred gesture region definition from paper to custom app after data collection ended
- Observed gesture trajectory against target region in individual head model during analysis phase for DV3 (Accuracy) measurement



Walkthrough of Data Collection for Pinch in gesture shape
Red moving dot = index finger, blue dot = thumb

Effect of Interaction Space Choice (RQ1) & Increasing Region Density (RQ2) in Midair Vs. Onskin

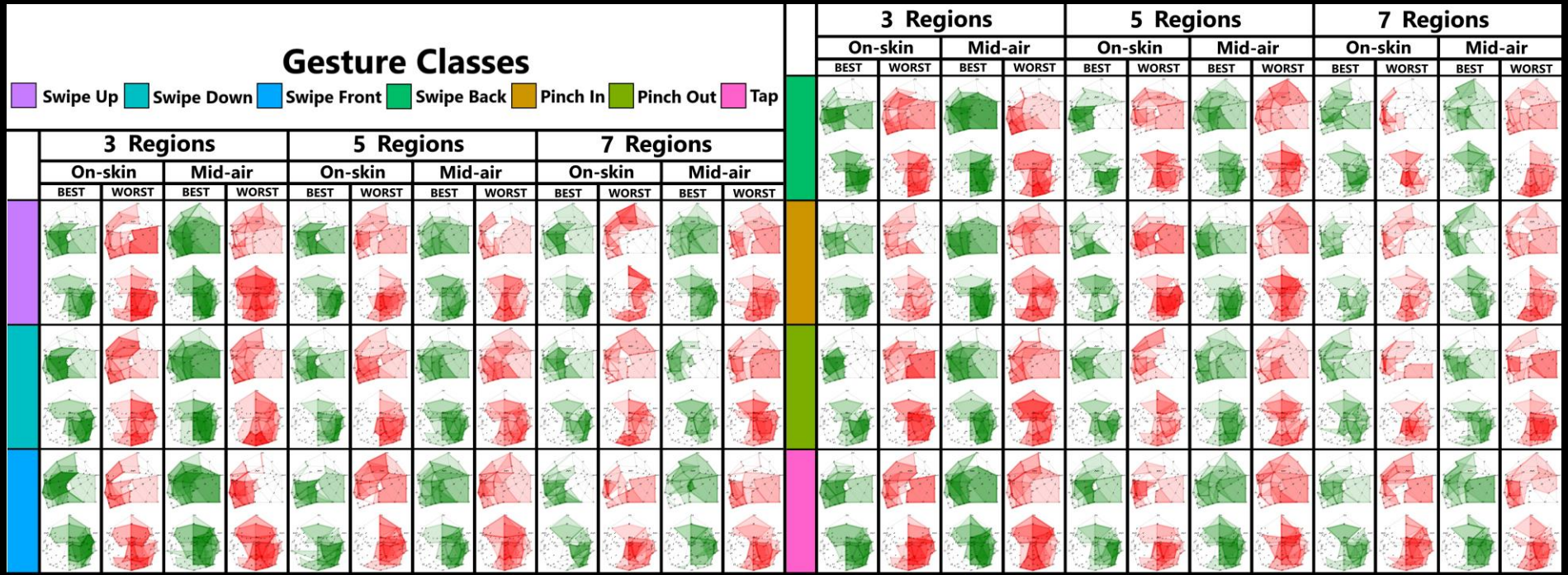




RQ1 &RQ2 Findings

- For 3 regions, mid-air and on-skin gestures have similar performance
  - midair gestures – longer pathlength but faster, and similar accuracy, preferred by end-users
- Regions >= 5, preference shifts towards onskin gesture because of tactile feedback.
- Gesture region threshold:
  - 3 Midair regions
  - 5 Onskin regions
- Individual gesture detail on RQ1 & RQ2 can be found on our paper

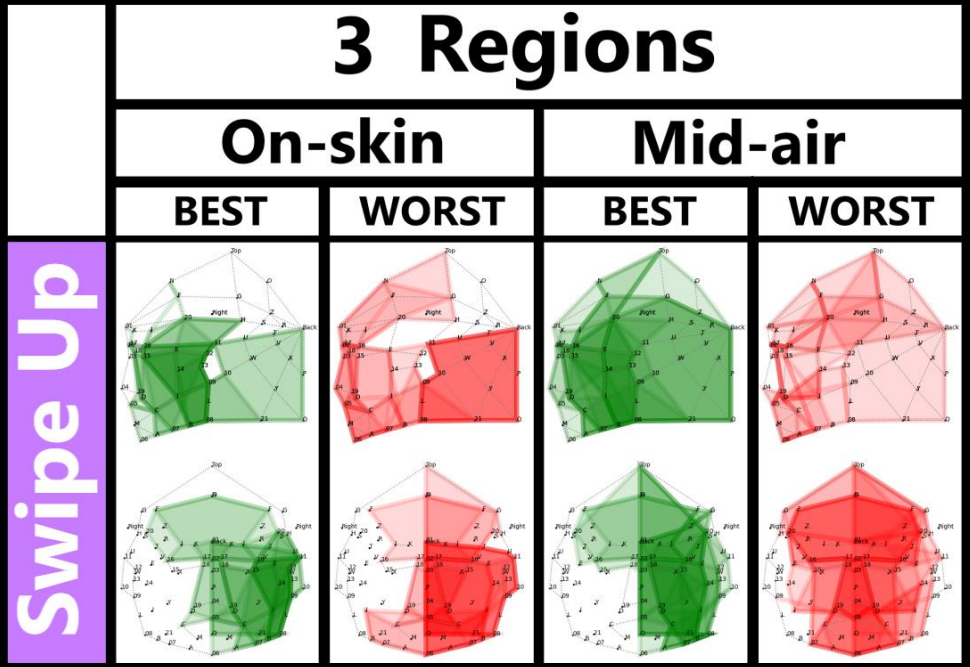
RQ3 : Overlap in Region Boundary Selection for Most and Least Preferred Regions

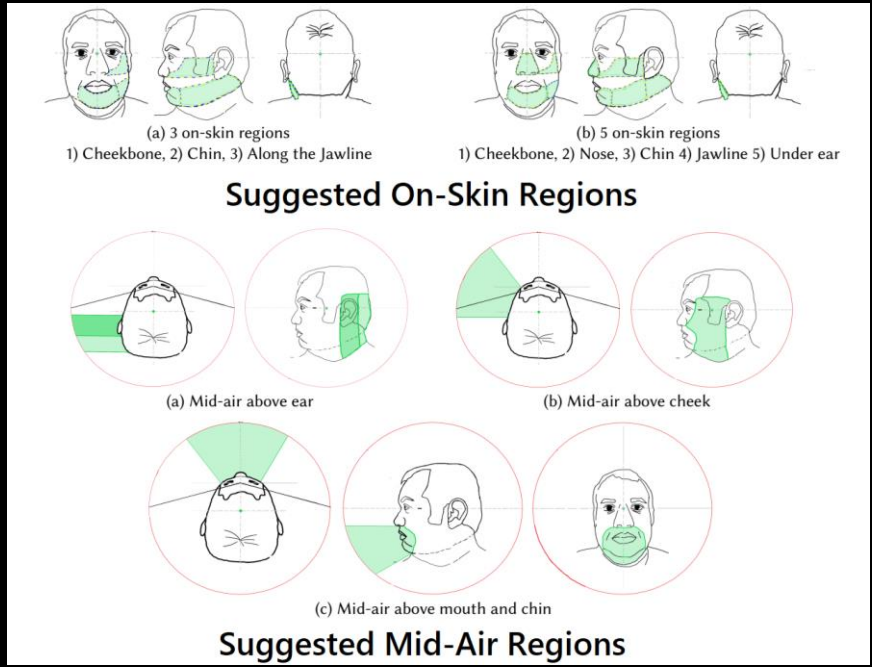


Green = Most preferred Region

Red = Least Preferred Region

RQ3: Recommended Region Boundaries

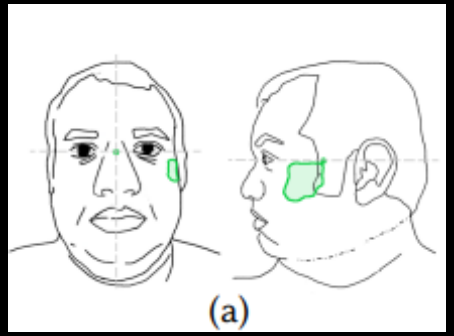




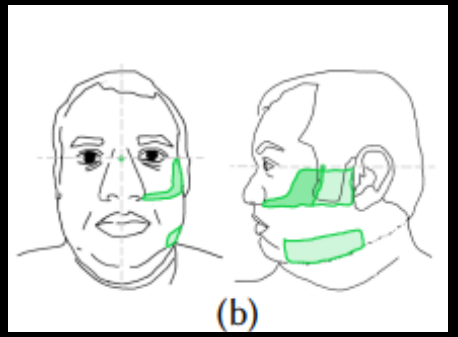
- Onskin –most frequent tasks should be mapped to cheekbone
- Mid-air –most frequent tasks should be mapped region from temple to jaw through cheek area (b)

Qualitative Observations: Onskin Regions

All Region Density

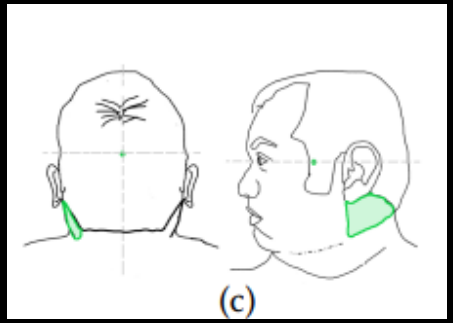


Preferred : taps & vertical swipes

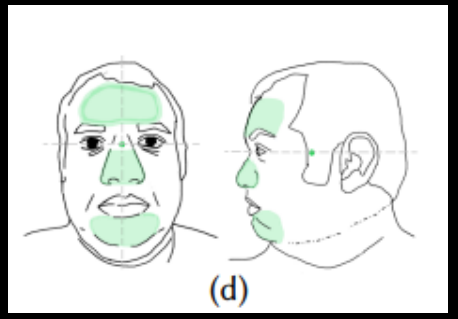


Preferred : horizontal pinch & swipe;
Region Preference:
Cheekbone > Jawline

>=5 Region Density

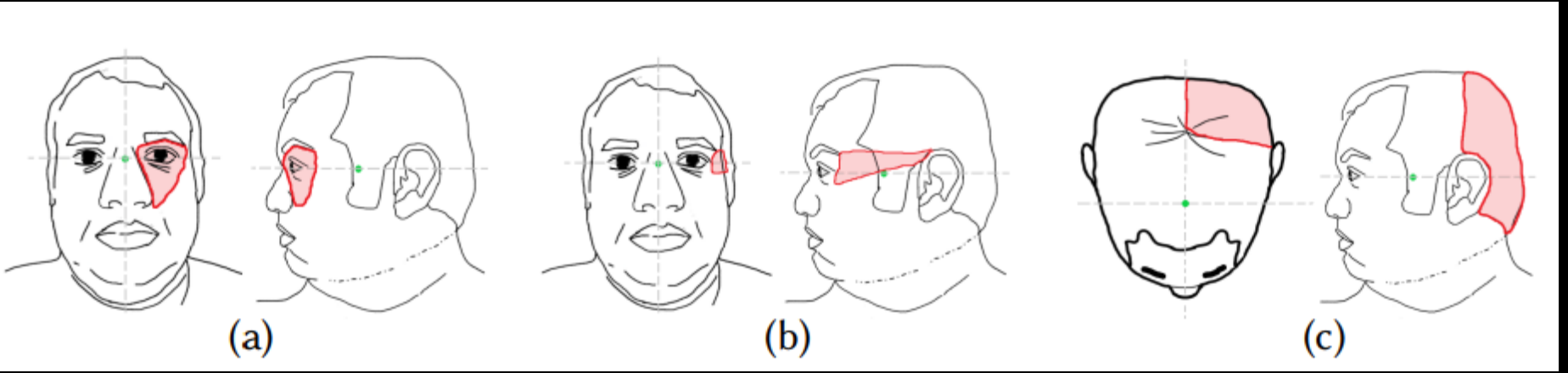


Preferred : taps & vertical pinch,
uncomfortable bending for other gestures



Region Preference:
Chin > Nose > Forehead

Qualitative Observations: Onskin Regions



- On-skin gestures around **A**, **B** and **C** should be avoided
- **B** can act as a buffer between region below and above the temple.
- Restrictions not applicable for mid-air gestures

Limitations

- Gesture region distribution and preference needs to be explored in the wild
  - Real world scenarios like Standing, Walking
- Gesture + Region and Task mapping needs to be explored for various scenarios
  - Controlling smart home environments
  - Public Display Interaction
  - AR / VR Interaction

Exploring Unimanual Around Ear Off-Device Gestures for Earables

Shaikh Shawon Arefin Shimon, University of Waterloo, Ontario, Canada
ssarefin@uwaterloo.ca | www.shawonarefin.com

Ali Neshati, Ontario Tech University, Ontario, Canada
Junwei Sun, Huawei Human-Machine Interaction Lab, Ontario, Canada
Qiang Xu, Huawei Human-Machine Interaction Lab, Ontario, Canada
Jian Zhao, University of Waterloo, Ontario, Canada

Waterloo Visualization Group

WATERLOO HCI
CHERITON SCHOOL OF COMPUTER SCIENCE

UNIVERSITY OF WATERLOO

References / Image Courtesy

[1]: MAF: Exploring Mobile Acoustic Field for Hand-to-Face Gesture Interactions, CHI'24
https://tachen-cs.github.io/data/papers/2024-chi-maf/chi24-maf.pdf
https://doi.org/10.1145/3613904.3642437

[2]: Beyond Audio: Towards a Design Space of Headphones as a Site for Interaction and Sensing, DIS'23
https://dl.acm.org/doi/abs/10.1145/3563657.3596022

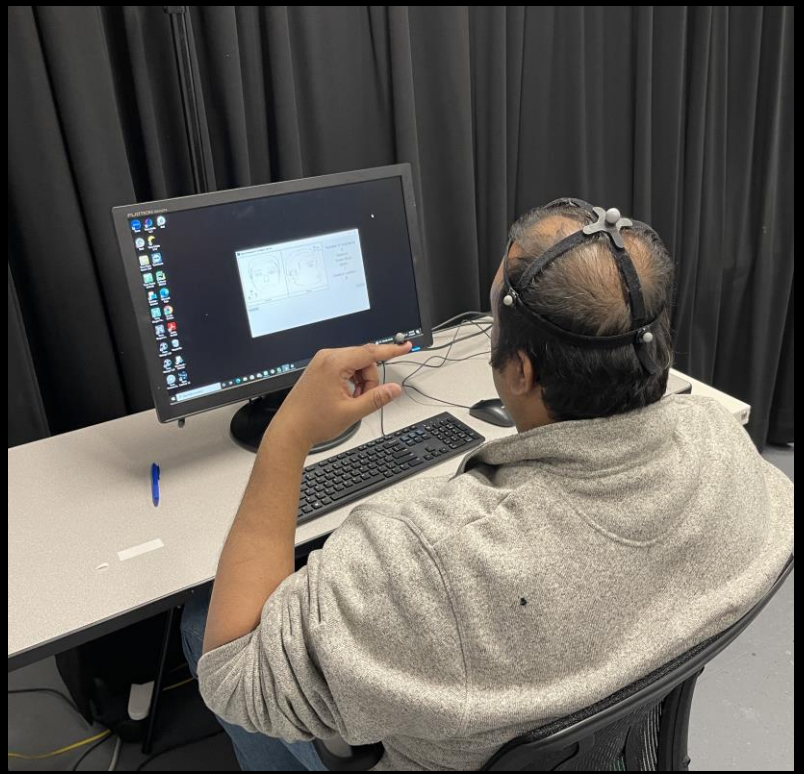
Additional Slides

Expansion of Our Work

- Our work published in IMWUT/Ubicomp 2024.
- Sato and Ameshka et. Al. from Keio University, Japan expanded our work by developing gesture recognition system for in-ear and ear-hook devices using IMU.
  - Their work published in MobileHCI2024

Apparatus

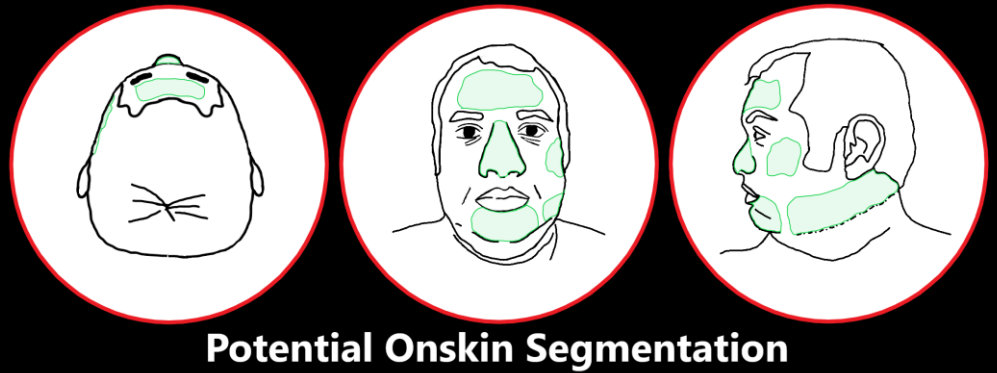
- All gesture motions started from designated position (white tape) on the table.
- Delimiter gesture after motion started from white tape
- Metric measurement started when fingers reached 30cm from head
- Gesture ended when participants put hand back on table.



Research Gap: Region-Based Off-Device, Around-Ear Unimanual Inputs

- Segmenting viable input space for reusing manual gestures has not been covered by prior research

- Benefits of reusing Tap/Pinch/Swipe gestures:
  - Expanding input set by Gesture Shape X Location
  - Small Number of Shapes = Higher gesture Memorability



Potential Onskin Segmentation

Potential Midair Segmentation