FlowRing: Integrated Microgesture and Surface Interaction Ring for Versatile XR Input

ISHAN CHATTERJEE*, University of Washington, USA
JIEXIN DING*, University of Washington, USA
ANANDGHAN WAGHMARE*, University of Washington, USA
JOSEPH BREDA, University of Washington, USA
YUQUAN DENG, University of Washington, USA
BO LIU, University of Washington, USA
YUNTAO WANG, Tsinghua University, China
SHWETAK PATEL, University of Washington, USA

Fig. 1. FlowRing enables both precise on-surface interactions on a desktop, pant, and palm surface as well as quick, subtle free-hand microgestures on the fingertip, facilitating constantly available, ad hoc interactions. This type of peripheral can help to control cross-context mobile devices, such as mixed reality glasses where interfaces range from precise desktop experiences when stationary to simple, discrete on-the-go widgets.

*The authors contribute equally to this paper.

Authors' Contact Information: Ishan Chatterjee, Paul G. Allen School of Computer Science & Engineering, University of Washington, Seattle, Washington, USA, ichat@cs.washington.edu; Jiexin Ding, Global Innovation Exchange, University of Washington, Seattle, Washington, USA, jxding@uw.edu; Anandghan Waghmare, Paul G. Allen School for Computer Science & Engineering, University of Washington, Seattle, Washington, USA, anandw@cs.washington.edu; Joseph Breda, Paul G. Allen School for Computer Science & Engineering, University of Washington, Seattle, Washington, USA, joebreda@cs.washington.edu; Yuquan Deng, Paul G. Allen School for Computer Science & Engineering, University of Washington, Seattle, Washington, USA, yuquand@cs.washington.edu; Bo Liu, Global Innovation Exchange, University of Washington, Seattle, Washington, USA, boliu97@uw.edu; Yuntao Wang, Department of Computer Science and Technology, Tsinghua University, Beijing, China, yuntaowang@tsinghua.edu.cn; Shwetak Patel, Paul G. Allen School for Computer Science & Engineering, University of Washington, Seattle, Washington, USA, shwetak@cs.washington.edu.

This work is licensed under a Creative Commons Attribution 4.0 International License.

As Extended Reality (XR) advances, a device has the potential to be used across contexts from immersive productivity at a desk to on-the-go, public scenarios. Existing input solutions lack the versatility to provide both high-throughput, mouse-grade input and subtle, ergonomic interaction. We introduce FlowRing, a novel ring-form device that combines microgestures with precise 2D mouse-like input on surfaces. FlowRing supports five microgestures for discreet interaction and 2D input for richer tasks, using an optical flow sensor, skin-contact microphone, and IMU at the base of the finger. In a study with 11 participants, FlowRing achieved 93.6% microgesture recognition accuracy across sessions and 85.2% across unseen users, rising to 90.1% with just four gesture set examples from a new user. A separate 2D Fitts' law study demonstrated its effectiveness for continuous input on various surfaces. FlowRing emerges as a versatile, user-friendly solution for the future of interactive technology.

CCS Concepts: • **Human-centered computing** → **Interaction devices**; **Mixed / augmented reality**; **Ubiquitous and mobile devices**.

Additional Key Words and Phrases: input device, XR, spatial computing, text tagging

ACM Reference Format:
Ishan Chatterjee, Jiexin Ding, Anandghan Waghmare, Joseph Breda, Yuquan Deng, Bo Liu, Yuntao Wang, and Shwetak Patel. 2025. FlowRing: Integrated Microgesture and Surface Interaction Ring for Versatile XR Input. *Proc. ACM Hum.-Comput. Interact.* 9, 5, Article MHCI010 (September 2025), 28 pages. https://doi.org/10.1145/3743706

1 Introduction

Researchers have long considered the potential for Extended Reality (XR) to one day become daily, pervasive wearables, providing users with a seamless computing fabric that adapts to their varying contexts and computing needs [23, 30, 47, 53]. For example, when seated at a desk, an XR device can deliver immersive, desktop-grade productivity applications. While on the couch, the device might adopt a television-like interface for watching media. Then, while on-the-go, it could transition to an in-context heads-up display. While former devices are often pigeonholed to one or the other end of the spectrum (for instance, Meta Quest Pro [35] for immersive use cases and North Focals' [5] for on-the-go scenarios), we are starting to see examples of increasingly versatile devices promising to deliver immersive displays in glasses-like form factors (e.g. XReal Air [58], Rokid Joy [41]) with many lighter weight goggles devices and display-capable glasses devices anticipated to come to the market in the next few years [2, 24, 34]. For instance, Meta's recent Orion Prototype demonstrates a variety of stationary and mobile use cases via its world-locked, seventy-degree displays within a glasses form factor [18]. XReal's Air 2 Pro feature electrochromic dimming in a glasses form factor allowing the user to move between VR and AR environments [59]. While massive strides have been made on lighter weight head-mounted display technology, to realize this vision of ubiquitous, all-day XR, input devices must be also adaptable enough to support this broad range of scenarios: from high throughput, precise control when the user is stationary, to on-the-go, subtle interaction, all while being comfortable and socially appropriate for extended use and wear (whether the user is actively using the display or not).

Current input devices, however, lack this necessary versatility. In stationary, productivity environments, headsets such as the Apple Vision Pro [21] and Meta Quest Pro [35] have integrated traditional peripherals like mice and trackpads, which are the gold standard for their ergonomic design and high throughput. Despite their effectiveness in stationary contexts, these input devices are impractical for mobile scenarios due to their lack of portability.

© 2025 Copyright held by the owner/author(s).
ACM 2573-0142/2025/9-ARTMHCI010
https://doi.org/10.1145/3743706

Conversely, mobile XR tasks (i.e. notifications, music control) have seen the adoption of more compact input solutions, exemplified by devices like North Focals' joystick ring [5] and Snapchat Spectacles' 1D temple touchpad [45]. These devices provide subtle and simple input mechanisms that are suitable for on-the-go use but fall short of offering the richer, more complex interactions necessary when users are stationary or engaged in tasks requiring higher throughput. Other common input methods, such as camera-based tracking for free-space hand gestures or controller input, introduce additional challenges, including reduced ergonomic comfort during extended use [12, 17, 39], social acceptability concerns in public settings, and increased power consumption and bulk for headsets.

In response to these limitations, we introduce FlowRing, a novel ring-form factor device designed to bridge the gap between the high-throughput, ergonomic input of stationary, surface-bound interactions and the subtle, portable control needed for all-day, mobile XR. To achieve this versatility and portability for ubiquitous XR, FlowRing must simultaneously meet multiple design criteria. First, FlowRing supports a dual-mode interaction paradigm: on-the-go input is facilitated by five subtle finger micro-gestures (four cardinal directional swipes and a tap), while richer, higher-throughput interactions are enabled through 2D continuous input and tapping on surfaces, akin to a familiar mouse interaction. Both of these interactions are motivated by literature on user preference, comfort, and social acceptability [1, 7, 16, 20, 33, 38, 43, 44, 56]. Second, as a wearable, FlowRing can operate on multiple always-available surfaces, such as the user's own desk, thigh or palm. FlowRing can automatically recognize these surfaces for driving contextual interactions, as well as seamlessly transition between surfaces and free-space gestures in real-time, offering a more portable and adaptable interaction experience than traditional mouse devices. Finally, FlowRing's design prioritizes ergonomics and social acceptability. Unlike other microgesture and surface interaction devices that are positioned at the fingertip or require multiple rings [28, 32, 37, 42, 43, 49, 62], FlowRing is worn at the base of the index finger, resembling conventional jewelry. This placement not only ensures greater comfort during extended use but also makes FlowRing more discreet and socially appropriate for a variety of settings, enhancing its utility as a ubiquitous input device for XR [16, 20, 44].

FlowRing's interaction capabilities are powered by a combination of sensing technologies: a miniature optical flow sensor, a skin-contact microphone, and an inertial measurement unit (IMU). These modalities complement each other, with the optical flow sensor and IMU capturing finger motion and the contact microphone detecting finger contact. This synergy allows FlowRing to provide a robust and responsive input experience that is well-suited for a wide range of XR scenarios.

In this paper, we seek to provide a versatile, discreet, and ergonomic input device towards ubiquitous XR input. Towards this goal, we offer several contributions. First, we propose FlowRing, and integrated and practically-designed ring device and accompanying processing algorithms that provide a versatile multi-function controller for mobile XR input. FlowRing enables discrete microgestures, continuous 2D input, surface recognition, and seamless clutching between these functions for on-the-go input. Second, we evaluate FlowRing's functions. We conduct an 11-person user study demonstrating the effectiveness of its microgesture input capabilities across users with different hand sizes and skin tones, as well as a 10-participant Fitts' law evaluation of continuous 2D surface interaction. We discuss the merits and limitations of FlowRing's performance across surfaces and lighting. Finally, we contextualize our interaction techniques within a set of example scenarios, demonstrating the potential of mobile, seamless, cross-surface interaction.

2 Interactions

As users change environments and activities, the surfaces available to them may change. In this paper, we target four commonly available surfaces for FlowRing use: a tabletop, pant leg, palm,

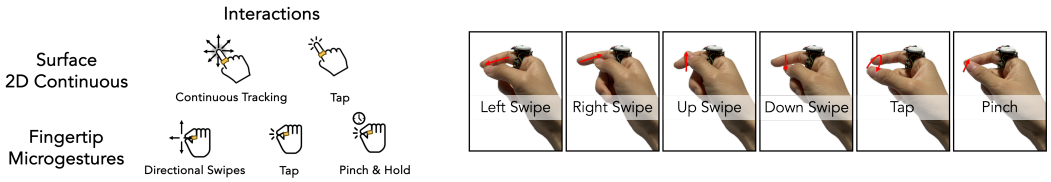

Fig. 2. The left depicts a summary of the interactions afforded by FlowRing. The right shows the discrete microgestures in greater detail. Please see supplementary Video Figure for further visual description of interactions.

and fingertip. These surfaces span different affordances along multiple dimensions, including their size, orientation, and number of hands required for interaction (see Fig. 1). A large surface like a desktop would be available while seated and permit expressive, mouse-like interactions. Similarly, the thigh area of a pant leg may be less precise but offers a similarly large area for continuous cursor control. A medium surface like the hand palm affords touch gestures, with one hand performing 2D continuous interactions on the other. Finally, a small surface such as a fingertip can enable subtle one-handed interactions, like discrete microgestures, which can be performed in standing or on-the-go scenarios.

Because future spatial computing devices will be used across scenarios and environments, we similarly target FlowRing to *seamlessly* support input modalities across these dimensions. FlowRing is designed to recognize these surfaces, allowing modality switches to occur without friction. Broadly, FlowRing supports two classes of interactions: 2D continuous surface input and microgestures, which we describe below.

2.1 2D Continuous Surface Input for Expressive Interaction

For scenarios featuring a richer user interface and larger surface, our system offers rich input through a trackpad-like 2D continuous input mechanism. Users can conveniently employ a desk while seated, use their clothing for spontaneous input while standing or lounging, or choose their palm for an ad hoc additional control surface.

Importantly, despite being mounted on the finger base, our system can enable mouse-like interaction since it supports, not only index finger dragging, but also a tap-to-select gesture.This poses a technical challenge to maintain cursor stability during finger lift and attack (see Sec. 5.1), and previous 2D interaction systems [25] do not support this capability. By enabling complete 2D continuous interactions across these surfaces, the system can also implicitly support other simpler interactions on surfaces, such as touchscreen-like swipe interactions and discrete gestures.

2.2 Microgestures for Rapid Interaction

The smallest surface our system supports is the fingertip. Given the size constraints of this interaction space, we target thumb-to-index finger microgestures, well suited for on-the-go and public use because they are subtle, ergonomic, and quick [1, 6, 7]. The system facilitates five distinct gestures: four cardinal directional swipes (left swipe: ${}_{t}\overset{\bullet}{\blacktriangleleft}({}_{i}\bullet_{m})$, right swipe: ${}_{t}\overset{\bullet}{\blacktriangleright}({}_{i}\bullet_{t})$, down swipe: ${}_{t}\overset{\bullet}{\blacktriangledown}({}_{i}\bullet_{t})$, up swipe: ${}_{t}\overset{\bullet}{\blacktriangle}({}_{i}\bullet_{t})$), enabling efficient navigation through hierarchical menus, and a tap gesture $({}_{t}\overset{\circ}{\blacktriangledown}({}_{i}\bullet_{t});{}_{t}\overset{\bullet}{\blacktriangle})$ for making selections [6]. These interactions can efficiently navigate the type of UIs found on smartglasses and other mobile UIs. The gestures are executed using the thumb on the near-tip region of the index finger; these two fingers' natural proximity and agility contribute to

the effortless execution of these gestures, enabling users to perform actions with one hand (see Figure 2).

Additionally, the system incorporates a stateful *pinch* ($_{t}\overset{\circ\bullet}{\blacktriangledown}(_{i}\bullet_{t})$), defined as the continuous detection of contact between the thumb and index finger [6]. As opposed to tap, pinch carries state, achieved by monitoring the touch-down and touch-up events of the thumb on the index finger independently [26]. The stateful pinch adds a temporal aspect to the tap interaction, facilitating more nuanced inputs such as long taps [3].

3 Related Work

Towards enabling always-available touch control for XR, FlowRing's goal is to enable *both* free-hand microgestures as well as mouse-like cursor control within the same set of sensors, which to our knowledge has not been previously demonstrated (see Table 1). Here we survey the set of devices that achieve each interaction separately — starting with focusing on continuous on-surface input, and then moving to microgesture interactions.

3.1 Continuous, On-Surface Input

Precise, continuous touch interactions on surfaces are a cornerstone of modern HCI interaction design. Pioneered by systems like [10], ubiquitous touchscreens and trackpads now enable precise 2D input and stateful touch segmentation across personal and mobile devices. While users have envisioned scaling these sensing modalities to enable world-scale touch [66], many common objects and surfaces are irregular or non-conductive. This has led researchers to consider other methods, such as conductive fibers and coatings with impedance measurement [40, 54, 65], surface acoustic measurements [11, 22], or optical methods [19]. However, with each of these methods, instrumenting a surface still limits touch interaction to that object itself.

Instrumenting the user offers both mobility and the potential for cross-surface input using the same sensing system. For XR interactions, this has most commonly been achieved by leveraging hand-tracking cameras on the headset [21, 35]. However, precise surface interaction via HMD cameras carries several drawbacks [57]: due to self-occlusion of the hand, surface touch segmentation can be difficult to achieve, vision-based systems are subject to lighting conditions, interaction is limited to the camera's field of view, and adding cameras to the headset increases bulk. Therefore, hand-wearable solutions have garnered intense interest.

These worn solutions can be categorized according to their surface sensing capabilities. The simplest devices perform touch segmentation, determining whether the finger is contacting a surface. For example, ElectroRing uses impedance sensing to determine surface touch state from an index ring but is limited to on-body or conductive grounded surfaces [26]. VibAware uses active bio-acoustics on a thumb ring and wristband to similarly determine contact and combines it with an IMU to achieve swiping touch gestures [28].

However, to enable high-fidelity applications, precise and continuous 2D tracking is required. Z-Ring uses bio-impedance sensing to perform 2D tracking but is limited to pre-calibrated, asymmetric conductive surfaces [51]. LightRing presents an index-mounted ring that can precisely move a cursor using a gyroscope and IR reflectance sensor [25]. However, this approach has a few limitations. As an inertial method, the palm must remain in place and cannot move freely across a surface, and calibration is required with each use to determine the inertial-to-display mapping. Additionally, a key function to control applications is missing; the user may use the ring for 2D targeting but cannot tap for selection.

With FlowRing, we target mouse-like capability, including both targeting and tapping. 3DTouch [37] and Magic Finger [62] explore this paradigm; however, both mount an optical sensor on

the fingerpad, limiting practicality and ergonomics. The recent MouseRing [42] is most similar in capability to our work. MouseRing uses one to two IMU rings with a kinematically-informed RNN to enable both 2D tracking and tapping, but like LightRing, it requires palm-anchored use. Additionally, at least one of the rings must be worn on the proximal phalanx closest to the fingertip, which departs from a ubiquitous form factor.

Through sensor fusion, FlowRing achieves mouse-like capabilities of both 2D continuous tracking and surface tapping from a location at the base of the finger, along with limited surface recognition capabilities. Additionally, to enable ubiquitous XR control even without a tabletop, FlowRing uniquely extends its interaction capability to also enable microgesture sensing.

3.2 Thumb-to-Index Microgestures

Finger-based microgestures can allow for ergonomic and subtle interaction with mobile ubiquitous computing systems [7, 33, 38, 56]. This has motivated a large body of work focused on thumb-to-index microgestures recognition to control devices, such as smartglasses, on- the-go (see Shilkrot et al. [44] and Vatavu and Bilius [50] for surveys). To allow for increased mobility compared to stationary methods [15, 29], wearable systems have been developed for single-hand thumb-to-index typing, such as TipText and DRG-Keyboard [31, 61], as well as for 1D and 2D swipe interactions, such as ThumbRing [49], DualRing [32], VibAware [28], and 'Ready, Set, Touch!' [43]. However, these systems require instrumentation on either the fingertip, thumb, or multiple points across the hand, reducing practicality compared to a single ring mounted at the base of the index finger. Previous preference data indicates that users prefer touching with their index finger as opposed to other indices [16, 50]. For microgestures, bringing the thumb to distal and middle phalanx pad is most comfortable and preferred [20]. Finally, users prefer wearing a single ring at the base of their fingers (over middle or distal phalanx) from a comfort, acceptance, and preference criteria [16], and dislike fingertip mounted systems that can disrupt daily activities [13]. Therefore, to achieve a user-friendly instantiation, for FlowRing we target achieving all interactions from a single ring at the base of index finger, with thumb-to-index microgestures executed on the index fingerpad.

Among these devices, TouchRing offers an always-available interface via finger-mounted printable electrodes and capacitive sensing for discrete microgestures, but it requires the thumb to be bent awkwardly to contact the ring surface [48]. Several recent approaches exhibit microgesture interaction similar to those of FlowRing. Both Cyclops Ring [8] and OptiRing [52] achieve pinching and swiping microgestures using ring-mounted cameras. Z-Ring enables 1D microgesture input through bio-impedance sensing with active RF injection from a wrist-mounted VNA into the index finger [51]. EFRing [9] instead uses a lower frequency capacitive projection method to achieve the same. FlowRing aims to achieve the same microgesture interaction but advances the state of the art by combining microgesture sensing with precise, continuous 2D input from a ring located at the base of the finger, offering a versatile solution that supports cross-context XR use.

4 FlowRing Prototype

4.1 Design Considerations

Toward the goal of realizing a device that can enable versatile and ergonomic input for all-day XR, we lay out four design criteria which FlowRing must achieve simultaneously:

(1) For precise control when stationary, FlowRing must realize 2D continuous input on a set of common surfaces. Index-on-surface interaction has been shown to be comfortable and natural [16, 42, 43]. Importantly, this control must be include not only tracking, but also a selection mechanism, unlike some previous single-ring work ([25]).

Table 1. A table comparing the most relevant prior work to FlowRing. To our knowledge, FlowRing is the first ring device that enables *both* continuous on-surface interaction and microgestures as well as the first device to facilitate seamless transitions across these input techniques and multiple surfaces. It also does so while adopting a common and ergonomic location on the base of the index finger.

Prior Work	Sensing Modality	Ubiquitous Form Factor?	Free-hand Microgestures	Continuous, On-Surface Tracking
3DTouch [37]	Optical flow, inertial	○ Finger pad	○	●
MouseRing [42]	Inertial	○ Proximal phalanx ring	○	●
Magic Finger [62]	Optical flow, camera	○ Finger pad	◒ Only pinch shown	●
VibAware [28]	Active acoustic	○ Thumb + wrist	●	○
DualRing [32]	Bio-impedance, inertial	○ Thumb + index ring	●	○
Ready Set Touch! [43]	Inertial	○ Nail-mounted	● Thumb nail mount	● Index nail mount
Z-Ring [51]	Bio-impedance	● Index ring	●	◒ Pre-calibrated conductive surfaces only
Acustico [14]	Acoustic	● Wristband	○	○ Tap on rigid surface only
AOFinger [60]	Acoustic, IR reflection	● Wristband	●	○
LightRing [25]	Inertial, IR reflection	● Index ring	○	● But no tap
ElectroRing [26]	Bio-impedance	● Index ring	◒ Pinch only	○
FlowRing (this work)	**Optical flow, inertial, acoustic**	● **Index Ring**	●	●

(2) To allow for mobile control, FlowRing must also realize a method ergonomic free-hand gestures. Given user preference around thumb-to-index microgestures on the index finger pad [20], we target these free-hand microgestures. As opposed to previous microgesture work [28, 32, 43, 51, 52, 60], our work must achieve this in addition to the continuous tracking, in a compact form factor using the same sensor set as for the tracking function.
(3) Towards enabling versatility, FlowRing must be able to implicitly translate between these functions without user intervention or explicit mode switching. This is not demonstrated in prior works.
(4) To be acceptable for all-day XR, FlowRing must achieve all the above functions according to established user preferences about all-day acceptability and comfort. Namely, FlowRing must achieve all functions in a single ring, mounted at the base of the index finger [13, 16], allowing it to stay low-profile when not in use. Ideally, the prototype also considers wireless and low-power operation.

4.2 Prototype

Our FlowRing design highlights practicality. We chose low-power sensors in low-profile mounting positions. FlowRing consists of three main sensing components: an optical flow sensor, a contact microphone, and a 6-DOF IMU. For the optical flow sensor, we adapted the Pixart PAT9130,[1] a component usually used in industrial applications to monitor shaft rotation. For our form factor, the large depth-of-field working distance was ideal for tracking centimeters away from the surface, and the integrated VCSEL enables a small total package size (4.6 mm by 4.4 mm by 0.3 mm). We mounted the chip on a custom power management and conditioning PCB.

We collect five features from the sensor at approximately 130 Hz: Δx and Δy values, quality of image keypoints, shutter duration, and average brightness of the auto-exposed frame. Together, shutter and brightness provide a proxy for surface z-distance; however, in practice, they are confounded by surface reflectivity, texture, and sensor angle.

We experimented with two alternative mounting positions for the sensor: (1) mounted to the radial side of the proximal phalanx rotated 45 degrees distally, and (2) mounted on the palmar (bottom) side of the proximal phalanx at a 10 degree angle distally. The first allows for thumb

[1] https://www.pixart.com/

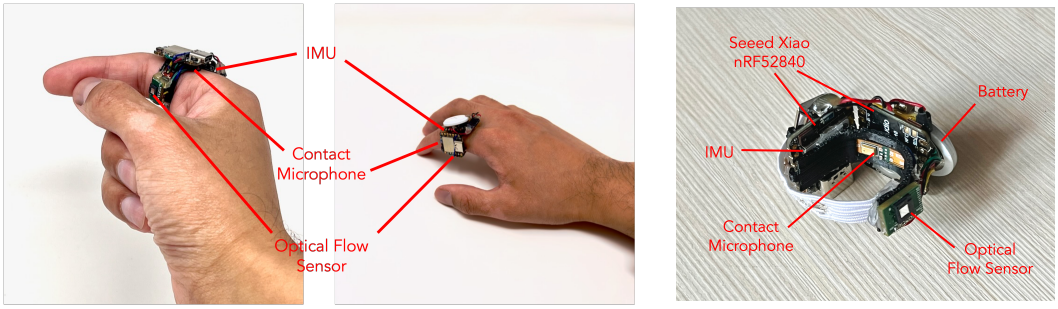

Fig. 3. FlowRing integrates multiple sensors on to a ring. Low-profile sensor placement is considered as part of the design.

microgestures on the radial side of the index finger, where the finger tip is more directly observable; however, for on-surface tracking, we found the steeper glancing angle of the sensor to the surface to be less sensitive for continuous motion. The second mounting position allows observation of the base of the thumb. During microgestures, this resulted in slightly less consistency of the Δx and Δy values due to human variances of hand position and the relatively smaller motion of base of the thumb to that of the fingertip during microgesturing. On the other hand, mounting the sensor more aligned to the surface enabled more robust continuous motion values. After experimentation, we opted for the latter mounting position because (1) our target microgesture set was discrete and therefore potentially less sensitive to inconsistencies, and (2) the second mounting position permitted a lower profile industrial design.

For the contact microphone, we leveraged Knowles V2S200D,[2] typically used in earbuds in conjunction with traditional microphones for enabling voice clarity during noisy conditions, such as wind noise. We positioned the microphone on the inside of the ring, with the package pressed against the dorsal (top) side of the user's finger. We felt that such positioning on the finger's bonier side could increase the vibration conduction signal from the finger pad to the microphone.

Although a higher SNR could be achieved with an analog microphone and separate amplifier, as in previous works [63, 64], we instead opted for a digital PDM microphone for low power and the lack of additional required circuitry, enabling a fully self-contained design. We tested various decimation ratios between a 4kHz and 16 kHz sample rate while rubbing our fingers and surfaces, and we settled on 4 kHz since it still captured the frequency bands of strongest vibrational power while balancing Bluetooth bandwidth and connection robustness.

Finally, we sample a radially mounted 6-axis IMU (ST LSM6DS3TR-C[3]) at approximately 130 Hz. Each sensor is streamed over Bluetooth LE to a laptop for modeling and inference. Our ring uses two Nordic nRF52840s due to their integrated BLE radio and PDM peripheral controller. A 120 mAh lithium-ion battery powers the device. FlowRing's ability to work completely wirelessly surpasses many prior works, allowing for greater user motion during data collection. It also demonstrates the feasibility of our downstream models to work with real-world challenges of mobile hardware, including increased latency, dropped packets, uneven sampling, and lower SNR sensors. All components are mounted atop a 3D-printed chassis, with an elastic fabric allowing for accommodation of users from ring size of 5 to 8 cm in circumference. The ring weighs 9 grams.

[2]https://www.knowles.com/V2S
[3]https://www.st.com/en/mems-and-sensors/lsm6ds3tr-c.html

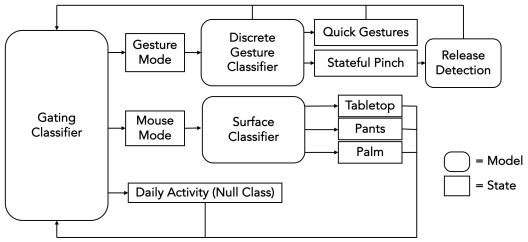

Fig. 4. FlowRing's high-level state architecture. A *gating classifier* rejects false positives from daily tasks and determines if the user performed a gesture or is holding their hand in a mouse-like configuration over a surface; it then routes the samples to the appropriate downstream classifier. If it is a gesture, the *discrete gesture classifier* then determines the type of gesture, i.e., a quick gesture like a tap or swipe, a longer gesture like a pinch and hold, or a negative sample. If it is a mouse-like action, the gating model can then engage a continuous 2D on-surface tracking scheme as well as trigger surface classification.

5 Input Pipeline Design

As noted, FlowRing enables two different input methods, 2D continuous on-surface tracking and free-hand microgestures. To selectively select between these methods, identify surfaces, and reject false positives from daily activities, we introduce a *gating classifier* (see Fig. 4). The following sections describe how we achieve each input method.

5.1 Continuous 2D On-Surface Tracking

Continuous 2D surface tracking requires fast and precise motion, but it also must function across a variety of surfaces, lighting conditions, orientations, and hand shapes. To balance speed and generalizability, we combine both heuristic-based and learned methods. We attempt to emulate the feel of a trackpad to leverage the years of familiarity users have with these types of devices. This task is technically challenging for several reasons: (1) to enable clutching (i.e. lifting the finger or hand to reset new control-display coordinate space), the model must quickly recognize both hand lift off a surface and re-engagement with surface-bound motion, (2) to enable tap selection where the motion sensor is mounted on the finger, the system must quickly recognize finger lift and suppress cursor motion, and (3) the system must support both small, precise finger motions and large display traversals in a limited surface space. The most closely related previous work in this space, LightRing [25], cannot achieve tap selection or clutching since it lacks an accurate measure of whether the user's finger is surface-bound and so cannot replace a mouse. As discussed below, the multi-sensor method is essential for the speed of clutching necessary achieve mouse-like function.

5.1.1 Overall Method. To enable cursor control, we follow the state diagram presented in Figure 5. We transition between four states: a latched (1) idle and (2) cursor motion state, and a spurious (3) clutching and (4) tap state. We describe each below:

5.1.2 Clutching Cursor Motion. Clutching is an interaction that enables unbounded cursor motion across a large display space. It occurs when the user moves their hand to shift the cursor towards a particular direction, wishes to continue to move the cursor in that direction, but runs out of ergonomic physical range of motion or encounters other physical obstructions to their motion. Enabling clutching is particularly important in our application area for three reasons. First, clutching facilitates cursor control across large field-of-view displays, such as those found in virtual environments, since the ergonomic range of motion of the hand affords a much smaller area than

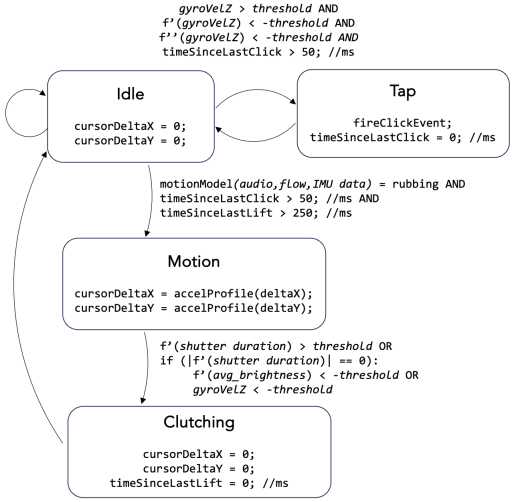

Fig. 5. State diagram of 2D surface interaction for cursor control. Transition conditions are listed next to the arrows.

the display. Second, it is important for limited physical spaces, such as an ad hoc control surface like the pant leg. Finally, it enhances ergonomics since the user can traverse a comparatively large display space with minimal motion of just the finger, wrist, and arm.

To help users apply their previous trackpad usage experience to the cursor control function of FlowRing, we allow users to clutch cursor motion based on whether their index finger is rubbing across any given surface. For real-time usage, a major latency challenge is that users expect perceptually immediate motion of the cursor upon moving their index finger across the surface as well as immediate cessation upon their finger's stopping or lifting from the surface. To achieve precision, users must be able to move the cursor precisely, even with a very slight and slow fingertip motion. For this reason, we chose to adopt a state architecture where we can (1) latch into a *MOTION* state based primarily on a low-latency, high-sensitivity model that detects the start of user finger rubbing, (2) latch out of *MOTION* quickly based on a heuristic, and (3) design an acceleration profile while in a *MOTION* state that suppresses no-motion drift but also enables slow-motion precision. We describe (1) and (2) in this subsection and (3) in the next (Sec.5.1.3).

5.1.3 Entering a Motion State. To anticipate whether a user intends to move the cursor, we use the rubbing of the index finger against a surface to clutch. For imperceptible cursor motion, we aim to achieve a max latency of approximately 100 ms total [4, 36] from start of hand motion to cursor motion. Given the latency of BLE packetization and transmission, this leaves a remaining processing budget of approx 65 ms, or approximately 260 samples of audio and 8 samples of flow and IMU data.

We first attempted to use a heuristic model based on the frequency information present in the audio data alone for this task, hypothesizing that rubbing against a surface should create distinct frequency changes. We did indeed see an increase in power in the 400 to 2000 Hz band compared to a hand kept still. However, in practice, we found it hard to separate distinct frequencies, so we instead opted to develop a lightweight, real-time model.

We collected data from 11 participants (see Sec. 5.2.1). Each user rubbed three surfaces within their vicinity — a desk, their own pants, and their own palm — for three 20-second sessions (remount between each session). Users were instructed to hold their hand in a mouse-like position. This yielded a total of 11 participants × 3 surfaces × 0.33 minute × 3 sessions = 33 total minutes of rubbing data. Each user was instructed to rub at any speed and direction. Since data collection occurred across many different settings, different participants used different desks, pants, and (obviously) own hands. The greatest number of participants who used the same desk was four.

Additionally, we collected negative data on the participants' holding their hand in the air and waving/shaking their hands for 30-second session for 3 sessions with remount, for a total of 11 participants × 2 actions × 0.5 minute × 3 sessions = 33 minutes of negative rubbing data. We also collected 3 sessions of 10 repetitions of each user tapping on the desk to add to the negative data. We trained an SVM that takes as input the fft of 256 samples (64 ms) of audio data. Experimenting with incorporating different sensor inputs did not change the accuracy significantly. In practice, we do not require extremely high model accuracy, rather only high sensitivity, due to latching into a *MOTION* state.

We chose to latch into a *MOTION* state instead of using model output alone to toggle *MOTION*. Model outputs on such short input windows are susceptible to spurious misclassifications. By latching into this state, we could also avoid additional latency associated with majority voting a sliding window across the model's output. We could also enable very precise per-pixel cursor control when the finger was moving very slowly since we could avoid fluctuating model output given minimal or no IMU or audio signal. By latching, this becomes a non-issue.

5.1.4 Exiting a Motion State. In our experimentation, we found exiting the motion state to have even stricter latency requirements than entering. This is due to high user perceptibly of cursor motion immediately prior to selection. As the user lifts their index finger to start a selection tap event, the cursor must immediately decouple from finger motion to avoid moving the cursor off a small target, for example. Therefore, to exit the motion state, we employ a low-latency heuristic on optical and IMU data to detect lift off the surface.

Specifically, we notice that the reflectivity of a given surface remains approximately constant as the user moves their hand for cursor control. The finger or hand lift results in a sudden change in the returned signal power to the optical flow sensor. This is most clearly seen by an increase in *shutter duration*. Therefore, we threshold the derivative of the most recent 4 samples (approx 30 ms) to determine lift. However, for users with large hands, where the sensor sits far from the surface, or for low reflectivity surfaces, we noticed that the shutter duration remained pinned to its maximum value. In that case, we threshold for a decrease on the derivative of *average frame brightness*.

A potential failure mode may be an unintended motion state exit when moving across a highly variegated surface (e.g., a checkerboarded surface). However, in practice, we did not encounter that issue even on patterned desks or on multi-colored paper prints. A final signal of user finger lift is high rotational velocity around the radial-ulnar axis (`gyroVelZ`). Any one of the preceding criteria triggers a fast decoupling of motion from movement.

5.1.5 During Cursor Motion. Once the cursor is in motion, we apply an acceleration profile as a control-display gain function to balance speed and precision. Slow hand motions result in small cursor displacements; quicker hand motions result in disproportionately larger cursor displacements. This allows for both precise motions and display traversal even in limited control area settings, e.g. the top of the thigh. Acceleration is applied on a per time-step basis. Thus, given $\mathbf{u} = \langle \Delta x, \Delta y \rangle$, the accelerated cursor displacement is defined by the equation $\mathbf{u}_{accel} = \mathbf{max}(\mathbf{round}(0.13 \times \|\mathbf{u}\|^2), 300) \cdot \hat{\mathbf{u}}$.

By applying this acceleration profile and rounding, the sensor's Δx and Δy noise are below 0.5 with no motion and therefore drop to zero after rounding. We constrain maximum cursor speed (acceleration multiplier = 0) for very high mouse motion to limit the user's losing their cursor on the screen. We apply the acceleration profile equally in any direction. Even if the sensor is slightly tilted with respect to the surface, we find that it does not change user perception.

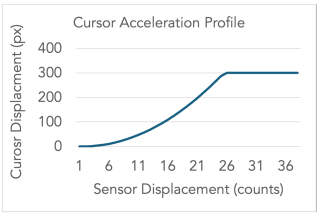

Fig. 6. Acceleration profile applied to cursor movement.

5.1.6 Selection. Finally, we use characteristics of the rotational velocity of the index finger to generate tap events. Specifically, we threshold on high rotational velocity around the radial-ulnar toward the palm, high rotational deceleration, and high rotational jerk. These features are present in the `gyroVelZ` signal when the finger strikes a surface.

We first prototyped heuristics around contact microphone tap events. However, we found finger motion to be a more robust heuristic across different textures, such as soft pants and hands as well as hard desks and walls. To prevent multiple tap events from triggering sequentially, we apply a cool down period of 50 ms. This still enables intentful double clicks that exceed 150 ms between two clicks.

5.2 Discrete Gesture Classification

For the smallest surface of fingertip our system detects thumb-to-index microgestures to control a simple, smartglasses-type UI. Our gesture set consists of five quick gestures: left swipe, right swipe, up swipe, down swipe and tap, and one stateful gesture, pinch.

5.2.1 Overall Method. We detect these five quick gestures and the onset of pinch by a neural network consisting of a 3-layer CNN and an LSTM. For the stateful pinch, an SVM model detects the release of the pinch. To improve the accuracy of detection, we fused data from a contact microphone, an optical flow sensor and a 6-DOF IMU. To evaluate, we performed a leave-one-out experiment, fine-tuning for new users, as well as an ablation study. We describe each in the following sections:

5.2.2 Data Collection and Segmentation. We recruited 13 participants for data collection, however retained data from 11 (5F, 6M) due to significant data loss during streaming in two instances. Their average hand size was 17.94 cm, average ring size was 6.25 cm, and their Fitzpatrick skin tone ranged from I to IV. Each participant spent around 1 hour on data collection activity, which was divided into three sessions. Each session consisted of approximately 5 minutes of discrete gestures data, 5 minutes of negative data, and 10 minutes of on-surface interaction data (see Section 5 for details). In between sessions, the participants were required to remove the ring and re-wear it to simulate different wearing situations. The negative data contains some aggressor actions in everyday life (hard negative: writing, clapping, typing, using the phone, hand clenching, and waving hand freely) and data when the hand is at rest (soft negative).

Because each microgesture is short and the duration of the action differs for each participant, segmenting positive samples is challenging. To solve this problem, we built a data collection game similar in style to "Dance Dance Revolution" with sliding prompt boxes. The participants performed the gesture within a 1.2 second "keep-in" sliding box (green box in Fig. 7), and centered their gesture on a short, 0.2 second prompt (red box in Fig. 7) within that larger box. This approach provided a couple benefits: as we strided our 1-second model input window (blue box in Fig. 7) across the "keep-in" interval, each observation was guaranteed to contain true positive gesture data, and this striding effectively provided time-shift augmentated data.

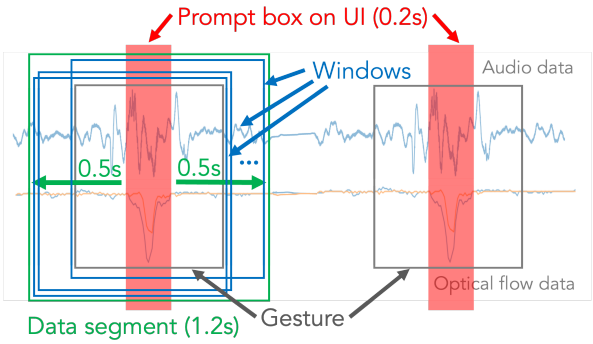

Fig. 7. We showed a prompt box on the user interface (UI) at the middle of each gesture when collecting data. Each data segment (green box) contains a prompt box of 0.2 seconds plus 0.5 seconds before and after it. A 1-sec sliding window is strided to generate multiple data samples from each segment.

5.2.3 Data Preprocessing. We aligned the three types of sensor data by timestamps and segmented the data into samples with a sliding window of 25 millisecond stride. After that, we performed feature extraction on each sample.

Fig. 8 shows an example of the raw data. First, we applied FFT (length of 512, overlap of 64) on the audio data, which produced a $n_{freq} * n_t$ matrix, where n_t represents a microgesture that

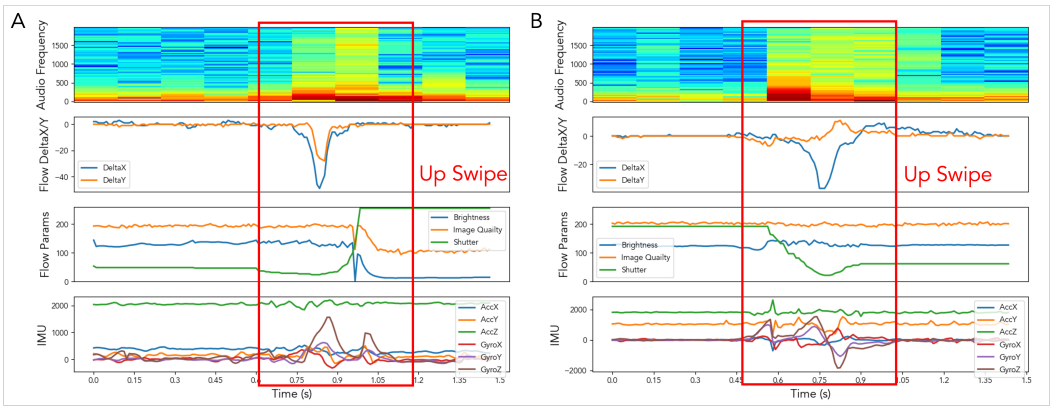

Fig. 8. The raw data from three types of sensors and the spectrogram of audio data. A: An up swipe performed by a participant with small hands. B: An up swipe performed by a participant with large hands. The amplitudes of their optical flow data differ, providing an example of an aspect our model needs to generalize across.

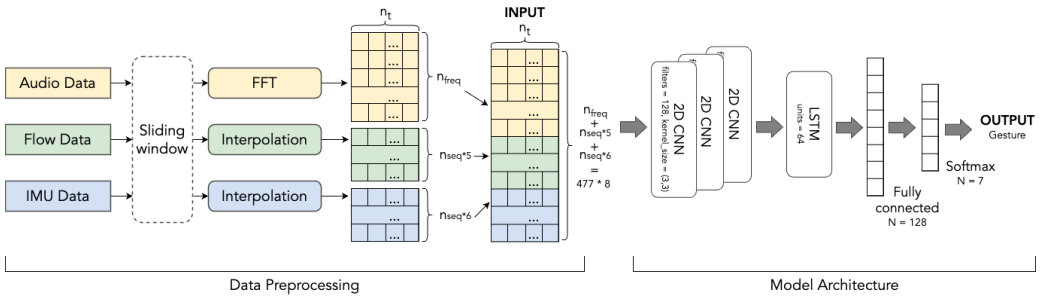

Fig. 9. Model input consists of features extracted from three sensing modalities (left). The quick gesture classification model contains three 2D CNN layers followed by a MaxPooling layer and an LSTM layer. The fully connected layer and softmax are then applied (right).

consists of n_t time sequences. For 5 types of optic flow data (Δx, Δy, image keypoints quality, shutter duration, average brightness) and 6 types of IMU data (6-DOF), we first interpolated the data to match lengths and used the average filter for noise reduction. Then, we reshaped each type of data into a $n_{seq} * n_t$ and concatenated them with the audio feature.

5.2.4 Quick Gesture Classification Model. The main part of the model is a three-layer CNN network and an LSTM network (see Fig. 9). The rationale for these choices is that 2D CNNs perform well when extracting features from 2D images, while LSTMs can capture temporal relationships in the sequence data. Each CNN layer has a depth of 128 and a kernel size of 3×3. This is followed by a maxpooling layer with a pool size of 1×2, where the size on the temporal channel is 1 to avoid losing the temporal information. We also added a dropout layer and a regularizer to prevent overfitting. We also experimented with Transformer architectures (single and dual attention layer), both produced similar results to the CNN-LSTM.

During training, we used a dynamic learning rate starting at 1e-5 and added early stopping to prevent overfitting. We used the sparse categorical cross-entropy as the loss. The model was trained on an RTX 4090 graphics card.

5.2.5 Stateful Pinch Detection. The fast gesture classification model detects the onset of a pinch. After it is detected, the state of the pinch is entered. Then, we use an SVM (C=1, kernel=rbf) model to detect the release of the finger, which is also the end of this pinch.

5.3 Gating and Surface Model

5.3.1 Gating Model. Given that our system offers two modes of input—free-hand and on-surface—we built a gating model to discern and respond to the user's specific interaction. This model processes input data streams from all sensors, including optical flow, acoustic, and IMU data, to categorize the input into three distinct categories: free-hand input, on-surface input, or no input. The purpose of this model is twofold: it aids in identifying the user's intended interaction and effectively filters out false positives when no participant interaction is detected.

To train this model, we repurpose the data collected from discrete gesture studies and continuous input experiments. We consolidate individual gesture classes into a unified *gesture mode* class, merging actions such as tapping, swiping, and others. Additionally, we combine instances of rubbing on pants, hands, and tabletop into a collective *continuous mode* class. Finally, we group non-input-related activities, such as writing, 'clapping, typing, phoning, and clenching, into a single "null" class to serve as hard negative data.

The gating model's architecture and training methodology closely mirror those used to train the discrete gesture classifier model. See Section 5.2.4 for in-depth information.

5.3.2 Surface Model. To support seamless cross-surface interaction, we built a surface model to distinguish the type of surface the participant interacted with. The model's input is the same as the gating model to reduce the number of data processing steps. The model's output is the three surfaces, "tabletop", "pants" and "palm", as well as "null", which is composed of soft and hard negative data (e.g., "typing" and "writing"). The fourth surface, fingertip, is returned if the gating model detects *gesture mode*. The architecture and training methodology are also similar to those used for the gating model.

6 Performance Evaluations

We present the performance metrics for each of the interactions FlowRing affords: 2D selection, thumb-to-index microgestures, and the ability to switch between these functions.

6.1 2D Selection Interaction Performance

To measure the performance of our pointing and selection system, we invited ten participants to try our system in a 2D Fitts'-style evaluation.

6.1.1 Participants. Ten participants' (2F, 8M) index finger ring circumference ranged from 6 to 7 cm (μ = 6.5 cm, σ = 0.46 cm) and hand lengths (palm to middle finger tip) ranged from 17 cm to 20 cm (μ = 18.4, σ = 1.29 cm). All participants had never previously used FlowRing for cursor control.

6.1.2 Experiment. We tested FlowRing in three conditions, i.e., on a desk surface, on the participants' own pant legs, and own palm. We use Wobbrock's FittsStudy application [55] to run an ISO 9241-9 compliant study on an 27-inch Dell Monitor U2719DC (screen resolution 2560 × 1440 pixels, 60 Hz refresh rate), where the system cursor is controlled using the preceding mouse algorithms running in real-time on an Alienware 17 R4 laptop.

Users performed six *Amplitude (A)* × *Width (W)* conditions with 2 levels of *W* (50 and 150 pixels = 1.1 and 3.3 cm) and 3 levels of *A* (250, 350, and 450 pixels = 5.7, 8.0, 10.2 cm) to produce nominal *Indices of Difficulty (ID)* of 1.415, 1.737, 2.0, 2.585, 3.0, 3.3219 bits (though univariate Crossman-corrected IDe's were used for throughput calculation). Participants were instructed to wear the ring on their right index finger, were shown the hand position for sliding and clicking, and were given a set of practice targets until they felt comfortable. We note that participants quickly adapted to FlowRing's interaction by applying their similar experiences with mouse or trackpad-like motion. No participant needed more than one minute before declaring they were ready to commence the trials. We collected one $A \times W$ block on each surface. The order of *ID* conditions was randomized, and the surface was counter-balanced.

Because our ring uniquely enables pointing *and* select functions, and due to the unavailability of previous pointing-only hardware like LightRing [25], there is no comparable state-of-the-art device to baseline against in a straightforward manner. We also collected as an upper-end an additional session on the laptop's built-in trackpad, despite its being a stationary and low-latency wired device. Due to the different affordances between our wearable and the trackpad, we provide it as a comparative data point rather than as a baseline condition to beat to allow researchers to calibrate future cross-study findings.[4]

[4]Due to the difficulty of reproducing hardware prototype systems, we encourage future researchers to do the same so results can be compared.

6.1.3 Results. All participants successfully completed all conditions, indicating that FlowRing can enable surface interaction in a user-independent manner. The average movement and selection time for a desk is 1637 ms (σ = 283), for pants is 1603 ms (σ = 353), and for palm is 2153 ms (σ = 543) (Fig. 10, center). Participants felt the system was quick and intuitive on desk and pants; a sense of system performance can be best understood from the Supplementary Video. However, performance was comparatively worse on the palm due to its small tracking size and noisier IR reflectivity. The palm therefore would likely be more appropriate for 2D touchscreen like manipulations rather than cursor control across a large screen. Two participants reported that given the roughness of the desk we used and the stickiness of their fingertip from sweat, they experienced friction, especially during upwards movements. All participants could clutch without issue. Most users were able to tap without issue. Two participants occasionally needed to tap multiple times on the desk to trigger selection; after being reminded to tap swiftly, they had no issue, and their data was retained. However, before correction, their cursor movement was more erratic during touchdown and thus stood as error rate outliers. Error rates were 6.25%, 4.58%, 9.52% for desk, pants, and palm respectively (Fig. 10, center).

We calculate throughput via a mean-of-means method [46], finding good agreement across users. The average TP_{avg} across users for the desk condition was 1.54 bits/sec (σ = 0.23), for pants was 1.64 bits/sec (σ = 0.27), and for palm was 0.86 bits/sec (σ = 0.21) (Fig. 10, left). Significant differences (ANOVA, followed by t-test with $p < 0.05$) were observed across nearly all pairwise results, except between desk and pants. There, a paired t-test indicated no significance ($p = 0.10$) as well as no correlation across participants. This result was moderately surprising since we expected the curved and soft nature of the pant leg to make cursor control more challenging. However, this result indicates the potential of FlowRing to work on non-traditional, ad hoc surfaces. For reference, the TP_{avg} of the built-in trackpad was 2.66 bits/sec (σ = 0.38), and the movement and selection time was 1101 ms (σ = 181). We also compare performance with respect to other ring-based cursor systems, though caution that throughput values across different experiments and set up can be inconsistent, especially those that do not follow the ISO 9241-9 standard. MouseRing reports a throughput of 11.98 bits/sec for touchpad, 10.69 bits/sec for a dual-ring configuration, and 7.22 bits/sec for a single ring worn on the middle phalanx of the index finger. Normalizing these values against our touchpad result, yields a normalized throughput of 2.37 bits/sec for two rings and 1.60 bits/sec for single ring. The latter result is similar to the performance of FlowRing, however unlike FlowRing, the MouseRing single-ring configuration is located on the less desired middle phalanx [16] and requires the palm to be fixed in place. LightRing reports throughput as 2.9 bits/sec for desk and 2.8 bits/sec for leg, however does not provide a standardized input device to allow for normalization of these values. Similar to MouseRing, FlowRing requires a fixed palm. Furthermore, LightRing cannot perform selection so users must use the other hand on the spacebar to perform the Fitts' test, further complicating direct comparison.

We assessed whether participant hand size correlated with throughput across three surfaces: *desk*, *pants*, and *palm*. Spearman's rank correlation revealed slight negative trends for all surfaces: $\rho = -0.57$ for desk, $\rho = -0.32$ for pants, and $\rho = -0.46$ for palm. These suggest that participants with larger hands may exhibit slightly lower throughput. Pearson correlations mirrored this trend ($r = -0.40$, -0.26, and -0.41, respectively), but none reached statistical significance (all $p > 0.2$).

6.1.4 Surface and Lighting. To survey how well FlowRing's 2D continuous tracking would work across a variety of everyday scenarios, we conducted a brief experiment across different surfaces and lighting conditions. To ensure repeatability, we attach FlowRing to a mock finger and conduct a drag test across 100 mm at a fixed velocity across a set of common surfaces. These surfaces varied in color, specularity, firmness, and microstructure. Specifically, we test white copy paper (control),

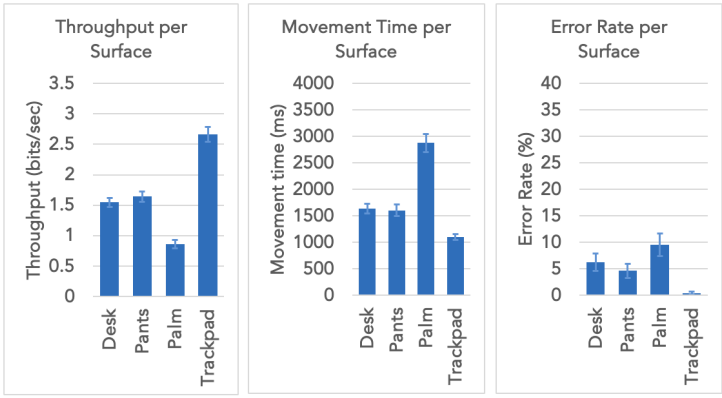

Fig. 10. Results from the Fitts' pointing and selection study. The larger desk and pants surface performed similarly, the smaller palm surface is slower for the same pointing task, potentially due to smaller control area. Error bars represent standard error. For comparison, we provide data from a stationary, wired trackpad.

white laminate wooden desk, black construction paper, black leather, white canvas, white glossy quartz counter, beige linen, black twill, blue denim, a plastic mirror, black rough plastic, glass phone screen (display off), forearm (Monk VI), forearm (Monk III), palm (Monk V), and palm (Monk III). We perform five trials for each surface, and report cursor travel normalized to the control case (white copy paper). The ring was kept at a fixed distance of 15 mm from the surface (approximately equal to median hand) and lighting was that of an indoor office. Results are presented in Figure 11. Performance was generally comparable across a number of fabrics, papers, wood, and quartz. Performance degraded on skin, especially that of the palm, in line with Fitts' results. The system was unusable on IR reflective and absorptive surfaces such as a mirror and rough dark plastic. Overall, these results indicate that visual color may not be exactly correlated with performance, but IR scattering properties and microstructure may play a role (especially given the difference between forearm and palm skin).

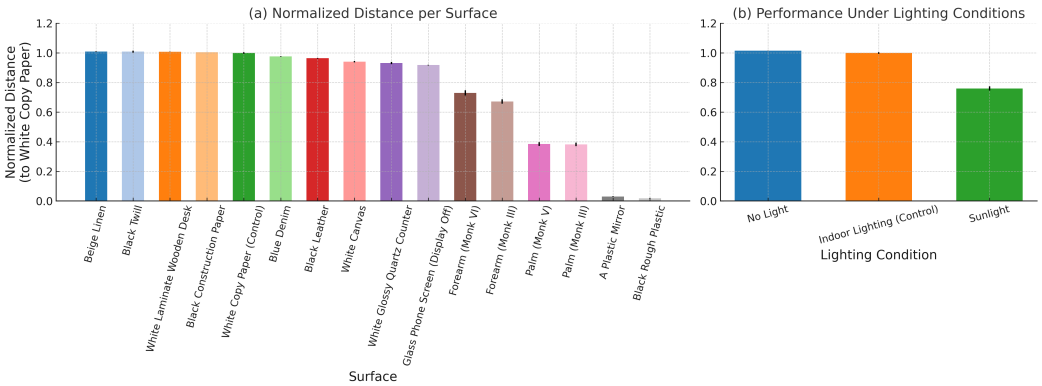

Fig. 11. Normalized distance traveled during the same task across different conditions. **(a)** Normalized distance across a variety of material surfaces, relative to white copy paper. **(b)** Normalized distance on white copy paper under different lighting conditions, relative to indoor lighting. Error bars represent standard error.

To investigate the effects of lighting on FlowRing's motion performance, we compared our control, white copy paper, across three lighting conditions: the aforementioned indoor lighting case, outdoor sunlight, and a dark room. As the sensor uses its own emitter, the dark room did not affect performance, however bright sunlight did decrease the measured normalize distance, likely because of conflict with the sensor's IR emitter. Across trials the standard error remains small, however, indicating that while throughput would certainly decrease in sunlight conditions, sensor noise remains low. We discuss possible mitigation methods in Sec. 7.2.2.

6.2 Microgesture Recognition Performance

The fast gesture classification model classifies the gestures into 7 categories: left swipe, right swipe, up swipe, down swipe, tap, the onset of a stateful pinch, and none of the above. Below, we first report the main gesture classification results of our model. Then, to investigate the potential of one-time calibration, we report results of a separate fine-tuning experiment we performed to measure the model's ability to learn new users' data, during which a small amount of data from a new user was added to the training set. Finally, we show results from an ablation study to observe the effects of different sensing modalities of data on the results.

6.2.1 Gesture Classification Results. For our main result, we generate three accuracies: within-session, leave-one-session-out (LOSO) and leave-one-user-out (LOUO) from the participant data discussed in the data collection section above (see Section 5.2.1). For the within-session setting, we mixed data from all users and all sessions and divided it into the training set and the testing set according to 80:20. In the LOSO setting, we used the data from two sessions for training and from the other session for testing. We repeated for all three sessions and calculated the average accuracy. In the LOUO setting, we used the data from ten participants for training and tested on the data from the left-out participant. We repeated this 11 times for each participant and calculated the average accuracy.

Table 2 shows the main results. The accuracy of the within-session setting is 93.86%, which shows that our model can learn the characteristics of different discrete gestures well. The LOSO result is slightly lower than the within-session setting, with an accuracy of 93.61% demonstrating that our model can also work well across sessions and may not need the per-session calibration. For LOUO, the average accuracy is 85.22%, with the highest accuracy of 96.75% (P2) and the lowest of 76.14% (note that P10 has the largest hand among participants), as shown in Figure 13. Beyond these outliers, we found no significant differences or general trends in performance between groups based on skin tone, ring size, or hand length. A Mann–Whitney U test revealed no effect of skin tone on performance ($U = 19.0$, $p = .41$), nor of ring size ($U = 7.0$, $p = .38$) or hand length ($U = 12.0$, $p = .79$). Our experiments with Transformer-based architectures (single and dual attention layers) yielded LOUO accuracies of 82.0% and 83.1% respectively.

Despite no significant general trend, hand sizes and finger lengths do vary across users, which may result in differences in the same gesture. For instance, as shown in Fig. 8, the data of one of the participants with smaller hands does not change in Δy as much as that of users with large hands when performing an up swipe. Therefore, we consider the possibility of a limited, one time calibration per user in the next section.

Table 2. Main results of the quick gesture classification model.

	Within-Session	LOSO	LOUO
Accuracy	93.86	93.61	85.22
F1-score	93.86	93.59	84.84

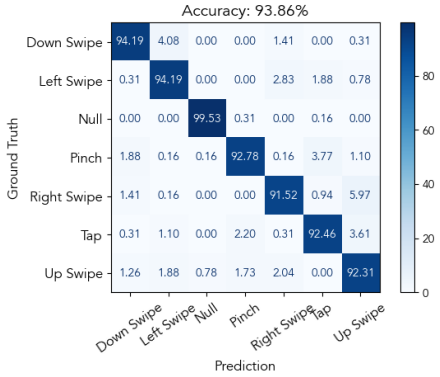

Fig. 12. The confusion matrix of the quick gesture classification model.

Limited Fine-Tuning for New Users. Rings are a personal device sized to a single user, therefore similar to FaceID or pHRTF ear scan for spatial audio, we are interested in exploring the potential of a minimal, one-time calibration for boosting accuracy. To investigate whether our model can learn the features of a new user's gestures from very few samples, we added a small amount of data from the new participant to the training set when doing LOUO. We used 10 participants' data from all three sessions plus a small amount of data from the left-out participant's first sessions for training and then tested on the data from remaining two sessions of this participant. We did the same process for each user and calculated the average accuracy. In two experiments, we randomly chose two repetitions and four repetitions of each gesture and added them to the training set. The results in Table 3 show that the accuracy of the two experiments improved by 3.32 and 4.93 percentage points, respectively, indicating that a very limited one-time calibration scheme can yield high accuracy even across remounts.

Table 3. The results of minimal new user calibration. Accuracy for leave one user out, inclusion of two gesture set examples from a new user, and of four.

	LOUO	LOUO +2 repetitions	LOUO + 4 repetitions
Accuracy	85.22	88.54	90.15
F1-score	84.84	88.41	90.01

Ablation Study. As shown in Table 4, we performed an ablation study to identify the contribution of each modality by removing the features extracted from each. We compared the results to the within-session setting.

After removing the audio features, accuracy drops marginally. This confirms our assumption that the audio data obtained from the contact microphone can help detect rubbing but is either not very useful for distinguishing between different gestures or is partially redundant to the IMU. The results after removing optical flow and IMU data also show that they contribute more to gesture classification, capturing motion as well as rubbing. As shown in Fig.8, there is a clear signal when the gesture is performed.

Result of Stateful Pinch Model. The negative data of this setting is the data when the hand is keeping still, and the positive data is the release data. We trained a simple SVM on the data from

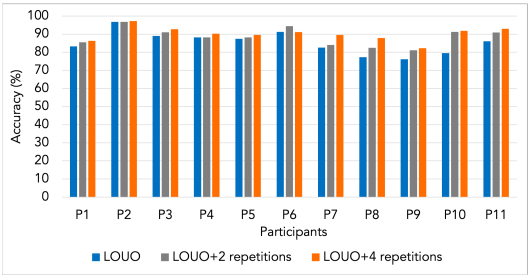

Fig. 13. The per-participant accuracy of the quick gesture classification model after adding two or four repetitions of the gesture set.

Table 4. Results of the ablation study.

	Within-Session	w/o Audio	w/o Optic flow	w/o IMU
Accuracy	93.86	92.24	86.27	87.62
F1-score	93.86	92.23	86.32	87.63

the first two sessions and tested it on data from the last session. The accuracy of our pinch release model is 94.2%. Since this is gated by the gesture classification model, we found that an inexpensive release model is sufficient for real-world use.

6.2.2 Gating and Surface Detection Performance. The evaluation method and dataset split are similar to the evaluation for discrete gesture classification (see Section 6.2).

The accuracy of the gating model under the within-session setting 99.21%. Results of LOUO and LOSO are slightly less accurate, but both exceed ninety percent (see Table 5). This shows that our gating model works well across users and sessions in real-world scenarios. The inference time of this model is 67 ms. During our real-time demo, we use a majority vote of 10 samples, enabling interaction mode switching within 0.67 sec.

The accuracy of the surface model under the within-session setting is 97.97%, and the result of LOSO is slightly less accurate. However, the accuracy of LOUO drops to only 85.76%, perhaps due to the variety in colors and materials of participants' pants and the colors and textures of their palms.

Table 5. Results of the gating and surface model.

	Within-Session	LOSO	LOUO
Gating model accuracy	99.21	97.30	93.82
Gating model F1-score	99.21	97.30	93.83
Surface model accuracy	97.97	94.70	85.76
Surface model F1-score	97.97	94.71	85.25

6.3 Real-time Performance

We evaluated the real-time accuracy of our approach in recognizing gestures and surfaces on six users who were not included in the training data. We mixed and shuffled all kinds of surfaces and microgestures, repeating each category five times. During the user study, participants played with

FlowRing for around 5 minutes to learn microgestures. Then they were asked to perform each surface or microgesture according to the instructions.

The average accuracy is 79.17% which is a little lower than the accuracy of the leave-one-user-out setting (85.22%). Considering that the final accuracy is the product of the accuracy of the gating model (93.82%) and the gesture model(85.22%) or the surface model (85.76%), this result can support the accuracy results shown in Section. 6.2.

We also interviewed participants for their subjective feedback. They mentioned that social acceptability and portability are better than traditional XR inputs such as the controller or mouse. However, the precision of the surface interaction still needs to be improved. In addition, one participant mentioned that the microgesture which were required to be performed on the palmar side of the index finger (defined as "front" in [6]) feels a bit uncomfortable.

7 Discussion

We proposed FlowRing, a discreet ring with the unique capability of being a highly versatile, multi-function input device. We next contextualize these functions in exemplary application scenarios, as well as discuss the limitations and future work of our approach.

7.1 Applications

FlowRing's interaction space extends seamlessly across surface sizes from desks to thigh to palm to fingertips. This continuum (see Fig. 1) can be mapped to support applications that span multiple dimensions, such as different postures, different UI affordances, and different devices. We build and demonstrate these applications in the accompanying Video Figure: https://youtu.be/efFffWrRg3E.

7.1.1 Cross-Posture Control. As users leverage a given computing device in different scenarios, their posture may change with use. For example, a user wearing an XR device may start at their desk. Since they have the greatest level of expressivity and precision on this surface, they may use precise pointer control with an application (e.g., Spotify's music player) in a desktop version, with high visual and target density. If they were to move to their couch, FlowRing would detect the user's thigh as the available interaction surface. The new surface would still allow for cursor control, but perhaps with slightly less precision than a perfectly flat, hard and smooth desktop surface. The headset could then show them a scaled Spotify UI with larger target elements and a cursor (e.g., like Apple iPad Pro's cursor experience). Finally, if they were to take a walk, the app might shrink to provide only key controls and information, such as a music miniplayer. To navigate the limited elements of the miniplayer, the individual could use discrete microgestures to select amongst the few interactables (Fig. 14).

7.1.2 Cross-UI Affordance Control. Within an application or digital environments, multiple controls can be mapped to multiple surfaces based on their affordances. For example, for a VR whiteboarding application, a user could draw on the canvas with a pencil tool. A 2D color circle picker could then let them select their desired inking color. Further, tool modes, such as eraser, line, circle, etc., could be available in a tool menu. To couple the UI elements to surfaces based on affordances, control of the canvas inking could be mapped to the largest and most precise surface: the tabletop. The ad hoc palm surface could enable a shortcut for quick color adjustments with finger drags. Finally, discrete swiping microgestures could provide a shortcut for quickly switching tools (Fig. 15).

7.1.3 Cross-Device Control. Users now commonly interact with multiple devices concurrently. With FlowRing, different devices can be associated with interactions on different surfaces. For example, users could use a mouse-like pointer on the desk surface to interact with their desktop VR or computer. Concurrently, they could control their connected earbuds with specific microgesture

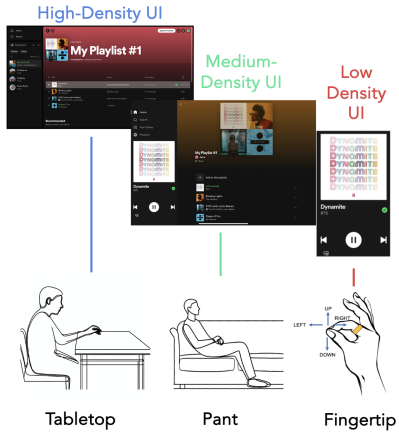

Fig. 14. FlowRing enables seamless input across different user postures. As FlowRing recognizes different surfaces, the user wearing an XR headset would be provided a desktop-like, high-density UI while seated at a desk with high precision, a tablet-like, medium-density UI while seated on the couch with gestures on pant, and a mobile-like, low-density UI while on the go with fingertip microgestures.

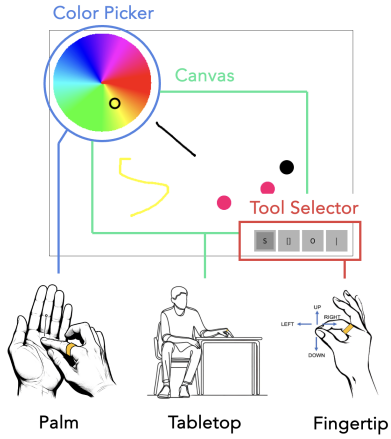

Fig. 15. UI elements can be mapped to surfaces according to affordance. FlowRing allows for color to be selected on the palm, the canvas to manipulated via tabletop swipes, and tools to be selected with fingertip microgestures.

controls (swipe left/right for prev/next song, up/down for volume control, tap for play/pause) and adjust their environment by linking swipes on the palm with the intensity and warmth of their smart lights (Fig. 16).

7.2 Limitations and Future Work

We look toward further characterizing our device, addressing limitations and improving our design in future iterations.

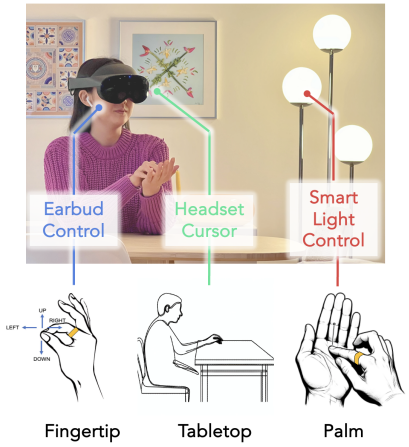

Fig. 16. Multiple devices can be associated with specific surfaces to enable seamless cross-device control. The intensity and warmth of the smart lighting can be controlled with palm swipes, the earbuds' music functions can be controlled via fingertip microgestures, and the headset's cursor can be controlled on the tabletop.

7.2.1 Hardware Power and Size Optimization. Though our prototype device demonstrates completely wireless functionality, there is room for improvement. For ease of maintaining bandwidth across various pilot experiments, our prototype used two microcontrollers and radios, but a future implementation could easily use one. Despite this, our ring consumed only 91 mW when constantly streaming data, or approximately 5 hours on the 120 mAh battery. This power consumption includes two MCUs, both power supplies for each, two BLE active radios, the optical flow chip and associated regulators (30 mW), the contact microphone (1.5 mW) and the IMU (1.0 mW). Power and form factor could be improved by consolidating these components. Unlike other optical-based devices [27, 52] that may require line-of-sight to the fingertips for function, we instead design FlowRing to keep sensors in a low-profile orientation. Fig. 17 shows an example of what a future system could look like with a custom flexible printed circuit (FPC).

Further power savings could be realized by duty-cycling the ring to wake only on distinct triggers, like a unique accelometer-based double tap or an incoming message on a head-worn display. Finally,

Fig. 17. Mock-up of a future FlowRing with a custom flex PCB.

investigating quantizing and shrinking the gating model to run on-device would allow the radios to remain off the vast majority of the time, waking only when needed.

7.2.2 Identifying Learning Effects and Fatigue Effects. Within our Fitts' evaluation, we performed only a single block of trials for each condition. It would be interesting to extend this evaluation to a longer time period to identify if users' performance would improve as they grew accustomed to the device or deteriorate if fatigue or other effects, such as sweaty fingertips, became a factor.

7.2.3 Longitudinal Study. We provide a first investigation and concept for a multi-surface ring. To allow the system to generalize to truly real-world use, a longitudinal, in-the-wild data collection and study would allow for a greater number and variety of observations across re-wears, surface types, motions, ambient lighting, and demographics. The fact that our ring is wireless lends itself toward such a data collection, however more engineering would need to be done to make the logging platform itself truly mobile. Generally, we found our flow sensor works well on the same surfaces that optical mice can function: Lambertian reflectivity without high IR absorption (e.g. worse performance on a clear glass table). This mainly has implications for on-skin interaction on textured areas such as palm as well as certain IR absorptive plastics. For surfaces that exhibit low IR reflectivity, a potential solution would be to adjust the gain accordingly to maintain tracking speed across these surfaces. For surfaces yielding high sensor noise, low-pass filtering could be applied, but would affect end-to-end responsiveness.

Based on previous user research, we target FlowRing to achieve interactions that are considered ergonomic—surface interaction [16, 43] and microgestures [7, 33, 38, 56]—as well as gesture set that is semantically aligned with UI affordances and familiar interfaces like touchscreens [7, 20, 50]. A longitudinal evaluation could more explicitly test across all common surfaces as well as provide an assessment of long-term comfort, ergonomic strain, and social acceptability.

7.2.4 3D Interaction Techniques. Spatial computing devices also feature 3D interactions and manipulations. While we do not specifically investigate such interactions techniques in this paper, FlowRing has the potential to allow for 3D manipulations with IMU + pinch detection. It also can be potentially used for on-surface 3D manipulation by using touchdown + finger yaw interaction techniques, such as those proposed in 3DTouch [37].

7.2.5 Practical Usability Aspects. While this research demonstrates the potential of microgestures plus surface interaction recognition using an on-body ring, several practical usability aspects warrant consideration:

Interaction Challenges for Users with Long Nails: Our investigation focused on interactions performed with the fingertip. Individuals with long fingernails may experience altered interaction mechanics with surfaces. The audio characteristics produced by a nail scraping a surface are likely distinct from those generated by skin contact, potentially impacting the accuracy of our models. Our user study did not include participants with long nails, and future research should explore the system's performance and required adaptations for this user group.

Potential for Fingertip Irritation: The proposed interaction paradigm involves repeated contact with external surfaces. Depending on the texture and material of these surfaces, prolonged or frequent interaction could lead to fingertip irritation or discomfort for some users. A long-term study is necessary to evaluate the potential for such effects and to explore mitigation strategies, such as providing feedback on interaction frequency or suggesting softer interaction surfaces.

Microgesture Accuracy: As our limited fine-tuning model for microgestures performs superior to our user-generalizable model, practically, users would need to perform a one-time calibration with at least four gesture set examples. While we expect this calibration procedure to be minimal (less than 30 seconds), user feedback would guide the acceptability of this procedure for real-world use.

Ring Orientation Dependency: The current sensor configuration necessitates a specific orientation of the ring on the user's finger for optimal performance. This requirement introduces a potential usability challenge, as users must consciously ensure correct orientation. To address this, future iterations could explore incorporating physical affordances, such as a distinct jewel or marking on the top of the ring, to provide a tactile or visual cue for proper alignment and encourage consistent wearing behavior.

Addressing these challenges in future research will be crucial for enhancing the practical usability of this technology.

8 Conclusion

This paper introduced FlowRing, a novel ring wearable designed to cater to the varied input needs of both on-the-move and immersive desktop contexts. Our investigation demonstrated that FlowRing, which features an optical flow sensor, skin-contact microphone, and IMU, can facilitate versatile input across tabletop, pants, palm, and fingertip. User studies substantiated its potential, showcasing promising gesture recognition accuracy across sessions and users. Furthermore, a successful 2D Fitts law test underscored the system's effectiveness in handling continuous input tasks. FlowRing represents a significant step forward in enhancing mixed reality interactions and holds promise for further advancements in this field.

Acknowledgments

The lead author is supported by National Defense Science and Engineering Graduate (NDSEG) Fellowship Program. This work is supported by Washington Research Foundation Entrepreneurship Endowment and Key Research and Development Program of Ningbo City under Grant No.2023Z062.

References

[1] Daniel L Ashbrook. 2010. *Enabling mobile microinteractions.* Georgia Institute of Technology.

[2] Tomislav Bezmalinovic. 2024. Meta CTO confirms work on lightweight mixed reality glasses. https://mixed-news.com/en/meta-cto-puffin/

[3] William Buxton et al. 1990. A three-state model of graphical input. In *Human-computer interaction-INTERACT*, Vol. 90. Citeseer, 449–456.

[4] Stuart K. Card, George G. Robertson, and Jock D. Mackinlay. 1991. The Information Visualizer, an Information Workspace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New Orleans, Louisiana, USA) *(CHI '91)*. Association for Computing Machinery, New York, NY, USA, 181–186. https://doi.org/10.1145/108844.108874

[5] Ashley Carman. 2019. North Focals glasses review: a 600 dollar smartwatch for your face. https://www.theverge.com/2019/2/14/18223593/focals-smart-glasses-north-review-specs-features-price

[6] Adrien Chaffangeon Caillet, Alix Goguey, and Laurence Nigay. 2023. μglyph: a microgesture notation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–28.

[7] Edwin Chan, Teddy Seyed, Wolfgang Stuerzlinger, Xing-Dong Yang, and Frank Maurer. 2016. User elicitation on single-hand microgestures. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 3403–3414.

[8] Liwei Chan, Yi-Ling Chen, Chi-Hao Hsieh, Rong-Hao Liang, and Bing-Yu Chen. 2015. Cyclopsring: Enabling whole-hand and context-aware interactions through a fisheye ring. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. 549–556.

[9] Taizhou Chen, Tianpei Li, Xingyu Yang, and Kening Zhu. 2023. EFRing: Enabling Thumb-to-Index-Finger Microgesture Interaction through Electric Field Sensing Using Single Smart Ring. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 4, Article 161 (jan 2023), 31 pages. https://doi.org/10.1145/3569478

[10] Paul Dietz and Darren Leigh. 2001. DiamondTouch: a multi-user touch technology. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology* (Orlando, Florida) *(UIST '01)*. Association for Computing Machinery, New York, NY, USA, 219–226. https://doi.org/10.1145/502348.502389

[11] Yuran Ding, Craig Shultz, and Chris Harrison. 2023. Surface I/O: Creating Devices with Functional Surface Geometry for Haptics and User Input. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–22.

[12] João Marcelo Evangelista Belo, Anna Maria Feit, Tiare Feuchtner, and Kaj Grønbæk. 2021. XRgonomics: facilitating the creation of ergonomic 3D interfaces. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11.

[13] Gauthier Robert Jean Faisandaz, Alix Goguey, Christophe Jouffrais, and Laurence Nigay. 2022. Keep in touch: Combining touch interaction with thumb-to-finger μgestures for people with visual impairment. In *Proceedings of the 2022 International Conference on Multimodal Interaction*. 105–116.

[14] Jun Gong, Aakar Gupta, and Hrvoje Benko. 2020. Acustico: surface tap detection and localization using wrist-based acoustic TDOA sensing. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. 406–419.

[15] Jun Gong, Yang Zhang, Xia Zhou, and Xing-Dong Yang. 2017. Pyro: Thumb-tip gesture recognition using pyroelectric infrared sensing. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. 553–563.

[16] Yizheng Gu, Chun Yu, Zhipeng Li, Weiqi Li, Shuchang Xu, Xiaoying Wei, and Yuanchun Shi. 2019. Accurate and low-latency sensing of touch contact on any surface with finger-worn IMU sensor. In *Proceedings of the 32nd annual ACM symposium on user interface software and technology*. 1059–1070.

[17] Jeffrey T Hansberger, Chao Peng, Shannon L Mathis, Vaidyanath Areyur Shanthakumar, Sarah C Meacham, Lizhou Cao, and Victoria R Blakely. 2017. Dispelling the gorilla arm syndrome: the viability of prolonged gesture interactions. In *Virtual, Augmented and Mixed Reality: 9th International Conference, VAMR 2017, Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9-14, 2017, Proceedings 9*. Springer, 505–520.

[18] David Heaney. 2024. *Meta's 'Orion' Prototype AR Glasses Have 70 Degree FOV And A Wireless Compute Puck*. https://www.uploadvr.com/meta-connect-2024-orion-prototype-ar-glasses/

[19] Steve Hodges, Shahram Izadi, Alex Butler, Alban Rrustemi, and Bill Buxton. 2007. ThinSight: versatile multi-touch sensing for thin form-factor displays. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology* (Newport, Rhode Island, USA) *(UIST '07)*. Association for Computing Machinery, New York, NY, USA, 259–268. https://doi.org/10.1145/1294211.1294258

[20] Da-Yuan Huang, Liwei Chan, Shuo Yang, Fan Wang, Rong-Hao Liang, De-Nian Yang, Yi-Ping Hung, and Bing-Yu Chen. 2016. Digitspace: Designing thumb-to-fingers touch interfaces for one-handed and eyes-free interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 1526–1537.

[21] Apple Inc. [n. d.]. Apple Vision Pro. https://www.apple.com/apple-vision-pro/

[22] Yasha Iravantchi, Yi Zhao, Kenrick Kin, and Alanson P Sample. 2023. SAWSense: Using Surface Acoustic Waves for Surface-bound Event Recognition. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[23] Ross Johnstone, Neil McDonnell, and Julie R Williamson. 2022. When virtuality surpasses reality: Possible futures of ubiquitous XR. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–8.

[24] Arjun Kharpal. 2024. Qualcomm says it's working on mixed reality smart glasses with Samsung and Google. https://www.cnbc.com/2024/09/05/qualcomm-working-on-mixed-reality-smart-glasses-with-google-samsung.html

[25] Wolf Kienzle and Ken Hinckley. 2014. LightRing: always-available 2D input on any surface. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 157–160.

[26] Wolf Kienzle, Eric Whitmire, Chris Rittaler, and Hrvoje Benko. 2021. Electroring: Subtle pinch and touch detection with a ring. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.

[27] David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. 2012. Digits: Freehand 3D Interactions Anywhere Using a Wrist-Worn Gloveless Sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA) *(UIST '12)*. Association for Computing Machinery, New York, NY, USA, 167–176. https://doi.org/10.1145/2380116.2380139

[28] Jina Kim, MinYung Kim, Woo Suk Lee, and Sang Ho Yoon. 2023. VibAware: Context-Aware Tap and Swipe Gestures Using Bio-Acoustic Sensing. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 609–610.

[29] Kenrick Kin, Chengde Wan, Ken Koh, Andrei Marin, Necati Cihan Camgöz, Yubo Zhang, Yujun Cai, Fedor Kovalev, Moshe Ben-Zacharia, Shannon Hoople, et al. 2024. STMG: A Machine Learning Microgesture Recognition System for Supporting Thumb-Based VR/AR Input. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–15.

[30] Pascal Knierim, Paweł W Woźniak, Yomna Abdelrahman, and Albrecht Schmidt. 2019. Exploring the potential of augmented reality in domestic environments. In *Proceedings of the 21st International Conference on human-computer interaction with mobile devices and services*. 1–12.

[31] Chen Liang, Chi Hsia, Chun Yu, Yukang Yan, Yuntao Wang, and Yuanchun Shi. 2023. DRG-Keyboard: Enabling Subtle Gesture Typing on the Fingertip with Dual IMU Rings. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–30.

[32] Chen Liang, Chun Yu, Yue Qin, Yuntao Wang, and Yuanchun Shi. 2021. DualRing: Enabling subtle and expressive hand interaction with dual IMU rings. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–27.

[33] Christian Loclair, Sean Gustafson, and Patrick Baudisch. 2010. PinchWatch: a wearable device for one-handed microinteractions. In *Proc. MobileHCI*, Vol. 10. Citeseer.

[34] Wayne Ma and Qianer Liu. 2024. https://www.theinformation.com/articles/apple-suspends-work-on-next-high-end-headset-focused-on-releasing-cheaper-model-in-late-2025?rc=o8rewz

[35] Meta. [n. d.]. Meta Quest Pro. https://www.meta.com/quest/quest-pro/

[36] Robert B. Miller. 1968. Response Time in Man-Computer Conversational Transactions. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I* (San Francisco, California) *(AFIPS '68 (Fall, part I))*. Association for Computing Machinery, New York, NY, USA, 267–277. https://doi.org/10.1145/1476589.1476628

[37] Anh Nguyen and Amy Banic. 2015. 3DTouch: A wearable 3D input device for 3D applications. In *2015 IEEE Virtual Reality (VR)*. 373–373. https://doi.org/10.1109/VR.2015.7223451

[38] Seo Young Oh, Boram Yoon, Hyung-il Kim, and Woontack Woo. 2020. Finger contact in gesture interaction improves time-domain input accuracy in HMD-based augmented reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[39] Eduardo GQ Palmeira, Alexandre Campos, Ígor A Moraes, Alexandre G de Siqueira, and Marcelo GG Ferreira. 2023. Quantifying the 'Gorilla Arm' Effect in a Virtual Reality Text Entry Task via Ray-Casting: A Preliminary Single-Subject Study. In *Proceedings of the 25th Symposium on Virtual and Augmented Reality*. 274–278.

[40] Patrick Parzer, Florian Perteneder, Kathrin Probst, Christian Rendl, Joanne Leong, Sarah Schuetz, Anita Vogl, Reinhard Schwoediauer, Martin Kaltenbrunner, Siegfried Bauer, and Michael Haller. 2018. RESi: A Highly Flexible, Pressure-Sensitive, Imperceptible Textile Interface Based on Resistive Yarns. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 745–756. https://doi.org/10.1145/3242587.3242664

[41] Rokid. [n. d.]. Rokid AR JOY. https://global.rokid.com/pages/rokid-ar-joy

[42] Xiyuan Shen, Chun Yu, Xutong Wang, Chen Liang, Haozhan Chen, and Yuanchun Shi. 2024. MouseRing: Always-available Touchpad Interaction with IMU Rings. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.

[43] Yilei Shi, Haimo Zhang, Kaixing Zhao, Jiashuo Cao, Mengmeng Sun, and Suranga Nanayakkara. 2020. Ready, steady, touch! sensing physical contact with a finger-mounted IMU. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–25.

[44] Roy Shilkrot, Jochen Huber, Jürgen Steimle, Suranga Nanayakkara, and Pattie Maes. 2015. Digital Digits: A Comprehensive Survey of Finger Augmentation Devices. *ACM Comput. Surv.* 48, 2, Article 30 (nov 2015), 29 pages. https://doi.org/10.1145/2828993

[45] Snapchat. [n. d.]. Spectacles by Snap Inc. - The Next Generation of Spectacles. https://www.spectacles.com/

[46] R. William Soukoreff and I. Scott MacKenzie. 2004. Towards a Standard for Pointing Device Evaluation, Perspectives on 27 Years of Fitts' Law Research in HCI. *Int. J. Hum.-Comput. Stud.* 61, 6 (dec 2004), 751–789. https://doi.org/10.1016/j.ijhcs.2004.09.001

[47] Ivan E Sutherland et al. 1965. The ultimate display. In *Proceedings of the IFIP Congress*, Vol. 2. New York, 506–508.

[48] Hsin-Ruey Tsai, Min-Chieh Hsiu, Jui-Chun Hsiao, Lee-Ting Huang, Mike Chen, and Yi-Ping Hung. 2016. TouchRing: subtle and always-available input using a multi-touch ring. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. 891–898.

[49] Hsin-Ruey Tsai, Cheng-Yuan Wu, Lee-Ting Huang, and Yi-Ping Hung. 2016. ThumbRing: private interactions using one-handed thumb motion input on finger segments. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. 791–798.

[50] Radu-Daniel Vatavu and Laura-Bianca Bilius. 2021. GestuRING: A Web-based Tool for Designing Gesture Input with Rings, Ring-Like, and Ring-Ready Devices. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 710–723. https://doi.org/10.1145/3472749.3474780

[51] Anandghan Waghmare, Youssef Ben Taleb, Ishan Chatterjee, Arjun Narendra, and Shwetak Patel. 2023. Z-Ring: Single-Point Bio-Impedance Sensing for Gesture, Touch, Object and User Recognition. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–18.

[52] Anandghan Waghmare, Roger Boldu, Eric Whitmire, and Wolf Kienzle. 2023. OptiRing: Low-Resolution Optical Sensing for Subtle Thumb-to-Index Micro-Interactions. In *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*. 1–13.

[53] Mark Weiser. 1999. The computer for the 21st century. *ACM SIGMOBILE mobile computing and communications review* 3, 3 (1999), 3–11.

[54] Michael Wessely, Ticha Sethapakdi, Carlos Castillo, Jackson C. Snowden, Ollie Hanton, Isabel P. S. Qamar, Mike Fraser, Anne Roudaut, and Stefanie Mueller. 2020. Sprayable User Interfaces: Prototyping Large-Scale Interactive Surfaces with Sensors and Displays. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376249

[55] Jacob O. Wobbrock, Kristen Shinohara, and Alex Jansen. 2011. The Effects of Task Dimensionality, Endpoint Deviation, Throughput Calculation, and Experiment Design on Pointing Measures and Models. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1639–1648. https://doi.org/10.1145/1978942.1979181

[56] Katrin Wolf. 2016. Microgestures—enabling gesture input with busy hands. *Peripheral Interaction: Challenges and Opportunities for HCI in the Periphery of Attention* (2016), 95–116.

[57] Robert Xiao, Julia Schwarz, Nick Throm, Andrew D. Wilson, and Hrvoje Benko. 2018. MRTouch: Adding Touch Input to Head-Mounted Mixed Reality. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (2018), 1653–1660. https://doi.org/10.1109/TVCG.2018.2794222

[58] XReal. [n. d.]. XReal Air. https://www.xreal.com/air/

[59] XREAL Inc. [n. d.]. XREAL Air 2 Pro. https://us.shop.xreal.com/products/xreal-air-2-pro.

[60] Chenhan Xu, Bing Zhou, Gurunandan Krishnan, and Shree Nayar. 2023. AO-Finger: Hands-free Fine-grained Finger Gesture Recognition via Acoustic-Optic Sensor Fusing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.

[61] Zheer Xu, Pui Chung Wong, Jun Gong, Te-Yen Wu, Aditya Shekhar Nittala, Xiaojun Bi, Jürgen Steimle, Hongbo Fu, Kening Zhu, and Xing-Dong Yang. 2019. Tiptext: Eyes-free text entry on a fingertip keyboard. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 883–899.

[62] Xing-Dong Yang, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. 2012. Magic finger: always-available input through finger instrumentation. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 147–156.

[63] Cheng Zhang, Anandghan Waghmare, Pranav Kundra, Yiming Pu, Scott Gilliland, Thomas Ploetz, Thad E. Starner, Omer T. Inan, and Gregory D. Abowd. 2017. FingerSound: Recognizing Unistroke Thumb Gestures Using a Ring. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 120 (sep 2017), 19 pages. https://doi.org/10.1145/3130985

[64] Cheng Zhang, Qiuyue Xue, Anandghan Waghmare, Ruichen Meng, Sumeet Jain, Yizeng Han, Xinyu Li, Kenneth Cunefare, Thomas Ploetz, Thad Starner, et al. 2018. FingerPing: Recognizing fine-grained hand poses using active acoustic on-body sensing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–10.

[65] Yang Zhang, Gierad Laput, and Chris Harrison. 2017. Electrick: Low-Cost Touch Sensing Using Electric Field Tomography. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3025453.3025842

[66] Yang Zhang, Chouchang Yang, Scott E Hudson, Chris Harrison, and Alanson Sample. 2018. Wall++ room-scale interactive and context-aware sensing. In *Proceedings of the 2018 chi conference on human factors in computing systems*. 1–15.