DualRing: Enabling Subtle and Expressive Hand Interaction with Dual IMU Rings

CHEN LIANG, CHUN YU*, YUE QIN, YUNTAO WANG, and YUANCHUN SHI, Department of Computer Science and Technology, Key Laboratory of Pervasive Computing, Ministry of Education, Tsinghua University, China

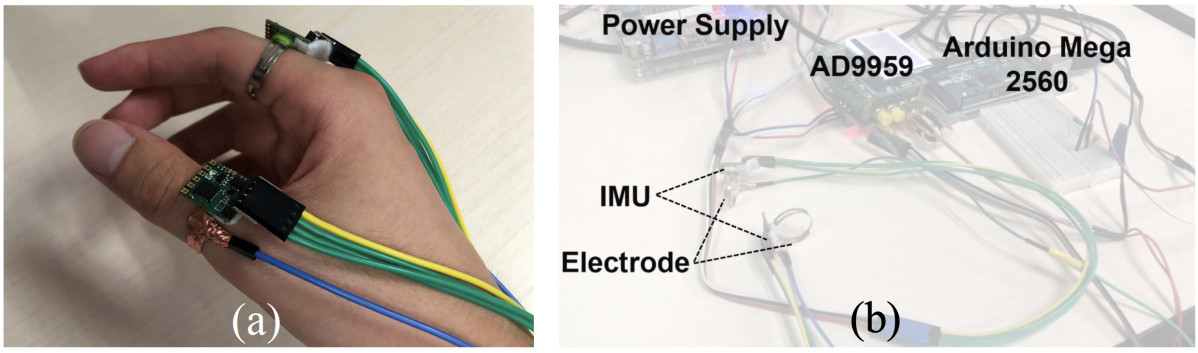


Fig. 1. The prototype of DualRing: (a) DualRing is worn on the user's thumb and index finger; (b)DualRing is composed of two IMUs and an external high-frequency AC circuit.

We present DualRing, a novel ring-form input device that can capture the state and movement of the user's hand and fingers. With two IMU rings attached to the user's thumb and index finger, DualRing can sense not only the absolute hand gesture relative to the ground but also the relative pose and movement among hand segments. To enable natural thumb-to-finger interaction, we develop a high-frequency AC circuit for on-body contact detection. Based on the sensing information of DualRing, we outline the interaction space and divide it into three sub-spaces: within-hand interaction, hand-to-surface interaction, and hand-to-object interaction. By analyzing the accuracy and performance of our system, we demonstrate the informational advantage of DualRing in sensing comprehensive hand gestures compared with single-ring-based solutions. Through the user study, we discovered the interaction space enabled by DualRing is favored by users for its usability, efficiency, and novelty.

CCS Concepts: • **Human-centered computing** → **Gestural input**; **Interaction devices**.

Additional Key Words and Phrases: ring form device, hand gesture sensing, hand interaction

*indicates the corresponding author.

Authors' address: Chen Liang, liang-c19@mails.tsinghua.edu.cn; Chun Yu, chunyu@tsinghua.edu.cn; Yue Qin, qiny19@mails.tsinghua.edu.cn; Yuntao Wang, yuntaowang@tsinghua.edu.cn; Yuanchun Shi, shiyc@tsinghua.edu.cn, Department of Computer Science and Technology, Key Laboratory of Pervasive Computing, Ministry of Education, Tsinghua University, China.

2474-9567/2021/9-ART115 $15.00
https://doi.org/10.1145/3478114

ACM Reference Format:
Chen Liang, Chun Yu, Yue Qin, Yuntao Wang, and Yuanchun Shi. 2021. DualRing: Enabling Subtle and Expressive Hand Interaction with Dual IMU Rings. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 115 (September 2021), 27 pages. https://doi.org/10.1145/3478114

1 INTRODUCTION

With the development of input technology, wearable ring-form devices are gaining popularity for their availability and portability. Equipped with different sensors (e.g., a camera [4], an infrared proximity sensor [20], and an inertial measurement sensor [27]), smart rings have the potential to capture rich hand and finger information, and thus are useful in various interaction scenarios such as VR, AR, IoT, and so on.

Compared with glove-like devices [47], nail-mounted devices [5], and touching foils attached on fingers [52, 53], ring-form devices are more lightweight and flexible and receive better social acceptance [27]. However, confined by the form factor and size, the sensing capability of a single ring is limited. It only supports a single interaction modality for the specific application (e.g., classifying a restricted set of gestures [4, 48], detecting a touching-down event[13, 27], etc.). Since there is always a trade-off between the simplicity of the form design and the sensing capability, finding a Pareto Optimality to enable rich hand input semantics while being less intrusive is of great importance.

In this paper, we propose DualRing, a ring-from input device compose of two inertial measurement units (IMUs) and a high-frequency AC circuit. With two IMU rings fixed to the user's thumb and index finger, DualRing can sense rich hand information, including the on-body contact signal, the orientation and pose of the hand, the relative movement between thumb and fingers, and the inertial features of the hand.

To detect the on-body contact signal, we build an AC circuit to measure the impedance between the thumb and the index finger by the phase delay relevant value (PDRV). To estimate the orientation and the relative movement of the thumb and the index finger, we first calculate the rotation matrix of each IMU, then calculate the relative rotation matrix. The technical difficulty is that the environmental magnetic interference is not negligible. To solve this problem, we develop a novel approximation method based on least square optimization to optimize the coordinates mapping. We also delve deep into designing statistical features in the time domain and relative features between the two IMUs to improve the accuracy of gesture recognition.

DualRing provides rich hand information to support a broad hand interaction space. We outline the interaction space and divide it into three sub-spaces: within-hand interaction, hand-to-surface interaction, and hand-to-object interaction. We propose the basic operants and interaction techniques in each subspace, demonstrating the strong sensing power of DualRing. We conduct an experiment to evaluate the performance of DualRing in recognizing different sets of gestures, as well as analyzing the salience of different features in different gesture sets. Our user study shows that users favor the usability, efficiency, and novelty provided by the interaction modalities and applications of DualRing.

To sum up, our contributions are three-fold:

- We propose and implement DualRing, a novel ring-form input device that can sense rich hand gestures and motions with a new sensing method hybridizing impedance measurement and dual IMUs.
- We outline the broad interaction space and the sub-spaces enabled by DualRing, showing its versatility for fine-grained and expressive hand gesture interactions along with applications.
- We show the informativeness and completeness of DualRing's sensing information compared with a single IMU ring and evaluate the performance of DualRing recognizing hand gestures through user studies. We also demonstrate that the interaction techniques enabled by DualRing is favored by users.

In the remainder of our paper, we first review previous work on ring-form devices, on-body sensing techniques, and hand/finger interaction. Then we illustrate the technical details and implementation of DualRing. We carry

on to outline the interaction space, dividing it into three sub-spaces, and designing the interaction techniques in each sub-space. We then conduct an experiment to analyze the recognition performance of DualRing under different gesture sets, followed by an informal user study to investigate the users' subjective preference of different gestures. Finally, we discuss important issues relative to deployment, algorithm optimization, and considerations in form determination.

2 RELATED WORK

In this section, we present related work as follows: First we summarize the previous exploration on ring-from devices, including the physical form, the sensing principle, and the design space. Then we present existing sensing techniques to detect on-body signals. Finally, we review existing methods to enable subtle hand interaction and full-hand gesture interaction.

2.1 Ring-Form Input Devices

Ring-form input devices are lightweight and portable, suitable for many interaction scenarios. They are often equipped with different sensors, such as a camera [4], an infrared proximity sensor [20], an inertial measurement sensor [27], an acoustic sensor [56], an electrical field sensor [48], and so on, for different sensing purposes in different applications. The sensing information generally includes the movement of fingers [20] and hands[4], the relative position between hand segments [40], the interaction events on the ring (such as touching and rotating) [9, 39], etc.

For example, Fukumoto and Suenaga [10] proposed the first interactive full-time wearable ring interface, the FingerRing, which senses the tapping event of five fingers and enables multi-finger typing on any surface with finger-worn accelerometers. LightRing [20] consists of an infrared proximity sensor for measuring finger flexion and a 1-axis gyroscope for measuring finger rotation, enabling 2D pointing input on any surface. ThumbRing [40] uses 2 IMUs, one connected to the user's thumb and one fixed to the user's wrist, to select discrete regions based on the relative pose. ERing [48], worn on the user's index finger, measures the difference in the electronic field to recognize different hand gestures. Ens et al. [9] proposed a hybrid method using the ring surface for touch input and a camera for finger location detection. TouchRing [39] enables subtle multi-touch input on the ring. Liu et al. [27] used an index-finger worn IMU ring to capture hand gestures and enable off-the-phone input for the visually impaired users. Coley et al. [8] investigated how the user can interact with a smart ring by touching or moving it.

However, most existing work on ring-form devices is designed task-specifically, only supporting a single interaction modality with limited channels of signal input. In our work, we aim to broaden the interaction space to support both within-hand interaction, hand-to-surface interaction, and hand-to-object interaction with a compact sensing scheme that can capture rich hand information.

2.2 On-body Sensing Techniques

The human body offers tremendous probabilities for interaction by providing novel input interfaces with self-bodied feedback. There are two main strategies for prototyping on-body interfaces.

The first strategy is to utilize the physical property of human body to transmit different types of signals [2], including acoustic signals (sound [15, 26, 51] and ultrasound [26, 31]), mechanical vibrations signals [1, 51], and electrical signals [34, 43]. Acoustic signals can either propagate above or through the human body. PUB [26] used an ultrasonic distance measurement unit to detect tapping signals on the arm and to measure the tapping location. Mujibiya et al. [31] proposed a novel sensing technique based on transdermal low-frequency ultrasound propagation to enable pressure-aware continuous touch sensing as well as various hand gestures on the human body. Mechanical vibration signals are mostly generated by contact, collision, and friction among body segments. Bitey [1] enables input via tooth clicks by a bone-conduction microphone. EarBuddy [51] uses the unmodified

microphones on commercial wireless earbuds to sense the vibration of tapping and sliding gestures on the face. Electrical signals incorporate measuring the impedance property of the human body or transmitting an AC signal or RF signal at a specific frequency. Acti-Touch [58] measures RF signal to detect body contact. Touche [34] proposed a novel Swept Frequency Capacitive Sensing technique to recognize complex body and hand gestures. BodyRC [43] sensed different body contact signals and body gestures with distributed high-frequency AC circuit.

The second strategy is to prototype the interface on or into the skin with a certain kind of skin-friendly material. Holz et al. [17] discussed the important issues for the implanted interface, including the input and output modality, the communication scheme, and the power supply. DuoSkin [18] introduced an efficient method for skin interface prototyping using gold leaf, enabling three types of interaction: input, display, and wireless communication. SkinMarks [44] proposed to install electrodes on human's body landmarks to detect the passive and active movements or contacts on those landmarks.

Among these two strategies, the former provides a more natural on-body interface and feedback with less mental and physical load, while the latter allows the user to model the interface with additional materials, enabling more interaction possibility. In our work, we adopt the first strategy to provide the most natural interaction experience to users.

2.3 Subtle Interaction within Hand

Researchers are interested in how subtle the user can perform a hand gesture. Subtle hand gestures mostly involve 1) subtle relative gestures among fingers [5, 11, 47] (such as the thumb slightly moving on the fingertip of the index finger, or a micro pinching gesture) and 2) subtle rotation or movement of the hand or certain hand segments [54].

Previous work has demonstrated the possibility of using pinching gesture for accurate input (e.g., cursor control [5], text input [47, 52, 53] and drone control [54]). FingerPad [5] enabled subtle touchpad interaction on the tip of the index finger with a nail-mounted magnetic sensor grid. DigiTouch [47] enabled subtle thumb-to-index-finger text input with a glove-based input device. TipText [53] and BiTipText [52] explored the feasibility of typing on a single-handed or two-handed fingertip QWERTY keyboard. Yau et al. [54] studied the capability of controlling a drone by hand orientation and pinching signal.

Previous work also explored the expressivity of micro hand gestures. FingerInput [37] discussed the completeness and extended the design space of thumb-to-finger micro-gestures. Pyro [11] used a thermal infrared sensor to capture near-field micro thumb-tip gestures, including drawing different patterns.

Different from previous work, where subtle thumb-to-index gestures were sensed with direct sensors either placed on the contact skin or fixed to the nails, we are the first to present a ring-form device worn on the proximal phalanx to enable subtle thumb-to-index gestures, which is of higher preference for users according to previous research [13].

2.4 Full-Hand Gesture Sensing and Recognition

Accurate hand gesture sensing is essential to enable rich and talented interaction techniques. Previous work on hand tracking and gesture sensing has researched full-hand gesture sensing solutions with different sensors, including cameras (RGB [19, 30, 55], IR [3, 46], and depth [36, 50]), EMG sensors [21, 28, 57], capacitive sensors [6], a millimeter-wave radar [12, 25, 41], magnetic field sensors [7], and electrical field sensors [48].

Among those techniques, vision-based methods were most frequently researched for their strong sensing capability to capture pixel-wise image data. A number of work has been proposed to detect the precise hand posture (e.g., the keypoints of the hand) from the raw image that captures the whole unobstructed hand in an ordinary perspective. These methods, including machine learning methods [30] and statistic-model-based methods [50] aimed to inference hand postures from the raw image by leveraging the structural constraints of

the hand [38], the local features from the image [32], and so on. The progress in both sensing algorithms and hardware leads to the success of commercial hand tracking and gesture recognition systems, including Leap Motion, Real Sense, and Kinect. Further, hand tracking becomes even harder if the camera is placed on mobile (e.g., a smartphone) or wearable (e.g., watches and glasses) devices, resulting in a restricted perspective that merely captures partial hands. For example, SkinMotion [42] and Back-Hand-Pose [49] proposed to mount the camera on a watch, looking along the back of the hand, to capture dorsum deformation and recover hand posture from the dorsum image (with/without optical markers). In other restricted scenarios for specific use, algorithms were optimzed to meet the need for certain applications. For example, HandSee [55] and Auth+Track [24] detected near-field and hand-to-phone gestures based on pre-defined features with modified front camera on smartphone. FaceSight [46] installed a down-looking IR camera on AR glasses to capture hand-to-face gestures by segmenting and recognizing hands based on face landmarks.

In addition to vision-based solutions, alternative techniques capture physiological or motional features of the user's hands and arms salient in hand gestures and motions leveraging wearable or near-hand sensors such as EMG sensors [21, 28, 57], electromagnetic sensors [7], millimeter-wave radars [12, 25, 41], IMUs [13, 23, 35], and so on. For example, changes in hand posture cause muscle deformation, thus leading to changes in the electromyogram signal on arm. Kim et al. [21] and Zhang et al. [57] recognized hand gestures with an EMG array on the arm while NeuroPose [28] proposed a neural network model to estimate 3D hand posture from the EMG signals. Electromagnetic sensing offers superior solutions to track absolute 3D position and motion of electromagnets. Finexus [7] attached electromagnetic probes to each fingertip to track the precise 3D motion of them. Soli [12, 25, 41] captured high-resolution signals with a near-hand a millimeter-wave radar and fed them to a neural network to detect both static and dynamic hand gestures.

Sensing methods based on inertial measurement units (IMUs) worth further illustration for their unique capability to capture micro vibration and subtle changes in attitude. Visual methods easily fail to detect slight, fast, and obstructed hand movements [13] due to limited resolution and motion blur, and suffer from robustness against the environment [46] (e.g., light condition). For example, when the user taps on the table with the index finger, it is hard to distinguish whether there is a contact or not [13], because the finger movement is fast and the difference between contact and non-contact is slight. Similar issues may occur to the above non-vision techniques. For example, electromagnetic tracking cannot efficiently detect contact signal (hand-to-hand or hand-to-object), while the EMG signals could be insignificant in detecting slight gestures. To this point, IMU-based complementary sensing techniques were widely explored. Gu et al. [13] combined a finger-worn IMU with a head-mounted camera to detect the touching event as well as the touching location. QwertyRing [14] enabled one finger typing with an index-finger-worn IMU ring by detecting hand-to-surface touching events and the rotation angles. Shi et al. [35] proposed an IMU-based method to detect whether the user's finger keeps contact with a static plane. Viband [23] recognized hand-to-hand and hand-to-object gestures based on high-frequency bio-acoustic signals capture by the IMU on a commodity smartwatch as well as enabling encoded vibration signal transmission via hand contact. Lu et al. [29] recognized hand-to-hand gestures with dual wrist-worn IMUs by measuring their relative attitude and detecting the synchronous vibration signals.

In our work, DualRing's uniqueness compared with existing IMU-based methods in sensing full-hand gestures are: 1) We attach IMUs on the thumb and the index finger, which are representatives of two main segments of the hand, thus can better represent a hand's internal posture irrelevant to the ground. 2) Since vibrations of the thumb and the index finger are sensed independently, fine-grained gestures involving vibrations in different hand segments can be recognized (e.g., detecting which finger is touching a surface). 3) By hybridizing impedance measurement and dual IMUs, DualRing has the potential to recognized complex gestures and their combinations with fused features.

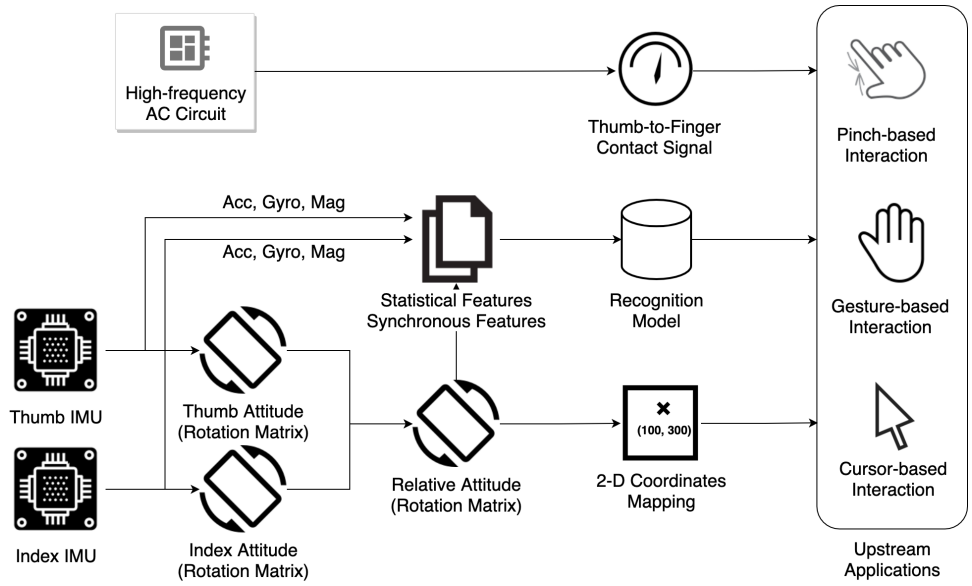


Fig. 2. System overview. DualRing is composed of two IMUs and a high-frequency AC circuit. By hybridizing the sensors, DualRing acquires high-level sensing information to support different modalities of upstream interaction.

3 DUALRING: TECHNICAL DETAILS AND IMPLEMENTATION

In this section, we introduce the sensing purposes as well as the technical details of DualRing, including the hardware setup and the sensing algorithms. We explain how to 1) achieve robust thumb-to-finger contact sensing, 2) calculate the relative attitude while dealing with the environmental magnetic interference, and 3) sense gestures and hand state in various scenarios based on synchronized IMU features.

3.1 Technical Overview

We first introduce the sensing purposes of DualRing before going into the technical details. The overall workflow is shown in Figure 2. DualRing is composed of two IMUs rings worn on the user's thumb and index finger and a high-frequency AC circuit to measure the impedance between the thumb and the index finger. The AC circuit detects thumb-to-finger contact information, including whether a thumb-to-finger touch happens and which finger the thumb is touching. The two IMUs represent the attitudes of the proximal phalanx of the thumb and the index finger independently. When combing the two attitudes, the relative attitude (represented in rotation matrix) between the thumb and the index finger can be acquired. With a parameterized approximation, the relative attitude is mapped to the 2D coordinates on the desired surface (e.g., the tip of the index finger), enabling a precise cursor mapping. Further, synchronized temporal and statistical features of the two IMUs (including the individual features and the relative features) can be extracted for gesture classification. We will illustrate the technical details below.

3.2 Thumb-to-finger Contact Detection

To enable sensitive and accurate thumb-to-finger interaction, we need a robust on-body touch detector. Since the human body is a good conductor to transmit high frequency AC signal, we build an AC circuit to detect touch events by measuring the impedance between the thumb and the index finger.

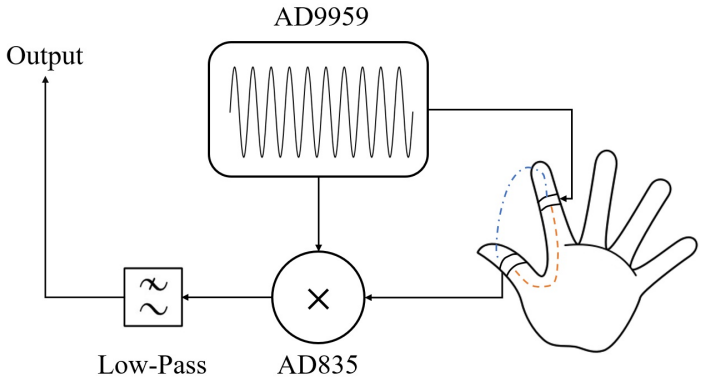


Fig. 3. Circuit Architecture. The blue dotted line and the red dotted line are two paths for high frequency AC signal transmission.

3.2.1 Sensing Principles. "Currents with a frequency high enough that wave nature must be taken into account" [33]. Previous work has proved the human body a good medium for high frequency AC signal transmission [33]. AC signals travel through different segments of the human body connected to a circuit. When a human changes their body pose, the total impedance of the body changes consequently (mainly due to changes in circuit structure and circuit component [43]), thus leading to changes in electrical properties of the whole circuit. On-body touch events can be recognized based on the signal amplitude and the phase delay because on-body touch will form another connected path in the circuit and change the circuit structure.

3.2.2 Circuit Implementation. We build an impedance measurement circuit to detect contact behavior. To measure the signal amplitude and the phase delay of the signal passing through the human body, we first created two synchronized copies of high frequency source signal s_{ref} and s_{mon}.

Assuming the reference signal is $s_{ref} = A_1 cos(\omega t + \phi_1)$, and the monitored signal passing through human body is $s_{mon} = A_2 cos(\omega t + \phi_2)$. By multiplying the reference signal and the monitored signal, we obtain a signal S composed of a high frequency component and a low frequency component.

$$S = s_{ref} * s_{mon} = A_1 A_2 cos(\omega t + \phi_1) cos(\omega t + \phi_2) \quad = \frac{A_1 A_2}{2}(cos(2\omega t + \phi_1 + \phi_2) + cos(\phi_1 - \phi_2)) \tag{1}$$

Then we pass S through a low pass filter and obtain the phase delay relevant value $PDRV = \frac{A_1 A_2}{2} cos(\phi_1 - \phi_2)$ of the reference signal and the monitored signal. By resampling the PDRV, we can obtain the peak amplitude of the monitored signal A_2 as well as the phase delay $\phi_1 - \phi_2$ of the two signals. In our work, we only use the signal amplitude to detect contact behavior. When the thumb touches different fingers, changes in the circuit component will lead to variations in signal amplitude. Therefore, the circuit cannot only detect contact behaviors but also distinguish which finger the thumb is touching.

The overall circuit architecture is shown in Figure 3.

3.2.3 Hardware Prototype. In our hardware prototype, we use an AD9959 synthesizer (4-channel, 500MHz DDS with 10-bit DACs) to generate two channels of synchronized 12.5MHz sinusoidal signal. One is used as the reference signal, while the other (the monitored signal) is transmitted into the body through two electrodes made of copper tape. As shown in Figure 1, the electrodes are connected to the proximal phalange of the thumb and the index finger, respectively. We use an AD835 4-quadrant multiplier to multiply the reference signal and the monitored signal. The output signal is amplified by 10 times with a post-op amplifier. We then connect the output

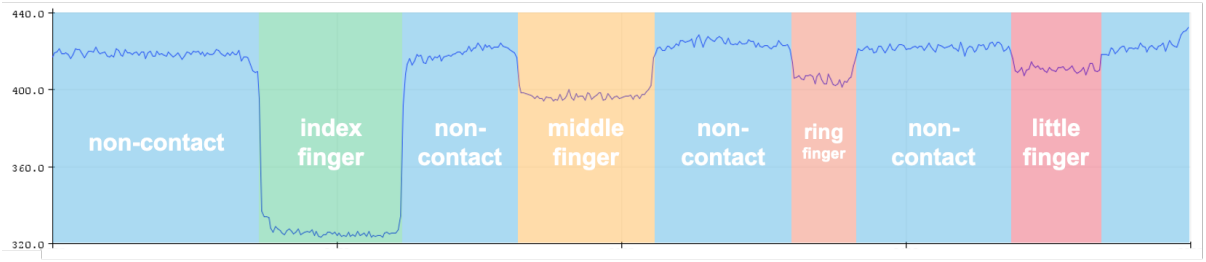


Fig. 4. A Sample of the output signal amplitude series of our high frequency AC circuit. Different signal amplitudes represent different contact states, as shown in different colors.

signal to a simple low pass Pi filter (composed of two 100 pF capacitors and a 100 μH inductor) to remove the high frequency components. Finally, we use an Arduino Mega 2560 to sample the amplitude of the filtered signal at 1 KHz, followed by a 100-frame-width sliding-window mean filter, and output the smoothed signal amplitude series. An output signal sample is shown in Figure 4. The signal is transmitted to a PC from the Arduino at 1 kHz through a USB serial port.

3.3 Relative Attitude Estimation

An inspiring motivation to use two 9-axis IMUs for hand pose sensing is that one of the IMUs provides reference coordinates, thus relative coordinates irrelevant to the ground, which better represent hand pose, can be naturally calculated from the two IMUs' data.

3.3.1 Pose Estimation. In our work, the attitude of the two IMUs is represented by Euler angles. Note that integral of the gyroscope data will result in the accumulative error in calculating the Euler angles, a Kalman filter [45] is applied to correct the roll and pitch angles (x-axis and y-axis) with the accelerameter and to correct the yaw angle (z-axis) with the geomagnetometer.

We use the direction cosine matrix [22] (DCM) to represent the attitude, or the rotation from the zero attitude (shown in Equation 2, ϕ, θ, ψ represent the roll, pitch, and yaw angle respectively, the order is yaw-pitch-roll).

$$DCM(\phi,\theta,\psi) = \begin{bmatrix} cos\theta cos\psi & sin\phi sin\theta cos\psi - cos\phi sin\psi & cos\phi sin\theta cos\psi + sin\phi sin\psi \\ cos\theta sin\psi & sin\phi sin\theta sin\psi + cos\phi cos\psi & cos\phi sin\theta sin\psi - sin\phi cos\psi \\ -sin\theta & sin\phi cos\theta & cos\phi cos\theta \end{bmatrix} \tag{2}$$

We denote the DCM of the reference IMU (e.g., the IMU on the index finger) M_r and the DCM of the dominant IMU (e.g., the IMU on the thumb) M_d, both of which share the same zero attitude. Then the relative rotation matrix from the dominant coordinates to the reference coordinates M_R can be represented as:

$$M_R = M_r^{-1} M_d \tag{3}$$

For any vector $v = (x_0, y_0, z_0)^T$ in the dominant coordinates, $M_R v = M_r^{-1} M_d v$ represents the same vector in the reference coordinates. Assuming $(e1, e2, e3) = I$ is the orthonormal base vectors in the dominant coordinates, we found that $(e1', e2', e3') = M_R$ is the orthonormal base vectors in the reference coordinates. For any vector v, it can be represented as the linear combination of $e1', e2', e3'$. The purpose is to figure out the linear factors that map M_R to the 2-dimensional orthogonal fingertip coordinates. Let (e_i, e_j) be the base vectors in the target 2-D coordinates, and v_0 be the orientation vector in the dominant coordinates. Then we can calculate the target (x, y) by:

$$(x, y)^T = (e_i, e_j)^T M_R v_0 \tag{4}$$

By sampling multiple data points, we can fit the undetermined coefficients e_i, e_j, v_0.

3.3.2 Dealing with Magnetic Interference. When the user's hand is near a strong magnetic field source (e.g., a laptop or a smartphone). the change in the magnetic field will lead to significant shift and delay in yaw angle calculation. Therefore, we modified our approximation strategy to fit the scenario, as illustrated below.

Note that (x, y) is not only a quadratic form, but also a linear combination of the coefficients of M_R with 9 DoFs. To improve the expressivity, we introduce additional parameters and reorganize the parameters into a simplified unrestricted linear form:

$$x = \sum_{1 \le i,j \le 3} p_{i,j} m_{i,j}, y = \sum_{1 \le i,j \le 3} q_{i,j} m_{i,j} \tag{5}$$

, where $p_{i,j}, q_{i,j}$ are the undetermined parameters, and $m_{i,j}$ are the elements in M_R. With such modification, the coefficient space is enlarged to 18 DoFs.

We adopt least square optimization with a regulation term to learn the parameters. Instead of directly learning the absolute coordinates (x, y), we aim to learn the difference, or the relative movement vector $(\Delta x, \Delta y)$, because relative movement is better perceived and performed by the user rather than absolute position on the fingertip. Assuming $S_h = \{(M1_i, M2_i), i = 1, 2, \cdots, k\}$ is the set of all the horizontal sampled pairs and S_v is the set of all the vertical sampled pairs. (A sampled pair $(M1_i, M2_i)$ is the relative DCM of the start point and the end point calculated from equation 3. A horizontal sample pair means the start point and the end point form a horizontal base vector $(1, 0)^T$, while a vertical sample pair is a pair representing a vertical base vector $(0, 1)^T$.) The objective loss function for optimization can be represented as:

$$\begin{aligned} L = & \sum_{(M1,M2) \in S_h} (((\sum_{1 \le i,j \le 3} p_{i,j}(m1_{i,j} - m2_{i,j})) - 1)^2 + ((\sum_{1 \le i,j \le 3} q_{i,j}(m1_{i,j} - m2_{i,j})) - 0)^2) + \\ & \sum_{(M1,M2) \in S_v} (((\sum_{1 \le i,j \le 3} p_{i,j}(m1_{i,j} - m2_{i,j})) - 0)^2 + ((\sum_{1 \le i,j \le 3} q_{i,j}(m1_{i,j} - m2_{i,j})) - 1)^2) + \\ & \lambda(\sum_{1 \le i,j \le 3} (p_{i,j}^2 + q_{i,j}^2)) \end{aligned} \tag{6}$$

, where $\lambda(\sum_{1 \le i,j \le 3} (p_{i,j}^2 + q_{i,j}^2))$ is a regulation term to prevent overfitting. By minimizing the loss function L, we acquire the optimized parameter $p_{i,j}, q_{i,j}$.

3.3.3 Prototypical Implementation. We used two MPU 9250 (9-axis IMU) for hand pose estimation. Each IMU is connected to a microcontroller on which a built-in Kalman filter is implemented. The filtered IMU data (3-axis acceleration, 3-axis gyroscope data, and 3-axis geomagnetic data for each IMU) is transmitted to a PC through UART (Universal Asynchronous Receiver/Transmitter) at 200Hz (115200 baud rate).

The algorithm pipeline on the PC was implemented in Python. We used the Numpy library for matrix calculation, and the Scipy library for parameter solving. The data receiving and the data processing processes are asynchronous. The FPS of the whole pipeline (including data receiving and data processing) is 200.

3.4 Hand Gesture Recognition based on Fused Features

DualRing provides synchronous data from the two IMUs and the AC circuit. To better utilize the multi-source data in recognizing different hand gestures, we further explore the feature design. We extract two categories of features from the raw data: the temporal statistical features and the relative and synchronous features among

IMUs. Let $(acc_1, gyr_1, ang_1, mag_1; acc_2, gyr_2, ang_2, mag_2)$ be a data frame f consisting of 3-axis acceleration, 3-axis angular velocity, 3-axis Euler angle and 3-axis geomagnetic field of the two IMUs.

3.4.1 Statistical Features in Time Domain. Temporal and statistical features are important in indicating a process with salient characteristics in time domain. We first calculate 9 statistical features: *mean, standard deviation, maximum, minimum, energy, skew, kurtosis, zero crossing rate, first-order difference*, and then use FFT to calculate the spectrum of a frame sequence $[f_1, f_2, \cdots, f_N]$ in each dimension (24 dimensions in total). The statistical features and spectrum is concatenated with the original frame sequence to form a fused feature vector.

3.4.2 Synchronous Features among IMUs. We are also interested in the mutual and relative features generated by the synchronous data from the two IMUs, including the correlation, the relative attitude and the spectrum of the differential signal.

(1) Relative attitude and rotation matrix. For each frame, the relative rotation matrix can be represented as $DCM(ang_1)^{-1} \cdot DCM(ang_2)$, and calculated from Equation 4. The rotation matrix is then flattened into a 1-dimensional feature vector.
(2) Spectrum of the differential signal. We first normalize the sensor data (acceleration, angular velocity and geomagnetic field) to the mutual ground coordinates with the attitude. We then calculate the signal difference between the two IMUs in each dimension and compute the spectrum in each dimension over a time interval.
(3) Cosine Similarity. For any two vectors v_1, v_2, their cosine similarity is $cos(v_1, v_2) = \frac{v_1 \cdot v_2}{|v_1||v_2|}$. We compute the cosine similarity of the two IMUs in time domain in each dimension.

3.4.3 Recognition Algorithm and Implementation. After obtaining the statistical features, the spectra, and the relative features, we train a random forest classifier (number of estimators = 500, max depth = 10) to recognize hand gestures. Note that because gestures from different gesture sets have diverse salience over different features, a feature selection phase is applied before determining the final model.

The feature extraction, feature selection, and model training phases were implemented in Python (Numpy library for feature extraction, Scipy and Sklearn libraries for feature selection and model training). The FPS of pre-processing a sample and output the prediction is 30 (9 ms for pre-processing and 24 ms for model prediction) running on a Intel i7-6700HQ CPU (2.6GHz, 1 thread).

4 DESIGN SPACE AND APPLICATIONS

DualRing provides four unique benefits for sensing hand pose and enhancing hand input. First, dual IMUs can naturally represent two sets of coordinates - the ground coordinates and the relative coordinates. The ground coordinates are used to represent gestures relative to the ground, while the relative coordinates represent the internal gestures of the hand irrelevant to the ground. Second, by incorporating a contact detection circuit, DualRing can support natural thumb-to-finger touch input. Third, by installing IMUs on both the thumb and the index finger, DualRing can sense both the interacting hand segments (e.g., the touching finger) as well as the non-interacting segments, thus has the potential to induce full-hand gestures. Fourth, the micro-vibration of the thumb and the index finger can be detected independently, from which more complex gestures (e.g., a gripping gesture) can be recognized. These benefits provide a rich interaction space for ring-based hand interaction. We divide the design space into: 1) within-hand interaction, 2) hand-to-surface interaction, and 3) hand-to-object interaction, as outlined in Figure 5.

4.1 Within-hand Interaction

Within-hand interaction mainly incorporates interactions between the thumb and the other four fingers. Since DualRing provides accurate orientation data of both the thumb and the index finger, as well as the contact

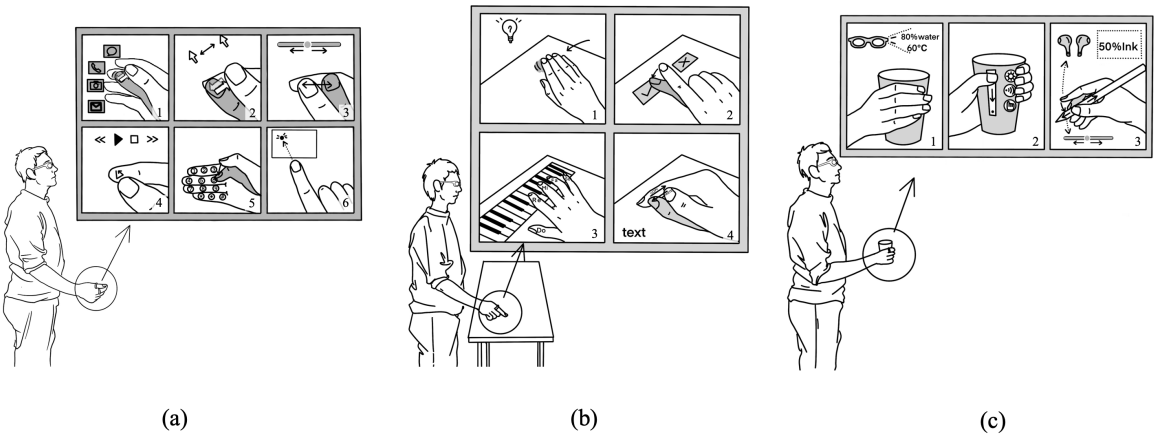


Fig. 5. Design space: (a) within-hand interaction, (b) hand-to-surface interaction, and (c) hand-to-object interaction.

signal between the thumb and other fingers, it can support various types of primitive thumb-to-finger operants, including pinching, continuous cursor control, region selection, sliding, and pattern drawing, as shown in Figure 5 (a).

4.1.1 Pinching. Pinching is the most natural within-hand gesture to perform a triggering. With DualRing, the user can perform a pinching gesture in three states: contact, non-contact and clicking. DualRing classifies the contact and the non-contact states by the sampled signal amplitude, while recognizing the clicking state by detecting a pulse in the signal spectrum. Moreover, since DualRing can saliently distinguish which finger is used for pinching by measuring signal amplitude, it enables the user to perform pinching with different fingers, which can be used to trigger different functions (Figure 5 (a) 1). In particular, DualRing can distinguish up to three types of pinching with different fingers: thumb-to-index-finger, thumb-to-middle-finger, and thumb-to- ring-finger. The amplitude variation of touching the little finger is not salient.

4.1.2 Continuous Cursor Control. DualRing detects the orientation and the movement of both the thumb and the index finger. With our parameterized approximation method, DualRing can perfectly detects instantaneous relative movement and map it to 2-dimensional coordinates, enabling accurate 2D continuous cursor control. Based on the accurate 2D cursor, DualRing can support various applications.

- Finger Pad. One of the direct and most powerful applications is a finger pad, which makes full use of the 2-D continuous movement information (Figure 5 (a) 2). The user's index finger is turned into a touchpad using their thumb to swipe on the index finger to perform a cursor input. With the support of the on-body contact detection circuit, all input modalities of an ordinary touchpad, including touch-down, clicking, and swiping, can be perfectly simulated by our fingerpad enabled by DualRing. In addition, the user can choose to enable multiple fingerpads on different fingers to control multiple sets of cursor (e.g., thumb-to-index-finger pad to control the movement and thumb-to-middle-finger pad to control the perspective in a video game).
- Finger Slider. The user can slide on the finger with the thumb to provide input to a 1-dimensional slider (Figure 5 (a) 3, e.g., controlling the brightness or volume). Moreover, the user can assign different slider functions to different fingers, so that he can control multiple sliders simultaneously.
- Pattern Drawing. Drawing a pattern is also a popular input method to convey rich semantics in a natural way. The user can map different patterns to the semantically-similar shortcuts. For example, when receiving a phone call, the user can draw a circle to answer the call, or a cross to reject it. A camera shortcut in AR

can be awakened by drawing a rectangle. The user can also draw a triangle to trigger "playing" for a music player (Figure 5 (a) 4)

4.1.3 Absolute Region Selection. Due to the magnetic interference and the delay of Kalman filtering, it is hard for DualRing to determine accurate absolute coordinates. However, based on the orientation data, DualRing can detect coarse-grained absolute coordinates, which benefits applications requiring discrete keyboard input. For example, in an dialing application, the user can put the four fingers together and touch different phalanxes of different fingers with the thumb, as dialing on a number pad (Figure 5 (a) 5). The user can also naturally type on the fingers as if typing on a T9 keyboard.

4.1.4 Ray Interaction. DualRing also enables interactions leveraging orientation and movement relative to the ground. The reference ring (e.g., the ring worn on the index finger) provides ground orientation in real world (just like a ray), while the thumb can generate discrete (e.g., touching the index finger) or parameterized (e.g., altering the thumb-index angle) input. Such interaction scheme is helpful especially in different 3D scenarios such as VR, AR, and IoT. For example, in a smart classroom, the teacher uses his index finger as a virtual pointer (Figure 5 (a) 6). The teacher can tap on the index finger with the thumb to switch the pointer mode, or change the thumb-index angle to zoom in / out. In VR or AR, ray interaction is also effective in pointing and element operation.

4.2 Hand-to-surface Interaction

When the user is interacting with a planar surface, usage of different hand gestures is a natural and effective way to express different semantics. Compared with a single index-finger-worn IMU ring, DualRing can sense the state and the movement of both the touching finger and the non-touching fingers. Therefore, it has the potential to detect the state of the whole hand and support a wider range of gesture set, as shown in Figure 5 (b).

4.2.1 On-surface Detection. Previous work has shown the feasibility to detect contact between the finger and a static surface by observing the micro-vibration with a nail-mounted accelerometer[35]. In our work, DualRing takes advantages of the two independent IMUs, which can sense the micro-vibration of both the thumb and the index finger respectively, thus has the potential to detect four types of contact: non-contact, index finger contact, thumb contact and palm contact. The enriched contact types allow for a wider range of interaction design. For example, in a sketchpad application, the user can turn a normal surface to a sketchpad – the user can use the index finger as a pen and the thumb as an eraser, while using palm contact gesture to copy and paste. This can be done by detecting a surface touch / contact signal as a trigger while mapping the independent attitude (or orientation) of the index finger and the thumb to the 2D coordinates (similar to ray interaction). The user can also use contact gestures in IoT scenarios, e.g., the lamp turning on when the user is seated and rests his palm on the desk (Figure 5 (b) 1).

4.2.2 Surface Gestures. DualRing can not only detect static hand state, but also detect dynamic gestures interacting with a surface. The most common hand-to-surface gesture is a tapping gesture. We demonstrate that tapping with five different fingers results in different features of DualRing's signal, thus is distinguishable. With such capability, DualRing enables five-finger tapping input on any surface, supporting a variety of applications such as any-surface typing (Figure 5 (b) 2), piano playing (Figure 5 (b) 3), and remote control.

4.2.3 Combination of Dominant Gestures and Controlling Gestures. DualRing provides information of both the touching finger and the non-touching fingers. We define the direct gestures interacting with the plane as dominant gestures (e.g., touching the surface with the index finger), while the gestures of the non-touching finger are noted as controlling gestures. By such definition, another intuitive exploration is the combination of the dominant gestures and the controlling gestures. The controlling gestures can be further divided into discrete gestures

(e.g., thumb tapping on index finger) and continuous gestures (e.g., thumb sliding on index finger, the angle between thumb and index finger, etc.). Such combination expands the design space of complex, multi-step, and multi-parameter gestures. For example, after touching a plane with the index finger, the user can slide backward with the thumb on the index finger to perform a "copy", while sliding forward to perform a "paste" (Figure 5 (b) 4).

4.3 Hand-to-object Interaction

Another important subspace enabled by DualRing is to interact with different objects in real-life scenarios. The gripping events and the gestures on the surface can be sensed, allowing the user to interact with ordinary objects, such as a bottle, a box, or a ball, as shown in Figure 5 (c).

4.3.1 Gripping Recognition. As an extension of detecting contact between the finger and static surfaces, we envision and demonstrate that the event of gripping an object while holding it in the air can be detected by DualRing. When the user is gripping a rigid object, the micro-vibration of the gripping fingers are likely to be synchronized. DualRing can capture the synchronous micro-vibration signals from the accelerameters to accurately recognize the gripping gestures. DualRing can not only detect whether the user is gripping an object, but also distinguish the shape of the gripped object by incorporating the orientation data. The gripping information benefits passive sensing and interaction in IoT and AR scenarios. For example, when the user is gripping and lifting an object, the information of it will be rendered in the AR glasses in real time (Figure 5 (c) 1, 3).

4.3.2 Interaction on the Object. Once DualRing detects the user is gripping an object, it turns the surface of the gripped object into an interactive interface. Basic operants such as clicking and sliding are available on the surface. Combination of the gripping object and on-surface gestures forms a large interaction space. For example, in a smart home scenario, the user can pick up a box-like object as a remote controller, swiping left / right to switch TV channels, up / down to control the volume. The user can pick up another bottle-like object, sliding up / down to control the brightness of the room(Figure 5 (c) 2).

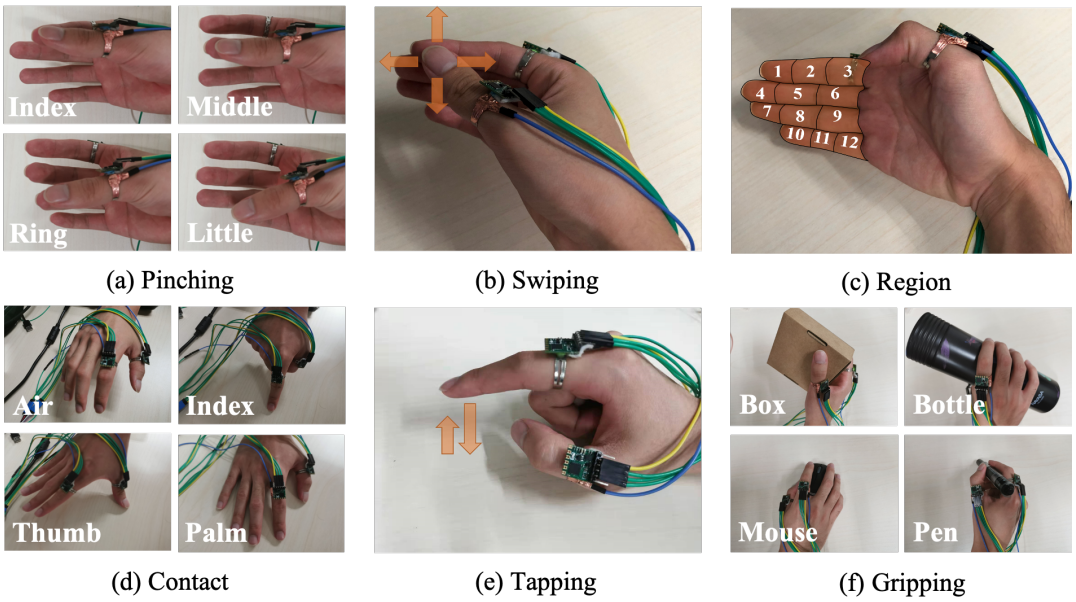
5 SYSTEM EVALUATION

In this section, we evaluate DualRing's performance in recognizing different hand gestures in the three sub design spaces. We demonstrate the informativeness and completeness of DualRing's sensing information compared with a single IMU ring.

5.1 Gesture Set

The gesture sets we wanted to evaluate includes:

- **Pinching gestures (within-hand).** The participant touches the pulp of different fingers (index finger, middle finger, ring finger, little finger) with the thumb, as shown in Figure 6 (a). 4 classes in total.
- **Swiping gestures (within-hand).** The participant swipes the thumb on the tip of the index finger in four directions (up, down, left, and right), as shown in Figure 6 (b). 4 classes in total.
- **Touching different finger segments (within-hand).** The participant touches different phalanxes (distal phalanx, middle phalanx, and proximal phalanx) of different fingers (index finger, middle finger, ring finger, little finger) with the thumb, as shown in Figure 6 (c). 12 classes in total.
- **Contact gestures (hand-to-surface).** As shown in Figure 6 (d) the participant performs four types of contact gestures (keeping contact with the plane with different hand segments): non-contact, index finger contact, thumb contact, and palm contact.
- **Tapping gestures (hand-to-surface).** The participant performs tapping gestures, as shown in Figure 6 (e), with the pulp of different fingers (non-tapping, thumb, index finger, middle finger, ring finger, and little finger). 6 classes in total.



(a) Pinching (b) Swiping (c) Region

(d) Contact (e) Tapping (f) Gripping

Fig. 6. The gesture sets for system evaluation. (a) Pinching gestures; (b) Swiping gestures; (c) Touching different finger segments; (d) Contact gestures; (e) Tapping gestures; (f) Gripping gestures.

- **Gripping gestures (hand-to-object).** The participant first performs a non-gripping gesture, then grips the following objects in order: box, bottle, mouse, and pen (shown in Figure 6 (f)). 5 classes in total.

5.2 Participants and Apparatus

We recruited 14 participants (8 males, 6 females) from the local campus. The average age was 21.0 (SD=2.04) and all participants were right-handed. We collect the data with the prototypical system described in Section 3. The data of the two IMUs and the AC circuit were fetched synchronously by a data collection thread. The sample rate of the whole system is 200 FPS and the delay between the two IMUs is always within 25 ms (5 frames).

5.3 Data Collection

At the beginning of the study, we instructed the participants to put on the rings properly, helping them to adjust the tightness so that the rings were neither too loose nor too tight. (Each ring is made with a flexible metal strip. The participants could adjust the tightness by bending the ends of the metal strip.) The data collection procedure consisted of 6 sessions. In each session, the participants were asked to collect hand gesture data corresponding to section 5.1 in order (6 gesture sets, 31 gestures in total).

We devised a graphical interactive program to guide the experiments. At the start of each session, a general text description was shown to introduce the gestures in the session. In the recording of a specific gesture, the participants were asked to perform the gesture while pressing the record key. For transient gestures (including swiping gestures and tapping gestures), the participants were required to perform the gesture while pressing the key simultaneously. To guide the procedure, a real-time video demonstration was shown on the screen, accompanied by a counter widget and a countdown widget indicating the remaining tasks and time.

For pinching gestures and region selection gestures, each participant performed each gesture 1 time, each lasting for 25 seconds. For contact gestures and gripping gestures, each participant performed each gesture 5 times, each lasting for 10 seconds. For tapping gestures and swiping gestures, each participant performed each

gesture 100 times, each taking 1 second. The whole data collection procedure for each took about 1 hour and each participant was paid 15 dollars for participation.

It is worth emphasizing that tapping and swiping are transient gestures occurring within 0.5s. We collected more trials (100) for these transient gestures to collect more diverse synchronized data. For the non-transient gestures, we collected fewer trials, but each trial lasting for a longer time (e.g, 10s). For the latter, we also required the participants to move their hand actively within a trial (e.g., rotating the hand, bending and slightly rubbing the finger) while preserving the performed gesture to enhance the data diversity. We then sampled each trial to multiple clips as data points. Therefore, we adopted a different number of trials for different types of gestures.

5.4 Results for Individual Datasets

We analyzed the recognition accuracy as well as the signal salience with respect to different gesture sets. For pinching gestures, we used a nearest neighbor algorithm for classification. For swiping gestures, we used a cosine similarity measurement for classification. For touching finger segments, contact gestures, tapping gestures, and gripping gestures, we adopted an random forest classifier as described in Section 3.4.3 and used all the features from Section 3.4 as feature input.

Table 1. Classification accuracy in different gesture set with different sensor enabled.

Gesture Set	Tapping (2 classes)	Tapping (6 classes)	Gripping	Contact
Dual IMU	97.65% (1.14)	82.71% (2.81)	82.35% (4.66)	93.24% (1.51)
Thumb IMU	95.82% (2.40)	59.47% (3.13)	63.92% (3.94)	67.47% (2.22)
Index IMU	96.87% (1.15)	66.53% (2.85)	61.81% (4.42)	62.74% (3.43)

5.4.1 Pinching Gestures. We recognized the pinching gesture only by the AC signal. Since impedance differed on different participants, we evaluated each participant individually based on their own data. For each participant, we sampled the first 10% frames (500 frames) for each gesture as the reference amplitudes. Then we classified the test samples sampled from the remaining frames based on a simple nearest neighbor algorithm. We randomly sampled (the start frame of each sample is randomly picked) 100 test samples (each sample is composed of 50 frames) from the remaining 90% frames (4500 frames) for each gesture. We classified each test sample by averaging the sample amplitudes and calculating the distance from each reference amplitude set of the same participant. The accuracy of the 4-class classification (non-pinching, index finger, middle finger, and ring finger) was 97.36% (SD=2.83), while the accuracy of the 5-class classification (the four classes above + little finger) was 91.11% (SD=4.02). The signal amplitude variation of touching the index finger, the middle finger, the ring finger, and the little finger are 27.39% (SD=5.60), 8.23% (SD=2.52), 3.97% (SD=1.88), 2.57% (SD=1.02), respectively, referenced to the non-pinching signal amplitude. The amplitude variation of the noise signal is 2.19% (SD=1.13).

5.4.2 Swiping Gestures. We used the AC signal to detect on-body contact and the two IMUs' data to estimate swiping direction. We evaluated each participant individually based on their calibration parameters. We first applied the parameter estimation method described in Section 3.2 to map the Euler angle to the 2-D coordinates using each participant's own calibration data. For any swiping vector v, we first computed the inner product with the participant's orthonormal base vectors v_1 and v_2. Then we classified the vector based on the cosine similarity. The accuracy of the classification is 99.37%. The average angular errors in the four directions and their deviations are shown in Table 2.

5.4.3 Touching Different Finger Segments. As we observed, due to structural differences (e.g., the length of different hand segments) and behavioral differences (e.g., some participants preferring stretching the thumb while the others preferring bending the fingers to perform a touch) among participants' hands, fusing the data

Table 2. The angular errors and the corresponding standard deviations of the swiping gesture in four directions.

Swiping Direction	Right	Left	Down	Up
Error	8.13°	10.55°	13.10°	12.32°
Standard Deviation	4.89°	4.52°	5.16°	5.23°

from different participants in the training phase would have a bad effect on prediction. For this consideration, we adopted a within-participant validation process to evaluate the classification accuracy. For each participant, we uniformly sampled 100 × 12 (classes) = 1200 samples (each sample containing 10 frames). We trained the random forest model on the temporally foremost 80% data from each gesture (960 samples) and tested it on the remaining 20% data (120 samples) for individual participants. We got a total accuracy of 89.57% (SD=5.01) without incorporating the AC signal, and an accuracy of 93.70% (SD=5.26) by incorporating the AC signal.

5.4.4 Contact Gestures. We applied a leave-one-user-out cross-validation process to evaluate the performance of contact gesture classification. We only used the IMU data for classification, since the variation of the AC signal is not significant. For each participant, we uniformly sampled 100 samples (each sample containing 80 frames) from each class (5 classes in total). The classification accuracy with respect to usage of different sensors is shown on Table 1. We saw a significant accuracy gap between using dual IMUs and dropping one of the IMUs (93.24% v.s. 67.47% / 62.74% at 80 frames). Compared with single IMU, DualRing had the potential to correctly recognize the contact gestures of the whole hand, though using the thumb-worn or index-finger-worn IMU can recognize the contact of specific finger to some extent.

5.4.5 Tapping Gestures. For tapping gesture recognition, we evaluated the accuracy of recognizing whether a tapping was performed (2 classes) as well as the accuracy of further distinguishing which finger is tapping (6 classes) with leave-one-user-out cross-validation. Only the IMU data is used in this session. For each participant, we preprocessed the data from each trial from each tapping class (100 trials × 5 classes in total). We first figured out the maximum value of the acceleration (the index of which is p) as the exact tapping point, then sampled the frames in the range of $(p - 20, p + 20]$ as the sampled data to ensure the alignment of transient gestures. We randomly sampled 500 samples (each sample contains 40 frames) from the non-tapping class as well as the peripheral frames from the tapping classes for 2-class evaluation and sampled 100 samples for 6-class evaluation to ensure model balance. Table 1 shows the accuracy of tapping classification and finger classification with respect to usage of different sensors. For the tapping classification, there was no significant improvement using two IMUs compared with using single IMU, because one IMU is enough to capture the vibration signal from a tapping attempt. However, DualRing significantly outperformed single IMU in classifying which finger is tapping (82.71% v.s. 59.47% / 66.53% at 40 frames). From the result, we found that although DualRing only attached IMUs to the thumb and the index finger, it had the potential to infer the state and movement of the other fingers, yielding a good classification result.

5.4.6 Gripping Gestures. We evaluate the recognition of gripping gestures with leave-one-out cross-validation. Only the IMU data is used in this session. For each participant, we uniformly sampled 100 samples (each sample contains 200 frames) from each class (5 classes in total). The recognition result is shown in Table 1. From the result, We observed significant accuracy improvement of using dual IMUs compared to using single IMU (82.35% v.s. 63.92% / 61.81% at 200 frames), demonstrating the informativeness of the data from dual IMUs in sensing gripping gestures.

5.5 Combination of Different Gesture Sets

As outlined in the interaction space, gestures from different gesture sets can be combined organically to represent more complex semantics (e.g., a gripping gesture followed by a tapping gesture). Though such a process can be

Table 3. Ablation Study: Three gesture sets were tested in our ablation study. For each gesture set, we removed specific features from the input vector, and measure the change in classification accuracy.

Gesture Set	- Raw	- Stats	- Spectrum	- Diff Spectrum	- Cosine Sim	- Orientation
Tapping Gestures (40 frames)	+1.50%	−0.26%	−4.39%	−1.13%	−1.51%	−0.73
Contact Gestures (80 frames)	−0.45%	−2.78%	−4.67%	−2.03%	−1.38%	−0.31%
Gripping Gestures (200 frames)	−2.16%	−10.42%	−4.42%	−7.98%	−5.76%	−6.10%

implemented in a step-wise manner, meaning at each step the model only recognizes gestures from individual set, we are also interested in how the model performs when combining gestures from different sets.

We conducted an experiment to explore DualRing's robustness across gesture sets. We merged the "contact gestures" (hand-to-surface interaction) and the "gripping gestures" (hand-to-object interaction), obtaining a combined dataset of 8 classes (the "in-air" class is mutual). For each participant, we uniformly sampled 100 samples (each sample contains 100 frames) from each class (8 classes in total). Results showed that leave-one-user-out cross-validation (100 frames, dual ring) yield an average accuracy of 79.72% (SD=5.46%), demonstrating our algorithm's robustness across different gesture sets.

We conducted another experiment to explore whether a sample from the tapping gesture set (transient) or the gripping gesture set (non-transient) would lead to a false positive issue. We found samples from the gripping set have a false positive rate of 2.75% (40 frames, dual ring) in the tapping set (e.g., recognized as tapping); while samples from the tapping set have a false positive rate of 7.40% (100 frames, dual ring) in the gripping set. The result showed that a transient gesture set and a non-transient gesture set do not have much interference on each other, which validates the feasibility of fusing transient gestures and non-transient gestures in recognition.

5.6 Feature Analysis

We conducted a brief ablation study to analyze the salience of different features in different gesture sets. The result is shown in Table 3. From the results, we concluded that the salience of different features vary in different gesture sets. For tapping gestures, we observed an increase in accuracy when removing the raw data, probably because the raw data contained interferential features. Removal of the independent spectra of the two IMUs caused the most significant decline in tapping gestures. For contact gestures, spectra of the two IMUs is most significant, probably because the spectrum of each IMU indicates the contact state of the thumb and the index finger respectively. For gripping gestures, statistical features, spectrum of the differential signal, and orientations had the most significant effect on performance. Spectrum of the differential signal may indicate the synchronization in micro-vibration between the two IMUs, while the orientation may represent the gripping gesture.

6 USER STUDY

In this section, we conducted a user study to understand users' subjective preference and comments on different hand gestures in our design space (Section 4).

6.1 Participants, Design, and Procedures

The participants of this study were the same as Section 5.2. We chose the same user group as the previous study to ensure that all participants understood the gesture set well.

We selected 8 representative gestures covering the three sub-spaces of our design space: pinching, finger cursor, finger sliding, region selection, finger pattern, ray pointing, tapping on a surface, gripping an object. We developed an interactive program for each gesture, either a real-time visual feedback demo (pinching, ray pointing, tapping on a surface, and gripping an object) or a simple demonstrative application (finger cursor, finger sliding, region selection, and finger pattern), as shown in Figure 7, to help participants better understand each

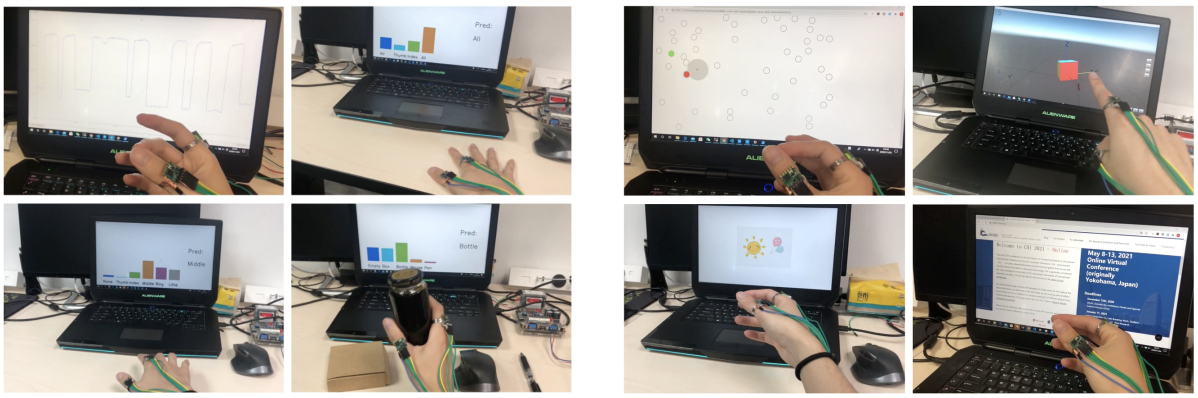


Fig. 7. Interactive programs for the user study. Left: Real-time visual feedback demos for 1) signal amplitude of pinching and 2) probability distribution of contact gestures, tapping gestures, and gripping gestures. Right: Demonstrative applications for 1) finger cursor, 2) ray pointing, 3) finger pattern, and 4) finger sliding.

gesture. We designed a questionnaire to evaluate each gesture, which contained questions of NASA TLX [16] and four additional questions about their willingness to use and whether these gestures are easy-to-use, fun-to-use, and convenient-to-use.

After introducing each gesture to the participants, including how to perform the gesture and the probable functional mappings of the gesture, we instructed them to perform each gesture correctly and used the gesture to interact with the corresponding demo. After performing all the gestures and experiencing the interactive demos, we asked each participant to fill in a questionnaire to rate each gesture in each aspect with a 7-point Likert score. At last, the participants provided subjective comments of all these gestures through an interview.

6.2 Results

Figure 8 shows the subjective ratings of different gestures in the ten dimensions. Higher score indicates lower mental demand, lower physical demand, lower temporal demand, higher efficiency, lower effort, lower frustration, easier to use, more fun to use, more convenient to use, and more willing to use. Results shows that participants generally enjoyed the 8 interaction techniques enabled by DualRing, with an average willing-to-use rating of 5.0. Pinching (5.6), finger cursor (4.8), finger sliding (5.9), ray pointing (5.4), surface tapping (5.6), and object gripping (5.2) were better received by the participants while region selection (3.6) and pattern drawing (3.8) were less favored. Among all gestures, pinching, finger sliding and surface tapping were most favored by participants for their low mental, physical, and temporal demand, their high efficiency, and the easiness and convenience for usage. Finger cursor was understood as a hard-to-use technique that requires high learning cost, extra concentration, and massive mental demand, though participants generally believed it could be an efficient, convenient, and fun gesture for experienced users. Most participants had an inferior rating on region selection and pattern drawing because the two gestures are regarded hard-to-accomplish, time-demanding and mentally-demanding.

We summarized the interview results and user's comments based on their understandings regarding different interaction spaces and gestures.

6.2.1 Thumb-to-Fingertip Gestures. All participants had a common agreement that pinching, finger cursor, finger sliding, and finger pattern were of the same type of gestures which only involve subtle finger movement, thus having low physical demand. However, the complexity of these gestures varies, leading to different degree of

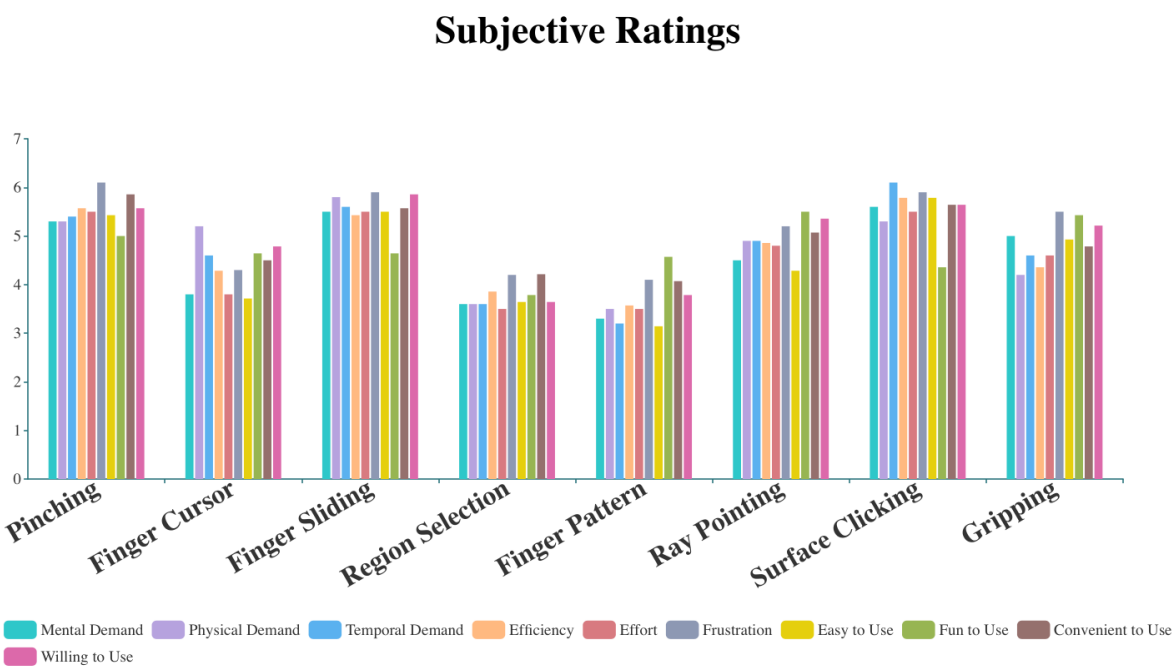


Fig. 8. Subjective ratings of different gestures in different aspects. Higher score indicates lower mental demand, lower physical demand, lower temporal demand, higher efficiency, lower effort, lower frustration, easier to use, more fun to use, more convenient to use, and more willing to use.

acceptance. Pinching and finger sliding are the easiest gestures requiring almost no learning cost or concentration, while having straightforward symbolic meanings (P5). So they were most favored. Finger cursor was regarded as easy to understand (P1, P2, P10-P12), hard to perfect (P2, P5, P12), and requiring extra concentration(P5, P12). *"Mastering finger cursor gives me a sense of achievement, especially when I am able to control the cursor precisely with subtle finger movement after becoming an expert,"* said P1. Finger pattern was generally regarded not as easily performed as pinching and sliding. But some participants loved it for the rich semantics that a symbolic gesture can convey. *"Assigning an operation with a similar symbolic pattern is easy to memorize. For example, drawing a right-oriented triangle to trigger play and a rectangle to trigger stop, which is identical with the graphic interface, is useful,"* commented P2.

6.2.2 Finger Region. Participants understood it as a gesture where the thumb touches different finger segments. It received lower acceptance mainly due to the difficulty to perform. P7 said *"I have a short thumb, so it is a hard for me to touch the proximal phalanx of the little finger."* The other 4 participants shared a similar problem with P7. However, they all agreed that if the gesture were constrained to easy-to-reach finger segments, such as the segments of the index finger and the middle finger, the gesture could become useful.

6.2.3 Hand-to-Surface Gestures. Participants were fascinated by the capability of DualRing to distinguish which finger touched the surface. P2 said, *"Tapping on the table is natural and easy to perform. I can't wait to imagine the applications leveraging this gesture, like playing the piano or typing on a ordinary desktop. It is a great pleasure."* Regarding the contact gestures, participants (P2, P4) thought they were closer to unconscious gestures and more suitable for passive control (e.g., deactivating the AR display when a rested hand is detected).

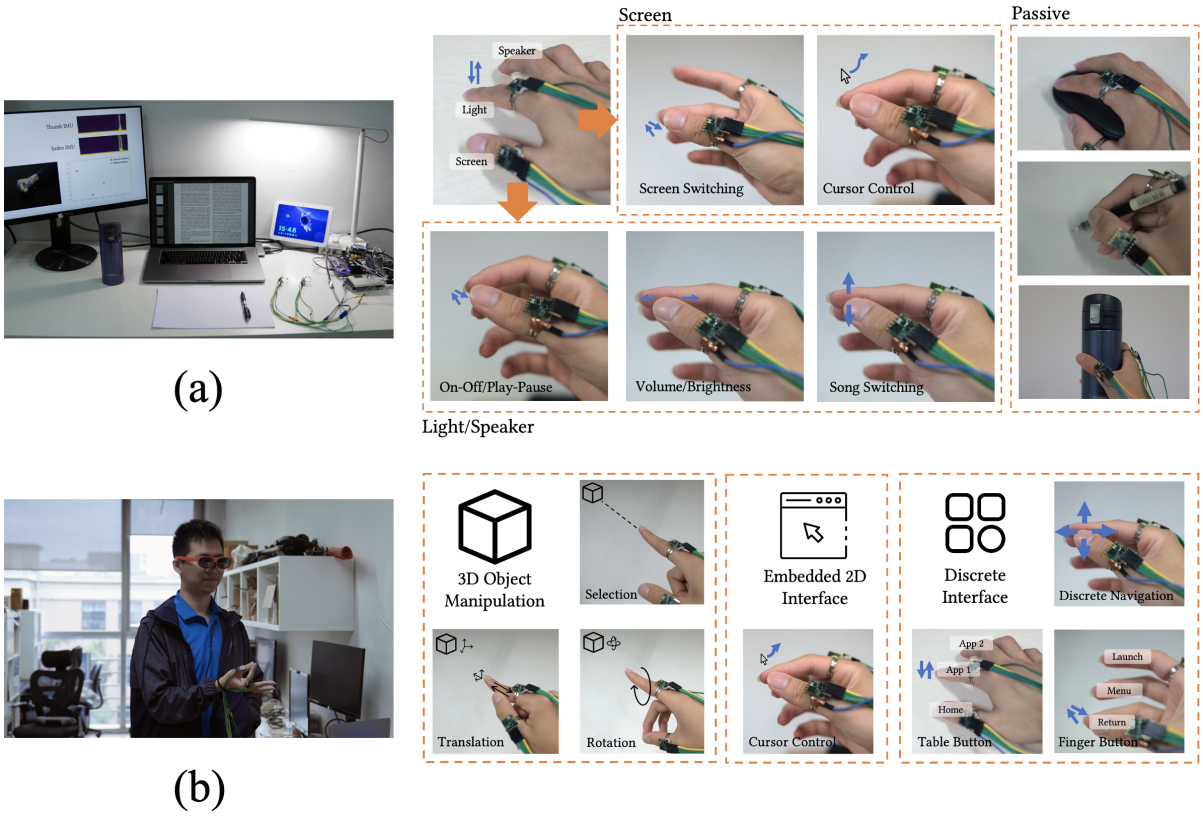


Fig. 9. Example application scenarios: (a) smart desk and (b) AR input.

6.2.4 Hand-to-Object Gestures. Participants described it as *"an effective way to enhance daily activities"*. P1 said, *"Gripping is quite a common and frequently-performed gesture. I would like the AR glasses to render information about an object once it is gripped, like the clock rendering detailed weather information when gripped. With such an intelligent grip, many things can be done as a grip."*

Generally, all gestures were well accepted by a large portion of participants. They remained curious about the applications supported by these gestures.

7 APPLICATION SCENARIOS

To illustrate the applicability of DualRing, we demonstrate two example scenarios involving different modalities from the design space, as shown in Figure 9.

7.1 Scenario Description

7.1.1 Smart Desk. The scenario setup, interaction modalities, and gestures are shown in Figure 9 (a). Alice is working in the office, sitting at a smart desk in which multiple IoT devices are controlled by DualRing in a unified manner. When she is seated with her palm resting on the desk for 5 seconds, the smart desk system automatically turns on, illuminating the main screen and enabling DualRing's gesture control.

During working, Alice follows a unified logic to control desk elements by performing different hand gestures. When she wants to start a control session, she taps on the tabletop with different fingers to navigate to the

corresponding element – thumb for screen control, index finger for light control, and middle finger for speaker control. Then she can perform within-hand gestures, such as pinching and swiping, to control the target device. When she wants to browse contents among screens, after entering the screen control mode, she turns her index finger into a touchpad, pinching with the thumb to simulate a click and swiping the thumb on the index finger to perform fine-grained cursor control. She can also perform a thumb-to-middle-finger pinch to switch the focusing screen or device. For light control, she performs a thumb-to-index-finger pinch to turn on (off) the light and slides on the index finger with the thumb to control the brightness. For speaker control, she pinches to start or pause playing music, slides on the index finger to control the volume, and swipes upward (downward) to switch to the previous (next) song.

In addition to active device control, Alice can receive passive health-related notifications from the smart desk taking advantage of DualRing's sensing capability to capture her hand status and activity. For example, a notification reminding her to drink water would be displayed on the screen accompanied by a message alarm when she has not drunk water (or a persistent gripping-bottle gesture has not been detected) for an hour. (The system is continuously detecting the gripping gestures described in section 5.1 after registration). Similarly, when she has been gripping the pen or the mouse for an hour, the system would send a notification reminding her to take a break.

DualRing's provides two unique benefits in the smart desk scenario. First, the unified control flow enabled by DualRing allows the user to control multiple devices seamlessly and efficiently using unified within-hand and hand-to-surface gestures without accessing their physical interfaces. Second, DualRing's passive activity monitoring capability plays an important role in daily healthcare, especially for the circumstances where people spend the entire day working at the desk.

7.1.2 AR Input. The scenario setup, interaction modalities, and gestures are shown in Figure 9 (b). Bob is controlling an AR interface with DualRing as a general input device which supports three types of control logic - 3D object interaction, embedded 2D UI interaction, and discrete interaction. In the 3D object manipulation APP, he first performs ray pointing to select a 3D object, then performs either a thumb-to-index pinch to move the object or a thumb-to-middle pinch to rotate the object. After a thumb-to-index pinch, he can use the ray to re-position the object, or slide the thumb along the index finger to move the object far or near. If he performs a thumb-to-middle pinch, he can use the ray to rotate the object. He releases the fingers when finishing manipulation. In the web browser APP, which is an embedded 2D interface in the 3D scene, he turns his index finger into a touchpad as described above to control the interface. He can also use discrete gestures to control the APPs which have a discrete layout (such as a grid or a list), including the home APP and the video player APP. In the home APP, he performs swiping gestures (in four directions) to navigate to the desired icon and perform a thumb-to-index pinch to launch the APP or a thumb-to-middle pinch to open the option menu. In the video player APP, he performs a pinch for start (pause), a right (left) swipe for fast forward (rewind), a upward (downward) swipe for the previous (next) video, and a continuous swipe after pinch for voice control. He performs a thumb-to-ring pinch to return to the previous page or the home. In addition, if he is around a horizontal plane (e.g., a desk), he can perform a tapping gesture for fast navigation (e.g., a thumb tap to the home APP).

Compared with existing AR input methods such as controller (e.g., NReal) and hand gesture (e.g., HoloLens), DualRing has three unique benefits for AR input. First, DualRing has the potential to support complex interactions, such as 3D object manipulation and 3D scene modeling which involves multiple sets of coordinates, in a unified and effective manner. Second, embedded UI interaction (e.g., browsing, chatting, and entering the password) with DualRing can be accomplished with secret and subtle gestures, which benefits social acceptance and privacy in public areas. Third, DualRing enables discrete interaction with low attention levels and mental load. For example, in the commuting scenario, simple gestures in the pocket help to accomplish common functions (e.g., switching

the music and reading notifications). Benefiting from a compact form and a strong sensing capability, DualRing has the potential to become an always-available general AR input device supporting multiple types of AR input.

7.2 Issues for Deployment

We discuss how the performance of individual recognition models and methods for different gesture sets and their combination (section 5) affects the usability of DualRing in the above scenarios. First, we observed that gestures based on original AC signal (pinching), individual attitude (ray pointing), and relative attitude (finger cursor and finger swiping), which is detected by determined algorithms instead of machine learning models, generally have satisfying performance for practical use (e.g., an accuracy of 97.36% for 4-class pinching, an accuracy of 99.37% for swiping). The average angular error for swiping was always within 15°, showing good usability in both gesture-based interaction (e.g., swiping in four directions) and cursor-based interaction (e.g., drawing a vertical line). For more complex gestures leveraging fused features, such as contact, tapping, and gripping, they could be deployed with certain design considerations for usability in our application scenarios. For example, the size of the tapping gesture set for the smart desk could be shrunk, avoiding assigning functions to gestures with confusion (e.g., ring finger and little finger). Further, recognition in our applications was designed sequentially (e.g, a tapping gesture followed by a swiping gesture in the smart desk scenario) and combined organically to avoid fusing different gesture sets for better detection accuracy. A label smoothing strategy could also be adopted on the gripping detection to avoid non-stable detection error. Moreover, our cross-set evaluation (section 5.5) demonstrated the feasibility to recognize different gesture sets simultaneously, enabling our smart desk to recognize tapping gestures for active control and gripping gesture for passive notification simultaneously. The usability can be further improved by 1) an optimization of the detection method (e.g., collecting more training data or using a better recognition model) and 2) a fine-grained interaction design.

8 LIMITATIONS AND DISCUSSION

In this section, we discuss the limitations as well as the potential issues related to the practical adoption and deployment of DualRing.

8.1 Form Factor

Currently, the on-body touch signal detection in our prototypical implementation is based on an external high frequency AC circuit, which lacks portability. Replacing the AC modules with an RF signal transmitter and a receiver could simplify the hardware structure and eliminate the wires while preserving the same sensing principle. In our work, the two electrodes are distributed separately on the thumb and the index finger, transmitting a 12.5 MHz signal into human body. The electrode material and placement, as well as the signal frequency, are topics worth further exploration. The two IMUs are currently wired to a PC, which limits the application scenarios. Using a bluetooth module for data transmission can relieve the wire constraint.

The ultimate form of DualRing we envision is that each ring is functionally symmetric, composed of a wireless RF signal transmitter, a receiver, and a bluetooth IMU module. The user can combine multiple rings, wearing them on different fingers to achieve different sensing purposes.

8.2 Interaction Space Design

Currently in our study, all gestures were designed by the authors through a brainstorming interview, mainly aiming to explore the design spaces and possibilities. From the user study, we found that some of the gestures and the corresponding interaction techniques needed detailed design and further improvement for practical use. For example, the touching region division could be more compact (e.g., distinguishing touching the edge or the center of a phalanx) while avoiding to cover hard-to-reach segments. Adapting dynamic cursor speed or bubble

cursor may also help to relieve participants as well as improving the efficiency. With better design, interaction techniques enabled by DualRing may bring better user experiences. In the future, we would recruit more users of different characteristics (e.g., age, culture, etc.) and conduct a study to evaluate and design each gesture and the corresponding interaction modality carefully, with special attention towards the balance of complexity and efficiency.

8.3 Dealing with Environmental Interference

Our current algorithm to detect on-body touch signal is based on a simple threshold method that only monitors the signal amplitude. However, the AC signal amplitude is affected by the user's physical state and various environmental factors such as air humidity, skin humidity, and so on, meaning that careful calibration is required before each use session. It is worthwhile to develop an algorithm to distinguish whether a change in amplitude is caused by an on-body touch event based on sequential information or information from other sensors, so that touch signal can be robustly detected.

Regarding the dual IMUs, as we observed, the magnetic data of them is easily affected by the surrounding magnetic interference. leading to bias and delay in Euler angle calculation. As a result, accurate estimation of relative pose is difficult. To solve this problem, a solution is to use signals that are not easily interfered (e.g., GPS) as reference signals.

8.4 Why DualRing?

Another important question to discuss is why we chose to use two IMUs and fixing them to the thumb and the index finger for gesture sensing. First, relative movement among hand segments, including the deliberately-performed gestures and the unconscious micro-vibration, can only be sensed with at least two sensors attached to different segments. Second, the thumb and the index finger are representatives of two main segments of the hand. When performing a common hand gesture, the index finger often drives the movement of the other three fingers. For example, when gripping an object, a person would unconsciously grip the object with the thumb and the index finger, while assisting with the other fingers. Based on these considerations, we believe the sensor installation of DualRing can provide comprehensive information for hand gesture estimation, which is also proven in our experiments.

9 CONCLUSION

We present DualRing, a novel ring-form device to sense rich hand information and enable a large interaction space. DualRing is composed of two IMU rings and a high frequency AC circuit. The two rings are worn on the user's thumb and index finger to sense the orientation and the movement of hand segments. The AC circuit is designed to detect on-body contact signal by measuring the impedance between the thumb and the index finger. We devise an algorithm to calculate the absolute and relative orientation of the thumb and the index finger, the thumb-to-finger contact signal, and various features among the two IMUs. Based on the rich sensing information, we enlarge the design space of ring-based hand interaction, and redivided it into three sub-spaces. By conducting an experiment to evaluate DualRing's performance in different gesture sets, we demonstrate the comprehensive sensing capability of DualRing compared with single-IMU-based solutions. Through a user study, we demonstrate that the interaction spaces and techniques enabled by DualRing are generally well accepted by users. We believe DualRing will inspire the exploration of ring-based hand gesture sensing technology to enable more expressive and more efficient hand interaction.

ACKNOWLEDGMENTS

This work is supported by the Natural Science Foundation of China under Grant No. 61521002, and National Key R&D Program of China No. 2019AAA0105200, and also by Beijing Key Lab of Networked Multimedia, the Institute for Guo Qiang, Tsinghua University, Institute for Artificial Intelligence, Tsinghua University (THUAI), and Beijing Academy of Artificial Intelligence (BAAI).

REFERENCES

[1] Daniel Ashbrook, Carlos Tejada, Dhwanit Mehta, Anthony Jiminez, Goudam Muralitharam, Sangeeta Gajendra, and Ross Tallents. 2016. Bitey: An Exploration of Tooth Click Gestures for Hands-Free User Interface Control. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 158–169. https://doi.org/10.1145/2935334.2935389

[2] Joanna Bergström and Kasper Hornbæk. 2019. Human–Computer Interaction on the Skin. *ACM Comput. Surv.* 52, 4, Article 77 (Aug. 2019), 14 pages. https://doi.org/10.1145/3332166

[3] Alex Butler, Shahram Izadi, and Steve Hodges. 2008. SideSight: Multi-"Touch" Interaction Around Small Devices. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology* (Monterey, CA, USA) *(UIST '08)*. ACM, New York, NY, USA, 201–204. https://doi.org/10.1145/1449715.1449746

[4] 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. https://doi.org/10.1145/2807442.2807450

[5] Liwei Chan, Rong-Hao Liang, Ming-Chang Tsai, Kai-Yin Cheng, Chao-Huai Su, Mike Y. Chen, Wen-Huang Cheng, and Bing-Yu Chen. 2013. FingerPad: Private and Subtle Interaction Using Fingertips. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 255–260. https://doi.org/10.1145/2501988.2502016

[6] Wook Chang, Kee Eung Kim, and Hyunjeong Lee. 2006. Recognition of Grip-Patterns by Using Capacitive Touch Sensors. 4 (2006), 2936–2941.

[7] Ke-Yu Chen, Shwetak N. Patel, and Sean Keller. 2016. Finexus: Tracking Precise Motions of Multiple Fingertips Using Magnetic Sensing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1504–1514. https://doi.org/10.1145/2858036.2858125

[8] Ashley Colley, Virve Inget, Inka Rantala, and Jonna Häkkilä. 2017. Investigating Interaction with a Ring Form Factor. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia* (Stuttgart, Germany) *(MUM '17)*. Association for Computing Machinery, New York, NY, USA, 107–111. https://doi.org/10.1145/3152832.3152870

[9] Barrett Ens, Ahmad Byagowi, Teng Han, Juan David Hincapié-Ramos, and Pourang Irani. 2016. Combining Ring Input with Hand Tracking for Precise, Natural Interaction with Spatial Analytic Interfaces. In *Proceedings of the 2016 Symposium on Spatial User Interaction* (Tokyo, Japan) *(SUI '16)*. Association for Computing Machinery, New York, NY, USA, 99–102. https://doi.org/10.1145/2983310.2985757

[10] Masaaki Fukumoto and Yasuhito Suenaga. 1994. "FingeRing": A Full-Time Wearable Interface. In *Conference Companion on Human Factors in Computing Systems* (Boston, Massachusetts, USA) *(CHI '94)*. Association for Computing Machinery, New York, NY, USA, 81–82. https://doi.org/10.1145/259963.260056

[11] 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* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 553–563. https://doi.org/10.1145/3126594.3126615

[12] Changzhan Gu and Jaime Lien. 2017. A Two-Tone Radar Sensor for Concurrent Detection of Absolute Distance and Relative Movement for Gesture Sensing. *IEEE Sensors Letters* 1, 3 (2017), 1–4. https://doi.org/10.1109/LSENS.2017.2696520

[13] 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* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1059–1070. https://doi.org/10.1145/3332165.3347947

[14] Yizheng Gu, Chun Yu, Zhipeng Li, Zhaoheng Li, Xiaoying Wei, and Yuanchun Shi. 2020. QwertyRing: Text Entry on Physical Surfaces Using a Ring. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 128 (Dec. 2020), 29 pages. https://doi.org/10.1145/3432204

[15] Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: Appropriating the Body as an Input Surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 453–462. https://doi.org/10.1145/1753326.1753394

[16] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[17] Christian Holz, Tovi Grossman, George Fitzmaurice, and Anne Agur. 2012. Implanted User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 503–512. https://doi.org/10.1145/2207676.2207745

[18] Hsin-Liu (Cindy) Kao, Christian Holz, Asta Roseway, Andres Calvo, and Chris Schmandt. 2016. DuoSkin: Rapidly Prototyping on-Skin User Interfaces Using Skin-Friendly Materials. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers* (Heidelberg, Germany) *(ISWC '16)*. Association for Computing Machinery, New York, NY, USA, 16–23. https://doi.org/10.1145/2971763.2971777

[19] Naoaki Kashiwagi, Yuta Sugiura, Natsuki Miyata, Mitsunori Tada, Maki Sugimoto, and Hideo Saito. 2017. Measuring Grasp Posture Using an Embedded Camera. In *Applications of Computer Vision Workshops*. 42–47.

[20] 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* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 157–160. https://doi.org/10.1145/2642918.2647376

[21] Jonghwa Kim, Stephan Mastnik, and Elisabeth André. 2008. EMG-Based Hand Gesture Recognition for Realtime Biosignal Interfacing. In *Proceedings of the 13th International Conference on Intelligent User Interfaces* (Gran Canaria, Spain) *(IUI '08)*. Association for Computing Machinery, New York, NY, USA, 30–39. https://doi.org/10.1145/1378773.1378778

[22] Roman V. Krems. 2018. *Direction Cosine Matrix*.

[23] Gierad Laput, Robert Xiao, and Chris Harrison. 2016. ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 321–333. https://doi.org/10.1145/2984511.2984582

[24] Chen Liang, Chun Yu, Xiaoying Wei, Xuhai Xu, Yongquan Hu, Yuntao Wang, and Yuanchun Shi. 2021. *Auth+Track: Enabling Authentication Free Interaction on Smartphone by Continuous User Tracking*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445624

[25] Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar. *ACM Trans. Graph.* 35, 4, Article 142 (July 2016), 19 pages. https://doi.org/10.1145/2897824.2925953

[26] Shu-Yang Lin, Chao-Huai Su, Kai-Yin Cheng, Rong-Hao Liang, Tzu-Hao Kuo, and Bing-Yu Chen. 2011. Pub - Point upon Body: Exploring Eyes-Free Interaction and Methods on an Arm. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. Association for Computing Machinery, New York, NY, USA, 481–488. https://doi.org/10.1145/2047196.2047259

[27] Guanhong Liu, Yizheng Gu, Yiwen Yin, Chun Yu, Yuntao Wang, Haipeng Mi, and Yuanchun Shi. 2020. Keep the Phone in Your Pocket: Enabling Smartphone Operation with an IMU Ring for Visually Impaired People. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 2, Article 58 (June 2020), 23 pages. https://doi.org/10.1145/3397308

[28] Yilin Liu, Shijia Zhang, and Mahanth Gowda. 2021. NeuroPose: 3D Hand Pose Tracking Using EMG Wearables. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) *(WWW '21)*. Association for Computing Machinery, New York, NY, USA, 1471–1482. https://doi.org/10.1145/3442381.3449890

[29] Yiqin Lu, Bingjian Huang, Chun Yu, Guahong Liu, and Yuanchun Shi. 2020. Designing and Evaluating Hand-to-Hand Gestures with Dual Commodity Wrist-Worn Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 1, Article 20 (March 2020), 27 pages. https://doi.org/10.1145/3380984

[30] Franziska Mueller, Florian Bernard, Oleksandr Sotnychenko, Dushyant Mehta, Srinath Sridhar, Dan Casas, and Christian Theobalt. 2018. Ganerated hands for real-time 3d hand tracking from monocular RGB. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 49–59.

[31] Adiyan Mujibiya, Xiang Cao, Desney S. Tan, Dan Morris, Shwetak N. Patel, and Jun Rekimoto. 2013. The Sound of Touch: On-Body Touch and Gesture Sensing Based on Transdermal Ultrasound Propagation. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces* (St. Andrews, Scotland, United Kingdom) *(ITS '13)*. Association for Computing Machinery, New York, NY, USA, 189–198. https://doi.org/10.1145/2512349.2512821

[32] Jakub Nalepa and Michal Kawulok. 2014. Fast and Accurate Hand Shape Classification. In *Beyond Databases, Architectures, and Structures*, Stanislaw Kozielski, Dariusz Mrozek, Pawel Kasprowski, Bozena Malysiak-Mrozek, and Daniel Kostrzewa (Eds.). Communications in Computer and Information Science, Vol. 424. Springer, 364–373. https://doi.org/10.1007/978-3-319-06932-6_35

[33] Pedro L. D. Peres, Ivanil S. Bonatti, and Amauri Lopes. 1998. Transmission Line Modeling: A Circuit Theory Approach. *Siam Review* 40, 2 (1998), 347–352.

[34] Munehiko Sato, Ivan Poupyrev, and Chris Harrison. 2012. Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 483–492. https://doi.org/10.1145/2207676.2207743

[35] 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. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 2, Article 59 (June 2020), 25 pages. https://doi.org/10.1145/3397309

[36] A. Sinha, C. Choi, and K. Ramani. 2016. DeepHand: Robust Hand Pose Estimation by Completing a Matrix Imputed with Deep Features. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 00. 4150–4158. https://doi.org/10.1109/CVPR.2016.450

[37] Mohamed Soliman, Franziska Mueller, Lena Hegemann, Joan Sol Roo, Christian Theobalt, and Jürgen Steimle. 2018. FingerInput: Capturing Expressive Single-Hand Thumb-to-Finger Microgestures. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces* (Tokyo, Japan) *(ISS '18)*. Association for Computing Machinery, New York, NY, USA, 177–187. https://doi.org/10.1145/3279778.3279799

[38] D. Tang, J. Taylor, P. Kohli, C. Keskin, T. Kim, and J. Shotton. 2015. Opening the Black Box: Hierarchical Sampling Optimization for Estimating Human Hand Pose. In *2015 IEEE International Conference on Computer Vision (ICCV)*. 3325–3333. https://doi.org/10.1109/ICCV.2015.380

[39] 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* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 891–898. https://doi.org/10.1145/2957265.2961860

[40] 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* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 791–798. https://doi.org/10.1145/2957265.2961859

[41] Saiwen Wang, Jie Song, Jaime Lien, Ivan Poupyrev, and Otmar Hilliges. 2016. Interacting with Soli: Exploring Fine-Grained Dynamic Gesture Recognition in the Radio-Frequency Spectrum. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 851–860. https://doi.org/10.1145/2984511.2984565

[42] Yuntao Wang, Ke Sun, Lu Sun, Chun Yu, and Yuanchun Shi. 2016. SkinMotion: What Does Skin Movement Tell Us?. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 914–917. https://doi.org/10.1145/2968219.2979132

[43] Y. Wang, C. Yu, L. Du, J. Huang, and Y. Shi. 2014. BodyRC: Exploring Interaction Modalities Using Human Body as Lossy Signal Transmission Medium. In *2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*. 260–267.

[44] Martin Weigel, Aditya Shekhar Nittala, Alex Olwal, and Jürgen Steimle. 2017. SkinMarks: Enabling Interactions on Body Landmarks Using Conformal Skin Electronics. 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, 3095–3105. https://doi.org/10.1145/3025453.3025704

[45] G. Welch. 2001. Kalman Filter. *Siggraph Tutorial* (2001).

[46] Yueting Weng, Chun Yu, Yingtian Shi, Yuhang Zhao, Yukang Yan, and Yuanchun Shi. 2021. *FaceSight: Enabling Hand-to-Face Gesture Interaction on AR Glasses with a Downward-Facing Camera Vision.* Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445484

[47] Eric Whitmire, Mohit Jain, Divye Jain, Greg Nelson, Ravi Karkar, Shwetak Patel, and Mayank Goel. 2017. DigiTouch: Reconfigurable Thumb-to-Finger Input and Text Entry on Head-Mounted Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 113 (Sept. 2017), 21 pages. https://doi.org/10.1145/3130978

[48] Mathias Wilhelm, Daniel Krakowczyk, Frank Trollmann, and Sahin Albayrak. 2015. ERing: Multiple Finger Gesture Recognition with One Ring Using an Electric Field. In *Proceedings of the 2nd International Workshop on Sensor-Based Activity Recognition and Interaction* (Rostock, Germany) *(iWOAR '15)*. Association for Computing Machinery, New York, NY, USA, Article 7, 6 pages. https://doi.org/10.1145/2790044.2790047

[49] Erwin Wu, Ye Yuan, Hui-Shyong Yeo, Aaron Quigley, Hideki Koike, and Kris M. Kitani. 2020. *Back-Hand-Pose: 3D Hand Pose Estimation for a Wrist-Worn Camera via Dorsum Deformation Network.* Association for Computing Machinery, New York, NY, USA, 1147–1160. https://doi.org/10.1145/3379337.3415897

[50] Chi Xu, Lakshmi Narasimhan Govindarajan, Yu Zhang, and Li Cheng. 2016. Lie-X: Depth Image Based Articulated Object Pose Estimation, Tracking, and Action Recognition on Lie Groups. *CoRR* abs/1609.03773 (2016). arXiv:1609.03773 http://arxiv.org/abs/1609.03773

[51] Xuhai Xu, Haitian Shi, Xin Yi, WenJia Liu, Yukang Yan, Yuanchun Shi, Alex Mariakakis, Jennifer Mankoff, and Anind K. Dey. 2020. EarBuddy: Enabling On-Face Interaction via Wireless Earbuds. 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, 1–14. https://doi.org/10.1145/3313831.3376836

[52] Zheer Xu, Weihao Chen, Dongyang Zhao, Jiehui Luo, Te-Yen Wu, Jun Gong, Sicheng Yin, Jialun Zhai, and Xing-Dong Yang. 2020. BiTipText: Bimanual Eyes-Free Text Entry on a Fingertip Keyboard. 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, 1–13. https://doi.org/10.1145/3313831.3376306

[53] 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* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 883–899. https://doi.org/10.1145/3332165.3347865

[54] Yui-Pan Yau, Lik Hang Lee, Zheng Li, Tristan Braud, Yi-Hsuan Ho, and Pan Hui. 2020. How Subtle Can It Get? A Trimodal Study of Ring-Sized Interfaces for One-Handed Drone Control. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 2, Article 63 (June 2020), 29 pages. https://doi.org/10.1145/3397319

[55] Chun Yu, Xiaoying Wei, Shubh Vachher, Yue Qin, Chen Liang, Yueting Weng, Yizheng Gu, and Yuanchun Shi. 2019. *HandSee: Enabling Full Hand Interaction on Smartphone with Front Camera-Based Stereo Vision.* Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300935

[56] 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 (Sept. 2017), 19 pages. https://doi.org/10.1145/3130985

[57] Xu Zhang, Xiang Chen, Wen-hui Wang, Ji-hai Yang, Vuokko Lantz, and Kong-qiao Wang. 2009. Hand Gesture Recognition and Virtual Game Control Based on 3D Accelerometer and EMG Sensors. In *Proceedings of the 14th International Conference on Intelligent User Interfaces* (Sanibel Island, Florida, USA) *(IUI '09)*. Association for Computing Machinery, New York, NY, USA, 401–406. https://doi.org/10.1145/1502650.1502708

[58] Yang Zhang, Wolf Kienzle, Yanjun Ma, Shiu S. Ng, Hrvoje Benko, and Chris Harrison. 2019. ActiTouch: Robust Touch Detection for On-Skin AR/VR Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1151–1159. https://doi.org/10.1145/3332165.3347869