WritingRing: Enabling Natural Handwriting Input with a Single IMU Ring

Zhe He
Department of Computer Science and Technology
Tsinghua University
Beijing, Beijing, China
hz23@mails.tsinghua.edu.cn

Zixuan Wang
Department of Computer Science and Technology
Tsinghua University
Beijing, China
w-zx21@mails.tsinghua.edu.cn

Chun Yu*
Department of Computer Science and Technology
Tsinghua University
Beijing, China
chunyu@tsinghua.edu.cn

Chengwen Zhang
Department of Computer Science and Technology
Tsinghua University
Beijing, China
zcwoctopus@gmail.com

Xiyuan Shen
Paul G. Allen School of Computer Science & Engineering
University of Washington
Seattle, Washington, USA
Department of Computer Science and Technology
Tsinghua University
Beijing, China
xyshen@uw.edu

Yuanchun Shi
Department of Computer science and Technology
Tsinghua University
Beijing, China
Qinghai University
Xining, China
shiyc@tsinghua.edu.cn

Figure 1: The Demonstration of WritingRing. (A) WritingRing supports natural handwriting on the surface using an IMU ring worn at the base of the index finger. (B) The handwriting trajectory reconstructed by WritingRing.

Abstract

Tracking continuous 2D sequential handwriting trajectories accurately using a single IMU ring is extremely challenging due to the significant displacement between the IMU's wearing position and the location of the tracked fingertip. We propose WritingRing, a system that uses a single IMU ring worn at the base of the finger to support natural handwriting input and provide real-time 2D trajectories. To achieve this, we first built a handwriting dataset using a touchpad and an IMU ring (N=20). Next, we improved the LSTM model by incorporating streaming input and a TCN network, significantly enhancing accuracy and computational efficiency, and achieving an average trajectory accuracy of 1.63mm. Real-time usability studies demonstrated that the system achieved 88.7% letter recognition accuracy and 68.2% word recognition accuracy, which reached 84.36% when restricting the output to words within a vocabulary of size 3000. WritingRing can also be embedded into existing ring systems, providing a natural and real-time solution for various applications.

*Corresponding author.

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3714066

CCS Concepts

• **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**; **Gestural input**; *Text input*.

Keywords

Ring Interaction, Handwriting, IMU, Touch Interface, Wearable

ACM Reference Format:

Zhe He, Zixuan Wang, Chun Yu, Chengwen Zhang, Xiyuan Shen, and Yuanchun Shi. 2025. WritingRing: Enabling Natural Handwriting Input with a Single IMU Ring. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3706598.3714066

1 INTRODUCTION

Smart rings have demonstrated significant potential to become the next generation of wearable devices due to their compact size, portability, comfort, and "always available" nature. Interactions based on smart rings have been widely explored, with researchers integrating various sensors, such as cameras[9, 25, 56], magnetic sensors[4, 10, 11, 38], Inertial Measurement Units (IMUs)[15, 36, 44, 45, 63], and proximity sensors[35, 55], to enable handwriting recognition, 2D pointing, and gesture recognition. However, smart rings still face challenges in terms of power consumption, wearability, and the naturalness of use. Their sensing capabilities remain limited, making it difficult to balance power efficiency with rich interaction functionality.

Therefore, in this paper, we focus on achieving handwriting input using a single IMU ring. Compared to other input methods utilizing smart rings, handwriting is more natural, expressive, and involves a lower learning curve for users. Additionally, IMU sensors meet the requirements of low power consumption and long-term operation, making them ideal for this task. Previous work often treated the finger as a rigid body, using in-air IMU trajectories to reconstruct finger movements[29, 57, 58]. However, in-air handwriting requires the hand to be suspended, which leads to significant fatigue and can not support continuous word input. In contrast, plane-based handwriting is much easier and allows contact and lift from the surface to define word boundaries. However, in plane-based handwriting, the shape of the user's finger changes, so it can no longer be considered a rigid body. Additionally, the displacement between the IMU's placement (usually at the base of the finger) and the predicted trajectory (typically at the fingertip) presents a significant challenge.

Based on this, we propose WritingRing, which utilizes a single IMU-equipped ring worn at the base of the finger to reconstruct 2D fingertip movement trajectories in real time during natural handwriting tasks. To the best of our knowledge, this is the first work to achieve real-time, cross-user handwriting recognition and reconstruction on the plane with high accuracy using a single IMU ring. To achieve this, we first collected an IMU handwriting dataset comprising 20 participants with a total duration of approximately 20 hours. To the best of our knowledge, this is the largest publicly available IMU ring handwriting dataset to date.

Next, we enhanced the traditional Long Short-Term Memory (LSTM)[17] model by integrating a Temporal Convolutional Network (TCN)[23] to better capture short-term motion information. Additionally, we improved the data slicing method by adopting a streaming approach for training on continuous data segments, and through an ablation study, we demonstrated the significant advantages of this training method for IMU data. Leveraging the algorithm's ability to comprehensively utilize both recent and historical temporal information, WritingRing achieved an average trajectory error of 1.63 mm. We directly fed the reconstructed trajectories into standard handwriting recognition software (Google IME), achieving an average letter recognition accuracy of 88.7% and an average word recognition accuracy of 68.2% (84.36% when restricting the output to words within a vocabulary of size 3000). Usability evaluation experiments showed that WritingRing has high recognition accuracy in real-world usage scenarios, and users expressed a strong willingness to use the system. Furthermore, we demonstrated that the algorithm can be directly deployed on the chips of existing rings, greatly reducing power consumption from signal transmission and showcasing its potential for on-device execution in current hardware.

In summary, our contributions are as follows:

(1) We introduced WritingRing, an efficient, natural, and portable handwriting input solution. It allows users to perform handwriting input on any surface in a natural way, achieving an average trajectory error of 1.63 mm, a letter recognition accuracy of 88.7% and a word recognition accuracy of 68.2% (84.36% when restricting the output to words within a vocabulary of size 3000).

(2) We created and publicly released what we believe to be the largest dataset of 2D handwriting input using an IMU ring. This dataset not only supports the implementation of handwriting functionality on existing smart rings but also serves as a valuable resource for evaluating time-series model performance.

(3) We explored the potential application space of WritingRing and demonstrated its high accuracy in real-world use through two usability studies, where users expressed a strong willingness to adopt the system.

2 RELATED WORK

2.1 Ring-based Input Techniques

The primary goal of WritingRing is to enable efficient and natural handwriting input on any surface by reconstructing the user's handwriting trajectory. To summarize previous research, we have compiled related work on ring-based input in Table 1.

Currently, ring-based input devices employ various sensors, primarily including cameras, pressure sensors, electromagnetic sensors, and IMUs. Devices like Magic Finger[56] and NailRing[25] use cameras mounted on the fingertip to capture finger movements and micro-gestures. CyclopsRing[9] features a fisheye camera mounted on the finger to observe the entire skin area of the hand, capturing both gestures and environmental information. Yuki Kubo's work[22] uses a small pressure sensor on the finger to simulate a touchpad for interaction. Systems like Finexus[11], AuraRing[38], Nenya[4], and uTrack[10] utilize electromagnetic sensors to detect finger position and movement for input tasks. While these approaches are intuitive, they have notable limitations: cameras consume significant power, making long-term use difficult; pressure sensors are bulkier compared to traditional rings, affecting daily wearability; and electromagnetic sensors often require multiple devices to work together (such as wearing wristbands or multi-finger setups), which can diminish portability and user experience.

Table 1: Overview of smart input rings.

Work	Sensor	Number	Position	Additional Device	Support Handwriting	Support Finger Tracking
Magic Finger[56]	Camera	1	Fingerpad	-	✗	✓
NailRing[25]	Camera	1	Fingertip	-	✗	✓
CyclopsRing[9]	Camera	1	Proximal Phalanx	-	✗	✓
Ring-type Device[22]	Pressure Board	1	Intermediate Phalanx	-	✗	✗
Finexus[11]	Electromagnet	4	Fingertip	Wristband	✓	✓
AuraRing[38]	Electromagnet	1	Proximal Phalanx	Wristband	✓	✓
Nenya[4]	Magnetic Sensors	1	Proximal Phalanx	Wristband	✗	✗
uTrack[10]	Magnetic Sensors	1	Intermediate phalanx	thumb Permanent Magnet	✗	✓
MouseRing[45]	IMU	1/2	Proximal phalanx (Intermediate phalanx)	-	✗	✓
ssLOTR[64]	IMU	5	Proximal Phalanx	Wristband	✓	✗
Anywhere Touch[36]	IMU	1	Fingertip	-	✗	✓
PeriSense[55]	Capacitive Proximity Sensor	1	Proximal Phalanx	-	✗	✗
QwertyRing[14]	IMU	1	Fingertip	-	✗	✗
Mouse on a Ring[63]	IMU	1	Proximal Phalanx	-	✗	✗
RotoSwype[15]	IMU	1	Proximal Phalanx	-	✓	✗
LightRing[20]	IMU + Infrared Sensor	1	Proximal Phalanx	Infrared Sensor	✗	✓
WritingRing	IMU	1	Proximal Phalanx	-	✓	✓

In contrast, IMU-based rings are lighter, smaller, and have lower power consumption. A considerable amount of research has explored IMU ring interactions. For instance, MouseRing[45] accurately reconstructs fingertip trajectories using two rings, enabling touchpad-like interactions. ssLOTR[64] employs self-supervised learning to achieve high-precision 3D finger motion tracking. Anywhere Touch[36] combines IMU data with anatomical models of the hand to compute finger joint movements via inverse kinematics. However, their methods use two or more rings or place rings away from the base of the finger, which are not aligned with daily usage habits.

Existing single-ring devices (worn at the base of the finger) have limited sensing capabilities. Although some studies have explored single-ring text input and 2D cursor control, few have managed to achieve continuous fingertip trajectory reconstruction and text input with a single ring. QwertyRing[14] captures continuous finger taps using a ring worn on the fingertip to simulate a virtual QWERTY keyboard, but it has a steep learning curve. Mouse on a Ring[63] and RotoSwype[15] use IMU sensors to track finger movements and translate them into cursor control, but both rely on mid-air interaction, which can lead to fatigue over extended use. LightRing[20] supports handwriting on a surface but is restricted by infrared sensors, which prevent its use with the wrist lifted from the desk, limiting the range of finger movements.

Compared to the above-mentioned works, our approach does not impose restrictions on wrist posture, allowing users to write in their preferred position on any surface, ensuring comfort and convenience. Additionally, we support high-precision fingertip trajectory reconstruction, enabling direct text input without the need for calibration. This ensures a positive user experience in real-world scenarios, encouraging consistent use and adoption of the device.

2.2 Trajectory Reconstruction Based on IMU

This section reviews traditional IMU-based trajectory reconstruction methods, techniques for reducing IMU data noise, and applications of IMU in trajectory reconstruction, highlighting the advantages of our approach.

There have been substantial applications focused on trajectory reconstruction using IMU (Inertial Measurement Unit) sensors [1, 42]. Traditional trajectory reconstruction methods aim to estimate the motion trajectory of IMU sensors in their local coordinate system. However, IMU sensors inherently suffer from intrinsic biases and random noise. Simply using raw IMU data leads to errors that accumulate over time and with movement [18, 53], making it challenging to achieve accurate, large-scale, long-term trajectory reconstructions. As a result, conventional positioning methods typically use IMUs as auxiliary components in positioning systems [32, 40, 62].

To address these limitations, various techniques have been proposed to enhance accuracy. For instance, some approaches utilize discriminators to identify stationary states and remove static biases [37] . Deep learning methods have also been employed, such as denoising altitude estimates using open-loop gyroscope data [7]. In addition to noise reduction [12, 13, 34, 39, 41, 61], some studies predict trajectories beyond the IMU itself. For instance, IMUs have been used to infer hand gestures through wearable rings or watches [27, 46, 60], classify handwritten characters [8, 43], and capture human motion [21]. However, these approaches have notable limitations. Some require a rigid-body linkage between the IMU and the target object [8], others rely on additional sensors for enhanced positioning [20, 43], and many use multiple IMUs [21, 46]. Moreover, certain end-to-end models only produce task-specific outputs [27, 60], such as handwriting recognition using language models [60], without reconstructing the trajectory itself.

In contrast, our work utilizes a single IMU sensor worn at the base of the finger to accurately predict the non-rigid-attached motion of the fingertip. We normalize the input rotation using gravity and adopt a hybrid model of Temporal Convolutional Networks (TCN) [23] and Long Short-Term Memory (LSTM) [17] to forecast fingertip velocity, instead of position, in a streaming fashion. With this method, we reach millimeter-level accuracy, making it effective for subsequent tasks such as handwriting.

2.3 Recognition of Finger Handwriting Based on Wearable Sensors

Handwriting recognition technologies often face a trade-off between accuracy and user convenience. We first review high-accuracy solutions that compromise comfort, followed by portable devices that prioritize convenience but have limitations in accuracy and generalizability. Finally, we introduce the advantages of our WritingRing compared to existing finger-worn IMU works.

In the field of handwriting recognition, high-accuracy solutions often compromise user convenience. For example, Tigrini et al. [49, 50, 51] use surface electromyography (sEMG) signals to recognize numbers. Singh and Chaturvedi [48] combine IMUs and sEMGs for letter recognition. However, these systems require multiple electrodes on the user's forearm or wrist, impacting comfort. Other advanced solutions, like data gloves [2, 54], radar sensors [24], ToF sensors [59], or head-mounted cameras [6], rely on external devices, reducing practicality for fast, everyday applications.

Portable and comfortable handwriting recognition devices like smartwatches, fingertip sensors, and smart rings offer high convenience. Li et al. [26] use IMU sensors in smartwatches for 91.4% word recognition accuracy but only support uppercase letters and require extensive personalized data collection for each user. Ardüser et al. [3] combine smartwatch motion data with audio signals, which can be affected by noisy environments. Fingernail pressure sensors used by Blumrosen et al. [5] achieved only about 80% accuracy. Liu et al. [28] use a portable ring with an embedded IMU, but the study was limited to one user, raising concerns about generalizability and real-world usability.

Previous works using finger-worn IMUs focus on in-air writing, treating the hand as a rigid body to reconstruct the IMU trajectory. Luo et al. [29] achieved 73% accuracy with DTW for character recognition, requiring pre-collected data for each user without any generalization ability. Younas et al. [57] used subjective evaluation of the writing results, which was later improved by Younas et al. [58] through a combination of KNN and DTW. However, to the best of our knowledge, all in-air writing approaches do not support continuous word input. Our work, in contrast, adopts a writing-on-plane approach. This introduces new challenges, as the sensor and fingertip are no longer rigidly attached, and the smaller range of motion amplifies the IMU noise. WritingRing addresses these challenges effectively.

We believe that existing research lacks handwriting input technology that simultaneously provides both high accuracy and user comfort while remaining practical for real-world applications. WritingRing, a single-ring design worn at the base of the finger, enables natural writing on a plane. This setting is identified as the most acceptable option in our user study due to its portability and comfort. To demonstrate its generalizability, we conduct real-time writing experiments with 24 users who are not part of the training set. These experiments include both character and word writing tasks. WritingRing achieves high recognition accuracy, and excellent generalizability, and supports cursive writing and word input, ensuring high input efficiency without compromising ease of use. This makes it a highly practical solution for everyday use, overcoming the limitations of previous methods.

3 WRITING RING DESIGN

The goal of WritingRing is to enable efficient, natural, and comfortable finger handwriting input on any surface. To understand what kind of experience users would find most natural, we conducted a survey that asked participants to imagine a well-developed technology that implemented handwriting functions and provide their opinions on it. This pilot study involved 32 participants (13 males and 19 females, aged from 18 to 60, M=28.91. Among them, 6 participants were older, and 6 were non-local participants (from Asia, North America, Europe, Africa, etc.).), with an average completion time of 6.51 minutes. Each participant received a reward of 2 dollars. We summarized our considerations and the actual user needs into the following two points.

3.1 Considerations for Wearing

DQ1: How many rings are users willing to wear?

We found that most users were reluctant to wear two or more rings. In our survey, 100% of the users are willing to wear one smart ring to use this technology, but only 40.6% are willing to wear two or more rings. The reasons they opposed wearing additional rings include inconvenience of wearing (21/32) / interference with daily activities (11/32) / impact on aesthetics (5/32). Although the accuracy of reconstructing finger movements using IMUs highly depends on the number of IMUs used [45], more rings mean a greater burden on the user in terms of wearability.

DQ2: What is the natural wearing position for users?

Nearly all participants (81.3%) in our survey indicated that wearing it at the base of the finger aligns with their daily habits. Meanwhile, only 28.1% of users were willing to wear the ring on the proximal phalanx or distal phalanx, even though these positions have been shown to provide more movement information[47]. The reasons for their reluctance include inconvenience of wearing (9/32) / interference with daily activities (22/32) / impact on aesthetics (1/32) / ease of falling off (6/32) .

Therefore, to increase the acceptance of this technology among users, we chose to use a single ring worn at the base of the index finger. Although this presents significant technical challenges, it is the most user-friendly approach, aligning with the natural habits of users.

3.2 Considerations for Handwriting

DQ3: What is the comfortable handwriting posture for users?

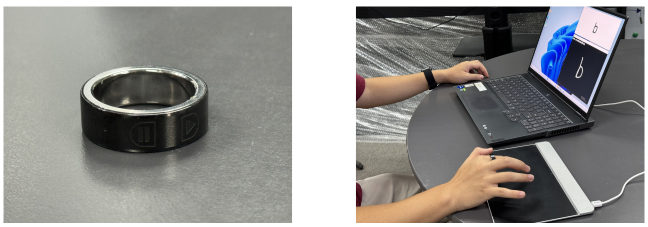
We observed that users naturally adopt two distinct postures when performing handwriting input: resting the wrist on the desk or keeping the wrist elevated (50% vs. 50% in the survey). The handwriting motion in these two postures corresponds to more finger joint flexion and extension in the first case, and full-hand movement in space in the second. To support users in writing in any natural way they prefer, both postures were taken into account in our design.

DQ4: Is visual feedback necessary?

Although handwriting can be an eyes-free process, about 60% of the respondents (19/32) in our survey expressed a preference for seeing the trajectory of their writing as they produce it. The reasons for this preference include greater confidence during input (10/19) / a more natural usage experience (7/19) / convenience to correct input (4/19). As a result, instead of performing direct end-to-end recognition of the user's handwriting, we first reconstruct the input trajectory and then recognize the content from this trajectory. The step of recognizing content from the trajectory has been extensively studied, so by accurately reconstructing the trajectory, we can effectively utilize existing handwriting recognition algorithms to complete the recognition process.

4 Data Collection

To implement our system which supports natural handwriting based on a single IMU ring worn at the base of the index finger, we conducted a data collection experiment to gather data on users performing natural finger handwriting tasks while wearing an IMU-equipped smart ring.



(a) The Smart Ring Used in the Experiment.

(b) The Participant is Inputting Letter 'b' in the Experiment.

Figure 2: Apparatus for data collection.

4.1 Apparatus

To ensure that users' behavior closely aligns with real-world usage scenarios, we opted not to connect the IMU via wires, as this could interfere with natural user behavior. Instead, we used a smart ring equipped with an IMU that communicates with a PC via Bluetooth. The IMU model is MPU9250, featuring 6 axes (see hardware design in the appendix). During the experiment, the IMU's frame rate was fixed at 200 fps. To capture the movement trajectory of the user's fingertip, we synchronized the recording with a Sensel Morph[1] pressure pad, set at a frame rate of 100 fps to capture pressure data. The experiment setup is shown in Fig. 2.

4.2 Participants

We recruited 20 participants (10 males and 10 females, aged 19 to 30, M = 23.48) from the university campus. All participants were right-handed, with an average right index finger length of 81.0 mm (SD = 12.5), which we believe represents a typical range of finger lengths in the general population.

4.3 Design and Procedure

At the start of the experiment, participants were instructed to wear the smart ring on the base of their right index finger. For participants with slimmer fingers, the ring was slightly loose when worn directly. To address this, we applied nano tape inside the ring to ensure a secure fit, preventing any loosening during the experiment. A marker was placed on the ring, which needed to face directly upward when the palm was laid flat. This ensured that the IMU's relative position to the finger remained consistent, with only minor variations that would not impact the experimental outcomes.

During the experiment, participants were asked to write either letters or words on the pressure pad while wearing the ring. Letters were case-sensitive, and words were randomly selected from the MacKenzie PhraseSet [30], with both connected writing and unconnected writing required. Based on the considerations discussed in the previous chapter, data was collected for both wrist-lifted and wrist-resting conditions. In this context, connected writing refers to allowing users to write an entire word in a single stroke without lifting their finger, while unconnected writing requires users to lift their finger between letters. Wrist lifted means the wrist remains elevated during the writing process, whereas wrist resting refers to the wrist being placed on the desk. Thus, each participant completed 6 writing tasks: 3 writing types (letters, connected writing words, unconnected writing words) $\times$ 2 wrist conditions (wrist lifted, wrist resting). Each task included approximately 5 blocks, resulting in a total of 600 words (including letters) being written. After completing each block, participants were given a 2-minute rest. The entire experiment lasted approximately 60 minutes, and each participant received 15 dollars as reward. The entire dataset comprises approximately 20 hours of handwriting data. All the data we collected can be found at https://huggingface.co/datasets/dBHz/WritingRing.

4.4 Data Pre-processing

Due to the presence of noise in the touchpad data, and considering that the finger's movement trajectory is relatively smooth, we applied exponential smoothing to the data on both the x and y axes before using it as the ground truth. The smoothing process is defined as follows:

$$S_t = \alpha \cdot X_t + (1 - \alpha) \cdot S_{t-1}$$

where S_t is the smoothed value at time t, X_t is the original data point, α is the smoothing factor, and S_{t-1} is the previous smoothed value. This smoothing helps reduce noise while preserving the overall trajectory. α is set to 0.5 in this experiment.

Additionally, due to a constant time offset between the IMU data timestamps and the touchpad data timestamps, we manually calibrated the time offset for each data point by aligning the IMU's peak values with the touchpad press timestamps. This process ensured that the time discrepancy between the IMU and touchpad data did not exceed 10 ms (2 frames). Furthermore, we used linear interpolation to map the touchpad data to the exact timestamps of each IMU sampling. After these adjustments, the time deviation between the IMU data and the touchpad data was kept below 10 ms, which can be considered instantaneous, with almost no delay[33].

5 WRITING RING ALGORITHM

5.1 Algorithm Pipeline

Figure 3 presents the pipeline of our algorithm. First, we apply a pose estimation algorithm to estimate the ring's orientation in space, followed by subtracting the gravitational component from the three-axis accelerometer data. This step reduces the influence of different orientations on the results and gets the acceleration

[1] https://morph.sensel.com/

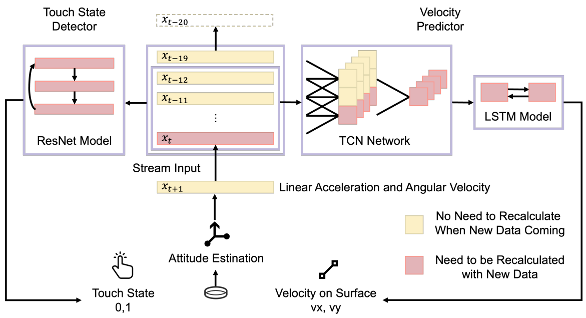


Figure 3: Algorithm pipeline for WritingRing.

(after gravity removal) and angular velocity as the input for the model. Next, we process the data in fixed-length windows, passing it through a carefully designed TCN network, ensuring that the temporal length is reduced to 1 with a convolutional stride of 1. The resulting data is then fed into an LSTM network, which predicts the velocity at the current moment.

In this framework, data can be streamed continuously, allowing us to maximize the reuse of previously computed results and minimize computational overhead during real-time usage. For example, as shown in Figure 3, since the stride is 1 when a new data frame arrives, only the red section needs to be recalculated, while the rest can utilize previously computed results. Additionally, we implemented a simple yet efficient touch state detector to precisely capture the moments of press and release during writing. Note that when the finger is in the air, the model output is meaningless. However, the continuous IMU input in the sequence helps the model learn more contextual information. Therefore, during training, we applied a mask to the model's output when there is no contact, meaning that only the data during contact will contribute to the model's loss. In actual use, we decide whether to use the current model's output based on the detection of finger press and lift events.

Finally, the reconstructed trajectory is fed into a handwriting recognition program. Most existing smart devices are equipped with handwriting input methods, and for our system, we chose the commonly used Google Input Method (Google IME)[2] as the recognition program.

5.2 Attitude Estimation and Linear Acceleration Calculation

To ensure consistent performance across various user postures, we opted to remove the gravity component from the accelerometer's output, as the gravity value is significantly larger (approximately 400% on average) than the linear acceleration generated during normal handwriting. By removing it, we improve the stability of the data, making it easier for the model to learn effectively.

To calculate the gravitational components along each axis, we first need to determine the orientation of the ring. For this, we used the Madgwick algorithm, commonly employed in inertial navigation [31]. This algorithm integrates angular velocity data and applies corrections using acceleration data. The gaining factor β were set to 0.041. By using this algorithm, we can compute the relative orientation of the ring with respect to the horizontal plane at any given moment. Once the orientation is known, the gravitational vector in the horizontal state can be rotated and subtracted accordingly. The equation is as follows:

$$g_{adjusted} = R(\theta) \cdot g_{horizontal}$$

$$a_{adjusted} = a_{raw} - g_{adjustud}$$

where $g_{adjusted}$ is the gravitational component to be subtracted, $R(\theta)$ is the rotation matrix derived from the ring's orientation, and $g_{horizontal}$ represents the gravity vector in the horizontal state. By subtracting $g_{adjusted}$ from the raw accelerometer reading a_{raw}, we can get the adjusted acceleration $a_{adjusted}$.

5.3 Trajectory Reconstruction

5.3.1 Model Architecture. Previous approaches often involve segmenting the data into small fragments, using neural networks to predict the IMU velocity at the end of each segment, and then integrating the velocity to obtain displacement. However, this method has several drawbacks. For instance, if the data fragments are too short, the model cannot capture long-term velocity information. On the other hand, if the fragments are too long, the computational efficiency decreases, and the data may contain irrelevant information for predicting the final velocity, which can lead to overfitting.

Therefore, instead of using segmented data as input, we designed a streaming model structure that uses the entire handwriting process as input, a method later validated as highly effective in our experiments. The most straightforward approach would be to pass the data through an LSTM network and predict velocity using a linear layer, forming a frame-to-frame architecture. However, this frame-to-frame prediction method is highly sensitive to temporal misalignment, as the data does not always perfectly align in time. Additionally, using a certain amount of future data improves the prediction of current displacement.

To address these issues, we designed our current algorithm pipeline. Data from a fixed-length window is processed by a Temporal Convolutional Network (TCN), which compresses the sequence into a single feature vector of length 1 with a dimension of 128. This vector is then fed into an LSTM network with a hidden size of 128, and the output is passed through a linear layer (with ReLU as the activation function) to predict the velocity in the x and y directions. This velocity corresponds to the fingertip movement speed at the midpoint of the data window. After testing, the optimal data window length was set to 13, resulting in a maximum delay of 6.5 frames, or 32.5 ms, which is acceptable for handwriting input.

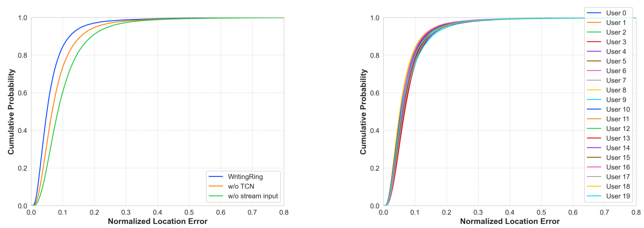
In this prediction process, the TCN network effectively captures motion variations within the current data segment, while the LSTM network leverages historical motion information through its hidden states to predict the current movement. This combination ensures accurate velocity prediction by incorporating both past and present information.

5.3.2 Model Training. Based on the analysis in the previous section, we adopted a streaming approach with longer data segments for training, rather than cutting the data into smaller fragments. In practical experiments, since a single block often takes 1-2 minutes to

[2]https://www.google.com/inputtools/

complete, we fixed the data segment length to 75 seconds (or 15,000 frames). For data shorter than this length, we padded the beginning with zeros, and for longer data, we symmetrically trimmed from both ends. This approach facilitates batch training.

We randomly divided the data (including both letters and words) into 70%, 10%, and 20% for the training, validation, and test sets, respectively. During training, the batch size was set to 16, the learning rate was 0.001, and we used the Adam optimizer. The loss function was Mean Squared Error (MSE), and the training process continued for 500 epochs.



(a) The performance of different models.

(b) The performance on different users.

Figure 4: The performance of different models and different users, presented through the cumulative distribution function.

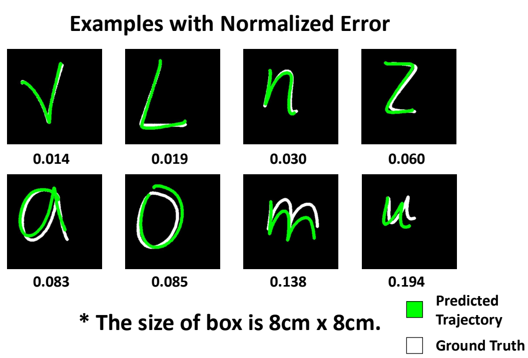
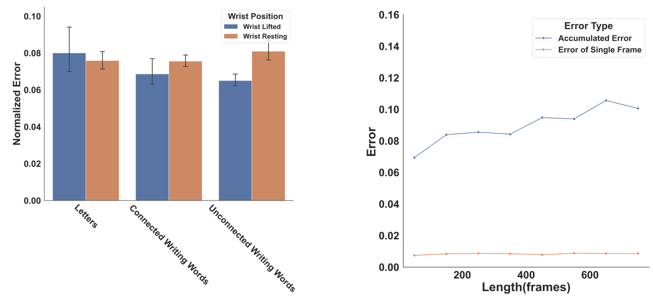


Figure 5: Samples from the dataset and their normalized error.

5.3.3 Model Performance. We used the normalized distance between each predicted point and ground truth as an evaluation metric for model performance. The formula is $\frac{\sqrt{(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2}}{L}$ where $\hat{x}_i$ and $\hat{y}_i$ are the predicted coordinates, x_i and y_i are the ground truth coordinates, and L represents the normalizing factor (typically the diagonal or a characteristic length of the writing area, here we use the diagonal of the bounding box of the trajectory). The normalization helps eliminate the influence of writing size on trajectory deviation.

As shown in Figure 4(a), the average error is 0.073, with over 80% of the errors being less than 0.1. Since this value might not be intuitive, we extracted several representative samples from the experimental results and annotated their respective errors, as shown in Figure 5. It can be observed that when the error is less than 0.060, the reconstructed trajectory closely matches the real trajectory in shape. When the error is less than 0.194, although there are deviations between the reconstructed and real trajectories, the letters in the trajectory are still recognizable. Additionally, we re-evaluated the error without normalization and found that the average error is 0.164 cm, achieving millimeter-level accuracy. Figure 6(a) shows that the WritingRing algorithm maintains good performance across different handwriting tasks (letter, connected word, and unconnected word) and various wrist positions.



(a) Normalized error under different settings.

(b) Accumulated and per-frame errors over time.

Figure 6: Writing error metrics under different conditions.

Due to errors in IMU readings, these errors accumulate over time and impact the accuracy of the reconstruction. As shown in Figure 6(b), the average normalized error (calculated as the average of data within every 100 frames) gradually increases with the number of consecutive predicted frames, but remains at a relatively low level overall. Furthermore, 99% trajectories have fewer than 300 frames (1.5 seconds). In addition, we also calculated the error for individual frames (i.e., the difference between the predicted and actual velocity), and found that this error does not change with the number of frames. This suggests that although errors accumulate over time as the number of frames increases, the impact is limited, and in most cases, the performance remains satisfactory.

Additionally, we used a leave-one-out approach (we retrained the model, where one participant was left out as the test set and the remaining participants were used as the training set) to evaluate user-independent results, which are visualized in Figure 4(b). As can be seen, while there are some differences in error across participants, the overall performance remains consistent without significant deviations. This suggests that the algorithm is generally applicable and effective for a wide range of users.

5.3.4 Ablation Study. To demonstrate the effectiveness of the specific design choices in our model, we conducted an ablation study, as shown in Figure 4(a). The stream input method had the most significant impact on performance. Compared to the common approach of cutting the dataset into small fragments, using longer segments for streaming input during training resulted in much better performance, reducing the mean error by 36.7%, from 0.116

to 0.073. We believe this is because small segments make it difficult for the model to learn the fundamental patterns of long-term movement, particularly the temporal relationships before and after each movement.

Additionally, using the TCN network to extract features from several adjacent frames was also proven to be effective, reducing the error by 23.1%, from 0.095 to 0.073. Regarding the choice of window length, we found that longer windows did not lead to further performance improvement. This may be because, beyond historical information, the temporal span that affects the current velocity or is influenced by the current velocity is relatively short. Therefore, we selected a TCN window length of 3 and a total window length of 13 frames, balancing computational efficiency and performance.

5.4 Touch Detection

To accurately reconstruct the user's trajectory, the system needs to detect whether the user's index finger is in contact with the surface. During training, we segmented the data into 0.1-second fragments, resulting in 20 frames of 6-channel data. We trained a deep learning model based on ResNet[16], which consists of three residual blocks with increasing sizes (8, 16, and 32 output channels, respectively). Batch normalization and Dropout were applied after each convolutional layer, followed by a fully connected layer for classification, enabling precise detection of finger contact events.

We classified each time window's data into four categories: contact with the surface, in the air, finger lift, and finger press. The four states are defined as follows:

(1) Contact with the surface: The finger inside the time window is continuously in contact with the surface.

(2) In the air: The finger inside the time window is not in contact with the surface.

(3) Finger lift: There is a transition from contact to non-contact within the time window.

(4) Finger press: There is a transition from non-contact to contact within the time window.

The inclusion of "in the air" and "contact with the surface" categories, beyond just detecting lifts and presses, was to enhance model robustness and minimize false triggers during usage. In real-world use, we only used the "finger lift" and "finger press" events to define the start and end of the trajectory. The model achieved a 98.3% accuracy in recognizing press and lift events.

5.5 Word Recognition

For letter recognition, we directly input the predicted trajectories into Google IME, a widely used handwriting recognition program deployable across multiple platforms such as computers and mobile devices. For word recognition, we distinguish between connected and unconnected writing. For connected writing, we directly input the trajectory into Google IME. For unconnected writing, since we do not predict the trajectory after the user lifts their finger, we first concatenate the discrete letters with fixed intervals to form a complete trajectory, which is then input into Google IME. For all inputs, Google IME returns several candidate words, and we retain the top 5 candidates as the final results. Based on this, users can input any letter, word, or even sentence, and can write sequences of arbitrary length.

6 REAL-TIME LETTER WRITING USABILITY EVALUATION

To further validate the practical usability of the system, we conducted two real-time usability evaluation studies , assessing whether the system truly provides a universal, efficient, and natural input experience. The first experiment is to test the basic capabilities of WritingRing, letter writing and recognition. The second experiment (presented in the next section) evaluates the system's performance in long-term continuous writing tasks by testing whole word input.

6.1 Participants and Apparatus

We recruited 12 participants (8 males and 4 females, aged 20 to 30, mean = 22.42) from a university campus, all of whom were right-handed. These participants did not overlap with those in the data collection experiment. Before the experiment began, each participant independently wore the smart ring, ensuring the marked point was facing upward, with minor misalignments permitted. For participants with slimmer fingers, we applied a layer of nano tape inside the ring to prevent it from loosening during the experiment. The system was run and visualized on a MacBook Pro 2021, and the touchpad used was a Sensel Morph.

6.2 Design and Procedure

Before the experiment began, participants were given approximately 5 minutes to familiarize themselves with the use of the smart ring. During this time, they were informed that they could write either with their wrist lifted or resting on the surface and were encouraged to choose the method that felt more comfortable and natural.

Afterward, participants completed two sets of comparison experiments. In one set, they wrote letters on a touchpad in a single stroke, including all uppercase and lowercase letters (excluding "i" and "j" as they cannot be written in one stroke). In the other set, participants performed the same letter-writing task on the desk surface while wearing the smart ring. Each letter was written four times. To counterbalance the learning effect, the order of the two experiments was randomized. A 2-minute break was given between each round.

After completing each set of tasks, participants filled out a subjective questionnaire and provided feedback regarding their experience. The experiment lasted approximately 40 minutes, and participants received 10 dollars for their participation.

6.3 Results

6.3.1 Accuracy. After removing recognition errors for letters that are identical in uppercase and lowercase (e.g., "x" and "X"), the average accuracy for all participants using the ring to input letters was 88.7% (SD = 0.056), while the average accuracy for input via the touchpad was 93.1% (SD = 0.037). The accuracy of ring-based input reached 95.3% of the touchpad input accuracy. The confusion matrix is shown in the appendix Fig. 13.

Common error types included vertical misalignment leading to stroke displacement, such as confusing "q," "a," and "d," or mixing up "b" and "p." Some letters, like "X", have complex stroke orders when written in a single stroke, making them harder for handwriting recognition systems to correctly interpret. Additionally, certain letters tend to introduce extra strokes when written in one motion, leading to misrecognition, such as "t" being recognized as "e". These types of errors accounted for 34.9% of all mistakes.

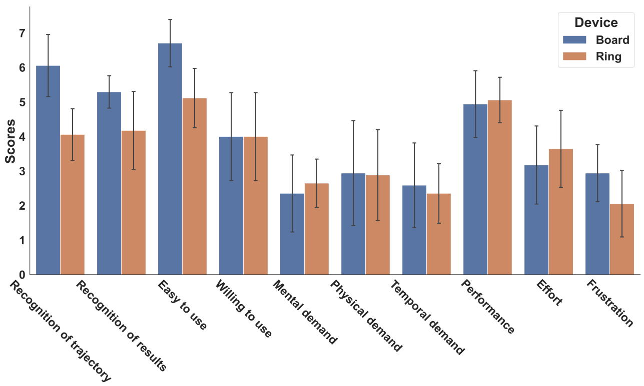


Figure 7: System scoring in the letter writing evaluation study.

6.3.2 Subjective Indicators. At the end of the experiment, we asked participants to rate the system on recognition accuracy, ease of learning, willingness to use, and NASA-TLX on a scale from 1 to 7. We did the Wilcoxon Signed-Rank Test to the results. The results, shown in Figure 7, indicate that the touchpad received significantly higher scores than WritingRing in terms of trajectory reconstruction and letter recognition accuracy, suggesting that the touchpad provided more precise trajectory data ($p < .05$, $Z = -3.06, -2.51$). Also, the touchpad was rated higher in ease of use ($p < .05$, $Z = -2.83$), as it required little to no learning effort from users. However, WritingRing also received relatively strong scores in these categories. The relatively low absolute values for mental demand, physical demand, and temporal demand, along with the high performance, indicate that using the ring for handwriting letters does not impose a greater burden compared to using a touchpad.

Interestingly, in terms of willingness to use, the ratings for the touchpad and WritingRing were comparable ($p = 0.76$, $Z = -1.49$). This suggests that, despite the touchpad's advantages in technical performance and ease of use, WritingRing's other features—such as portability and comfort—were appreciated by users, leading to a similar level of willingness to adopt it. Some participants commented, "While the touchpad performs better, the ring is very portable and allows me to use it anywhere."

In summary, even when compared to the touchpad, which served as the ground truth, WritingRing demonstrated impressive results, particularly in terms of user willingness to adopt it.

7 REAL-TIME WORD WRITING USABILITY EVALUATION

To evaluate the system's performance in long-term continuous writing tasks, we conducted a word-level text input experiment. Compared to individual letters, word input is more commonly used and can better reflect the system's practical usability in real-world scenarios.

7.1 Participants and Apparatus

We recruited 12 participants from a university campus (6 males and 6 females, aged 21 to 29, mean = 23.08), all of whom were right-handed and had not participated in any previous experiments. The devices used in the experiment (including the smart ring and the computer running the program) and the method of wearing the ring were the same as those in the previous experiments(see section 6.1).

7.2 Design and Procedure

We designed two methods for completing word input: unconnected writing and connected writing.

(1) Unconnected writing: In this setup, users are required to input individual letters one by one. After completing each word, they confirm the input by pressing a touch button.

(2) Connected writing: In this setup, users write the entire word continuously without lifting their fingers. After completing the word, they press a touch button to confirm the input.

The recognition algorithm remains the same as in Section 5.5. In both methods, the dots for the letters "i" and "j" are not required to be written, as we found that Google IME often recognizes the correct result even without the dots.

Before the experiment began, users were given 5 minutes to familiarize themselves with the usage of the smart ring. Then, users were asked to complete tasks for both input methods. To counterbalance the order effect, half of the participants started with the unconnected writing task, while the other half started with the connected writing task.

For each input method, users completed 5 rounds of tasks. In each round, participants were required to input 20 randomly selected words in lowercase (drawn from a list of the 3000 most common English words [19]). Errors could not be undone during the test. There was a break of at least 1 minute between each round of testing.

After completing one input method, users filled out a subjective questionnaire, which included the NASA-TLX scale and a subset of questions extracted from the SUS (System Usability Scale). Participants were also given the opportunity to provide qualitative feedback. The entire experiment lasted approximately 60 minutes, and each participant received a reward of 15 dollars.

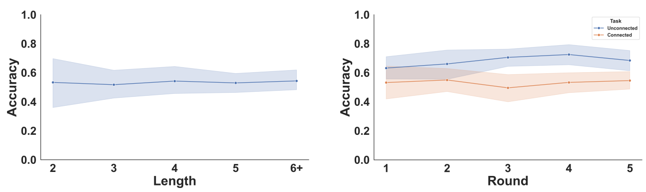
7.3 Results

Table 2: Word accuracy of different experiment setup.

Method	Top1 Accuracy	Top3 Accuracy	Top5 Accuracy
Unconnected Writing	68.16%(0.14)	79.96%(0.13)	82.08%(0.12)
Connected Writing	53.14%(0.15)	63.99%(0.14)	67.17%(0.14)
Unconnected Writing (Vocabulary Size 3000)	84.36%(0.17)	90.24%(0.13)	91.83%(0.13)
Connected Writing (Vocabulary Size 3000)	74.17%(0.12)	80.51%(0.10)	81.78%(0.09)

7.3.1 Accuracy. All accuracy results are shown in Table 2. In the unconnected writing setting, the average top-1 accuracy reached 68.16%. For comparison, the accuracy when performing the same task using a touchpad was 71.4% (using data from the data collection experiment). The connected writing setting requires users to continuously write a complete word, which allows us to assess the performance of the WritingRing algorithm on longer sequences. In this setting, the accuracy achieved was 53.14%.

Since Google IME does not include explicit auto-correction, it outputs any combination of letters. However, in certain scenarios, only words within the vocabulary need to be output. Therefore, we also calculated the accuracy when only words from the vocabulary (defined as the 3,000 most common words) are considered. When the output from Google IME is found in the vocabulary, it remains unchanged; if it is not in the vocabulary, the DTW method is used to find the word in the vocabulary that is closest in letter distance (the definition of letter distance is provided in Appendix Figure 14). With this setup, unconnected writing achieved an accuracy of 84.36%, while connected writing achieved 74.17%, both showing significant improvements. We believe that the recognition performance could be further improved if a dedicated recognition algorithm were trained specifically for the trajectories generated by WritingRing.



(a) Accuracy across different word lengths under connected writing.

(b) Accuracy across different rounds of tasks.

Figure 8: The accuracy statistics based on word length or writing rounds.

7.3.2 The Influence of Input Sequence Length. To further evaluate the performance of the WritingRing algorithm on long sequences, we analyzed the relationship between word length and accuracy under the connected writing setting, as shown in Figure 8(a). It can be observed that the standard error was relatively larger when the word length was 2. This may be due to the presence of words in the word list, such as ti, ii, il, etc., which are difficult to write in one continuous stroke. There were no significant differences in accuracy across different word lengths($\chi^2(4, N = 12) = 1.63, p = 0.803$).

7.3.3 Learning Efficiency. As shown in Figure 8(b), there is no significant change in accuracy across different rounds, which is also confirmed by the results of the Friedman test ($\chi^2(4, N = 12) = 4.02, p = 0.404$ for unconnected task) and $\chi^2(4, N = 12) = 1.33, p = 0.856$ for connected task. This indicates that WritingRing is very easy to learn, and even with a short learning period, users can achieve good performance.

7.3.4 Subjective Indicators. We used the NASA-TLX scale and a subset of questions extracted from the SUS for subjective evaluation, with the full results shown in Figure 9. As can be seen, although unconnected writing achieved higher accuracy, it scored lower than connected writing on all SUS metrics (including willing to use, confidence, easy to use, and easy to learn). Also, both approaches received relatively high scores overall, especially connected writing, indicating the system's high usability. Users generally felt their performance was good, but at the same time, they rated the physical demand as relatively high. This is understandable, as handwriting as an input method is not ideal for long-duration use.

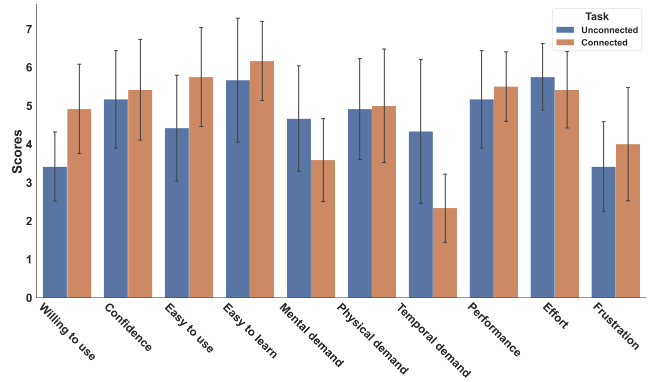


Figure 9: System scoring in the word writing evaluation study.

In addition to the scales, some users also provided positive feedback. "The system is easy to learn and easy to accept" (P10), "I didn't have much confidence at first, but I was surprised when it recognized my input" (P12). Although users found that unconnected writing had higher accuracy, "Writing letter by letter is much more accurate" (P4, P5), more users (11/12) preferred connected writing because "Writing continuously is more user-friendly" (P1), "Continuous writing is convenient and fast, and the recognition is almost as good as letter-by-letter writing" (P9). Finally, some users mentioned, "My hand gets tired after writing for a while" (P2, P4, P8, P10), which aligns with the results from the NASA-TLX scale. We believe this is due to the nature of handwriting itself. In real-world usage, users are unlikely to input words for long periods of time as they did in the experiment.

8 APPLICATION

(a) VR/AR Input.

(b) Smart Devices Control in IoT.

(c) Quick Commands on Mobile Devices.

Figure 10: Multiple application scenarios of WritingRing.

8.1 VR/AR Input

Since WritingRing interactions can occur on any surface, it is particularly well-suited for VR and AR applications. It can leverage the display interfaces of VR and AR to show handwriting trajectories, enabling users to input text on any surface. This addresses the challenge of text entry in VR and AR environments, especially for non-standard inputs like passwords, which are difficult to handle with existing methods. Even when no physical surface is available, users can use their other hand as a surface for input.

Moreover, because this input method does not rely on camera-based hand tracking, it significantly reduces power consumption related to cameras and hand tracking in current VR/AR systems. This makes WritingRing a highly efficient solution for text input in immersive environments.

8.2 Smart Devices Control in IoT

In current Internet of Things (IoT) environments, there are often many devices that, even when connected to a control terminal, require cumbersome interaction processes (such as menu navigation) to operate. WritingRing offers a more efficient solution by allowing users to quickly input commands, such as the first letter of a device or action, to execute the desired control. Since a ring is typically worn at all times, users can effortlessly control any device from anywhere in the home, reducing operational complexity. Additionally, for smart devices with displays (such as smart TVs), WritingRing can also facilitate easy text input, streamlining interactions with these devices. Since such inputs typically involve a limited number of commands, good recognition performance can be achieved by restricting the size of the vocabulary.

8.3 Quick Commands on Mobile Devices

In addition to supporting letter and word input, WritingRing, due to its trajectory-based input, can also accommodate any custom gestures by incorporating gesture classification algorithms [52]. This allows users to input letters, words, or custom gestures to perform quick actions on their mobile devices. For instance, the letter "p" or a double-tap gesture could be mapped to the "play" function.

Moreover, since interactions on mobile devices are often enumerable and context-specific, with the support of large language models (LLMs), users can express their intent with minimal letter input. For example, in a chat scenario, when responding to a yes/no question, the user could simply input the letter "Y," and the LLM could generate the rest of the response, providing additional context or reasons. This makes interactions more efficient and intuitive.

9 DISCUSSION

9.1 Writing Content Outside of the Dataset

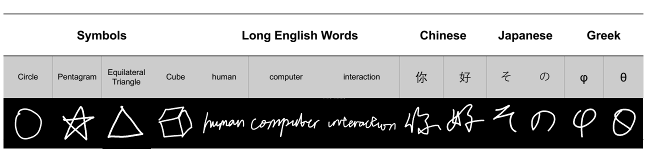


Figure 11: Handwriting cases of symbols, long words, and text in different languages, written with WritingRing.

Compared to most IMU-based handwriting recognition works that use classification algorithms, one of the main advantages of our trajectory reconstruction approach is that it is decoupled from the recognition system. This allows for the writing of content outside the dataset. We invited a participant who was not in the dataset and had not participated in the experiments to write various content, including lines, symbols, long words, and text in different languages, as shown in Figure 11.

From the figure, we can see that the participant successfully completed the writing task. Even for longer words or more complex symbols, the system achieved good performance. This capability means that users who wish to input content outside the dataset no longer need to retrain the model; they can simply integrate the corresponding handwriting recognition algorithm. For user-defined symbols, few-shot learning techniques can be employed using algorithms such as DTW or $Q [52] to adapt to new content.

9.2 Computational Efficiency

This section demonstrates that the computational overhead of our algorithm is sufficiently low, allowing it to run directly on existing ring systems. Due to the engineering complexity, we have not yet implemented the algorithm directly within the ring, but our calculations show that this is feasible.

For the trajectory reconstruction component, the model has a parameter size of 44.0 KB and requires 2.64 MFLOPs. For the touch detection component, the model's parameter size is 210.1 KB, requiring 8.37 MFLOPs. The ring's RAM capacity is 512 KB, with an additional 1 MB of SiP Flash and 8 MB of SiP PSRAM, which is sufficient to accommodate the required parameters. The ring's CPU is an ARM Cortex-M4F 32-bit, running at 96 MHz, theoretically capable of processing 8.72 frames per second. By employing techniques such as model quantization, pruning, and adding specialized compute chips, this number can be increased severalfold, making the system fully capable of meeting the demands of real-world use.

Power consumption is also a critical consideration for our system. Since users spend most of their time in a non-interactive state, we can run the touch state detection algorithm at a very low frame rate continuously. Once a press gesture is detected, the trajectory reconstruction algorithm can be activated, thus saving power. Alternatively, the touch sensor on the ring could be used as a trigger button to activate the system, further reducing power consumption. This approach ensures that the system only uses full computational resources when necessary, optimizing energy efficiency.

9.3 Personalized Calibration

We observed that different users have varying handwriting habits and finger usage styles. Although the differences between users were not particularly large in the leave-one-out experiments, these variations can still impact the overall experience to some extent. To address this, we can consider implementing personalized calibration methods when users first begin using the smart ring.

One potential solution is to have users write a short piece of text freely with their finger on their phone. During this process, we can collect IMU signal features and ground truth data (such as writing speed, stroke size, and whether they connect letters). These features can be used to explicitly adjust for user-specific factors (e.g., inferring finger length) or implicitly, by creating a personalized user embedding using a black-box model. This personalization step would help optimize the existing general model for each individual user, enhancing the overall handwriting recognition performance.

9.4 Towards Making the Use Experience More Natural

Although in this work we have placed significant emphasis on enabling natural handwriting for users, we believe there are still opportunities to make the entire system even more natural.

First, since the focus of this work is on the long-term stable reconstruction of the user's continuous handwriting trajectory and subsequent recognition, we require users to complete the writing of letters or words continuously. For native language writing, continuous writing is not typically a difficult task. However, there are certain characters that are challenging to write continuously (e.g., the English letters "i" and "j"). Due to the lack of ground truth for finger movements in the air, this work is currently unable to perfectly address this issue, which remains a limitation. Given the existing research on predicting in-air IMU trajectories, we believe that this problem can be solved in the future.

In future work, we plan to address this by incorporating hand tracking methods, such as using VR systems or motion capture systems like OptiTrack[3], to capture the ground truth of finger movements in mid-air. This will enable us to predict mid-air trajectories and create a more seamless and natural writing experience.

Meanwhile, the placement of the ring is worth discussing. Although we positioned it on the base of the index finger, which is generally the most suitable location, some users expressed a preference to wear the ring on different fingers, particularly for cultural reasons. Writing with non-index fingers is a potential task; while it may be less dexterous compared to the index finger, it could still accomplish handwriting tasks. This requires additional data to support. Different configurations (e.g., wearing varying numbers of rings on different fingers or positions) lead to varying effects. The extent to which users are willing to sacrifice recognition accuracy for a more comfortable wearing experience should be addressed in future work.

Another consideration regarding the ring's wearability is the orientation. In this study, we standardized the ring orientation for all participants to ensure consistency in the IMU's positioning. However, this may not always feel natural in everyday use. To address this, two potential solutions include inferring the gravitational direction from extended IMU data to deduce the ring's orientation or using data augmentation techniques. These issues will be addressed in our future work.

10 CONCLUSION

We propose WritingRing, a system that utilizes a single IMU ring worn at the base of the finger to reconstruct 2D fingertip trajectories in real time during natural handwriting scenarios. To the best of our knowledge, this is the first work to achieve real-time, cross-user handwriting recognition and reconstruction on the plane with high accuracy using a single IMU ring. We have built and plan to publicly release what we believe to be the largest handwriting dataset for IMU rings (N=20). On this dataset, our algorithm achieved an average trajectory reconstruction accuracy of 1.63mm. We made improvements to the original LSTM model through the use of streaming input and the addition of a TCN model, which we validated via ablation studies. We further demonstrated WritingRing's high usability in practical applications through two real-time usability evaluation studies, achieving 88.7% letter recognition accuracy and 68.2% word recognition accuracy (84.36% when restricting the output to words within a vocabulary of size 3000). Finally, we explored potential applications for WritingRing, hoping it can play a key role in various scenarios such as VR, AR, and IoT, becoming an integral part of everyday life.

[3]https://optitrack.com/

Acknowledgments

This work is supported by the National Key Research and Development Plan of China under Grant No. 2024YFB4505500 & 2024YFB450 5502, Beijing Key Lab of Networked Multimedia, Institute for Artificial Intelligence, Tsinghua University (THUAI), Beijing National Research Center for Information Science and Technology (BNRist), 2025 Key Technological Innovation Program of Ningbo City under Grant No. 2022Z080, Beijing Municipal Science and Technology Commission, Administrative Commission of Zhongguancun Science Park No.Z221100006722018, and Science and Technology Innovation Key R&D Program of Chongqing.

References

[1] Norhafizan Ahmad, Raja Ariffin Raja Ghazilla, Nazirah M. Khairi, and Vijayabaskar Kasi. 2013. Reviews on Various Inertial Measurement Unit (IMU) Sensor Applications. In *International Journal of Signal Processing Systems*. 256–262. doi:10.12720/ijsps.1.2.256-262

[2] Christoph Amma, Dirk Gehrig, and Tanja Schultz. 2010. Airwriting recognition using wearable motion sensors. In *Proceedings of the 1st Augmented Human international Conference*. 1–8.

[3] Luca Ardüser, Pascal Bissig, Philipp Brandes, and Roger Wattenhofer. 2016. Recognizing text using motion data from a smartwatch. In *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. IEEE, 1–6.

[4] Daniel Ashbrook, Patrick Baudisch, and Sean White. 2011. Nenya: subtle and eyes-free mobile input with a magnetically-tracked finger ring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2043–2046.

[5] Gaddi Blumrosen, Katsuyuki Sakuma, John Jeremy Rice, and John Knickerbocker. 2020. Back to finger-writing: Fingertip writing technology based on pressure sensing. *Ieee Access* 8 (2020), 35455–35468.

[6] Helene Brashear, Thad Starner, Paul Lukowicz, and Holger Junker. 2003. Using multiple sensors for mobile sign language recognition. In *Seventh IEEE International Symposium on Wearable Computers, 2003. Proceedings*. IEEE Computer Society, 45–45.

[7] Martin Brossard, Silvere Bonnabel, and Axel Barrau. 2020. Denoising IMU Gyroscopes with Deep Learning for Open-Loop Attitude Estimation. doi:10.48550/arXiv.2002.10718 arXiv:2002.10718 [cs, stat]

[8] Yanling Bu, Lei Xie, Yafeng Yin, Chuyu Wang, Jingyi Ning, Jiannong Cao, and Sanglu Lu. 2022. Handwriting-Assistant: Reconstructing Continuous Strokes with Millimeter-level Accuracy via Attachable Inertial Sensors. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4 (2022), 146:1–146:25. doi:10.1145/3494956

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

[10] Ke-Yu Chen, Kent Lyons, Sean White, and Shwetak Patel. 2013. uTrack: 3D input using two magnetic sensors. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*. 237–244.

[11] 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*. 1504–1514.

[12] I Arun Faisal, T Waluyo Purboyo, and A Siswo Raharjo Ansori. 2019. A review of accelerometer sensor and gyroscope sensor in IMU sensors on motion capture. *J. Eng. Appl. Sci* 15, 3 (2019), 826–829.

[13] Eric Foxlin. 1996. Inertial head-tracker sensor fusion by a complementary separate-bias Kalman filter. In *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*. IEEE, 185–194.

[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. doi:10.1145/3432204

[15] Aakar Gupta, Cheng Ji, Hui-Shyong Yeo, Aaron Quigley, and Daniel Vogel. 2019. Rotoswype: Word-gesture typing using a ring. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–12.

[16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770–778.

[17] S Hochreiter. 1997. Long Short-term Memory. *Neural Computation MIT-Press* (1997).

[18] Krzysztof Jaskot and Artur Babiarz. 2010. The inertial measurement unit for detection of position. *Przeglad Elektrotechniczny* 86, 11A (2010), 323–333.

[19] Josh Kaufman. 2021. google-10000-english. https://github.com/first20hours/google-10000-english.

[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*. 157–160.

[21] Weisheng Kong, Salvatore Sessa, Sarah Cosentino, Massimiliano Zecca, Kohei Saito, Chunbao Wang, Usama Imtiaz, Zhuohua Lin, Luca Bartolomeo, Hiroyuki Ishii, et al. 2013. Development of a real-time IMU-based motion capture system for gait rehabilitation. In *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2100–2105.

[22] Yuki Kubo. 2022. Ring-type Indirect Pointing Device for Large Displays using Three-axis Pressure Sensor. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*. 1–2.

[23] Colin Lea, Michael D Flynn, Rene Vidal, Austin Reiter, and Gregory D Hager. 2017. Temporal convolutional networks for action segmentation and detection. In *proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 156–165.

[24] Seong Kyu Leem, Faheem Khan, and Sung Ho Cho. 2019. Detecting mid-air gestures for digit writing with radio sensors and a CNN. *IEEE Transactions on Instrumentation and Measurement* 69, 4 (2019), 1066–1081.

[25] Tianyu Li, Yue Liu, Shining Ma, Mingwei Hu, Tong Liu, and Weitao Song. 2022. NailRing: An Intelligent Ring for Recognizing Micro-gestures in Mixed Reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 178–186.

[26] Yujia Li, Kaisheng Yao, and Geoffrey Zweig. 2015. Feedback-based handwriting recognition from inertial sensor data for wearable devices. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2269–2273.

[27] Xinye Lin, Yixin Chen, Xiao-Wen Chang, Xue Liu, and Xiaodong Wang. 2018. Show: Smart handwriting on watches. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 4 (2018), 1–23.

[28] Zhe-Ting Liu, Davy PY Wong, and Pai H Chou. 2020. An IMU-based wearable ring for on-surface handwriting recognition. In *2020 international symposium on VLSI design, automation and test (VLSI-DAT)*. IEEE, 1–4.

[29] Yuqi Luo, Jiang Liu, and Shigeru Shimamoto. 2021. Wearable air-writing recognition system employing dynamic time warping. In *2021 IEEE 18th annual consumer communications & networking conference (CCNC)*. IEEE, 1–6.

[30] I Scott MacKenzie and R William Soukoreff. 2003. Phrase sets for evaluating text entry techniques. In *CHI'03 extended abstracts on Human factors in computing systems*. 754–755.

[31] Sebastian Madgwick et al. 2010. An efficient orientation filter for inertial and inertial/magnetic sensor arrays. *Report x-io and University of Bristol (UK)* 25 (2010), 113–118.

[32] Alvin Marquez, Brinda Tank, Sunil Kumar Meghani, Sabbir Ahmed, and Kemal Tepe. 2017. Accurate UWB and IMU Based Indoor Localization for Autonomous Robots. In *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*. 1–4. doi:10.1109/CCECE.2017.7946751

[33] Robert B. Miller. 1968. Response time in man-computer conversational transactions. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I* (San Francisco, California) *(AFIPS '68 (Fall, part I))*. Association for Computing Machinery, New York, NY, USA, 267–277. doi:10.1145/1476589.1476628

[34] Faraz M Mirzaei and Stergios I Roumeliotis. 2008. A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation. *IEEE transactions on robotics* 24, 5 (2008), 1143–1156.

[35] Shahriar Nirjon, Jeremy Gummeson, Dan Gelb, and Kyu-Han Kim. 2015. Typingring: A wearable ring platform for text input. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. 227–239.

[36] Ju Young Oh, Jun Lee, Joong Ho Lee, and Ji Hyung Park. 2017. Anywheretouch: Finger tracking method on arbitrary surface using nailed-mounted imu for mobile hmd. In *HCI International 2017–Posters' Extended Abstracts: 19th International Conference, HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017, Proceedings, Part I 19*. Springer, 185–191.

[37] Tse-Yu Pan, Chih-Hsuan Kuo, Hou-Tim Liu, and Min-Chun Hu. 2019. Handwriting Trajectory Reconstruction Using Low-Cost IMU. *IEEE Transactions on Emerging Topics in Computational Intelligence* 3, 3 (2019), 261–270. doi:10.1109/TETCI.2018.2803777

[38] Farshid Salemi Parizi, Eric Whitmire, and Shwetak Patel. 2019. Auraring: Precise electromagnetic finger tracking. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 4 (2019), 1–28.

[39] Raouf Rasoulzadeh and Alireza Mohammad Shahri. 2017. Accuracy improvement of a multi-MEMS inertial measurement unit by using an iterative UFIR filter. In *2017 European Navigation Conference (ENC)*. IEEE, 279–286.

[40] Jose A Rios and Elecia White. 2002. Fusion filter algorithm enhancements for a MEMS GPS/IMU. In *Proceedings of the 2002 National Technical Meeting of The Institute of Navigation*. 126–137.

[41] Gaspare Santaera, Emanuele Luberto, Alessandro Serio, Marco Gabiccini, and Antonio Bicchi. 2015. Low-Cost, Fast and Accurate Reconstruction of Robotic and Human Postures via IMU Measurements. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. 2728–2735. doi:10.1109/ICRA.2015.7139569

[42] Aparajita Saraf, Seungwhan Moon, and Andrea Madotto. 2023. A Survey of Datasets, Applications, and Models for IMU Sensor Signals. In *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*. 1–5. doi:10.1109/ICASSPW59220.2023.10193365

[43] Maximilian Schrapel, Max-Ludwig Stadler, and Michael Rohs. 2018. Pentelligence: Combining pen tip motion and writing sounds for handwritten digit recognition. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–11.

[44] Junxiao Shen, Roger Boldu, Arpit Kalla, Michael Glueck, Hemant Bhaskar Surale, and Amy Karlson. 2024. RingGesture: A Ring-Based Mid-Air Gesture Typing System Powered by a Deep-Learning Word Prediction Framework. *IEEE Transactions on Visualization and Computer Graphics* 30, 11 (Sept. 2024), 7441–7451. doi:10.1109/TVCG.2024.3456179

[45] Xiyuan Shen, Chun Yu, Xutong Wang, Chen Liang, Haozhan Chen, and Yuanchun Shi. 2024. MouseRing: Always-available Touchpad Interaction with IMU Rings. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 412, 19 pages. doi:10.1145/3613904.3642225

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

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

[48] Shashank Kumar Singh and Amrita Chaturvedi. 2023. Leveraging deep feature learning for wearable sensors based handwritten character recognition. *Biomedical Signal Processing and Control* 80 (2023), 104198.

[49] Andrea Tigrini, Simone Ranaldi, Alessandro Mengarelli, Federica Verdini, Mara Scattolini, Rami Mobarak, Sandro Fioretti, Silvia Conforto, and Laura Burattini. 2024. Automatic Handwriting Recognition with a Minimal EMG Electrodes Setup: A Preliminary Investigation. In *2024 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. IEEE, 1–6.

[50] Andrea Tigrini, Simone Ranaldi, Federica Verdini, Rami Mobarak, Mara Scattolini, Silvia Conforto, Maurizio Schmid, Laura Burattini, Ennio Gambi, Sandro Fioretti, et al. 2024. Intelligent Human–Computer Interaction: Combined Wrist and Forearm Myoelectric Signals for Handwriting Recognition. *Bioengineering* 11, 5 (2024), 458.

[51] Andrea Tigrini, Federica Verdini, Mara Scattolini, Federico Barbarossa, Laura Burattini, Micaela Morettini, Sandro Fioretti, and Alessandro Mengarelli. 2023. Handwritten digits recognition from sEMG: Electrodes location and feature selection. *IEEE Access* 11 (2023), 58006–58015.

[52] Radu-Daniel Vatavu, Lisa Anthony, and Jacob O Wobbrock. 2018. $ Q: A super-quick, articulation-invariant stroke-gesture recognizer for low-resource devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–12.

[53] John H Wall and David M Bevly. 2006. Characterization of inertial sensor measurements for navigation performance analysis. In *Proceedings of the 19th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2006)*. 2678–2685.

[54] 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. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–21.

[55] Mathias Wilhelm, Daniel Krakowczyk, and Sahin Albayrak. 2020. PeriSense: ring-based multi-finger gesture interaction utilizing capacitive proximity sensing. *Sensors* 20, 14 (2020), 3990.

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

[57] Junaid Younas, Hector Margarito, Sizhen Bian, and Paul Lukowicz. 2020. Finger air writing-movement reconstruction with low-cost imu sensor. In *MobiQuitous 2020-17th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*. 69–75.

[58] Junaid Younas, Hector Margarito, and Paul Lukowicz. 2022. Fairwrite-movement reconstruction and recognition using a low-cost imu. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE, 298–303.
[59] Jiajin Zhang, Guoying Peng, Hongyu Yang, Chao Tan, Yaqing Tan, and Hui Bai. 2023. Real-Time Finger-Writing Character Recognition via ToF Sensors on Edge Deep Learning. *Electronics* 12, 3 (2023), 685.
[60] Qian Zhang, Dong Wang, Run Zhao, Yinggang Yu, and JiaZhen Jing. 2021. Write, attend and spell: Streaming end-to-end free-style handwriting recognition using smartwatches. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–25.
[61] He Zhao and Zheyao Wang. 2011. Motion measurement using inertial sensors, ultrasonic sensors, and magnetometers with extended kalman filter for data fusion. *IEEE Sensors Journal* 12, 5 (2011), 943–953.
[62] Yingwei Zhao. 2017. Applying Time-Differenced Carrier Phase in Nondifferential GPS/IMU Tightly Coupled Navigation Systems to Improve the Positioning Performance. *IEEE Transactions on Vehicular Technology* 66, 2 (2017), 992–1003. doi:10.1109/TVT.2016.2558206
[63] Yuliang Zhao, Xianshou Ren, Chao Lian, Kunyu Han, Liming Xin, and Wen J Li. 2021. Mouse on a Ring: A Mouse Action Scheme Based on IMU and Multi-Level Decision Algorithm. *IEEE Sensors Journal* 21, 18 (2021), 20512–20520.
[64] Hao Zhou, Taiting Lu, Yilin Liu, Shijia Zhang, and Mahanth Gowda. 2022. Learning on the rings: Self-supervised 3d finger motion tracking using wearable sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–31.

A Complete Survey Questionnaire

This questionnaire asks you to imagine a new wearable device: a smart ring that you can wear on your finger. Its key feature is that it can track your fingertip's movements on any surface in real-time and accurately recognize what you are writing. We assume that this ring can seamlessly connect to any existing smart devices (such as computers, phones, smartwatches, smart home devices, etc.). Additionally, we assume that its recognition performance is highly accurate, with virtually no delay. For example, after connecting it to your smart TV, you can use it to write numbers to change channels or write letters or words to select a program.

After imagining this device and scenario, please proceed to answer the following questions:

(1) Basic Information:
- Name:
- Gender:
- Age:

(2) Do you usually wear rings? (If no, please enter 0; if yes, please enter the number of rings worn)
(3) Please rate your familiarity with smart rings (1-7, where 1 is completely unfamiliar, 7 is daily use, and 4 is heard of but not purchased).
(4) To achieve such functionality, how many smart rings are you willing to wear at most?
(5) What are the reasons for your reluctance to wear more smart rings?
(6) To achieve such functionality, on which part of the finger would you be willing to wear the ring? (Base of the finger (third phalanx), Middle part of the finger (second phalanx), Tip of the finger (first phalanx))
(7) What are the reasons for your reluctance to wear the ring on other parts of the finger?
(8) Please use your finger to randomly write a letter or a word on a nearby flat surface, assuming you are using this device for text input. After completing the writing, indicate whether your wrist was suspended in the air or resting on the surface.
(9) Do you think you needed visual feedback during your recent writing process, that is, would you like to see the trajectory of what you are writing in real time? Please note that even if the input trajectory is not displayed, this system can still output your writing results to the device.
(10) What are your reasons for wanting to see the trajectory? (If (9) answered yes)
(11) What are your reasons for not needing to see the trajectory? (If (9) answered no)
(12) Please use your imagination and describe a scenario where you hope this device could be useful (e.g., using it for quick commands on a smartphone, controlling smart light switches, teaching children how to write, etc.).

B Hardware Prototype Design

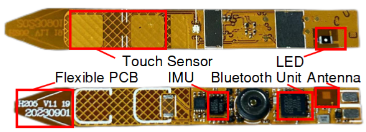


Figure 12: The hardware prototype flexible PCB used in this work.

Here is a photo of our hardware prototype's Flexible PCB. In addition to the 6-axis IMU sensor with a maximum frequency of 200 fps, it also includes two touch sensors, an LED light, a Bluetooth chip, and an antenna. The IMU model used is MPU9250, and the processing unit is an ARM Cortex-M4F 32-bit, running at 96 MHz. During use, the Flexible PCB is wrapped around a fixed-size metal inner ring, forming the IMU smart ring. The user experience of wearing this ring is nearly identical to that of a regular ring.

C The Whole Confusion Matrix of the Results in Usability Study

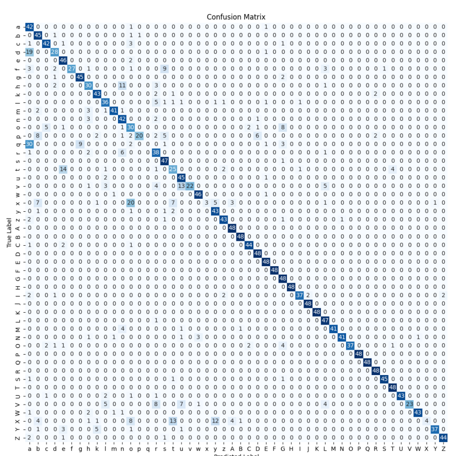


Figure 13: The confusion matrix of usability evaluation study results.

D The Letter Distance Matrix

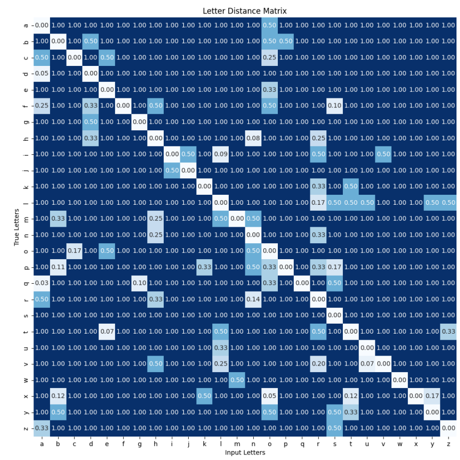


Figure 14: This matrix represents the distance between each pair of letters, with letters closer together having a smaller distance.

This matrix is mostly obtained by taking the inverse of the matrix in Figure 13, meaning that the distance between more easily confusable letters is closer, with some values adjusted.