ViWatch: Harness Vibrations for Finger Interactions with Commodity Smartwatches

Wenqiang Chen
wc5qd@virginia.edu
University of Virginia

John Stankovic
jas9f@virginia.edu
University of Virginia

ABSTRACT

Wearable devices like smartwatches and smart wristbands have gained substantial popularity in recent years. While enabling a variety of computing applications, such devices are not always convenient to interact with because of the limited size of the touchscreen. A wide variety of approaches have been considered to improve user experiences, ranging from using customized RF sensors, to multiple sensors in smartwatches. These solutions have limitations related to the characteristics of their technology. We propose ViWatch (Vibration Watch), which harnesses vibrations with an IMU sensor on commodity smartwatches to enable fine-grained finger interactions. We detect subtle finger vibrations from noise and design a novel adversarial neural network to mitigate human body variations. ViWatch is able to recognize finger typing and writing induced vibrations with accuracy, even when users are in different states of motion or in noisy environments.

CCS CONCEPTS

• **Human-centered computing** → **Text input**.

KEYWORDS

On-body interface, Finger activity recognition, Vibration sensing

ACM Reference Format:
Wenqiang Chen and John Stankovic. 2022. ViWatch: Harness Vibrations for Finger Interactions with Commodity Smartwatches. In *13th ACM Workshop on Wireless of the Students, by the Students, and for the Students Workshop (S3'22), October 17, 2022, Sydney, NSW, Australia.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3556563.3558532

Figure 1: Harness vibrations for finger tapping and writing systems with commodity smartwatches.

1 INTRODUCTION

Smart wearable devices, such as smartwatches, smart wristbands, and smart glasses, refer to what can be directly worn on the body or integrated into clothes and accessories, often interacting with the cloud through software support. With the miniaturization of computer chips, the improvement of circuit adaptability, and the popularization of mobile networks, wearable devices have become pervasive in the industry and are promising computing platforms. With the significant increase in the availability of smartwatches in 2012, more and more wearable devices were launched, gradually forming an independent industry, and the speed of development is amazing. Tractica predicts that the shipment of wearable devices will rise from 118 million in 2016 to 430 million in 2022.

Although wearables have huge potential value and application prospects, wearable devices are emerging as new products, and many new problems have arisen that affect their large-scale use. Among them, poor user experience caused by weak interactive functions has become one of the main problems. For example, products such as smart wristbands do not have touch screens, and small screens such as those on smartwatches result in poor performance of traditional touch sensing methods. This makes human-computer interaction applications extremely dependent on mobile phones and computers. Therefore, an effective interaction scheme will play a huge role in promoting the popularization of smart wearable devices.

The core idea of the current human-computer interaction approaches for wearable devices is to use various technical schemes to realize effective information input through intelligent perception. In addition to traditional speech recognition and image processing technologies, in recent years, researchers have also used a series of technologies such as radio frequency signals, acoustic signals, and inertial measurement units (IMUs) to achieve human-computer interaction. These solutions all have limitations related to the characteristics of

S3'22, October 17, 2022, Sydney, NSW, Australia

ACM ISBN 978-1-4503-9526-7/22/10.
https://doi.org/10.1145/3556563.3558532

their technology. For example, the performance of a scheme based on radio frequency signals (RF) is greatly affected by changes in the environment, and RF signals require a large signal receiving and sending end units; Acoustic sensing requires multiple pairs of microphones and speakers, and the performance is strongly affected by the placement positions and ambient noise. These problems make the application scenarios significantly restricted.

We propose ViWatch (Vibration Watch), which harnesses vibrations for fine-grained finger interactions using a single Inertial Measurement Unit (IMU) sensor on commodity smartwatches. On the one hand, we looked at the body areas surrounding the smartwatch as a potential input surface, specifically the skin on the back of the hand. Tapping on different regions of the back of the hand generates unique vibrations, and these regions (e.g., hand knuckles) can serve as natural "buttons" for tapping interactions, as is shown in Figure 1. On the other hand, we observe that the micro finger movement causes vibrations: When users' index fingers write different numbers/letters, the finger will produce different vibrations, which propagate through the hand to the smartwatch. With a IMU in a commodity smartwatch, we can recognize the user's finger writing in the air through vibrations, as is shown in Figure 1. The tapping and writing input can be sent to many smart devices, such as smartglasses, smart TVs, and ubiquitous smart devices, through wireless connections.

However, developing these fine-grained on-body interface systems has several challenges. First, it is difficult to detect subtle, fine-grained, and distinctive vibrations from noise, such as human activities. Second, users write in the air continuously; therefore, segmenting a single number or a letter for labeling and training is complicated. Third, different users have different hand shapes, writing and tapping in various fashions. To cope with these challenges, we first detect vibrations with severe signal processing methods [4–6]. Then, inspired by the Recurrent Neural Aligner [2, 7], ViWatch classifies continuous finger writing numbers and letters. Furthermore, to make ViWatch work across different users, we build a neural network with adequate regularization to mitigate overfitting on different training users [8]. Last, we design a refinement and calibration scheme to improve the performance during users' daily usage with transfer learning and adversarial learning.

We built the finger writing and tapping systems as standalone end-to-end systems using commercial smartwatches. Our implementation achieves a real-time finger input without noticeable latency. We also evaluated the real-time version of ViWatch in real-world scenarios under various disturbances, such as smartwatch types, wearing positions of wristbands, tapping fingers, tapping strengths, writing fashions, arm orientations, the user states, temporal stability, and different environments. The accuracy shows that ViWatch is usable under various disturbances (above 90% of finger writing and above 95% of finger tapping). User experience studies show that ViWatch is accurate, robust, and user-friendly under different representative applications. We have posted two demo videos on YouTube: Finger writing (https://youtu.be/aAEPv8KJ1Jk) and finger tapping (https://youtu.be/N5-ggvy2qfI).

Notably, existing IMU-based hand tracking systems can only track coarse-grained handwriting since they require the users to move their hand in a large motion so that the smartwatch on the wrist could have enough movement for distance calculation. In order to detect fine-grained finger movement, existing systems required on-finger sensors, such as customized rings and gloves. It is notable that commodity devices are easily manufactured in large volumes and have become affordable. Thus, a commodity smartwatch to recognize fine-grained finger writing is more practical and easily-accessed than customized devices. Some existing work on hand gesture classification, which leverages an IMU in smartwatches for detecting input, only works for static coarse-grained hand gestures, but not for continuous fine-grained finger movements. ViWatch requires no cumbersome instrumentation of the hands or fingers, and works with a single IMU sensor in commodity wrist-worn devices. Furthermore, ViWatch achieves continuous micro finger movements. Additionally, our system works across different users in different states of motion or in noisy environments.

To summarize, our main contribution is: By detecting finger vibrations and modeling continuous vibrations through large user studies, we build fine-grained finger vibration interfaces using a commodity IMU sensor, thus providing a new usable interaction capability for commodity wearable devices with tiny/no touchscreens.

2 VIWATCH

2.1 Finger Activity Detection

Capturing finger activity induced vibrations on the hand is challenging. First, an IMU sensor is designed to detect large hand motions rather than weak vibrations. Also, the finger activity induced vibration captured by an IMU sensor is easily affected by ordinary human activities, which corrupts the vibration feature from finger activities. Furthermore, unlike active vibrations from modulated signals, finger movement vibrations are passive and unmodulated, and comprised of a variety of frequencies which make the feature extraction and pattern matching much more challenging. To cope with these challenges, we first sum up vibration energy from six IMU axes and use an energy-based double threshold to segment the vibration signals, and then use a General Cross Correlation (GCC)-based algorithm [9] to align the segments. Then, ViWatch utilizes a 10Hz Butterworth filter to remove the noise caused by body motion and designs activate gestures to further prevent false positives from daily activities. Additionally, we extract fusion weighted features from both the time and frequency domains, based on proposed position-sensitive points and position-relevant points [1].

2.2 Finger Activity Recognition

Although there are many machine learning algorithms, we focus on solving two Challenges: First, finger writing signals can be continuous; how do we label and train models for continuous vibration signals. Second, how do we train a deep learning model aiming for a good generalization ability across many users with different variations.

To solve these challenges, we designed a novel Gated Recurrent Unit and connectionist temporal classification (GRU-CTC) network and data augmentation scheme to allow users to write numbers and letters in both discrete and continuous fashion [3]. Then, we utilize the Siamese network for finger tapping classification based on a CNN model with adequate regularization to mitigate over-fitting on different training users.

2.3 Refinement and Calibration

Although we have designed deep neural networks for finger writing and tapping on many users as described in the previous section, these general models are only as good as their training data. If the labeled training dataset fails to cover a considerable diversity in the population, the models trained on it may encounter generalization difficulties. In this section, we demonstrate that the end user experience may be further improved via user-specific adaptation of the classification model. Inspired by online learning and domain adaptation, we designed two calibration modes to continuously improve the model by using the data generated from the new user: Auto calibration mode and User intervention mode. In the auto calibration mode, we continuously improve the model using the new user's daily usage without them noticing. However, these daily generated data have no labels. Thus, we utilized an unsupervised domain adversarial neural network (DANN) to match those human-based variations (domains). Unfortunately, it is impractical for DANN to separate hundreds of domains with cross-entropy loss because it was designed for the two-domain adaptation problem. To address this problem, we modified the DANN and optimized its domain discriminator with Siamese contrastive training. In the user intervention mode, we improve the mode by requiring the new user to provide new labeled data to fine-tune the model. Also, we implement a spell check to further improve the input accuracy.

3 DEMONSTRATION

We will play two demo videos to demonstrate our system: Finger writing (https://youtu.be/aAEPv8KJ1Jk) and finger tapping (https://youtu.be/N5-ggvy2qfI). In the video of finger writing, the video shows the screen of smart glasses, displaying the results of finger writing. The video first shows writing numbers. Then it shows letter writing for movie searching on smart glasses. In the video of finger tapping, we developed several representative exemplar applications using ViWatch as the input surface. For example, we switch slides and zoom in or zoom out the screen. Also, we built remote controls for smartglasses to switch menus, play videos, and adjust volumes. By tapping on the hand coupled with the watch on the wrist, we can play games on the TV or solely on the watch. Furthermore, a simple tap on the skin provides a shortcut to any app we need. We can also control the smartphone camera remotely: tap our fingers to take a photo, snap to take a video, and tap on the hand to switch different cameras. We can also pick up or end a phone call without interacting directly with the phone. This system does not require any initial training process or calibration before the first usage, and it is robust in real world deployment. Any user can wear the smartwatch and immediately begin using the system out of the box. As you can see in the video, ViWatch has great performance with different arm orientations and works whether users are standing, seated or lying down. Users can use it with different tapping strengths. Users can use any finger to tap on the skin and they don't need to worry if the smartwatch changes positions on the wrist. The versatility of this system even allows users to use it while walking and with a wet hand. This innovative system also works across different types of smartwatches. Besides the demo videos, we will demonstrate the system live and let the workshop attendances experience the smartwatch systems.

REFERENCES

[1] Wenqiang Chen, Lin Chen, Yandao Huang, Xinyu Zhang, Lu Wang, Rukhsana Ruby, and Kaishun Wu. 2019. Taprint: Secure text input for commodity smart wristbands. In *ACM Mobicom*. 1–16.

[2] Wenqiang Chen, Lin Chen, Meiyi Ma, Farshid Salemi Parizi, Shwetak Patel, and John Stankovic. 2021. ViFin: Harness Passive Vibration to Continuous Micro Finger Writing with a Commodity Smartwatch. *ACM IMWUT/Ubicomp* 5, 1 (2021), 1–25.

[3] Wenqiang Chen, Lin Chen, Meiyi Ma, Farshid Salemi Parizi, Patel Shwetak, and John Stankovic. 2020. Continuous micro finger writing recognition with a commodity smartwatch: demo abstract. In *ACM SenSys*. 603–604.

[4] Wenqiang Chen, Lin Chen, Kenneth Wan, and John Stankovic. 2020. A smartwatch product provides on-body tapping gestures recognition: demo abstract. In *ACM SenSys*. 589–590.

[5] Wenqiang Chen, Maoning Guan, Yandao Huang, Lu Wang, Rukhsana Ruby, Wen Hu, and Kaishun Wu. 2018. Vitype: A cost efficient on-body typing system through vibration. In *IEEE SECON*. 1–9.

[6] Wenqiang Chen, Maoning Guan, Yandao Huang, Lu Wang, Rukhsana Ruby, Wen Hu, and Kaishun Wu. 2019. A Low Latency On-Body Typing System through Single Vibration Sensor. *IEEE TMC* 19, 11 (2019), 2520–2532.

[7] Wenqiang Chen, Shupei Lin, Elizabeth Thompson, and John Stankovic. 2021. SenseCollect: We Need Efficient Ways to Collect On-body Sensor-based Human Activity Data! *ACM IMWUT/Ubicomp* 5, 3 (2021), 1–27.

[8] Wenqiang Chen, Ziqi Wang, Pengrui Quan, Zhencan Peng, Shupei Lin, Mani Srivastava, and John Stankovic. 2022. Making Vibration-based On-body Interaction Robust. In *ACM/IEEE ICCPS*. IEEE, 300–301.

[9] Kaishun Wu, Yandao Huang, Wenqiang Chen, Lin Chen, Xinyu Zhang, Lu Wang, and Rukhsana Ruby. 2020. Power saving and secure text input for commodity smart watches. *IEEE TMC* 20, 6 (2020), 2281–2296.