Making Vibration-based On-body Interaction Robust

Wenqiang Chen
University of Virginia
wc5qd@virginia.edu

Ziqi Wang
University of California, Los Angeles
wangzq312@g.ucla.edu

Pengrui Quan
University of California, Los Angeles
prquan@g.ucla.edu

Zhencan Peng, Shupei Lin
VibInt AI Limited
shupei.lin@vibint.ai

Mani Srivastava
University of California, Los Angeles
mbs@ucla.edu

John Stankovic
University of Virginia
stankovic@cs.virginia.edu

2022 ACM/IEEE 13th International Conference on Cyber-Physical Systems (ICCPS) | 978-1-6654-0967-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCPS54341.2022.00041

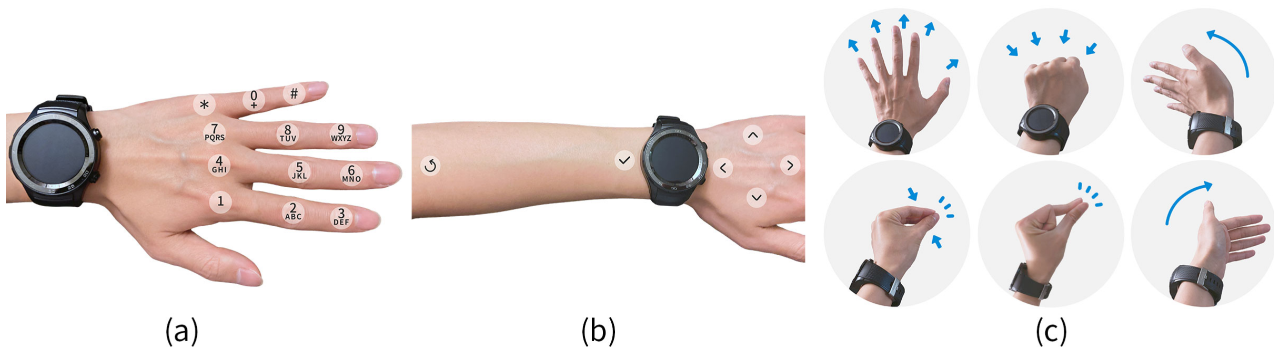

Figure 1: Three keyboards for vibration-based on-body interaction. (a) dial keyboard, (b) direction keyboard, (c) one-hand control

ABSTRACT

Wearable devices like smartwatches and smart wristbands have gained substantial popularity in recent years. However, due to the limited size of the touch screens, smartwatches typically have a poor interactive experience for users. Recently, new technology has converted the human body into a virtual interface using finger activity induced vibrations. However, these solutions fail to meet expectations during real-world deployments, e.g., system performance significantly degrades due to human-based variations, such as hand shapes, tapping forces, and device positions. To mitigate these human-based variations, we collected a dataset of 114 users, built a deep-learning model, and designed a novel Siamese domain adversarial training algorithm. In this way, we implement a robust system which works at accuracy (97%) across different hand shapes, finger activity strengths, and smartwatch positions on the wrist. We have posted a demo video on YouTube (https://youtu.be/N5-ggvy2qfI).

KEYWORDS

On-body Interaction, Vibration Sensing, Domain Adaptation, Unsupervised Adversarial Training

1 INTRODUCTION

Smart wearable devices, especially smartwatches and fitness wristbands, have become pervasive in the industry and are promising computing platforms [5, 7]. However, by necessity, a smartwatch is relatively tiny compared to traditional computing devices (e.g., laptops and smartphones). Input technologies for traditional computing devices cannot be easily replicated on wearable devices because of their size disparities. As a modern representative of wearable devices, a smartwatch uses a built-in capacitive screen as the surface for both input and output. Unfortunately, the relatively small screen size limits the richness of interactions. For example, "fat-finger" errors may not be a significant problem on smartphone screens. However, this problem is significantly exaggerated on a smartwatch.

Many researchers have sought to use the body areas surrounding the smartwatch as an extended input surface using sensors on their customized devices, such as laser, electromagnetic, camera, and vibration sensors. However, these customized devices remain as prototypes and are not widely deployed. Therefore, some researchers used commodity smartwatches to achieve on-body interaction through acoustic or ultrasonic sensors. Unfortunately, these solutions are sensitive to surrounding noise or have high power consumption. Most recently, a new on-body tapping interaction technique, Taprint [2], has gained much attention. This technique only used a single inertial measurement unit (IMU), which is already in commodity smartwatches and has relatively low power consumption [2]. This technique enables the IMU sensor to detect and recognize tapping-induced vibrations at different hand locations. Then, users can tap on different parts of the hand and use them as shortcuts to launch smartwatch apps, remote controllers for smart glasses, and joysticks to play games. Because of the efficiency and accessibility of this technology, many companies have started to build new smartwatches featuring this technology this year, such as the Apple Watch with AssistiveTouch, the Huawei Watch 3, and the Madgaze Watch.

Although this on-body tapping technology receives enthusiastic market responses, it unfortunately still has many real-world challenges [4, 6]. For example, different users have different hand shapes. To ensure the system performance, users have to provide data for initial training or calibration before the first usage, which is exhausting and not user-friendly [2, 3]. Worst of all, system performance significantly degrades in the real world when users perform finger activities with different strengths or when the smartwatch slips to a different locations on the wrist.

These challenges raise the question we try to tackle in this work: can we make this vibration-based on-body interaction system robust to those human-based variations without any personal initial training/calibration? In recent years, human-based variation problems have been investigated in IMU signal recognition of large-scale movements such as human activities [1]. However, to the best of our knowledge, there is no work that studies the human-based variation problems during the recognition of fine-grained finger movements such as tapping vibrations. Toward this end, we propose a system, namely ViWatch, to make the vibration-based on-body interaction robust for deployment in the real world.

2 VIWATCH

We classify the vibration-based on-body interaction scenarios into three categories: dial keyboard, direction keyboard, and one-hand control (see Figure 1): Figure 1 (a) maps 12 knuckles into a dial keyboard. Users can use this keyboard to dial numbers and type sentences. Figure 1 (b) has four direction "buttons" on the back of the hand and two "buttons" on the arm. This direction keyboard can control a wide variety of applications, such as playing games or switching menus. Figure 1 (c) shows six one-hand gestures. Users can open the palm or make a fist to zoom in and zoom out a car GPS map; swing the palm to the left/right to switch TV channels, music or slides; pinching three fingers to take a photo and snapping the fingers to take a video.

However, it is challenging to develop ViWatch. In order to make the system work without requiring users to collect and label initial training data before first usage, we recruited 114 volunteers to tap on various locations of their bodies. This large amount of volunteers' data covers different hand shapes, different tapping strengths, and different smartwatch positions on the wrist. With this multi-user dataset, we then designed a deep learning model to train a general model. However, this general model comes with a natural trade-off: contrary to individual models for particular users, an "average" model that works for all users covers more human-based variations, but may have a lower per-user accuracy. Further, while we have taken measures to battle over fitting, the accuracy for completely new (unseen) users may still suffer if the training data collected from volunteers is insufficient and does not cover the unseen users' data characteristics. Inspired by online learning and domain adaptation, we question whether we can continuously improve the model by using the data generated from new users' 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 separated hundreds of domains with cross-entropy loss because it was designed for two domain adaptation problem. To address this problem, we modified the DANN and optimized its domain discriminator with Siamese contrastive training. With this series of methods, we make vibration-based on-body interaction robust in the real world.

We have implemented ViWatch as a standalone application program on a commodity Android smartwatch, Huawei Watch 2. ViWatch utilizes the built-in accelerometer and gyroscope and acquires the motion readings through existing Android Wear APIs to detect the on-body tapping induced vibrations. The sampling rate through the APIs is 100 Hz. We trained the neural network models in Pytorch 1.5.1 on a desktop computer. PyTorch supports an end-to-end workflow from Python model training to Android model deployment (via the PyTorch Android API). After training the model, we implement all the components of our system including signal processing and neural network classification on a COTS smartwatch to classify the on-body tapping in real-time. To collect users' unlabeled data during daily usage for updating models, we used network socket with IP addresses to send collected data from the smartwatch to the server and send back updated models to the smartwatch. Our implementation achieves a real-time on-body tapping input without noticeable latency (0.2 seconds).

3 DEMONSTRATION

We will play a demo video to demonstrate our system (https://youtu.be/N5-ggvy2qfI). In this video, 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 hands-free. 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 will provide us with 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 is 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 video, we will demonstrate the system live on how the smartwatch interacts and controls the computer.

REFERENCES

[1] Youngjae Chang, Akhil Mathur, Anton Isopoussu, Junehwa Song, and Fahim Kawsar. 2020. A systematic study of unsupervised domain adaptation for robust human-activity recognition. *Proc. of the ACM IWMUT* 4, 1 (2020), 1–30.

[2] 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 *The 25th Annual International Conference on Mobile Computing and Networking*. 1–16.

[3] 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. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–25.

[4] Wenqiang Chen, Lin Chen, Kenneth Wan, and John Stankovic. 2020. A smartwatch product provides on-body tapping gestures recognition: demo abstract. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*. 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 *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 1–9.

[6] Wenqiang Chen, Yanming Lian, Lu Wang, Rukhsana Ruby, Wen Hu, and Kaishun Wu. 2017. Virtual keyboard for wearable wristbands. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*. 1–2.

[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! *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–27.