VRTwitch: Enabling Micro-motions in VR with Radar Sensing

Ryo Hajika
University of Auckland
Auckland, New Zealand
ryo.hajika@auckland.ac.nz

Tamil Selvan Gunasekaran
University of Auckland
Auckland, New Zealand
themastergts007@gmail.com

Alaeddin Nassani
University of Auckland
Auckland, New Zealand
alaeddin.nassani@auckland.ac.nz

Yun Suen Pai
Keio University Graduate School of Media Design
Yokohama, Japan
pai@kmd.keio.ac.jp

Mark Billinghurst
University of Auckland
Auckland, New Zealand
mark.billinghurst@auckland.ac.nz

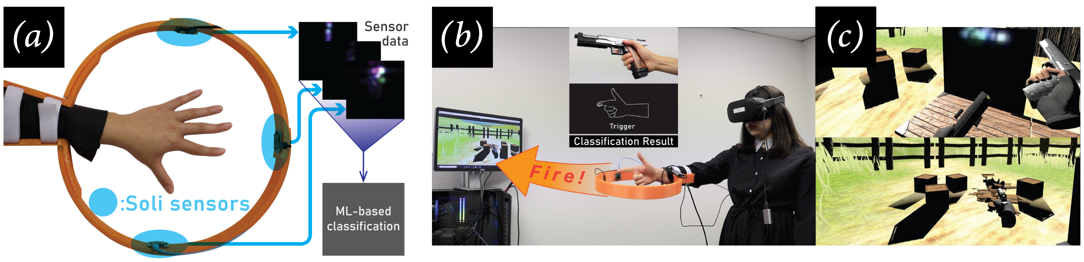

Figure 1: The VRTwitch setup; a) VRTwitch controller has three Soli sensors mounted on a 3D-printed housing, b) a player performs a micro gesture to manipulate a gun in the VR experience, c) screen captures of the demo experience

ABSTRACT

Micro-motions are often difficult to incorporate in Virtual Reality (VR) while macro-motions are a popular interaction method, due to technological limitations with VR tracking methods. In this poster, we introduce VRTwitch, a forearm-mounted wearable device that is able to sense micro hand motions. VRTwitch uses an array of reconfigurable miniaturized radar sensors placed around the hand to capture subtle finger movements for gesture detection towards enhanced interaction in VR space. We created a simple interactive VR shooting game that requires precise finger motion for virtual gun manipulation as a demonstration.

CCS CONCEPTS

• **Human-centered computing** → **Interface design prototyping**.

KEYWORDS

Project Soli, Micro-motion, Gesture recognition, Virtual Reality

ACM Reference Format:

Ryo Hajika, Tamil Selvan Gunasekaran, Alaeddin Nassani, Yun Suen Pai, and Mark Billinghurst. 2021. VRTwitch: Enabling Micro-motions in VR with Radar Sensing. In *SIGGRAPH Asia 2021 Posters (SA '21 Posters), December 14–17, 2021, Tokyo, Japan.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3476124.3488650

SA '21 Posters, December 14–17, 2021, Tokyo, Japan

ACM ISBN 978-1-4503-8687-6/21/12.
https://doi.org/10.1145/3476124.3488650

1 INTRODUCTION

User interfaces in Virtual Reality (VR) typically use motion-tracked controllers or hand-tracking for full-body interaction. However, when minute finger movements are required, VR tracking solutions are often not able to detect these motions. The current solutions for hand tracking, such as the use of Leap Motion[1] or the Oculus Quest[2] hand tracking, offer limited detection and tracking of small finger twitches. In addition, finger occlusion remains a key issue, especially for camera-based systems.

We present VRTwitch, a hand-tracking method that uses an array of Google Soli radar sensors to track hand micro-motions. Radar is able to penetrate through many materials, is robust to environmental changes and physical abuse, while requires very low power. We designed a forearm-mounted wearable device that mounts up to three radar sensors to form a radar array for precise tracking. The signal gathered from the radars that are placed on different axes are logged, trained and classified using a custom deep neural network (DNN) model.

The contributions of our work are twofold: 1) Development of a wearable device to track micro hand motion in VR with a novel radar sensor array for robust, precise and occlusion-free micro-motion tracking, and 2) A demo VR experience requires micro hand gesture input.

[1] https://www.ultraleap.com/product/leap-motion-controller/
[2] https://www.oculus.com/quest/

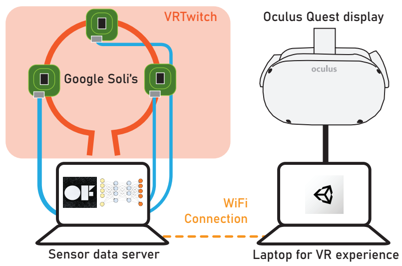

Figure 2: System Design

2 RELATED WORKS

Micro gestures can be defined as short time interruptions for primary tasks [Ashbrook 2010]. Wolf et al. [Wolf et al. 2011] defines them as gestures that do not draw significant attention, yet our work looks into gestures that are highly precise and controlled motions using small muscle groups. The closest related work is Counterpoint [Ens et al. 2018] which uses both macro and micro gestures for AR interaction. However, their system uses a single wrist mounted radar that limits gesture tracking to a single axis. By using an array of radar sensors, we significantly increase the tracking robustness to three axes, allowing finger gestures to be sensed more accurately from all directions.

3 IMPLEMENTATION

Each Soli in the sensor array is tethered to a laptop, and range-Doppler map from the sensors are streamed to a server machine by the ZeroMQ protocol that infer a gesture using a DNN model. We employed three radar sensors, each mounted at different angles and facing the user's hand. We designed a wearable device that straps to the forearm while leaving the hand completely free from contact unlike glove-based wearables, as shown in Figure 1. It overall weights only 450 grams.

A DNN model is used to recognise the micro gestures of the person wearing the twitch device. Our target gesture set consists of 6 gestures based on real life gun operation, as illustrated in Figure 3. To train the DNN model on the gesture set, we asked each participant to put on the device and perform each gesture 20 times over a period of 1 hour 20 minutes. They were also requested to occasionally remove and put on the device again to introduce more variability for model robustness. We recruited 12 participant (7 male, 5 female) with mean age 26.2. The participant had to wear the hand mount and perform each gesture displayed in the screen for 20 times. The gesture set was recorded and exported as a numpy file. After preprocessing we had a dataset of 1410 samples. A single radar sensor data stream contains range-Doppler map in a manner of 32 * 32 pixels two-dimensional image for 16 frames from 6 Channels. For gesture classification, we stitch frame-by-frame data by the three sensor inputs for the convenience in data processing. A final input to the model can be represented as 32 * 96 * 16 * 6 pixels worth of data. We referred a DNN model [Hayashi et al. 2021] that combines both a convolutional neural network (CNN) and Long-Short Term Memory network (LSTM) model for spatial and temporal information analysis. We adapted the Adam Optimizer and categorical cross entropy loss function with a learning rate of 0.001 and 80 epochs of training. For the training, the dataset we obtained was split at 80:20 ratio in random order for training and testing (282 samples for testing). After the training, the model achieved gesture classification accuracy of 92.7%, with the predicted output streamed to a VR application developed with Unity3D[3] for the interaction described later.

Figure 3: Gesture Sets with relation to a prop pistol

3.1 VR Interaction

For VRTwitch, we modelled a virtual gun in Unity3D that animates according to all the trained gestures realistically. After the virtual gun is loaded, the user will cock the hammer and toggle the safety before they can start firing. After they empty the magazine, the slide will lock in place since the gun is empty. They can then press the magazine release to drop the magazine before reloading a new one. The slide lock is then pressed to reset the slide and load the next bullet. Figure 3 illustrates the gestures that are used to perform these actions. On top of the VR shooting experience, other potential applications, including a VR-aided surgery simulation or precision machine assembling training, in which require very precise finger motions needs to be tracked.

4 CONCLUSION

We propose VRTwitch, a wearable device that supplements VR interactions with fine finger motion. As a following project, we aim to design a less obtrusive sensor mount with different placement of the sensors for better user experience. Investigation on the real-time hand posture reconstruction using multiple Soli's is also a potential future work towards precise hand and finger tracking.

REFERENCES

Daniel Lee Ashbrook. 2010. *Enabling mobile microinteractions.* Ph.D. Dissertation. Georgia Institute of Technology.

Barrett Ens, Aaron Quigley, Hui-Shyong Yeo, Pourang Irani, Thammathip Piumsomboon, and Mark Billinghurst. 2018. Counterpoint: Exploring mixed-scale gesture interaction for ar applications. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–6.

Eiji Hayashi, Jaime Lien, Nicholas Gillian, Leonardo Giusti, Dave Weber, Jin Yamanaka, Lauren Bedal, and Ivan Poupyrev. 2021. *RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor.* Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445367

Katrin Wolf, Anja Naumann, Michael Rohs, and Jörg Müller. 2011. A taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-Dependent requirements. In *IFIP conference on human-computer interaction*. Springer, 559–575.

[3] https://unity.com/