Demo: *SMART*: Screen-based Gesture Recognition on Commodity Mobile Devices

Zimo Liao[1], Zhicheng Luo[2], Qianyi Huang[2,5], Linfeng Zhang[3], Fan Wu[1], Qian Zhang[4], Yi Wang[2,5], Guihai Chen[1]

[1]Shanghai Jiao Tong University

[2]Southern University of Science and Technology [3]Tsinghua University

[4]Hong Kong University of Science and Technology [5]Peng Cheng Laboratory

{zimo_liao, wu-fan}@sjtu.edu.cn, zhichluo.chn@gmail.com, huangqy@sustech.edu.cn,

zhang-lf19@mails.tsinghua.edu.cn, qianzh@cse.ust.hk, wy@ieee.org, gchen@cs.sjtu.edu.cn

ABSTRACT

In-air gesture control extends a touch screen and enables contactless interaction, thus has become a popular research direction in the past few years. Prior work has implemented this functionality based on cameras, acoustic signals, and Wi-Fi via existing hardware on commercial devices. However, these methods have low user acceptance. Solutions based on cameras and acoustic signals raise privacy concerns, while WiFi-based solutions are vulnerable to background noise. As a result, these methods are not commercialized and recent flagship smartphones have implemented in-air gesture recognition by adding extra hardware on-board, such as mmWave radar and depth camera. The question is, can we support in-air gesture control on legacy devices without any hardware modifications?

In this demo, we design and implement *SMART*, an in-air gesture recognition system leveraging the screen and ambient light sensor (ALS), which are ordinary modalities on mobile devices. We implement *SMART* on a tablet. Results show that *SMART* can recognize 9 types of frequently used in-air gestures with an average accuracy of 96.1%.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Gestural input**.

ACM MobiCom '21, January 31–February 4, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-8342-4/22/01…$15.00

https://doi.org/10.1145/3447993.3511174

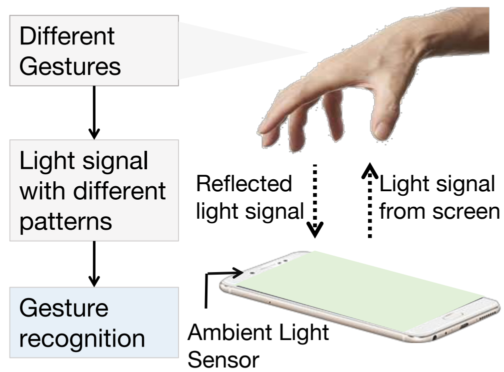


Figure 1: "Screen-Hand-ALS" light path. Light from screen is reflected by the hovering hand, and the ALS can sense the intensity of the reflected light. We analyze the received light signal and recognize different gesture.

KEYWORDS

Gesture recognition; visible light sensing; device-free;non-intrusive visible communication

1 INTRODUCTION

Gesture control is a natural and user-friendly way to interact with devices. It extends the traditional keyboard/touch screen and provides users with great freedom. In home scenarios, smart TV can be directly controlled with gestures, instead of using a remote controller; when driving, the driver can adjust the volume of music using simple gestures, which is less distracting than using touch screens or buttons. Besides, gesture control prevents our hands from physically touching any devices which may carry harmful viruses. This is of vital importance for devices in public areas.

Although prior works have implemented gesture recognition via hardware on commercial devices like cameras [5], microphones [12, 15], and Wi-Fi radios [9, 11, 14], none of them have as yet been commercialized on mobile devices. Solutions based on cameras and microphones raise privacy concerns, resulting in low user acceptance. Solutions based on Wi-Fi mainly rely on specialized NIC models (e.g., Intel 5300) and thus lack generality. Recently, several flagship smartphones have been released on the market and they are equipped with specialized hardware to support in-air gesture recognition. For example, Google Pixel 4 [3] relies on Soli [2], a 60GHz mmWave radar, to sense human gestures in the air; Huawei Mate 30 Pro [4] supports a similar functionality, but it relies on an extra depth camera on the front panel.

We observe that we can leverage the "Screen-Hand-ALS" light path to recognize hand gestures, as shown in Figure 1. When a user is performing hand gestures over the screen, the light signal transmitted from the screen is reflected by hand to the ALS on the mobile phone. The amplitude of the reflected light signal received by ALS is relative to the position of the user's hand. Thus, it is possible to infer the hand gesture through analyzing the time-series of ALS readings. Screen and ALS are both ordinary modalities on mobile devices. Thus, the solution is compatible with commercial-off-the-shelf mobile devices.

In this demo, we design and implement *SMART*, which leverages the screen on mobile devices for air gesture recognition. We design the screen update mechanism (the transmitter side) and the gesture recognition framework (the receiver side). We use a tablet as the transmitter and an ALS as the receiver.

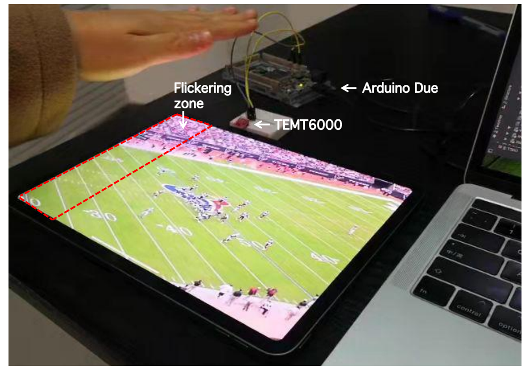
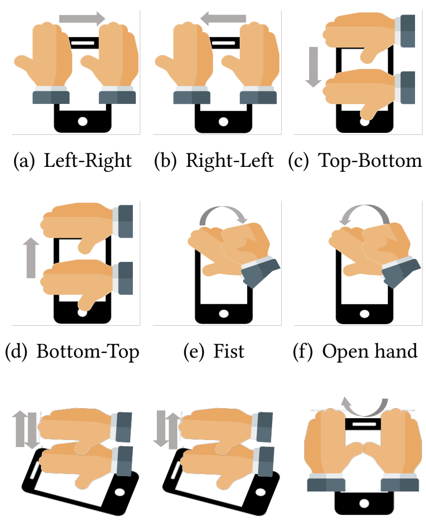


Figure 2: *SMART* Demonstration.



(a) Left-Right (b) Right-Left (c) Top-Bottom

(d) Bottom-Top (e) Fist (f) Open hand

(g) Up-Down (h) Down-Up (i) FlipPalm

Figure 3: Nine gestures of *SMART* .

2 *SMART* DESIGN

SMART leverages the "Screen-Hand-ALS" light path to implement in-air gesture control on legacy mobile devices. The design of *SMART* mainly contains the transmitter side and the receiver side.

Transmitter. *SMART* embeds spatial information into the light signal while preserving the viewing experience. *SMART* designs a mechanism based on the CIE 1931 XYZ color space to decouple the original frame into a pair of switching, complementary frames. Different from the ALS, human eyes have persistence-of-vision effect, thus the fusion result of complementary frames looks the same as the original frame. Different blinking blocks are arranged at different positions on the screen to convey spatial information. In order to overcome the restrictions of the screen refresh rate, *SMART* exploits the line-by-line screen refresh mechanism [13] to provide high-frequency signals, which are above the frequency that human eyes can perceive.

Receiver. *SMART* proposes a framework to recognize gestures from low-quality ALS data. The signal quality of the received light is poor, as the light from the screen is attenuated during propagation and reflection, while the noise level is high. We carefully segment the signal part with high signal to noise ratio (SNR) part and extract distinguishing features for the gesture recognition task. After feature extraction, *SMART* builds a lightweight classifier for gesture recognition.

3 DEMO SETUP

A demonstration scenario is shown in Figure 2. In this demo, we implement *SMART* on a commercial off-the-shelf tablet, i.e., iPad Pro with an 11-inch screen. As the operating system restrains the operation access to the screen driver, we use pre-processed videos to emulate the switching between complementary frames. The blinking blocks are positioned on one side of the screen. The width of the blinking zone is about 5cm, which can fit onto the screens of the majority of mobile phones[6]. By default, the brightness of the screen is 100%.

We use a standalone ambient light sensor (i.e., TEMT6000) as the receiver since the operating system also restricts the sampling rate of light sensors on commercial off-the-shelf devices [10]. ALS is connected to an Arduino DUE micro-controller as the receiver. We place the ALS just above the screen to emulate the relative position between screen and light sensor on commercial devices. The default sampling rate of ALS is set to 250Hz, since the integration time of most ALSs are below 4ms [1, 7, 8]. Users perform gestures at approximately 10cm above the screen.

As shown in Figure 3, *SMART* can recognize 9 types of frequently used in-air gestures (i.e., "LeftRight", "RightLeft", "TopBottom", "BottomTop", "Fist", "Openhand", "UpDown", "DownUp", "FlipPalm") under different scenarios.

ACKNOWLEDGMENTS

This work was supported in part by the Key-Area Research and Development Program of Guangdong Province No. 2020B 0101390001, in part by China NSF grant No. 62002150, 62025204, and 61972252, in part by Alibaba Group through Alibaba Innovation Research Program. The opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies or the government. Q. Huang is the corresponding author of this paper.

REFERENCES

[1] 2019. APDS-9253-001. https://docs.broadcom.com/docs/APDS-9253-001-DS
[2] 2019. Google project soli. https://www.google.com/atap/project-soli/
[3] 2021. Google pixel 4. https://store.google.com/product/pixel_4a
[4] 2021. Huawei mate30. https://consumer.huawei.com/en/phones/mate30/
[5] 2021. Leap Motion,. https://www.leapmotion.com/
[6] 2021. Smartphone unit shipments worldwide by screen size from 2018 to 2022. https://www.statista.com/statistics/684294/global-smartphone-shipments-by-screen-size/
[7] 2021. TSL2540 Ambient Light Sensor. https://ams.com/zh/tsl2540
[8] 2021. TSL2740 Ambient Light Sensor. https://ams.com/tsl2740
[9] Heba Abdelnasser, Moustafa Youssef, and Khaled A Harras. 2015. Wigest: A ubiquitous wifi-based gesture recognition system. (2015), 1472–1480.
[10] L. Li, P. Hu, C. Peng, J. Shen, and A. F. Zhao. 2014. Epsilon: A Visible Light Based Positioning System. *USENIX Association* (2014).
[11] Li Sun, Souvik Sen, Dimitrios Koutsonikolas, and Kyu-Han Kim. 2015. Widraw: Enabling hands-free drawing in the air on commodity wifi devices. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. 77–89.
[12] W. Wei, A. X. Liu, and S. Ke. 2016. Device-free gesture tracking using acoustic signals. In *International Conference on Mobile Computing & Networking*.
[13] Z. Yang, J Zhang, Z. Wang, and Z. Qian. 2018. Lightweight Display-to-device Communication Using Electromagnetic Radiation and FM Radio. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2018).
[14] N. Yu, W. Wang, A. Liu, and L. Kong. 2018. QGesture: Quantifying Gesture Distance and Direction with WiFi Signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2018).
[15] S. Yun, Y. C. Chen, H. Zheng, L. Qiu, and W. Mao. 2017. Strata: Fine-Grained Acoustic-based Device-Free Tracking. In *International Conference on Mobile Systems*.