Gesture Recognition on Mobile Devices

CS5760 : User Interface and Human-Computer Interaction

Soham Sheth, College of computing, Michigan Technological University,

Houghton ,MI USA.

ssheth1@mtu.edu

Abstract

This paper explores the air gesture recognition technology that allows users to control their smartphones or tablets without physically touching the screen by using hand gestures in the air. This technology uses a combination of hardware, such as cameras, proximity sensors, and accelerometers, and software, including machine learning algorithms, to track the user's hand movements and interpret them as commands.Previous work has implemented this functionality using cameras, acoustic signals, and Wi-Fi on existing hardware in commercial devices, but these methods have low user acceptance due to privacy concerns and vulnerability to background noise. Recent flagship smartphones have added extra hardware, such as mmWave radar and depth cameras, to support in-air gesture recognition. However, this raises the question of whether in-air gesture control can be supported on legacy devices without any hardware modifications.This paper tries to explain the alternative option without hardware modifications

Background

Touch screen human-computer interaction is a widely used technology on mobile devices that allows users to interact with their devices through physical touch on the screen. This technology uses capacitive touch sensors that detect changes in the electrical field when a user's finger or stylus comes into contact with the screen.
Touch screen human-computer interaction enables a variety of actions, including tapping, swiping, pinching, and dragging, which allows users to navigate through menus, select items, and perform actions within apps. This technology has become a ubiquitous part of modern mobile devices, enabling users to interact with their devices in an intuitive and natural way.

However the technology is becoming outdated and in general people are shifting from touch screen devices to air gesture recognition devices.

Introduction

Gesture control offers a natural and user-friendly method of interacting with devices, providing users with increased flexibility beyond traditional keyboard and touch screen input. For example, in home scenarios, a smart TV can be controlled directly with gestures, eliminating the need for a remote controller. While driving, the driver can easily adjust the volume of music with simple gestures, which is less distracting than using touch screens or buttons.

The question arises can we find any better solution without hardware modification and more security friendly ? To address this question, this work proposes an in-air gesture recognition system that leverages the screen and ambient light sensor (ALS), which are ordinary modalities on mobile devices. The system uses a screen display mechanism to embed spatial information and preserve the viewing experience on the transmitter side and a framework to recognize gestures from low-quality ALS readings on the receiver side.Overall, air gesture recognition has the potential to revolutionize the way we interact with our mobile devices, but there are still challenges that need to be overcome to make it a widely adopted technology.

In addition to convenience, gesture control also offers a hygienic advantage by eliminating the need for physical contact with devices, which can carry harmful viruses. This is especially important in public areas, such as self-service machines at airports or vending machines in shopping malls. Overall, gesture control is a promising technology that has the potential to enhance our interactions with devices in various contexts.

Gesture Control on Mobile Devices

No gesture recognition technology has yet been commercialized for mobile devices, despite earlier works having implemented it via hardware on commercial devices including cameras, microphones, and Wi-Fi radios. The privacy concerns raised by solutions based on cameras and microphones make them less popular with users.

Wi-Fi-based solutions are vulnerable to background noise, such as the movement of people or objects in the users' surroundings, because Wi-Fi signals have low spatial resolution. Additionally, Wi-Fi-based solutions lack generality because they primarily rely on specialist NIC types (such the Intel 5300).

Lately, a number of high-end smartphones hit the market, each with specialized hardware to facilitate in-air gesture recognition. For instance, the LG Q8 ThinQ has an onboard ToF camera to support in-air gesture recognition. Google Pixel 4 uses Soli, a 60GHz mmWave radar, to detect human gestures in the air. Huawei Mate 30 Pro supports a similar functionality but relies on an additional depth camera on the front panel. This begs the question: Can we support gesture recognition on old hardware without changing the hardware?

Gesture Control using Ambient Light Sensor

ALS (Ambient Light Sensor) can detect gestures on a mobile screen by measuring the changes in light intensity caused by the user's hand movements. The ambient light sensor is typically located on the front of the device, near the screen, and is used to adjust the brightness of the screen based on the surrounding lighting conditions.

When the user interacts with the screen by performing a gesture, such as tapping, swiping, or zooming, their hand interrupts the light that is emitted from the screen and reflected back to the ambient light sensor. This causes a change in the light intensity that is detected by the sensor.

The ALS captures this change in light intensity and converts it into an electrical signal. The signal is then processed by the device's software to determine the type of gesture that has been performed.

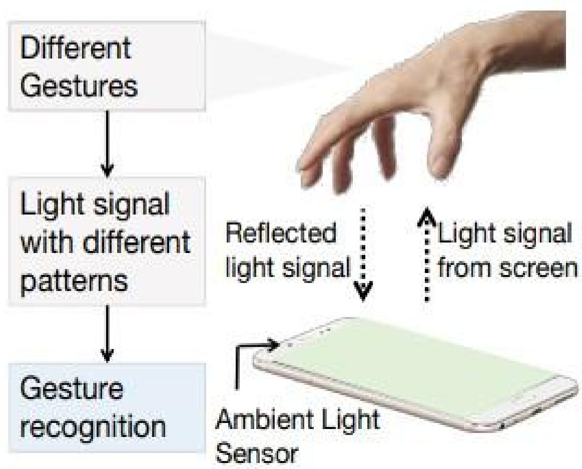

Figure 1: “Screen-Hand-ALS” light path.

(Light from the 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 gestures.)

There are several approaches that can be used to detect gestures using ALS. One common method is to use machine learning algorithms to recognize specific patterns of changes in light intensity associated with different types of gestures. The algorithms can be trained using large datasets of hand movements and corresponding light intensity measurements.

Another approach is to use computer vision techniques to analyze the video feed from the front-facing camera, which can capture the user's hand movements directly. This approach can be more accurate than ALS-based methods, but it requires more processing power and may consume more battery life.

Overall, ASL-based gesture recognition can be a useful and low-cost way to enable gesture-based interaction on mobile devices, especially for applications that do not require high accuracy or complex hand movements.
When a user makes hand gestures over the screen, the light signal emitted from the screen is reflected by the user's hand and detected by the ALS (Ambient Light Sensor)”on the mobile device.

The amplitude of the reflected light signal received by the ALS is indicative of the position of the user's hand, allowing us to analyze the time-series of ALS readings to infer the hand gesture.

Both the screen and ALS are standard modalities on mobile devices. The ALS is commonly found on mobile phones, tablets, and smartwatches, and is utilized to sense ambient light intensity and adjust screen brightness accordingly. Therefore, our proposed solution is compatible with commercially available mobile devices. Unlike cameras, ALS only captures ambient light intensity, which contains minimal sensitive information.

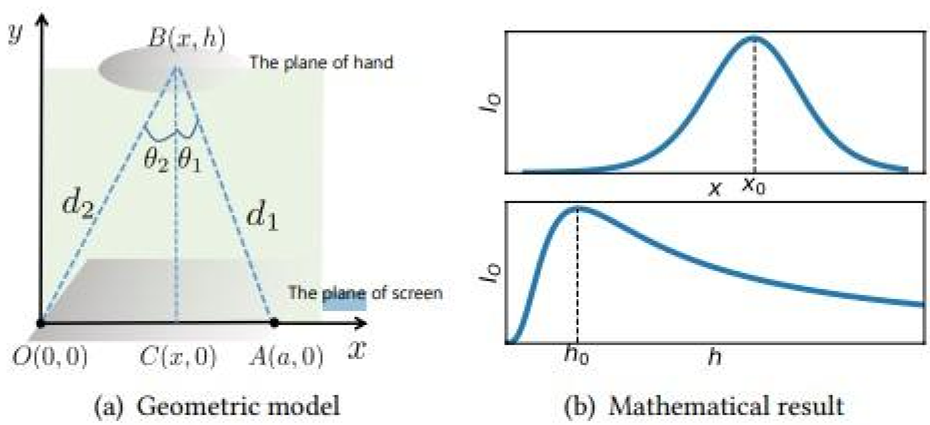

Fig 2 : Geometric Model and mathematical results

Here from the above diagram we see the geometric model where OA represents the plane of screen and, B is the plane of hand. Different signal patterns depending on the hand movement is shown in diagram 2(b) mathematical results.

Gestures Recognised :

The ambient light sensor (ALS) in a mobile device is typically located on the front of the device, usually near the top. The sensor consists of a photodiode that measures the amount of ambient light entering the sensor. This data is then processed by the device's software to adjust the display's brightness and color temperature to match the surrounding lighting conditions.

To recognize air gestures using ALS in a mobile device, a separate transmitter and receiver system is not required. Instead, the device's ALS can be used in combination with other sensors, such as cameras and infrared sensors, to capture the necessary data to recognize and classify hand air gestures.

The ALS in the mobile device can detect changes in ambient light levels caused by hand movements, such as waving or swiping, which can be used as a trigger to initiate the gesture recognition algorithms. The data captured by the ALS is then combined with data from other sensors to recognize and classify the hand air gestures.

The gesture recognition algorithms in the device's software can be trained using machine learning techniques, which involve training the software to recognize patterns in the data captured by the sensors that correspond to specific gestures. Once trained, the software can accurately recognize and classify hand gestures based on the data captured by the sensors.

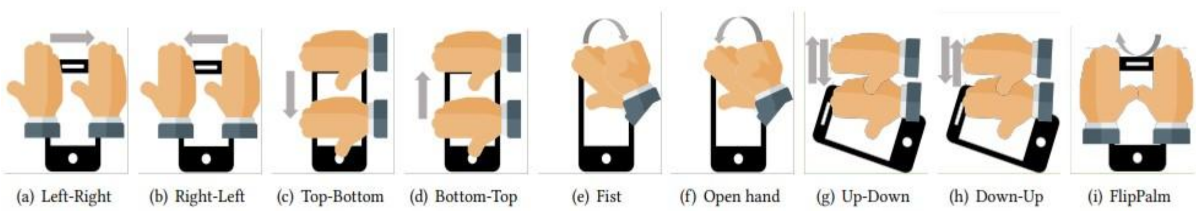

Fig 3. Nine Gestures of application

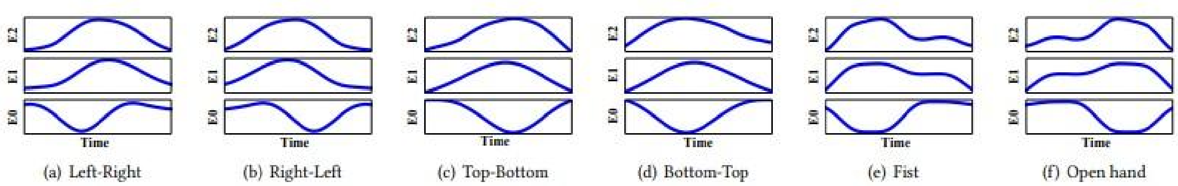

Fig 4. Signal patterns for several gestures.

Machine learning Algorithms

Here are several machine learning algorithms that can be used to train the gesture recognition algorithms for recognizing hand gestures captured by sensors in mobile devices. Some commonly used machine learning algorithms include:

1. Convolutional Neural Networks (CNNs): These are commonly used for image recognition tasks, and can be applied to hand gesture recognition by processing the image data captured by cameras or other sensors.
2. Recurrent Neural Networks (RNNs): These are useful for processing sequential data, such as time-series data captured by sensors, and can be used for recognizing gesture patterns over time.
3. Support Vector Machines (SVMs): These are used for classification tasks and can be applied to gesture recognition by classifying hand gestures based on features extracted from sensor data
4. Hidden Markov Models (HMMs): These are used for modeling sequential data and can be used for gesture recognition by modeling the temporal structure of hand gesture patterns captured by sensors.
5. K-Nearest Neighbors Algorithm(KNN): KNN can be used as a classification algorithm to classify different gestures based on their features. The KNN algorithm works by finding the k nearest neighbors of a new data point based on their distance, and then assigning the label of the majority of those neighbors to the new data point.

The choice of algorithm depends on several factors, such as the type of sensor data being used, the complexity of the hand gestures being recognized, and the computational resources available on the mobile device.

Evaluation metrics

The Evaluation metrics are used to evaluate the performance of the machine learning classifier. Previous studies use different evaluation metrics to evaluate the different aspects of the classifier. Some of the popular examples of evaluation metrics include Accuracy, Precision, Recall, and F1-Score. Each of these evaluation metrics can be calculated using True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) from the confusion metrics. The accuracy measures the fraction of correct predictions over total predictions. The fraction of correctly predicted positive overall positives is called Precision, and Recall is the proportion of positive instances correctly identified by the classifier. Whereas F1-Score is the harmonic mean of Precision and Recall, i.e., the F1-score will be high only if the classifier has high Precision and Recall.

$$Accuarcy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN}$$

Conclusion :

We can support gesture recognition on old hardware without changing the hardware with the help of Ambient Light Sensor (ALS) and some machine learning algorithms. Instead of using costly hardware devices and specifications without much modifications we can surely do the gesture recognitions but there is more experimental work required for finalizing and implementing this technology in our daily usage.Some required work is mentioned below.

To recognize and classify hand gestures, we need to decide which machine learning algorithm we need to use. Machine learning algorithms such as Convolutional Neural Networks (CNNs), K-Nearest Neighbor (KNN), Recurrent Neural Networks (RNNs), Support Vector Machines (SVMs), and Hidden Markov Models (HMMs) can be trained using the data captured by sensors. The choice of algorithm depends on various factors, such as the type of sensor data, the complexity of the gestures, and the computational resources available on the mobile device.

Overall, the combination of sensor data and machine learning algorithms has enabled mobile devices to accurately recognize and classify hand air gestures, allowing for more intuitive and efficient interaction with mobile devices without any modification in the hardware part.

Future Scope :

There is still significant work that can be done in the area of hand gesture recognition using sensors in mobile devices. Here are some potential areas of research:

1. Improving accuracy: While the accuracy of hand gesture recognition systems has improved significantly in recent years, there is still room for improvement. Researchers could explore new machine learning algorithms, data processing techniques, or feature extraction methods to improve the accuracy of recognition systems.
2. Real-time processing: Real-time processing is essential for many applications of hand gesture recognition, such as gaming or virtual reality. Researchers could explore techniques to optimize processing speed and reduce latency, such as hardware acceleration or more efficient algorithms.
3. User experience: Hand gesture recognition has the potential to provide a more natural and intuitive way of interacting with mobile devices, but it is important to ensure that users find the experience comfortable and enjoyable. Researchers could explore techniques to improve the ergonomics of hand gestures or to provide haptic feedback to users.
4. Generalizability: Many current hand gesture recognition systems are designed to recognize specific sets of gestures. Researchers could explore techniques to make these systems more generalizable, for example by training them on larger datasets or by developing transfer learning techniques.

Overall, there is significant potential for research in the area of hand gesture recognition using sensors in mobile devices. By addressing the limitations of current techniques and developing new approaches, researchers can create more accurate, robust, and user-friendly gesture recognition systems that enable more intuitive and efficient interaction with mobile devices without the modification in the hardware.

References :

[1]2021. Zimo Liao, Zhicheng Luo, Qianyi Huang, Linfeng Zhang, Fan Wu, Qian Zhang, and Yi Wang. SMART: screen-based gesture recognition on commodity mobile devices.
[2] 2018. Introduction and use of Android Framebuffer. https://kaustav.space/2018-09-22/introduction-to-android-framebuffer
[3] 2019. APDS-9253-001. https://docs.broadcom.com/docs/APDS-9253-001-DS
[4] 2019. Google project soli. https://www.google.com/atap/project-soli/
[5] 2021. CIE 1931 color space. https://en.wikipedia.org/wiki/CIE_1931_color_space
[6] 2021. Google pixel 4. https://store.google.com/product/pixel_4a
[7] 2021. Huawei mate30. https://consumer.huawei.com/en/phones/mate30/
[8] 2021. Leap Motion,. https://www.leapmotion.com/
[9] 2021. LG G8 ThinQ. https://www.lg.com/us/cell-phones/lg-LMG820UM1-attg8-thinq
[10] 2021. Measuring Device Power. https://source.android.com/devices/tech/power/ device
[11] 2021. Smartphone unit shipments worldwide by screen size from 2018 to 2022. https://www.statista.com/statistics/684294/global-smartphone-shipmentsby-screen-size/
[12] 2021. TSL2540 Ambient Light Sensor. https://ams.com/zh/tsl2540
[13] 2021. TSL2740 Ambient Light Sensor. https://ams.com/tsl2740
[14] Heba Abdelnasser, Moustafa Youssef, and Khaled A Harras. 2015. Wigest: A ubiquitous wifi-based gesture recognition system. (2015), 1472–1480.
[15] J. R. Barry. 1997. Wireless Infrared Communications. PROCEEDINGS- IEEE 85, 2 (1997), 265–298.
[16] Sidhant Gupta, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. Soundwave: using the doppler effect to sense gestures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 1911–1914.
[17] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, and Andrew W. Fitzgibbon. 2011. KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In Acm Symposium on User Interface Software & Technology.
[18] M. Izz, Z. Li, H. Liu, Y. Chen, and F. Li. 2016. Uber-in-light: Unobtrusive visible light communication leveraging complementary color channel. In IEEE INFOCOM 2016.
[19] B. Kellogg, V. Talla, and S. Gollakota. 2014. Bringing gesture recognition to all devices. In Usenix Conference on Networked Systems Design & Implementation.
[20] L. Li, P. Hu, C. Peng, J. Shen, and A. F. Zhao. 2014. Epsilon: A Visible Light Based Positioning System. USENIX Association (2014).
[21] Tianxing Li, Chuankai An, Zhao Tian, Andrew T Campbell, and Xia Zhou. 2015. Human sensing using visible light communication. In Proceedings of the 21st Annual International Conference on Mobile Computing and Networking. 331–344.
[22] Tianxing Li, Chuankai An, Xinran Xiao, Andrew T Campbell, and Xia Zhou. 2015. Real-time screen-camera communication behind any scene. In Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services. 197–211.

[23] T. Li, Q. Liu, and X. Zhou. 2017. Practical Human Sensing in the Light. GetMobile: Mobile Computing and Communications 20, 4 (2017), 28–33.
[24] Tianxing Li, Xi Xiong, Yifei Xie, George Hito, Xing-Dong Yang, and Xia Zhou. 2017. Reconstructing hand poses using visible light. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1, 3 (2017), 1–20.
[25] Yichen Li, Tianxing Li, Ruchir A Patel, Xing-Dong Yang, and Xia Zhou. 2018. Self-powered gesture recognition with ambient light. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology. 595–608.
[26] Dong Ma, Guohao Lan, Mahbub Hassan, Wen Hu, Mushfika B Upama, Ashraf Uddin, and Moustafa Youssef. 2019. Solargest: Ubiquitous and battery-free gesture recognition using solar cells. In The 25th Annual International Conference on Mobile Computing and Networking. 1–15