Expanding One-Handed Input Vocabulary for Smartphone Using In-Air Gesture of Index Finger Captured by Rear Camera

Yusuke Sei
University of Tsukuba
Tsukuba, Ibaraki, Japan
sei@iplab.cs.tsukuba.ac.jp

Buntarou Shizuki
University of Tsukuba
Tsukuba, Ibaraki, Japan
shizuki@cs.tsukuba.ac.jp

ABSTRACT

This paper presents an in-air gestural input method using the index finger of the hand gripping a smartphone. Our method allows the user to interact with the smartphone through a bending gesture and swiping gesture performed around the back of the device. Therefore, the user can interact with the smartphone without occluding the screen during one-handed use. Moreover, since our method uses the built-in rear camera, it can be easily introduced to off-the-shelf smartphones without additional sensors. To demonstrate our method, we built a prototype system using an off-the-shelf smartphone. In a pilot study to measure the gesture detection accuracy of our system, we obtained the total detection accuracy of 66.17%.

CCS CONCEPTS

• **Human-centered computing → Gestural input**; **Interactive systems and tools**.

KEYWORDS

one-handed interaction, back of device interaction, gesture recognition, mobile device

ACM Reference Format:
Yusuke Sei and Buntarou Shizuki. 2021. Expanding One-Handed Input Vocabulary for Smartphone Using In-Air Gesture of Index Finger Captured by Rear Camera. In *Asian CHI Symposium 2021 (Asian CHI Symposium 2021), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3429360.3468181

1 INTRODUCTION

Back-of-device (BoD) interaction, in which the back of a mobile device is used as an interaction space, is one approach to enrich mobile interaction; expanding the interaction space to the back enables a user to operate a smartphone without occluding the screen with the finger during one-handed use. Previous techniques enabling BoD interactions used built-in sensors [5, 7, 8, 10] or customized devices [1, 7, 9]. We also previously proposed two similar types of input methods. One is based on a combination of index finger gestures performed around the back of the smartphone and thumb touch gestures performed on the screen [3]. The other relies on gestures of the index finger on and around the back of the smartphone [4]. These methods expand one-handed input vocabulary by exploiting the three-dimensional space on and above the back of a device for index finger gestures. However, the methods require the user to wear a ring-like device to track the index finger.

In this paper, we present an in-air gestural input method using the index finger of the hand gripping the smartphone. Our method allows the user to interact with the smartphone through index finger gestures performed around the back of the device (Fig. 1). Therefore, the user can interact with the smartphone without occluding the screen during one-handed use. In our method, index finger gestures are detected through the built-in rear camera on the smartphone, following methods in previous studies [2, 6] for detecting hand gestures performed around the smartphone. Since our method uses the built-in rear camera, it can be easily introduced to off-the-shelf smartphones without additional sensors.

Our method expands the smartphone input vocabulary by adding a bending gesture (Fig. 1a) and a swiping gesture (Fig. 1b), performed using the index finger of the hand gripping the smartphone. In our method, it is necessary to distinguish in-air gestures of the index finger from incidental hand movements. Therefore, we use the posture keeping the index finger still in the air for 300 ms captured by the rear camera as a trigger of gesture detection. With this design, the method can be used simply by extending the index finger in the air around the rear camera and then performing in-air gestures.

As one application, our method serves as a shortcut to launch smartphone applications (Fig. 2). The user can perform a bending gesture to take a screenshot (Fig. 2a) and a swiping gesture to switch on/off the rear flashlight (Fig. 2b). While the user generally employs some operations including touch gestures to launch these applications, shortcuts using in-air gestures can be performed without occluding the screen. Therefore, the user can launch these applications while browsing content such as maps and documents.

2 IMPLEMENTATION

We built a prototype system using a Samsung Galaxy S7 edge (150.8 × 72.6 × 8.3 mm) to demonstrate our method. The rear camera of this smartphone is located in the center of the upper part of the back. The system detects in-air gestures of the index finger from RGB images captured by the rear camera.

2.1 Skin Color Segmentation

Skin color segmentation is performed to detect the area of the index finger in the RGB image. For the segmentation, we used the algorithm of Song et al. [6] to obtain a binary image with the index finger shape.

Asian CHI Symposium 2021, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8203-8/21/05.
https://doi.org/10.1145/3429360.3468181

Figure 1: Overview of our method, which provides input vocabulary through in-air gestures of the index finger around the back of a smartphone. The user can perform a) a bending gesture and b) a swiping gesture to interact with the smartphone.

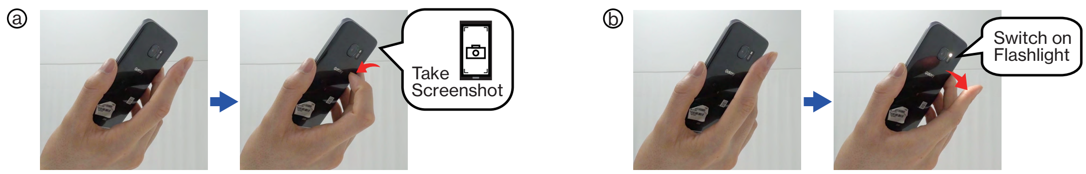

Figure 2: One application: shortcuts to launch smartphone applications. The user can a) take a screenshot using a bending gesture and b) switch on/off a flashlight on the back of the smartphone using a swiping gesture.

2.2 Detection of Gestures

To detect gestures, we focused on the tilt and movement of the index finger in the binary image (Fig. 3). The tilt of the index finger can be acquired as the inclination of a straight line L fitted to white pixels using least squares. In addition, to acquire the movement of the index finger, we used the center of gravity of white pixels. Our system acquires the center of gravity C before a gesture is performed, and C' while a gesture is performed. To determine whether a gesture was performed, our system uses the speed of C'.

To classify two gestures, we used L and V. L is acquired before a gesture was performed. V is the vector from C to C'. Our system uses the smaller angle θ produced by L and V. Since θ of the bending gesture (Fig. 3a) is smaller than that of the swiping gesture (Fig. 3b), our system detects gestures based on the threshold of θ that can be adjusted by the user.

3 PILOT STUDY

We conducted a pilot study to measure the gesture detection accuracy of our system. Four participants (three males; 22–24 years of age, M = 22.8; one left-handed) were recruited. We asked participants to perform the bending gestures and the swiping gestures using the index finger after extending it in the air while gripping the smartphone with their dominant hand. During the task, the smartphone vibrated to inform participants of the start of gesture detection. Participants performed the task under four lighting conditions: in the office room with the blinds down and the lights on (Indoor Light-On) or off (Indoor Light-Off); and outside in the sun (Outdoor In-Sun) or in the shade (Outdoor In-Shade) on a sunny day. One session consisted of 40 tasks (= 2 gestures × 20 repetitions) performed in a random order. Each participant performed a practice session and two evaluation sessions. In the practice session, each participant adjusted the threshold of θ. In total, we collected detection results for 1,280 tasks (= 2 gestures × 20 repetitions × 2 evaluate sessions × 4 lighting conditions × 4 participants).

Fig. 4 shows the confusion matrix for gesture detection. We considered the results as a failure if white pixels in the binary image disappeared during a gesture. The total detection accuracy was 66.17%, and the detection accuracy for each lighting condition was as follows: 76.56% for Indoor Light-On, 62.19% for Indoor Light-Off, 65.00% for Outdoor In-Sun, and 60.94% for Outdoor In-Shade.

4 DISCUSSIONS

The rate of detection failure was high (15.23%), especially under the Indoor Light-Off, Outdoor In-Sun, and Outdoor In-Shade conditions. These results suggest that the gesture detection of our system is affected by lighting conditions. However, such detection failure is not necessarily serious because the user could interact with the smartphone as intended by performing the gesture again.

We observed that many errors, including false-positives and false-negatives, were caused by the direction of the gesture and position of the index finger. For example, the system failed to detect gestures when the user performed gestures in directions different from our design, since the direction observed from the smartphone heavily depends on the grip posture. Moreover, the system failed to capture the overall movement of the gesture when the index finger was close to the edge of the camera's field of view. To solve these problems, we plan to modify our system to better inform the user of the space where gestures can be performed. One possibility is providing vibration feedback when the index finger moves into the area where the user can start a gesture. If wide-angle cameras are commonly installed in smartphones in the future, using the wide field of view would be another solution.

Our system captures the index finger gestures using the rear camera. However, the position and angle of view of the rear camera depend on the smartphone model. Therefore, the position where the user performs gestures depends on the smartphone. In the future, we will conduct a user study to evaluate the performance of our method using models other than the Samsung Galaxy S7 edge.

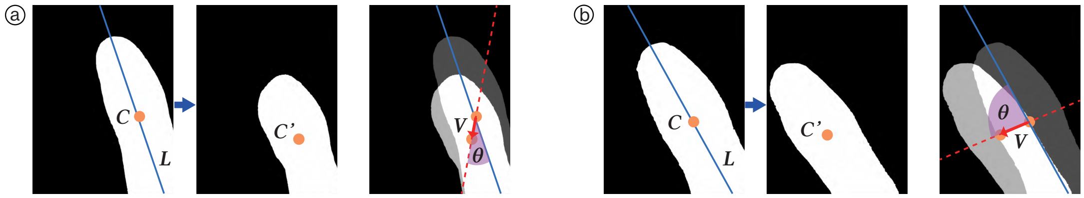

Figure 3: Binary images of a) a bending gesture and b) a swiping gesture captured by the rear camera. To detect gestures, our system acquires the fitting line L and center of gravity C before a gesture is performed, and the center of gravity C' while the gesture is performed. Our system detects gestures from the smaller angle θ between L and the vector V from C to C'.

a

	Bending	Swiping	Failure
Bending	388	124	128
Swiping	114	459	67

b

	Indoor Light-On			Indoor Light-Off			Outdoor In-Sun			Outdoor In-Shade		
	Bending	Swiping	Failure	Bending	Swiping	Failure	Bending	Swiping	Failure	Bending	Swiping	Failure
Bending	107	35	18	90	29	41	94	35	31	97	25	38
Swiping	18	138	4	32	109	19	23	114	23	41	98	21

Figure 4: Confusion matrix for gesture detection a) for all data and b) under four lighting conditions. We considered the results as a failure if white pixels in the binary image disappeared during a gesture.

5 CONCLUSION

We presented an in-air gestural input method to expand the one-handed input vocabulary of smartphones, using a bending gesture and a swiping gesture performed around the back of the smartphone. Our method allows the user to operate the smartphone without occluding the screen during one-handed use and can be easily introduced to off-the-shelf smartphones without additional sensors. To demonstrate our method, we built a prototype system that detects the gestures from images captured by the built-in rear camera. Moreover, we conducted a pilot study to measure the gesture detection accuracy and obtained a total detection accuracy of 66.17%. The results showed that the low detection accuracy was caused by lighting conditions and the user's grip posture of the smartphone. In the future, we will improve our system and conduct a user study to evaluate the performance of our method for several smartphone models.

REFERENCES

[1] Hiroyuki Hakoda, Yoshitomo Fukatsu, Buntarou Shizuki, and Jiro Tanaka. 2015. Back-of-Device Interaction Based on the Range of Motion of the Index Finger. In *OzCHI '15*. ACM, 202–206.

[2] Akira Ishii, Takuya Adachi, Keigo Shima, Shuta Nakamae, Buntarou Shizuki, and Shin Takahashi. 2017. FistPointer: Target Selection Technique Using Mid-Air Interaction for Mobile VR Environment. In *CHI EA '17*. ACM, 474.

[3] Yusuke Sei, Minto Funakoshi, and Buntarou Shizuki. 2019. Expanding One-Handed Touch Input Vocabulary Using Index Finger on and Above Back-of-Device. In *OzCHI '19*. ACM, 585–589.

[4] Yusuke Sei, Minto Funakoshi, and Buntarou Shizuki. 2020. Expanding Input Vocabulary Using Index Finger on and Above Back of Smartphone. In *AsianCHI '20*. ACM, 6.

[5] Karsten Seipp and Kate Devlin. 2014. BackPat: One-Handed off-Screen Patting Gestures. In *MobileHCI '14*. ACM, 77–80.

[6] Jie Song, Gábor Sörös, Fabrizio Pece, Sean Ryan Fanello, Shahram Izadi, Cem Keskin, and Otmar Hilliges. 2014. In-air Gestures Around Unmodified Mobile Devices. In *UIST '14*. ACM, 319–329.

[7] Pui Chung Wong, Hongbo Fu, and Kening Zhu. 2016. Back-Mirror: Back-of-Device One-Handed Interaction on Smartphones. In *SA '16*. ACM, 5:1–5.

[8] Xiang Xiao, Teng Han, and Jingtao Wang. 2013. LensGesture: Augmenting Mobile Interactions with Back-of-Device Finger Gestures. In *ICMI '13*. ACM, 287–294.

[9] Wataru Yamada, Hiroyuki Manabe, and Daizo Ikeda. 2018. CamTrackPoint: Camera-Based Pointing Stick Using Transmitted Light Through Finger. In *UIST '18*. ACM, 313–320.

[10] Cheng Zhang, Anhong Guo, Dingtian Zhang, Yang Li, Caleb Southern, Rosa I. Arriaga, and Gregory D. Abowd. 2016. Beyond the Touchscreen: An Exploration of Extending Interactions on Commodity Smartphones. *ACM Trans. Interact. Intell. Syst.* 6, 2, Article 16 (2016), 23 pages.