FGFlick: Augmenting Single-Finger Input Vocabulary for Smartphones with Simultaneous Finger and Gaze Flicks

Yuki Yamato(✉), Yutaro Suzuki, and Shin Takahashi

University of Tsukuba, Tsukuba, Ibaraki, Japan
{yamato,ysuzuki,shin}@iplab.cs.tsukuba.ac.jp

Abstract. *FGFlick* is an interactive technique featuring simultaneous single-finger operation and a gaze. The user flicks a smartphone and moves their gaze linearly. FGFlick thus augments the single-finger input vocabulary. As a result of the evaluation of the FGFlick gestures, we achieved success rates of 84.0%.

Keywords: FGFlick · Single-finger input vocabulary · Flick gesture · Gaze input · Interaction techniques · Mobile devices

1 Introduction

Smartphone users favor one-handed operation [4], but it causes the limited thumb reach problem—distant bezels and physical buttons are hard to reach [1]. The one-handed operation also restricts multi-finger inputs. It is necessary to increase the single-finger input vocabulary to allow one-handed operation with a stable grip.

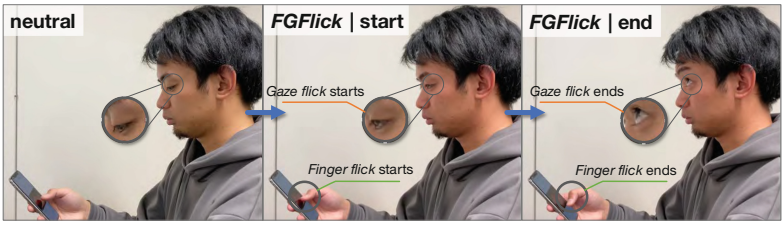

Fig. 1. FGFlick overview. The user flicks with a finger on a smartphone while he also moves his gaze linearly.

In this work, we developed *FGFlick*, an interactive technique that simultaneously uses a single-finger operation and an explicit gaze (Fig. 1). The user flicks a smartphone and moves their gaze linearly. By combining finger and gaze flicks, FGFlick augments the single-finger input vocabulary. Here, we briefly describe the design and implementation of FGFlick.

Published by Springer Nature Switzerland AG 2021
C. Ardito et al. (Eds.): INTERACT 2021, LNCS 12936, pp. 421–425, 2021.
https://doi.org/10.1007/978-3-030-85607-6_50

2 Related Work

Previous work proposed combined hand and eye movements to improve the manipulation of laptop computers and wearable devices [3,8]. The works of Pfeuffer et al. [5–7] forwarded a touch to a gazed point on a tablet; this permitted point selection and cursor manipulation around the point. They sought principally to attain unreachable areas of the screen. However, we sought to augment the input vocabulary.

The work of Elleuch et al. [2] proposed a method to interact with mobile devices using a combination of gazes and static hand gestures; we employed finger flicks instead. Wang et al. [9] developed the BlyncSync gesture set (synchronous blinks and touches); we employed simple gaze gestures.

3 FGFlick

FGFlick is an interactive technique featuring simultaneous single-finger operation and a gaze (Fig. 1). The user flicks a smartphone while moving the gaze. We have designated this gaze gesture as a "gaze flick". Because finger and gaze flicks must be performed simultaneously, they are distinguished from conventional flicks and involuntary eye movements. A gaze flick is a linear gaze. Compared with compound gazes, a gaze flick is rapid and does not require visual guidance. By combining the directions of finger and gaze flicks, FGFlick augments the single-finger input vocabulary of smartphones.

FGFlick has various interaction examples. As an alternative to a physical button, an FGFlick gesture is used to capture screenshots; the user is not required to hold the power button and the volume button simultaneously to get screenshots, as one of the traditional ways. FGFlick can also be employed to manipulate unreachable targets. We assign FGFlick gestures in the same direction to "Back" and "Open". The user can navigate back to a photo gallery from a screen displaying a selected image by executing *Flick-Right-Gaze-Right*. The user can open the Notification Center by executing *Flick-Left-Gaze-Left*.

4 Prototype Implementation

We used an iPhone X and ARKit[1] to implement eye-tracking and the FGFlick recognition system.

4.1 Recognition of FGFlick Gestures

The system starts FGFlick recognition when it detects a finger flick. Figure 2 shows the finger flick and simultaneous gaze flick. When the system detects a finger flick, it notes the start and end times. It then calculates the start and end times of a possible gaze flick: $G_{\text{start}} = F_{\text{start}} - \mathit{offset}_{\text{start}}$, and $G_{\text{end}} =$

[1] https://developer.apple.com/documentation/arkit.

- F_{start} / F_{end}: the start and the end position of the finger flick on the screen.
- G_{start}: the start position of the gaze flick before $offset_{start}$ from F_{start}.
- G_{end}: the end position of the gaze flick after $offset_{end}$ from F_{end}.
- $offset_{start}$ / $offset_{end}$: the offset between F_{start} and G_{start} and between F_{end} and G_{end}.

Fig. 2. A part of FGFlick recognition parameters.

$F_{end} + offset_{end}$. Based on preliminary experiments, we set $offset_{start}$ to 0 ms and $offset_{end}$ to 66.8 ms.

Next, the system decides whether the finger flick is an FGFlick gesture or a conventional flick, depending on whether the gaze flick is explicit or simply an involuntary eye movement. In preliminary experiments, we found that a gaze flick (an explicit gesture) was longer than involuntary eye movements during conventional flicks. Thus, we distinguish gaze flicks by the length of eye movement during the finger flick (d_{gaze}; the linear distance travelled from G_{start} to G_{end}). If the length is shorter than a threshold (d_{th}), the system regards the eye movement as involuntary and the finger flick is thus a conventional flick; otherwise, the system regards the eye movement as a gaze flick. We set d_{th} to 218 pt based on preliminary experiments.

The FGFlick gestures are classified by the directions of the finger and gaze flicks. For finger flicks, the system calculates the angle (θ_{finger}) of F_{start} to F_{end} and assigns it to one of four directions. For gaze flicks, the system calculates the angle (θ_{gaze}) of G_{start} to G_{end} and assigns it to one of four directions.

4.2 Preliminary Evaluation

We evaluated the accuracy of the FGFlick recognition system. We enlisted eight volunteers (P1–P8; mean age, 23.0 years [range, 21–25 years]; six males). We asked them to perform 20 types of FGFlick gestures, four finger flicks (up, down, left, right) multiplied by four gaze flicks (up, down, left, right) + "Keep." All volunteers grasped an iPhone X and performed each FGFlick gesture once in one session; they completed five sessions in total. We thus obtained 800 gestures: $= 8_{volunteers} \times 5_{sessions} \times 20_{gestures}$.

Figure 3 shows the confusion matrix. The success rate was 84.0%. "Down" was often mistaken for "Keep" because many volunteers positioned the device lower than the face. Thus, they initially looked down and lacked the ability to gaze "Down" further. We plan to change the "Down" distance threshold.

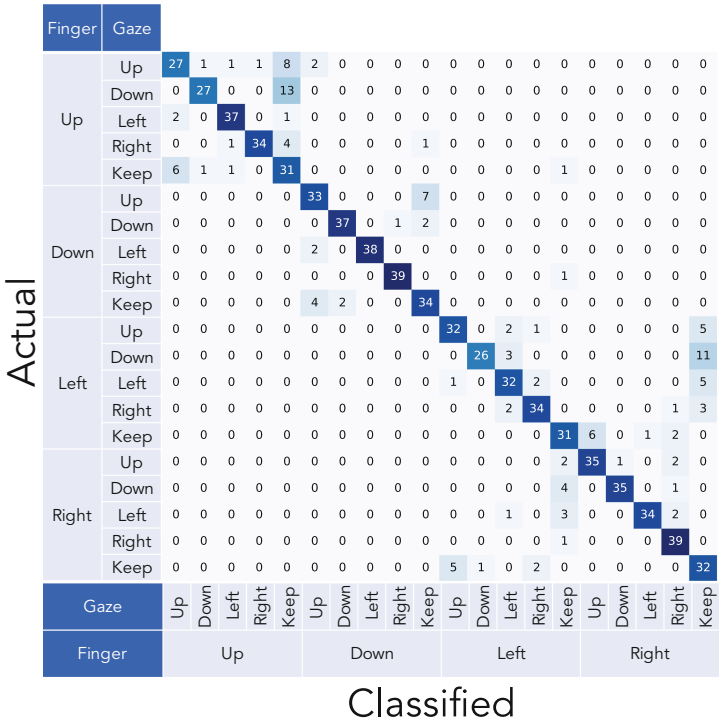

Fig. 3. The confusion matrix of the FGFlick gestures.

5 Conclusion and Future Work

FGFlick is an interactive technique featuring simultaneous single-finger operation and a gaze. The user flicks a smartphone and moves their gaze linearly. FGFlick thus augments the single-finger input vocabulary. We will further evaluate accuracy and usability, then explore whether users develop eye fatigue.

References

1. Bergstrom-Lehtovirta, J., Oulasvirta, A.: Modeling the functional area of the thumb on mobile touchscreen surfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2014, pp. 1991–2000. Association for Computing Machinery, New York (2014). https://doi.org/10.1145/2556288.2557354
2. Elleuch, H., Wali, A., Samet, A., Alimi, A.M.: Interacting with mobile devices by fusion eye and hand gestures recognition systems based on decision tree approach. In: Verikas, A., Radeva, P., Nikolaev, D.P., Zhang, W., Zhou, J. (eds.) Ninth International Conference on Machine Vision (ICMV 2016), vol. 10341, pp. 62–66. International Society for Optics and Photonics, SPIE (2017). https://doi.org/10.1117/12.2269010

3. Nagamatsu, T., Yamamoto, M., Sato, H.: MobiGaze: development of a gaze interface for handheld mobile devices. In: CHI 2010 Extended Abstracts on Human Factors in Computing Systems, CHI EA 2010, pp. 3349–3354. Association for Computing Machinery, New York (2010). https://doi.org/10.1145/1753846.1753983
4. Ng, A., Brewster, S.A., Williamson, J.H.: Investigating the effects of encumbrance on one- and two- handed interactions with mobile devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2014, pp. 1981–1990. Association for Computing Machinery, New York (2014). https://doi.org/10.1145/2556288.2557312
5. Pfeuffer, K., Alexander, J., Chong, M.K., Gellersen, H.: Gaze-touch: combining gaze with multi-touch for interaction on the same surface. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, UIST 2014, pp. 509–518. Association for Computing Machinery, New York (2014). https://doi.org/10.1145/2642918.2647397
6. Pfeuffer, K., Alexander, J., Chong, M.K., Zhang, Y., Gellersen, H.: Gaze-shifting: direct-indirect input with pen and touch modulated by gaze. In: Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology, UIST 2015, pp. 373–383. Association for Computing Machinery, New York (2015). https://doi.org/10.1145/2807442.2807460
7. Pfeuffer, K., Gellersen, H.: Gaze and touch interaction on tablets. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST 2016, pp. 301–311. Association for Computing Machinery, New York (2016). https://doi.org/10.1145/2984511.2984514
8. Rivu, S., Abdrabou, Y., Mayer, T., Pfeuffer, K., Alt, F.: GazeButton: enhancing buttons with eye gaze interactions. In: Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications, ETRA 2019, pp. 1–7. Association for Computing Machinery, New York (2019). https://doi.org/10.1145/3317956.3318154
9. Wang, B., Grossman, T.: BlyncSync: enabling multimodal smartwatch gestures with synchronous touch and blink. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI 2020, pp. 1–14. Association for Computing Machinery, New York (2020). https://doi.org/10.1145/3313831.3376132