

# スマートフォンの背面カメラを用いて認識可能な 空中人差し指ジェスチャによるショートカットの検討

清 佑輔\* 志築 文太郎†

**概要.** 我々はスマートフォン背面の空中にて実行される人差し指によるジェスチャ（空中人差し指ジェスチャ）を、アプリケーションを起動するためのショートカットに用いることを検討している。このショートカットにより、ユーザは画面上部に親指を伸ばすことなく、ステータスバーに収められている機能を実行できる。今回、我々は背面カメラの前にて実行可能な7種類の空中人差し指ジェスチャをショートカットに用いた。さらに、畳み込みニューラルネットワークを用いて空中人差し指ジェスチャを認識するシステムを試作した。



図 1. 提案手法の概要。ユーザは空中にて人差し指によるジェスチャを実行することにより、通知の表示および端末設定の変更を行える。

## 1 はじめに

スマートフォンの片手使用時にて、ユーザが把持姿勢を変えることなく親指を用いて操作できる領域は限られる [1, 6]。よって、ユーザが親指の届きにくい領域に対して操作を行う際には、把持姿勢の変更を必要とする。一方で、多くのスマートフォンにて、通知 [3] を表示する等、ユーザは親指が届きにくい画面上部に位置するステータスバーに収められている機能を実行する必要がある。その際、不安定な把持姿勢を伴う。しかしながら、多くのユーザは片手のみを用いてスマートフォンを操作する傾向がある [5]。そのため、片手使用時にこれらの機能を容易に実行できるような操作が求められる。

我々は、スマートフォン背面の空中にて実行される人差し指によるジェスチャ（以降、空中人差し指ジェスチャ）を、アプリケーションを起動するため

のショートカットとして用いることを検討している（図 1）。このショートカットにより、ユーザは画面上部に親指を伸ばすことなく、通知の表示等の、ステータスバーに収められている機能を実行できる。

本研究では、畳み込みニューラルネットワーク（Convolutional Neural Network, 以降, CNN）を用いて空中人差し指ジェスチャを認識するシステムを試作した。このシステムは、スマートフォンの背面カメラ画像から空中人差し指ジェスチャを認識する。そのため、本ショートカットは背面カメラを備えた既製のスマートフォンに容易に導入される。

## 2 関連研究

スマートフォンの背面カメラを用いて認識可能なジェスチャはこれまでも研究されている。例として、背面カメラを指にて覆うジェスチャ [11, 10] および背面に対する指によるジェスチャ [9] が挙げられる。さらに、我々のジェスチャに似ているものとして、Oh および Hong [7] の、空中にて実行される人差し指ジェスチャが挙げられる。Oh および Hong は、このジェスチャを 3 次元の UI の操作に用いた。一方で、我々は、ジェスチャをアプリケーションを起動するためのショートカットに用いる。

また、内蔵カメラを用いて認識可能な、スマートフォン周辺にて実行されるジェスチャも研究されている [8, 4, 12]。これらはスマートフォンを把持した手と異なる手のジェスチャである一方で、我々のジェスチャはスマートフォンを把持している手の指によるジェスチャである。

## 3 提案手法

我々は、図 2 に示す 7 種類の空中人差し指ジェスチャを用いることとした。これらは、スマートフォンを把持しながら背面カメラの前にて実行可能なジェスチャおよびそれを 2 回繰り返すジェスチャである。

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

\* 筑波大学情報理工学位プログラム

† 筑波大学システム情報系

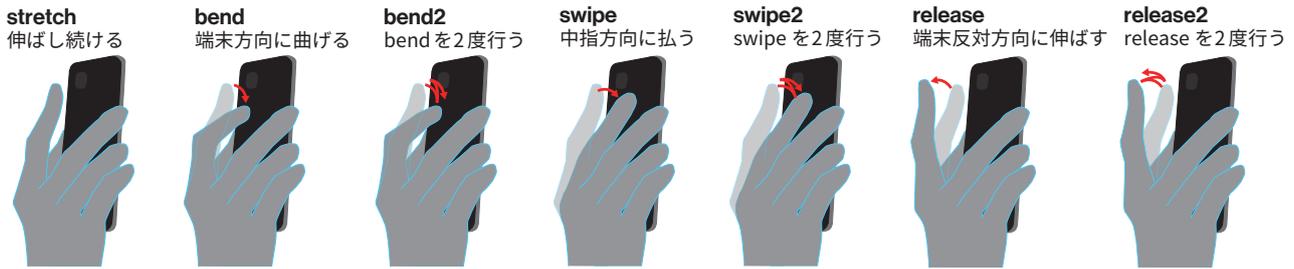


図 2. 7 種類の空中人差し指ジェスチャ.

本ショートカットの利用例として、通知の表示および端末設定の変更のショートカットを実装した(図1). 例えば、ユーザは、bendを実行することによって通知を表示できる. さらに、releaseを実行することによって Wi-Fi 接続機能の ON/OFF を切り替えられる. 本ショートカットを用いることにより、ユーザは届きにくい画面上部に親指を伸ばすことなく、これらの機能を使用できる.

#### 4 ジェスチャ認識システム

スマートフォンの背面カメラにて取得した画像から、空中人差し指ジェスチャを認識するシステムを試作した. 実装においては、スマートフォンとして Google Pixel 4 (147.1 mm × 68.8 mm × 8.2 mm, 5.7 インチ, Android 11) を用いた. 本実装は背面カメラから 240 × 320 pixels の画像を 50 fps にて取得する.

空中人差し指ジェスチャの認識は 700 ms 間にて取得した 35 枚の画像に基づく. まず、本システムは、35 枚の画像に対して肌の領域の抽出を行う. この処理には、Song ら [8] のアルゴリズムが使用される. さらに、肌の領域を抽出した画像に対してグレイスケール化の処理を行う. これらの処理を行った後、Motion History Image [2] (以下、MHI) を作製する. MHI は、図 3a に示すように、動作履歴を輝度値にて表現する画像である. 本システムでは、MHI から空中人差し指ジェスチャを認識するために、CNN を用いた認識モデルを使用する. 我々は、3 個の畳み込み層および 2 個の全結合層から成る認識モデルを作製した. このモデルは、240 × 320 pixels の MHI を入力として、認識したジェスチャの種類またはジェスチャが未実行であるかを出力する.

#### 5 予備実験

本システムの認識精度を計測するために実験を行った. 本実験では、4 人の実験参加者から取得したデータを用いて認識モデルを作製した. このデータは、参加者がジェスチャを行った 840 試行 (7 種類のジェスチャ × 30 回 × 4 人) およびスマートフォンを把持したままジェスチャを行わなかった 120 試

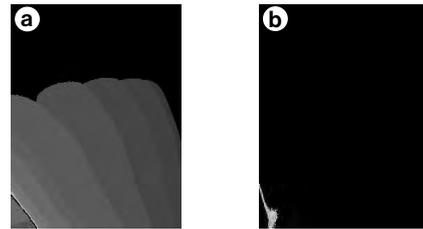


図 3. 試作システムにて作製された MHI. a: カメラの画角内にて swipe が実行された場合. b: カメラの画角外にて swipe が実行された場合.

行 (30 回 × 4 人) にて作製された MHI である. 5 分割交差検証の結果、認識精度は分割したデータ間にて、平均 81.9% (SD = 1.43%) であった.

認識精度が低い原因として、認識モデルの学習時に、指の動作をうまく捉えられていない MHI (図 3b) が正解データに含まれていたことがあげられる. これは、背面カメラの画角外にてジェスチャが実行された場合に作製される. 今後、指の動作をうまく捉えられていない MHI を学習データから取り除くことによって、認識精度の向上を試みる. そのために、画角外にてジェスチャが実行された試行を MHI の画素値から認識する処理を導入する.

#### 6 おわりに

本稿では、スマートフォン背面にて実行される空中人差し指ジェスチャを用いたショートカットを示した. また、7 種類の空中人差し指ジェスチャ、およびその利用例を示した. さらに、CNN を用いて、空中人差し指ジェスチャを認識するシステムを実装した.

ただし実験にて、実験参加者が空中人差し指ジェスチャを実行した際に、把持姿勢を変更する機会があった. この把持姿勢の変更が使いやすさへ影響を与える可能性がある. 今後は、使いやすさとして把持姿勢の安定度および実行速度に関して、本ショートカットを、画面上部からのスワイプ操作を必要とする既存操作と比較して評価する.

## 参考文献

- [1] J. Bergstrom-Lehtovirta and A. Oulasvirta. Modeling the Functional Area of the Thumb on Mobile Touchscreen Surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 1991–2000, New York, NY, USA, 2014. ACM.
- [2] A. F. Bobick and J. W. Davis. The Recognition of Human Movement Using Temporal Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):257–267, 2001.
- [3] Google LLC. Notifications Overview - Android Developers, 2021. <https://developer.android.com/guide/topics/ui/notifiers/notifications> (最終参照日: 2021年11月16日) .
- [4] A. Ishii, T. Adachi, K. Shima, S. Nakamae, B. Shizuki, and S. Takahashi. FistPointer: Target Selection Technique Using Mid-Air Interaction for Mobile VR Environment. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, p. 474, New York, NY, USA, 2017. ACM.
- [5] A. Karlson, B. Bederson, and J. Contreras-Vidal. Studies in One-Handed Mobile Design: Habit, Desire and Agility. In *Proceedings of 4th ERCIM Workshop on User Interfaces for All*, UI4ALL, pp. 1–10. Citeseer, 2006.
- [6] H. V. Le, S. Mayer, P. Bader, and N. Henze. Fingers' Range and Comfortable Area for One-Handed Smartphone Interaction Beyond the Touchscreen. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 31:1–12, New York, NY, USA, 2018. ACM.
- [7] B.-H. Oh and K.-S. Hong. Finger Gesture-based Three-Dimension Mobile User Interaction Using a Rear-facing Camera. *International Journal of Multimedia and Ubiquitous Engineering*, 8:119–130, Sept. 2013.
- [8] J. Song, G. Sörös, F. Pece, S. R. Fanello, S. Izadi, C. Keskin, and O. Hilliges. In-air Gestures Around Unmodified Mobile Devices. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, pp. 319–329, New York, NY, USA, 2014. ACM.
- [9] P. C. Wong, H. Fu, and K. Zhu. Back-Mirror: Back-of-Device One-Handed Interaction on Smartphones. In *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications*, SA '16, pp. 10:1–5, New York, NY, USA, 2016. ACM.
- [10] X. Xiao, T. Han, and J. Wang. LensGesture: Augmenting Mobile Interactions with Back-of-Device Finger Gestures. In *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*, ICMI '13, pp. 287–294, New York, NY, USA, 2013. ACM.
- [11] W. Yamada, H. Manabe, and D. Ikeda. Cam-TrackPoint: Camera-Based Pointing Stick Using Transmitted Light Through Finger. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, pp. 313–320, New York, NY, USA, 2018. ACM.
- [12] X.-D. Yang, K. Hasan, N. Bruce, and P. Irani. Surround-See: Enabling Peripheral Vision on Smartphones during Active Use. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology*, UIST '13, pp. 291–300, New York, NY, USA, 2013. ACM.