

人差し指のホバーを利用した仮名文字入力手法の提案

加藤 進吾* 國分 晴利* 志築 文太郎†

概要. 現在販売されている XR デバイスにおける文字入力手法は、手の位置の制限あるいは文字入力デバイスの把持が必要である。本研究では手首装着型デバイスを用いた人差し指のホバーによる文字入力手法を提案する。手首装着型デバイスを用いることで、文字入力の際の腕の位置が制限されず、さらに手にデバイスを持つ必要がなくなる。また、文字入力にホバーを用いることで、ユーザの移動範囲を抑えることができる。なお、スマートフォンのユーザにとって既知である 12 個のキーからなる仮名テンキーを基に文字入力システムを構築する。これにより、新規なキー配置を用いた文字入力手法と比較して、文字入力手法に慣れるまでの時間を必要としない。なお、本システムを将来的にスマートウォッチに組み込むことで、場所にとらわれず XR デバイス上で文字入力ができる。

1 はじめに

近年、拡張現実および仮想現実の市場拡大に伴い、AR グラスや MR グラスといった XR デバイスが注目を集めている。現在販売されている XR デバイスでは、視野角内に表示された仮想キーボードあるいはスマートフォンを用いて文字入力を行う。前者は視野角内に手を位置させて操作する必要があり、後者は手にデバイスを持つ必要がある。また、XR デバイス上にて使用可能な文字入力手法として、LeapMotion を用いた文字入力 [2, 5, 3] がある。LeapMotion を用いた文字入力手法は、デバイスを特定の位置に設置する必要があり、その有効範囲においてのみ使用できる。そのほか、指を用いたトグル入力およびフリック入力がある。トグル入力は 1 文字あたり 1-5 回のタップ動作が必要であり、フリック入力は手首を広範囲に動かす必要がある。

これらの課題を解決するために、手首装着型デバイスを用いた人差し指のホバーによる文字入力手法を提案する。同デバイスを用いることで、文字入力の際に腕の位置が制限されず、さらに手にデバイスを持つ必要がなくなる。また、将来的に同システムをスマートウォッチに組み込むことで、場所にとらわれず XR デバイス上で文字入力ができる。本手法はスマートウォッチを装着した手により文字入力を行うため、もう一方の手を占有しないというメリットがある。本手法ではスマートフォンのユーザにとって既知である 12 個のキーからなる仮名テンキーを基に文字入力を行う。そのため、キー配置を新規に提案する研究 [7, 8] と比較して、文字入力手法に慣れるまでの時間を必要としない。

2 関連研究

本節では、手の姿勢を認識するシステム、および指のホバーを利用した文字入力手法に関する研究を述べる。

2.1 手の姿勢を認識するシステム

手首装着型デバイスを用いて手の姿勢を認識するシステムとして、Back-Hand-Pose [6] がある。同システムは、手首装着型デバイスに搭載された RGB カメラから手の甲の画像を取得し、独自のニューラルネットワークに入力することにより 19 箇所の関節角度を回帰により推定する。なお、同システムはスマートウォッチへの搭載を想定して構築されている。本研究の文字入力手法もスマートウォッチ上での動作を想定し、かつ手の姿勢を入力に用いるため、同研究を参考にしてデバイスを作製する。一方で、本研究では仮名テンキーにおける各キーの選択を識別できれば文字入力が行えるため、回帰ではなく分類を行う。

2.2 指のホバーを利用した文字入力手法

塚田らは、指のホバーを利用した文字入力手法を提案した [9]。同手法では、ユーザは仮名テンキーにおける各キー上にて指を一定時間留めることにより子音選択を行い、その後十字形に展開される母音キーにて同操作により母音選択を行う。手の姿勢は LeapMotion を用いて認識される。一方、本手法では指のホバー時間に応じて同一キー上で母音が遷移する。また、[9] ではユーザは肩から先を動かして文字入力するが、本手法では手首を動かして文字入力を行う。これらの設計により、本手法はユーザの移動範囲を抑える。

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 1. 作製したデバイス。

3 提案手法

本節では、提案手法の実装に用いたデバイス、機械学習モデルを述べた後、本手法の操作方法を述べる。

3.1 デバイス

Back-Hand-Pose [6] を参考にして、図 1 に示す手首装着型デバイスを作製した。土台は 3D プリンタ (Ultimaker3 Extended) にて PLA フィラメントを用いて作製した。土台の指先側に 30fps で動作する 180° 魚眼カメラ (ELP FHD01M-USB CAM) を設置し、手の RGB 画像を取得できるようにした。

3.2 機械学習モデル

本手法では、仮名テンキーにおける各キーに人差し指を向けて文字入力を行うため、人差し指が向けられたキーをデバイスの取得画像から識別する必要がある。よって、識別モデルを作成するため、印刷した A2 サイズのテンキー画像を壁に貼り、前腕の位置を固定した状態で各キーに人差し指を向けた際の画像を 2000 枚ずつ収集した。なお、本手法では手の甲を上に向けた状態で、各キーに対して指のみではなく手首ごと向けて文字入力を行うため、同状態における画像を収集した。収集した 24000 枚の画像にホールドアウト法 (学習データ: テストデータ = 8:2) を用いて、機械学習を行った。Singhal らの研究 [4] では空中における各指によるタップ動作の識別精度を 4 種類の機械学習手法で比較し、深層学習による識別精度が最も高いという結果が得られた。本手法においても空中における手の姿勢を識別する必要があるため、同じように深層学習を用いる。識別には、先行研究 [6] で使用されているため ResNet18 [1] を用いた。

3.3 操作方法

作成したモデルを Unity 上で動作させ、実際にホバー入力を行う様子を図 2 に示す。予測値が低いキーは黒色で表示され、高いキーは赤色にて表示される。キーに人差し指を向け、予測値が高い状態に



図 2. ホバー入力を行う様子。「く」を入力するために「か」キーに人差し指を向け、1 秒間ホバーして表示文字を「き」、「く」と変化させた。この後に指をキー外に向けてホバーを終了することで、左側に表示された入力文字列に「く」が追加される。

て 1 秒間ホバーすることで、「か」であれば「き」、「く」、「け」、「こ」、「か」のように表示文字が変化する。指をキー外に向けた際にホバー終了とみなし、左側に表示された入力文字列に表示文字が追加される。また、指を移動する際に意図せず文字入力することを防ぐため、キーに指を向けてから 0.5 秒間は選択キーとして扱わない。そのため、あ段を入力するには 0.5 秒以上ホバーし、い段に変化する前にホバーを終了して文字入力を行う。

4 まとめと今後の予定

XR デバイスにおける文字入力手法の課題を解決するため、手首装着型デバイスを用いた人差し指のホバーによる仮名文字入力手法を提案した。現在までに文字入力システムを構築した。今後は、選択キーの有効判定および表示文字の変化までに必要な時間を調整し、ユーザビリティの向上を図る。また、画像データの増加および最適な学習済みモデルの選定により、選択キーの識別向上に努める。さらに、収集した画像データに輝度・回転加工を施してデータ拡張を行い、再度機械学習を行う。これにより、デバイスの使用環境および装着の仕方に対する汎化性能を高める。

参考文献

- [1] K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [2] K. Komiya and T. Nakajima. A New Japanese Input Method for Virtual Reality Applications. In M. Kurosu ed., *Human-Computer Interaction. Interaction Technologies*, Vol. 10903, pp. 43–55, Cham, 2018. Springer International Publishing.

- [3] M. A. Rahim, J. Shin, and M. R. Islam. Gestural Flick Input-Based Non-Touch Interface for Character Input. *Vis. Comput.*, 36(8):1559–1572, 2020.
- [4] Y. Singhal, R. H. Noeske, A. Bhardwaj, and J. R. Kim. Improving Finger Stroke Recognition Rate for Eyes-Free Mid-Air Typing in VR. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.
- [5] Y. Tomaru and T. Tanaka. Proposal of Character Input Method in VR Environment using Leap Motion. In *2019 8th International Congress on Advanced Applied Informatics (IIAI-AAI)*, pp. 1037–1038, 2019.
- [6] E. Wu, Y. Yuan, H.-S. Yeo, A. Quigley, H. Koike, and K. M. Kitani. Back-Hand-Pose: 3D Hand Pose Estimation for a Wrist-Worn Camera via Dorsum Deformation Network. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, UIST '20, pp. 1147–1160, New York, NY, USA, 2020. Association for Computing Machinery.
- [7] 細野 敬太, 笹倉 万里子, 田邊 浩亨, 川上 武志. Leap Motion を用いたジェスチャ操作による文字入力方法の提案. 人工知能学会全国大会論文集, 28:2E14in.
- [8] 井上 賢人, 梅澤 猛, 大澤 範高. 両手の指と掌の接触を用いたかな文字入力手法の検討. 第 81 回全国大会講演論文集, 2019.
- [9] 塚田 敏彦, 三和田 靖彦. 生産工程で使える非接触データ入力手法の検討. 総合技術研究所研究報告, No. 23, 2021.