

ヘッドマウントディスプレイ使用時の文字入力手法

荻谷 太平^{*1} 荒堀 喜貴^{*2} 権藤 克彦^{*3}

WIP: Effective Text Input for Commodity Head Mounted Displays

Taihei Ogitani^{*1}, Yoshitaka Arahori^{*2} and Katsuhiko Gondow^{*3}

Abstract – We propose an effective text-input system for commodity head mounted displays (HMDs). Despite the wide-spread use of HMDs in virtual reality (VR) and augmented reality (AR) applications, existing text-input systems for HMDs are either (1)inaccurate, (2)dependent on special devices, or (3)require non-small physical space. This paper presents our work-in-progress to address all of these problems. Our system, called PalmFlick, reduces the HMD text-input problem into the problem of flicking to select characters on a palmtop virtual panel. We outline the key design choices for PalmFlick to achieve high accuracy on commodity HMDs.

Keywords : Text input, Head mounted display, Virtual reality, Augmented Reality, Human interface

1. 概要

近年、ヘッドマウントディスプレイ（HMD）を用いた仮想現実（VR）や拡張現実（AR）のコンテンツの発展が著しい。一方で、HMD 使用時の文字入力の手法は、精度が低い、特殊なデバイスが必要、もしくは入力に広い空間が必要などの理由で適切なものがあるとはいえない。本論文では、カメラを使った手のひらの認識による文字入力の手法を提案し、提案手法が導入しやすく、かつ入力精度の高い手法であることを示す。

2. 導入

近年、HMD を用いたコンテンツの発展が著しい。その一方で、HMD 使用時の文字入力手法には適切なものがあるとはいえない。HMD を用いたコンテンツには VR コンテンツと AR コンテンツが存在する。VR コンテンツでは、ユーザの視界は HMD に覆われ、外の景色は一切見られなくなる事が多い。このため、キーボードや手持ちのコントローラを使った入力手法では入力のために HMD を外す、コンテンツ内でデバイスを表示するなどをする必要があり、入力を困難にし没入感の低下を招く。AR コンテンツは、その特性上歩きながら利用したり、外出先で利用することが今後考

えられる。この際、文字入力のために特定のデバイスを携帯することは非合理的である。従って、HMD 使用時の文字入力には特定の手持ちのデバイスを必要としないことが求められる。

本論文では HMD を SNS や情報検索、ゲームなどに利用すると考え、文字入力の対象を SNS 等の短いコミュニケーションや検索語の入力と想定している。

既存手法として、音声認識やアイトラッキングがある。音声認識は発声する必要があるため、電車内など状況によっては不適切である。アイトラッキングは視線を動かす必要上、没入感の低下を招くため特に VR では適切でない。また、カメラを用いたハンドトラッキングの手法も多数研究されている ([1], [3], [7], [8])。これらは精度が低いか、特殊な機材を使用して精度を確保している。また、精度の確保のために現実空間、もしくは仮想空間を大きく占有する必要があるものもある ([8])。手法のこうした特徴は、実用上の障害となる。

本論文では上記の問題を解決するため、安価かつ市場入手可能なトラッキングカメラを用いた省スペース文字入力手法 *PalmFlick* を提案する。本手法は片手のひらを入力面として、反対の手の指でフリック入力をする。トラッキングする関節数を減らし、トラッキングの少ない関節を用いて関節のトラッキングを最適化することで、精度の低いトラッキングカメラでも入力精度を高く保てると考えた。

本手法により、以下の 3 点の貢献が予測できる。

- 機材の導入障壁の低減
- 身体感覚による精度の向上
- 占有空間の縮小

手法の有用性を示すため、タイプ速度、エラー率とそれぞれの向上率について既存手法との比較実験を予

*1: 東京工業大学 情報理工学専攻 計算工学専攻

*2: 東京工業大学 情報工学系 情報工学コース

*3: 東京工業大学 情報工学系 情報工学コース / 情報工学系 知能情報コース

*1: Graduate School of Information Science and Engineering, Tokyo Institute of Technology

*2: Dept. of Computer Science Major in Computer Science, Tokyo Institute of Technology

*3: Dept. of Computer Science Major in Computer Science / Major in Artificial Intelligence, Tokyo Institute of Technology

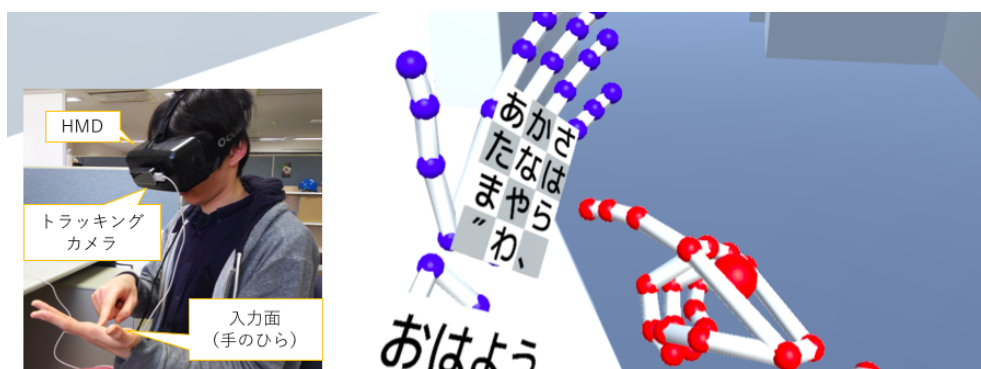


図1 *PalmFlick* の VR での使用イメージ

Fig.1 Using simulation of *PalmFlick*

定している。

今後は、実験による手法の有用性の証明とともに、さらなる入力精度の向上や多言語への対応により実用に即した手法とすることを目指す。

本論文の構成を示す。第3節ではハンドトラッキングに関する既存手法について述べる。第4節では既存手法の問題点から本論文の問題設定を述べ、第5節ではその問題を解決する提案手法と貢献について述べる。第6節では今後行う実験と、解決済みでない問題について述べ、第7節で結論、第8節は参考文献とする。

3. 関連研究・手法

3.1 手に関連する研究

人間の手の指は、特定の指だけを動かそうとすると他の指も同時に動いてしまう事がある。この現象について報告した先行研究がある^([1],[6])。この研究結果から、ジェスチャーを用いた手法の構築の際は、曲げる指の組み合わせを考慮するべきだといえる。

また、先行研究^[2]での実験の結果から、ハンドトラッキングを用いて手のひらを入力面とした際、十分な精度を確保するためには入力面の分割数は縦横4以下であることが必要だと分かっている。

3.2 文字入力手法

HMD 使用時の文字入力手法には音声認識やアイトラッキングなどがある。音声認識は音声を発声する必要があるため電車内などの状況下では不適切であり、アイトラッキングは入力のために視界を動かす必要があり没入感の低下につながる。いずれも本手法の想定する状況とは差異がある。

HMD に最適化されたジェスチャーを使う文字入力手法は、2段階バーチャルキーボード (2VK)^[7] や PalmType^[3]、Drum Keys^[8] などが挙げられる。

2VK はカメラでトラッキングした手の2次元移動と5本指の曲げ伸ばしで文字を入力する。前項で述べたように指によっては一本だけ曲げることが難しく、それによるトラックミスが問題である。PalmType は片

手首に赤外線センサを装着し、手のひらを疑似 QW-ERTY キーボードとして入力する。手のひらの領域を細かく区切るため、トラックミスが問題となる。Drum Keys はセンサと手持ちコントローラにより両手をトラッキングし、そこから伸びるバチで拡大されたキーボードを仮想空間内で叩くことで入力する。この手法は現実、仮想問わず空間を広く必要とするため没入感や屋外での使用に問題がある。

4. 問題設定

表1は、既存手法と提案手法の比較である。既存手法は精度の高いデバイスを使うことで文字入力精度を高めているが、そうしたデバイスの導入は商業的には困難であり HMD 普及の障壁になる。本手法は、精度が低い機材でも精度の高い文字入力ができることを目標としている。また、入力のために現実空間や仮想空間を広く必要とすると、周囲への影響や没入感の低下を招く問題がある。本手法では、手のひら以上の大きさの空間を要しないことで問題を解決した。

表1 関連手法と提案手法の比較
Table 1 Comparison between related method and proposed method

手法	トラック精度 ¹	機材導入障壁 ²	占有空間の大きさ ³
2VK ^[7]	△	○	△
PalmType ^[3]	×	△	○
Drum Keys ^[8]	○	×	×
提案手法	○ ⁴	○	○

5. 提案手法

第4節の問題を受けて、低精度カメラでの省スペース文字入力手法 *PalmFlick* を提案する。

PalmFlick は、片手のひらを入力面として逆の手の人差し指によるフリック入力で文字を入力する。両手

1: カメラ1台でのトラックを仮定

2: 高価もしくは独自の機材を必要とするかどうか

3: 現実、仮想現実を問わず、入力に必要な空間の大きさ

4: 目標値

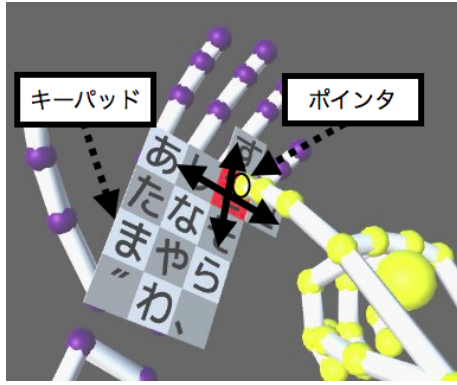


図2 PalmFlick のシステム構成要素
Fig.2 System component of PalmFlick

のトラッキングにはカメラを用い、HMD の前面に取り付けられているか内蔵されていることを想定する。

図1は、PalmFlickをVRコンテンツで利用するイメージである。文字入力の際にユーザの手のポーズを認識し、同様のポーズを取った手の3DCGをレンダリングする。その上で一方の手のひらにキーパッドを表示することで、没入感を損なわず文字入力をさせられる。ARコンテンツでは、単純にARHMDを通して見える一方の手のひらにキーパッドを合成表示する。

図2は、提案手法の入力時のシステム構成要素である。図中のキーパッドは入力面となる手の手のひらに表示され、反対の手の人差し指の指先をキーに触れて入力するポインタとする。キーパッドとポインタの両者のトラッキングに必要なのは、本質的には手のひらの位置と角度、人差し指の指先の位置のみなので、位置の取得とその補正に使用しない指や関節のトラッキングミスは無視できる。入力中はキーパッド側の手のひらの位置は動かず、ポインタ側の人差し指のポーズは変化しないと仮定できるため、比較的トラッキングミスの少ない手首や人差し指の第三関節の位置で補正し入力ミスを減らせる。以上から、提案手法は検知精度の低いカメラでも入力精度を高められると予測できる。

5.1 効果

HMD使用時の入力手法への効果として、以下の3つがある。

機材の導入障壁の低減

機材の導入障壁を低くした。特定のコントローラや高価な検知機材の利用により文字入力の精度を高める既存手法^{[3], [8]}と違い、本手法では導入の難しい機材ではなく安価で一般市場で入手可能なカメラを取り付ける（あるいは取り付けられているカメラを利用する）ことのみで、HMD使用時の文字入力が可能となる。

身体感覚による精度の向上

身体感覚を利用し、VR空間内での位置関係や距離

感をつかみやすくした。また、手のひらへのタップの精度は手のひらを見なくても大きく落ちない^[4]ため、手のひらを利用しない手法と比較してユーザのブラインドタッチの習熟が速くなることが期待できる。

占有空間の縮小

文字入力に必要な空間を縮小した。ある程度画面内を占有する装置や画面を表示する既存手法^{[7], [8]}があるが、これらはVRへの没入感の低下や、ユーザの周辺環境への悪影響などの問題につながる。本手法では、手のひらと同様のサイズのキーパッドを表示するに留めることで、この問題を防ぐことが期待できる。

以上の3点から、本手法はHMD使用時の文字入力手法として既存手法と比較して大幅に実用に即したものだといえる。

6. 今後の展望と課題

6.1 実験

本手法の有用性を示すため、被験者実験を実施する予定である。HMD使用時の文字入力手法として実用に耐えうるかを検証するため、入力手法のタイプ速度とエラー率、それぞれの向上率を評価する。また、精度の低いカメラでの入力精度が本手法は優れていることを示すため、同条件での既存手法との比較を行う。

6.1.1 装置

HMDにはOculus Rift CV1を使用する。Oculus Riftには手をトラッキングするためのカメラは内蔵されていないため、Leap Motionをトラッキングカメラとして利用する。

6.1.2 計画

被験者には1週間の間、HMD使用時の文字入力の試行を行わせる。1回の試行は、20字程度の文字入力とする。各被験者に対し、 i 日目1日当たりのタイプ速度 WPM_i 、エラー率 TER_i と、1週間全体でのタイプ速度、エラー率の向上率 I_{WPM} 、 I_{TER} を計測、被験者毎に比較評価する。

Drum Keys^[8]とPalmType^[3]を既存のHMD使用時の文字入力手法として提案手法と比較する。両者は、Leap Motionを用いた実験に利用できるよう再実装する。

6.1.3 評価及び手法

以下の4点を評価し手法を比較する。

タイプ速度 WPM_i

$$WPM_i = \frac{\sum_{j \in Task_i} \frac{S_{i,j}-1}{T_{i,j}} \times 60 \times \frac{1}{5} \times (1 - U_{i,j})}{|Task_i|} \quad (1)$$

1日の実験における未修正の入力文字を除く平均WPM（単語/分）を評価する。この手法は^[9]を参考にしている。

なお、 $Task_i$ は i 日目の試行の集合であり、 $S_{i,j}$ 、 $T_{i,j}$ 、 $U_{i,j}$ はそれぞれ、 i 日目の試行 j における入力文字数、試行の完了に要した時間（秒）、試行終了時に誤っている文字の割合である。

また、*Fitts* の法則 ([5]) を元に、各手法の潜在的なタイプ速度を比較する。

エラー率 TER_i

$$TER_i = \frac{IF_i + INF_i}{TS_i} \quad (2)$$

1 日の実験における修正済み、未修正のエラー率の合計値を評価する。この手法は [9] を参考にしている。

IF_i はバックスペースにより修正された文字数の、 INF_i は試行終了時に誤っている文字数の、 TS_i はバックスペースを除く全ての入力文字数の、それぞれ i 日目の合計である。

タイプ速度の向上率 I_{WPM} 、エラー率の向上率 I_{TER}

最小二乗法を用いて、全被験者の 7 日間の WPM 、 TER の記録について傾きと切片をとり、それらを比較する。

6.2 今後の改善点

6.2.1 AR 環境での評価

本手法は AR 環境下でも有用だと考えるが、実際に有用かは、AR 環境で本手法を用いた評価実験が必要である。

6.2.2 入力機能の向上

提案手法では、トラックする部位は片手のひらと片手の人差し指のみである。これらのトラッキングを更に最適化することで、トラックミスの軽減や低機能なカメラでの運用が可能になる。

また、キーボードのレイアウトを、より手のひらでの入力に適したものに改善する事も考えられる。

6.2.3 多言語、記号への対応

HMD の普及を考えれば、多国語への対応は非常に重要である。フリック入力上で再現できるキー数 ($60 = 12 \times 5$) で表現できる言語であれば本手法に適用可能である。

加えて、絵文字や記号への対応のためのキーパッドレイアウトの調整も今後必要である。

6.3 未解決の課題・疑問

- 実験手法は文字入力の評価の妥当な手段か
- *Fitts* の法則モデルはヒューマンインタフェースの領域で認められているか
- ユーザの能力の向上率の有効な求め方について
- 比較に使用する手法は妥当か
- 没入感を評価する手法はあるか

7. 結論

HMD 使用時の文字入力の問題に対して、片手のひらを入力面としたハンドトラッキングを用いる手法を

提案した。この手法は一般的に購入可能かつ比較的安価なデバイスを用いて実現できるよう設計されているため、導入のコストが低い。また VR 空間内にも存在するユーザ自身の手のひらを利用し、手のひら以上の領域を必要としないことから VR 空間での文字入力に適していると考えられる。

今後は、手法の改良と被験者実験を通して手法の有用性を示すことを目標に研究を進める。

8. 参考文献

参考文献

- [1] S. Sridhar, A. M. Feit, et al.: Investigating the Dexterity of Multi-Finger Input for Mid-Air Text Entry.; In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15). ACM, New York, NY, USA, 3643-3652. , 2015.
- [2] N. Dezfili, M. Khalilbeigi, et al.: PalmRC: imaginary palm-based remote control for eyes-free television interaction.; In Proceedings of the 10th European Conference on Interactive TV and Video (EuroITV '12). ACM, New York, NY, USA, 27-34., 2012.
- [3] C. Wang, W. Chu, et al.: PalmType: Using Palms as Keyboards for Smart Glasses.; In Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15). ACM, New York, NY, USA, 153-160. 2015.
- [4] C. Harrison, H. Benko, and A. D. Wilson.: OmniTouch: wearable multitouch interaction everywhere.; In Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11). ACM, New York, NY, USA, 441-450. 2011.
- [5] I. S. MacKenzie.: Fitts' Law as a Research and Design Tool in Human-Computer Interaction.; In HUMAN COMPUTER INTERACTION, 1992, Volume 7,91-139.Copyright Q 1992, Lawrence Erlbaum Associates, Inc. 1992.
- [6] Schieber, M. H.: Individuated finger movements of rhesus monkeys: a means of quantifying the independence of the digits;. J Neurophysiology 65, 6 (1991), 1381-91.
- [7] S. Kim and G. J. Kim.: Using keyboards with head mounted displays.; In Proceedings of the 2004 ACM SIGGRAPH international conference on Virtual Reality continuum and its applications in industry (VRCAI '04). ACM, New York, NY, USA, 336-343. 2004.
- [8] Diaz J:Google 's Daydream Labs Explores Typing On Virtual Drums. : <https://www.androidheadlines.com/2016/05/googles-daydream-labs-explores-typing-virtual-drums.html> (2017/7/18)
- [9] I. S. MacKenzie and K. Tanaka-Ishii.: Text Entry Systems; Mobility, Accessibility, Universality. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. 2007.