

# VR環境下における身体背面部への タップ操作による入力領域拡張

國枝 直希<sup>1,a)</sup> 真鍋 宏幸<sup>1,b)</sup>

**概要：**現在のVRアプリケーションで利用される操作の一つに、ハンドトラッキングによるジェスチャがあるが、この操作手法にはHMDの内蔵カメラに手が映る範囲でしか利用できない問題や、入力語彙が少ないなどの問題がある。本稿では、これらの問題を解決するため、身体背面部などのあまり利用されていない領域を活用し、身体背面部をVR空間上のUI操作の入力領域として用いる手法を提案した。提案手法の実装を行い、腰や後頭部などの身体背面部へのタップを検出できることを確認した。



図1 提案手法の様子

## 1. はじめに

教育や娯楽など、さまざまな分野でVRの利用が広がっている。現在のVRコンテンツで利用されている操作方法の一つとして、HMDに搭載されたカメラを利用したハンドトラッキング機能によるジェスチャがある。HMDの内蔵カメラの視野内に映った手の動きを検出、認識することでジェスチャに対応した操作を行っているが、内蔵カメラに手が明瞭に映っていないと利用できないため、操作に利用できる領域が内蔵カメラがついている前方側に限られてしまう問題がある。その結果、内蔵カメラの視野外である身体後方や背面部などの領域が入力に利用されることはほとんどない。また、ジェスチャ操作での入力語彙を拡張するには、認識するジェスチャの種類を増やすなどの方法があるが、識別するジェスチャの種類が増えるほどユーザの意図しない動作に対して操作が行われてしまう問題（ミダスタッチ問題）や、ジェスチャの複雑化が発生してしま

うため、入力語彙が限られてしまう問題がある。

我々は、身体後方や背面部などのあまり利用されていない領域をタップ入力可能な領域として活用することで、VR空間上での入力語彙を増やすことができると考え、HMDの内蔵カメラを用いたハンドトラッキングと手首に装着したIMUを組み合わせた手法を提案する。

## 2. 関連研究

本研究は、ボディトラッキングなどの身体の各部位の位置や姿勢を推定する研究、身体を入力デバイスとして用いることで入力語彙を拡張する研究、身体後方でハンドトラッキングを行う研究などに関連する。

### 2.1 ボディトラッキング

Caoらは、カメラで撮影した画像、動画内の人物に対してリアルタイムに姿勢推定を行うシステムを提案している[1]。このシステムを利用することで、カメラの情報のみでボディトラッキングを行うことができるが、ユーザの全身の動きが隠れることなくトラッキングできるようにするには、カメラの死角がないように複数のカメラを設置する必要がある。

Casermanらは、HTC ViveとVive Trackerを用いることで低遅延、高精度のボディトラッキングを行っている[2]。しかし、この方法ではトラッキングを行いたい部位毎にVive Trackerを装着し、Vive Trackerを認識するためのベースステーションを設置する必要がある。

カメラやトラッカーなどの外部機器を用いずに全身トラッキングを行う研究として、HMDと左右の手に持ったコントローラの3点の情報から強化学習によってユーザ

<sup>1</sup> 芝浦工業大学

<sup>a)</sup> al20082@shibaura-it.ac.jp

<sup>b)</sup> manabehirokyu@acm.org

のアバタ全身の姿勢を推定する研究が行われている [3][4]. これらの研究では、頭部と両手の3点の座標のみで全身の動きを再現することができるものの、ユーザとのインタラクションは環境とのインタラクションにとどまっている.

## 2.2 身体を入力デバイスとして用いる研究

Harrison らは、腕をタップした際の振動を、センサを搭載したアームバンドで検知することで、腕を入力デバイスとして用いる手法を提案している [5]. しかし、この手法では、アームバンドを装着した部分の延長線上しか入力に用いることができない問題がある.

また、小林らは、身体へのタップジェスチャの回数パターン組み合わせによってコミュニケーションを行うシステムを提案した [6]. しかし、タップジェスチャの回数のみで多数の表現を認識するには多くのタップ回数と時間を必要とする問題がある.

## 2.3 身体後方でのハンドトラッキング

Wei らは、スマートウォッチに搭載された6軸IMUを利用して肘、手首の位置をRNNによって推論する手法を提案した [7]. この研究では9軸IMU(加速度、ジャイロ、地磁気)を利用しなくともトラッキングが行えることを示したが、肘との相対位置として手首の位置を推定しているため、肘の誤差の影響により手首の絶対位置の誤差が大きくなる問題や、胴体や肩が移動していない状況下でしか利用できない制約がある.

Streli らは、VR空間上でHMDのハンドトラッキング範囲外のオブジェクトに対して手による操作(つかむ、引く、移動させるなど)を行うHOOVを提案した [8]. この研究では、HMDに搭載されたカメラでのトラッキング範囲内ではカメラによって、範囲外でのトラッキングは過去のトラッキング情報や手首に装着したIMUのセンサ情報などから機械学習的に推論することで、6軸IMUでも高い精度で身体後方でのハンドトラッキングを行っている. しかし、この手法ではピンチ操作を検出できる範囲は拡張されたものの、ユーザの入力はピンチ操作のみであり、操作の種類は拡張されていない.

## 3. 提案手法

本研究では、手首のトラッキングによって身体のどの部位がタップされたかを検出し、検出された部位と対応する操作(カーソル移動、決定など)を行うことで、手でタップができる身体の部位を入力領域として利用するシステムを提案する.

本手法では、HOOV [8] と同様に、左右の手首にIMUを装着し、HMDによるトラッキング範囲内ではHMD、範囲外では過去のトラッキング情報やIMUによって手首のトラッキングを行う. トラッキングと同時にタップの検出

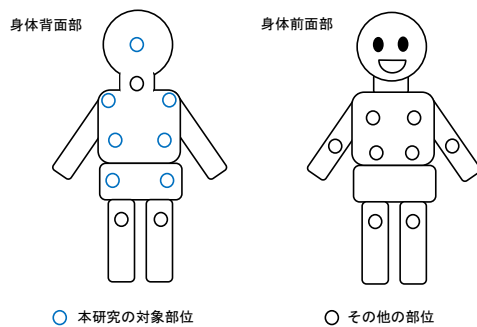


図2 検出を行う部位

やタップした部位の推定を行い、部位に対応した操作を事前に割り当てることでタップジェスチャによる操作を可能にする. 本システムでは、HOOVのトラッキング手法を用いるものの、身体背面部へのタップ操作を行うことで、HOOVで活用されていない領域を入力に用いることができる.

ただし、HOOVよりも手が身体後方にある時間が長いことが予想されるため、IMUによるトラッキング精度が低下する可能性がある. この問題については、機械学習モデルの改良を行うとともに、HMDによるハンドトラッキング範囲内での操作に提案手法を加える形でUIを実装し、手が身体後方にある時間が短くなるように工夫することで、十分な精度の達成を目指す.

右腰をタップした際の提案手法の様子を図1に示す. また、図2に、提案手法でカバーできると考えられる部位を示す. 今回の実装では青で示す部位(後頭部、肩、背中、腰)を検出対象とするが、それら以外の身体背面部の部位や、HMD内蔵カメラの視野から外れることの多い腹部など身体前面部位(図中、黒)を含むあらゆる身体の部位に対応することができる.

## 4. 実装

ユーザの左右の手首に装着するIMUセンサとしてリストバンドに装着したJoy-Conを、HMDとしてMeta Quest 2を用いて実装を行った. 手の座標や姿勢を推論するモデルはHOOV [8] と同様のものをPyTorchを用いて実装し、Vive Tracker 3.0を利用して左右の手首の3次元座標と姿勢情報を取得することで、データセットを作成した. タップについては、Joy-Conに搭載されたIMUの3軸加速度センサから得られた3次元ベクトルから閾値を設定して検出する. また、各部位の座標については、立っている状態でのHMDのUnity上の座標をもとにユーザの身長を算出し、身長をもとに統計的な数値からユーザの体格と各部位の相対的な位置を設定した. タップした部位の推定は、タップ検出を行った際に手が最も近くにある部位をタップした部位としている. Quest 2向けのVRアプリケーションの実装はUnity 2021で行い、ソケット通信を介してPythonで

作成したアプリケーションと VR アプリケーション間で情報を送受信している。Python アプリケーションは、Core i9-9900K, GTX 960, 32GB RAM を搭載した PC 上で動作している。

これらの実装を行った後、左右の手でそれぞれタップした部位を VR 空間上に表示するアプリケーションを作成した。その結果、左手、右手それぞれにおいて、今回検出対象とした後頭部、肩、背中、腰の各 4 か所ずつに対してタップした部位の検出が行えることを確認した。

## 5. 議論

本章では、本手法に関する課題や将来的な拡張について議論を行う。

### 5.1 ユーザの姿勢変化や体格による影響

現在の実装では、HMD の座標から身長を算出し、その身長を元に相対的な部位の位置を決定している。したがって、ユーザが背中を曲げる、頭を傾けるなど大きく姿勢を変えた状態で本システムの利用を行った場合や、ユーザの体格がその身長での統計的体格と異なる場合、正確な検出ができないことが考えられる。この問題については、ユーザごとの事前のパラメータ入力や、頭と両手首の 3 点トラッキングをベースにした機械学習によるボディトラッキング、追加でトラッカーなどを取り付けることで改善する可能性がある。

### 5.2 ハンドトラッキングの精度の影響

先行研究 [8] と同様の機械学習モデルを用いたが、Ground Truth を OptiTrack から Vive Tracker 3.0 に変えている点や、学習データの規模の違い、座標系を合わせるためのキャリブレーションの精度などが影響して精度が落ちている可能性がある。これらの点については、より精密なキャリブレーションを行い、学習データを追加で集めることで改善すると考えられる。また、今回のようなハンドトラッキング手法を用いる場合、精度を上げたとしても数 cm から 10 数 cm 程度の誤差が発生する。このような誤差により、腰、背中、肩などの大まかな部位の検出は高い精度で行うことができるが、1 つの部位を複数の領域に分割するような細かな検出を行うことは非常に難しいと予想される。

### 5.3 身体後方でのハンドトラッキングの必要性

部位推定に身体後方でのハンドトラッキングが必要なのかを確認するため、HMD のカメラの視界から外れる直前の手の座標情報をもとに推定を行った。その結果、理想的な経路で手を動かすことで、機械学習による推論がなくとも検出が可能であることが分かった。しかし、今後検出部位数を増やしていくことを考えると、身体前方の部位と取り違えてしまう問題や、理想的でない経路で手を動かした

際に経路付近の別の部位と誤認識する問題があり、身体後方でのハンドトラッキングは必要と考えられる。

## 5.4 検出部位数の増加による拡張性

本稿では後頭部、肩、背中、腰それぞれの左右の計 7 か所の部位に対して検出を行ったが、検出部位を増やすことによって、手法そのものは身体の前側（胸、腹など）にも適用することが可能である。特に、ユーザが座っている際には身体背面部は利用が難しいため、身体の前側を利用できるように拡張することで、入力語彙が増加するだけでなく、別のユースケースにも本手法が適用できることが期待される。

## 6. まとめ

自身の身体背面部をタップすることで入力を行う VR 環境向け操作手法を提案した。提案手法の実装を行い、身体背面部へのタップを検出できることを確認した。また、今後の改善に向けた議論を行った。

## 参考文献

- [1] Cao, Z., Simon, T., Wei, S.-E. and Sheikh, Y.: Real-time Multi-Person 2D Pose Estimation Using Part Affinity Fields, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017).
- [2] Caserman, P., Garcia-Agundez, A., Konrad, R., Göbel, S. and Steinmetz, R.: Real-time body tracking in virtual reality using a Vive tracker, *Virtual Reality*, Vol. 23, pp. 155–168 (2019).
- [3] Jiang, J., Strel, P., Qiu, H., Fender, A., Laich, L., Snape, P. and Holz, C.: AvatarPoser: Articulated Full-Body Pose Tracking from Sparse Motion Sensing, *Proceedings of European Conference on Computer Vision*, Springer (2022).
- [4] Winkler, A., Won, J. and Ye, Y.: QuestSim: Human Motion Tracking from Sparse Sensors with Simulated Avatars, *SIGGRAPH Asia 2022 Conference Papers*, SA '22, New York, NY, USA, Association for Computing Machinery, (online), DOI: 10.1145/3550469.3555411 (2022).
- [5] Harrison, C., Tan, D. and Morris, D.: Skinput: Appropriating the Body as an Input Surface, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, p. 453–462 (online), DOI: 10.1145/1753326.1753394 (2010).
- [6] 小林優維, 呉健朗, 大和佑輝, 宮田章裕ほか: 身体へのタップジェスチャでコミュニケーションを行うシステムの基礎検討, マルチメディア, 分散協調とモバイルシンポジウム 2019 論文集, Vol. 2019, pp. 105–110 (2019).
- [7] Wei, W., Kurita, K., Kuang, J. and Gao, A.: Real-Time 3D Arm Motion Tracking Using the 6-axis IMU Sensor of a Smartwatch, *2021 IEEE 17th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 1–4 (online), DOI: 10.1109/BSN51625.2021.9507012 (2021).
- [8] Strel, P., Armani, R., Cheng, Y. F. and Holz, C.: HOOV: Hand Out-Of-View Tracking for Proprioceptive Interaction Using Inertial Sensing, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, (online), DOI: 10.1145/3544548.3581468 (2023).