

# インソール型センサーによるリアルタイム姿勢推定

近江 俊樹\* Erwin Wu† 小池 英樹\*

**概要.** 既存の三次元姿勢推定技術は高価な設備や限定的な環境、もしくは、多数のウェアラブルセンサーを取り付けるような大掛かりな環境を必要としていた。しかし、様々なアプリケーションにおいて姿勢推定技術を最大限活用するためには、より普遍的な設備で姿勢推定を行う必要がある。そこで我々は、ウェアラブルな姿勢推定技術として、インソール型センサー単体での三次元姿勢推定技術を開発したが、胴体や下半身に比べて腕や手首の推定精度が低いことが分かった。そこで本論文では、インソール型センサーが取得する各足の足圧分布と6自由度加速度信号および、リストバンド型センサーが取得する6自由度加速度信号を用いて三次元姿勢推定を行うシステムを提案する。

## 1 はじめに

現在、三次元姿勢推定技術は様々なアプリケーション [3?] に用いられている。しかし、従来の姿勢推定技術は、高価な設備や光学マーカーといった特定のシステム [1, 2] に依存している。近年、単眼のRGB(D)カメラによるマーカーレスシステム [8] が提案されているが、マルチカメラシステムによる姿勢推定の精度には及ばない。さらに、カメラを用いたシステムでは、カメラがオクルージョンなく対象の身体を捉えている必要がある。

一方、環境に依存しないウェアラブルセンサーを利用した姿勢推定 [7] が提案されている。しかしながら、既存のウェアラブルな姿勢推定手法 [5, 6] の多くは、6つ以上のセンサーを取り付ける必要があり、不自然で煩わしく、時には危険な場合もある。

このような課題に対処するために、我々はインソール型センサーのみを用いたリアルタイム三次元姿勢推定手法を開発した。これまで足圧分布と姿勢に関する研究 [4] が行われているが、足のダイナミクスから姿勢推定を行う研究は行われていない。

これまでの我々の研究から、インソール型センサーのみを用いた姿勢推定では胴体や下半身に比べて腕や手首の姿勢推定精度が低いことが分かっている。そこで本論文では、腕や手首の情報を補うために、インソール型センサー<sup>1</sup>とリストバンド型センサー<sup>2</sup>を用いたリアルタイム三次元姿勢推定システムを提案する。

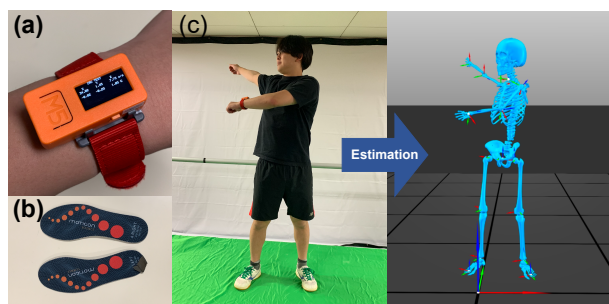


図 1. センサーと提案システムの概要。(a) リストバンド型センサー。(b) インソール型センサー。(c) 取得したデータを元に推論を行う。

関節	部位
1	頭
2	首
3	肩
4	肘
5	手首
6	腹部
7	股関節
8	膝
9	足首
10	つま先

表 1. 関節と部位の対応表

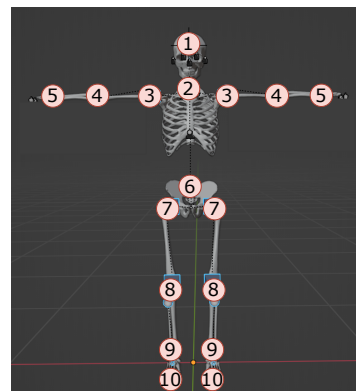


図 2. 関節位置のイメージ図

## 2 実装

インソール型センサーおよびリストバンド型センサーによる姿勢推定の効果を検証するために、姿勢推定システムを構築した。

### 2.1 セットアップ

本システムでは、まず、図 1(b) に示す、16 チャンネルの圧力センサーと IMU センサーが埋め込まれ

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.

\* 東京工業大学

† Huawei Research Japan

<sup>1</sup> <https://moticon.com/opengo/sensor-insoles>

<sup>2</sup> <https://m5stack.com/>

表 2. 各関節項目ごとのリストバンド型センサーがある場合とない場合に対する PMJPE(mm) と PMJAE(degree) の計算結果.

関節点	全身 (17点)		腕・手首以外 (11点)		腕・手首 (6点)		左半身 (7点)		右半身 (7点)	
	MPJPE	MPJAE	MPJPE	MPJAE	MPJPE	MPJAE	MPJPE	MPJAE	MPJPE	MPJAE
インソール型センサーのみ	91.5	10.3	58.1	4.9	152.7	20.1	98.6	11.4	87.6	11.4
提案手法	95.1	10.6	61.5	5.1	156.5	20.6	101.5	11.8	92.5	11.7

ているインソール型センサーを靴の中に入れる。さらに、図 1(a) に示す、IMU センサーが内蔵されたリストバンド型デバイスを左手首につける。これらを用いて、各足の圧力分布と各足および左手首の 6 自由度加速度信号を取得し、三次元姿勢推定を行う。

## 2.2 推論

推定する三次元姿勢は、図 2 に示す各関節とその角度である。各関節の位置と部位との対応を表 1 に示す。また、姿勢推定に用いるモデルは、図 3 に示すように Transformer である。入力には 15 シーケンス長の 3 つの 6 自由度加速度信号と 2 つの 16 チャンネルの足圧データである。

## 3 実験

リストバンド型センサーの導入による姿勢推定の性能の変化を調べるために、インソール型センサーを用いた姿勢推定と、インソール型センサーとリストバンド型センサーを用いた姿勢推定の推定精度の比較を行った。比較指標として、MPJPE および MPJAE を使用した。また、腕・手首の姿勢推定を評価するために、全関節点だけではなく、腕・手首の関節点のみでの評価および腕・手首以外の関節点での評価も行う。ただし、腕・手首の関節点は肩、肘、手首の 6 点である。さらに、リストバンド型センサーを片腕だけにつける影響を調べるために、体の中心の関節点を除いた左半身と右半身のそれぞれ 7 点に対する評価も行う。

本実験では、ラジオ体操運動時の姿勢推定を行う。ラジオ体操はその場で行う全身運動であり、下半身に対して上半身の動きの種類が多く、左右非対称な動きが含まれる体操運動である。三次元姿勢推定モデルの学習・評価に用いる関節の三次元回転量および三次元座標はマーカーレスモーションキャプチャ<sup>3</sup>を使用して取得した。

実験結果を表 2 に示す。

## 4 考察

実験結果から、提案システムの改善点に関する知見が得られた。

<sup>3</sup> <https://qualisys.archivetips.com/>

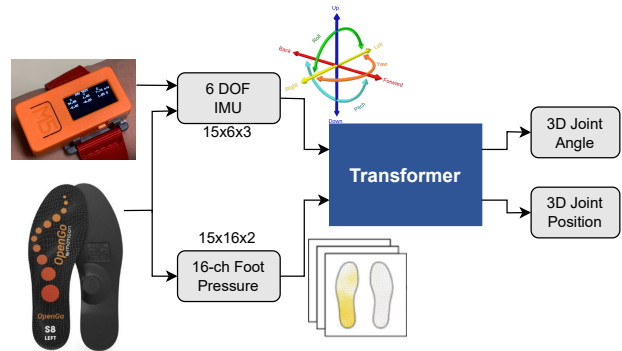


図 3. モデル概要図

表 2 から、すべての項目で、提案手法による姿勢推定は、インソール型センサーのみによる姿勢推定に比べてわずかに低い精度となった。また、リストバンド型センサーの有無に関わらず、腕・手首の関節点の推定精度は他の関節点の推定精度に比べて低いことが分かる。さらに、リストバンド型センサーは左腕だけに装着しているにも関わらず、右半身と左半身とで推論精度に大きな差異が生じていないことが分かる。

以上の結果から、本システムにおける推論はリストバンド型センサーから得られるデータではなく、インソール型センサーから得られるデータに依拠していると考えられる。これは学習および推論に用いたデータセットに原因があるものと考えられる。本実験に用いたデータセットはラジオ体操を行っている時の計測データで構成されているが、ラジオ体操には主に上半身を動かす種目が多く、中には下半身をほとんど動かさない種目が存在している。そのため、インソール型センサーで取得したデータとリストバンド型センサーで取得したデータを組み合わせるとき、推論に有効な特徴量を学習できなかったのではないかと考えられる。

この問題に対処するために、より多様な運動から計測されたデータセットを学習に用いる必要がある。特に、歩行やスキップなどの足と腕が連動する運動を取り入れる必要があると考えられる。

## 5 結論

本研究では、リアルタイム姿勢推定システムとして、インソール型センサーとリストバンド型センサーを用いた手法を提案した。実験結果から、本手法を改善させる知見を得ることができた。

## 謝辞

本研究はJST CREST JPMJCR17A3およびJST ムーンショット型研究開発事業 JPMJMS2012 の支援を受けている。

## References

- [1] H. Joo, H. Liu, L. Tan, L. Gui, B. Nabbe, I. Matthews, T. Kanade, S. Nobuhara, and Y. Sheikh. Panoptic Studio: A Massively Multiview System for Social Motion Capture. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 3334–3342, 2015.
- [2] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars eds., *Computer Vision – ECCV 2014*, pp. 740–755, Cham, 2014. Springer International Publishing.
- [3] N. Sarafianos, B. Boteanu, B. Ionescu, and I. A. Kakadiaris. 3D Human Pose Estimation. *Comput. Vis. Image Underst.*, 152(C):1–20, nov 2016.
- [4] J. Scott, B. Ravichandran, C. Funk, R. T. Collins, and Y. Liu. From Image to Stability: Learning Dynamics from Human Pose. In *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXIII*, p. 536–554, Berlin, Heidelberg, 2020. Springer-Verlag.
- [5] X. Yi, Y. Zhou, M. Habermann, S. Shimada, V. Golyanik, C. Theobalt, and F. Xu. Physical Inertial Poser (PIP): Physics-aware Real-time Human Motion Tracking from Sparse Inertial Sensors. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022.
- [6] X. Yi, Y. Zhou, and F. Xu. TransPose: Real-Time 3D Human Translation and Pose Estimation with Six Inertial Sensors. *ACM Trans. Graph.*, 40(4), jul 2021.
- [7] Z. Zhang, C. Wang, W. Qin, and W. Zeng. Fusing Wearable IMUs With Multi-View Images for Human Pose Estimation: A Geometric Approach. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2197–2206, 2020.
- [8] Q. Zhao, C. Zheng, M. Liu, P. Wang, and C. Chen. PoseFormerV2: Exploring Frequency Domain for Efficient and Robust 3D Human Pose Estimation. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8877–8886, 2023.

## 未来ビジョン

本論文ではインソール型センサーを用いたリアルタイム姿勢推定システムを提案した。

論文中でも述べているように、従来の姿勢推定システムは、特定のカメラシステムが導入されたモーションキャプチャスタジオであったり、多くのセンサーを身につけなければならないシステムであったりと、特別な環境でしか行えないものであった。特に、カメラを用いたシステムは高性能である一方で、その値段の高さやシステムを構築する環境の用意など、その導入のハードルの高さから多くの人は利用することができなかった。

近年、ウェアラブルデバイスを用いた姿勢推定に関する研究が盛んに行われているが、その中でもインソール型デバイスは、カメラやスマートフォンに比べて安価に購入することができ、多くの人が普段使っている履き物に導入することができるデバイスである。今後も

インソール型センサーを用いた姿勢推定システムを拡張していくが、その一つとして、インソール型センサーを用いたシステムで高性能な姿勢推定システムを構成し、安価でウェアラブルなモーションキャプチャアプリケーションを開発したいと考えている。我々が目指しているのは、誰もが、手軽に、どこでもモーションキャプチャを行うことができる世界である。