

大局座標で測位可能な慣性型モーションキャプチャーの開発

Development of Inertial-based Motion Capture Capable of Positioning in Global Coordinate System

渡部 大基[†]

Taiki Watanabe

プリマ オキ ディッキ アルディアンシャー[†]

Oky Dicky Ardiansyah Prima

1. はじめに

モーションキャプチャー (Motion Capture; MoCap) は、身体の動きを 3 次元的に定量化する技術であり、スポーツやリハビリテーションなどの分野で利用されている。MoCap は、光学マーカー型と深度カメラ型、そして慣性型の 3 つに大別できる。慣性型 MoCap は携帯性が高いため、広域の運動計測に有効であるが、一般に得られたモーションデータが局所座標系で記述されるため、移動距離や移動速度を把握することが困難であった。近年発売された同型の MoCap (Notch) では、大局座標でのモーションデータの出力が可能になったが、ジャンプやスキップなどのような足元が地面に接していない場合の運動において当該データを出力できないという課題がある。本研究では、広域でのモーションデータを取得することを念頭において、Notch の出力データに Long Short-Term Memory (LSTM) を適用し、多彩な走行運動におけるモーションデータを大局座標で取得することを試みる。

2. 関連研究

歩行者の自律航法 (Pedestrian Dead Reckoning; PDR) は、慣性計測装置 (Inertial Measurement Unit; IMU) を身体に取り付けることで、体の動作の開始点からの相対的な動作の変化からモーションデータを求める手法である。当該手法は、歩行速度を求めるとともに、歩行者の位置の検出も行える。近年、PDR における精度の向上に深層学習による手法が導入され、様々な歩容の速度に対して頑健に推定できるようになった [1]。現在のように、多くのスマートフォンなどの携帯デバイスには IMU が搭載されているため、スマートフォンによるモーションデータの取得が容易になる。さらにスマートフォンとスマートウォッチを組み合わせることで、より高精度にモーションデータを取得することも可能である [2]。

慣性型 MoCap では、複数の IMU を体の関節に装着し、各 IMU の加速度と角速度、方位角の諸情報をもとに骨格モデルの動きに置き換えることで、当該モデルの位置 (x, y, z)・姿勢 (pitch, yaw, roll) を推定することが可能である。様々な慣性型 MoCap は開発されているが、その中で Notch (Wearnotch 社) はワイヤレス化した防水 IMU を採用することで歩行や水泳におけるモーションデータの取得を可能にした [3]。さらに取得したモーションデータを大局座標で記録できることで、被計測者の位置とその移動速度を確認できる。ただし、ジャンプやスキップなどのような足元が地面に接していない場合の運動において、位置および移動速度を正しく計測できないという欠点がある。

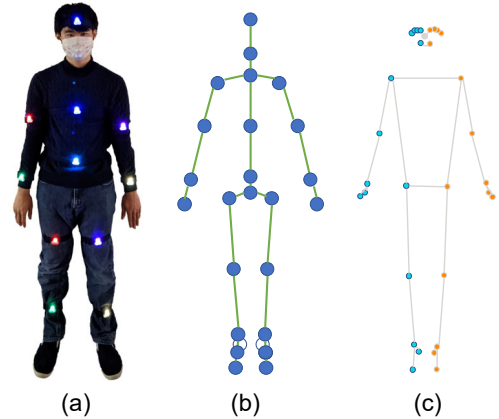


図 1 Notch の装着様子(a)と作成したモーションデータ (b) Notch の出力, (c) MediaPipe の出力

表 1 本研究で作成したモーションデータセット

動作	距離	学習・検証用	評価用
歩行(移動)	10 m	4 回	-
	6 m	10 回	4 回
歩行(静止)	0	3 回	-
スキップ	6 m	9 回	4 回
歩行→ジャンプ→歩行	6 m	7 回	2 回
合計		33 回	10 回

3. モーションデータセットの作成

本研究では、被験者の身体 11ヶ所に Notch を利用して、歩行・スキップ・ジャンプの三種類の 500Hz のモーションデータとそれらの動画を記録した。動画は、スキップとジャンプにおいて正しく出力できないモーションデータを補完するためのものである。ここで、動画からの補完となるモーションデータを作成するために、3次元人物姿勢推定ライブラリ MediaPipe を利用した [4]。図 1 には、身体に装着した Notch と作成したモーションデータの一例を示す。ここで、各 Notch が異なる色を発光するのは、骨格モデルとの対応をとるためのものである。表 1 は、作成したデータセットの諸元を示す。学習・検証用のモーションデータは、次節に述べるニューラルネットワークを構築するために用いられる。評価用のモーションデータは、学習したニューラルネットワークを利用して推論した被験者の位置と移動速度を評価するためのものである。なお、評価用のモーションデータとして 0m と 10m 走行のモーションデータを採用しなかったのは、6m 走行のモーションデータには静止と移動する歩行の動作が含まれているからである。

[†] 岩手県立大学ソフトウェア情報学研究科

Graduate School of Software and Information Science,
Iwate Prefectural University

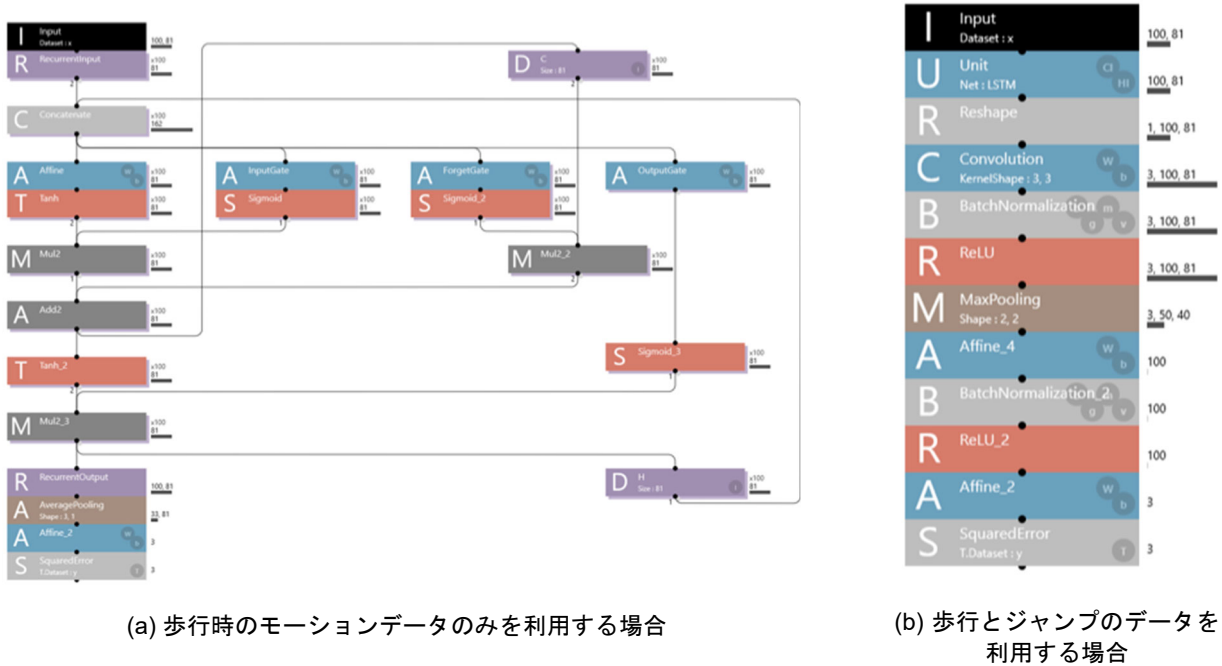


図 2 本研究で使用したニューラルネットワークの構成

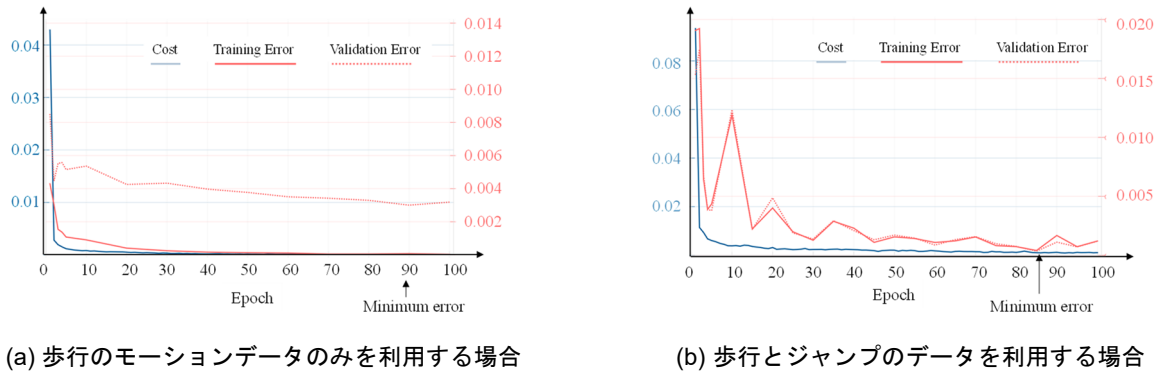


図 3 学習曲線

4. 局所座標系から大局座標への変換

本研究では、様々な動作におけるモーションデータを大局座標で表現するために、LSTM を利用して身体の各関節のモーションデータと身体の移動速度との関連付けを試みる。具体的には以下に示す。

a. 歩行のモーションデータのみを利用する場合

通常の歩行におけるモーションデータを Notch で取得し、LSTM の学習データとする。ただし、LSTM による推論を効率化するために、データのサンプリングレートを 500Hz から 100Hz にリサンプリングする。なお、身体の移動速度は Notch による推定したものを利用する。図 2(a) には、ニューラルネットワークの構成を示す。

b. 歩行とジャンプのデータを利用する場合

歩行とジャンプの動作を含む動画データセットに対して、3 次元人物姿勢推定を行い、身体の各関節のモーションデータを作成する。ここで、身体の移動速度は動画のフレー

ム長とその再生時間から求められる。これらのデータを利用して上記 a. の学習済みのニューラルネットワークを転移学習する。図 2(b) には、ニューラルネットワークの構成を示す。ブロック U (Unit) は学習済みのニューラルネットワークの該当部分である。

学習では Sony Neural Network Console (NNC) を利用し、最適化アルゴリズムを Adam、バッチサイズを 128、Loss 関数を Squared Error とした。学習環境としては、Intel Core i9-9900 CPU 3.10GHz、64GB メモリ、そして NVIDIA GeForce RTX 2080Ti GDDR6 11GB×2-SLI である。図 3 には、歩行のモーションデータのみを利用する場合 (a) と、歩行とジャンプのデータを利用する場合 (b) のそれぞれの学習曲線を示す。前者は Epoch 90 で、後者は Epoch 85 でバリデーションエラーが最も小さい。本研究では、これらの Epoch の時点でネットワークパラメータを保存し、それぞれの学習済ニューラルネットワークを出力した。

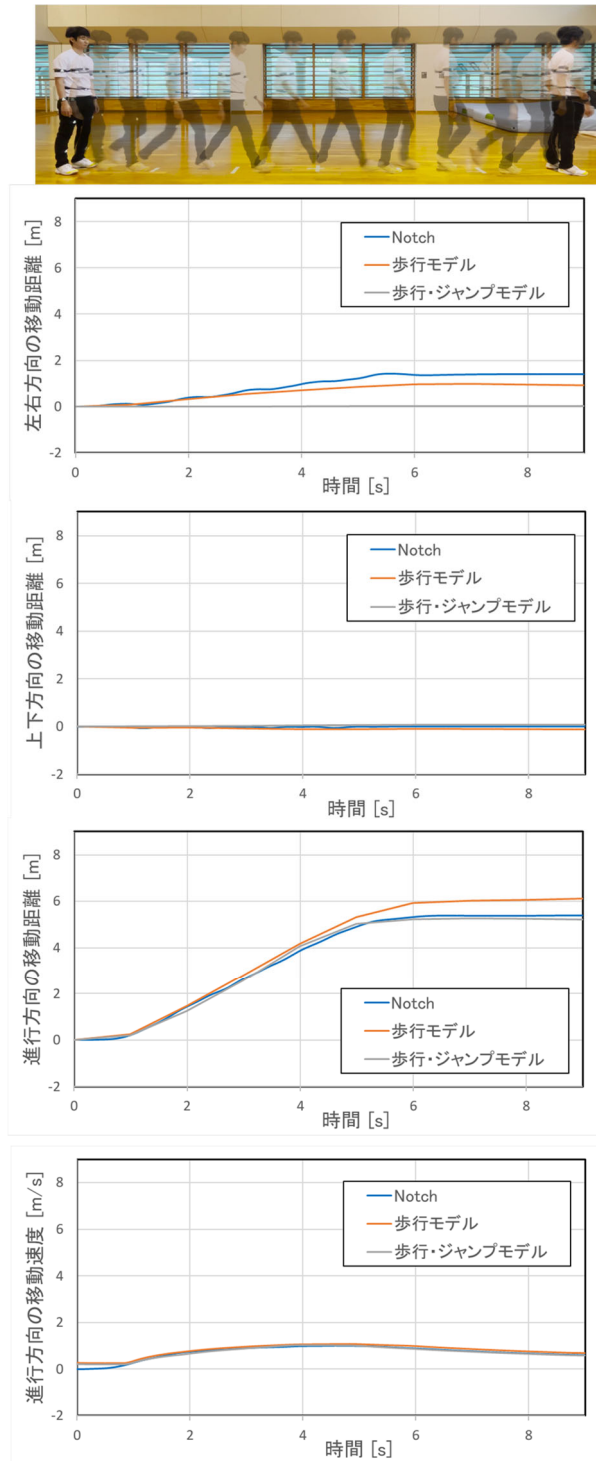


図 4 歩行時の被験者の位置と移動速度

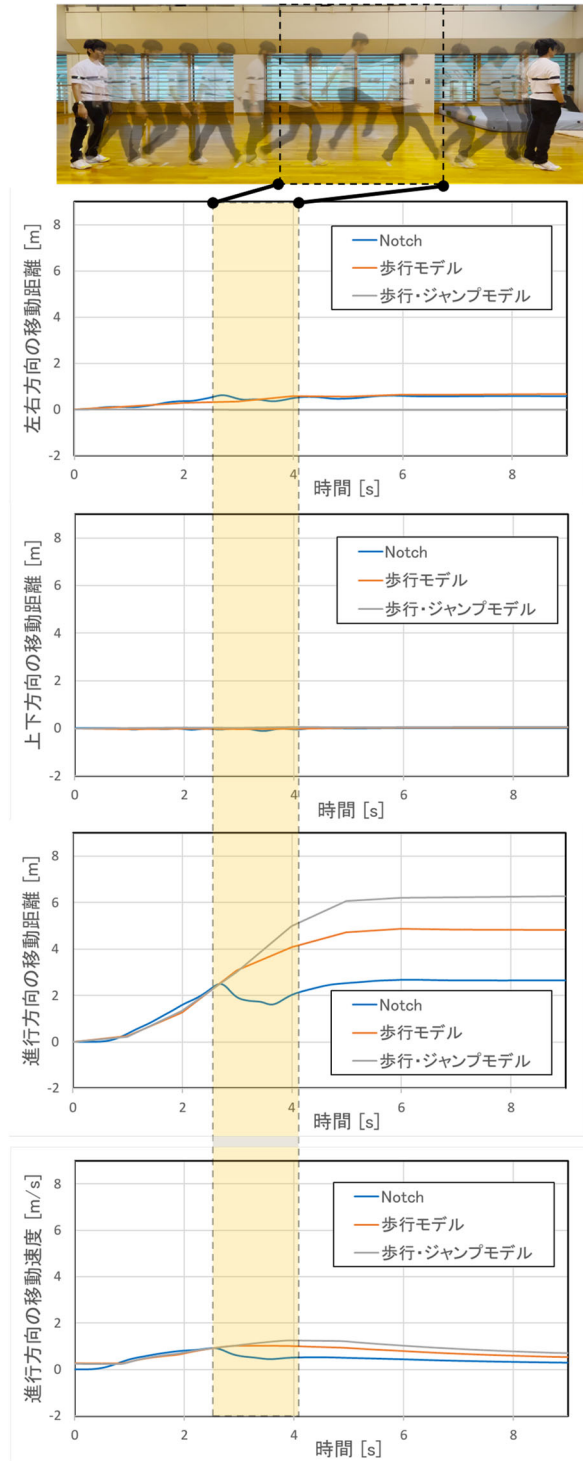


図 5 歩行とジャンプ時の被験者の位置と移動速度

5. 結果

本研究で作成したニューラルネットワークを用いて、評価用のモーションデータから被験者の位置と移動速度の推論を行った。図 4 と図 5 には、被験者の位置と移動速度の一例を示す。Notch によって得られた結果に対して、本研究で構築したニューラルネットワークによって推論された結果について次のようなことが分かった。

まず、図 4 のような足元が地面に接している場合の運動において、被験者の位置と移動速度について顕著な差が見られなかった。ただし、本評価用は 6m の直進歩行のモーションデータであるため、歩行モデルのよって被験者の位置がより正確に推論されることが分かった。一方、左右方向の移動について、歩行・ジャンプモデルが最も直進歩行

を表しているが、これには当該モデルの学習データが左右方向のモーションをゼロとしたことによるものである。

次に、図 5 のような歩行の間にジャンプが発生する場合の運動において、Notch の結果と本研究のモデルの結果との間に顕著な差が見られた。図中の点線で囲まれた箇所はジャンプが発生する区間を示している。当該区間の進行方向のグラフから、Notch がジャンプによって変化する被験者の位置を計測できないことが分かった。一方、本研究のモデルは、ジャンプの間にも被験者の移動を推論できることが分かる。さらに、歩行・ジャンプモデルから推論した結果から、6m の走行区間を正確に推論することが分かった。また、移動速度においても、本研究のモデルによってジャンプ中における速度の低下がないことが確認できた。

6. おわりに

本研究では、LSTM を利用して慣性型 MoCap を改善して被計測者の位置と移動速度を表現できる大局座標モーションデータを実現した。LSTM の学習には、歩行のみとジャンプを含めた歩行のモーションデータを利用したが、どの学習データを用いても、ジャンプの間にも MoCap 被計測者の移動を推論できることを確認した。ただし、ジャンプを含めた歩行のモーションデータを学習データとして利用すれば、より正確なモーションデータを出力できることが分かった。このことから、様々な運動のモーションデータを学習データとして収集できれば、より多様な運動に対応できるモーションデータの出力を期待できる。今後も、本提案のアプローチは現実社会で実用的に利用できるように、さらに研究を続けていきたい。

参考文献

- [1] 吉田 拓人, 野崎 惇登, 廣井 慧, 梶 克彦, 米澤 拓郎, 河口 信夫, “歩容に頑健な PDR のための 深層学習を用いた歩行速度推定手法”, 研究報告ユビキタスコンピューティングシステム (UBI), 2019-UBI-63, 3, pp. 1-7 (2019)
- [2] T. Wakaizumi and N. Togawa, “Carrying-mode Free Indoor Positioning Using Smartphone and Smartwatch and Its Evaluations,”
- [3] Journal of Information Processing, 30, pp. 52–65 (2022)
- [4] Notch Interfaces Inc, <https://wearnotch.com>, 確認日(2022/06/23)
- [5] C. Lugaresi et al., “MediaPipe: A Framework for Perceiving and Processing Reality,” Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR), pp. 1-4 (2019)
- [6] V. Bazarevsky, I. Grishchenko, K. Raveendran, T. Zhu, F. Zhang, and M. Grundmann, “BlazePose: On-device Real-time Body Pose tracking,” CVPR Workshop on Computer Vision for Augmented and Virtual Reality, pp. 1-4 (2020)