

## センサデータと骨格データを組み合わせた作業行動推定手法の提案

小村皓大<sup>†</sup> 佐藤祐紀<sup>†</sup> 猪股一步希<sup>†</sup> 堀川三好<sup>†</sup> 岡本東<sup>†</sup>岩手県立大学大学院ソフトウェア情報学研究科<sup>†</sup>

## 1. はじめに

近年、生産物流現場においてウェアラブルやカメラデバイスから得られる情報をもとに、作業者の行動推定を行うことが普及している。ここで、作業者の行動分析とは、モーションデータや姿勢推定から作業の種類、場所や時間を可視化することを意味する。しかしながら、現場で収集されるデータは、欠損等を含み不正確なものが多く、行動推定において精度向上のための工夫が必要である。

本研究では、簡易センサデバイスからモーション・位置データ、動画から骨格データを取得し、マルチモーダル学習を用いた作業者の行動推定を提案する。動画から得られる骨格データの収集においては、工場内の障害物や作業者同士の交差により欠損データが多く生じる。また、センサデバイスからは、安定したデータ取得が可能なものの、詳細な行動推定を行うためには複数センサを体の各部位に装着する必要がある。

本稿では、簡易センサデバイスを作業者に1つ装着して得られるセンサ・位置データ、および動画から得られる骨格データを用いて行動推定を行う。骨格データは、グラフ構造を用いた機械学習モデルを適用する。提案手法の有用性を検証するために行った実験結果から、簡易デバイスと動画から行動推定を行うことで、精度向上が見込まれるだけでなく、データの欠損等、不正確なデータが多い場合にも、精度低下を抑制する効果があることを報告する。

## 2. 関連研究

簡易センサデバイスを用いた作業者の行動推定の先行研究として、猪股らは、動作情報のみでなく位置情報を考慮した行動推定を提案<sup>[1]</sup>し、精度向上を報告した。しかしながら、膨大なセンシングデータの訓練データを作成するための工夫が課題となり、動画からのアノテーションを必要とした。

Proposal of a Method for Estimating Worker Behavior by Combining Sensor Data and Skeletal Data

<sup>†</sup>Kodai Komura, <sup>†</sup>Ibuki Inomata, <sup>†</sup>Yuki Sato, <sup>†</sup>Mitsuyoshi Horikawa, <sup>†</sup>Azuma Okamoto

<sup>†</sup>Graduate School of Software and Information Science, Iwate Prefectural University

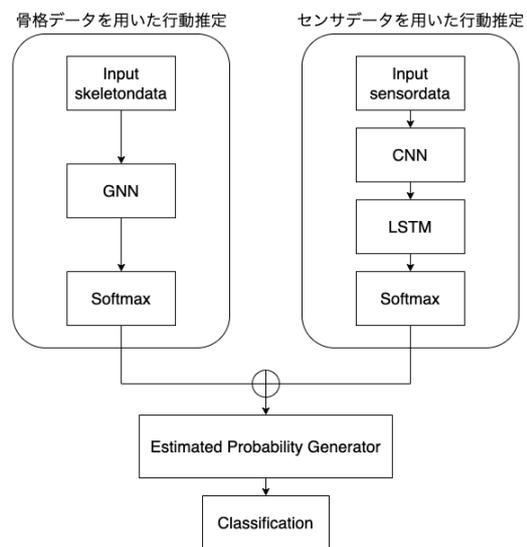


図1 行動推定モデル

また、近年、動画を用いた行動推定は、骨格データを元にグラフ構造化し、機械学習に適用したGNN (Graph Neural Networks) が注目されている。特に、時系列を考慮したモデル (GC-LSTM) や複数の重みや係数から結果を結合するモデル<sup>[2]</sup> (GAT) 等が高い精度を実現している。

## 3. 作業者の行動推定

## 3.1 収集データ

本研究では、センサデータの収集のために、㈱イーアールアイと開発したスマートタグを用いる。スマートタグは、約300ミリ秒毎に位置 (周辺ビーコンの電波強度) と動作情報 (加速度・ジャイロ値) の収集が可能である。電波の性質上、必ずしも一定間隔で収集はできないが、データが欠損する頻度は少ない。また、動画からの骨格データ抽出は、FPSに依存するが一定間隔で取得可能である。しかしながら、障害物等や骨格推定精度の問題からデータの欠損が多く生じる。

## 3.2 行動推定モデル

本研究で用いる行動推定モデルのアーキテクチャを図1に示す。

センサデータを用いた行動推定には、位置と動作情報の特徴量を入力値として、複数の機械学習モデルを比較した先行研究<sup>[1]</sup>を参考にCNN-

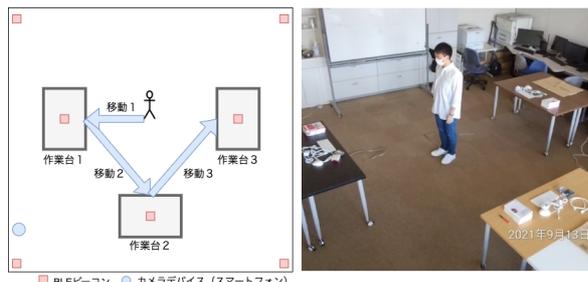


図2 実験環境

LSTMを用いる。また、骨格データを用いた行動推定として、複数の機械学習モデルを比較する。機械学習モデルとして、LSTM (Long short-term memory) と GNN の派生である以下の3つを取り上げる。

- GCN : Graph Convolutional Networks
- GC-LSTM : Graph Convolutional Long short-term memory
- GAT : Graph Attention Networks

Estimated Probability Generator では、各モデルからの推定確率  $P(skeleton)$ ,  $P(sensor)$  をもとに、両データの異なる取得間隔を考慮しつつ、推定係数  $w$  を用いて、新たな推定確率  $P(n)$  を生成し、行動を推定する。

$$P(n) = P(skeleton) \times w + P(sensor) \times (1 - w)$$

#### 4. 検証実験

##### 4.1 実験概要

生産現場を想定し、図2の環境で実験を行う。被験者はズボンの右ポケットにスマートタグを入れ、作業台への移動と各作業台にて「組み立て・梱包・検査」の計12手順(約5分)を1セットとして実施する。

##### 4.2 データセット

収集したデータセットは39であり、センサデータは約38,000件、骨格データは約253,000件である。また、収集された骨格データの欠損率は全体の17%であった。欠損データの比率による影響を確認するため、骨格データの欠損をランダムに発生させ、欠損率が30%、50%、70%、80%、90%のデータセットを生成した。訓練データにおける行動ラベルは、撮影された動画から手動でアノテーションを行う。

##### 4.3 実験結果

リーブワンアウト法による精度比較を行った。評価尺度として、F値(適合率と再現率の調和平均)の平均値(以降、F値)を用いる。表1に提案手法の有無によるF値の一例、図3に欠損率による推定精度の推移を示す。また、センサデータを用いた行動推定はF値の平均値が0.69であった。欠損率が高い場合でも、本提案手法の精度

表1 提案手法の有無によるF値の比較

	骨格データのみでの行動推定		センサと骨格データを用いた行動推定(GAT)
	LSTM	GAT	
17%	0.57	0.8	0.898
30%	0.45	0.768	0.888
50%	0.415	0.738	0.878
70%	0.411	0.672	0.85
80%	0.407	0.524	0.794
90%	0.392	0.464	0.774

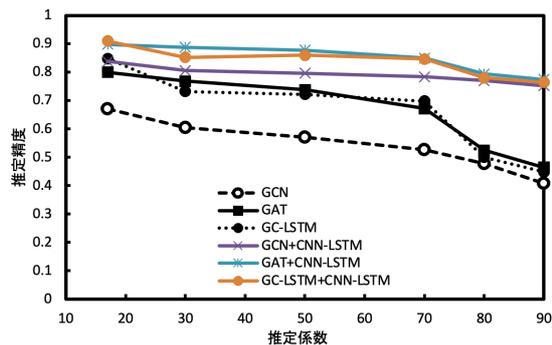


図3 欠損率による推定精度の推移

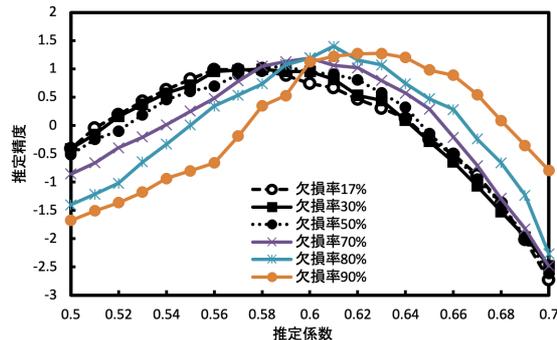


図4 推定係数  $w$  による推定精度の推移

が0.75を下回ることにはなかった。

3.2節にある数式の推定係数  $w$  による精度の違いについては、図4のように、骨格データの欠損率が大きくなるにつれ、推定精度が最大になる時の推定係数  $w$  も大きくなった。

#### 5. おわりに

本稿では、センサデータと骨格データを組み合わせた作業行動推定手法を提案した。検証実験では、本提案が骨格データの欠損の課題に対して、有用性があることが明らかになった。

今後は、実際の工場現場における実証実験を通して、現場で生じる骨格データの欠損パターンに、本提案が有用であることを検証していく。

#### 参考文献

- [1] 猪股一歩希他：“製造工程における作業者の行動分析手法の提案”，情報処理学会，第83回全国講演論文集(2021)
- [2] Veličković Petar 他：“Graph Attention Networks”，ICLR 2018. 12 pages, 2 figures (2018)