

ウェアラブルセンサーと行動認識モデルによる組み立て作業習熟度の可視化

鈴木悠斗¹⁾ Dharma Pranata²⁾ 中野貴行²⁾ 青木 崇浩¹⁾²⁾
山本泰生²⁾ 西村雅史¹⁾³⁾

1. はじめに

自動化が進む現代工場においても、未だ人手作業は数多く残されており、今後も必要な要素である。そして人手作業の課題として技能伝承があげられる。この課題に対して作業現場では、作業初心者に対する研修をより効率的に行うために、作業に向いている人や習熟が速い人を判断する仕組みが求められている。

本研究では慣性計測装置(IMU)やマイクといったウェアラブルセンサーから得られる特徴量に対して組み立て作業行動の認識を行う。作業行動認識には画像データが用いられることが多いがウェアラブルセンサーを装着し、情報を収集する手法は、画像データが捉えることができない細かな動きの違いを捉えることができる[1]。

被験者が作業に慣れるに従い、行動が安定することはセンサーデータ自体にもある程度現れるが、あらかじめ習熟した動きをモデル化しておけば出力から得られる尤度を主な指標として、次元削減のアルゴリズムである t-SNE(t-Distributed Stochastic Neighbor Embedding)を適用し、より詳細に習熟度を可視化することができることが先行研究として分かっている[2]。しかし、先行研究では作業習熟度の可視化にとどまり客観的に数値化することはできていなかった。そのため本研究では新たな指標を導入しこの値が作業習熟度を数値化するのに有効であるを検証する。また、行動作業行動認識において音情報と慣性情報を用いた場合と同等の性能が出ている骨格点の特徴量と比較をすることによりウェアラブルセンサーが習熟度の可視化に有効的であるか検証を行う。

2. 作業行動認識処理のフロー

作業行動認識のフローを図 1 に示す。

作業者は両手首と頭部に IMU を、右手首にマイクを装着する。また、作業の様子は 3 台のカメラ(正面、右、左)を用いて撮影をする。IMU から収集されるデータは 10msec 毎で x, y, z の 3 軸の加速度(3次元)と角速度(3次元)の計 18 次元(6次元+3次元)の特徴量が得られる。マイクから得られた音響特徴量はメル周波数ケプストラム分析を行い、12次元の MFCC を 10msec 毎に抽出する。それぞれ得られた特徴量を MS-TCN(Multi-Stage Temporal

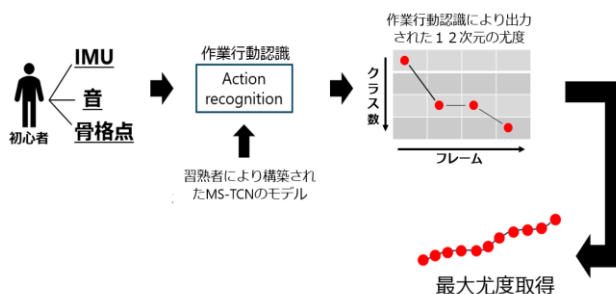


図 1 作業行動認識処理のフロー

Convolutional Network)[3]に入力し、10msec のフレーム単位で行動認識を行う。対象作業では 12 種類の作業に対して認識を行う。MS-TCN による作業推定は入力データに対しフレーム毎に各行動クラスの尤度を出力する。そのため、本研究では MS-TCN の出力は 12 次元の対数尤度系列が出力される。そして出力された、12 次元の対数尤度の系列から最大となる対数尤度のパスを取得し、これを習熟度の定量化指標として扱う。

3. 習熟度の可視化

本研究で習熟度可視化に用いる t-SNE は学習データにより得られる値が変わってしまい、客観的に数値化することが難しい。そのため本研究ではどのような空間にマッピングされていても共通の数値で習熟度を示すために t-SNE の値を正規化する。

3.1 習熟時データの重心からの距離(=SD)

対数尤度に対し t-SNE を適用し正規化した例を図 2 に示す。図 2 の作業回数は水色(作業 1 回目)から紫色(作業 40 回目)まで順に表している。また、左の図は学習データの分布を表しており、右の図はテストデータの分布を表している。また、右の図において習熟時の重心(学習データにおける作業回数 30~40 回の重心)を黒点で表している。習熟時データの重心とテストデータのユークリッド距離(Skilled Distance: 以下, SD)を求め、SD と作業回数

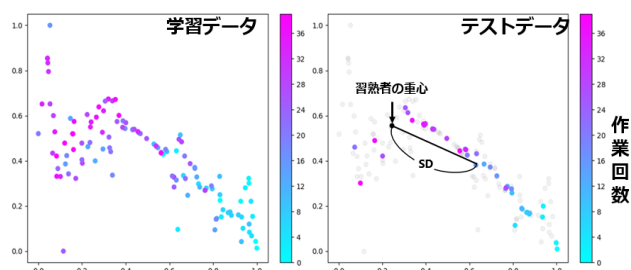


図 2 慣性特徴量を使用した際の t-SNE による可視化例と SD (テストデータと習熟時データの重心との距離)

1) 静岡大学情報学部

Graduate School Of Integrated Science and Technology, Shizuoka University

2) ヤマハ発動機株式会社生産技術部

Manufacturing Technology Div, Yamaha Motor Co., Ltd.

3) 愛知産業大学造形学部

Faculty of Architecture and Design, Aichi Sangyo University, Okacho, Okazaki-shi, Aichi, 444-0005, Japan

の相関関係を調べた。

4. 実験

実際の工場での新人教育用の組み立て作業工程をベースとし、12種類の行動からなる部品組み立て作業を設計した。そして20代男女4名(A, B, C, D)が各40回作業を実施した。習熟モデルあるMS-TCNの構築は作業員3名分の30~40回目の作業を習熟したデータとして計30データを使用し学習する。テストデータは学習に使用しなかった作業員1名をテストデータとして使用する。また、実験はウェアラブルセンサーの有効性を検証するために固定視座3台のカメラから画像認識により取得した42次元の骨格点の特徴量と比較し、被験者A, B, C, Dに対し交差検証を行う。

4.1 作業行動の認識精度

作業員(A, B, C, D)に対し交差検証を行い、MS-TCNの作業行動の認識精度をフレーム単位のF値で評価した結果を表1に示す。

表1 各特徴量の認識精度比較

特徴量	F1-score
音(12次元)	0.806
慣性(18次元)	0.704
骨格(42次元)	0.895

4.2 SDと作業回数との相関関係

SDが習熟度を現すことができるか分析を行うために作業回数とSDの相関関係について検証した各特徴量を使用し、作業員(A, B, C, D)に対して作業回数10回毎にSDの平均を取りSDの変化を表した図を図3に示す。

また作業回数とSDの相関係数を表2に示す。

表2 SDと作業回数との相関係数

特徴量	A	B	C	D	平均
慣性(18次元)	-0.581	-0.683	-0.867	-0.875	-0.752
音(12次元)	-0.419	-0.735	-0.864	-0.899	-0.729
骨格(42次元)	0.632	-0.441	-0.192	-0.560	-0.140

5. 考察

表2から慣性情報と音情報を使用することにより作業回数とSDの間で負の相関関係を示した。

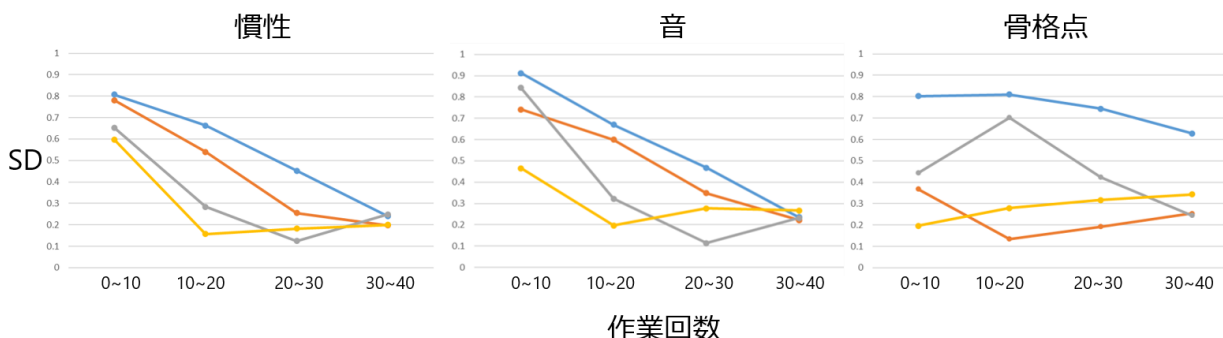


図3 各特徴量における作業回数毎のSDの変化

同様に図3では慣性情報と音情報を使用した場合にはSDの値が0.3以下になるとSDは安定する傾向がみられた。特に作業員Aに着目するとSDと作業回数の相関関係は見られないが、図3では作業初修時(作業0~10回)のSDの値が大きく作業初修以降(作業10回以降)ではSDの値が0.3以下で安定していることが分かる。このことより作業員A他の対象者より習熟が早い推測できる。これらのことよりSDという指標は作業習熟度の客観的な数値として有効であると考えられる。

一方、表2と図3の両方において骨格点を使用した場合にはSDによる習熟度可視化を行うことができていないことが分かる。また、図3の骨格に着目するとSDの値は作業回数に関わらず一定である傾向がある。また、表1において骨格点を使用した場合には他の特徴量と比較し高い認識精度を示している。これは画像データによる骨格点の特徴量を使用した行動認識では大まかに作業行動単位で特徴を捉えることはできるが、習熟時と初修時点の間にある作業行動内の細かな動きの違いを捉えることができなかつたためと考えられる。

6. おわりに

t-SNEの分布上においてSDという指標を導入したところ、作業回数との相関が大きく、この値が作業習熟度を数値化するのに有効であることがわかった。また、今回使用した慣性計測装置(IMU)やマイクといったウェアラブルセンサーを作業員に装着することにより、画像データを用いて抽出した骨格点を使用する場合より作業員の細かな変化を捉え習熟度を効果的に可視化できることがわかった。

参考文献

- [1]Tomoya Hirade ,et.al. “Visualization of the Proficiency Level of Assembly Work Using Inertial and Sound Sensors” 2024 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing
- [2]Enrique Garcia-Ceja, et.al. “Multi-viewstacking for activity recognition withsound and accelerometer data”,Inf.Fusion.40, pp45-46, 2018
- [3]Yazan Abu Farha ,et.al. “MSTCN:Multi-Stage Temporal ConvolutionalNetwork for Action Segmentaion”, IEEEConference on Computer vision and PatternRecognition(CVPR),2019.