

# 慣性情報と音情報を用いた組み立て作業行動の異常検知

平出知也<sup>†</sup> 塩野由紀<sup>‡</sup> 白澤怜樹<sup>‡</sup> 中野貴行<sup>‡</sup>

青木崇浩<sup>‡</sup> 山本泰生<sup>†</sup> 西村雅史<sup>†</sup>

静岡大学総合科学技術研究科<sup>†</sup> ヤマハ発動機生産技術部<sup>‡</sup>

## 1. はじめに

製造現場では、生産ラインの自動化が進む一方で、作業品質の保持、障害対応などを目的として人手での作業が数多く残されている。生産ラインにおいて、製品の品質不良などが発生し欠陥品などができると、顧客の信頼の喪失、欠陥品の廃棄といった大きな損害が発生する。少なからず発生するヒューマンエラーを防止するため、製造現場では、作業者の異常作業行動を検出する技術が求められている。

作業行動の認識には画像ベースの手法が用いられることが多い。対して、対象者にセンサを装着し、情報を収集するセンサベースの手法[1]は、行動認識で利用されることは少ないが、画像に比べてデータサイズが小さいことや、画像で追いきれない対象者の細かな動きを観測できるメリットがある。

本稿では慣性計測装置(IMU)で得られる慣性情報に加え、作業中に発生する音を特徴量として捉え、作業者の行動を、DNNを用いて自動推定すると共に、定型作業においてしばしば発生する「作業飛ばし」を異常作業行動として自動検知する方法について検討を行う。

## 2. 提案手法

提案手法の概要を図1に示す。

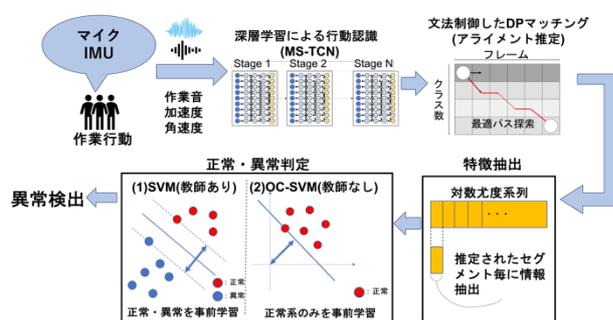


図1 提案手法の概要図

## 2.1 MS-TCNに基づく作業行動認識

作業者は両手首と頭部にIMUを、右手首にマイクを装着しており、3台のIMUから3軸の加速度と角速度(合計6次元×3台)を10msec毎に取得する。一方、マイクから得られる音響特徴量は、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。これらを統合させた計30次元の特徴をMS-TCN(Multi-Stage Temporal Convolutional Network)[2]に入力し、10msecのフレーム単位で行動認識を行う。ここでは、12種類の作業行動を認識対象としているため、12次元の対数尤度系列が出力される。

## 2.2 アライメント推定とセグメント内情報抽出

MS-TCNから出力された12種類の行動の対数尤度系列に対して、動的計画法(DP)を適用し、事前に用意した正常行動の文法に沿った行動分節情報(アライメント)を推定する。この情報に基づき、各行動セグメント内の尤度関連情報(平均・分散・最大値・中央値・最小値)を抽出し、これを正常・異常の特徴付けに用いるものとする。本稿では、(a)推定されたアライメントに従って、行動分割する方法、(b)全行動データを時間等分割する方法、(c)行動分割を行わない方法の3つの方法を比較する。以下に行動分割の概要を図2として示す。

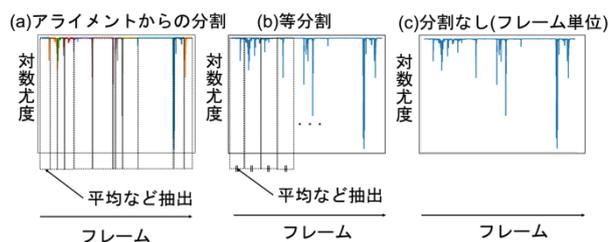


図2 行動分割の例。(破線が分割位置を示す)

## 2.3 異常作業の自動検出

各行動セグメントから抽出した特徴量を用い、(1)事前に高頻度に起きる異常パターンも学習しておく教師ありの手法(SVM)と、(2)正常パターンのみを事前学習し、異常行動の自動検出を行う手法(One-Class SVM)の2通りの方法について性能評価を行う。

Abnormality Detection of Assembly Operations Using Inertial and Sound Information

Tomoya Hirade<sup>†</sup>, Yuki Shiono<sup>‡</sup>, Takayuki Nakano<sup>‡</sup>, Takahiro Aoki<sup>‡</sup>, Yoshitaka Yamamoto<sup>†</sup>, Masafumi Nishimura<sup>†</sup>

<sup>†</sup>Graduate School Of Integrated Science and Technology, Shizuoka University

<sup>‡</sup>Manufacturing Technology Center, Yamaha Motor Co.,Ltd.

### 3. 実験

#### 3.1 対象作業データ

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。一連の作業には平均約100秒、各行動には平均1~30秒の時間を要していた。20代男女4名が作業をそれぞれ50回実施し、合計200回分の正常データを収集した。また、異常作業データとして対象者4名の内2名が、図3に示す動作を飛ばした作業をそれぞれ10回ずつ、合計で60回分を収集した。

正常作業 : 0 1 2 3 4 5 6 7 8 9 10 11

異常箇所 : 0 1 2 3 4 5 6 7 8 9 10 11  
(作業飛ばし)



図3 作業飛ばしの対象とした動作例

#### 3.2 作業行動の認識性能

MS-TCNでは、正常データのみを用い、学習データを3名分、評価データを1名分として、交差検証を行い、フレーム単位のF値で評価した(表1)。

表1 MS-TCNによる作業認識性能

使用した特徴量	F1-score
音 + (加速度+角速度:3箇所)	0.91

#### 3.3 異常行動の検出性能

異常検知の従来手法として、文法制約を与えたDPを行った場合に得られる累積尤度を用いた手法が考えられる。この累積尤度を用い、ホテリング理論に基づき閾値を設定し、異常検知を行う。これを従来手法とする。

異常検知実験において、2節提案手法において述べた、方法1(教師あり:SVM)、方法2(教師なし:One-Class SVM)と、行動分割法(a), (b), (c)の全ての組み合わせで性能評価を行い、従来手法と比較する。評価指標は、正解率、適合率、再現

率、F値を用いて評価する。以下に、異常検知実験の実験結果を表2に示す。

#### 3.4 考察

累積尤度のみを基準とする場合に比べ、提案手法は遥かに高い性能を示した。また、高頻度で起きる異常をあらかじめ想定し、異常行動もSVMで事前学習した場合が当然ながら最も異常検知性能が高い。また、この場合は事前の行動分割処理が重要で、等分割や分割なしでは十分な性能が出ていない。一方、異常行動の事前学習を行わないOne-Class SVMベースの方法でも、等分割と組み合わせた場合に、先の教師ありの方法に近い性能を出せることがわかった。なお、ここでは正確な行動分割推定は等分割よりも悪影響を与える結果となっているが、これは異常行動の分割位置推定の影響を事前に学習できないからだと考えている。

#### 4. おわりに

音情報と慣性情報を組み合わせてMS-TCNに入力することで高い精度で作業行動を分節化できることを確認した。異常検知において、誤りパターンを想定できない場合にも、等分割処理を行えば、分割誤りの影響を大きく受けないこともあり、結果として、誤りパターンの事前学習なしでも、教師ありの場合に近い異常検知性能を達成できることがわかった。今後は教師なし、教師ありの両面において、更なる性能改善を目指す。

#### 参考文献

- [1] Enrique Garcia-Ceja, et.al., “Multi-view stacking for activity recognition with sound and accelerometer data”, Inf.Fusion. 40, pp45-46, 2018
- [2] Yazan Abu Farha, Juergen Gall, “MS-TCN:Multi-Stage Temporal Convolutional Network for Action Segmentation”, IEEE Conference on Computer vision and Pattern Recognition(CVPR), 2019.

表2 異常検知実験結果

モデル	手法	特徴量	適合率	再現率	F値	正解率
		累積尤度	0.67	0.12	0.20	0.54
SVM	(a)アライメント	max	0.82	0.92	<b>0.86</b>	<b>0.85</b>
OC-SVM	(a)アライメント	mean	0.55	1.00	0.71	0.78
SVM	(b)10分割セグメント	mean	0.53	1.00	0.69	0.55
OC-SVM	(b)10分割セグメント	mean	0.75	0.94	0.83	<b>0.85</b>
SVM	(c)セグメントなし		0.13	1.00	0.24	0.57
OC-SVM	(c)セグメントなし		0.00	0.00	0.00	0.50