

骨格検知技術を活用した標準作業の自動判別

西畑宏祐^{*1}, 山田理華子^{*1}, 野村正明^{*2}, 福庭広巳^{*3}, 玉井康弘^{*4}

Automatic Identification of Standard Work Using Skeletal Detection Technology

Kosuke Nishibata^{*1}, Rikako Yamada^{*1}, Masaaki Nomura^{*2}
Hiroshi Fukuba^{*3}, Yasuhiro Tamai^{*4}

要旨

豊田合成では工場の自動化が進む中で人作業工程も多く残っている。この残存する工程における改善活動は、現地現物が基本であるが、すべての作業を人が確認することは工数の観点から困難であった。

そこで本稿では人の動きを定量化する骨格検知技術を活用し自動で作業を判別することで、監督者の作業確認工数を削減する技術を開発したので報告する。

Abstract

While factory automation has been progressing at Toyoda Gosei, many human-operated processes remain. Kaizen activities in the remaining human-operated processes are based on the “genchi-genbutsu” (actual work on site) method, but from the viewpoint of person-hours it has been difficult for humans to check all the standard work.

In this paper, we report on the development of a technology to reduce the person-hours required for supervisors to check work tasks by automatically identifying tasks using a skeletal detection technology that quantifies human movements.

1. はじめに

豊田合成では、先進国の少子高齢化や新興国の労務費高騰といった社会情勢に対応するため、IoT/AI活用による生産現場のモノづくり革新に向けた活動を進めている。近年の半導体の影響による生産数の増減やコロナ禍の影響で人の入れ変わりも増加しており、効果的な人作業の品質維持・向上が求められている。その中で、新たなアプローチとして、人の動きを定量化する骨格検知技術を活用した開発を進めている。本稿では骨格検知技術を活用し、標準作業が正しく行われているか否かを自動で判別する技術を確認し、作業改善に繋がったので報告する。

2. 背景

2-1. 対象製品

カーテンエアバッグは、衝突の際に側面の窓を覆うように膨らみ頭部を保護する機能があり高い品質が求められている。樹脂リテーナー、エアバッグ、テープで構成され(図-1)、人作業で組み付けを行い、最後に検査機で確認することで品質維持を行っている。

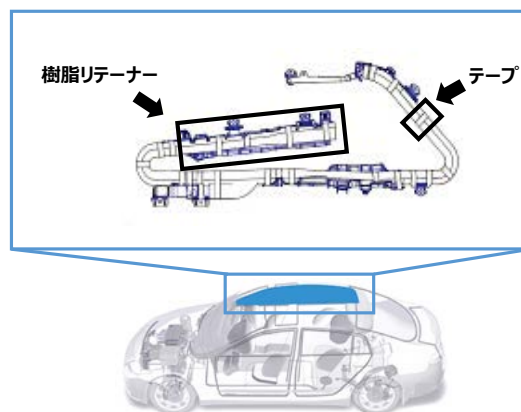


図-1 カーテンエアバッグ

*1 生産技術統括部 IoT 推進室

*2 マシンエンジニアリング部 設備開発室

*3 SS 製造部 エアバッグ第3課

*4 SS 製造部 製造技術室

2-2. 工程の課題

製造現場では、人作業に対する品質の維持・向上の取り組みとして、管理監督者が定期的を目視での作業確認を実施している。近年、人の入れ変わりが増えたことで、新人技能員が増え、監督者による作業確認工数増加が問題となっていた。しかし、作業を確認するために監督者が現地で目視確認する必要があるため、工数の低減が課題となっていた。

3. 対策の方針

監督者の作業確認工数増加は、目視でしか技能員の作業の善し悪しが判断できないことを原因と捉え、目視の代わりに自動で作業判別する技術が必要だと考えた。

そこで作業中の動きを定量化することで、サイクル毎に標準作業が正しく実施されていることを自動で判別することにより、監督者の工数を低減することを目標に下記3つのプロセスで開発を行った。

- 1) 作業サイクルの切り出し
- 2) 作業の定量化
- 3) 標準作業の自動判別

実工程での判別精度 90% 以上で運用を狙うものとした。

4. 作業サイクルの切り出し

技能員の作業を認識するにあたり作業サイクルの開始・終了のトリガを取得する必要がある。設備からのトリガを取得するには設備改造が必要であり、今後の展開性も考慮し設備信号の取得以外の手法を検討した。そこで組み付け作業の前後では部品を手を取るために、技能員の手が必ず特定の箇所を通ることに着目し、画像処理技術の活用を検討した。

画像は R, G, B の3つの値からなるピクセルが多数合わさって構成されている。そこで手が通りすぎる箇所に枠を設定し、その枠内の R, G, B 値の平均値を算出し、時系列データ化を行う。手が通り過ぎる際には、枠内の色合いが変わることから R, G, B の平均値も変化する。枠内の R, G, B 平均値の時系列データに対して閾値管理を行うことで、特定の箇所を手が通ったことの検知が可能となった。これを組み合わせることで、サイクルの切り出しのための開始・終了のトリガとして活用した(図-2)。

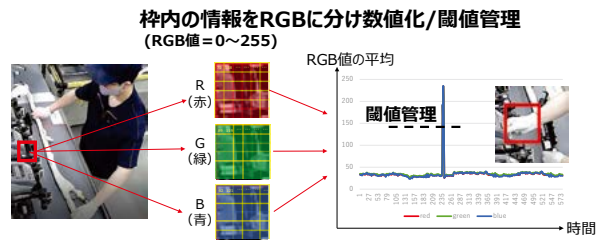


図-2 画像処理によるトリガ取得

5. 作業の定量化

5-1. 人の動きの定量化手法の検討

作業を自動で判別するためには、人の動きを定量化する技術が必要であるため、ベンチマークを実施した。マーカーを体の関節部位に取り付け、複数のカメラでその位置を捉えるマーカー方式、画像情報からあらかじめ学習したデータを元に、人の関節点の位置を予測する AI 骨格検知方式の2つの技術を検討したが、現場の技能員がマーカーを取り付けて作業をすることは現実的でないことから、カメラの設置のみで活用できる AI 骨格検知手法を選定した(表-1)。

表-1 人の動きの定量化技術について

	マーカー方式	AI 骨格検知方式
特徴	マーカーを撮影し、体、工具の動き(座標)を取得	人間の骨格を AI が予測し動き(座標)を取得
メリット	分解能 0.1mm で精度が高い	マーカーをセットする実用がなく、工程への適用が容易
デメリット	マーカーをセットする必要があり、工程で適用することは非現実的	マーカー方式と比べ、精度が低い※静止状態でも座標が変動

5-2. AI 骨格検知手法の検討

次に AI 骨格検知手法について、商用利用可能なオープンソースの調査を実施した。今回、Mediapipe¹⁾、YOLOv7²⁾、ViTPose³⁾ の3手法について着目し、検知可能な人数、関節部位の種類や処理速度から YOLOv7 を選定した(表-2)。

表-2 骨格検知 AI ベンチマーク

手法	検知部位		人物検知精度		骨格点検知精度		処理速度
	全身骨格	手指骨格	一人	複数人	一人	複数人	
Mediapipe	可	可	不可	不可	低	不可	△
YOLOv7	可	不可	高	高	高	高	◎
ViTPose	可	可	高	中	高	中	○

YOLOv7は入力画像内の人物の骨格17点の位置を予測し、画像の左上を原点としたx, y座標で出力する(図-3)。

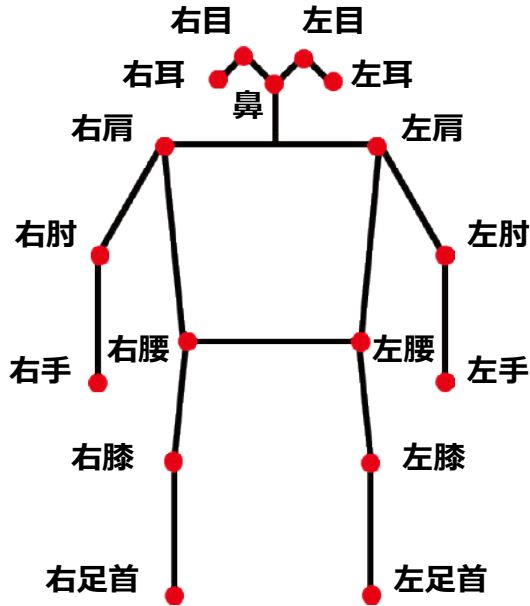
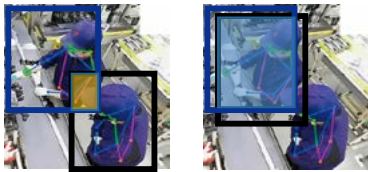


図-3 抽出できる骨格点

活用において、作業中の技能員以外の人物が画像内に映る場合に骨格点が混同する課題があった。そこで作業中の技能員を確実に検知し、追従するために、直前の人物位置からの変化が小さい方を紐づけることで対象の技能員の追従機能を追加し対応した(図-4)。

直前の技能員の位置(青枠)と重なる領域を比較



重なりが大きい(位置変化が小さい)方を技能員として紐づけることで、正確に技能員を追従



図-4 作業中の技能員の追従

次に、サイクルの開始から終了までの作業について、YOLOv7により骨格点を予測することで作業サイクル中の動きの時系列データ化を行った(図-5)。すべての作業は手元作業のため、上半身の姿勢が特に重要であることから両手首、両肘、両肩の骨格点を活用した。

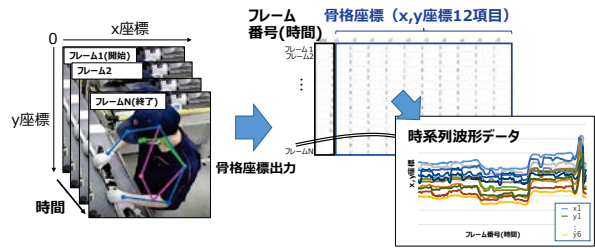


図-5 作業中の骨格点の時系列データ化

6. 標準作業の自動判別

6-1. 対象作業

カーテンエアバッグ組み付け工程において、樹脂リテーナーとエアバッグをテープで巻きつけて組み付けを行うが、巻き付けた後の剥がれ防止のために、テープ巻き後に擦るという作業がある(図-6)。手元作業で監督者が確認しづらく、工数を要するため、この作業を対象に検討を実施した。

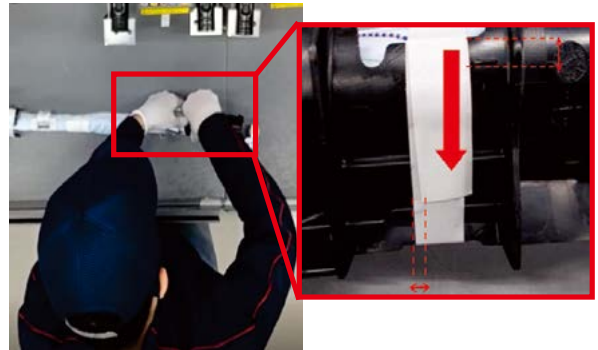
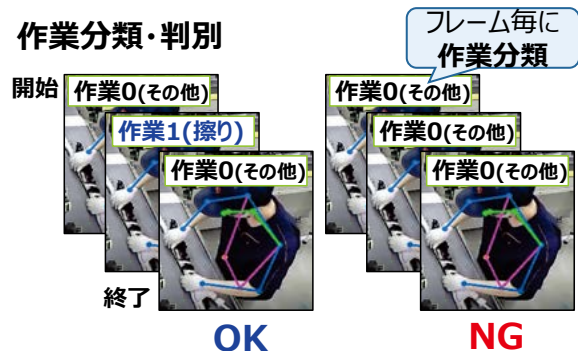


図-6 テープ擦り作業

6-2. 作業分類モデルの構築

骨格データを活用しフレーム毎に作業分類を行い、擦り作業の検知有無によって、OK/NGの判別を検討した(図-7)。

作業分類・判別



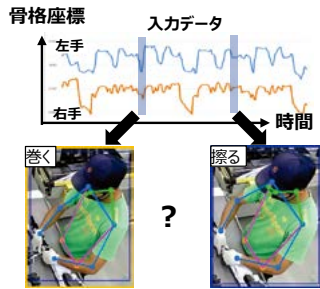
擦り作業の有無でOK/NG判別

図-7 フレーム毎の作業判別

ここで、組み付け工程では手元作業が多くあり、同様の姿勢でも別の作業を行うことがあり、瞬間的な姿勢だけでは判断がつかない。

そこで、少し前の動作を含めた一連の動作を考慮する必要がある。入力データを一定時間の骨格データとし（図-8）、フレーム毎に擦り作業か擦り以外の作業かの分類を行う時系列波形の2クラス分類としてAIを活用した作業分類モデルを構築した。

瞬間的には作業姿勢が似ており、作業の分類ができない



一連の動作で判断することで同様の姿勢でも分類可能

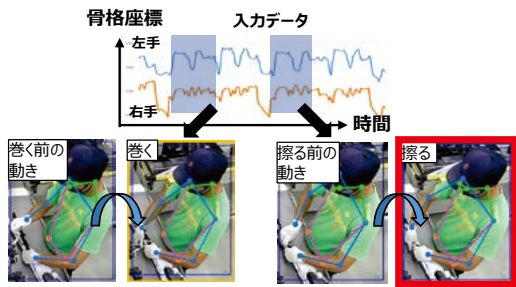


図-8 前後の動作で判別

作業分類手法としては時系列波形分類手法である1次元畳み込みニューラルネットワーク⁴⁾を活用した。擦り作業をラベル1、擦り以外の作業をラベル0としてフレーム毎にラベル付けを行った（図-9）。説明変数を過去一定時間（window幅）の骨格座標データとし、モデル構造を畳み込み層、プーリング層、全結合層、出力層で構成した（図-10）。

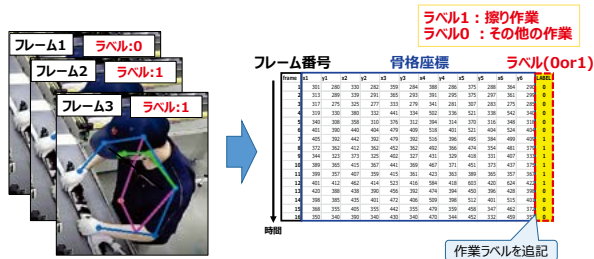


図-9 作業ラベル追記

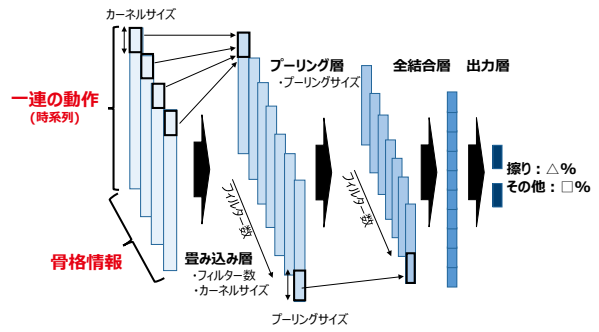


図-10 1D-CNNを活用した作業分類AI構築

特定の技能員の正常作業50サイクル分を収集し、42サイクルを学習データ、8サイクルをテストデータとしてモデル1の学習を行った。結果、精度（ACC）は88%であった。混同行列で評価したところ、擦り作業をうまく分類できていないものが約11%あった（表-3）。

表-3 モデル1精度

		予測	
		擦り作業	その他の作業
観測	擦り作業	18	42
	その他の作業	1	317

ここで、擦り作業以外の時に擦り作業と誤分類されたものは標準不履行の見逃しにつながるため、モデルの精度向上を検討した。誤分類のあった作業動画を確認したところ、作業自体は実施されていたが、作業位置のバラツキがみられた。そこで、作業位置バラツキへの対策として各骨格点を左肩基準の相対座標に変換を行い、再度モデル2の学習を行った（図-11）。



図-11 骨格点の相対処理

結果ACCは92%と向上したが誤分類は8%であったためさらなる精度向上を検討した（表-4）。

表-4 モデル2精度

		予測	
		擦り作業	その他の作業
観測	擦り作業	33	27
	その他の作業	2	316

作業自体に異常な部分はみられなかったため、モデル構造の改善を検討した。入力データから作業の特徴を学習しきれていないと考察し、畳み込み層、プーリング層を追加し（図-12）、再度モデル3の学習を実施した。結果、分類精度は98%となった（表-5）。

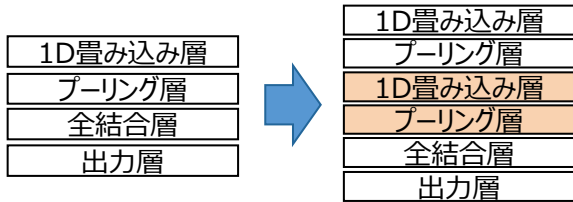


図-12 モデル構造の変更

表-5 モデル3精度

		予測	
		擦り作業	その他の作業
観測	擦り作業	55	5
	その他の作業	2	316

6-3. 作業分類モデルの運用

ここで、モデル精度は向上しているものの、誤分類の可能性は残っている。AIの予測精度100%を確保することは困難であることから、作業分類モデルの運用方法での精度向上を検討した。実際の擦り作業は瞬間的に行われるものではなく、一連の動作で成り立つことから、直近の複数フレームの分類結果を元に、最終的な擦り作業の実施有無の判別を検討した。作業分類モデルの結果が擦り作業と分類した場合でも、直近の複数フレームでその他の作業と分類されていれば、その他の作業として判定することで誤分類に対応する運用方法を制定した（図-13）。

直近の複数フレームの分類結果を元に最終的な擦り作業かその他の作業かを判定
例)326フレーム目の判定結果

フレーム番号	...	319	320	321	322	323	324	325	326	327	...
予測ラベル	...	その他の作業	その他の作業	その他の作業	その他の作業	その他の作業	その他の作業	その他の作業	擦り作業	その他の作業	...
判定(多数決)	その他の作業	その他の作業	その他の作業	その他の作業	...

直近の複数フレームでは「その他の作業」の予測が大多数なので、「擦り作業」と予測されたとしても「その他の作業」と判定

図-13 作業分類モデルによる判定手順

6-4. 量産工程での精度検証

以上のモデルと運用方法をもって、実工程でリアルタイムに精度検証を実施した。結果正解率93.5%となり、目標の90%以上は達成したがラボでの検討と比較し、実工程では精度が低下した（表-6）。

表-6 実工程での検証結果

		予測	
		擦り作業	その他の作業
観測	擦り作業	121	28
	その他の作業	1	300

正常な作業については検知できたが、テープ巻き中のもたつきなど、手首までの動きは同じでも、指先の動きが少し異なるときに誤判別があった。明らかな標準不履行は検知できることから、監督者の作業確認の工数削減可能となる技術を確立したが、実工程での細かな違いを識別する上では更なる精度向上の余地が残った。

7. 今後の進め方

実工程での判別精度を向上させるにあたり、指先の細かい違いも判別することが求められる。手指骨格の検出が可能なViTPoseの活用や、2視点からの骨格情報の活用などで精度向上を狙う。

また、自動化が進むとともに、日々作業改善が行われており、標準作業自体が変更することも考えられることから、短期間で高精度な作業分類モデルを構築することが求められる。今回の事例から、骨格点の相対座標変換や畳み込み層やプーリング層の追加がモデル精度に大きく寄与することから、前処理やパラメータ調整の重要性を感じた。今回蓄積した骨格点を活用した作業分類モデル精度向上のノウハウを元に、モデル構造や各種パラメータを網羅的に振り分けて複数のモデルを構築し、最も精度の高いモデルを採用することで、効率的に高精度なモデルを構築し、短期間で導入できるように標準化を進める。

8. まとめ

骨格検知 AI と作業分類モデルの 2 つの AI を組み合わせ、量産工程で 90% 以上の精度で標準作業の自動判別が達成できたことで、人作業に関わる様々な用途に活用できる基盤技術を確立できた。また 100% の精度を確保することが難しい AI を実工程で活用するためには、運用面での工夫など、適切に扱うことが重要であると感じた。

自動化が進む中で、今後も人作業工程は残ることが考えられるため、骨格検知技術を活用したシステムの開発・展開を進める。

参考文献

- 1) Mediapipe,
<https://github.com/google/mediapipe>
- 2) YOLOv7,
<https://github.com/WongKinYiu/yolov7>
- 3) ViTPose,
<https://github.com/ViTAE-Transformer/ViTPose>
- 4) François Chollet, 「Python と Keras によるディープラーニング」, マイナビ出版, 2018 P236-244

著 者



西畑宏祐



山田理華子



野村正明



福庭広巳



玉井康弘