

染色加工スケジュールリング工程における動画像を用いた行動認識

三村 亮[†] 柴田 楓也[‡] 手取屋 友雅[‡] 上田 芳弘[‡] 坂本 一磨[‡] 池田 理玖[†]
 公立小松大学大学院サステイナブルシステム科学研究科[†] 公立小松大学生産システム科学部[‡]

1. はじめに

画像処理技術や、センシング性能の向上などのデジタル技術の発展により、多種多様なデータの取得や、蓄積、分析などが可能となった。また、製造業などの工場内のデジタル化も推進され、IoT化や機械学習を用いた自動化による業務効率化も進められている。既存研究では、RFID (Radio Frequency Identification) を用いて、取得した位置データから非効率な作業の分析や不審者の発見を行う研究[1]、モーションキャプチャーと機械学習を用いて作業者の行動分析を行う研究[2]が行われている。

多くの製造業では、人手による作業工程が未だに多く存在する。人手に依存した作業では、人為的なミスによる作業ロスが生じるなど、生産性の向上を行うことは難しい。また、突発的な受注や生産計画の変更に対応が困難であるという課題がある。現在、共同研究を進めている繊維加工企業では特に、生産のボトルネックとなっている染色加工におけるスケジュールリング工程の改善が重要となっている。しかし、作業分析には、多量のデータが必要なため、人手によるデータ収集は困難である。そのため、機械学習を用いた作業分類によるデータ収集の効率化が有効である。また、画像認識では、対象人物以外の背景が行動認識に影響を与える問題[3]がある。そこで、本研究では、カメラにより撮影した作業中の人物の動画像から MMpose[4] による骨格推定を行い、得られた骨格情報を基に LSTM[5] による作業分類を行う。

2. 提案手法

本システムの概要を図1に示す。本提案手法は、骨格推定機能、作業特徴学習機能、作業分類機能で構成される。

2.1 撮影方法

スケジュールリングにおける作業の動画撮影は、1つの固定カメラ(Webcam AW-V11)を用いて行う。

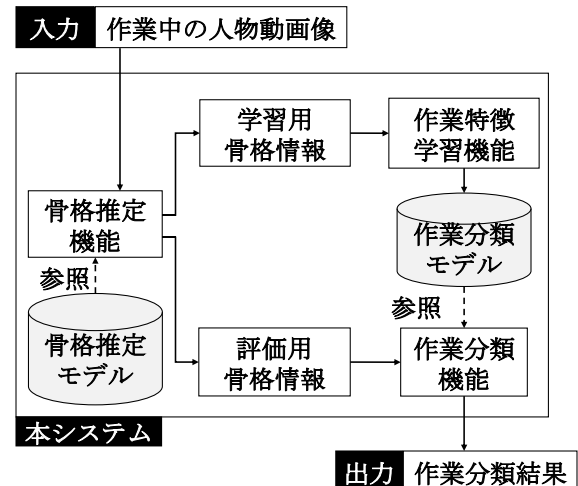


図1 本システムの概要

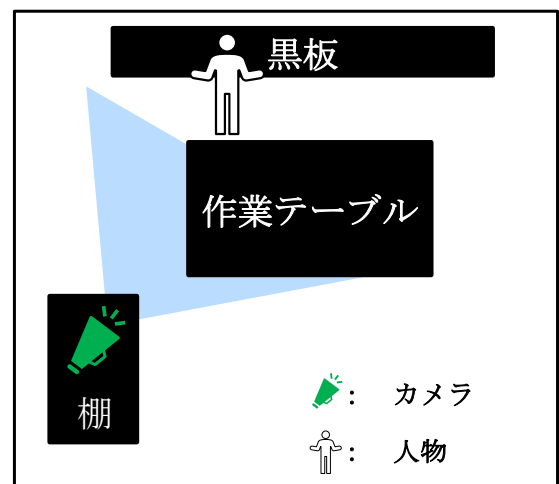


図2 撮影方法イメージ

画角や設置場所のイメージを図2に示す。カメラは事務所内の棚の上に設置し、黒板前の人物の行動を撮影する。

2.2 骨格推定機能

本機能では、MMposeを用いて、作業中の人物の動画像を入力として、動画像内の作業者の骨格推定を行い、骨格情報を出力する。なお、骨格情報は、13個のキーポイントがそれぞれ、画像の左上を原点とした2次元座標で表される。

2.3 作業特徴学習機能

本機能では、作業者の骨格情報を基に、対象者の「鼻」のキーポイントを原点とした12個のキーポイントの座標と、その座標の時間変化を特徴量としてLSTMを用いて学習を行い、作業分

Human Activity Recognition by Image Sequence for Dyeing Schedule Processes

[†]Aki Mimura, Riku Ikeda

Graduate School of Sustainable Systems Science, Komatsu University

[‡]Fuya Shibata, Yuga Tedoriya, Yoshihiro Ueda, Kazuma Sakamoto

Faculty of Production Systems Engineering and Sciences, Komatsu University

類モデルを構築する。

2. 4 作業分類機能

本機能では、構築した作業分類モデルを参照し、骨格情報を基に作業分類を行う。

3. 実験

本実験では、提案手法の有効性を評価するため、工場内で撮影したスケジューリング工程中の作業者の動画像を用いて、LSTMによる作業分類を行う。

3. 1 データセット

本実験で使用するデータセットは、繊維加工企業の工場内で撮影を行ったスケジューリング工程中の作業者の動画像である。また、スケジューリングに関係する動作が大きく分けて黒板を確認する動作、黒板に文字を書き込む動作、黒板の文字を消す動作、書類の整理の4種類である。しかし、本実験では、黒板の文字を書く動作と消す動作が酷似しているため、1つの行動として扱う。そのため、本実験の分類カテゴリは、「黒板確認」、「黒板編集」、「書類整理」、「その他」の4カテゴリとした。データセットの数は、カテゴリごとに学習用データが120件、評価用データが30件である。

3. 2 結果と考察

表1、表2に作業認識の結果を示す。表1の作業認識結果において、各行は正解とした実際の作業を示し、各列はLSTMによる予測した作業を示す。表2の評価結果は、それぞれの作業における適合率、再現率、F値を示す。

作業ごとのF値を見ると、「その他」の値が一番高く、0.889であった。これは、体の向きや、各キーポイント座標の変化がほかの作業と大きく異なるためだと考えられる。たとえば、「その他」では、黒板前を通過するだけの動作や、PCを操作している動作などであり、他の作業と異なった特徴量を持っているためだと考えられる。一方で、「黒板確認」のF値は0.774となった。これは、複数の行動を同時に行っていることが原因であると考えられる。「黒板確認」では、黒板を見ると同時に、手に所持している書類やテープなどの道具を動かす動作やヘルメットを脱ぐ動作などが確認でき、その手の動きなどが誤分類の要因であると考えられる。これは、複数の動作の認識を1つのラベルとして扱った結果であるため、同時に複数の動作を分けて認識できるシステムを構築することで改善できると考えられる。

4. おわりに

本稿では、骨格情報を用いたLSTMによる作業分類を行うシステムを提案し、実験により有効

表1 作業認識結果(混同行列)

予測 真	黒板 確認	黒板 編集	書類 整理	その他
黒板 確認	24	2	1	3
黒板 編集	4	24	1	1
書類 整理	3	1	25	1
その他	1	0	1	28

表2 評価結果

作業	適合率	再現率	F 値
黒板確認	0.750	0.800	0.774
黒板編集	0.889	0.800	0.842
書類整理	0.893	0.833	0.862
その他	0.848	0.933	0.889
平均	0.845	0.842	0.842

性を検証した。評価実験として、「黒板確認」、「黒板編集」、「書類整理」、「その他」の4カテゴリの作業分類を行ったところ、F値0.842の精度で作業を認識することができた。

今後の予定として、作業のカテゴリの増加や、分類精度の向上を目指す。また、今回の手法では、人物の判別は困難だったため、画像認識やセンサなどを本手法と併用することを検討する。加えて、今回、協力していただいた工場以外の場所でのデータ収集を行い、汎用性の検証を行うことを検討する。

参考文献

- [1] 服部, 豊嶋, 板倉, 吉田, 折原: 時系列の位置データを用いた人間行動分析, 第21回人工知能学会全国大会論文集, 人工知能学会, (2007)。
- [2] 川根, 伊集院, 杉, 中嶋, 山田: 機械学習による作業者のモーション分析の一考察, Webインテリジェンスとインタラクション研究会予稿集, Webインテリジェンスとインタラクション研究会, Vol. 17, No. 6, pp. 29-30, (2021)。
- [3] Y. He, S. Shirakabe, Y. Sand H. Kataoka.: Human Action Recognition without Human. *European Conference on Computer Vision (ECCV)*, (2016)。
- [4] MMpose Contributors.: OpenMMLab Pose Estimation Toolbox and Benchmark, 入手先<<https://github.com/open-mmlab/mmpose>> (参照 2024-1-12)。
- [5] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, Vol. 9, No. 8, pp. 1735-1780, (1997)。