

建設現場を想定したヘルメット模様による 人物の識別と追跡に関する研究

井上晴可[†] 梅原喜政[‡] 田中成典^{††} 今井龍一^{†‡} 神谷大介^{†‡} 中畑光貴^{††} 島野寛己^{††}

大阪経済大学情報社会学部[†] 関西大学先端科学技術推進機構[‡] 関西大学総合情報学部^{††}

法政大学デザイン工学部^{†‡} 琉球大学工学部^{†‡} 関西大学大学院総合情報学研究科^{††}

1. はじめに

近年、建設現場では、作業員と建設機械の接触事故や高所からの墜落事故が多発[1]しており、安全管理を支援する映像・画像処理のICT技術の導入が求められている。例えば、作業員に対して警告や注意喚起を行うことで安全管理に貢献すると考えられているが、その情報を作業員全員へ伝達するには人物を識別する必要がある。そのため、Society 5.0に向けた取り組み[2]が目ざされており、IoTやAIなどの先端技術が様々な現場において活用されようとしている。例えば、先行研究[3]では、単体カメラの映像から深層学習を用いて作業員のヘルメットに貼付した模様を識別することで人物を特定する手法が提案されている。しかし、単体カメラでは、オクルージョンが発生した場合や施工範囲が広い場合に人物を特定できない。そこで、本研究では、複数のカメラを用いた人物識別手法を提案する。

2. 研究の概要

本システム（図1）は、1) 模様識別機能、2) 撮影距離による補正機能、3) 時系列による修正機能により構成される。入力データは人物を撮影した動画、出力データは人物の識別結果とする。

2.1 模様識別機能

本機能では、物体認識手法であるYOLO[4]を

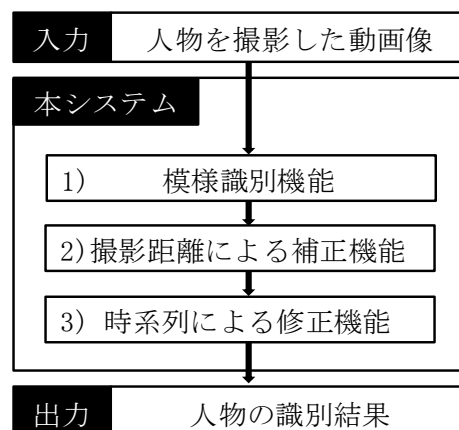


図1 処理フロー



図2 模様の例

用いて、人物を撮影した動画からヘルメット領域を取得する。そして、その領域に画像分類手法であるVGG16[5]を適用し、模様（図2）を識別する。

2.2 撮影距離による補正機能

本機能では、各カメラから人物までの距離を加味して模様の識別結果を補正する。まず、建設現場を俯瞰した現場の図面から各カメラの位置を指定し、その点を母点としたボロノイ分割によりボロノイ図を作成する。次に、動画内の人物位置を射影変換し、ボロノイ図に重ねる。そして、ボロノイ境界に基づいて人物がどの領域内にいるかを判定し、カメラからボロノイ領域までの距離に応じて模様の識別結果に重みづけを行う。最後に、その集計結果から各模様の合計値が最も高いものを採用することで識別結果を補正する。

Research for Identifying and Tracking Person Using Helmet Code Assuming Construction Sites

[†] Haruka Inoue

Faculty of Information Technology and Social Sciences, Osaka University of Economics

[‡] Yoshimasa Umehara

Organization for Research and Development of Innovative Science and Technology, Kansai University

^{††} Shigenori Tanaka and Hiroki Shimano

Faculty of Informatics, Kansai University

^{†‡} Ryuichi Imai

Faculty of Engineering and Design, Hosei University

^{†‡} Daisuke Kamiya

Faculty of Engineering, University of the Ryukyus

^{††} Koki Nakahata

Graduate School of Informatics, Kansai University

2. 3 時系列による修正機能

本機能では、前述の補正機能で複数の識別結果が推定された場合や誤識別された場合に対応するために、映像のフレーム間の時系列解析により2つの修正処理を適用することで識別結果を洗練する。

1つ目の修正処理では、識別結果を時系列順に並べ、識別結果が一意でないフレームAを検出する。そして、フレームAから過去に遡り、複数の識別結果の内いずれかと合致するフレームBの識別結果を採用する。そして、識別結果を一つに絞り込む。2つ目の修正処理では、全てのフレームを対象に、過去420フレーム分の識別結果を用いて多数決を行い、その結果を注目フレームの識別結果として採用することで、同一人物に対する誤識別を修正する。

3. 実証実験

本実験では、複数のカメラを用いた人物識別手法の有用性を確認する。

3. 1 実験内容

本実験(図3)では、2つの実験を行う。まず、実験1では、高さ4.5m、人物から5m離れた地点にカメラを設置し、17模様を対象に静止した人物を撮影する。そして、識別精度の高い模様を選定する。次に、実験2では、高さ4.1m、人物から8.5m~18.5m離れた2地点にカメラを設置し、実験1にてF値が高かった5模様を対象に4m×10mの長方形の外周を1分間歩行する人物を撮影する。撮影した動画像に対して提案手法の各機能を適用することで、模様ごとの識別精度を比較する。

3. 2 結果と考察

実験1の結果を表1に示す。表1に示すとおり、識別精度が高い上位5模様は222、???, △△△, UUU と NNN となった。そして、これらの模様を対象として実験2を評価した。

実験2の結果を表2に示す。表2の網掛け部に示すとおり、機能AのF値は0.36、機能Bは0.39、機能Cは0.71となり、提案手法の有用性を確認した。全体のF値を比較すると、機能Cは機能Aより0.35高く、撮影距離による補正機能と時系列による修正機能を適用することにより、識別精度を向上できることがわかった。撮影距離による補正機能では、複数カメラを使用してカメラから人物のいるボロノイ領域までの距離に応じて重みづけを行うため、距離の遠い人物への識別精度が向上したと考えられる。一方、時系列による修正機能では、複数フレームを対象として多数決を行うため、識別結果が一意でないフレームや誤識別となるフレームの識別精度が大きく向上したと考えられる。



図3 実験環境

表1 模様の識別精度 (一部抜粋)

順位	模様	適合率	再現率	F 値
1	2 2 2	0.70	0.71	0.71
2	???	0.73	0.39	0.51
3	△△△	0.45	0.58	0.51
4	UUU	0.46	0.50	0.48
5	NNN	0.61	0.38	0.46

表2 各手法適用後の F 値

提案手法	2 2 2	???	△△△	UUU	NNN	全体
機能 A	0.38	0.46	0.31	0.23	0.38	0.36
機能 B	0.42	0.43	0.37	0.33	0.41	0.39
機能 C	0.83	0.80	0.72	0.53	0.63	0.71

【凡例】機能A：模様識別機能，機能B：撮影距離による補正機能，機能C：撮影距離による補正機能+時系列による修正機能

4. おわりに

本研究では、複数カメラを用いた人物識別手法を提案し、実証実験を通じて提案手法の有用性を確認した。今後は、実現場においても提案手法を適用することで、提案手法の実用性を確認する。

参考文献

- [1] 厚生労働省：労働災害統計，<<https://anzeninfo.mhlw.go.jp/user/anzen/tok/anst00.htm>>，(入手 2021.1.8)。
- [2] 内閣府：Society 5.0，<https://www8.cao.go.jp/cstp/societ y5_0/>，(入手 2021.1.8)。
- [3] 今井龍一，神谷大介，井上晴可，田中成典，坂本一磨，藤井琢哉，菊池英一，伊藤誠：深層学習を用いた建設現場における人物識別に関する研究，土木学会論文集 F3 (土木情報学)，土木学会，Vol.75, No.2, pp. I_57-I_68, 2019。
- [4] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A.: You Only Look Once: Unified, Real-Time Object Detection, *Computer Vision and Pattern Recognition*, IEEE, Vol.29, No.2, pp.779-788, 2016.
- [5] Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, <<https://arxiv.org/pdf/1409.1556.pdf>>，(入手 2021.1.8)。