

ビデオカメラを用いた人物属性の推定に関する研究

井上晴可[†] 梅原喜政[‡] 今井龍一^{††} 神谷大介^{‡‡} 田中成典^{‡‡‡} 米田寿樹^{‡‡‡}

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[‡] 法政大学デザイン工学部^{††}

琉球大学工学部^{‡‡} 関西大学総合情報学部^{‡‡‡}

1. はじめに

近年、我が国では、Society 5.0[1]の提唱を契機に IoT や AI などの先端技術を活用する機会が増加している。画像処理分野では、動画画像から人物の顔領域を取得し、深層学習を用いて顔領域を解析することで人物の性別や年代といった属性情報を推定する研究[2]が注目されている。この属性情報を利用することで、効果的な広告掲示や商業施設での物品配置といった営業戦略に活用できる。しかし、カメラとの距離が遠く顔の解像度が低下する場合やカメラに正対せずに顔が写らない場合に属性情報を推定できない課題が残る。そこで、本研究では、人物の顔領域だけではなく、全身の特徴から属性を推定することで安定して性別と年代を推定する手法を提案する。

2. 研究の概要

本システム(図1)は、人物認識機能、性別推定機能と年代推定機能により構成される。入力データは人物の全身の動画画像、出力データは人物属性の推定結果とする。

2.1 人物認識機能

本機能では、物体検出手法である YOLO (You only Look Once) [3]を用いて、動画画像から人物の領域を取得し、その領域を人物画像として切り出す。本システムでは、この人物画像を画像分類手法である VGG19[4]に適用して人物の性別と年代を推定する。そのため、VGG19 の入力層に合わせて人物画像を 224×224 ピクセルに収まるように縮小し、余白部分は黒にする。

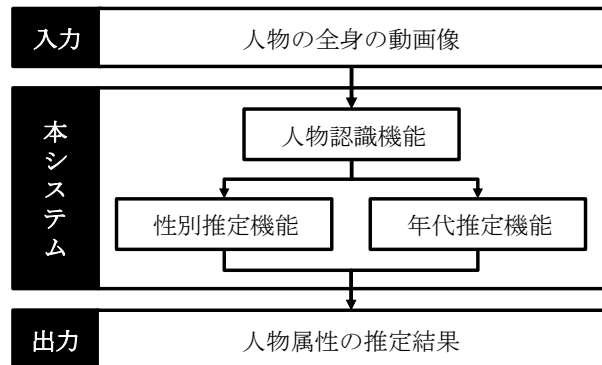


図1 処理フロー



図2 骨格情報を付与した人物画像

2.2 性別推定機能

本機能では、深層学習を用いて人物画像の性別を推定する。性別推定モデルは、男性と女性の人物画像を VGG19 により学習することで構築する。

2.3 年代推定機能

本機能では、深層学習を用いて人物画像の年代を推定する。推定対象の年代は、マーケティング分野において用いられる年齢区分[5]である 12 歳以下、13~19 歳、20~34 歳、35~49 歳と 50 歳以上の 5 つとする。年代推定モデルは、人物画像に骨格情報を重畳した画像を VGG19 で学習して構築する。本研究では、年代により歩行時の姿勢が異なる点に着目し、人物の骨格推定手法である OpenPose[6]を用いて推定した骨格情報を人物画像に重畳(図2)することで、年代の推定精度向上を図る。

Research for Extracting Person Posture Using Point Cloud Data

[†] 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

^{††} Ryuichi Imai

Faculty of Engineering and Design, Hosei University

^{‡‡} Daisuke Kamiya

Faculty of Engineering, University of the Ryukyus

^{‡‡‡} Shigenori Tanaka and Toshiki Yoneda

Faculty of Informatics, Kansai University

3. 実証実験

本実験では、人物の全身を撮影した動画像に提案手法を適用することで、提案手法の属性推定精度を確認する。

3.1 実験内容

学習データは、市街地にて撮影した歩行者の人物画像を各属性 300 枚、合計 2,100 枚とする。評価データは、同地点にて撮影した歩行者の画像を各属性 100 枚、合計 700 枚とする。正解データは、目視により画像に映った人物の属性を評価することで作成する。そして、提案手法の推定結果と正解データを比較し、適合率、再現率と F 値で評価する。なお、年代推定機能では骨格情報の重畳の有用性を評価するために、骨格情報を付与しない場合の年代推定モデルを生成し、骨格情報を付与した場合の年代推定モデルとの精度を比較する。

3.2 結果と考察

実験結果を表 1 と表 2 に示す。まず、性別の推定精度 (表 1) を確認すると、男性の F 値は 0.87、女性の F 値は 0.84 であることから、提案手法により性別を高精度に推定できることがわかった。しかし、女性が帽子を被っている場合や短髪で半袖のシャツとデニムパンツのような中性的な服装をしている場合に誤推定する傾向が見られた。

次に、年代の推定精度 (表 2) を確認すると、骨格情報無しの F 値が 0.92、骨格情報有りの F 値が 0.88 であることから、提案手法は骨格情報を含まない年代推定に比べると精度が低下したが、年代を高精度に推定できることがわかった。年代別に比較すると、12 歳以下の再現率が 1.00、0.99 と高いのは、幼稚園や小学校などの制服をこの年代の特徴として学習しているためだと考えられる。一方で、20~34 歳と 35~49 歳の再現率が低い傾向が見られた。これは、上記の年代の着用している服装が図 3 に示すように色や形の差が少ないことやスーツを着用している人物が多いことが要因となり、誤分類していると考えられる。そのため、上記のような服装の人物画像を学習データに追加することで、精度の向上が期待できる。

4. おわりに

本研究では、人物の全身の特徴に基づいた人物属性手法を提案し、実証実験により提案手法の有用性を確認した。今後は、提案手法を学習データとは別の時期および場所において評価することで、提案手法の汎用性を評価する。

参考文献

[1] 内閣府: Society 5.0, 内閣府 (オンライン), 入手先

表 1 実験結果 (性別)

性別	適合率	再現率	F 値
男	0.80	0.95	0.87
女	0.94	0.76	0.84

表 2 実験結果 (年代)

骨格情報	年代	適合率	再現率	F 値
無	12 歳以下	0.93	1.00	0.96
	13-19 歳	0.97	0.98	0.98
	20-34 歳	0.87	0.87	0.87
	35-49 歳	0.92	0.84	0.88
	50 歳以上	0.93	0.93	0.93
	全年代	0.92	0.92	0.92
有	12 歳以下	0.96	0.99	0.98
	13-19 歳	0.97	0.95	0.96
	20-34 歳	0.83	0.85	0.84
	35-49 歳	0.87	0.74	0.80
	50 歳以上	0.83	0.92	0.87
	全年代	0.89	0.89	0.88



図 3 参考画像

(https://www8.cao.go.jp/cstp/society5_0/) (参照 2022-1-7)

- [2] Levi, G. and Hassner, T.: Age and Gender Classification using Convolutional Neural Networks, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.34-42 (2015).
- [3] Bochkovskiy, A., Wang, Y. and Liao, M.: YOLOv4: Optimal Speed and Accuracy of Object Detection, arXiv(online), available from (<https://arxiv.org/pdf/2004.10934.pdf>) (accessed 2022-1-7).
- [4] Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv(online), available from (<https://arxiv.org/pdf/1409.1556.pdf>) (accessed 2022-1-7).
- [5] onca: 広告や販促のターゲティングにおける「C 層・T 層・F 層・M 層」とは, onca (オンライン), 入手先 (<https://onca.co.jp/column/column/p4398/>) (参照 2022-1-7).
- [6] Cao, Z., Simon, T., Wei, S. and Sheikh, Y.: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.7291-7299 (2017).