Vol.2018-HCI-177 No.31
2018/3/17

歩行動作に基づく人物識別に関する研究

櫛崎 翔太　　田内 康　　水上 嘉樹

山口大学理工学研究科 〒755-8611 山口県宇部市常盤台 2-16-1

E-mail: {w029vg, tauchi, mizu}@yamaguchi-u.ac.jp

あらまし 歩容識別は歩行動作に基づき人物識別を行う技術である。本研究では、推定姿勢に基づいたモデルベース歩容識別手法を提案する。提案手法は、姿勢推定のためのニューラルネットワーク、推定姿勢の整形処理部、歩容識別のためのニューラルネットワークの3つから構成されている。姿勢推定のためのニューラルネットワークは動画像から人物の姿勢を推定するためのものであり、Cao らによる OpenPose を採用している。姿勢整形処理部では、推定姿勢に含まれる脚関節の左右ラベリングの修正、推定できなかった関節座標の補間、身長に基づいた関節座標の正規化が行われる。歩容識別のためのニューラルネットワークでは、整形姿勢を入力として人物識別を行う。基礎的な実験を行い、姿勢整形処理の有効性、および、歩容識別ニューラルネットワークの構成方法について検討する。

キーワード 歩容識別，姿勢推定，ニューラルネットワーク

A Study on Person Identification Based on Walking Motion

Syota KUSHIZAKI　Yasushi TAUCHI　and　Yoshiki MIZUKAMI

Graduate School of Science and Engineering, Yamaguchi University

2-16-1 Tokiwadai, Ube, Yamaguchi, 755-8611 Japan

E-mail: {w029vg, tauchi, mizu}@yamaguchi-u.ac.jp

Abstract Gait identification is a technique for person identification based on walking motion. In this research, we propose a model based gait identification method based on estimated pose. The proposed method consists of three units, neural networks for pose estimation, an estimated pose refining unit, and a neural network for gait identification. The neural network for pose estimation (OpenPose by Cao et al., 2017) is for estimating the pose of a person from a set of sequential images. The estimated pose refining unit corrects the right and left labeling of the legs' joints included in the estimated pose, interpolates the joint coordinates which could not be estimated, and normalizes the joint coordinates based on the person's height. Finally, the neural network for gait identification classifies the person to be identified. We investigate the effectiveness of estimated pose refinement and the performance of the different neural network architectures for gait identification.

Keyword gait identification, pose estimation, neural network

1. はじめに

遠距離から撮影された動画像に基づいた生体認証として、歩容識別が注目されている。防犯カメラをはじめ、我々の身の回りにはカメラが多数設置されており、カメラの映像を元に人物の認証を行うケースがある。例として、防犯カメラの映像に写る人物を確認することで人物の過去の脚取りを辿り、事件捜査の手助けになることがあげられる。国内では、実際に事件捜査の証拠のひとつとして、歩容鑑定の結果が用いられたケース[1]もあり、新たな生体認証の選択肢として期待されている。

2. 従来研究

歩容識別はアピアランスベース歩容識別手法とモデルベース歩容識別手法に分けられる。杉浦ら[2]のアピアランスベース歩容識別手法は、撮影画像から抽出されたシルエットに基づいて歩容特徴抽出を行い歩容識別を行う。高田ら[3]のモデルベース歩容識別手法は、

ⓒ 2018 Information Processing Society of Japan

撮影画像から抽出された推定姿勢に基づいて歩容特徴抽出を行い歩容識別を行う。また、歩容識別にニューラルネットワークを用いた研究[4]も報告されている。

モデルベース歩容識別手法の前処理として、歩行者の姿勢を推定する必要がある。そのために３つの姿勢推定手法が提案されてきた。１つ目は人が画像１枚１枚に対して手動で姿勢検出を行う手法[5]、２つ目は機械学習を用いて画像内の関節位置の検出を行い、それらを接続して姿勢推定を行う手法[6]、３つ目は画像のもつカラー情報に加えて、深度情報を用いる手法[3]である。これらの姿勢推定手法は、それぞれに問題点がある。手動姿勢推定では、1 枚あたりの処理に時間と要すること、関節位置を検出する熟練した技術が必要であることがあげられる。機械学習による姿勢推定では、検出した関節を接続する際の接続候補の組み合わせが多く、処理時間が増大してしまうことがあげられる。深度情報を用いる手法は、姿勢推定のために赤外線を用いるため、太陽光の影響を受けやすい屋外での適用が困難である。一方で、近年、ニューラルネットワークを用いた指定推定手法が注目されている。

以上の問題点を踏まえて、本研究では Cao らによって提案された深層学習姿勢推定手法 OpenPose に着目する[7]。OpenPose では、関節検出時に関節位置および方向を表す Part Affinity Field(PAF)を出力する。PAF に基づいて関節の接続探索を行うことによって、探索速度が飛躍的に向上し、リアルタイム姿勢推定を実現している。

3. 提案手法

提案手法を４つのフェーズに分割して図 2 に示す。フェーズ１では、入力画像の撮影を行う。フェーズ２では､ニューラルネットワークによる姿勢推定を行う。以降、フェーズ２で使用するニューラルネットワークを姿勢推定 NN と呼ぶ｡フェーズ３では､姿勢推定 NN により出力された推定姿勢の整形処理を行う。フェーズ４では、ニューラルネットワークによる歩容識別を行う。以降、フェーズ４で使用するニューラルネットワークを歩容識別 NN と呼ぶ。

3.1 入力画像の撮影

入力画像は人が歩行する様子を側面から撮影したものを用いる｡カメラは 1 視点の RGB カメラのみを用いており､全方位カメラや赤外線センサは使用しない。撮影された入力画像例を図 3 に示す。図 3 は連続した 20 フレームを示しており､それぞれ左上が 1 フレーム目、右上が 10 フレーム目、左下が 11 フレーム目、右下が 20 フレーム目である。

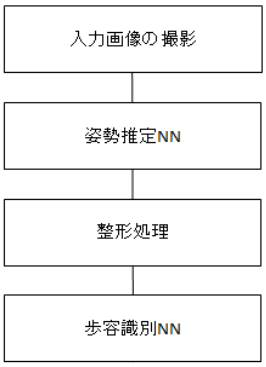


図 2 提案手法の流れ

図 3　撮影した入力画像例

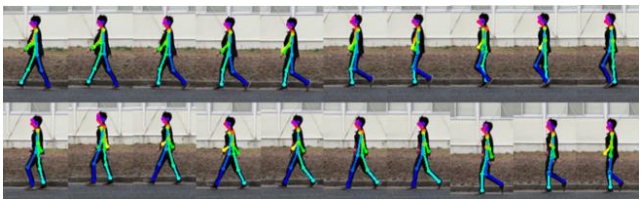


図 4　姿勢推定 NN による姿勢推定例

3.2 姿勢推定 NN

姿勢推定 NN により、撮影画像に対して姿勢推定を行う。本研究では Cao らによる OpenPose[7]を姿勢推定 NN として用いる。図 3 に対して姿勢推定 NN を適用し、姿勢推定を行った結果を図 4 に示す。各フレームの並びは図 3 と同様である。

3.3 推定姿勢の整形処理

姿勢推定 NN として OpenPose を用いることで､高精度かつ高速な姿勢推定が可能となっているが、得られた推定姿勢は 2 つの改善すべき点がある。一つ目は、脚関節ラベルの左右割当て時の不安定性である。これは遠方から脚部を撮影した際に、不明瞭な見え方に起因して本来右脚（左脚）に属する関節を左脚（右脚）に属したものとしてラベリングしてしまう問題である。図 5 に示した該当例は、図 4 中の 17 フレーム目から 19 フレーム目を切り出したものである｡17 フレームのラベリング結果では左脚が青色、右脚が水色として分類されているが､18 フレーム以降では左右のラベリングが反転している。そこで、体軸よりも前方向にある脚を前方脚、後方向にある脚を後方脚と見なし、再ラベリングする。二つ目は、関節の検出失敗である。図

ⓒ 2018 Information Processing Society of Japan

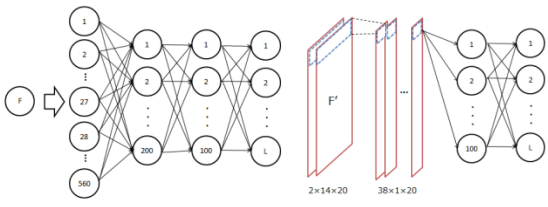
4 の 11 フレーム目から、胴体に隠れた右手（黄色）が検出できていないことがわかる。検出に失敗した関節の座標値は(0,0)となる。このとき、基準として検出に失敗した関節座標値を J_N とおき、時系列的に過去にある検出成功関節座標値を J_1、未来にある検出成功関節座標値を J_2 とし、J_1 および J_2 の中間値を J_N の座標とする。中間値の計算ができない場合、同ラベル内で時系列的に最も近い関節座標値をそのまま関節座標値として使用する。

歩容識別問題では、関節の周期的な変化が重要な歩容特徴となる。関節検出およびラベリングに失敗した推定結果が含まれる入力データを歩容識別 NN の学習に用いると、正しく歩容特徴を学習することが困難となる。推定結果の修正後、遠近へのスケール対応のため、推定結果スケルトンの高さ方向のスケールを一律とする。

図 5 姿勢推定時の両脚ラベリングの不安定性

3.4 歩容識別 NN

歩容識別問題のために適切な歩容識別 NN を検討するため、ネットワークモデルを 2 種類用意する。1 つ目の歩容識別 NN として、全結合層を用いたニューラルネットワークを使用し、これを全結合型 NN とする（図 6(a)）。2 つ目の歩容識別 NN として、畳み込み層および全結合層を用いたニューラルネットワークを使用し、これを畳み込み型 NN とする（図 6(b)）。結合重みの初期値として、標準偏差が 0.01 のガウス分布を用いる。学習率の調整には AdaGrad[8]を用いる。また、各層の終わりには Batch Normalization を用いる。



(a) 全結合型 NN　　(b) 畳み込み型 NN

図 6 提案する歩容識別 NN の構成

4. 実験と結果

検証に参加した被撮影者は 8 人である。連続する 20 フレーム分の入力画像に対して 3.2 節で示した姿勢推定 NN の処理を適用し推定結果を得る。次に、推定姿勢に対して 3.3 節で示した整形処理を行い、訓練データおよびテストデータを生成する。最終的に得られた訓練データ数、テストデータ数を表 1 に示す。

表 1 訓練データ数およびテストデータ数

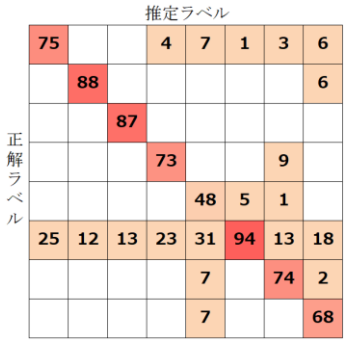
	訓練データ数	テストデータ数
被撮影者1	220	79
被撮影者2	253	172
被撮影者3	182	156
被撮影者4	196	132
被撮影者5	256	147
被撮影者6	271	203
被撮影者7	206	153
被撮影者8	142	109
合計	1726	1151

本実験では、整形処理の有効性を確認するために、整形処理を用いた場合と用いない場合の識別率を示す。さらに、全結合型 NN および畳み込み型 NN のそれぞれを用いた場合についても検証する。よって、整形処理の有無、および、ネットワークの構成から 4 通りの識別率を比較検討することとする。

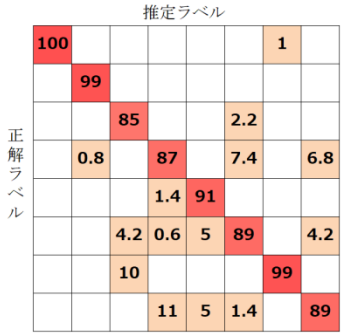
全結合型 NN に対して、整形処理を行わずに作成した入力データにより学習、および、テストを行った識別率を図 7(a)に示す。また、整形処理を行い作成した入力データにより 5 回の学習により作成したそれぞれの学習モデルのテストデータに対する識別率を図 7(b)に示す。同様の条件で畳み込み型 NN に対して、学習、およびテストを行った結果を図 7(c)、図 7(d)に示す。また、それぞれの結果の平均識別率を表 2 に示す。

実験では、2 つの結果が示された。1 つ目の結果は、全結合型 NN および畳み込み型 NN のどちらにおいても、整形処理を行った入力データを用いた歩容識別 NN の識別結果が 10%以上の識別率向上を示した。このことから、姿勢推定 NN の推定結果を歩容識別 NN の入力データとするための整形処理の有用性が示された。一方、2 つ目の結果では、全結合型 NN、および畳み込み型 NN の識別率はどちらも約 92%であり、ネットワーク構成の変化による優位性は示されなかった。これについては、入力画像の撮影角度追加や被撮影者の人数増加に関する検討が必要であると考えられる。

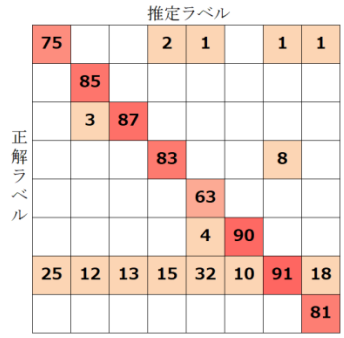
ⓒ 2018 Information Processing Society of Japan



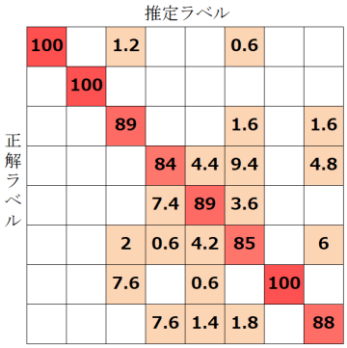
(a) 整形処理を行わない場合の全結合型 NN



(b) 整形処理を行った場合の全結合型 NN



(c) 整形処理を行わない場合の畳み込み型 NN



(d) 整形処理を行った場合の畳み込み型 NN

図 7 歩容識別時に得られたラベル毎の識別率[%]

表 2 平均識別率[%]

	整形処理を行った入力データ	整形処理を行った入力データ
全結合型NN	約76	約92
畳み込み型NN	約82	約92

5. おわりに

本研究では、モデルベース歩容識別問題に対して、２つのニューラルネットワークを用いた基礎的な実験を行った。提案手法は全４フェーズで構成した。フェーズ１にて、入力画像の撮影を行った。フェーズ２にて、姿勢推定 NN による姿勢推定を行った。フェーズ３にて、推定結果の整形処理を行った。フェーズ４にて、歩容識別 NN による歩容識別を行った。フェーズ４で用いた歩容識別 NN では２種類のニューラルネットワークである、全結合層のみを用いた全結合型 NN、および畳み込み層と全結合層を組み合わせた畳み込み型 NN を提案した。

参考文献

[1] Japan Business Press
http://jbpress.ismedia.jp/articles/-/47148

[2] 杉浦 一成，槇原 靖．八木康史 全方位カメラを用いた複数方向の観測による歩容認証 情報処理学会論文誌 In コンピュータビジョンとイメージメディア Vol. 1 No. 2 76-85,2008

[3] 高田 憲一，北須賀 輝明，有次 正義 マーカレスモーションキャプチャ装置を用いた歩容による個人識別法の提案 情報処理学会研究報告
In Vol.2012-UBI-35 No.9 ,2012

[4] Eman Fares Al Mashagba Human Identification by Gait Using Time Delay Neural Networks ComputerandInformation Science; Vol. 8, No.4; 2015

[5] Peter K. Larsen Erik B. Simonsen and Niels Lynnerup Gait Analysis in Forensic Medicine In TECHNICAL NOTE ,2008

[6] Leonid Pishchulin，Mykhaylo Andriluka，Peter Gehler ，Bernt Schiele PoseletConditionedPictorialStructures In CVPR,2013

[7] Zhe Cao Tomas Simon Shih-En Wei Yaser Sheikh RealtimeMulti-Person2DPoseEstimationusingPartAffinity Fields In CVPR ,2017

[8]AdaGrad - Adaptive Subgradient Methods
https://cs.stanford.edu/~ppasupat/a9online/1107.html

ⓒ 2018 Information Processing Society of Japan