

人物属性を考慮した歩容認証

守脇 幸佑^{1,a)} 村松 大吾^{1,b)} 八木 康史^{1,c)}

概要: 生体特徴を用いた個人認証手法である生体認証の中でも、映像解析に基づいた歩容認証は、他の手法と異なる特性を有しているため、犯罪捜査目的等を中心に注目が集まっている。本論文では、畳み込みニューラルネットワーク (CNN) を用いて、性別や年齢などの属性情報を考慮した歩容認証の手法を提案する。1枚の歩容エネルギー画像 (GEI) を入力とし、その入力データに関する複数の情報をラベルとして学習させることで、人物属性などの情報を考慮した歩容認証手法を提案する。大規模歩容データベースを用いた、評価実験により、提案手法が従来手法より高い認証性能を実現できることを確認した。

キーワード: 歩容認証, 深層学習, 人物属性, 年齢, 性別

1. はじめに

生体認証に利用される生体特徴は、DNA や指紋、虹彩、歩容といった様々な種類がある。その中でも歩容特徴は他の多くの生体特徴と異なり、対象人物の協力を必要とせず、また離れた位置からでも取得可能であるため防犯カメラを用いた犯罪捜査への利用が期待できる。実際に海外で犯罪捜査に歩容認証が利用されており [1]、日本においても犯罪捜査のための有力な情報として活用され始めている。平成 26 年には警察白書に新しい捜査手法として個人の歩容特徴から個人識別を可能にする技術が紹介されており [2]、これまで盛んに研究がされている。歩容認証の初期研究から生体運動学に基づく手法 [3] も提案されているが、近年では歩容を画像ではなく映像として認識する、映像解析に基づいた歩容認証に注目が集まっている。

歩容認証とは歩容 (人間の歩いている姿) を用いて人物認証を行う技術であるが、人物の行動に起因する特徴を利用した手法であるため、認証に利用する特徴がほかの生体特徴よりもばらつきやすい、という問題がある。また、認証の際に比較する歩容間で撮影角度や路面状況、対象人物の歩く速度や外見 (服装、靴、持ち物の有無等) といった条件の違いは認証精度の低下を引き起こす要因になる。これは条件の違いにより実際の歩行が変化する場合や、また歩行が変化しなくとも、服装や撮影角度の関係で対象者の歩きの見えが変化し、映像から抽出しうる特徴が変わるため

である。滑りやすい路面状況では、人は滑らないように注意をして歩くため、乾燥した道路を歩く場合と比較すると歩き方そのものが変わってしまう。路面状況が同じで、同じように歩いている場合でもスーツを着ている場合とダウンジャケットを着ている場合では、見えが異なる。また、全く同じように歩いている人物であっても、観測角度が変わると見えが異なり、観測できる情報が変わる。例えば正面から撮影した歩容からは人物の体の幅や足の左右方向の開き具合といった特徴が分かるが、側面から撮影したものではそれらの特徴は分からず、代わりに歩幅や腕の振り幅といった特徴を得ることができる。実環境で歩容認証を行う際には、自由に歩いている人物を対象とすることが多いため、撮影角度が異なる、という場合が発生しやすい。したがって、撮影角度の違いに対する頑健性を実現することは、歩容認証において重要な課題である。

本論文では観測視点の違いに頑健な特徴を抽出可能な手法に人物属性を考慮する仕組みを導入した歩容認証の手法を提案し、異なる観測視点間でも高い認証精度を可能とする手法を構築した。人物属性として、本論文では性別と年齢に注目する。性別や年齢はソフトバイオメトリクスと呼ばれ、それら単体では高い認証精度を有しないが、ある程度は個人を絞り込むことが可能な情報で、認証精度の向上が期待できると考えた。また、観測角度の違いに頑健な認証手法には、方向変換モデル [4] や MvDA [5] など、多くの手法が候補として考えられる。しかし、多くの手法は対象となる歩行映像が照合の際にどの観測視点から撮影されたかという、観測視点に関する情報を必要とするため、これらの方法は、防犯カメラ映像を対象とした認証問題のよ

¹ 大阪大学
Osaka University

a) moriwaki@am.sanken.osaka-u.ac.jp

b) muramatsu@@am.sanken.osaka-u.ac.jp

c) yagi@am.sanken.osaka-u.ac.jp

うな、認証対象人物が自由に歩いている場面での利用を考えた場合には制約が大きい。そのため、観測視点情報がなくとも認証可能な手法が好ましい。そこで、我々は認証の際に観測視点情報が不要な手法に着目する。具体的な歩容認証の手法として、近年画像認識の分野で大きな成功を取っている深層学習 [6] に基づく歩容認証手法 [7] に注目した。深層学習は十分な数の学習データがあれば 1 つの枠組みで特徴抽出と認識が同時に実現可能である。更に、画像データの学習に適している畳み込みニューラルネットワーク (CNN : Convolutional Neural Network) を用いることで画像特有の空間的近接性が考慮され、高精度の認識が可能となる。CNN は画像分類 [8]、動画分類 [9]、行動認識 [10] など多くのタスクにおいてその有用性が示されており、生体認証に関しても静脈認証 [11] や顔認証 [12] でも利用されている。

提案手法の有効性を確認するために複数の観測視点からの歩容データを含む OU-ISIR large population dataset (OULP) [13] のサブセットを用いて検証を行った。このサブセットには、1912 人の被験者データが含まれており、被験者に対して 4 つの異なる観測視点の歩行データが利用できる。OULP を用いた評価実験では、属性情報を考慮する提案手法が、従来の手法をよりも高い精度を実現できた。

2. 関連研究

2.1 複数の情報を用いた認証

複数の情報を用いて精度の改善を試みる手法がある。村松らの研究 [14] では 1 つの歩行映像から顔や歩容などの異なるモダリティ特徴を抽出し、更に多視点特徴も考慮することで認証精度の向上を図っている。また、Kale ら [15] は観測視点に頑健な歩容認証と顔認証手法を階層的、もしくはスコアレベルで統合する手法を提案している。Hofmann ら [16] は、単純な背景差分の代わりに重みをもたせた背景レイヤーを取得することで、背景の影響を除去した顔画像に加えて従来の GEI よりも高精度なアルファ GEI を抽出して統合する手法を提案している。

2.2 歩容の表現手法

歩容認証には、その用途から防犯カメラや定点カメラ等から取得された映像を用いることが一般的である。それらから撮られた映像をどのように処理して歩容特徴を抽出するかは、モデルに基づく表現と見えに基づく表現に大別される [17] が、モデルに基づく表現は低解像度画像に対しては精度よく当てはめることが難しいことから、現在は見えに基づく表現を用いたものが主流である。特に、近年の歩容認証では服の色やテクスチャの影響を受けない、シルエットに基づく表現がよく用いられている。その中でも最も頻繁に用いられているのが平均シルエット [18] としても知られている、歩容エネルギー画像 (GEI : Gait Energy

Image) [19] であり、撮影された映像から得られる歩容のシルエット系列を一歩行周期分で平均することで得られる。GEI は体型などの静的な成分と腕や脚の動きといった動的な成分が混合された特徴である。一方、静的な成分に影響を及ぼす服装や荷物による影響を軽減するために、動的な成分により重点を置いた歩容の特徴表現も提案されている。その典型的な例が、Wang らの 1/4 歩行周期中の歩行姿勢に対して色付け (例えば、両脚支持相を赤、片足支持相を青など) したシルエット輪郭を足し合わせた crono-gait image [20] や、Lam らのシルエット輪郭に対するフローの有無を含めたバイナリ画像列を 1 歩行周期で平均化した画像を用いる gait flow image [21] などである。

図 1 は異なる二人の人物の GEI を示したものである。この図より、人物が異なると、その GEI が大きく異なることが確認できる。また、図 1 の左側の (A) はある 17 歳男性の GEI であり、(B) は 6 歳の女性の GEI であるが、これらの性別や年齢の情報も GEI には含まれていると考えられる。事実、GEI から年齢や性別を推定する手法 [22] が提案されているため、本研究課題では、GEI を対象とした手法を考えることとする。

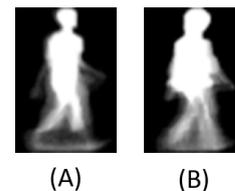


図 1 GEI の例 (A) が 17 歳の男性、(B) が 6 歳の女性

2.3 深層学習を用いた歩容認証

深層学習を歩容認証に適用した事例として、武村らの研究 [7] では同一被験者の複数方向から撮影した GEI を学習データとして学習させることで、観測視点の情報は不要でかつ、視点変化に頑健な歩容認証を可能とする GEINet が提案されている。また、Hossain らの研究 [23] では防犯カメラからの映像を制約付きボルツマンマシンに与えて抽出された特徴量を利用する手法を提案している。そのどちらも学習の際に与えるラベルは単一の種類のもので、今回のように複数のラベルを与えた場合の学習については行われていない。Wu らの研究 [24] では、複数の CNN 構造を提案・比較するとともに、それらを組み合わせた手法を提案している。一方で、年齢や性別を人物認証に用いる研究 [25] も報告されている。この手法では本提案手法とは異なり、人物認証と年齢、性別のマルチタスクを考えている。

3. Parallel-GEINet

本論文では武村らの論文 [7] により提案されている視点変化に頑健な歩容認証の CNN である GEINet をベースと

し、人物属性を考慮する手法として、1枚の入力画像から複数の人物属性ラベルを別々に学習させ、入力データに対する特徴量を出力するfc3層の出力を統合し、類似度を出力する歩容認証を提案する。この手法は、GEINetをパラレルに組み合わせた構造となることより、本手法をParallel-GEINetと呼ぶ。

3.1 GEINet : 視点変化に頑健な歩容認証

GEINetは同一被験者の複数方向から撮影したGEIを学習させることで、視点変化に頑健な歩容認証を可能とするものである。異なる観測視点の歩容特徴について、人物ラベルのみを正解データとして与え学習させることで、異なる観測視点の歩容特徴間でも、共通な特徴を見つけることができ、観測視点の違いに頑健な歩容認証が実現できる。本研究ではこのGEINetを基にして新たなCNNを構築するため、まずはこのGEINetについて詳しく説明する。

3.1.1 GEINetの構造

GEINetの構造を図2に示す。GEINetはGEIを入力として受け取り、その人物IDに対応するラベルを出力する構造となっている。GEINetは全部で8層のCNNで、はじめの6層は畳み込み層-プーリング層-正規化層の3つ組が2つ連なって構成されている。

畳み込み層は、畳み込み演算により入力データの形状を維持したまま次の層にデータを出力する層であり、画像の空間的情報を保持することが可能となる。プーリング層は畳み込み層から受け取ったデータに対してその空間サイズを小さくする役割を持つ。正規化層では受け取ったデータの値が決められた範囲（今回は0.0~1.0）に収まるように変換する層である。

続くfc3層はユニット数1024の全結合層で、活性化関数にはReLUを用いる。全結合層fc4層は学習データに含まれる歩容特徴の人物数と同じN個のユニットを持つ。学習時にはfc4層にソフトマックス関数を適用し、入力画像に対する被験者ごとの類似度を出力する。具体的に言うと、最終層のi番目のユニットはi番目の被験者との類似度を表すことになる。GEINetで設定されている各層の値を表1に示す。

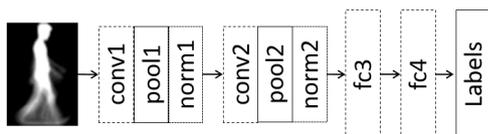


図2 GEINetの構造

3.1.2 学習

GEINetの学習の際にはその重みパラメータ w を M 個の訓練用データを用いて最適化する。訓練用データには、

表1 GEINetの各層の設定

Layer	#Kernel	Size/Stride	Act	Pooling
conv1	18	$7 \times 7 \times 1/1$	ReLU	
pool1		$2 \times 2/2$		Max pooling
conv2	45	$5 \times 5 \times 18/1$	ReLU	
pool2		$3 \times 3/2$		Max pooling

N人分の歩容特徴が含まれているとする。m番目の入力画像 I_m を与えた時のfc4層の出力を $V_{fc4}(I_m; w)$ 、fc4層の出力ベクトルのn番目の要素のソフトマックスを $V'_n(I_m; w)$ と定義する。学習誤差は以下に示す交差エントロピーとし、これを最小化することによって学習を行う。

$$L_1(w) = - \sum_{m=1}^M \sum_{n=1}^N \delta_{y_m n} \log V'_n(I_m; w) \quad (1)$$

ただし $\delta_{y_m n}$ はクロネッカーのデルタ、 y_m は入力画像 I_m の人物IDを表す。式1に示す関数を最適化するパラメータを学習することで、ネットワークは、学習人物のIDに対応するラベルのみに高い値を出力するようになる。

3.1.3 認証

認証時には、学習したGEINetを特徴抽出器として利用する。一般的な物体認識問題等では、学習時と認識時には正解ラベルが共通であるが、人物認識の問題においてそのような設定は一般的ではない、したがって、学習したネットワークの出力ラベルにより人物を認識するのではなく、ネットワークを特徴抽出器として用い、抽出した特徴を比較することで、認識を行う。認識対象として入力される2つのデータをプローブとギャラリーとし、それぞれのGEI、 I_p および I_g をGEINetに入力した時のfc4層の出力ベクトルを $V_{fc4}(I_p; w)$ 、 $V_{fc4}(I_g; w)$ を抽出した特徴量とする。そして、抽出された特徴量の差のL2ノルムを計算することで、特徴間の相違度を定める。相違度の計算式は以下のように表される。

$$dist(I_p, I_g; w) = \|V_{fc4}(I_p; w) - V_{fc4}(I_g; w)\|_2 \quad (2)$$

他人同士である2枚のGEIを比較する際には、GEINetによって抽出されるそれらの特徴の差のL2ノルムは大きく、逆に本人同士では小さくなるのが期待される。そこで、これを利用し、プローブとギャラリーのGEIの組み合わせにおいて最も特徴の差が小さい、つまりL2ノルムが小さい組み合わせを本人同士と判断する。

3.2 Parallel-GEINet

3.2.1 入力データ

本手法ではGEINetと同様にGEIを入力とし、その画像サイズも 88×128 ピクセルとした。GEIは、個人認証に有効な特徴であるのみならず、年齢や性別推定にも利用されている特徴であるため、GEIを入力とすることで、一

つの入力データから人物 ID や年齢、性別といった異なるラベルを学習することが可能となるためである。

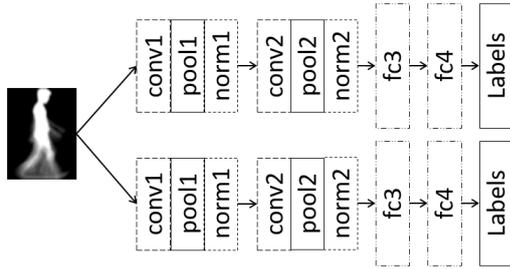


図 3 Parallel-GEINet の 2 ラベルの場合の構成

3.2.2 正解ラベル

今回 Parallel-GEINet に 1 つの入力に対して正解ラベルとして与えるラベルは、2 ラベルの場合は人物 ID とその年齢、人物 ID とその性別、の二種類を用いた。また、3 ラベルの場合には人物 ID と年齢及び性別の 3 種類を正解ラベルとして用いた。各ラベルの持つクラス数は、人物 ID では認証の際に用いる被験者の人数分（後述の評価実験の際には 956 人の被験者に対して認証を行うので 956 クラス）、年齢ラベルはその認証を行う集団に含まれる年齢の数だけ存在し、性別ラベルは男性か女性かの 2 クラスである。

3.2.3 ネットワーク構造

Parallel-GEINet の構造は前述の GEINet を複数個並列に組み合わせた CNN で、例として 2 つの情報を与えて学習させる 2 ラベルの構造を図 3 に示す。1 枚の GEI の入力に対し、学習時に与える情報の種類数だけ GEINet を並べ、その各々で 1 種類のラベルに関しての学習を行う。それゆえ Parallel-GEINet を構成する各 GEINet の構造や各層の設定は前述のものと同様だが、最終層である fc4 の持つユニット数のみ、そのラベルに対応するクラス数となるよう変更する。

3.2.4 学習

Parallel-GEINet を構成している各 GEINet の学習はそれぞれ独立に行う。入力される ID が y_m である被験者の GEI を I_m 、年齢ラベルを a_m 、性別ラベルを g_m とする。fc4 層の出力ベクトルの n 番目の要素のソフトマックスを $V'_n(I_m; w)$ で定義すると、ID に関する損失関数は

$$L_{ID}(\omega) = - \sum_{m=1}^M \sum_{n=1}^{N_{ID}} \delta_{y_m n} \log V'_n(I_m; w) \quad (3)$$

と表され、この交差エントロピーを最小化することによって ID に関する学習を行う。ここで N_{ID} は学習被験者に含まれる人物数である。同様に、年齢に関する損失関数は、年齢のクラス数を N_a とすると、

$$L_a(\omega) = - \sum_{m=1}^M \sum_{n=1}^{N_a} \delta_{a_m n} \log V'_n(I_m; w) \quad (4)$$

性別に関する損失関数は、性別のクラス数を N_g とすると

$$L_g(\omega) = - \sum_{m=1}^M \sum_{n=1}^{N_g} \delta_{g_m n} \log V'_n(I_m; w) \quad (5)$$

で表され、最小化することによって学習を行う。

また、学習により求まる w は各ラベルによって異なるのでそれらを

$$\hat{\omega}_{ID} = \underset{\omega}{\operatorname{argmin}} L_{ID}(\omega) \quad (6)$$

$$\hat{\omega}_a = \underset{\omega}{\operatorname{argmin}} L_a(\omega) \quad (7)$$

$$\hat{\omega}_g = \underset{\omega}{\operatorname{argmin}} L_g(\omega) \quad (8)$$

とする。

3.2.5 認識

あるプローブ I_p とギャラリー I_g のペアが与えられたとき、これらの特徴はそれぞれを Parallel-GEINet に入力として与えて順伝搬したときの各 GEINet における fc3 層の出力ベクトルの和を結合したものとす。プローブ I_p に対して、人物 ID のネットワークの fc3 層の出力を $\nu_{fc3}(I_p; \hat{\omega}_{ID})$ 、年齢用ネットワークの fc3 層の出力を $\zeta_{fc3}(I_p; \hat{\omega}_a)$ 、性別用ネットワークの fc3 層の出力を $\eta_{fc3}(I_p; \hat{\omega}_g)$ としたとき、抽出される特徴は $V_{fusion}(I_p)$ は、例えば ID と年齢の組み合わせは

$$V_{fusion}(I_p; \hat{\omega}_{ID}, \hat{\omega}_a) = [\nu_{fc3}^T(I_p; \hat{\omega}_{ID}), \zeta_{fc3}^T(I_p; \hat{\omega}_a)]^T \quad (9)$$

とする。ここで、各特徴は列ベクトルで表現されているとし、T は転置を意味する。

同様にギャラリー I_g についても $V_{fusion}(I_g)$ を求める。これによりギャラリーとプローブの相違度は二つの特徴ベクトルの差の L2 ノルムとして求められ、式 10 計算できる。

$$\operatorname{dist}(I_p, I_g; \hat{\omega}_{ID}, \hat{\omega}_a) = \|V_{fusion}(I_p) - V_{fusion}(I_g)\|_2 \quad (10)$$

4. 評価実験

4.1 実験設定

4.1.1 データ分割

本実験では OULP 中の 1912 人分の歩容画像を実験に使用した。各被験者に対してそれぞれ異なる角度 (55, 65, 75, 85 度) から撮影したデータが 2 つずつ、合計 8 データ存在する。図 4 に各撮影角度での GEI の例を示す。

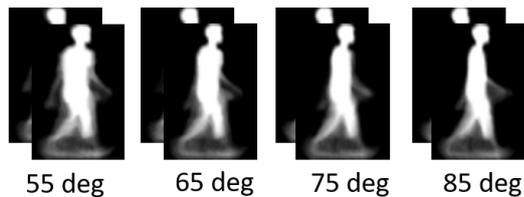


図 4 OULP に含まれる 4 つの撮影角度からの歩容シルエット画像例

先行研究との比較のため、データベースに含まれる 1912 人の被験者を互いに重複しないような 956 人ずつの 2 つのグループに分け、一方の被験者グループを学習データ、もう片方のグループをテストデータとし、2-分割交差検証を行う。ただし、分割方法によるデータの偏りを考慮するためにグループ分割を 5 回行い、それぞれに対して交差検証を行った。1 人の被験者に関して GEI が 8 データ存在するので、一回のグループ分割では学習とテスト共に 956×8 の 7,648 データを用いることになる。

この実験設定は、VTM with transformation consistency measure (TCM+) [26], quality-dependent TVM (wQVTM) [27], GEINet [7] などの既存手法と同一のプロトコルを用いている。ただし、既存手法では、年齢や性別に関する情報を利用しないのに対し、提案手法では、年齢及び性別ラベルを学習時のみ利用している。

4.1.2 学習

学習時におけるミニバッチサイズは先行研究 [7] と合わせるため 239 とし、学習率の初期値は 0.02, 総イテレーション数は 50,000 とした。また、学習率は 10,000 イテレーションごとに $\gamma = \sqrt[3]{0.01}$ 倍にして小さくなるようにした。GEINet 及び Parallel-GEINet の学習とテストに関しては深層学習のフレームワークである Caffe を用い、GPU として NVIDIA Quadro K600 を使用した。

4.1.3 1 対 N 認証評価

テスト用グループに属する被験者について、各角度の GEI がそれぞれ 2 枚ずつ存在するため、その片方をプローブ、もう一方をギャラリーとする。各角度のプローブ合計 956 枚の中の 1 枚 1 枚に対して、55 度、65 度、75 度、85 度のそれぞれのギャラリーに対し、最も相違度の小さいギャラリー画像を選ぶ。そして選ばれた最も相違度の小さい特徴が本人同士の特徴である割合、Rank-1 認証率を計算する。

4.1.4 1 対 1 認証評価

Parallel-GEINet について、他人受入率 (FAR: False accept rate) と本人受入率 (FRR: False rejection rate) を計算し、等価エラー率 (EER: Equal error rate) を求めた。FAR は相違度の閾値を変化させていった際に、その時の閾値に対して、他人であるにも関わらず本人であると Parallel-GEINet が判断した割合を表しており、FRR はそ

表 2 人物 ID と年齢について学習した Parallel-GEINet の Rank-1 認証率 [%] の比較

ギャラリーの角度	手法	プローブ画像の角度			
		55	65	75	85
55 度	Parallel-GEINet	<u>96.9</u>	<u>96.9</u>	<u>93.5</u>	<u>85.5</u>
	GEINet	94.4	93.4	89.8	81.9
	TCM+		79.9	70.8	54.5
	wQVTM		78.3	64.0	48.6
65 度	Parallel-GEINet	<u>95.6</u>	<u>96.7</u>	<u>95.9</u>	<u>94.4</u>
	GEINet	93.5	94.6	93.8	92.4
	TCM+	81.7		79.2	70.2
	wQVTM	81.5		79.2	67.5
75 度	Parallel-GEINet	<u>92.1</u>	<u>95.7</u>	<u>96.8</u>	<u>96.2</u>
	GEINet	89.5	93.9	95.3	94.5
	TCM+	71.9	80.0		79.0
	wQVTM	70.2	80.0		78.2
85 度	Parallel-GEINet	<u>86.2</u>	<u>93.9</u>	<u>96.0</u>	<u>96.5</u>
	GEINet	81.7	93.3	95.3	95.5
	TCM+	53.7	73.0	79.4	
	wQVTM	51.1	68.5	79.0	

の閾値に対して入力 GEI が本人である場合でも他人と判断される割合を表す。EER は FAR と FRR が同じ値となる閾値におけるエラー率である。

4.2 実験結果

4.2.1 Rank-1 認証率の比較

Parallel-GEINet に ID と年齢情報を与えて学習させた場合の Rank-1 認証精度と他の手法の認証精度を比較した結果を表 2 に示す。表において、最も良い値を下線付きで示した。

結果から、年齢と人物 ID を与えて学習させた Parallel-GEINet の認証率がどの観測視点の組み合わせにおいても既存の手法 [26] [27] [7] よりも高い精度を実現できていることがわかる。

次に、Parallel-GEINet に人物 ID とその被験者の性別情報の 2 つを与えて認証に用いた場合と、人物 ID, 性別, 年齢の 3 つを与えて認証を行った場合の Rank-1 認証率を表 3 に示す。比較対象として人物 ID と年齢について学習した Parallel-GEINet と GEINet の Rank-1 認証率もまとめる。プローブとギャラリーの各観測視点の組み合わせにおいて最も良い値を下線付きで示している。

いずれの場合も GEINet と比較して認証精度は向上している。人物 ID に加えて年齢と性別情報を与えた場合の Parallel-GEINet の Rank-1 認証率が、16 組の角度の組み合わせのうち 10 組で最も高いものとなっているが、先述した人物 ID と年齢についてのみ学習した Parallel-GEINet からの大幅な精度向上は見られなかった。これらの組み合わせであり精度が向上しなかった原因としては性別と年齢を識別する特徴に相関があり、認証の際にそれらを組み合わせても人物推定に有効な特徴がそれほど増加しなかつ

表 3 人物 ID, 年齢, 性別情報について学習した Parallel-GEINet の Rank-1 認証率 [%] の比較

ギャラリーの角度	手法	プローブ画像の角度			
		55	65	75	85
55 度	Parallel-GEINet				
	ID+年齢+性別	96.7	<u>96.9</u>	<u>94.0</u>	85.8
	ID+性別	96.2	95.1	90.6	<u>85.9</u>
	ID+年齢	<u>96.9</u>	<u>96.9</u>	93.5	85.5
	GEINet	94.4	93.4	89.8	81.9
65 度	Parallel-GEINet				
	ID+年齢+性別	<u>95.7</u>	<u>96.8</u>	<u>96.0</u>	<u>94.7</u>
	ID+性別	94.4	95.8	94.5	93.5
	ID+年齢	95.6	96.7	95.9	94.4
	GEINet	93.5	94.6	93.8	92.4
75 度	Parallel-GEINet				
	ID+年齢+性別	92.4	95.6	<u>96.8</u>	<u>96.5</u>
	ID+性別	<u>93.8</u>	94.1	95.3	95.2
	ID+年齢	92.1	<u>95.7</u>	<u>96.8</u>	96.2
	GEINet	89.5	93.9	95.3	94.5
85 度	Parallel-GEINet				
	ID+年齢+性別	86.2	93.4	<u>96.0</u>	<u>96.7</u>
	ID+性別	<u>89.3</u>	<u>96.1</u>	<u>96.0</u>	95.9
	ID+年齢	86.2	93.9	<u>96.0</u>	96.5
	GEINet	81.7	93.3	95.3	95.5

表 4 人物 ID と年齢について学習した Parallel-GEINet の EER[%] の比較

ギャラリーの角度	手法	プローブ画像の角度			
		55	65	75	85
55 度	Parallel-GEINet	2.0	1.6	2.7	4.1
	GEINet	<u>1.6</u>	<u>1.5</u>	<u>2.0</u>	3.3
	TCM+		3.2	4.0	5.7
	wQVTM		3.6	4.8	6.5
65 度	Parallel-GEINet	2.0	1.3	2.1	2.5
	GEINet	<u>1.7</u>	<u>1.4</u>	<u>1.5</u>	<u>2.0</u>
	TCM+	3.0		3.4	4.2
	wQVTM	3.5		3.4	5.1
75 度	Parallel-GEINet	2.9	2.0	1.7	1.8
	GEINet	<u>2.3</u>	<u>1.2</u>	<u>1.4</u>	<u>1.5</u>
	TCM+	4.0	3.4		3.8
	wQVTM	4.7	3.7		3.8
85 度	Parallel-GEINet	4.1	2.4	2.3	1.7
	GEINet	<u>3.4</u>	<u>1.8</u>	<u>1.5</u>	<u>1.5</u>
	TCM+	5.5	4.4	3.7	
	wQVTM	6.5	4.9	3.7	

た, といった要因が考えられる.

4.2.2 1 対 1 認証の比較

表 4 に 1 対 1 認証精度をまとめる. EER に関しては GEINet と比較して, 精度が劣化する結果となった. また, 表 5 に示した人物 ID と年齢, 性別情報を与えた場合の EER に関しても, GEINet より優れている箇所はあるものの劇的な精度改善は確認できなかった.

5. おわりに

本研究では, CNN を用いて人物属性を考慮した歩容認証の手法を提案した. 1 枚の GEI を入力とし, その被験者の

表 5 人物 ID, 年齢, 性別情報について学習した Parallel-GEINet の EER[%] の比較

ギャラリーの角度	手法	プローブ画像の角度			
		55	65	75	85
55 度	Parallel-GEINet				
	ID+年齢+性別	<u>1.5</u>	<u>1.4</u>	2.1	3.5
	ID+性別	2.3	2.1	3.3	4.8
	ID+年齢	2.0	1.6	2.7	4.1
	GEINet	1.6	1.5	<u>2.0</u>	<u>3.3</u>
65 度	Parallel-GEINet				
	ID+年齢+性別	<u>1.7</u>	<u>1.2</u>	<u>1.3</u>	<u>2.0</u>
	ID+性別	2.6	1.7	2.5	3.1
	ID+年齢	2.0	1.3	2.1	2.5
	GEINet	<u>1.7</u>	1.4	1.5	<u>2.0</u>
75 度	Parallel-GEINet				
	ID+年齢+性別	<u>2.0</u>	1.4	<u>1.3</u>	1.6
	ID+性別	3.7	2.5	2.1	2.3
	ID+年齢	2.9	2.0	1.7	1.8
	GEINet	2.3	<u>1.2</u>	1.4	<u>1.5</u>
85 度	Parallel-GEINet				
	ID+年齢+性別	3.5	<u>1.7</u>	<u>1.4</u>	<u>1.4</u>
	ID+性別	4.7	2.7	2.4	1.9
	ID+年齢	4.1	2.4	2.3	1.7
	GEINet	<u>3.4</u>	1.8	1.5	1.5

ID に加えて年齢や性別情報を並列に学習するネットワークを提案し, それを特徴抽出器として用いることで, 年齢や性別を考慮した特徴抽出器を実現し, 歩容認証に用いた. 大規模歩容データベースを用いた評価実験では, Rank-1 認証精度に関して既存手法よりも高い性能を達成することができた. 本論文での評価に用いたデータベースに含まれる観測視点の角度差は最大 30 度である. この角度差は非常に大きな差とは言えないものの, 現在の犯罪捜査においては秘匿撮影により観測視点と比較的近い歩容画像が鑑定対象となることが多いため, この視点差での認証精度の向上は犯罪捜査において非常に価値があると考えられる. ただし, 1 対 1 認証においては提案手法は既存手法の精度を改善できた, とは言えない. 1 対 N 認証において認証精度が向上していることより, 提案手法は本人と他人を識別する能力は改善しているものと思われる. それにも関わらず, 1 対 1 認証の精度改善がみられないのは, 今回考慮した属性ラベルの特性が, データに大きく依存して変わり, それにより計算される相違度が, プローブにより大きく変わっていることが考えられる. このような場合には, 対象データ毎に相違度を正規化することで精度改善が見込めるため, 今後は正規化等の導入により 1 対 1 認証の精度改善を目指していく. また本論文では人物属性として年齢, 性別の二種類しか用いなかったが, ほかに有効な属性情報があると考えられるため, 他の属性情報の考慮も今後進めていきたい.

謝辞 本研究は JSPS 基盤研究 (A)15H01693 の助成を受けたものである.

参考文献

- [1] Imed Bouchrika, Michela Goffredo, John Carter, and Mark Nixon. On using gait in forensic biometrics. *Journal of Forensic Sciences*, Vol. 56, pp. 882 – 889, 2011.
- [2] 警察庁. 平成 26 年度警察白書. 2014.
- [3] G. Johansson. Visual motion perception. *Scientific American*.
- [4] Y. Mukaigawa T. Echigo Y. Makihara, R. Sagawa and Y. Yagi. Gait recognition using a view transformation model in the frequency domain. *Proc.9th European Conference on Computer Vision*, pp. 151–163, 2006.
- [5] D. Muramatsu A. Mansur, Y. Makihara and Y. Yagi. Cross-view gait recognition using view-dependent discriminative analysis. *Int.Joint Conf. Biometrics (IJCB 2014)*, pp. 1–8, 2014.
- [6] S. Osindero G.E. Hinton and Y.-W. Teh. A fast learning algorithm for deep belief nets. *Neural Comput*, Vol. 18, No. 7, pp. 1527–1554, 206.
- [7] 武村紀子, 白神康平, 榎原靖, 村松大吾, 越後富夫, 八木康史. 畳み込みニューラルネットワークを用いた視点変化に頑健な歩容認証. *The 9th IAPR International Conference on Biometrics*, 2016.
- [8] M. Yang S. Ji, W. Xu and K. Yu. 3D convolutional neural network for human action recognition. *IEEE Trans.Pattern Anal.Mach.Intell.*, Vol. 35, No. 1, pp. 221–231, 2013.
- [9] S. Shetty T. Leung R. Sukthankar A. Karpathy, G. Toderici and L. Fei-Fei. Large-scale video classification with convolutional neural networks. *2014IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1725–1732, 2014.
- [10] M. Ranzato Y. Taigman, M. Yang and L. Wolf. Deep-face: Closing the gap to human-level performance in face verification. *2014IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1701–1708, 2014.
- [11] 塩土亮平, 福見稔, 伊藤桃代, 伊藤伸一. 深層学習を用いた手首筋電の個人認証. 電気学会研究会資料, pp. 57–61, 2017.
- [12] Gpu/fpga で深層学習/顔認識の性能向上——nec が画像認識ソフトを強化, 確認日:2017/02/20. <https://it.impressbm.co.jp/articles/-/14771>.
- [13] Y. Makihara H. Iwama, M. Okumura and Y. Yagi. The ou-isir gait database comprising the large population dataset and performance evaluation of gait recognition. *IEEE trans. Inf. forensics and Security*, Vol. 7, No. 5, pp. 1511–1521, 2012.
- [14] 村松大吾, 岩間晴之, 木村卓弘, 榎原靖. 1 歩行映像から取得される複数特徴を用いた個人認証. 電子情報通信学会論文誌, pp. 1–14, 2014.
- [15] A. Roy-Chowdhury A. Kale and R. Chellappa. Fusion of gait and face for human identification. *Proc.of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing 2004 (ICASSP' 04)*, Vol. 5, pp. 901–904, 2004.
- [16] AN. Rajagopalan M. Hofmann, S.M. Schmidt and G. Rigoll. Combined face and gait recognition using alpha matte preprocessing. *Proc. of the 5th IAPR Int. Conf. on Biometrics*, pp. 1–8, 2012.
- [17] M.S. Nixon T.N. Tan and R. Chellappa. Human identification based on gait. *Int.Series on Biometrics, Springer-Verlag*, 2005.
- [18] Z. Liu and S. Sarkar. Simplest representation yet for gait recognition: Averaged silhouette. *Proc. 17th International Conference on Pattern Recognition*, Vol. 1, pp. 211–214, 2004.
- [19] J. Han and B. Bhanu. Individual recognition using gait energy image. *IEEE Trans. Pattern Anal.Mach.Intell.*, Vol. 28, No. 2, pp. 316–322, 2006.
- [20] L. Wang J. Pu C. Wang, J. Zhang and X. Yuan. Human identification using temporal information preserving gait template. *IEEE Trans. Pattern Anal.Mach.Intell.*, Vol. 34, No. 11, pp. 2164–2176, 2012.
- [21] K.H. Cheung T.H.W. Lam and J.N.K. Liu. Gait flow image: A silhouette-based gait representation for human identification. *Pattern Recognit*, Vol. 44, pp. 973–987, 2011.
- [22] H. Iwama Y. Makihara, M. Okumura and Y. Yagi. Gait-based age estimation using a whole-generation gait database. *Proc. of the Int. Joint Conf. on Biometrics (IJCB 2011), Paper*, No. 195, pp. 1–6, 2012.
- [23] E. Hossain and G. Chetty. Multimodal feature learning for gait biometric based human identity recognition. *IEEE Trans. Multimedia*, Vol. 17, No. 11, pp. 1960–1968, 2015.
- [24] L. Wang Z. Wu, Y. Huang and T. Tan. A comprehensive study on cross-view gait based human identification with deep cnns. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 1, No. 1, pp. 1–10, 2016.
- [25] N. Guil F. de la Torre R. Medina-Carnicer M.J. Marin-Jimenez, F.M. Castro. Deep multi-task learning for gait-based biometrics. *IEEE International Conference on Image Processing*, 2017.
- [26] Y. Makihara Daigo Muramatsu and Y. Yagi. Cross-view gait recognition by fusion of multiple transformation consistency measures. *IET Biometrics*, Vol. 4, No. 2, pp. 62–73, 2015.
- [27] Y. Makihara Daigo Muramatsu and Y. Yagi. View transformation model incorporating quality measures for cross-view gait recognition. *Cybernetics, IEEE Transactions*, Vol. 46, No. 7, pp. 1602–1615, 2016.