

画像間差分と CNN を用いた低解像度画像による歩容認証

Gait recognition with low resolution images using inter-frame difference and CNN

○ 瀬戸口 直輝 (九大) 中嶋 一斗 (九大) 呉 桐 (九大)
 岩下友美 (NASA/Caltech JPL) 正 倉爪 亮 (九大)

Naoki SETOGUCHI, Kyushu University, setoguchi@irvs.ait.kyushu-u.ac.jp
 Kazuto NAKASHIMA, Kyushu University
 Wu TONG, Kyushu University
 Yumi IWASHITA, NASA/Caltech JPL
 Ryo KURAZUME, Kyushu University

Gait-based person recognition has received an increasing amount of attentions for monitoring and surveillance applications. One of issues in gait recognition is that it is difficult to recognize people with high performance, in case that the resolution of captured images is too low. To deal with this problem, this paper presents SFDEINet, which uses Signed Frame Difference Energy Image (SFDEI) as input images. SFDEI has an advantage of explicit representation of motion of walking people, by changing the size of frame difference. To take the size of frame difference into account, SFDEINet is designed to have convolution layers in parallel, followed by fusing outputs of each layer. We will show the SFDEINet is more suitable for SFDEI than GEINet in experiments.

Key Words: Gait recognition, Low resolution recognition, SFDEI, CNN

1 はじめに

近年, 世界的なテロ行為の過激化などを背景に, 生体情報を用いた個人識別は, 本人確認手段として重要性が高まっている. 生体情報の一つである歩容は, 識別対象者に特別な動作を行わせることなく, かつ非接触で取得可能であるという特徴がある. そのため, 歩容を用いた個人識別手法は犯罪捜査や防犯などの応用が期待されている.

歩容認証で高い性能を示す特徴の一つとして, Gait Energy Image (GEI) がある. GEI は一歩行周期分のシルエットの平均画像として得られるが, 撮影機器が低解像度機器である場合や, 機器と識別対象者との距離が遠く, 画像中の人物領域が小さくなる場合は, 得られる画像の解像度が低下するため, 識別性能も低下する問題があった. そこで, 我々は現在までに, 新たな歩行特徴として, GEI に画像間差分の平均を組み合わせた Signed Frame Difference Energy Image (SFDEI)[5] を提案している. 文献 [5] の GEINet を使用した実験では, 解像度が著しく低い画像を入力した場合, SFDEI が GEI よりも高い性能をもつことが報告されている. しかし, GEINet の構造は GEI に最適化されたものであり, SFDEI が個人間の歩行速度の差異を陽に表すという長所を活かしたネットワークとは言い難い.

そこで, 本稿では, GEINet[1] の畳み込み層を並列化して, 時間間隔 Δt の異なる複数の SFDEI を同時に学習するように拡張させた SFDEINet を提案する. 実験では, OU-ISIR Large Population Dataset[2] を用いて, GEINet と提案手法の比較を行い, 提案手法により, 低解像度画像を入力した場合の識別性能が向上することを示す.

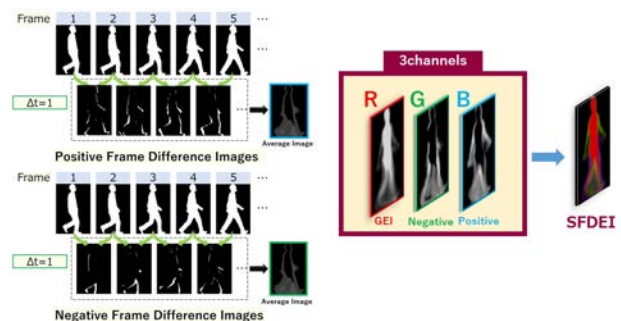


Fig.1 SFDEI

2 SFDEINet

SFDEI とは歩行画像列から一定間隔で画像間差分を求め, その値が正か負のいずれの値をもつかによって分けた 2 つの平均差分画像を, GEI に 3 チャンネルで組み合わせたものである. SFDEI を生成するまでの過程を図 1 に示す. SFDEI は歩行画像列の画像間差分をとる時間間隔 Δt を可変にすることで, 個人間の歩行速度の差異を陽に表現できるという長所がある. 例えば Δt の値を大きくするにつれて, 歩行速度が速い人ほど, 腕振りなどの動作情報の変化が大きくなる. その一方で, どのような軌跡で腕が振られたか等の動作情報が失われてしまう可能性がある. また, 文献 [5] では GEINet で抽出した複数の特徴を単純に結合して, 識別器である Nearest Neighbor に入力していたが, 単純な結合が最適であるとは限らない.

Table 1 Top-1 accuracy of combined features extracted by GEINet[%]

	結合した SFDEI の Δt の組み合わせ								
	1~2	1~3	1~4	1~5	1~6	1~7	1~8	1~9	1~10
解像度 (128×88)	92.5	92.8	92.8	93.2	92.8	92.7	92.6	92.5	92.1
解像度 (96×66)	92.5	92.8	92.9	93.0	92.7	92.7	92.4	92.4	92.0
解像度 (64×44)	92.5	93.0	92.9	93.0	92.9	92.7	92.6	92.5	92.1
解像度 (32×22)	92.1	92.4	92.2	92.2	92.6	92.4	91.9	91.8	92.7
解像度 (16×11)	84.5	85.4	85.6	85.4	85.5	85.5	85.5	85.6	85.25

Table 2 Top-1 accuracy of features extracted by SFDEINet[%]

	入力した SFDEI の Δt の組								
	1,2	1,3	1,4	1,5	1,6	1,7	1,8	1,9	1,10
解像度 (128×88)	93.8	92.9	92.4	93.2	92.9	93.0	92.6	92.5	92.0
解像度 (96×66)	93.9	92.4	92.7	92.8	93.0	93.2	92.6	92.3	91.9
解像度 (64×44)	93.5	92.9	92.5	93.3	93.2	92.9	92.5	92.4	91.8
解像度 (32×22)	93.5	92.5	91.9	92.9	92.5	93.2	91.9	92.1	91.5
解像度 (16×11)	85.5	86.9	85.0	87.7	86.1	86.7	86.5	88.1	86.9

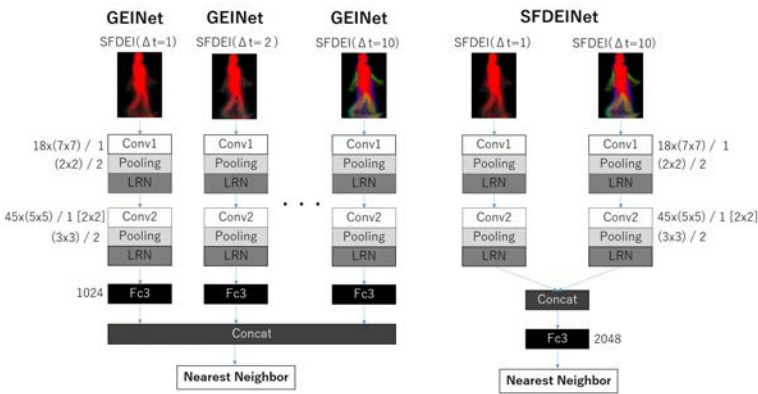


Fig.2 Difference of Identification method between GEINet and SFDEINet

conv : Number of filter×(filter size) / stride [padding size]
 pooling : (filter size) / stride
 Fc : Number of output node

そこで、畳み込み層を2並列にし、それぞれに Δt が異なる SFDEI を入力することで、歩行速度の差異と動作情報を合わせて学習でき、かつ人が介さずとも高性能な特徴が抽出できるよう GEINet を拡張した。GEINet は2層の畳み込み層 (Conv1, Conv2) と2層の全結合層 (Fc3, Fc4) から成る。そこから以下のように変更を加えた。

- 2層の畳み込み層を2並列化し、その出力を Fc3 層の前で結合 (学習するパラメータは独立)
- Fc3 層のパラメータ数を 2048 に変更

各畳み込み層のフィルタサイズ、チャンネル数、ストライドなどは GEINet と同一である。初期値について、重みは HeNormal、バイアスは0で固定している。また、畳み込み層からの出力に対して、Max pooling と Local Response Normalization (LRN) を適用している。

SFDEINet の畳み込み層の並列数と Fc3 層のパラメータ数は性能がピークとなる値に設定した。Fc4 層のパラメータ数は、後述

する実験設定に合わせて 956 としている。識別では、SFDEINet を特徴抽出器として用いて、抽出した特徴を Nearest Neighbor に入力することで識別を行う。図2に SFDEINet の構造の詳細を示す。

3 実験

3.1 実験設定

実験では、OU-ISIR Large Population Dataset[2] を用いた。このデータセットは約 4000 人の歩容データが含まれている。本実験では文献 [3] にあるように、1912 名を用いた方法で評価する。このデータセットはサブセットとして 10 通り (CV01~CV10)、各 956 名に、1 人につき 2 シーケンス分 (Gallery, Probe) のデータが存在し、その中には歩行方向が 55, 65, 75, 85 度の 4 通りが含まれている。本実験では歩行方向が 55 度のデータを使用し、学習に CV01、識別に CV02 を用いる。また、解像度が 128×88 の画像を、96×66, 64×44, 32×22, 16×11 の解像度に変更して、低解像度画像の識別性能を調べる。それぞれの画像解像度の一例を図3に示す。

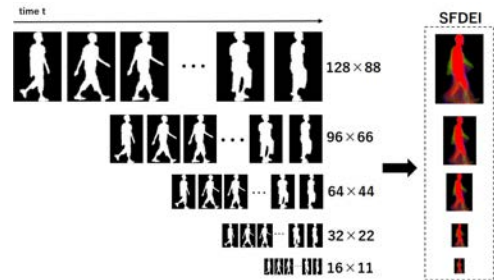


Fig.3 Changing of image resolution

学習時における学習データの総枚数は、956 名 × 2 シーケンス × 5 の計 9560 枚で、その中には、上述した計 5 種類の解像度の画像が含まれている。識別では CV02 の 956 名の Gallery と Probe を、学習済みのモデルに入力して特徴抽出を行う、その際 Fc3 層からの出力を特徴ベクトルとみなして、それらを Nearest Neighbor を用いて比較する。ここで、歩容認証では、データベ

スに登録されている画像は高解像度画像であると想定しているため、Nearest Neighbor に入力する教師データは 128×88 に固定し、評価データのみ解像度を変更したものを入力する。

GEINet を用いた実験では、各 Δt の SFDEI それぞれで特徴抽出を行い、識別器である Nearest Neighbor へ入力する前に、単純に特徴ベクトルを結合する。今回の実験では Δt を 1~10 と設定し、 $\Delta t=1 \sim N$ ($N=1,2,\dots,10$) の計 10 通りの組み合わせで特徴結合を行う。この実験は文献 [5] でも行われているため、実験結果は文献 [5] から一部引用する。

SFDEINet を用いた実験でも同様に Δt を 1~10 と設定する。GEINet での実験と異なる点は、特徴結合は行わず、SFDEINet への入力に用いる SFDEI のペアの組み合わせのみを変更した点である。また $\Delta t = 1, N$ ($N=1,2,\dots,10$) の計 10 通りの組み合わせで特徴抽出を行った。

3.2 実験結果

GEINet を用いた場合の結果を表 1 に、SFDEINet を用いた場合の結果を表 2 に示す。表の通り、 16×11 という解像度が著しく低い画像が入力されたときの最も高い識別率が、GEINet を用いて特徴抽出し、それを結合した場合は 85.6% であるのに対し、SFDEINet を用いて特徴抽出した場合は 88.1% となり、2.5% 性能が向上している。なお、従来の GEI のみを用いた場合は、80.3% であった。

4 まとめと今後の予定

本報告では、GEINet をベースに、畳み込み層を並列化して、異なる Δt で生成された SFDEI を同時に学習できるように拡張した SFDEINet を提案した。

OU-ISIR Large Population Dataset を用いた実験では、SFDEINet を用いることで、 16×11 という低解像度画像において、GEINet と比較して識別性能が向上することを示した。

今後は、まず SFDEINet のパラメータチューニングを行う。その後、SFDEI の Δt の値を今回設定したものより、更に大きく設定し、学習と識別に使う SFDEI の組を変えて、識別性能の変化を確かめる実験を行う。また、DNN を用いた最新の歩容認証手法において、低解像度画像を入力とした場合の識別率を検証することで、提案手法の優位性を確認する。

謝辞

本研究は文部科学省科学研究費補助金基盤研究 (A) (課題番号 JP26249029) の支援を受けた

参考文献

- [1] K. Shiraga, et al., "GEINet: View-Invariant Gait Recognition Using a Convolutional Neural Network", Int. Conf. on Biometrics (ICB 2016), No. O19, pp. 1-8, 2016.
- [2] Y. Makihara, et al., "The OU-ISIR Gait Database Comprising the Treadmill Dataset," IPSJ Trans. on Computer Vision and Applications, Vol. 4, pp. 53-62, 2012.
- [3] "http://www.am.sanken.osaka-u.ac.jp/BiometricDB/dataset/GaitLP/Benchmarks.html"
- [4] Noriko Takemura, Yasushi Makihara, Daigo Muramatsu, Tomio Echigo, Yasushi Yagi, "On Input/Output Architectures for Convolutional Neural Network-Based Cross-View Gait Recognition", IEEE Transactions on Circuits and Systems for Video Technology, vol.28, no.1, 2018.
- [5] 呉 桐, 中嶋 一斗, 瀬戸口 直輝, 岩下 友美, 倉爪 亮, 正負フレーム間差分を用いた深層学習による歩容認証, 画像の認識理解シンポジウム (MIRU2017), pp.PS1-19, 2017.8.8