

H-051

歩行の向きに依存しない多視点人物歩容認識

Gait Recognition Independent of Walking-Direction Using Multiple Cameras

朱 暁東[†] 高橋 友和[†] 井手 一郎[†]
 Xiaodong ZHU Tomokazu TAKAHASHI Ichiro IDE

目加田 慶人[‡] 村瀬 洋[†]
 Yoshito MEKADA Hiroshi MURASE

1 はじめに

近年、生体情報を用いた個人認証技術への関心が高まっている。認証に用いられる人間の生体情報としては、声紋、指紋、手の平の静脈、顔、歩容（歩き方）などがあげられる。この中で、歩容は人物がカメラから離れている場合でも認証が可能であるという点が特徴であり、部外者の侵入検知や、不審人物の認識などへの応用も期待されている。

人物歩容認識の従来手法としては、モデルベースの手法とアピランスベースの手法 [1] が提案されている。我々はこのうち、アピランスベースの手法に着目する。Han ら [2] は歩容を表す新たな特徴として GEI (Gait Energy Image) を使用し、個人認識を行う手法を提案した。これらの手法は単一カメラで、しかも人物の歩く方向が固定されている状態を仮定している。これに対し、Makihara ら [3] は撮影角度の変化に対応するために角度変化モデル VTM (View Transition Model) [4] を用い、歩容データに対する時間方向の FFT により得られる特徴を使用した個人認識を行った。しかし、この手法の場合でも、収集された歩容データと認識対象歩容データ間の撮影角度の差が大きければ、認識精度は下がる。

そこで、本稿では複数視点からの入力を用いることで、撮影角度の変化に、よりロバストな新たな人物歩容認識手法を提案する。歩容特徴として GEI を用い、VTM を用いることによって、さまざまな撮影角度からの歩容を辞書データとして生成する。そして、認識の際に複数のカメラ入力を利用することより、認識性能の向上を図る。提案手法を用いた歩容認識の実験を行い、認識性能の向上を確認した。

2 提案手法

2.1 手法の概略

本手法は学習段階と認識段階の二つの段階からなる。学習段階では、まず様々な人物の歩容データから GEI を特徴として抽出し、VTM を作成する。それから、VTM を用いて、認識対象の人物の辞書データを作成する。認識段階では、複数カメラから得られた GEI と辞書データを照合することで人物を認識する (図 1)。

2.2 Gait Energy Image

本研究で用いる歩容データは大きさの正規化が施されたシルエット画像 (二値画像) 列である。各人物、各

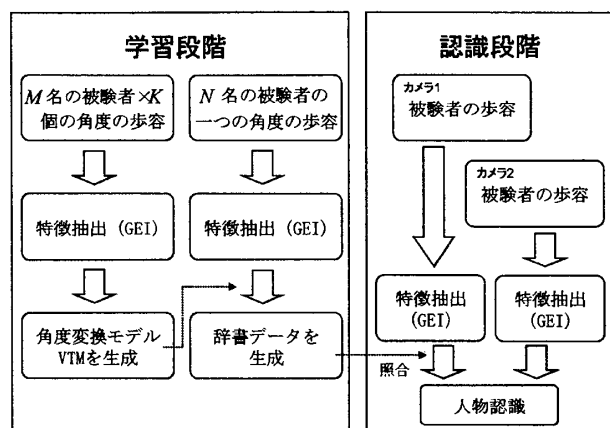


図 1: 提案手法の流れ (2台カメラを用いる場合)

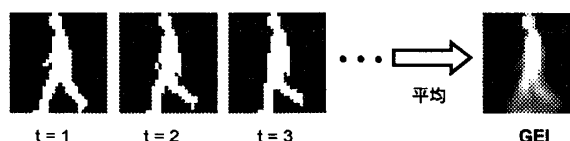


図 2: 時間方向の平均画像 (GEI)

撮影角度毎に、画像列を時間方向に平均することによって、GEI を得る (図 2)。

2.3 View Transition Model

ある被験者の撮影角度 θ_i の歩容特徴ベクトルを \mathbf{x}_{θ_i} としたとき、異なる角度 θ_j の歩容特徴ベクトル $\hat{\mathbf{x}}_{\theta_j}$ は以下の式で生成できる。これが VTM による角度変換である。

$$\hat{\mathbf{x}}_{\theta_j} = T_{i \rightarrow j} \mathbf{x}_{\theta_i} \quad (1)$$

$T_{i \rightarrow j}$ は、事前に他の人物の様々な角度の歩容データから作成する。以上の手法によって、変換された GEI と実際の GEI を図 3 に示す。VTM 作成の詳細は [4] に譲る。

2.4 学習段階

はじめに、 M 名の人物、 K 個の撮影角度からの GEI を用意して角度変換モデル VTM を作成する。次に、認識対象 (N 名) の辞書データを作成する。このとき、ある特定の角度 θ_i からの GEI のみを取得し、前に作成した VTM によって、それ以外の角度 θ_j ($j = 0, 1, 2, \dots, K, j \neq i$) の GEI を生成する。

[†] 名古屋大学 大学院情報科学研究科

[‡] 中京大学 生命システム工学部

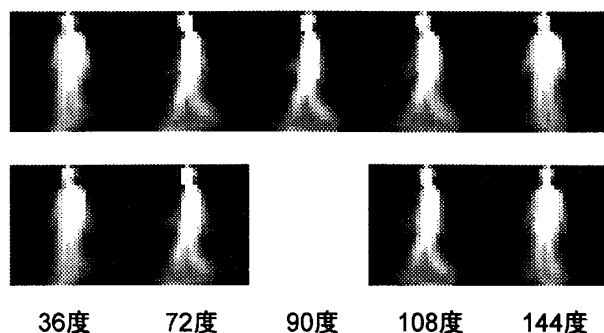


図 3: 実際の GEI(上) と 90 度の GEI から変換された GEI(下)

2.5 認識段階

まず、 C 台のカメラから撮影されたシルエット画像列から、それぞれ GEI を作成し、得られた入力データ y_c と人物 p 、向き θ の辞書データ $\hat{x}_{p,\theta+\alpha_c}$ の間の距離 2 種類を、以下のように定義する。2 種類定義したのは比較のためである。

$$d_{p,\theta}^{(\text{sum})} = \sum_{c=1}^C \|y_c - \hat{x}_{p,\theta+\alpha_c}\| \quad (2)$$

$$d_{p,\theta}^{(\text{min})} = \min_c \|y_c - \hat{x}_{p,\theta+\alpha_c}\| \quad (3)$$

ただし、 α_c はカメラ c の撮影角度である。そして、次の式により、認識結果 \hat{p} を得る。

$$\hat{p} = \arg \min_p \min_{\theta} d_{p,\theta}^{(A)} \quad (4)$$

ここで、 A が sum の場合と min の場合を次節の実験により比較する。

3 実験

3.1 実験条件

データベース (CASIA[6]) 中の 20 人分の画像を使い、撮影角度が 90 度 (真横) の歩容データから異なる 10 種の角度 ($0^\circ, 18^\circ, \dots, 72^\circ, 108^\circ, \dots, 180^\circ$) への VTM を作成した。VTM 作成とは異なる 65 人分のデータの撮影角度 90 度のデータに対して、VTM を用いてそれ以外の撮影角度の GEI を作成し、認識実験を行った。

3.2 結果と考察

認識時に 1 台のカメラのみを用いた場合の結果を (1 CAM)、角度差 90 度で 2 台カメラを用いて式 (2) と式 (3) を用いた場合の認識率をそれぞれ (2 CAMS+SUM)、(2 CAMS+MIN) として図 4 に示す。1 台のカメラを用いた場合には、入力された歩行の向きが辞書データ作成の際に与えた撮影角度である 90 度から離れるにつれて、認識率が低下している。それに対し、提案手法である 2 台カメラの用いた場合には、入力された歩行の向きが $0^\circ, 180^\circ$ の場合にも高い認識率を得ることができ、全体的にも認識率の向上が見られた。しかしながら、歩行の向きが $54^\circ, 72^\circ, 108^\circ, 126^\circ$ の場合には、1 台のカメラを用いた時と比較して認識率の低下が見られた。また、min と sum を比較すると、平

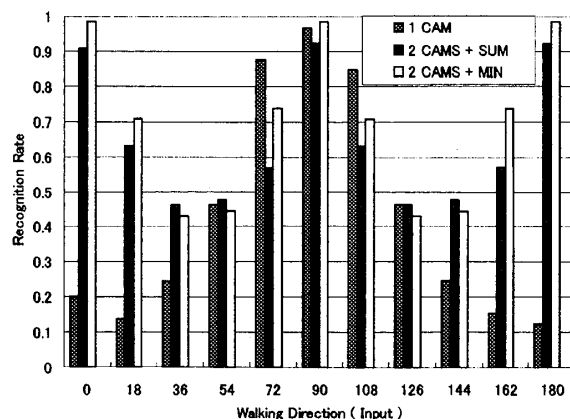


図 4: 2 台カメラを用いた実験結果

均的には min を用いた方が高い認識率が得られたが、sum を用いた方が良い場合もあった。このことから、2 台のカメラの入力をより適切に組み合わせて認識する手法の検討が課題として挙げられる。

4 まとめ

本稿では、複数カメラ入力の利用によって、学習時と認識時の撮影角度に変化があってもロバストに人物歩容を認識する手法を提案した。今後の課題としては、複数の入力を適切に組み合わせる方法の検討、辞書データ作成においても複数カメラ入力を用いることによる更なる認識性能の向上を検討している。また、VTM 以外の角度変換モデル、GEI 以外の特徴の検討も行いたい。

謝辞 日頃より熱心にご討論頂く名古屋大学村瀬研究室諸氏に感謝する。本研究の一部は日本学術振興会科学研究費補助金、21世紀CEOプログラム「社会情報基盤のための音声・映像の知的統合」による。

参考文献

- [1] H.Murase et al., "Moving object recognition in eigenspace representation: gait analysis and lip reading", Pattern Recognition Letters, Vol.17, pp.155-162, (1996)
- [2] J.Han, B.Bhanu., "Individual Recognition Using Gait Energy Image", IEEE Trans., Vol.28, No.2, pp.316-322, (2006)
- [3] Y.Makihara et al., "Gait Identification Using a View Transformation Model in a Frequency Domain", CVIM2006, pp.117-124, (2006)
- [4] A.Utsumi, N.Tesutani, "Adaptation of appearance model for human tracking using geometrical pixel value distributions", Proc. of the 6th Asian Conf. on Computer Vision, Vol.2, pp.795-799, (2004)
- [5] D. Tan et al., "Efficient Night Gait Recognition Based on Template Matching", Proc. ICPR2006, pp.1000-1003, (2006)