

歩容認証の高精度化のための 人物領域の抽出精度の評価

中村 浩一朗^{1,a)} 内海 ゆづ子^{1,b)} 岩村 雅一^{1,c)} 黄瀬 浩一^{1,d)} 槇原 靖^{2,e)} 村松 大吾^{2,f)}
八木 康史^{2,g)}

概要：防犯カメラによる犯罪捜査への応用を目的とした歩容認証が注目を集めている。歩容認証の手法は、画像上の人物領域を抽出して得られる歩容シルエット画像を用いる方法が現在の主流となっている。人物領域を手動で抽出すると負担がかかるため、自動で抽出する手法として、標準歩容モデルを用いた人物領域抽出が提案されている。この手法では、抽出する対象となる人物と類似する標準歩容モデルを選択し、適用することで、歩容の認証精度が向上することが示されている。しかし、人物領域の抽出精度の評価は限定的である。そこで、本研究では、標準歩容モデルの適用により、人物領域の抽出精度がどのように変化するかを調査した。また、類似する標準歩容モデルの選択方法の改善により、人物領域の抽出精度を更に向上させる手法を提案する。実験において、提案手法を用いたときが最も高い人物領域抽出精度を示した。

1. はじめに

人の生体情報を利用して個人を認証する生体認証技術の一つに歩容認証がある。歩容認証は、人の歩く様子を用いて認証を行うため、他の生体認証と比べて、遠距離から撮影された低解像度の映像でも人物を認証できるという利点がある。防犯カメラで撮影される動画は低解像度であるため、このような利点から防犯カメラによる犯罪捜査への応用が期待されており、既に警察機関において歩容に基づく人物鑑定システムの実証実験が実施されている [1]。また、刑事裁判において歩容認証を利用した鑑定結果の証明力が日本の裁判所に認められた事例もある [2]。

歩容認証の手法は、モデルベースとアピアランスベースの2手法に大きく分けることができる [3]。モデルベースの手法では、入力画像にモデルを当てはめることで、人の体型や動きといった特徴を抽出する。しかし、この手法は、低解像度の画像に対しては、モデルの当てはめに失敗しや

すいという欠点がある。そのため、防犯カメラで撮影された低解像度の動画による犯罪捜査には不向きである。それに対して、アピアランスベースの手法は、入力画像から直接特徴を抽出するため、防犯カメラで撮影された低解像度の画像でも特徴を抽出できる。

アピアランスベースの手法で用いられる特徴の中でも、歩容シルエット画像から特徴抽出を行う方法が現在の主流となっている [3][4]。しかし、歩容シルエット画像を得るために、画像上の人物領域を手動で抽出する作業は大変負担がかかる。そのため、人物領域を自動で高精度に抽出するセグメンテーション手法が求められている。歩容認証を目的としたセグメンテーション手法として、標準歩容モデルを用いた人物領域抽出 [5] がある。これは、背景差分に基づくグラフカット領域分割 [6] に、人の歩行の様子を表したシルエット画像列である標準歩容モデルを適用したものである。実験において、標準歩容モデルを適用することによって、歩容の認証精度が向上することが示されている。しかし、人物領域の抽出精度の評価は限定的であり、抽出精度の向上が歩容認証の高精度化に寄与しているかは十分には検証されていない。また、手法 [5] で選択される標準歩容モデルは、上半身は入力画像の前景領域と類似しているが、足元は大きくずれており、その影響により人物領域の足元の部分が適切に抽出されていない場合があった。

そこで、本稿では標準歩容モデルを適用したとき、人物領域の抽出精度がどのように変化するかを調査した。また、標準歩容モデルの選択が正しく行われるように手法を

¹ 大阪府立大学 〒 599-8531 堺市中区学園町 1-1
Osaka Prefecture University 1-1, Gakuencho, Naka, Sakai,
Osaka 599-8531, Japan
² 大阪大学 〒 565-0871 大阪府吹田市山田丘 1-1
Osaka University 1-1, Yamadaoka, Suita, Osaka 565-0871,
Japan
a) nakamura@m.cs.osakafu-u.ac.jp
b) yuzuko@cs.osakafu-u.ac.jp
c) masa@cs.osakafu-u.ac.jp
d) kise@cs.osakafu-u.ac.jp
e) makihara@am.sanken.osaka-u.ac.jp
f) muramatsu@am.sanken.osaka-u.ac.jp
g) yagi@am.sanken.osaka-u.ac.jp

改善し、そのときの人物領域の抽出精度と歩容の認証精度を評価した。

2. 関連研究

歩容認証を目的として、画像上の人物領域を自動で高精度に抽出する手法が求められている。人物領域の抽出には、画像に対し、領域分割を行うセグメンテーション手法を用いる。近年、セグメンテーション問題をエネルギー最小化問題として解く手法が提案されている。そのような手法として、Level Sets[7] や、グラフカット [8] が挙げられる。Level Sets は、境界線に対してのエネルギー関数を作成し、エネルギー関数が小さくなるように境界線を変化させる手法である。そのため、Level Sets は、局所解しか求めることができず、人物領域の抽出に用いた場合、適切に抽出できない場合がある。それに対して、グラフカットは、各領域からエネルギー関数を定義し、それらの大域解を求めることが可能である。

グラフカットを用いた手法として、Rother らは前景と背景の色分布をセグメンテーション結果から再学習し、繰り返しセグメンテーションをする Grab Cut[9] を提案している。Grab Cut は、最初に、前景を囲む矩形を入力し、矩形の外側を背景と仮定する。それに基づき、グラフカットを用いて、セグメンテーションを行い、混合ガウス分布を用いて、色分布を学習する。この処理を繰り返すことで精度の高いセグメンテーションを行う。

また、Li らは、グラフカットを基にした手法として、Lazy Snapping[10] を提案している。Lazy Snapping は、高速化を行うため、前処理として、watershed アルゴリズム [11] を用いて、画像のセグメンテーションを行っている。

槇原は、標準歩容モデルを用いた人物領域抽出 [5] を提案している。Grab Cut や Lazy Snapping は、セグメンテーションの際、抽出する対象を限定していないのに対して、手法 [5] は、抽出する対象を歩行する人物のみに限定している。そのため、人の歩行の様子を表したシルエット画像列である標準歩容モデルをグラフカットを用いた手法 [6] に適用して人物領域の抽出を行っている。このことから、手法 [5] は、本稿が想定している防犯カメラを用いた犯罪捜査への活用にも最も適していると考えられる。

3. 従来手法

本章では、提案手法の基となる標準歩容モデルを用いた人物領域抽出 [5] について説明する。標準歩容モデルとは、図 1 に示す人の歩行の様子を表したシルエットの時系列画像である。手法 [5] は、背景差分に基づくグラフカット領域分割 [6] に標準歩容モデルを適用したものである。標準歩容モデルを適用することで、歩容の認証精度が向上することが示されている。

手法 [5] による人物領域の抽出の流れを図 2 に示す。手

法 [5] では、防犯カメラで撮影された画像上から抽出する対象人物を矩形で切り出したものを入力画像とする。そして、入力画像の前景と背景の領域の情報（データ項）、入力画像の前景と背景の境界の情報（平滑化項）、人の形の情報（標準歩容モデル）を用いて、人物領域を抽出する。このとき、使用する標準歩容モデルは、抽出する対象となる人物と体型や姿勢が類似するものを選択する必要がある。以降、3.1 節で標準歩容モデルの選択について述べた後、3.2 節で、人物領域の抽出の方法について述べる。

3.1 標準歩容モデルの選択

使用する標準歩容モデルは抽出する対象となる人物と体型や姿勢が類似するものを選択する必要がある。入力画像と標準歩容モデルは時系列画像であり、位相に大きなずれが生じないようにするため、時系列に沿って、標準歩容モデルを選択する。このとき、人の歩くペースや開始の周期は異なるため、単純に時系列画像を比較することはできない。そのため、1 フレームごとに入力画像と標準歩容モデルとの比較を行い相違度を求め、最終的に DP マッチングを用いる。

標準歩容モデルの選択を行う前に、標準歩容モデルを作成しておく必要がある。標準歩容モデルに用いるそれぞれの人物の一周期分の歩容を表す時系列画像を用意し、手動で人物領域を抽出し、図 1 に示す歩容シルエットを作成する。

標準歩容モデルの選択の詳細について説明する。標準歩容モデルの選択の際、図 3 に示す入力画像の背景差分をし



図 1 標準歩容モデル

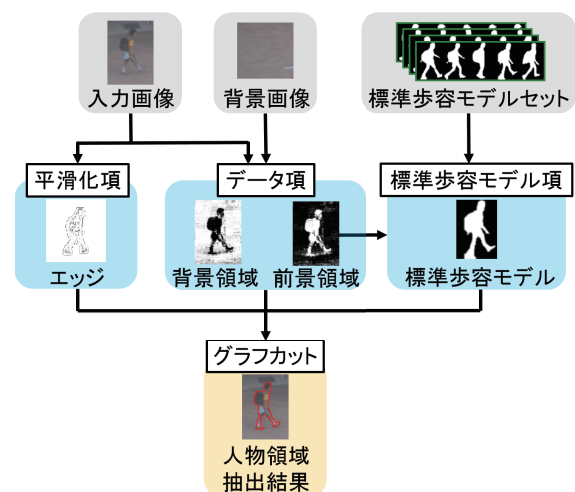


図 2 人物領域の抽出の流れ



図 3 入力画像の前景領域



図 4 入力画像の前景領域と標準歩容モデルの重ね合わせ

て得られる前景領域を用いる。このときの背景差分に用いる背景画像は、人物のいない背景画像を複数枚用意しておき、それらの各画素の画素値の平均をとることで作成する。

次に、入力画像の前景領域と標準歩容モデルの相違度の計算について説明する。相違度の計算には、谷本距離 [12] を用いる。標準歩容モデルと入力画像の前景領域を図 4 に示すように、重ね合わせ、その重なり具合を相違度とする。図 4 は、黒色の部分が入力画像の前景領域、赤色の部分が標準歩容モデルである。このとき、標準歩容モデルのスケールや座標を変更し、それぞれの相違度を求める。標準歩容モデルと入力画像の前景領域を重ね合わせる際、それぞれに外接矩形で正規化を行う。標準歩容モデルの人数を M とし、 m 番目 ($m = 1, \dots, M$) の人物のシルエットの歩容系列から、高さを正規化して位置合わせしたシルエット系列 $g_m(\phi, S) \in \mathbf{R}^{W \times H}$ を生成し、標準歩容モデルとする。ここで W, H はそれぞれ画像の幅と高さである。また、 $S = [s, s_x, s_y]^T$ として、 $\phi = (\phi_1, \phi_2, \dots, \phi_l)$ は位相、 $s = (s_1, s_2, \dots, s_r)$ はスケール、 $s_x = (s_{x_1}, s_{x_2}, \dots, s_{x_t})$ は x 方向の並進、 $s_y = (s_{y_1}, s_{y_2}, \dots, s_{y_v})$ は y 方向の並進を表す。1 人あたりの標準歩容モデルの総枚数を L 枚とすると、 $L = lrtv$ となる。同様に、入力画像の前景領域についても、正規化を行う。 n フレーム目 ($n = 1, \dots, N$) の入力画像の前景領域に対して、外接矩形が標準歩容モデルと同じ大きさになるように、正規化した画像を $f(n) \in \mathbf{R}^{W \times H}$ として表す。標準歩容モデル、入力画像の前景領域のそれぞれ x, y 座標の画素値を $I(g_m(\phi, S), x, y)$, $I(f(n), x, y)$ とし、入力画像の前景領域 $f(n)$ と標準歩容モデル $g_m(\phi, S)$ の谷本距離 d_1 は以下のように表される。

$$d_1 = 1 - \frac{\sum_{(x,y)} \min\{I(f(n), x, y), I(g_m(\phi, S), x, y)\}}{\sum_{(x,y)} \max\{I(f(n), x, y), I(g_m(\phi, S), x, y)\}} \quad (1)$$

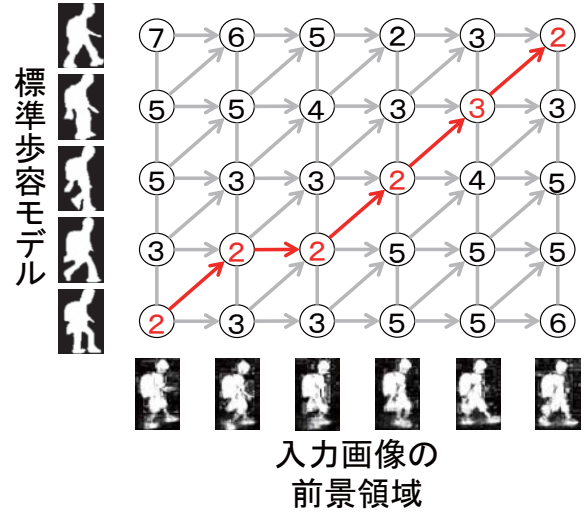


図 5 DP マッチングの概要

谷本距離を用いて得られた相違度を基に、DP マッチングを用いて、標準歩容モデルを選択する。DP マッチングを用いることで時系列に沿った標準歩容モデルの選択ができる。図 5 に DP マッチングの概要を示す。数字は相違度を表しており、赤い矢印は最適経路である。まず、各フレームの入力画像の前景領域とある人物から得られた標準歩容モデルの全てのフレームについて、相違度を計算する。そして相違度の和を DP マッチングの累積コストとし、累積コストが最小となる経路を求める。これを全ての標準歩容モデルの人物で計算して、累積コストが最小となる人物を選択することで、入力画像の各フレームについて類似した標準歩容モデルが求められる。

3.2 人物領域の抽出

人物領域の抽出は、入力画像の前景と背景の領域の情報（データ項）、入力画像の前景と背景の境界の情報（平滑化項）、人の形の情報（標準歩容モデル）を用いて行う。入力画像の各画素 q が背景に含まれると仮定するとき ($X_q = BG$)、または、 q が前景に含まれると仮定するとき ($X_q = FG$) に応じて定義される以下のエネルギー最小化問題として定式化される。

$$E(X) = \sum_{q \in Q} g_q(X_q) + \sum_{(p,q) \in E} h_{pq}(X_p, X_q) + \sum_{q \in Q} g_s(X_q) \quad (2)$$

ここで、 X は入力画像の全画素 Q に対するラベル X_q の集合を表し、 E は全ての隣接画素の集合を表す。また、第 1 項がデータ項、第 2 項が平滑化項、第 3 項が標準歩容モデル項である。

手法 [5] では、式 (2) のエネルギー最小化問題として解く手法として、グラフカット [8] を用いる。グラフカットは、各領域からエネルギー関数を定義し、それらの大域解を求めることが可能であるという利点がある。また、隣り合ったピクセルに滑らかさなどの制約条件を付ける問題に

適しており、手法 [5] でも、平滑化項が制約条件を担っている。以降、それぞれの項について説明する。

3.2.1 データ項

データ項は、入力画像の画素 q が背景に含まれると仮定したとき ($X_q = BG$) は、背景データを用いて、 q と背景の類似度を表現する。同様に、入力画像の画素 q が前景に含まれると仮定したとき ($X_q = FG$) は、前景データを用いて、 q と前景の類似度を表す。

データ項での、背景、前景の類似度表現方法について説明する。背景、前景の類似度計算では、背景データ、前景データをモデル化し、それを基に類似度を算出する。背景データは、入力画像に対する照明条件が異なる複数枚の背景画像から得られる背景の色情報である。まず、複数枚の背景画像を用いて画素 q に対する背景の画素値の平均 $\mu_{bg,q}$ と共分散行列 $\Sigma_{bg,q}$ を算出し、入力画像の画素 q の画素値 c_q に対するマハラノビス距離 $d_{bg,q}$ を以下のように計算する。

$$d_{bg,q} = (c_q - \mu_{bg,q})^T \Sigma_{bg,q}^{-1} (c_q - \mu_{bg,q}) \quad (3)$$

算出したマハラノビス距離に基づいて、入力画像の画素 q に対する背景の類似度を以下のように定義する。

$$g_q(X_q = BG) = \omega_{bg} \exp(-\kappa_{bg} d_{bg,q}) \quad (4)$$

ここで、 ω_{bg} , κ_{bg} は背景データ項の重みを制御するハイパーパラメータである。

次に、前景データについて説明する。前景は、入力画像によって、それぞれ異なるため、あらかじめモデル化することができない。そこで、入力画像に応じて、前景データを作成する。入力画像の画素 q と背景データで用いた複数枚の背景画像の平均 $\mu_{bg,q}$ の差分から、前景領域候補を抽出する。そして、前景候補領域の画素に対して、 k 平均法を適用することで k 個のクラスに分類し、各クラスをガウス分布で表現したものを重ね合わせた混合ガウス分布で表現する。そして混合ガウス分布の k 番目のガウス分布に対する平均 $\mu_{fg,k}$ と共分散行列 $\Sigma_{fg,k}$ から、入力画像の画素 q の画素値 c_q に対するマハラノビス距離 $d_{fg,k}$ を以下のように計算される。

$$d_{fg,k} = (c_q - \mu_{fg,k})^T \Sigma_{fg,k}^{-1} (c_q - \mu_{fg,k}) \quad (5)$$

そして、前景データ項 $g_q(X_q = FG)$ は、最小のマハラノビス距離を与える混合ガウス分布の要素に対する距離を用いて、以下のように定義する。

$$g_q(X_q = FG) = \omega_{fg} \exp(-\kappa_{fg} \min_k d_{fg,k}) \quad (6)$$

ここで、 ω_{fg} , κ_{fg} は前景データの重みを制御するハイパーパラメータである。

3.2.2 平滑化項

平滑化項は、人物領域を滑らかに抽出するために、隣接二点間での前景・背景ラベルの変化を抑制する項である。隣接二点間で画素値の変化が大きくなるところが境界となるようにし、隣接する画素 p , q の色 c_p , c_q における平滑化項を以下のように設定する。

$$h_{pq}(X_p, X_q) = \begin{cases} 0 & (X_p = X_q) \\ \omega_{sm} \exp(-\kappa_{sm} \frac{|c_q - c_p|^2}{|c_q + c_p|^2 + \epsilon}) & (X_p \neq X_q) \end{cases} \quad (7)$$

ここで、 ω_{sm} , κ_{sm} , ϵ は平滑化の重みを制御するハイパーパラメータである。

3.2.3 標準歩容モデル項

標準歩容モデル項は、3.1 節で得られた入力画像の前景領域と類似する標準歩容モデルの形状との類似度を表す。まず、入力画像の各フレームに対する標準歩容モデルのシルエットに対して、符号付き距離場 $d_{Q,sh}$ を計算し、標準歩容モデルのシルエットの形状を求める。符号付き距離場とは、各画素と標準歩容モデルの人物領域の一番近いエッジとの距離を格納している距離場に対して、各画素が標準歩容モデルの人物領域の内部か外部にあるかで異なる符号をつけたものである。符号付き距離場 $d_{Q,sh}$ を用いて、標準歩容モデルのシルエットの形状との類似度を以下のように定義する。

$$g_{sh}(X_q = BG) = \omega_{sh} \frac{1}{1 + \exp(-\kappa_{sh} d_{q,sh})} \quad (8)$$

$$g_{sh}(X_q = FG) = \omega_{sh} \frac{1}{1 + \exp(\kappa_{sh} d_{q,sh})} \quad (9)$$

ここで、 κ_{sh} , ω_{sh} は標準歩容モデル項の重みを表すハイパーパラメータである。

4. 提案手法

標準歩容モデルを用いた人物領域抽出 [5] において、図 6(a) に示すように、赤色の部分の選択される標準歩容モデルが、上半身は入力画像の人物領域と類似しているが、足元は大きくずれており、それに影響されて、図 6(b) に示すように、人物領域の足元の部分が適切に抽出されていない場合があった。本稿では、入力画像の前景領域と標準歩容モデルの谷本距離 [12] を用いた相違度の計算方法に問題点があると仮定し、相違度の計算方法の改善を提案する。

手法 [5] で、特に足元の領域抽出が失敗する原因は、足元や胴体などのパーツによって、人物領域の面積に差があるためではないかと考えられる。図 7 に、本稿で用いた人物画像を示す。このように、一般的に人物の全身を撮影した画像では、下半身に比べて上半身の面積の方が大きくなる。式 (1) では、全ての画素に対して同じ重みで類似度を計算しているため、上半身の相違度が強く影響し、適切な標準歩容モデルが選択されなかった。よって、全身の相違



(a) 標準歩容モデルの選択 (b) 人物領域の抽出結果

図 6 人物領域の抽出の失敗例



図 7 防犯カメラで全身を撮影した画像

度が等しく影響を与えるようにすれば、適切な標準歩容モデルが選択されると考えられる。

提案手法では、入力画像の前景領域と標準歩容モデルの相違度を計算する際、足元の影響が大きくなるように相違度の計算に重み付けを行う。提案手法では、標準歩容モデル、入力画像の前景領域のそれぞれの x, y 座標の画素値を $I(g_m(\phi, S), x, y)$, $I(f(n), x, y)$ とし、前景領域 $f(n)$ と標準歩容モデル $g_m(\phi, S)$ の相違度 d_2 を重み $\omega(x, y)$ を用いて以下のように定義する。

$$d_2 = 1 - \frac{\sum_{(x,y)} \omega(x, y) \min\{I(f(n), x, y), I(g_m(\phi, S), x, y)\}}{\sum_{(x,y)} \omega(x, y) \max\{I(f(n), x, y), I(g_m(\phi, S), x, y)\}} \quad (10)$$

式 (10) を用いることで、全身の相違度が、体の部位の大きさを考慮した形で扱われ、適切な標準歩容モデルが選択され、人物領域の抽出精度が向上すると考えられる。

5. 実験

標準歩容モデルを用いた人物領域抽出 [5] において、標準歩容モデルを用いることで、人物領域の抽出精度がどのように変化し、歩容認証の高精度化に寄与しているのかを調査した。以下では、手法 [5] で標準歩容モデルを用いずに人物領域を抽出する手法をモデル無し手法とし、手法 [5] を従来手法とする。また、提案手法を用いたときの人物領域の抽出精度と歩容認証の精度の調査を行った。

5.1 実験条件

実験には、小学校に設置された 2 台の防犯カメラで撮影した動画画像を用いた。各動画画像のスナップショットを図 8 に示す。また、動画画像の詳細は表 1 に示す。カメラ A、カメラ B でそれぞれ 56 人の歩行者を撮影し、このうち 50



(a) カメラ A (b) カメラ B

図 8 防犯カメラの映像のスナップショット

表 1 防犯カメラの映像の詳細

解像度 [画素]	800 × 450
フレームレート [fps]	12
歩行者の高さ [画素]	約 100(カメラ A), 約 80(カメラ B)

人を入力画像、6 人を標準歩容モデルとして用いた。入力画像として用いた画像は 1 人あたり 21~41 枚である。また、標準歩容モデルの枚数は 1 人あたり 9~13 枚である。また、人物領域の抽出精度を評価するため、カメラ A で撮影された 5 人分の入力画像に対し、図 9 のように、人物領域の正解データを手動で作成した。

人物領域を抽出する際、DP マッチングで選択された、入力画像の前景領域と類似する標準歩容モデルを用いて人物領域の抽出結果を出力し、歩容特徴を抽出した。提案手法では、式 (10) において、画像の下部分 20% の相違度を計算する際、 $\omega(x, y) = 7.5$ 、それ以外は $\omega(x, y) = 1.0$ となるように設定した。人物領域の抽出精度の評価には、人物領域の抽出結果と正解データを比較し、人物領域の抽出に失敗した画素の総数を i 、入力画像の画素の総数を u として、人物領域の抽出の誤差 e を以下のように定義する。

$$e = \frac{i}{u} \quad (11)$$

歩容認証の際、人物領域により得られたシルエット列から、高さを正規化して位置合わせをした正規化シルエット画像列を作成した。次に正規化シルエット列から周期検出を行い、歩容特徴として平均シルエットである Gait energy image[4] を作成した。得られた歩容特徴から、ユークリッド距離に基づいた照合 [1] を行った。カメラ A とカメラ B では歩行者の歩行方向が異なるため、カメラ A で撮影された動画を回転させ、カメラ B で撮影された動画と歩行方向を同じにしたものを用いた。カメラ A で撮影した入力画像から抽出される歩容特徴を Probe、カメラ B で撮影した入力画像から抽出される歩容特徴を Gallery とし、50 対 50 の認証を行った。50 人分の入力画像に対し、正しく認証される割合を歩容認証の一位認証率とした。



図 9 正解データ

表 2 人物領域の抽出誤差 [%]

人物	モデル無し手法	従来手法	提案手法
1	0.291	0.169	0.147
2	0.091	0.089	0.088
3	0.161	0.156	0.155
4	0.206	0.187	0.190
5	0.172	0.173	0.169
平均	0.184	0.154	0.149



(a) 従来手法



(b) 提案手法

図 10 従来手法と提案手法で選択された標準歩容モデルと抽出結果の比較

5.2 人物領域の抽出精度と歩容認証の精度の評価

モデル無し手法，従来手法，提案手法を用いたときの人物領域の抽出誤差と歩容認証の一位認証率を調査した。モデル無し手法，従来手法，提案手法を用いたときの人物領域の抽出誤差を，人物領域の正解データを作成した人物ごとと，その平均をそれぞれ表 2 に示す。表 2 の平均の抽出誤差から，モデル無し手法よりも，従来手法の方が抽出の精度が向上した。このことから，人物領域の抽出に標準歩容モデルを用いることは有用であると考えられる。また，提案手法を用いたときの人物領域の抽出誤差が最も低くなっていた。この原因として，従来手法では，足元がずれていて，適切に標準歩容モデルが選択されていなかったものが，提案手法を用いることによって，適切な標準歩容モデルが選択されるようになっていたことが考えられる。その一例を図 5.2 に示す。このように，標準歩容モデルの選択の改善により，人物領域の足元がより適切に抽出されていた。

また，モデル無し手法と，従来手法，提案手法のそれぞ



図 11 上半身のみ的人物領域

れの歩容認証の一位認証率を表 3 に示す。提案手法を用いたとき，従来手法と比較して，人物領域の抽出の精度は向上したが，歩容認証の一位認証率は低下していた。提案手法を用いることによって，主に人物領域の足元が適切に抽出されるようになっていたことから，人物領域の足元の抽出は歩容認証の高精度化に寄与していないと考えられる。

表 3 歩容認証の一位認証率 [%]

モデル無し手法	従来手法	提案手法
62	70	64

5.3 考察

提案手法を用いることで，標準歩容モデルを適切に選択できるようになり，モデル無し手法，従来手法と比較して，主に人物領域の足元が適切に抽出されるようになった。しかし，歩容認証の一位認証率は従来手法と比較して低下していた。この原因として，人物領域の足元の抽出が歩容認証の高精度化に寄与していないことが考えられる。よって，それぞれの手法で抽出された人物領域の画像の下部分 30%を黒塗りし，上半身部分のみを用いて歩容認証の一位認証率を評価した。歩容認証に用いた人物領域の例を図 11 に示す。このときのそれぞれの手法ごとの歩容認証の一位認証率が，実験の認証精度と近い値であれば，人物領域の下半身の抽出は，歩容認証の高精度化に寄与していないといえると考えられる。

モデル無し手法，従来手法，提案手法のそれぞれで抽出された人物領域の上半身部分を用いた歩容認証の一位認証率を表 4 に示す。それぞれの手法ごとの間に大きな変化はなく，人物領域に足元の部分を付加することで，実験の結果のように，それぞれの手法ごとの間に変化が生じたと考えられる。このことから，歩容認証の精度は人物領域の足元の抽出精度にも影響を受けていると考えられる。そのため，提案手法は，従来手法と比較して，人物領域の抽出精度は向上した。しかし，歩容認証の精度は低下した原因については，歩容特徴や認証方法に問題があると考えられ，更なる考察が必要である。

6. おわりに

本稿では，歩容認証を目的とした人物領域の抽出において，標準歩容モデルを適用したとき，人物領域の抽出精度がどのように変化するかを調査した。また，標準歩容モデル

表 4 上半身部分のみを用いたときの歩容認証の一位認証率 [%]

モデル無し手法	従来手法	提案手法
60	62	60

の選択が正しく行われるように手法を改善し、そのときの人物領域の抽出精度と歩容の認証精度を評価した。その結果、標準歩容モデルを適用することで、人物領域の抽出精度が向上していた。しかし、人物領域の抽出精度が向上したにもかかわらず、歩容の認証精度が低下する場合があった。本稿では、その原因として、人物領域の足元の抽出が、歩容認証の高精度化に寄与していないと仮定し、人物領域の上半身部分のみを用いて歩容認証を行った。その結果、それぞれの手法ごとの歩容の認証精度の間に大きな変化はなかったため、歩容の認証精度は人物領域の足元の抽出精度にも影響を受けていると考えられる。そのため、歩容特徴や認証方法に問題があると考えられ、更なる考察が必要である。

謝辞 本研究は、JSPS 基盤研究 (A)JP15H01693 の助成を受けたものである。

参考文献

- [1] 岩間晴之, 村松大吾, 榎原靖, 八木康史: 犯罪捜査支援のための歩容に基づく人物鑑定システム, 情報処理学会研究報告コンピュータビジョンとイメージメディア (CVIM), Vol. 2013-CVIM-186, No. 3, pp. 1–10 (2013).
- [2] 黒沢健至: 防犯カメラ映像の解析技術, セイフティエンジニアリング, Vol. 43, 公益財団法人総合安全工学研究所, pp. 21–25 (2016).
- [3] 東山侑真, 榎原靖, 西野恒, 八木康史: 様々な歩行状況下における歩容認証手法の性能評価, 情報処理学会研究報告コンピュータビジョンとイメージメディア (CVIM), Vol. 2013-CVIM-187, No. 10, pp. 1–8 (2013).
- [4] Han, J. and Bhanu, B.: Individual recognition using gait energy image, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 2, pp. 316–322 (2006).
- [5] Makiyara, Y., Tanoue, T., Muramatsu, D., Yagi, Y., Mori, S., Utsumi, Y., Iwamura, M. and Kise, K.: Individuality-preserving silhouette extraction for gait recognition, *IPSJ Transactions on Computer Vision and Applications*, Vol. 7, pp. 74–78 (2015).
- [6] Makiyara, Y. and Yagi, Y.: Silhouette extraction based on iterative spatio-temporal local color transformation and graph-cut segmentation, *Proceedings of 19th International Conference on Pattern Recognition*, pp. 1–4 (2008).
- [7] Sussman, M., Smereka, P. and Osher, S.: A level set approach for computing solutions to incompressible two-phase flow, *Journal of Computational Physics*, Vol. 114, No. 1, pp. 146–159 (1994).
- [8] Boykov, Y. Y. and Jolly, M.-P.: Interactive graph cuts for optimal boundary & region segmentation of objects in ND images, *Proceedings. Eighth IEEE International Conference on Computer Vision*, Vol. 1, pp. 105–112 (2001).
- [9] Rother, C., Kolmogorov, V. and Blake, A.: Grabcut: Interactive foreground extraction using iterated graph cuts, *ACM Transactions on Graphics (TOG)*, Vol. 23,

- No. 3, ACM, pp. 309–314 (2004).
- [10] Li, Y., Sun, J., Tang, C.-K. and Shum, H.-Y.: Lazy snapping, *ACM Transactions on Graphics (ToG)*, Vol. 23, No. 3, pp. 303–308 (2004).
- [11] Vincent, L. and Soille, P.: Watersheds in digital spaces: an efficient algorithm based on immersion simulations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 6, pp. 583–598 (1991).
- [12] Sloan, K. R. and Tanimoto, S. L.: Progressive refinement of raster images, *IEEE Transactions on Computers*, Vol. 28, No. 11, pp. 871–874 (1979).