

犯罪捜査支援のための歩容に基づく人物鑑定システム

岩間 晴之^{1,a)} 村松 大吾¹ 楨原 靖¹ 八木 康史¹

概要: 近年、歩容に基づく人物認証技術の犯罪捜査応用が注目を集め、実際に活用されはじめている。監視カメラ映像で観測された犯人の歩容特徴から対象人物を鑑定するためには、当該分野の専門的知識及びスキルが必要となるため、従来の犯罪捜査においては、歩容認証を専門とする研究者などの歩容の専門家にそれを依頼してきた。しかし、より迅速かつ効率的な犯罪捜査を行うためには、歩容の専門家ではない犯罪捜査員が、手で即時に人物鑑定結果を得ることが望ましい。そこで本研究では、そのような歩容の非専門家による使用を前提とした、世界初の歩容に基づく人物鑑定システムを構築した。本システムは、最先端の歩容認証技術が実装されているだけでなく、GUIに基づく簡易な操作インタフェースを備え、非専門家であっても、専門家と同様の人物鑑定結果を、簡易な操作手順によって得ることができるよう設計されている。実験では、一人の歩容の専門家及び10人の非専門家を被験者とし、本システムによる模擬鑑定実験を行った。結果として、非専門家が行った合計50組の鑑定のうち、46組の鑑定で専門家と同様の結果を得ることができ、本システムの有用性を確認することができた。

1. はじめに

現在、監視カメラの設置台数は増加の一途をたどり、監視カメラ映像からの人物認証技術は犯罪捜査において重要な役割を占めるようになりつつある。特に、人の歩き方から人物を認証する歩容認証技術は、遠距離から撮像した低解像度な映像からでも人物を認証できるという長所を持ち、近年その犯罪捜査応用が注目を集めている。実際、これまで様々な歩容認証技術がコンピュータビジョンの研究分野において提案されてきている [1], [2], [3], [4], [5]。さらに、イギリスでは歩容に基づく手掛かりが犯罪の証拠として法廷で認められた実例があり [6]、また日本においても、犯罪捜査のための有力な情報として活用されはじめている。

犯罪捜査における、コンピュータビジョン技術に基づく歩容認証技術の主たる役割は、犯行現場付近で撮影された映像中の犯人と、その他の監視カメラまたは犯罪捜査員により秘匿撮影された容疑者、もしくは犯罪捜査機関のデータベースに登録された前科者とを、歩容の特徴に基づいて定量的に鑑定することである。特に、犯罪捜査応用において注意すべき点は、その鑑定結果が犯罪捜査方針全体に大きな影響を及ぼし、さらには裁判における証拠ともなり得るという点であり、歩容に基づく鑑定結果は十分に合理性を持ったものでなくてはならない。従って、鑑定者は歩容

認証処理の各過程（対象の選択、追跡、領域分割、特徴抽出、特徴照合、といった過程）を、逐次目視で確認、さらにはその良否や正誤を判断する必要がある（例えば、対象を正しく選択しているか、歩容特徴に予期せぬノイズが入っていないか、などの確認）。

この点を考慮すると、全自動のソフトウェア処理によって鑑定結果を得るのは望ましくないといえ^{*1}、歩容認証の各過程の処理を行う自動処理と、各処理に対する目視確認や判断の結果をフィードバックしながら歩容認証処理を進めていけるような、対話的な枠組みで鑑定が行われるのが望ましいといえる。そのような対話的な鑑定作業を行うためには、歩容認証の各過程に関する専門的知識が必要不可欠となる。例えば、どのような自動処理を、どのようなプログラムで実装し、どのような順番で行い、またどのような観点でその結果を確認しながら処理を進めていけば良いかなど、これら全てを理解・実行できなければならない。

しかし、歩容認証は、指紋や顔などと違って比較的新しい技術であり [7]、現状では大半の犯罪捜査員はそのような専門的知識を有しておらず、そういった鑑定作業を自力で行うのは困難である。また、その作業を支援するためのパッケージ化されたソフトウェア等も存在していない。結果として、犯罪捜査員はその作業を、歩容認証を専門とする研究者のような歩容の専門家に依頼することになる。そのような歩容に基づく人物鑑定作業の流れの一例を図1に

¹ 大阪大学 産業科学研究所
The Institute of Scientific and Industrial Research, Osaka University

^{a)} iwama@am.sanken.osaka-u.ac.jp

^{*1} 常に完全な鑑定結果の出力を保証する自動処理ソフトウェアは存在しない。

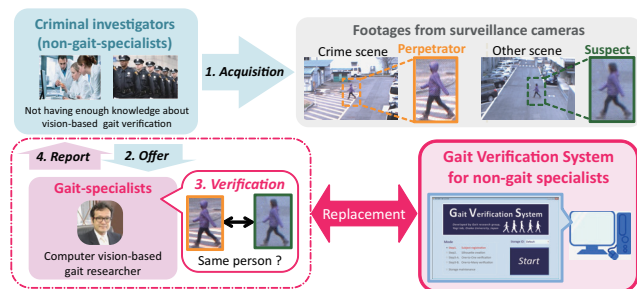


図 1 歩容に基づく人物鑑定作業の流れの一例と構築したシステムの目的。まず犯罪捜査員は犯人と容疑者がそれぞれ映った映像を入手する。次に、犯罪捜査員はその両者の歩容に基づく人物鑑定を、歩容認証を専門とする研究者のような歩容の専門家に依頼する。そして依頼を受けた専門家は、鑑定結果を犯罪捜査員に報告し、犯罪捜査員はその結果を受けて操作の次の方針を決めることになる(例えば、その容疑者に対する捜査を更に行う必要がある、または捜査対象から外す、など)。本システム構築の目的は、この作業の流れにおける歩容の専門家の置き換えである。

示す。実際に、我々の研究室においては、ここ数年間、一カ月に一件以上の頻度で、歩容に基づく人物鑑定依頼が日本の警察機関から寄せられている。

このようにして、専門家に鑑定を依頼することは、犯罪捜査を行う上で妥当な選択肢の一つであるといえるが、犯罪捜査の度に毎回そのような依頼をすることは、コストと時間の観点から非効率である。特に、時間という要素は多くの場合、犯罪の初動捜査において重要な要素であることを考えると、犯罪捜査員が自らの手で、即時に、歩容に基づく人物鑑定結果を得られるような枠組みが構築されることが望ましい。

そこで本研究では、そういった犯罪捜査を支援するための、歩容に基づく人物鑑定システムを構築した。本システムの目的は、従来の鑑定作業の流れにおける、歩容の専門家の置き換え(図 1)である。即ち、歩容に基づく人物鑑定の概略と本システムの使い方に関する一通りの講習を受けた犯罪捜査員が、説明書に沿って本システムを操作するだけで、専門家と同等の鑑定結果を得ることができるようになることを目的とする。我々の知る限り、本システムは世界初の歩容に基づく人物鑑定システムである*2。

犯人を鑑定するシナリオは様々に考えられるが、本研究では以下の二つのシナリオに着目する。

- (1) 個別鑑定: 犯人と犯罪捜査員が捜査対象とみなす容疑者との間での個別鑑定
- (2) 一括鑑定: 犯人と犯罪捜査機関のデータベースに登録された前科者達との間での一括鑑定

上記において、技術的には本来、一括鑑定は個別鑑定を包含するが、実用における利便性を考慮し、これらを別々に扱っている。

*2 大学や他の研究機関において使用されている個々の研究用プログラムなどは除く。

本システムは前記の二つのシナリオに対応する二種類の鑑定結果を出力し、鑑定処理の全過程、即ち鑑定対象の選択から歩容特徴の照合までを支援する。本システムの入力と出力を図 2 に示す。また、本システムの特徴を以下に示す。

● **非専門家向けインターフェース**

本システムでは全ての操作を GUI を介して行い、ユーザが計算機に対する素養を持ってさえいれば、対話的に鑑定処理を進めていけるよう設計されている。

● **最先端の歩容解析**

鑑定のための歩容特徴として、現在最も幅広く使用されている、Gait Energy Image (GEI) [1] を用いる。GEI は大規模データベース [8] を用いた性能評価において、高い識別能力を持つことが報告されている [9]。さらに、方向変換モデル [2], [10] の導入により、鑑定対象の間で観測方向(対象の見え方の方向)が異なる場合の鑑定も行うことができる。

● **客観的な出力**

鑑定結果は、鑑定された対象の組が同一人物であることを示す事後確率の形で出力され、これは鑑定された対象間の歩容特徴の距離に基づいて計算される。個別の鑑定機能においては、ユーザが指定する一組の鑑定対象(プローブとギャラリー)に関する事後確率が出力される。また、一括の鑑定機能においては、ユーザが指定するプローブと、システムに登録されている全ギャラリーの間で事後確率が計算され、それらの中で、事後確率の大きい順に、上位 10 位までの鑑定結果が出力される。ここで、本システムは事後確率という客観的数値を出力するだけであり、鑑定対象の組が同一人物であるか否かを最終判断するシステムではない、ということに注意されたい。

● **様々な状況への適用性**

鑑定される対象が撮影された状況はしばしば異なる。例えば、対象が撮影された空間解像度、フレームレート、観測方向、対象の服装や荷物などの状況は異なる場合が多い。本システムではそういった状況の差異を考慮し、各状況に合わせた事後確率を出力する。

本稿では、まず 2 章で歩容に基づく鑑定処理の枠組みを説明し、3 章で構築したシステムの詳細に関して述べる。さらに 4 章でシステムの有効性を実験結果によって示し、5 章で考察を行う。最後に 6 章で本研究についてまとめる。

2. 歩容に基づく人物鑑定処理の枠組み

本章では、本システムに実装されている、歩容に基づく人物鑑定処理の枠組みを示す。鑑定処理においては、鑑定される対象の組(プローブとギャラリー)の歩容特徴が計算、及び比較され、その特徴間の類似度に基づいて、その対象同士が同一人物であることを示す事後確率が計算される。

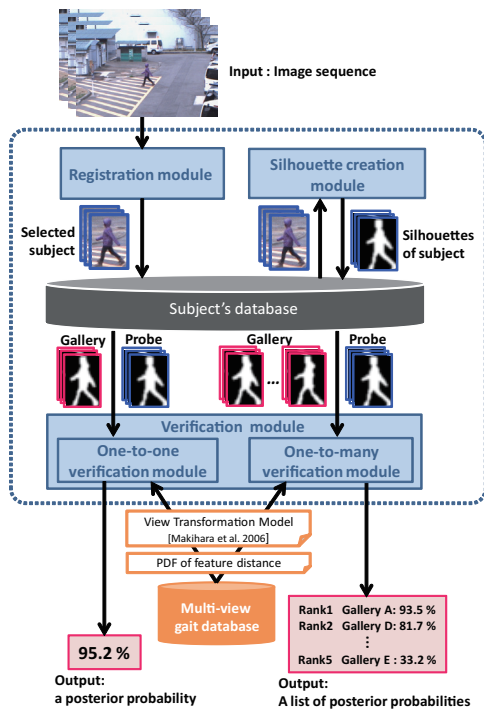


図 2 構築したシステムのデータの流れ。

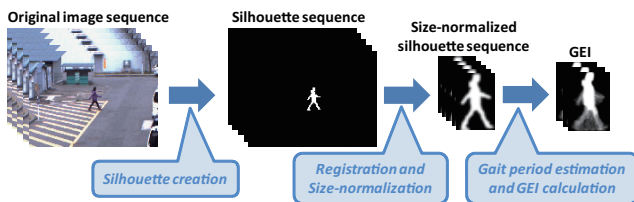


図 3 GEI 生成の流れ。

2.1 歩容特徴と距離計算

本システムで歩容特徴として用いる GEI の生成の流れを図 3 に示す。

N フレームの画像から構成される画像列における歩行周期を M フレーム ($M \leq N$) とすると、その画像列において、区間の先頭フレームを 1 フレームずつずらして考えると、合計で $N - M + 1$ 個の異なる歩行周期区間を設定することができる。そして、プローブとギャラリーの画像列における歩行周期区間数をそれぞれ L_p , L_g , また第 k 区間における GEI を $\mathbf{x}_{p,k}$, $\mathbf{x}_{g,k}$ とし、プローブとギャラリーの特徴間距離 D を次式で定義する。

$$D = \min_{i,j} \|\mathbf{x}_{p,i} - \mathbf{x}_{g,j}\|_2, \quad (1)$$

ここで、 i ($1 \leq i \leq L_p$), j ($1 \leq j \leq L_g$) はそれぞれプローブとギャラリーの画像列中の区間番号、 $\|\cdot\|_2$ はユークリッド距離を表す。

2.2 観測方向や服装の差異が存在する場合の照合

鑑定対象の組 (プローブとギャラリー) の状況間に差異が存在する場合、その差異が歩容特徴の識別能力に影響を与える可能性がある。特に、対象の観測方向、服装の種類や

荷物の有無といった因子の差異は識別能力を大幅に低下させることが報告されている [11], [12]。さらに、そのような差異がある下でそのまま鑑定することは、鑑定そのものの妥当性を損なう恐れがある。

従って、本研究では観測方向の差異に対しては、歩容特徴の方向変換モデル [2] に基づく歩容特徴照合の枠組みを取り入れ、また服装や荷物の差異に対しては、特徴のマスクによる部分領域に基づく歩容特徴照合の枠組みを取り入れることで、これらの問題を解決する。即ち、本枠組みでは、式 (1) に示した距離の計算に用いられる GEI \mathbf{x} は、方向変換処理とそれに続く特徴のマスク処理が施された GEI \mathbf{x}' と書き改められる。

ここで、鑑定の際には、鑑定対象の組それぞれに対し、観測方向とマスク領域を適切に指定する必要があるが、本システムにおけるそれらの指定方法は、3 章にて詳細に示す。

2.3 状況に合わせた事後確率

鑑定する対象間の特徴距離を計算した後、その距離に基づき鑑定対象の組が同一人物である確率を表す事後確率を、鑑定対象それぞれの状況 (カメラの設定、観測方向、服装・荷物など) を考慮した上で計算する。鑑定するプローブとギャラリーの状況を表す変数をそれぞれ I_p , I_g とすると、両者の特徴間距離 D が計算された下での事後確率は次式で計算される。

$$P(X=1|D; I_p, I_g) = \frac{P(D|X=1; I_p, I_g)P(X=1)}{\sum_X P(D|X; I_p, I_g)P(X)}, \quad (2)$$

ここで、 X はプローブとギャラリーが同一人物であるか ($X=1$)、異なるか ($X=0$) を表す確率変数である。

2.3.1 状況の定義

式 (2) に示す事後確率を計算するためには、状況の組み合わせに応じた、特徴間の距離に関する確率密度関数 ($P(D|X; I_p, I_g)$, $X \in \{0, 1\}$) を学習しておくことが必要となる*3。ここで、鑑定対象間のあらゆる状況の組み合わせを考えると、無限に近い組み合わせの数となるが、それら全組み合わせに対して十分な学習データを準備するのは不可能である。従って、考慮する状況は、妥当性を損なわない程度に近似的に (離散的に) 定義される必要があり、その妥当性と学習データ収集の現実性のトレードオフを考慮した上で、注意深く定義されなければならない。

上記議論を考慮した上で、本研究では、映像中の対象の大きさ、フレームレート、観測方向、及び特徴のマスク領域、という四つを、確率密度関数に影響を与える因子として考える。ただし、対象の大きさに関しては、シルエットの正規化サイズ (GEI のテンプレートサイズ) を考えることとする。そして、これら四つの因子をさらに、表 1 に示

*3 式 (2) の計算には事前確率 $P(X)$ の設定も必要となるが、本システムでは全ての状況に対し、 $P(X=1) = P(X=0) = 0.5$ と設定する。

表 1 因子の定義

因子	考慮する設定の種類
GEI の大きさ (テンプレートの大きさ)	22×32 [pixels], 44×64 [pixels]
フレームレート	5, 7.5, 15, 30 [fps]
観測角度	12 方位角 ×2 俯角 (図 8 に各角度の画像を例示)
特徴のマスク	11×16 [blocks] (図 10(a) を参照)

表 2 各モジュールにおける代表的機能と処理モード

機能	処理モード
登録モジュール	
画像列の入力	手動
対象の選択	半自動 (対話的)
対象の登録	手動
シルエット生成モジュール	
シルエットの生成	半自動 (対話的)
シルエットの保存	手動
個別/一括鑑定モジュール	
特徴の計算	自動
特徴のマスキング	手動
事後確率の計算	自動

すように離散化して考える。ここで、マスク領域に関しては、図 10(a) に示すようなブロック単位のマスク領域を考えることに注意されたい。本研究ではこれらの因子を状況として定義し、それらの組み合わせを考慮した確率密度関数を学習する。

鑑定においては、プローブとギャラリそれぞれの状況をその都度指定する必要があるが、本システムにおけるその指定方法は 3 章にて詳細に示す。

2.4 方向変換モデルと状況に合わせた確率密度関数の学習

本研究では、二つのデータセットから構成される多視点歩容データベースを用い、方向変換モデル及び状況に合わせた確率密度関数をそれぞれ学習する。これらのデータセットは、トレッドミル上の被験者の歩行を 24 台のカメラ (12 方位角 ×2 俯角) により撮像した、正規化シルエット画像列の集合から構成される。この 24 台のカメラの観測方向は、表 1 に示す観測方向に対応する。また、正規化シルエットのサイズは 88×128 [pixels] であり、撮像フレームレートは 60 [fps] である。方向変換モデルの学習用データセットは被験者 82 人の合計 114 画像列から構成され、状況に合わせた確率密度関数の学習用データセットは被験者 103 人の合計 206 画像列から構成される。

これらのデータセットそれぞれに対して、正規化シルエット画像列のリサイズとリサンプリングを行うことで、異なる正規化サイズ (22×32, 44×64 [pixels])、及び異なるフレームレート (5, 7.5, 15, 30 [fps]) の正規化シルエット画像列を生成し、それらを拡張データセットとして学習に利用する。

3. 歩容に基づく人物鑑定システム

3.1 システムの概要

本システムはシステムに登録されたプローブとギャラリの画像列それぞれで抽出される歩容特徴に基づき、両者を鑑定する機能を提供する。なお、本稿では便宜上、犯人をプローブ、容疑者もしくは前科者をギャラリと見做す。

本システムは一つのデータベースと四つのモジュールから構成され、それぞれのモジュールを、登録モジュール、シルエット生成モジュール、個別鑑定モジュール、一括鑑定モジュールと称する。これらの間でのデータの流れを図 2 に、各モジュールの主な機能を表 2 に示す。また、四つのモジュールのインターフェースダイアログを図 4(b), (c), 図 5, 及び図 6 に示す。これらのダイアログは図 4(a) に示すメインメニューダイアログ (システムのホームダイアログ) から起動する。

3.2 各機能の処理モード

1 章で述べた通り、犯罪捜査応用において、鑑定処理の結果は犯罪の重要な証拠となり得るため、その結果は十分な妥当性を持たなければならない。従って、鑑定処理の各過程は鑑定者 (ユーザ) によって適宜確認及び判断されなければならない。さらに、場合によっては、人間の手動処理がどうしても必要になる、または自動処理よりも効率が良くなることもあり得る。例えば、複数の対象が映っている画像列から所望の対象 (犯人や容疑者) を指定する場合、それを指定できるのは、それを把握しているユーザだけである。また別の例として、シルエット生成過程におけるノイズ除去処理は、画像全体に対する包括的な自動ノイズ除去処理を行うよりも、手動によるノイズ領域指定 (例えばマウス操作による該当領域の指定) とその指定領域の局所的な自動ノイズ除去処理を組み合わせ対話的に行う方が効率的である。こういった点を十分に考慮し、我々は、システムの使用性、処理効率性、及び処理の妥当性、という観点から、各モジュールの各機能の処理モードを表 2 に示すように、自動、半自動 (対話的)、手動、のいずれかに設定した。

3.3 システムのモジュールとユーザの操作

本節では各モジュールにおけるシステムの動作とユーザの操作を述べる。



図 4 メインメニュー、登録モジュール、シルエット生成モジュールのインターフェースダイアログ。(A) 入力画像列、(B) 表示されている画像のフレーム番号、(C) 画像列閲覧制御パネル、(D) 対象の切り抜き画像、(E) 対象のシルエットの輪郭線 (赤線)。

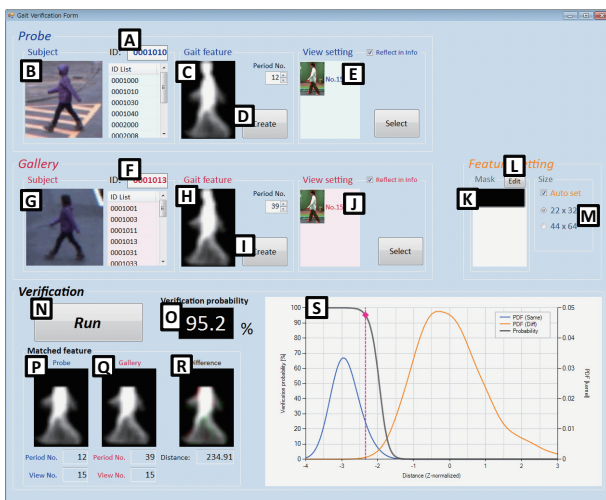


図 5 個別鑑定モジュールのインターフェースダイアログ。(A) 選択されたプローブの ID、(B) プローブの切り抜き動画、(C) プローブの GEI、(D) プローブの GEI 生成ボタン、(E) プローブの観測方向、(F) 選択されたギャラリーの ID、(G) ギャラリーの切り抜き動画、(H) ギャラリーの GEI、(I) ギャラリーの GEI 生成ボタン、(J) ギャラリーの観測方向、(K) マスク領域 (黒い領域がマスク領域に対応)、(L) マスク領域編集ダイアログ (図 10 (b)) 表示ボタン、(M) GEI のサイズ、(N) 鑑定実行ボタン、(O) 事後確率、(P) 方向変換及びマスク処理されたプローブの GEI、(Q) 方向変換及びマスク処理されたギャラリーの GEI、(R) プローブとギャラリーの GEI の差異 (差異がある領域をカラーで表示)、(S) プローブとギャラリーの状況に関する特徴間距離の確率密度関数。

3.3.1 登録モジュール

このモジュールでは対象のシステムへの登録を行う。ユーザはまず対象が映った画像列をシステムに入力し、その画像列が図 4(b) に示すようにダイアログに表示される。次にユーザは対象が真っ直ぐ安定的に歩いている区間 (定常歩行区間) を指定し、さらに対象を囲む矩形領域を各フレームに対して指定していく。ここで、定常歩行区間の指定においては、ユーザはその区間の最初と最後のフレームを指定する。また矩形領域の設定においては、まず区間の

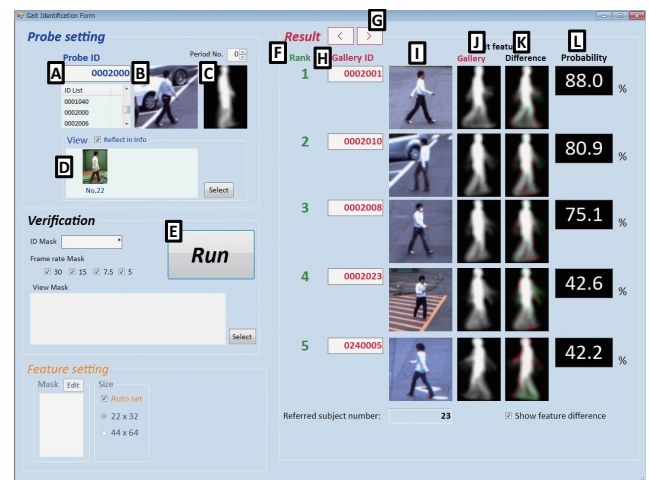


図 6 一括鑑定モジュールのインターフェースダイアログ。(A) 選択されたプローブの ID、(B) プローブの切り抜き動画、(C) プローブの GEI、(D) プローブの観測方向、(E) 鑑定実行ボタン、(F) 事後確率によるランク、(G) ランクリストの表示切り替えボタン (ランク 1 位から 5 位までの表示と 6 位から 10 位までの表示の切り替え)、(H) ギャラリーの ID、(I) ギャラリーの切り抜き動画、(J) ギャラリーの GEI、(K) プローブとギャラリーの GEI の差異、(L) 事後確率。

最初と最後のフレームにおいて、手動 (マウスによる矩形描画) で設定し、残りの区間の矩形領域は自動補間機能により、自動的に設定する (自動補間機能はダイアログのコンテキストメニューから選択する)。

続いてユーザは、その対象をシステムのデータベースに登録する。登録の際には、その対象の ID、種別 (プローブ、ギャラリー、背景^{*4}のいずれか)、フレームレート、及び観測方向を入力した上で登録する。これらの情報の入力登録モジュールのファイルメニューから開かれる登録情報入力ダイアログ (図 7) 上で行う。ここで、フレームレートは図 7(C) に示すコンボボックス内の選択肢から選択し、また観測方向は図 8 に示す観測方向選択ダイアログ (図 7(D))

*4 ここでいう背景とは主に移動物体を含まない画像列のことであり、シルエット生成の際に用いられる。

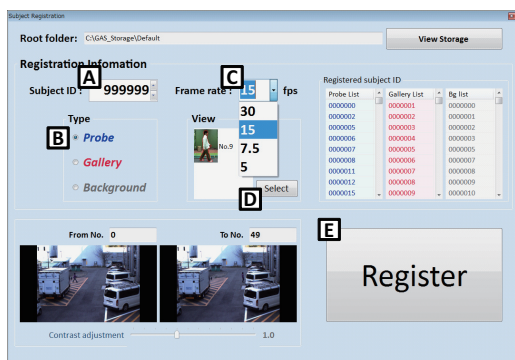


図 7 登録対象の情報入力ダイアログ. (A) 対象 ID, (B) 対象の種類, (C) フレームレート, (D) 観測方向選択ダイアログ (図 8) を開くボタン, (E) 登録実行ボタン.

のボタンにより開かれる) を介して指定する*5.

3.3.2 シルエット生成モジュール

このモジュールではシステムに登録された対象のシルエットを生成する. ユーザはまずデータベースから対象を選択 (ID リストから ID を選択) し, その対象が画面上に表示される. 表示画面上において, 対象のシルエットは図 4(E) に示すように, 輪郭線 (赤線) の形で原画像にオーバーレイ表示される (初期状態においてはシルエットは生成されていないため輪郭線は表示されない).

ユーザは各フレームに対し, 対象のシルエットを, マウス操作に基づく手動ペイント機能, 及び背景差分やグラフカットに基づく領域分割 [13] といった自動機能*6を使ってシルエットを生成していく. これらの機能の処理結果を図 9(a)~(c) に示す. また, 他にも面積フィルタリング処理, 領域膨張・収縮処理といった自動機能も使用することができる. こういった自動機能は全て画面のコンテキストメニューから選択することができ, また各自動処理におけるパラメータ (閾値など) はユーザが調整することができる.

シルエット生成の典型的な手順の一例は次の通りとなる. まずユーザは背景差分処理を行い, その差分結果を初期値としてグラフカットに基づく領域分割を行う. そして最後にユーザは各フレームにおけるシルエットの品質を目視で確認し, 手動ペイント機能により適宜シルエットを修正する. この手順により生成されたシルエットの例を図 9(d) に示す.

全フレームに対するシルエットを生成後, ユーザはそのシルエットをシステムのデータベースへ保存する.

3.3.3 個別鑑定モジュール

このモジュールではシステムに登録されたプローブと

*5 技術的には, シーンに対してカメラが正しく校正されていれば観測方向を自動的に推定することは可能であるが, そのような校正には専門的な知識が必要となるため, 本システムが想定するユーザにはほとんど不可能であるといえる. この点に関して, 5 章にてより詳細に論ずる.

*6 ただし, これらの機能を使うためには, そのシーンに対応する背景画像を事前に登録しておく必要がある.

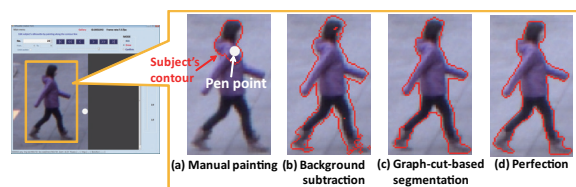
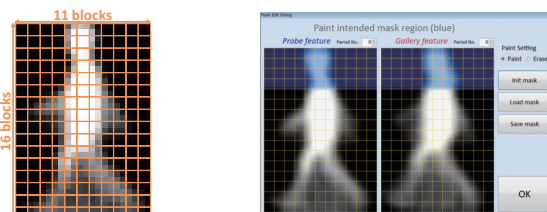


図 9 シルエット生成ダイアログにおける処理結果の例. (a) 手動ペイント処理の結果, (b) 背景差分処理の結果, (c) グラフカットに基づく前景領域分割の結果, (d) 完成したシルエット.



(a) ブロック単位のマスク (b) マスク領域編集ダイアログ

図 10 ブロック単位のマスクの定義とマスク領域編集ダイアログ. マスク領域編集ダイアログにおいて, 鑑定される GEI の組がダイアログ上に表示され (左側: プローブ, 右側: ギャラリー半透明の青色ブロックがユーザによりマスク領域として指定されたブロックを表している).

ギャラリー間の個別鑑定を実行する. ユーザはまず鑑定を行う一組のプローブとギャラリーの組をデータベースから選択 (リストから ID を指定) する. 次に, 図 5(D), 及び (I) に示すボタンを押し, 選択されたプローブとギャラリーの GEI を生成する. この時, 生成される GEI のサイズはプローブとギャラリーのシルエットの高さの小さい方を基準とし, そのシルエットの高さよりも小さくなるよう自動的に決定される (例えば, 小さい方のシルエットの高さが 50 [pixels] であれば, 生成される GEI のサイズは 22×32 [pixels] となる). なお, この時決定された GEI のサイズは図 5(M) に示すラジオボックスで確認できる.

ユーザは続いて, 比較する GEI の外れ値 (例えば一方の対象がヘルメットやフードを被っている, バックパックを背負っている, といった明らかな服装や荷物の状況の差異) を確認し, そういった外れ値を含む領域に対するマスク領域を手動で設定する. このマスク領域の指定は, 図 10(b) に示すマスク領域編集ダイアログ (図 5 (L) に示すボタンにより開かれる) を介して行う. ユーザはこのダイアログ上で, 所望のブロックをマウスでポインティングすることでマスク領域を指定する.

最後に, ユーザは図 5(N) に示すボタンを押し, 鑑定を実行する. そして, 鑑定結果となる事後確率がダイアログ上に表示される (図 5(O)).

3.3.4 一括鑑定モジュール

このモジュールではシステムに登録されたプローブと複数のギャラリー間の一括鑑定を実行する. ユーザはまず鑑定を行うプローブを選択し, 次に図 6 (E) に示すボタンを押して, 選択されたプローブとシステムに登録され全ギャラ



図 8 観測方向選択ダイアログ. ユーザはこのダイアログに表示されている各観測方向の画像を見て、入力した観測方向に最も近いものを選択する。

り間での鑑別を実行する。そして、鑑別の結果は、事後確率の高い順に、上位 10 位までのギャラリーとその事後確率などが、事後確率に関して降順に並べられたリスト (図 6 (L)) としてダイアログ右側に表示される。

3.4 実装

本システムは、VisualStudio2008 開発環境の下、C#, C++/CLI, 及び C++ の混合プラットフォーム上で開発されている。また、画像処理には OpenCV (version 2.3) ライブラリを用いている。

4. 実験

構築したシステムを用い、11 人の被験者による模擬鑑別実験を行った。11 人の被験者のうち、一人は歩容の専門家 (かつシステムの開発者)、残りの 10 人は非専門家 (研究室の学生や技術スタッフ) であり、本実験の目的は、システムに関する簡単な講習を受け、システムの説明書を読んだ非専門家が、専門家と同様の鑑別結果を、妥当な操作時間内で得ることができるかを確認することであり、本実験では個別鑑別のシナリオを想定してそれを評価した*7。

4.1 システムの講習

模擬実験の前に、非専門家は、歩容に基づく人物鑑別の概略と、システムの説明書に従ったシステムの使い方 (鑑別手順) に関する約 90 分の講習を受けた。次に、非専門家は、練習用のデータセットを用い、専門家から助言を受けながら、3~5 時間に渡る鑑別の練習を行った。

4.2 模擬鑑別に用いたデータセット

模擬鑑別実験には五つのデータセット (プローブとギャ

ラリーの画像列の集合) を用いた。各データセットのサンプル画像を表 4 に、また詳細を表 3 に示す。なお、各データセットにおいて、ギャラリー間での背丈や体形の差は小さく、これらは比較的難しい鑑別データセットになっていることに注意されたい。

4.3 模擬鑑別結果

s s

4.3.1 事後確率

被験者毎の模擬鑑別結果 (事後確率) を表 5 に示す。この表において、被験者 #1 が歩容の専門家であり、この結果では、全てのデータセットに対して、同一人物の組に対しては事後確率が極めて高く (89.9%以上)、別人の組に対しては低くなっており (35.4%以下)、妥当な鑑別結果を得ていることがわかる

次に、非専門家の被験者の結果に目を向け、まず、プローブとギャラリー間での服装の差異を含むデータセット A, B の結果に着目する。データセット A に関しては、プローブ a はフードを被っている一方、ギャラリー b と c は被っていない状態であるため、該当領域をマスクして鑑別を行う必要がある。このデータセットに対しては、全ての被験者が該当領域をマスク領域として指定することができ、結果として、表 5 に示すように、全ての非専門家が専門家と同様の結果を得ることができた。

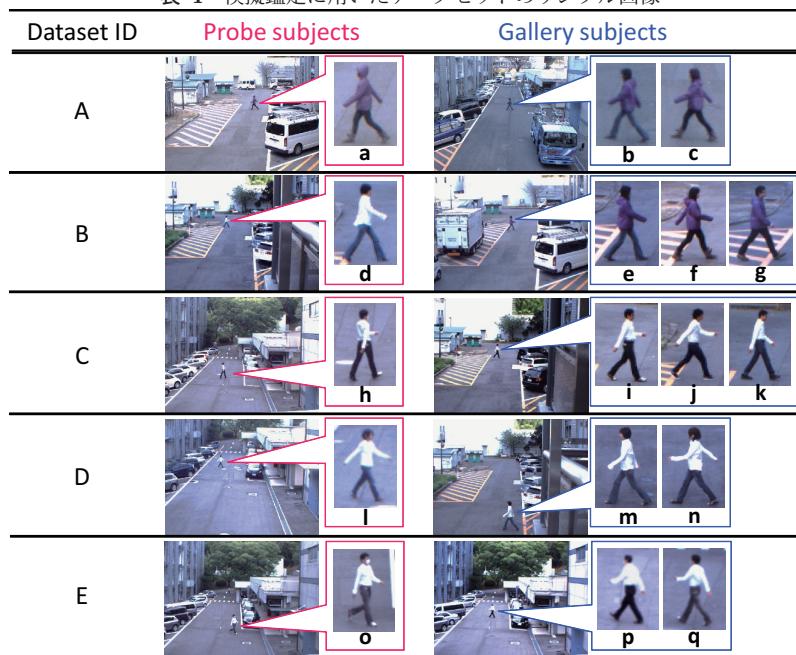
データセット B に関しては、プローブ d が薄手のブラウスを着ているのに対し、ギャラリー e, f, g は厚手のフード付きジャケットを着ているため、上半身領域 (腰から首にかけての領域) をマスクして鑑別を行う必要がある。このデータセットに対しては、10 人の非専門家のうち 8 人は該当領域をマスクできたが、二人 (#6 と #9) はフード領域にのみマスク領域を設定した。結果として、表 5 に示すように、その二人が出力した事後確率は、他の被験者に比べ少し低い値となっている。

*7 ユーザ操作という観点では、一括鑑別における操作は個別鑑別とほぼ同じであるため、本実験では個別鑑別のシナリオのみで評価を行った。

表 3 模擬鑑定に用いたデータセットの詳細

データセット ID	対象 ID	対象の高さ (概算) [pixel]	歩行区間の画像枚数 (概算)	フレームレート [fps]	観測方向
A	a	60	70	15	左側方
	b	50	65	15	左側方
	c	50	70	15	左側方
B	d	80	40	7.5	右側方
	e	95	30	7.5	右側方
	f	95	35	7.5	右側方
	g	95	30	7.5	右側方
C	h	65-75	90	15	右斜め後方
	i	90	65	15	右側方
	j	90	60	15	右側方
	k	90	80	15	右側方
D	l	50-55	60	15	右斜め前方
	m	120	50	15	左側方
	n	120	50	15	左側方
E	o	85-98	55	7.5	右斜め前方
	p	55-60	55	15	左斜め後方
	q	55-60	60	15	左斜め後方

表 4 模擬鑑定に用いたデータセットのサンプル画像



次に、プローブとギャラリー間で観測方向が異なるデータセット C, D, E の結果に着目する。表 5 において、データセット C, E に関しては全ての非専門家が専門家と同様の結果を得ていることがわかる。実際、これらのデータセットに対しては、全ての非専門家が全ての対象の観測方向を正しく設定できた。

一方、データセット D に関しては、二人の非専門家 (#6 と #9) が観測方向、特に俯角の種類を選択を誤った*8。このように観測方向を誤ると、誤った方向変換モデルが鑑定

対象に適用されるため、特に方向変換という観点から、鑑定の妥当性は大きく損なわれる。結果として、この二人から得られた事後確率は、他の非専門家の結果よりも低い値になっていることがわかる。

全体として、被験者 #6 と #9 の結果 (前記 4 組の鑑定結果) は、マスク領域と観測方向の選択に関して、事前の練習がもう少し必要であることを示しているものの、非専門家が行ったほぼ全ての鑑定の組 (50 組の鑑定に対し 46 組) において、専門家と同様の鑑定結果を得ることができた。

4.3.2 操作時間

今回各被験者が一組の鑑定に要した合計時間 (プローブ

*8 データセット D に関しては正しい観測方向は図 8 に示す camera 15 (俯角無し) であるが、被験者 #6 と #9 の選んだものは camera 11 (俯角有り) であった。

表 5 被験者毎の模擬鑑定結果（事後確率）[%]

データセット ID	対象 ID			被験者 ID										
	プローブ	ギャラリ	真値	1	2	3	4	5	6	7	8	9	10	11
A	a	b	別人	0.4	20.5	4.3	0.2	4.2	0.5	0.9	0.7	2.2	0.9	1.6
		c	同一人物	95.2	90.2	96.7	93.4	80.4	83.4	97.3	97.6	89.4	94.5	99.5
		e	同一人物	96.0	96.7	95.1	97.7	92.7	76.7	96.5	88.4	87.0	96.7	93.0
B	d	f	別人	11.0	30.0	10.6	18.0	4.0	0.3	23.7	17.9	1.7	0.2	3.9
		g	別人	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		i	同一人物	89.9	89.4	89.1	77.6	90.5	90.5	83.9	83.6	86.5	90.0	88.6
C	h	j	別人	3.8	7.7	15.3	3.9	7.7	4.9	3.8	5.7	4.4	4.8	3.3
		k	別人	7.9	8.8	13.4	8.0	11.7	14.0	11.6	7.5	11.2	8.4	12.3
		l	別人	35.4	37.0	54.9	29.0	30.4	22.3	36.2	35.5	25.9	36.7	35.7
D	l	m	別人	35.4	37.0	54.9	29.0	30.4	22.3	36.2	35.5	25.9	36.7	35.7
		n	同一人物	91.5	81.0	89.1	73.5	85.0	46.0	73.3	74.4	65.9	72.9	85.6
		o	別人	32.5	33.5	29.1	39.5	27.1	31.4	34.6	30.3	33.4	30.5	30.5
E	o	p	別人	32.5	33.5	29.1	39.5	27.1	31.4	34.6	30.3	33.4	30.5	30.5
		q	同一人物	93.3	88.1	89.8	86.6	88.3	80.4	87.3	91.1	74.5	87.5	88.9

とギャラリそれぞれに対し登録モジュールを2回、シルエット生成モジュールを2回、及び個別鑑定モジュールを1回使用)は最も短い場合で24分、最も長い場合でも100分であり、これらの結果は各被験者が妥当な時間内で鑑定結果を得ることができたことを示しているといえる。

5. 考察

5.1 システム操作におけるユーザ間のばらつき

前章では、模擬鑑定実験の結果において、ほとんど全ての鑑定の組に対して、専門家と非専門家が同様の結果を得たことを示した。しかし一方、妥当な鑑定を行えなかった結果を含め、各被験者間で事後確率の出力値にばらつきが存在していることも事実である。このようなばらつきは主に、定常歩行区間、観測方向、及びマスク領域といったユーザの判断を伴う設定部分の差異に起因している。従って、そういったばらつきを低減するため、ユーザ判断を支援するような半自動(対話的)機能を導入することは有意義であると考えられる。

まず一つの対策として、歩行軌跡から定常歩行区間と観測方向を自動的に推定することが考えられる。例えば、シーンに対して校正済みのカメラとシーンの歩行平面の間の平面拘束を利用すれば、対象の歩行平面上の足下点の系列を歩行軌跡として抽出するのは比較的容易である[14],[15]。ただし、カメラを校正するためには、校正対象を用いた(カメラが設置されている)現地での校正作業が不可欠であり、それは簡単な作業では無い。非専門家にはほとんど不可能な作業である。本システムがそういった歩行軌跡の推定機能を含んでいないのは、上記の理由からである。

別のカメラ校正方法として、シーンの平行線(建物、道路、回廊など)から推定される消失点に基づいてカメラ校正を行う方法も提案されている[16]。この方法ではスケールに関しては不定性が残るものの、スケールは定常歩行区

間や観測方向を推定するのに不要であるため、この手法は有効であるといえる。このような方法によって自動的に推定された定常歩行区間や観測方向が“候補”としてユーザに提示され、ユーザは必要に応じてそれを修正する、という対話的な操作の枠組みを導入すれば、歩行区間と観測方向の設定におけるユーザ間のばらつきは減少するものと考えられる。

マスク領域の設定に関しては、歩容認証においてマスクすべき領域を自動的に推定するような手法が数多く提案されてきている(例えば、服装変化に頑強な歩容特徴[3]、服装変化を考慮した部分領域に基づく歩容認証[12]、バックパック領域の検出[17]、手荷物の検出[18])。こういった手法を応用し、自動的に推定されたマスク領域が候補領域としてユーザに提示され、ユーザは必要に応じてそれを手動で修正するという対話的なマスク領域設定の枠組みを導入すれば、マスク領域の設定におけるユーザ間のばらつきも減少するものと考えられる。

このようにして、自動処理を取り入れた対話的な機能がユーザ間の判断のばらつきを抑制し、結果として事後確率の出力値のばらつきも低減することが期待できる。

5.2 方向変換モデルの性能改善

表5における、観測方向が異なる組の鑑定結果(データセットC,D,E)が示すように、本システムに実装されている方向変換モデルは基本的には良好に動作しているものの、観測方向が同じ組の鑑定結果(データセットA,B)に比べ、同一人物鑑定時の事後確率と別人鑑定時の事後確率の差が小さくなる傾向にある、即ち、本人と他人の識別性能が低下している。一方、方向変換を介した歩容認証では、対象の複数の観測方向の特徴を追加することで、認証性能が向上することが報告されている[2],[19]。従って、異なる観測方向の組に対する鑑定性能は、対象の複数の観測方向が映った映像をさらに収集して鑑定に用いることで改善

するものだと考えられる。

また、異なる観測方向の組に対する鑑定性能の低下は、システムで定義した24観測方向(図8)と実際の対象の観測方向が厳密には一致しない場合にも引き起こされる。この問題に対しては、任意視点方向変換モデル[20]の導入が有効であると考えられる。

6. おわりに

本研究では、世界初となる犯罪捜査支援のための歩容に基づく人物鑑定システムを構築した。本システムは、歩容の非専門家である犯罪捜査員が、システムの使い方を覚えるだけで、専門家と同様の人物鑑定の結果を得ることができるように設計されている。また本システムは、鑑定される対象同士が同一人物であることを示す事後確率を、対象のサイズ、フレームレート、観測方向、服装といった状況を考慮した上で計算し、鑑定結果として出力する。実験では、本システムを使用した模擬鑑定実験を行い、ほぼ全ての鑑定データセットに対し、歩容の非専門家である10人の被験者が、歩容の専門家と同様の結果を、妥当な時間内に得ることができることを確認した。

謝辞

本研究はJSPS基盤研究(S)21220003及び文部科学省の科学技術戦略推進費による「安全・安心な社会のための犯罪・テロ対策技術等を実用化するプログラム」の助成を受けたものである。

参考文献

- [1] Han, J. and Bhanu, B.: Individual Recognition Using Gait Energy Image, *Trans. on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 2, pp. 316–322 (2006).
- [2] Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T. and Yagi, Y.: Gait Recognition Using a View Transformation Model in the Frequency Domain, *Proc. of the 9th European Conf. on Computer Vision*, Graz, Austria, pp. 151–163 (2006).
- [3] Bashir, K., Xiang, T. and Gong, S.: Gait recognition without subject cooperation, *Pattern Recognition Letters*, Vol. 31, No. 13, pp. 2052–2060 (2010).
- [4] Lam, T. H. W., Cheung, K. H. and Liu, J. N. K.: Gait flow image: A silhouette-based gait representation for human identification, *Pattern Recognition*, Vol. 44, pp. 973–987 (2011).
- [5] Wang, C., Zhang, J., Wang, L., Pu, J. and Yuan, X.: Human Identification Using Temporal Information Preserving Gait Template, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 11, pp. 2164–2176 (2012).
- [6] How biometrics could change security, BBC (online), available from http://news.bbc.co.uk/2/hi/programmes/click_online/7702065.stm (accessed 2013-01-25).
- [7] Nixon, M. S. and Carter, J. N.: Automatic recognition by gait, *Proceedings of the IEEE*, Vol. 94, No. 11, pp. 2013–2024 (2006).
- [8] The OU-ISIR Gait Database, Large Population Dataset, Osaka University (online), available from <http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitLP.html> (accessed 2013-02-01).
- [9] Iwama, H., Okumura, M., Makihara, Y. and Yagi, Y.: The OU-ISIR Gait Database Comprising the Large Population Dataset and Performance Evaluation of Gait Recognition, *IEEE Transactions on Information Forensics and Security*, Vol. 7, No. 5, pp. 1511–1521 (2012).
- [10] Utsumi, A. and Tetsutani, N.: Adaptation of appearance model for human tracking using geometrical pixel value distributions, *Proc. of the 6th Asian Conf. on Computer Vision*, Vol. 2, pp. 794–799 (2004).
- [11] Yu, S., Tan, D. and Tan, T.: Modelling the Effect of View Angle Variation on Appearance-Based Gait Recognition, *Proc. of 7th Asian Conf. on Computer Vision*, Vol. 1, pp. 807–816 (2006).
- [12] Hossain, M. A., Makihara, Y., Wang, J. and Yagi, Y.: Clothing-Invariant Gait Identification using Part-based Clothing Categorization and Adaptive Weight Control, *Pattern Recognition*, Vol. 43, No. 6, pp. 2281–2291 (2010).
- [13] Boykov, Y. and Jolly, M.: Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images, *Proc. of Int. Conf. on Computer Vision*, Vol. 1, pp. 105–112 (2001).
- [14] Hild, M.: Estimation of 3D Motion Trajectory and Velocity from Monocular Image Sequences in the Context of Human Gait Recognition, *Proc. of the 17th Int. Conf. on Pattern Recognition*, Vol. 1, pp. 231–235 (2004).
- [15] Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T. and Yagi, Y.: Adaptation to Walking Direction Changes for Gait Identification, *Proc. of the 18th Int. Conf. on Pattern Recognition*, Vol. 2, Hong Kong, China, pp. 96–99 (2006).
- [16] Bazin, J.-C., Seo, Y., Demonceaux, C., Vasseur, P., Ikeuchi, K., Kweon, I. and Pollefeys, M.: Globally Optimal Line Clustering and Vanishing Point Estimation in Manhattan World, *Proc. of the 25th IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 638–645 (2012).
- [17] Decann, B. and Ross, A.: Gait curves for human recognition, backpack detection, and silhouette correction in a nighttime environment, *Proc. of the SPIE, Biometric Technology for Human Identification VII*, Vol. 7667, pp. 76670Q–76670Q–13 (2010).
- [18] Damen, D. and Hogg, D.: Detecting Carried Objects from Sequences of Walking Pedestrians, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 6, pp. 1056–1067 (2012).
- [19] Kusakunniran, W., Wu, Q., Zhang, J. and Li, H.: Support Vector Regression for Multi-View Gait Recognition based on Local Motion Feature Selection, *Proc. of IEEE computer society conferene on Computer Vision and Pattern Recognition 2010*, San Francisco, CA, USA, pp. 1–8 (2010).
- [20] Muramatsu, D., Shiraishi, A., Makihara, Y. and Yagi, Y.: Arbitrary View Transformation Model for Gait Person Authentication, *Proc. of the 5th IEEE Int. Conf. on Biometrics: Theory, Applications and Systems*, No. Paper ID 39, Washington D.C., USA, pp. 1–6 (2012).