

# 周波数領域における方向変換モデルを用いた歩容認証

榎原 靖<sup>†</sup> 佐川 立昌<sup>†</sup> 向川 康博<sup>†</sup>  
越後 富夫<sup>††</sup> 八木 康史<sup>†</sup>

近年、カメラから離れた場所での個人認証手法として、歩容認証が注目を集めている。歩容認証の問題点として、歩行方向やカメラの視線方向の変化にともなう見えの変化があげられる。本研究では、周波数領域特徴と方向変換モデルを用いた歩容認証手法を提案する。最初に時系列の歩容シルエットで構成される時空間ボリュームから、フーリエ解析によって周波数領域特徴を抽出する。次に、複数被験者の複数方向の特徴を用いて、方向変換モデルを学習する。認証時には、方向変換モデルによって、登録人物の特徴を入力特徴と同じ方向の特徴に変換して、照合を行う。最後に、15度刻みの24方向の歩行に対する個人認証実験を行い、本手法の有効性を確認した。

## Gait Identification Using a View Transformation Model in the Frequency Domain

YASUSHI MAKIHARA,<sup>†</sup> RYUSUKE SAGAWA,<sup>†</sup> YASUHIRO MUKAIGAWA,<sup>†</sup>  
TOMIO ECHIGO<sup>†</sup> and YASUSHI YAGI<sup>†</sup>

Gait identification have recently gained attention as methods of individual identification at a distance from a camera. However, appearance changes due to view direction changes cause difficulties for gait identification. We propose a method of gait identification using frequency-domain features and a view transformation model. We first extract frequency-domain features from a spatio-temporal gait silhouette volume. Next, our view transformation model is obtained with a training set of multiple persons from multiple view directions. In an identification phase, the model transforms gallery features into the same view direction as that of a probe feature. Experiments involving gait identification from 24 view directions demonstrate the effectiveness of our method.

### 1. はじめに

近年、監視システムやアクセスコントロールのための個人認証技術の必要性が高まっており、人間の生体情報に基づくバイオメトリクスによる認証技術が注目されている。バイオメトリクスには、声紋、指紋、手の平の静脈、虹彩、顔画像、歩容等を利用したものがある。この中で、歩容認証はカメラから離れた場所での認証が可能であるという点で、広域監視への応用が期待されている。

現在の歩容認証手法は、主にモデルベース<sup>2),7),18),19),21),22)</sup> とアピアランスベースの手

法<sup>3),6),8)-10),12),15),16)</sup>に分かれる。

モデルベースの方法では入力画像にモデルをあてはめることで、人の体形や動き等の特徴を抽出する手法である。Urtasun ら<sup>19)</sup>は解剖学的な人体モデルを用いて関節角度の抽出を行い、Spencer ら<sup>18)</sup>は抽出された関節点の周期的な位置情報をもとに、方向変化に対して不変な関節角度の抽出を行った。これらのモデルベースの手法は、特に低解像度の画像に対しては、モデルあてはめや特徴抽出に失敗する場合がある。

アピアランスベースの手法は、画像を直接解析して人物モデルを用いずに特徴を抽出する手法である。Sarkar ら<sup>16)</sup>は歩容認証手法のための基準手法として、シルエット画像列どうしの直接照合手法を提案した。Sagawa ら<sup>15)</sup>は歩容画像列を時間方向に積み重ねることで時空間ボリューム( $x-y-t$  volume)を作成し、ボリュームから抽出した特徴により照合を行った。これらのアピアランスベースの手法における問題の1つと

<sup>†</sup> 大阪大学産業科学研究所  
The Institute of Scientific and Industrial Research,  
Osaka University

<sup>††</sup> 大阪大学電気通信大学情報工学科  
Department of Engineering Informatics, Osaka Electro-  
Communication University

して、歩行方向や観測方向の変化による見えの変化があげられる．実際、BenAbdelkader<sup>1)</sup>やYuら<sup>23)</sup>は、歩容認証における入力データと参照データの観測方向の違いによる影響を解析しており、その違いによって認証性能が低下することを示している．

方向変化への対応手法としては、パラメトリック固有空間法<sup>24)</sup>によって方向および個人を認識する方法や、視体積交差法によって仮想的な方向の歩容画像を合成する手法<sup>17)</sup>が考えられる．しかし、そのためには認証対象となる全被験者の多数方向からの同期画像が必要となる．また、1方向からの特徴のみを利用する方法としては、歩容を矢状面内運動と仮定して、透視投影によって方向を変換する手法<sup>5)</sup>や、直線歩行中のカメラ視線変化を利用する手法<sup>4)</sup>が提案されている．しかし、これらの手法では、入力方向と登録方向の差が大きくなると仮定が成り立たなくなる．

一方、歩容認証以外の研究分野では、方向変化への対応手法が数多く提案されている．Mukaigawaら<sup>11)</sup>は顔の姿勢や表情変化に対する画像生成を行い、Utsumiら<sup>20)</sup>は人物の様々な姿勢や方向に対する画像変換を行う方向変換モデル (View Transformation Model, 以下 VTM と呼ぶ) を提案している．

そこで、本研究では、図 1 に示すように、最初に学習用の被験者に対する全方向の歩容特徴を取得して、それをもとにあらかじめ VTM を学習する．また、認識対象の被験者については、限定された少数の参照方向のみの歩容特徴が得られているものとする．そして、認識時には任意方向の歩容特徴のプロブ (入力) に対して、参照方向の歩容特徴から VTM によってプロブと同じ方向に変換した歩容特徴をギャラリー (辞書) として照合を行うことで、方向変化に対応する．

しかし、Utsumi らの手法<sup>20)</sup>は静止画像間の変換手法であり、歩容認証のように動画像を対象とする場合には、方向変換や照合を行う前に、同一姿勢どうしの静止画を対応させるための処理、すなわち、動画像

どうしを時間軸上でシフトさせて同期がとれるシフト量を求める処理が必要であり、そのフレーム同期のための計算量が多くなるという問題点がある．そこで、本研究では最初に、歩容シルエットを時間軸上に積み重ねることで作成される歩容シルエットボリューム (Gait Silhouette Volume, 以下 GSV と呼ぶ) から、時間非依存の周波数領域特徴を抽出し、その特徴に対して方向変換や照合を行うことで、フレーム同期処理を省略して、問題を単純化する．

本論文の構成を以下に示す．まず、2章で GSV の作成方法について述べ、3章で GSV の照合方法について述べる．次に4章で VTM による方向変化への対応方法を述べる．5章では、様々な方向に対する歩容認証実験について述べる．最後に6章で結論と今後の課題について述べる．

## 2. GSV の作成

### 2.1 歩容シルエットの抽出

最初に、背景差分を用いて歩容シルエットを抽出する．その際、通常のカラークメラによる背景差分においては、図 2 (b) に示すように影や照明変動によって抽出に失敗することがある．そこで、代わりに赤外線カメラ (NEC TH1702MX) を用いることで、温度画像による背景差分を行う．本赤外線カメラの画像サイズは 320 pixel × 240 pixel で、フレームレートは 30 fps である．図 2 (c) は赤外線カメラによって撮影された入力画像であり、背景と比較して人物領域の温度が高

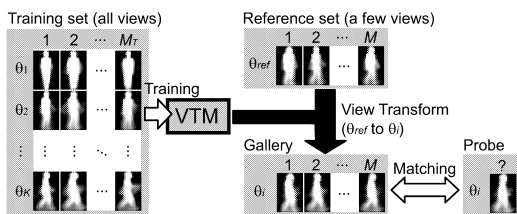


図 1 方向変換の問題設定  
Fig. 1 Problem setting of view transformation.

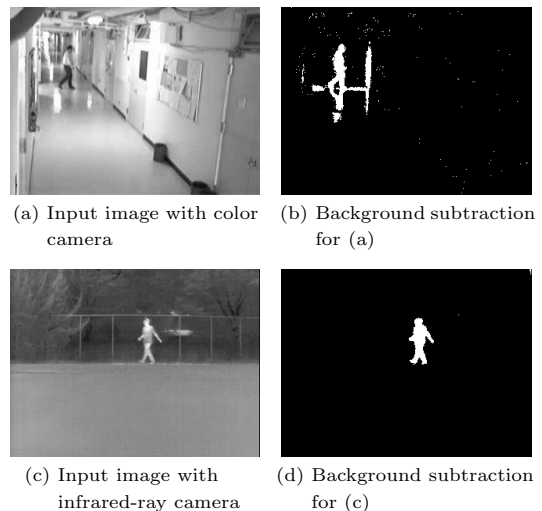


図 2 カラーカメラと赤外線カメラによる背景差分の比較 ((c) において、明るさが温度の高さに対応)  
Fig. 2 Comparison of background subtraction by color camera and that by infrared-ray camera (In (c), brighter value indicates higher temperature).

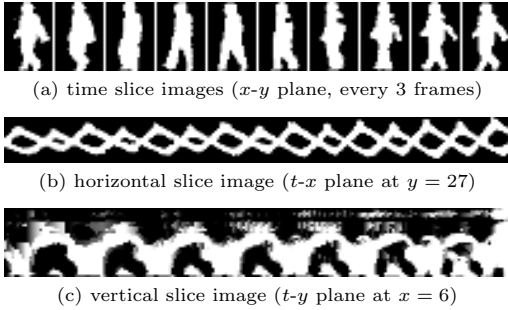


図 3 GSV の例

Fig. 3 An example of GSV.

くなっていることが分かる．よって，図 2 (d) のような鮮明な歩容シルエットが得られる．また，ここでは簡単化のために，画像中に存在する人物は 1 人であるとして，最大面積を持つ領域を人物領域として保持するものとする．

**2.2 歩容シルエットのスケールと位置合わせ**  
次に，抽出された歩容シルエットのスケールと位置合わせを行う．まず，フレームごとに領域の上端・下端・水平方向の中心を取得する．ここで，水平方向の中心は，領域に属する画素の水平位置の中央値として定義する．次に，それらに対して 30 フレームの移動平均フィルタをかけ，平滑化された高さが 30 pixel になるように，アスペクト比を保ったままスケールする．そして，平滑化された水平方向の中心が画像中心になるように位置合わせして，画像サイズ 20 pixel × 30 pixel の歩容シルエット画像列を作成する．最後に，それらを時間方向に重ね合わせることで，GSV を作成する．

作成された GSV の時間断面 ( $x$ - $y$  断面)，水平断面 ( $t$ - $x$  断面)，垂直断面 ( $t$ - $y$  断面) 画像を図 3 に示す．図 3 (b)，(c) より歩容の周期性の存在を容易に確認することができる．

### 3. GSV の照合

#### 3.1 歩行周期の検出

照合の第 1 段階として，歩行周期の検出を行う．歩行周期  $N_{gait}$  は GSV の時間軸方向の正規化自己相関を計算することによって，以下のように算出する．

$$N_{gait} = \arg \max_{N \in [N_{min}, N_{max}]} C(N) \quad (1)$$

$$C(N) = \frac{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n)g(x,y,n+N)}{\sqrt{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n)^2} \sqrt{\sum_{x,y} \sum_{n=0}^{T(N)} g(x,y,n+N)^2}} \quad (2)$$

$$T(N) = N_{total} - N - 1 \quad (3)$$

ここで  $C(N)$  は  $N$  フレーム分シフトしたときの GSV の正規化自己相関であり， $g(x,y,n)$  は  $n$  フレーム目の空間位置  $(x,y)$  における GSV の値で，歩容シルエットのスケール時の線形補間によって  $[0,1]$  の範囲内の中間階調値をとるものとする．また， $N_{total}$  は GSV に含まれる総フレーム数である．今回は，走行・早歩き・牛歩といったような通常の歩行から外れるものについては取り扱わないので，歩行周期の範囲は通常の歩行における個人差を考慮して実験的に  $N_{min} = 20$ ， $N_{max} = 40$  とする．

#### 3.2 周波数領域特徴の抽出

1 章でも述べたように，照合や方向変換の際にフレーム同期の問題を避けるために，歩行周期に基づく周波数領域特徴を歩容特徴として利用する．最初に全体シーケンス  $S$  から歩行周期  $N_{gait}$  フレームごとにサブシーケンス  $\{S_i\}$  ( $i = 1, 2, \dots, N_{sub}$ ) を取り出す．ここで， $i$  番目のサブシーケンス  $S_i$  のフレームの範囲は  $[iN_{gait}, (i+1)N_{gait} - 1]$  となる．次に，各サブシーケンスに対して，時間軸方向の離散フーリエ変換 (DFT) を計算して，さらに歩行周期  $N_{gait}$  で正規化された振幅スペクトルを以下のように計算する．

$$G_i(x,y,k) = \sum_{n=iN_{gait}}^{(i+1)N_{gait}-1} g(x,y,n)e^{-j\omega_0kn} \quad (4)$$

$$A_i(x,y,k) = \frac{1}{N_{gait}} |G_i(x,y,k)|. \quad (5)$$

ここで  $\omega_0$  は歩行周期  $N_{gait}$  に対応する基本角周波数であり， $G_i(x,y,k)$  は DFT の  $k$  倍周波数に対応する成分であり， $A_i(x,y,k)$  は歩行周期  $N_{gait}$  で正規化された  $G_i(x,y,k)$  に対する振幅スペクトルである．

ここで，高周波成分 (実験的に  $k > 2$  とする) の DFT は，低周波成分の DFT と比較してノイズの割合が大きくなるため，歩容特徴から除く．よって，歩容特徴としては  $A_i(x,y,k)$  ( $k = 0, 1, 2$ ) が用いられ，その次元  $N_A$  は  $20 \times 30 \times 3 = 1800$  となる．

図 4 に様々な方向に対して抽出した歩容特徴を示す．方向  $\theta$  は図 5 のように光軸と歩行方向のなす角で定義され，本論文での単位は度とする．歩容特徴は被験者ごとに方向間で幅広く変化しており，また各方向に対して被験者間でもある程度の変化が存在している．さらに，方向間での歩容特徴の変化はすべての被験者に対して似た傾向を示していることが分かる．これより，方向変化による歩容特徴の変化は，個人差とは独立に VTM によって表現可能であるものと考えられる．

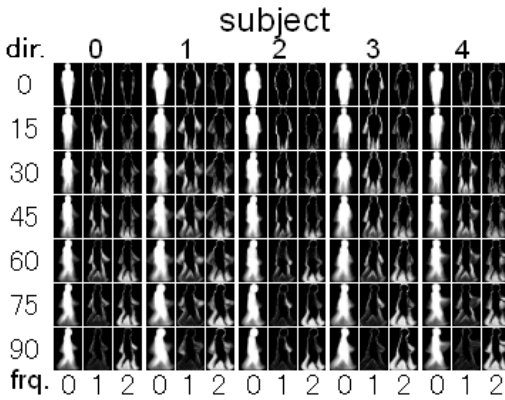


図4 各被験者の各方向（15度刻み）に対する歩容特徴

Fig. 4 Gait features for some subjects from each view (every 15 deg).

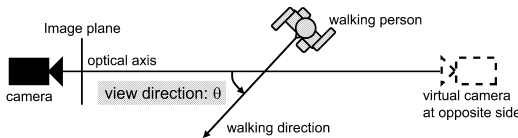


図5 歩行方向  $\theta$  の定義（真上図）

Fig. 5 Definition of view direction  $\theta$  at top view.

### 3.3 照合基準

最初に、ある2つのサブシーケンス  $S_i, S_j$  間の照合基準を定義する。ここで、サブシーケンス  $S_i$  の振幅スペクトル  $A_i(x, y, k)$  の各要素をベクトルとして並べた  $N_A$  次元の特徴ベクトルを  $a(S_i)$  として定義する。照合基準  $d(S_i, S_j)$  は、特徴ベクトル間のユークリッド距離として以下のように定義する。

$$d(S_i, S_j) = \|a(S_i) - a(S_j)\|. \quad (6)$$

次に、2つの全体シーケンス間の照合基準について述べる。ここで、プローブ（入力）とギャラリー（辞書）に対する全体シーケンスをそれぞれ  $S_P, S_G$  として定義して、そのサブシーケンスをそれぞれ  $\{S_{P_i}\}$  ( $i = 1, 2, \dots$ ),  $\{S_{G_j}\}$  ( $j = 1, 2, \dots$ ) として定義する。そして、外れ値やばらつきを考慮して、照合基準  $D(S_P, S_G)$  を、プローブの各サブシーケンス  $S_{P_i}$  に対して計算したギャラリーの全サブシーケンス  $\{S_{G_j}\}$  ( $j = 1, 2, \dots$ ) との最小距離の中央値として以下のように定義する。

$$D(S_P, S_G) = \text{Median}_i[\min_j\{d(S_{P_i}, S_{G_j})\}]. \quad (7)$$

## 4. 方向変換モデル

### 4.1 方向変換モデルの学習

ここでは Utsumi ら<sup>20)</sup> の方法に基づいて、VTM の定式化について簡単に触れる。ただし、Utsumi らが

VTM を静止画像に適用しているのに対して、我々は GSV から周波数領域特徴を抽出することで、VTM を周波数領域に拡張している点が異なることに注意されたい。

最初に方向  $\theta$  を  $K$  個の方向に離散化し、 $m$  番目の被験者の  $k$  番目の方向に対する  $N_A$  次元の歩容特徴ベクトルを  $a_{\theta_k}^m$  として定義する。ここで、学習セットとして  $M$  人の被験者の  $K$  方向の特徴ベクトルが得られたとすると、その特徴ベクトルを、各行が各方向を、各列が各被験者に対応するように並べた行列を作成して、それを特異値分解 (SVD) によって以下のように分解する。

$$\begin{bmatrix} a_{\theta_1}^1 & \cdots & a_{\theta_1}^M \\ \vdots & \ddots & \vdots \\ a_{\theta_K}^1 & \cdots & a_{\theta_K}^M \end{bmatrix} = USV^T = \begin{bmatrix} P_{\theta_1} \\ \vdots \\ P_{\theta_K} \end{bmatrix} \begin{bmatrix} v^1 & \cdots & v^M \end{bmatrix}. \quad (8)$$

ここで、 $U$  は  $KN_A \times M$  の直交行列、 $V$  は  $M \times M$  の直交行列、 $S$  は特異値を並べた  $M \times M$  の対角行列、 $P_{\theta_k}$  は行列  $US$  を構成する  $N_A \times M$  の部分行列、 $v^m$  は  $M$  次元の列ベクトルである。

ベクトル  $v^m$  は方向に対して独立な  $m$  番目の被験者に固有な特徴ベクトルである。また、部分行列  $P_{\theta_k}$  は被験者に固有なベクトル  $v$  を方向  $\theta_k$  の特徴に変換する行列であり、各被験者に対して共通、すなわち、被験者に対して独立である。よって、 $m$  番目の被験者の方向  $\theta_i$  の特徴ベクトル  $a_{\theta_i}^m$  は以下で表される。

$$a_{\theta_i}^m = P_{\theta_i} v^m. \quad (9)$$

これより、参照方向  $\theta_{ref}$  から別の方向  $\theta_i$  への特徴ベクトルの変換は、以下のように得られる。

$$\hat{a}_{\theta_i}^m = P_{\theta_i} P_{\theta_{ref}}^+ a_{\theta_{ref}}^m. \quad (10)$$

ここで、 $P_{\theta_{ref}}^+$  は  $P_{\theta_{ref}}$  の擬似逆行列である。ただし、歩容シルエットの画像面に対して直交する方向の動きは縮退するので、1方向のみを用いた変換では不十分となることがある。たとえば、図4の正面方向の特徴  $a_{\theta_0}^m$  から側面方向の特徴  $a_{\theta_{90}}^m$  を推定することは、我々人間にとっても困難である。よって、複数方向の特徴（その方向を  $\theta_{ref}(1), \dots, \theta_{ref}(k)$  とする）が得られた場合には、以下のようにして、方向  $\theta_i$  に対するより正確な特徴に変換することができる。

$$\hat{a}_{\theta_i}^m = P_{\theta_i} \begin{bmatrix} P_{\theta_{ref}(1)} \\ \vdots \\ P_{\theta_{ref}(k)} \end{bmatrix}^+ \begin{bmatrix} a_{\theta_{ref}(1)}^m \\ \vdots \\ a_{\theta_{ref}(k)}^m \end{bmatrix}. \quad (11)$$

また、これまでの定式化においては、特徴全体の間

の変換を行っているが、実際には 2.2 節で説明しているように高さによる正規化を行っているため、頭・手・膝等の体の部分はどの方向に対しても同じ垂直位置に現れることになる。よって、ある垂直位置  $y_i$  から別の垂直位置  $y_j$  ( $\neq y_i$ ) への変換はないものとして、方向変換モデルを垂直位置  $y_i$  ごとに別々に定義するものとする。

#### 4.2 幾何モデルに基づく参照方向の追加

前節で述べた方向変換モデルにおいて、学習セットにノイズが含まれる場合には、方向変換行列に誤差が伝播し、変換特徴にも誤差が生じることとなる。そこで、本節ではそのような変換誤差を軽減するための、幾何モデルに基づく参照方向の追加方法について述べる。

被験者がカメラから十分に離れて観測されて弱中心投影が仮定できる場合には、図 5 のように反対側にある仮想的なカメラで観測される方向  $\theta_{ref}$  のシルエット画像 (以下  $I_{opp}(\theta_{ref})$  とする) は、方向  $\theta_{ref}$  の元のシルエット画像 (以下  $I(\theta_{ref})$  とする) の鏡像 (左右反転画像) となる。さらに、 $I_{opp}(\theta_{ref})$  と  $I(\theta_{ref}+180)$  は明らかに同じ画像となるので、 $I(\theta_{ref}+180)$  は  $I(\theta_{ref})$  の鏡像となる (以下、pers. とする)。

また、弱中心投影に加えて、歩容の左右対称性が仮定できる場合には、歩容シルエット  $I(360 - \theta_{ref})$  は  $I(\theta_{ref})$  の左右反転画像になる (以下、sym. とする)。加えて、これら 2 つの仮定をあわせて考えると、 $I(\theta_{ref})$  と  $I(180 - \theta_{ref})$  が同じシルエット画像になる (以下、pers.+sym. とする)。

よって、参照方向  $\theta_{ref}$  の実在の特徴が得られると、方向  $(180 - \theta_{ref})$  に対して同じ特徴を、方向  $(\theta_{ref} + 180)$  と  $(360 - \theta_{ref})$  に対して左右反転特徴を仮想的に追加することができる。

ただし、これらの幾何モデルはあくまでも仮定に基づくものであるため、仮想的に追加する特徴と実在の特徴とは必ずしも一致しない。たとえば、歩容の左右対称性の仮定は、被験者によっては成り立たないことがあり、弱中心投影の仮定は、腕の振りや足の運び等の歩容の前後運動が興行き方向に現れる正面方向 (0 度と 180 度) 付近ではあてはまり難くなる。

そこで、最初にノイズによる変換誤差を軽減するために、上述の仮想的な参照方向の追加を行う。そのうえで、実在の参照方向が複数存在する場合で、かつその実在の参照方向と仮想的な追加方向が重なったと

きには、実在の参照方向の特徴を優先して用いることで、幾何モデルの仮定による変換誤差を軽減する。

## 5. 実験

### 5.1 データセット

実験には 20 人の被験者の 15 度刻みの 24 方向からなる 736 の歩容シーケンスを用いた。各歩容シーケンスは、10~20 歩の直線歩行で構成されている。VTM の学習セットには、10 人の被験者の 24 方向からなる 240 シーケンスを用いた。ギャラリーには、20 人の被験者の  $k$  参照方向からなる  $20k$  シーケンスを用いた。プローブには、VTM の学習に用いた被験者を除く 10 人の残りのシーケンスを用いた。各シーケンスの方向については、歩行者の軌跡を利用することで比較的容易に求まるものとして、事前に与えておく。

また、今回比較に用いる手法を以下にまとめる。

- NT: 参照特徴を変換せずにそのまま利用
- PP: 矢状面の透視投影<sup>5)</sup>
- VTM: 方向変換モデル
- VTM+: 幾何モデルによる参照方向を追加した VTM

ここで、参照方向が複数ある場合に、NT においては、プローブ方向に最も近い参照方向の特徴を利用して、PP においては、矢状面と前頭面 (矢状面に直交する垂直平面) の内で入力方向が近い方の面を基準面として透視投影変換するものとする。

### 5.2 方向変換の精度評価

最初に、方向変換の具体例として、1 参照方向 (0, 45, 90 度) と 2 参照方向 (0 と 90 度) から VTM+ によって変換した特徴を図 6 に示す。これより、全体として、変換特徴が各方向に対して元の特徴 (図 6(a)) に近いものとなっていることが分かる。

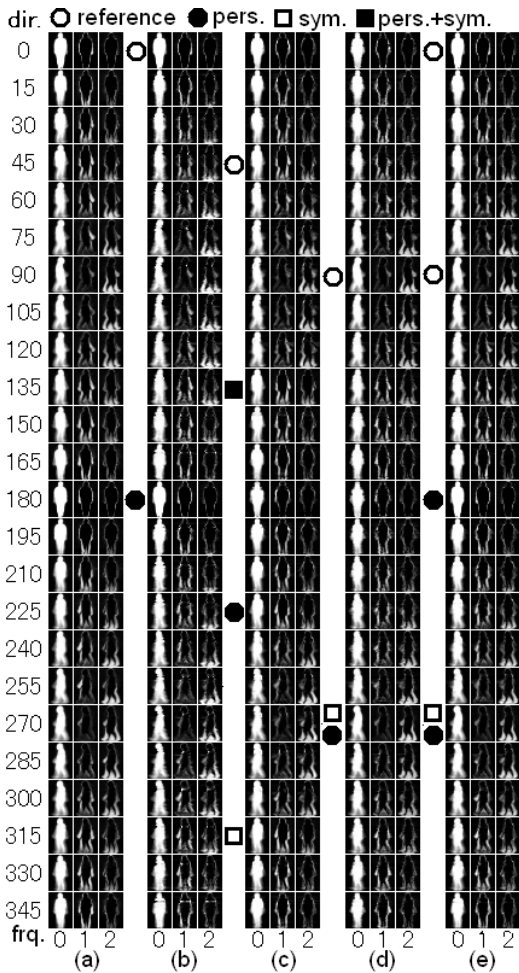
次に、方向変換の精度を評価するために、元の特徴  $a$  と変換特徴  $\hat{a}$  との自乗平均平方根誤差 (RMSE) を各比較手法に対して方向ごとに計算したものを図 7 に示す。ここで、VTM については、学習に用いた被験者に対する方向変換 (以下、VTM (tr.) とする) とそれ以外の被験者に対する方向変換の精度の違いを調べるために、別々に誤差を求めた。

全体の傾向について見てみると、NT や PP では、参照方向もしくは幾何モデルによる追加方向 から

方向  $\theta_{ref}$  は実際のカメラに対して定義され、実際のカメラおよび仮想的なカメラの両方に対して共通に使われる。

たとえば 0 度と 180 度に対する実在の特徴が存在する場合に、実在の 0 度特徴と 180 度の鏡像による仮想的な 0 度の特徴が重なる。

たとえば図 7(a) の  $\theta_{ref} = 0$  に対する 180 度や、図 7(c) の  $\theta_{ref} = 90$  に対する 270 度が相当。



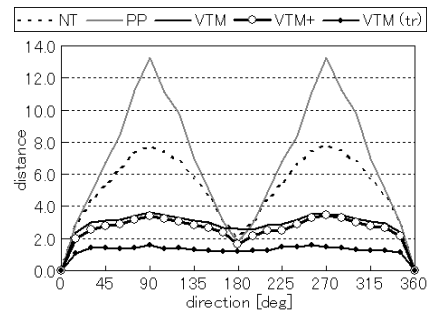
(a): original feature, (b)–(e): transformed features from 0, 45, 90, and 0 and 90 deg. reference

図 6 変換特徴

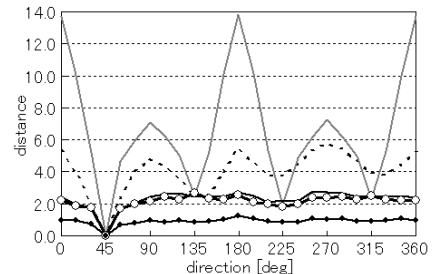
Fig. 6 Transformed features.

離れるに従って、誤差が大きくなっている．これに対して、VTM(tr.)では、いずれの方向に対してもほぼ一定の小さな誤差で抑えられている．また、VTMでは、VTM(tr.)に比べて2倍程度の誤差になっているものの、いずれの方向に対してもほぼ一定の誤差に抑えられている．ただし、幾何モデルによる追加方向に対しては、NTやPPよりも誤差が大きくなっている場合がある．一方、VTM+では、それらの方向においてもNTやPPと同程度の変換誤差になっており、全体としてVTMよりも小さい変換誤差に抑えられていることから、参照方向の追加が有効であることが分かる．

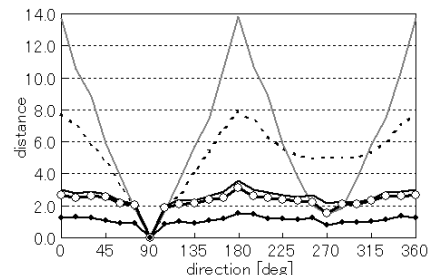
ただし、参照方向に対する直交方向については変換誤差がある程度大きくなっていることに注意されたい．



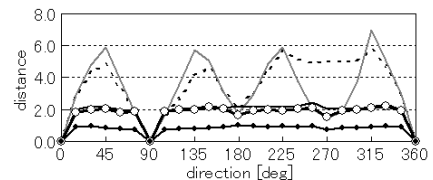
(a)  $\theta_{ref} = 0$



(b)  $\theta_{ref} = 45$



(c)  $\theta_{ref} = 90$



(d)  $\theta_{ref}(1) = 0, \theta_{ref}(2) = 90$

図 7 変換特徴の自乗平均平方根誤差 (RMSE)

Fig. 7 Root mean squared errors (RMSE) of transformed features.

たとえば、参照方向が0度の場合には、腕の振りや足の運び等の前後運動が画像面に直交して縮退しているため、図6(b)において側面方向(90度や270度)付近の変換特徴が不正確になっており、図7(a)においても側面方向付近で変換誤差が大きくなっている．また、参照方向が90度の場合には、前後運動についてはよく観測されるものの、体の幅等が画像面に直交して縮退する．よって、図6(d)において正面方向(0度や180

度) 付近の変換特徴が不正確になっており, 図 7(c) においても正面方向付近での変換誤差が大きくなっている.

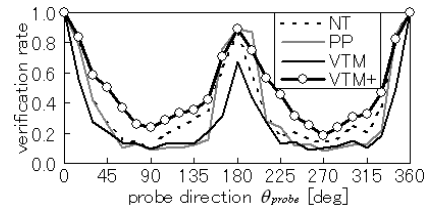
これに対して, 参照方向が 45 度の場合は, 直交方向 (135 度と 315 度) に対する特徴が, 歩容の左右対称性の仮定によって追加されている. そのため, 画像面に直交して縮退する動きが補われるため, 図 6(c) において, いずれの方向に対しても比較的正確な変換が行われており, 図 7(b) においても, 参照方向が 0 度や 90 度の場合 (図 7(a), (c)) と比べて全体的に変換誤差が小さくなっていることが分かる. 同様に, 参照方向が 0・90 度の 2 方向の場合にも, 直交方向が互いに補われているので, 図 6(e) において正確な変換が行われており, 図 7(d) においても変換誤差が小さくなっていることが分かる.

### 5.3 歩容認証の性能評価

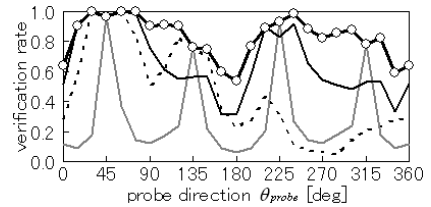
認証実験は, ランダムに選択された異なる 20 組の学習セットに対して行い, その平均性能を受信者操作特性 (ROC) 曲線<sup>14)</sup> における 10% 誤報率に対応する照合率によって評価した. 一例として, 参照方向 0, 45, 90 度の 1 方向, および 0, 90 度の 2 方向から変換した特徴をギャラリーとした場合の結果を図 8 に示す.

PP については, 参照方向もしくは幾何モデルによる追加方向 については高い照合率になっているものの, それらの方向から離れるにつれて照合率が大幅に低下している. また, VTM については, 参照方向が 0 度の場合 (図 8(a)) 以外では, 参照方向や追加方向から離れた方向に対する照合率が NT や PP よりも高い場合が多く, 平均的には NT や PP よりも高い照合率になっている. ただし, 追加方向に対しては, いずれも PP より低い照合率となっている.

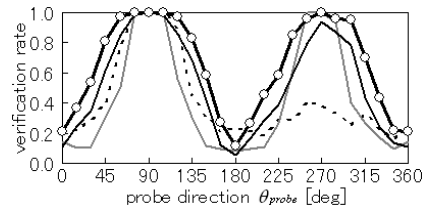
これに対して, VTM+ では追加方向に対しても PP と同等の照合率になっており, 全体として最も高い照合率を達成しているため, 幾何モデルによる追加方向が有効であることが分かる. 特に, 参照方向が 45 度の場合 (図 8(b)) は, 幾何モデルによる追加が 3 方向 (135, 225, 315 度) に対して行われるため, 参照方向が 0 度や 90 度の場合 (図 8(a), (c)) と比べて, 全体的に高い照合率になっており, 平均で 84% の照合率を達成している. また, 0, 90 度の 2 方向を参照方向とした場合 (図 8(d)) には, さらに全体的に照合率が上がっており, 照合率の悪いプローブの方向でも 80% 以上の照合率で, 平均では約 94% の照合率になっ



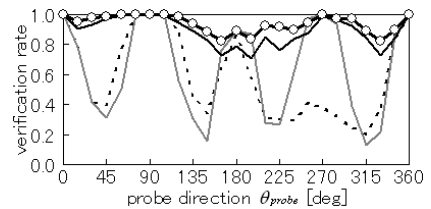
(a)  $\theta_{ref} = 0$



(b)  $\theta_{ref} = 45$



(c)  $\theta_{ref} = 90$



(d)  $\theta_{ref}(1) = 0, \theta_{ref}(2) = 90$

図 8 10% 誤報率時の照合率

Fig. 8 Verification rate at 10% false alarm.

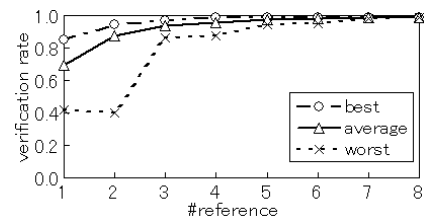


図 9 参照方向数に対する平均照合率

Fig. 9 Averaged verification rate for the number of references.

ている.

最後に, 参照方向数と平均照合率の関係を 図 9 に示す. 同図において, 各参照方向数に対して, 平均照合率が最高・最低となる参照方向の組合せを利用した結果が各々 best, worst で表されており, 全組合せの平均の結果が average で表されている. ここで, best

たとえば 図 8(a) の  $\theta_{ref} = 0$  に対する 180 度や 図 8(c) の  $\theta_{ref} = 90$  に対する 270 度が相当.

は登録時の歩行者の動線がある程度限定可能な場合に想定されるものであり、2方向で約94%の照合率を達成している。また、最低の場合(worst)でも3方向あれば、約87%の照合率を達成していることが分かる。これより、少数方向であっても方向変換を精度良く行うことができ、様々な方向の歩行に対して高い照合率を得られることが分かる。

## 6. おわりに

本論文では周波数領域における方向変換モデル(VTM)を用いた様々な方向に対する歩容認証について述べた。最初に、歩容シルエットの時空間ボリュームから、フーリエ解析によって振幅スペクトルを歩容特徴として抽出した。次に、複数被験者の複数方向の特徴からなる学習セットを特異値分解することでVTMを学習して、少数の参照方向の特徴をもとにした方向変換を定式化した。また、方向変換の精度を上げるための幾何モデルに基づく参照方向の追加方法について述べた。実験では、20人の被験者の24方向を含む739シーケンスを用いて歩容認証を行い、10%誤報時の照合率によって性能を評価した。結果として、VTMによって少数方向から他の方向への変換を精度良く行うことで、従来法の透視投影による方法よりも高い認証性能を達成した。具体的には、45度の1参照方向に対して平均84%の照合率を、0度と90度の2参照方向に対して平均94%の照合率が得られた。

また、今回の実験で得られた認証性能を検討するために他の手法について着目してみると、Sarkarら<sup>16)</sup>が、側面に近い方向から撮影したギャラリーとプローブ(方向差が30度程度)に対する歩容認証実験において、10%誤報に対して96%の照合率となったことを報告しており、方向変化等を含む顔認証用のデータベースFRVT2002<sup>13)</sup>においては、最も認証性能の良かった手法で10%誤報に対して96%の照合率となったことを報告している。これらの照合率は本手法の2方向参照時の平均照合率94%に近い値であり、本研究で扱った方向変化の大きさを考慮すると、本手法が有効な歩容認証手法であると考えられ、顔認証と同様に歩容認証が有効な認証手段となりうるものと考えられる。

ただし、監視システムやアクセスコントロールといった目的での認証としては、本論文で示した精度ではまだ不十分であると考えられる。そこで、さらなる認証性能の向上を目指すためには、たとえば、最初に歩容認証によって認識対象の候補を限定しておき、次に限定された人物の顔にカメラをズームして顔認証を行うことで、全体として低誤報率・高照合率の認証性

能を得る方法が考えられる。よって、このような複数のバイオメトリクスの統合手法が、今後の課題の1つとしてあげられる。

またそれ以外にも、今後取り組むべき課題として、以下があげられる。

- カメラのチルト方向の変化への対応
- 一般的なデータベース(たとえば HumanID Gait Challenge Problem Datasets<sup>16)</sup>)に対する認証実験

## 参 考 文 献

- 1) BenAbdelkader, C.: Gait as a Biometric For Person Identification in Video, Ph.D. Thesis, Maryland Univ. (2002).
- 2) Cunado, D., Nixon, M. and Carter, J.: Automatic Extraction and Description of Human Gait Models for Recognition Purposes, *Computer Vision and Image Understanding*, Vol.90, No.1, pp.1-41 (2003).
- 3) Cuntoor, N., Kale, A. and Chellappa, R.: Combining Multiple Evidences for Gait Recognition, *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Vol.3, pp.33-36 (2003).
- 4) Han, J., Bhanu, B. and Roy-Chowdhury, A.: A STUDY ON VIEW-INSENSITIVE GAIT RECOGNITION, *Proc. IEEE Int. Conf. on Image Processing*, Vol.3, pp.297-300 (2005).
- 5) Kale, A., Roy-Chowdhury, A. and Chellappa, R.: Towards a View Invariant Gait Recognition Algorithm, *Proc. IEEE Conf. on Advanced Video and Signal Based Surveillance*, pp.143-150 (2003).
- 6) Kobayashi, T. and Otsu, N.: Action and Simultaneous Multiple-Person Identification Using Cubic Higher-Order Local Auto-Correlation, *Proc. 17th Int. Conf. on Pattern Recognition*, Vol.3, pp.741-744 (2004).
- 7) Lee, L.: Gait Analysis for Classification, Ph.D. Thesis, Massachusetts Institute of Technology (2002).
- 8) Liu, Y., Collins, R. and Tsin, Y.: Gait sequence analysis using frieze patterns, *Proc. 7th European Conf. on Computer Vision*, Vol.2, pp.657-671 (2002).
- 9) Liu, Z. and Sarkar, S.: Simplest Representation Yet for Gait Recognition: Averaged Silhouette, *Proc. 17th Int. Conf. on Pattern Recognition*, Vol.1, pp.211-214 (2004).
- 10) Mowbray, S. and Nixon, M.: Automatic Gait Recognition via Fourier Descriptors of Deformable Objects, *Proc. IEEE Conf. on Ad-*



- vanced Video and Signal Based Surveillance*, pp.566–573 (2003).
- 11) Mukaigawa, Y., Nakamura, Y. and Ohta, Y.: Face Synthesis with Arbitrary Pose and Expression from Several Images — An integration of Image-based and Model-based Approach, *Proc. 3rd Asian Conf. on Computer Vision*, Vol.1, pp.680–687 (1998).
  - 12) Murase, H. and Sakai, R.: Moving Object Recognition in Eigenspace Representation: Gait Analysis and Lip Reading, *Pattern Recognition Letters*, Vol.17, pp.155–162 (1996).
  - 13) Phillips, P., Blackburn, D., Bone, M., Grother, P., Micheals, R. and Tabassi, E.: Face Recognition Vendor Test (2002). <http://www.frvt.org>
  - 14) Phillips, P., Moon, H., Rizvi, S. and Rauss, P.: The FERET Evaluation Methodology for Face-Recognition Algorithms, *Trans. of Pattern Analysis and Machine Intelligence*, Vol.22, No.10, pp.1090–1104 (2000).
  - 15) Sagawa, R., Makihara, Y., Echigo, T. and Yagi, Y.: Matching Gait Image Sequences in the Frequency Domain for Tracking People at a Distance, *Proc. 7th Asian Conference on Computer Vision*, Vol.2, pp.141–150 (2006).
  - 16) Sarkar, S., Phillips, J., Liu, Z., Vega, I., Grother, P. and Bowyer, K.: The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis, *Trans. Pattern Analysis and Machine Intelligence*, Vol.27, No.2, pp.162–177 (2005).
  - 17) Shakhnarovich, G., Lee, L. and Darrell, T.: Integrated Face and Gait Recognition from Multiple Views, *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Vol.1, pp.439–446 (2001).
  - 18) Spencer, N. and Carter, J.: TOWARDS POSE INVARIANT GAIT RECONSTRUCTION, *Proc. IEEE Int. Conf. on Image Processing*, Vol.3, pp.261–264 (2005).
  - 19) Urtasun, R. and Fua, P.: 3D Tracking for Gait Characterization and Recognition, *Proc. 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pp.17–22 (2004).
  - 20) Utsumi, A. and Tetsutani, N.: Adaptation of appearance model for human tracking using geometrical pixel value distributions, *Proc. 6th Asian Conf. on Computer Vision*, Vol.2, pp.794–799 (2004).
  - 21) Wagg, D. and Nixon, M.: On Automated Model-Based Extraction and Analysis of Gait, *Proc. 6th IEEE Int. Conf. on Automatic Face and Gesture Recognition*, pp.11–16 (2004).
  - 22) Yam, C., Nixon, M. and Carter, J.: Automated Person Recognition by Walking and Running via Model-based Approaches, *Pattern Recognition*, Vol.37, No.5, pp.1057–1072 (2004).
  - 23) Yu, S., Tan, D. and Tan, T.: Modelling the Effect of View Angle Variation on Appearance-Based Gait Recognition, *Proc. 7th Asian Conf. on Computer Vision*, Vol.1, pp.807–816 (2006).
  - 24) 村瀬 洋 : Nayar, S. : 2次元照合による3次元物体認識—パラメトリック固有空間法, *電子情報通信学会論文誌*, Vol.J77-D-II, No.11, pp.2179–2187 (1994).

(平成 18 年 5 月 5 日受付)

(平成 18 年 11 月 10 日採録)

(担当編集委員 太田 直哉)



榎原 靖 (正会員)

2001年大阪大学工学部応用理工学科卒業。2002年同大学大学院工学研究科電子制御機械工学専攻博士前期課程修了。2005年同専攻博士後期課程修了。博士(工学)。同年同大学産業科学研究所特任助手, 2006年同研究所助手となり現在に至る。物体認識, 色恒常性, 歩容認証に関する研究に従事。日本機械学会第79期ロボティクス・メカトロニクス部門ベストプレゼンテーション表彰受賞。日本ロボット学会第18回研究奨励賞受賞。情報処理学会平成18年度山下記念研究賞受賞。日本ロボット学会, 日本機械学会各会員。



佐川 立昌 (正会員)

1998年京都大学工学部情報工学科卒業。2000年東京大学大学院工学系研究科情報工学専攻修士課程修了。2003年東京大学大学院工学系研究科電子情報工学専攻博士課程修了。現在, 大阪大学産業科学研究所助手。実時間視覚処理と物体のモデリングの研究に従事。博士(工学)。日本ロボット学会, 電子情報通信学会, IEEE各会員。



向川 康博 (正会員)

1997年筑波大学大学院博士課程工学研究科修了。同年～2002年岡山大学助手。2003～2004年10月筑波大学講師。2004年11月より大阪大学助教授。コンピュータビジョン、複合現実感の研究に従事。博士(工学)。電子情報通信学会、日本バーチャルリアリティ学会、IEEE各会員。



越後 富夫 (正会員)

1980年大阪府立大学工学部電気工学科卒業。1982年同大学大学院修士課程修了。2003年大阪大学大学院基礎工学研究科博士後期課程修了。博士(工学)。1982年日本アイ・ピー・エム(株)入社。同社東京基礎研究所においてロボットビジョン、知能映像メディアの研究に従事。2003年大阪大学客員教授。2006年大阪電気通信大学情報通信工学部情報工学科教授。医用画像処理、映像検索の研究に従事。電子情報通信学会、日本ロボット学会各会員。



八木 康史 (正会員)

1983年大阪大学基礎工学部制御工学科卒業。1985年同大学大学院修士課程修了。同年三菱電機(株)入社。同社産業システム研究所にてロボットビジョンの研究に従事。1990年大阪大学基礎工学部情報工学科助手。同学部システム工学科講師。同大学院助教授を経て、2003年より同大学産業科学研究所教授。1995～1996年英オックスフォード大学客員研究員。2002年仏ピカルディー大学招聘助教授。全方位視覚センシング、画像理解、知能ロボットに関する研究に従事。1996年度電子情報通信学会論文賞、2003年ACM VRST2003 Honorable Mention Award、IEEE、電子情報通信学会、日本ロボット学会各会員。博士(工学)。