

## 個人識別のための歩行動作特徴の抽出と評価

吉村 巧朗 † 桐島 俊之 ††  
† 奈良工業高等専門学校 †† 奈良工業高等専門学校

概要：近年、人の身体的特徴や動作を用いるバイオメトリクス認証が提案され、窃盗・紛失の危険性が低い認証技術として実用化されている。現在では新たなバイオメトリクス技術として歩行動作から特徴を抽出する方法が研究され、実現すれば、人への負担が少ない認証技術になると期待されている。過去に行われた研究の多くでは、歩行動作そのものではなく、歩行者を撮影した動画の各フレームからシルエットや姿勢特徴を抽出し、フレーム毎にそれらのマッチングを取る手法が多い。そこで本論文では、歩行動作中に含まれる周波数成分を用いることで歩行動作自体を特徴として抽出する方法を提案する。また、評価実験により特徴が認証に応用できるか検討する。

## Extraction and evaluation of human gait features for person identification

Takuro Yoshimura † Toshiyuki Kirishima ††  
†Nara National College of Technology NNCT  
††Nara National College of Technology NNCT

Abstract: In recent years, the method of extracting the feature from walk operation is studied as new biometric technology, and it is expected to decrease the burdens to people. Many of researches done in the past extracted the silhouette and the posture feature from each frame in the video. But it is not feature of walk operation itself. In this paper, it is proposed the method of extract the walk operation itself as a feature by using the frequency component contained during walk operation. And it is examined whether the feature can apply to attestation by evaluation experiment.

### 1 まえがき

近年、人の身体的特徴や動作を用いるバイオメトリクス認証が提案され、窃盗・紛失の危険性が低い認証技術として注目されている。バイオメトリクスには、カメラで撮影された顔画像から特徴を抽出し個人識別を行う顔認識や、手のひらや指中にある静脈パターンを読み取り、それを特徴として個人識別を行う静脈認証がある。このようなバイオメトリクスはアクセスコントロールの一種として既に実用化されている。ただし顔認識や静脈認証は特徴を抽出する際に、識別対象者の協力を必要とする。顔認識

では、高い認識率を得るためにある程度高解像度の顔画像が必要である。よって識別対象者はカメラの方を向く、ある程度カメラに近づく等の行動をする必要がある。静脈認証では、静脈パターンの取得に専用の読み取り装置を用いる場合が多い。そのため識別対象者には手や指で読み取り装置に触れてもらう必要がある。

そこで現在、映像中に存在する人の歩行動作より特徴を抽出し、個人識別を行う手法が注目され始めている。歩行動作により個人識別を行う研究は近年始められたもので、どの程度の精度が実現可能なの

かは未だよく分かっていないが、特別な機器を使用しない、識別対象者にかかる負担が少ない等の利点から、実現が強く期待されている。これまでに行われている関連研究には、歩行者を撮影した動画中の各フレームから歩行者のシルエット画像を抽出し、画像認識手法に基づき標準パターンと照合させることで、個人識別を行う手法がある [1]。また、モーションキャプチャで計測した歩行動作をモデル化し、そのモデルを元に歩行動作を解析する手法も提案されている [2]。また、将来的に研究が進み新しい手法が提案された際、その比較対照となるベースラインアルゴリズムも提案・公開されている [3]。しかし、これらの多くは歩行動作自体より各フレームでの姿勢を重視している。

本論文では歩行者の姿勢画像そのものではなく、歩行動画像中に含まれる周波数成分に着目する。歩行動作は手足を前後に動かす周期的な運動である。そのため歩行動作中の周波数成分には、歩行動作に関する様々な特徴が含まれていると期待できる。そこで本研究では、歩行動作中に含まれる周波数成分が個人識別の特徴として利用出来るか検討する。

本論文では 2 章にて従来手法であるベースラインアルゴリズムについて述べる。3 章では本論文で提案する特徴を抽出する方法を述べる。4 章では 3 章で抽出した特徴の評価実験を行い、提案する特徴が個人識別に有意であるか検討する。5 章でまとめを述べる。

## 2 従来手法

この章では、既存手法の一つであるシルエットテンプレートマッチング [3] を紹介する。シルエットテンプレートマッチングとは、将来的に研究が進み新たな認識手法が提案された際、その比較対象となるべくして考えられたアルゴリズムである。

### 2.1 用語の定義

シルエットテンプレートマッチングの説明に先立ち、次の用語を定義する。図 1(a) のような原画像に対し、歩行者が映っていない図 1(b) のような画像を背景画像とする。また、図 1(a) 中の図 1(c) のような歩行者が映っている領域を歩行者領域とする。さらに、図 1(c) の歩行者領域に対して図 1(d) のような、歩行者に相当する領域を黒く塗りつぶした画像をシルエットと呼ぶ。

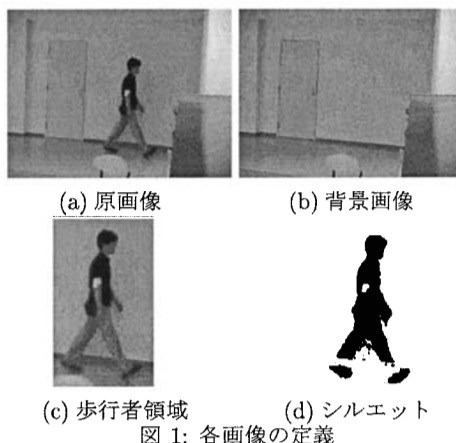


図 1: 各画像の定義

### 2.2 処理手順

シルエットテンプレートマッチングの手順は次の通りである。まず、動画中の各フレームで歩行者領域の抽出を行う。次に抽出された各歩行者領域からそれぞれ歩行者のシルエットを抽出する。そして、各シルエットの画素数より歩行周期を求める。最後に入力動画とデータベース上の動画間の類似度を求める。この類似度の高いものから順に、第 1 位、第 2 位…という認識結果になる。以下、各手順について説明する。

### 2.3 歩行者領域の抽出

初めに歩行者領域を動画中の各フレームから抽出する。歩行者がある程度一定の速度で一定の方向に移動していると仮定すれば、歩行者領域はフレームが進むにつれ線形的に移動する。よって、いくつかのフレームで歩行者領域が既知であれば、他のフレーム中にある歩行者領域の位置を推定することが出来る。シルエットテンプレートマッチングでは、最初、中間、最後の 3 フレーム中の歩行者領域を元に、各フレーム中にある歩行者領域を求める。ただし元となる 3 フレーム中の歩行者領域は、手動で切り出す。

### 2.4 シルエットの抽出

次に歩行者領域から、シルエットを抽出する。あらかじめ背景画像を用意しておき、歩行者領域と背景画像中の対応する画素間でマハラノビス距離  $D_m$  を計算する。画素  $i$ 、画素  $j$  間のマハラノビス距離

$D_m$  は式 (1) で求められる。

$$D_m(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^p \frac{(x_i - y_i)^2}{\sigma_i^2}} \quad (1)$$

ここで  $x_i, y_i$  は各画素が持つ画素値であり、 $p$  は画素が持つ色要素の数である。例えば RGB 空間の画素では  $p = 3$  となる。また  $\sigma_i$  は  $x_i$  の集合における標準偏差である。マハラノビス距離を元に、歩行者に対応する画素のみが黒になるよう手動でしきい値を決めて二値化する。その後、シルエットの高さが 128 ピクセル、幅が 88 ピクセルとなるよう画像を拡大・縮小する。

## 2.5 歩行周期の抽出

このステップではまずシルエット中の黒画素数を各フレームで計測する。人が手足を広げている間はシルエット中の黒画素数が増加する。また、手足をたたんだ直立状態の間は黒画素数が減少する。よって縦軸に黒画素数、横軸にフレーム数を取ると正弦波状のグラフになる。歩行動作の一周期の間には、手足を広げる、手足をたたむという動作を 2 回行う。そのため、正弦波状のグラフでの 2 周期分が歩行周期の 1 周期分に相当する。よって 2 周期毎にフレーム数を計測していき、その中間値を歩行周期  $N_{gait}$  とする。

## 2.6 類似度の計算

データベース上の動画画像  $S_g = \{S_g(1), \dots, S_g(N)\}$  と入力動画画像  $S_p = \{S_p(1), \dots, S_p(M)\}$  間の類似度を計算する。まず、式 (2)、式 (3) でフレーム間類似度  $FrameSim$  と相関値  $Corr$  を定義する。

$$FrameSim(S_p(i), S_g(j)) = \frac{Num(S_p(i) \cap S_g(j))}{Num(S_p(i) \cup S_g(j))} \quad (2)$$

$$Corr(S_{pk}, S_g)(l) = \sum_{j=0}^{N_{gait}-1} FrameSim(S_p(k+j), S_g(l+j)) \quad (3)$$

ここで  $S_{pk} = \{S_p(kN_{gait}, \dots, (k+1)N_{gait})\}$  である。また、 $Num$  はシルエットの画素の数を返す関数である。そして、式 (2)、式 (3) を元に式 (4) で示す類似度を計算し、評価値とする。

$$Sim(S_p, S_g) = Median_k(max_l Corr(S_{pk}, S_g)(l)) \quad (4)$$

ここで、 $Median, max$  は与えられたデータ群からそれぞれ中間値、最大値を抽出する演算子である。

## 3 歩行動作特徴の抽出

本章では本論文で提案する特徴の抽出法を述べる。

### 3.1 特徴ベクトル抽出処理手順

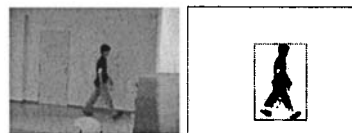
本論文では歩行動作の周波数成分を特徴ベクトルとして用いる。特徴ベクトルを抽出する手順は、まず最初に動画中の各フレームからシルエットを抽出する。次に、シルエットをいくつかの区間に分割し、それぞれの区間でシルエットの平均幅を求める。そしてシルエット平均幅を区間別にまとめて時系列データにする。最後に FFT を用いて各区間のシルエット平均幅の周波数成分をそれぞれ抽出する。ここで抽出された周波数成分を特徴ベクトルとする。

### 3.2 シルエット抽出

シルエットの抽出では、ある程度の精度を確保するために半手動でシルエットの抽出を行う。まず、あらかじめ人が映っていない背景画像を用意する。次に、背景画像と各フレーム間のそれぞれ対応する画素でユークリッド距離  $D_e$  を計算する。画素  $\vec{x}$ 、画素  $\vec{y}$  間のユークリッド距離  $D_e$  は式 (5) で求められる。

$$D_e(\vec{x}, \vec{y}) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2} \quad (5)$$

式 (1) 同様、 $p$  は画素が持つ色要素の数である。そして、ユークリッド距離を元にシルエットを黒、背景を白として 2 値化する。この時のしきい値は動画画像毎に適切な値を手動で設定する。さらにノイズ除去のため、 $3 \times 3$  のメディアンフィルタをかける。メディアンフィルタは注目画素を中心とする  $n \times n$  の範囲の画素を画素値でソートし、その  $\frac{n \times n}{2}$  番目の値を注目画素の画素値とするフィルタである。最後に、2 値化された画像から歩行者のシルエットを含む領域を手動で切り出す。この時、白画素しか含まない上端下端の方形領域は自動で除去する。以上の処理により、図 2(a) のような画像から、図 2(b) 中の枠で囲われた領域のようなシルエットを抽出できる。



(a) 原画像 (b) シルエット抽出  
図 2: シルエット抽出

### 3.3 シルエット平均幅

シルエットを図3のようにいくつかの区間に分割し、各区間でそれぞれシルエットの平均幅  $W_s$  を求める。分割する区間の数は、人の頭身が大体8頭身程度であると仮定して、その倍の16とする。またそれに従い、分割した区間は足下から順に0～15の区間番号を付ける。



図3: シルエットの分割

一区間の長さを1とすると、区間番号  $k$  のシルエット平均幅  $W_s(k)$  は式(6)で求める。

$$W_s(k) = \frac{1}{l} \sum_{i=0}^{l-1} w((k \times l) + i) \quad (6)$$

ここで  $w(j)$  はシルエット中の  $j$  番目の走査線におけるシルエットの幅である。ただし、最下端の走査線を0番目として扱う。またシルエットの幅は、走査線中の左端の黒画素と右端の黒画素間の距離とする。

### 3.4 周波数成分の抽出

まずシルエット平均幅  $W_s$  を各フレームの各区間でそれぞれ求め、区間毎に時系列データとしてまとめる。すると16区間分のシルエット平均幅の時系列データが得られる。そこで次にシルエット平均幅の時系列データを高速フーリエ変換 (FFT) にかける。そしてFFTの出力よりパワースペクトルを求め、特徴ベクトルとする。以上の処理で抽出できる特徴ベクトルの一例を図4に示す。

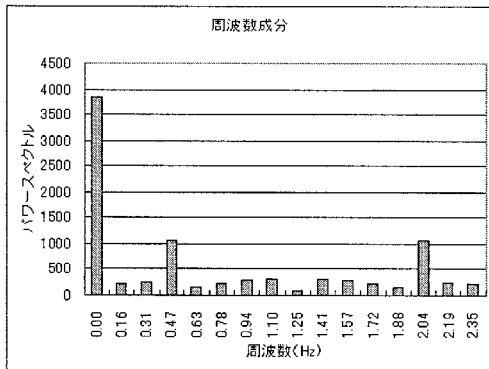


図4: 特徴ベクトル例

図4と同様の特徴ベクトルが分割区間の数である16個分抽出できる。

## 4 評価実験

前節で述べた手法によって抽出される特徴により、実際に個人識別が可能か実験した。実験内容は、まず歩行者を撮影した動画を複数個用意し、それぞれ特徴ベクトルを抽出した。そして、特徴ベクトル間の評価値とF検定の二つの方法で評価した。最後にその結果について考察した。以下、実験に使用した動画像と評価方法、結果と考察について述べる。

### 4.1 実験用歩行動画像について

実験にはそれぞれ異なる歩行者を撮影した動画を9個用いた。歩行者の撮影は、図5のような状況で行った。撮影場所は屋内で、背景となる廊下は一面白色だった。照明はカメラの背後にある窓から入る天然光と、天井の蛍光灯の二種の光源で構成されていた。

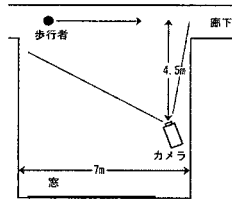


図5: 撮影状況

### 4.2 特徴間の評価値

特徴の評価方法として、二つの方法を用いた。一つは、各動画像から得られた特徴ベクトル間の評価値を求める方法を用いた。もし同一の被験者間において高く、異なる被験者間で低い評価値が得られれば、評価値を元に個人を分離できる。そのため、この特徴は個人識別に用いられると言える。本実験では、評価値として各動画像から抽出した特徴ベクトルの相関係数を用いた。特徴を  $M, M'$  とすると相関係数  $C$  は式(7)から求められる。

$$C(M, M') = \frac{\langle M, M' \rangle}{\|M\| \times \|M'\|} \quad (7)$$

ここで、 $\langle M, M' \rangle$  は特徴ベクトル間の内積で、式(8)で求められる。また、 $\|M\|, \|M'\|$  はそれぞれ  $M, M'$  のノルムで、特徴ベクトルの次元数を  $N$  とすると式(9)で求められる。

$$\langle M, M' \rangle = \sum_{i=0}^N M_i \times M'_i \quad (8)$$

$$\|M\| = \sqrt{\sum_{i=0}^N M_i^2} \quad (9)$$

### 4.3 F 検定

もう一つはF検定による方法を用いた。F検定は分散分析の一種であり、二つのデータ群が同一の母集団に属するという帰無仮説を棄却する検定である。すなわち、入力したデータ群が異なる母集団に属する可能性があるということを証明できる。F検定では式(10)で表されるF値を元に、式(11)で示すF分布を用いて棄却域を求める。

$$F(\phi_1, \phi_2) = \frac{\sigma_1^2/\phi_1}{\sigma_2^2/\phi_2} \quad (10)$$

$$f_{\phi_1, \phi_2}(F) = \frac{\phi_1^{\phi_1/2} \phi_2^{\phi_2/2}}{B(\phi_1/2, \phi_2/2)} \frac{F^{\phi_1/2-1}}{(\phi_1 F + \phi_2)^{(\phi_1+\phi_2)/2}} \quad (11)$$

ここで、 $\sigma_1^2, \sigma_2^2$  は検定にけるデータ群の分散、 $\phi_1, \phi_2$  は自由度である。n個のデータを持つデータ群  $\vec{x}$  の分散  $\sigma^2$  は、 $\vec{x}$  の平均を  $\bar{x}$  とすると式(12)で求められる。また、Bはベータ関数と呼ばれるもので、式(13)で定義される。

$$\sigma^2 = \frac{1}{n} \sum_{i=0}^n (\bar{x} - x_i)^2 \quad (12)$$

$$B(\lambda_1, \lambda_2) = \int_0^1 x^{\lambda_1-1} (1-x)^{\lambda_2-1} dx \quad (13)$$

本実験では、特徴ベクトル群中にある要因の効果と残差についてF検定を行った。ここで言う特徴ベクトル群とは、各動画像から抽出した特徴ベクトルを区間にまとめたものである。今回の場合、要因の効果とは被験者の違いであり、それ以外の効果によるばらつきが残差である。特徴ベクトル中にあるデータのばらつき  $G$  は要因の効果のばらつき  $G_f$  と残差のばらつき  $G_o$  の和と考えると、式(14)が得られる。また、特徴ベクトル中にあるデータのばらつき  $G$  自体は式(15)で表せる。

$$G(k, l) = G_f(k, l) + G_o(k, l) \quad (14)$$

$$G(k, l) = V_f(k, l) - \frac{1}{N \times M} \sum_{i=1}^N \sum_{j=1}^M V_f(i, j) \quad (15)$$

ここで、 $V_f(k, l)$  ( $k = 1, 2, \dots, N, l = 1, 2, \dots, M$ ) は特徴ベクトル  $k$  の  $l$  番目の要素を表す。また、要因の

効果のばらつき  $G_f$  は要因の平均値に代表されると考えると、 $G_f$  と  $G_o$  は式(16)式(17)で表される。

$$G_f(k, l) = \frac{1}{M} \sum_{j=1}^M V_f(k, j) \quad (16)$$

$$G_o(k, l) = G(k, l) - G_f(k, l) \quad (17)$$

以上のことを踏まえ本実験では、特徴ベクトル群から  $G_f, G_o$  を求めた。そしてそれぞれの分散を求め、式(10)でF値を計算し、F検定を行った。要因の効果と残差の母集団が同一でないと判断できれば、要因の効果は有意であると言える。この検定を各区間の特徴ベクトル群で行い、どの区間がより有意であるかを調べた。

### 4.4 実験結果および考察

前述の実験により16区間それぞれで評価値とF値が得られた。16箇所ある特徴ベクトルのF値より、対応する棄却域を求めたところ、図6のような結果になった。

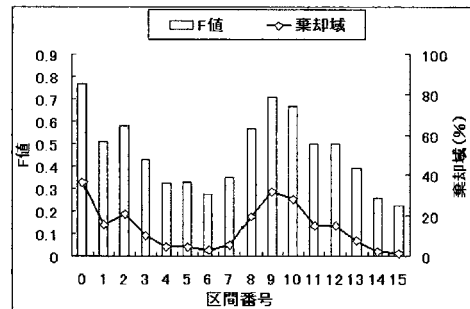
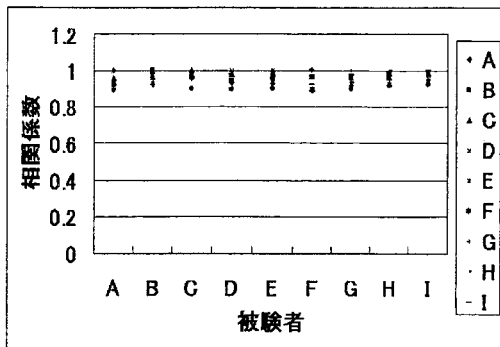


図6: F値と棄却域

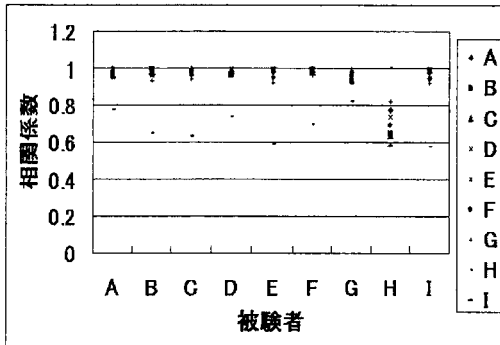
図6で示した棄却域からは、どの区間でも有意な差があるとは言えない。これについては抽出した周波数成分をすべて用いて評価値を計算したため、個人性に関わらない特徴（歩行動作そのものを示す特徴など）が存在し、結果に影響を与えた可能性がある。しかし、図6では0, 9番の区間付近においてF値が比較的高い値を示していることから、0, 9番にあたる足、腕部で比較差があることがわかる。このことから、被験者間には手足の動きに違いがあり、個人性を示す特徴が比較的に顕著に現れていると推察できる。

そこで、0,9番区間での評価値を図7に示す。図中でA~Iは各被験者を表す。





(a) 区間番号 0



(b) 区間番号 9

図 7: 被験者間の評価値

図 7 を見る限り、どの被験者間でも高い評価値となることが多く、F 値同様に有意な差は見られない。これに関しても F 値同様に個人性にかかわらない特徴の影響があると思われる。今回は被験者数が 9 人と少ないこともあり、この時点で個人識別の可・不可は判断できないが、評価値が一様に高ければ、識別が困難になると予測できる。よってこの結果の一因だと思われる、個人性にかかわらない特徴を除去する必要性が示唆できる。

#### 4.5 今後の課題

前節の実験結果より、今後の課題を考える。まず今回の実験では被験者の母数が少なすぎるため、特徴が個人識別に利用可能か不可能かという結論には至らなかった。そのため被験者数を増やして再実験し、個人識別に利用可能・不可能の判断を行う必要がある。次に実験結果より、提案した特徴には個人性に関わらない歩行動作の特徴が含まれていると推測された。個人性に関わらない特徴が存在していれば、個人識別の際に誤りを発生させる要因となりえる。そのため個人性に関わらない特徴がどのように存在するのか調査し、それを取り除く必要がある。これに関して、個人性は歩行動作中の癖など、細かい動

きに多く含まれると仮定する。すると、特徴ベクトル中の低周波成分を除去し、高周波成分を重視した評価を行えば、個人性に関係の薄い特徴は除去できると考える。

## 5 まとめ

本研究では、歩行動作の周波数成分を抽出し、個人識別に利用可能か検討した。

特徴抽出に使用する入力には歩行者を撮影した動画とする。周波数成分の抽出法は、まず歩行動作中の各フレームよりシルエットを抽出する。次にシルエットを 16 区間に分割し、各区間毎にシルエット平均幅を求める。そしてシルエット平均幅を各フレームで求めて区間毎に時系列データとしてまとめる。最後にシルエット平均幅の時系列データを FFT にかけて、周波数成分を抽出する。この周波数成分を特徴ベクトルとする。

提案した特徴ベクトルが個人識別に利用可能か調べるため、評価実験において評価値の計算と F 検定を行った。その結果、特徴中に個人性に関わらない歩行動作の特徴が含まれている可能性が示唆された。また F 検定の結果より、腕部に比較的顕著に個人性が含まれているであろうことが分かった。しかし実験に用いた被験者数が少なすぎたため、今後は被験者数を増やし実際に利用可能かどうかを明らかにすることが課題である。

## 参考文献

- [1] Zongyi Liu, Sudeep Saekar, "Improved Gait Recognition by Gait Dynamics Normalization", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.28, No.6, June 2006, pp863-876
- [2] 高崎 秀明, 稲垣 伸吉, 鈴木 達也, "リンクモデルの確率的切り替えに基づく歩行動作のモデル化と解析", ロボティクス・メカトロニクス講演会 2006, 2P1-A13, May 26-28, 2006, 早稲田
- [3] Sudeep Sarkar, P.Jonathon Philips, Zongyi Liu, Isidro Robledo Vega, Patrick Grother, Kevin W.Bowyer, "The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol27, No2, FEBRUARY 2005, pp.162-177