Vol.2012-UBI-35 No.9
Vol.2012-EC-25 No.9
2012/7/14

# マーカーレスモーションキャプチャ装置を用いた歩容による個人識別法の検討

高田 憲一[1]　北須賀 輝明[1]　有次 正義[1]

**概要**：本研究では，マーカーレスモーションキャプチャ装置として Kinect for Xbox360 を用いて，歩容によって家庭内で個人識別する方法を検討する．カメラに映る歩行の向きが変わると見え方が変わるが，モーションキャプチャ装置から 3 次元の骨格データが取得できるため，それを辞書データと同じ向きの歩容に変換することで，個人識別できると考え調査した．識別には機械学習 (サポートベクタマシン) を使用し，10 分割交差検定で評価する．ひとつの向きの歩容で識別したときには 10 人のデータで 82.4%の精度で識別できた．カメラの光軸に対して 45 度の歩容を 90 度の歩容に見えるように変換し識別したが，精度は低かった．そのため，様々な角度の組合せで同時に歩容を撮影したところ，90 度の歩容では骨格のずれが大きいことがわかった．

**キーワード**：歩容, 個人識別, 機械学習, モーションキャプチャ, Kinect

# A Consideration of an Individual Identification Method by Gait Using a Markerless Motion Capture Device

Kenichi Takada[1]　Teruaki Kitasuka[1]　Masayoshi Aritsugi[1]

***Abstract:*** In this work, we examine to recognize household members by gait using Kinect for Xbox360 as a markerless motion capture device. We thought and examined that could recognize members with gait of a various direction by converting to look like the gait of the same direction as the dictionary data, because we could acquire of skleton data as three dimensional data. We show 10 people identification accuracy of 82.4% with a support vector machine classifier in gaits of one direction. It was converted to look like the walking gait of 45 degrees to 90 degrees from the optical axis of camera, individual identification accuracy was low. Therefore, we compared the two gaits that were captured at the same time. And it was revealed that the gap of skeleton data was big by the gait of 90 degrees.

***Keywords:*** gait, individual identification, machine learning, motion capture, Kinect

## 1. はじめに

家電の性能が向上したことで，個人に合わせたサービスを家電で提供しようという動きがある．たとえば，テレビでの番組推薦や広告の提供 [1][2]，エアコンで個人に合わせた温度管理などがある．しかし，そのためには個人を識別する方法が必要となる．本研究では個人を識別する方法として，Microsoft 社製のカメラ画像や深度画像が取得可能なゲームコントローラ Kinect[3] を用いた歩容による個人識別を提案する．歩容とは歩いている時の身体運動の様子を指しており，見た目に表れる歩き方である．この方法では Kinect のカメラの範囲内で両足で一歩ずつ (以降，1 周期と呼ぶ) 歩くだけで個人識別するため，ユーザに特別な入力を求めることなく識別することができる．また，カメラを使用して個人を識別する際に問題となるプライバシについては，Kinect から得られる骨格データと呼ばれる体の 20 個のポイントだけに注目して識別することで軽減する．

[1] 熊本大学 大学院 自然科学研究科
Graduate School of Science and Technology, Kumamoto University

© 2012 Information Processing Society of Japan

歩容認証 [4][5] は歩容をバイオメトリクスとして利用して個人を認証する技術のひとつである．認証は本人であることを確認して，その人がある権利を行使できる状態にすることで，識別は複数の人の中から本人を見つけて区別することである．歩容認証は，もともとカメラから離れていて顔による識別ができない場合に識別することを想定された認証方法で，現在は犯罪捜査にも使われている．この方法では，カメラに映る歩行の向きが変わると見え方が変わるため，予め同じ向きの辞書データを用意するか，3 次元モデルを作成しそれを使うことで同じ向きの辞書データを用意するなどの方法を用いている．

Kinect for Xbox360 は，ジェスチャ認識機能や音声認識機能を有する Xbox360 用のコントローラである．本研究では Kinect for Windows SDK beta 2 を使用した．Kinect は RGB カメラと深度カメラ，マルチアレイマイク，プロセッサを持つ．深度カメラは，800mm から 4000mm の範囲を mm オーダーで Kinect からの距離画像を撮影する．ユーザのシルエットや骨格データは，深度カメラから取得したデータを基に Kinect 内のプロセッサで推定される．

本研究では，歩容認証がカメラを使いながら顔画像を用いずに識別できることに注目し，Kinect を使うことで骨格データだけで個人を識別することを目標としている．本研究では，家族内での使用を想定しているが，データ収集には実際の家族を使用していない．家族内には体形が似ている人物がいる可能性があると考えたことから，体形の情報を今回の研究では識別に使用していない．また，骨格データのみを用いることで服装の違いによる影響も小さいと考えられる．そして，カメラを使用しているが顔などの画像をまったく扱わないことで，プライバシへの配慮をしている．歩行方向の違いによる見え方の違いに対しては，骨格データは 3 次元データとして取得できるので見え方を変換することで対応することを目標としている．

以降，2 章では，既存研究である八木らの歩容認証について紹介する．3 章では，マーカーレスモーションキャプチャ装置として使用した Kinect について紹介する．4 章では提案手法であるマーカーレスモーションキャプチャ装置を使用した歩容認証について示す．5 章では提案手法を用いた歩容認証の結果を示す．6 章では本稿のまとめと今後の課題について述べる．

## 2. 歩容認証

この章では，歩容認証 (Gait recognition) に関する既存の研究を述べる．現在いたるところに監視カメラが設置されるようになったが，監視カメラの画質ではカメラから遠く離れた人の顔を識別できないことが多い．このような場合に，歩容認証技術が使われる．八木らの研究 [4] では，歩容のシルエットのデータを歩行周期ごとに時間軸で離散フーリエ変換し，振幅スペクトルを計算し，画像特徴としている．0 次に姿勢や体形，1 次には左右非対称な動き，2 次には左右対称な動きが特徴として表れる．この方法では，被験者 143 名に対し，98.2 %の高い精度を示している．この方法では，同じ方向の辞書データを用意するか，歩容データの 3 次元モデルを必要とする．これは，ユーザが負担と感じる可能性がある．

八木らの認証 [4] の手順は次のようになる．なお，図 1 から図 5 は著者らが全方位ではないカメラを用いて行った再現画像である．

(1) 歩容シルエットの抽出
歩行者のシルエットを抽出するために，背景差分を使う．ただし，背景差分のみでは影などの影響を受けて，人物以外の領域も前景として抽出してしまう．そのため，影を除去をする．図 1(b) に図 1(a) からシルエットを抽出した例を示す．

(2) 位置あわせとスケーリング
比較するために，画像サイズが 20 × 30pixel のシルエット画像を作成する．人物領域が水平方向の中心に来るように位置あわせし，20 × 30pixel の画像に収縮する．位置あわせとスケーリングした画像を図 1(c) に示す．

(3) 歩行周期の算出
歩容特徴には，歩容の周期性を利用した周波数領域特徴を用いる．その際に必要となる歩行周期を算出する．

(4) 周波数領域特徴の抽出
次に歩行周期に基づく周波数領域の抽出方法について述べる．周期ごとに時間方向の離散フーリエ変換 (DFT) を計算する．高周波数の特徴には雑音が含まれるため，0，1，2Hz の特徴を抽出し，これがテストデータとなる．複数周期のすべてに同じ処理を施す．図 1(d) はフーリエ変換する前の 1 周期分の歩容のシルエットデータを示し，図 1(e) には図 1(d) をフーリエ変換した結果を示している．図 1(e) は，左から 0，1，2Hz となっている．

(5) 特徴の照合
辞書データとテストデータの差の大きさを複数周期のそれぞれで求める．各辞書データごとにこの処理をして，それぞれで複数周期分の中から中央値を使い，差がもっとも小さくなる辞書データの被験者を識別結果とする．中央値を使うのは外れ値による影響を小さくするためである．

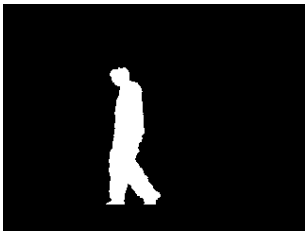
## 3. マーカーレスモーションキャプチャ装置

### 3.1 Kinect

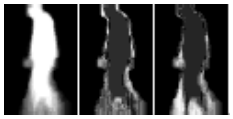
本研究では，マーカーレスモーションキャプチャ装置を使用した歩容データによる個人識別をする．現在市販されている低価格のモーションキャプチャができるデバイスとして Microsoft 社製の Kinect for Xbox360 と Kinect for

© 2012 Information Processing Society of Japan


(a) カラー画像


(b) シルエット抽出


(c) スケーリング


(e) フーリエ変換


(d) 1 周期分の歩容

**図 1** 既存研究

Windows, Asus 社製の Xtion Pro および Xtion Pro Live がある．本研究では，デバイスには Kinect for Xbox360 を用いた．Kinect は 2010 年に Microsoft 社から発売された Xbox360 用のコントローラである．名前は，動力学を意味する”Kinetics”と，つながりを表す”Connection”を組み合わせている．Kinect はカラーカメラ，深度センサ，マルチアレイマイクを持ち，マーカーの必要ないモーションキャプチャを可能とするデバイスである．

2011 年 6 月から公式に WindowsPC 向けの開発環境 Kinect for Windows SDK の beta 版を公開している．2012 年 2 月 1 日から Kinect for Windows が発売され，統合開発環境は Kinect for Windows SDK となった．Kinect for Windows は，Near モードと呼ばれる新しい機能が加わり，PC 用に使用することを目的としている．

開発環境には Kinect for Windows SDK beta 2 を用いた．Kinect for Windows SDK による開発には，Windows 7 以降のマイクロソフトの OS が必要で，開発言語として C++と C#が使用できる．また，ハードウェアは Dual-core2.66Ghz 以上の CPU，2GB 以上の RAM，USB 2.0 以上の USB が必要になる．

### 3.2 Kinect から取得できるデータ

Kinect から取得できるデータは大きく分けてカラー画像と深度データ，骨格データ，3 次元音声データの 4 つである．

- カラー画像
  640 × 480 pixel もしくは 1280 × 960 pixel のカラー画像を取得できる．640 × 480 pixel で取得した例を図 2 に示す．ただし，図 2 の顔部分はプライバシの問題のため後から処理をした．


**図 2** カラー画像 640 × 480

- 深度データ（プレイヤーデータ）
  320 × 240 pixel もしくは 640 × 480 pixel の深度データを取得できるが，プレイヤーデータと呼ばれる人のシルエットデータ取得時は 320 × 240 pixel しか取得できない．測定可能な深度は 800 mm から 4000 mm 程度である．深度データを図 3 に，プレイヤーデータを図 4 に示す．図 3 はカメラからの距離が大きくなるほど黒くなるようにしている．ただし，4000 mm 以上離れているところは白くなっている．
  プレイヤーデータは人のシルエットを取得でき，最大で 7 人まで識別できる．
- 骨格データ
  頭，肩中心，両肩，両肘，両手首，両手，脊柱，尻中心，両尻，両ひざ，両足首，両足の合計 20 箇所の部位で $(x, y, z)$ の 3 次元データを取得できる．骨格データは，同時に最大で 2 人まで取得できる．$x$，$y$ はそれぞれ画像の横方向，縦方向の座標で，どちらも 0 から 1 で出力される．$(x, y)$=(0, 0) なら画像の右上，(1, 1)

© 2012 Information Processing Society of Japan



**図 3** 深度画像 320 × 240



**図 4** プレイヤーデータ 320 × 240

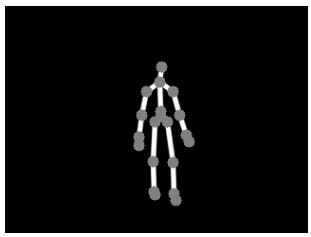


**図 5** 骨格データ

**表 1** 骨格データ

| skeleton position | x | y | z [mm] |
|---|---|---|---|
| HIP CENTER (尻 中心) | 0.4864 | 0.4813 | 3367 |
| SPINE (脊柱) | 0.4887 | 0.4606 | 3425 |
| SHOULDER CENTER (肩 中心) | 0.4905 | 0.3338 | 3445 |
| HEAD (頭) | 0.4834 | 0.2659 | 3392 |
| SHOULDER LEFT (肩 左) | 0.4454 | 0.3746 | 3452 |
| ELBOW LEFT (ひじ 左) | 0.4252 | 0.4789 | 3415 |
| WRIST LEFT (手首 左) | 0.4021 | 0.5650 | 3330 |
| HAND LEFT (手 左) | 0.3925 | 0.5954 | 3283 |
| SHOULDER RIGHT (肩 右) | 0.5342 | 0.3745 | 3411 |
| ELBOW RIGHT (ひじ 右) | 0.5491 | 0.4802 | 3402 |
| WRIST RIGHT (手首 右) | 0.5588 | 0.5756 | 3361 |
| HAND RIGHT (手 右) | 0.4383 | 0.8532 | 3273 |
| HIP LEFT (尻 左) | 0.4656 | 0.5075 | 3362 |
| KNEE LEFT (ひざ 左) | 0.5598 | 0.6136 | 3342 |
| ANKLE LEFT (足首 左) | 0.4435 | 0.8230 | 3317 |
| FOOT LEFT (足 左) | 0.4477 | 0.6874 | 3319 |
| HIP RIGHT (尻 右) | 0.5055 | 0.5085 | 3342 |
| KNEE RIGHT (ひざ 右) | 0.5132 | 0.6818 | 3300 |
| ANKLE RIGHT (足首 右) | 0.5098 | 0.8147 | 3279 |
| FOOT RIGHT (足 右) | 0.5071 | 0.8297 | 3182 |

なら左下となる．$z$ は深度データと同じく 800 mm から 4000 mm の範囲で出力される．
図 5 は表 1 のデータを基に OpenCV で示した画像である．

- 3 次元音声データ
マルチアレイマイクから取得できる音声データである．音の出た方向もわかる．本研究では用いない．

## 4. 歩容による個人識別法

### 4.1 概要

本研究は，Kinect により取得できる骨格データのみを利用し，個人を識別する．そして，個人識別には歩行中の各部位の周波数特徴を利用する．また，角度変換をすることで，あらゆる角度からの歩容にも対応することを目標とする．

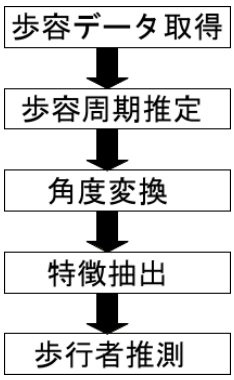



**図 6** 処理の流れ

八木ら [4] が歩容認証に用いる識別法との違いを述べる．八木らの識別法は，まずカメラ画像からシルエット情報を抽出し，1 周期分のシルエットをトリミングとリサイズして離散フーリエ変換をする．そして，カメラで撮影した複数周期分の識別対象画像すべてに対してこの処理を施し，辞書データと比較し，各辞書データごとに複数周期分の中の中央値を用いて，個人を識別する．また，カメラに映る歩行の方向が変わるごとに，辞書データを変えるか，数十台のカメラにより同時に撮影し，3D モデルを作成しておく．それに対し，本研究では，シルエット情報ではなく，Kinect から得られる骨格データを離散フーリエ変換する．ただし，Kinect は撮影できる範囲が狭いため，複数周期分の歩容をとることが難しい．そのため，1 周期分のみで識別する．カメラから見える体の向きが変わった場合は，Kinect では奥行きのデータも取得できるため，辞書データと同じ角度で歩いているように変換することで対応する．こうして，高価な装置や特別な場所を必要とせず複数の角度の歩容に対応することを目指す．処理の流れは図 6 のようになり，次の節からそれぞれについて説明する．

### 4.2 角度変換

本研究では，Kinect を使うことで得られる骨格データを使用するため，奥行きのデータを含めた 3 次元の情報がわかる．さらにシルエットを使用せず，取得したデータを辞書データと同じ方向の歩容データへと変換することができると考えられる．この節では，角度を変換する方法を 3 つの段階に分けて説明する．

変換する前の斜めに歩いた歩容データの例を図 7 に示す．右足，右腕をグレーで表し，それ以外の体の部分を白で表している．図 7 でグレーになっているのが，左足，左腕に見えるのは骨格データが左右反転した状態で取得されるためである．左から右へ行くごとにフレームが 5 ずつ進

© 2012 Information Processing Society of Japan

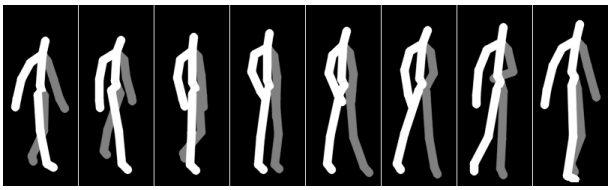


**図 7** 斜め歩容データ

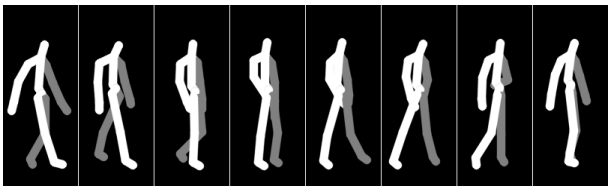


**図 9** 角度変換後の歩容データ

んでいる．

図 8 に，頭のポイントの動きを上から見た図を示す．

骨格データは，$x, y$ はそれぞれ 0 から 1 で，$z$ は mm オーダーで取得されるため，回転処理をする前にまず，スケールをあわせる必要がある．また，カメラの座標系は，人の目と同じで遠くになるほど見ることのできる範囲が広くなるので，近くのものが大きく，遠くのものが小さく見えるようになっている．変換するための式を式 (1) に示す．変換前の座標を $(x, y, z)$, 変換後の座標を $(x', y', z')$ とする．また，$(cx, cy)$ は画面の中心座標とする．

$$\begin{cases} x' = (x - cx) * (0.996 * z + 264.6) \\ y' = (y - cy) * (0.750 * z + 192.4) \\ z' = z \end{cases} \quad (1)$$

次に，歩行方向を辞書データと同じ方向に変換するために回転処理をする．$y$ 軸方向の回転処理は式 (2) を使う．変換前の座標を $(x', y', z')$, 変換後の座標を $(x'', y'', z'')$, 回転させる角度を $\theta$ とする．

$$\begin{pmatrix} x'' \\ y'' \\ z'' \end{pmatrix} = \begin{pmatrix} cos\theta & 0 & sin\theta \\ 0 & 1 & 0 \\ -sin\theta & 0 & cos\theta \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \quad (2)$$

最後に，辞書データがカメラの座標系になっているため，カメラの座標系に変換する．式 (1) を逆にしたものが変換式となる．以上の処理の結果が図 9 となる．処理結果である図 9 は足や左腕などが不自然に見えるが，識別に使うのは各点の周波数特徴のみなので影響ないと考えられる．

図 10 に処理後の頭のポイントの動きを上から見た図を示している．

### 4.3 特徴抽出

使用する特徴データには，周波数特徴のみを使用する．周波数特徴を抽出するために離散フーリエ変換する．フーリエ変換するために，1 周期分の歩容を見つける必要があるが，既存研究のように何周期分ものデータをとることができないため，周期性を見つけることができなかった．そのため本研究では，歩行状態の検出や，1 周期分の歩容を見つけることは手動でする．

離散フーリエ変換は式 (3) のようになる．$P$ はデータ数，$N$ はフレーム番号，$f_N$ は $N$ フレーム目の骨格データの値，$M$ が次数をそれぞれ表す．

**表 2** 実験環境

| | |
|---|---|
| CPU | Intel(R) Core(TM) i7-2600K CPU @ 3.40GHz |
| RAM | 8.00GB |
| OS | Windows 7 Professional Service Pack 1 (64bit) |
| 開発環境 | Microsoft Visual Studio 2010 Ultimate<br>Kinect for Windows SDK beta 2 |
| カメラ | Kinect for Xbox360 |

$$a_M = \frac{1}{P} \sum_{N=0}^{P-1} (f_N e^{-j(2\pi/P)MN}) \quad (3)$$

$M$ は 1 から 20 まで変化させ，$a_1$ から $a_{20}$ の 20 個の特徴とした．これらの特徴が，骨格データの 20 個の部位の 3 軸それぞれから求められるので，20 × 3 × 20=1200 個の特徴となる．

### 4.4 識別方法

抽出した特徴データを使って識別する方法として，機械学習を用いた．機械学習には，WEKA[8] を使用し，サポートベクタマシン (Support Vector Machine) により識別する．

## 5. 実験

この章では，歩容による個人識別法の検討のためにした，2 種類の実験について述べる．1 つ目は，図 11 に示すように，カメラの光軸に対して 90 度の歩容データのみを使用して，Kinect から得られるデータを利用して個人識別ができることを示す実験をした．2 つ目は，辞書データと異なる向きの歩容データに回転処理をすることによって，識別することができるかを確認するための実験をした．

### 5.1 実験環境

実験に用いた環境を表 2 に示す．

### 5.2 同じ向きの歩容による識別

まず，データの収集方法を述べる．各被験者にカメラから約 3m 離れたところを，図 11 のように横切っている歩容データを Kinect で撮影した．今回の実験に使用したカメラはフレームレートは約 30fps，シルエットデータの解像度は 320 × 240pixel で撮影した．歩容データの撮影は，1 日に 2 往復分取得した．これを繰り返し，1 人当たり 9 から 22 回分の歩容データを取得した．日を分けて撮影した

© 2012 Information Processing Society of Japan

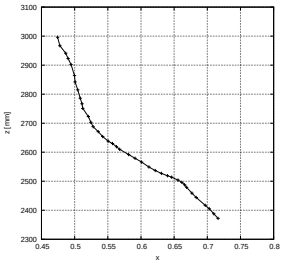


**図 8** 斜め歩容データ 頭の $(x, z)$ 座標

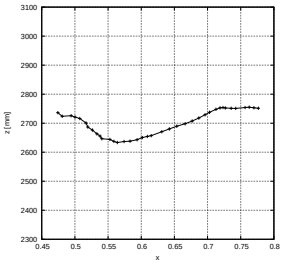


**図 10** 変換後の歩容データ 頭の $(x, z)$ 座標

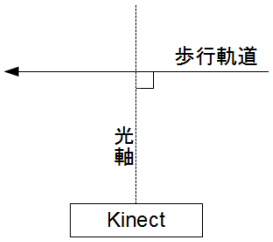



**図 11** 90 度の歩容

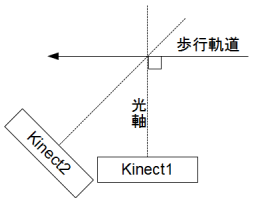



**図 12** 45, 90 度の歩容

**表 3** 同じ向きの歩容による識別結果

| | データ数 | 正答数 | 精度 [%] |
|---|---|---|---|
| 被験者 1 | 20 | 18 | 90.0 |
| 被験者 2 | 20 | 19 | 95.0 |
| 被験者 3 | 22 | 20 | 90.9 |
| 被験者 4 | 20 | 13 | 65.0 |
| 被験者 5 | 18 | 15 | 83.3 |
| 被験者 6 | 14 | 13 | 92.9 |
| 被験者 7 | 17 | 15 | 88.2 |
| 被験者 8 | 12 | 6 | 50.0 |
| 被験者 9 | 9 | 7 | 77.8 |
| 被験者 10 | 13 | 10 | 76.9 |
| | 165 | 136 | 82.4 |

のは，服装の違いや日ごとの被験者の状態の違いを考慮するためである．10 人の被験者から歩容データを収集した．

集めたデータから抽出した特徴を使い，サポートベクタマシンで 10 分割交差検定により精度を調べた．識別結果を表 3 に示す．

10 人の被験者の歩容により識別をした結果，精度は 82.4%となった．この識別精度は，テレビのリモコンにつけた加速度センサによる識別 [9] で 2 人から 4 人の家族で 70%から 92%の識別精度で十分であると評価されていることから，テレビの番組推薦などに使う場合に必要な精度を満たしていると考えられる．また，データ数の少ない被験者の精度がよくないので，データ数を増やすことにより更なる精度の向上が期待できる．

### 5.3 異なる向きの歩容による識別

次に，辞書と異なる方向の歩容に関する実験について述べる．はじめに，カメラの光軸に対して 45 度の角度の歩容を撮影し，提案手法により光軸に対し 90 度の歩容になるように変換し識別した．斜めに歩いていたデータを変換した識別結果を表 4 に示す．辞書データには，5.2 節で収集したカメラの前を横切ったデータの被験者 1 から 4 を使用し，テストデータには斜めに歩いたデータを変換したものを使用した．被験者数は 4 人で，1 人あたり 4 回分のデータを取得した．

**表 4** 識別結果

| | | 識別結果 | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 識別精度 [%] |
| 実 | 被験者 1 | 0 | 3 | 0 | 1 | 0.0 |
| 被 | 被験者 2 | 0 | 4 | 0 | 0 | 100.0 |
| 験 | 被験者 3 | 0 | 0 | 0 | 4 | 0.0 |
| 者 | 被験者 4 | 0 | 1 | 0 | 3 | 75.0 |
| | | | | | | 43.8 |

識別の精度は 43.8 %となった．この識別精度はかなり低く，正しく識別されていないと考えられる．

### 5.4 異なる向きの歩容についての軌道評価

変換が正しく行われているか確かめるため，2 台の Kinect を使い 1 つの歩容を図 12 のように，角度を変えて同時に撮影し提案手法により変換して，軌道を比較した．

データを集めるために 2 台の Kinect を使用し，30, 45, 60, 75, 90 度の歩容をそれぞれ 2 つずつ同時に撮影した．フレームレートは約 20fps とし，同期は取らなかった．

2 台の Kinect により同時に撮影された同じ部位の 2 本の軌道を評価するために，relative root mean square distance(RRMSD)[10] を使用する．評価する 2 つの軌道のリストを $P : \{p_i\}, Q : \{q_i\}, i = 1, ..., N$ とする．

RRMSD を求める前に，P と Q の位置を調整する必要がある．調整するための式を式 (4) に示す．ここで，変換前の P の座標が $(p_{ix}, p_{iy}, p_{iz})$，P の変換後の座標が $(p'_{ix}, p'_{iy}, p'_{iz})$，リスト P の平均が $(\overline{p_x}, \overline{p_y}, \overline{p_z})$，リスト Q の平均が $(\overline{q_x}, \overline{q_y}, \overline{q_z})$ とする．

$$\begin{aligned} p'_{ix} &= p_{ix} + \overline{q_x} - \overline{p_x} \\ p'_{iy} &= p_{iy} + \overline{q_y} - \overline{p_y} \\ p'_{iz} &= p_{iz} + \overline{q_z} - \overline{p_z} \end{aligned} \tag{4}$$

RRMSD を求めるための式を式 (5) に示す．

© 2012 Information Processing Society of Japan

表 5 RRMSD

| 角度 | $x$ | $y$ | $z$ | RRMSD |
|---|---|---|---|---|
| 90-75 | 42.1 | 35.2 | 37.5 | 0.073 |
| 90-45 | 41.6 | 41.2 | 53.1 | 0.081 |
| 75-75 | 39.6 | 30.4 | 31.1 | 0.057 |
| 75-60 | 36.7 | 33.8 | 31.3 | 0.059 |
| 75-45 | 37.7 | 28.3 | 32.5 | 0.062 |
| 75-30 | 37.5 | 27.4 | 31.6 | 0.062 |
| 45-45 | 39.3 | 27.3 | 29.6 | 0.057 |

$$RRMSD = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}||p_i - q_i||^2}}{Diag\,(P \cup Q)} \tag{5}$$

$$Diag(P \cup Q) = \sqrt{(\max_j(r_{jx}) - \min_j(r_{jx}))^2 + (\max_j(r_{jy}) - \min_j(r_{jy}))^2 + \overline{(\max_j(r_{jz}) - \min_j(r_{jz}))^2}} \tag{6}$$

$r_j$ は，$P \cup Q : \{r_j\}, j = 1, ..., 2N$ で，$r_{xj}$ と $r_{yj}, r_{zj}$ は，$r_j$ の $x, y, z$ の値である．

各データを，光軸に対して 90 度の歩容に見えるように変換した歩行軌道を使い，RRMSD を求めた結果を表 5 に示す．表 5 の 90-75 などの角度は，90 度と 75 度で同時に撮影したデータを意味する．

表 5 の $x$, $y$, $z$ は式 (5) の分子だけを $x$, $y$, $z$ のそれぞれで求めたものである．また，75-75 や 45-45 は，同じ向きの歩容でどのくらい RRMSD に違いが生まれるのかを示している．

表 5 より，2 つの角度差が大きくなるほどに RRMSD の値が大きくなっていることがわかる．その中でも，90 度と 45 度が特に大きく，90 度にかかわるものが全体的に大きいことがわかる．そこで，90 度で撮影された骨格データを見てみると，他の角度で撮影したものに比べて正しい位置から外れやすいことがわかった．そのため，辞書データに 90 度の向きの歩容データが向いていない可能性があることがわかった．

## 6. おわりに

本研究では，マーカーレスでモーションキャプチャが可能な装置である Kinect を用いた歩容による個人識別法を検討した．検討した手法では，特徴抽出のために骨格データを離散フーリエ変換する．また，辞書データと歩行方向が異なる場合の識別には，骨格データを回転処理して辞書データと同じ方向の歩容データとする．そして，抽出された特徴を機械学習のサポートベクタマシンを利用して識別する．

識別結果は被験者数 10 人とし，1 人あたりのデータ数が 9 個から 22 個で，合計 165 個のデータを使って評価したところ，識別精度の平均は 82.4 %となった．そして，データ数の少ない被験者の精度が低い傾向がみられた．この結果は，テレビなどの家電が個人に合わせたサービスを提供するための個人識別機能として使うのに十分な精度であると考えられる．

また，辞書データと異なる向きでの歩容での識別として，光軸に対して 45 度の歩容データを取得して変換したが，識別結果は 43.8%とかなり低い精度となった．そのため，辞書データと異なる向きでの歩容による識別に対応するための調査として，異なる 2 つの角度で同時に撮影された歩行軌道を取得し，比較，評価した．それにより，90 度の歩容データで特に差が大きいことがわかった．そのため，他方向の歩容データによる識別をする場合には，辞書データの向きを 90 度以外の角度にすることによって，識別精度が上がると期待できる．

今後の課題として，まずカメラの光軸に対して 90 度でない歩容データの収集とそれを用いた評価があげられる．そして，そのデータを使用して，辞書データを作り精度の向上が見られるか，回転処理によって辞書データと異なる歩行方向でも個人識別ができるかを確認する．また，システム化するために，取得したデータから歩いていることを検知することと，1 周期分の歩行を見つけることの 2 つを自動化する必要がある．

### 参考文献

[1] Tivo, http://www.tivo.com/, (2012 年 6 月確認).
[2] Orca Interactive, http://www.orcainteractive.com/, (2012 年 6 月確認).
[3] Kinect，http://www.xbox.com/ja-JP/kinect, (2012 年 6 月確認).
[4] 杉浦一成, 槇原靖, 八木康史: 全方位カメラを用いた複数方向観測による歩容認証. 情報処理学会研究報告, vol.2007-CVIM-159, pp.187–194 (2007).
[5] 白石明, 槇原靖, 八木康史: 三次元歩容データを用いた任意視点変換モデルによる歩容認証. 情報処理学会研究報告, vol.2011-CVIM-175, no.34, pp.1–8, 2011.
[6] Kinect for Windows, 入手先 ⟨http://www.microsoft.com/en-us/kinectforwindows/⟩, (2012 年 6 月確認).
[7] Kinect wiki, http://www.kinect-wiki.info/, (2012 年 6 月確認).
[8] weka, 入手先 ⟨http://www.cs.waikato.ac.nz/ml/weka/⟩, (2012 年 6 月確認).
[9] Keng-hao Chang, Jeffrey Hightower, Branislav Kveton : Inferring Identity using Accelerometers in Television Remote Controls. Proceedings of 7th International Conference on Pervasive Computing (Pervasive 2007), pp.151–167 (2007).
[10] Frederic Jean, Robert Bergevin and Alexandra Branzan-Albu: Computing and Evaluating View-normalized Body Part Trajectories. Image and Vision Computing, vol.27, No.9, pp.1272–1284, August 2009.

© 2012 Information Processing Society of Japan