研究速報

Kinect からの歩行動作による個人識別

Personal Identification Based on Skeletal Movement Using Kinect

宮島 春菜^{†*}, 正会員 山本 正信[†]Haruna Miyajima^{†*} and Masanobu Yamamoto[†]

Abstract The Microsoft Kinect motion sensing device can capture the skeletal movement of a pedestrian. A skeleton in motion draws an orbit in high dimensional pose parameter space. The Karhunen-Loève expansion can project this high dimensional orbit into two-dimensional eigenspace. A method is proposed to differentiate pedestrians by using their respective traces in the eigenspace. This uses the nearest neighbor rule to identify pedestrians on the basis of the cosine similarity between the traces. In our experiments, this method gave a personal identification rate of 34 pedestrians.

キーワード：キネクト，歩行，個人識別，固有空間

1. ま え が き

歩行は個人認証のための有力な情報源である。これまでの歩行認証はビデオカメラからの映像を利用することが多かった。しかし、ビデオ映像はカメラと歩行者との相対的な位置・姿勢関係に依存するため、相対的關係が変われば認証手法の前提条件が成り立たなくなる。

歩行者を骨格モデルで表すことができれば、身体の姿勢は各部位の上位部位に対する相対角度で表すことができるため、センサと歩行者との相対的關係に依存しない認証方法が可能である。近年、Kinect センサの登場により身体の3次元骨格モデルが容易に得られるようになった。Kinectは赤外線センサにより、距離画像を経由して骨格モデルを得るため、センサの視点や歩行者の衣服に依存せず、夜間でも個人認証が可能と考えられる。

骨格モデルには、身体の構造情報と運動情報が含まれている。これまで提案されてきた手法は、もっぱら身体寸法などの構造情報を特徴として利用するものが多かった¹⁾²⁾⁴⁾⁶⁾。運動情報は歩幅や速度⁴⁾、胴体の中心から四肢への距離⁶⁾などの利用が試みられているが、これらは全身の運動情報の一部である。これに対し、Munsell³⁾らは、歩行中の全身の姿勢角を照合させることにより認証を行っている。時間軸上の照合は、照合開始時刻を探索過程に加える必要があ

り、計算に手間がかかる。DTWなどの計算法もあるが、Munsell³⁾らは、DCT係数を使った周波数領域での照合を行っている。しかし、同じ個人の歩行でも、加速や減速による歩容の揺らぎがあれば、DCT係数のシフトを招き認証性能に悪影響を及ぼす。

骨格モデルの運動は、姿勢パラメータ空間内で軌道を描く。歩行は周期運動であるのでこの軌道は閉じた軌道となる。歩行の加速や減速などの揺らぎは、この軌道内の位置の変化に過ぎない。したがって、この軌道がパラメータ空間内の図形として表されれば、歩行による個人認証は図形の照合問題となる。

身体動作のパラメータ空間は広大であるため、OhnoとYamamoto⁵⁾は、KL展開により2次元のパラメータ空間に射影し、2次元図形化された軌道の比較により、ラジオ体操の動作識別を行った。本論文では、同様に歩行の軌道を2次元パラメータ空間に射影し、射影された軌道の照合により個人識別を行う。

身体寸法を用いた個人識別では、90%前後以上の高い識別率が示されている。ただし、歩行者は10名程度であった。本論文では、提案手法が先行研究と同程度の識別率を確保できるか否かの確認と、10名以上の歩行者に対する識別率の評価を行う。

2. Kinect による歩行動作の測定

歩行者はKinectに正面を向け、カメラ奥から手前へ歩行した。Kinectの測定範囲は、0.8 m～4mであり、歩数にして4～5歩である。Kinectによるモーションキャプチャでは、図1左のように頭部や手先、足先のような端点を含む20個の関節位置を測定することができる。このうち、頭

2013年7月26日受付，2013年8月27日再受付，2013年9月18日採録

†新潟大学 工学部 情報工学科

(〒950-2181 新潟県新潟市西区五十嵐二の町 8050, TEL 025-262-7428)

* キヤノンイメージングシステムズ (株)

† Faculty of Engineering Niigata University

(8050, Ninotyou, Ikarashi, Nishi-ku, Niigata-shi, Niigata 950-2181, Japan)

* CANON Imaging Systems Inc.

図 1 Kinect で測定した骨格 (左: 全骨格, 右: 使用した骨格)
Skeleton from Kinect (Left: full skeleton, Right: skeleton
used in this paper).

部と足先は、身体ができるだけ Kinect に接近するまで測定したとき、視野から外れることが多いため使用しなかった。また、身体が Kinect から遠ざかったとき、手は明瞭に捉えられないため手先も使用しなかった。図 1 右が使用した関節である。この図には 15 個の関節と、関節を端点とする 14 本の線分が示されている。この線分は身体部位の体軸である。

関節で繋がった身体部位は、身体末端に近い部位を子部位、身体中心に近い部位を親部位とする。身体部位にはそれぞれ関節位置を原点とする固有の座標系が付けられており、部位の姿勢は親部位に対する座標系の姿勢で表す。Kinect for Windows SDK v1.5 から、各部位の姿勢が回転行列として出力できるようになった。各部位の体軸は座標軸の z 軸としているので、回転行列の 3 列目が体軸ベクトルである。体軸ベクトルの要素を部位の順に並べて構成されたベクトルを姿勢ベクトルとよび、

$$\mathbf{q}_t = (q_{t1}, q_{t2}, \dots, q_{tm})^T \quad (1)$$

と表す。ここで、 t は時刻、 m はパラメータ数 (本研究では $m = \text{使用する体軸の数} \times 3 \text{次元} = 42$) を表す。Kinect は毎秒 30 回のビデオレートで姿勢を測定できるので、 t をフレームともよぶ。

3. 歩行動作の固有空間

姿勢ベクトルの時系列、すなわち歩行動作は姿勢パラメータ空間内に軌道を描く。軌道の比較により個人認証を行うために、軌道をパラメータ空間内に記録する。パラメータ空間をメッシュに区分けし、メッシュを投票箱とみなし、軌道がメッシュを通過すれば 1 票を投じる。本研究では姿勢ベクトルは 42 次元であるため、パラメータ空間を 1 次元あたり 40 区画に量子化するとすれば、全空間では 40^{42} 個のメッシュを必要とする。この規模のメッシュは現実的ではないので、KL 展開を使用し、軌道の特徴を損なわずに低次元のパラメータ空間に射影する。

姿勢ベクトルの平均値

$$\mathbf{c} = \frac{1}{N} \sum_{t=1}^N \mathbf{q}_t \quad (2)$$

図 2 固有空間上の歩行動作軌道 (左), 補間した歩行動作軌道 (右)

Trace of gait on eigenspace (Left), trace after interpolation (Right).

を求める。ここで N は歩行時間を表すフレーム数。

姿勢ベクトルからその平均値 $\mathbf{c}$ を差し引いたベクトルから、行列 Q を構成する。

$$Q = (\mathbf{q}_1 - \mathbf{c}, \mathbf{q}_2 - \mathbf{c}, \dots, \mathbf{q}_N - \mathbf{c}) \quad (3)$$

共分散行列 QQ^T の固有値・単位固有ベクトルを求め、固有値の大きい順から $k (\leq m)$ 個の固有値に対する単位固有ベクトルを $\mathbf{e}_i (i = 1, 2, \dots, k)$ とする。行列 E を

$$E = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k) \quad (4)$$

とし、 E を使って姿勢 $\mathbf{q}_t$ を低次元空間の姿勢 $\mathbf{r}_t$ に射影する。この低次元空間は固有ベクトルの張る空間であるので固有空間とよぶ。

$$\mathbf{r}_t = E^T \mathbf{q}_t \quad (5)$$

本論文では、2 次元固有空間 (以下、固有平面とよぶ) に射影する。各軸を 40 区画に区分けするとすれば、 40×40 個のメッシュが得られ、軌道は固有平面上の画像に図形として描かれる。

図 2 左に画像上に投影された軌跡を示す。軌道が通過した画素がドットで示されている。Kinect のサンプリングレートに比べて動作が速いと、隣り合ったフレームでの姿勢が画像上で離れた位置に投影される。本来、軌道は連続して描かれるはずであるので、2 フレーム間の姿勢の投影が 8 近傍でない画素間は軌道を補間する。線形補間により補間された軌道を図 2 右に示す。

4. 個人識別

固有平面上の画像に描かれた歩行動作軌道を使って個人の識別を行う。

歩行動作からの個人認証法の実用化を想定したとき、教示用の歩行データを多くは用意できない。センサの前を 1 ～ 2 度歩行したデータを教示用歩行とすることになる。したがって、この歩行をプロトタイプとした最近傍決定法による識別が現実的と考えられる。

個人クラスのプロトタイプを $\mathbf{p}_i$ とする。 $\mathbf{p}_i$ は画像上の歩行動作軌道であり、 i は個人クラスを表す。 $\mathbf{x}$ を同じく画像上に描かれた未知の歩行軌道としたとき、識別は、コサ

イン類似度を使って,

$$k = \arg \left(\max_i \frac{\mathbf{x} \cdot \mathbf{p}_i}{\|\mathbf{x}\| \|\mathbf{p}_i\|} \right) \quad (6)$$

として得られた k が歩行者 $\mathbf{x}$ の識別名となる.

5. 実験

34名のボランティアに協力してもらい歩行データを得た. 内訳は男性32名, 女性2名であり, 年齢は18歳から23歳までの青年である. Kinect センサに向かって3回歩行し, このうちの2回を教示用歩行, 残りの1回をテスト用歩行とした.

まず, 体格や性別に偏らない姿勢の特徴空間を得るために, 34名の中から小柄な男性と女性を1名ずつ, 大柄な男性と女性を1名ずつ, 計4名を選んだ. ただし, 2名の女性は男性ほどの体格の大きな差はなかった. 彼らの歩行から固有平面への変換行列 E を求めた. 累積寄与率は0.488

表 1 100% の識別率を達成した混同小行列
Confusion submatrix showing complete identification.

	25	26	27	28	29	30	31	32	33	34
25	0.48	0.15	0.30	0.30	0.21	0.37	0.29	0.21	0.25	0.25
26	0.29	0.36	0.25	0.19	0.33	0.30	0.35	0.31	0.15	0.17
27	0.44	0.43	0.51	0.27	0.29	0.26	0.25	0.29	0.30	0.21
28	0.32	0.25	0.34	0.42	0.18	0.27	0.28	0.12	0.23	0.14
29	0.35	0.20	0.26	0.20	0.42	0.26	0.29	0.21	0.26	0.29
30	0.35	0.17	0.24	0.38	0.23	0.43	0.30	0.10	0.17	0.19
31	0.22	0.15	0.25	0.30	0.20	0.30	0.34	0.20	0.30	0.20
32	0.27	0.26	0.23	0.10	0.33	0.33	0.26	0.43	0.16	0.16
33	0.25	0.18	0.25	0.27	0.20	0.17	0.25	0.19	0.52	0.17
34	0.23	0.25	0.19	0.24	0.14	0.21	0.27	0.14	0.24	0.43

表 2 50% の識別率に終わった混同小行列
Confusion submatrix resulting into low identification rate.

	12	13	14	15	16	17	18	19	20	21
12	0.3	0.26	0.3	0.32	0.27	0.25	0.24	0.17	0.36	0.31
13	0.21	0.41	0.27	0.32	0.24	0.2	0.19	0.18	0.28	0.16
14	0.19	0.27	0.46	0.39	0.24	0.41	0.24	0.29	0.34	0.25
15	0.23	0.17	0.37	0.48	0.38	0.23	0.25	0.29	0.28	0.28
16	0.15	0.15	0.3	0.4	0.32	0.21	0.29	0.12	0.23	0.13
17	0.1	0.1	0.22	0.17	0.2	0.24	0.16	0.27	0.28	0.39
18	0.15	0.13	0.19	0.2	0.25	0.15	0.29	0.17	0.1	0.21
19	0.16	0.12	0.18	0.14	0.08	0.11	0.07	0.31	0.11	0.13
20	0.29	0.27	0.26	0.26	0.18	0.14	0.21	0.16	0.29	0.29
21	0.25	0.2	0.4	0.3	0.29	0.31	0.25	0.34	0.33	0.31

図 3 34名の混同行列と n 名の混同小行列
 $n \times n$ confusion submatrices from 34×34 confusion matrix.

であった. 彼らの歩行を変換行列により固有平面に射影し, それぞれの次元の最大値と最小値を求め, 固有平面をこの範囲内で 40×40 メッシュに画像化した.

個々の歩行者について, 2回の教示用歩行を固有平面上に重ねて描き, プロトタイプを作成する. 残りの1回をテスト用歩行とし, 軌道をテスト用画像に描く. テスト用画像とプロトタイプとのコサイン類似度を求める. その結果を混同行列 (confusion matrix) に記載する. 34名の 34×34 の混同行列から10名分を抜き出した混同小行列を表1に示す. 行番号はテスト用の個人名, 列番号はプロトタイプの個人名である. 各行ごとに類似度が最大となる列名がテスト歩行者の識別名となる. 最大値を太字で示している. この10名は100%の識別率を達成し, 本手法は先行研究の手法と比べて遜色がない.

しかし, 他の10名を選んだときは識別率は50%になった. このときの混同小行列を表2に示す. 識別率は個人の選び方により偏る可能性がある. この偏りは, 識別する人数が少ないほど大きくなる. 例えば, 34名の識別を行ったとき, 1名の誤識別があれば, $1/34$ の識別率低下で済むが, 2名の識別を行ったとき, 1名の誤識別があれば識別率は $1/2$ も低下する.

そこで, 同じ人数ごとに幾つかのグループについて識別率を求め平均化する. n 人のグループは, 1番から n 番の被験者について, 混同小行列を34名の混同行列から抜き出し識別率を求める. 次に2番から $n+1$ 番までの被験者について, 混同小行列を抜き出し同様に識別率を求め, この操作を $33-n$ 番から34番までの被験者の混同小行列を抜き出すまで続ける. 図3に34名の混同行列から抜き出した n 人の混同小行列を図示する. この結果, $35-n$ 個の識別率が得られそれらの平均識別率を n 名の識別率とする. 識別人数が少ないほどグループ数が多くなり, 平均操作により偏りが少なくなると考えられる.

歩行者人数に対する識別率を図4に示す. 横軸は人数, 縦軸は識別率である. 各自3回の歩行のうち最初の歩行をテスト歩行としたときの識別率を $walk\#1$, 2番目の歩行に対しては $walk\#2$, 3番目の歩行に対しては $walk\#3$, 全ての歩行の平均識別率を $average$ で示す.

先行研究の手法¹⁾²⁾⁴⁾⁶⁾が, 10名程度で90%以上の高い識別率を維持しているのに比べ, 本手法の性能は良くない. 提案手法の識別率を向上させるための方法を, 実験結果から考察する.

固有平面を構成した際の累積寄与率0.488は高くない. 累積寄与率が0.9を超えるのは第9固有値以降であった. したがって, 歩行の特徴を精密に捉えるためには, 少なくとも9次元の固有空間が必要である. しかし, 1次元あたり40区画に量子化するとすれば, 9次元では 40^9 のメッシュを必要とする. これでもまだ現実的な大きさではない. そこで, 第 $2i-1$ 第 $2i$ 固有ベクトルの張る固有平面を $i=1, 2, \dots, 5$ として5枚構成し, 本実験で使用した固有平面 ($i=1$) を

図 4 歩行人数と識別率
Identification rate of pedestrians.

含めて残りの固有平面の利用法を考えれば良い。

一方、表 2 では、識別に失敗した歩行者 12 のコサイン類似度は、4 番目の高さであった。同様に、歩行者 16, 17, 20, 21 のコサイン類似度は、それぞれ 2, 4, 2, 3 番目の高さである。このことは、識別を一旦保留し、上位 4 番目以上の候補から、例えば残りの固有平面を利用するなどして候補を絞ることも考えられる。

6. む す び

Kinect から得られる骨格モデルを使い、歩行動作から個人識別を行った。提案手法は、多次元姿勢パラメータ空間内の歩行軌道を、固有平面に KL 展開を利用して射影した軌道を利用するものである。

今後の課題として、動作特徴に身体寸法などの構造特徴を加えたマルチモーダルな識別法も必要となろう。

〔文 献〕

- 1) R. M. Araujo, G. Graña and V. Andersson: "Towards skeleton biometric identification using the microsoft Kinect sensor", SAC'13 28th Annual ACM Symposium on Applied Computing, pp.21-26 (2013)
- 2) 真部雄介, 齋藤隆輝, 嶋田弦, 菅原研: "歩行・顔・身体ソフトバイオメトリック特徴を用いた正面観測個人認証", 知能と情報, 24, 5, pp.988-1001 (2012)
- 3) B. C. Munsell, A. Temlyakov, C. Qu and S. Wang: "Person identification using full-body motion and anthropometric biometrics from Kinect videos", ECCV'12 12th European Conference on Computer Vision - Part III, pp.91-100 (2012)
- 4) J. Preis, M. Kessel, M. Werner and C. Linnhoff-Popien: "Gait Recognition with Kinect", 1st International Workshop on Kinect in Pervasive Computing, Newcastle (2012)
- 5) H. Ohno and M. Yamamoto: "Gesture recognition using character recognition techniques on two-dimensional eigenspace", ICCV'99 7th International Conference on Computer Vision, pp.151-156 (1999)

- 6) A. Sinha, K. Chakravarty and B. Bhowmick: "Person identification using skeleton information from Kinect", ACHI'13 6th International Conference on Advances in Computer-Human Interactions, pp.101-108 (2013)

みやじま はるな
宮島 春菜 2013 年、新潟大学工学部情報工学科卒業。同年、ギャノンイメージングシステムズ (株) 入社。在学中はパターン認識の研究に従事。

やまもと まさのぶ
山本 正信 1973 年、九州工業大学工学部制御工学科卒業。1975 年、東京工業大学大学院修士課程修了。同年、電総研 (現産総研) 入所。動画像処理、コンピュータビジョン等の研究に従事。1989 年～1990 年、カナダ国立研究協議会招聘研究員。現在、新潟大学工学部情報工学科教授。工博。正会員。