Computer Security Symposium 2017
23 - 25 October 2017

深度センサによる歩容特徴量を用いた個人識別・追跡方式の提案

森 駿文[1,a)]　菊池 浩明[1,b)]

概要：現在、歩容識別の研究がいくつかなされているが、既存の研究では認証としての用途しか考えられておらず、識別精度を上げる方向でしか研究がなされてこなかった。しかし歩容情報は指紋などの多くの生体情報と異なり対象者が意識することなくデータが収集できるといった特徴がある。歩容情報は人の動線推定などに活用できると考えられるが、対象者の同意なくデータ収集を行うためにはプライバシーに関わる情報を極力排除したり意図的に識別精度を下げることが望ましい。そこで、本研究では Microsoft 社の Kinect を用いて人物のスケルトンデータを取得し、使用するデータにより歩容識別の精度がどれほど変わるかを検証した。

キーワード：歩容識別，Kinect, プライバシー

Takafumi Mori[1,a)]　Hiroaki Kikuchi[1,b)]

1. はじめに

近年，不正アクセス防止などのため高度で安全な個人認証の必要性が高まっている．しかし，パスワードなどの記憶情報を用いた古典的な認証方式では，秘密情報をユーザが忘れてしまったり第三者に漏洩してしまう危険性がある．そこで，個人のユニークな生体情報を利用したバイオメトリクスが広まってきている．本研究ではその中でも人の歩行の様子を用いて認証を行う歩容認証のに焦点をあてる．

2017 年 5 月に個人情報が改正され，生体情報の特徴量などは個人識別符号とみなされ，厳格な安全管理措置が必要となっている．加えて，カメラ画像などにより，顧客の追跡や動線を解析し，店舗内の顧客の流れを特定することで商品の配置を変更するなどビッグデータとしての活用も可能であると考えられる．

歩容認証にはコンピュータビジョンを用いたもの [1] やウェアラブルセンサーを用いたものなどがある．コンピュータビジョンを用いた方式は他の認証方式と異なり，対象者から離れた位置にいても認証行為が可能であるという特徴があり，広域の監視が期待されている．また離れた位置から観測が可能であるという特徴から，対象者に意識させることなく一方的な観測が可能である．これらの特徴により我々は，歩容情報は既存の認証としての用途だけでなく個人の識別に活用できるのではないかと考えている．人物識別は，求められる精度はさほど高くない代わりにプライバシーに配慮しなければならないなどといった認証と異なる特徴がある．

よって本研究では，識別の手段として歩容情報を検討する．Microsoft 社の Kinect v2 を用いて数名の歩容情報を取得し，それらの値の分布を基にして識別に有効な特徴量は何かを明らかにする．

結果として，本論文で定義した特徴量のうち使用するものを変更することで EER が 0.25 から 0.50 まで変化することがわかった．

2. 識別と認証の違い

利用者認証とは事前にシステムに登録しておいたユーザ情報と対象者が同一人物であるかを判定することである．これに対して識別とは別時刻の 2 つの対象が同一であるかを判定することである．認証は主に，システムにログオンする際や建物に入る際などに正当な利用者であることを証明することに使われているが，識別の場合は定点センサを

1 明治大学大学院先端数理科学研究科
Meiji University Graduate School of Advanced Mathematical Science
a) cs172059@meiji.ac.jp
b) kikn@meiji.ac.jp

© 2017 Information Processing Society of Japan

表 1　識別と認証の違い

	人物識別	生体認証
対象者の協力	なし	あり
求められる精度	低い	高い
マッチング人数	$m:n$	$1:n$
プライバシー配慮	必要	不要
オプトアウト可能	必要	不要
攻撃パターン	自分ではない誰かになりすます	特定の人物になりすます

設置しておき，同じ人がいつ何回通ったかなどの統計情報を取得することができる．

この 2 つの間には次のような差がある．

- 認証は必ず対象者の意思が必要となるため本人の協力はあるが，識別は対象者の意図と無関係に行われる．識別結果を統計情報として扱う場合は多少の誤差は許容できるため，生体認証ほどの精度は求められない．
- 人物識別においては対象者の意思に関係なく個人の特徴量を取得することになるため，プライバシーへの配慮が必要．
- 情報の提供を拒否したい人はオプトアウトできることが望ましい．
- 認証は一人の対象者がテンプレートに登録されている人のうちどれに当てはまるかを特定するという問題なので，テンプレートが登録されている人数を n とすると $1:n$ のマッチングとなる．識別では同時に現れる複数人が，過去にデータベースに登録されたどの人と一致するか，もしくは登録されていない新規の人物であるのかを特定するという問題になる．従って，現在写っている人数を m，テンプレートに登録されている人数を n とすると，対象は m で，クラスタリング候補は登録されているテンプレート n との $m:n$ のマッチングとなる．そして新規の対象が来た場合は n が増加する．
- 認証の脅威は特定の人物になりすまし，システムを不正利用する攻撃である．識別の脅威は自分がその場にいることを知られたくない人が，誰でもいいので他人になりすます攻撃が考えられる．

以上の人物識別と生体認証の違いについて，**表 1** にまとめた．

3. 関連研究

歩容認証に用いる特徴量としてはシルエット画像がよく使われる．中でも，Han らが 2006 年に提案した歩容エネルギー画像 (GEI: Gait Energy Image)[3] がある．これは歩行のシルエット画像の 1 周期分の平均をとるというものであり，1 周期の全フレームの画像を比較する場合と比べ，処理の高速化や使用ストレージの削減，ノイズへの耐性などが期待できる

これを用いた研究の一つとして，Shiraga らの GEINet[1] がある．彼らは近年画像処理の分野において注目されている畳み込みニューラルネットワーク (CNN: Convolutional Neural Network) を用いて，GEI 画像を分類することで高精度な歩容認証を可能とした．

シルエット画像以外の歩容情報として，2015 年の Anderson らによる，Kinect により取得した骨格情報を用いた歩容認証の手法 [4] や，2013 年の Igual らによる，深度情報を用いた性別推定手法 [5] がある．

しかしこれらは全て認証や属性推定が目的であり，人物識別という観点で行われた研究は存在しない．

4. 実験

4.1 実験目的

本研究の目的を以下に示す．

- 人物識別に利用可能な歩容に関する特徴量を検討すること
- 使用する特徴量を変えることによりどの程度の識別率になるかを調査すること
- 使用した特徴量の安定性や他人との差異が十分にあるかを調査すること

4.2 実験機材

4.2.1 Kinect v2

Kinect v2 とは，Microsoft 社から発売されたモーションキャプチャデバイスである．これを用いることでユーザの身体の動きによりゲーム中におけるプレイヤーの操作を行うことを実現した．NUI(Natural User Interface) と呼ばれるインタフェースの一つである [8]．

Kinect には RGB カメラ，深度センサ，マイクなどが搭載されており，体の関節の 3 次元座標を推定し，その変化に基づいて人の動きを認識する．関節の座標データはスケルトンデータと呼ばれており，Kinect SDK を用いて参照することができる．Kinect v2 の仕様は**表 2**[7] の通りである．Kinect v2 により推定できる関節の情報を**図 1** に示す．

表 2　Kinect v2 のスペック

	スペック
RGB 解像度	1920×1080
RGB データのフレームレート	30fps
深度 解像度	512×424
深度データのフレームレート	30fps
人物領域	6 人
骨格推定	6 人
推定可能な関節	25 個
骨格推定可能距離	0.5〜4.5m

© 2017 Information Processing Society of Japan

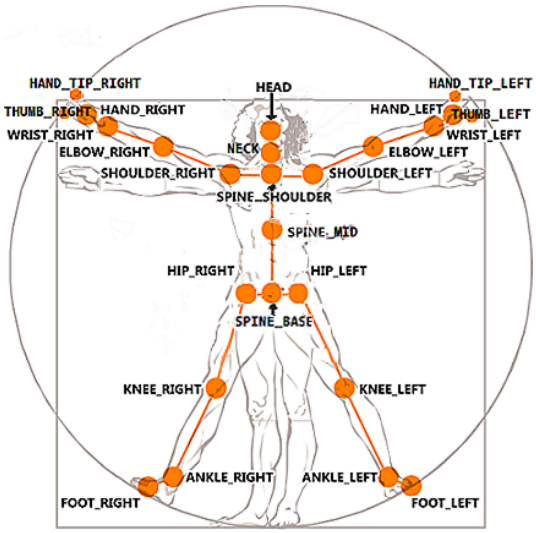

図 1　Kinect v2 により認識可能な関節 [6]

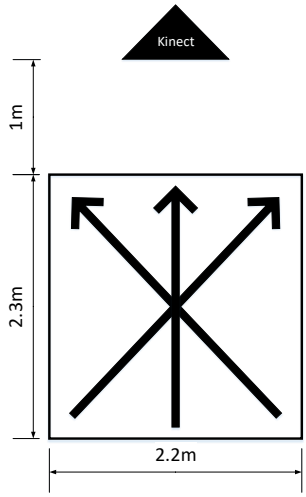

図 2　データ採取環境

表 3　データ採取環境

項目	値
開始日	2017/08/05
終了日	2017/08/17
人数	10 人
一人あたりの歩行回数	6 回

表 4　被験者の情報

項目	最小	最大
身長	165cm	170cm
年齢	20	51

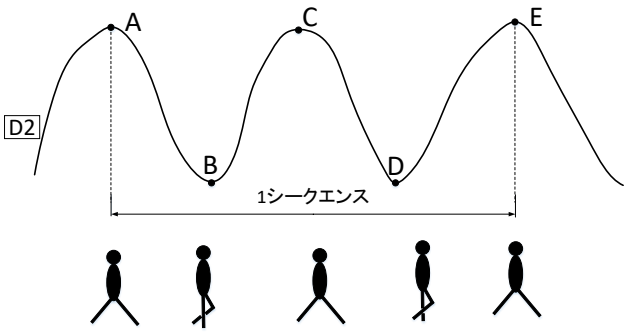

図 3　1 シークエンスのサンプル

4.3　実験方法

4.3.1　データ採取

データ採取の環境を表 3 に示す．10 人の被験者にはそれぞれ U_1〜U_{10} の被験者 ID を振った．被験者の情報を表 4 に示す．

データ採取は実験室内の図 2 の空間で行い，図の 3 つの矢印の方向の 2 回ずつ，計 6 回歩いてもらい，データを取得した．

4.3.2　前処理

採取したデータから 1 周期分を抜き出す．1 周期を判別する基準には足首間の距離を用いる．本研究で，は歩行の 1 周期とは片方の足が前に出た時刻から再びその足が前に出るまでの間とする．

まず足首間の距離の時系列データに対して，2 フレームの移動平均を求め平滑化をする．各時点において前後 2 点との大小を比較しピーク値を求める．1 つ目のピークから 3 つ目のピークまでのフレームを抜き出し，1 周期分とする．これを 1 シークエンスと呼ぶ．

1 シークエンスの例を図 3 に図示する．この例の場合 1 つ目のピークである A 点から 3 つ目のピークである E 点までのフレームが 1 シークエンスとなる．

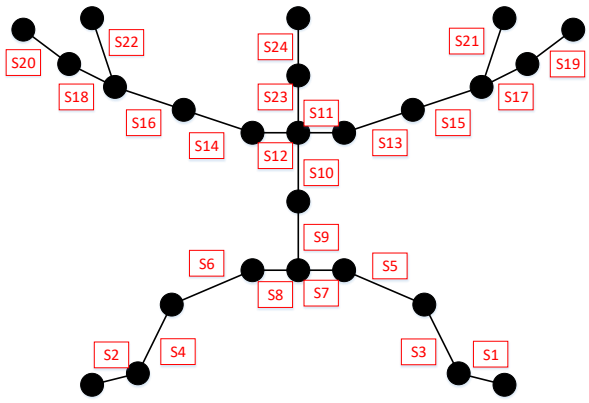

図 4　静的距離の簡略図

4.3.3　特徴量の種類

次の 4 種類の特徴量を考える．

4.3.3.1　静的距離

図 1 の 25 個の関節の内，隣接している関節間の距離である．静的距離を表 5 と図 4 に示す．これは骨格の長さを表す．

© 2017 Information Processing Society of Japan

表 5　静的距離リスト

ID	特徴名	平均値	標準偏差	中央値	最大値
S_1	FootL-AnkleL	11.76	2.29	11.58	17.71
S_2	FootR-AnkleR	11.88	2.08	11.86	17.05
S_3	AnkleL-KneeL	37.60	3.57	38.00	45.26
S_4	AnkleR-KneeR	37.46	3.60	37.86	45.33
S_5	KneeL-HipL	35.77	3.76	35.63	49.77
S_6	KneeR-HipR	35.96	3.79	35.76	51.38
S_7	HipL-SpineBase	8.21	1.44	8.52	11.66
S_8	HipR-SpineBase	7.71	1.56	8.15	10.76
S_9	SpineBase-SpineMid	31.10	1.01	31.05	39.50
S_{10}	SpineMid-SpineShoulder	22.80	0.68	22.76	28.10
S_{11}	SpineShoulder-ShoulderL	17.24	1.15	17.29	21.61
S_{12}	SpineShoulder-ShoulderR	17.58	1.43	17.69	22.46
S_{13}	ShoulderL-ElbowL	25.65	2.02	25.65	31.21
S_{14}	ShoulderR-ElbowR	25.57	2.32	25.41	31.60
S_{15}	ElbowL-WristL	23.92	2.18	23.43	31.37
S_{16}	ElbowR-WristR	24.28	2.46	23.66	32.33
S_{17}	WristL-HandL	7.52	1.71	7.62	12.56
S_{18}	WristR-HandR	7.71	1.72	7.77	15.08
$S19$	HandL-HandTipL	7.01	1.93	7.37	11.50
S_{20}	HandR-HandTipR	7.08	1.88	7.40	15.67
S_{21}	WristL-ThumbL	9.18	2.54	9.01	19.28
S_{22}	WristR-ThumbR	9.43	2.45	9.39	18.73
S_{23}	SpineShoulder-Neck	7.46	0.21	7.45	9.10
S_{24}	Neck-Head	14.68	1.02	14.70	18.25

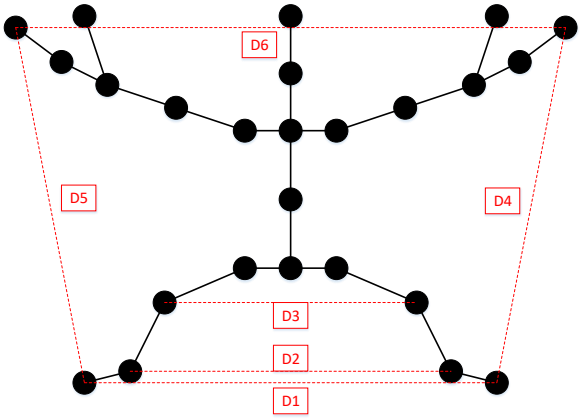

図 5　動的距離の簡略図

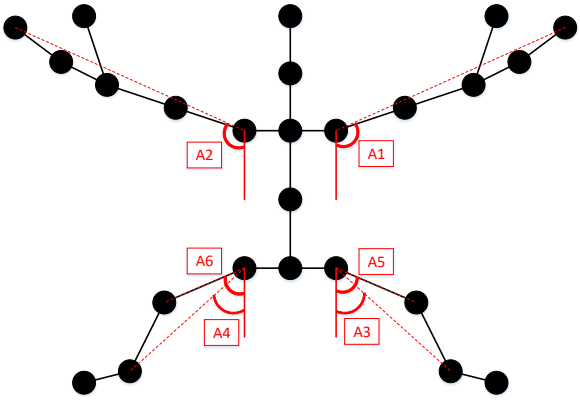

図 6　関節角度の簡略図

4.3.3.2　動的距離

歩行の際に，周期的に変動すると考えられる関節同士の距離である．動的距離を表 6 と図 5 に示す．

4.3.3.3　関節角度

歩行の際に，周期的に変動すると考えられる腕や足の角度である．ここでの角度とは，ある 2 つの関節を結ぶ (真下に対する) 角度である．関節角度のリストを表 7 に列挙し，図 6 に図示する．

4.3.4　後処理

4.3.2 節で抽出した各シークエンスにおいて 4.3.3 節のそれぞれの特徴量を計算する．特徴量は，各シークエンス毎に平均値，中央値，最大値の 3 方法で求める．

4.3.5　識別

単一の特徴量のみで識別を行う場合は，対象者 i のシークエンス k における，ある特徴量

$$f \in \{S_1, ..., S_{24}, D_1, ..., D_6, A_1, ..., A_6\}$$

の統計値を

$$\mu(f_{i,k}), median(f_{i,k}), max(f_{i,k})$$

© 2017 Information Processing Society of Japan

表 6　動的距離リスト

ID	特徴名	平均値	標準偏差	中央値	最大値
D_1	FootL-FootR	32.40	14.61	32.10	74.51
D_2	AnkleL-AnkleRigh	32.37	12.75	33.29	73.94
D_3	KneeL-KneeR	23.19	4.96	23.46	37.72
D_4	FootL-HandTipL	71.12	9.96	70.51	106.90
D_5	FootR-HandTipR	69.18	11.02	67.48	119.61
D_6	HandTipL-HandTipR	47.20	9.96	46.46	93.01

表 7　関節角度リスト

ID	特徴名	平均値	標準偏差	中央値	最大値
A_1	ShoulderL-HandTipL	13.41	6.73	13.00	40.16
A_2	ShoulderR-HandTipR	12.92	7.29	12.33	64.09
A_3	HipL-AnkeL	18.95	10.46	16.87	45.94
A_4	HipR-AnkleR	18.62	10.26	16.62	43.44
A_5	HipL-KneeL	13.76	7.68	12.17	38.60
A_6	HipR-KneeR	13.83	7.91	12.03	41.30

閾値を θ とすると，$\mu(f_{i,k})$ と $\mu(f_{j,k'})$ を用いは識別は次のように行う．

$$same(i,j) = \begin{cases} T & \text{if } |\mu(f_{i,k}) - \mu(f_{j,k'})| \leq \theta \\ F & \text{otherwise} \end{cases}$$

中央値，最大値についても同様に行う．

複数の特徴量を組み合わせる際はそれぞれの特徴量を要素とするベクトルのユークリッド距離で判断する．例えば，2 つの特徴量 f と g を用いて識別を行う場合は次のように識別する．

$$same(i,j) = \begin{cases} T & \text{if } \sqrt{(\mu(f_{i,k}) - \mu(f_{j,k'}))^2 + (\mu(g_{i,k}) - \mu(g_{j,k'}))^2} \leq \theta \\ F & \text{otherwise} \end{cases}$$

中央値，最大値についても同様に計算を行う．また，特徴量を 3 つ以上使う際も同様にユークリッド距離を用いる．

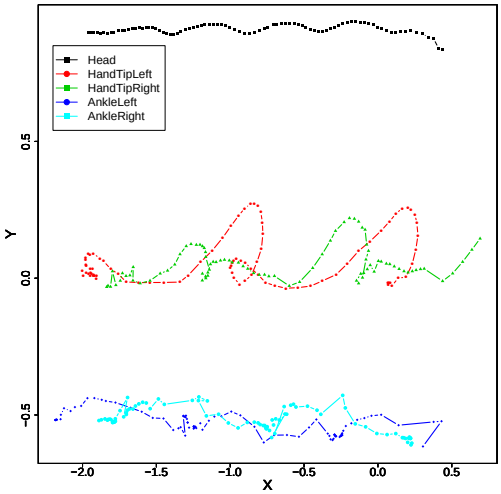

図 7　カメラ方向から見た関節の時系列データ

4.4　実験結果

4.4.1　データ採取

採取した各関節の 3 次元座標の時系列データを図 **7** に示す．歩行をカメラ方向から見た場合の 5 つの代表的な関節の時系列データである．

4.4.2　前処理

前処理の結果，抽出された 1 シークエンスの例を図 **8** に示す．2 点間での平滑化されたデータを赤で示す．

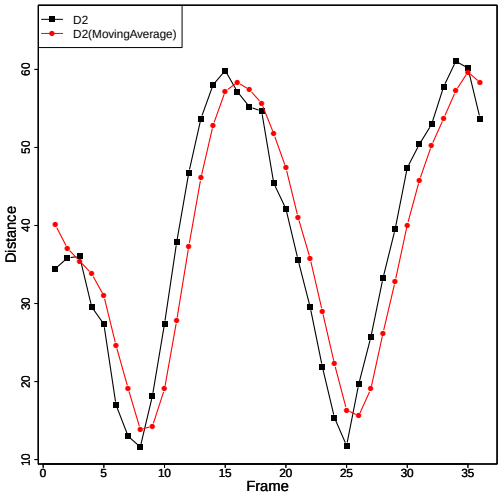

図 8　1 シークエンスの D_2 の例

4.4.3　後処理

抽出した特徴量の例として，ある被験者のいくつかのシークエンスの D_5 のデータを表 **8** に示し，U_6 の D_5 の特徴抽出方法ごとの分布を図 **9** に示す．分布はどの統計量も大差ないが，Median が外れ値の影響を受けていないことがわかる．また，全被験者の $\mu(D_5)$ の分布を図 **10** に示す．

4.4.4　識別

例として $\mu(D_5)$ について，本人同士で測った距離と他人

© 2017 Information Processing Society of Japan

表 8　特徴量の例

被験者 ID	シークエンス	$\mu(D_5)$	$median(D_5)$	$max(D_5)$
U_1	1	65.4	63.9	77.0
U_1	2	62.8	63.0	70.2
U_2	1	76.3	73.1	100.4
U_2	2	75.6	75.3	104.5
U_5	1	62.7	61.6	68.5
U_5	2	60.8	60.8	71.9

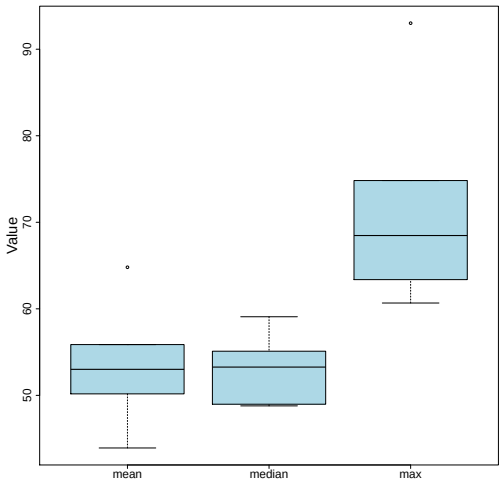

図 9　U_6 の D_5 の統計量分布

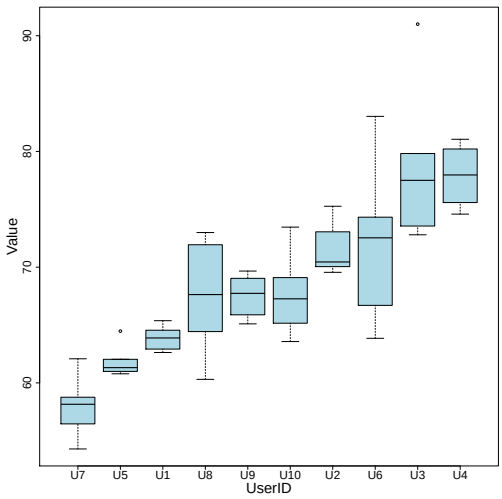

図 10　全被験者の $\mu(D_5)$ の分布

同士で測った距離のヒストグラムを図 11 に示す．

また，閾値を変化させた場合の他人受入率 (FAR) と本人拒否率 (FRR) のトレードオフを示した受信者操作特性 (ROC) 曲線について，図 12 に示す．$\mu(S_6)$,$median(D_1)$, $max(A_2)$ の 3 つの特徴量を示している．この 3 つの特徴量の中では $median(S_6)$ が最も優れた特徴量である．

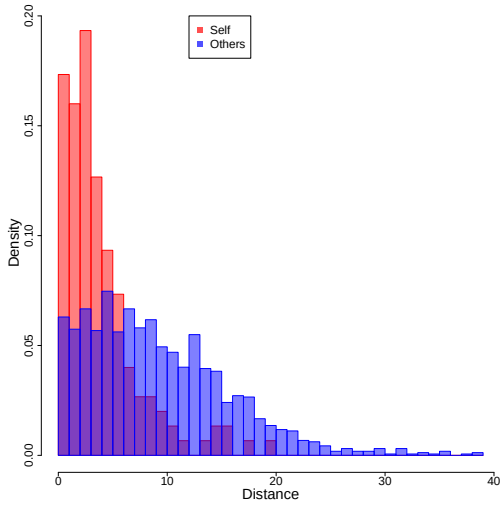

図 11　$\mu(D_5)$ の本人と他人の距離の比較)

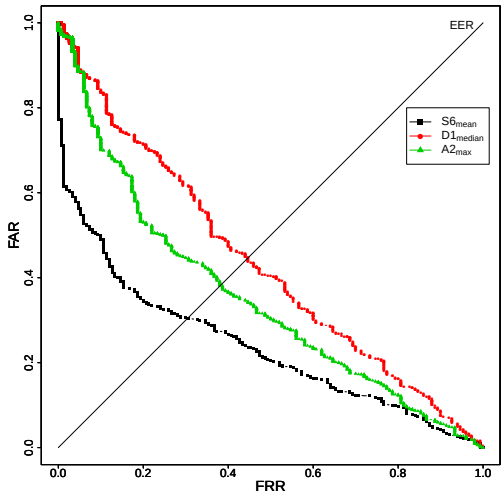

図 12　特徴量 $\mu(S_6)$,$median(D_1)$, $max(A_2)$ の ROC 曲線

表 9　EER の Top10 の特徴量

特徴量	分類	特徴量の取り方	EER
$\mu(D_5)$	動的距離	平均値	0.29
$max(D_5)$	動的距離	最大値	0.29
$\mu(S_6)$	静的距離	平均値	0.30
$median(D_5)$	動的距離	中央値	0.30
$\mu(S_5)$	静的距離	平均値	0.31
$median(A_4)$	関節角度	中央値	0.31
$median(S5)$	静的距離	中央値	0.31
$median(S_6)$	静的距離	中央値	0.31
$\mu(D_4)$	動的距離	平均値	0.32
$\mu(A_4)$	関節角度	平均値	0.32

5.　評価・分析

5.1　特徴量の比較

誤差 EER の値小さい順に上位 10 個の特徴量を表 9 に示す．

© 2017 Information Processing Society of Japan

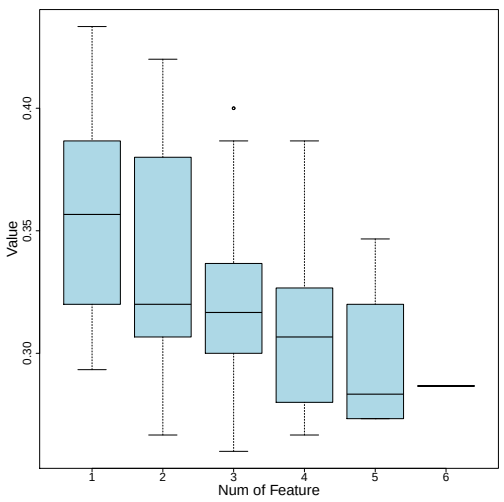

図 13　特徴量の組み合わせ数と EER の関係

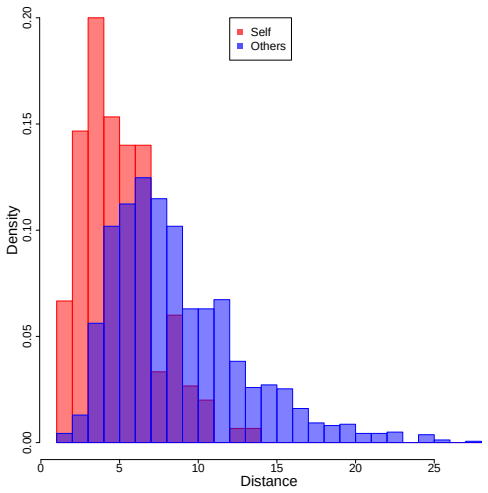

図 14　動的距離 (Max) のときに 6 つの特徴量を全て組み合わせた祭の本人と他人の距離の比較

5.2　特徴量の組み合わせ

動的距離と関節角度について，それぞれの特徴量を組み合わせた際の EER を計算した．誤差の小さい順に並べたときの上位 10 個を**表 10** に，下位 10 個を**表 11** に示す．Max の場合の動的距離について，組み合わせる特徴量の数とそのときの EER の関係を**図 13** に示す．組み合わせる特徴量を増やすことで EER は単調に減少しているが 5 個で頭打ちとなっている．

また，Max の動的距離の 6 つの特徴量を全て組み合わせた場合の，本人同士の距離と他人同士の距離のヒストグラムを**図 14** に示す．

5.3　次元削減

例として A1〜A6 を要素とする特徴ベクトルを作成し，それを多次元尺度法 (MDS: Multi-Dimensional Scaling) でマッピングした結果を**図 15** に示す．

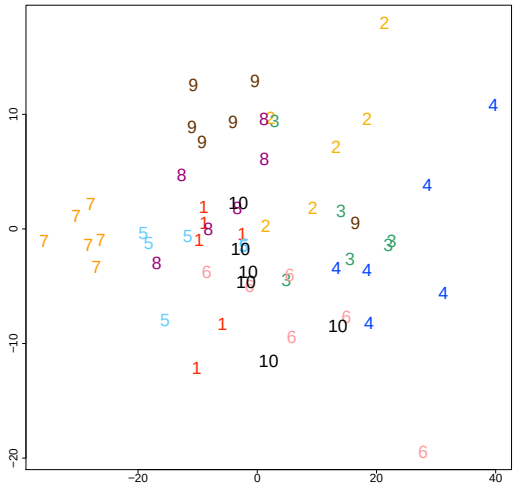

図 15　Max の場合の動的距離の MDS

6.　考察

図 9 の外れ値の原因は，Kinect の誤測定と考える．なぜならば，Kinect v2 は直接見える部分しかトラッキングできないため，複数人が同時に現れたり，手などの一部が体に隠れるなどが原因で正確な情報が取得できない可能性がある．正確に取得できない関節の座標は推定で補うことができるが，本研究ではその推定値も実測値と同様に扱ってしまったため外れ値が増加したと考えられる．

表 9，表 10 において，後処理の最大値が他の平均などよりも悪かったのは，最大値の場合 1 シークエンスの中に一つでも外れ値で大きな値があると影響を大きく受けてしまうためである．

表 9，表 10 より，動的距離が識別に有効である可能性が高いことがわかる．これは，静的距離は体の大きさの影響を大きく受け，関節角度は動きの大きさの影響を大きく受けるのに対し，動的距離は体の大きさと動きの大きさ両方の影響を受けるのでより個人間の差が大きくなるためであると考えられる．

表 11 より，関節角度，特に A_5,A_6 を使用すると EER が大きくなった．これは，人によってかかとや足先を上げる高さにバラつきがあっても，膝を上げる角度は身長の影響も受けにくく，個人差が現れにくいので識別率が低いと考えられる．

7.　おわりに

本研究では，特徴量を提案し Kinect v2 を用いて採取した歩容データから実際に特徴量を抽出する実験を行った．実験の結果提案手法は人物識別に使用できる可能性があることがわかった．また，組み合わせる特徴量を変えることで EER が 0.25 から 0.50 まで変化することが可能であった．

© 2017 Information Processing Society of Japan

表 10　特徴量を組み合わせた際の EER の上位 10

組み合わせた特徴量	分類	特徴量の取り方	EER
$\mu(D_3)$, $\mu(D_5)$	動的距離	平均値	0.25
$\mu(D_3)$, $\mu(D_4)$, $\mu(D_5)$	動的距離	平均値	0.25
$\mu(D_2)$, $\mu(D_3)$, $\mu(D_4)$, $\mu(D_5)$	動的距離	平均値	0.25
$\mu(D_3)$, $\mu(D_4)$, $\mu(D_5)$, $\mu(D_6)$	動的距離	平均値	0.25
$\mu(D_1)$, $\mu(D_3)$, $\mu(D_4)$, $\mu(D_5)$, $\mu(D_6)$	動的距離	平均値	0.25
$\mu(D_2)$, $\mu(D_3)$, $\mu(D_4)$, $\mu(D_5)$, $\mu(D_6)$	動的距離	平均値	0.25
$\mu(D_2)$, $\mu(D_3)$, $\mu(D_5)$	動的距離	平均値	0.26
$\mu(D_1)$, $\mu(D_2)$, $\mu(D_3)$, $\mu(D_4)$, $\mu(D_5)$	動的距離	平均値	0.26
$\mu(D_4)$, $\mu(D_5)$, $\mu(D_6)$	動的距離	平均値	0.26
$\mu(D_1)$, $\mu(D_2)$, $\mu(D_3)$, $\mu(D_4)$, $\mu(D_5)$, $\mu(D_6)$	動的距離	平均値	0.26

表 11　特徴量を組み合わせた際の EER の下位 10 個

組み合わせた特徴量	分類	特徴量の取り方	EER
$\mu(A_6)$	関節角度	平均値	0.50
$median(A_6)$	関節角度	中央値	0.48
$median(A_5)$, $\mu(A_6)$	関節角度	平均値	0.47
$median(A_2)$,	関節角度	中央値	0.47
$median(A_2)$, $median(A_5)$	関節角度	中央値	0.46
$max(A_5)$	関節角度	最大値	0.45
$max(A_5)$,	関節角度	平均値	0.45
$median(A_2)$, $median(A_5)$, $median(A_6)$	関節角度	中央値	0.45
$median(A_2)$, $median(A_6)$	関節角度	中央値	0.45
$median(D_1)$	動的距離	中央値	0.45

今後の課題としては，統計情報として人物識別を行うにはどの程度の識別率があればよいのかを調査すること，プライバシーを保護しつつ人物識別を行うにはどの特徴量を用いて識別率をどの程度にしたらよいのかを検討することなどがある．

参考文献

[1] K. Shiraga, Y. Makihara, D. Muramatsu, T. Echigo, Y. Yagi, “GEINet: View-Invariant Gait Recognition Using a Convolutional Neural Network””, In Proc. of the 8th IAPR Int. Conf. on Biometrics (ICB 2016), pp.1-8, Halmstad, Sweden, Jun 2016

[2] S. D. Bakchy, M. R. Islam, A. Sayeed, “Human identification on the basis of gait analysis using Kohonen self-organizing mapping technique” 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Dec 2016

[3] J. Han and B. Bhanu, “Individual recognition using gait energy image,” IEEE Trans. Pattern Anal. Mach. Intell., vol.28, no.2, pp.316322, 2006.

[4] V. Andersson and R. Araujo, “Person Identification Using Anthropometric and Gait Data from Kinect Sensor,” In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015.

[5] L. Igual, À. Lapedriza, R. Borràs, “Robust gait-based gender classification using depth cameras”, EURASIP Journal on Image and Video Processing 2013, 2013.

[6] “JointType Enumeration - MSDN - Microsoft”, (https://msdn.microsoft.com/ja-jp/library/microsoft.kinect.jointtype.aspx,2017 年 8 月参照)

[7] “Kinect ハードウェア - Microsoft Developer”, (https://developer.microsoft.com/ja-jp/windows/kinect/hardware, 2017 年 8 月参照)

[8] 石丸 大輔, 姜 玄浩，岩村 惠市，“キネクトの骨格情報による非日常ポーズ認証”, SCIS2015, Jan 2015.

© 2017 Information Processing Society of Japan