

歩容データからの属性暴露リスクについて

三好 駿, 森 駿文, 菊池 浩明

明治大学

研究背景

歩容（歩き方の特徴）

- ・ 2016年 個人情報保護法改正
- ・ 歩容データも個人識別符号とみなされる

先行研究

- ・ 森らの研究[1]でEER=0.048の精度で歩容識別できることが証明された
- ・ 橋口らの研究[2]では歩行データからパーキンソン病の早期診断の成果が出ている

[1]森駿文, 菊池浩明, 歩容データのDTW距離に基づく個人識別における複数部位のフュージョン手法, コンピュータセキュリティシンポジウム (CSS 2018), pp. 1-8, 2018.

[2]橋口裕徳, 堀宏有, 廣部祐樹, 沢田裕之, 稲葉 彰, 織茂智之, 三宅美博, 歩行軌道データに基づくパーキンソン病の早期診断システム, 第28回自律分散システム・シンポジウム資料(2016), pp.155-158

課題と研究目的

課題

- ・歩容データから性別や年齢，病歴などの個人に関わる属性が暴露されるリスクがあるのではないか[2]

研究目的

- ・歩容から属性推定のリスクがあるのかを明らかにする
- ・本研究では性別に着目
- ・最も識別される特徴量を明らかにする

[2]橋口裕徳, 堀宏有, 廣部祐樹, 沢田裕之, 稲葉 彰, 織茂智之, 三宅美博, 歩行軌道データに基づくパーキンソン病の早期診断システム, 第28回自律分散システム・シンポジウム資料(2016), pp.155-158

問題点と解決策

問題点

- データ
 - 歩容のビッグデータがない
- 性別推定
 - 男女の歩き方にどのような差があるのか明らかではない

解決策

- データの取得
 - Kinectを用いてデータを取得する
 - 本研究では、121名から歩容データを取得
- 可視化プログラムの実装
 - 観測に基づき、独自の特徴量を提案する

歩容データの取得

取得方法と条件

- Kinect v2(Microsoft)を使用
- 明治大学中野キャンパス多目的室
- 床から0.9mの位置にKinect
- 被験者に2.5m歩行してもらう
- 121名に6回歩行してもらい測定

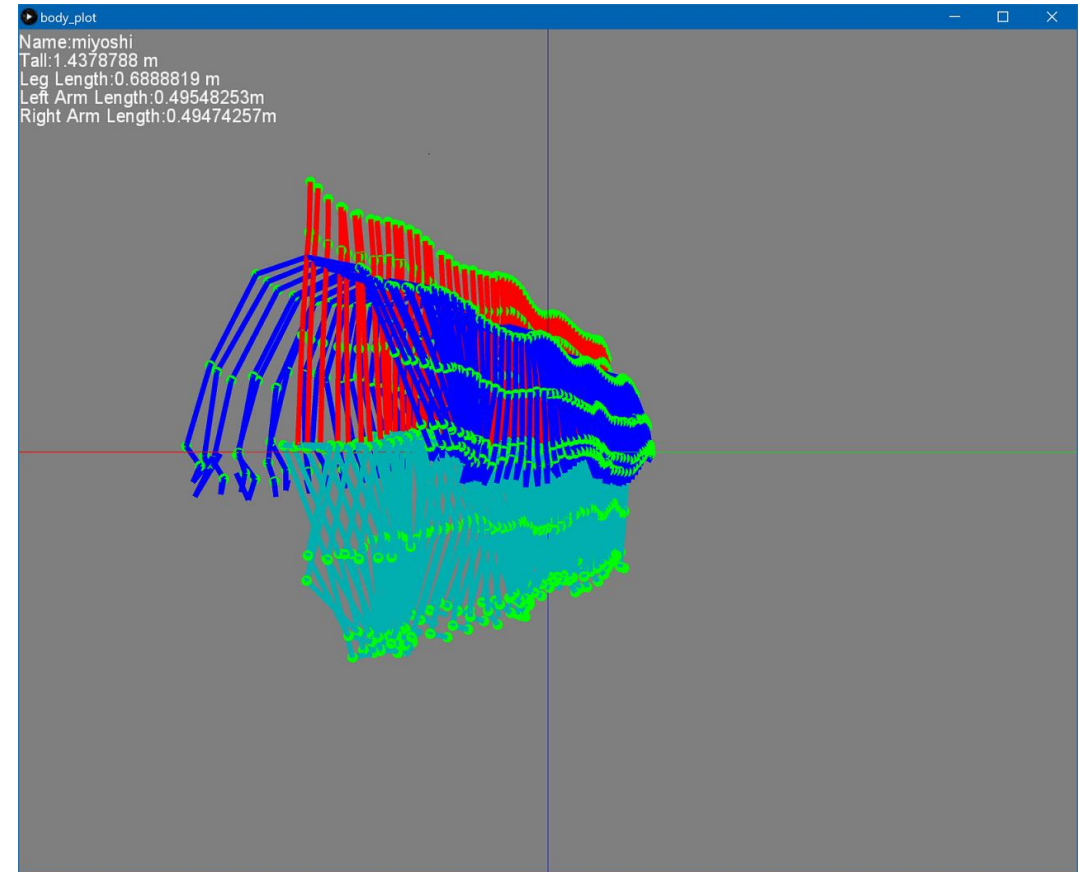
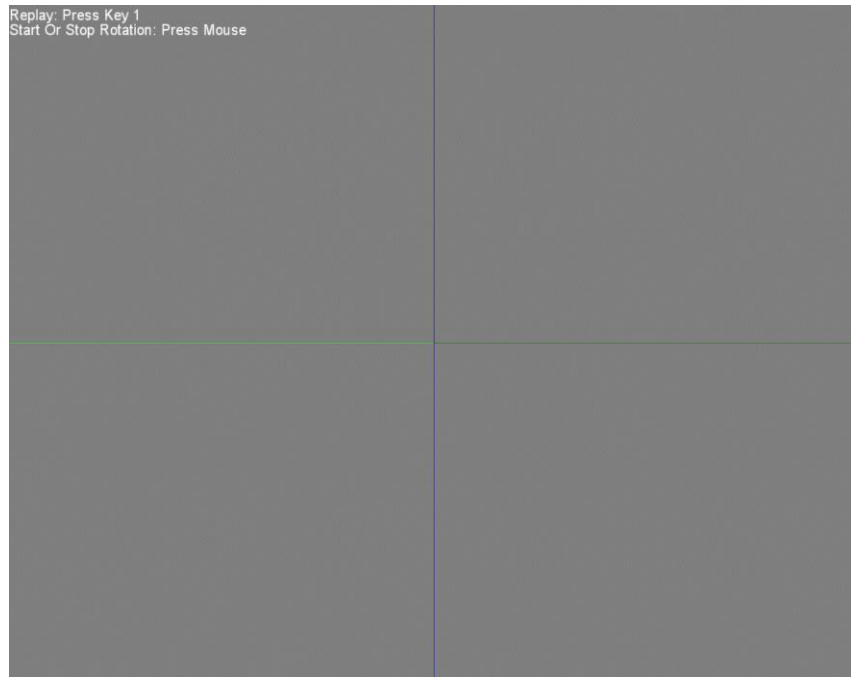


被験者	男	女
人数	77	44
年齢	18~57歳	18~70歳
データ数	462	262

可視化プログラム

可視化プログラムの実装

- Processingを使用
 - 歩容データのアニメーション
 - 移動履歴の累積プロット

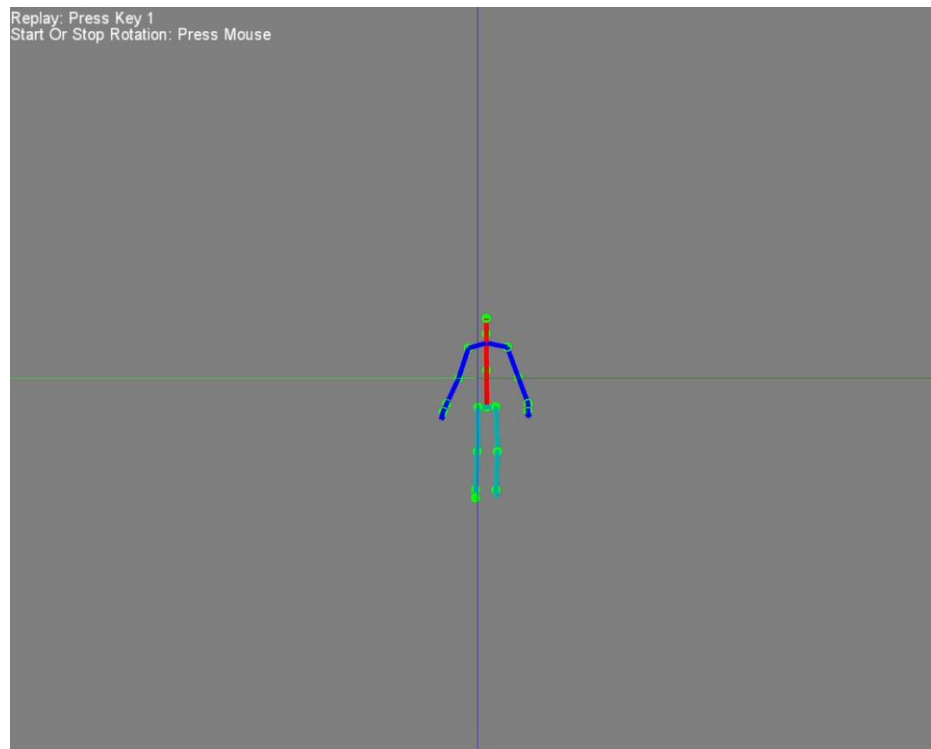


男女の歩容

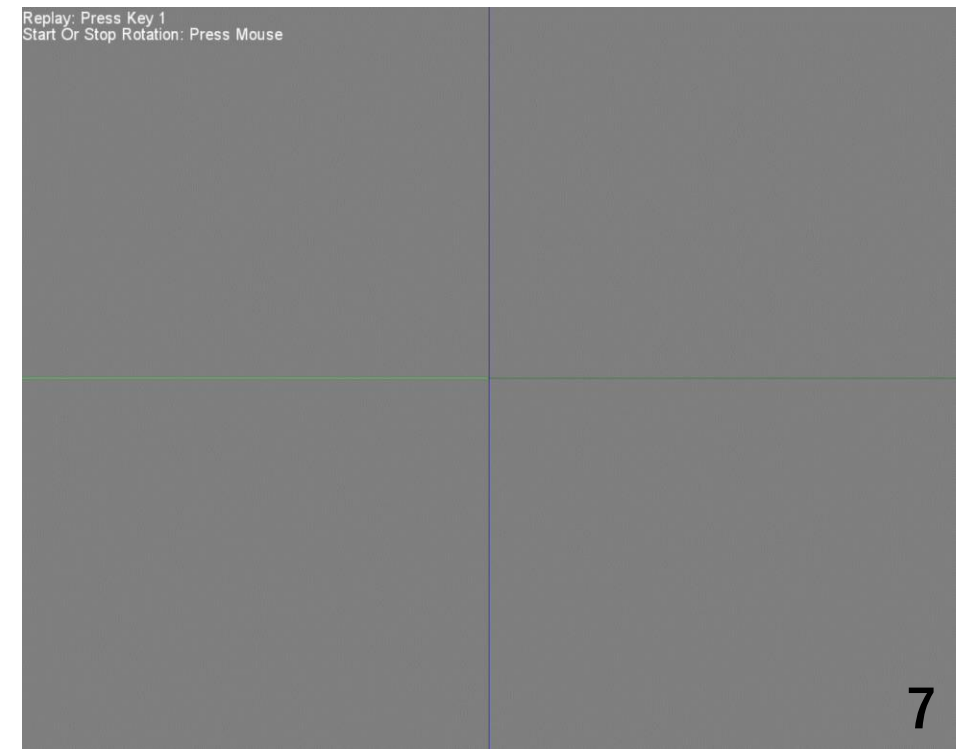
可視化プログラムによる観測

- ・男女による歩き方の違いが見られた

女性



男性

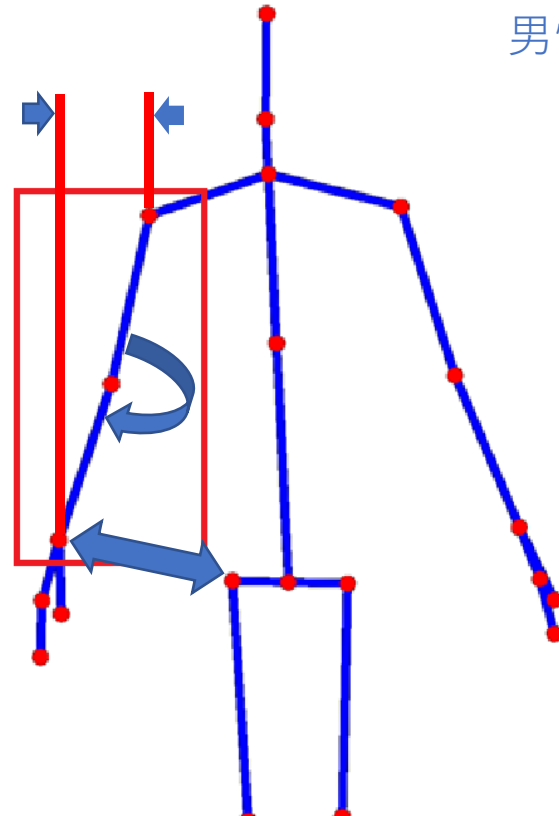


特徴量の提案

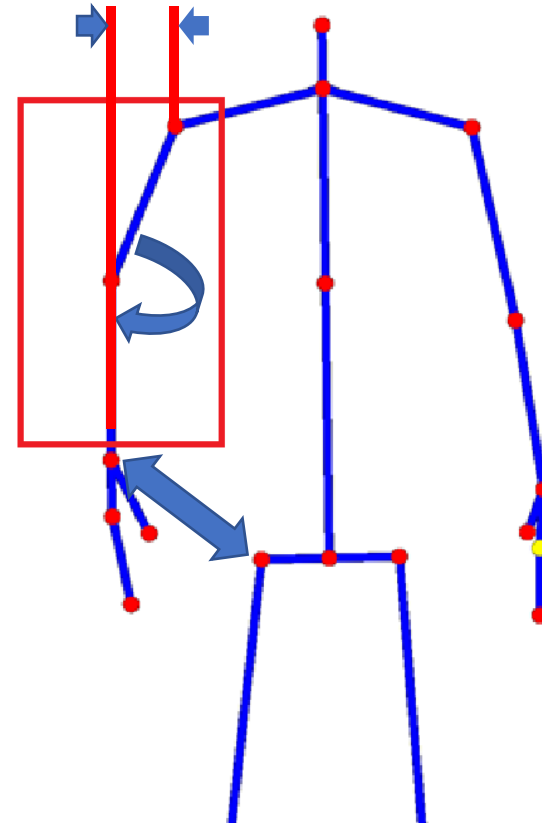
動的特徴量の提案

- ・可視化された歩容を観察し、男女の歩き方が顕著だった部分に着目した

女性



男性



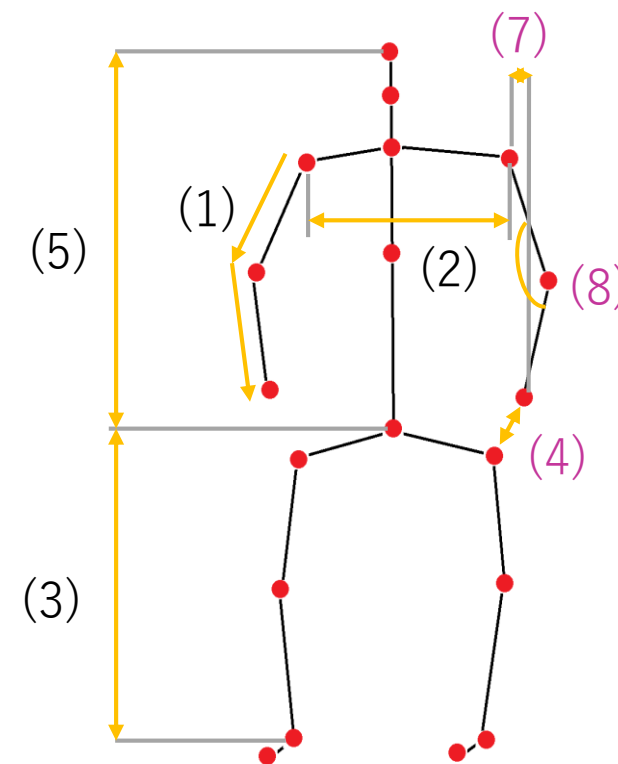
特徴量の定義

特徴量

- ・ 動的特徴量3つと、身体的な静的特徴量5つを選定
- ・ 合計で8つの特徴量を定義する

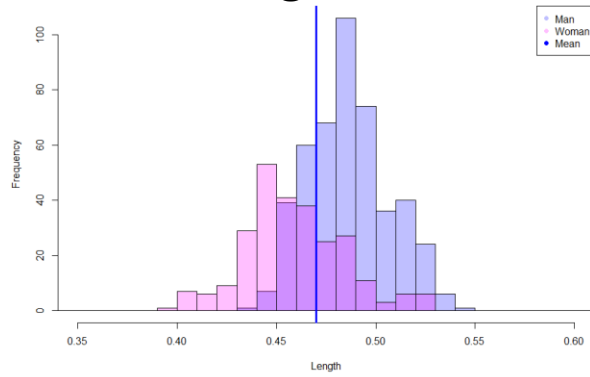
動的特徴量	歩行中に値が変化する特徴量
静的特徴量	歩行中に値が変化しない特徴量

特徴量	定義
(1)Arm Length	SHOULDER-WRIST
(2)Shoulder	SHOULDER_R-SHOULDER_L
(3)Leg	HIP-ANKLE
(4)Hip Wrist	HIP-WRIST
(5)Upper Body	HEAD-SPINE_BASE
(6)Tall	Upper Body + Leg
(7)Shoulder Wrist	SHOULDER-WRIST
(8)Arm Angle	ELBOWの角度

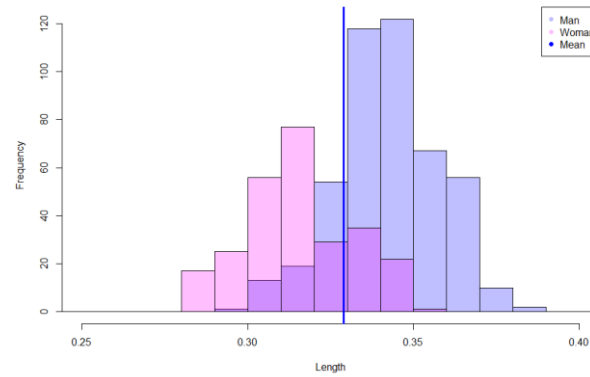


特徴量の分布

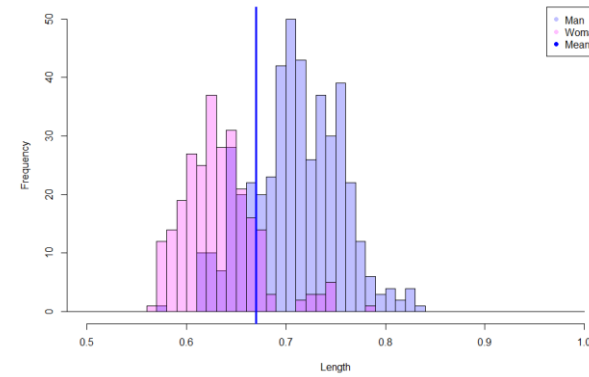
(1) Arm Length



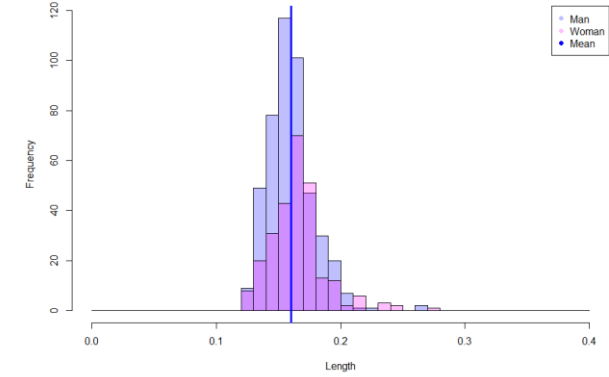
(2) Shoulder



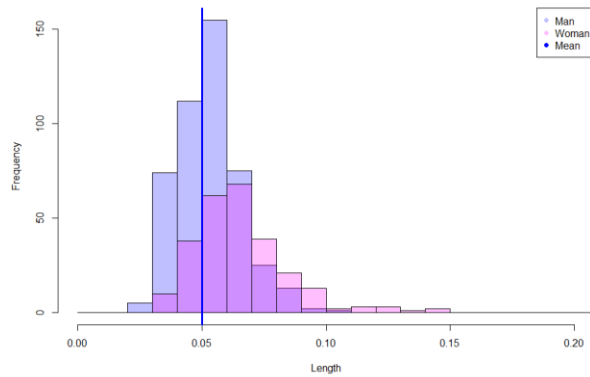
(3) Leg



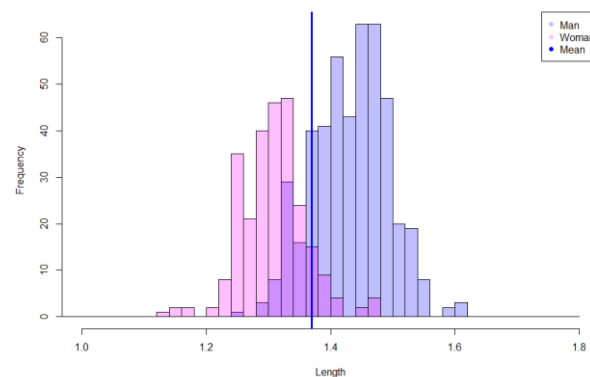
(4) Hip Wrist



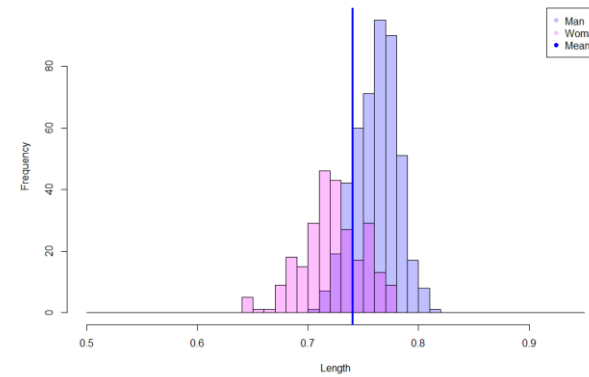
(5) Upper Body



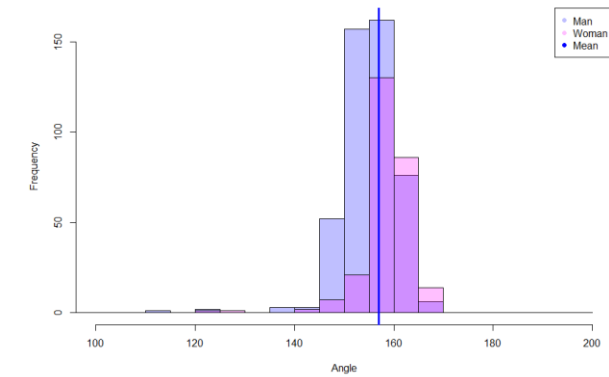
(6) Tall



(7) Shoulder Wrist



(8) Arm Angle



分布の誤差

誤差の原因

- ・男女ともに異性らしい歩き方をする被験者が見られた

女性



男性



性別推定率

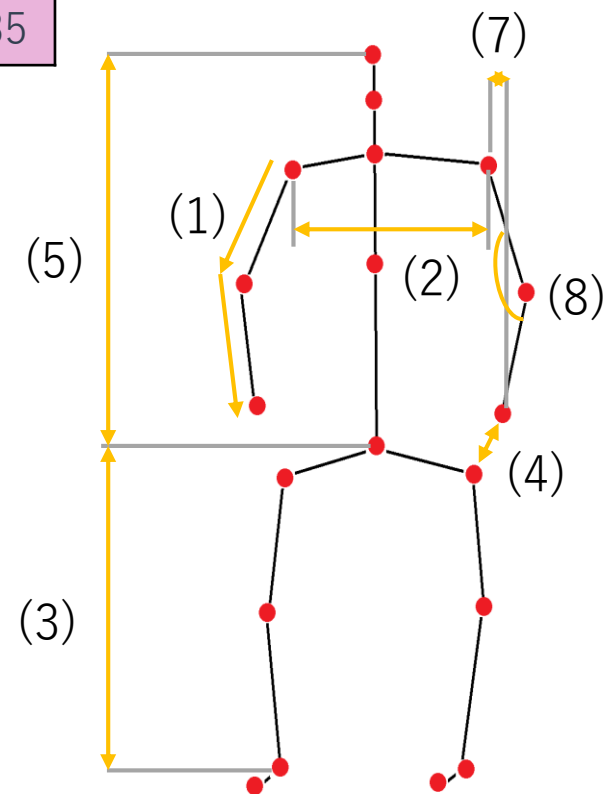
定義

- 推定率 = $\frac{a+d}{a+b+c+d}$
- 右表の数値はArm Lengthの値

	真の性別	
	男	女
推定性別		
男	a:385	b:28
女	c:76	d:235

結果

特徴量	平均 男	平均 女	推定率(%)
(1)Arm Length	0.48	0.45	85.63
(2)Shoulder	0.34	0.31	82.04
(3)Leg	0.70	0.63	80.24
(4)Hip Wrist	0.15	0.16	79.97
(5)Upper Body	0.76	0.72	72.92
(6)Tall	1.43	1.30	69.06
(7)Shoulder Wrist	0.05	0.06	65.60
(8)Arm Angle	155.23	158.65	59.94



精度の向上

特徴量の統合

- ・ 推定率を向上させるため、複数の特徴量を統合
 - 統合する特徴量の個数をkとする
 - 各特徴量を推定率の高い順にソート
 - 上位k個の特徴量のうち過半数の推定性別を統合時の推定性別とする

例：k=3の時

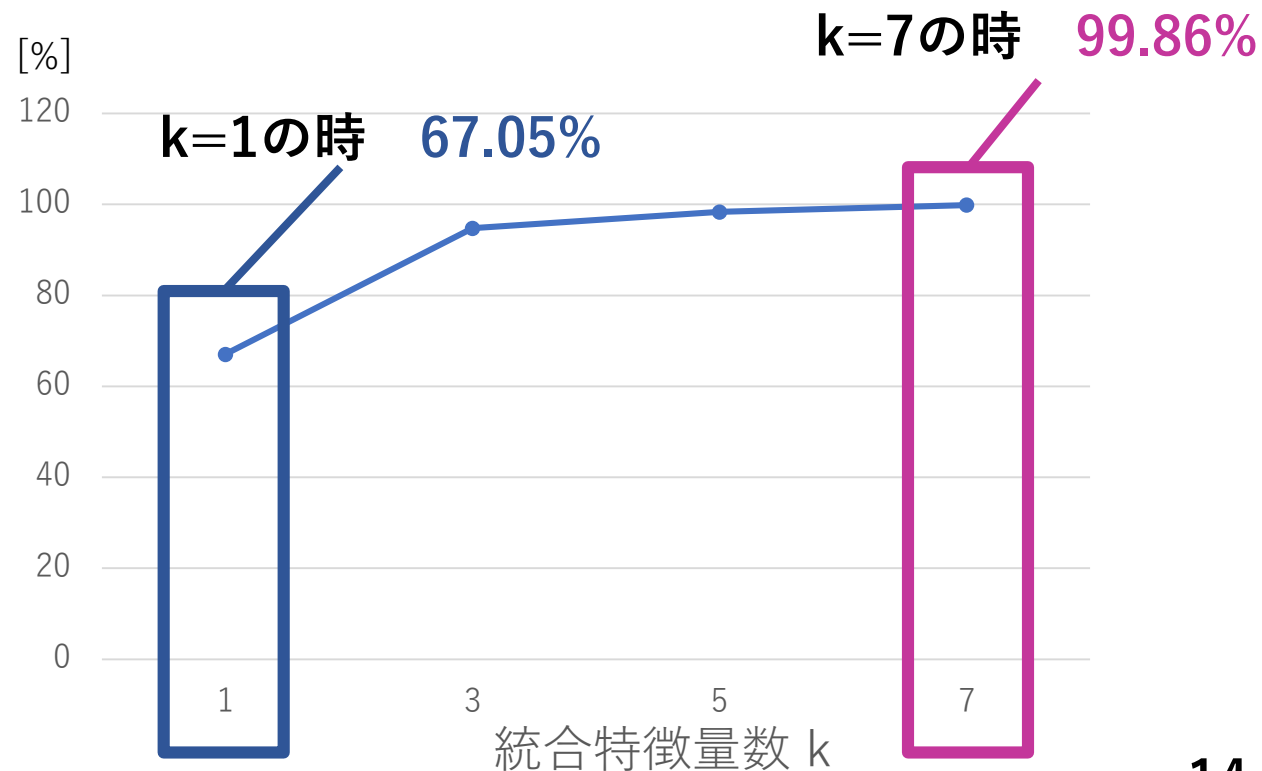
特徴量	推定性別	統合性別
(1)Arm Length	M	M
(2)Shoulder	M	
(3)Leg	F	

統合推定結果

結果

- ・ kを増やすと推定率は向上した
- ・ これは各特徴量が独立しているため向上していると考えられる

k	特徴量	推定率(%)
3	(1)~(3)	94.75
5	(1)~(5)	98.34
7	(1)~(7)	99.86



おわりに

結論

- ・本研究では、歩容を用いた性別推定手法を提案し、実験により精度評価を行った
- ・実験の結果、腕の長さ(1)Arm Lengthが最も性別を識別されやすい特徴量であることが明らかになった
- ・さらに7個の特徴量を組み合わせることにより、99.86%の精度で性別の推定が可能であることを示した
- ・歩容には性別暴露のリスクがあることを明らかにした

今後の課題

- ・本研究では、学習データとテストデータを分けずに評価を行ってしまった