

特別研究報告書

フットステップデータの収集と多変量解析を
用いた個人識別能力の分析

指導教員 美濃 導彦教授

京都大学工学部情報学科

藤井 崇介

平成16年2月10日(木)

フットステップデータの収集と多変量解析を用いた個人識別能力の分析

藤井 崇介

内容梗概

ユビキタスネットワークの発展に向け、様々な機器をネットワークに接続し、ユーザへサービスを提供するための研究が行われている。このようなユビキタス環境ではユーザごとに適応したきめ細かなサービスの提供が望まれており、そのためにユーザの個人識別が必要不可欠となっている。一方、ユビキタス環境における個人識別ではユーザに煩わしさを感じさせないことが望まれている。フットステップデータを用いた個人識別手法はユーザへの負担が少ない手法として期待されている。本研究ではフットステップデータについて基礎的な知見を得ることを目的とする。

現在ユビキタス環境では個人識別手法としてRFID タグなどの徽章を用いるものや、虹彩、指紋などの個人属性を基に識別するもの、パスワード入力によるものなどが研究、あるいは実用段階にある。しかしながらこれらの手法では物体の携帯や機器の操作が必要であるため、ユーザにとって物理的負担となっている。また、虹彩、指紋を利用した個人識別手法では個人属性の情報を与えるため、プライバシーの観点から心理的負担となっている。それに対し、床に配置した圧力センサの上をユーザが踏み通り抜けることで歩行中の1歩の間に床にかかる力であるフットステップデータを取得し、識別を行う手法はユーザにとって物理的負担、心理的負担が少ないと考えられる。

フットステップデータを利用した個人識別に関する先行研究においては個人識別の可能性が示されている。しかしながら、識別に影響を与えると考えられるデータ収集環境や被験者について十分な考察が示されていない。そこで本研究では、まず収集する環境がフットステップデータにどのような影響を与えるかについて調べ、適切なフットステップデータ収集環境を構築する。次に収集されたデータについてその識別能力の検討を行う。

本研究では縦 50cm 横 50cm 高さ 3cm の台はかりを用いてフットステップデータを収集する。比較する収集環境として、先行研究で行われている2つの状況、

- 台はかりをそのまま床に置く。これにより計測前後で段差が生じる。
- 台はかりの周囲に台はかりと同じ高さの物体を配置。計測前後に段差が生

しない。

について比較を行う。得られるデータから先行研究に用いられている特徴量を求め、その分布について仮説検定を行った。その結果、この2つの状況で得られるデータの特徴量は異なる分布をしていることが分かった。

また台はかりに対する被験者の足の接地位置への意識がデータの特徴量の分布に与える影響を調べる。そのために複数枚の台はかりを利用することで被験者の設置位置への意識を低減させる環境を構築する。次の2つの条件、

- 1枚の台はかりを明示し、被験者に接地位置を意識させる。
- 接地位置を意識させないようにする。

で得られるデータについて比較を行う。データから特徴量を求め、その分布について仮説検定を行った。その結果、2つの条件で得られるデータの特徴量は異なる分布をしていることが分かった。これにより計測前後で床の高さが変わらず、被験者が足の接地位置を意識しないフットステップデータ収集環境を構築する必要があることが分かった。

この収集環境で得られるフットステップデータの識別能力について調べる。各個人のデータのバラつき方と個人の間バラつき方を調べるために、まずデータの分散共分散行列を求め、主成分分析により分散について多くの情報を持つ軸を選び出す。そして識別能力について議論するためにフットステップデータの分布を次のように仮定する。

- 全フットステップデータの母集団は正規分布であり、その平均と分散共分散行列は標本のものとも一致する。
- ユーザ毎のデータの分布は平均を除いて同一の正規分布であり、ユーザ毎の平均は全フットステップデータの母集団の分布に基づく。

この仮定の下でモンテカルロ法によりユーザのデータの分布を設定する。設定したユーザのもつデータの分布の和領域に対する重複領域の割合より識別能力を評価する。その結果、検出率95%、誤識別率5%で10人程度の個人識別が可能であることが示された。

また体重が識別に与える影響を調べるために、体重情報を除いたフットステップデータの識別能力を検討する。その結果、体重情報を除くと誤識別率が大幅に大きくなり、体重がフットステップデータを用いた個人識別において重要な要因となっていることが分かった。しかしながら検出率95%、誤識別率5%で3人程度の個人識別が可能であることも示された。

Footstep Data Collecting System and Multi-variate Analysis of User Identification Performance

Shusuke FUJII

Abstract

The goal of this research is to obtain basic knowledge of footstep features whether they are applicable or not to the user identification system .

With the expansion of ubiquitous computing, many appliances will be connected via network. Already several researches have been done for appliances to offer multiple services to the users. Since the users would like to access their favorite services easily, the appliances should identify which user is accessing them. At the same time user identification should not be troublesome. Now many user identification systems are proposed and used, for example: RFID (radio frequency identification) tags, iris recognition, finger print matching, and password checking. In those systems users should have or handle an appliance, which could be troublesome in operation. Moreover, iris recognition or finger print matching could be mental burden for the users, since their sensitive personal data could be stored and mishandled in such systems. User's footstep features identification has a little load and no mental burden for the users since they only walk on the floor with pressure sensors.

In the previous research on footstep features, possibility of identifying users based on their walking was discussed. The difference of weight could have influence on footstep data, but the distribution of the subjects' weight was not analyzed in the previous research. The method to collect data could have influence on footstep analysis. In this report, first we discuss how to collect footstep data and we create the proper system to collect footstep data. Then, we analyze user identification performance of the footstep features.

To collect the footstep data, we use the weighting plate with four load cells. This plate is the size of 50cm x 50cm x 3cm. We collect footstep data in the following two scenarios based on the previous research:

- The plate is set on the floor. There is height difference between the floor and the surface of the plate.
- The object with the same height as the plate is set around. There is no

height difference.

Then we evaluate distributions of the features obtained from the collected data. These features are the same as in previous research. In the consequence, it is shown that the features obtained in these scenarios have different distributions.

We examine the possible effects of users' being aware of the position when they step on the plate. In order not to make them aware of the position where they step, we create the system using many plates to collect footstep data. We collect footstep data in the following two scenarios:

- The users are aware of the position of the specific plate.
- The users are not aware of the position.

Then we evaluate distributions of features obtained from the collected data. In the consequence, we show that the features obtained in these two scenarios have different the distributions. As a result the footstep data should be collected with the condition that there is no height difference and that the users are not aware of the plate position.

We collect footstep data with the above condition and analyze user identification performance using these data. A procedure to analyze the user identification is as follows. We calculate the covariance matrix from the collected data, and project the data into three dimensional space using the principle component analysis. Then we hypothesize the distributions of footstep data as :

- Footstep data of large users population has a normal distribution so the mean values and covariance should be similar for small groups of users.
- All users' footstep data have similar covariance but different mean values.

Finally, we assume some distributions of user in the above constraints and analyze the user identification performance from the overlapped domain of footstep data. As a result, the possibility of user identification with the above condition is that detection rate is 95%, error rate is 5%, for ten users.

To examine the influence of users' weight, we analyze footstep data for similar conditions. As a result, the error rate rises much higher and it is revealed that the weight is important for the user identification.

フットステップデータの収集と多変量解析を用いた個人識別能力の分析

目次

1	緒論	1
2	背景	2
2.1	ユビキタス環境における個人識別の問題点	2
2.2	フットステップデータを利用した個人識別手法	2
3	フットステップデータ収集システムの構築	4
3.1	実験装置	4
3.2	フットステップ収集環境がフットステップデータに与える影響	5
3.2.1	段差が与える影響	5
3.2.2	接地位置を意識することによる影響	8
3.2.3	複数枚の台はかりを利用した時のフットステップデータ抽出手順	11
4	多変量解析を用いた特徴量の識別能力の分析	17
4.1	フットステップデータの識別能力	17
4.2	体重以外の特徴量を用いた場合	22
4.3	考察	25
5	結論	25
	謝辞	27
	参考文献	27
	付録	A-1
A.1	補足資料	A-1
A.2	確率楕円	A-5

1 緒論

近年，ネットワーク技術の発達により，PCのみならず家電製品やロボットなど様々な機器がネットワークに接続されるようになってきている．そして，これらの機器がお互いに情報を通信しあうことによって，ユーザが高度なサービスを簡単に利用できるユビキタス環境の実現に向けて研究が行われている．このような環境では，ユーザごとに適応したきめ細かなサービスが提供されることが望ましい．そのためにはどのユーザが，いつ，どこで，どの機器を使っているか，などのユーザについての情報を各機器が把握しなければならず，個人識別が必要不可欠となる．

個人識別では，それぞれ目的に応じて適用される手法が異なる．例えば，家庭の電子錠や銀行口座の管理など高いセキュリティが求められる場合では，ユーザがカード等の物体を携帯することによる負担やパスワードを入力するといったような機器操作による負担は一般的に許容されている．一方，家庭において家電製品が個人に適応したサービスを提供をする場合などには，ユーザにあまり煩わしさを感じさせないことが望まれている．

このようなユーザへの負担が少ない個人識別手法として，人間の日常的な動作である歩行に着目し，歩行中の1歩の間に床にかかる力であるフットステップデータをもとに識別する手法が提案されている [1][2]．しかしながら従来研究においては，フットステップデータの持つ性質やその識別性能などについて十分に明らかにされてこなかった．そこで本研究ではフットステップデータを個人識別に用いるための基礎的な知見を得ることを目的とする．まずフットステップデータを収集する環境の違いによりデータがどのような影響を受けるかについて調べ，定常的な歩行動作中のフットステップデータを収集する環境を構築する．次に収集されたデータについてその識別能力の検討を行う．

以下，第2章ではユビキタス環境における個人識別とその課題について検討し，フットステップデータを利用した個人識別手法の特性について述べる．第3章では歩行の収集方法がフットステップデータに与える影響を考察し，それに基づいて構築したフットステップデータ収集システムについて述べる．第4章では実験によって得られたフットステップデータについて評価を行う．第5章では本研究のまとめを行う．

2 背景

2.1 ユビキタス環境における個人識別の問題点

一般にユビキタス環境においてサービスを提供する場合には、ユーザにとって負担が少ない個人識別システムが望まれている。ユビキタス環境の個人識別はRFID(radio frequency identification) タグや赤外線タグなどの徽章を用いるもの [3], 虹彩, 指紋, 顔画像, 音声などの個人の属性を基に識別するもの [4][5], パスワード入力によるものなどが実用化に向けて研究が進められ, あるいは既に実用に供されている。RFID タグや赤外線タグを利用したものは, それぞれのタグが固有の情報を保持しているため, ユーザがタグを携帯していれば当該ユーザを直接特定することが可能であるという利点を持っている。しかしユーザがタグを常に携帯する必要がある。一方で, 虹彩や指紋, 顔画像, 音声などの個人属性を用いた識別システムは, 認証機器を設置するだけなのでユーザにとって物体の携帯という負担はないが, 識別を行うためにユーザ自らが機器の前に移動し, 操作を行うという負担がある。また, 個人属性の情報をを用いるためプライバシーの観点からユーザに心理的負担を与える。またパスワード入力ではユーザが機器の操作を行わなければならないのに加え, さらにユーザがパスワードを記憶していなければならないという負担もある。

これらのことからユーザに低負担な識別システムを実現するには, ユーザに特別な行為は要求しないという物理的負担の低減と, 個人識別に用いるユーザの個人属性は少なくするという心理的負担の低減が重要であるということができる。フットステップデータを利用した個人識別手法では, 床に配置した圧力センサの上をユーザが踏み, 通り抜けるだけで個人識別ができるので, ユーザの物理的な負担がない。また, 識別に用いる個人属性はユーザの歩行に関するデータのみなので, ユーザにとって心理的負担が少ないと考えられる。次節ではフットステップデータを利用した個人識別手法について述べる。

2.2 フットステップデータを利用した個人識別手法

人間が歩行し, 圧力センサを片足で踏み, 通り抜けた時に得られる力の時系列データ例を図1に示す。横軸は時間を示しており, 縦軸は各時間で床にかかる垂直方向の力の大きさを示している。人間の歩行と床に対する足の付き方の関係を図を基に考える。まず床に踵が着地した直後に衝撃による力が作用して前

半の極大値が現れる．ついで次第に足裏全体が床に着地して重心が移動する間に床にかかる力が弱くなり極小値が現れる．その後踵の方から次第に床から離れるにつれて床にかかる力が大きくなり，爪先で床を蹴り上げる時に後半の極大値が現れる．最後に足が床から離れるにつれ床にかかる力が弱まり，足が床から離れた時に力は0となる．以上に述べた，足が着地してから離れるまでの力の，前半の極大値を前半極大値，極小値を中間極小値，後半の極大値を後半極大値と呼ぶことにする．またそれぞれの着地開始からの時間を前半極大時間，中間極小時間，後半極大時間と呼ぶことにする．また着地開始から離れるまでの時間を接地時間と呼ぶことにする．本研究では図1に示したような，歩行時に片方の足が床に着地を開始してから離れるまでの間に床面に与える垂直方向の力の変化を表す時系列データのことをフットステップデータと呼ぶ．

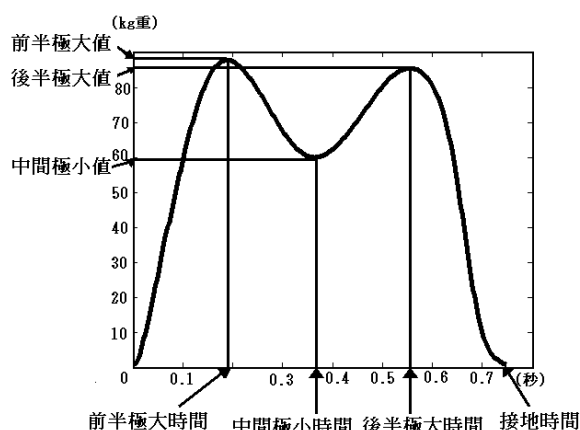


図1: フットステップデータ例

Addleseeらはフットステップデータを用いた個人識別手法を提案した [1]．彼らの研究ではフットステップデータを収集するために縦・横 50cm の 9 枚の板を 3 × 3 の正方形に床の上に配置し，9 枚の板のうち中心の板の下四隅と床との間に圧力センサとしてロードセルを設置し，残りの 8 枚は中心の板と床からの高さが等しくなるようにしている．板の上にはカーペットが敷き詰められ，周りには傾斜路が付けられている．被験者には中心の板の境界を踏まないように，中心の板を踏んで歩くように指示が与えられた．このようにして収集したデータに隠れマルコフモデルを用いた認識手法を適用した結果，被験者 15 人に対して認識率 91%を得た．また Orr らは縦・横 50cm からなる 1 枚の板の下四隅と床との間にロードセルを設置し，板上にカーペットを敷き，被験者にはこの上を

踏んで通り抜けるように指示を与えた．このようにして収集したフットステップデータから前半・後半の極大値やその間にある極小値，時間などの特徴量を抽出し，最近傍法を用いることによって個人識別を行った結果，被験者 15 人に対して認識率 93%を得た [2]．

これらの先行研究より，フットステップデータによる個人識別の可能性が示されている．しかしながら少数の標本で認識率をもとめる場合，一般的に被験者の選び方によって結果は異なる．たとえば被験者の体重が異なれば，床にかかる力の大きさも変わるため，被験者間の体重のバラつきが認識率に影響を与えることが推測される．

また，Addlesee らの研究と Orr らの研究とでは，データを収集した環境が異なっている．Addlesee らの研究ではロードセルを設置してデータを計測する床の表面の高さと周囲の床との高さと同じになるように環境が構築されているが，Orr らの研究では計測する床の表面と周囲の床との高さが異なっている．このことによって歩行形態が変化する可能性があるが，その影響について詳細に調査されてはいない．フットステップデータが持つ個人識別能力を分析するにあたっては，データを収集する環境の与える影響についても精査する必要がある．

3 フットステップデータ収集システムの構築

3.1 実験装置

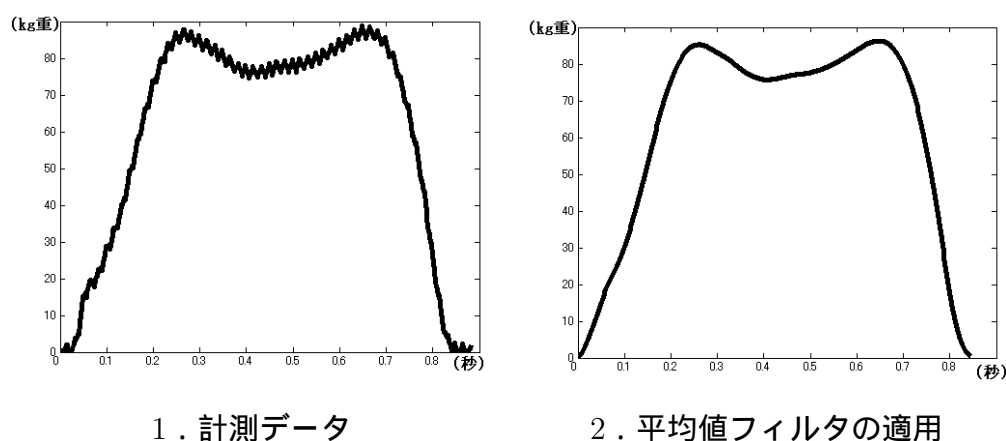


図 2: ノイズとその処理

本節では，フットステップデータを収集するのに必要な機材について述べる．本研究では，人のフットステップデータを収集するため，ティアック電子計測社製の低床ロードセル式台はかり (TL-PS) を使用した．これは縦 50cm，横 50cm，

厚さ 3cm の匡体中に 4 つのロードセルが埋め込まれたものである．定格容量は 200kg 重であり，誤差 200g 重の範囲で計測が可能である．原理的には Addsee や Orr の研究と同様にロードセルを用いるものであるが，はかりとして構成されているためデータ収集時の環境の設定が容易である．台はかりに力が加わるとその力の台はかり表面の垂直方向の成分にそれに比例した電圧が台はかりから出力される．出力された電圧は CONTEC 社製の PC ボード (ADI12-16) を通じ PC 内に取り込まれる．ボード上から PC 内に取り込むドライバは CONTEC 社製の API-PAC(W32) を用いた．使用した PC の CPU は Pentium4(1700MHz) で OS は Windows2000 である．

フットステップデータは毎秒 500 個のサンプリング間隔で 10 秒間収集し，片足が台はかりに力を与えている時のデータのみを切り出した．図 2-1 にこれらの機材を利用して得られるフットステップデータ例を示す．計測された生データには 60Hz の周期的なノイズが乗っていたため，平均値フィルタを用いることにより，平均化処理を行った．その結果を図 2-2 に示す．

3.2 フットステップ収集環境がフットステップデータに与える影響

3.2.1 段差が与える影響

Orr らの研究ではフットステップデータを計測する床の表面の高さが周囲の床より高い状況でデータを収集している．それに対し Addsee らの研究では計測する床の表面の高さが周囲の床の高さと等しくなるよう考慮された環境でデータを収集しているが，その効果についての言及はない．そこで計測する床の段差がフットステップデータに与える影響について調べるため以下のような実験を行った．

定常的な歩行におけるフットステップデータを収集するために，4 名の被験者に対し約 2m 手前から歩き始め，台はかりを片足で踏み通り抜けるように指示を行った．また本研究では履物の効果を取り除くため，全ての収録で被験者は裸足で歩くように指示を行った．このとき段差の影響を調べるため次の 2 つの条件を設定した．

- (i) 台はかりを実験室の床の上に設置．台はかりの表面の高さが周囲の床の高さよりも台はかりの厚み 3cm 分だけ高くなる．
- (ii) 台はかりを実験室の床の上に設置．台はかりに対して歩行経路の前後に各 1 枚ずつ，台はかりと同じ厚み 3cm を持つ縦 150cm 横 80cm の板を台

はかりの端と短辺が接するように設置．台はかりで力を計測する前後で床の高さが一定となる．

表 1: 各特徴量の平均値

被験者	前半極大値 (kg 重)		中間極小値 (kg 重)		後半極大値 (kg 重)	
	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)
a	60.5	60.6	52.0	53.0	58.3	60.9
b	78.7	83.0	53.6	51.9	68.8	73.7
c	75.0	60.6	43.2	43.1	53.4	57.7
d	57.9	58.3	45.6	48.4	58.0	62.3

	前半極大時間 (秒)		中間極小時間 (秒)		後半極大時間 (秒)		接地時間 (秒)	
	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)
a	0.172	0.167	0.330	0.292	0.628	0.584	0.872	0.828
b	0.158	0.142	0.327	0.300	0.579	0.520	0.771	0.690
c	0.166	0.166	0.355	0.330	0.567	0.525	0.742	0.685
d	0.197	0.225	0.411	0.390	0.590	0.565	0.785	0.769

表 2: 各特徴量の標準偏差

	前半極大値 (kg 重)		中間極小値 (kg 重)		後半極大値 (kg 重)	
	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)
a	1.54	1.22	0.85	1.43	1.28	1.16
b	4.32	3.19	2.72	2.15	1.70	3.21
c	1.76	3.19	3.48	1.85	2.53	2.05
d	2.05	1.10	1.32	0.83	1.64	1.14

	前半極大時間 (秒)		中間極小時間 (秒)		後半極大時間 (秒)		接地時間 (秒)	
	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)	条件 (i)	条件 (ii)
a	0.018	0.034	0.089	0.035	0.065	0.021	0.026	0.018
b	0.018	0.008	0.060	0.034	0.034	0.022	0.036	0.018
c	0.016	0.013	0.044	0.020	0.029	0.020	0.084	0.023
d	0.014	0.011	0.049	0.025	0.028	0.017	0.033	0.016

被験者 4 名 (a, b, c, d) に対して上記の 2 つの条件でフットステップデータの収録を行った．実験では各被験者毎にそれぞれの条件に対して左右の足各 10 回

ずつ台はかりを踏むように歩行し、計 160 個のフットステップデータを得た。

Orr らの研究で特徴量として用いている前半極大値、中間極小値、後半極大値、前半極大時間、中間極小時間、後半極大時間、接地時間に対して (i)(ii) の条件でどのような違いがあるか比較した。それぞれの特徴量について各個人における平均を表 1 に、各個人における標準偏差を表 2 に示す。

段差の影響が無ければ各特徴量の分布関数は条件 (i)(ii) で同一のはずである。そこで分布関数を正規分布と仮定し、各被験者毎に各特徴量の母分散が等しいと仮説を立て、危険率 5% で検定を行った。この結果を表 3 に示す。また、分散について仮説が棄却されなかった場合について、母平均が等しいと仮説を立て危険率 5% で検定を行った。この結果を表 4 に示す。

表 3: 母分散に関する検定結果。ただし表中は (棄却数)/(対象数) を表す

前半極大値	中間極小値	後半極大値	前半極大時間	中間極小時間	後半極大時間	接地時間
3/4	3/4	1/4	2/4	4/4	4/4	4/4

表 4: 平均値に関する検定結果。ただし表中は (棄却数)/(対象数) を表す

前半極大値	中間極小値	後半極大値	前半極大時間	中間極小時間	後半極大時間	接地時間
1/1	1/1	3/3	0/2	0/0	0/0	0/0

表 3 の結果より、検定を行った 28 例のうち 21 例で分散が等しいという仮説が棄却された。また表 4 より分散が等しいという仮説が採択された 7 例も 5 例で平均値が等しいという仮説が棄却された。このことから段差の存在がフットステップデータに影響を与えることが分かった。

分散が等しくないので検定はできないが、表 1 を見ると中間極小時間、後半極大時間、接地時間は段差がある時の方が無い時に比べて長くなっている。これは段差がある時は床から台はかりに上がる、降りるといった動作が加わるため、通常の歩行動作が妨げられていることが原因と考えられる。また後半極大値は段差がある時の方が無い時に比べて小さくなっており、台はかりの上から床に降りる動作では、被験者が床を強く蹴りだす必要がないことが原因であると考えられる。

このように計測時に段差が存在する場合は歩行の形態が変容している。上記の実験により、自然で日常的な歩行動作中のフットステップデータを収集するには台はかりの表面と周囲の高さが一定になるようにデータの収集環境を構築しなければならないことが分かる。

(ii) の条件もとで 20 代から 50 代の男性 32 名女性 10 名合計 42 名の被験者に対し実験を行った。被験者はそれぞれの足で各 10 回ずつ台はかりを踏むように歩行し、その結果 840 個のフットステップデータを得た。この結果による各特徴量の全体の平均、標準偏差を表 5、6 に示す。また、各被験者毎について前半極大値、中間極小値、後半極大値のレンジを求めた。レンジとは各特徴量について最大の要素と最小の要素の差のことである。各特徴量について被験者全体のレンジの平均と標準偏差を表 7 に示す。

表 5: 各特徴量の平均

前半極大値 (kg 重)	中間極小値 (kg 重)	後半極大値 (kg 重)
68.0	53.8	67.2

前半極大時間 (秒)	中間極小時間 (秒)	後半極大時間 (秒)	接地時間 (秒)
0.195	0.343	0.531	0.727

表 6: 各特徴量の標準偏差

前半極大値 (kg 重)	中間極小値 (kg 重)	後半極大値 (kg 重)
2.09	1.30	1.65

前半極大時間 (秒)	中間極小時間 (秒)	後半極大時間 (秒)	接地時間 (秒)
0.016	0.025	0.023	0.021

表 7: 極値のレンジの平均と標準偏差

	前半極大値 (kg 重)	中間極小値 (kg 重)	後半極大値 (kg 重)
平均	5.90	3.88	4.84
標準偏差	3.22	1.92	1.90

3.2.2 接地位置を意識することによる影響

前節で行った実験では台はかりを 1 枚だけ使用したため、確実にその 1 枚を踏まなければならないという意識が被験者に働く。実際に実験では被験者が台はかりの直前で一瞬止まったり、歩幅を急激に変化させる状況が観察され、また被験者からも足の接地位置を意識してしまったという意見があった。この意識により自然な歩行動作が妨げられ、フットステップデータに影響を与える可能性がある。このような接地位置への意識を取り除くために以下のような収集環境を構築した。

図 3 のように直線状に 9 枚の台はかりを互いに接するように実験室の床上に

配置する．この時実験室の床と台はかりの表面は台はかりの厚み 3cm 分だけ高さに差がある．各台はかりの上に台はかりと同じ 50cm × 50cm のカーペットを敷き，台はかりの間の境界が被験者に見えないようにする．これにより被験者は台はかりの上を直線状に歩行する場合には足の接地位置を意識することが無いものと考えられる．9 枚の台はかりから出力される電圧は同期して PC 内に取り込まれ記録される．

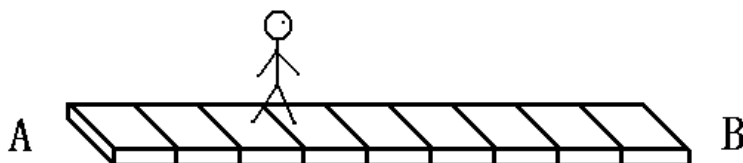


図 3: 9 枚の台はかりを利用したデータ収録方法

被験者が足の接地位置を意識した場合と意識しない場合のフットステップデータを比較するために 20 代から 30 代の被験者 26 名に対し，条件 (iii)(iv) のもとで実験を行った．

- (iii) 図 3 の 9 枚の台はかりのうち中心の 1 枚の周りに目印用のテープを貼付．被験者は A 端から B 端に向かって歩き，目印のついた台はかりの上を片足で踏み歩行するように指示．
- (iv) 図 3 で A 端から歩き始め，1 枚目の台はかりを片足のみで踏み歩行する．そのまま台はかりの上を直線状に歩行し，端まで進むとそのまま B 端におりる．B 端からまた台はかりに上がりそのまま直線状に歩行し A 端まで往復するように被験者に指示．

条件 (iii) のもと被験者が台はかりの中心を左右の足で各 10 回ずつ踏み歩行するデータを得たが，台はかりの端を踏むなどして左右どちらかの足で 10 回のフットステップデータを得られない被験者がいたため，その被験者を除いた 20 名計 400 個のフットステップデータを実験に用いた．

条件 (iv) では 1 回の計測時間を 10 秒とし，被験者はそれぞれ 4 回台はかりの上を往復した．このとき 9 枚の台はかりの出力を次節に述べる方法で統合し，497 個のフットステップデータを得た．この中から条件 (iii) で実験対象とした 20 名の被験者のフットステップデータをそれぞれ 10 個ずつ，計 200 個のフットステップデータを選び，(iii) で得られたデータとの比較を行った．

前節同様に，前半極大値，中間極小値，後半極大値，前半極大時間，中間極

表 8: 各特徴量の平均値例

被験者	前半極大値 (kg 重)		中間極小値 (kg 重)		後半極大値 (kg 重)	
	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)
1	51.0	52.9	40.3	37.3	57.3	58.0
2	75.9	72.5	54.3	55.2	74.7	73.3
3	67.6	65.2	46.8	48.2	69.6	69.6

	前半極大時間 (秒)		中間極小時間 (秒)		後半極大時間 (秒)		接地時間 (秒)	
	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)
1	0.183	0.160	0.317	0.318	0.530	0.495	0.715	0.666
2	0.174	0.175	0.330	0.329	0.482	0.481	0.638	0.640
3	0.165	0.179	0.307	0.319	0.520	0.524	0.679	0.697

表 9: 各特徴量の標準偏差例

	前半極大値 (kg 重)		中間極小値 (kg 重)		後半極大値 (kg 重)	
	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)
1	2.03	2.69	2.07	1.70	1.79	4.43
2	3.07	0.84	1.57	1.05	1.85	2.23
3	3.42	1.35	2.41	0.84	2.75	1.90

	前半極大時間 (秒)		中間極小時間 (秒)		後半極大時間 (秒)		接地時間 (秒)	
	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)
1	0.030	0.019	0.024	0.019	0.033	0.018	0.041	0.026
2	0.009	0.005	0.011	0.010	0.010	0.011	0.012	0.010
3	0.010	0.010	0.018	0.007	0.010	0.009	0.011	0.013

小時間，後半極大時間，接地時間を特徴量として比較した．各条件におけるそれぞれの特徴量について表 8 に各個人における平均例を，表 9 に標準偏差例を示す．20 名の平均と標準偏差は付録 A.1 に示す．被験者の意識がフットステップデータに影響を与えなければ，各特徴量の分布関数は条件 (iii)(iv) のもとで同一のはずである．そこで分布関数を正規分布であると仮定し，母分散が等しいと仮説を立て危険率 5% で検定を行った．この結果を表 10 に示す．また，分散について仮説が採択された場合について，母平均が等しいと仮説を立て危険率 5% で検定を行った．この結果を表 11 に示す．

表 10: 母分散に関する検定結果．ただし表中は (棄却数)/(対象数) を表す

前半極大値	中間極小値	後半極大値	前半極大時間	中間極小時間	後半極大時間	接地時間
9/20	7/20	5/20	3/20	8/20	4/20	4/20

表 11: 平均値に関する検定結果．ただし表中は (棄却数)/(対象数) を表す

前半極大値	中間極小値	後半極大値	前半極大時間	中間極小時間	後半極大時間	接地時間
5/11	7/13	2/15	8/17	1/12	10/16	10/16

表 10 の結果より，検定を行った 140 例のうち 40 例で分散が等しいという仮説が棄却された．また表 11 より分散が等しいという仮説が採択された 100 例も 43 例で平均値が等しいという仮説が棄却された．このことから被験者が接地位置を意識した場合フットステップデータに影響を与えることが分かった．

表 11 を見ると，前半極大値の分散が異なる被験者が多く，後半極大時間や接地時間における平均値が異なる被験者が多いことが分かる．このことは台はかりに対する被験者の意識が歩行動作に影響を与えたためだと考えられる．

3.2.3 複数枚の台はかりを利用した時のフットステップデータ抽出手順

本節では前節 (iv) の条件もとの 9 枚の台はかりの出力データからフットステップデータを抽出する手順を示す．

本来フットステップデータは 1 枚の台はかりを片足のみが踏むことを前提としている．このため複数の台はかりを利用したフットステップデータ収集では次の 2 つの場合で問題が生じる．

第 1 の問題は 1 つの台はかりに両足が乗る場合である．歩行時には両足ともに床に接地している時間があり，両足とも同一の台はかりに接地しているときに，このような問題が発生する．この場合は 1 つの台はかりの出力に両足の力が重畳されて計測されるので，片足のみのフットステップデータを得ることは難しい．

第 2 の問題は台はかりの境界を踏む場合である．この場合には片足が垂直方向に床にかける力が 2 枚の台はかりに分かれるため，1 枚の台はかりの出力のみではフットステップデータを得ることができない．しかしながら 2 枚の台はかりの境界に置かれた物体の垂直方向に床にかける力は，各台はかりの値を加算することで得られるため，隣り合う 2 枚の台はかりの出力からフットステップデータを得ることができる．但し，この場合 2 枚の台はかりを合わせて 1 枚の台はかりとみなしていることになり，両足が同時に 2 枚の台はかりの中に乗っ

ている場面があると第1の問題と同様の理由でフットステップデータを得ることができない。

このように1枚の台はかりの出力データあるいは隣り合う2枚の台はかりの出力を加算したデータにはフットステップデータでないものが含まれているため、その中からフットステップデータを区別し選択しなければならない。

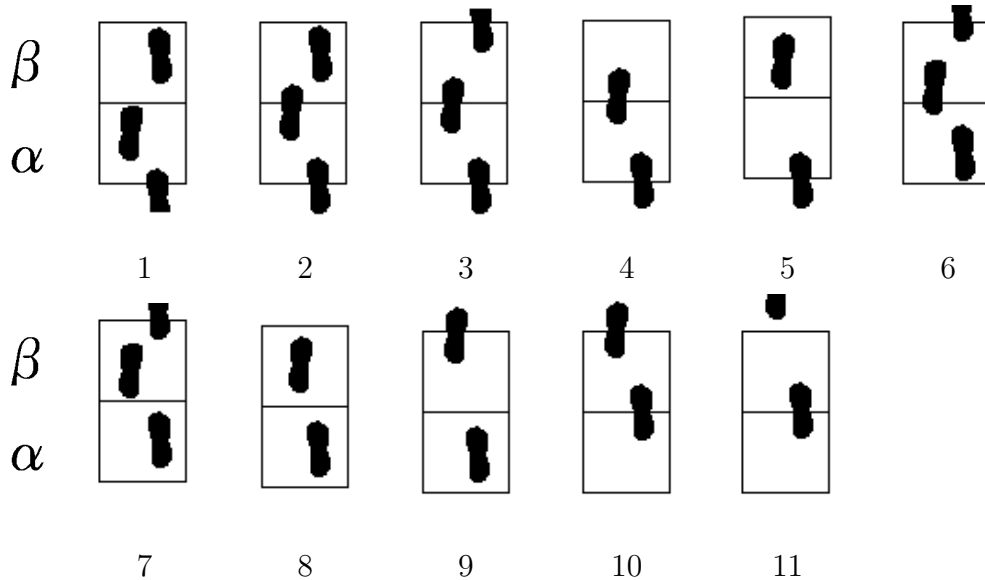


図4: 台はかりの踏み方1

まずどのようなデータが得られるかを調べるために、4名の被験者に対し3.2.2節の条件(iv)のもとで実験を行った。被験者はそれぞれ4回往復し、各台はかり毎に10秒間床にかかる力の時系列データを収集した。収集したデータ及び収集時の振る舞いから、

- 1枚目の台はかりは片足のみで踏まれる。
- 歩行中両足とも1枚の台はかりの内部に接地する場合は見られない。
- 足が接地されない台はかりはない。

との知見を得た。被験者の足の大きさを25cm、歩幅を被験者毎にほぼ一定であるという仮定をおくと、得られた知見より被験者の歩幅は約40cm~70cmとなる。2つの台はかりに着目した時、この条件のもとでの足の付き方は図4に示す11種類に限られる。図中下方の台はかりを α 、上方の台はかりを β とする。被験者は α から β に向かって歩行するものとする。このときの台はかりの出力について検討する。

1枚の台はかりの出力がそのままフットステップデータとなる場合を考える。

図4の1~11の踏み方のうち、1枚の台はかりの踏み方として出現するのは図5のa~fに示す通りである。この中でaの踏み方に対応する出力がフットステップデータとなる。この踏み方をしているのは図4では1の β 、5の β 、7の α 、8の α, β 、9の α である。

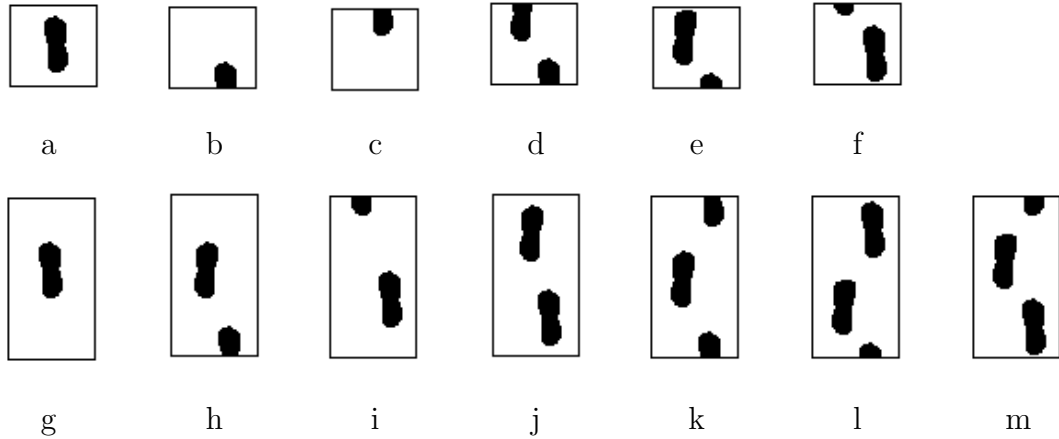


図5: 台はかりの踏み方2

次に2枚の台はかりの出力を加算したものがフットステップデータとなる場合を考える。1~11の踏み方のうち、2枚の台はかり全体を1枚の台はかりとみなした時の踏み方を図5のg~mに示す。この中でgの踏み方に対応する出力がフットステップデータとなる。図4中の11がこの踏み方に該当する。

片足の付き方は台はかりの内部を踏む場合か境界を踏む場合しかないので、片足の1回の接地での踏み方は足全体で踏む(図5のa)、爪先で踏む(図5のb)、踵で踏む(図5のc)の3通りに分けられる。よってd~mの踏み方は全てa~cの踏み方を組み合わせることで表現できる。このことからd~mの踏み方をする時に台はかりから出力されるデータは、a~cの台はかりから出力されるデータを組み合わせることで表現できると思われる。表12にd~mの踏み方に対するa~cの踏み方の組み合わせを示す。また図6にa~cの踏み方に対応したデータ例とa~cのデータ例を組み合わせることができるd~mの踏み方に対応したデータ例を示す。但し実線が出力データで点線が組み合わせられたa~cのデータ例である。また、2枚の台はかりを1枚とみなした時にgはaと、hはeと、fはiと台はかりの踏み方が一致するので、それぞれ同じパターンの出力となる。

a~mのパターンのうちパターンa(g)が正しいフットステップデータに対応する。大量のデータからパターンa(g)の全てを視認によって抽出するのは実験者に大きな負担となる。そこで3.2.1節で得られた840個のフットステップデー

タの統計量を基にパターン a(g) を自動化で抽出した．抽出手法は以下の通りである．図 7 に手法の処理の流れを示す．

表 12: 台はかりの踏み方の組み合わせ

踏み方	d	e(h)	f(i)	j	k	l	m
基本ステップ	b,c	b,a	a,c	a,a	b,a,c	b,a,a	a,a,c

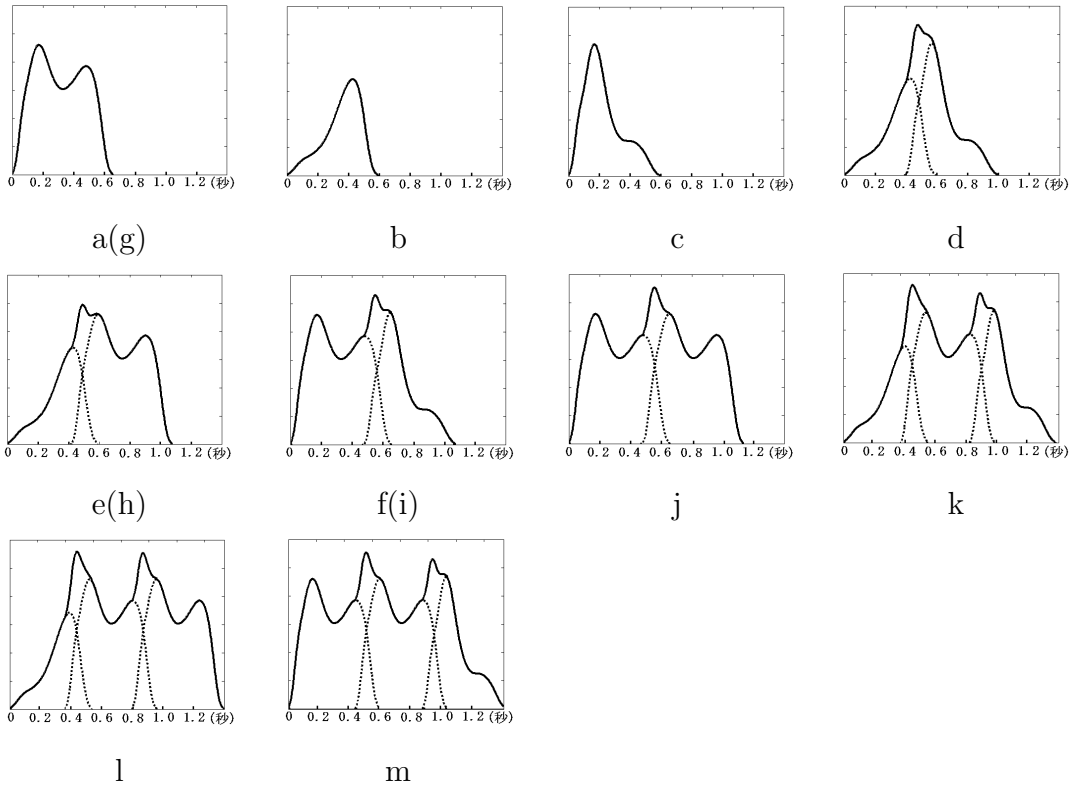


図 6: データ例

1. 接地時間による抽出

図 6 の j ~ m のデータは 1 枚とみなした台はかりの内部に両足が乗っている時に計測されるデータで，正しいフットステップデータに比べて接地時間が長い．また，b で爪先が少しだけしか入っていない場合や，c で踵が少しだけしか入っていない場合は接地時間が短い．このように得られたデータには接地時間からフットステップデータでないと分かるものがあり，これらを除去する．840 個のフットステップデータ中，最短接地時間は 0.58 秒，最長接地時間は 0.93 秒であったことから接地時間 0.55 秒以下のデータおよび 0.95 秒以上のデータはフットステップデータでないと除去する．この抽出方法を適用すると，e(h)，f(i) の一部も接地時間が長いいため除去されることがある．

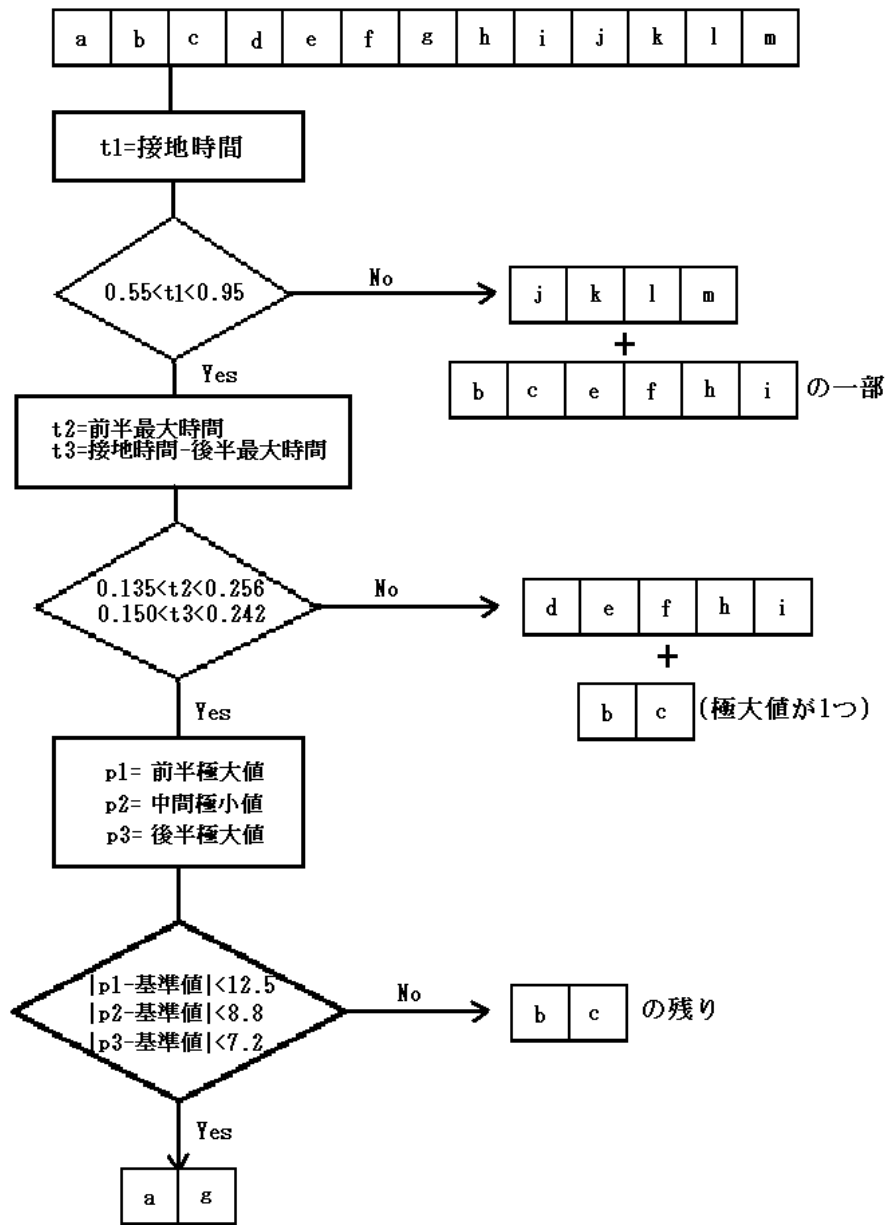


図 7: 自動化の手順

2. 前半極大時間，後半極大時間と接地時間の差による抽出

3.2.1 節の結果から前半極大時間は接地時間の半分より小さく，後半極大時間は接地時間の半分より大きいことが示されている．この知見を用いてフットステップデータを抽出することを考える．しかしながら図 6 の b, c の一部や d の場合には極大値が 1 つしか存在せず，j, l, m の場合には 3 つ以上の極大値が存在するため，前半極大や後半極大を見出すことが難しい．そこで時系列デー

データを接地時間の前半分と後半分に分割して、データの前半側の最大値を前半最大値、その時の時間を前半最大時間、後半側の最大値を後半最大値、その時の時間を後半最大時間と呼ぶこととする。また前半最大時間と後半最大時間の間の最小となる値を中間最小値、その時の時間を中間最小時間と呼ぶこととする。前半最大値、中間最小値、後半最大値は入力データがフットステップデータの時はそれぞれ前半極小値、中間極小値、後半極大値と一致する。

このとき図6のe(h)は前半部分で台はかりの境界を踏んでいるデータが重畳されているため、前半最大時間が長くなっている。また図6のf(i)では後半部分で境界を踏んだデータが重畳されているので、後半最大時間と接地時間の差が長くなっている。このように前半最大時間や後半最大時間と接地時間の差からフットステップデータでないと分かるものを除去する。

3.2.1節で得られた840個のデータから求められた前半極大時間の分布が正規分布であると仮定し、信頼係数95%において推定を行うと信頼区間は0.135～0.256秒となった。また、後半極大値時間と接地時間の差の分布について同様に正規分布と仮定し、信頼係数95%で区間推定を行うと0.150～0.242秒となった。フットステップデータであるならば、前半最大時間は0.135～0.256秒の区間に、後半最大時間と接地時間の差は0.150～0.242秒の区間に入る可能性が大きい。そこでこの区間に入らないデータを除去する。

図6のdのように極大値が1つしかないデータやb,cのうち極大値が1つしかないデータでは、前半最大時間、後半最大時間を求めるとどちらか一方は接地時間の半分となる。なぜなら、前半側に極大値がある場合は後半側は単調減少していることになる。そのため接地時間の半分が後半最大時間となる。後半側に極大値がある場合は前半側が単調増加しているため、接地時間の半分が前半最大時間となる。前半最大時間が接地時間の半分の時、1で述べた抽出方法よりデータの接地時間は0.55秒より長いので、前半最大時間は0.275秒より長くなる。この時には前半最大時間の信頼区間に該当しない。また後半最大時間が接地時間の半分の時、後半最大時間と接地時間の差も接地時間の半分となる。この時には後半最大時間と接地時間の差の信頼区間に該当しない。よってdのように極大値が1つしかないデータやb,cのうち極大値が1つしかないデータも除去できる。

3. 前半極大値、中間極小値、後半極大値のレンジによる抽出

図6のbで踵が少しだけ外に出ているデータやcで爪先が少しだけ外に出て

いるデータでは時間に大きな影響はなく、データの極値にその影響が見られる。このように極値からフットステップデータでないと分かるものを除去する。

しかし被験者によってそれぞれの極値の大きさは異なるので、一定の閾値を設定することはできない。そこで各極値の抽出の基準となる値を求め、その値と比べ一定の範囲に入るものを抽出する。基準となる値を得るためにデータ収録時に被験者に台はかり外の実験室の床から歩き始め1枚目の台はかりを片方の足のみで踏み歩行するように指示する。この1歩目は段差の影響を受けてはいるが、フットステップデータとして切り出すことが可能である。このデータから得られた極値をそれぞれの極値の基準値とした。また、範囲の閾値にはレンジを用いた。前節で得られた42人の各極値のレンジが正規分布であると仮定し、信頼係数95%で区間推定を行うと、前半極大値のレンジは12.5、後半極大値のレンジは8.8、中間極小値のレンジは7.2となった。各極値の基準値からそれに対応する極値のレンジの推定区間に入るものをフットステップデータとして抽出する。

前節の条件(iv)で得られた2094個のデータに対してこの抽出方法でフットステップデータの抽出を行い、その結果508個のデータが得られた。この508個のデータを視認したところ、パターンe(h)及びf(i)のデータが11個含まれていた。この11個のデータは境界を踏むことで重畳されるデータの影響が小さいため閾値処理では除去できなかったものである。この11個のデータを除き497個のフットステップデータを得た。前節ではここで得られた497個のデータを検定に用いている。自動化により収集可能であるフットステップデータが捨てられている可能性があるが、誤データを減らすことを目的としたため、この点に関しては考慮しないことにした。

4 多変量解析を用いた特徴量の識別能力の分析

4.1 フットステップデータの識別能力

本章ではフットステップデータが持つ識別能力について検討する。

前章までの結果から自然な定常的歩行であると考えられる3.2.2節の条件(iv)のもとで収集されたデータについて分析を行う。得られた497個のフットステップデータのうち、26名のフットステップデータをそれぞれ10個ずつ、計260個のフットステップデータを分析対象とする。フットステップデータはそれぞれ接地時間が異なるためそれらを直接比較することができない。フットステップ

データの接地時間は最長 0.95 秒であるため，接地時間以降の値を 0 として足の着地開始時刻から 1 秒間，サンプリング間隔が毎秒 500 個であるので，結局 500 次元のデータをフットステップデータとして分析を行うことにする．

フットステップデータの識別能力は各個人毎のデータのバラつき方と，個人間のバラつき方の関係による．このことを模式的に示した図 8 を用いて説明する．被験者 a と被験者 b に帰属するデータが二次元上に分布しており，それぞれとで示す．1 では a と b のデータが混在して識別は容易ではない．しかし 2 は a と b の平均位置が 1 と同じであるが，それぞれの分散が小さいので識別可能である．3 の a と b の分散は 1 と同じであるが，a と b の平均位置が離れたため識別可能である．このようにフットステップデータの識別能力について調べるには，フットステップデータのバラつき方，すなわち平均と分散を知る必要がある．

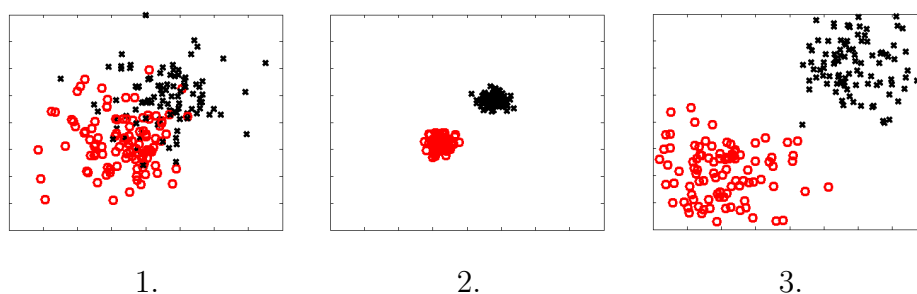
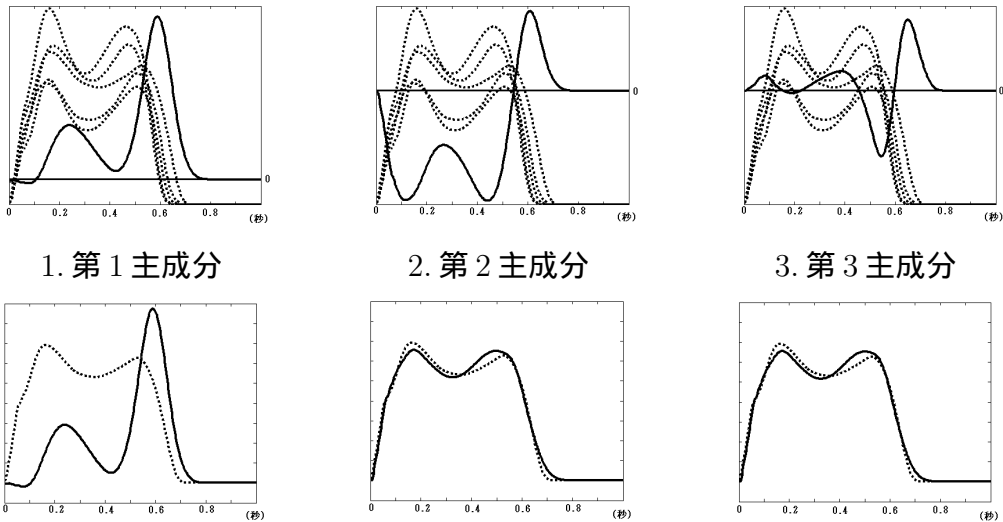


図 8: データ分布の模式図

そこで分散共分散行列を求め主成分分析を行い，分散について多くの情報を持つ軸で識別能力について調べる．主成分分析の結果，第 1 主成分の寄与率は 56.1%，第 2 主成分の寄与率は 29.9%，第 3 主成分の寄与率は 6.9% となり，ここまでの寄与率を合計すると合計で 90% を超える．そこで第 3 主成分までの 3 軸を用い，3 次元データとして分析する．図 9 の 1~3 はフットステップデータと各主成分の関係を示すため，フットステップデータ例と各主成分を重ねて描画したグラフである．ただし，点線はフットステップデータ例，実線は各主成分を表す．また 4~6 に各主成分の寄与を見るために第 1 主成分から第 3 主成分までの近似を示す．

フットステップデータの識別能力を議論するために，仮想的なユーザの集合を考え，



4. 第 1 主成分までの近似 5. 第 2 主成分までの近似 6. 第 3 主成分までの近似

図 9: 主成分分析

- 各ユーザのデータが全体としてどのように分布しているか。
- ユーザごとのデータがどのように分布しているか。

について定める．各ユーザがそれぞれの分布関数に応じてフットステップデータを生成するとする．被験者 26 名の 260 個のフットステップデータの平均 G_1 ，分散共分散行列 Σ_1 をユーザ集合に属するユーザの生成する全フットステップデータの母集団の平均，分散とみなし，分布関数が正規分布 $N(G_1, \Sigma_1)$ であると仮定する．次に簡単のためにユーザ毎のデータの分布がユーザ毎に平均を除いて同一であるとする．この分布関数を求めるために，被験者毎にデータの平均を求め，平均位置が原点となるように平行移動し分散共分散行列 Σ_2 を求める．この分散共分散行列 Σ_2 の値を持つ正規分布 $N(0, \Sigma_2)$ を各ユーザの平均周辺分布と仮定する．そして利用環境で識別対象となる n 人のユーザに対し，それぞれのユーザのフットステップデータの平均 $g_i (i = 1, \dots, n)$ の分布が正規分布 $N(G_1, \Sigma_1)$ に従うものとする．このときユーザ i のフットステップデータは正規分布 $N(g_i, \Sigma_2)$ に従うとモデル化する．

次に識別について検討する．本研究では簡単のため正規分布での確率楕円を用いることにする．確率楕円について詳細は付録 A.2 で述べるが，確率楕円とは分布の平均と分散共分散行列が与えられた時，その分布関数の積分値が確率 p となるように定められた楕円体のことである．得られたデータが確率楕円 e_i

に属する時そのデータがユーザ i に帰属するものと考えことにし、このときの p を検出率と呼ぶことにする。各ユーザごとに平均 g_i の周辺に分散共分散行列を Σ_2 として確率楕円 e_i を設ける。あるフットステップデータが観測された時、そのデータは次の3つの場合に分けられる。

1. データが1つの確率楕円のみ属する。
2. データがどの確率楕円にも属さない。
3. データが2つ以上の確率楕円に属する。

1の場合属する楕円 e_i のユーザが i のフットステップデータであると識別する。2の場合は該当するユーザが識別対象内に存在しないと見え、識別しない。この発生確率は検出率を p とすると、近似的には $1 - p$ と見積もることができる。本研究では3を誤識別と考え、これが発生する割合を評価する。

各ユーザの確率楕円の和集合により求まる領域の体積を総体積、2つ以上の確率楕円が重なっている部分の体積を重複体積と呼ぶことにする。このとき、総体積に比べて重複体積が大きい時が誤識別の確率が高くなる。そこでこの値を重複度と呼び、本研究ではユーザ数に対する重複度を用いて識別能力を評価する。

$$\text{重複度} = \frac{\text{重複体積}}{\text{総体積}} \quad (1)$$

重複度を求めるためには、任意の位置の確率楕円の体積と重複体積を求める必要がある。しかし解析的に確率楕円の重複体積を求めるのは難しい。そこでモンテカルロ法を適用した。モンテカルロ法を用いれば乱数を用いた確率計算として体積を求めることができる。すなわち一定の空間上にランダムにちりばめられた点から、条件を満たす点の割合を求め、その割合と空間の体積の積を取ることによって、条件を満たす体積を求めることが可能である。

以下本研究での具体的手順を示す。まず26名の被験者から得られたフットステップデータの分散共分散行列を求め主成分分析を行う。本研究では分散共分散行列の上位3つの固有値に対応する固有ベクトルをフットステップデータの主成分とする。3つの主成分に対するフットステップデータの平均 G_1 、分散共分散行列 Σ_1 は次の通りである。これを用いて母集団の分布関数 $N(G_1, \Sigma_1)$ を求める。

$$G_1 = \begin{pmatrix} 576 \\ -762 \\ -19.7 \end{pmatrix} \quad \Sigma_1 = \begin{pmatrix} 31900 & 0 & 0 \\ 0 & 17000 & 0 \\ 0 & 0 & 3943 \end{pmatrix}$$

次にフットステップデータの被験者毎の平均位置を求め、被験者毎の平均が原点になるように平行移動したデータについて分散共分散行列 Σ_2 を求めた。これを用いて各ユーザが持つ分布関数 $N(0, \Sigma_2)$ とする。

$$\Sigma_2 = \begin{pmatrix} 4218 & 2277 & 713 \\ 2277 & 1700 & 795 \\ 713 & 795 & 2093 \end{pmatrix}$$

ユーザの平均 g_i を $N(G_1, \Sigma_1)$ に基づいてランダムに配置し、各ユーザの重心 g_i を決定する。 n 人のユーザの確率楕円の総体積 V と重複体積 M をモンテカルロ法を用いて求め、重複度 P を、

$$P = \frac{M}{V} \quad (2)$$

により求める。

シミュレーションでは検出率を $p = 0.95$ と設定した。すなわち全データに対して5%の確率で検出できないものとする。ユーザ数は2人~10人とした。図10にユーザ数が5人の時に生成される確率楕円例を示す。3つの軸はそれぞれ第1主成分、第2主成分、第3主成分を表す。それぞれのユーザ数に対してユーザを1000通り配置し、各配置毎に重複度を求めた。各ユーザ数に対する重複度の平均を表13に示す。

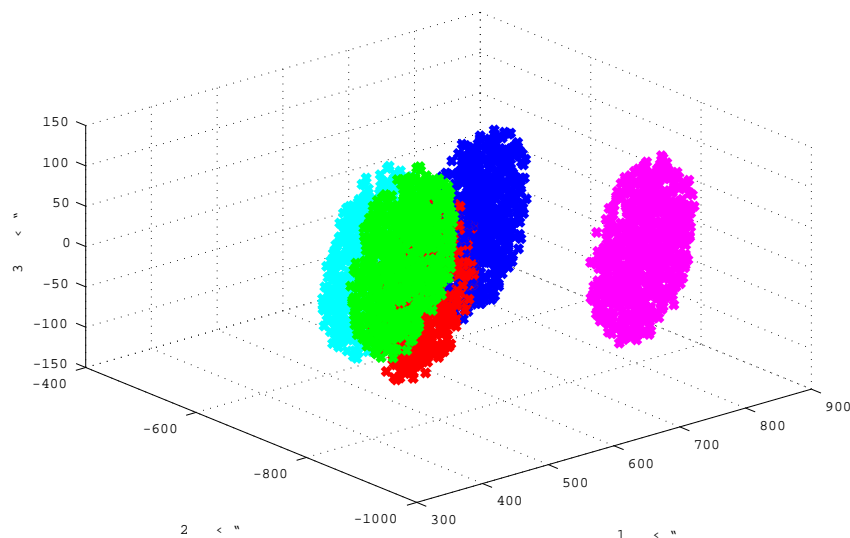


図 10: 確率楕円例

重複度は全確率楕円の占める体積に対して2つ以上の確率楕円が重なる体積

表 13: 重複度の平均

ユーザ数	2	3	4	5	6	7	8	9	10
重複度 (%)	1.3	2.3	3.6	4.8	6.5	6.9	7.4	8.8	10.1

の割合を表すものなので，本研究ではこの値の半分の値を誤識別率とする．表 13 より 10 人以上のユーザ数に対して誤識別率が 5%を超えることが分かる．

4.2 体重以外の特徴量を用いた場合

床にかかる力は体重の影響を受けている．よって，ユーザ間の体重のバラつきが重複度に影響を及ぼすことが予想される．もし体重のみが識別の要因となるならば，同じ体重のユーザを識別できないことになる．そこで本節では体重の違いが識別に与える影響について検討する．

体重の影響を調べるには実際に同じ体重の被験者のフットステップデータを比較することが望ましいが，同じ体重の複数の被験者のデータを収集するのは困難である．そこで前節で分析を行ったデータを体重で除し，前節の結果と比較することで性能を調べることにする．実際には体重の違いが単にフットステップデータの値に比例するだけでなく，歩行形態の変化をもたらすことも考えられるが，本研究では考慮しないことにする．

分析対象となるデータは収集時に被験者の体重を計測していない．そこで体重に代わる特徴量を見出し，フットステップデータをこの値で除することにする．体重に直接関係がある特徴量には前半極大値，中間極小値，後半極大値がある．歩行中に体重が変化することはないので，これらの特徴量のうち最も変動が少ないものを用いることにする．本節では 3.2.2 節の結果中，最も (標準偏差/平均) の低い中間極小値を用いる．すなわちフットステップデータを被験者それぞれの中間極小値の平均で除することで同体重の被験者のデータとみなす．

中間極小値で除することで正規化されたフットステップデータに対し，前節同様に分散共分散行列を求め，主成分分析を行った．第 1 主成分の寄与率は 75.9%，第 2 主成分の寄与率は 11.7%，第 3 主成分の寄与率は 5.7%であり，ここまでの寄与率の合計が 90%を超える．そこで第 3 主成分までの 3 軸を用い 3 次元データとして分析する．図 11 の 1~3 は正規化後のフットステップデータと各主成分の関係を示すため，フットステップデータ例と各主成分を重畳したグラフである．ただし，点線は正規化後のフットステップデータ例，実線は各主成分を表

す．また，4～6に各主成分の寄与を見るために第1主成分から第3主成分までの近似を示す．

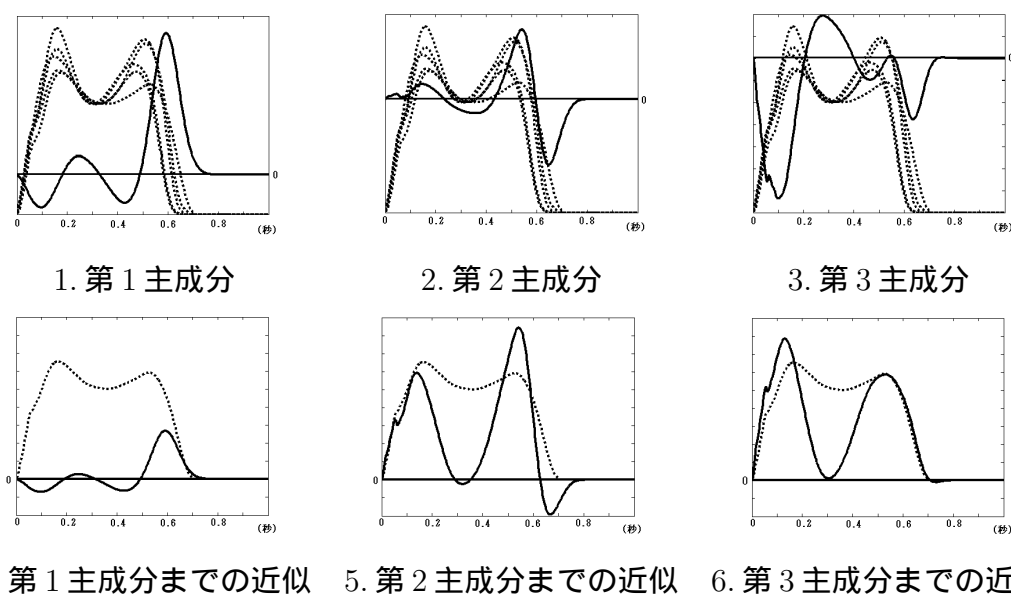


図 11: 正規化後の主成分分析

前節の図9と図11を比較すると第1主成分同士は形状は類似しているが，第2主成分同士は形状は類似していない．また前節の第3主成分と本節の第2主成分の形状が正負を逆転した形で類似している．このことを確かめるため，前節で求めた各主成分と本節で求めた各主成分について内積を求めた．表14は各主成分ごとの内積の値であり，縦軸は前節で求めた主成分を表し，横軸は本節で求めた主成分を表す．この表より前節の第1主成分と本節の第1主成分同士は相関が強いことが分かる．また前節の第2主成分と本節の第2主成分の相関が弱いのに比べ，前節の第3主成分と本節の第2主成分の相関が強いことも分かる．このことから第1主成分は正規化した影響が少なく，体重に依存する情報がほとんどないことがわかる．また前節の第2主成分が体重に関する主成分であり，中間極小値で正規化することでその成分が低減されたと考えることができる．

また図9の4～6と図11の4～6を比較すると，前節では第2主成分までの近似によりフットステップデータを十分に表現できているが，本節では第3主成分までの近似においても十分にフットステップデータを表現しているとはいえない．このことは前節での第2主成分がフットステップデータのパターン全体に

表 14: 主成分同士の内積

前節 \ 本節	第 1 主成分	第 2 主成分	第 3 主成分
第 1 主成分	0.89	0.09	-0.17
第 2 主成分	0.45	-0.26	0.37
第 3 主成分	-0.06	-0.95	-0.26

影響を与えており，正規化によってその効果が減少したためと考えられる．正規化により 2 次の成分が減少していることから，体重成分が識別において重要な要因となっており，正規化により識別性能が低下することが予想される．

前節と同様の手順で，ユーザのモデルを作成する．フットステップデータに対し主成分分析を行った上位 3 次元の主成分に対するフットステップデータの平均 G'_1 ，分散共分散行列 Σ'_1 は次の通りである．これを用いて母集団の分布関数 $N(G'_1, \Sigma'_1)$ を求める．

$$G'_1 = \begin{pmatrix} 2.82 \\ 4.62 \\ -6.10 \end{pmatrix} \quad \Sigma'_1 = \begin{pmatrix} 7.59 & 0 & 0 \\ 0 & 1.17 & 0 \\ 0 & 0 & 0.57 \end{pmatrix}$$

また各個人の分散共分散行列 Σ'_2 は次の通りである．これを用いて各個人がもつ分布関数 $N(0, \Sigma'_2)$ とする．

$$\Sigma'_2 = \begin{pmatrix} 1.522 & -0.280 & 0.193 \\ -0.280 & 0.631 & -0.037 \\ 0.193 & -0.037 & 0.294 \end{pmatrix}$$

ユーザの平均 $g'_i (i = 1, \dots, n)$ を $N(G'_1, \Sigma'_1)$ に基づいてランダムに配置し，各ユーザの重心 g'_i を決定する． n 人のユーザの確率楕円の総体積 V' と重複体積の総和 M' をモンテカルロ法より求め，これらより重複度 P' を求める．

シミュレーションでは検出率を $p = 0.95$ と設定した．ユーザ数を 2 人～10 人とし，それぞれの場合について 1000 通りのユーザを配置して，各配置毎に重複度を求めた．各ユーザ数に対する重複度の平均を表 15 に示す．

表 15: 正規化後のユーザ数と重複度の関係

ユーザ数	2	3	4	5	6	7	8	9	10
重複度	5.8	10.8	16.2	20.6	23.4	27.3	28.7	32.3	33.7

正規化前に比べて各ユーザ数に対する重複度が大きくなり，ユーザ数 3 人以

上で誤識別率 5%を超える．このことから，体重情報を除くと大幅に誤識別率が上昇することが分かる．

4.3 考察

本章の実験から 10 人程度のユーザ数に対して検出率 95%では誤識別率 5%の範囲内で個人識別が可能であることが分かった．このことからフットステップデータを利用した個人識別手法は少人数のオフィスや家庭内で適用できると思われる．また，個人識別においては体重が重要な情報を持っているため，同じ体重の人では識別率がかなり下がるが，ごく少人数の場合でのみ識別可能であることが分かった．

本章の実験における被験者は実験者と同じビル内の 20 代から 30 代の男性 19 名，女性 7 名を対象としたため，無作為な被験者の抽出とはいえない．そこで識別影響の大きい体重の分布に関して一般的な日本人の調査による結果と比較する．表 16 に一般的な人における (体重の標準偏差/体重) の値と [6]，本研究における被験者の (中間極小値の標準偏差/中間極小値) を示す．ただし若者は 18 ~ 29 歳の男女各 200 名のデータを対象としている．この結果より実験における

表 16: 体重のバラつき方の比較

	標準偏差/体重		標準偏差/中間極小値
若者 (男)	0.133	被験者 (男)	0.119
若者 (女)	0.118	被験者 (女)	0.130

被験者の分布は一般的な若者の分布と大きな差があるとは言えない．よって母集団を 20 代から 30 代の日本人として想定する場合には，この結果には妥当性があると考えられる．

5 結論

本報告書では，フットステップデータを利用した個人識別手法について検討するために，まずフットステップデータを収集する環境の違いによりデータがどのような影響を受けるかについて調べ，定常的な歩行動作中のフットステップデータを収集するのに適切な環境を構築した．次に収集されたデータについてその識別能力の検討を行った．

収集環境の違いがデータに与える影響を先行研究で用いられているフットス

テップデータの特徴量をもとに調べた。まず先行研究において実施されている 2 種類の条件，

- (i) 台はかりをそのまま床に置く。これにより計測前後で段差が生じる。
- (ii) 台はかりの周囲に台はかりと同じ高さの物体を配置し，計測前後で段差が生じないようにする。

についてデータを収集し統計的な有意差について検定を行った。その結果，条件 (i)(ii) ではフットステップデータの特徴量の分布が異なることを示した。

次に台はかりに対する被験者の足の接地位置への意識がフットステップデータの特徴量の分布に与える影響を調べた。9 枚の台はかりを連続して配置し，計測することで被験者の接地位置への意識を低減させる環境を構築した。次の条件，

- (iii) 1 枚の台はかりを明示し，被験者に接地位置を意識させる。
- (iv) 接地位置を意識させないようにする。

に対してデータを収集し統計的な有意差について検定を行った。その結果，条件 (iii)(iv) ではフットステップデータの特徴量の分布が異なることを示した。これにより定常的な歩行時のフットステップデータを収集するには，計測時の前後で床の高さが変わらず，被験者が足の接地位置を意識しないフットステップデータ収集環境を作る必要があることが分かった。

このような収集環境のもとで収集したフットステップデータに対し，その識別能力を調べるために，各個人毎のデータのバラつき方と個人間のバラつき方を調べた。まずデータの分散共分散行列を求め，主成分分析により合計の寄与率が 90%以上となる上位 3 軸で分析を行うことにした。次にユーザが生成する全フットステップデータの母集団の分布が本実験で得られた全フットステップデータの平均，分散共分散行列を持つ正規分布であると仮定した。また，各ユーザが生成するフットステップデータの分布も正規分布と仮定し，その分散共分散行列が本実験での各被験者ごとのフットステップデータの平均的な分散共分散行列であると仮定した。この各ユーザの分布から確率楕円を計算し，その重複度から識別性能を調べた。その結果，10 人程度のユーザ数に対して検出率 95%では誤識別率 5%の範囲で個人識別が可能であることを示した。

また体重が識別に与える影響を調べるために，体重情報を除いたフットステップデータの識別能力を検討した。その結果，体重情報を除くと大幅に識別率が低下し，体重情報が個人識別の重要な要因となっているが，その他に識別に関わる情報が含まれていることも分かった。

本研究では被験者は裸足であり、データ収集において定常的で一様な歩行になるように条件をそろえて実験を行った。しかし日常的な歩行では履物が異なる場合もあれば、状況によって歩行速度が一様でない場合もある。このような条件のもとでは、本研究における手法が適用可能であるとはいえない。オフィスや家庭のような実環境での利用に向けて、ユーザの設定や利用状況に応じたデータの収集、分析を行う必要がある。

謝辞

本研究を行うにあたり多くの御教示、熱心な御指導を賜りました美濃導彦教授に深く感謝いたします。本研究について数多く助言を頂きました河原研究室の沢田篤史助教授に深く感謝いたします。また、いつも大変お世話になり、本報告書の作成においても御指導をいただきました美濃研究室の関氏をはじめとするNAグループ並びに美濃研究室の皆様、実験に御協力くださった河原研究室の皆様、コンテンツ作成室の皆様に深くお礼を申し上げます。

参考文献

- [1] M.D. Addlesee, A. Jones, F. Livesey, and F. Samaria.: ORL Active Floor, *IEEE Personal Communications*, Vol.4, No.5, pp. 35-41 (Oct. 1997).
- [2] R.J. Orr and G.D. Abowd.: The Smart Floor - A mechanism for natural user identification and tracking, *In Proc. 2000 Conf. Human Factors in Computing Systems(CHI 2000)*, New York, USA, (2000).
- [3] A. Hopper, A Harter and T Blackie.: The Active Badge System, *INTER-CHI93.*, Amsterdam, (Apr, 1993).
- [4] 南 敏: 顔画像による個人識別の技術, *システム / 制御 / 情報*, Vol.35, No.7, pp.415-422, (1991).
- [5] 古井定熙: 音声による個人識別の技術, *システム / 制御 / 情報*, Vol.35, No.7, pp.408-414, (1991).
- [6] 生命工学技術研究所: 設計のための人体寸法データ集, *人間生活工学研究センター*, pp.216.(1996).

付録

A.1 補足資料

表 A.1: 各特徴量の平均値

被験者	前半極大値 (kg 重)		中間極小値 (kg 重)		後半極大値 (kg 重)	
	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)
1	51.0	52.9	40.3	37.3	57.3	58.0
2	75.9	72.5	54.3	55.2	74.7	73.3
3	67.6	65.2	46.8	48.2	69.6	69.6
4	74.8	74.1	56.6	57.2	75.7	75.0
5	86.5	82.0	64.8	62.7	83.8	86.8
6	54.5	55.7	40.3	37.5	51.9	52.7
7	68.5	67.7	52.9	51.3	64.1	64.7
8	86.0	82.3	52.9	52.4	83.4	87.8
9	63.8	61.8	51.3	49.8	72.2	71.5
10	89.7	80.5	55.7	57.7	78.9	79.0
11	74.7	70.1	49.0	50.8	61.8	63.9
12	71.1	73.5	59.3	56.2	75.2	75.3
13	71.4	71.9	47.3	46.7	69.7	69.9
14	83.9	84.8	69.7	69.7	85.6	84.4
15	54.0	51.1	39.9	39.3	51.5	50.9
16	67.2	70.6	48.4	45.5	72.5	73.9
17	55.4	54.9	39.7	40.3	52.3	52.3
18	70.2	71.3	57.9	56.2	69.6	70.4
19	68.5	69.6	47.2	45.5	67.7	66.4
20	61.5	63.7	49.3	47.4	58.6	58.1
全体	69.8	68.8	51.2	50.3	68.8	69.2

被験者	前半極大時間 (秒)		中間極小時間 (秒)		後半極大時間 (秒)		接地時間 (秒)	
	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)
1	0.183	0.160	0.317	0.318	0.530	0.495	0.715	0.666
2	0.174	0.175	0.330	0.329	0.482	0.481	0.638	0.640
3	0.165	0.179	0.307	0.319	0.520	0.524	0.679	0.697
4	0.176	0.172	0.336	0.329	0.529	0.519	0.696	0.687
5	0.176	0.182	0.364	0.348	0.572	0.544	0.776	0.753
6	0.188	0.161	0.326	0.296	0.508	0.461	0.697	0.633
7	0.172	0.176	0.342	0.349	0.534	0.535	0.750	0.740
8	0.135	0.141	0.285	0.275	0.449	0.434	0.601	0.598
9	0.225	0.183	0.388	0.357	0.594	0.547	0.766	0.710
10	0.154	0.174	0.337	0.326	0.555	0.522	0.717	0.683
11	0.132	0.147	0.290	0.297	0.457	0.469	0.640	0.647
12	0.148	0.145	0.296	0.285	0.484	0.466	0.644	0.615
13	0.150	0.150	0.290	0.290	0.466	0.468	0.628	0.626
14	0.184	0.180	0.330	0.323	0.514	0.497	0.689	0.675
15	0.145	0.151	0.287	0.287	0.456	0.428	0.637	0.599
16	0.145	0.136	0.300	0.302	0.465	0.459	0.619	0.607
17	0.175	0.165	0.324	0.313	0.459	0.454	0.606	0.614
18	0.178	0.176	0.317	0.327	0.515	0.518	0.693	0.688
19	0.155	0.156	0.316	0.325	0.515	0.496	0.683	0.661
20	0.184	0.170	0.341	0.335	0.533	0.510	0.726	0.698
全体	0.167	0.164	0.321	0.317	0.507	0.491	0.680	0.662

表 A.2: 各特徴量の標準偏差

	前半極大値 (kg 重)		中間極小値 (kg 重)		後半極大値 (kg 重)	
	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)
1	2.03	2.69	2.07	1.70	1.79	4.43
2	3.07	0.84	1.57	1.05	1.85	2.23
3	3.42	1.35	2.41	0.84	2.75	1.90
4	2.70	2.12	1.77	1.95	1.08	2.01
5	4.22	3.07	2.21	2.27	2.06	3.06
6	1.04	1.60	0.98	1.38	1.43	2.59
7	2.82	2.29	1.70	1.28	2.67	1.88
8	4.52	3.43	2.49	1.60	2.37	2.54
9	2.80	2.03	0.85	1.03	1.36	1.74
10	3.59	4.73	2.43	3.98	2.44	2.94
11	5.81	2.77	2.64	1.09	2.77	0.79
12	3.97	5.89	1.91	3.23	1.79	1.17
13	2.54	3.03	0.94	1.39	1.58	2.60
14	2.49	1.75	1.41	1.78	1.78	2.50
15	3.50	3.70	1.66	2.07	1.94	2.20
16	3.30	5.89	2.61	2.08	3.38	3.83
17	1.36	2.00	1.57	1.10	1.74	1.78
18	1.29	0.94	2.66	1.57	2.49	1.30
19	2.60	2.86	2.95	2.80	2.10	2.49
20	2.22	2.28	1.09	2.29	1.10	2.35
全体	11.38	10.09	8.21	8.47	10.55	11.00

	前半極大時間 (秒)		中間極小時間 (秒)		後半極大時間 (秒)		接地時間 (秒)	
	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)	条件 (iii)	条件 (iv)
1	0.030	0.019	0.024	0.019	0.033	0.018	0.041	0.026
2	0.009	0.005	0.011	0.010	0.010	0.011	0.012	0.010
3	0.010	0.010	0.018	0.007	0.010	0.009	0.011	0.013
4	0.010	0.018	0.012	0.015	0.020	0.025	0.024	0.031
5	0.013	0.024	0.032	0.030	0.020	0.025	0.027	0.028
6	0.012	0.010	0.020	0.018	0.020	0.020	0.019	0.027
7	0.012	0.018	0.026	0.016	0.044	0.013	0.033	0.013
8	0.006	0.005	0.010	0.007	0.008	0.011	0.014	0.013
9	0.014	0.014	0.027	0.021	0.017	0.012	0.021	0.017
10	0.009	0.025	0.035	0.028	0.028	0.046	0.029	0.049
11	0.012	0.011	0.020	0.010	0.013	0.013	0.013	0.010
12	0.026	0.038	0.017	0.023	0.009	0.022	0.010	0.026
13	0.006	0.008	0.011	0.010	0.010	0.013	0.016	0.013
14	0.008	0.009	0.010	0.012	0.013	0.017	0.012	0.018
15	0.014	0.012	0.016	0.013	0.023	0.021	0.023	0.018
16	0.013	0.009	0.022	0.008	0.014	0.012	0.012	0.012
17	0.011	0.009	0.015	0.017	0.019	0.027	0.021	0.032
18	0.013	0.013	0.024	0.032	0.040	0.014	0.012	0.013
19	0.011	0.016	0.022	0.023	0.019	0.027	0.024	0.026
20	0.011	0.020	0.014	0.024	0.018	0.039	0.020	0.040
全体	0.025	0.021	0.033	0.029	0.046	0.041	0.055	0.050

A.2 確率楕円

k 次元の確率変数 x_1, \dots, x_k の平均を m_1, \dots, m_k とし, それらの分散を σ_j^2 , 共分散を $\sigma_{ij} (i, j = 1, \dots, k)$ とおく. この時, 分散共分散行列 Σ を,

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1k} \\ \vdots & \ddots & \vdots \\ \sigma_{k1} & \cdots & \sigma_k^2 \end{bmatrix} \quad (\text{A.1})$$

とする. また \mathbf{x}, \mathbf{m} を,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, \mathbf{m} = \begin{bmatrix} m_1 \\ \vdots \\ m_k \end{bmatrix} \quad (\text{A.2})$$

とおく. このとき楕円面

$$\frac{1}{k}(\mathbf{x} - \mathbf{m})^t \Sigma^{-1}(\mathbf{x} - \mathbf{m}) = 1 \quad (\text{A.3})$$

を集中楕円面と呼ぶ. ただし Σ^{-1} は Σ の逆行列を, $(\mathbf{x} - \mathbf{m})^t$ は $(\mathbf{x} - \mathbf{m})$ の転置行列を表す. この時集合 A の確率を $P(A)$ で表すと, 任意の $p (0 \leq p \leq 1)$ に対して,

$$P\left(\frac{1}{k}(\mathbf{x} - \mathbf{m})^t \Sigma^{-1}(\mathbf{x} - \mathbf{m}) < \lambda^2\right) = p \quad (\text{A.4})$$

となるような $\lambda (\lambda \geq 0)$ が存在する. この λ に対して,

$$\frac{1}{k}(\mathbf{x} - \mathbf{m})^t \Sigma^{-1}(\mathbf{x} - \mathbf{m}) < \lambda^2 \quad (\text{A.5})$$

を確率 p における確率楕円という. 確率 p が与えられた時の λ を決定する.

$k = 3$ の場合について考える. 簡単のため $m_1 = m_2 = m_3 = 0$ とすると, (A.3) で表される集中楕円面は,

$$\frac{1}{3}\mathbf{x}^t \Sigma^{-1}\mathbf{x} = 1 \quad (\text{A.6})$$

とかける. 3つの軸のそれぞれ半分の長さを $Q, R, S (Q \geq R \geq S)$ とする. 上式を楕円の標準形

$$\left(\frac{y_1}{Q}\right)^2 + \left(\frac{y_2}{R}\right)^2 + \left(\frac{y_3}{S}\right)^2 = 1 \quad (\text{A.7})$$

の形に変換することを考える．これは \mathbf{C} を直交行列するとき，

$$\mathbf{C}^t \Sigma \mathbf{C} = \mathbf{L} = \begin{bmatrix} l_1 & 0 & 0 \\ 0 & l_2 & 0 \\ 0 & 0 & l_3 \end{bmatrix} \quad (\text{A.8})$$

として， $\mathbf{y} = \mathbf{C}^t \mathbf{x}$ なる直交変換を (A.6) に対して行えばよい．ただし l_1, l_2, l_3 は Σ の固有値であり $l_1 \geq l_2 \geq l_3$ とする．また

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \mathbf{C}^t \mathbf{x} \quad (\text{A.9})$$

とする．(A.6) に (A.9) を代入すると

$$\frac{1}{3} \mathbf{x}^t \Sigma^{-1} \mathbf{x} = \frac{1}{3} \mathbf{y}^t (\mathbf{C}^t \Sigma^{-1} \mathbf{C}) \mathbf{y} = 1 \quad (\text{A.10})$$

$$\mathbf{C}^t \Sigma^{-1} \mathbf{C} = (\mathbf{C}^t \Sigma \mathbf{C})^{-1} = \mathbf{L}^{-1} = \begin{bmatrix} \frac{1}{l_1} & 0 & 0 \\ 0 & \frac{1}{l_2} & 0 \\ 0 & 0 & \frac{1}{l_3} \end{bmatrix} \quad (\text{A.11})$$

よって，

$$\frac{1}{3} \mathbf{x}^t \Sigma^{-1} \mathbf{x} = \frac{1}{3} \mathbf{y}^t \mathbf{L}^{-1} \mathbf{y} = \left(\frac{y_1}{\sqrt{3l_1}} \right)^2 + \left(\frac{y_2}{\sqrt{3l_2}} \right)^2 + \left(\frac{y_3}{\sqrt{3l_3}} \right)^2 = 1 \quad (\text{A.12})$$

となる．このことから

$$Q = \sqrt{3l_1}, \quad R = \sqrt{3l_2}, \quad S = \sqrt{3l_3} \quad (\text{A.13})$$

であることが分かる．これにより式 (A.6) が表す楕円体の体積 V_1 は，

$$V_1 = \frac{4}{3} \pi QRS = 4\pi \sqrt{3l_1 l_2 l_3} \quad (\text{A.14})$$

となる．

また， l_1, l_2, l_3 は Σ の固有値なので，

$$\begin{vmatrix} \sigma_1^2 - l & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 - l & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 - l \end{vmatrix} = 0 \quad (\text{A.15})$$

を l について解くと l_1, l_2, l_3 が求まる。(但し $\sigma_{12} = \sigma_{21}, \sigma_{23} = \sigma_{32}, \sigma_{31} = \sigma_{13}$)

この方程式は,

$$\begin{aligned} -l^3 + (\sigma_1^2 + \sigma_2^2 + \sigma_3^2)l^2 + (\sigma_{12}^2 + \sigma_{23}^2 + \sigma_{31}^2 - \sigma_2^2\sigma_3^2)l \\ + \sigma_1^2\sigma_2^2\sigma_3^2 - \sigma_1^2\sigma_{23}^2 - \sigma_2^2\sigma_{31}^2 - \sigma_3^2\sigma_{12}^2 + 2\sigma_{12}\sigma_{23}\sigma_{31} = 0 \end{aligned} \quad (\text{A.16})$$

となるので 3 次方程式の解の公式より,

$$l_1 l_2 l_3 = \sigma_1^2 \sigma_2^2 \sigma_3^2 - \sigma_1^2 \sigma_{23}^2 - \sigma_2^2 \sigma_{31}^2 - \sigma_3^2 \sigma_{12}^2 + 2\sigma_{12}\sigma_{23}\sigma_{31} \quad (\text{A.17})$$

となる.

ある変数 \mathbf{x} が,

$$\frac{1}{3}\mathbf{x}^t \Sigma^{-1} \mathbf{x} = r^2 \quad (\text{A.18})$$

の楕円体の内部に入る確率 $\varphi(r)$ を考える. このとき楕円面 $\frac{1}{3}\mathbf{x}^t \Sigma^{-1} \mathbf{x} = r^2$ は直交変換により,

$$\left(\frac{y_1}{Q}\right)^2 + \left(\frac{y_2}{R}\right)^2 + \left(\frac{y_3}{S}\right)^2 = r^2 \quad (\text{A.19})$$

$$\left(\frac{y_1}{rQ}\right)^2 + \left(\frac{y_2}{rR}\right)^2 + \left(\frac{y_3}{rS}\right)^2 = 1 \quad (\text{A.20})$$

となる. これは (A.7) の軸をそれぞれ r 倍した時, (A.19) で表される式の軸の長さになる. よってその時の体積 V は,

$$V = \frac{4}{3}\pi QRSr^3 = 4\pi\sqrt{3l_1 l_2 l_3} r^3 \quad (\text{A.21})$$

ここで \mathbf{x} の分布が 3 次元正規分布

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}\mathbf{x}^t \Sigma^{-1} \mathbf{x}} \quad (\text{A.22})$$

とする. この時,

$$\begin{aligned} |\Sigma|^{\frac{1}{2}} &= \sqrt{\sigma_1^2 \sigma_2^2 \sigma_3^2 - \sigma_1^2 \sigma_{23}^2 - \sigma_2^2 \sigma_{31}^2 - \sigma_3^2 \sigma_{12}^2 + 2\sigma_{12}\sigma_{23}\sigma_{31}} \\ &= \sqrt{3l_1 l_2 l_3} \end{aligned} \quad (\text{A.23})$$

である. (A.21) より, V を r の関数とみなせば,

$$dV = 12\pi\sqrt{3l_1 l_2 l_3} r^2 dr \quad (\text{A.24})$$

となるので，楕円内部 ($V(r)$) の確率 $\varphi(r)$ は，

$$\begin{aligned}\varphi(r) &= \int_{\mathbf{x} \in V} f(\mathbf{x}) dV \\ &= \int_{\mathbf{y} \in V} \frac{1}{(2\pi)^{\frac{3}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{1}{2} \mathbf{y}^t \mathbf{L}^{-1} \mathbf{y}} dV \\ &= \int_0^r \frac{1}{(2\pi)^{\frac{3}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{3}{2} r^2} \cdot \left(12\pi \sqrt{3l_1 l_2 l_3} r^2 \right) dr \\ &= \int_0^r 3\sqrt{\frac{6}{\pi}} r^2 e^{-\frac{3}{2} r^2} dr\end{aligned}\tag{A.25}$$

となる．よって $p = \varphi(r)$ の解として確率 p に対応する確率楕円のパラメータ λ を求めることができる．