

床に固定した複数の加速度センサの 歩行センシングによる履物と人物の推定

林田 望海¹ 勝間 亮¹

概要：現在、空港で飛行機に乗る際や建物の出入りをする際の様々な認証システムが開発されている。その例として、パスワードやICカード、指紋認証などがある。しかし、これらの認証システムは特徴量を認識するために文字の入力が必要であったり、カードや手をかざす等の特別な動作を必要とする。一方、人の歩く動作には多くの要素を持ち、その人特有の歩き方を認識することで移動中に人物を特定でき、先に述べた認証のみを目的とした動作を行わなくてよくなると期待されている。本研究では、複数の振動センサを設置した床を靴を履いたまま歩くことにより、複数個所で得た加速度データから履物および人物を識別することを目的とする。本稿では、センサの個数と識別精度の関係を調査する実験の成果を報告する。実験では、振動センサを木製の床に4箇所設置し、その上を3人の被験者が2種類の履物（スリッパ、スニーカー）で歩行し、それにより生じる床の振動を検出し、畳み込みニューラルネットワークを用いて識別精度の検証を行った。実験結果は、センサ1つの場合、履物の識別は90%~94%であり、人物の識別では86%~99%であった。センサ4つの場合、履物の識別は96%であり、人物の識別は97%であった。履物の識別はセンサを組み合わせて精度が向上することが分かったが、人物の識別は必ずしもそうはならなかった。

1. はじめに

近年、空港で飛行機に乗る際や会社の建物に出入りをする際などあらゆる場面でセキュリティ対策のために本人確認が求められるようになっており、それとともに様々な認証システムが開発されている。その代表例としてパスワード認証が挙げられる。パスワード認証とは、事前に登録されたものとユーザーが入力したものが一致すれば本人であるとみなされる認証システムである。パスワードは主に数字や記号の羅列によって構成され、一般的に文字数が多いほどセキュリティは高くなる。しかし、ユーザーはそれを記憶し入力する必要があるため、長すぎるパスワードは安全ではあるがユーザーにストレスを与えることになる。このようにパスワード認証は利用者の記憶に基づく認証システムであるのに対し、鍵やICカードといった利用者のみが持っているものによって認証をする所持物に基づく認証システムや指紋認証や虹彩認証といった利用者の生体情報に基づいた認証システムなども多く研究されている。所持物に基づく認証システムの利点は、記憶や入力という煩わしさを一切なくしたことである。しかし、もちろん利用の際は必ず所持していなければならず、もし落とすなどして紛失し

てしまった場合には本人が利用できなくなるだけでなく、拾った人に利用されるなどリスクも大きい。一方で指紋認証など身体の一部を用いた生体認証は紛失やなりすましの可能性が低く、セキュリティが高いため多くの場面で利用されることが期待される。しかし、これらの認証システムはICカードの認証と同じで特徴量を認識するために手をかざす等の特別な動作を必要とする。

一方で、人の歩く動作には、身長や体重などの体格による歩幅や踏み込みの強さ、年齢や習慣的な行動による姿勢や歩行速度など多くの要素を持っており、その人特有の歩き方（歩容）を認識することで、特別な動作を必要とせず、認証対象の人物に意識させることなく人物を特定することができる。既存研究では、人の歩行を撮影して映像から特徴を得たり、圧力センサを敷き詰めた床を素足で歩くことで足形などの特徴量を得たりすることで人物を特定する方法がある [3], [4]。しかし、映像では撮影するときの明るさや撮影する角度に精度が左右される。また、センサを敷き詰めるには相当な数のセンサが必要となりコストが大きくなる。そのため、できる限り少ないセンサ数で、靴を脱ぐ動作を必要とせず、死角のない安定した歩容認証が求められる。

本研究では靴を履いたまま、振動センサを設置した床を歩くことにより数少ないセンサで人の歩幅や踏み込み強度

¹ 大阪府立大学
Osaka Prefecture University, Sakai, Osaka 599-8531, Japan

などの特徴量を得ることで人物を特定することを目的とする。そのために、本論文では振動センサが取得する加速度データ系列から畳み込みニューラルネットワーク (CNN) で人物の特定を行う際に、用いるセンサの個数と識別精度の関係を実測実験により明らかにする。また、歩容の違いは履物の種類にも起因するため、スリッパとスニーカーの2種類の履物を想定し、それらの識別率とセンサの個数の関係についての実験結果も報告する。様々な特徴の中から履物に注目した理由の一つとして、日本では屋内と屋外で靴を履き替える文化があり土足厳禁の場所は多く、屋内用と屋外用の履物の識別をすることで、外靴のまま屋内に入ろうとする人に対して自動で警告を与える等のシステムに利用できることが挙げられる。その他、関係者以外立ち入り禁止の区画での侵入者の検知などへの利用も期待できる。

2. 関連研究

歩容には歩幅や歩行速度、姿勢、体重移動などその人特有の特徴が多く含まれる。歩容ベースの人物特定はこれらの様々な特徴をそれに適したセンサで検出して行う。本章では歩容認証システムの研究例を挙げ、それらが解決できていない課題と本研究の位置付けについて述べる。

2.1 映像による歩容認証

映像を用いた歩容認証の手法は多数提案されている [1], [2]。しかし、これらの認証システムは歩行方向や観測方向の変化により認証精度が左右されるという弱点が存在する。また、方向変化への対応手法として、全方位カメラから得られる複数方向の歩容画像を用いた歩容認証システムが杉浦らにより提案されている [3]。これらの映像による歩容認証システムは、対象人物が撮影位置から遠い場合やマスクやヘルメットなどで顔が見えない状況においても利用ができる。しかし、明るさや歩行者の服装など外的要因により容易に認証精度が変化するという課題が挙げられる。また、身体全体が映るように広範囲の動画を撮影する必要があるため、個人認証に用いる特徴を抽出する処理も複雑である。

2.2 圧力センサによる歩容認証

マット型圧力センサを用いた足底圧変化による歩容認証システムが山川らにより提案されている [4]。この認証システムでは、床にマット型圧力センサを設置し、その上を靴を履かずに歩行を行ったときの、マット上の圧力変化を歩行データとして取得する。そのデータから圧力が最大になる時間や踏み込み、踏み抜き時間のような特徴を用いて人物の特定を行う。マット型圧力センサは、マット上の圧力分布を数 mm 単位で得ることができ、体重移動など詳細な情報が検出することができるため、より高度な認証が可能である。しかし、足底圧変化を正確に得るためには裸足

で歩く必要がある。また、歩く範囲全体に圧力センサを敷き詰める必要があり、導入コストが高くなるという点が課題として挙げられる。

2.3 本研究の位置付け

2.1 節や 2.2 節のように様々な歩容認証システムが提案されているが、映像による歩容認証はあらゆる死角により精度が左右される。圧力センサによる歩容認証では裸足である必要があり、またセンサの数も多くなる。そこで、靴を履いたままできる限り少ない数のセンサで、死角のない安定した歩容認証システムが求められる。柏本らは床に張り付けた振動センサを用いた位置推定手法を提案している [5]。このシステムは、床に複数の振動センサを設置し、ドアを開閉したときの振動、椅子に着席・離席したときの振動など、ユーザの位置を特定できる振動を検出することによって位置推定を行う。この研究を応用し、床に設置した振動センサからその人特有の歩行の振動を検出することにより人物を特定する方法を提案する。

3. 想定環境

本研究のねらいは、床に設置した振動センサで取得したデータから歩行の振動を検出し、その振動の特徴から履物や人物を識別することである。本章ではその想定環境について述べる。

3.1 床の定義

本研究では建物への改札のような出入りを想定し、大人が2歩以上歩くことができる領域で歩行の振動の検出をする。歩行者はその範囲を一定方向に歩き、領域を歩き始める時刻 t_{start} から歩き終わる t_{end} まで常に振動を検出する。その間1人しか領域内に立ち入らないようにする。

3.2 振動センサの定義

振動センサとして十分に高いサンプリングレートの3軸加速度センサを用いる。 n 個の振動センサ $s_1, s_2, s_3, \dots, s_n$ を使用し、進行方向に y 軸、右方向に x 軸、上方向に z 軸となるように床に固定する。床上の歩行以外による振動はないものとする。

3.3 歩行者の定義

m 人の歩行者 $a_1, a_2, a_3, \dots, a_m$ が存在し、それら全員を識別対象者とする。歩行者はそれぞれ履物としてスリッパもしくはスニーカー (図1) のどちらかを装着し、リズム等に影響されない自然体での歩行を行う。

3.4 入力情報

1歩目の振動を検出する直前時刻 t_1 から2歩目の振動を検出した後の時刻 t_2 までの加速度センサ s_i の3軸の加速度



図 1 スリッパとスニーカー

値 $s_{i,x}(t)$, $s_{i,y}(t)$, $s_{i,z}(t)$ を入力とする. ただし, $1 \leq i \leq n$, $t_{start} \leq t_1 \leq t_2 \leq t_{end}$, $t \in \{t_1, t_2\}$.

3.5 出力情報

履物識別においては, スリッパかスニーカーかのどちらであるかを出力する. 人物識別においては, $a_1, a_2, a_3, \dots, a_m$ のいずれに該当するかを出力とする.

3.6 識別器

履物および人物の識別には CNN を用いる.

4. 実験

本章では, 歩行者の履物の識別と, その人物の識別を行う実験について説明する. そのうえで, 使用するセンサの個数を変化させたときの識別成功率の変化を明らかにする.

4.1 実験環境

実際に履物や人物の識別をする歩行者の情報や使用するセンサの数, 設置位置, 入力情報について述べる.

4.1.1 歩行者

本実験では体格の違う 3 人の被験者に同じ 2 種類の履物 (スリッパとスニーカー) を履いてもらい, 板の上を一定方向にそれぞれ 50 回ずつ歩行を行ってもらった. 被験者の体格を表 1 に示す.

表 1 被験者情報

	a_1	a_2	a_3
身長 (cm)	170	178	170
体重 (kg)	57	52	68

4.1.2 センサの数と位置

振動センサはモノワイヤレス株式会社の加速度センサー無線タグ TWELITE 2525A-トワイライトニコニコというサンプリング数 100Hz の加速度センサを用いた [6]. 振動検出領域の床は 90cm × 180cm × 2cm の振動を伝えやすい木製の板をとし, 振動センサ 4 個 s_1, s_2, s_3, s_4 をそれぞれ領域の左手前, 右手前, 左奥, 右奥の四隅に設置した (図 2).

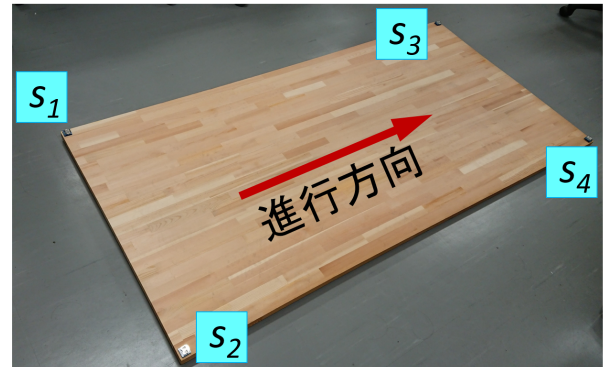


図 2 センサを設置した板

4.1.3 入力情報

2 歩以上の連続した歩行のデータを扱うことによって歩行速度などの情報を得られる. そこで 1 歩目の振動を検出した 0.15 秒前を t_1 とし, そこから約 1 秒後を t_2 としたサンプリング数 100 に切り取った 100 行 × 3 列のデータを 1 つ分のデータとして扱った. この 1 秒という値は 2 歩の振動を含み 3 歩目の振動を含まないように定めた値である. 図 3 に s_2 で実際に得られた被験者 a_2 のスリッパによる歩行のデータを示す.

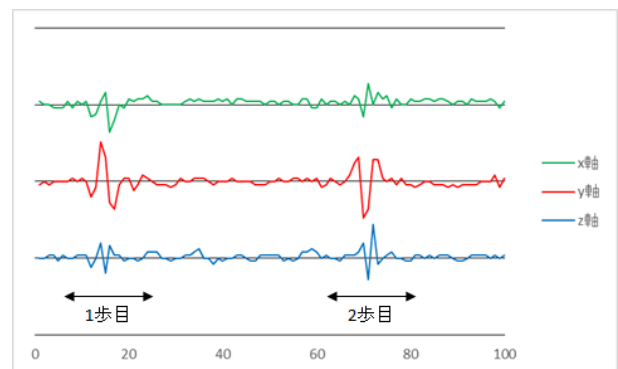


図 3 被験者 a_2 のスリッパのデータの例

4.1.4 識別

本実験では機械学習ツールである Neural Network Console の CNN を用いた [8]. 図 4 にセンサ 2 つを用いる場合の人物を識別をする CNN の構成を示す. 図 4 のように入力層 (Input), 畳み込み層 (Convolution), 正規化層 (ReLU, Tanh,), プーリング層 (AveragePooling), 全結合層 (Affine), 出力層 (Softmax, CategoricalCrossEntropy) の計 14 層で構成した. 履物の識別の場合は出力以外は同じ構成の CNN を使用し, 出力層の Softmax を Sigmoid に, CategoricalCrossEntropy を BinaryCrossEntropy に変更する. また, 履物の識別と人物の識別でデータも同じデータを用いた. データ数は各被験者スリッパとスニーカー 50 回ずつの計 300 個で行い, 母数が少ないため 5 分割交差検証を用いる. 5 分割交差検証とはデータセットを

5つに分割し、そのうちの4つを学習データとして使用して学習し、残りの1つで評価を行い、それを全パターン行う方法である。

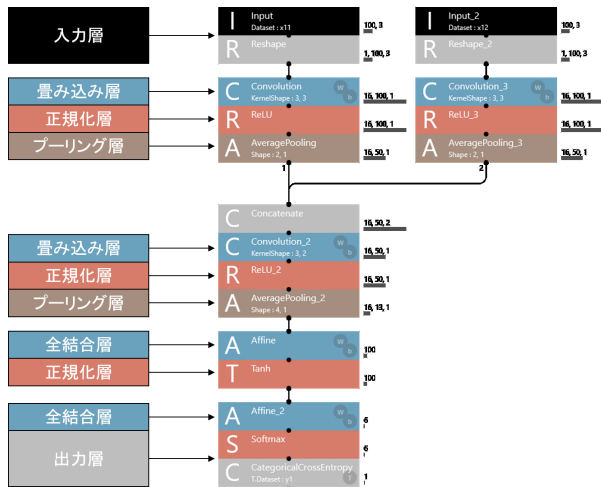


図4 センサ2つ用いた場合の人物識別 CNN 構成

4.2 実験方法

本実験は使用するセンサの個数と識別精度の関係を実測により調査することを目的とする。また、データから何かしらの特徴を検出し入力情報に追加することで識別精度の向上を図れるかの検証も行った。

4.2.1 複数のセンサデータによる推定精度

センサの個数と識別精度の関係を調査するために、CNNの入力情報にセンサ1つを用いる場合 (s_1, s_2, s_3, s_4), センサ2つを用いる場合 (s_1 と s_2, s_1 と s_3, s_1 と s_4, s_2 と s_3, s_2 と s_4, s_3 と s_4), センサ3つを用いる場合 (s_1 と s_2 と s_3, s_1 と s_2 と s_4, s_1 と s_3 と s_4, s_2 と s_3 と s_4), センサ4つを用いる場合 (s_1 と s_2 と s_3 と s_4) の全15パターンで識別を行った。

4.2.2 床を擦る特徴検出による推定精度

一般にスリッパは脚が上がらず床を擦ることが多いそれは床の振動にも表れ、図5に見られるようにスニーカーよりスリッパが振動時間が長いことがある。そこで履物の識別を行うにあたって、スリッパに見られる床を擦る特徴を検出し入力情報に追加することによって、生のデータで学習をするより識別の精度が向上するのではないかと考えた。以下では床を擦る特徴の検出方法を述べる。

まず各センサの x y z 軸のデータ値 $s_i.x(t), s_i.y(t), s_i.z(t)$ を扱うにあたって1つの数値にするために(1)式を用いて二乗平均平方根 $s_i.RMS(t)$ を求める。

$$s_i.RMS(t) = \sqrt{\frac{s_i.x(t)^2 + s_i.y(t)^2 + s_i.z(t)^2}{3}} \quad (1)$$

$s_i.RMS(t)$ が閾値 v_1 以上の数が一定値 v_2 以上、つまり大きく振動してる時間が長い場合は1を、そうでない場合

は0を入力情報に追加する。ただし v_1 は実験で得た300個のデータの $s_i.RMS(t)$ の平均値とし、 v_2 は v_1 以上であった数の平均値とする。取得した加速度データを使って適切な値を調査した結果、 $v_1 = 0.02943, v_2 = 10$ であった。

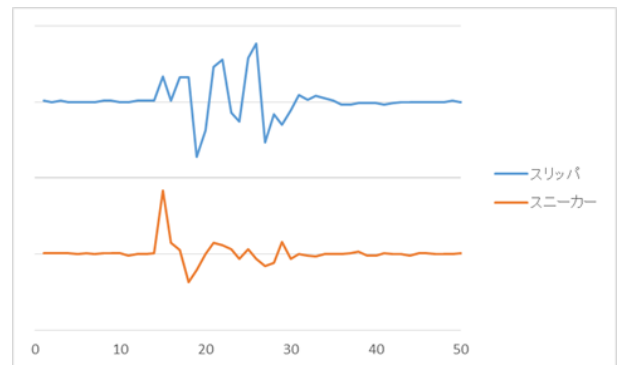


図5 スリッパとスニーカーのデータの違いの例

4.3 実験結果

4.3.1 複数のセンサデータによる識別精度

入力情報にセンサ1つを用いた場合、2つ用いた場合、3つ用いた場合、4つ用いた場合の識別結果をそれぞれ表2~9にまとめた。

表2 センサ1つの場合の履物識別結果

センサ位置		スリッパ'	スニーカー'
s_1	スリッパ	140	10
	スニーカー	8	142
s_2	スリッパ	132	18
	スニーカー	13	137
s_3	スリッパ	134	16
	スニーカー	7	143
s_4	スリッパ	138	12
	スニーカー	5	145

表3 センサ1つの場合の人物識別結果

センサ位置		a_1'	a_2'	a_3'
s_1	a_1	98	0	2
	a_2	1	99	0
	a_3	7	1	92
s_2	a_1	100	0	0
	a_2	0	100	0
	a_3	0	3	97
s_3	a_1	84	10	6
	a_2	8	88	4
	a_3	8	7	85
s_4	a_1	96	3	1
	a_2	2	93	5
	a_3	0	4	96

入力情報にセンサ1つを用いた場合、2つ用いた場合、3

表 4 センサ 2 つの場合の履物識別結果

センサ位置		スリッパ'	スニーカー'
s_{1,s_2}	スリッパ	143	7
	スニーカー	5	145
s_{1,s_2}	スリッパ	138	12
	スニーカー	5	145
s_{1,s_4}	スリッパ	140	10
	スニーカー	4	146
s_{2,s_3}	スリッパ	140	10
	スニーカー	5	145
s_{2,s_4}	スリッパ	139	11
	スニーカー	5	145
s_{3,s_4}	スリッパ	139	11
	スニーカー	7	143

表 5 センサ 2 つの場合の人物識別結果

センサ位置		a_1'	a_2'	a_3'
s_{1,s_2}	a_1	99	1	0
	a_2	0	100	0
	a_3	0	1	99
s_{1,s_3}	a_1	87	3	10
	a_2	3	97	0
	a_3	11	4	85
s_{1,s_4}	a_1	99	0	1
	a_2	0	98	2
	a_3	0	2	98
s_{2,s_3}	a_1	96	3	1
	a_2	3	97	0
	a_3	2	4	94
s_{2,s_4}	a_1	99	1	0
	a_2	1	99	0
	a_3	1	3	96
s_{3,s_4}	a_1	93	3	4
	a_2	6	91	3
	a_3	0	5	95

表 6 センサ 3 つの場合の履物識別結果

センサ位置		スリッパ'	スニーカー'
s_{1,s_2,s_3}	スリッパ	139	11
	スニーカー	3	147
s_{1,s_2,s_4}	スリッパ	143	7
	スニーカー	4	146
s_{1,s_3,s_4}	スリッパ	138	12
	スニーカー	5	145
s_{2,s_3,s_4}	スリッパ	142	8
	スニーカー	5	145

つ用いた場合、4つ用いた場合の識別精度を表 10 にまとめました。

4.3.2 床を擦る特徴による推定精度

スリッパの床を擦る特徴を入力情報に追加した場合のスリッパとスニーカーの識別結果を表 11 にまとめた。また、被験者 3 人の識別の結果を同様に表 12 にまとめた。

表 7 センサ 3 つの場合の人物識別結果

センサ位置		a_1'	a_2'	a_3'
s_{1,s_2,s_3}	a_1	96	3	1
	a_2	1	99	0
	a_3	2	3	95
s_{1,s_2,s_4}	a_1	100	0	0
	a_2	0	100	0
	a_3	2	2	96
s_{1,s_3,s_4}	a_1	97	0	3
	a_2	1	98	1
	a_3	4	2	94
s_{2,s_3,s_4}	a_1	98	0	2
	a_2	1	98	1
	a_3	1	3	96

表 8 センサ 4 つの場合の履物識別結果

	スリッパ'	スニーカー'
スリッパ	141	9
スニーカー	4	146

表 9 センサ 4 つの場合の人物識別結果

	a_1'	a_2'	a_3'
a_1	98	1	1
a_2	2	97	1
a_3	2	2	96

表 10 センサの組合せによる識別精度

センサ位置	履物	人物
s_1	0.940	0.963
s_2	0.897	0.990
s_3	0.923	0.857
s_4	0.943	0.950
平均	0.926	0.940
s_{1,s_2}	0.960	0.993
s_{1,s_3}	0.943	0.897
s_{1,s_4}	0.953	0.983
s_{2,s_3}	0.950	0.957
s_{2,s_4}	0.947	0.980
s_{3,s_4}	0.940	0.930
平均	0.949	0.957
s_{1,s_2,s_3}	0.953	0.967
s_{1,s_2,s_4}	0.963	0.987
s_{1,s_3,s_4}	0.943	0.963
s_{2,s_3,s_4}	0.957	0.973
平均	0.954	0.973
s_{1,s_2,s_3,s_4}	0.957	0.970

表 11 センサ 4 つの場合の履物識別結果 (特徴あり)

	スリッパ'	スニーカー'
スリッパ	142	8
スニーカー	4	146

さらにスリッパの床を擦る特徴を入力情報に追加した場合としない場合の識別の精度を表 13 にまとめた。

表 12 センサ 4 つの場合人物識別結果 (特徴あり)

	a_1'	a_2'	a_3'
a_1	98	1	1
a_2	1	98	1
a_3	2	2	96

表 13 センサ 4 つの場合の識別精度

	履物	人物
特徴なし	0.957	0.970
特徴あり	0.960	0.973

5. 考察

5.1 センサの数と位置による識別精度

表 10 のセンサの数と識別精度の関係を図 6 に示す。図 6 より使用するセンサの数ごとの精度の平均を見ると、履物の推定ではセンサ 1 つの場合から順に 92.6%, 94.9%, 95.4%, 95.7% の識別成功率だった。また人物の推定はセンサ 1 つの場合から順に 94.0%, 95.7%, 97.3%, 97.0% の識別成功率だった。このことからセンサの数を増やすほど精度が良くなるのが分かる。また、表 10 より履物の推定では s_1 と s_2 を組み合わせたとき 96% となり、最も精度が低い s_2 と組み合わせたにも関わらず、 s_1 単体で識別したときの 94% よりも精度が向上した。このように履物の推定では複数センサの推定において、最も精度が低い s_2 を含めてもセンサ 1 つで推定するより精度が向上することが確認できた。一方で、人物の推定では s_1 と s_3 を組み合わせたとき 90% となり、最も精度が低い s_3 と組み合わせると、 s_1 単体で識別したときの 96% より精度が下がった。このように人物の推定では複数センサの推定において精度の低い s_3 を含むと推定精度が下がるという違いが見られた。また、履物識別における s_2 の判定や、人物識別における s_3 の判定のようにセンサの設置位置によっても識別精度は左右されることがわかったためセンサの数を増やすだけでなく設置位置の調整も必要と考えられる。

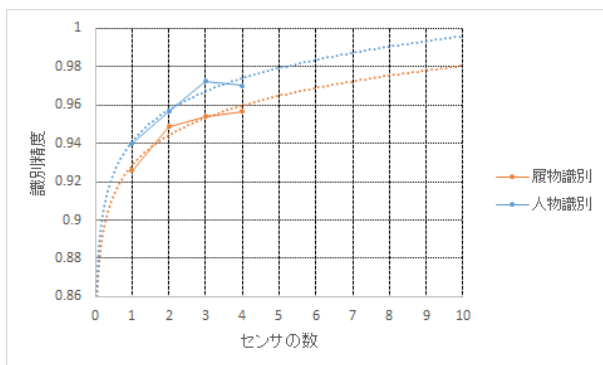


図 6 センサの数と識別精度の関係

5.2 床を擦る特徴検出による識別精度

表 13 よりスリッパの床を擦る特徴を入力情報に追加した場合、履物の識別と人物の識別ともに識別成功率が 0.3% 向上しているがこれは誤認識が 1 個減っただけであり、精度を向上させる特徴量と言えない。今回は床を擦る特徴を床の振動から視覚的に見える特徴 (大きく振動している時間の長さ) として抽出した。しかしこのような視覚的に得られる特徴は識別器内ですでに抽出できる可能性が高い。そのため精度にほとんど影響を与えなかったと考えられる。精度を上げる入力情報を追加するには、振動の平均値や分散値などの視覚的に得られない情報や振動以外から得られる情報を追加する必要があると考えられる。

6. まとめ

本論文では、振動センサを用いた歩行による履物および人物の推定を行う実験を通して、センサの個数と推定精度の関係を調査した。実験結果は、センサ 1 つの場合、履物の推定は 90%~94%, 人物の推定は 86%~99% であった。センサ 4 つの場合、履物の推定は 96%, 人物の推定は 97% であった。履物の推定は複数センサを組み合わせることで精度が向上することが確認できた。一方で、人物の推定は複数センサを組み合わせると精度の悪いセンサデータにより推定精度が下がることが確認できた。推定精度を上げるために、これら 2 種の性質の解析と、歩行速度や踏み込み強度以外の視覚的に見えない特徴量の抽出やセンサ位置や個数の調整などの検討が必要であると考えられる。

参考文献

- [1] 村松大吾, 岩間晴之, 木村卓弘, 榎原靖, 八木康史: “歩行映像から取得される複数特徴を用いた個人認証,” 電子情報通信学会論文誌, Vol.J97-A, No.12, pp. 735-748, (2014)
- [2] 松本圭祐, 真部雄介, 菅原研次: “骨格情報を用いた歩容認証における有効なカメラ配置の検討,” *Proc. of JSAI2013*, 2H1-5, (2013).
- [3] 杉浦一成, 榎原靖, 八木 康史: “全方位カメラを用いた複数方向の観測による歩容認証,” 情報処理学会論文誌 コンピュータビジョンとイメージメディア, Vol. 49, No. 2(CVIM 22), pp.76-85, (2018).
- [4] 山川剛史, 谷口和彦, 木綿俊男, 小橋昌司, 近藤克哉, 畑豊: “足底情報による動的バイオメトリクス個人認証,” *Proc. of 23rd Fuzzy System Symposium*, (2007).
- [5] 柏本幸俊, 荒川豊, 安本慶一: “床に貼り付けた振動センサによる屋内位置推定手法,” *IPSJ SIG Technical Report*, Vol. 2015-MBL-74, No.31, pp. 135-138, (2015).
- [6] 2019 Mono Wireless Inc: TWELITE 2525A, <https://mono-wireless.com/jp/products/TWE-Lite-2525A/>
- [7] Shijia Pan, Ningning Wang, Yuqiu Qian, Irem Velibeyoglu, Haeyoung Noh, Pei Zhang: “Indoor Person Identification through Footstep Induced Structural Vibration,” *Proc. of the 16th Int'l Workshop on Mobile Computing Systems and Applications*, pp. 81-86, (2015).
- [8] Sony Network Communications Inc: Neural Network Console, <https://support.dl.sony.com/docs-ja/>