

センサ内蔵型モバイル機器を用いた ジェスチャ認識に関する一考察

村尾和哉^{†1} 寺田 努^{†2} 矢野 愛^{†3}
松倉隆一^{†3} 西尾章治郎^{†1}

近年、計算機の小型化に伴い加速度センサなどの小型センサを内蔵するデバイスが数多く登場しその応用に注目が集まっている。特に携帯電話や家庭用ゲーム機のコントローラにジェスチャ認識技術を用いることで、スクロールなどの直観的操作や立体図形描画などの従来のインタフェースでは困難であった入力が可能になる。しかし、デバイスに搭載されているセンサは基本的に1個であり、複数個のセンサを搭載した場合の精度への影響に関する研究は行われていない。また認識可能なジェスチャの種類が認識精度に与える影響に関する研究も少ない。そこで、本研究では9個の加速度センサを配置した盤面を用いて27種類のジェスチャのデータを採取し、デバイスに内蔵されたセンサの個数や配置およびジェスチャの種類によるジェスチャ認識精度への影響に関する調査を行う。

A Study of Gesture Recognition with Sensor-contained Mobile Devices

KAZUYA MURAO,^{†1} TSUTOMU TERADA,^{†2} AI YANO,^{†3}
RYUICHI MATSUKURA^{†3} and SHOJIRO NISHIO^{†1}

Recently, along with a downsizing of computers, there has been increasing interests in sensor-contained devices and their applications. Especially, mobile phones or video game controllers using gesture recognition technologies would enable intuitive operations such as scroll and drawing 3D objects. However, the number of sensors installed on the device is basically one, and influences on the recognition accuracy in the case of using multiple sensors has not been tackled on. Moreover, the influence of the variety of gestures has not been examined. In this paper, we hold experiment capturing 27 kinds of gestures with 9 accelerometers and gyroscopes, and investigate influences on the recognition accuracy by changing the number and the place of sensors and the variety of gestures.

1. はじめに

近年、マイクロエレクトロニクスの発展による計算機の小型化に伴い、加速度センサなどの小型センサを内蔵するデバイスが数多く登場し、その応用が注目されている。特に iPhone や Android 端末などの携帯電話や Wii や PS3 などの家庭用ゲーム機のコントローラにも加速度センサが内蔵され、そのセンサの値から端末の傾きや動きなどを検出してユーザの直観的な操作や複雑な入力を実現している。

市販されているデバイスに搭載されている加速度センサは消費電力や実装面積の観点から基本的に1つであり、複数のセンサを搭載しているものは少ない。ユーザが携帯電話をポケットに入れたり、腕に装着した状態で生活し、その運動を認識するなど携帯電話自体が回転などの細かな動きをしない場合は1個のセンサで高い精度が得られるが、携帯電話を把持した状態で左右に振ったり傾けたりするようなジェスチャ動作を認識するには複数のセンサによる精度への影響があると考える。しかし、センサを複数個搭載することによる認識精度の影響を調査した研究は筆者らの知る限り行われていない。複数のセンサを用いることで認識精度の向上が得られるのであれば、デバイス開発者はハードウェア面の制約と認識精度のトレードオフを考えることができる。

また、デバイスを把持しながら行うジェスチャはさまざま考えられる。基本的に認識するジェスチャの種類が増加すれば認識精度は低下するが、厳密に言うと認識精度は同一動作のデータ空間的な近さである再現性能と異なる動作のデータ空間的な散らばりである判別性能で決まる。センサの種類によっては認識困難なジェスチャも存在すると考えられるが、多数のジェスチャを取り扱う研究も筆者らの知る限り行われていない。

そこで本研究では、9個の加速度センサと角速度センサを配置した盤面を用いて27種類のジェスチャのデータを採取し、すべてのセンサの組合せにおける認識精度を計測し、センサの個数および配置場所による影響を調査する。また、27種類のジェスチャについても認識精度の観点から考察を行う。

以下、2章で関連研究について述べ、3章で実験について述べ、4章で実験結果とその考

^{†1} 大阪大学大学院情報科学研究科マルチメディア工学専攻

Department of Multimedia Engineering, Graduate School of Information Science and Technology, Osaka University

^{†2} 神戸大学大学院工学研究科電気電子工学専攻

Department of Electrical and Electronic Engineering, Graduate School of Engineering, Kobe University

^{†3} 富士通研究所ヒューマンセントリックコンピューティング研究所

Fujitsu Laboratories Ltd., Human-centric Computing Laboratory

表 1 デバイス内のセンサ
 Table 1 Sensors installed on the devices

メーカー	デバイス	センサの数	センサのメーカー
Apple Inc.	iPod	3 軸加速度 1 個	伊仏 STMicroelectronics
Apple Inc.	iPhone	3 軸加速度 1 個	伊仏 STMicroelectronics
米 T-Mobile	T-Mobile G1	3 軸加速度 1 個 3 軸地磁気 1 個	旭化成マイクロシステム
Nintendo	Wii リモコン	3 軸加速度 1 個	伊仏 STMicroelectronics
Nintendo	ヌンチャク	3 軸加速度 1 個	米 Analog Devices, Inc.
Nintendo	Wii モーションプラス	2 軸角速度 1 軸角速度	米 INVENSENSE EPSON TOYOCOM

察を示し、5章で本研究をまとめる。

2. ジェスチャ認識精度への影響の要因

本節ではジェスチャ認識精度へ影響を与える要因として考えられる加速度センサの個数とその配置および認識するジェスチャの種類について述べる。このほかにデバイスの形状も要因として考えられるが、ジェスチャ認識精度向上という目的のために携帯電話などのデバイスの形状を変更することは現実的ではないため本研究では議論しない。

2.1 加速度センサの数および配置

市販されているデバイスに内蔵されているセンサの種類、個数および製造メーカーを表 1 に示す。Apple Inc. の iPod や iPhone、Nintendo の Wii リモコンやヌンチャクは 3 軸加速度センサを 1 つ搭載している。Wii リモコンに接続する拡張デバイスとして Wii モーションプラスが発売されており、2 軸の角速度センサと 1 軸の角速度センサを組み合わせることで 3 軸の角速度を検出している。これにより卓球の球にカーブをかける手首のひねりなども検出できる。また、T-Mobile の Android 携帯電話 T-Mobile G1 には 3 軸加速度センサと 3 軸地磁気センサが 1 つになったセンサが 1 個搭載されている。このように、搭載されているセンサは基本的に 1 個であり、同一のセンサを複数個搭載しているデバイスは筆者らの知る限り販売されていない。

また図 1 から図 4 に iPod、iPhone、Wii リモコン、T-Mobile G1 内に搭載されているセンサの位置を示す。ただし、T-Mobile G1 の加速度センサは地磁気センサと同一チップである。これらのデバイス内のセンサの位置は決まっておらず、特に同じ Apple Inc. 製のデバイスである iPod と iPhone のセンサの位置も統一されておらず、センサの位置は基盤の実装



図 1 iPod
 Fig.1 iPod

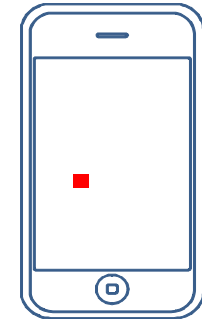


図 2 iPhone
 Fig.2 iPhone



図 3 Wii リモコン
 Fig.3 Wii remote

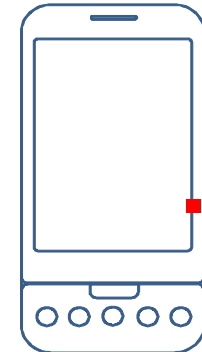


図 4 T-Mobile G1
 Fig.4 T-Mobile G1

の都合による場所が大きいと考えられる。

ここで、筆者らの先行研究¹⁾では、ウェアラブルコンピューティング環境における複数の装着型加速度センサを用いた行動認識システムにおいて、両手、両足、腰の 5 か所のセンサのすべての組合せにおける認識精度を計測している。その結果、センサの個数が減少すると認識精度は低下し、同数のセンサを装着していてもセンサの組合せによって精度が異なる結果が得られている。

しかし、このような腕や足など各センサが独立した動きをする環境ではセンサ配置による

影響は大きいですが、1つのデバイス内に複数のセンサを搭載する環境におけるセンサの個数や位置によるジェスチャ認識精度への影響を評価した研究はこれまでになく、本研究により、センサの個数や位置を決定する際に消費電力や実装面積や配線等実装上の都合だけでなく、認識精度も考慮できると考える。

2.2 認識対象となるジェスチャ

ジェスチャ認識精度に影響を与える要因としてセンサの個数や配置のほか認識対象となるジェスチャが存在する。これまでに加速度センサを用いたジェスチャ認識技術を応用したさまざまなシステムが提案されている。例えば柏木ら²⁾は情報の送り手と受け手側で同じ動作を行うことで直接情報の受渡を行うシステムを提案している。このシステムでは円、四角、三角の3種類の動作で実験を行っている。

また、金ら³⁾は1個の3軸加速度センサを用いて携帯電話の格納動作を認識し、携帯電話の所有場所を判定する手法を提案している。この手法では携帯電話の格納場所としてズボンの前ポケット、ズボンの後ろポケット、胸ポケット、ジャケットのポケット、首に下げている5か所を想定し、手から格納場所への格納動作および格納場所からの取り出し動作を分類している。

Liuら⁴⁾はNokia研究所が提唱する直線や円などの8種類の図形を1個の3軸加速度センサで認識している。この研究では長期間のサンプリングにより8名の被験者から4000以上のサンプルを収集している。認識アルゴリズムはダイナミックタイムワーピング⁵⁾を採用しており、学習したデータを逐次更新することで98.6%の精度を達成している。

このように、携帯電話のようなセンサ内蔵型モバイル機器を用いたジェスチャ認識に関する研究は数多く行われているが、扱うジェスチャの数が少なく、ジェスチャの数が増えた場合の影響に関する取り組みは行われていない。

ここで、ジェスチャ認識を支援するツールとしてWesteynらが提案するGeorgia Tech Gesture Toolkit⁶⁾がある。Getsure Toolkitは既存のHMM(Hidden Markov Model)ツールキット⁷⁾をHMMの知識がないユーザにも簡単に利用できるようにしたものである。Gesture Toolkitの応用例が4つ述べられているが、その1つに手首と肘に装着した3軸加速度センサを用いたジェスチャ認識があり、「やすりをかける」や「のこぎりで切る」、「ねじを回す」など10種類のジェスチャを93.3%の精度で認識している。しかし、この研究で扱われているジェスチャはある一定時間継続される動作も含まれている。本研究で扱うジェスチャは1回きりの動作を想定する。

3. 実験

本節では加速度センサを用いたジェスチャ認識精度に関する実験について述べる。

3.1 実験環境

実験では8名の被験者に携帯電話やPDAの利用を想定して作成した図5に示す9個のセンサを配置した盤面を用いて表2に示す27種類のジェスチャを各10回ずつ計270サンプル採取した。盤面のサイズはW117×H155×D16(mm)で、重さは200gである。使用したセンサはWireless T社製WAA-006⁸⁾で、3軸加速度センサおよび3軸角速度センサを搭載している。被験者は全員右手にデバイスを持ってジェスチャを行った。サンプリング周波数は50Hzである。一般的に多くのデバイスに搭載されているセンサは加速度センサであるが、角速度センサの有効性も検証するために本研究では加速度3軸と角速度3軸の計6軸のデータを採取した。ジェスチャを被験者に教示する際は表2に示す言葉や図ではなく実際の動作を示して行っており、個人差はあるが被験者の解釈の違いによる影響は少ない。

3.2 特徴量抽出

ジェスチャ認識を行う場合、ジェスチャ部分の波形を切り取る必要がある。本研究では被験者が各ジェスチャ間に静止させ、過去数秒間の平均値と現在値の乖離から装着者の動きを検出し、ジェスチャ開始点の検出を行う。具体的には、現在時刻を $t = T$ とすると、次式に示すように、時刻 $t = T$ から過去10サンプル(200ミリ秒)のセンシングデータの移動平均 $\bar{x}(T)$ と現在のセンシングデータ $x(T)$ の差の絶対値が閾値 ε 以上である場合に、装着者に動きがあると判定する。

$$\bar{x}(T) = \frac{1}{10} \sum_{t=T-9}^T x(t) \quad (1)$$

この移動平均 $\bar{x}(t) \pm \varepsilon$ の領域をEpsilon tubeと呼ぶ。一度波形がEpsilon tube外に出てから再びEpsilon tube内に戻り、その後200ミリ秒間滞在し続けるまでの波形をジェスチャとして検出する。予備実験より静止時においてセンサが出力するノイズが50mG程度であることを確認したため、本章では $\varepsilon = 100$ [mG]とした。

切り取ったジェスチャ波形の平均値および分散値を抽出し特徴量とした。各センサは加速度センサの場合3軸であり、6次元の特徴量が得られ、 n 個のセンサを用いる場合、特徴量は $6n$ 次元となる。3軸角速度の場合も同様に $6n$ 次元となり、加速度と角速度の両方を用いる場合は $12n$ 次元となる。



図5 9個のセンサを配置した実験用盤面
 Fig.5 An experimental board with nine sensors.

さらに、加速度と角速度はスケールが異なり等価に扱うことが出来ないため、次に従い正規化し、D次元の特徴ベクトル $Z(t) = (z_1(t), z_2(t), \dots, z_D(t))$ (平均0, 分散1) を得る。

$$Z(t) = \frac{X(t) - M}{S} \quad (2)$$

ここで $M = (m_1, m_2, \dots, m_D)$ および $S = (s_1, s_2, \dots, s_D)$ は正規化前のデータ $X(t) = (x_1(t), x_2(t), \dots, x_N(t))$ の各成分の平均 $m_i = \frac{1}{N} \sum_{t=1}^N x_i(t)$ および標準偏差 $s_i = \sqrt{\frac{1}{N} \sum_{t=1}^N \{x_i(t) - m_i\}^2}$ である。リアルタイムで認識する際は、学習データの平均および標準偏差を用いて認識するデータを正規化する。

各ジェスチャから得られた特徴量にジェスチャラベルを付与して評価用データセットを作成し、次節で述べるクラスタリング精度によって各ジェスチャの分類精度を計測する。

3.3 クラスタリング精度

クラスタリングはあるデータセット内のデータを距離の近いもの同士でグループ化するこ

表2 ジェスチャー一覧
 Table 2 List of Gestures

番号	ジェスチャ	図
1	前に傾ける	
2	奥に傾ける	
3	左に傾ける	
4	右に傾ける	
5	上側を2回タップする	
6	右側を2回タップする	
7	左に2回振る	
8	右に2回振る	
9	カードをきる	
10	ふるいにかける感じで側面を叩く	
11	すくいあげる	
12	机の上にトランプを広げる	
13	広げたトランプをしまう	
14	長い側面を机にトントンする	
15	基盤面に平行に机にトントンする	
16	基盤面を2回ノックする	
17	ひっくり返す	
18	机の上で時計回りに1回転させる	
19	上に移動する	
20	下に移動する	
21	左に移動する	
22	右に移動する	
23	手前に移動する	
24	奥に移動する	
25	丸	
26	三角	
27	四角	

とでデータ構造を明らかにする手法である。クラスタリング精度⁹⁾ はデータの属するクラスおよびクラス内に属するデータの割合からデータセットの判別性能を測る手法であり、純粋にデータの距離によって分類するため認識器に依存しない結果が得られる。クラスタリング精度を用いることでクラスタリング後のデータの分散度の観点から特徴量を順位付けすることができる。本節ではクラスタリング手法として文献⁹⁾ で採用されている k-means を用いた。クラス数は 100 とした。

理想的に完全な判別性能を有する場合、1 つのクラスに属するデータは全て同じジェスチャのものとなる。これは特徴量を各ジェスチャごとに完全に分離可能で認識器の入力として適していることを意味する。最悪の場合、それぞれのクラスに属するデータの各ジェスチャのラベルの割合がデータ全体に対する各ジェスチャのラベルのデータの割合（事前確率）と一致する。これは特徴量を各ジェスチャごとに全く分類できておらず認識器の入力として適していないことを意味する。

クラスタリング精度では、はじめに各クラス i における各ジェスチャのラベル j のデータ分布

$$p_{i,j} = \frac{|C_{i,j}|}{\sum_j |C_{i,j}|} \quad (3)$$

を調べる。ここで、 $C_{i,j}$ はクラス i に属するデータのうちジェスチャのラベルが j であるデータの集合である。続いて次式に従いこれらの重みづけ和をとることで各ジェスチャ j のクラスタリング精度 p_j を得る。

$$p_j = \frac{|\sum_i p_{i,j} C_{i,j}|}{\sum_i |C_{i,j}|} \quad (4)$$

上式において、クラスタリング精度が 1 に近い値となるジェスチャが属するクラスデータの大部分はそのジェスチャのデータであり、判別性能が高いことを意味する。また、各クラスデータのデータサンプル数を重みとして用いることで、小さなクラスの影響を小さくしている。

ここで、ジェスチャ認識アルゴリズムとして波形自体を伸縮して類似度を計算するダイナミックタイムワーピング (DTW)⁹⁾ も提案されている。「時計回りの円を描く」と「反時計回りの円を描く」というようなジェスチャは特徴量に変換すると誤認識するため DTW を用いる必要があるが、本研究で扱うジェスチャは始点と終点が一致して経路だけが異なるものは含んでいないため特徴量に変換する手法を採用した。

4. 結 果

4.1 センサの個数による分類精度への影響

各センサの個数における「加速度のみを利用した場合」、「角速度のみを利用した場合」、「加速度と角速度を利用した場合」の平均クラスタリング精度を表 3 および表 4 に示す。表 3 はジェスチャを行っている波形全体から抽出した特徴量を用い、表 4 はジェスチャ開始 1 秒間の波形から抽出した特徴量を用いた場合の結果である。

波形全体と開始 1 秒間を比較するとほぼすべてにおいて開始 1 秒間の特徴量を利用した場合が優位であることがわかる。これは、ジェスチャ終了部分の動作は元に戻る位置に依存し、ばらつきが大きいと考えられる。

また、センサの個数に注目すると、僅かずつではあるが使用センサ数が増加するにつれて精度は上昇している。精度の上昇率をみると 1 個から 2 個に変化した場合が大きい。

さらに利用するセンサに注目すると、角速度のみを用いた場合の精度は加速度センサと比較して劣っている。これは 4.2 節でも述べるが衝撃を与えたり、平行移動させるジェスチャは角速度に現れないためであると考えられる。一方、加速度は複数利用するとセンサ間の値の大きさの違いから角速度と似た成分を得られるため精度が高い。両者を利用した場合は衝撃、平行移動、回転のすべてを検出できるためさらに精度が改善している。

結果より、開始 1 秒間のデータを利用した結果の方が優位であるため、以降の結果は開始 1 秒間のデータを利用した場合のみを示す。

4.2 センサの配置による分類精度への影響

9 個のセンサの組合せ $2^9 - 1 = 511$ 通り、センサ 1 個の組合せ 9 通り、センサ 2 個の組合せ ${}^9C_2 = 36$ 通りのそれぞれ精度上位 5 位のセンサ配置およびクラスタリング精度を表 5、表 6、表 7 に示す。ただし、1 つもセンサを利用しない組合せは除いている。表 5 は加速度センサのみ、表 6 は角速度センサのみ、表 7 は両方を利用した場合である。センサの配置を 9 ビットで表現しており、下位からのビットと図 5 のセンサの番号が対応しており、1 が有り、0 が無しを示す。例えば 001000010 はセンサ 2 とセンサ 7 の 2 個配置していることを示す。

結果より、1 個配置する場合は盤面上側に配置した場合が上位に来ていることが分かる。これは、ジェスチャの大半を盤面の右下を把持した状態でっており、支点から距離があることで大きな成分が得られるためと考える。

また 2 個配置する場合は互いの距離が開くように配置する場合が上位に来ている。これ

表 3 各センサの個数におけるクラスタリング精度 (波形全体)
 Table 3 Cluster precision for each number of sensors. (Whole waveform)

	センサの個数 [個]								
	1	2	3	4	5	6	7	8	9
加速度	0.927	0.941	0.944	0.945	0.947	0.947	0.947	0.948	0.948
角速度	0.830	0.837	0.841	0.845	0.848	0.850	0.852	0.851	0.857
両方	0.966	0.969	0.972	0.973	0.973	0.973	0.972	0.974	0.977

表 4 各センサの個数におけるクラスタリング精度 (ジェスチャ開始 1 秒間)
 Table 4 Cluster precision for each number of sensors. (For first one second)

	センサの個数 [個]								
	1	2	3	4	5	6	7	8	9
加速度	0.935	0.948	0.951	0.952	0.953	0.954	0.954	0.953	0.952
角速度	0.850	0.859	0.864	0.865	0.869	0.871	0.872	0.874	0.874
両方	0.970	0.973	0.974	0.975	0.975	0.975	0.975	0.975	0.969

は、互いの距離が近いと似たデータしか取得できず冗長となってしまうためと考える。

全体については盤面に分散させた配置が上位にきていることがわかる。

4.1 節および 4.2 節の結果から、基本的にセンサの個数は 1 個で高い精度が得られることがわかる。その際は支点となる位置から離れた位置が有効である。しかし、センサが 2 個になるデメリットを考慮しても 1% 程度の精度向上を求めるのであれば 2 個使用した方がよい。3 個以上センサを使用しても精度は僅かしか向上しないが、僅かでも精度向上を求めるのであれば多くのセンサを利用することも有効である。その際は、盤面に分散した配置が有効である。

4.3 ジェスチャの種類による分類精度への影響

4.2 節の結果で得た全体で最も精度の良い組合せ、1 個のとき最も精度の良い組合せ、2 個のとき最も精度の良い組合せにおける各ジェスチャのクラスタリング精度を表 8 に示す。この結果はジェスチャ開始 1 秒間の特徴量を利用した場合のものである。結果より、加速度センサのみを用いた場合、ジェスチャ 12, 13, 19~24 の精度が他のジェスチャと比較して劣っている。これらのジェスチャの共通点としていずれも平行移動していることが挙げられる。平行移動時の加速度の値は小さくなるため、分類精度が低下すると考えられる。

角速度に関しても同様に平行移動であるジェスチャ 19~24 の精度が悪い。これは平行移動は角速度成分に現れないためである。しかし、12 と 13 の精度は加速度と比較して優位で

表 5 クラスタリング精度が上位 5 位のセンサ配置 (加速度・ジェスチャ開始 1 秒間)
 Table 5 Sensor allocations of top-5 in Cluster precision. (Accelerometers/For first one second)

		順位				
		1	2	3	4	5
全体	個数	4	5	6	6	7
	配置	110100100	110110100	010111110	101010111	111110110
	精度	0.964	0.964	0.962	0.962	0.962
1 個	配置	000100000	100000000	010000000	000010000	001000000
	精度	0.948	0.943	0.941	0.939	0.938
2 個	配置	000010100	001100000	001000100	010010000	010000100
	精度	0.956	0.956	0.955	0.954	0.953

表 6 クラスタリングが上位 5 位のセンサ配置 (角速度・ジェスチャ開始 1 秒間)
 Table 6 Sensor allocations of top-5 in Cluster precision. (Gyroscopes/For first one second)

		順位				
		1	2	3	4	5
全体	個数	4	3	3	2	3
	配置	001100110	000001011	000101010	010000010	001001010
	精度	0.899	0.896	0.895	0.894	0.891
1 個	配置	000000010	000010000	000000100	000001000	100000000
	精度	0.877	0.852	0.851	0.848	0.848
2 個	個数	010000010	000001010	000100010	100000010	000000011
	精度	0.894	0.886	0.882	0.880	0.879

表 7 クラスタリング精度が上位 5 位のセンサ配置 (加速度と角速度・ジェスチャ開始 1 秒間)
 Table 7 Sensor allocations of top-5 in Cluster precision. (Accelerometers+Gyroscopes/For first one second)

		順位				
		1	2	3	4	5
全体	個数	2	5	8	3	4
	配置	010010000	111000110	111111110	100001100	110101000
	精度	0.983	0.982	0.982	0.982	0.981
1 個	配置	100000000	000100000	000010000	000001000	000000010
	精度	0.980	0.976	0.973	0.973	0.970
2 個	配置	010010000	000011000	010000100	000000011	100010000
	精度	0.983	0.979	0.977	0.977	0.977

ある。これはトランプを広げたりしまう動作は手首を中心とした扇状で回転成分が含まれるため検出できたと考える。逆にジェスチャ 5, 6, 14, 15, 16 のような叩く場合に得られる衝撃は加速度成分に現れ、角速度成分に現れないため、加速度で高く角速度で低い精度となっている。

表 8 各ジェスチャのクラスタリング精度 (各被験者の平均)
 Table 8 Cluster precision for each gesture. (Individual average)

ジェスチャ	全体で 1 位			1 個配置で 1 位			2 個配置で 1 位		
	加速度	角速度	両方	加速度	角速度	両方	加速度	角速度	両方
1	1.000	0.860	0.994	1.000	0.912	0.991	1.000	0.905	0.994
2	1.000	0.843	1.000	0.990	0.893	1.000	0.973	0.898	1.000
3	1.000	0.904	1.000	1.000	0.868	1.000	1.000	0.879	1.000
4	1.000	0.947	1.000	1.000	0.913	1.000	1.000	0.910	1.000
5	0.970	0.852	0.979	0.956	0.840	0.974	0.959	0.828	0.979
6	0.968	0.850	0.986	0.949	0.829	0.980	0.963	0.807	0.986
7	0.992	0.906	1.000	1.000	0.916	0.994	0.969	0.908	1.000
8	0.992	0.928	1.000	1.000	0.947	0.994	0.967	0.936	1.000
9	1.000	0.926	1.000	1.000	0.916	1.000	1.000	0.908	1.000
10	0.992	0.975	1.000	0.975	0.933	1.000	0.979	0.956	1.000
11	0.984	0.948	0.991	0.980	0.927	0.983	0.989	0.940	0.991
12	0.794	0.984	0.994	0.734	0.975	1.000	0.812	0.994	0.994
13	0.795	0.990	1.000	0.735	0.990	1.000	0.807	1.000	1.000
14	1.000	0.923	1.000	1.000	0.814	1.000	1.000	0.955	1.000
15	0.992	0.873	0.994	0.983	0.873	0.994	0.973	0.881	0.994
16	1.000	0.867	1.000	0.960	0.766	0.985	0.985	0.842	1.000
17	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
18	1.000	1.000	1.000	0.996	1.000	1.000	1.000	1.000	1.000
19	0.957	0.791	0.953	0.923	0.689	0.952	0.901	0.798	0.953
20	0.945	0.814	0.954	0.910	0.772	0.959	0.901	0.820	0.954
21	0.923	0.782	0.943	0.873	0.737	0.910	0.915	0.742	0.943
22	0.913	0.846	0.940	0.878	0.785	0.907	0.920	0.769	0.940
23	0.934	0.764	0.913	0.915	0.743	0.896	0.913	0.775	0.913
24	0.915	0.731	0.901	0.893	0.725	0.868	0.887	0.717	0.901
25	0.992	1.000	1.000	0.968	1.000	1.000	1.000	1.000	1.000
26	1.000	0.982	1.000	1.000	0.955	0.990	1.000	0.982	1.000
27	0.982	0.982	1.000	0.973	0.957	0.989	0.991	0.980	1.000
平均	0.964	0.899	0.983	0.948	0.877	0.976	0.956	0.894	0.983

これらに対し、加速度と角速度を併用した場合は両者の長所を活かしており、加速度および角速度で精度が劣っていたジェスチャに対して精度が改善している。

ところで、ここまでの実験結果は個人のデータのみを用いて分類精度を計測した。しかし、ジェスチャ認識を利用するアプリケーションを販売する際に、あらかじめ利用者でない

表 9 センサ 9 個配置時の各ジェスチャのクラスタリング精度 (全被験者の混合データ)
 Table 9 Cluster precision for each gesture with nine sensors. (Mixture of all subjects)

ジェスチャ	加速度	角速度	両方
1	0.868	0.386	0.864
2	0.570	0.434	0.707
3	0.895	0.596	0.932
4	0.812	0.699	0.851
5	0.362	0.185	0.323
6	0.473	0.209	0.611
7	0.727	0.701	0.678
8	0.615	0.646	0.570
9	0.874	0.236	0.823
10	0.521	0.524	0.645
11	0.534	0.379	0.466
12	0.500	0.815	0.822
13	0.487	0.846	0.842
14	1.000	0.212	1.000
15	0.739	0.356	0.734
16	0.522	0.211	0.461
17	0.987	1.000	1.000
18	0.782	1.000	1.000
19	0.583	0.178	0.529
20	0.460	0.208	0.429
21	0.412	0.377	0.514
22	0.441	0.422	0.478
23	0.527	0.301	0.524
24	0.480	0.212	0.455
25	0.655	0.909	0.901
26	0.685	0.651	0.774
27	0.428	0.612	0.670
平均	0.627	0.493	0.689

人から採取したデータを用いて認識することが考えられる。そこで、全ての被験者のデータを 1 つにして、つまり利用者以外の人間のデータ混じっているデータを用いて計測したクラスタリング精度を表 9 に示す。結果より表 8 と比較して利用するセンサ間の傾向は同様であるが、精度が大幅に低下している。これはジェスチャデータの個人差が大きく、他のジェスチャとの分類が困難になっているためであり、利用者のジェスチャデータを学習しないと

認識精度は非常に低くなることを意味している。そのため、実際の利用では利用者のデータを重点的に学習することでジェスチャ認識精度は向上させる必要があると考える。

最後に、これらのジェスチャの分類精度が実用上十分であるかの議論はアプリケーションに依存するが、ジェスチャを用いた機器の操作では誤認識によって意図しない操作が行われる場合があるため、認識結果に自信がない時は何もリアクションを起こさないなどの工夫をすることでユーザへの負担を軽減できると考える。

5. ま と め

本研究ではセンサ内蔵型デバイス内のセンサの個数および配置、ジェスチャの種類が認識精度に与える影響を実験により計測した。実験では9個の加速度センサと角速度センサを格子状に配置した盤面を用いて27種類のジェスチャのデータを収集しクラスタリング精度を求めた。結果より加速度センサと角速度センサそれぞれに得意なジェスチャと不得意なジェスチャが存在し、特に加速度センサは叩くなどの衝撃のあるジェスチャは得意で、ゆっくりと平行に移動する動作が不得意である。逆に角速度は回転するジェスチャが得意で、衝撃が苦手なため、両者を併用することで精度の向上を確認した。しかし、センサの種類によって得意不得意なジェスチャが存在するため、アプリケーション開発者は不得意なジェスチャを採用しないなどの工夫が必要である。

また、ジェスチャ部分全体の波形を用いて認識するよりも、ジェスチャ開始数秒の波形のみを用いた方がジェスチャ終了時のばらつきを含まずに高い精度を得ることが分かった。さらに、センサ配置はジェスチャの支点である把持位置から遠く、かつ複数配置する場合は隣接させない方が有効であると分かった。

本研究により、従来まで基盤の実装面積や配線、消費電力等の都合でセンサを搭載方法が決まっていたが、認識精度の観点から複数センサの利用も選択肢となるため、デバイスの開発の幅が広がると考える。

現在は加速度センサと角速度は同一の場所に配置することを想定しているが、両者を独立して扱いさらに効率的な配置について考察を行う予定である。

謝辞 本研究の一部は、科学研究費補助金基盤基盤 (A)(17200006)、特定領域研究 (21013034) および特別研究員奨励費 (21・249) の支援によるものである。ここに記して謝意を表す。

参 考 文 献

- 1) Muraio, K., Terada, T., Takegawa, Y., and Nishio, S.: A Context-Aware System that Changes Sensor Combinations Considering Energy Consumption, in *Proceedings of International Conference on Pervasive Computing (Pervasive 2008)*, pp.197–212 (May 2008).
- 2) 柏木晶範, 本間弘一, 福地健太郎, 小池英樹: ジェスチャを真似ることによる情報の受け渡し, インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2007), pp. 147–148 (Dec. 2007).
- 3) 金 春善, 藤波香織: 格納動作パターンを用いた携帯電話の所有場所判定, 情報処理学会研究報告 (コピキタスコンピューティングシステム研究会 2009-UBI-21), Vol.2009, No.17, pp.23–30 (Mar. 2009).
- 4) Liu, J., Wang, Z., Zhong, L., Wickramasuriya, J., and Vasudevan, V.: uWave: Accelerometer-based Personalized Gesture Recognition and its Applications, in *Proceedings of IEEE International Conference on Pervasive Computing and Communication (PerCom 2009)*, pp.1–9 (Mar. 2009).
- 5) Myers, C. S. and Rabiner, L. R.: A Comparative Study of Several Dynamic Time-warping Algorithms for Connected Word Recognition, *The Bell System Technical Journal*, Vol. 60, pp. 1389–1409 (Sept. 1981).
- 6) Westeyn, T., Brashear, H., Atrash, A., and Starner, T.: Georgia Tech Gesture Toolkit: Supporting Experiments in Gesture Recognition, in *Proceedings of International Conference on Multimodal Interfaces (ICMI '03)*, pp. 85–92 (Nov. 2003).
- 7) HTK Hidden Markov Model Toolkit home page: <http://htk.eng.kcam.ac.uk/>.
- 8) Wireless Technologies, Inc.: <http://www.wireless-t.jp/>.
- 9) Huynh, T. and Schiele, B.: Analyzing Features for Activity Recognition, in *Proceedings of the Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies (sOcEuSAI 2005)*, pp. 159–163 (Oct. 2005).