

ポケット内のスマートフォンによる 両足ジェスチャ認識手法の提案と分析

田村 柁優紀^{†1} 中村 聡史^{†1}

概要：スマートフォンは身近なデバイスの一つとして日常の様々な場面で利用されているが、画面を見ながらのタッチ操作は、赤ちゃんを抱っこして手がふさがっている場合や料理して手が汚れていて画面に触りたくない場合などには不向きである。そこで本研究では、ポケットに入れたスマートフォンを用いて両足のジェスチャ認識を行うため、加速度や角加速度、音圧を利用した手法を提案する。またプロトタイプシステムを実装し、使用実験からその有用性を検証する。

キーワード：ジェスチャ認識，両足，スマートフォン，デバイス操作，SVM，DP Matching

1. はじめに

近年、スマートフォンは身近なデバイスの一つとして日常の様々な場面で利用されている。総務省によると平成26年時点でのスマートフォンの普及率は64.2%とされている[1]。また、スマートフォンはWebサイトの閲覧や、ゲームアプリの実行、メールの送信などの連絡手段、音楽の聴取などで広く用いられており、多くの人々に浸透していると言える。

スマートフォンは、主にタップやスワイプ、長押しなどによって操作する。これらの操作は目と手の両方を使用することで、直感的な対象の操作を可能にしている。しかし、荷物などで手がふさがっている状況、手が汚れていてスマートフォンに触りたくない状況や、赤ちゃんを抱いている状況などにおいては操作が困難になるという問題がある。このような問題の改善案として音声認識によるデバイス操作も考えられるが、この方法には繁華街や居酒屋などの賑やかな場所や、図書館などの発声し辛い場所、赤ちゃんを寝かしつけているような状況では利用できない。

先述した問題の改善するため、Scottら[2]はスマートフォンに搭載されている加速度センサを利用し、ズボンのポケットにスマートフォンを入れた状態で足の動きをセンシングすることで、足でのジェスチャ入力とデバイス操作を可能とする手法を提案している。また実験により、スマートフォンを入れるポケットの位置によっては高いジェスチャ認識精度を出すことを明らかにしている。しかし、この研究では実際に足のジェスチャによりデバイス操作を行うシステムを実装している訳ではなく、また、片足のみのジェスチャ認識を対象としているものであった。そこで本稿では、加速度以外にも、スマートフォンに搭載されている角加速度センサやマイクといった多様なセンサを用いることにより、一方のポケットから、両足でのジェスチャ認識が可能であるかを検証する。

具体的には、片足のポケットに入れたスマートフォンで

センシングした加速度、角加速度、音圧を用いて、両足でのジェスチャ認識を行う。また、本システムはリアルタイムでの利用を想定し、ジェスチャ認識の判定に使用する際の最適なフレーム数の検証を行う。さらに、その検証により得られた結果を基にプロトタイプシステムの実装を行い、使用実験を行う事で本提案手法の有用性を検証する。

本研究の構成を以下に述べる。2章では、行動推定や、ジェスチャ認識に関する研究について述べ、本研究の位置付けを行う。3章では提案手法とそれを用いたプロトタイプシステムの実装について述べ、4章では認識精度の算出を行う。5章ではプロトタイプシステムの実装を行い、6章では、そのプロトタイプシステムを用いた使用実験により提案手法の有用性を検証する。7章では実験結果に対する考察を行い、8章にてまとめと今後の課題と展望について述べる。

2. 関連研究

センシング機器を取り付けたユーザのセンサデータから行動推定を行う研究は様々行われている。村尾らの研究[3][4]では、手首、腰、足首の3か所に加速度センサを装着し、それぞれの箇所で取得した加速度の値でSVMとDP Matching、またその二つを組み合わせた提案手法による推定精度を比較することで、行動推定に適した認識手法の検討や行動推定を行っている。大内らの研究[5]では、環境にデバイスを設置せず、センシングデバイスを胸ポケットに入れた状態での加速度センサの値をもとに、「歩行」「作業」「安静」の3状態を95%以上の精度で推定している。河内らの研究[6][7]では、身体上の5か所の既定の場所のうちどこに携帯電話が格納されているかを、歩行中に判定可能なシステムの実装を行っている。一方本研究では、推定を行うだけでなくスマートフォンで取得したセンサの値を用いてリアルタイムにジェスチャを認識し、デバイス操作を行う点で異なっている。

^{†1} 明治大学
Meiji University

センシングデバイスを用いてジェスチャ認識を行う研究は様々行われている。奥村らの研究[8]では、靴底にセンシングデバイスを取り付けることで、歩行中でもジェスチャによりデバイス操作可能なシステムの実装を行っている。吉田らの研究[9]では、3次元空間において、小型の飛行ロボットを用いた人間の身体性を拡張する手法の提案をしている。また小型の飛行ロボットを、自身の体の一部であるかのように操作可能にするジェスチャの提案も行っている。富林らの研究[10]では、DJの行動範囲が制限される問題に対して、装着型無線加速度センサを用いたジェスチャ認識を用いて、機器から離れた場所においてもDJ操作を可能とするシステムの提案を行った。FingerPad[11]は、親指で人差し指に書いた文字やジェスチャを認識させ、その認識結果を基にデバイス操作可能な個人向けかつ繊細なインタラクションシステムを提案しており、座った状態で操作を行った場合横幅1.2mmのターゲットを93%の精度で操作する事が可能であった。しかし、これらの研究ではシステム用のデバイスを使用しているため、手軽にシステムを利用することが出来ない。一方本研究では、システム用の装置を用いることなく、既存のデバイスを用いたジェスチャ認識を行っている点でこれらの研究とは異なる。

加藤らの研究[12]では、携帯型情報機器におけるペンジェスチャ入力UIにおいて、デジタルインク入力とペンジェスチャ入力の区別が難しいという問題に対し、画面上にボタンを配置する手法とペンの停留を用いる手法の2つの手法を用いることで、デジタルインク入力とペンジェスチャ入力の区別を実現する手法の提案を行っている。また石原らの研究[13]では、手に持った状態の携帯端末から取得した加速度の値に対してDP Matchingを用いることにより手軽かつ高精度の個人認証を可能とする3D動作認証を提案し、そのシステムを1か月以上にわたって動作させた場合、本人拒否率を10%以下にできることを明らかにした。しかし、これらの研究では、手や目を必要とするジェスチャ認識であるため、手や目がふさがっている際には利用することが出来ない。一方本研究では、ズボンのポケットにスマートフォンを入れた状態で、目や手を用いずに両足によるジェスチャでデバイス操作を行っている点で異なる。

3. 提案手法

3.1 事前調査

両足のジェスチャ認識においてどのようなジェスチャが適切か、またを調査するための事前調査を行った。

Scottらの研究[1]においては「踵の上下運動」、「つま先の上下運動」、「踵を軸としてつま先を回転させる運動」、「つま先を軸として踵を回転させる運動」の4種類のジェスチャを実装している。ここで、ジェスチャを行う際に重心の

変化が起きることにより、角加速度、加速度の値が変化しやすくと考える。そこで、その4種類の中から、最も体の重心が変動しやすくと、筆者が判断した踵の上下運動を、実験協力者8人に行ってもらった。また、ジェスチャ時の複数のセンサ値を保持したデータを基にジェスチャにおけるデータ変動の傾向分析を行った。

分析を行った結果、両足による踵の上下運動は体の重心が動き易く角加速度、加速度の値も変化し易い傾向にあった。ここで、踵の上下運動の種類には、力を入れる位置により、重心が動くものとそうでないものの2種類が存在した。この両者のジェスチャの力の動かし方を考察したところ、歩くように踵を上下すると重心が動きやすくなる傾向にあった。そこで、歩くように踵の上下するジェスチャを採用する。

一方、生活音などを限りなく排除し、ズボンとの衣擦れ音のみを取得するため周波数を分析した結果、ほとんどのズボンにおいてスマートフォンとの衣擦れ音の周波数が90Hz以下であり、他の音声で90Hz以上となる傾向にあった。そのためこの周波数を利用した認識が可能になると考えられる。

3.2 両足のジェスチャによる操作手法

本研究では片足のポケットにスマートフォンを入れた状態で、両足でのジェスチャ認識を可能とする手法を提案し、システムの有用性を検証する。具体的には、事前調査から体の重心の変化が大きかった踵の上下運動によるジェスチャのみを使用し、そのジェスチャをユーザが行った際に、左右どちらの足によりジェスチャが行われたかをどの程度識別可能か調査する。スマートフォンを入れた足とそうでない足によるジェスチャを行うため、ポケット内のスマートフォンとズボンとの衣擦れによって発生する衣擦れ音の波形や大きさも異なると考えられる。そこで、マイクのセンサを用いることが両足のジェスチャ認識に与える認識精度への影響を検証する。そして、ジェスチャ認識に用いるフレーム数を表すウィンドウ幅の違いによる認識精度の変化について分析を行い、本研究において適切なウィンドウ幅の決定を行う。

なお認識においては、モバイル端末におけるジェスチャ認識の研究[19]で利用されているDP Matchingと、機械学習の1種であるSVM(サポートベクターマシン)などを用い、認識精度の良いものを採用する。

以降では、適切なセンサおよびジェスチャ認識手法について比較実験を実施することにより明らかにする。

4. 認識精度の評価実験

4.1 実験目的

両足でのジェスチャ認識が可能かどうかを検証するため、認識精度の評価実験を行う。そこで、データセット構

築用の Android 用センシングアプリを開発した。本センシングアプリは、随時加速度や角加速度、音圧などについて記録を行うものである。

本システムを用いてデータセットを構築し、関連研究で用いられていた SVM と DP Matching の 2 つの認識手法によるジェスチャ認識精度をそれぞれ検証し、左右どちらの足でジェスチャが行われたかをどの程度識別可能かについて調査を行う。なお、この実験では 1 種類のジェスチャを左足と右足で行い、ジェスチャの識別が出来るか調査すると同時に、本研究に適している認識手法を明らかにする。

4.2 データセット構築

19 歳～23 歳の実験協力者 8 人に、データセット構築用のセンシングアプリを起動中のスマートフォンを右足の前ポケットに入れたその状態で、タブレット端末を用いて実験協力者にジェスチャ課題を提示し、ジェスチャを行ってもらった。ここで、全員の踵の上下運動のジェスチャを統一するため、言葉によるジェスチャの説明と、実際にジェスチャを行う姿を見せることにより教示した。実験協力者のズボンは、ジェスチャ認識を用いたシステムを日常的に利用することを想定しているため、実験協力者がその日履いていたズボンをそのまま着用してもらい実験を行った。

図 1 は実験での課題提示の様子を示している。(ア)の画面では、ジェスチャを何秒後にどちらの足でのジェスチャを行うかについて提示を行い、(イ)の画面では、現在何回目の課題であるか、どちらの足による課題であるかを提示している。実験協力者には実験の際に、(ア)の時には待機



(ア) 待機画面



(イ) タスク提示画面

図 1 実験におけるタスク提示方法

してもらい、一定の待機時間が経過した後に自動で (イ) に切りかわり、それと同時にジェスチャを行うよう実験開始前に指示した。そして、(イ) に切り替わってから一定の課題提示時間が経過した後、(ア) に自動で切り替わる。この試行を計 150 回行ってもらい、両足によるジェスチャのログを取得した。ここで、(イ) に切り変わってすぐにジェスチャが出来るように、(ア) では何秒後にどちらの足でジェスチャを行うかを提示している。なお、提案システムは両手が塞がっている状態を想定しているため、実験協力者には本システムの使用例の 1 つとして考えている赤ちゃんを抱えている状態を再現するため、赤ちゃんと同程度の重さの物を両手に抱え得るように持ってもらった。

データセット構築では、タブレット端末の課題提示時間から 75 フレーム分のデータを保存している。そのデータをシステムのウィンドウ幅に応じて学習データの長さを変更して使用している。また、事前調査により実験を始めてから数回のデータは、ジェスチャ入力が提示時間から遅れているケースが多かった。そのため、今回は慣れを考慮して両足ともに 10 回目のデータから 49 回目のデータを用いてデータセットを構築している。

4.3 性能評価

本研究においてユーザが行ったジェスチャを正確に認識することが重要であるため、再現率と適合率の両方が重要である。そのため、DPM や SVM のシステムの評価はまず再現率 (Recall) と適合率 (Precision) を基にそれぞれのジェスチャの F 値を求めることでシステムの評価を行う。ここで右足による踵の上下運動の F 値(F_r)は以下の様に求めることが可能である。

$$F_r = \frac{2 * Recall * Precision}{Recall + Precision}$$

ジェスチャ毎の F 値が求められた後、3 種類のクラスの F 値の平均を算出した値 F_{all} をシステム全体の評価指標として用いた。 F_{all} の値が高い程ジェスチャ認識の精度が高いことを表す。

本研究では、古いデータで新しいデータを識別した際のジェスチャ認識精度が重要であるため、構築したデータセット内に登録されている各ジェスチャ 40 個のジェスチャデータの中で、最初の試行から 20 個分のデータを学習データとして、残りの 20 個のデータをテストデータとして、F 値の算出を行う。

本研究においては、村尾ら[3]が用いていた加速度センサと、体の重心変化が顕著に表れると筆者が予想した角加速度センサと、マイクにより取得した音の三つのセンサを用いてジェスチャ認識を行う。

4.4 SVM を用いた認識精度の評価実験

4.4.1 SVM による認識

本研究で用いる特徴ベクトルの種類は角加速度 3 軸と加

速度 3 軸の組み合わせと、音圧とそれらのデータの最大値・最小値の組み合わせの合計 31 パターンの特徴ベクトルを用意した。ここで音圧のデータの最小値は何も音を拾っていない時を表すと考えたため、音圧は最大値のみを用いている。特徴ベクトルは以下の通りである。

- A) 角加速度 3 軸 (最大値) 3 次元
- B) 角加速度 3 軸 (最小値) 3 次元
- C) 角加速度 3 軸 (最大値・最小値) 6 次元
- D) 加速度 3 軸 (最大値) 3 次元
- E) 加速度 3 軸 (最小値) 3 次元
- F) 加速度 3 軸 (最大値・最小値) 6 次元
- G) 角加速度 3 軸 (最大値), 加速度 3 軸 (最大値) 6 次元
- H) 角加速度 3 軸 (最大値), 加速度 3 軸 (最小値) 6 次元
- I) 角加速度 3 軸 (最大値), 加速度 3 軸 (最大値・最小値) 9 次元
- J) 角加速度 3 軸 (最小値), 加速度 3 軸 (最大値) 6 次元
- K) 角加速度 3 軸 (最小値), 加速度 3 軸 (最小値) 6 次元
- L) 角加速度 3 軸 (最小値), 加速度 3 軸 (最大値・最小値) 9 次元
- M) 角加速度 3 軸 (最大値・最小値), 加速度 3 軸 (最

表 1 ウィンドウ幅の変化による F_{all} の値 (F_{all} の取りうる最大値は 1)

特徴ベクトル	15	20	25	30	35	40
A	0.6293	0.8258	0.8944	0.8856	0.8797	0.8738
B	0.6280	0.8080	0.8837	0.8725	0.8823	0.8805
C	0.7568	0.9001	0.9305	0.9134	0.9089	0.9151
D	0.5879	0.7545	0.7787	0.7350	0.7966	0.7829
E	0.6218	0.7915	0.8549	0.8727	0.8639	0.8654
F	0.6812	0.8269	0.8494	0.8615	0.8694	0.8659
G	0.6260	0.7912	0.8686	0.8938	0.8877	0.8854
H	0.6706	0.8333	0.8953	0.8953	0.8967	0.8880
I	0.7173	0.8244	0.8862	0.8862	0.9025	0.9157
J	0.7021	0.8256	0.8776	0.8776	0.8987	0.9034
K	0.7083	0.8494	0.8969	0.8494	0.9038	0.9168
L	0.7308	0.8582	0.8930	0.8930	0.9082	0.9088
M	0.7584	0.9044	0.9306	0.9177	0.9088	0.9165
N	0.7659	0.9004	0.9327	0.9240	0.9152	0.4950
O	0.7679	0.8516	0.9029	0.9168	0.9106	0.8864
P	0.4639	0.5205	0.5266	0.5260	0.4916	0.8841
Q	0.6561	0.8235	0.8794	0.8777	0.8882	0.8864
R	0.7164	0.8043	0.8621	0.8657	0.8754	0.8841
S	0.7562	0.8810	0.9132	0.9173	0.9171	0.9187
T	0.6199	0.7583	0.7690	0.7910	0.7929	0.7937
U	0.6691	0.7547	0.8692	0.8735	0.8534	0.8629
V	0.7036	0.8278	0.8767	0.8767	0.8686	0.8571
W	0.6587	0.8033	0.8632	0.8632	0.8846	0.8805
X	0.6743	0.8305	0.8858	0.8927	0.9025	0.8811
Y	0.6997	0.8407	0.8905	0.8998	0.8928	0.8880
Z	0.7088	0.8347	0.8741	0.8851	0.8768	0.8883
AA	0.7143	0.8363	0.8977	0.9164	0.9037	0.9061
BB	0.7238	0.8394	0.9013	0.9051	0.8969	0.9096
CC	0.73123	0.8352	0.8983	0.8942	0.9087	0.9003
DD	0.7472	0.8660	0.9132	0.9071	0.9069	0.9046
EE	0.7509	0.8424	0.9039	0.9189	0.9122	0.9167

- 大値) 9 次元
- N) 角加速度 3 軸 (最大値・最小値), 加速度 3 軸 (最小値) 9 次元
- O) 角加速度 3 軸 (最大値・最小値), 加速度 3 軸 (最大値・最小値) 12 次元
- P) 音圧 (最大値) 1 次元
- Q) ~ EE) については A ~ O までの組み合わせに音圧の最大値を加えた特徴ベクトルである。

4.5 SVM による実験結果

表 1 が 31 パターンの特徴ベクトルによるウィンドウ幅 15, 20, 25, 30, 35, 40 のそれぞれのパターンにおける F 値の結果である。結果としては、ウィンドウ幅 25 における角加速度 3 軸 (最大値・最小値), 加速度 3 軸 (最小値) の 9 次元による特徴ベクトルで作成した分類器による精度が最も高い傾向となった。

4.5.1 ウィンドウ幅による認識精度の違い

ウィンドウ幅における認識精度の結果としては、ウィンドウ幅 25 の時に F 値が最も高くなる傾向にあった。ウィンドウ幅 20, 15 の時の認識精度は期待した程高い値は得られなかった。

4.6 DP Matching を用いた認識精度の評価実験

DP Matching は非線形伸縮パターンマッチであるため、本システムでセンシングしたセンサ値の波形のノイズの除去を行う。そして、センサ値の絶対値の最大によって全てのセンサ値を除算することで正規化を行ったあと離散化した。その後、処理を施したデータに対して DP Matching によりクラスタリングをウィンドウ幅 30 で行った。その結果が表 2 である。この結果から正解のデータを正しく分類出来た割合が 6 割程度となってしまっているため、SVM と比べて高い精度を得ることは出来なかった。

4.7 認識精度の評価実験に対する考察

SVM の N のウィンドウ幅 25 による特徴ベクトルにより作成した分類器での認識精度が、本評価実験において一番高い認識精度であった。そのため、プロトタイプシステムには N の特徴ベクトルにより作成した分類器を用いた実装を行う。また、本研究においては音圧の最大値を用いたことで、全体としては精度が下がる傾向になったが、2 人の

表 2 DPMatching による F 値

センサ	30
角加速度 X	0.6468
角加速度 Y	0.4836
角加速度 Z	0.6590
加速度 X	0.6760
加速度 Y	0.7277
加速度 Z	0.6646
全体	0.6429

実験協力者では音圧の最大値を分類器に含めることで 1% 前後精度が上がる傾向にあった。

5. プロトタイプシステム

第 1 章で述べたように、スマートフォンには様々なセンサが搭載されている。そこで、本システムのジェスチャ認識において重要なセンサを選出するにあたり、様々なセンサ (角加速度, 加速度, マイク) によるセンシングを可能とするシステムの実装を行った。

4 章の結果に基づき最もジェスチャの認識精度が高かった SVM を用いたプロトタイプシステムを実装した。なお、SVM での認識には Libsvm [27] を用いた。Libsvm のカーネルには線形カーネルを使用し、サポートベクタ分類器の種類には C-SVM クラス分類器である C-SVC を使用している。ジェスチャ認識を開始するタイミングについては、あるセンサの軸の数値が閾値を超えた場合にジェスチャ認識を開始する方法と、一定時間ごとにジェスチャ認識を行う方法などが考えられる。しかし、本プロトタイプシステムでは、前者の方法によりジェスチャ認識を開始している。閾値の決定については、ジェスチャログを基にどのセンサの軸を用いて、どの数値を閾値とするかを手動で決定している。

3.1 での事前調査から、マイクによるセンシングでは、高速フーリエ級数変換により 0~90Hz の周波数帯域のみの音圧を抽出し、その周波数帯域における最大値のみを本システムでは用いている。サンプリング周波数は 30Hz 前後であり、全てのセンサデータを同じタイミングで取得している。また、第 4 章の結果から取得した 25 回分のデータを用いてジェスチャ認識を行っている。ジェスチャ認識には、センサのサンプリング周波数が 30Hz より低いセンサも含まれている。しかし、システムのサンプリング周波数を低いサンプリング周波数のセンサに合わせると、サンプリング周波数の高いセンサの変化が取得出来なくなると考え、値に変化がない場合には変化前の値を用いている。また、本プロトタイプシステムでは、事前調査において体の重心変化が大きい傾向にあった踵の上下運動のジェスチャのみを識別する。

同じく第 4 章の結果より、プロトタイプシステムではウィンドウ幅 25 で認識を行う。そのため、ジェスチャ認識では、ジェスチャ認識開始のためのセンサ値が閾値を超えた場合、閾値を超える直前の 9 回分のセンサデータと、直後 16 回分のセンサデータの計 25 回分のセンサデータを基にジェスチャ認識を行っている。

1 回しか行われていないジェスチャを誤認識により 2 回分として認識しないようにジェスチャ認識の終了時点から 20 フレーム分 (約 0.67 秒) をインターバルとして設定した。

システムの構成と条件は以下の通りである。

- スマートフォン：XPERIA Z4
 - OS：Android 5.0.2
 - RAM：3GB
 - ROM：32GB
 - CPU：Snapdragon 810
- 収納場所：前ポケット
- プログラミング言語：Java
- センサ：角加速度，加速度，マイク
- ウィンドウ幅：25
- 教師データ：20 個ずつ

6. 使用実験

6.1 実験目的

実験協力者にプロトタイプシステムを使用してもらい、認識精度，ページめくりまでにかかる時間などの観点からシステムの有用性の検証を行う。

6.2 実験手順

使用実験では，19 歳～23 歳の実験協力者 8 人を対象に行った。まず，プロトタイプシステムに用いる学習データを両足 20 個ずつ取得するため，第 4 章で行ったデータセット構築時と同様に，75 回分ジェスチャを行ってもらった。その後，75 回分のジェスチャデータを取得したスマートフォンを，ポケットから取り出し，PC にそのデータを取り込み，プロトタイプシステムにこの学習データ 20 個分をセットし，起動させた状態でスマートフォンを被験者のポケットに戻した。そして，使用実験終了直後にアンケートに答えてもらうことを伝え，アンケート内容の説明を行った。その説明が終了した後，それぞれの足によるジェスチャでどのような操作が行われるかを説明し，実際にジェスチャを行ってもらうことで操作の確認をした。また，操作の確認とともにジェスチャ認識開始の閾値の設定に問題がないかを確認した。問題があると回答された場合には，閾値の調整を手動で行った。それらの確認が終了した後，足によるジェスチャ認識を用いて，タブレットを操作し，電子書籍を読んでもらった。また，読書が終了した直後にアンケートに回答してもらった。

6.3 アンケート内容

図 2 がアンケート用紙の実例である。アンケートでは，システムに対する意見と感想を調査するため，アンケート項目には以下の 4 つを選定した。

- 質問 1：ページめくりまでにかかる時間をどのように感じましたか
- 質問 2：認識精度についてどのように感じましたか
- 質問 3：今度もこのシステムを使いたいと感じましたか
- 質問 4：体感認識率

質問 1～3 に対しては，それぞれ 5 段階のリッカート尺度 (-2～2) により評価してもらった。そして，それぞれの

アンケートで「早い」「良い」「使いたいと感じた」と答えた場合に最大値，逆の答えの場合に最小値となるようにした。「体感認識率」の質問に対しては，使用実験を通してジェスチャの認識精度が体感で何%位と感じたかを記入してもらった。また自由記述欄を用意し，システムに対する意見や感想を記述してもらった。

6.4 実験環境

ジェスチャの認識には第 5 章で説明したプロトタイプシステムを用いた。また，電子書籍を読むためのデバイスは XPERIA Z4 Tablet を使用し，アプリケーションには Amazon Kindle を用いた。

本プロトタイプシステムを用いて，XPERIA Z4 Tablet を遠隔から操作するため，スマートフォンのテザリング機能を用いて操作対象と同じネットワークへの接続を行った。また，ADB(Android Debug Bridge)に存在する TCP/IP 通信機能と，タッチ操作コマンドを用いて，電子書籍のページめくりを実現するプロトシステムの実装を行った。

なお，一般的な電子書籍はページの左側をタップすることでページを進め，右側をタップすることでページを戻すため，左足によるジェスチャでページを進め，右足によるジェスチャでページを戻すという操作を全員の被験者に行ってもらった。

読んでもらった漫画については，約 1000 冊の漫画の中からジャンルを問わず，読みたい漫画を 1 巻選択してもらった。これは筆者が本を選んだ場合，ユーザ間でその本の好き嫌いが分かれることでアンケートに影響が表れる可能

使用実験用アンケート				
氏名 _____				
漫画を読み終えた後，以下の質問に対してご回答よろしくお願ひ致します。 1～3 はあてはまる項目に○を付けて下さい。				
1、ページめくりまでにかかる時間をどのように感じましたか				
早い	少し早い	ちょうどよい	少し遅い	遅い
2、認識精度についてどのように感じましたか				
良い	少し良い	どちらとも言えない	少し悪い	悪い
3、今後もこのシステムを使用したいと感じましたか				
感じた	やや感じた	どちらとも言えない	やや感じない	感じない
4、体感認識率 _____%				
自由記述欄				

図 2 アンケート用紙

表3 アンケート結果

実験協力者	質問 1	質問 2	質問 3	質問 4
P 1	- 1.00	1.00	1.00	85.0%
P 2	- 1.00	-1.00	1.00	80.0%
P 3	- 2.00	-1.00	-1.00	75.0%
P 4	0.00	-1.00	1.00	70.0%
P 5	- 1.00	1.00	1.00	90.0%
P 6	- 1.00	1.00	0.00	90.0%
P 7	- 1.00	1.00	1.00	70.0%
P 8	- 1.00	-2.00	-2.00	25.0%
全体	- 1.00	-0.125	0.250	73.1%

性を考慮し、その影響を取り除くためである。

6.5 実験結果

表3が実験協力者8名(P 1~P 8)によるアンケート結果である。質問1においては、ページめくりまでの時間が少し遅いと感じている傾向にあり、質問2では認識精度についてはどちらとも言えないという傾向にある。質問3の今後でも使用したいかどうかについてもどちらとも言えないという傾向にある。質問4の体感認識率の結果としては、73.1%であった。実験協力者P1~P7の実験協力者は70%~90%の範囲で回答しているが、P8のみ25%と極端に低い回答をしていた。被験者の自由記述には、「フィードバックが欲しかった」、「連続入力出来ないようになっていてやりにくかった」、「人によって読むペースが違うと思うので、それを踏まえたフィードバック時間の設定が出来ると良い」などといった回答がみられた。

7. 考察

本研究の目的は、片足のポケットに入れたスマートフォンでセンシングした加速度、角加速度、音声を用いて、両足でのジェスチャ認識が可能かを検証することであった。その結果としてはウィンドウ幅25の時の角加速度3軸(最大値・最小値)、加速度3軸(最小値)の9次元による特徴ベクトルで作成した分類器による認識精度が最も高い結果となった。また、音圧を加えることで全体的に精度は減少する傾向にあった。しかし、8人中2人は音圧の最大値を加えることにより、精度が約1%向上する傾向にあった。このように精度が上がった事例を参考にすることで、音圧を加えることにより認識精度が上がる可能性がある。そのため、今後の研究において、精度が上がった条件などを分析する予定である。

また、本システムはリアルタイムでの利用を想定しているため、ジェスチャ認識の判定に使用する際の最適なフレーム数の検証を行った。その結果として、第4章の結果からウィンドウ幅25の時の認識精度が最も高い傾向にあった。しかし、その検証により得られた結果を基に実装した

プロトタイプシステムを用いた使用実験からページめくりまでにかかる時間は遅いと感じている傾向にあった。この要因としては、ウィンドウ幅が適切ではなく、長かったためと考えられる。今後は、ページめくりまでにかかる時間はさらに短くするため、ウィンドウ幅を短くする必要があると考えられる。

使用実験の結果からは、認識精度については良くも悪くも無いと感じている傾向にあるが、第4章でのF値(0.9327)と比較すると、体感認識率は73.1%とかなり低い結果となった。また、極端に低く回答していたP8の被験者を除いた場合でも80.0%と期待した程高い結果とはならなかった。P8が極端に悪い精度に感じているのは、ジェスチャ自体の大きさがかなり小さかったため、認識精度自体がかなり落ちてしまったことが挙げられる。またP8の被験者の自由記述から、「人によって読むペースが違うのでそれを踏まえたフィードバック時間の設定が出来ると良い。」と偏見を与えてしまうことを考慮して伝えていなかったインターバルの存在に気づき、そのインターバルの長さが良くなかったと答えていた。また他の被験者の中にも「連続でジェスチャ入力出来ないように設定してやりにくかった」といった意見がみられた。そのため、インターバルの時間の設定が長すぎたため、第4章で算出されたF値よりP8の被験者では極端に低い体感認識精度の評価が行われ、全体としても認識精度が下がってしまったと考えられる。

今後ともこのシステムを使用したいかどうかについては、実験協力者の8人中5人が今後でも使用したいとやや感じていることから、提案手法の有用性がある程度示されたと考えられる。

8. まとめと今後の展望

本研究では、片足のポケットに入れたスマートフォンでセンシングした加速度、角加速度、音声を用い、両足でのジェスチャ認識が可能かを検証した。また、本システムはリアルタイムでの利用を想定し、ジェスチャ認識の判定に使用する際の最適なフレーム数の検証を行った。さらに、

その検証により得られた結果を基にプロトタイプシステムの実装を行い、そのプロトタイプシステムを用いた使用実験から本提案手法の有用性を検証した。

その結果をもとに認識精度の評価実験では、加速度の3軸と角加速度の3軸と音圧と最大値・最小値を組み合わせた31パターンから本提案システムに適する特徴ベクトルの検証を行った。その結果、 F_{all} の値が最も大きかった組み合わせはウィンドウ幅25の時の角加速度3軸(最大値・最小値)、加速度3軸(最小値)の9次元の特徴ベクトルにより作成した分類器であり、F値は0.9327であった。

使用実験の結果をみると、現在のプロトタイプシステムによるページめくりまでにかかる時間は遅いと感じられる傾向にあった。この要因としては、ウィンドウ幅が長かったためと考えられる。そのため、今回用いたウィンドウ幅を25からさらに短くすることで、ページめくりまでにかかる時間の改善を行う予定である。

現在のプロトタイプシステムによる体感認識精度は精度評価実験の値に比べて73.1%と低い値となった。この要因としては、システムの制約などにより、第4章での認識精度の評価実験よりも悪くなってしまったことが考えられる。

今後の課題としては、ジェスチャの種類を増やしていくのと同時に、どのくらいの種類のジェスチャを識別可能であるかを検証する。なお、この検証を通して増やしたジェスチャの認識が可能であった場合には、ジェスチャの種類を増加させ、使用実験を行う予定である。

また本研究では、足場が安定している場所では精度の算出を行えていないため、足場の不安定な場所においても認識精度の調査を行う。この調査を通して認識精度の悪化がみられた場合には、そのような状況でも足の識別を実現できる方法について検討を行っていく予定である。

本研究における使用実験では、最初に学習データを取得するために75回ジェスチャを行ってもらい、そのデータを用いて使用実験を行った。これはデータセット構築時に150回のジェスチャを行ってもらった為、使用実験ではその半分の75回ジェスチャを行ってもらい使用実験における学習データとして使用した。しかし、一度学習データをデータセットに登録することで、次にシステムを使用するときには、その学習データを用いることが出来る可能性がある。また、今回の様にジェスチャの種類が少ない場合には、他人の学習データを利用しても高い精度が出ることも考えられる。そのため、使用したいジェスチャを5回未満のジェスチャ行ってもらい、自分のジェスチャに近い他の人の学習データを用いることでジェスチャ認識可能なシステムの実装についても検討する。

謝辞

本研究の一部はJST CREST, JST ACCEL, 明治大学重点研

究Aの支援を受けたものである。

参考文献

- [1] 総務省.
<http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h27/html/nc372110.html>
- [2] Scott, J., Dearman, D., Yatani, K., Truong, N.. Sensing foot gestures from the pocket. In Proc. ACM UIST '10 (2010), 2010 p.199-208.
- [3] 村尾和哉, Laerhoven, K., 寺田務, 西尾章治郎. センサのピーク値を用いた状況認識手法とその評価. 除法処理学会研究報告, 2010 vol. 51, no. 3, p.1068-1077.
- [4] 村尾和哉, 寺田務. 加速度センサの定常性による動作認識手法. 情報処理学会論文誌, 2011 vol. 52, no.6, p.1968-1979.
- [5] 大内一成, 土井美和子. Activity Analyzer: 携帯電話搭載センサによるリアルタイム生活行動認識システム. 情報処理学会研究報告, 2011 vol. 2011-UBI-30 no. 3, p.1-8.
- [6] 河内智志, 藤波香織. 携帯電話の格納場所の判定 ~鞆への格納状態の認識~. 第74回全国大会講演論文集, 2012 p.437-438.
- [7] 河内智志, 薛媛, 藤波香織. 携帯端末の身体上格納場所判定機能のスマートフォンへの実装. インタラクシオン 2011, 2011 p.531-534.
- [8] 奥村典明. モバイル環境における足入力インタフェース, 筑波大学大学院博士課程 システム情報工学研究科修士論文.
- [9] 吉田成朗, 鳴海拓志, 橋本直, 谷川智洋, 稲見昌彦, 五十嵐健夫, 廣瀬通孝. ジェスチャ操作型飛行ロボットによる身体性の拡張. 情報処理学会シンポジウム論文集, 2012 IEXB-45
- [10] 田中正次, 村松茂, 山下茂. 9段数7次陽的Runge-Kutta法の最適化について. 情報処理学会論文誌. 1992, vol. 33, no. 12, p. 1512-1526.
- [11] Chan, L., Liang, R., Tsai, M., Cheng, K., Su, C., Mike Y., Chen, Cheng W., Chen B.. FingerPad: Private and Subtle Interaction Using Fingertips. In Proc. ACM UIST '13 (2013), 2013, p.255-260.
- [12] 加藤直樹, 大美賀かおり, 中川正樹. 携帯型ペン入力情報機器におけるペンジェスチャ入力指示インタフェース, 情報処理学会論文誌. 2000, vol. 41, no. 9, p.2413-2422.
- [13] 石原進, 太田雅敏, 行方エリキ, 水野忠則. 端末自体の動きを用いた携帯端末向け個人認証. 情報処理学会論文誌, 2005 vol. 46, no. 12, p.2997-3007.
- [14] libsvm: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>