|  |
| --- |
| 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 |
|  |
| 田村柾優紀 |
|  |
| **概要**：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が，画面を見ながらの操作は，手がふさがっている場合や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のセンシングを行い，両足でのジェスチャ認識による操作を可能にするプロトタイプシステムの実装を行った．その結果，認識手法にSVMを用いることである程度高い精度で両足のかかとの上げ下げのジェスチャを認識することが出来ることを明らかとした．  |
|  |
| **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，DPMatching |
|  |

# はじめに

## suma-sumスマートフォンの普及

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%である[1]．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，多くの人に浸透していると言えるだろう．

## suma-sumスマートフォン操作の問題点suma-sum

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や，手が汚れていてスマートフォンに触りたくない状況や，赤子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に問題がある．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では利用できないという問題が存在する．

また先述した問題の改善案として，Scottらの研究[2]では，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し，ポケットにスマートフォンを入れた足でのジェスチャ入力によるデバイス操作の提案と検証を行っている．その結果，スマートフォンを入れるポケットの位置によって高いジェスチャ認識精度を出すことに成功している．このことから，ポケットに入れたスマートフォンにより足の動きをセンシングすることで，片足によるジェスチャが認識可能であることと考えられる．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としての実装や，足によるデバイス操作としての有用性についての調査は行われていない．またこの研究では，片足のみのジェスチャ認識であったが，仮に両足によるジェスチャ入力が可能となれば，左右の特性を活かした操作の実現や，表現できるジェスチャ数の増加が考えられる．そのため，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磁気センサといった多様なセンサを用いて，多角的にジェスチャ認識を行う事で，両足でのジェスチャ認識の可能性として検討を行う．

## suma-sum本研究の目的

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，音声を用いて，両足でのジェスチャ認識を可能とすることが出来るかを検証する．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．

本論文の構成を以下に述べる．第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．第3章ではプロトタイプシステムに関する説明などの提案手法について述べ，第4章では認識精度の検証を行う．第5章では全体に対する考察を行い，第6章ではまとめと今後の課題と展望について述べる．

# 関連研究

## 行動推定に関する研究

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DPMatching，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し，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．米田らの研究[8]では，気圧センサを用いて，屋外状態では軌道上移動の気圧推移比較を用いた現在地推定を，屋内状態では気圧値の観測による階層推定手法を提案している．階層推定はかなり高い精度で認識可能であった．また，現在地推定については75％の確率において推定位置とGPS測位による位置との距離が0.5km以内で推定可能な手法の提案を行った．

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リアルタイムに用いてジェスチャを認識し，デバイス操作する点で異なっている．

## ジェスチャ認識に関する研究

　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．

### システム用のデバイスを用いたジェスチャ認識に関する研究

奥村らの研究[9]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を行えるシステムの実装を行っている． 吉田らの研究[10]では，3次元空間で，小型の飛行ロボットを用いて人間の身体性を拡張するための提案をしている．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できるようなジェスチャの提案も行っている．冨林らの研究[11]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[12]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．Humantenna[13]は，実世界の自分の全身をアンテナとして用いてリアルタイムなインタラクションを行い，自分の全身を使ったジェスチャ認識を実現するシステムを提案している．Hanらの研究[14]では，モバイルインタラクションにおける効果的な足ジェスチャとしてキックジェスチャをXbox Kinect Cameraによりジェスチャの識別を可能とする手法を実現した．吉村[15]らの研究では，Xbox Kinectに存在する”Kinect-Kamehameha”というゲームにおいて実装されているジェスチャの認識率を，SVMを用いることでジェスチャの認識率改善を行う研究を行い，そのジェスチャの認識率の改善に成功した．Cyclops[16]では，魚眼レンズを自分の体の中央に取り付け，その魚眼レンズで得られた情報によりジェスチャ認識可能なシステムである．MimiSense[17]は口を開ける，口を閉じる，及び顎を左に動かすなどの下顎運動によりジェスチャを行い，デバイスを操作可能とするシステムを提案している．Ubi-Finger[18]は，手指のジェスチャを用いて情報機器や情報家電機器の操作を実現する指装着型のウェアラブルデバイスを提案し，実際にUbi-Fingerを用いて実世界機器を操作する実験を通して，システムに対する評価実験を行っている．山本らの研究[19]では，ジョギング時における入力インタフェースとして，足ステップによるメニュー制御システムの提案と実装を行っており，使用実験から本提案システムの有用性を検証した．澤田らの研究[20]では，手の甲に取り付けた3次元加速度センサから運動量として，加速度の変化，回転力，回転方向の分布などを求め，標準パターンとのマッチングを行う事により，10種類程度のジェスチャを100％近い識別率で認識出来ることが分かった．

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ることにより，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．

### 既存のデバイスを用いたジェスチャ認識に関する研究

加藤らの研究[21]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，ディ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，ディ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た．石原らの研究[22]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DPMatching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を行った場合，本人拒否率を10%以下にできることを確かめた．呉らの研究[23]では肌色領域，髪色領域の一次二次モーメントに基づいて入力画像にある頭部の3次元姿勢を推定し，連続DPMatchingにより，頭部ジェスチャのスポッティング認識を実現し，良好な認識結果が得られた．西村らの研究[24]では動作者にデータグローブ等の接触型センサやマーカを装着させることなく，人間の身振り手振りをカメラにより取得した動画像を基にジェスチャ認識をおこなう手法に着目し，モデルの自動切り出し法と認識判定法を提案し，オンライン教示可能なシステムを作成することにより動作者に適応可能とした．見上らの研究[25]では，タッチパッドや加速度センサなどを備えた携帯端末（スマートフォンなど）の普及に伴って，携帯端末における認証方式も様々な物が提供されている．その認証手法は仕様環境やデバイスの特性を考慮しなければならない問題に着目し，その解決策としてタッチパネルと加速度センサを用いたジェスチャ認識とその入力方式の提案と考察を行っている．玉城らの研究[26]ではカメラ動画による3次元ハンドジェスチャ認識には前腕回旋動作時に認識精度が低下するという問題について，手指の輪郭特徴量情報に爪の位置情報を追加する事で解決した．

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，手や目を使うことが出来ない際には利用することが出来ない．そこで本研究では，ズボンのポケットにスマートフォンを入れている状態で，目と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う点で異なる．

# 提案手法

## 事前調査

　どのようなジェスチャであれば，両足によるジェスチャが認識させることが可能となるかを調査するため事前調査を行った．

J.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，その4種類の中から，体の重心が一番変動しやすいと，筆者が判断した踵の上下運動を，実験協力者8人にジェスチャを行ってもらった．そしてこの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い傾向にあった．また踵の上下運動の種類には，力を入れる位置により，重心が動くような種類と，重心が全く動かないような種類の2種類が存在した．歩くように踵を上下すると重心が動きやすくなることから，このジェスチャを行う際に歩くように踵の上下運動をするように指示する事で，踵の上下運動のジェスチャを統一することが可能と考えられる．

また，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，他の音声が90Hz以上となる傾向にあった．

## 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ることで，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を調査する．右足と左足によるジェスチャであるため，ポケット内のスマートフォンとズボンとの衣擦れの仕方が異なるため，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，マイクのセンサが両足によるジェスチャ認識において有用であると考えられる．そのため，マイクのセンサを用いることによって両足のジェスチャ認識における認識精度への影響を検証する．そして，ウィンドウ幅の変化による認識精度の変化の分析を行い，本研究において適切なウィンドウ幅の決定を行う．

プロトタイプシステムの実装においては，モバイル端末におけるジェスチャ認識の研究[19]で利用されたDPMatchingと，機械学習の1種であるSVM（サポートベクターマシーン）などのジェスチャ認識手法の中から本研究に適した手法を認識精度の比較実験により明らかにする．そして，実験結果から決定した手法を用いたリアルタイムな両足によるジェスチャ認識を可能にするシステムの実現を目指す．そのため，実験協力者がジェスチャを行ったときのセンサデータを用いてデータセット構築を行う．その後，認識手法の比較実験によりどのセンサが両足でのジェスチャ認識において重要であるかについて認識精度の評価実験により検証する．

# 認識精度の評価実験

## 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，まず実験協力者に両足によるジェスチャを行ってもらい，データセットを構築した．次に，そのデータセットを用いて，関連研究において用いられていたSVMとDPMatchingの二つの認識手法によるジェスチャ認識の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が識別可能かを調査する．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．また，この調査を通して本研究に適している認識手法を決定する．

## データセット構築

データセットの構築では，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たうえで，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するため，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おり，（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と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）の画面の時には待機してもらい，一定の待機時間が経過した後に自動で（イ）の画面に切りかわり，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）の画面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）の画面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）の画面に切り変わってすぐにジェスチャが出来るように，待機画面（ア）では何秒



1. 待機画面



1. タスク提示画面

図1　実験におけるタスク提示方法

後にどちらの足でジェスチャを行うかを提示した．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子を抱いている状態を想定し，赤子と同程度の重さの物を両手に抱え得るように持つことで，この状態を仮想的に再現している．ま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，一つの学習データとして用いており，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学習データとして使用している．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ため，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．

分類するクラスは，左足による踵の上下運動のジェスチャと，右足による踵の上下運動のジェスチャ，待機状態の3種類である．また他の学習データと同じように，10問目の課題を待機している状態のデータから20問分を待機状態の学習データとして使用している．

その他の実験詳細は以下の通りである．

* 実験協力者：19～23歳の8人
* 収納場所：右足の前ポケット
* ジェスチャ種類：踵の上下運動
* 出題方法：タブレット
* 出題数：ランダムに150問
* 課題提示時間：1秒間
* 待機間隔：5秒間
* タブレット端末：XPERIA Z4 Tablet
	+ OS：Android 5.0.2
	+ RAM：3GB
	+ ROM：32GB
	+ CPU：Snapdragon 810

## 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値を求めることにより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種類のジェスチャの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．

本研究においては，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，データセット構築内の各ジェスチャ40個ずつ登録されているジェスチャデータの中に分割し，最初の試行の20個分のデータで残りの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．

## SVMを用いた認識精度の検証実験

### 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，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．ここで音声データのそしてその特徴ベクトルは以下の通りである．

* 1. 角加速度3軸（最大値）3次元
	2. 角加速度3軸（最小値）3次元
	3. 角加速度3軸（最小値最大値）6次元
	4. 加速度3軸（最大値）3次元
	5. 加速度3軸（最小値）3次元
	6. 加速度3軸（最大値・最小値）6次元
	7. 角加速度3軸（最大値），加速度3軸（最大値）
	6次元

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 |

* 1. 角加速度3軸（最大値），加速度3軸（最小値）
	6次元
	2. 角加速度3軸（最大値），加速度3軸（最大値・最小値）　9次元
	3. 角加速度3軸（最小値），加速度3軸（最大値）
	6次元
	4. 角加速度3軸（最小値），加速度3軸（最小値）
	6次元
	5. 角加速度3軸（最小値），加速度3軸（最大値・最小値）9次元
	6. 角加速度3軸（最大値・最小値），加速度3軸（最大値）9次元
	7. 角加速度3軸（最大値・最小値），加速度3軸（最小値）9次元
	8. 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
	9. 音（最大値）1次元
	10. 角加速度3軸（最大値），音圧（最大値）4次元
	11. 角加速度3軸（最小値），音圧（最大値）4次元
	12. 角加速度3軸（最小値最大値），音圧（最大値）7次元
	13. 加速度3軸（最大値），音圧（最大値）4次元

表2　DPMatchingによるF値

|  |  |
| --- | --- |
|  | **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 |

* 1. 加速度3軸（最小値），音圧（最大値）4次元
	2. 加速度3軸（最大値・最小値），音圧（最大値）7次元
	3. 角加速度3軸（最大値），加速度3軸（最大値），音圧（最大値）7次元
	4. 角加速度3軸（最大値），加速度3軸（最小値），音圧（最大値）7次元
	5. 角加速度3軸（最大値），加速度3軸（最大値・最小値），音圧（最大値）　10次元
	6. 角加速度3軸（最小値），加速度3軸（最大値），音圧（最大値）7次元
	7. 角加速度3軸（最小値），加速度3軸（最小値）
	音圧（最大値），7次元
	8. 角加速度3軸（最小値），加速度3軸（最大値・最小値），音圧（最大値）10次元
	9. 角加速度3軸（最大値・最小値），加速度3軸（最大値），音圧（最大値）10次元
	10. 角加速度3軸（最大値・最小値），加速度3軸（最小値），音圧（最大値）10次元
	11. 角加速度3軸（最大値・最小値），加速度3軸（　最大値・最小値），音圧（最大値）13次元

### 実験結果

### ウィンドウ幅による認識精度の違い

本プロトタイプシステムにおいて，適切なウィンドウ幅を決定するため，ウィンドウ幅が15，20，25，30，35，40フレームの6パターンをA，B，Cの特徴ベクトルにより作成した分類器の中で$F\_{all}$の値が一番高いウィンドウ幅を検証した．その結果，ウィンドウ幅25でのF値が良い傾向にあった．また，ウィンドウ幅を25から減少させると認識精度が減少する傾向にあった．

## DPMatchingを用いた認識精度の検証実験

### DPMatchingによる認識

　DPMatching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，センサ値の最大の絶対値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．この後，処理を施したデータに対してDPMatching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．

## 認識精度の検証実験の考察

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Aの特徴ベクトルにより作成した分類器を用いて実装を行う．本研究においては音圧を用いたことで，全体としては精度が下がる傾向になったが，2人の実験協力者はマイク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．

# プロトタイプシステム

4章の結果に基づき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，一定時間ごとに認識させる方法などが考えられるが，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めている．

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．そのため，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，地磁気）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．角加速度センサはXYZ軸のそれぞれを中心とした際の回転速度を取得するセンサである．以降X軸を中心とした際の回転速度を角加速度X，同じようにY軸では角加速度Y，Z軸では角加速度Zと表記する．

前節での事前調査から，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デシベル値の抽出を行い，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．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が，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．

　認識までの過程では，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10回分のセンサデータと，直後20回分のセンサデータの計30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．

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，地磁気
* ウィンドウ幅：30回分
* 教師データ：20個ずつ

# 考察

　第4章における精度算出の実験ではある程度高い精度が得られた．しかし，今回の実験ではプロトタイプシステムを用いての使用実験を行う事が出来ていない．そのため，使用実験を行う事により，本研究におけるシステムの認識精度として十分であるかを検証する必要があると考えられ

る．

　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を可能とすることが出来るかを検証する事であった．

しかし，本研究においては音圧を用いたことで，全体としては精度が下がる傾向になったが，2人の実験協力者はマイク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．そのため，この様に精度が上がった事例を参考



図2　スマートフォンのXYZ軸

にしてジェスチャを行うことで，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検証する予定である．

# まとめと今後の展望

本研究では，事前調査により認識させるジェスチャを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に絞り，ジェスチャがどちらかの足によるものかが，どの程度の精度で識別可能かを調査した．

その結果，認識精度の評価実験では，踵の上下運動がどちらの足により行われたかをある程度高い精度で認識させることに成功した．しかし，実際にユーザに使用してもらっての評価実験を行う事が出来ていないため，今回の認識精度が実際にどの程度本提案システムの精度として十分な精度かどうかが検証されていない．そのため，今後使用実験を行い，今回の認識精度が提案システムの精度として十分な精度であるかを検証する予定である．

プロトタイプシステムの実装においては，ジェスチャ認識手法の中から本研究に適した手法を認識精度の実験により明らかにし，その実験結果から決定した手法を用いてリアルタイムな両足によるジェスチャ認識を可能とするシステムの実現を目指した．そのため，実験協力者がジェスチャを行ったときのセンサデータを取得し，データセット構築を行い，認識手法の比較実験を行い，どのセンサが両足でのジェスチャ認識において重要であるかについて検証を行うため，加速度の3軸と角加速度の3軸と分散値・平均知を組み合わせた9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，一番$F\_{all}$の値が一番大きかった組み合わせは角加速度3軸，加速度3軸の平均値と分散値（12次元）という特徴ベクトルにより作成した分類器であった．

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，使用実験により意見感想を得る予定である．

本研究においては，足場が安定している場所でしか精度の算出を行えていないため，足場の不安定な場所など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対策について検討を行っていく予定である．

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目指す予定である．

現在は静止状態での環境においてジェスチャ認識を行っているため，足場の不安定な場所でのジェスチャ認識が出来ていない制約が存在する．そのため，今後の研究においてそのような足場が不安定な場所でのジェスチャ認識についても行う予定である．

本研究における，精度評価の実験においてはまず，加速度と角加速度センサを特徴ベクトルに用いることで，両足のジェスチャ認識の精度を検証したが，足を動かすことでズボンとスマートフォンの衣擦れ音が，スマートフォンをポケットに入れていない方の足でジェスチャを行った際にも起こると予想しているため，今後は音についてもジェスチャ認識に用いる予定である．

　**謝辞**本論文を執筆するにあたり，指導教員である中村聡史准教授，また研究室メンバーは研究についてのアドバイスやお手伝いや相談に乗ってくださり，学業面以外にも，精神面でも助けて下さいました．そういったこともあり，心より感謝申し上げます．

**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.. 寺田努，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　～鞄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. インタラクション2011, 2011 p.531-534
8. 米田圭祐, 望月祐洋, 西尾信彦. 気圧センシングを用いた行動認識手法. 研究報告モバイルコンピューティングとユビキタス通信, 2014-MBL-70，2014 Vol. 14，P. 1-8
9. 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
10. 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 1EXB-45
11. 富林豊, 竹川佳成, 寺田努, 塚本昌彦. 装着型無線加速度センサを用いたウェアラブルDJシステム. インタラクてぃぶシステムとソフトウェアXVI日本ソフトウェア科学会　WISS2008, 2008 p.51-56.
12. 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
13. Cohn, G., Morris, D., Patel, S., Tan, D.. Humantenna：Using the Body as anAntena for Real-Time Whole-BodyInteration. CHI 2012，2012，p.1901-1910
14. Han, T., Alexander, J., Karnik, A., Irani, P., Subramanian, S.. Kick: investigating the use of kick gestures for mobile interactions. In Proc．ACM UIST ’11 (2011), 2011, p.29–32
15. 吉村拓哉, 橋本剛. 機械学習を用いたジェスチャ認識精度向上方法の研究. ゲームプログラミングワークショップ2012論文集, 2012 p.167-170
16. Chan, L., Hsieh C., Chen Y., Yang, S., Huang D., Liang, R.， Chen B.. Cyclops: Wearable and Single-Piece Full-Body Gesture Input Devices. In Proc．[CHI '15](http://chi2015.acm.org/) Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems，2015, p.3001-3009
17. 安藤宗孝, 久保勇貴, 志築文太郎, 高橋伸. Mimi Sense:：外耳道内の気圧変化を利用した下顎運動認識認識システム. WISS 2016 2016, p. 1-6.
18. 塚田浩二, 安村通晃. Ubi-Finger：モバイル指向ジェスチャの研究. 情報処理学会論文誌, 2002, vol. 43, no. 12, p.3675-3684
19. 山本哲也, 寺田努, 塚本昌彦, 義久智樹. ジョギング時における情報機器利用のための足ステップ入力方式. 情報処理学会論文誌, 2009 vol. 50, no.12, p2881-2888
20. 澤田秀之，橋本周司. 加速度センサを用いたジェスチャ認識と音楽制御への応用. 電子情報通信学会論文誌，基礎・境界 J79-A(2)，1996，p.452-459
21. 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
22. 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
23. 呉海元, 小林弘知, 陳謙, 塩山忠義, 島田哲夫. 色彩動画像からの頭部ジェスチャ認識システム. 情報処理学会論文誌，vol37，no.6，p.1234-1242
24. 西村拓一, 向井理朗, 野崎俊輔, 岡隆一. 動作者適応のためのンライン教示可能なジェスチャ動画像のスポッティング認識システム. 電気情報通信学会論文誌, 1998 II-情報処理 J81-D-2(8)，p.1820-1830
25. 見上一憲, 林原尚浩. タッチパネルと加速度センサを用いた携帯端末向けジェスチャ認証とその入力方式の提案. 情報処理学会研究報告, CSEC-56，2012，vol. 2012-DPS-150 no. 8，p. 1-7
26. 玉城絵美, 味八木祟, 暦本純一. インタラクションシステムのための高次元な3次元ハンドジェスチャ認識手法. 情報処理学会論文誌, 2010 vol. 51, no. 2, p.229-239
27. libsvm：https://www.csie.ntu.edu.tw/~cjlin/libsvm/

**著者紹介**

|  |
| --- |
| **田村柾優紀**明治大学　総合数理学部　先端メディアサイエンス学科　所属 |