

## 推薦研究論文

## 携帯端末の利用環境に依存しない端末把持姿勢認識手法

朴 燦鎬<sup>1,a)</sup> 小川 剛史<sup>2,b)</sup>

受付日 2015年7月15日, 採録日 2015年11月30日

**概要:** ユーザの状況により, 携帯端末の様々な把持姿勢が存在する. 携帯端末に適切なインタフェースを提示するためには, その把持姿勢を認識することが重要である. これまで把持姿勢を認識する研究がさかんに行われているが, 新たなセンサを追加しなければ適用することができない手法が多い. また, 携帯端末の内蔵センサの入力情報を用いて把持姿勢を認識する研究も行われているが, 把持姿勢を即座に認識することが困難であり, 部屋にいる場合や乗り物に乗っている場合など, 携帯端末の利用環境の違いが考慮されていない. 本研究では携帯端末の利用環境に依存しない端末の把持姿勢認識を実現するため, 端末の内蔵センサであるジャイロセンサおよび加速度センサのデータ, タッチスクリーンへの入力情報を用いた把持姿勢認識手法に関する検討を行った. 本論文では, 機械学習の1つであるサポートベクターマシン(SVM)による把持姿勢認識を行い, 室内と電車内で実施した実験結果について述べる.

**キーワード:** 把持姿勢認識, タッチスクリーン, ジャイロセンサ, 加速度センサ, サポートベクターマシン

## A Grasp Recognition Method of Smartphone Independent of Users' Situations

CHANHO PARK<sup>1,a)</sup> TAKEFUMI OGAWA<sup>2,b)</sup>

Received: July 15, 2015, Accepted: November 30, 2015

**Abstract:** There are various hand postures of smartphones according to the users' situations. In order to provide appropriate interface, it is important to recognize hand postures when users are using the smartphones. Although there are many studies on estimation of grasp posture, most of studies require additional instrumentation of a smartphone. Others have also proposed techniques for recognizing grasp posture using the built-in sensor of a smartphone. However, these studies are difficult to recognize the grasp posture immediately, and do not consider the users' situations. To realize grasp recognition independent of users' situations, we consider using the sensor data built-in gyroscope and accelerometer, and touchscreen information. In this paper, we use support vector machine (SVM), which is one of the machine learning system, to classify grasp postures. And we also describe the result of the experiments when users are using the devices in the room and on the train.

**Keywords:** grasp recognition, touchscreen, gyroscope, accelerometer, support vector machine

## 1. はじめに

スマートフォンやタブレットなどの携帯端末の普及にと  
もない, 電話やメッセージ機能だけでなく, ウェブ検索や  
ゲームなど様々な作業が場所に依存せずできるようになっ

た. 端末の操作は, 片手で行うこともあれば, 両手で行う  
こともあり, その把持姿勢はユーザがおかれている状況に  
よって異なる. たとえば, 一方の手に荷物を持っている状  
態や, 電車内でつり革を掴んだ状態で携帯端末を操作する  
場合には空いた手でしか操作ができないが, ベンチや電車  
の座席に座って操作する場合には片手でも, 両手でも操  
作できる (図 1). 操作できる手が異なれば, 入力インタ  
フェースもそれに応じて変更することが有効であると考え  
られ, これまで携帯端末の把持姿勢を認識する手法に関す  
る研究が行われてきた.

従来の把持姿勢認識に関する研究 [1], [9], [10] では, ユー

<sup>1</sup> 東京大学大学院工学系研究科  
Graduate school of Engineering, The University of Tokyo,  
Bunkyo, Tokyo 113-0033, Japan

<sup>2</sup> 東京大学情報基盤センター  
Information Technology Center, The University of Tokyo,  
Bunkyo, Tokyo 113-8658, Japan

a) c.park@cml.t.u-tokyo.ac.jp

b) ogawa@nc.u-tokyo.ac.jp

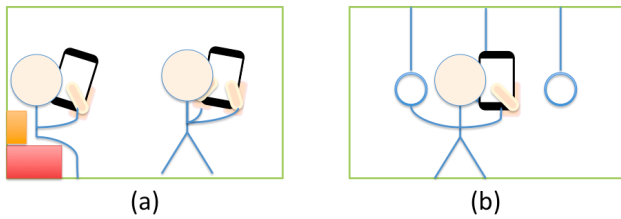


図1 ユーザがおかれている状況による様々な把持姿勢パターン：(a) 片手操作と両手操作；(b) 片手操作のみ

Fig. 1 Grasp postures according to the users' situations: (a) One and two-handed interaction; (b) One-handed interaction.

ザが端末のどの部分を触れているのかを直接検出するために、静電容量センサを用いており、一般に普及している端末に手法を適用することは容易ではない。端末の内蔵センサのみを利用した研究 [3] も存在するが、認識できる把持姿勢の種類が少なかったり、把持姿勢が変化することを即座には認識できないといった問題がある。さらに、日常生活において、我々は室内だけでなく屋外や移動中の乗り物内など、いつでもどこでも携帯端末を利用しているが、従来研究では屋内と乗り物内といったように異なる場所での利用については検討されていない。

そこで本論文では、端末の内蔵センサのみを用いて、室内や乗り物内といった端末の利用環境が変化しても、端末の把持姿勢を認識できる手法を提案する。具体的には、内蔵センサとして加速度センサ、ジャイロセンサ、タッチスクリーンから得られるデータに基づき機械学習により把持姿勢を認識する。端末の利用環境としては、ふだんの利用を想定した室内と通勤通学時などでの利用を想定した電車内の2種類を想定している。

以下、2章では、把持姿勢認識の関連研究を紹介する。3章では提案する把持姿勢認識手法について述べ、4章では把持姿勢認識精度を検証するために行った実験について述べる。5章において被験者実験の結果について考察し、最後に6章にて、本論文のまとめについて述べる。

## 2. 関連研究

本章では、端末の把持姿勢認識に関する従来研究について述べ、提案手法の位置付けを明確にする。表1に従来手法と提案手法の比較を示す。

### 2.1 把持姿勢認識に関する研究

スマートフォンなどの携帯端末の把持姿勢認識に関する研究 [1], [2], [9], [10], [11], [12] がさかに行われている。Kimらは、携帯端末における代表的なアプリケーションである「電話」、「メッセージ」、「カメラ」、「動画」、「ゲーム」に対して8種類の把持姿勢を定義し、加速度と端末に触れた手の形状データから把持姿勢を認識する手法を提案している [9]。手の形状データの取得には、端末に追加した64

個の静電容量センサのデータを用いており、高精度な認識を実現している。Chengらは、静電容量センサのデータから認識した把持姿勢に基づいて携帯端末の表示を自動的に切り替えるインターフェースを提案している [1], [11]。加速度センサを用いてユーザが把持姿勢を変更したタイミングを検出し、画面表示を切り替えに利用している。これらの研究は、端末に標準では搭載されていないセンサの利用を前提としており、一般に普及している端末をそのまま利用することはできない。

静電容量センサ以外にも新たなセンサを端末に追加して把持姿勢を認識する研究が存在する。Onoらは、物体の持ち方や持つ強さによって物体表面における音の伝わり方が異なることに着目し、物体に取り付けた振動スピーカとピエゾマイクを用いて物体の把持姿勢を認識する手法を提案している [2]。7種類の把持姿勢と3種類の把持力を区別できることが報告されている。実験ではユーザが座った場合のみの認識精度が示されているが、立って把持するなど利用状況が変わると端末へ加わる荷重分布が変化して音の伝わり方が変化するため、認識精度が低下する可能性があることや、静電容量センサを用いた研究と同様に、新たなセンサが必要であることが課題となっている。

本研究では、携帯端末の把持姿勢を認識するために、新たなセンサは使わず、内蔵センサのみを用いることとしている。また、把持した手の形状から把持姿勢を認識するのではなく、自然に携帯端末のスクリーンをタップしたときのタップ情報と端末の動きによるジャイロセンサと加速度センサの値に関する情報を用いて把持姿勢を認識する。

### 2.2 内蔵センサを用いた把持姿勢認識に関する研究

携帯端末の内蔵されたセンサのみを利用して把持姿勢を認識する手法も提案されている。Goelらは、タッチスクリーン上でのタップ操作とスワイプ操作から把持姿勢を認識するシステムを構築している [3]。ジャイロセンサから得られた端末の傾きデータ、タッチスクリーンに触れている指の面積とスライド操作における指先の軌跡を特徴量とし、「片手操作」2種類と「両手操作」1種類の操作方法を区別すべき把持姿勢としている。把持姿勢判定には、利用者が5回以上連続して端末を操作した際のセンサデータを用いるため、把持姿勢を変更した際に即座にその変更を認識することは困難であったり、両手操作の場合には持ち手と打ち手の認識ができないといった制限がある。

本研究においても、端末の内蔵センサのみを用いるという点に関してはこれらの研究と同様であるが、アプリケーション実行時や利用中などにタッチスクリーン上で1回の入力操作が行われれば、即座に把持姿勢を認識する点が異なる。また、認識する把持姿勢を、持ち手と打ち手、左右の認識もできるように図3に示す5通りとしていることや、実環境での利用を想定して、部屋と電車内で座った状態や立っ

表 1 従来手法と提案手法の比較

Table 1 Comparison of the conventional method and the proposed method.

把持姿勢認識手法		Kim et al [9]	Touch & Activate [2]	iRotateGrasp [1], [11]	GripSense [3]	提案手法
内蔵センサ	タッチスクリーン	×	×	×	○	○
	ジャイロセンサ	×	×	×	○	○
	加速度センサ	×	×	○	×	○
追加デバイス		静電容量センサ, 加速度センサ	スピーカ, マイク	静電容量センサ	×	×
識別器		NB, SVM	SVM	SVM	閾値	SVM
把持姿勢		8 種類	7 種類	3 種類	4 種類	5 種類
持ち手判定		×	○	○	○	○
打ち手判定		×	×	×	△	○
判定タイミング		把持する	把持する	把持する	タップ, スライド 操作 (5 回以上)	タップ操作 (1 回)
利用環境		室内 (座る)	室内 (座る)	室内 (座る, 立つ, 寝転がる)	室内 (座る)	室内 (座る, 立つ), 電車内

た状態での認識について検討している点でも異なっている。

### 3. 提案手法

本研究では、新たなセンサを追加せず、携帯端末の内蔵センサのみを用いて、利用環境における把持姿勢を判定する。システムの概要を図 2 に示す。

提案システムは、ユーザが携帯端末のスクリーンをタッチした際の内蔵センサの計測値を取得し、それらの計測データから特徴量を抽出する。特徴量には、端末の把持姿勢認識を実現している文献 [3] を参考にジャイロセンサの値とスクリーンにおけるタッチ情報から得られる特徴と、ユーザのコンテキスト推定 [4], [5], [6] などによく利用される端末の加速度センサの値から得られる特徴を採用する。識別器にはサポートベクターマシン (SVM, Support Vector Machine) を用い、抽出した特徴量を学習させる。

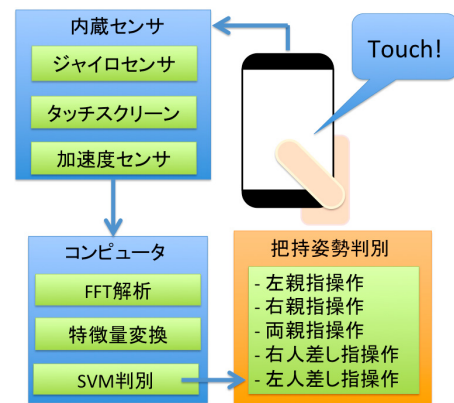


図 2 システムの概要

Fig. 2 System configuration.

#### 3.1 認識する把持姿勢

認識する端末の把持姿勢は、通常の操作でよく見られる 5 パターンとした。図 3 に認識する把持姿勢を示す。

片手操作は、片方の手で携帯端末を把持し、把持した手の親指で行う操作である。つまり、片手操作は左親指操作 (Left Thumb: L-Th) と右親指操作 (Right Thumb: R-Th) の 2 パターンとなる。両手操作は一方の手で携帯端末を把持しながら、他方の手の人差し指で操作する場合 (Right Index finger: R-In, Left Index finger: L-In) と両手で携帯端末を把持し、親指で端末を操作する場合 (Both Thumbs: B-Th) の 3 パターンとした。本論文では、この 5 パターンの把持姿勢について認識する。

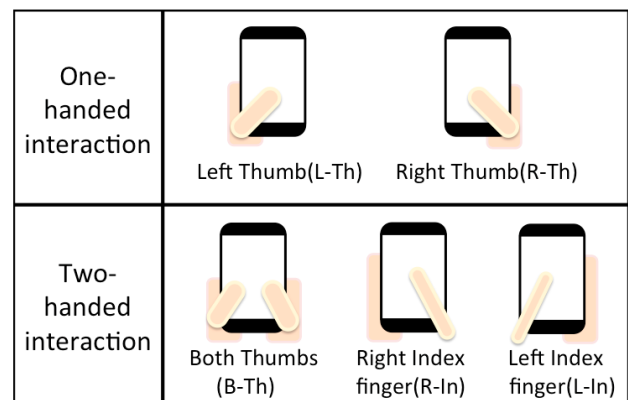


図 3 把持姿勢セット

Fig. 3 Grasp postures set.

#### 3.2 把持姿勢認識

##### 3.2.1 データの取得

スクリーンのタッチイベントをトリガとして各データを

取得する。タッチスクリーンに関する情報は、タッチ座標 (x 座標, y 座標) とタッチ面積 (pixels) を記録する。また、内蔵センサであるジャイロセンサと加速度センサについては、サンプリング周波数 50 Hz で値を計測し、タッチイベントの前後 0.2 秒間のデータを記録する。

表 2 特徴ベクトル  
Table 2 Feature vectors.

特徴ベクトル	ジャイロスコープ	タッチスクリーン	加速度	線形加速度	次元
A	○	○	×	×	69
B	○	○	○	×	130
C	○	○	×	○	130
D	○	○	○	○	191

加速度センサの計測値については、ハイパスフィルタ (HPF, High Pass Filter) を適用した値についても利用することとした。HPF を通過したデータは重力加速度成分が除去された端末自身の動きに関する加速度であると考えられ、本論文ではこの加速度を線形加速度と呼ぶ。

### 3.2.2 特徴量抽出

タッチスクリーンからのデータは、ユーザの把持姿勢によりデータが変化すると思われる。特に、スクリーンをタップする部分の座標によりデバイスの動きが異なったり、前回にタップした部分の座標が何かによって、次にタップする際のデバイスの動きが変化すると思われる。本研究では、ユーザがタッチした座標と前回タッチした座標、およびその座標差と距離、タッチ面積の 8 次元を特徴量として用いた。これは文献 [3] で用いられている特徴量を参考にしている。

内蔵センサ (ジャイロセンサおよび加速度センサ) のデータはタップ操作による端末の動き (回転) や、タップ力、手ぶれ、外部環境からの雑音に関する情報が含まれている。

ウェアラブルセンシングに基づくユーザの行動認識研究 [5], [6], [7], [13], [14], [15] では、サンプリングしたセンサ値だけでなく、平均・標準偏差・相関係数・最大値・最小値などが特徴量として用いられている。本研究では、タッチした瞬間に内蔵センサ (3 軸) から取得したデータと低周波数領域と高周波数領域におけるデータ (LPF と HPF を適用するため、タッチした瞬間と 20 ミリ秒後のセンサデータを用いて計算) が各 3 次元、タップ前後の連続した 11 個のデータ (タップした時刻を基準にして前後 5 フレームのデータ) を用いて平均、標準偏差、相関係数、RMS、軸間の相関係数、最大値、最小値が各 3 次元ずつ、3 軸合成値の平均が 1 次元で計 31 次元を特徴量とする。また、内蔵センサのデータに関しては、周波数成分分析により、周波数に関する特徴量を抽出する。

周波数成分分析は、行動認識や音声認識などの分野においてもよく利用されており [4], [5], [6], [7], 把持姿勢推定に関する研究 [2] でも利用されている。本研究では、タッチした瞬間と前後のデータの中で 8 個のデータ (FFT 変換のため、2 の  $n$  乗のデータ) を用いて FFT 変換を行い、周波数成分から抽出できる特徴量を用いた予備実験を行った結果、把持姿勢認識に大きく寄与すると考えられる周波

数成分の振幅に関するもの (振幅の最大値, 2 番目に大きい値, 2 番目に大きい値のインデックス, 軸間の相関係数が各 3 次元) を特徴量として用いた。周波数成分分析は低周波数領域と高周波数領域におけるデータに対しても適用し、振幅に関する情報 (振幅の最大値, 標準偏差, 相関係数が各 3 次元) を加えて計 30 次元とした。

以上、タッチスクリーンから 8 次元, 内蔵センサ 1 つに対して 61 次元の特徴量を用いることになる。表 2 のように使用するセンサに応じて、特徴ベクトル A から D を定義した。すべての特徴量を用いる場合 (特徴ベクトル D) は最大 191 次元となる。

### 3.2.3 SVM による識別

本研究では、把持姿勢推定のため、機械学習手法の 1 つである SVM (Support Vector Machines) を用いて、前節で述べた特徴量と把持状態の関係を学習し、判別を行った。線形 SVM の 1vs1 法を複数使い、多値分類を行った。実装に Python を用いて SVM ライブラリの 1 つである LIBSVM [8] を使用した。識別のため、取得した特徴量から LIBSVM 形式データに変換し、スケール調整を行っている。RBF カーネル関数を用い、最適なパラメータ  $C$  と  $\gamma$  を求めるために、交差検定からグリッド探索を行った。訓練データに対してモデルを生成し、最後にテストデータに適用した。SVM の判別により 5 パターンの把持姿勢の中から 1 つのパターンに分類する。

## 3.3 システム構成

携帯端末には Samsung Galaxy Note3 を用い、標準で内蔵されているジャイロセンサと加速度センサからの計測値を利用する。OS は Android OS バージョン 4.3, 重さは 168 グラム, 画面サイズは 5.7 インチ (解像度: 1,080 × 1,920 ピクセル) である。

取得したデータからの特徴量抽出と機械学習および把持姿勢認識は Apple 社の MacBook Air (CPU: Intel Core i5, RAM: 4 GB) を用いてオフライン処理にて実施した。

## 4. 評価実験

特徴量の違いによる把持姿勢の認識精度の違いを検証するため、表 2 に示す特徴ベクトルを用いて判別実験を行った。まず、実験 1 では予備実験として室内における座位状態での認識精度を、そして実験 2 では室内、電車内におけ

る座位、立位状態での認識精度を検証する。

#### 4.1 実験タスク

携帯端末のスクリーンをタップした際に、各センサのデータを記録するアプリケーションを用意した。アプリケーションの画面例を図4に示す。スクリーンを縦8個、横5個のセルに分割し、ランダムに選択したセルに円を表示する。ユーザが円をタップすると、その瞬間のセンサ値を記録し、0.5秒後に次の円をランダムに選択した場所に表示する。すべてのセルに円を表示するまで、以上の手続きを繰り返す。

#### 4.2 実験1：室内環境における認識実験

室内における座位状態での把持姿勢認識精度を検証する。被験者に室内の椅子に座らせ4.1節のアプリケーションを起動し、図3の5パターンの把持姿勢でタスクを実施した。被験者は、肘を机などに固定しないで、タップ操作を行った。各タスクでは、図4のスクリーン上の40個のセルをランダムな順序で被験者にタッチさせることで1試行とした。一度、携帯端末を把持すると1試行中は、ほぼ同じ持ち方で入力したデータとなるが、1回入力するごとに端末をおいてもよいと指示した。被験者には、各把持姿勢に対して実験タスクを5回試行させたため、合計で25試行実施し、タスクを5回試行した後は、10分程度の休憩時間を取った。

訓練データには、各把持姿勢における200回分(5通りの把持姿勢で合計1,000回分)の入力データがあり、学習モデルを決定する際の適切なパラメータ(C,  $\gamma$ )を決定するために、各把持姿勢における入力データを5等分し、5分割交差検定を行った(図5)。このようにして構築した学習モデルを用いて、学習には用いていないテストデータ(1,000回分の入力データ)を識別し、その認識精度を検証した。

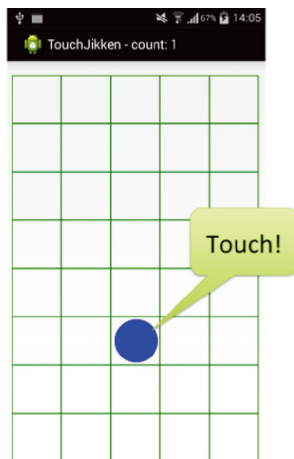


図4 実験アプリケーション  
Fig. 4 An application for experiments.

#### 4.3 実験1の結果と考察

特徴ベクトルごとの認識精度と順位を表3に示す。すべての特徴ベクトルにおいて把持姿勢の認識精度が90%以上(91.8~95.8%)となった。また、特徴ベクトルB, C, Dが特徴ベクトルAより高い認識精度を示しているため、加速度に関する情報が認識に有効であると考えられる。特に加速度に関する情報を用いた特徴ベクトルBがより高い認識精度を示している。3軸加速度センサの値に占める重力加速度成分は大きく、各軸における加速度の変化量の多くは線形加速度成分ではなく重力加速度成分となっていることから、特徴ベクトルBは端末の向きに関する情報が含まれていると考えられる。室内で座っている状態では、端末自身の動きより、端末の向きに関する情報が把持姿勢認識により寄与していることが分かった。

表4に各把持姿勢の認識結果を示す。表の縦軸は被験者が行った実際の把持姿勢、横軸は認識した把持姿勢を示している。多くの姿勢に対して90%以上の認識率を示しているが、片手操作の左親指操作(L-Th)の場合は他の姿勢よりも認識率が低くなっており、特に両親指操作(B-Th)と誤る場合が多かった。両親指操作では、実験に利用した5.7インチスクリーンの上部をタッチする際、指が届かず

データ1	データ2	データ3	データ4	データ5
L-Th	L-Th	L-Th	L-Th	L-Th
R-Th	R-Th	R-Th	R-Th	R-Th
B-Th	B-Th	B-Th	B-Th	B-Th
R-In	R-In	R-In	R-In	R-In
L-In	L-In	L-In	L-In	L-In

□ : 1試行(タッチ40回分のデータ)

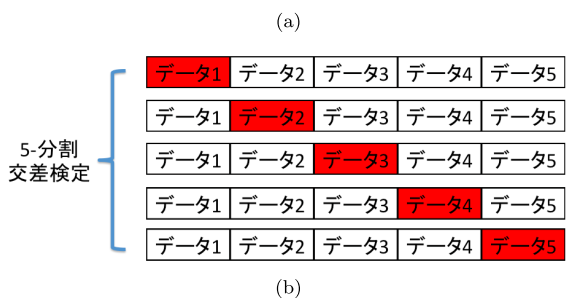


図5 訓練データセット：(a) 実験タスクとデータセットの分割；(b) 5分割交差検定

Fig. 5 Training data set: (a) Experiment task and Division of the data set; (b) 5-fold cross validation.

表3 特徴ベクトルによる認識精度と順位

Table 3 Grasp recognition accuracy according to feature vectors.

特徴ベクトル	A	B	C	D
認識精度 (%)	91.8	95.8	93.2	95.4
認識精度の高い順位	4	1	3	2

表 4 実験1の把持姿勢認識結果：(a) 特徴ベクトル A；(b) 特徴ベクトル B；(c) 特徴ベクトル C；(d) 特徴ベクトル D

Table 4 The result of Experiment1; grasp recognition under 5 grasp postures: (a) Feature vector A; (b) Feature vector B; (c) Feature vector C; (d) Feature vector D.

特徴ベクトル A	認識率 (%)				
把持姿勢	L-Th	R-Th	B-Th	R-In	L-In
L-Th	89.5	1	9	0	0.5
R-Th	1	95.5	3.5	0	0
B-Th	1.5	0	87.5	6	5
R-In	0	1	2.5	93.5	3
L-In	0	0	0	7	93

(a)

特徴ベクトル B	認識率 (%)				
把持姿勢	L-Th	R-Th	B-Th	R-In	L-In
L-Th	91.5	0.5	8	0	0
R-Th	0.5	96	3.5	0	0
B-Th	1	0	98	1	0
R-In	0	0.5	0.5	95.5	3.5
L-In	0	0	0.5	1.5	98

(b)

特徴ベクトル C	認識率 (%)				
把持姿勢	L-Th	R-Th	B-Th	R-In	L-In
L-Th	88	1.5	10	0	0.5
R-Th	0	96.5	3.5	0	0
B-Th	1.5	0	92	4	2.5
R-In	0	1	2.5	93.5	3
L-In	0	0	0.5	3.5	96

(c)

特徴ベクトル D	認識率 (%)				
把持姿勢	L-Th	R-Th	B-Th	R-In	L-In
L-Th	91	0	9	0	0
R-Th	0	97	3	0	0
B-Th	1	0	98.5	0.5	0
R-In	0	0	1	95	4
L-In	0	0	0.5	4	95.5

(d)

デバイスが動くことがあったことが影響していると考えられる。スクリーン上でタッチする場所によっても把持姿勢認識の精度が異なるため、把持姿勢認識を行う領域についても今後検討する必要がある。

#### 4.4 実験2：利用環境における認識実験

ユーザの利用環境に非依存な把持姿勢認識を実現するため、室内における座位状態に加え、室内における立位状態、また、電車内での座位、立位状態における把持姿勢認識精度を検証した。図6に電車内で想定する利用状況を示す4パターン（座って操作、ドアの横にもたれて操作、つり革を持ちながら操作、立って操作）を考慮した。特に電車内で立っている場合は、通常、片手ではつり革を掴んでいる

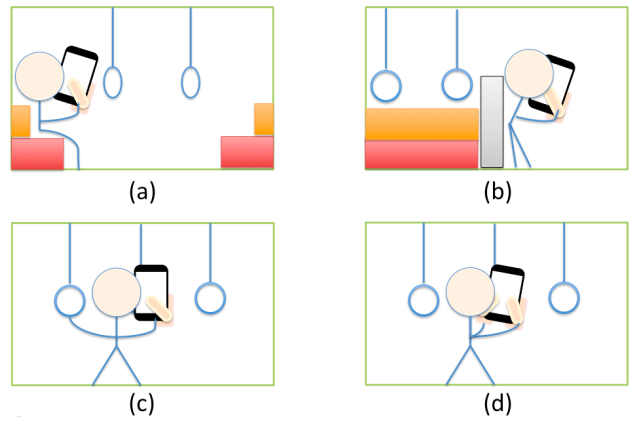


図 6 電車内での状況：(a) 座る；(b) ドアの横にもたれる；(c) つり革を持つ；(d) 立つ

Fig. 6 Using the devices on the train: (a) Sit on the chair; (b) Lean on the wall; (c) Hold a strap; (d) Stand.

表 5 電車内でのデータ

Table 5 The sensor data when users are using on the train.

状況	進行方向	データ数	場所
座る	右	1,000 個	山手線 (大崎～西日暮里)
ドアの横にもたれる	前	1,000 個	
つり革を持つ	右	1,000 個	
立つ			

ことが多いと考えられるため、右（左）手でつり革を持って左（右）手で操作する把持姿勢を L-Th (R-Th) とする。また、車内で何にも掴まらずに立ち、両手で携帯端末を操作する状況を想定し、B-Th, R-In, L-In の 3 通りを考慮した。電車内で取得したデータを表 5 に示す。データはすべて電車の走行中に取得した。訓練データは 4.2 節と同じ室内座位状態でのデータを用い、表 5 のデータをテストデータとして利用した。

#### 4.5 実験2の結果と考察

実験結果を表 6 に示す。各特徴ベクトルに対する把持姿勢の認識精度は、室内の座っている状態に比べて低くなった。特に、室内の立っている場合には、特徴ベクトル C が高い認識精度を示している。座っている状態と立っている状態で端末の持ち方に違いがあるかどうかを検証するため人差し指操作時の端末の傾きについて調査した。端末に内蔵された加速度センサの 3 軸と被験者が端末を把持しているときの端末の向きを図 7 に示す。重力加速度の x 軸成分は非常に小さい値であったため、 $0\text{m/s}^2$  と近似して、重力加速度の y 軸成分と z 軸成分から被験者が端末を把持している際の水平面からの傾き  $\theta$  を式 (1) で求めると、座位状態で約 29 度、立位状態で約 19 度となり、被験者の姿勢により、携帯端末を見る視線方向と端末の傾きが変化していることが分かった。

$$\theta = \tan^{-1} \left( \frac{\text{重力加速度の y 軸成分}}{\text{重力加速度の z 軸成分}} \right) \quad (1)$$

表 6 実験 2 の特徴ベクトルによる把持姿勢認識結果；(a) 認識精度；(b) 順位

Table 6 The result of Experiment2; (a) grasp recognition accuracy according to feature vectors; (b) ranking.

認識精度 (%)		特徴ベクトル			
場所	状況	A	B	C	D
室内	座る	91.8	95.8	93.2	95.4
	立つ	89.3	75.9	90.4	79.2
電車内	座る	80.4	87.2	86.4	88.7
	ドアの横にもたれる	81.4	87.3	82.4	86.7
	つり革を持つ/立つ	82.0	82.8	84.9	83.8

(a)

認識精度の順位		特徴ベクトル			
場所	状況	A	B	C	D
室内	座る	4	1	3	2
	立つ	2	4	1	3
電車内	座る	4	2	3	1
	ドアの横にもたれる	4	1	3	2
	つり革を持つ/立つ	4	3	1	2

(b)

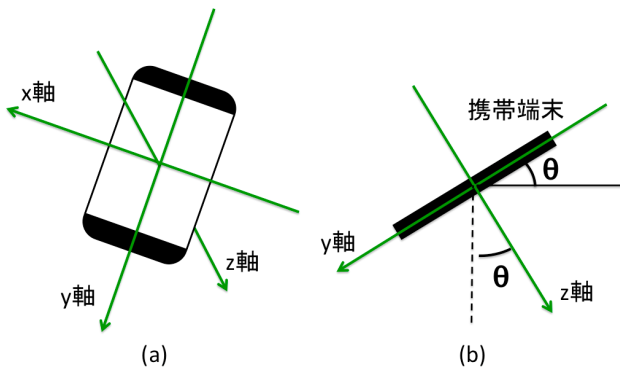


図 7 加速度センサの軸：(a) 内蔵された加速度センサの 3 軸；(b) 端末の側面図と傾き

Fig. 7 Axis of accelerometer: (a) 3 axis of built-in accelerometer; (b) Side view of smartphone and slope.

訓練データが座位状態で、テストデータが立位状態の場合は端末の向きが異なるため、重力加速度を含む特徴ベクトル B と D の認識精度が低くなり、端末自身の動きに関する特徴量は似ているため、特徴ベクトル C の方が有効であると考えられる。

電車内の場合は、室内環境にはない揺れといった外界からの影響が把持姿勢の認識精度を低下させていると考えられる。

被験者の状況によって認識精度の高い特徴ベクトルが異なっている。おおむね加速度に関する情報を利用した場合が認識精度を高めていると考えられるが、特徴量の決定方法については今後検討する予定である。

## 5. 考察

### 5.1 認識精度の改善

#### 5.1.1 センサ値に関する考察

各軸のデータに対して特徴量を考えたが、各軸のデータを合成した 3 軸合成値についてはセンサからの生データのみ特徴量として用いた。ユーザがスクリーンをタップする際のデバイスの動きによりセンサ値が変化するが、3 軸合成値の絶対値から動きの大きさが分かる。スクリーンをタップする際の端末の動きの大きさは、把持姿勢により異なると考えられるが、3 軸合成値の前後のデータを用いた様々な特徴量を検討することで、認識精度の改善が期待できる。

本研究では内蔵センサから得られる情報を用いることで利用環境が異なっても比較的高い認識精度を実現することができた。しかし、これらの情報には把持姿勢の違いによるデバイスの動きやスクリーンをタッチする力の強さだけでなく、外部環境からの雑音が入る可能性が高い。たとえば、電車内で端末操作を行う状況を考えると、電車の揺れや進行方向への加速度影響によりセンサ値が変化し、訓練データとして用いた室内で座っている状態での操作と比べデータが大きく異なると考えられる。今後、室内座位状態での操作と外部環境に影響がある場所との差を小さくするため、室内座位状態でのデータと実環境でのデータの差を除去し、変動幅が小さい差分を用いて把持姿勢を推定し判別精度を検証する予定である。

#### 5.1.2 両親指操作に関する考察

今回用いた 5 パターンの把持姿勢の中で両親指操作と他の 4 パターンの操作との間での誤認識が比較的多く見受けられた。特に、両親指で操作する場合には、端末の上部をタップする際、指が届かずデバイスが動くことがあったことが影響し、片手操作の左親指操作と同じシチュエーションだと判定した。

実験の際、被験者には図 3 で示した 5 種類の把持姿勢で携帯端末を操作するように指示したが、携帯端末のどの位置に指を添え把持するかといった詳細な指示は与えず、操作に利用する指を指示するのみで、自由に被験者に携帯端末を持たせた。被験者の両親指操作時の把持姿勢を確認したところ、図 8(a) のように端末の同じ高さの場所に両手を添えるのではなく、図 8(b) のように左手よりも少し高い位置に右手を添えていたことが分かった。

実験端末が画面サイズが 5.7 インチのタブレット (Phablet) であったことから、被験者は指が届きづらい場所を少なくするために手の高さを変えていたと考えられる。しかし、スクリーンの左上は依然として、操作しづらい場所となっており、指の届きづらさという点では左親指操作 (L-Th) と両親指操作 (R-Th) で同様であったと推察され、タップ時の特徴量が似た値となることで、誤認識が発生し

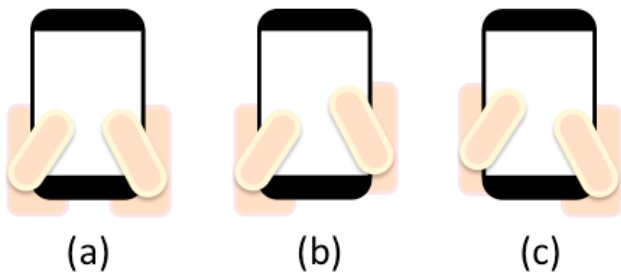


図 8 両親指操作 (B-Th) のパターン：(a) 両手の位置が一致；(b) 右手の位置が高い；(c) 左手の位置が高い

Fig. 8 Patterns of Both Thumbs Interaction: (a) Position of both hands is a match; (b) Position of the right hand is high; (c) Position of the left hand is high.

たとえられる。一方、スクリーン右上の操作は、右手を少し上に添えているために、親指が届きやすくなることで、右親指操作 (R-Th) を両親指操作 (B-Th) と誤認識する割合が減少したと考えられる。ふだん画面サイズが大きいタブレットを利用しているユーザは図 8 の 3 通りの把持姿勢で両親指操作を行うことが分かった。このような同じ操作指であっても把持姿勢が異なる場合を考慮して、把持姿勢認識の精度向上に向けた検討する必要がある。

また、実験タスクでは、スクリーン上にランダムに円が登場し繰り返してタップする操作であったが、このようなタスクの場合は、把持した手を端末から離さずに操作した方が時間を短くし、疲労を減らすことができる。しかし、実際にユーザがスクリーンが大きい端末で操作を行うときには、把持した手を離したり、把持姿勢を変える傾向がある [16]。特に、ユーザが両親指操作を行った後に (キーボードインタフェース) 上部をタップする際には、片手を離して、その手で上部をタップすることが多い。今後、両親指操作のパターンについてさらに考察し、他の把持姿勢との同じシチュエーションだと判定する誤認識に対する対策を考える必要がある。

## 5.2 被験者について

本論文における評価実験は 1 名の被験者データを用いて判別精度を検証し、同一人物のデータを学習に用いれば端末の把持姿勢を高い確率で識別することが可能であることが分かった。一方、他人の学習モデルで把持姿勢を認識することが可能とすれば、システムの導入コストを大きく低減することができ、システムの有用性が向上すると考えられる。今後は複数の被験者に対して実験を行うことで、ある特定の被験者の訓練データから他のユーザの把持姿勢も認識できるかどうかを検証するとともに、被験者に依存しない把持姿勢認識を実現するために有効な特徴量について検討を進める予定である。

## 5.3 把持姿勢認識のタイミング

提案手法では、スクリーンのタップ操作で把持姿勢を推

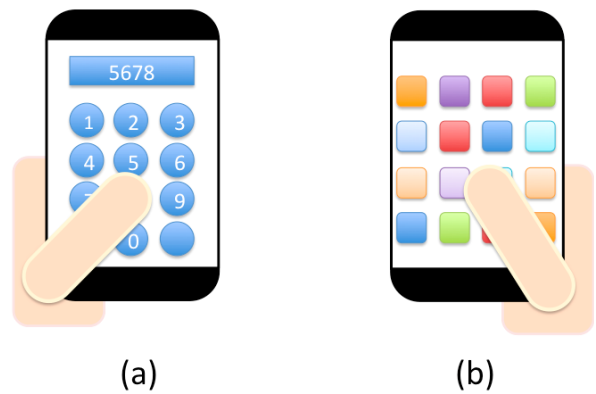


図 9 把持姿勢認識のタイミング：(a) 4桁ロック解除；(b) アイコン選択

Fig. 9 The timing of Grasp recognition: (a) 4-digit unlock; (b) Click icon.

定するため、タップ操作を行った後で適切な入力インタフェースを提示することが考えられる。たとえば、端末のロックに用いられる 4 桁の数字を入力する際のデータやアプリケーションを実行するためにアイコンをタップする際のデータから端末の把持姿勢を推定し、その把持姿勢に応じた適切な入力インタフェースを提示することが考えられる (図 9)。以上のような例を実現するために、どのような把持姿勢に対してもユーザがタップしやすい部分を考慮する、もしくは認識精度が高い部分のみ使用することが考えられる。

## 5.4 オフライン処理の妥当性

本研究では、携帯端末に内蔵されているセンサのデータを用いて、図 3 の想定する把持姿勢をどの程度認識することが可能であるかを検証するために、携帯端末上で取得したデータを外部のノート PC で処理した。しかし、最終的には携帯端末で把持姿勢を認識する必要があり、オンラインで処理できる必要がある。本研究では、機械学習を用いて把持姿勢とセンサ値の関係を学習し、スクリーンをタッチした際のセンサ値から把持姿勢を認識しているが、実運用の際に携帯端末で行うべき処理は、センサデータの取得、特徴量抽出、把持姿勢認識となる。把持姿勢認識には、1 回のタッチに関して、座標などタッチスクリーンから得られるデータと、タッチした瞬間を基準として前後 100 ミリ秒間の加速度センサおよびジャイロセンサのデータを利用するため、データ取得に 100 ミリ秒必要である。特徴量抽出では、取得したデータの FFT 解析やデータの平均、標準偏差などを計算するために約 70 ミリ秒、学習モデルにテストデータを投入し、把持姿勢を認識するために約 15 ミリ秒必要となる。これらの処理時間は、データ取得を除き、オフライン処理したノート PC 上で計測した時間で、タッチ操作から約 200 ミリ秒後に把持姿勢を認識できていることになる。



携帯端末上で同様の処理を行った場合については、ノート PC より多くの時間を必要とすることが予想されるが、ほぼリアルタイムでの認識が可能であると思われる。今後、携帯端末上で動作するプロトタイプを実装して確認する予定である。

## 6. おわりに

本論文では、携帯端末の内蔵センサのみを用いて SVM より 5 パターンの把持姿勢を認識する手法を提案した。端末把持姿勢の認識実験を行い、室内で座っている状態での認識精度は 91.8~95.8%、室内で立っている状態と電車内での状態での認識精度は 75.9~90.4% となった。特に、ジャイロセンサとタッチスクリーン情報に加えて、加速度センサの情報が判別に有効であることが分かった。今後は、より多くの状況で加速度センサからのデータを取得し推定精度を検証する。また被験者を追加して、個人差による影響について調査する予定である。

**謝辞** 本研究の一部は日本学術振興会科学研究費補助金基盤研究 (C) (25330227) の研究助成によるものである。ここに記して謝意を表す。

## 参考文献

- [1] Cheng, L.-P., Lee, M.-H., Wu, C.-Y., Hsiao, F.-i., Liu, Y.-t., Liang, H.-S., Chiu, Y.-C., Lee, M.-S. and Chen, M.Y.: iRotateGrasp: Automatic Screen Rotation based on Grasp of Mobile Devices, *Proc. CHI'13*, pp.3051-3054, ACM (2013).
- [2] Ono, M., Shizuki, B. and Tanaka, J.: Touch & Activate: Adding Interactivity to Existing Objects using Active Acoustic Sensing, *Proc. UIST'13*, pp.31-40, ACM (2013).
- [3] Goel, M., Wobbrock, J.O. and Patel, S.N.: GripSense: Using Built-In Sensors to Detect Hand Posture and Pressure on Commodity Mobile Phones, *Proc. UIST'12*, pp.545-554, ACM (2012).
- [4] Bao, L. and Intille, S.S.: Activity recognition from user-annotated acceleration data, *Proc. 2nd International Conference on Pervasive Computing*, pp.1-17 (2004).
- [5] Sun, L., Zhang, D., Li, B., Guo, B. and Li, S.: Activity Recognition on an Accelerometer Embedded Mobile Phone, *Ubiquitous Intelligence and Computing*, pp.548-562 (2010).
- [6] 太田和也, 岩崎正裕, 藤波香織: 携帯端末を用いた行動認識における端末格納場所情報を用いた認識パラメータ更生法に関する検討, 情報処理学会マルチメディア, 分散, 協調とモバイル (DICOMO2013) シンポジウム, pp.641-646 (2013).
- [7] Martinez, J., Perez, H., Escamilla, E. and Suzuki, M.M.: Speaker recognition using Mel frequency Cepstral Coefficients (MFCC) and Vector quantization (VQ) techniques, *Electrical Communications and Computers (CONIELECOMP), IEEE 2012 22nd International Conference*, pp.248-251 (2012).
- [8] Chang, C.-C. and Lin, C.-J.: Libsvm: A library for support vector machines, *ACM Trans. Intelligent Systems and Technology*, Vol.2, No.3, pp.27:1-27:27, May (2011).
- [9] Kim, K.-E., Chang, W., Cho, S.-J., Shim, J., Lee H.,

Park, J., Lee, Y. and Kim, S.: Handgrip pattern recognition for mobile user interfaces, *18th conference on Innovative applications of artificial intelligence Volume 2, IAAI'06*, pp.1789-1794 (2006).

- [10] Taylor, B. and Bove, V.: Graspables: grasp-recognition as a user interface, *Proc. 27th international conference on Human factors in computing systems, Proc. CHI'09*, pp.917-926, ACM (2009).
- [11] Cheng, L.-P., Liang, H.-S., Wu, C.-Y. and Chen, M.: iGrasp: Grasp-based Adaptive Keyboard for Mobile Devices, *Proc. CHI'13*, ACM (2013).
- [12] He, Z. and Jin, L.: Gesture recognition based on 3D accelerometer for cell phones interaction, *IEEE Circuits and Systems*, pp.217-220 (2008).
- [13] Chodhury, T., Borriello, G., Consolvo, S., Haehenl, D., Harrison, B., Hemingway, B., Hightower, J., Klasnja, P., Koscher, K., LaMarca, A., Landay, J.-A., LeGrand, L., Lester, J., Rahimi, A., Rea, A. and Wyatt, D.: The Mobile Sensing Platform: An Embedded Activity Recognition System, *Pervasive Computing*, pp.32-41, IEEE (2008).
- [14] Van Laerhoven, K. and Cakmakci, O.: What shall we teach our pants? Wearable Computers, *IEEE 4th International Symposium*, pp.77-83 (2000).
- [15] Schmidt, A., Aidoo, K., Takaluoma, A., Tuomela, U., Van Laerhoven, K. and de Velde, W.: Advanced Interaction in Context, *1st International Symposium on Handheld and Ubiquitous Computing*, Springer Berlin, Heidelberg, pp.89-101 (1999).
- [16] Hooper, S.: The Rise of the Phablet: Designing for Larger Phones, available from (<http://www.uxmatters.com/mt/archives/2014/11/the-rise-of-the-phablet-designing-for-larger-phones.php>).



朴 燦鎬 (学生会員)

1990 年生。2014 年東京大学工学部電子情報工学科卒業。2015 年同大学大学院工学系研究科電気系工学専攻修士課程在学中。



小川 剛史 (正会員)

1974 年生。1997 年大阪大学工学部情報システム工学科卒業。1999 年同大学大学院工学系研究科博士前期課程修了。2000 年同研究科博士後期課程中退後、同大学サイバーメディアセンター助手。2007 年東京大学情報基盤センター講師、2010 年同准教授となり、現在に至る。拡張現実感、ヒューマンインタフェース、グループウェア等に関する研究に従事。博士 (情報科学)。