

携帯端末の利用環境に依存しない端末把持姿勢推定に関する 一検討

朴燦鎬^{†1} 小川剛史^{†2}

スマートフォンなどの携帯端末に、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、携帯端末に搭載されたジャイロセンサのデータやタッチスクリーンへの入力情報などを用いて、携帯端末の把持姿勢を推定する研究が行われているが、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。一方、携帯端末の加速度センサを用いて、ユーザの行動を認識する研究が盛んに行われている。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推定を実現するため、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推定手法に関する検討を行った。本稿では、機械学習による把持姿勢推定手法と、室内と電車内で実施した実験結果について述べる。

An Study on Grasp Recognition of a Smartphone Independent of Users' Situations

CHANHO PARK^{†1} TAKEFUMI OGAWA^{†2}

In order to provide the appropriate interface, it is important to recognize the grasp posture on using smartphones. There are many studies on estimation of grasp posture using the input information like touchscreen and gyroscope data equipped with the device. However, these studies do not consider the users' situations. On the other hand, there are many studies on recognition of users' contexts by using an accelerometer. To realize grasp recognition independent of users' situation, we consider using the sensor data built in gyroscope and accelerometer, and touchscreen information. In order to evaluate our system by machine learning, we also describe the result of the experiments that users are using the devices in the room and the train.

1. はじめに

スマートフォンやタブレット端末の普及に伴い、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。

端末の把持姿勢はユーザの置かれている状況にも存在する。例えば、方の手に荷物を持っていたり、電車内でつり革を掴んだ状態で携帯端末を操作したい場合には片手操作となり、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。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判別する手法に関する研究が行われてきた。

把持姿勢に関する研究[1][9][10]では、端末のどの部分を把持し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、既存の端末に変更を加えずに適用することができない。また、端末に内蔵されるジャイロセンサや加速度センサを用いて把持姿勢を推定する研究も存在するが、室内や電車内などユーザの場所が変化すると把持姿勢の推定精度も変化すると考えられる。

そこで、本研究では携帯端末の動きや、タッチイベントなどを検出するときに良く用いられるジャイロセンサとタッチパネルからの情報とともに、行動認識の分野で良く使

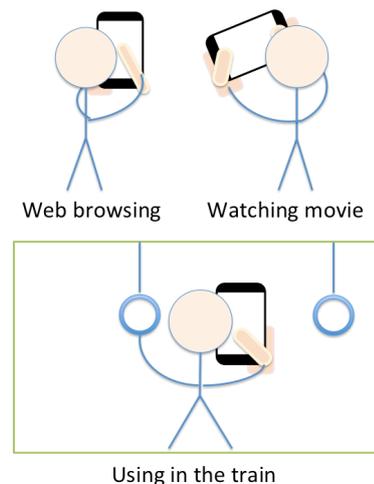


図1 いろいろな状況での把持姿勢パターン

Figure 1 Many patterns How users are grasping the devices

われている加速度データの情報をうまく改善することで、把持姿勢認識を行う。そして、室内で座っている状態での操作だけでなく、立っている時、電車内での操作も考慮し、実環境への応用が期待できる。

2章で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章にて本稿のまとめと今後の課題について述べる。

^{†1} 東京大学大学院工学系研究科
Graduate School of Engineering, The University of Tokyo
^{†2} 東京大学情報基盤センター
Information Technology Center, The University of Tokyo

2. 関連研究

2.1 把持姿勢認識に関する研究

携帯端末の把持姿勢認識に関する研究が盛んに行われている。

Cheng ら[1][11]は静電容量センサを用いて把持姿勢に応じて携帯端末の表示を自動的に切り替わる。多くの静電容量センサを端末に装着して把持姿勢の推定を行うため、高精度な把持姿勢認識が可能であるが、コストが高く、新しいハードウェアの構成が必要である。そして、把持した手を離してもそのままのインタフェース（画面の配置など）で操作したいが、この手法だと把持するときのインタフェースのみ考えている。

その他のセンサを用いた研究例[2][9][10][12]もある。Ono らは、スピーカとマイクを物体に貼り付けることによって、物体を把持する手の姿勢を判別する手法を提案している[2]。しかし、端末へ加わる荷重分布が体の姿勢により違う点や、適用可能な範囲が振動スピーカの周囲に限定されているなど、実環境利用における課題が残っている。

また、端末に内蔵されたセンサを使用した手法も提案されている。Goel らは、タップ操作とスワイプ操作を行う際にどのような把持姿勢をしているのかを判断するシステムを構築した[3]。この手法では、把持姿勢を推定するために、ジャイロセンサから得た端末の傾きデータ、タッチスクリーンに入力した際の指先がスクリーンに触れている面積とスライド操作の形状を特徴量として用いている。実際にどのぐらいの精度で判定を行うか実験を行い、5 回以上の連続なタッチデータで判別率を示した。追加のデバイスを必要としない手法だが、5 回以上のタッチから判別するようなシステムになっていて短時間で判別ができない。そして、認識するための方法を決めといて行っているため、判別性能と実環境利用においてまた課題が残っている。

本研究では、新しいセンサは使わずに、内蔵されたセンサのみを用いて把持姿勢認識を行う。そして、文献[3]で用いられているジャイロセンサとタッチスクリーンの情報に加えて、加速度センサの情報を用いる点が異なる。

2.2 加速度データを用いたユーザのコンテキスト推定に関する研究

続いて、加速度データを用いたユーザのコンテキスト推定に関する研究を紹介する。行動認識の分野ではユーザの動きを判定するために加速度センサを用いた研究[4][5][6][13][14][15]が盛んに行われている。Laerhoven らは、ユーザの動きを区分するため、ウェアラブルデバイスに加速度センサを用いて歩く、階段に上がるなどのユーザのコンテキストを推定した[14]。そして、Schmidt らは、加速度の情報から、デバイスを把持しているか、テーブルに置いてあるか、鞆の中に入っているかなどの行動認識を区分した[15]。例のように行動認識の分野では加速度データを用

いていろいろな成果が出ている。

2.3 加速度データを用いた把持姿勢認識に関する研究

また、加速度データを用いた把持姿勢認識に関する研究[9][10][12]もある。Kim ら[9]は、加速度データと、静電容量センサに触れる手の形状情報から操作方法を判別する提案をしている。しかし、把持する手の左右は判別できず、新しいセンサを使用しているため一般的ではない。Taylor ら[10]は、実際の物体などを把持する姿勢を想定し、加速度のデータを用いてユーザの把持姿勢変化を推定したが、対象が携帯端末ではなく、新しいハードウェアで構成されている。

3. 提案手法

関連研究では、ユーザが端末を手にするだけで把持姿勢を推定しているが、新しいセンサなどハードウェアの追加が必要となることがある。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把持姿勢を判定する。実装したシステムの概要を図 3 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などで良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。

3.1 識別する把持姿勢

本手法で認識する端末の把持姿勢は先行研究も参考にし、通常の操作で良く見られる 5 パターンとした。図 4 に識別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、同じ手の親指で操作することを意味する。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 パターンがある。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パターン(Right Index finger: R-In, Left Index finger: L-In)と両手でモバイル端末を把持し、両方の手の親指で端末を操作するパターン(Both Thumb: B-Th)がある。

3.2 把持姿勢認識

3.2.1 データの取得

提案手法では、ユーザによる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 50Hz で値を計測しており、タッチイベントが発生した時刻を中心に 0.2 秒間のデータを記録する。

表1 特徴ベクトル
 Table1 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

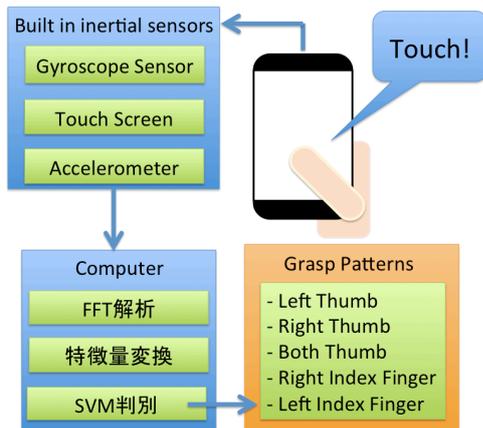


図3 システムの概要
 Figure3 System Configuration

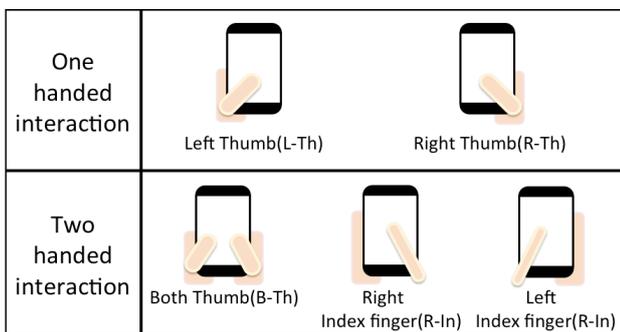


図4 把持姿勢認識セット
 Figure4 Grasp recognition set

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る。本稿ではこの加速度を線形加速度と呼ぶ。

3.2.2 特徴量抽出

3.2.1 節で述べたデータから以下のように特徴量を決定する。

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。特徴量として、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 次元を用いた。

内蔵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。

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が各 3 次元、タップ前後の 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 次元になる。

また、内蔵センサのデータについては周波数成分分析を行う。周波数成分分析は、行動認識や音声認識などの分野では、データの周波数成分が特徴量としてよく利用されており [6][7]、把持姿勢推定を行う研究 [2] にも用いられている。本研究では、周波数成分分析に関する特徴量を用いる。

具体的には、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、周波数成分の中で、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 次元ずつ) 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も用いて振幅に関する情報 (振幅の最大値、標準偏差、相関係数が各 3 次元ずつ) で計 30 次元になる。

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 1 のように使用するセンサを決めることで、特徴ベクトル A~D を考えることができ、すべての特徴量を用いる場合 (特徴ベクトル D)、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、機械学習手法の一つである 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 ライブラリの一つである 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た。RBF カーネル関数を用い、最適なパラメータ C と γ 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対して適用した。SVM の判別に

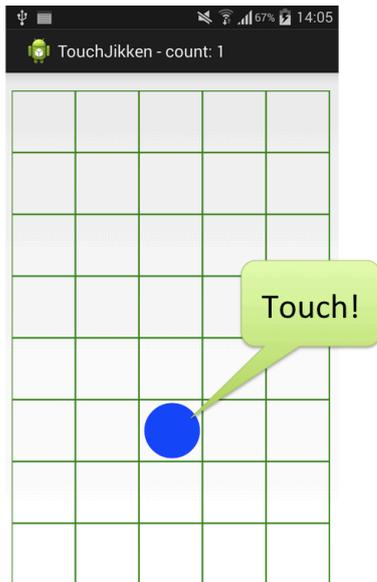


図5 アプリタスク
 Figure5 Application Task

より5パターンの把持姿勢の中で1つを判別する。

3.3 システム構成

携帯端末には 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168g, 画面サイズは5.7インチ(会場度1080×1920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判別はApple社のMacBook Air(CPU: Intel Core i5, RAM: 4GB)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判別性能の違いを検証するため、表1に示す特徴ベクトルを用いて判別実験を行った。実験1では予備実験として室内に置ける座位状態での判別性能を、実験2では室内、電車内に置ける座位、立位状態での判別性能を検証する。

4.1 実験タスク

携帯端末のスクリーンタップした際に、その把持姿勢を推定するため、スクリーンをタッチした際の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5に示す。スクリーンを縦8個、横5個のマス目に分割し、ランダムに選択したマスに円を表示する。ユーザが円をタップすると、その瞬間のセンサの値を記録し、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。すべてのマスに円を表示するまで、以上の手続きを繰り返し行う。

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判別性能を検証する。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4の5パターンの把持姿勢でタスクを実施した。

表2 特徴ベクトルによる認識精度と順位

Table2 grasp recognition accuracy according to feature vectors and ranking

特徴ベクトル	A	B	C	D
認識精度(%)	91.8	95.8	92.2	95.4
判別率の高い順位	4	1	3	2

表3 実験1の把持姿勢判別結果:

- (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

Table3 The result of Experiment1; grasp recognition with 5 hand postures.

- (a) Feature vector A; (b) Feature vector B;
 (c) Feature vector C; (d) Feature vector D
 (a)

特徴ベクトル A	判別率 (%)				
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

(b)

特徴ベクトル B	判別率 (%)				
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

(c)

特徴ベクトル C	判別率 (%)				
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

(d)

特徴ベクトル D	判別率 (%)				
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

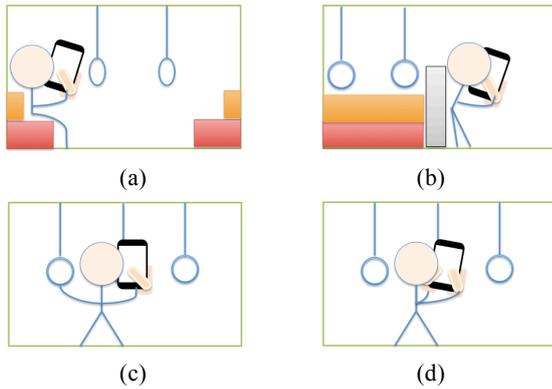


図6 電車の中での状況
 (a) 座る; (b) ドアの横に寄りかかる;
 (c) つり革を持つ; (d) 立つ
 Figure6 Using the devices in the train
 (a) Sit on the chair; (b) Lean on the wall;
 (c) Hold a strap; (d) Stand

各タスクでは、スクリーン上の 40 カ所をランダムにそれぞれ 5 回タッチさせ、把持姿勢 5 パターン分を合わせたタッチ 1000 回分を 1 データセットとし、訓練データおよびテストデータとしてそれぞれ 1 セット用意した。

4.3 実験結果と考察

特徴ベクトルごとの認識精度と順位を表 2 に示す。結果として、特徴ベクトルに対する把持姿勢の認識精度が 90% 以上 (91.8~95.8%) となった。このことは、個人での使用を目的とする端末把持姿勢手法としての実用可能性が考えられる。

また、特徴ベクトル B, C, D が特徴ベクトル A より高い判別率を示しているため、加速度に関する情報が判別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 がより高い判別性能を示している。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が、特徴ベクトル B における判別性能が高かった理由のひとつと考えている。

表 3 に各把持姿勢判別結果を示す。多くの姿勢に対して 90% 以上の判別率を示しているが、片手操作の左親指操作 (L-Th) の場合は他の姿勢より低くなっている。左親指操作を両親指操作 (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。タッチするスクリーンの場所によっても把持姿勢判別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、初期検討として、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

表 4 電車内でのデータ

Table4 The data using in the train

状態	進行方向	データ数	場所
座る	右	1000 個	山手線 (大崎 ~ 西日暮里)
ドアの横に寄りかかる	前	1000 個	
つり革を持つ、立つ	右	1000 個	

表 5 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

Table5 The result of Experiment2; (a) grasp recognition accuracy according to feature vectors; (b) ranking

(a)

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	座る	80.4	87.2	86.4	88.7
	ドアに寄りかかる	81.4	87.3	82.4	86.7
	つり革を持つ、立つ	82.0	82.8	84.9	83.8

(b)

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	座る	4	2	3	1
	ドアに寄りかかる	4	1	3	2
	つり革を持つ、立つ	4	3	1	2

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、ドアの横によりかかって操作、つり革を持ちながら操作、立って操作) を考慮した。電車の中でのテストデータを表 4 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 4.3 節と同じデータを用いた。

4.5 実験結果と考察

実験結果は表 5 に示す。各特徴ベクトルに対する把持姿勢の認識精度は、室内に比べて低かった。これは、理想な環境に比べ、端末が不安定で、電車の進行方向による、揺れによる影響などを受けるので、把持姿勢の判別精度が低下すると考えられる。

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。概ね加速度に関する情報を利用した方が認識精度を高めることが可能であると考えられるが、特徴量の決定方法について今後検討する予定である。

5. おわりに

本稿では、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推定した。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考察する。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。

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. In Proc. CHI'13, ACM, 2013.
- 2) Ono, M., Shizuki, B., and Tanaka, J. Touch & Activate: Adding Interactivity to Existing Objects using Active Acoustic Sensing. In Proc. UIST'13, 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, pp545-554, 2012.
- 4) Bao, L., and Intille, S.S.: Activity recognition from user-annotated acceleration data, Proc. the 2nd International Conference on Pervasive Computing, pp.1-17, 2004.
- 5) Sun, L., Zhang, D., Li, B., Guo, B., 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., Suzuki, M.M. : Speaker recognition using Mel frequency Cepstral Coefficients (MFCC) and Vector quantization (VQ) techniques. Electrical Communications and Computers (CONIELECOMP) , 2012 22nd International Conference, IEEE, pp.248-251, 2012.
- 8) Chang, C.-C., and Lin, C.-J. Libsvm: A library for support vector machines. ACM Transactions on 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. In Proceedings of the 18th conference on Innovative applications of artificial intelligence Volume 2, IAAI'06, pp. 1789-1794. AAAI Press, 2006.
- 10) Taylor, B., and Bove, V., Graspables: grasp-recognition as a user interface. In Proceedings of the 27th international conference on Human factors in computing systems, CHI '09, pp. 917-926, New York, NY, USA, ACM, 2009.
- 11) Cheng, L.-P., Liang, H.-S., Wu, C.-Y., Chen, M., iGrasp: Grasp-based Adaptive Keyboard for Mobile Devices. In 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, pp217-220, 2008.
- 13) Chodhury, T., Consolvo, S., Harrison, B., et al. The Mobile Sensing Platform: An Embedded Activity Recognition System. Pervasive Computing, IEEE 7, pp32-41, 2008.

14) Van Laerhoven, K., and Cakmakci, O. What shall we teach our pants? Wearable Computers, The Fourth International Symposium on, pp77-83, 2000.

15) Schmidt, A., Aidoo, K., Takaluoma, A., Tuomela, U., Van Laerhoven, K., and de Velde, W. Advanced Interaction in Context. In H.-W. Gellersen, ed., Handheld and Ubiquitous Computing. Springer Berlin, Heidelberg, pp89-101, 1999.