

ユーザの姿勢を考慮した 携帯端末の把持姿勢認識に関する一検討

朴 燦鎬¹ 小川 剛史²

概要: いつでもどこでも利用できる携帯端末であるが、携帯端末に適切なインタフェースを提示するため、ユーザの姿勢を考慮した把持姿勢を認識することが重要である。ユーザの姿勢を考慮した把持姿勢認識は、重力加速度に基づくシステムが一般的に普及されているが、例えばウェブを見ながら横になると画面が意図せず回転し操作が中断される。また、ジェスチャ認識により画面を提示する手法も存在するが、新しいジェスチャを学ぶ必要がある。そこで本研究では、より自由な姿勢で携帯端末を利用できるインタフェースの実現を目的として、ユーザの姿勢を考慮した携帯端末の把持姿勢認識を提案する。携帯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て、サポートベクターマシンにより 12 種類の把持姿勢を認識した。

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 Study on Grasp Recognition of Smartphone with Considering Users' Orientations

PARK CHANHO¹ OGAWA TAKEFUMI²

Abstract: We can use smartphone anytime and anywhere. It is important to recognize hand postures with considering users' orientations in order to provide appropriate interface. Although there are many studies on gravity-based grasp recognition that has been applied in commodity smartphone, it is difficult to use without worrying about the orientation. For example, if user is lying down while searching the web, screen orientation is rotated unintentionally, so operation is interrupted. Others have also proposed techniques by recognition gestures. However, these techniques require the user to learn new gestures. In this study, we propose grasp recognition of smartphones with considering users' orientations in order to realize the interface that we can use the smartphone with more free orientations. Using the sensor data built-in gyroscope and accelerometer, and touchscreen information, we classify 12 grasp postures by support vector machine.

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

1. はじめに

スマートフォンなどの携帯端末が普及し、いつでもどこでも素早く様々な情報を得ることが可能になった。携帯端末に内蔵された複数のセンサを用いることで端末の位置や向きなどを検出でき、横に長いコンテンツを見たい場合に

は、端末を横にして把持することで自動的に画面を切り替えることができる [1][2]。一方、日常生活では、ベッドで横になりながらウェブを見たりするなど、持ち方は同じでも姿勢が異なる場合が存在する。姿勢に応じて適切なインタフェースが決まるため、ユーザの姿勢を考慮した把持姿勢認識は重要である (図 1)。

ユーザの姿勢を考慮した把持姿勢認識に関する従来研究では、加速度センサのセンサ値から重力方向を検出して画面を切り替えるシステムが提案されている [1][2]。ユーザ

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

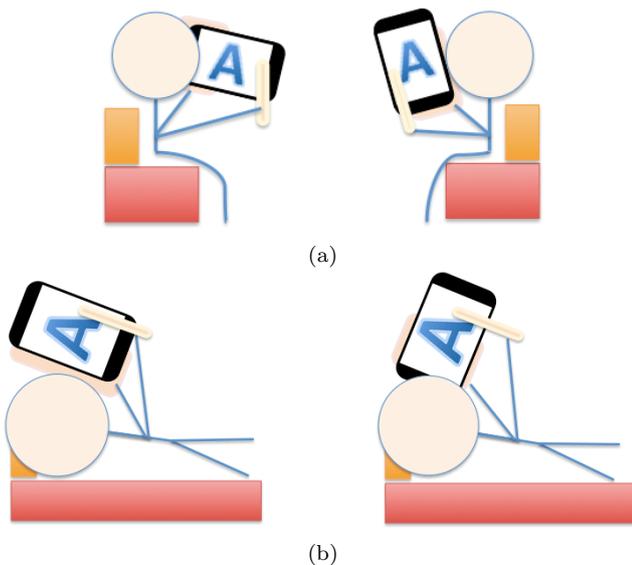


図 1 ユーザの利用環境および姿勢に応じる携帯端末の把持姿勢: (a) 横画面操作と縦画面操作; (b) 横になりながら操作

Fig. 1 Grasp postures of Smartphone according to the users' situations and positions: (a) Landscape mode and Portrait mode; (b) Users are using the smartphones while lying in bed

が横(縦)画面にして操作したい場合、端末を横(縦)向きに把持するだけで自動的に画面の表示が回転し、アイコンなども再配置される。しかし、例えば縦画面でウェブを見ながらベッドに横になると画面が意図せず回転して操作が中断されるなど、姿勢を気にせず利用することは困難である。文献 [4] では、ユーザが端末のどの部分に触れているのかを検出するために静電容量センサを用いており、端末に新たなセンサを追加する必要がある。したがって、一般に普及している端末をそのまま利用することができない。

我々の研究グループでは、今まで携帯端末の内蔵センサ(ジャイロセンサ、加速度センサ)とタッチスクリーンから得られるデータに基づいて把持姿勢を認識し、携帯端末の利用環境に非依存な把持姿勢認識手法としての可能性について検討してきた [15]。本稿では、ユーザの姿勢を考慮した把持姿勢認識を検討するため、端末を縦画面と横画面にして座っている状態で操作する場合だけでなく、ユーザが横になっている状態での操作も想定して把持姿勢認識を行う。特に、ユーザの姿勢が変わることで端末の向きがそれぞれ異なるが、端末の向きに関する情報が含まれている加速度センサの値を特徴量として用いることで、認識精度の改善ができるかを検討する。

以下、2章では、関連研究について紹介する。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、ユーザの姿勢を考慮した把持姿勢認識精度を検証するために行った評価実験について述べる。最後に5章にて、本論文のまとめについて述べる。

2. 関連研究

携帯端末のユーザの姿勢を考慮した把持姿勢認識の手法 [1][2][3][4] が提案されている。Hinckley ら [2] は、傾きセンサ(2軸加速度センサ)を用いて端末の傾き角度に応じて画面を提示するシステムを提案した。傾き角度を 45° ごとに閾値を決め、「Portrait」、「Portrait(Upside Down)」、「Landscape(Left)」、「Landscape(Right)」の4種類で分類し画面を配置した。現在のスマートフォンなどの多くの端末には加速度センサなどで検出した重力加速度に基づくシステムとして構築されている。しかし、ユーザが横になっている状態などで端末操作を行うと画面が不自然に回転してしまう問題点が存在する。

内蔵センサ以外にも新たな構成を追加して把持姿勢認識を行う研究 [3][4] も存在する。iRotate[3] は端末に搭載されたカメラからの映像を用いて顔認識を行うことでユーザの姿勢を考慮した把持姿勢を認識している。ユーザの顔の向きに合わせた画面の提示ができるが、端末の動きによってカメラ画像にブレが発生した場合や前方カメラを指で覆ってしまった場合、顔がカメラの画角に入らない場合などには正確な顔認識が難しい。iRotateGrasp[4] 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フェースを提案している。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るに利用している。ユーザの姿勢を考慮して把持姿勢認識を行い、それに伴う画面の表示は可能であるが、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。

ジェスチャ認識による画面の提示システム [5][6] も提案されている。Ording ら [5] は、二本の指でスクリーンの上で回転するようなジェスチャを行うことで画面を回転させるインタフェースを構築した。また、Sensor synaesthesia[6] ではユーザがスクリーンをタッチしながら端末を回転させると、画面の表示を保持する操作を実現している。しかし、画面の表示を変えるために、ユーザは意識的なタップ操作やスライド操作などの新しいジェスチャを学ぶ必要がある。

本研究では、ユーザの姿勢を考慮した携帯端末の把持姿勢を認識するため、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。特に、新しいジェスチャではなく、自然に携帯端末のスクリーンをタップした時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。

3. 提案手法

本研究では、携帯端末の内蔵センサのみを用いて、把持姿勢を認識する。システムの概要を図2に示す。具体的に

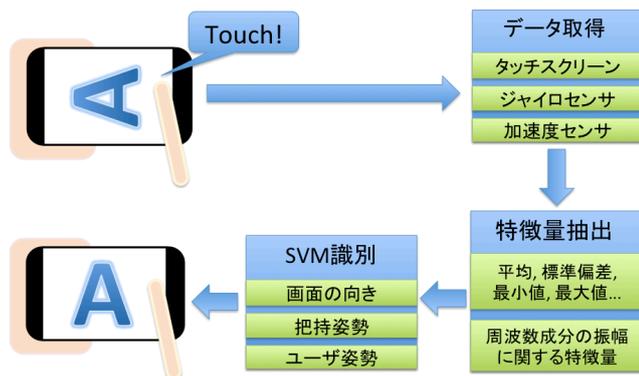


図 2 システムの概要
Fig. 2 System Configuration

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。特徴量は文献 [15] で用いたものを基本的に利用しているが、本システムでは 1 回のタッチでユーザの姿勢と端末の把持姿勢を認識するため、連続するタッチに関するデータは考慮しない。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する。

3.1 ユーザの姿勢を考慮した把持姿勢

ユーザの姿勢として、椅子に座っている状態 (座位状態) および体の右側 (左側) を下にしてソファに寝ている状態 (側臥位状態) を、端末の把持姿勢として、右手 (左手) で端末を持ち、左人差し指 (右人差し指) で画面を操作する状態を考慮して、図 3 に示した 12 通りの組み合わせを提案システムで認識すべき、ユーザの姿勢を考慮した端末の把持姿勢とする。側臥位状態では、端末の持ち易さを考慮し、体の下側になる手で端末を持つこととしている。

例えば、把持姿勢 R-Po1 は、右手 (R: Right) による画面操作を行う状況で、端末はユーザから見て縦画面 (Po: Portrait), ユーザは座位状態 (1) であることを示し、L-LaR2 は、左手 (L: Left) による画面操作を行う状況で、端末はユーザから見て横画面 (縦画面の状態から端末を右に 90 度回転した状態 LaR: Landscape Right), ユーザは側臥位状態 (2) であることを示している。

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として、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、タッチイベントの前後 0.2 秒間のデータを記録する。

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

タッチスクリーンからのデータは、端末操作中の自然なタップ操作 1 回で把持姿勢を認識するため、ユーザがタッチした座標とタッチ面積の 3 次元を特徴量として用いた。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 [9][10][12][13][14] 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、行動認識や把持姿勢認識などの分野においてもよく利用されている [7][8][9][10]。我々の研究グループでは、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て把持姿勢認識に有効に寄与していることが分かり [15]、本研究でも用いることとした。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 3 次元、内蔵センサ 1 つに対して 61 次元の特徴量となる。表 1 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 186 次元である。

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。実装では、Python で記述された SVM ライブラリの一つである LIBSVM [11] 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

ユーザ姿勢	縦画面(Po)		横画面(LaL)		横画面(LaR)	
	右手操作(R)	左手操作(L)	右手操作(R)	左手操作(L)	右手操作(R)	左手操作(L)
座位状態						
側臥位状態						

図 3 ユーザの姿勢を考慮した把持姿勢

Fig. 3 Grasp postures based on user's positions

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 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適用した。SVM の判別により 12 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(モデル名: SC-01F) 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してデータを取得する。OS は Android OS バージョン 5.0, 重さは 168 グラム, 画面サイズは 5.7 インチ (解像度: 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認識は 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認識精度を検証するために、表 1 に示す特徴ベクトルを用いて認識実験を行った。被験者は 23~25 歳の男性 5 名で、12 種類の端末把持姿勢におけるタップ操作時のデータを取得した。取得したデータに基づいて提案手法により把持姿勢を認識する。具体的には、実験 1 では各被験者のデータごとに交差検定を行うことで、個人使用環境における評価を、実験 2 では全被験者のデータを使用して交差検定を行うことで、複数人使用環境における評価を行う。

4.1 実験タスク

実験では、まず、被験者に 12 種類の把持姿勢について説明し、図 4 に示すデータ記録用アプリケーションを用いて、各把持姿勢におけるデータを取得する。実験アプリ

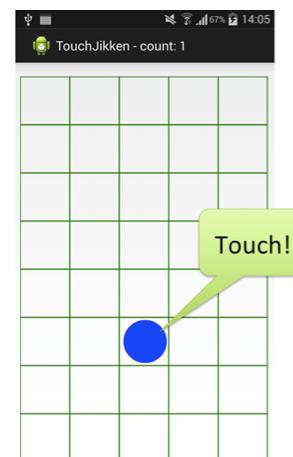


図 4 実験アプリケーション

Fig. 4 An application for experiments

ケーションでは、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。被験者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、同様の処理を繰り返す。

以上のタスクを、各把持姿勢について 5 回実施した。把持姿勢を変更する際には、適宜休憩時間をとった。すべての被験者が各タスクを行うことで得られるデータの数は、12000 個のタッチデータ (12 種類の把持姿勢 × 40 箇所のセル × 5 回タッチ × 5 被験者) となる。

4.2 実験 1: 個人使用環境における把持姿勢認識精度

個人使用環境におけるユーザの姿勢を考慮した把持姿勢認識精度を検証する。各被験者の 2400 個のタッチデータに対して、5-分割交差検定を行った。

表 2 に被験者ごとの特徴ベクトルによる認識精度を示す。どの被験者も特徴ベクトル D が一番高い認識精度を示している。また、特徴ベクトル B, C, D が特徴ベクトル A

表 1 特徴ベクトル
Table 1 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			×	×	64
B				×	125
C			×		125
D					186

表 2 実験 1 の結果: 個人使用環境における把持姿勢認識精度

Table 2 The result of Experiment1: Grasp recognition accuracy in personal use environment

被験者	認識精度 (%)			
	特徴ベクトル			
	A	B	C	D
被験者 1	88.08	99.71	99.0	99.75
被験者 2	85.5	99.88	99.08	99.96
被験者 3	83.58	99.88	97.5	99.88
被験者 4	92.13	100	98.0	100
被験者 5	85.29	100	96.75	100

より高い認識精度 (96.75~100%) 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。

文献 [15] では、ジャイロセンサの値と線形加速度の値は端末の動きに関する情報で、加速度センサの値は端末の傾きに関する情報であると考えられ、利用環境により重要な特徴量がそれぞれ異なった。

一方、ユーザの姿勢を考慮する場合には、ジャイロセンサと線形加速度の値だけでなく、重力加速度の情報が含まれている加速度センサの値も把持姿勢認識に有効に作用したと考えられる。

4.3 実験 2: 複数人使用環境における把持姿勢認識精度

複数人使用環境におけるユーザの姿勢を考慮した把持姿勢認識精度を検証する。各被験者の 2400 個のタッチデータ (各把持姿勢に対して 5 回試行) から 480 個のタッチデータ (1 試行) を用いて、合計 2400 個のデータ (480 個 × 5 被験者) を用いて 5-分割交差検定を行った。表 3 に特徴ベクトルによる認識精度を示す。

複数人使用環境では、前節の個人使用環境の場合と比較して認識精度が低下した。特に、特徴ベクトル A と C の場合、より顕著に低下している。端末の動きに関する情報から把持姿勢を認識する場合、複数人のデータを用いると把持姿勢を認識することが難しい。一方、特徴ベクトル B、D の場合は複数人使用環境においても高い認識精度を示している。ユーザの姿勢を考慮した把持姿勢認識では、端末の動きと向きに関する情報が有効に作用し認識精度を高くしているものと考えられる。

最も認識率の高かった特徴ベクトル D を用いた場合の把持姿勢認識に関する混同行列を表 4 に示す。ユーザの

表 3 実験 2 の結果: 複数人使用環境における把持姿勢認識精度

Table 3 The result of Experiment2: Grasp recognition accuracy in general use environment

把持姿勢	認識精度 (%)			
	特徴ベクトル			
	A	B	C	D
12 種類	50.79	97.08	77.0	97.71

姿勢を区別すると 12 種類の把持姿勢となるが、端末に表示する画面の向きを基準に区別すると図 2 に示すように、「Portrait」、「Landscape Left」、「Landscape Right」の 3 種類の画面表示の中で、どの表示を選択すればよいか判定できればよい。つまり、各画面モードの 4 種類の把持姿勢を同一であると考えることが可能で、その認識精度は 99.7% の認識精度となった。

5. おわりに

ユーザは携帯端末をいつでもどこでも利用できるが、姿勢を変えることで画面が意図せず回転してしまい、端末操作が中断される問題点があった。本稿では、ユーザの姿勢を考慮した把持姿勢を認識するため、携帯端末の内蔵センサである加速度センサとジャイロセンサ、タッチスクリーンから得られるデータを用いて、12 種類の把持姿勢を SVM により認識する手法を提案した。特徴ベクトル D による認識精度は、個人利用環境において 99.75~100%、複数人 (5 人) 利用環境において 5-分割交差検定を行った結果、97.71% となった。特に、加速度センサの値に関する特等量がユーザの姿勢を考慮した把持姿勢認識に有効であることが分かった。

今後は、リアルタイム上で動く画面自動切り替えシステムを作り、オンラインでの認識実験を行う予定である。

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] Bartlett, J.F.: Rock 'n' Scroll is Here to Stay. IEEE Computer Graphics and Applications, 40-45, 2000.
- [2] Hinckley, K., Pierce, J., Sinclair, M., and Horvitz, E.: Sensing techniques for mobile interaction. In Proc. UIST'00, 91-100, 2000.

表 4 実験 2: 特徴ベクトル D による把持姿勢認識の混同行列
Table 4 Confusion Matrix on grasp recognition using features vector D

把持姿勢	Portrait				Landscape_Left				Landscape_Right			
	R-Po1	R-Po2	L-Po1	L-Po2	R-LaL1	R-LaL2	L-LaL1	L-LaL2	R-LaR1	R-LaR2	L-LaR1	L-LaR2
R-Po1	193	0	7	0	0	0	0	0	0	0	0	0
R-Po2	0	200	0	0	0	0	0	0	0	0	0	0
L-Po1	9	0	191	0	0	0	0	0	0	0	0	0
L-Po2	0	0	0	198	0	0	0	0	2	0	0	0
R-LaL1	0	0	0	0	192	0	8	0	0	0	0	0
R-LaL2	0	0	0	0	0	200	0	0	0	0	0	0
L-LaL1	0	0	0	0	9	0	191	0	0	0	0	0
L-LaL2	0	0	0	0	0	0	0	199	0	1	0	0
R-LaR1	0	0	0	2	0	0	0	0	191	0	7	0
R-LaR2	1	0	1	0	0	0	0	0	0	198	0	0
L-LaR1	0	0	0	0	0	0	0	0	8	0	192	0
L-LaR2	0	0	0	0	0	0	0	0	0	0	0	200

- [3] Cheng, L.-P., Hsiao, F.-I., Liu, Y., and Chen, M.Y.: iRotate: Automatic Screen Rotation based on Face Orientation. In Proc. CHI '12, 2203-2210, 2012.
- [4] 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, pp.3051-3054, 2013.
- [5] Ording, B., Van Os, M., and Chaudhri, I. Screen Rotation Gestures on a Portable Multifunction Device. US Patent 7978182, 12 July, 2011
- [6] Hinckley, K., and Song, H. Sensor synaesthesia: touch in motion, and motion in touch. Proceedings of the 2011 annual conference on Human factors in computing systems, ACM (2011), 801-810.
- [7] Ono, M., Shizuki, B., and Tanaka, J.: Touch & Activate: Adding Interactivity to Existing Objects using Active Acoustic Sensing. In Proc. UIST'13, ACM, pp.31-40, 2013.
- [8] Bao, L., and Intille, S.S.: Activity recognition from user-annotated acceleration data. In Proc. the 2nd International Conference on Pervasive Computing, pp.1-17, 2004.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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, IEEE, pp32-41, 2008.
- [13] Van Laerhoven, K., and Cakmakci, O.: What shall we teach our pants? Wearable Computers, The 4th International Symposium, IEEE, pp77-83, 2000.
- [14] Schmidt, A., Aidoo, K., Takaluoma, A., Tuomela, U., Van Laerhoven, K., and de Velde, W.: Advanced Interaction in Context. the 1st International Symposium on Handheld and Ubiquitous Computing. Springer Berlin, Heidelberg, pp.89-101, 1999.
- [15] 朴 燦鎬, 小川 剛史: 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, 情報処理学会論文誌 (デジタルコンテンツ), Vol. 4, No. 1, 2016 (掲載予定).