

ユーザの身体的特徴情報を用いた 行動認識モデルの学習手法

前川 卓也^{1,†1,a)} 渡部 晋治^{1,†2}

受付日 2011年10月7日, 採録日 2012年4月2日

概要: 健康管理, 見守り, ライフログングといったさまざまな実世界アプリケーションの実現は, 高齢化社会や生活習慣病といった現代社会がかかえるさまざまな問題に対処するうえで重要である. 特に, 実環境を観測するセンサを用いた行動認識技術は, これらの実世界アプリケーションの基盤的技術の1つとされている. 本論文では, エンドユーザの教師ラベルあり/ラベルなしセンサデータを用いることなく, そのエンドユーザに応じた行動認識モデルを構築する手法を提案する. 提案手法は, 性別や身長といったユーザの身体的特徴情報を用いることで, あらかじめ用意した多数の他のユーザのセンサデータから, エンドユーザのセンサデータと類似しているであろうデータを, 機械学習のアプローチを用いて発見する. そして, その類似ユーザのラベルありセンサデータを用いて, エンドユーザの行動モデルを学習する. すなわち, エンドユーザは, 自身のセンサデータを集めたり, そのデータにラベルを付与したりする必要がない. 評価実験では, 40人の実験参加者から収集した約100時間のセンサデータを用いて提案手法の有効性を検証した.

キーワード: センサ, Web 検索, クエリフリーサーチ, 日常物

Learning Activity Recognition Models with Physical Characteristics Data of Users

TAKUYA MAEKAWA^{1,†1,a)} SHINJI WATANABE^{1,†2}

Received: October 7, 2011, Accepted: April 2, 2012

Abstract: Realizing real world applications such as supporting the care of the elderly, fitness monitoring, and lifelogging is important to solve various problems of modern society such as the aging of society and adult diseases. In particular, activity recognition techniques using sensors that monitor real environments is one of the most important tasks in pervasive computing applications. This paper proposes a new activity recognition method that constructs activity models specific to an end user without need for her labeled/unlabeled sensor data. The method finds other users whose sensor data prepared in advance may be similar to those of the end user by employing information about the end user's physical characteristics such as height and gender. Then, we model the end user's activities by using the labeled sensor data from the similar users. Therefore, our method does not require the end user to collect and label her training sensor data. We confirmed the effectiveness of our method by using 100 hours of sensor data obtained from 40 participants.

Keywords: sensor, web search, query free search, daily object

1. はじめに

高齢化社会や生活習慣病に関する関心の高まりから, 実環境に設置されたセンサを用いて人の生活を観測・理解する研究がさかんに行われている. 特に, センサを用いた行

¹ NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories, Souraku, Kyoto
619-0237, Japan

^{†1} 現在, 大阪大学大学院情報科学研究科

^{†2} 現在, Mitsubishi Electric Research Laboratories (MERL)

^{a)} maekawa@ist.osaka-u.ac.jp

動認識技術は、健康管理、ライフログ、見守りといったさまざまな実世界アプリケーションの基盤的技術の1つである。たとえば、遠隔地の高齢者が行っている行動をセンサにより自動的に認識することで、コストの低い見守りのアプリケーションを実現することなどが期待される。人の行動は、その人の身体の部位の動きに関係することから、身体に装着した加速度センサを用いて身体のさまざまな部位の特徴的な動きを観測し、行動認識を行う研究が数多く行われている [1], [11]。たとえば、人が「歩く」行動をするとき、手と足に周期的な動きが見られるだろう。本論文でも、同様に身体に装着した加速度センサに注目する。しかし、これまでの多くの研究は教師ありの機械学習手法を用いて、行動を認識するモデルを学習しているため、エンドユーザがラベルありトレーニングデータを自身で用意する必要があった。また、他の研究では、エンドユーザに応じた行動モデルを、そのユーザの少量のセンサデータと、他のユーザのセンサデータを用いて構築している。しかし、多くの場合、エンドユーザによるデータ収集やラベリングを必要とする。

本論文で提案する手法は、エンドユーザのラベルありまたはラベルなしセンサデータを用いずに、エンドユーザに応じた行動モデルを構築する。ここで、あるエンドユーザ（以降では目標ユーザと呼ぶ）の行動モデルを構築するとする。このとき、数多くのその他のユーザ（以降では元ユーザと呼ぶ）のラベルありトレーニングデータをあらかじめ用意しておく。提案手法は、目標ユーザと元ユーザの性別や身長といった身体的特徴情報を用いて、あらかじめセンサデータを用意した元ユーザから、目標ユーザのセンサデータと類似しているであろうデータを持つ元ユーザを発見する。そして、その類似元ユーザのラベルありセンサデータを用いて、目標ユーザの行動モデルを学習する。たとえば、目標ユーザの「歩く」の行動モデルを構築するとき、目標ユーザの「歩く」行動データに類似しているであろう元ユーザを、それらの身体的特徴情報を用いて発見する。そして、目標ユーザの「歩く」の行動モデルを、類似元ユーザの「歩く」行動データを用いて学習する。

一般的に、多量のトレーニングデータを用いることで、良質の行動モデルを学習できる。これは、多くのトレーニングデータを用いることで、さまざまなパターンの行動を学習できるためである。しかし、ある目標ユーザの行動モデルを構築する際、目標ユーザのセンサデータと大きく異なる元ユーザのセンサデータを用いることで、それらのセンサデータの不適合により、目標ユーザの行動認識精度に悪影響を及ぼすと考える。たとえば、右利きのユーザの「歯磨き」行動のセンサデータは、左利きのユーザのものと大きく異なるだろう。そこで提案手法は、ある程度選定したトレーニングデータからモデルを学習する。さらに、目標ユーザのテストデータを認識する際、学習されたモデ

ルをテストデータに適応させることにより、より高精度な認識を目指す。以上のように、目標ユーザは自身の身体的特徴情報のみしか用意する必要がないため、目標ユーザにかかる負担は小さい。

2. 関連研究

エンドユーザによるラベリングの負担を低減するための研究について紹介する。Ohmuraらは、maximum-likelihood linear regression (MLLR) や maximum a posteriori (MAP) 適応 [4], [6] を用いて、エンドユーザの少量のセンサデータだけでエンドユーザに適応した行動モデルを構築している [9]。Forsterらは、認識エラーを自動的に検知するための脳波計を備えた適応的なジェスチャー認識システムを実現している [3]。検知された認識エラーの情報を用いて、自動的に認識モデルを再トレーニングすることを特徴とする。Stikicらは、アクティブ学習を用いて、ラベリングの負担を軽減しつつ、高い認識精度を達成している [13]。Huynhらは、生成モデルと分類モデルを組み合わせることで、ラベリングのコストを低減させている [5]。具体的には、生成モデルを用いてラベルなしセンサデータをクラスタリングしたあと、少ない数のラベルから構築した分類モデルを組み合わせることによりクラスタリングの精度を向上させる。Liuらは、ジェスチャー認識のタスクにおいて、過去に得られたジェスチャーのテンプレート群の中から、ユーザが行った短時間のジェスチャーと類似したものを選び、その類似テンプレートをジェスチャー認識に用いる手法を提案している [7]。このように、典型的な動作をユーザに実施させることで、認識精度を向上させることができるだろう。一方、提案手法ではその代わりに身体的特徴の入力をユーザに求めている。典型的な動作を実施させる手法の方が高い精度を達成できることも考えられるが、実施のコストは身体的特徴の入力に比べて大きいだろう。また、情報通信技術に不慣れなユーザにもその動作を正しく実施させる必要がある。動作が正しく実施されなかった場合、逆に精度は下がると考えられる。また、ユビキタスセンサを用いた行動認識研究においても、ラベリングの負担の低減を目的とするものは多い。van Kasterenらは、家に設置された多数のユビキタスセンサを想定し、転移学習のアプローチにより他の家から得られたラベルありセンサデータを用いて対象とする家の行動モデルを構築している [14]。一方、本論文で提案する手法は、ユーザのラベルあり/なしセンサデータを用いる代わりに身体的特徴情報を用いることで、ユーザの負担を減らすことを特徴とする。我々の知る限り、ユーザの身体的特徴情報を用いて行動のモデルを構築する手法はない。

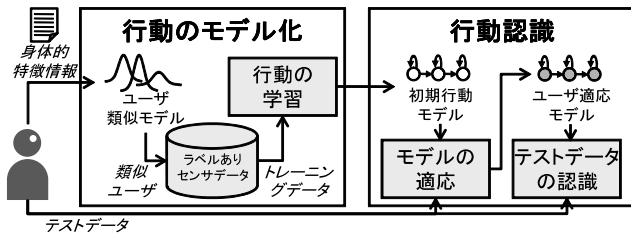


図 1 行動のモデル化と認識の概要

Fig. 1 Architecture of our recognition method.

3. 提案手法

3.1 提案手法の概要

提案手法は、「前準備」、「行動のモデル化」、「行動認識」の3つの手順から構成される。「前準備」では、まず、元ユーザからあらかじめ取得したラベルありセンサ（加速度）データから、元ユーザの行動間の類似度を計算する。たとえば、元ユーザ A と B の「歩く」行動間の類似度（行動類似度）を計算する。この計算を、ユーザ A と B のすべての行動のクラスにおいて行う。さらに、すべての組合せの2人の元ユーザ間においても類似度計算を行う。次に、それぞれの行動クラスにおいて、行動類似度とユーザの身体的特徴から求めた属性の関係を学習する。すなわち、2人のユーザの身体的特徴情報から求めた属性を用いて、それらのユーザのある行動クラスにおける行動類似度を推定するモデルを学習する。以降では、このようなモデルを「ユーザ類似モデル」と呼ぶ。たとえば、「歩く」行動のユーザ類似モデルは、2人のユーザの性別、身長、体重などの情報を用いて、その2人のユーザの「歩く」行動の類似度を推定する。身体的特徴情報は、身長や年齢などのユーザの基本的な情報を含むとする。また、「卓球をする」行動の利き手や「皿を洗う」行動の日常生活における頻度といった、認識したい行動に関するユーザの情報も含むとする。

上記の手順は目標ユーザの行動のモデル化および認識のための前準備である。図 1 に、「行動のモデル化」と「行動認識」の手順の概略を示す。「行動のモデル化」では、行動のクラスごとに、そのクラスに対応するユーザ類似モデルを用いて、目標ユーザの行動（センサデータ）に類似しているであろう元ユーザを求める。そして、そのクラスの行動モデルを、類似元ユーザのラベルありセンサデータを用いて学習する。このように構築されたモデルを、「初期行動モデル」と呼ぶ。「行動認識」では、まず目標ユーザのラベルなしテストデータを用いて、初期行動モデルを目標ユーザに適応させる。そして、適応させたモデル（ユーザ適応モデル）を用いてテストデータの認識を行う。3つの手順の詳細を以下に示す。

3.2 前準備

行動のクラスごとに、2人のユーザの身体的特徴情報を

ユーザ (基底- 対象)	年齢				性別			行動 類似度
	基底 ユーザ	対象 ユーザ	差	比	基底 ユーザ	対象 ユーザ	差	
A-B	29	32	3	1.1	男	男	0	0.3
A-C	29	50	21	1.7	男	男	0	0.1
...								
Z-Y	36	36	0	1.0	男	女	1	0.01

属性 答え

図 2 ある行動クラスにおける属性と行動類似度のペア群

Fig. 2 Attributes and activity similarities of a certain activity class.

用いて、それらのユーザの行動類似度を推定するモデルを構築する。そのため、2人のユーザの身体的特徴情報から、行動類似度を推定するための属性を計算する。図 2 に示すように、2人の元ユーザのそれぞれの組合せごとに、属性と答え（2人の行動の行動類似度）のペアを求める。たとえば、図 2 の最初のペアでは、元ユーザ A を基底ユーザ、元ユーザ B を対象ユーザとし、それらの身体的特徴情報から属性を、ラベルありセンサデータから行動類似度を計算する。そして、計算したペア群を用いてユーザ類似モデルを学習する。以下に手順の詳細を説明する。

3.2.1 属性の計算

2人のユーザの身体的特徴情報から属性を計算する。ここでは、数値とクラスの2種類の身体的特徴情報を想定している。数値の情報は、ユーザの身長、体重、年齢などの情報を含む。2人のユーザの数値情報からは4つの属性が計算される。たとえば、基底ユーザ A と対象ユーザ B の年齢の情報からは、基底ユーザの年齢、対象ユーザの年齢、対象ユーザと基底ユーザの年齢の差、基底ユーザと対象ユーザの年齢の比が図 2 に示すように計算される。このような属性を、体重や年齢などの数値情報ごとに計算する。

クラスの情報は、ユーザの性別、利き手、スポーツ経験などの情報を含む。たとえば、性別の情報は「男」や「女」などの値を、スポーツ経験の情報は「あり」や「ある程度」や「なし」などの値を持つ。2人のユーザのクラス情報からは3つの属性が計算される。たとえば、基底ユーザ A と対象ユーザ B の性別の情報からは、基底ユーザの性別、対象ユーザの性別、対象ユーザと基底ユーザの性別の差が図 2 に示すように計算される。差を求める際、最大の値が1、最低の値が0となるように正規化した数値をクラスの値に対応させ、それらの差を求める。たとえばスポーツ経験の場合、「あり」を1、「ある程度」を0.5、「なし」を0とし、差を求める。このような属性を、利き手やスポーツ経験などのクラス情報ごとに計算する。

3.2.2 行動類似度の計算

元ユーザのセンサデータから特徴抽出を行ったあと、抽出した特徴を用いて元ユーザ間の行動類似度を計算する。既存の行動認識研究における手法を基に特徴抽出を行う。

ここでは時系列の加速度データを想定しているため、スライディング時間窓ごとに特徴ベクトルを求める。64 サンプルの時間窓内で求めたFFT成分を基に特徴を計算する。特徴として、平均、エネルギー、主要周波数を用いた。平均はセンサ（身体の部位）の姿勢を表す特徴である。エネルギーは、動きの強度を表す特徴であり、立った状態などと、歩行などの強度の異なる行動を区別できる [1], [15]。エネルギーは、FFT成分ごとの振幅の二乗を合計し、さらにウィンドウサイズで除算することで正規化したものである。ただし、その合計からFFTの直流成分は除いている。主要周波数は、最大の値を持つFFT成分の周波数であり、異なる周波数の周期的動作を区別するために用いる [8]。身体に装着されたすべての加速度センサのすべての軸のデータから上記のようにして抽出された特徴を連結し、特徴ベクトルを作成する。

元ユーザ間の行動類似度は、GMM (Gaussian mixture model) を用いて計算する。元ユーザ A と B の「歩く」行動の行動類似度を計算するとする。ユーザ A を基底ユーザ、ユーザ B を対象ユーザとし、対象ユーザの「歩く」行動のセンサデータと、基底ユーザの「歩く」行動のGMM間の類似度を計算する。まず、ユーザ A の「歩く」行動のセンサデータから抽出された特徴ベクトルを用いて、ユーザ A の「歩く」行動をGMMによりモデル化する。GMMのパラメータ推定にはEMアルゴリズムを用いた [2]。そして、対象ユーザ B の「歩く」行動のセンサデータから抽出された特徴ベクトルごとにGMM尤度を求める。そして、対象ユーザ B の「歩く」行動のすべての特徴ベクトルの尤度の平均を、基底ユーザ A と対象ユーザ B の「歩く」行動の行動類似度とする。上記の処理を、2人の元ユーザのすべての組合せ、および、すべての行動クラスにおいて行う。

3.2.3 ユーザ類似モデルの学習

上述したように、2人の元ユーザの身体的特徴情報とセンサデータから、属性とある行動の行動類似度のペアを作成する。そして、2人の元ユーザのすべての組合せにおいてペアを作成し、そのペア群をトレーニングデータとして回帰のためのSMO (sequential minimal optimization) アルゴリズム [12] を用いてユーザ類似モデルを学習する。そして、すべての行動クラスごとにモデルを学習する。このモデルにより、2人のユーザの身体的特徴情報から、その2人のユーザのある行動クラスにおける行動類似度を推定できる。

3.3 行動のモデル化

行動のクラスごとに、その行動を認識するためのモデルを構築する。そのために、まず目標ユーザは自身の端末（携帯電話など）を用いて自身の身体的特徴情報を入力とする。そして、その身体的特徴情報とユーザ類似モデル

により、行動モデルの学習に用いるラベルありセンサデータを自動的に発見する。ここで、トレーニングデータの発見とモデルの学習は計算機（サーバなど）上でを行い、センサデータの認識はユーザの端末上で行うことを想定している。すなわち、ユーザの端末と計算機間で、ユーザが入力した身体的特徴情報や学習モデルのやりとりが何らかの通信手段を用いて行われる。以下ではある行動クラスに注目して、その行動モデルを構築する手順を説明する。

3.3.1 類似元ユーザの発見

注目する行動クラスに対応するユーザ類似モデルを用いて、目標ユーザとそれぞれの元ユーザ間の行動類似度を推定する。そして、推定された類似度を基に、元ユーザのランキングを作成する。

3.3.2 行動の学習

注目している行動クラスに関する上位 n 類似元ユーザのラベルありセンサデータを用いて、目標ユーザの行動モデルを構築する。最適な n の値に関しては、評価実験において検証する。観測変数が特徴ベクトルに対応する left-to-right HMM (hidden Markov model) を用いて注目している行動クラスのモデル化を行う。このとき、HMMの各状態の出力確率を混合ガウス分布を用いて表現するとする。特徴ベクトルの計算方法に関しては、3.2.2項を参照してほしい。また、HMMのパラメータ推定にはBaum-Welchアルゴリズム [10] を用いた。以降では、類似元ユーザのラベルありデータから作成されたモデルを初期行動モデルと呼ぶ。そして、認識したい行動のクラスごとに、初期行動モデルを学習する。

3.4 行動認識

目標ユーザから取得したラベルなしテストデータを、その目標ユーザの行動モデルを用いて認識する。すなわち、それぞれの時間窓ごとに特徴ベクトルを抽出し、それらを適切な行動クラスに分類する。その前に、テストデータを用いて初期行動モデルを目標ユーザに適応させる。そして、そのユーザ適応モデルを用いてテストデータの認識を行う。ここでは、HMM内のガウス分布の平均を線形変換するMLLR適応を行う。すなわち、初期行動モデル (HMM) の出力分布をテストデータの分布に合うようシフトする。適応された分布の平均 $\hat{\mu}$ は以下のように表される。

$$\begin{aligned}\hat{\mu} &= \mathbf{A}\mu + \mathbf{b} \\ &= \mathbf{W}\xi.\end{aligned}$$

ここで、 μ は初期モデルの分布の平均ベクトル、 \mathbf{A} は $k \times k$ の変換行列、 k は特徴ベクトルの次元数、 \mathbf{b} はバイアスベクトル、 \mathbf{W} は $k \times (k+1)$ の変換行列であり、 $\mathbf{W} = [\mathbf{b} \ \mathbf{A}]$ のように分解される。また、 ξ は拡張平均ベクトルである $\xi = [\mu_1 \ \mu_2 \ \dots \ \mu_k]^T$ 。初期モデルとテストデータの mismatches を減らすような \mathbf{W} をEMアルゴリズムを用いて推

定する。最も単純な MLLR 適応の方法としては、あるユーザのすべての行動モデルに含まれるすべてのガウス分布をシフトする 1 つの変換行列を推定し、その行列を用いてすべてのガウス分布をシフトする方法が考えられる。

ここで、初期モデルをより厳密に適応させるため、それぞれのモデル (HMM) に含まれるガウス分布やガウス分布のグループごとにガウス分布の平均を変換させることが考えられる。その方法について簡単に説明する。目標ユーザのテストデータにはラベルが付与されていないため、まず初期行動モデルを用いてそのテストデータを認識し、その認識結果に基づき厳密な適応を行う。初期行動モデルを用いた認識により、どの特徴ベクトルがどのモデル (HMM) の状態からの出力に対応するか分かる。すなわち、HMM の状態ごとに、その状態の混合ガウス分布とその状態の出力に対応する特徴ベクトル群との mismatches を小さくする変換行列を求めることができる。しかし、初期行動モデルを用いた認識結果はある程度のエラーを含むことが多い。また、それぞれの状態に対応するテストデータが少ない場合、その少ないデータに対して過剰適合を起こしてしまうため、認識精度の低下を引き起こす。

そこで、より柔軟な適応を行うため、音声認識の分野などで用いられている回帰木を用いて適応を行う。回帰木を用いることで、テストデータの量に応じて適応の対象とするガウス分布のグループを自動的に決定することができる。まず、すべての初期行動モデルに含まれるガウス分布をそれぞれのノードで 2 つの群にクラスタリングするように二分回帰木を構成する。木の葉ノードはガウス分布の最終的なグループに対応する。図 3 は、二分回帰木の例である。根ノード G_1 は、目標ユーザのすべての初期行動モデルのガウス分布に対応し、根ノードはそのガウス分布を 2 つのグループ G_2 と G_3 に分割する。このとき、分布間のユークリッド距離を用いて分割を行う。そして、指定した数の葉ノードが作成されるまで分割を繰り返す。すなわち、類似した分布どうしが階層的にグループ化されるように木が構築される。図 3 の木を用いて説明する。木のそれぞれ

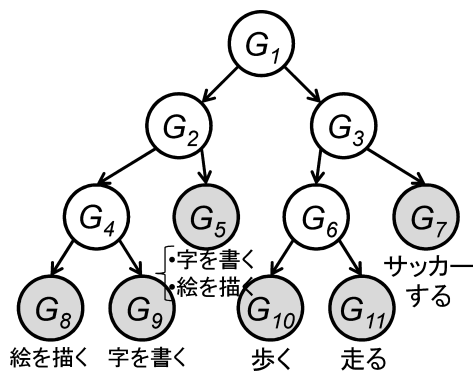


図 3 回帰木の例

Fig. 3 Example binary regression tree.

の葉ノードは、その下部に示されているラベルに対応する初期行動モデルのガウス分布によって占められているとする。たとえば、 G_{10} は「歩く」行動モデルのガウス分布によって占められている。また、類似した行動クラスのガウス分布が同じノードに含まれることもある。たとえば、 G_3 は、「歩く」、「走る」、「サッカーをする」の行動モデルのガウス分布によって占められている。

このような回帰木を用いることで、テストデータの量に応じた柔軟な適応を行うことができる。たとえば、テストデータの量が非常に少ない場合は、根ノードに対応するガウス分布群 (初期行動モデルのすべてのガウス分布群) に着目して、それらを 1 つの変換行列を用いて大まかな変換を行う。一方、十分にテストデータの量がある場合は、より深いノードに着目して、それぞれのノードごとにノードに対応するガウス分布群に対してより厳密な変換を行う。すなわち、テストデータの量に応じて、変換を行うノードの数を柔軟に変えることができる。テストデータの量が多いほど、数多くのノードに対して変換を行えるため、厳密な適応を行うことができる。回帰木を用いた適応のさらなる詳細に関しては、文献 [6] を参照してほしい。

4. 評価実験

40 人の実験参加者から取得した約 100 時間のセンサデータを用いて提案手法の評価を行う。

4.1 データセット

それぞれの参加者には両手首、腰、右太腿に 3 軸加速度センサを搭載した無線センサノードを装着してもらった。これまでの多くの既存研究において、これらの部位に装着した加速度センサの有効性が確かめられているため [1], [5], 本研究でもこれらの部位に装着したセンサノードを用いた。そして、参加者らは、表 1 に示した行動をランダムな順で行うセッションを計 10 セッション行った。表 1 の行動のうち、いくつかについて詳細に説明する。行動 J では、参加者はキーボードでいくつかの文字の入力を行った。行動 K と L では、いくつかの文章をノートもしくは白板に記述した。そして、40 人の参加者から 9871 のラベル付き行動を取得した。

また、参加者に対して、それぞれの身体的特徴情報の記入も行ってもらった。身長、体重、性別といった基本的な

表 1 実験で行われた行動のリスト

Table 1 Activities performed in our experiment.

A	立つ	F	階段を降りる	K	白板に書く
B	歩く	G	自転車をこぐ	L	ノートに書く
C	走る	H	歯を磨く	M	卓球をする
D	座る	I	皿洗いをする	N	掃除機を使う
E	階段を昇る	J	PC を使う		

表 2 実験で用いた参加者の身体的特徴情報

Table 2 Physical characteristics information used in our experiment.

身体的特徴	値	身体的特徴	値
性別	{ 男, 女 }	年齢	数値
身長	数値	体重	数値
利き手 (書き)	{ 右, 左 }	利き手 (卓球)	{ 右, 左 }
利き手 (歯磨き)	{ 右, 両方, 左 }	利き手 (掃除機)	{ 右, 両方, 左 }
卓球経験	{ あり, ある程度, なし }	習字経験	{ あり, なし }
ブラインドタッチの可否	{ 可, ある程度, 不可 }	皿洗いの頻度	{ 毎日, 時々, 稀に, ない }
自転車に乗る頻度	{ 毎日, 時々, 稀に }	掃除機のタイプ	{ キャニスタ, ハンディ }
自転車のタイプ	{ シティサイクル, 折りたたみ }		

身体的特徴情報や、それぞれの行動の利き手などの情報を記入してもらった。また、参加者が実験において用いたモノの情報（自転車や掃除機のタイプなど）についても記入してもらった。実験では、被験者ごとに行動で使うモノのタイプを指定していた。たとえば行動 G では、シティサイクルか折りたたみ自転車のいずれかを使ってもらった。行動 N では、キャニスタ型掃除機とハンディ掃除機のいずれかを使ってもらった。表 2 に、今回の実験で用いた身体的特徴情報を示している。

4.2 評価方法

認識手法を ‘leave-one-participant-out’ 交差検定を用いて評価する。すなわち、1 人の参加者を目標ユーザ、残りの参加者を元ユーザとし、交差検定を行う。評価尺度としてはエラー率を用いた。エラー率は $1.0 - F\text{-measure} = 1.0 - \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$ のように表される。Precision と recall は、それぞれの特徴ベクトルの分類結果から求めたものである。提案手法の性能を検証するため、以下の 7 つの手法を用いて評価を行う。

- ユーザ依存モデル (DPN)：目標ユーザの行動モデルを、そのユーザのラベルありトレーニングデータを用いて学習する。この手法はユーザによるラベル付け作業を必要とする。
- ランダム選択 (RND)：ランダムに選択した元ユーザのラベルありデータを用いて、目標ユーザの行動モデルを学習する。n が 39 のとき、すべての元ユーザのデータを用いて学習する。n が大きいほど、より多くの数の元ユーザのさまざまな行動パターンを行動モデルに学習させることができる。
- ランダム選択と MLLR 適応 (RND+MLLR)：ランダムに選んだ n の元ユーザのデータから行動モデルを学習したあと、MLLR 適応を用いて目標ユーザにモデルを適応させる。具体的には、目標ユーザのすべての初期行動モデルに含まれるすべてのガウス分布をシフトする 1 つの変換行

列を推定し、その行列を用いてすべてのガウス分布をシフトする。

- ランダム選択と回帰木を用いた MLLR 適応 (RND+Tree)：ランダムに選んだ n の元ユーザのデータから行動モデルを学習したあと、回帰木を用いた MLLR 適応により目標ユーザにモデルを適応させる。
- 身体的特徴を用いた選択 (PC)：3.2 節で述べたユーザ類似モデルを用いて n の元ユーザを選択し、そのラベルありデータを用いて目標ユーザの行動モデルを学習する。
- 身体的特徴を用いた選択と MLLR 適応 (PC+MLLR)：ユーザ類似モデルを用いて選んだ n の元ユーザのデータから行動モデルを学習したあと、MLLR 適応を用いて目標ユーザにモデルを適応させる。
- 身体的特徴を用いた選択と回帰木を用いた MLLR 適応 (PC+Tree)：提案手法に対応する。

4.3 類似ユーザの推定精度

行動認識の精度に関して説明する前に、ユーザ類似モデルを用いた類似元ユーザの推定精度に関して簡単に説明する。提案手法は、ある元ユーザと目標ユーザの身体的特徴情報を用いて、それらの行動類似度を推定する。そして、推定した類似度を基に元ユーザのランキングを作成する。ここでは、推定されたランキングと実際のランキング（正解のランキング）を比較し、評価する。正解のランキングは、目標ユーザとそれぞれの元ユーザのセンサデータを実際に比較して類似度を計算することによって求める。評価指標としては、推定された上位 n 類似元ユーザの正解率を用いる。すなわち、上位 n 類似ユーザとして推定された元ユーザのうち、実際に上位 n 位以内であった元ユーザの比率とする。図 4 は、n を変化させたときの提案手法の平均正解率の推移を示す。また、ランダムに上位 n 類似ユーザを選ぶ手法の正解率の推移も示している。たとえば n が 2 のとき、ランダムに類似ユーザを選ぶ手法の正解率は $2/39 = 0.0513$ となる。図ではすべての n において提案

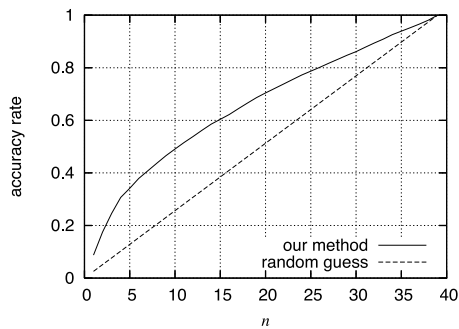


図 4 類似ユーザ推定精度の遷移

Fig. 4 Average accuracy rate of similar user estimation.

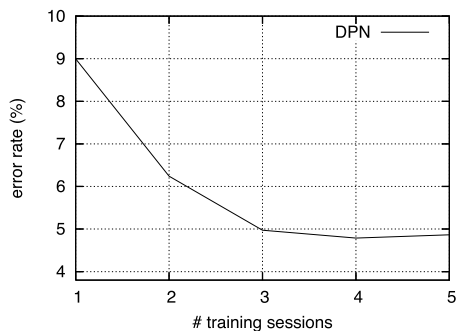


図 5 DPN 手法におけるエラー率の遷移

Fig. 5 Transition of error rates of DPN method.

手法がランダムな手法を上回っている。たとえば、上位 10 類似元ユーザを推定するとき、すなわち n が 10 のとき、提案手法は約 50% の正解率で上位 10 類似元ユーザを推定できていた。

4.4 評価結果

まず、DPN 手法の精度に関して述べる。図 5 は、トレーニングデータとして用いるセッションの数を変化させたときのエラー率の推移を示す。たとえば、トレーニングデータのセッションの数が 2 のとき、DPN は目標ユーザの行動モデルを、そのユーザの最初の 2 セッションのラベルありデータを用いて学習し、残りの 8 セッションのデータで認識を行う。トレーニングデータのセッションの数が 4 のとき、4.79% の低いエラー率を達成している。目標ユーザのラベルありデータなしで、このエラー率を達成することを目標とする。

図 6 に、 n を変化させたときの DPN 以外の 6 手法のエラー率の変化を示す。まず、適応手法を用いていない RND と PC 手法に注目する。基本的に、 n が増加すればエラー率は減少する。これは、トレーニングデータの増加にともない、行動モデルにさまざまな行動パターンを取り込めるためである。 n が 39 のとき RND と PC のエラー率は等しくなるのは、これらの手法がすべての元ユーザのラベルありデータを用いて学習を行っているためである。 n が 39 のとき RND と PC は 6.85% のエラー率を達成しており、

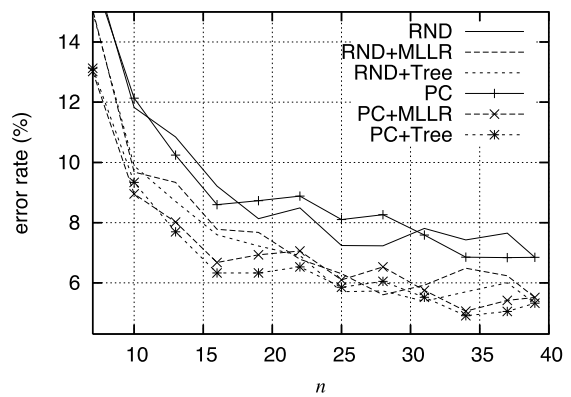


図 6 各認識手法におけるエラー率 (error rate) の遷移

Fig. 6 Transitions of error rates of recognition methods.

RND の最低のエラー率だった。一方、PC の n が 39 より小さいときの最低のエラー率は 6.84% だった。すなわち、PC 手法は、すべての元ユーザのラベルありデータを用いなくても、すべての元ユーザのデータを用いたときと同程度の精度を達成していた。

次に、RND+MLLR と PC+MLLR 手法に注目する。ほとんどの n において、PC+MLLR は RND+MLLR より低いエラー率を達成している。 n が 39 のとき RND+MLLR のエラー率は最も低く、5.52% だった。一方、 n が 34 のとき PC+MLLR のエラー率は最も低く、5.06% を達成しており、RND+MLLR と比べて 0.46% 低かった。PC 手法の最低のエラー率は RND とほぼ同じだったものの、PC+MLLR の最低のエラー率は RND+MLLR に比べて低かった。MLLR 適応を用いることにより、ランダム選択の手法ではエラー率が平均して 1.54%、ユーザ類似モデルを用いた手法では 2.35% 低下していた。すなわち、MLLR 適応により、ランダム選択の手法に比べてユーザ類似モデルを用いた手法のエラー率を効果的に低下させることができていた。これは、ランダム選択の手法によって選ばれた元ユーザの特徴ベクトルに、目標ユーザの特徴ベクトルに比べて大きく異なるものが含まれていたためと考える。目標ユーザの特徴ベクトルと大きく異なる特徴ベクトルから導出されたガウス分布は、分布の適応に悪影響を与えられられる。一方、ユーザ類似モデルを用いた手法が選択した元ユーザの特徴ベクトルには、目標ユーザの特徴ベクトルと大きく異なるものが少なかったと考えられる。すなわち、トレーニングデータを適切に選択することで、適応手法の効果を増大させることができたと考えられる。

最後に、RND+Tree と PC+Tree に注目する。 n が 39 のとき、RND+Tree は最も低いエラー率を達成しており、そのエラー率は 5.32% だった。一方、 n が 34 のとき、PC+Tree は最も低いエラー率を達成しており、そのエラー率は 4.91% だった。これは、RND+Tree と比べて 0.41% 低かった。すなわち、ユーザ類似モデルを用いることでエラー率を 0.41% 低下させることができた。また、PC+Tree の最

も低いエラー率は、PC+MLLR のものと比べて 0.15% 低かった。一方で、PC+Tree は DPN 手法より低いエラー率 (4.79%) を達成することができなかった。しかし、最も単純な手法である RND のエラー率 6.85% に比べて、エラー率が 1.94% 低下していた。すなわち、RND 手法に比べて 28.3% のエラー低減を達成していた ($1.94/6.85 = 0.283$)。

ここで、図 6 に示すように、エラー率の遷移はそれほど安定していない。特に、ランダム選択を用いた手法の遷移は安定していないといえる。すなわち、元ユーザの選択方法にエラー率は大きく影響を受けるといえるだろう。しかし、 n が 39 より小さいとき、ランダム選択を用いたどの手法のエラー率も、 n が 39 のときのエラー率に比べて高くなっている。一方、 n が 39 より小さいときの PC+MLLR と PC+Tree の最低エラー率は、 n が 39 のときのエラー率を下回っている。これは、ユーザ類似モデルを用いて適切に元ユーザを選択することができたためと考える。

表 3 に、PC+Tree の認識結果の混合行列を示す。この行列は特徴ベクトルごとの分類結果に基づいて求めたものである。この行列から、ほとんどの行動が高い精度で認識されていることが分かる。また、「歩く」、「階段を昇る」、「階段を降りる」の行動の識別に誤りが多いことが分かる。これは、これらの行動が非常に類似していたためである。

ここで、実際の行動認識システムにおいて適切な n をどのように決定するかについて簡単に述べる。あらかじめ用意したすべての元ユーザのラベルありデータを用いて交差検定を行うことで、適切な n は簡単に決定できると考える。

4.5 身体的特徴情報の効果

目標ユーザのセンサデータに類似した元ユーザのセンサデータの量が十分でないとき、目標ユーザの行動認識精度も高くはならないだろう。たとえば、「卓球をする」行動において左利きの参加者の数は 8 人だったが、39 のときの RND+Tree 手法におけるそれらの参加者の n がエラー率は 6.78% だった。これは、右利きの参加者 (32 人) のエラー率 4.98% に比べて大幅に悪かった。また、「掃除機を使う」行動において 5 人のユーザがハンディ掃除機を使っていたが、 n が 39 のときのそれらの参加者のエラー率は 7.31% だった。これも、キャニスタ型掃除機を用いた参加者 (35 人) のエラー率 5.06% に比べて大幅に悪かった。前節では、ユーザ類似モデルを用いることでエラー率を低減させることができることを確認した。ここでは、目標ユーザの身体的特徴に応じたユーザ類似モデルの影響について検証する。PC+Tree 手法を用いたとき、ハンディ掃除機を使っていた参加者のエラー率は平均して 1.60% 低減されたが、キャニスタ型掃除機を用いた参加者のエラー率は平均して 0.45% しか低減されなかった。また、PC+Tree 手法を用いたとき、左利きの参加者のエラー率は平均して 1.51% 低減されたが、右利きの参加者のエラー率は平均して

0.34% しか低減されなかった。すなわち、マイナーな身体的特徴を持つ目標ユーザであっても、提案手法はそのユーザの行動をある程度高い精度で認識することができる。

4.6 考察

[センサの装着条件とそのスペック]

本論文では、我々が開発したセンサノードを用いて、ノードの装着条件を固定してセンサデータを収集した。今後、よりロバストな行動推定を実現するためには、トレーニングデータの量を増やす必要があるだろう。しかし、本論文で用いたものと同じセンサを同じ部位に装着した大多数のユーザからセンサデータを収集することは容易ではない。このような問題は、センサの装着条件とスペックを、身体的特徴情報と同様に扱うことで解決できると考える。すなわち、ユーザがトレーニングデータを収集する際、その身体的特徴情報に加えて、センサの装着条件とスペックの情報も収集しておく。これにより、センサの装着条件とスペックも考慮して、学習に用いるトレーニングデータを選択することができるだろう。

[モデルの学習に関する計算コスト]

類似元ユーザの発見には、回帰のための SMO アルゴリズムを用いており、目標ユーザのある行動に関する類似元ユーザを発見するのにかかる時間は、約 10 ミリ秒だった。一方、類似元ユーザを発見したあとに、34 人の類似元ユーザのトレーニングデータを用いて、ある目標ユーザの初期行動モデルを作成するのにかかる時間は、約 500 秒だった。この結果から、元ユーザ (のトレーニングデータ) が定期的に追加されるような環境においても、その追加ごとに目標ユーザの類似元ユーザを低い計算コストで再度発見することができる。そして、前回の処理の結果から類似元ユーザが変更されたときのみ、行動モデルを再学習する。これにより、低い計算コストで目標ユーザのモデルを定期的に再アップデートできるだろう。

[トレーニングデータの追加と精度の低下]

現在は、我々の実験施設で取得したセンサデータのみをトレーニングデータとして用いている。今後、行動認識精度をより高めるには、エンドユーザ (目標ユーザ) の認識結果とセンサデータをトレーニングデータとして追加することが考えられる。このとき、その認識結果にはある程度の誤りが含まれていると考えられるため、逆に認識精度の低下を引き起こす恐れもあるだろう。しかしこのような問題は、ある程度のラベルミスを考慮したモデルの学習手法を用いることで緩和できると考える。たとえば音声認識の分野では、認識結果の信頼度を用いて、信頼性の高い認識結果のみを学習に用いる研究が行われている [16]。

[身体的特徴の設定]

類似元ユーザの発見には、15 種類の身体的特徴に関する情報を用いている。表 2 に示すように、身体的特徴には数

表 3 PC+Tree 手法の混合行列 ($n = 34$). 値はパーセンテージ
 Table 3 Confusion matrix of PC+Tree results. The values are percentages.

	A: 立つ	B: 歩く	C: 走る	D: 座る	E: 階段を昇る	F: 階段を降りる	G: 自転車こぐ	H: 歯を磨く	I: 皿洗いをする	J: PCを使う	K: 白板に書く	L: ノートに書く	M: 卓球をする	N: 掃除機を使う
A	91.0	5	1.5	0	0	0	0.1	0.2	0.1	0.3	0.3	0.1	0.8	0.6
B	0.2	89.8	0.3	1.4	1.1	2	1.1	0.5	0.4	0.3	1	0.2	1	0.6
C	0.1	0.2	99.4	0	0	0	0.2	0	0	0	0	0	0.1	0
D	0	1.4	0	98.1	0	0.1	0	0	0	0.3	0	0	0	0
E	0	10.3	0.2	0.1	88.5	0.2	0	0	0	0	0	0	0.7	0
F	0	9.4	0.7	0	0	88.9	0	0	0	0	0	0	0.9	0
G	0.1	0.2	0.1	0	0	0	99.6	0	0	0	0	0	0	0
H	0	0.1	0	0	0	0	0	96.5	2.6	0	0.7	0	0	0
I	0.1	0	0	0	0	0	0	0.1	99.7	0	0	0	0	0
J	0	0.1	0	0	0	0	0	0	0	94.6	0	5.3	0	0
K	0	0.2	0	0	0	0	0	0.2	0	0	99.3	0.1	0	0.1
L	0.1	0.2	0	0.1	0	0	0	0	0	3.7	0	95.8	0	0
M	0.2	0.4	0.2	0	0	0	0.1	0.6	0.3	0	0.3	0	97.3	0.4
N	0.6	0.6	0	0.3	0	0	0	4.3	1	0.1	0.6	0.1	1.3	91.2

値とクラスの2種類の情報があり, 自身の身体的特徴情報に対応する値を, 自身の端末のキーなどを用いて, 入力・選択することになるだろう. 現在の実装では15項目の情報を入力するだけでよいため, その手間は十分に小さいと考える. しかし, 認識する行動の種類が増えれば, それに応じてユーザが設定しなければならない身体的特徴の量も増える. 設定する身体的特徴の量が増えれば, 行動認識システムを導入する際の障壁となる. このような問題は, 身体的特徴の入力方法を工夫することで, 解決できると考えられる. 表2に示すように, 身体的特徴には2種類の情報があり, その中でもクラスに対応する身体的特徴の数が多く, この特徴の入力の手間を減らすことが必要である. 一方, 実験結果から明らかになったように, 左利きなどの一般的でないクラスの身体的特徴を持つユーザに対して, 提案手法は特に有効だった. すなわち, そのようなユーザに積極的に身体的特徴情報を入力してもらうように指示すればよいだろう. 一方, その入力がなかったユーザの身体的特徴情報には, 一般的な値(右利きなど)をデフォルト値として用いる. これにより, 身体的特徴を設定する全体のコストを大きく下げることができる.

5. おわりに

本論文では, ユーザの身体的特徴情報を用いて, エンドユーザの行動モデルを構築する新しい手法の提案を行った. 提案手法では, そのエンドユーザのセンサーデータと類似していると考えられる他のユーザを, それらのユーザの身体的特徴情報から推定し, 推定されたユーザのラベルあり

データを用いてエンドユーザの行動モデルを学習する. すなわち, 提案手法はエンドユーザのラベルありセンサーデータを必要としない. さらに, 学習した行動モデルをエンドユーザのラベルなしテストデータに自動的に適応させることで, エンドユーザの行動の認識精度を向上させる. 評価実験では, 類似ユーザを適切に選択する手法とモデル適応手法を用いることで効果的に認識精度を向上できることを確認した. さらに, 類似ユーザを選択する手法により, モデル適応手法の効果を増大させることも確認した. 今後は, 普及が進みつつある携帯電話や腕時計, 運動靴などに埋め込まれているセンサから得られたデータなどを用いて, 提案手法の有効性を検証する予定である.

参考文献

- [1] Bao, L. and Intille, S.: Activity recognition from user-annotated acceleration data, *Pervasive 2004*, pp.1–17 (2004).
- [2] Dempster, A., Laird, N., Rubin, D., et al.: Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol.39, No.1, pp.1–38 (1977).
- [3] Forster, K., Biasucci, A., Chavarriaga, R., Millan, J. del R., Roggen, D. and Tröster, G.: On the use of brain decoded signals for online user adaptive gesture recognition systems, *Pervasive 2010*, pp.427–444 (2010).
- [4] Gauvain, J. and Lee, C.: Maximum a posteriori estimation for multivariate Gaussian mixture observations of Markov chains, *IEEE Trans. Speech and Audio Processing*, Vol.2, No.2, pp.291–298 (2002).
- [5] Huynh, T. and Schiele, B.: Towards less supervision in activity recognition from wearable sensors, *Int'l Symp. on Wearable Computers*, pp.3–10 (2006).

- [6] Leggetter, C. and Woodland, P.: Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models, *Computer Speech & Language*, Vol.9, No.2, pp.171–185 (1995).
- [7] Liu, J., Zhong, L., Wickramasuriya, J. and Vasudevan, V.: uWave: Accelerometer-based personalized gesture recognition and its applications, *Pervasive and Mobile Computing*, Vol.5, pp.657–675 (2009).
- [8] Maekawa, T., Yanagisawa, Y., Kishino, Y., Ishiguro, K., Kamei, K., Sakurai, Y. and Okadome, T.: Object-based activity recognition with heterogeneous sensors on wrist, *Pervasive 2010*, pp.246–264 (2010).
- [9] Ohmura, R., Hashida, N. and Imai, M.: Preliminary evaluation of personal adaptation techniques in accelerometer-based activity recognition, *Int'l Symp. on Wearable Computers: Late Breaking Results* (2009).
- [10] Rabiner, L.: A tutorial on hidden Markov models and selected applications in speech recognition, *Proc. IEEE*, Vol.77, No.2, pp.257–286 (1989).
- [11] Ravi, N., Dandekar, N., Mysore, P. and Littman, M.: Activity recognition from accelerometer data, *IAAI 2005*, Vol.20, No.3, pp.1541–1546 (2005).
- [12] Shevade, S., Keerthi, S., Bhattacharyya, C. and Murthy, K.: Improvements to the SMO algorithm for SVM regression, *IEEE Trans. Neural Networks*, Vol.11, No.5, pp.1188–1193 (2002).
- [13] Stikic, M., Van Laerhoven, K. and Schiele, B.: Exploring semi-supervised and active learning for activity recognition, *Int'l Symp. on Wearable Computers*, pp.81–88 (2008).
- [14] van Kasteren, T., Englebienne, G. and Kröse, B.: Transferring Knowledge of Activity Recognition across Sensor Networks, *Pervasive 2010*, pp.283–300 (2010).
- [15] Welk, G. and Differding, J.: The utility of the DigiWalker step counter to assess daily physical activity patterns, *Medicine & Science in Sports & Exercise*, Vol.32, No.9, pp.S481–S488 (2000).
- [16] Wessel, F., Schluter, R., Macherey, K. and Ney, H.: Confidence measures for large vocabulary continuous speech recognition, *IEEE Trans. Speech and Audio Processing*, Vol.9, No.3, pp.288–298 (2001).



渡部 晋治

平成 11 年早稲田大学理工学部物理学卒業，平成 13 年同大学大学院修士課程修了。同年日本電信電話（株）入社。NTT コミュニケーション科学基礎研究所をへて，平成 24 年より Mitsubishi Electric Research Laboratories (MERL) 所属。博士（工学）。音声認識を中心とする音声言語情報処理・機械学習の研究に従事。平成 16 年電子情報通信学会論文賞，平成 18 年日本音響学会独創研究奨励賞板倉記念，電気通信普及財団テレコムシステム技術賞等受賞。ASJ, IEEE, IEICE 各会員。



前川 卓也 （正会員）

平成 16 年大阪大学大学院情報科学研究科マルチメディア工学専攻博士前期課程修了。平成 18 年同大学院情報科学研究科博士後期課程修了。同年日本電信電話株式会社入社。平成 24 年 4 月より大阪大学大学院情報科学研究科准教授。博士（情報科学）。本会平成 22 年度山下記念研究賞等受賞。ユビキタスコンピューティング，Web 情報処理の研究に興味を持つ。ACM，電気学会，日本データベース学会各会員。