情報処理学会論文誌 Vol.56 No.1 23–34 (Jan. 2015)

車両走行時の振動発生モデルに基づく移動手段識別方式

大橋 洋輝[1,a)] 秋山 高行[1,b)] 山本 正明[1,c)] 佐藤 暁子[1,d)]

受付日 2014年3月31日, 採録日 2014年10月8日

概要：本論文では，スマートフォン（スマホ）に標準的に搭載されている GPS 受信機と加速度センサから得られるデータを用いて，スマホ保持者の移動手段を識別する手法を提案する．特に，特徴が類似しており，従来手法では識別が困難であった車とバイクを高精度に識別する手法を提案する．車両走行時に振動が発生するメカニズムをモデル化し，それに応じた特徴量を適切に選択することで，この識別を実現する．評価実験を行ったところ，80%以上の精度で車とバイクの識別が可能であることを確認した．

キーワード：Intelligent Transport Systems（ITS），プローブ交通情報，スマートフォン，移動手段識別，機械学習

Modality Classification Method Based on the Model of Vibration Generation while Vehicles are Running

Hiroki Ohashi[1,a)] Takayuki Akiyama[1,b)] Masaaki Yamamoto[1,c)] Akiko Sato[1,d)]

Received: March 31, 2014, Accepted: October 8, 2014

Abstract: This paper proposes a method for classifying modes of transportation using GPS data and acceleration data, which are easily obtained with smartphones since they are usually equipped with GPS receivers and accelerometers. We especially focus on distinction between cars and motorbikes. The distinction has been a difficult problem owing to their similarity. We achieve this distinction by choosing appropriate features based on a model of vibration generation while vehicles are running. We conducted an experiment to evaluate the classification accuracy and obtained the accuracy of over 80%.

Keywords: Intelligent Transport Systems (ITS), probe traffic information, smartphone, modality classification, machine learning

1. はじめに

近年新興国では都市化にともなって重度の交通渋滞が発生している．これにより，市民の利便性低下だけでなく排ガスの増加による環境負荷増加や騒音問題など，様々な社会問題が引き起こされている．交通渋滞を解消するためには，まず現在の都市内の交通利用状況を広範囲にわたって把握し，それに応じた的確な渋滞解消策を立案することが必要である．とりわけ新興国のいくつかの都市では，バイクの利用率が非常に高いという特有の交通事情があるため，このような都市では車だけでなくバイクの交通利用状況を正確に把握することが重要である．バイクは車の間をぬって走行できること，バイクの方が比較的狭い路地もよく走行すること，時間帯によってバイク専用のレーンができることがあり通行可能路が異なることがあること，などの理由から状況によって車の交通流とバイクの交通流は必ずしも同じではないからである．

交通利用状況を把握するために，日本をはじめとする先進国ではタクシーなどの車両（プローブカー）に搭載された車載 GPS 受信機によって得られる位置情報を利用する方法（タクシープローブシステム）[1], [2] や，VICS（Vehicle Information and Communication System）に代表される

1 株式会社日立製作所中央研究所
Hitachi Ltd., Central Research Laboratory, Kokubunji, Tokyo 185–8601, Japan
a) hiroki.ohashi.uo@hitachi.com
b) takayuki.akiyama.hv@hitachi.com
c) masaaki.yamamoto.af@hitachi.com
d) akiko.sato.td@hitachi.com

© 2015 Information Processing Society of Japan

ような路側のセンサを用いる方法 [3], [4], [5]，またそれらを統合した方法 [6] などが研究され，実用化されている．ところが，タクシーを用いる方法だけでは当然バイクの交通状況を把握することはできない．また路側センサを用いる方法はセンサの設置コストが大きいため，都市内の広範囲をカバーすることは難しい．したがってこれらの方法をそのまま新興国に適用することは困難であるといえる．

そこで本研究グループではスマートフォン（以下，スマホと呼ぶ）を用いて交通情報を生成するシステム（以下，スマホプローブシステムと呼ぶ）を提案している．スマホには通常 GPS 受信機が搭載されており，これを用いてスマホの位置情報を得ることができる．またスマホは車・バイク・歩行など，どんな移動手段を用いているときでも携行可能である．したがって，スマホ保持者が車やバイクに乗って移動しているときの位置情報を収集することで，車やバイクの交通状況を推定することができると期待できる．さらに，スマホは近年新興国においても普及が進んできており [7]，これを用いれば道路インフラ側への投資コストが不要になるというメリットがある．

しかしスマホプローブシステムを実現するためには課題がある．スマホは常時携行されていることが多いため，スマホで収集されるデータの中には様々な移動手段利用時のデータが含まれうる．したがって，スマホで収集する位置情報を利用して車やバイクなどの交通状況を推定するためには，どの時間帯にどのような移動手段を利用していたかを識別する必要がある．そこで本研究では，スマホに標準的に搭載されている GPS 受信機および加速度センサのデータを用いて，スマホ保持者の移動手段を自動で識別する手法を提案する．とりわけ，新興国での交通情報生成システムへの適用のため，車とバイクを高精度に識別する手法について提案する[*1]．

以下ではまず 2 章で，スマホをはじめとする携帯端末に搭載されているセンサを用いた移動手段識別手法に関する先行研究を紹介し，本研究の位置づけを述べる．次に 3 章で本研究の提案手法について説明する．4 章では手法の有効性の評価のために行った実験について紹介し，その評価結果と考察を述べる．最後に 5 章で本論文のまとめを述べる．

2. 先行研究と本研究の位置づけ

本章ではまず 2.1 節で，移動手段識別に関する先行研究について紹介し，それを受けて 2.2 節で本研究の位置づけを述べる．

[*1] 本研究は既報 [8], [9], [10] で発表した内容を再度精査したうえで，新興国であるベトナム/ハノイ市の現地住民を被験者として実施した実験による精度評価の結果を改めて報告するものである．

2.1 先行研究

これまでにも交通行動調査や，ユーザ状態に応じたサービス提供（コンテンツアウェアサービス），人の生活行動を記録するライフログなどへの適用を目指して，携帯端末に搭載されているセンサを用いた移動手段の識別手法の研究が数多くなされてきた．

先行研究において使用されているセンサを見ると，GPS 受信機を用いるもの [11], [12], [13], [14], [15], [16], [17], [18], [19]，加速度センサを用いるもの [20], [21], [22]，その両方を用いるもの [23], [24], [25], [26], [27], [28], [29], [30] があり，携帯端末搭載のセンサを用いた移動手段識別の研究ではほとんどすべての研究で，少なくとも GPS 受信機または加速度センサのどちらかが用いられていることが分かる．これはこれらのセンサが最も一般的にスマホなどに搭載されているためと考えられる．識別対象の拡大や識別の高精度化のために，GPS 受信機や加速度センサに加え，マイクを用いる方法 [23], [26], [31], [32]，電車やバスの路線情報や駅の位置などの GIS（Geographic Information System）情報を用いる方法 [19], [27]，車やバイクなどを保有しているかどうかという情報を用いる方法 [25]，GPS データの欠損を補うために携帯電話基地局からの信号を用いる方法 [28], [30] なども研究されている．

GPS データを用いるものでは，速度およびその統計量（最大速度，平均速度，分散など）が識別のための特徴量として，最もよく用いられている．これは移動手段ごとに走行速度帯が異なることによるものである．これに加え，文献 [11], [13], [16], [17] では，歩行・自転車・車や電車などの乗り物では加速度が異なることに注目し，速度を微分して得られる加速度を用いている．一方文献 [15], [17], [18], [19] では，電車などでは進行方向の変化がなだらかであるのに対し，他の移動手段では急速な変化もありうることに注目し，GPS データから計算できる方向変化を特徴量として用いている．

加速度センサを用いる研究では，加速度の統計量に加え，パワースペクトラムや自己相関関数など周波数領域の特徴量が用いられることが多い．これは特に，歩行やランニングが周期的な運動であることに着目したものである．歩行の際，1 秒間に 2 歩程度歩く人が多いので，周波数変換を行うと 2 Hz 前後の周波数帯に強いピークが現れ，これを検出することで歩行が識別できる．文献 [21], [22], [24] では加速度をさらに詳細に分析し，その鉛直方向成分や水平方向成分を抽出してから利用することで識別精度を高める手法を提案している．

識別手法としては，SVM [17], [26] やベイジアンネットワーク [25]，ニューラルネットワーク [21], [24]，決定木 [11], [28], [30] など様々なものが用いられている．Zhang ら [17] は，まず静止・歩行・自転車・乗り物を速度や進行方向の変化量で識別した後，さらに乗り物の種類に関して

© 2015 Information Processing Society of Japan

SVM による識別を行うという 2 段階のアプローチを用いることで，車・バス・トラム・電車などを 90%程度の識別率で識別可能であることを示した．

2.2　本研究の位置づけ

2.1 節で示したように，これまでにも様々な特徴量および識別手法が提案され，多くの成功を収めているが，車とバイクの識別を扱った研究は文献 [25], [28], [30] などに限られる．一方で 1 章で述べたように，新興国の各都市への適用を考えると，主要な移動手段であるバイクの識別ができることが必須となる．

Feng ら [25] の手法は，90%以上の高い精度で識別が可能である．ところが，Feng ら [25] の手法においては，車やバイクなどを保有しているかどうかという情報が用いられており，これを手入力する手間を考えると，あまりこの手法は実用的であるとはいえない．また，車やバイクの保有状況以外の特徴量としては，速度，加速度，GPS 測位精度などに関する特徴量が使用されている．特に，一定時間内で計算したバイクの平均走行速度 v_b のうち 88.53%が $v_1 < v_b < v_2$ となるのに対し，車の平均速度 v_c については 63.44%が $v_2 < v_c < v_3$ となると報告されており，このような速度領域の違いが識別精度に貢献しているものと考えられる．しかし，渋滞の激しい新興国の道路においては，バイク・車ともに平均時速が 15 km 以下などの小さな値になることが頻繁に起こりうる．また，混雑している状況下ではバイクが車の間をぬって走行するため，バイクの方が車より平均速度が高くなることもある．このような環境下では，文献 [25] で提案された手法は必ずしも有効であるとはいえない．

Nitsche ら [28] や Widhalm ら [30] の方法では，加速度データと GPS データなどの位置情報以外のデータは前提としておらず，識別精度も 80%程度と良好な結果が得られている．これらの研究では，特徴量として速度や加速度，角速度に関する統計量や加速度データのパワースペクトラムが用いられている．ところが，上述のとおり新興国の車とバイクにおいては速度をはじめとするこれらの特徴量は類似していることが多いため，文献 [28] や [30] の手法をそのまま適用しても十分な精度が得られないと考えられる．**図 1**，**図 2**，**図 3** には，実際に新興国で予備実験としてデータ収集を行った際に得られたデータにおける速度や加速度などに関する特徴量を示す[*2]．図 1，図 2，図 3 から，これらの特徴量においては十分な識別精度を得るほどの明確な違いは見られないことが分かる．

また，2.1 節で述べた他の手法を車とバイクの識別の問題に適用することも考えられる．しかし，それらの中で主に用いられている走行速度や加速度などの特徴量も，図 1，図 2，図 3 に示すように，車とバイクでは特徴が類似していることが分かる．したがって，これらの手法をそのまま車とバイクの識別問題に適用するのは困難であると考えられる．

以上のような状況に鑑み，本研究は，一般にスマホに搭載されている加速度センサと GPS 受信機から得られるデータのみに基づいて，新興国で主要な交通手段となっている車とバイクを高精度に識別する手法を提案する．本研究では，加速度データや GPS データを単純に用いるのではなく，車両が走行する際の振動発生に関わる要因について考察し，これに基づいて適切な特徴量を選択する．その

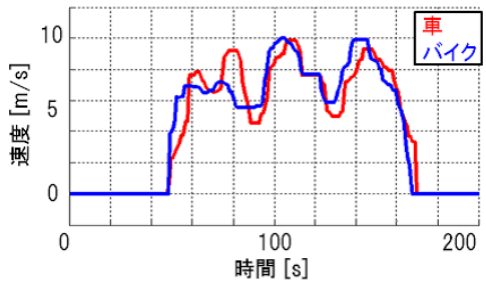


図 1　速度の比較．同一道路で同一時刻に収集したデータを示している

Fig. 1　Comparison of velocity. The data were collected on identical road at identical time.

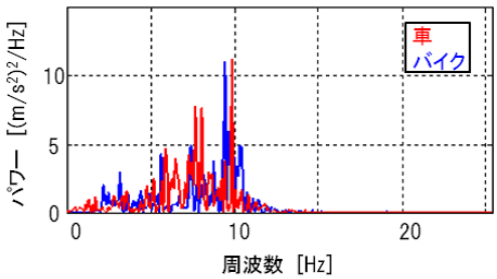


図 2　加速度データに FFT を適用した値の比較

Fig. 2　Comparison of frequency of acceleration data.

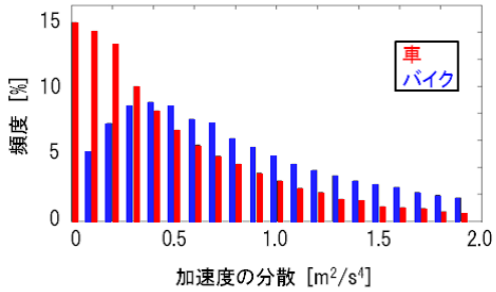


図 3　加速度データの分散値の比較．5 秒ごとに計算した分散値の頻度分布を示している

Fig. 3　Comparison of variance of acceleration data. Frequency distribution of variance calculated every 5 sec is shown.

*2　予備実験は，ベトナム/ハノイ市，インド/バンガロール市，インド/ハイデラバード市で行った．加速度データのサンプリングレートは 50 Hz，GPS データのサンプリングレートは 1 Hz とし，現地で実験協力者を得てデータ収集を行った．スマホの持ち方に関しては制約を設けなかった．

© 2015 Information Processing Society of Japan

うえで，その特徴量に基づいて車とバイクを識別する方法を導出する．必要最小限の特徴量を利用するため，過学習が生じず，高い識別精度が達成できると期待できる．先にあげたような先行研究では，移動手段の違いに起因する観測値の差異をこのような物理的なモデルに基づいてとらえたものはない．このようなモデルを導入したうえで，それに応じた特徴量選択を行い，識別器を構築することが本研究の主たる貢献である．

3. 提案手法

本章では，スマホに一般的に搭載されている GPS 受信機と加速度センサから得られるデータのみを用いて，車・バイクを識別する手法を提案する．まず 3.1 節で提案手法の全体構成を述べる．次に 3.2 節で，本研究の主眼である走行中の車両における振動発生モデルに基づく車とバイクの識別手法について詳しく説明する．

3.1 全体構成

本研究で提案する移動手段識別手法の全体構成を図 4 に示す．まず前処理として GPS 情報が取得できていないデータや，加速度が所定のサンプリングレート以上でとれていないデータを識別対象外として除去する．次に先行研究にならい，加速度データの周期性に注目して歩行を検出する．具体的には，加速度データの自己相関関数を計算し，そのピークを検出する．この検出されたピークが歩行の周期に対応する所定の範囲内にある場合に歩行として検出する．続いて静止区間を識別する．ここでいう静止区間とは，オフィスで仕事をしたり食事をしたりするなど，移動をしていない状態のことを指す．具体的には，静止区間では加速度センサの値にほとんど変動が見られないため，加速度センサの値の分散値を計算し，閾値以下の区間を静止区間として検出する．次に，静止でも歩行でもない区間で一定以上の長さを持つ区間を乗車区間とする．そしてここで検出された乗車区間に対して，車かバイクかの識別を行う．この方法の詳細については 3.2 節で説明する．なお，本研究では電車やバスなどは識別の対象としておらず，これらのデータは処理対象のデータに含まれていないと想定している．このようにして識別した一連のデータに対し，最後に移動手段の遷移の妥当性に基づいて識別データを補正する．たとえば，5 分ごとに車からバイクへまたバイクから車へなどと乗り換えることは少ない．したがって本提案手法では，このような移動手段の遷移はないものと仮定して識別結果を補正する．具体的には，直近一定個の識別結果を参照し，それらの加重平均を用いて識別結果の補正を行う．なお，本研究ではデータを 5 分ごとの区間に区切り，この 5 分間の区間で 1 つの識別結果を出力することとする．

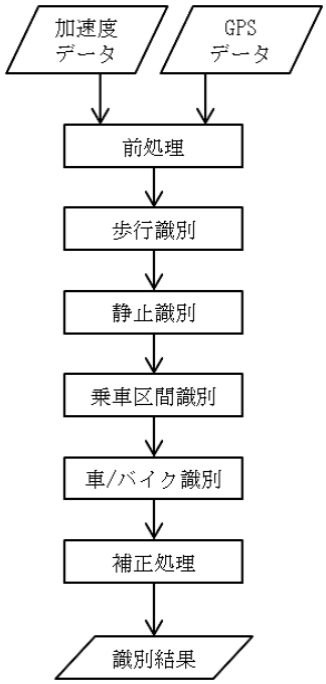


図 4 識別手法の全体のフロー
Fig. 4 Flow of proposed method.

3.2 車/バイクの識別

本節では車両走行時の振動発生モデルに基づいて車とバイクを識別する手法について述べる．まず提案する振動発生モデルについて述べ，次にそれに基づいて具体的に識別を行う方法について説明する．

3.2.1 振動発生モデル

本研究では，車両走行時にスマホで観測される振動は (i) 路面から受ける振動（Vib_R とおく），(ii) エンジンに起因する振動（Vib_E とおく），の 2 種類の振動の重ね合わせによって得られると考える．ここで，本研究における振動とは，具体的には振動の振幅を指す．

車両の走行速度と振動の関係については，既往研究で走行速度が速いほど観測される振動も大きいことが報告されている [33], [34]．そこで車両の走行速度を表すパラメータ Vel を導入し，$Vib_R \propto Vel$ とおく．さらに Vib_R に関して，路面に存在する凹凸の密度が高いほど単位時間あたりに観測される振動は大きくなると考えられる．そこで，路面の凹凸の密度を表すパラメータ $Roughness$ を導入し，$Vib_R \propto Roughness$ とおく．また，車両に振動が生じたとき，実際にスマホで観測される振動の大きさは車体の安定性の影響を受けると考えられる．車体の安定性とは，振動を吸収する機能である座席のサスペンションの性能や，車輪数，車高，車重などによって決定される．安定性が高いほど観測される振動は小さくなると考えられるので，車体の安定性を表すパラメータ $Stability$ を導入し，$Vib_R \propto Stability^{-1}$ とおく．

以上より，スマホで観測される振動を Vib とおくと，車

© 2015 Information Processing Society of Japan

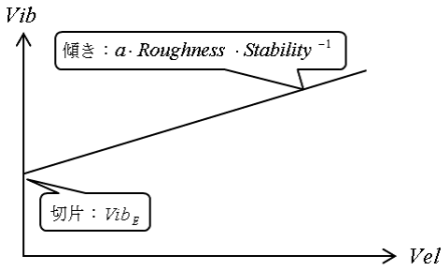


図 5 モデルの表す速度と振動の関係
Fig. 5 Relation between velocity and vibration that the model represents.

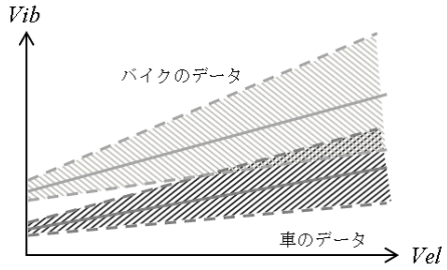


図 6 予想される車のデータとバイクのデータの存在領域の違い
Fig. 6 Expected difference in existence region of data from cars and data from motorbikes.

両走行時の振動発生モデルとして次式を得る．

$$\begin{aligned} Vib &= Vib_R + Vib_E \\ &= a \cdot Roughness \cdot Stability^{-1} \cdot Vel + Vib_E \end{aligned} \quad (1)$$

a は比例定数である．

路面状況は道路ごとに異なるため，パラメータ *Roughness* はレーザ計を用いるなどして，何らかの形で前もって計測しておけることが望ましい．その上で識別の際には，対象車両が走行した道路を GPS データから同定し，その道路の *Roughness* を与えることで，よりモデルに即した形の精緻な識別が可能となる．しかし本論文では，簡略化のため同一地域内では路面状況は類似していると仮定して，パラメータ *Roughness* は一定であると見なす．

さらに，エンジンに起因して生じる振動の走行状態による変分は微小なものとして無視すると，式 (1) は**図 5** のように速度と振動の 2 次元平面内での直線を表すことが分かる．

車とバイクでは車体の安定性や観測されるエンジン振動の大きさが有意に異なると考えられる．具体的には，バイクの方が車体の安定性は低いため，振動はバイクの方が大きい傾向があると考えられる．またエンジン振動そのものは車の方が大きくても，エンジンとの距離やエンジン周辺の機構の違いにより，運転者（や同乗者）が観測する振動はバイクの方が大きい傾向がある．これは，図 5 中の傾きや切片の値がバイクの方が大きいことを示唆する．この差異によって，車のデータとバイクのデータでは，**図 6** のように速度と振動で構成される空間内での存在領域に差異が生じると考えられる．ここで，図 6 中でデータの存在領域がモデルで表される直線を中心として上下にばらついているのは，実データではセンサノイズに加え路面状況の違いなどのノイズ要因があると考えられるためである．本研究ではこの速度と振動で構成される空間内での存在領域の差異を利用して車のデータとバイクのデータを識別する．

3.2.2 識別手法

本項では，スマホのセンサデータを利用して速度と振動で構成される空間内で車とバイクを識別する具体的な手法について説明する．

まずセンサデータから速度と振動を計算する方法を説明する．速度は GPS データによる位置情報および時間情報を用いて，位置の時間差分として計算できる．振動は加速度センサによって計測されることが多いため [35]，振動（の振幅）は加速度データの分散値で表現する．ここではさらに，池谷・長ら [21], [24] の方法を参考に，加速度データを鉛直方向と進行方向横方向に分解してからそれぞれの方向の分散値を計算することとする．二輪車であるバイクは四輪車である車に比べ，特に進行方向に対し横方向の振動が生じやすいと考えられるため，このようにすることで識別精度の向上が見込める．

識別にはベイズ理論を用い，確率分布の計算にはヒストグラム密度推定法を用いる．ヒストグラム密度推定法は，学習データの分布が特定の分布に従わない場合でも確率密度の学習が可能であり，さらに，1 度学習をしてしまえば学習用データを破棄でき，識別計算も高速にできるという特徴がある．そのため本研究のようにデータ点の分布の形状が明白でなく，かつ大量のデータを扱うような場合に好適である．

次にベイズ理論を本研究の車とバイクの識別問題に適用する方法について説明する．まず学習ステップで，移動手段が既知であるデータ（学習用データ）を用いて，移動手段が車（バイク）であったときに観測値 $\mathbf{x}$ が得られる条件付き確率 $p(\mathbf{x}|C)$ ($p(\mathbf{x}|B)$) をヒストグラム密度推定法を利用してそれぞれ求めておく．ここで $\mathbf{x}$ は速度，加速度の鉛直成分の分散，加速度の進行方向横方向の分散，からなる 3 次元空間上のベクトルを表し，C は移動手段が車であるという事象，B は移動手段がバイクであるという事象である．

移動手段が未知のデータ集合 $\mathbf{X} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N}\}$ が得られたときには，次式により移動手段の識別を行う．

$$mode = \arg\max_{m \in \{C,B\}} p(m|\mathbf{X}) \quad (2)$$

ここで，*mode* は識別結果の移動手段であり，事後確率 $p(C|\mathbf{X})$，$p(B|\mathbf{X})$ は次式により計算できる．

$$p(C|\mathbf{X}) = \frac{p(C)\prod_{i=1}^{N} p(\mathbf{x_i}|C)}{p(\mathbf{X})} \quad (3)$$

© 2015 Information Processing Society of Japan

$$p(B|\mathbf{X}) = \frac{p(B)\prod_{i=1}^{N} p(\mathbf{x_i}|B)}{p(\mathbf{X})} \tag{4}$$

$p(\mathbf{x}|C)$ および $p(\mathbf{x}|B)$ は学習ステップで計算したものを用い，事前確率 $p(C)$ および $p(B)$ は対象地域の車両保有台数に関する統計調査データから決定する．また，データ点を 1 点ずつ個別に識別すると瞬間的なノイズなどの影響により誤識別が起こりやすくなるため，式 (3)，(4) のようにいくつかの点（たとえば 5 分間分のデータなど）をまとめた集合として事後確率を計算し，比較するものとする．ここで，式 (3)，(4) で計算される事後確率のうち大きい方の値が小さいとき（すなわち事後確率の差が小さいとき）は，識別結果の信頼性が低いと考えられる．そのため事後確率の大きい方の値を閾値処理し，閾値以下のデータに関しては識別を行わないこととする（「どちらか不明」という判断をしていることに相当）．

4. 評価

提案した手法の有効性を評価するため実験を行った．なお，本評価実験では車とバイクの識別精度に焦点を絞って評価を行った．以下 4.1 節で実験設定についてまとめ，4.2 節で評価結果を紹介し，4.3 で結果に対する考察を述べる．

4.1 実験設定

表 1 に実験データの概要を示す．実験は新興国であるベトナム/ハノイ市で行った．実験期間は 2012 年の 9 月から約 2 週間で，実験参加者は現地協力会社を通して募集したハノイ市在住者 20 人である．実験に使用する端末は実験者側で用意し，私用のスマホは用いなかった．実験では車のデータを 49 時間分，バイクのデータを 498 時間分収集できた．これらのほかに，歩行時などのデータが 224 時間あり，総時間は 771 時間であった．3.1 節で述べたように，本研究では電車やバスは識別の対象としていないため，これらのデータはあらかじめ除外した．データ収集用の Android アプリを開発し，GPS データと加速度データを時刻付きで収集した．加速度データ・GPS データは可能な限り多くサンプリングすることとし，それぞれおよそ 1 Hz，50 Hz でサンプリングを行った．収集したデータのうち，GPS 情報が受信できていないデータや加速度データが取得できていないデータなどは識別の対象外として除外した．

被験者には通常どおりの日常生活を送ってもらい，移動時にアプリを起動してデータ収集を行ってもらうようにした．アプリ起動後に移動手段を入力できるようにし，これから利用する移動手段を移動前に入力してもらうことで，移動手段の正解値を記録した．なお，実験期間中，1 人の被験者が車に乗ることもあればバイクに乗ることも多くあったため，各端末には複数の移動手段のデータが含まれていた．端末の保持方法に制約は設けず，ポケットに入れる，バッグに入れるなど自由とした．データ収集アプリ起動中にはスマホの操作は行わないようにした．

評価の際には，車の数とバイクの数がおおむね同数になるように，バイクのデータをランダムに 10 分割した．その 10 分割されたバイクのデータと，車のデータ，その他のデータを 1 組として，10 組のデータセットそれぞれに対し，10-fold のクロスバリデーションを行い，その平均値を用いて精度評価を行った．

表 1　実験データの概要
Table 1　Summary of the experimental data.

実験場所	ベトナム/ハノイ
実験期間	2012 年 9 月 21 日～2012 年 10 月 4 日
被験者数	20 人
実験端末	HTC One X (Android OS 4.0)：10 台 Samusung Galaxy S3 (Android OS 4.1.2)：10 台
収集データ量	合計 771 時間 ＜内訳＞ 車：49 時間 バイク：498 時間 その他：224 時間
データ収集時の サンプリングレート	加速度データ：50 Hz GPS データ：1 Hz
有効データ*3量	合計 439 時間

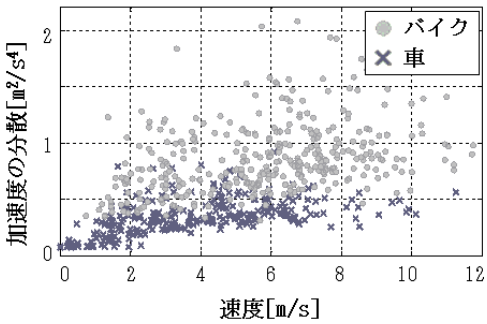


図 7　実験データにおける速度と振動（加速度の分散）の関係
Fig. 7　The relationship between velocity and vibration (variance of acceleration) of the experimental data.

4.2 結果

図 7 は速度と振動（加速度の分散値で表現）の平面上に実験で収集したデータをプロットした図である．ここでは車・バイクそれぞれ約 30 分間分ずつのデータを示している．図 7 から，実験データにおいても図 6 で示した想定のように，車とバイクの存在領域に差異が見られることが確認できる．

*3　加速度データも GPS データも取得できているデータ．

© 2015 Information Processing Society of Japan

表 2　適合率
Table 2　Precision.

		推定	
		車	バイク
真値	車	81.1% [±3.3%] (1304/1608)	4.9% [±0.8%] (78/1586)
	バイク	11.5% [±4.5%] (185/1608)	82.3% [±3.1%] (1305/151)
	その他	7.4% [±2.1%] (119/1608)	12.8% [±1.5%] (203/1586)

表 3　再現率
Table 3　Recall.

車	バイク
55.6% [±11.6%] (1304/2346)	53.8% [±7.2%] (1305/2425)

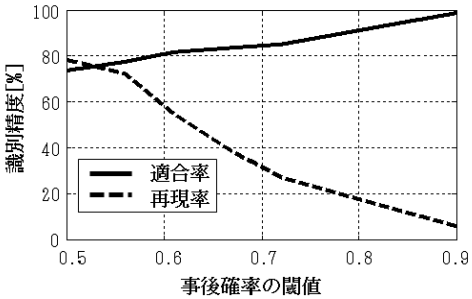


図 8　適合率・再現率の関係
Fig. 8　The relationship between precision and recall.

実験データに提案手法を適用した際の評価結果を，**表 2**，**表 3** に示す．表中の括弧内の数字は，識別単位である 5 分間の区間の数を示している．表 2 に示すように，車・バイクのいずれの移動手段についても 80%以上の適合率を達成できている．このことから，本提案手法が新興国における車とバイクの識別問題に適用可能であることが確認できる．表 3 に示す再現率の値が適合率に比べ低くなっているのは，3 章で述べたように式 (3)，(4) で計算される事後確率の高い方の値に関する閾値処理を行い，事後確率が閾値以下の値を持つデータに関しては識別を行っていないためである．事後確率の閾値の具体的な値としては 0.60 を用いた．

図 8 に示すように，適合率と再現率はトレードオフの関係にある．本研究のように，識別結果を利用した交通情報生成を想定する場合，再現率はサンプル数を増やすことでカバーできるが，適合率は生成する交通情報の信頼性に関わるため，再現率よりも適合率を重視すべきである．一方でサンプル数を増やすことが難しい場合などは，閾値を低く設定することで高い再現率を達成することができる．

図 9　実験データのカバー範囲
Fig. 9　Spatial coverage of the experimental data.

4.3　考察

4.3.1　データ量

収集したデータ量が少ないとサンプルに偏りが生じるため，精度評価の信頼性が低下してしまう．そこで，サンプルに偏りが生じる要因という観点から，被験者の人数・実験期間・地理的な網羅性，の 3 点を基準に，データの量について考察を加える．

被験者としては，性別や年齢などにおいて様々なプロファイルの被験者を網羅できることが望ましい．そのような観点からは 20 人という被験者数は必ずしも十分であるとはいえず，より多数の被験者を募集することが好ましい．一方で 2 週間という実験期間に関しては，実験期間中に 2 度の週末を含むことができ，平日の通勤・帰宅などの決まったルートが想定されるデータのみが対象とはならないため，精度評価のための期間としては十分であると考えられる．**図 9** は収集した GPS データの軌跡の一部を地図上にプロットしたものを示す．図 9 はハノイ市の中心部の地図を示している．これより，ハノイ収集したデータは，ハノイ市中心部の様々な位置を網羅的にカバーできていることが確認できる．したがって，地理的な網羅性という観点からも収集したデータ量に不足はないと考えられる．以上より，被験者の人数という点でやや改善の余地があるものの，基本的な精度評価のためのデータ量としては十分な量が確保できているといえる．

4.3.2　誤識別の原因

まず真値が車でもバイクでもないにもかかわらず，車あるいはバイクと誤識別している場合について述べる．本研究では初めに歩行検出や静止検出を行い，それらで検出されなかった区間を乗車区間とし，その区間に対し車/バイクの識別を適用している．今回の実験では移動時のみに計測を行ったため，この誤識別の原因は歩行の検出漏れであると考えられる．

本研究では歩行の際に生じる周期的なリズムを加速度の自己相関関数のピーク値により検出している．歩行が検出できないのは，このような特徴的なリズムが検出できなかった場合であると考えられる．本実験では，スマホの保

© 2015 Information Processing Society of Japan

図 10 真値はバイクだが車と誤識別していた道路
Fig. 10 The data of motorbikes collected on this road were misclassified into cars.

図 11 真値は車だがバイクと誤識別していた道路
Fig. 11 The data of cars collected on this road were misclassified into motorbikes.

持方法に制約を設けなかったため，たとえばクッション性の高い鞄の中にスマホを入れて歩行している場合などに，振動が吸収され特徴が表れにくくなることが考えられる．現在歩行検出の際には，自己相関関数のピーク値が一定の閾値以上のものだけを対象とし，それが歩行に相当する周期の範囲内にあるかを判定しているが，全体の加速度の振れ幅の傾向などから，この閾値を動的に変更することで精度向上が見込める．

次に車をバイクと誤識別した場合およびバイクを車と誤識別した場合について述べる．本研究では車とバイクの識別は走行速度と振動との関係に基づいて行っている．そのため，車であるにもかかわらず走行速度に対する振動値が大きかったり，反対にバイクであるにもかかわらず振動値が小さかったりすると誤識別の要因となる．

このように，全体の傾向とは異なる振動値が観測される原因としては路面状況の違いが考えられる．3 章で述べたように，本研究で提案する式 (1) の振動発生モデルにおける路面の凹凸の密度を表すパラメータ *Roughness* は，現在同一地域内では一定であると仮定している．4.2 節で述べたように，この仮定の下でも一定の識別精度が得られてはいるものの，実際には同一地域内でも路面状況が異なる個所があるため，そのような個所で誤識別が生じているものと考えられる．そこで，誤識別の生じている箇所について路面状況の調査を行った．

図 10 はバイクを車と誤識別していた道路の例である．平均的な道路と比べ舗装状況が良いため，バイクでも振動値が小さかったためであると考えられる．**図 11** は車をバイクと誤識別していた道路の例である．舗装状況はあまり良くなく，他の車両や道路際の障害物などを避けながら走行していたため，車でも振動値が大きくなったためであると考えられる．

図 12 は同一車両でこれらの道路を走行したときの速度と振動の関係を表す図である．図 10 の道路の方が，実際に走行速度に対する振動値が小さい傾向があることが分かる．このように道路ごとに異なる路面状況について，対象地域で調査を行う，またはデータから推定を行うことによって，式 (1) における *Roughness* を与えることができればさらなる精度の向上が見込める．

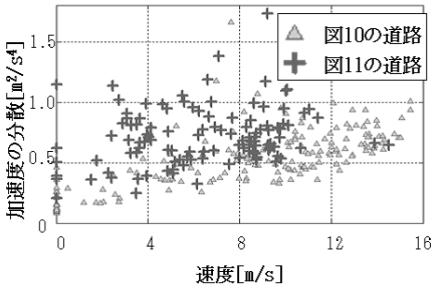


図 12 路面状況の異なる 2 つの道路における速度と振動の関係の違い
Fig. 12 Different relationship of velocity and vibration between two dataset collected on two different road.

表 4 車をバイクと誤識別したデータ数の端末 ID 別割合
Table 4 Misclassification of cars as motorbikes.

端末 ID	全体に占める誤識別の割合
1	56.3%
2	41.7%
3	2.1%

表 5 バイクを車と誤識別したデータ数の端末 ID 別割合（表 4 と同じ ID のものは同一端末を示す）
Table 5 Misclassification of motorbikes as cars. The IDs appear in Table 4 as well denote the identical devices.

端末 ID	全体に占める誤識別の割合
4	30.6%
5	20.7%
2	10.8%
6	9.0%
7	7.2%
8	5.4%
1	4.5%
9	4.5%
3	3.6%
10	3.6%

表 4，表 5 は，評価の際に生じた誤識別の数を端末 ID ごとに数えた結果である．これらから，誤識別は特定の端末に偏って発生していることが分かる．たとえば，車をバイクと誤識別したものに関しては，誤識別数の多い上位 2 つの端末で誤識別全体の 98.0%を，バイクを車と誤識別し

© 2015 Information Processing Society of Japan

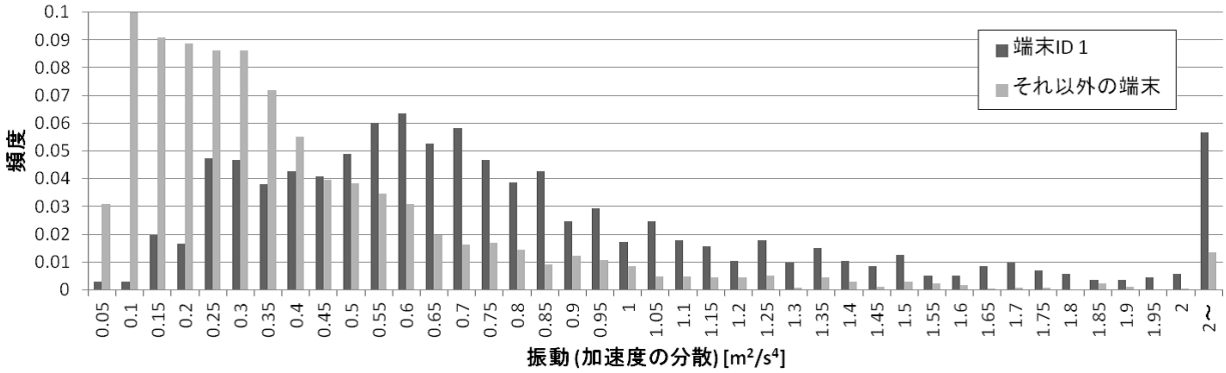


図 13　端末 ID 1（車をバイクと誤識別した端末）における車データと，その他の端末における車データの振動値の分布の比較

Fig. 13　Comparison of distribution of vibration values between one device (ID 1) and the others.

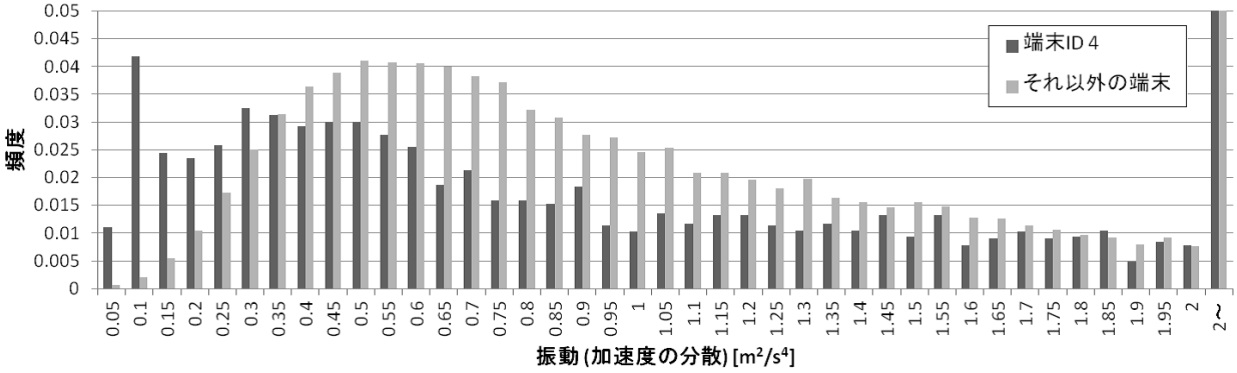


図 14　端末 ID 4（バイクを車と誤識別した端末）におけるバイクデータと，その他の端末におけるバイクデータの振動値の分布の比較

Fig. 14　Comparison of distribution of vibration values between one device (ID 4) and the others.

たものに関しては，同 51.3%を占めている．

表 4，表 5 において最も誤識別が多かった端末（それぞれ端末 ID 1 と端末 ID 4）とその他の端末との振動値の分布の違いを**図 13**，**図 14** に示す．ここで，図 13，図 14 における振動値は，ハノイ市中心部における定常走行と見なせる時速 25 km～時速 35 km のデータだけを抽出してから算出した．これは，3 章で述べたように速度と振動の大きさには相関関係があるため，その影響を除いた振動値の比較を行うためである，図 13 から，ID 1 の端末ではその他の端末に比べ，車に乗っている際でも振動値が大きい傾向があることが分かる．逆に，図 14 から，ID 4 の端末ではその他の端末に比べ，バイクに乗っている際でも振動値が小さい傾向があることが分かる．これらのような，通常の端末との振動値の違いがこれらの端末で誤識別が生じやすい原因であると考えられる．

これらの端末で他の端末と異なる振動値が得られる理由の 1 つとして，よく利用する道路の整備状況の違いが考えられる．被験者によって生活圏は異なるため，よく利用する道路も被験者ごとに異なることになる．たとえばいつも整備状況の悪い道路を通って通勤するような人の場合は，車で通勤していたとしてもその間に観測される振動は相対的に大きくなり，誤識別の原因となりうる．また，端末の保持方法による違いも原因として考えられる．たとえばクッション性の高いバッグの中に端末を入れて移動するような人の場合は，バイクで移動していたとしてもその間に観測される振動は相対的に小さくなり，誤識別の原因となりうる．

ふだんよく利用する道路やスマホの保持方法は，人による違いはあるものの，同一人物内では一定の傾向があると考えられる．ゆえに，一定期間以上継続してデータ収集をすることができれば，振動の大小のバイアス傾向を学習させて個人ごとに識別基準を調整することで，精度向上につなげることができると考えられる．

4.3.3　識別手法の比較

表 6 に識別手法別の適合率の評価結果を示す．k-Nearest Neighbor 法や Support Vector Machine など，他の識別手法でも 80%に近い高い精度での識別ができているが，本研究で使用したヒストグラム密度推定法を利用したベイズの

© 2015 Information Processing Society of Japan

表 6　識別手法ごとの精度の比較
Table 6　Comparison on classification methods.

識別手法	適合率
ベイズ（ヒストグラム密度推定法）	81.1%
KNN	78.9%
SVM	78.0%

表 7　予備実験のデータ概要
Table 7　Summary of the preliminary experimental data.

実験場所	ベトナム/ハノイ，Xuan Thuy 通り
実験期間	2014 年 3 月 11 日〜2014 年 4 月 2 日
被験者数	54 人（車 8 人，バイク 46 人）
収集データ量	車：199 データ バイク：3,188 データ
GPS データのサンプリングレート	1 Hz

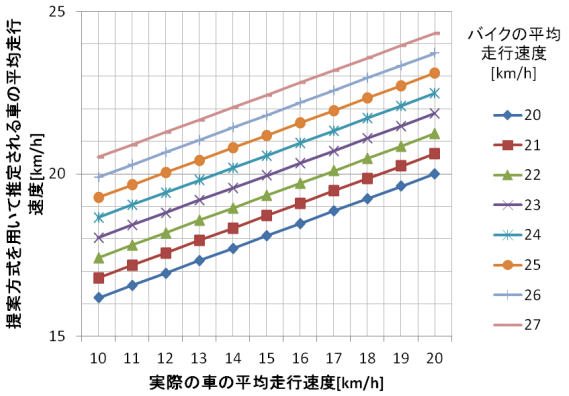


図 15　実際の走行速度と，提案手法で達成した識別精度において推定される走行速度の関係
Fig. 15　Relationship between actual velocities and estimated velocities calculated by applying the proposed classification method.

手法が最も識別精度が高いという結果が得られた．特徴空間内での点のばらつきを，確率分布としてより正確にとらえることができているためであると考えられる．拡張性という観点からも，地域によって異なる車やバイクの保有台数などの事前知識を導入しやすいベイズの手法は，より実用的であると考えられる．

4.3.4　達成した識別精度と実アプリケーションへの応用可能性

1 章で述べたように，本研究は交通手段ごとの交通状況を正確に把握できるスマホプローブシステム実現のための車とバイクの識別技術の確立を目的としている．交通状況把握の具体的な例としては，交通渋滞状況のレベルの把握があげられる．たとえば日本道路交通情報センター[*4]では，一般道において平均走行速度が時速 10 km 以下の場合を「渋滞」，時速 20 km 以下の場合を「混雑」と定義し，情報配信を行っている．

車とバイクが混在する交通流においては，まず車のデータとバイクのデータを本提案手法のような方法で区別した後，それぞれのデータから平均走行速度を計算する必要がある．たとえば混雑時においてはバイクは車の間をぬって走行できるため，バイクの方が走行速度が速くなるといった場合があるからである．実際に，ハノイ市においてよく渋滞が発生する道路の 1 つである Xuan Thuy 通りで予備実験を行い，**表 7** に示すようなデータを収集したところ，バイクの平均速度は時速 21.3 km となったのに対し，車の平均速度は時速 12.4 km となった．このような違いがあるため，車とバイクの識別において誤識別が生じると，（上記の場合では走行速度の速いバイクと走行速度の遅い車が混ざってしまい）推定される走行速度，そしてそれを基に生成される渋滞情報に誤りが混入してしまうことになる．

そこで，このような新興国においての典型的な状況における本提案手法の応用可能性を調べるため，車やバイクの平均走行速度として様々な値を仮定し，本研究で達成した識別精度のときに推定される速度がどの程度になるかをシミュレーションにより検討する．具体的には，実際の車の走行速度を v_{c_act}，実際のバイクの走行速度を v_{b_act}，車の推定精度（適合率）を p_{p_c}，真値はバイクであるが車であると誤識別する確率を $p_{m_{bc}}$，車であると推定した数を n_c，バイクであると推定した数を n_b とおくと，本提案手法を適用した結果計算できる車の推定速度 v_{c_est} は次式 (5) で計算できる．

$$\begin{aligned} v_{c_est} &= \frac{v_c n_c p_{p_c} + v_b n_b p_{m_{bc}}}{n_c p_{p_c} + n_b p_{m_{bc}}} \\ &= \frac{v_c p_{p_c} + v_b k p_{m_{bc}}}{p_{p_c} + k p_{m_{bc}}} \end{aligned} \tag{5}$$

この推定速度 v_{c_est} と実際の走行速度である v_{c_act} や v_{b_act} との関係を計算する．ここで $k = n_c/n_b$ は，車であると推定した数とバイクであると推定した数の比である．シミュレーションにおいては k の値は何らかの形で事前に与えなければならない．そこで，ハノイ市における車とバイクの保有状況[*5]を参考に，ここでは $k = 4{,}038/350 \simeq 11.5$ とした．また，4.2 節で述べた精度評価結果より，$p_{p_c} = 0.811$，$p_{m_{bc}} = 0.115$ とした．

シミュレーション結果を**図 15** に示す．図 15 から，たとえば上述したようにバイクの実際の平均速度が時速 21.3 km，車の実際の平均速度が時速 12.4 km であったときには，推定される車の平均走行速度は約時速 18 km となり，上記日本道路交通情報センターの定義する渋滞情報区分のレベルでは正しい情報生成ができることが分かる．逆に，この渋滞情報区分のレベルを超えない範囲から逆算すると，実際の車の平均走行速度が時速 10 km 程度のときには，バイクの平均走行速度が時速 25 km 程度であっても正

*4　http://www.jartic.or.jp/jartic_web/others/faqs.html

*5　http://www.jica.go.jp/project/vietnam/016/activities/

© 2015 Information Processing Society of Japan

しい渋滞情報の生成ができるのに対し，車の平均走行速度が時速 15 km 程度のときには，バイクの平均走行速度が時速 24 km 以上になると正しい渋滞情報の生成ができなくなることが分かる．このように，そのときどきの車・バイク各々の走行速度によって正確な渋滞情報生成に必要な識別精度は異なるが，本研究で達成した識別精度のときには，上述のような典型的なケースにおける渋滞情報生成は正しく生成できることから，本研究の提案する識別手法は実応用の観点からも有効であるといえる．

5. まとめ

本研究ではスマホプローブシステムへの適用を目指して，スマホ搭載のセンサ情報を利用して車とバイクを識別する手法を提案した．特徴が類似しており，従来識別が困難であった車とバイクについて，走行時の振動発生モデルを導入することにより，高精度に識別を行う手法を提案した．提案した手法を実験によって評価したところ，適合率 80%以上，再現率 50%以上を達成しており，提案手法の有効性を確認した．今後は，スマホ保持方法の違いや路面状況の違いなどへの対応を検討していく予定である．

参考文献

[1] 熊谷正俊，蛭田智昭，奥出真理子，横田孝義：特徴空間軌跡の追跡による動的交通状況予測，情報処理学会論文誌，Vol.53, No.1, pp.243–250 (2012).

[2] Wang, W., Fushiki, T., Zhang, H., Kimita, K. and Mizuta, H.: A Probe Car System for Beijing, *12th World Congress on Intelligent Transport Systems* (2005).

[3] 森田綽之，割田　博：動的な交通状況の予測と交通情報生成，国際交通安全学会誌，Vol.31, No.1, pp.24–30 (2006).

[4] 山根憲一郎：VICS データを利用した渋滞傾向情報に関する検討，土木計画学研究・論文集，Vol.29, pp.171–174 (2004).

[5] 舟橋賢二，西村茂樹，堀口良太，赤羽弘和，桑原雅夫，小根山裕之：VICS 蓄積データを用いた旅行時間短期予測手法に関する研究，土木計画学研究・論文集，Vol.27, pp.176–179 (2003).

[6] 王　立暁，姜　美蘭，森川高行：プローブカーデータと VICS データの融合による旅行時間推計に関する研究，土木計画学研究・論文集，Vol.23, No.4, pp.1011–1018 (2006).

[7] The Motley Fool: Report: Southeast Asia Smartphone Market Grows 61% on Heels of Android Popularity (2013).

[8] 大橋洋輝，秋山高行，佐藤暁子：車両走行時の振動発生モデルに基づく移動手段識別方式，電子情報通信学会総合大会講演論文集 (2013).

[9] 大橋洋輝，秋山高行，佐藤暁子：車両走行時の振動発生モデルに基づく移動手段識別方式，第 47 回土木計画学研究発表会・講演集 (2013).

[10] Ohashi, H., Akiyama, T., Yamamoto, M. and Sato, A.: Modality Classification Method Based on the Model of Vibration Generation while Vehicles are Running, *Proc. 6th ACM SIGSPATIAL International Workshop on Computational Transportation Science* (2013).

[11] Zheng, Y., Liu, L., Wang, L. and Xie, X.: Learning Transportation Mode from Raw GPS Data for Geographic Applications on the Web, *Proc. 17th International Conference on World Wide Web* (2008).

[12] Bohte, W. and Maat, K.: Deriving and Validating Trip Destinations and Modes for Multi-day GPS-based Travel Surveys: A Large-scale Application in the Netherlands, *International Conference on Survey Methods in Transport: Harmonisation and Data Comparability* (2008).

[13] Sun, Z. and Ban, X.: Vehicle classification using GPS data, *Transportation Research Part C*, Vol.37, pp.102–117 (2013).

[14] Bohte, W., Maat, K. and Quak, W.: A Method for Deriving Trip Destinations and Modes for GPS-based Travel Surveys, *Research in Urban Series* (2008).

[15] 松本修一：GPS 携帯を活用した行動調査に関する基礎的研究，*KEIO SFC JOURNAL*, Vol.9, No.1, pp.21–28 (2009).

[16] Bolbol, A., Cheng, T., Tsapakis, I. and Haworth, J.: Inferring hybrid transportation modes from sparse GPS data using a mobing window SVM classification, *Computers, Environment and Urban Systems* (2012).

[17] Zhang, L., Dalyot, S., Eggert, D. and Sester, M.: Multi-stage Approach to Travel-mode Segmentation and Classification of GPS Traces, *ISPRS Workshop on Geospatial Data Infrastructure: From Data Acquisition and Updating to Smarter Services* (2011).

[18] 堀口良太，長岡　亨，畑　成年：GPS 携帯電話による大規模パーソンプローブ調査のためのトリップ情報抽出手法に関する研究，土木計画学研究・講演集 (2006).

[19] Stenneth, L., Wolfson, O., Yu, P.S. and Xu, B.: Transportation Mode Detection using Mobile Phones and GIS Information, *Proc. 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (2011).

[20] Long, X., Yin, B. and Aarts, R.: Single-Accelerometer-Based Daily Physical Activity Classification, *31st Annual International Conference of the IEEE EMBS* (2009).

[21] 池谷直紀，菊池匡晃，長　健太，服部正典：3 軸加速度センサを用いた移動状況推定方式，信学技報，電子情報通信学会 (2008).

[22] Hemminki, S., Nurmi, P. and Tarkoma, S.: Accelerometer-Based Transportation Mode Detection on Smartphones, *Proc. 11th ACM Conference on Embedded Networked Sensor Systems* (2013).

[23] 小林亜令，岩本健嗣，西山　智：釈迦：携帯電話を用いたユーザ移動状態推定方式，情報処理学会論文誌，Vol.50, No.1, pp.193–208 (2009).

[24] Cho, K., Iketani, N., Setoguchi, H. and Hattori, M.: Human Activity Recognizer for Mobile Devices with Multiple Sensors, *Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing* (2009).

[25] Feng, T. and Timmermans, H.J.: Transportation mode recognition using GPS and accelerometer data, *Transportation Research Part C*, Vol.37, pp.118–130 (2013).

[26] Hato, E.: Development of behavioral context addressable loggers in the shell for travel-activity analysis, *Transportation Research Part C*, Vol.18, pp.55–67 (2010).

[27] Chen, C., Gong, H., Lawson, C. and Bialostozky, E.: Evaluating the feasibility of a passive travel survey collection in a complex urban environment: Lessons learned from the New York City case study, *Transportation Research Part A*, Vol.44, pp.830–840 (2010).

[28] Nitsche, P., Widhalm, P., Breuss, S., Brandle, N. and Maurer, P.: Supporting large-scale travel surveys with smartphones - A practical approach, *Transportation Research Part C* (2013).

© 2015 Information Processing Society of Japan

[29] Reddy, S., Burke, J., Estrin, D., Hansen, M. and Srivastava, M.: Determining Transportation Mode On Mobile Phones, *12th IEEE International Symposium on Wearable Computers, ISWC 2008* (2008).
[30] Widhalm, P., Nitsche, P., Breuss, S. and Brandle, N.: Tranport Mode Detection with Realistic Smartphone Sensor Data, *21st International Conference on Pattern Recognition*, pp.573–576 (2012).
[31] Kobayashi, A., Iwamoto, T. and Nishiyama, S.: UME: Method for Estimating User Movement Using an Acceleration Sensor, *International Symposium on Applications and the Internet* (2008).
[32] Kobayashi, A., Muramatsu, S., Kamisaka, D., Watanabe, T., Minamikawa, A., Iwamoto, T. and Yokoyama, H.: Shaka: User Movement Estimation Considering Reliability, Power Saving, and Latency Using Mobile Phone, *IEICE Trans. Information and Systems*, Vol.E94-D, No.6, pp.1153–1163 (2011).
[33] 石井博典，藤野陽三，水野祐介，貝戸清之：営業車両の走行時の車両振動を用いた起動モニタリングシステム（TIMS）の開発，土木学会論文集 F，Vol.64, pp.44–61 (2008).
[34] 峰岸順一，上野慎一郎，橋本喜正：車軸の加速度測定による舗装路面の段差評価と振動の関係，東京都土木技術支援・人材育成センター年報，pp.65–71 (2012).
[35] 穀山　渉：振動加速度標準及び角振動標準に関する調査研究，産総研計量標準報告，Vol.9, pp.49–74 (2014).

大橋 洋輝　（正会員）

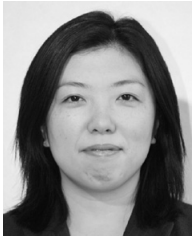
2011 年京都大学大学院情報学研究科修士課程修了．同年（株）日立製作所中央研究所入社．知能システム研究部所属．現在，機械学習・パターン認識を用いたセンサデータ解析およびその応用に関する研究に従事．電子情報通信学会会員．

秋山 高行　（正会員）

2008 年東京大学大学院工学研究科修士課程修了．同年（株）日立製作所中央研究所入社．知能システム研究部所属．現在，センサデータによる人流・交通流分析とシミュレーションの研究に従事．

山本 正明

2005 年三重大学大学院工学研究科博士前期課程修了．同年（株）日立製作所中央研究所入社．知能システム研究部所属．現在，センサデータ応用に関する研究・開発に従事．電子情報通信学会会員．2005 年度映像情報メディア学会丹羽高柳賞，2013 年度電子情報通信学会通信ソサイエティ論文賞．



佐藤 暁子　（正会員）

1998 年東京大学大学院工学研究科修士課程修了．同年（株）日立製作所中央研究所入社．知能システム研究部所属．現在，持続可能な都市交通に関する研究・開発に従事．

© 2015 Information Processing Society of Japan