

Recursive Autoencoder にもとづいた移動軌跡からの特徴量自動抽出手法の提案

Feature Extraction from Movement Trajectory Based on Recursive Autoencoder

河野 慎[♡] 遠藤 結城[◇] 戸田 浩之[▲]
小池 義昌[◇] 植田 一博[♡]

Makoto KAWANO Yuki ENDO
Hiroyuki Toda Yoshimasa KOIKE
Kazuhiro Ueda

本研究では、ユーザのコンテキスト解析研究の一つである移動軌跡にもとづいた移動手段推定に着目する。移動手段推定を行う従来手法では、移動手段の推定に有効な移動軌跡の特徴量を経験にもとづき人手で設計していた。しかし人手による有効な特徴量設計は時間がかかり、また全てを網羅することは難しい。そこで特徴量を自動で抽出することが出来る表現学習の一つである深層学習に着目する。深層学習は、ニューラルネットワークを多層にすることで、従来よりも表現力を向上させたものであり、本研究では関連手法の一つである recursive autoencoder を用いる。本研究では、この recursive autoencoder に次元圧縮調節と正解ラベル学習の機能を加えた N-transformed supervised recursive autoencoder を提案する。Microsoft GeoLife データセットを用いて本手法の評価実験を行い、従来の手法より高い推定精度を示し、本手法によって有効な特徴量を自動抽出できることを示す。

This paper focuses on transportation mode estimation from movement trajectories, which is one of the research topics of users' context analysis. In order to estimate transportation modes, the traditional approaches have designed effective features based on researchers' prior knowledge. However, manually engineering features is time-consuming and cannot cover all effective features. To address this issue, we adopt deep learning, which is a kind of representation learning and can automatically extract features from input data. Deep learning is a method for learning neural networks that have multi-layer structures with high expressiveness, and we use recursive autoencoder that is one of the deep neural networks. We propose a recursive autoencoder equipped with functions of adjusting the dimensions of output features and supervised learning, called N-transformed supervised recursive autoencoder. In our experiments with the Microsoft GeoLife dataset, we show that the proposed method can extract effective features, which result in higher accuracy than the existing method.

♡ 非会員 東京大学大学院情報学環・学際情報学府
9759014385@mail.ecc.u-tokyo.ac.jp

◇ 非会員 日本電信電話株式会社 NTT サービスエボリューション研究所

▲ 正会員 日本電信電話株式会社 NTT サービスエボリューション研究所

1. はじめに

近年、スマートフォンや様々なセンサ、SNS などの普及により、ユーザを取り巻く環境に関する情報を取得することが可能になってきている。これらの情報を用いてユーザのコンテキストを解析し、ユーザの状況を推定や判定する研究は多く取り組まれてきた。例えば、GPS では測定が難しい建物内におけるユーザの位置を Wi-Fi のアクセスポイント [17] やセンサノード [10] を用いて推定するものやユーザの行動をウェアラブルデバイスによって判定するもの [9]、スマートフォンのジャイロセンサ [15] や SNS のチェックイン機能 [23] を用いてユーザの移動先を予測するものなどがあげられる。

本研究は、ユーザのコンテキスト理解研究の一つである GPS や Wi-Fi 測位された移動軌跡からの移動手段判定に着目する。移動手段判定とは、移動軌跡を適当な区間で分割し（移動軌跡分割）、各区間においてユーザがバスや車などで移動しているのか、歩いているのかなどの移動手段を推定するタスク（移動手段推定）である [26]。移動手段判定が実現することで、様々なアプリケーションへの応用が可能となる。例えば、ユーザ自身が過去の行動履歴を確認できるライフログサービスや、ユーザの移動手段を把握し、ユーザの行動を支援するパーソナルアシスタントサービスなどへの応用が期待できる。

GPS 等のセンサ情報からユーザのコンテキストを理解する場合、一般には教師ありのマルチクラス分類問題を機械学習によって解くことが多い [5, 12]。機械学習はユーザの行動を判定するには有効な手法であるが、判定精度を向上させるためには生データから特徴量を人が設計する必要がある。この工程は feature engineering [1] と呼ばれ、移動手段推定においても、feature engineering を行う必要がある。従来手法では、速度や加速度などの基本的な特徴量 [26] とこれらの基本的な特徴量を組み合わせた heading change rate¹ などの応用的な特徴量 [24] を人間が経験にもとづいて設計し、移動手段推定を行った。その結果、基本的な特徴量に加えて応用的な特徴量を用いることによる精度向上が確認されている。このことから、人間の行動は単純な特徴だけでは区別することが難しく、複雑な特徴も必要であると考えられる。しかしこの特徴が複雑になるほど設計は難しくなり、移動手段推定に有効な特徴量を網羅することは難しい。

そこで、本研究では応用的な特徴量を基本的な特徴量から自動抽出することで、この問題の解決を試みる。特徴量を自動抽出する手法として、表現学習 [2] があげられる。表現学習は与えられたデータからタスクに応じた特徴量を自動抽出する技術であり、中でもニューラルネットワークを拡張した深層学習 [1] が近年注目されている。深層学習は画像認識 [7] や音声認識 [18]、自然言語処理 [19] の分野で適用され、従来の feature engineering によって設計された特徴量を用いた場合よりも高い精度を達成している。本研究では移動手段推定のための特徴量を深層学習によって抽出する。しかし多くの深層学習の実装手法は、入力ベクトルの次元を固定する必要があるため、単純に移動軌跡に適用することは難しい。なぜならば、移動軌跡によってその移動時間が異なることから、測位点の数、すなわち入力ベクトルの次元数も異なってしまうからである。

この異なる次元数に対応するため、本研究では入力ベクトルの次元数を固定せずに利用できる recursive autoencoder (RAE) [4, 16] に着目する。RAE を用いることで異なる大きさの移動軌跡を入力ベクトルとして扱うことが可能となる。RAE の特徴は入力ベクトルの次元を圧縮することであるが、入力ベクトルによっては次元圧縮を過度に行ってしまう傾向にある。次元圧縮が過度に行われた場合、特徴量の表現力が低下してしまうことがある。特に移動軌跡の場合、移動手段によっては入力ベクトルの次元数が非常に大きいものがあるため、この問題が顕著になる。また、一般的に autoencoder は教師なし学習であるが、Socher ら [19] に

¹ 移動軌跡における方向転換の割合

表 1: Zheng らによる提案特徴量

カテゴリ	特徴量
basic feature [26] 10 次元	セグメント長
	速度平均
	速度期待値
	速度分散
	上位三つの速度
	上位三つの加速度
advanced feature [24] 3 次元	速度変化率
	停止率
	方向変化率

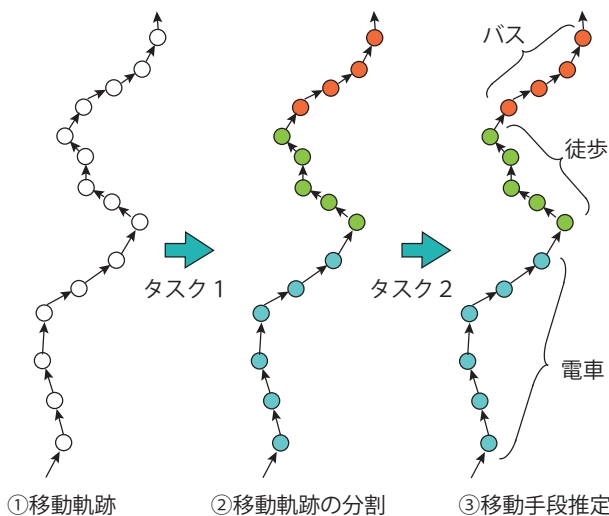


図 1: 移動手段判定の流れ。丸点は測位点を表している。タスク 1 において、移動軌跡に含まれる各測位点の速度や加速度からセグメント (各セグメントは何らかの移動手段を表す) に分割する。そしてタスク 2 においてセグメントの移動手段を推定する。

よって、RAE は教師あり学習に拡張されている。RAE を教師あり学習にすることで、移動手段をより精度良く推定しやすくなる特徴量の抽出が可能になる。

以上のことから本研究では、次元圧縮の調節と教師あり学習が可能である *n*-transformed supervised recursive autoencoder (NTS-RAE) を提案し、移動手段推定に適用する。提案手法による移動手段推定の精度を評価するため、本研究では Microsoft Research によって公開されている GeoLife のデータセット²を用いた。本実験では、まずこのデータセットに提案手法を適用し、特徴量を自動抽出する。その後、feature engineering によって設計された特徴量と提案手法による特徴量のそれぞれを使った場合の移動手段推定の精度を評価する。

本研究の貢献を以下に示す。

- RAE にもとづいた手法の移動軌跡への適用とそのための入力ベクトルの工夫
- 次元圧縮調節と教師あり学習が可能な *n*-transformed supervised recursive autoencoder の提案
- 加工されていない移動軌跡から自動抽出された特徴量を利用することによる移動手段推定の評価

本稿の構成は以下の通りである。2 章で移動手段判定に関する従来の手法と問題意識について述べ、3 章では提案手法について説明を行う。4 章では評価を行い、5 章では関連研究を紹介する。最後に 6 章で本研究のまとめと今後の展望を述べる。

2. 移動手段判定

本章では、移動手段判定の流れ (図 1) と従来の手法 [24, 26] の問題意識について説明する。Zheng らによって GPS ロガーを用いてユーザから一定期間内収集された GPS データである移動軌跡から移動手段を判定するためには大きく分けて次の二つのタスクを行う必要がある。

タスク 1. 移動軌跡をセグメントと呼ばれる単位で分割

タスク 2. セグメントの移動手段を推定

それぞれのタスクの詳細については以下で述べる。

2.1 移動軌跡の分割

移動軌跡は複数の測位点から形成され、各測位点は緯度、経度、時刻情報をもつ。Zheng ら [26] は移動軌跡の連続した測位点から

速度、加速度を算出し、これらの数値が急激に変化した瞬間の測位点において移動手段が切り替わると仮定した。移動手段が切り替わった測位点から次に移動手段が切り替わる測位点までの移動軌跡をセグメントと呼ぶ。この仮定のもと実験を行った結果、約 90% の精度でセグメントを分割することが可能であると報告されている。そこで本研究では、この移動軌跡をセグメントに分割するタスクとして Zheng らの手法を用いることを想定し、次節のタスク 2 に取り組む。

2.2 移動手段推定

タスク 1 で分割された個々のセグメントは、車や徒歩など一つの手段で移動されていると考える。この前提のもと、タスク 2 では各セグメントの移動軌跡から特徴量を抽出し、機械学習を用いてセグメントに付与されている移動手段を推定する。機械学習には様々な手法が提案されているが、いずれの手法も利用する際に特徴量を人が設計したものである。Zheng らも support vector machine (SVM) や decision tree (DT) を用いるために、速度や加速度などの基本的な特徴量である basic feature [26] を設計し、さらにこの basic feature を組み合わせた複雑な特徴量の advanced feature [24] を設計している (表 1)。Basic feature と advanced feature を組み合わせた結果、basic feature のみで移動手段推定をした場合に比べ、精度の向上が報告されている。このことから特徴量抽出は機械学習を用いる際に非常に重要であると言える。しかし、移動手段推定で用いる移動軌跡がもつ情報は少なく、有効な特徴量抽出は難しい。このため速度などの基本特徴を組み合わせた複雑な特徴量は Zheng らが設計した basic feature や advanced feature 以外にも存在し、それを網羅しきれていないと考えた。そこで特徴量を自動で抽出することができる表現学習に着目し、速度などの基本的な特徴量を自動で組み合わせることで網羅しきれていない特徴量抽出を行う。

3. 深層学習による特徴量抽出

入力ベクトルから移動手段推定に有効な特徴を自動抽出する手法として、深層学習があげられる。深層学習はニューラルネットワークの隠れ層が多層となったものであり、従来のニューラルネットワークに比べ、高い表現力をもつ。

この深層学習による特徴量抽出に用いる入力ベクトルについて定義する。先行研究の知見から速度や時間、距離が移動手段を区別する重要な要素であり、またこれらを組み合わせた特徴量がユーザの行動分類においても重要であると本研究は仮定する。この仮定のもと、本研究では Zheng らの手法によって分割されたセグメントを書く移動手段推定に割り当てるタスクを考える。そこでセグメントに含まれている測位点 $point = \{point_1, point_2, \dots, point_i, \dots, point_{m+1}\}$ (ただし $point_i = \{\text{緯度}, \text{経度}, \text{時刻情報}\}$) において、連続する測位点 $point_i$ と $point_{i+1}$ の緯度、経度、時刻情報を用いて、距離 ($Distance_i$)、時間 ($Time_i$)、速度 ($Velocity_i$) を算出し、3 次元

²<http://research.microsoft.com/en-us/projects/GeoLife/>

ベクトル

$$gps_i = (Distance_i, Time_i, Velocity_i) \quad (1)$$

と表す。各測位点が3次元ベクトル gps_j をとっているため、 $m+1$ 点の測位点が含まれているセグメントは $3m$ 次元ベクトルとして表される。本研究では、このベクトルを入力ベクトルとする。

しかし、この入力ベクトルに深層学習の一般的な手法を適用するのは難しい。例えば、深層学習を実現する代表的な手法の一つとして **autoencoder** があげられる。**Autoencoder** は入力ベクトルを圧縮して特徴量を抽出し、圧縮した特徴量から復元が可能な特徴量を抽出できる性質をもつ。そして、**autoencoder** の拡張である **stacked autoencoder (SDA)** [3] をはじめとした多くの深層学習の手法は、入力データの次元数を一意に指定しなければならない。³ 移動手段推定において本研究が用いる入力ベクトルは $3m$ 次元ベクトルであるが、これはセグメントに含まれている測位点の個数に依存している。そのため、次元を一意に決めることができず、SDA などの手法を移動手段推定に適用することができない。

3.1 Recursive Autoencoder

前述の問題を解決するため、本研究では **recursive autoencoder (RAE)** [4, 16] を用いる。RAE は Socher ら [19] により、自然言語処理において応用されている。自然言語において、一文に含まれる単語数は文によって異なっている。このため一文を入力ベクトルとして扱った場合、文によってその次元数は異なるが、Socher らは RAE を用いることで対処している。したがって、自然言語と同様に入力ベクトルであるセグメントに含まれる測位点の個数が異なる本研究において、この RAE は有効な手法であるといえる。

図 2 に RAE を移動軌跡に適用した概要を示す。一つのセグメントに含まれる m 個の測位点間の特徴量を $gps = \{gps_1, gps_2, \dots, gps_m\}$ とする。RAE は二分木構造をもち、最小の二分木は $(p \rightarrow c_1, c_2)$ と表すことができ、図 2 に当てはめた場合、次のように示される： $((GPS_1 \rightarrow gps_1, gps_2), (GPS_2 \rightarrow gps_3, gps_4), (GPS_3 \rightarrow GPS_1, GPS_2))$ 。 $GPS_j (1 \leq j \leq m-1)$ はニューラルネットワークにおける隠れ層にあたり、 gps_i と同じ次元のベクトルである。図 2 の GPS_1 は、

$$p = f(W^{(1)}[c_1; c_2] + b^{(1)}) \quad (2)$$

によって求まる。重みは $W^{(1)} \in \mathcal{R}^{d \times 2d}$ であり、 d は測位点のベクトル次元数を表しており、本研究では $d = 3$ となる。 $b^{(1)}$ はバイアスであり、活性化関数には \tanh や sigmoid などを用いる。次にこの d 次元ベクトルが有効的に次元圧縮行われているかを評価するために復元層で、

$$[c'_1; c'_2] = W^{(2)}p + b^{(2)} \quad (3)$$

によって復元を行う。そして入力ベクトル $[c_1; c_2]$ と復元されたベクトル $[c'_1; c'_2]$ の誤差を

$$E_{rec} = \frac{1}{2} \|[c_1; c_2] - [c'_1; c'_2]\|^2 \quad (4)$$

で求める。そしてこの誤差が最小になるように、確率的勾配法を用いて重み W とバイアス b を最適化する。RAE は、ある木構造において全ての入力と復元の誤差を最小化することが目的となる。そこで、 gps から構築可能な木構造全ての集合を $A(gps)$ で表し、 $T(y)$ を木構造 y から全ての最小となる二分木 s を返す関数とする。つまり、

$$RAE(gps) = \arg \min_{y \in A(gps)} \sum_{s \in T(y)} E_{rec}([c_1; c_2]_s) \quad (5)$$

³例えば入力データが画像である場合、画像のサイズを $28 * 28 = 784$ ピクセルと指定しなければならない。

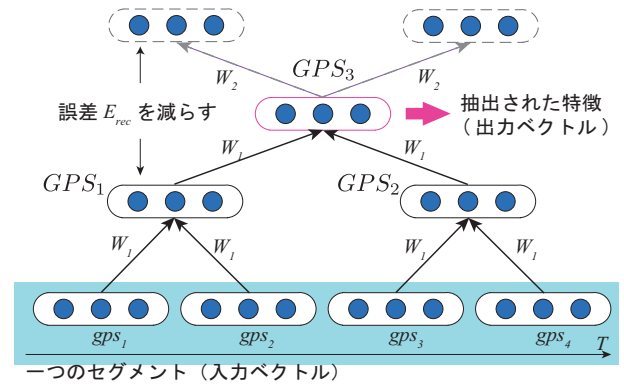


図 2: Recursive autoencoder の概要。一つのセグメントに含まれる各測位点は3次元で表現される。二つの任意の測位点において誤差 E_{rec} を算出し、最小となる測位点の組合せを選ぶ。最終的に木構造 RAE が決定するまで繰り返し行われる。

のように、誤差が最小となる木構造 $RAE(x)$ を採用する。以上のように、セグメント毎に次元数が異なる本研究において RAE を適用することが可能であることがわかる。

3.2 N-Transformed Supervised Recursive Autoencoder

式 (2) や図 2 のように各セグメントに RAE を適用した場合、入力するセグメントの $3m$ 次元ベクトルは3次元に圧縮される。セグメントに含まれる測位点間の個数 n が小さい場合は次元圧縮量は小さいが、大きい場合は次元圧縮量が大きくなってしまふ。その結果、入力ベクトルに含まれていた情報の多くが失われてしまふ問題が存在する。また RAE は教師なし学習であるため、抽出される特徴量は移動手段に固有なものではなく、汎用的なものとなっている。

これらの問題を解決するため、**n-transformed supervised recursive autoencoder (NTS-RAE)** を提案する。NTS-RAE は、次元圧縮の度を任意で決めることができる次元特徴圧縮調節と、移動手段推定により適合するための教師あり学習 [19] という二つの機能をもつ。

次元特徴圧縮調節。 一つのセグメントにおいて、 gps_i はそれ以前の $gps_1 \sim gps_{i-1}$ によって決定されると仮定し、図 3 に示すように、最小となる部分木を $(GPS_i \rightarrow GPS_{i-1}, gps_i)$ に固定する。次にセグメントの移動軌跡に乱数生成された n 次元ベクトルを加えて入力ベクトルとする。これによって出力される次元数を任意の n 次元に設定することが可能となる。出力されるベクトルの次元圧縮の調節が可能になることで、特徴の表現力を不必要に落とさずに入力データを圧縮することができる。

教師あり学習。 移動手段特有の特徴を抽出させやすくするため、正解ラベルを用いた学習を組み合わせる。Socher ら [19] が提案する手法では、全ての p ベクトルにおいて学習を行っている。本研究では計算量を抑えるため、最後に出力される GPS ベクトルのみ学習を行う。以下の式によって学習が行われる。

$$d(p) = \text{softmax}(W^{\text{label}} p) \quad (6)$$

p は最終的に出力される n 次元ベクトルであり、セグメントの特徴量を表現している。移動手段のラベルが K 種類あった場合、 $d \in \mathcal{R}^K$ は K 次元多項分布となっており、 $\sum_{k=1}^K d_k = 1$ (d_k は多項分布 d の k 番目の移動手段であると推定された確率を示す) が成り立っている。また $d_k = p(k|[c_1; c_2])$ は条件付き確率であることから、誤差はクロスエントロピー誤差となり、以下の式で算出される。

$$E_{CE}(p, t) = - \sum_{k=1}^K t_k \log d_k(p) \quad (7)$$

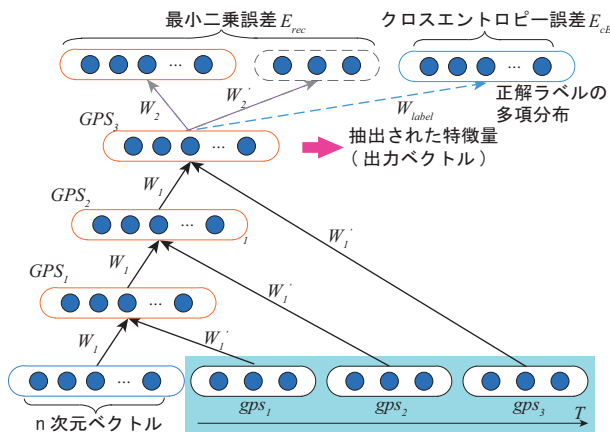


図 3: 提案手法 : n-transformed supervised recursive autoencoder. 従来の recursive autoencoder の木構造を図のように固定し、入力層に乱数生成された n 次元ベクトルを追加する。さらに教師あり学習するために、最小二乗誤差に加えて、出力される推定ラベルの多項分布と、正解ラベルの多項分布とのクロスエントロピー誤差の両方を最小化するようにモデルを学習する。

t_k は正解ラベルの多項分布 t における k 番目の要素を表しており、具体的には 2 つ目のラベルが正解であった場合、 $t = (0, 1, 0, 0, \dots)$ と表される。算出された E_{cE} と式 (4) を組み合わせて、

$$E([\mathbf{c}_1; \mathbf{c}_2], \mathbf{p}, t) = (1 - \alpha)E_{rec}([\mathbf{c}_1; \mathbf{c}_2]) + \alpha E_{cE}(\mathbf{p}, t) \quad (8)$$

という誤差が最小になるように、パラメータを最適化する。なお α は教師あり学習の割合を調節するパラメータであり、 $\alpha = 0.0$ は学習なし、 $\alpha = 1.0$ は全て学習を意味している。

4. 評価実験

提案手法によって、各セグメントにおける特徴を自動で抽出し、この抽出された特徴をもとに移動手段を推定する。この工程を評価するため、実験 1) 提案手法によって自動抽出された特徴量を用いた移動手段推定の精度評価と実験 2) 提案手法で追加した二つの機能による精度の変化について検証を行った。

4.1 データセット

本実験ではデータセットに Microsoft GeoLife データセット [24, 25, 26] を用いた。GeoLife データセットは GPS データの移動軌跡であり、65 人のユーザから 10ヶ月間収集された。このデータセットに含まれているセグメント数は 9111 個となっている。また各セグメントにはユーザによってラベリングされており、そのラベリング (移動手段) は 11 種類 (表 3) となっている。この GeoLife データセットを全セグメントでランダムシャッフルした後、訓練用 (Train)、批准用 (Validation)、検証用 (Test) の三つに Train : Validation : Test = 7 : 1 : 2 に分割して用いた。

4.2 実験 1: 移動手段推定の精度

提案手法によって自動抽出された特徴量を用いた移動手段推定の精度評価を行う。Zheng らによって提案された basic feature [26] と自動抽出された特徴量の以下の組み合わせによる精度の比較をそれぞれ行った。

- basic feature
- 自動抽出された特徴量
- basic feature + 自動抽出された特徴量

パラメータについては Train データを用いて学習し、Validation データにより最適化を行う。最適化されたパラメータで Test データを予測し、その精度の評価を行う。なお精度は、

$$\text{精度} = \frac{\text{正解した移動手段推定数}}{\text{全ての移動手段推定数}}$$

表 2: 自動抽出特徴の精度

データ	分類器		
	SVM	DT	LR
手動抽出特徴 [26]	45.3%	75.0%	65.0%
自動抽出特徴	48.9%	44.8%	50.1%
手動抽出 + 自動抽出特徴	55.6%	76.6%	69.9%

表 3: 移動手段別の推定精度

移動手段		特徴量	
名前	データ数	手動設計	手動設計 + 自動抽出
Taxi	497	32%	35%
Walk	3749	88%	89%
Bus	1800	66%	68%
Car	774	64%	67%
Bike	1540	78%	79%
Subway	571	62%	65%
Train	154	67%	80%
Airplane	14	50%	100%
Run	3	0%	0%
Boat	7	0%	0%
Motorcycle	2	0%	0%
マイクロ平均		75%	77%
マクロ平均	計 9111	46%	53%

によって求める。

3 種類の教師あり学習である SVM, DT, そしてロジスティックス回帰 (LR) を用いて、各セグメントの移動手段推定を行った。なお SVM のカーネルには RBF カーネルを用い、その他のパラメータやその他の手法のパラメータについてはグリッドサーチによって決定した。

実験の結果を表 2 と表 3 に示す。表 2 は移動手段全てにおける精度を示したものである。表 2 において、Zheng らの basic feature を DT に用いた精度は 75% である。この Zheng らの basic feature に加え、本提案手法によって自動抽出された特徴量を加えたところ、DT による精度は 76.6% であり、Zheng らの実験結果よりも 1.6% の精度の向上が確認できた。またその他の分類器 (SVM, LR) においても精度向上が確認できた。以上のことから提案手法により、手動抽出できていなかった特徴量が抽出でき、移動手段をより精度良く推定できることがわかった。

次に表 3 は移動手段別に精度を算出したものである。分類器には、最も精度が良かった DT を用いた。Taxi/Car それぞれにおいて、basic feature よりも 3% 精度が向上しており、本手法によって抽出された特徴量がこのタクシーと車を分類する特徴を捉えていると考えられる。その他の移動手段についても、本手法によって抽出された特徴量を組み合わせることで精度が向上している。

また表 3 に移動手段ごとに算出した精度の平均であるマクロ平均と、全てのセグメントの精度を算出するマイクロ平均を示す。二つの平均両方で basic feature のみよりも提案手法によって自動抽出された特徴量を組み合わせた場合のほうが精度が高くなっている。以上のことから今まで推定しにくかった移動手段が、本手法によって自動抽出された特徴量を組み合わせることで、より細かく推定可能であるといえる。

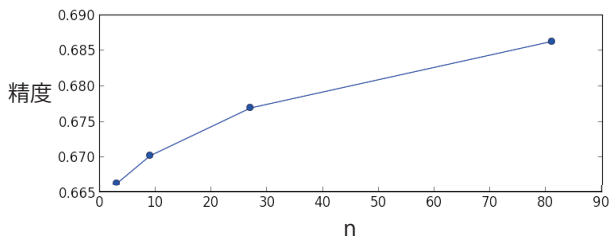


図 4: 出力ベクトルの次元数 n を変えたときの精度. 圧縮の度合 n を変えたときの精度の変化を表す. n が大きくなるほど入力ベクトルが圧縮される度合は減り, 精度が向上している.

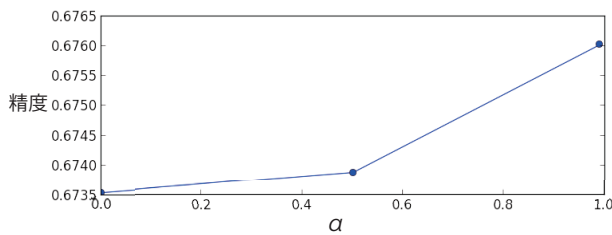


図 5: 教師あり学習度合 α を変えたときの精度. 教師あり学習の度合 α を変えたときの精度の変化を表す. α が大きくなるほど教師あり学習になり, 精度が向上している.

4.3 実験 2: 提案手法の機能評価

提案した自動抽出手法の機能それぞれについて評価⁴を行った.

N次元特徴圧縮調節. 出力されるベクトルの次元数の増加による移動手段推定の精度変化を検証した. 従来の RAE の場合, $n=3$ となり, 本実験では, $n = \{3, 9, 27, 81\}$ の 4 種類の数値を扱った. 実験結果を図 4 に示す. 図 4 から n の値が増加するにつれて, 移動手段推定の精度も向上していることがわかる. このことから入力ベクトル次元圧縮の調節によって有効な特徴力を抽出することが可能であることがわかる.

教師あり学習. 正解ラベルの学習による移動手段推定の精度変化を検証した. 学習度合として $\alpha = \{0.01, 0.5, 1.0\}$ を用いた. 実験の結果を図 5 に示す. 図 5 において, α の値が 1.0 に近づくにつれて精度が向上していることがわかる. このことから, 正解ラベルを用いて学習させることで, より移動手段推定しやすくなる特徴を抽出できていることが期待できる.

4.4 議論

提案手法によって, 従来よりも細かく移動手段を推定することが可能であることを確認した. 本研究では抽出する特徴量の次元数を 3 次元から 81 次元の範囲で設定し, 実験を行った. 特徴量の次元数を 81 次元以上に増やすことで, さらなる精度の向上が期待される. 今後の課題としてパラメータの探索範囲を広げて実験を行う予定である. 特徴量の次元数に加え, 学習回数なども細かくチューニングができていないため, これも今後の課題である. また Socher ら [19] は教師あり学習ではなく, 半教師あり学習にすることにより, ラベルなしデータも利用することで精度が向上すると述べている. 本手法においてラベルなしデータを利用することも今後の展望としてあげられる. 提案手法によって移動軌跡から特徴量を抽出したが, その他のコンテキスト解析研究における特徴量抽出に応用することが今後の展望としてあげられる.

5. 関連研究

センサ情報からユーザのコンテキストを解析・推定する研究は数多く存在する. GPS では測定が難しい建物内におけるユーザの位置を推定するため, Saha ら [17] や Sun ら [21] は複数のアクセ

スポイントを用いて, 接続されたアクセスポイントまでの距離などからユーザがいる部屋の特定位置の推定を行っている. その他にも Kaltioka ら [10] や Xu ら [22] は MICAz mote⁵ などのセンサノードを室内に設置することで, ユーザの室内における位置の推定を行っている. またユーザの行動を推定する研究としてのユーザにウェアラブルセンサやデバイスを装着させ, また物に RFID タグを貼り付けることによって, ユーザがどの道具を持って何をしているのか推定する Helaoui ら [9] や Stikic ら [20], Longstaff ら [13] のものがあげられる. その他には Cho ら [5] は LBSN を用いて, Liao ら [12] は GPS データを用いてユーザの活動を推定する研究などがあげられる. 近年の多くのスマートフォンに搭載されているジャイロセンサ [15] や SNS のチェックイン機能 [23] を用いてユーザの移動先を予測する研究などが行われている. これら多くの研究は GPS 以外のセンサから得られるデータも利用している一方で, 本研究は GPS センサのみを用いてユーザの移動手段を推定するタスクに着目している. 利用するセンサの数を減らすことで, デバイスの負荷を減らすことが可能である.

GPS センサから得られるデータのみを用いた移動手段推定の既存研究として, Zheng らは GPS データから成る移動軌跡からまず 10 種類の特徴量である basic feature [26] を設計し, さらに basic feature を組み合わせた advanced feature [24] を設計している. そして設計した特徴量を用いて, ユーザが車や電車など移動に利用している手段を推定している. 特徴量を増やすことで, 推定精度を向上させることができていた. しかしながら, タスクに有効な様々な特徴量を人手で設計するのは限界がある. この特徴量の設計を自動で行う手法として深層学習があげられる. 深層学習の関連手法は画像認識, 音声認識や自然言語の分野においても効果的に適用されているものの [6, 7, 8, 11, 14, 18], 移動手段への適用例はあまり存在しない. 本研究においては移動軌跡に深層学習を適用し, 特徴量を自動で抽出することで, 精度の向上を確認した.

6. おわりに

本研究では, 移動軌跡からの移動手段推定を行った. 従来は特徴量の設計を feature engineering のみに頼っていた一方で, 我々は特徴量を自動で抽出できる深層学習を移動手段推定に適用した. 移動軌跡を入力データとするため, 深層学習の実装手法の一つである recursive autoencoder を拡張し, 次元圧縮調節と正解ラベル学習の機能をもつ n-transformed supervised recursive autoencoder を提案した. 提案手法によって抽出した特徴量を用いて移動手段推定を行い, 実データセットを用いて評価を行った. その結果, basic feature と提案手法によって抽出した特徴量を組み合わせることで推定精度を向上させることができた. 今後の課題としては, advanced feature との比較, 半教師あり学習への拡張などがあげられる. また, 提案手法を移動手段以外のコンテキスト推定に応用することも期待される.

【文献】

- [1] Bengio, Y.: Learning deep architectures for AI, *Foundations and trends® in Machine Learning*, Vol. 2, No. 1, pp. 1–127 (2009).
- [2] Bengio, Y., Courville, A. and Vincent, P.: Representation learning: A review and new perspectives, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 35, No. 8, pp. 1798–1828 (2013).
- [3] Bengio, Y., Lamblin, P., Popovici, D. and Larochelle, H.: Greedy layer-wise training of deep networks, *Advances in neural information processing systems*, Vol. 19, p. 153 (2007).
- [4] Bengio, Y. and LeCun, Y.: Scaling learning algorithms towards AI, *Large-scale kernel machines*, Vol. 34, pp. 1–41 (2007).

⁵<http://www.openautomation.net/page/productos/id/22/title/MICAz-2.4-GHz>

⁴実験効率化のため, パラメータのグリッドサーチはせず, 固定した.

- [5] Cho, E., Myers, S. A. and Leskovec, J.: Friendship and mobility: user movement in location-based social networks, *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, pp. 1082–1090 (2011).
- [6] Collobert, R. and Weston, J.: A unified architecture for natural language processing: Deep neural networks with multi-task learning, *Proceedings of the 25th international conference on Machine learning*, ACM, pp. 160–167 (2008).
- [7] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K. and Fei-Fei, L.: Imagenet: A large-scale hierarchical image database, *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, IEEE, pp. 248–255 (2009).
- [8] Ghoshal, A., Swietojanski, P. and Renals, S.: Multilingual training of deep neural networks, *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, IEEE, pp. 7319–7323 (2013).
- [9] Helaoui, R., Riboni, D. and Stuckenschmidt, H.: A probabilistic ontological framework for the recognition of multilevel human activities, *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, ACM, pp. 345–354 (2013).
- [10] Kaltiokallio, O., Bocca, M. and Patwari, N.: Enhancing the accuracy of radio tomographic imaging using channel diversity, *Mobile Adhoc and Sensor Systems (MASS), 2012 IEEE 9th International Conference on*, IEEE, pp. 254–262 (2012).
- [11] Krizhevsky, A., Sutskever, I. and Hinton, G. E.: Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, pp. 1097–1105 (2012).
- [12] Liao, L., Fox, D. and Kautz, H.: Extracting places and activities from gps traces using hierarchical conditional random fields, *The International Journal of Robotics Research*, Vol. 26, No. 1, pp. 119–134 (2007).
- [13] Longstaff, B., Reddy, S. and Estrin, D.: Improving activity classification for health applications on mobile devices using active and semi-supervised learning, *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on-NO PERMISSIONS*, IEEE, pp. 1–7 (2010).
- [14] Mikolov, T., Karafát, M., Burget, L., Cernocký, J. and Khudanpur, S.: Recurrent neural network based language model., *INTERSPEECH*, pp. 1045–1048 (2010).
- [15] Ouyang, R. W., Srivastava, A., Prabahar, P., Roy Choudhury, R., Addicott, M. and McClernon, F. J.: If you see something, swipe towards it: crowdsourced event localization using smartphones, *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, ACM, pp. 23–32 (2013).
- [16] Poultney, C., Chopra, S., Cun, Y. L. et al.: Efficient learning of sparse representations with an energy-based model, *Advances in neural information processing systems*, pp. 1137–1144 (2006).
- [17] Saha, S., Chaudhuri, K., Sanghi, D. and Bhagwat, P.: Location determination of a mobile device using IEEE 802.11 b access point signals, *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*, Vol. 3, IEEE, pp. 1987–1992 (2003).
- [18] Seide, F., Li, G. and Yu, D.: Conversational Speech Transcription Using Context-Dependent Deep Neural Networks., *Interspeech*, pp. 437–440 (2011).
- [19] Socher, R., Pennington, J., Huang, E. H., Ng, A. Y. and Manning, C. D.: Semi-supervised recursive autoencoders for predicting sentiment distributions, *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, pp. 151–161 (2011).
- [20] Stikic, M., Van Laerhoven, K. and Schiele, B.: Exploring semi-supervised and active learning for activity recognition, *Wearable Computers, 2008. ISWC 2008. 12th IEEE International Symposium on*, IEEE, pp. 81–88 (2008).
- [21] Sun, G., Chen, J., Guo, W. and Liu, K. R.: Signal processing techniques in network-aided positioning: a survey of state-of-the-art positioning designs, *Signal Processing Magazine, IEEE*, Vol. 22, No. 4, pp. 12–23 (2005).
- [22] Xu, C., Firner, B., Moore, R. S., Zhang, Y., Trappe, W., Howard, R., Zhang, F. and An, N.: Sepl: Indoor device-free multi-subject counting and localization using radio signal strength, *Proceedings of the 12th international conference on Information Processing in Sensor Networks*, ACM, pp. 79–90 (2013).
- [23] Yang, D., Zhang, D., Yu, Z. and Yu, Z.: Fine-grained preference-aware location search leveraging crowdsourced digital footprints from LBSNs, *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, ACM, pp. 479–488 (2013).
- [24] Zheng, Y., Chen, Y., Li, Q., Xie, X. and Ma, W.-Y.: Understanding transportation modes based on GPS data for web applications, *ACM Transactions on the Web (TWEB)*, Vol. 4, No. 1, p. 1 (2010).
- [25] Zheng, Y., Li, Q., Chen, Y., Xie, X. and Ma, W.-Y.: Understanding mobility based on GPS data, *Proceedings of the 10th international conference on Ubiquitous computing*, ACM, pp. 312–321 (2008).
- [26] Zheng, Y., Liu, L., Wang, L. and Xie, X.: Learning transportation mode from raw gps data for geographic applications on the web, *Proceedings of the 17th international conference on World Wide Web*, ACM, pp. 247–256 (2008).

河野 慎 Makoto KAWANO

東京大学大学院学際情報学府在学中。2010年慶應義塾大学環境情報学部環境情報学科卒業。同年東京大学大学院学際情報学府修士課程に入学。主に、ユビキタスコンピューティングシステム、サイバーフィジカルシステム、データマイニングの研究に従事。

遠藤 結城 Yuki ENDO

日本電信電話株式会社, NTT サービスエボリューション研究所所属。2012年筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻博士前期課程修了。同年日本電信電話株式会社に入社。以来、データマイニングの研究開発に従事。情報処理学会、電子情報通信学会、各会員。

戸田 浩之 Hiroyuki TODA

日本電信電話株式会社, NTT サービスエボリューション研究所所属。1997年名古屋大学工学部材料プロセス工学科卒業。1999年同大学院工学研究科材料プロセス工学専攻博士課程前期課程修了。同年日本電信電話株式会社に入社。以来、情報検索、データマイニングの研究開発に従事。2007年筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻博士後期課程修了。博士(工学)。ACM、情報処理学会、電子情報通信学会、各会員。

小池 義昌 Yoshimasa KOIKE

日本電信電話株式会社, NTT サービスエボリューション研究所所属。1989年東北大学大学院材料化学専攻博士前期課程修了後、日本電信電話株式会社に入社。以来、パターン認識の研究、遠隔教育システムの研究開発、地域情報検索サービスの研究に従事。

植田 一博 Kazuhiro Ueda

東京大学大学院総合文化研究科教授(大学院情報学環を兼任)。1988年東京大学教養学部卒業。1993年東京大学大学院総合文化研究科博士課程修了。博士(学術)。創造性研究、日本伝統芸能の技の分析、意思決定研究などに従事。第7回ドコモ・モバイル・サイエンス賞・奨励賞(2008年)、日本認知科学会論文賞(2004年、2007年、2012年)などを受賞。人工知能学会、電子情報通信学会、日本認知科学会、日本心理学会、日本認知心理学会、行動経済学会、社会情報学会、Cognitive Science Society 各会員。