

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

河野 慎[†] 遠藤 結城^{††} 戸田 浩之^{††} 小池 義昌^{††} 植田 一博[†]

[†] 東京大学大学院情報学環・学際情報学府 〒113-8654 東京都文京区本郷7丁目3番1号

^{††} 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

〒239-0847 神奈川県横須賀市光の丘1丁目1番

E-mail: [†]qq146403@iii.u-tokyo.ac.jp

あらまし 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をする。移動手段推定を行う際に、従来の手法では移動軌跡から人が経験にもとづいて分類に有効な特徴量を考えていた。しかし人手による有効な特徴量抽出は時間がかかり、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 representation learning の一つである deep learning に着目する。Deep learning は、ニューラルネットワークを多層にし、従来よりも表現力を向上させたものであり、本研究では実装手法の一つである recursive autoencoder を用いた。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 データセットに本手法を適用させた結果、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した。

キーワード Deep Learning, Recursive Autoencoder, 移動手段推定, 移動軌跡

Makoto KAWANO[†], Yuki ENDO^{††}, Hiroyuki TODA^{††}, Yoshimasa KOIKE^{††},

and Kazuhiro UEDA[†]

[†] Interfaculty Initiative in Information Studies
and Graduate School of Interdisciplinary Information Studies

University of Tokyo

Hongo, Bunkyo, Tokyo, Japan

^{††} NTT service evolution laboratories, NTT Corporation

Yokosuka, Kanagawa, Japan

E-mail: [†]qq146403@iii.u-tokyo.ac.jp

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、ユーザの状況の推定や判定をする研究は多くされている。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 [19] 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 [16] や SNS のチェックイン機能 [26] を用いてユーザの移動先を予測するものなどが挙げられる。

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その移動手段を推定するタスク（移動手段推定）である [29]。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に応用することが可能となる。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解くことが多い [5] [13]。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のデータから特徴量

を人が設計する必要がある。この工程は feature engineering [1] と呼ばれる。移動手段判定の一部である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 [29] とこれらの基本的な特徴量を組み合わせた heading change rate^(注1) などの応用的な特徴量 [27] 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手法を用いることで、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 (representation learning) [2] があげられる。Representation learning は与えられたデータから次元数を変える特徴量抽出技術の総称であり、中でもニューラルネットワークを拡張した深層学習 (deep learning) [1] が近年注目されている。Deep learning は画像認識 [7] や音声認識 [20]、自然言語処理 [21] 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 deep learning によって抽出する。しかし多くの deep learning 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。移動軌跡は、移動手段によって所要時間が異なるため、移動手段ごとの測位点の数も異なる。そのため測位点にもとづいた入力ベクトルの場合、移動手段によって次元が異なってしまう。

この異なる次元に対応するため、本研究では入力値の次元数を固定せずに利用できる recursive autoencoder (RAE) [4] [17] 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 [21] によって、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^(注2)を用いた。本実験では、まずこのデータセットに

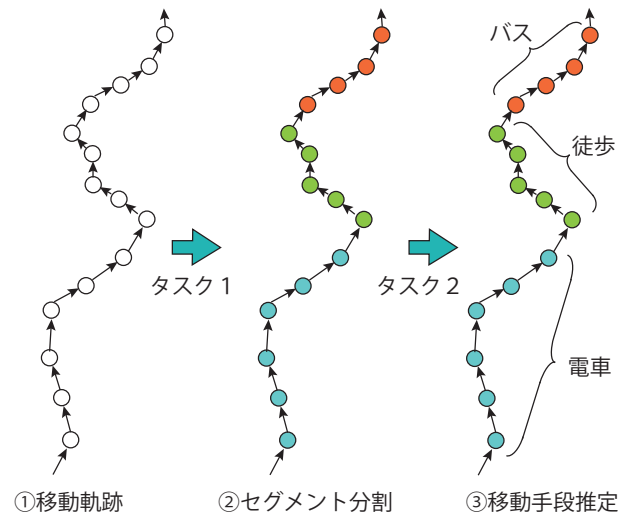


図1 移動手段判定の流れ。丸点は測位点を表している。タスク1において、各測位点の速度や加速度からセグメント（各セグメントは何らかの移動手段を表す）に分割する。そしてタスク2においてセグメントの移動手段を推定する。

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 [29] [27] の問題意識について説明する。Zheng らは、GPS ロガーを用いてユーザから移動軌跡（GPS データ）を一定期間内収集した。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 [29] 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を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

(注1)：移動軌跡における方向転換の割合

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

表 1 Zheng らによる提案特徴量

カテゴリ	特徴量
basic feature [29] 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 [27] 3 次元	速度変化率 停止率 方向変化率

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 で分割した全てのセグメントには、車や徒歩などの移動手段の一つが正解ラベルとして付与されている。タスク 2 では特徴量を抽出し、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する必要がある。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 [29] 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 [27] 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量は少なく、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 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 representation learning 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. Deep Learning 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、deep learning が挙げられる。Deep learning 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 deep learning 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 らによって分割されたセグメントをデータセットとして扱う。そこでセグメントに含まれている測位点 $point = \{point_1, point_2, \dots, point_i, \dots, point_n\}$

(ただし $point_i = \{\text{緯度, 経度, 時刻情報}\}$) 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 次元ベクトル

$$gps_j = (Distance_j, Time_j, Velocity_j)$$

と表す。各測位点が 3 次元ベクトル gps_j をととなっているため、 n 点の測位点が含まれているセグメントは $3n$ 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、この入力ベクトルを deep learning の一般的な手法に提供するのは難しい。例えば、deep learning を実現する代表的な手法の一つとして autoencoder が挙げられる。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 [3] をはじめとした多くの deep learning 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。^(注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 $3n$ 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] [17] を用いる。RAE は Socher ら [21] 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、一文に含まれてい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_i 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、

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

によって求まる。重みは $W^{(1)} \in \mathcal{R}^{n \times 2n}$ であり、 n は測位点のベクトル次元数を表しており、本研究では $n = 3$ 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 n 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、

(注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。

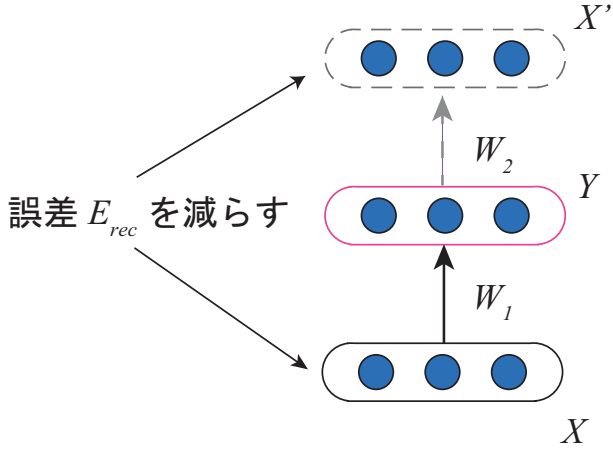


図 2 Recursive AutoEncoder の概要. 一つのセグメントに含まれる各測位点は 3 次元で表現される. 二つの任意の測位点において誤差 E_{rec} を算出し, 最小となる測位点の組合せを選ぶ. 最終的に木構造 RAE が決定するまで繰り返し行われる.

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

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

$$E_{rec} = \frac{1}{2} ||[c_1; c_2] - [c'_1; c'_2]||^2 \quad (3)$$

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,

$$RAE(gps) = \arg \min_{y \in A(gps)} \sum_{s \in T(y)} E_{rec}([c_1; c_2]_s) \quad (4)$$

のように, 最もその誤差が小さい木構造 $RAE(x)$ 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 (1) 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 $3n$ 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う. 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う問題が存在する.

こ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, 次元圧縮の度合を任意で決めることができる N 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 [21] という二つの機能をもつ.

N 次元特徴圧縮調節. 一つのセグメントにおいて, gps_i はそれまでの $gps_1 \sim gps_{i-1}$ 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 ($GPS_i \rightarrow GPS_{i-1}, gps_i$)

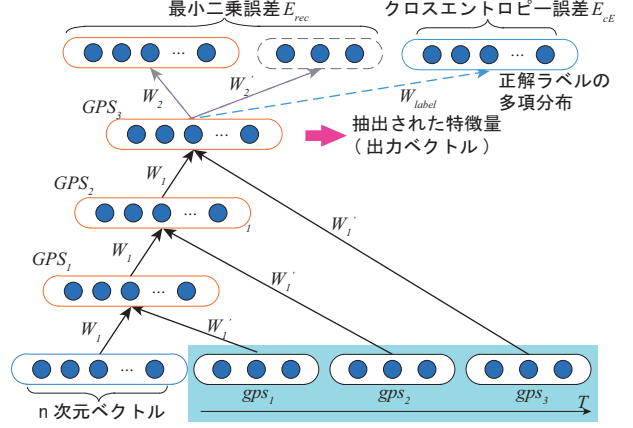


図 3 提案手法: N-Transformed Supervised Recursive AutoEncoder. 従来の Recursive Autoencoder の入力値に乱数生成された n 次元ベクトルを追加する. そして木構造は, 図のように固定する. また教師あり学習をするため, 最終的に出力される GPS ベクトルと正解ラベルの多項分布ベクトルでクロスエントロピー誤差を算出する.

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 [21] 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.

$$d(p) = \text{softmax}(W^{\text{label}}p) \quad (5)$$

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. 正解ラベルが 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, $\sum_{k=1}^K d_k = 1$ 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.

$$E_{cE}(p, t) = - \sum_{k=1}^K t_k \log d_k(p) \quad (6)$$

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, $t = (0, 1, 0, 0, \dots)$ というように表される. 算出された E_{cE} と式 (3) を組み合わせて, 最終的に出力されるベクトルのみ,

$$E([c_1; c_2], p, t) = (1 - \alpha)E_{rec}([c_1; c_2]) + \alpha E_{cE}(p, t) \quad (7)$$

という誤差を用いる. なお α は教師あり学習の度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元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 [27] [28] [29] を用いた。Microsoft 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 Microsoft 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 [29] 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、

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

によって求める。

3 種類の機械学習である SVM, DT, そして Logistic Regression (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

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 [29]	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	
	手動抽出	手動抽出 + 自動抽出
Taxi	32%	35%
Walk	88%	89%
Bus	66%	68%
Car	64%	67%
Bike	78%	79%
Subway	62%	65%
Train	67%	80%
Airplane	50%	100%
Run	0%	0%
Boat	0%	0%
Motorcycle	0%	0%
マイクロ平均	75%	77%
マクロ平均	46%	53%

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載せる。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。

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。以上のことから、自然言語処理で利用されていた従来の教師ありへの拡張が移動軌跡でも有効であることを確認できた。

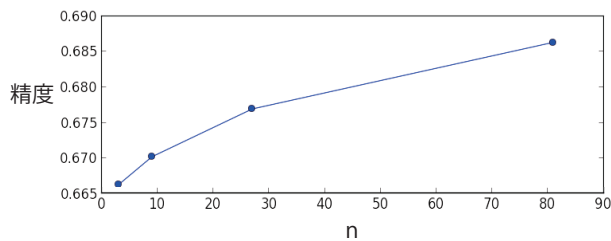


図4 圧縮度合 n 値を変えたときの精度。圧縮の度合 n を変えたときの精度の変化を表す。 n が大きくなるほど圧縮度合は減り、精度が向上している。

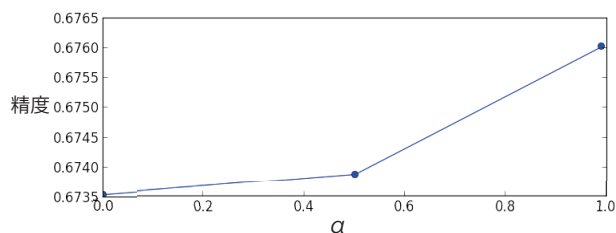


図5 教師あり学習度合 α を変えたときの精度。教師あり学習の度合 α を変えたときの精度の変化を表す。 α が大きくなるほど教師あり学習になり、精度が向上している。

4.4 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。精度が1.6%向上することで、GPS 取得サービスを利用している約100万人いた場合、約1.6万人程度のユーザの移動手段を推定することが可能となる。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 [21] 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挙げられる。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挙げられる。

5. 関連研究

まずユーザのコンテキストの解析と推定を行う研究について説明を行う。その後、deep learning が適用されている研究について説明を行う。

5.1 コンテキスト解析研究

位置推定。 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研究は数多くなされている。Saha ら [19] や Sun ら [23] 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制の推定を行っている。その他にも Kaltioika ら [10] や Xu ら [25]

は MICAz mote^(注4)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特定を行っている。

行動推定。 Helaoui ら [9] や Stikic ら [22], Longstaff ら [14] は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している。その他には Cho ら [5] は LBSN を用いて、Liao ら [13] 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挙げられる。

移動先推定。 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 [16] や SNS のチェックイン機能 [26] 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。

今回取り組んだ移動手段推定と紹介したコンテキスト解析研究には、いずれも特徴量が必要であるという共通点があることから、提案手法の応用が期待される。

5.2 Deep learning を用いた研究

画像認識。 Le ら [12] によれば、Youtube^(注5)から無作為に画像を選出し、ニューラルネットワークに適用したところ、猫の顔を学習した。また Deng ら [7] は一般物体認識コンテストである ILSVR2012 において、従来の feature engineering による画像認識よりも10%のエラー差をつけて優勝した。この他にも Krizhevsky ら [11] によって用いられていることから、deep learning は非常に有効な手法であると言える。

音声認識。 Seido ら [20] によって word error rate で従来の手法よりも deep learning の一種である deep neural network - hidden markov model(DNN-HMM) を用いたことによって認識の精度を10%前後改善された。その他にも、Ghoshal ら [8] による多言語をデータセットに使う手法などが音声認識に応用されている。

自然言語処理。 Socher ら [21] の recursive autoencoder など言語構成性のモデル化や Mikolov ら [15] の時系列に着目して提案された recurrent neural network など言語モデルの構築、Wang ら [24] の構造予測など、様々な観点で deep learning が応用されている。その他にも Collobert ら [6] や Ranzato ら [18] によって自然言語処理においても deep learning が用いられており、有効な手法であることがわかる。

Deep learning はいずれの研究においても様々な手法が提案・拡張されており、同様に提案手法が改良されることによる精度の向上が期待される。

6. ま と め

本研究では、移動軌跡からの移動手段推定を行った。移動手段推定において多くの研究で用いられる feature engineering による有効な特徴量の設計は難しい。そこで、特徴量を自動で抽出できる deep learning 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、deep learning 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

(注4) : <http://www.openautomation.net/page/productos/id/22/title/MICAz-2.4-GHz>

(注5) : <http://youtube.com/>

ラベル学習の機能をもつ N-transformed supervised recursive autoencoder を提案した。本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、評価を行った。その結果、既存手法に比べて精度を 1.6%向上させることができ、各機能の有効性も確認した。今後の課題としては、半教師あり学習への拡張、パラメータのチューニング等が挙げられ、その他のコンテキスト解析研究への応用が期待される。

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] 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 multitask 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] Le, Q. V.: Building high-level features using large scale unsupervised learning, *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, IEEE, pp. 8595–8598 (2013).
- [13] 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).
- [14] 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).
- [15] Mikolov, T., Karafiát, M., Burget, L., Cernocký, J. and Khudanpur, S.: Recurrent neural network based language model., *INTERSPEECH*, pp. 1045–1048 (2010).
- [16] Ouyang, R. W., Srivastava, A., Prabakar, 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).
- [17] 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).
- [18] Ranzato, M. and Szummer, M.: Semi-supervised learning of compact document representations with deep networks, *Proceedings of the 25th international conference on Machine learning*, ACM, pp. 792–799 (2008).
- [19] 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).
- [20] Seide, F., Li, G. and Yu, D.: Conversational Speech Transcription Using Context-Dependent Deep Neural Networks., *Interspeech*, pp. 437–440 (2011).
- [21] 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).
- [22] 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).
- [23] 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).
- [24] Wang, S. and Manning, C.: Fast dropout training, *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, pp. 118–126 (2013).
- [25] Xu, C., Firner, B., Moore, R. S., Zhang, Y., Trappe, W., Howard, R., Zhang, F. and An, N.: Scpl: 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).
- [26] 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).
- [27] 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).
- [28] 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).
- [29] 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).