

トピックモデルに基づく人行動分析技術

Human Action Analysis based on Topic Model

八田 俊之
Toshiyuki Hatta

三輪 祥太郎
Shotaro Miwa

1. はじめに

近年、スマートフォンやウェアラブルセンサの普及により、加速度センサを用いた人の行動センシングが活発になっている。また、IoT (Internet of Things) や M2M (Machine to Machine) による家電・産業機器のネットワーク化に伴い、日常生活における人の行動を予測した機器の制御を目指した研究開発が盛んになっている。行動予測のためには、認識対象となる行動を限定しない行動モデルが重要となる。

従来の行動認識研究において、行動モデルは識別器として構築されてきた。識別器としての行動モデルには、対象行動のラベル付けされたデータベース作成と、このデータベースの学習(教師あり学習)が必要である。しかし、ラベル付き行動データベースの作成には多大な手間がかかるだけでなく、そもそも日常生活における行動について対象行動を定義すること自体が困難である。

本稿では上記問題解決のため、教師なし学習の一つであるトピックモデルに基づく行動分析モデルを提案する。トピックモデルは自然言語処理分野で利用されており、単語の頻度分布で表された文書から、特定の単語の頻度分布で表されるトピックを推定する。トピックは料理・政治・スポーツといったジャンルに相当し、文書は各トピックの混合として表現できる。

提案モデルでは、行動によって生成されたセンサ値を単語、時間窓で区切られた行動を文書として、行動にトピックモデルを適用する。これにより、センサ値の頻度分布で表された行動から、特定のセンサ値の頻度分布で表される行動トピックを推定する。行動トピックは基本的な行動のジャンルに相当し、行動は各行動トピックの混合として表現できると考えられる。

以下、2章では、本稿で提案するトピックモデルに基づく行動分析モデルについて述べる。3章では、提案モデルにより行動データベースから推定された行動トピックについて述べる。4章では、提案モデルにより推定された行動トピックが意味する行動のジャンルについて考察する。5章では、結論について述べる。

2. トピックモデルに基づく行動分析

2.1 概要

本稿では、代表的なトピックモデルである潜在的ディリクレ配分法(LDA; Latent Dirichlet Allocation) [2]を用いる。LDAのグラフィカルモデルを図1に示す。

提案モデルにおいて、行動は D 個の時間窓で区切られており、区切られた行動 $d \in \{1, \dots, D\}$ によって生成されたセンサ値集合を $\mathbf{w}_d = \{w_{d,n}\}_{n=1}^{N_d}$ とする。ここで、 $w_{d,n}$ は行動 d における n 番目のセンサ値であり、1から V までの自然数をとる。 N_d は行動 d の時間窓長である。

また、行動 d は固有の行動トピック分布 θ_d によって表され、行動トピック j は固有のセンサ値分布 ϕ_j によって表される。このとき、センサ値 $w_{d,n}$ の生成過程は以下のようになる。

- (1) ディリクレ事前分布 $\text{Dir}(\alpha)$ から、行動 d を表す行動トピック分布 θ_d を生成する ($d \in \{1, \dots, D\}$)。
- (2) ディリクレ事前分布 $\text{Dir}(\beta)$ から、行動トピック j を表すセンサ値分布 ϕ_j を生成する ($j \in \{1, \dots, T\}$)。
- (3) 行動 d において、多項分布 $\text{Mult}(\theta_d)$ から、 n 番目の行動トピック $z_{d,n}$ を生成する ($n \in \{1, \dots, N_d\}$)。
- (4) 行動 d において、多項分布 $\text{Mult}(\phi_{z_{d,n}})$ から、 n 番目のセンサ値 $w_{d,n}$ を生成する ($n \in \{1, \dots, N_d\}$)。

ここで、 $\alpha = \{\alpha_j\}_{j=1}^T$ と $\beta = \{\beta_v\}_{v=1}^V$ はそれぞれ θ_d と $\phi_{z_{d,n}}$ が従うディリクレ事前分布のパラメータ(ハイパーパラメータ)である。また、行動トピック $z_{d,n}$ はセンサ値 $w_{d,n}$ を生成する行動トピックであり、1から T までの自然数をとる。

上記過程によって生成されたセンサ値 $w_{d,n}$ の頻度分布は、 T 個の行動トピックを表すセンサ値分布 ϕ_j の、行動トピック分布 θ_d を混合割合とした混合分布となる(図2)。このように、提案モデルを用いることで、行動を各行動トピックの混合として表現できる。

2.2 行動トピックを表すセンサ値分布の推定

行動トピック j を表すセンサ値分布 ϕ_j は、行動トピック集合 $\mathbf{z} = \{\{z_{d,n}\}_{n=1}^{N_d}\}_{d=1}^D$ およびハイパーパラメータより推定できる。行動トピック集合 \mathbf{z} について、本稿ではギブスサンプリング法[3]により推定する。行動トピック $z_{d,n}$ のサンプリング確率を以下に示す。

$$P(z_{d,n} = j | w_{d,n} = v, \mathbf{w}_{\setminus i}, \mathbf{z}_{\setminus i}, \alpha, \beta) \propto \frac{C_{j \setminus i}^d + \alpha_j}{C_{\setminus i}^d + \alpha} \cdot \frac{C_{v \setminus i}^j + \beta_v}{C_{\setminus i}^j + \beta} \quad (1)$$

ここで、 C_j^d は時間窓 d における行動トピック j の生成回数、 C_v^j は行動トピック j におけるセンサ値 v の生成回数であり、行動トピック \mathbf{z} により得られる。また、 $C^d = \sum_j C_j^d$ 、 $C^j = \sum_{v=1}^V C_v^j$ 、 $\alpha = \sum_{j=1}^T \alpha_j$ 、 $\beta = \sum_{v=1}^V \beta_v$ であり、 $\setminus i$ は時間窓 d の n 番目のセンサ値を除いた時の回数もしくは集合を表す。

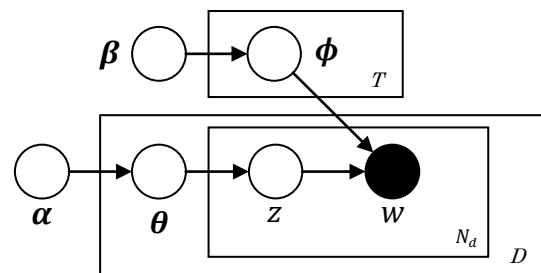


図1 LDAグラフィカルモデル

†三菱電機株式会社

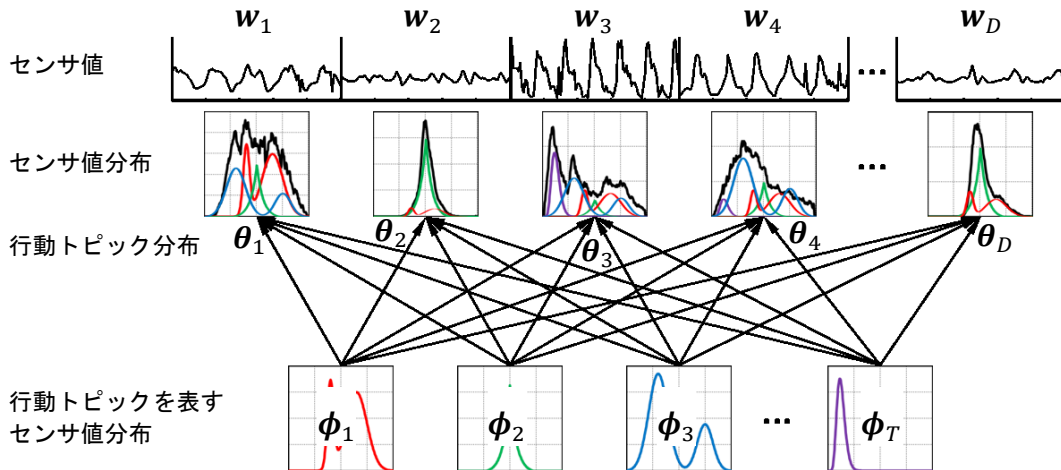


図 2 トピックモデルに基づく行動モデル

また、本稿ではハイパーパラメータ α, β について、不動点反復法[4]により行動トピック \mathbf{z} に対して最適化する。最適化の更新式を以下に示す。

$$\alpha_j \leftarrow \alpha_j \frac{\sum_{d=1}^D \Psi(C_j^d + \alpha_j) - D\Psi(\alpha_j)}{\sum_{d=1}^D \Psi(C^d + \alpha) - D\Psi(\alpha)} \quad (2)$$

$$\beta_v \leftarrow \beta_v \frac{\sum_{j=1}^T \Psi(C_v^j + \beta_v) - T\Psi(\beta_v)}{\sum_{j=1}^T \Psi(C^j + \beta) - T\Psi(\beta)} \quad (3)$$

ここで、 $\Psi(\cdot)$ はディガンマ関数である。

行動トピック \mathbf{z} のサンプリングとハイパーパラメータ α, β の最適化を交互に繰り返した後、行動トピック j のセンサ値分布 $\phi_j = \{\phi_{j,v}\}_{v=1}^V$ は以下の式を用いて推定できる。

$$\hat{\phi}_{j,v} = \frac{C_v^j + \beta_v}{C^j + \beta} \quad (4)$$

なお、行動トピック分布 θ_a は以下のように推定できる。

$$\hat{\theta}_{a,j} = \frac{C_j^d + \beta_v}{C^d + \beta} \quad (5)$$

2.3 各センサ値を生成した行動トピックの推定

行動トピック分布 θ_a は行動 d の時間窓長に含まれる行動トピックの混合割合であり、時間分解能の粗い行動の表現である。前節で推定した行動トピック集合 \mathbf{z} は、より密な時間分解能の行動の表現であるが、計算コストが問題となる。

本稿では、 T 個のセンサ値分布 $\hat{\phi}_j$ を事前に推定することで、より小さな計算コストで、より密な時間分解能の行動の表現を得る。具体的には EM アルゴリズムを用いて、センサ値集合 \mathbf{w}_d の頻度分布に占める T 個のセンサ値分布 $\hat{\phi}_j$ の混合割合として θ_a を推定する。この過程において、各行動トピックが各時刻におけるセンサ値 $w_{d,n}$ を生成した確率 $\psi_{d,n} = \{\psi_{d,n,j}\}_{j=1}^T$ を推定できる。手順を以下に示す。

(1) 行動 d における行動トピック分布の推定値 $\hat{\theta}_d$ の初期値を設定する。

(2) E-Step: 行動 d において、各行動トピックが n 番目のセンサ値 $w_{d,n}$ を生成した確率 $\psi_{d,n}$ を推定する。

$$\hat{\psi}_{d,n,j} = \frac{\hat{\theta}_{d,j} \hat{\phi}_{j,w_{d,n}}}{\sum_{k=1}^T \hat{\theta}_{d,k} \hat{\phi}_{k,w_{d,n}}} \quad (6)$$

(3) M-Step: 時間窓 d における行動トピック分布の推定値 $\hat{\theta}_d$ を推定する。

$$\hat{\theta}_{d,j} = \frac{1}{N} \sum_{n=1}^N \hat{\psi}_{d,n,j} \quad (7)$$

(4) 手順 (2), (3) を規定の回数繰り返す。

以上の手順により、各行動トピックがセンサ値 $w_{d,n}$ を生成した確率 $\hat{\psi}_{d,n}$ を得る。

3. 行動トピックを表すセンサ値分布の推定実験

本章では、提案モデルを用いて、加速度センサにより収集した行動データベースから行動トピックを推定する。また、提案モデルにより推定した行動トピックを表すセンサ値分布について、被験者への依存性を検証する。

3.1 実験方法

本稿では、スマートフォン (iPod touch) 搭載の 3 軸加速度センサを用いて、男性 10 名の日常的な行動を模擬した行動データベースを作成した。スマートフォンは被験者のズボンの右後方ポケットに挿入した。センサ値はサンプリング周波数 20 Hz で収集し、各フレームで 3 軸加速度の値からノルムを算出し、0.08 g から 2.00 g まで 0.04 g 刻みで量子化した。

本稿では行動データベースに収集する行動として、以下に示す歩行時の行動を対象とした。括弧内は継続時間と収集回数を示す。なお、収集した各行動は時間窓長 5 秒として区切った。また、時間窓長 $N_d = 100$ である。

- ・止まっている (20 秒×1回)
- ・歩いている (20 秒×1回)
- ・走っている (20 秒×1回)
- ・足踏みしている (20 秒×1回)
- ・その場で回っている (左右回り各 10 秒×1回)
- ・直径 1.5 m の円上を歩いている (左右回り各 10 秒×1回)
- ・階段を上がっている (5 秒×2回)
- ・階段を下りている (5 秒×2回)
- ・その場でジャンプしている (10 秒×2回)

また、行動トピックを表すセンサ値分布の推定において、ギブスサンプリング 1 回と、不動点反復法の繰返し 50 回

を交互に 10000 回繰り返した。これらの回数は推定値が十分収束する回数として設定した。また、経験的に行動トピック数 $T = 5$ とした。

3.2 実験結果

推定された行動トピックを表すセンサ値分布について、被験者 3 名の結果を例として図 3 に示す。図 3 において各列は、同一の被験者に関する行動トピックを表すセンサ値分布である。各行動トピックを表すセンサ値分布形状は、それぞれ異なる特徴を持つことが分かる。また各行は、異なる被験者間において、類似したセンサ値分布形状を持つと見られる行動トピックである。

以上より、提案モデルを用いることで、異なる被験者間で、それぞれ類似したセンサ値分布形状を持つと見られる 5 つの行動トピックを推定できることが分かった。

4. 行動トピックが相当する行動のジャンル考察

前章において、加速度センサによる行動データベースから、5 つの行動トピックを推定した。本章では、各時刻のセンサ値を生成した行動トピックを推定し、これらの行動トピックが相当する行動のジャンルについて、同時刻の被験者の映像と比較することで考察する。

4.1 実験方法

3 章の実験において、ウェブカメラにより加速度センサと同期した被験者の映像を収集した。また 2.3 節で述べた方法により、前節で推定した各行動トピックが各時刻におけるセンサ値を生成した確率について推定した。EM アルゴリズムの繰り返し回数は、十分に収束する回数として 20 回とした。

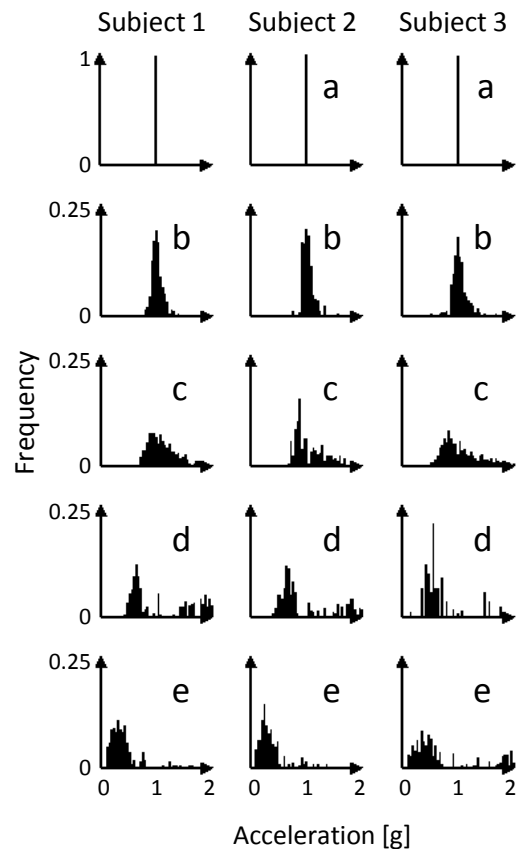


図 3 行動トピックを表すセンサ値分布の例

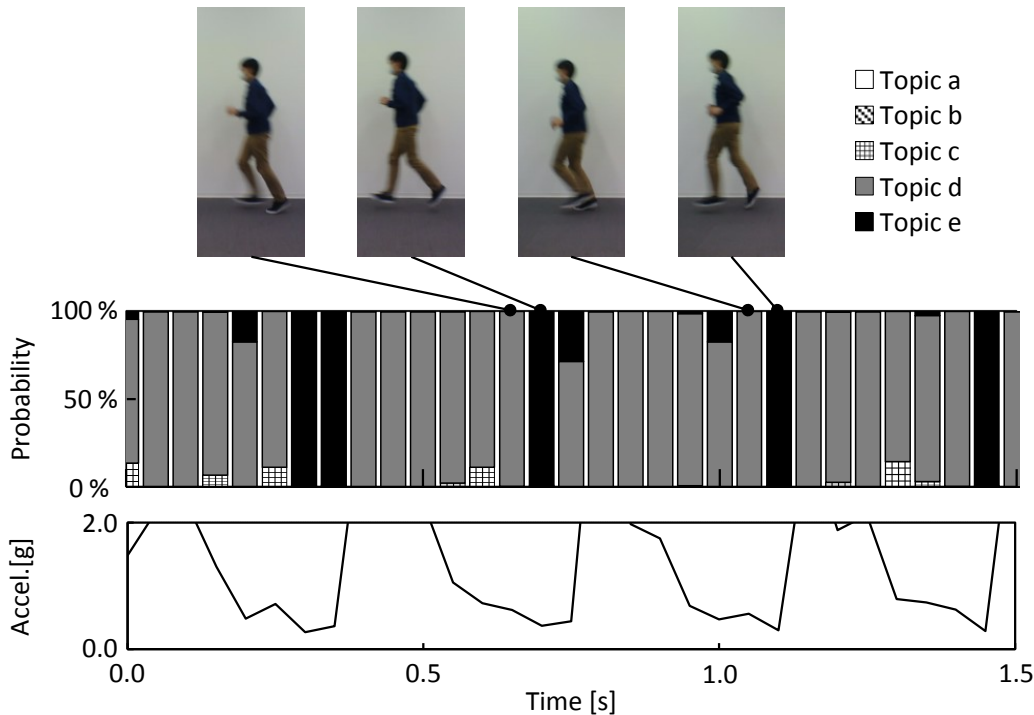


図 4 行動「走っている」における行動トピックの遷移

4.2 実験結果

結果より、各行動トピックについて、センサ値を生成した確率が高くなる時刻における被験者の状態を表 1 にまとめた。表 1 より、各行動トピックは基本的な運動状態に相当することが分かる。

結果の例として、被験者 1 の行動「走っている」より 1.5 秒間を図 4 に示す。表 1 に示すように、行動トピック d の割合が高い時刻では身体は浮いており、行動トピック e の割合が高い時刻では身体は着地している。また、行動トピック d と行動トピック e は交互に高い確率を示している。つまり、被験者 1 の行動「走っている」を、行動トピック d 「着地している」と行動トピック e 「浮いている」が交互に表れる行動として表現できていることが分かる。

以上より、提案モデルを用いることで、行動を基本的な運動状態の遷移として表現できることが分かった。

表 1 各行動トピックが相当する行動のジャンル

行動トピック	被験者の状態
a	静止している
b	前後左右に動揺している
c	足踏みしている
d	着地している
e	浮いている

5. まとめ

本稿では、加速度センサを用いた人の行動予測を目的として、教師なし手法であるトピックモデルに基づく行動分析モデルを提案した。

3 章では、スマートフォン搭載の加速度センサにより収集された行動データベースに対して提案モデルを適用した。結果、提案モデルを用いることで、異なる被験者において、それぞれ類似したセンサ値分布形状を持つと見られる 5 つの行動トピックを推定できることが分かった。

4 章では、各時刻のセンサ値を生成した行動トピックを推定し、これらの行動トピックが相当する行動のジャンルについて、同時刻の被験者の映像と比較することで考察した。結果、加速度センサによる行動データベースから推定された行動トピックは、基本的な運動状態に相当することが分かった。また、提案モデルを用いることで、行動を基本的な運動状態の遷移として表現できることが分かった。

以上より、提案モデルは認識対象となる行動を限定しない行動モデルとして有用であり、日常行動における人の行動を予測した機器の制御を実現する可能性があることが分かった。

本稿では特定の行動に基づく行動データベースを用いて実験を行った。今後、より多くの行動を含んだ行動データベースに対する提案モデルの適用を検討する。

参考文献

[1] Óscar D. Lara and Miguel A. Labrador, “A survey on human activity recognition using wearable sensors,” *IEEE Communications Surveys and Tutorials*, vol. 15, no. 3, pp. 1192-1209, 2013.

[2] David M. Blei, Andrew Y. Ng, and Michael I. Jordan, “Latent Dirichlet allocation,” *Journal of Machine Learning Research*, vol. 3, pp. 993-1022, 2003.

[3] Thomas L. Griffiths and Mark Steyvers, “Finding scientific topics,” in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, pp. 5228-5235, 2004.

[4] Thomas P. Minka, “Estimating a Dirichlet distribution,” Technical report, Massachusetts Institute of Technology, vol. 2000, pp. 1-13, 2000.