

1-14 環境認識ライフログからの行動パターン解析 による類似性・イベント検出

奥原研究室
1415048 福嶋 瑞希

1. 序論

従来のライフログは手間がかかり、GPS を使用するためプライバシーの侵害という問題を引き起こす。本研究は、多くの人に広く受け入れられるライフログとして、手間がかからず GPS を使用しないライフログデータの取得を行うことと、取得したデータから類似性やイベント性を検出することを目的とする。個人情報保護に着目したライフログのため、スマートグラスと画像認識 API を用いたリアルタイム視界情報取得アプリケーションを開発・使用し、ビッグデータ構築・多変量解析を用いて、ライフログの行動パターンの類似性・イベント検出を行う。

2. ライフログとスマートグラス

ライフログとは、個人の活動に関するログである [1]。手間のかからないライフログのため、メガネ型ウェアラブルデバイスであるシースルーモバイルビューアー MOVERIO BT-300 を使用する。また、画像認識 API を使用することで、MOVERIO のカメラ画像がもつ情報をテキストで取得できる。

3. 行動識別

KH Coder[2] と R 言語を用いて、多変量解析からライフログデータの行動を識別し、自己組織化マップから時系列を可視化する。入力層と出力層の 2 つに分かれ、入力層から出力層への入力を x ，出力層のニューロンの番号を i ，参照ベクトルを m_i と定義し、入力ベクトルと各ニューロンの参照ベクトルとの距離で出力層のニューロンを競合させ、勝者ニューロン c の参照ベクトルと入力ベクトルを近づける学習を行う [3]。

$$c = \arg \min_i \{ \|x - m_i\| \} \quad (1)$$

また、勝者ニューロンに近いニューロンも参照ベクトルを同様に近づけ、 h_{ci} は勝者ニューロンとの距離によりガウス関数で減衰する係数、 r_i は i 番目のニューロンの出力層上での位置、 r_c は勝者ニューロンの出力層上での位置、 $\alpha(t)$ は学習率係数、 $\sigma^2(t)$ は学習半径、 t は学習回数である。

$$m_i(t+1) = m_i(t) + h_{ci}(t) \cdot \{x(t) - m_i(t)\} \quad (2)$$

$$h_{ci} = \alpha(t) \cdot \exp \frac{-\|r_c - r_i\|^2}{2\sigma^2(t)} \quad (3)$$

4. 提案手法

提案するシステムはデータ取得部と行動識別部から構成される。データ取得部で、視界情報をテキストに変換することにより、データ量を削減しプライバシーに配慮したライフログデータを取得する。データを取得するために、視界情報である MOVERIO のカメラ画像を一定時間ごとに自動で取得し、画像認識 API を用いて、画像情報をテキストに変換・記録するアプリケーションを開発する。行動識別部で、記録したデータを整理し、多変量解析を用いて解析・比較を行うことで、行動を識別し周期性を検出する。

5. シミュレーション結果ならびに考察

データを構成する行動の検出、類似する行動とそうではないイベント性のある行動を検出することができた。また、自己組織化マップの解析から、同じ行動でも視界に写る物体の違いから行動の類似性やイベント性を検出できた。この結果から、同じ行動でも使用する場所や物体の変化によって別行動として認識させることができるため GPS を使用せず、ライフログデータに位置情報を付加できると考えられる。

6. 結論と今後の課題

個人情報保護に着目したライフログデータから類似性やイベント性を検出できた。今後の課題として、ユーザーの取得タイミングを組み込んだライフログデータ取得アプリケーションの開発、長時間のライフログデータ取得にデバイスが耐えうるようプログラムを改善することが挙げられる。

参考文献

- [1] 新保史生，“ライフログの定義と法的責任 個人の行動履歴を営利目的で利用することの妥当性”，情報管理, Vol. 53, No. 6, pp. 295–310, 2010.
- [2] 樋口耕一，“テキスト型データの計量的分析：2つのアプローチの峻別と統合”，理論と方法, Vol. 19, No. 1, pp. 101–115, 2004.
- [3] 岡晋之介，“自己組織化マップを用いた気象要素の分類と予測”，<http://www.gifu-nct.ac.jp/elec/deguchi/sotsuron/oka/oka.html>, 閲覧日 2018,1,7.