

2019 年 2 月 19 日

環境・生体データからの 勾配・制約を考慮した粒子群最適化による 行動パターン解析

山本 聖也

富山県立大学 情報基盤工学講座

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

1. はじめに

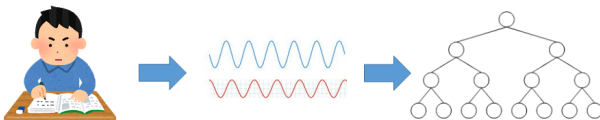
2/17

本研究の目的

行動パターンの分析のためのN進木の分類表現と，粒子群最適化 (**Particle Swarm Optimization : PSO**) による決定木のパラメータ推定アルゴリズムの提案．ならびに勾配・制約を考慮した **PSO** の探索能力の向上．

本研究の概要

- 環境・生体データ収集システムの開発
- N進木の分類表現による行動パターン解析
- 勾配・制約を考慮した **PSO** によるクラスタリングの提案



2. ライフログと各種センサ

3/17

環境・生体データ収集機器

マイコン (**Arduino** と **Raspberry Pi 3.0**) とセンサ類を組み合わせ
て独自のライフログの測定機器を開発する。

収集データ一覧

- 環境センサ「**GPS**(緯度, 経度, 海拔), 温度, 湿度, 気圧, 照度, 人感, 加速度 (3 軸), 角速度 (3 軸), 磁気コンパス (3 軸)」
- 生体センサ「体温, 心拍, **GSR**」

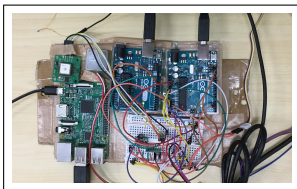


Figure 1: データ収集機器

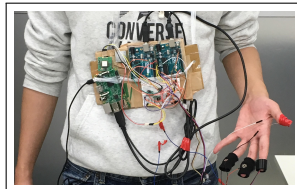


Figure 2: 装着時

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

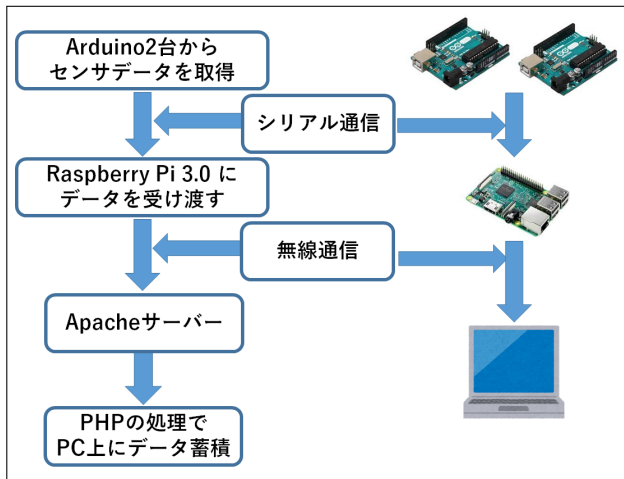


Figure 3: 環境・生体データ収集機器の詳細

3. センサデータからの行動識別

5/17

KH Coder による従来手法

テキストから指定した抽出後に対する多変量解析

- 階層的クラスター分析
- 多次元尺度構成法

実験環境ならびに収集されたデータ例

時間帯:ある平日の 13 時から 18 時

場所:研究室, 廊下, 階段, トイレ, 学外, 自宅

行動・状態:静止, 水平移動, 昇降移動

Table 1: ラベル付けしたデータ

LON36.70638	LAT137.095490	HEI-91.9	T20.07	HUM34.92	PRE1019.04	rig646	zinkan0
LON36.70627	LAT137.095470	HEI-97.3	T20.02	HUM34.96	PRE1019.01	rig477	zinkan0
LON36.70620	LAT137.095630	HEI-102.1	T19.98	HUM35.10	PRE1019.00	rig668	zinkan0
LON36.70619	LAT137.095600	HEI-106.4	T19.91	HUM35.19	PRE1019.07	rig514	zinkan0

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

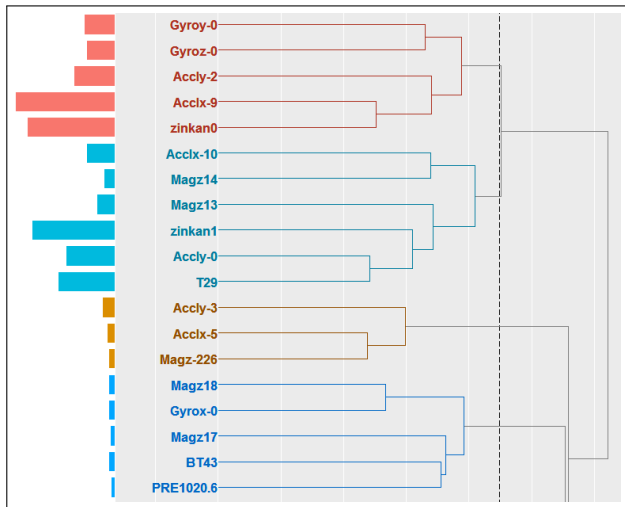


Figure 4: 環境・生体データの階層的クラスター分析

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

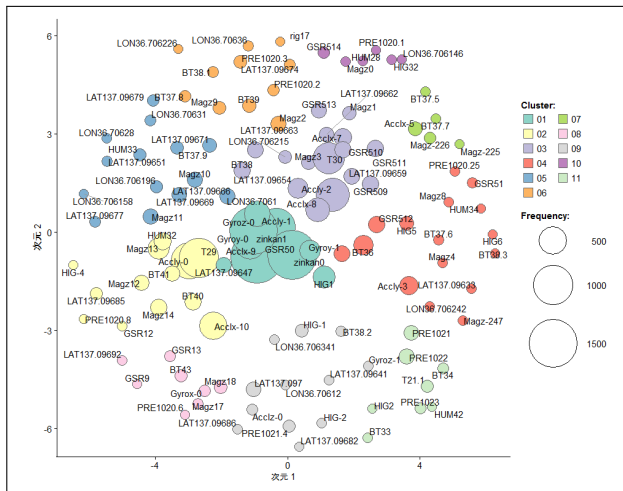


Figure 5: 環境・生体データの多次元尺度構成

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

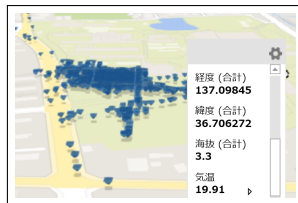


Figure 6: センサデータのマッピング

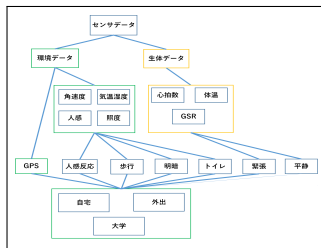


Figure 7: N 進木の分類表現

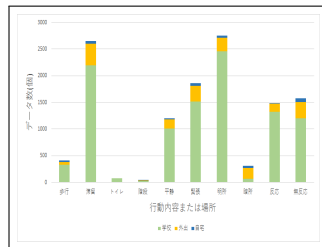


Figure 8: 分類結果

粒子群最適化 (PSO)

PSO によるクラスタリングアルゴリズムを提案する前段階として、**PSO** に勾配と制約を考慮したハイブリッドなモデルを示す。

連続時間 PSO + 勾配情報

$$\frac{dx^p(t)}{dt} = c \int_0^t e^{-a(t-\tau)} [F^p(x^p(\tau), \tau) + C(x^p(\tau), \tau) - \nabla E(x^p(\tau), \tau)] \tau$$

$$\frac{d^2 x^p(t)}{dt^2} + a \frac{dx^p(t)}{dt} = c [F^p(x^p(t), t) + C(x^p(t), t) - \nabla E(x^p(t), t)]$$

個体: p , パラメータ: c, a , 関数: F, C , 勾配関数: ∇E , ステップ数: τ, t

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

連続時間 PSO + 勾配情報 + 制約

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

$$\begin{aligned} \min \quad & E(x) \\ \text{subj.to} \quad & p_i \leq x_i \leq q_i, i = 1, \dots, n \end{aligned}$$

を解くために、非線形変数変換モデルを作成する.

$$x_i = f_i(y_i) = \frac{q_i + p_i \exp(-y_i)}{1 + \exp(-y_i)}$$

制約条件付問題に代入し、 x を消去すると

$$\min E(f(y))$$

を得ることができる. よって,

$$\begin{aligned} \frac{d^2 y^p(t)}{dt^2} + a \frac{dy^p(t)}{dt} = & c[F^p(y^p(t), t) + C(y^p(t), t) \\ & - \nabla E(y^p(t), t)] \end{aligned}$$

従来手法との比較

Griewank :多峰性を持つ関数 :

$$f(x_1, \dots, x_n) = 1 + 1/4000 \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos(x_i/\sqrt{i}), \quad f_{\min}(0, \dots, 0) = 0$$

Booth :単峰性を持つ関数 :

$$f(x_1, x_2) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2, \quad f_{\min}(1, -3) = 0$$

1. 多峰性と単峰性の違いによる影響の比較
2. 粒子の数による影響の比較
3. 制約条件を考慮した場合

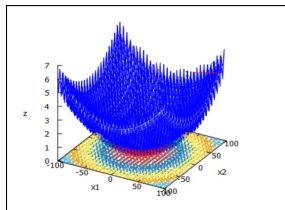


Figure 9: Griewank

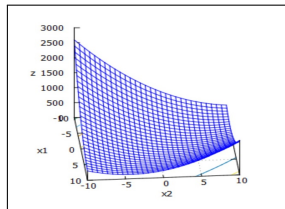


Figure 10: Booth function

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

1. 多峰性と単峰性の違いによる影響の比較

各パラメータ : $c_1 \cdot c_2 = 1.4$, $c_3 = 0.1$, $a = 1.0$, $\Delta T = 0.9$, $N = 10$

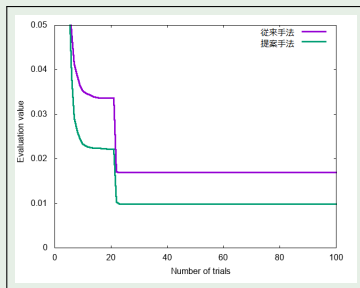


Figure 11: Griewank 実行結果

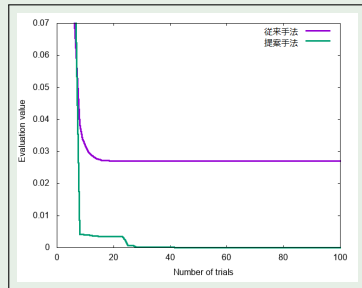


Figure 12: Booth 実行結果

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

2. 粒子の数による影響の比較

評価関数:Booth

各パラメータ : $c_1 \cdot c_2=1.4$, $c_3=0.1$, $a=1.0$, $\Delta T=0.9$

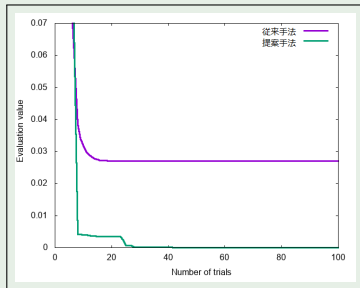


Figure 13: N=10 の場合

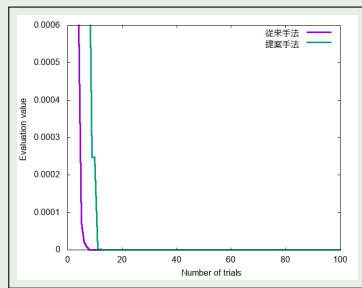


Figure 14: N=100 の場合

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

3. 制約条件を考慮した場合

評価関数: **Booth**

制約 $-10 < x_1 < -5, -10 < x_2 < -5, f_{min}(-5, -5) = 884$

各パラメータ : $c_1 \cdot c_2 = 1.4, c_3 = 0.1, a = 1.0, \Delta T = 0.9, N = 10$

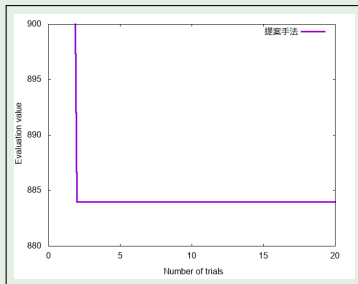


Figure 15: 制約条件を考慮した場合

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

PSO によるクラスタリング

PSO でクラスタリングを行う際には、適合度関数 f によって、全てのパターン X_p に対してユークリッド距離を計算し、最も近い重心を持つクラスターに X_p を割り当てる。

$$E(x) = f(x_i, M_i) = w_1 \bar{d}_{\max} + w_2 (R_{\max} - d_{\min}(x_i)) + w_3 J_e$$

粒子の番号 : i , クラスターの番号 : p , クラスター重心ベクトル : V ,
クラスタ重心の可能なセット数 : x ,
データセット内の最大の特徴量 : R_{\max} ,
粒子のクラスターへのパターンの割り当てを表す行列 : M ,
任意のクラスタ間のユークリッド距離 : d

¹Oded Maimon, Lior Rokach, "Soft Computing for Knowledge Discovery and Data Mining", Springer, Israel, 2005.

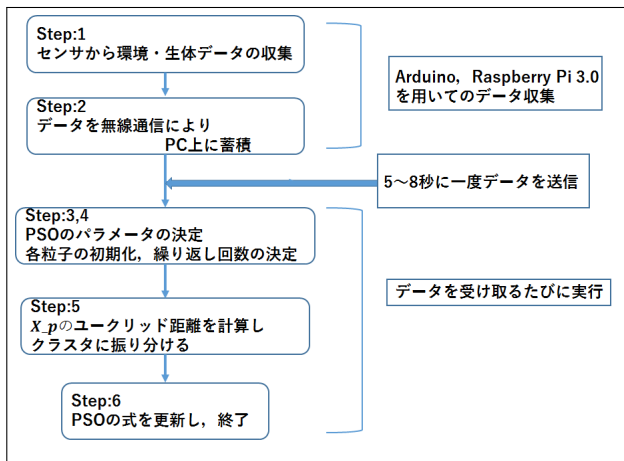


Figure 16: 提案手法の全体のフロー

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

1. はじめに
2. ライフログと各種センサ
3. センサデータからの行動識別
4. 提案手法
5. おわりに

まとめ

- 環境・生体データの収集機器を開発
- 従来手法によるクラスタリング
- 勾配・制約を考慮した **PSO** の有効性の確認
- **PSO** によるクラスタリング・アルゴリズムの提案

今後の課題

提案したアルゴリズムを用いたクラスタリングを行い、有効性の比較、検証を行う必要がある。