

逐次的カーネル密度推定法を用いた4次元画像データ中の細胞の自動検出・追跡手法

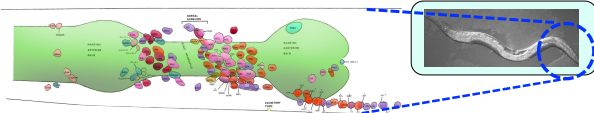
徳永旭将^{1,2}, 広瀬修^{1,3}, 川口翔太郎³, 豊島有^{1,5}, 寺本孝行^{1,6}, 久下小百合^{1,6}, 佐藤博文^{1,5}, 池端久貴^{1,4}, 逸見 昌之^{1,2}, 石原健^{1,6}, 飯野雄一^{1,5}, 吉田亮^{1,2,4,7}

¹(独)科学技術振興機構 CREST ²統計数理研究所 ³金沢大学理工研究領域 ⁴総合研究大学院大学統計科学 ⁵東京大学大学院理学系研究科 ⁶九州大学大学院理学研究院 ⁷(独)科学技術振興機構 ERATO

線虫C.エレガンスの神経系は、302個の神経細胞から構成されており、全てのシナプス結合の構造が同定されている。4D(空間3次元+時間1次元)計測技術を利用して、複数の「生きた」神経細胞の活動状態を同時に測定することで、神経系による空間認識や運動制御、化学走性など、神経回路の動作原理を明らかにすることを目的している。

線虫の神経系

- 302個の神経からなる
- すべての神経間結合が既知(5000の化学シナプス, 600の電気シナプス)
- ゲノムが進んでおり、突然変異体(遺伝子ノックアウト体)の作成が容易

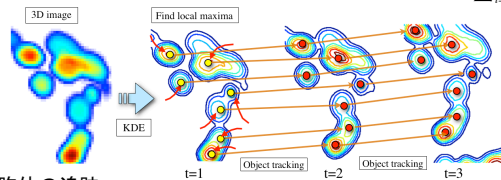


4次元顕微鏡動画に基づく神経細胞のカルシウムイオン分布の定量化

神経細胞のカルシウムイオン濃度の時空間分布を顕微鏡で計測し、取得された動画から個々の細胞の神経活動を定量化する。画像中のオブジェクト(細胞体)は高密度に分布し、形状は不均一、さらには時空間的な変化を伴う。このような画像情報から細胞の個数を数え上げ、位置を同定し、オブジェクトの時空間変化を追跡する問題に取り組んでいる。

STEP1: 前処理(3D画像の平滑化)

- 3Dイメージの密度関数を推定する ボクセルの座標: $x_i \in \mathbb{R}^3 (i=1, \dots, n)$
- ピークの数~細胞体の個数 正規化された輝度: $w_i \geq 0, \sum_{i=1}^n w_i = 1$
- 局所最大点~細胞体の位置 3D画像のカーネル密度関数: $p(x) = \sum_{i=1}^n w_i k_h(x - x_i)$



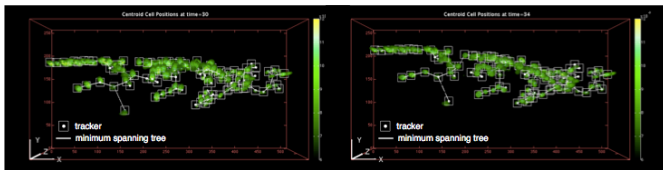
STEP3: 細胞体の追跡

実験中に検体が動くため、細胞体の位置を追跡しながら、カルシウムイオン濃度を推定するアルゴリズムが必要となる。プリム法により最小全域木を生成し、 $t-1$ 時点から t 時点の細胞の位置の遷移をマルコフ確率場で記述し、最近傍の細胞は同じように動くというpriorを設定。各時点において以下の最適化問題を解き、細胞体の位置を追跡しながら、カルシウムイオン濃度の時空間分布を推定する。

- 初期時点の細胞位置を推定 $\Psi_0 = (\psi_0^1, \dots, \psi_0^m) \in \mathbb{R}^{3 \times m}$
- フレーム毎に、以下の最適化問題を解く($t=1:n$)

$$\begin{aligned} \text{maximize } p(\Psi_t | \Psi_{t-1}, I_t) &\propto \prod_{i=1}^m \exp\left(\gamma \log \sum_{j=1}^m k_h(\psi_{t-1,j} - x_i) w_{i,j}\right) \\ &\times \prod_{k \in C_j} \exp\left(-\lambda_{k,j} \left\| \psi_{t-1,j} - \psi_{t-1,k} \right\|_2\right) \\ &= \text{Image force} \times \text{Constrain on the move } \Psi_{t-1} \rightarrow \Psi_t \end{aligned}$$

尤度最大化(image force)に加え、細胞の入れ替わりを防ぐために、細胞の時空間共変動を反映する拘束条件を導入。細胞同士の距離が近いほど、強い拘束条件を加えることにより、密集した細胞が激しく動くデータに対しても、入れ替わりを回避しながら自動追跡を行うことが可能となった。



細胞のレジストレーション

STEP1からSTEP4で同定されたオブジェクトには、神経細胞のアノテーションが付けられていない。そこで、エキスパートによってアノテーションが付けられた標準画像をもとに、ターゲット画像に含まれる細胞の名前を自動的に決定するためのレジストレーション法の開発を行っている。具体的には、以下の重み付き2部グラフのマッチング問題を解くことで、標準画像とターゲット画像の細胞の最適な対応関係を決定する。

2部グラフマッチング

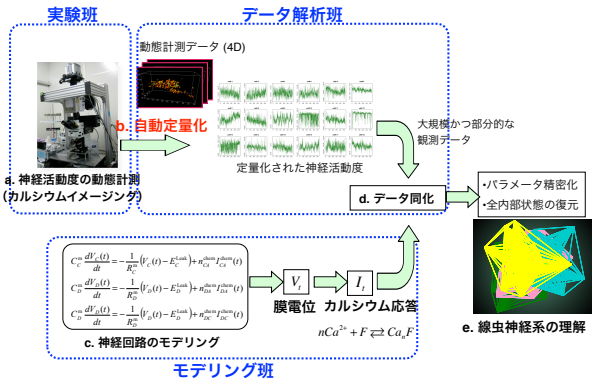
$$\begin{aligned} \min \sum_{i,j} w_{ij} x_{ij} \\ \text{subject to} \\ x_{ij} \geq 0 \\ \forall i: \sum_j x_{ij} \leq 1 \\ \forall j: \sum_i x_{ij} \leq 1 \\ \sum_i \sum_j x_{ij} \leq \lambda \end{aligned}$$

■ 距離
■ 体積・形状
■ 細胞同士の近接関係

マッチング数を制御

標準画像(細胞名あり) ターゲット画像(細胞名なし)

細胞同士のマッチング結果



STEP2: 細胞数の数え上げ・位置の同定

- ローカルモードの探索(EMアルゴリズム)をM個並列に実行する。
- 複数の探索軌道が近接すると、一方の周辺のカーネル関数を除去。
- ブロックされた探索軌道は押し出されて結果的に軌道が逸れる。
- 反発力を調節し(最終的にはゼロ)、複数のローカルモードを効率良く探索。

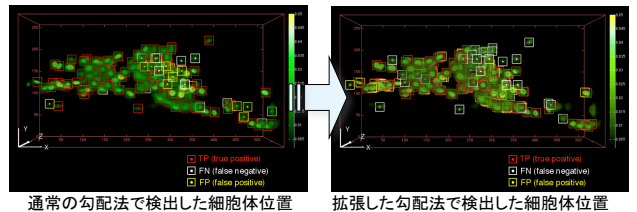
$$\psi_j^{new} = \sum_{i=1}^n \alpha_i (\psi_j^{old}) x_i \quad \alpha_i (\psi_j^{old}) \propto k_h(\psi_j^{old} - x_i) w_i \times \prod_{l \neq j} I(x_i \in S_l[\psi_l^{old}])$$

Conventional hill climbing method
異なる初期値からスタート。探索軌道は相対的に確率が高い同領域にトラップ

Repulsive parallel hill climbing method
探索軌道を反発させ、異なる領域に向かうような作用を加える。



通常の勾配法ではマイナーモードを検出できないが(左:細胞数の推定値=35)、反発を加えながら勾配法を並列実行することで、ほぼ全てのピークを検出できるようになる(右:細胞数の推定値=45)。



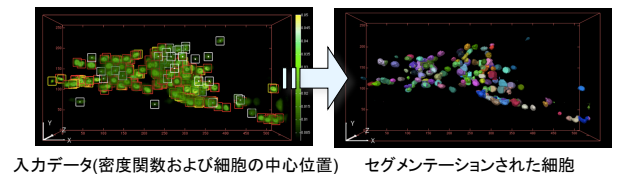
STEP4: 細胞のセグメンテーション

各神経細胞の活動度を定量化するために、細胞の領域を分画し、その輝度値から活動度を算出する。STEP1及びSTEP2で得られたカーネル密度関数と細胞の中心位置情報をもとに、細胞間の境界を推定するためのセグメンテーション法の開発を行った。

■ 分配率を計算

$$\Pr(i \in \diamond) \propto \sum_{j \in \diamond} k_h(x_i - x_j) w_j$$

■ 確率の大きいROIに分配

$$\begin{aligned} \diamond \leftarrow \diamond + \{i\} \\ \text{if } \Pr(i \in \diamond) < \Pr(i \in \heartsuit) \text{ and} \\ \Pr(i \in \heartsuit) \propto \sum_{j \in \heartsuit} k_h(x_i - x_j) w_j \quad \Pr(i \in \heartsuit) \geq \text{threshold} \end{aligned}$$


参考文献

Terumasa Tokunaga, Osamu Hirose, Shotaro Kawaguchi, Yu Toyoshima, Takayuki Teramoto, Hisaki Ikebata, Sayuri Kuge, Takeshi Ishihara, Yuichi Iino and Ryo Yoshida, Automated detection and tracking of many cells by using 4D live-cell tracking data (2014), doi:10.1093/bioinformatics/btu271, *Bioinformatics*, in press.