

論文の内容の要旨

論文題目

Uncovering the functional implications of spatiotemporal neural activity patterns

(神経集団の時空間発火パターンの機能的役割の解明)

氏名 渡辺啓太

はじめに

神経細胞は自分がつながっている他の神経細胞から一定以上の強さの入力を受け取ると発火し、後続の神経細胞に情報を伝える。このような個々の神経細胞の働きについての理解に比べ、これらの神経細胞の通信が総体として、どの様に情報を表現しているかについては未知の部分が多い。しかし、多細胞同時記録法の発展により、ある順序を伴った複数の神経細胞の発火、つまりシーケンスの繰り返しとその役割の一部を担っていることが示唆されている。こうした情報表現に対する研究は、同時記録技術も発展により成し遂げられてきた。それに伴い、こうした高次元化するデータからの特徴抽出の手法が提案されている。しかしながらそうした手法は様々な問題を抱えている。本博士論文の主要な貢献はそれらの問題点を解決する新規の解析手法の提案である。本手法では編集類似度と呼ばれる自然言語処理分野において用いられていた尺度を神経活動に適用可能なように拡張することを提案している。提案手法と従来手法を人工・実データを用いて比較した。また、複数の実データ解析においてはこれまでの手法で観察されなかった新しいパターンの検出に成功した。

編集類似度を用いた神経活動の比較

本研究では編集類似度を用いた神経活動の比較を行うことにより、シーケンスの検出を目指す。比較の対象となる神経活動は特定の長さの時間窓をスライドさせることにより神経活動を表す行列を切り取ったものを用いた。編集類似度とは2つの文字列を一方の文字列を他の文字列に一致させるのに必要なギャップ文字の挿入、文字の一致を取る等の文字列操作の回数を元に定義する尺度である。本研究では動的計画法[15]を用いた編集類似度計算アルゴリズムを、神経活動を表すバイナリの行列に対して適用可能なよう拡張した手法を用いる。動的計画法はある問題を連続する複数の小問題に分割し、小問題の解を利用して元の問題の解を求める手法である。これは2つの神経活動の類似度を比較するために、各々の部分神経活動同士の類似度を計算した結果を用いることに対応する。具体的な計算は、図1左に示すようなテーブルの各マス目の値を左下から右上に埋めていくことで行う。図左の#記号は空の部分神経活動を表している。なお、図右に示す2つの行列は例として比較する神経活動であり、

$W(t_k) = (r_1(t_k), r_2(t_k), \dots, r_L(t_k))$ $W(t'_k) = (r_1(t'_k), r_2(t'_k), \dots, r_L(t'_k))$ に対応する列ベクトルが特定の時刻の各神経細胞の発火の様子を表している。このベクトルの行は神経細胞の番号に対応しており、ベクトル中の0, 1はそれぞれ発火の有無を表現している。また、ここで各マスの計算に用いた更新式は式(1)である。ここで、 ϵ_j^i は図1左のテーブルの(i, j)要素の値を表す。

$$\epsilon_j^i = \max \begin{cases} \epsilon_j^{i-1} \\ \epsilon_{j-1}^i \\ \epsilon_{j-1}^{i-1} + r_i(t_k) \cdot r_j(t'_k) \end{cases} \quad (1)$$

$r_5(t_k)$	0	1	2	2	2	3
$r_4(t_k)$	0	1	2	2	2	3
$r_3(t_k)$	0	1	1	2	2	2
$r_2(t_k)$	0	1	1	1	1	1
$r_1(t_k)$	0	1	1	1	1	1
#	0	0	0	0	0	0
#	$r_1(t_k)$	$r_2(t_k)$	$r_3(t_k)$	$r_4(t_k)$	$r_5(t_k)$	

図1：(右)例として評価する2つの神経活動。各行列の行が神経細胞のIDを、列が各時刻を表す。グレーのマスは2つの神経活動に共通するパターンを表す。(左)動的計画法を用いた神経活動間の類似度評価の例。式(1)に従って二次元の表を左下から右上まで埋めていくことによりもっとも多数の発火の一致を取る神経活動のアラインメントを求めることができる。

$r_1(t_k)$	$r_2(t_k)$	$r_3(t_k)$	$r_4(t_k)$	$r_5(t_k)$
0	1	0	1	0
1	0	0	1	1
0	0	1	0	0
$r_1(t'_k)$	$r_2(t'_k)$	$r_3(t'_k)$	$r_4(t'_k)$	$r_5(t'_k)$
0	0	0	0	0
1	1	0	0	1
0	0	1	0	0

また、テーブルを埋める際に、式(2)に示すようにどのマスが後続のマスを決定したかを記録する事ができる。この情報をもとに最上段再右列の要素がどのマスの値から来ているのかを評価することにより図1左図中の黒矢印を得る。この矢印の各要素が各神経活動に対してギャップを挿入する、読み飛ばす、一致を取るといった編集操作を施すことに対応するため、これから2つの神経活動中に共通して存在する時系列成分を抽出することができる。

なお、実際の解析に際してはセグメント中の最適なパターン開始位置を求めると共にGotohらの線形ギャップ罰則を拡張した漸化式を用いている。

手法の概説

前節では、異なる時間窓から抽出された2つの神経活動を比較するための類似度を提案した。この類似度が、期待通り発火列に対して高い類似度を与えているならば、各神経活動に対応する点は距離空間上での近接点となる。この場合、検出したい繰り返し出現するシーケンスはこの空間中において密に分布している点の集合に対応する（図1C）。ただし、本来この空間を正しく構成するためには任意のペアの類似度を計算する必要があるが、計算時間の観点から類似度が高くなりうるペアのみを比較している。こうした密度によって定義されるクラスタを抽出する手法として本研究ではOPTICSを用いた。さらに、このアルゴリズムは1つの時間窓に対して複数種類のシーケンスのみが含まれる状況では異なるクラスタを同一視してしまうという問題が存在するため、ここで得られた結果を、グラフ上からのコミュニティ検出の手法であるCOPRAで更に処理をすることによりクラスタを抽出している。なお、クラスタリングの結果を制御するパラメタは、モジュラリティを用いて最適化している。また、ここで得られたクラスタ中に存在する発火列を抽出するため、反復共通配列計算アルゴリズムを用いた。このアルゴリズムにより抽出された発火列が検出された繰り返しシーケンスである（図1G）。

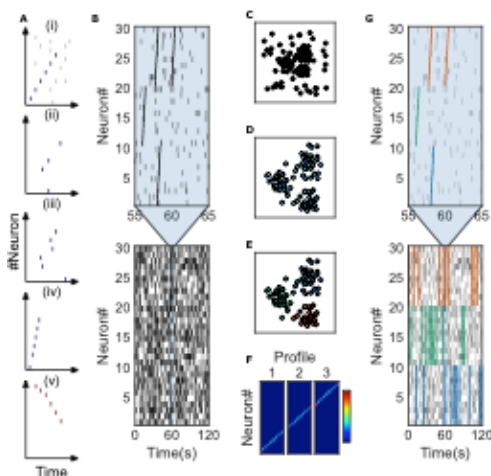


図1：提案手法の概要 本手法では（A）に示すような（I）多数の、パターンと無関係な発火が存在する状況で、発火の抜け（II）、発火位置のズレ（III）パターンの伸縮等を考慮しつつ、神経活動データ中に繰り返し出現するパターンを検出する手法を提案する。（B）シーケンスを含む神経活動の模式図（C-E）提案する類似尺度によって得られるセグメント間の距離を表す特徴空間（D）OPTICSによるノイズ除去の模式図（E）COPRAによるクラスタの分離の模式図（F）抽出された各クラスタ内の共通時系列成分（G）抽出された共通パターンと実際の神経活動との類似

人工データを用いた手法の性能評価

提案手法のシーケンス検出能力を評価するために、人工データを用いてPCA/ICAを用いた手法との比較を行った。

比較に用いた人工データには5種類の重複しないパターンが2Hzのポアソン過程で発火する100個の神経細胞からなる活動データに対して20回ランダムな位置に埋め込まれている（図2（A））。それらのデータをそれぞれの手法を用いてクラスタリングし、性能を評価したのが図2Bである。この図が示すとおり提案手法は既存手法と比較してノイズに対して頑健なクラスタリング性能を有することが確認できる。また、図2Dに示すような、同じ神経細胞を共有する複数の時系列パターンの弁別を行った。

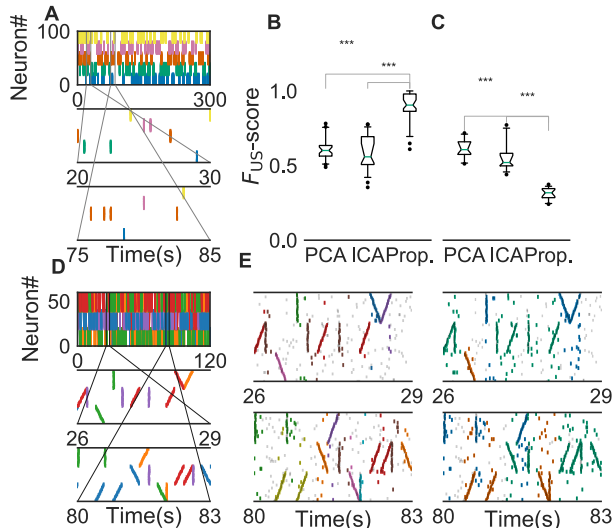


図2：人工データを用いた提案手法の性能検証 (A) 性能評価に用いたデータセットの例。異なる色のパターンは異なる種類のパターンに対応する。可視化のためにノイズは表示していない (B) Fスコアで比較した各手法のパフォーマンスProp. が提案手法。 (C) 各パターンの時系列構造を崩した場合のパフォーマンス比較。提案手法のスコアが大きく落ちていることから、本手法が時系列構造に着目して各パターンを弁別していることがわかる。 (D) 異なる時系列構造を持ったデータセットの例。 (E) 提案手法ではギャップに対する罰則の値を調整することにより、これらのパターンを弁別するか (左) 同一視するか (右) をコントロールすることができる。

応用例1：Prifrontal Cortex (PFC)

提案手法の実データに対する応用の例として二次元アリーナを探索するタスク中のラットのPFCから記録された神経活動データに対して本手法を適用した。その結果図3に示すような異なる種類のシーケンスの検出に成功した。

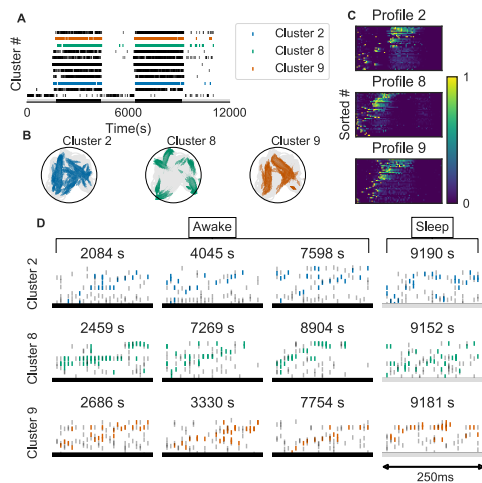


図3：PFCから検出されたシーケンスの例 (A) 検出したパターンの出現時刻 (B) クラスタ2, 8, 9に対応するパターンがラットの行動中のどの位置で観察されたかを示す。各クラスターに対応する神経活動はラットの特定の行動に対応している様子が確認できる。 (C) 各クラスターから抽出された共通神経活動。それぞれ独立に発火位置の重心でソートされている。 (D) 行動時、睡眠時に観察されたシーケンスの例。色付けされている発火がシーケンスに対応する。

応用例2：海馬

提案手法をラットの海馬CA1から記録されたデータに対して適用した。これはラットが線形迷路を移動中に計測されたものであり、87個の神経細胞から約20分間記録されたものである。このデータからはラットの場所に対応した神経活動パターンの抽出に成功した (図4)。検出したパターン中には行動中、停止中双方のパターンが観察された (図5) ことから本手法では時間スケールの異なるパターンを同時に検出できることが確認できた。

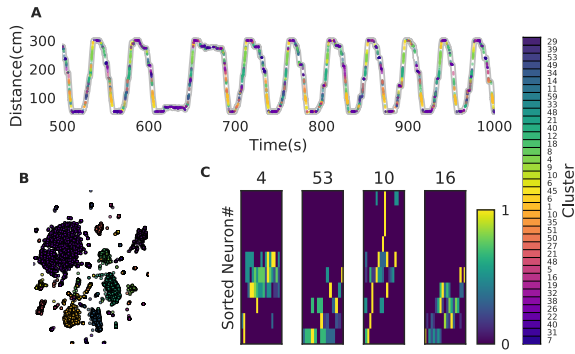


図4：海馬からのシーケンス検出
(A) 各シーケンスはラットの行動中の特定の位置で観察される。縦軸は線形迷路の位置，横軸は時間を表す。(B) tSNE を用いて可視化した特徴空間 (C) 共通時系列パターンの例

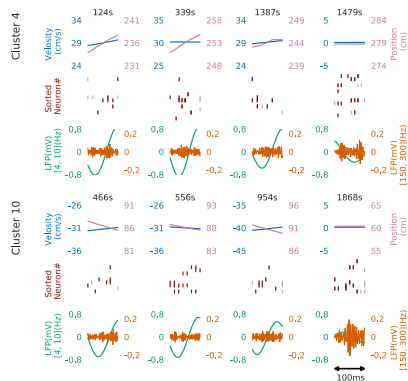


図5：検出されたシーケンスの例
2つのクラスタに属するシーケンスの例。各クラスタについて上段がラットの行動を，中段がその際の神経活動を，下段がその際の LFP 活動を表す。

応用例3：Amygdala

Amygdala と海馬の同時記録データからのシーケンス検出を試みた。このデータではラットは線形迷路走行中に Airpuff 刺激を提示される。先行研究では、この刺激を受ける場所の周辺でラットの移動速度が遅くなることが報告されている。

データを提案手法及びテンプレートマッチングを用いて解析したところ、テンプレートマッチング（図6右）では検出できなかった、これまで未報告のラットの移動加速度に負の相関を持つパターンの検出に成功した。

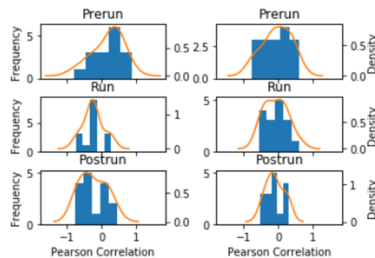


図6：ラットの移動測度と負の相関を持つパターンの検出
提案手法（左）とテンプレートマッチング（右）を用いて検出した神経活動パターンとラットの加速度との間の相関係数。n=4, 総セッション数 17。