

## 解 説

## 情報量最大化原理による皮質神経回路のモデル化

田 中 琢 真

滋賀大学データサイエンス学部\*

## Modeling Cortical Neural Circuits with the Infomax Principle

Takuma Tanaka

Faculty of Data Science, Shiga University\*

## 概要

大脳皮質の動作原理は未解明だ。様々な動作原理の候補が提案されているが、そのうち情報量最大化原理 (infomax principle) は皮質の性質のある部分を説明する。本解説では、Linsker (1988) や Bell & Sejnowski (1995, 1997) の古典的な結果を解説し、筆者らの提案したリカレント情報量最大化原理 (recurrent infomax; Tanaka *et al.*, 2009) を紹介する。情報量最大化原理に基づくモデルの発展や応用も述べる。特に、深層学習, reservoir computing, 神経回路の学習則との関連を詳述する。

## 1. はじめに

大脳皮質は哺乳類の高度な知的情報処理の中樞だと考えられている。現在、大脳皮質の動作原理の解明は、おもに実験神経科学 (解剖学・生理学) の観点から進められている。解剖学と生理学が目指すのは大脳皮質の回路の構造と機能の解明だ。解剖学的には大脳皮質には、層構造がある・コラム構造がある・領野ごとに視床核との相互結合がある<sup>1,2)</sup>など、ある程度整然とした構築の原則があることが明らかになっているが、基本回路はまだ十分解明されていない。また、すでに解剖学的に確立された回路でも、よい実験系が確立されねば生理学的機能が明らかにならず、したがってその回路が何を実現しているのかは不明瞭なことが多い (たとえば、小脳の生理についての我々の知識のほとんどは前庭動眼反射という制御しやすい実験系に依存している)。それは、回路を虚心坦懐に眺めても、神経回路は (たとえば系統発生的な経緯によってたまたま本来の用途以外に転用されていて) 本質的な機能がわかりにくいからだ。

このような場合には、実験からの接近よりも、理論的な考察が回路の動作原理の理解に役立つことがある。つまり、大脳が行っているはずの処理を仮定し、この仮定を原理としてどのような回路が存在するはずかを考察する。解剖学的・生理学的制約は脇に置いて、こ

の「存在するはずの回路」と実際の回路を比較することで回路の本質的な面を明らかにする。

大脳が行う処理は何だろうか？ それは端的には情報処理だ。大脳が情報処理のための装置ならば、大脳の情報処理能力は高くなっているはずだ。だから、情報処理能力が高い回路はどのような回路で、どのような性質を持つかを考察するのは有意義だろう。

情報処理には様々な側面があるから、「情報処理能力が高い」にはもう少し限定が必要だ。情報処理のためには何より情報を正確に伝達し、保持する必要があるから、大脳の回路は情報の伝達・保持能力が高くなっていると考えよう。これが情報量最大化原理の考え方だ。

本解説では、情報量最大化原理とそれに基づく中枢神経系・大脳皮質回路の理解を紹介する。第 2 節は情報量最大化原理の源流と感覚系への適用について述べる。第 3 節は情報量最大化原理の拡張とその他の皮質領野への適用について述べる。第 4 節は情報量最大化原理に基づくモデルを、その他の関連するモデルと比較する。

## 2. 情報量最大化原理

情報の観点による神経科学・心理学の再構築は古くから試みられてきた。Barlow (1959)<sup>3)</sup>は、この試みが感覚の受容は感覚印象の連合と秩序化であると述べた経験論哲学の認識論にさかのぼると述べている。情報理論の確立後に限定しても、神経科学・心理学への情報の

\* 〒 522-8522 滋賀県彦根市馬場 1-1-1

概念の導入と情報に基づく神経科学・心理学の再構築は Shannon (1948) による情報理論の提案<sup>4)</sup>から数年で始まった。神経科学では, Jacobson (1950, 1951)<sup>5, 6)</sup> による感覚入力の情報量の推定が最初期の研究の一つだ。心理学への初期の導入は, 実験的応用よりも概念的な方面からなされた<sup>7, 8)</sup>。Attneave (1954)<sup>8)</sup>では感覚情報の冗長性が強調されている。Barlow (1959)<sup>3)</sup>は脳が段階的に(階層性のある処理で)冗長性を削減するモデルを提示し, さらに情報理論的に最適な表現を生成する(optimal coding)のが感覚情報処理だとしてモデル化することを提唱している。この論文では Craik (1943)<sup>9)</sup>を引いて, 情報理論的に最適な表現は環境のモデルになっていることも議論されている。また, この論文では末梢から中枢へ(感覚器から皮質へ)進むにつれて細胞数が増えるので, 情報のチャネルは増えるが発火率は下がるような表現がされていると推測されている。発火率の順応現象や側抑制もできるだけ小さなコストによる情報の表現と解釈できることが主張されている。これらの研究はまだ萌芽的だったが, 以下で見るように, その後の研究の射程範囲をほぼ予言していた。

情報量最大化原理 (infomax principle) という用語を提案し, 現在の数理モデルの基礎を作ったのは Linsker (1988) だ<sup>10)</sup>。Linsker は, Hebb 則に従うモデルが視覚路の細胞に類似の応答特性を示すようになることを示した。また, このモデルでは出力が入力の情報をできるだけ多く伝達するようになっていることを指摘した。そして, 一般に情報伝達効率の高い回路がどのような特性を持つはずかを議論した。特に, 情報量最大化原理に従う回路はノイズが小さければ出力の冗長性を下げ, ノイズが大きければ冗長性を高めるだろうと論じた。この研究は, 情報の観点から感覚系のモデルを構築できる可能性を示唆したことで大きな影響を及ぼした。情報量最大化原理から導かれる神経細胞の動作特性については本特集の島崎 (2018)<sup>11)</sup>を参照されたい。

フィードフォワード型のネットワークで情報伝達効率を最適化するモデルを構築したのは Bell & Sejnowski (1995, 1997)<sup>12, 13)</sup>だ。考察されているのは出力  $y_i$  が入力  $x_i$  によって

$$y_i = f \left( \sum_{j=1}^n W_{ij} x_j \right) \quad (1)$$

で表されるフィードフォワード型ネットワークだ。情報量最大化原理を仮定すれば, 出力が入力に対して持っている情報つまり入出力間の相互情報量

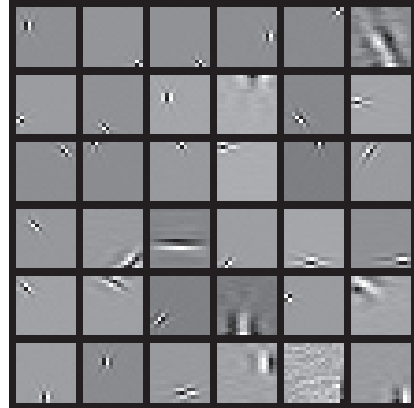


Fig. 1 Simple cell-like basis vectors obtained with the ICA of natural scenes.

$$I(\mathbf{x}; \mathbf{y}) = H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x}) \quad (2)$$

が最大化されているはずだ。Bell & Sejnowski はノイズが無視できる, つまり  $H(\mathbf{y}|\mathbf{x}) = 0$  として出力のエントロピー  $H(\mathbf{y})$  を最大化した。これは, 様々な入力  $\mathbf{x}$  から出力  $\mathbf{y}$  を生成しつつエントロピー  $H(\mathbf{y})$  が大きくなるように結合行列  $\mathbf{W}$  をアップデートすれば達成できる。入力  $\mathbf{x}$  として画像の小部分を与えると, 出力  $y_i$  は一次視覚野の単純型細胞と類似した線分に対する選択性を獲得する (Fig. 1)。すなわち, 画像の中の特定の部位に特定の傾きの線分があるときに  $y_i$  は大きな値をとる。これは, 単純型細胞は画像の情報を効率的に表現する要素であることを示唆する。

Bell & Sejnowski のアルゴリズムは独立成分分析 (independent component analysis : ICA) の代表的なアルゴリズムでもある。このアルゴリズムは, 既述のとおり, ある種の活性化関数を持つフィードフォワードネットワークにおける情報量最大化として解釈できるが, 一方で, 出力の確率分布としてある種の分布を仮定した場合の最尤推定としても解釈できる。単純型細胞と類似の特徴を抽出する出力が出現するのは確率分布として裾の長い分布 (優ガウス分布) を仮定した場合に限られる。言い換えれば, 出力がたいていは 0 に近く, まれに大きな値をとると仮定したときのみこのモデルは単純型細胞と類似した特徴を抽出するようになる。

これは疎表現 (スパースコーディング) と密接に関連する。疎表現の基本的な考え方は, 「神経系はできるだけ少数の発火でできるだけ正確に情報を表現する」だ。実際, Lennie (2003) によれば皮質の細胞の発火率は

非常に低く、1 Hz 以下だ<sup>14)</sup>。疎表現と情報量最大化原理は類似したアイデアだ。疎表現を明示的にモデル化したのが Olshausen & Field (1996, 1997)<sup>15, 16)</sup>だ。このモデルでは、入力画像を  $\mathbf{x}$ 、単純型細胞に対応する出力を  $\mathbf{y}$  で表し、目的関数

$$E = \frac{1}{2} \|\mathbf{x} - \mathbf{A}\mathbf{y}\|_2^2 + \lambda \|\mathbf{y}\|_1 \quad (3)$$

を最小化するように  $\mathbf{y}$  と  $\mathbf{A}$  を最適化する。ここで、 $\|\cdot\|_n$  は  $L_n$  ノルムを表す。この目的関数の最小化は低い発火率でできるだけ入力の情報量を正確に表現することに対応する。確率密度が  $\exp(-E)$  に比例する確率モデルだと考えて、誤差  $\mathbf{x} - \mathbf{A}\mathbf{y}$  が正規分布に従い、パラメータ  $\mathbf{y}$  の  $L_1$  正則化を行っていると考えてもよい。このモデルでも  $y_i$  は単純型細胞のような線分を抽出する特性を獲得する。Bell & Sejnowski のモデルと対応させると、 $\mathbf{A} = \mathbf{W}^{-1}$  に近い（ただし、このモデルでは  $\mathbf{y}$  のスケールに不定性があるので、通常最適化の中で  $\mathbf{y}$  のノルムを 1 に固定するなどの操作を行う）。目的関数  $E$  の第一項は情報の保持の正確さを表す項で、第二項は発火率の低さを表す項と考えてもよい。

疎表現の考えでは、脳は少数の要素で感覚入力を表現する装置だ。言い換えれば、脳は世界が少数の要素からなっていると仮定している。この仮定は私たちの住む世界については確かに正しい。たとえば、視覚的世界は、少数の要素で表現できる。視野の中の多くの部分は隣と同じ色で、同じ明るさだ。色や明るさの変化する箇所は少ないから、変化する場所だけを表現すれば少数の要素で視覚的世界を表現できる。これが単純型細胞の作る表現だ。

情報量最大化原理と疎表現の考え方で複雑型細胞<sup>17, 18)</sup>も説明できることがわかっている。情報量最大化に基づくモデルでは、出力が二層あるネットワークで画像を入力とし、出力の情報量を最大化すると出力の一層目に単純型細胞、二層目に複雑型細胞が出現する<sup>19)</sup>。また、疎表現に基づくモデルでは、単純型細胞の活動の相関をできるだけ少数の要素で表現するのが複雑型細胞だ<sup>20, 21)</sup>。これらのモデルでは複雑型細胞の持つ方位のコントラストによる応答の増強<sup>22)</sup>や超複雑型細胞の特性も説明できる。

情報量最大化原理・疎表現の考え方は視覚系だけではなく、聴覚系の理解でも有効であることがわかっている。一層のネットワークで画像の代わりに音圧時系列を入力として情報量最大化・疎表現を行うと、聴神経のような特定周波数の音に対する選択性が出現する<sup>23)</sup>。二層のネットワークでは脳幹聴覚関連領域や皮質聴覚野で見られる FM 音に対する応答や、基本周波数  $f$  だけ

でなく  $2f, 3f, \dots$  にも応答する細胞も出現する<sup>24~26)</sup>。

情報量最大化の観点からの複雑型細胞の位置づけは次のようになっている。単純型細胞と違って、複雑型細胞は線分の詳細な位置によらない応答を示す。言い換えれば、複雑型細胞は線分の位置の情報を捨てる。このように考えると、複雑型細胞は情報量最大化の意味では不利な表現であるように思われる。しかし、式 1 で表される単純型細胞に対応するユニットを、

$$y_i^+ = \begin{cases} y_i & y_i > 0 \\ 0 & y_i \leq 0 \end{cases}, \quad (4)$$

$$y_i^- = \begin{cases} -y_i & y_i < 0 \\ 0 & y_i \geq 0 \end{cases}, \quad (5)$$

のように二つのユニットに分割し、この入力を受ける出力層の情報量を最大化すると、ある条件下では出力は  $y_i^+ + y_i^- = |y_i|$  の線形結合になる。特にはっきりこの傾向が出るのは  $y_i^+$  や  $y_i^-$  の線形結合以外に  $y_i$  そのものも出力になる場合で、生理学的に解釈するならば、単純型細胞と複雑型細胞の両方を一次視覚野の出力とみなすならば、複雑型細胞はあえて線分の詳細な位置を捨てて画像について単純型細胞とは違った側面を表現するようになることを意味する。

### 3. リカレント情報量最大化原理

ここまで見た情報量最大化原理に基づくモデルは、フィードフォワードネットワークの情報伝達能力を最大化するとどうなるかを調べていた。これらは感覚系のフィードフォワード的な側面に焦点を当てたモデルだ。実際、一次感覚野は視床からのフィードフォワード入力に駆動される。しかし皮質ではフィードバック入力がきわめて多いことも知られている。たとえば、一次視覚野でも視床からの主たる入力部位である 4C $\alpha$  層と 4C $\beta$  層でも外側膝状体背側核由来の非対称性シナプスは全体の 10% 以下だ<sup>27)</sup>。だから、フィードバック入力がある状況での最適な情報処理がどのようなものであるかを考える必要がある。

細胞間に相互結合があるネットワークではどのように情報を保持するのが最適だろうか。相互結合のあるネットワークでは、ある瞬間の回路の状態（発火状態）を回路が次の瞬間もよく記憶できていけばよいだろう。つまり、ある瞬間の発火パターンの情報が次の瞬間によく保持されていれば情報の保持効率が高いと定義しよう。より形式的には、時刻  $t$  の発火パターン  $\mathbf{x}_t$  と次の瞬間の発火パターン  $\mathbf{x}_{t+1}$  の間の相互情報量

$$\begin{aligned} I(\mathbf{x}_t; \mathbf{x}_{t+1}) &= H(\mathbf{x}_t) + H(\mathbf{x}_{t+1}) - H(\mathbf{x}_t, \mathbf{x}_{t+1}) \\ &= H(\mathbf{x}_{t+1}) - H(\mathbf{x}_{t+1} | \mathbf{x}_t) \end{aligned} \quad (6)$$

を情報の保持効率の尺度としよう。

相次ぐ二つの時刻の状態間の相互情報量を最大化するとどのような発火パターンが出てくるだろうか。たとえば、ある瞬間の発火パターンと同一の発火パターンが次の瞬間にも続くような回路ならば、情報が保持される。これはすべての細胞に強い自己結合がある回路だ。しかしこのような回路だけが相互情報量の高い回路ではない。たとえば、細胞 1 からは細胞 2 だけに強い正の結合があり、細胞 2 からは細胞 3 だけに強い正の結合があるような回路も相互情報量の高い回路だ。細胞 2 が発火していれば前の瞬間には細胞 1 が発火していたことがわかり、細胞 3 が発火していれば前の瞬間には細胞 2 が発火していたことがわかるからだ。つまり、強い一対一結合があつて、発火の連鎖が成立するような回路が相互情報量の高い回路だとわかる。しかし本当にこの議論は正しいだろうか？ それを確かめるためには、相互情報量を最大化する学習則を導き、出てくるダイナミクスを調べればよい。

神経回路の情報量については Tsukada らや Ay らの先駆的な仕事がある<sup>28-30)</sup>。Tsukada *et al.* (1975) は単一神経細胞モデルへの入力スパイク時系列と出力スパイク時系列の相互情報量を求めた<sup>28)</sup>。Wennekers & Ay (2005) は<sup>30)</sup>はユニット間の transfer entropy が大きい系が多細胞同時記録で観察されるのと類似の現象を示すことを報告した。これらの研究では、神経回路の学習則は導かれていない。フィードフォワードネットワークと比べてリカレントネットワークで学習則を導くのは難しい。Tanaka *et al.* (2009) は近似的に学習則を導いた<sup>31)</sup>。このモデルではリカレントネットワークの中で情報量最大化を行うのでこれをリカレント情報量最大化 (recurrent infomax) モデルと呼ぶ。以下、このモデルについて説明する。

このモデルでは細胞  $i$  の時刻  $t$  の状態  $x_i(t)$  は発火 (1) か非発火 (0) のいずれかをとる。時刻  $t$  の細胞の発火状態は  $t-1$  における発火状態から決まる。時刻  $t$  の細胞  $i$  の発火確率を

$$p_i(t) = \frac{p_{\max}}{1 + \exp\left(h_i - \sum_{j=1}^N W_{ij} x_j(t-1)\right)} \quad (7)$$

で定める。閾値  $h_i$  は  $p_i(t)$  の時間平均があらかじめ設定した値  $\bar{p}$  に近くなるように調整する。疎表現の考えに従い、 $\bar{p}$  は 0.01 や 0.05 などの低い値に設定する。

$p_{\max}$  は発火率の最大値 (発火の信頼性) を決めるパラメタだ。通常のシグモイド型二値ニューロンモデルでは  $p_{\max} = 1$  だが、ここではこのパラメタは 0.5 や 0.95 などに設定する。このパラメタを導入した理由は二つだ。一つは、細胞の挙動の不確実性を表現できるからだ。もう一つは、 $p_{\max} = 1$  だと結合強度  $W_{ij}$  が大きいときモデルがほぼ決定論的になり、モデルの挙動が不安定になるからだ。このモデルでは相互情報量  $I(\mathbf{x}_t; \mathbf{x}_{t+1})$  を  $\mathbf{x}_t$  や  $\mathbf{x}_{t+1}$  が正規分布に従うとして近似する (もちろん  $\mathbf{x}_t$  と  $\mathbf{x}_{t+1}$  の各要素は 0 か 1 の値しかとらないのでこれはかなり粗い近似だ)。そのほかにも高次モーメントを無視するなど様々な近似を導入すると学習則を導ける。残念ながら、この学習則は局所的ではないため神経細胞には実行できない。この問題点の解決法を第 4 節で述べる。

$N = 50$ ,  $\bar{p} = 0.05$ ,  $p_{\max} = 0.95$  でのシミュレーション結果を Fig. 2 に示す。Fig. 2A1 は初期状態でのモデルの細胞の発火活動のラスタプロット、A2 は学習が十分進んだ後のラスタプロットだ。Fig. 2B1 と B2 にそれぞれの結合強度行列を図示した。学習後には、強い結合はごくわずかで、ほとんど一対一結合になっていることがわかる。つまり、発火の連鎖が出現する。実際 Fig. 2A2 では同一の発火パターンが繰り返す発火連鎖があることがわかる。また、発火連鎖の産物として、同じ発火パターンが繰り返し発生することがわかる (同一の発火パターンを線でつないだ)。特に、同じ発火パターンの発生回数を数えると、ランダムな回路に比べて圧倒的に多いことがわかる。これは Hebb の提唱した cell assembly や phase sequence, Abeles の報告した synfire chain に近い<sup>32-37)</sup>。

一対一結合が生ずるのは、先に考察したように情報量の意味では一対一結合が最適だからだ。相互情報量に基づいて説明すると次のようになる。このモデルでは細胞の発火率が固定されているのでエントロピー  $H(\mathbf{x}_{t+1})$  の上限は決まっている。一対一の強い結合があれば、ある細胞が発火したら次の瞬間に確実に特定の細胞が発火する。これによって条件付きエントロピー  $H(\mathbf{x}_{t+1} | \mathbf{x}_t)$  が下がる。学習の結果できあがる結合は疎だが、結合を疎にする項はないことに注意しよう。発火率が疎なとき、情報保持能力を高めると、結合に何ら制約がなくとも一対一結合が最適になる。

以上は細胞の信頼性が高い ( $p_{\max} = 0.95$ ) 場合の結果だ。細胞の信頼性が低ければどうだろうか。信頼性が低ければ、複数の細胞が同じ情報を重複して伝達するだろう。たとえばもし  $p_{\max} = 0.5$  ならば、強い一対一結合があつても、シナプス後細胞はシナプス前細胞

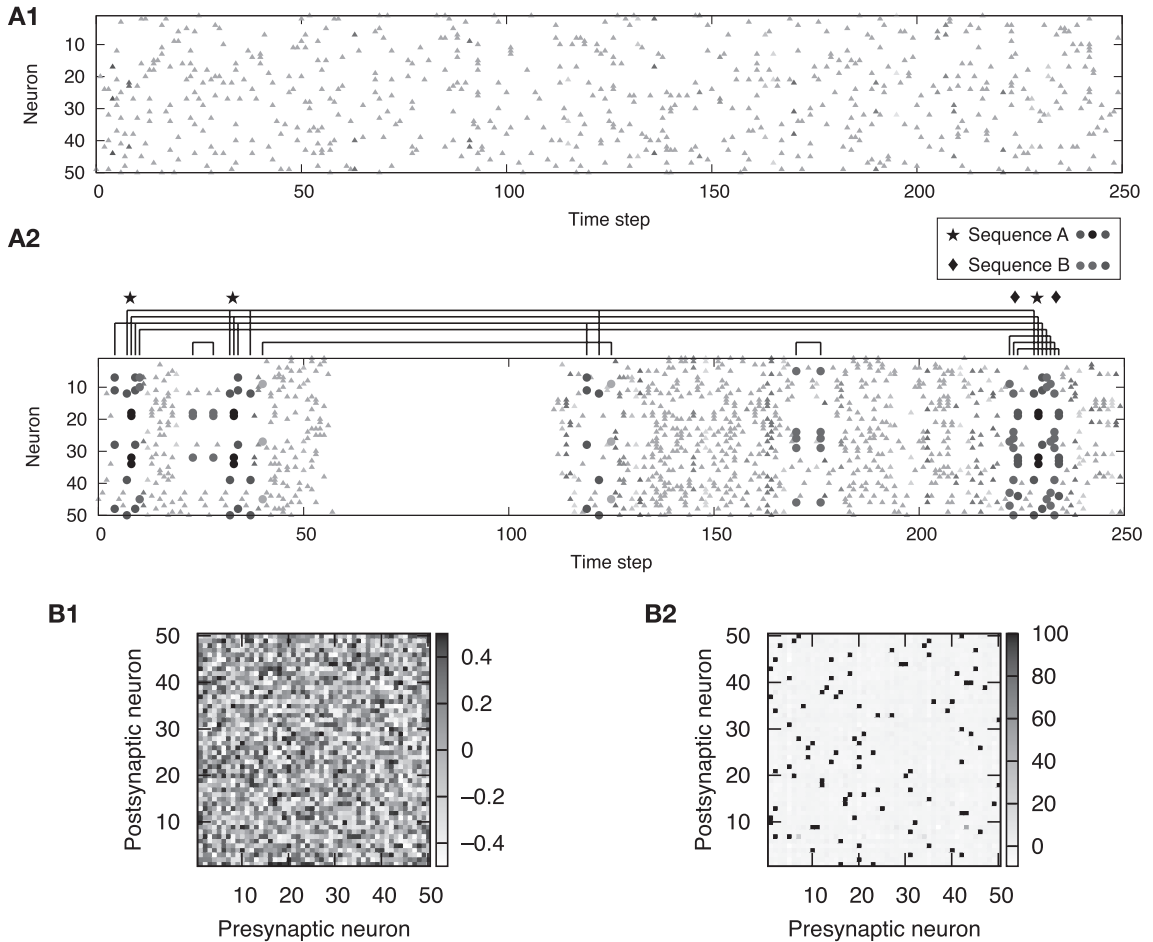


Fig. 2 Spike sequences in the recurrent infomax model with  $N = 50$ ,  $p_{\max} = 0.95$ , and  $\bar{p} = 0.05$ . (A1,A2) Raster plots before and after the learning. (B1,B2) Connection weight matrices before and after the learning.

の半分の確率でしか発火できない。ならば、一対一結合ではなく、一つの細胞からは二つの細胞に強い結合が送られ、また一つの細胞には二つの細胞から強い結合を送られるような構造になっているはずだ。Fig. 3Aは結合強度、Bはラスタプロットを示す。実際に結合強度はほぼ二対二の結合になっていることがわかる。発火活動がしばらく続いた後で停まることを繰り返す。この発火活動が始まってから終わるまでに延べ何個の細胞が発火したかを数え、その分布をグラフにしたのがFig. 3Cだ。冪 $-1.5$ の冪分布に従っていることがわかる。これはBeggs & Plenz (2003)の報告したneuronal avalancheの再現になっている<sup>38,39</sup>。彼らは、皮質のスライス培養の自発発火活動の統計的性質を調べ、一連の自発発火活動のサイズが冪 $-1.5$ の冪分布

に従うことを示した。彼らは、これは皮質が自己組織臨界状態<sup>40</sup>にあるため生じているのだと論じ、さらにこれがフィードフォワードネットワークで情報が最適に伝達される場合に対応することを示した<sup>39</sup>。これはリカレント情報量最大化モデルの結果と一致する。

さらに、リカレント情報量最大化モデルでは、一部の細胞の出力を画像のピクセルの値で固定し様々な画像を入力として与えると、その他の細胞は単純型細胞のような選択性を持つことや、一定パターンの入力の時系列を与えると発火時系列を埋め込めることがわかった。繰り返し一定の時系列刺激を与えることで皮質スライス内に一定の発火パターンが埋め込まれることは実験的にも示されている<sup>41</sup>。

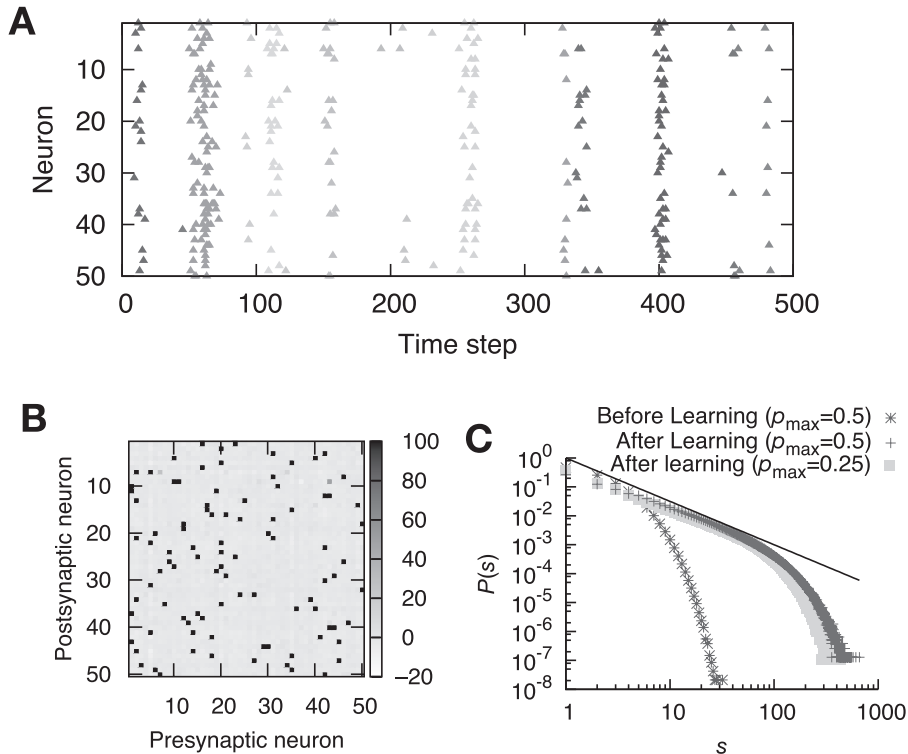


Fig. 3 Neuronal avalanche-like activity in the recurrent infomax model with  $N = 50$ ,  $p_{\max} = 0.5$ , and  $\bar{p} = 0.01$ . (A) Raster plot. (B) Connection weight matrix. (C) The size distribution of avalanche.

#### 4. 終わりに

以上紹介した研究から、情報量最大化原理は感覚系の細胞の特性を説明でき、それ以外の皮質領野の発火活動も説明できることが明らかになってきた。本節では、情報量最大化原理や疎表現に基づく結果とほかの研究との関連について述べる。

まず、近年急速な進歩を遂げた深層学習との関連を述べる。しばしば深層学習では、畳み込みを単純型細胞、プーリングを複雑型細胞になぞらえる。実際、画像を入力とした場合、入力層に近い畳み込み層には線分に対する選択性が出てくることが多い。プーリング層では平行移動された線分に対する選択性が生ずるため、複雑型細胞に似た特性が生ずる。しかし、実際の複雑型細胞や、先行研究<sup>19-21)</sup>での複雑型細胞に対応するユニットは、単に単純型細胞をまとめただけの応答を示すわけではない。周辺抑制や方位選択性のある抑制が見られる。深層学習ではプーリング層では結合を固定することで学習を簡略化するが、以上の結果は、

プーリング層についても何らかの学習を行う精緻化が可能であることを示唆する。

続いて、reservoir computing との関連を述べる。Reservoir computing の枠組みでは、リカレントネットワークの結合強度は固定し、リカレントネットワークから出力へつながるフィードフォワードネットワークのみ学習を行う<sup>42, 43)</sup>。リカレントネットワークの結合強度が適切に設定されていれば、様々な課題を実行できることが示されている。Reservoir computing では、リカレントネットワークがカオス領域と安定領域の境界 (edge of chaos) にあるときに計算能力がもっとも高くなることが知られている<sup>44)</sup>。具体的には、結合行列の最大固有値の絶対値を 1 に設定する (最大リアプノフ指数を 1 に設定する) のがよいとされる。この特殊例として、一対一結合があるリカレントネットワークの計算能力が優れていることが知られている<sup>45)</sup>。これはリカレント情報量最大化モデルに発火連鎖が出現するのに対応する。Beggs & Plenz (2004) によれば、neuronal avalanche が出現するのは一個の細胞の発火

が平均でちょうど一個の細胞を発火させるときだ<sup>46)</sup>。もし一個の細胞が一個より多くの細胞の発火を引き起こせば活動は不安定になり、一個未満の細胞しか発火させなければ情報は消える。これは reservoir computing でダイナミクスのリアプノフ指数が1付近 (edge of chaos) にあるのが最適であるのに類似した結果だ。言い換えれば、リカレント情報量最大化モデルは回路のダイナミクスを edge of chaos に置き、計算能力を高めている。この観察に基づいて、Iwade *et al.* は最近、学習後のリカレント情報量最大化モデルそのものが reservoir として使えることを示した<sup>47)</sup>。

最後に、リカレント情報量最大化モデルの改良について述べる。Hayakawa *et al.* (2014) はいくつかの近似を導入することでリカレント情報量最大化モデルと同様の挙動を示す局所的な学習則を導いた<sup>48)</sup>。学習則に入っているのはある種の相互抑制 (細胞間の不必要な相関を下げる) で、フィードフォワードネットワークにおいて情報量最大化原理に基づくモデルと同様の結果を出す局所的学習則と部分的に類似している<sup>49-52)</sup>。

#### 参 考 文 献

- 1) Cho, R.H., Segawa, S., Okamoto, K., Mizuno, A., Kaneko, T. (2004): Intracellularly labeled pyramidal neurons in the cortical areas projecting to the spinal cord. II. Intra- and juxtacolumnar projection of pyramidal neurons to corticospinal neurons, *Neurosci. Res.*, Vol.50, No.4, pp.395-410.
- 2) Murray Sherman, S. (2016): Thalamus plays a central role in ongoing cortical functioning, *Nature Neuroscience*, Vol.19, No.4, p.533.
- 3) Barlow, H.B. (1959): Sensory mechanisms, the reduction of redundancy, and intelligence. In *The mechanisation of thought processes*, pp.535-539, Her Majesty's Stationery Office.
- 4) Shannon, C.E. (1948): A mathematical theory of communication, *Bell System Technical Journal*, Vol.27, pp.379-423, 623-656.
- 5) Jacobson, H. (1950): The informational capacity of the human ear, *Science*, Vol.112, No.2901, pp.143-144.
- 6) Jacobson, H. (1951): The informational capacity of the human eye, *Science*, Vol.113, No.2933, pp.292-293.
- 7) Miller, G.A. (1953): What is information measurement?, *American Psychologist*, Vol.8, No.1, p.3.
- 8) Attneave, F. (1954): Some informational aspects of visual perception, *Psychological Review*, Vol.61, No.3, pp.183-193.
- 9) Craik, K. (1943): The nature of explanation, Cambridge University, Cambridge UK.
- 10) Linsker, R. (1988): Self-organization in a perceptual network, *Computer*, Vol.21, No.3, pp.105-117.
- 11) 島崎秀昭 (2018): 認識と行動の適応原理, *日本神経回路学会誌*, Vol.25, No.3, pp.86-103.
- 12) Bell, A.J., Sejnowski, T.J. (1995): An information-maximization approach to blind separation and blind deconvolution, *Neural Comput.*, Vol.7, No.6, pp.1129-1159.
- 13) Bell, A.J., Sejnowski, T.J. (1997): The "independent components" of natural scenes are edge filters, *Vision Res.*, Vol.37, No.23, pp.3327-3338.
- 14) Lennie, P. (2003): The cost of cortical computation, *Current Biology*, Vol.13, No.6, pp.493-497.
- 15) Olshausen, B.A., Field, D.J. (1996): Emergence of simple-cell receptive field properties by learning a sparse code for natural images, *Nature*, Vol.381, No.6583, pp.607-609.
- 16) Olshausen, B.A., Field, D.J. (1997): Sparse coding with an overcomplete basis set: a strategy employed by V1?, *Vision Res.*, Vol.37, No.23, pp.3311-3325.
- 17) Hubel, D.H., Wiesel, T.N. (1959): Receptive fields of single neurones in the cat's striate cortex, *J. Physiol. (Lond.)*, Vol.148, pp.574-591.
- 18) Hubel, D.H., Wiesel, T.N. (1962): Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.*, Vol.160, No.1, pp.106-154.
- 19) Tanaka, T., Nakamura, K. (2013): Information maximization principle explains the emergence of complex cell-like neurons, *Frontiers in Computational Neuroscience*, Vol.7, p.165.
- 20) Karklin, Y., Lewicki, M.S. (2005): A hierarchical Bayesian model for learning nonlinear statistical regularities in nonstationary natural signals, *Neural Comput.*, Vol.17, No.2, pp.397-423.
- 21) Karklin, Y., Lewicki, M.S. (2009): Emergence of complex cell properties by learning to generalize in natural scenes, *Nature*, Vol.457, No.7225, pp.83-86.
- 22) Jones, H.E., Wang, W., Sillito, A.M. (2002): Spatial organization and magnitude of orientation contrast interactions in primate V1, *J. Neurophysiol.*, Vol.88, No.5, pp.2796-2808.
- 23) Lewicki, M.S. (2002): Efficient coding of natural sounds, *Nat. Neurosci.*, Vol.5, No.4, pp.356-363.
- 24) Karklin, Y., Ekanadham, C., Simoncelli, E.P. (2012): Hierarchical spike coding of sound.



- In *Advances in Neural Information Processing Systems*, pp.3032–3040.
- 25) Terashima, H., Hosoya, H., Tani, T., Ichinohe, N., Okada, M. (2013): Sparse coding of harmonic vocalization in monkey auditory cortex, *Neurocomputing*, Vol.103, pp.14–21.
  - 26) Tanaka, T. (2016): Information maximization in a feedforward network replicates the stimulus preference of the medial geniculate and the auditory cortex. In Hirose, A., Ozawa, S., Doya, K., Ikeda, K., Lee, M., Liu, D., eds., *Neural Information Processing: 23rd International Conference, ICONIP 2016, Kyoto, Japan, October 16–21, 2016, Proceedings, Part IV*, pp.183–190, Springer International Publishing.
  - 27) Latawiec, D., Martin, K.A.C., Meskenaite, V. (2000): Termination of the geniculocortical projection in the striate cortex of macaque monkey: a quantitative immunoelectron microscopic study, *Journal of Comparative Neurology*, Vol.419, No.3, pp.306–319.
  - 28) Tsukada, M., Ishii, N., Sato, R. (1975): Temporal pattern discrimination of impulse sequences in the computer-simulated nerve cells, *Biol. Cybern.*, Vol.17, pp.19–28.
  - 29) Ay, N. (2002): Locality of global stochastic interaction in directed acyclic networks, *Neural Comput.*, Vol.14, No.12, pp.2959–2980.
  - 30) Wennekers, T., Ay, N. (2005): Finite state automata resulting from temporal information maximization and a temporal learning rule, *Neural Comput.*, Vol.17, No.10, pp.2258–2290.
  - 31) Tanaka, T., Kaneko, T., Aoyagi, T. (2009): Recurrent infomax generates cell assemblies, neuronal avalanches, and simple cell-like selectivity, *Neural Comput.*, Vol.21, No.4, pp.1038–1067.
  - 32) Hebb, D.O. (1949): *The Organization of Behavior; a Neuropsychological Theory*, Wiley, New York.
  - 33) Abeles, M. (1991): *Corticonics*. Cambridge Univ. Press, Cambridge.
  - 34) Skaggs, W.E., McNaughton, B.L. (1996): Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience, *Science*, Vol.271, No.5257, pp.1870–1873.
  - 35) Louie, K., Wilson, M.A. (2001): Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep, *Neuron*, Vol.29, No.1, pp.145–156.
  - 36) Cossart, R., Aronov, D., Yuste, R. (2003): Attractor dynamics of network UP states in the neocortex, *Nature*, Vol.423, No.6937, pp.283–288.
  - 37) Sakurai, Y., Takahashi, S. (2006): Dynamic synchrony of firing in the monkey prefrontal cortex during working-memory tasks, *J. Neurosci.*, Vol.26, No.40, pp.10141–10153.
  - 38) Plenz, D., Aertsen, A. (1996): Neural dynamics in cortex-striatum co-cultures–II, Spatiotemporal characteristics of neuronal activity, *Neuroscience*, Vol.70, No.4, pp.893–924.
  - 39) Beggs, J.M., Plenz, D. (2003): Neuronal avalanches in neocortical circuits, *Journal of Neuroscience*, Vol.23, No.35, pp.11167–11177.
  - 40) Bak, P., Tang, C., Wiesenfeld, K. (1987): Self-organized criticality: An explanation of the 1/f noise, *Phys. Rev. Lett.*, Vol.59, No.4, pp.381–384.
  - 41) Johnson, H.A., Goel, A., Buonomano, D.V. (2010): Neural dynamics of in vitro cortical networks reflects experienced temporal patterns, *Nature Neuroscience*, Vol.13, No.8, p.917.
  - 42) Maass, W., Natschläger, T., Markram, H. (2002): Real-time computing without stable states: A new framework for neural computation based on perturbations, *Neural Computation*, Vol.14, No.11, pp.2531–2560.
  - 43) Jaeger, H., Haas, H. (2004): Harnessing non-linearity: Predicting chaotic systems and saving energy in wireless communication, *Science*, Vol.304, No.5667, pp.78–80.
  - 44) Boedecker, J., Obst, O., Lizier, J.T., Mayer, N.M., Asada, M. (2012): Information processing in echo state networks at the edge of chaos, *Theory in Biosciences*, Vol.131, No.3, pp.205–213.
  - 45) Rodan, A., Tino, P. (2011): Minimum complexity echo state network, *IEEE Transactions on Neural Networks*, Vol.22, No.1, pp.131–144.
  - 46) Beggs, J.M., Plenz, D. (2004): Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures, *J. Neurosci.*, Vol.24, No.22, pp.5216–5229.
  - 47) Iwade, H., Nakajima, K., Tanaka, T., Aoyagi, T. (2018): Effect of recurrent infomax on the information processing capability of input-driven recurrent neural networks. (submitted).
  - 48) Hayakawa, T., Kaneko, T., Aoyagi, T. (2014): A biologically plausible learning rule for the infomax on recurrent neural networks, *Frontiers in Computational Neuroscience*, Vol.8, p.143.



- 49) Földiák, P. (1990): Forming sparse representations by local anti-Hebbian learning, *Biol. Cybern.*, Vol.64, No.2, pp.165–170.
- 50) Földiák, P. (1991): Learning invariance from transformation sequences, *Neural Comput.*, Vol.3, pp.194–200.
- 51) Falconbridge, M.S., Stamps, R.L., Badcock, D.R. (2006): A simple Hebbian/anti-Hebbian network learns the sparse, independent components of natural images, *Neural Comput.*, Vol.18, No.2, pp.415–429.
- 52) Tanaka, T., Aoyagi, T., Kaneko, T. (2012): Replicating receptive fields of simple and complex cells in primary visual cortex in a neuronal network model with temporal and population sparseness and reliability, *Neural Computation*, Vol.24, No.10, pp.2700–2725.