

論文の内容の要旨

論文題目

Computational Model for Auditory Cortex: An Analogy to Visual Cortex

(大脳皮質聴覚野の計算論的モデル：視覚野との類比)

氏 名

寺島 裕貴

我々の世界認識を支える視覚・聴覚の情報処理様式はそれぞれ高度に発達しているが、双方の主要な計算を一手に担っている器官が大脳皮質である。大脳皮質は様々な領野に分かれ、およそ知能に関わるあらゆる機能を別個に実現している一方で、解剖学的には一様な特徴を持つために何らかの計算原理の存在が示唆されてきた。大脳皮質の本質を捉えた計算論的モデルを最終的に実現するには、この多様性と一様性に一貫した解釈を与える必要がある。

既存の計算論的モデルは主に単一領野、特に一次視覚野（V1）に対象を絞って成功を収めてきた。しかし大脳皮質全体の計算原理を解明するには、複数の領野を同時に説明できる単一モデルを構築する必要がある。本論文では特に、最も生理学的な知見が豊富な領野である V1 と一次聴覚野（A1）に焦点を当てる。A1 の神経生理学的な特徴は V1 に比べ乱雑で解釈し辛く、そのモデルは V1 に比べ極めて未発達だった。その一方で A1 と V1 は類似の領野だとも考えられてお

り、何が共通で何が異なるのか、という問いが A1 のモデル化において本質的な意味を持つ。

本論文の目的は、V1 と何が共通で何が異なるのかという問いに対して仮説を提唱し、A1 を計算論的にモデル化することにある。論文全体を通しての仮説は、V1 と A1 は学習則を共有しつつ異なる入力（自然刺激）に適応している、というものである。単純細胞受容野・地図構造・複雑細胞という V1 でよく知られる 3 つの観点において、一見異なるように見える A1 の特徴にアナロジーを見出す。具体的には、自然画像を学習する V1 モデルの学習則だけをそのまま使い、代わりに自然音を学習させることで A1 の特徴が再現可能であることを示す。この結果は単に A1 のモデルを与えるのみならず、大脳皮質全体の計算論的理解を進めるだろう。

本論の前に、共通する背景知識として V1 と A1 の基本的な性質を紹介する。その上で、まず V1 と A1 の類似点として視覚系と聴覚系の階層的な構造における位置付けや解剖書きの構造を挙げる。一方、両者の対照的な性質を 3 つの観点（受容野の局在性・地図の乱雑さ・複雑細胞）から議論する。

本論文では、この類似点と非類似点に対して、次の統一的な解釈を提唱する：V1 と A1 は学習則を共有しており、その意味で非常によく似ている。しかし、自然界における刺激、すなわち自然画像と自然音の統計的性質が明確に異なるために、各々へ適応した結果である神経生理学的特徴を観察した場合には対照的に見える。

最初に示す解析では、自然刺激データベースから相関行列を作成し、自然画像と自然音がどのように異なる統計的特徴を持つのかを例示する。自然画像から作成した相関行列は対角成分のみが顕著で、視野空間上で非常に局所的な統計的依存性を示す。

一方、自然音は局所相関のみならず非局所的な構造をも示す。特に音声はハーモニクスを豊富に含み、周波数空間上で離れた点の間にも強い依存関係がある。これら異なる特徴を持つ入力に対して同一の学習則で適応した場合、対照的な学習結果が得られると予想される。

第 1 のモデルは、刺激に対してほぼ線形に応答する細胞（V1 では単純細胞として知られる）

を扱う。V1 単純細胞の応答パターン（受容野）は視野空間で局在しているが、A1 細胞ではその限りではなく複数の離れた周波数に反応することがある。この対照的な特徴は両領野の計算原理が異なることを示唆しているのだろうか？

V1 の計算論的研究において、スパースコーディングモデルは単純細胞受容野が自然画像統計性を学習した結果として得られることを示してきた。本研究ではこのモデルの学習則をそのまま用い、学習対象を自然画像でなく自然音とすることで、A1 で見られる非局在受容野を再現できることを示す。また、A1 受容野で見られる複数ピーク周波数間のハーモニックな関係も学習される。この結果は、学習則の共有という仮説を支持するものである。

第 2 のモデルは地図構造、すなわち皮質表面に神経細胞反応特性がどのような空間分布を示すかに焦点を当てる。V1 と A1 は、視野空間と周波数空間に対応する地図構造であるレチノトピーとトノトピーをそれぞれ持つことが知られ、この類似性は両者の相同性の論拠のひとつだった。しかし近年の計測技術の進展により、微細スケールでは A1 地図のほうはかなり乱雑であることが明らかになった。この乱雑度の差は何に由来するのだろうか？

計算論的モデル研究は、滑らかな V1 地図が自然画像統計性を学習した結果であることを示唆してきた。我々は学習則の共有という仮説をもとに、この V1 地図モデル（トポグラフィック独立成分分析）の学習則を自然音に対して適用した。その結果、自然画像を用いた場合よりも乱雑な地図構造が得られた。この結果は、乱雑な A1 トノトピーは自然音に含まれる様々な周波数を統合するという音声処理目的に最適化された結果だと示唆する。

上記 2 モデルは受容野と地図について V1 と A1 間で知られる既知の対照を説明し、学習則の共有という仮説を支持した。第 3 のモデルでは、未だ議論されていない相同性をこの仮説をもとに演繹的に議論する。V1 では、単純細胞と対になる概念として複雑細胞という区分が古くから確立している。これは刺激に対して単純細胞のように方位選択性を示しつつも非線形な応答を示す細胞である。不思議な事に、他モダリティの感覚野において複雑細胞に相当する概念は議論

されてこなかった．実際のところ神経生理学的に直接アプローチするのは困難だが，我々は計算論的アプローチの可能性を示す．

V1 計算論モデルは，複雑細胞もまた自然画像の統計性と深く関係しており，さらに滑らかな地図構造との関係を示唆してきた．我々はここでも学習則の共有を念頭に，この学習モデル（トポグラフィック独立成分分析の過完備拡張）を自然音に適用した．モデル上で学習された A1「複雑細胞」の特性を検討した結果，その一部が近年 A1 で発見されたピッチ細胞に酷似する非線形応答を示すことを発見した．この結果は，A1 ピッチ細胞が V1 複雑細胞と計算論的に相同である可能性を示唆する．

本論文の目的は，V1 と類比しつつ，A1 を計算論的にモデル化することだった．両領野の何が共通で何が異なるのかという問いに対して我々は，学習則が同じで自然刺激の統計性が異なるという仮説を提唱した．一貫してこの立場に沿い，異なる 3 つの観点（単純細胞受容野・地図構造・複雑細胞）から，一見 V1 と異なるようにも見える A1 の神経生理学的特性が V1 モデルの学習則を自然音に適用することで説明できることを示した．

大脳皮質の既存モデルは，主に単一領野，特に V1 を対象として成功を収めてきた．しかし我々は，大脳皮質全体に普遍的な計算原理を理解するには複数領野を単一モデルで説明する必要があると主張した．本論文では特に，V1 向けに提案された学習モデルが A1 にも適用可能であることを示し，A1 のみならず大脳皮質自体の理解をも進めた．このように複数領野の類比を意識した計算論的モデル化は，大脳皮質の計算原理を理解しようとする試みにおいて有用である．