

Optimal Model Estimation in Vision

(視覚における最適モデル推定)

田嶋 達裕

1. 序論

ヒトの脳は、単に物理的な外的刺激を表現する以上のことを行っている。すなわち、刺激に埋め込まれた統計的構造を抽出し、その生成過程を推定し、将来の環境を予測することができる。ヒトをはじめとする動物にとって主要な感覚処理のひとつである視覚は、端的には網膜の神経応答に基づいて、上記のような複雑な外界の情報を解読する過程である。これまでに多くの先行研究によって、視覚あるいは視覚系の神経応答が、刺激の物理的特徴と関係づけられてきた。しかしながら、前述した高度な推定問題を含む視覚処理がどのような計算原理に基づいて行われているのかは、いまだに明らかでない部分が多い。本論文では、視覚処理の中心的な問題を、経験ベイズ推定におけるモデル選択としてとらえることで、様々な心理物理学および生理学的な現象を統一的に解釈するための理論的枠組みを構築した。

統計モデルの推定/選択は、確率的情報処理の核心を成す問題である。視覚情報は、3次元環境で様々な反射、屈折過程を経た光が網膜の神経細胞の反応を引き起こすことで、2次元の網膜像として与えられる。視覚処理は、限られた視覚神経応答の観測データから複雑な外界の状態を推定することであり、一般的に不良設定問題である。こうした不良設定問題を扱う上で有効な方法は、自然界の統計的特徴に基づいた事前知識を制約として取り入れることである。自然画像は人工的なノイズ刺激と異なり、様々な点で統計的な規則性を有している。この規則性を記述するものを、本論文では「統計モデル」あるいは単に「モデル」と呼ぶ。統計モデルの具体的内容としては、刺激のテクスチャ、カテゴリ、光学過程等、様々なものが含まれ得る。統計的推定においては、使用される統計モデルの精度は本質的な問題である。過去に多くの研究が、統計モデルに基づいた視覚情報処理の有用性を示してきたが、一方で、「統計モデル自体を神経システムがどのように推定/獲得し得るか」という問題については、あまり深く議論されてこなかった。本論文ではこの問題に焦点を当てた。

2. 統計モデルの神経符号化

まず理論的な立場から、神経系における知覚形成のモデルを提案した。ここでは、2種類の典型的な神経回路

モデル、すなわち均一/不均一な受容野特性を持った神経細胞集団のそれぞれについて、定式化と理論的解析を行った。

網膜や低次視覚野のように、同じような受容野構造を持った神経細胞の配列による刺激の表現を、ここでは「均一な」神経集団符号と呼ぶ。均一な集団符号の解読は、工学分野における画像修復問題の拡張としての「画像再構成」と解釈できる。ベイズ推定は、画像のノイズ除去において集中的に研究されてきた方法のひとつであり、自然画像における「滑らかさ」などの統計的性質を事前知識として導入することで、確率的観測過程に伴う不良設定性を解決できる。これはベイズの定理の形で次のように表現される：

$$P(s | r, \theta) = P(r | s)P(s | \theta) / P(r | \theta)$$

ここで、 s は刺激、 r は神経応答、 θ は刺激の統計モデルを決定するパラメータである(図 1)。例えば、自然画像の周波数空間での振幅スペクトルは冪則に従う事が知られており(図 2)、その指数がテクスチャなど画像特徴を反映することが分かっている。これを反映した統計モデルを事前分布 $P(s | \theta)$ とする。一方、 $P(r | s)$ は、この場合神経細胞の受容野と応答ノイズのモデルを与える。

経験ベイズの枠組では θ を観測データ r から推定する。その一つの方法は、 $P(r | \theta)$ を最大化するような θ を選択することである。工学的な画像修復では、 θ を単なる制約条件として捉え、刺激 s の推定精度に重点が置かれてきた。一方、生物の視覚処理においては、刺激そのものよりも、その統計モデル θ の推定を通してテクスチャ等

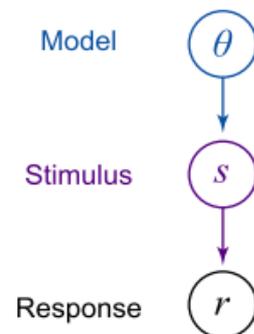


図 1 刺激の観測過程

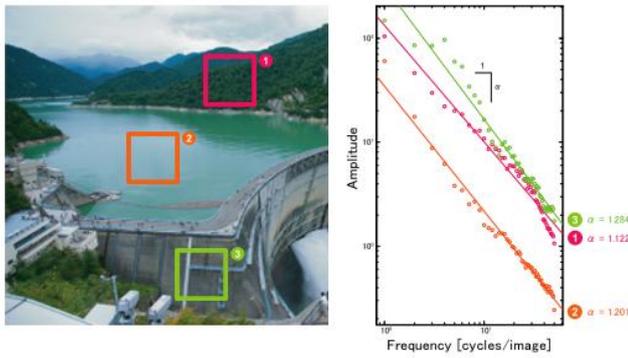


図 2 自然画像を表現する統計モデルの例

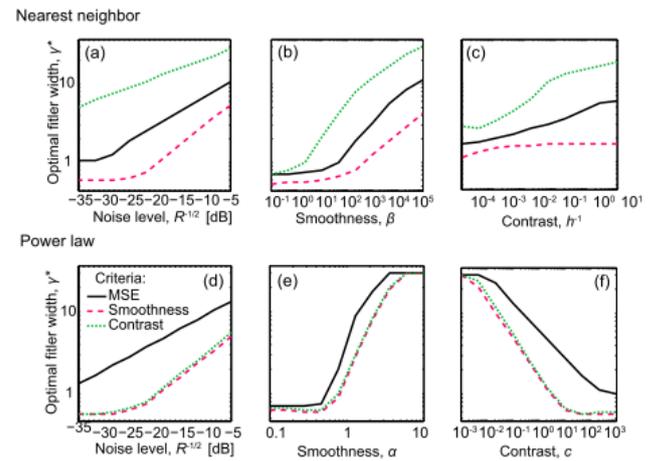


図 4 統計量と最適な受容野サイズ

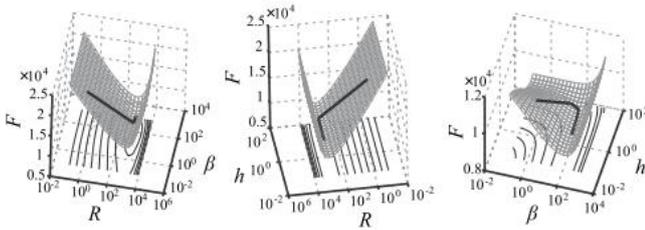


図 3 統計量の経験ベイズ推定

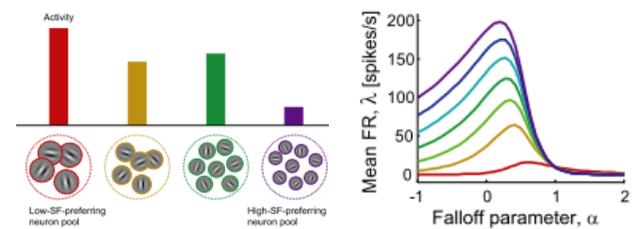


図 5 不均一な神経集団による統計量の表現

の高次情報を得る事のほうが重要と考えられる。そこで、本論文ではモデルの推定に焦点を当て、神経集団符号から刺激統計量を推定するアルゴリズムを導くとともに、その性能を確認した(図 3)。また、統計量の推定という目的に対して、どのような神経受容野が最も高い性能を示すかを、刺激統計量に関するフィッシャー情報量を用いて定量化した。この結果、最適受容野サイズの刺激依存性が理論的に導かれた(図 4)。

さらに、不均一な神経集団においても、刺激および神経受容野、神経ノイズのモデルに基づいて、統計量が推定できることを示した(図 5 左)。画像統計量の例として、冪則に従う振幅スペクトルの指数を考え、この画像統計量に対する低次視覚野の神経応答特性を予測した(図 5 右)。また、ここでも統計量の推定精度をフィッシャー情報量により定量化し、行動学的な刺激弁別の精度に用いられる d' および弁別閾との関係を導いた。

3. 実際の脳におけるモデルの符号化と復号化

次に、実証的な立場から理論を検証した。まず、刺激符号化の側面からの検証として、機能的核磁気共鳴画像(fMRI)による生理学的実験を行った。理論的解析からは、刺激の強度が小さい場合には、加算領域の大きな受容野を用いる事が、統計量および刺激の復号化精度を高める上で有利であることが示された。しかし、こうした受容野の変化を実際のヒトの脳で直接検証した例は過

去に報告されていない。そこで、中心刺激を高/低2通りのコントラストに設定して周辺効果の特性を検証した。

その際、応答変化の検出感度を高めるため、刺激呈示中に周辺コントラストを長周期のサイン波形に従って連続的に変化させる新たな手法を開発した。一般に、刺激が低いコントラストで呈示された場合、得られる賦活は高コントラストのときに比べて小さくなるが、特定の周波数成分に注目することで、視覚刺激に由来する応答を抽出できる。さらに、刺激コントラストを連続的に変化させることにより、単一の実験で様々な周辺コントラスト条件での応答を記録できる点で効率的な方法である。

実験の結果、ヒトの第一次視覚野において、中心コントラストに依存して促進的/抑制的な相反する2種類の周辺効果を、初めてヒトの脳において検出した。中心が高コントラストの条件下では周辺効果は抑制的であるのに対し、中心が低コントラストの条件下では、同じ位置に呈示された周辺刺激が促進的な性質を持つことが示された(図 6)。これは提案理論から予測される最適受容野の変化と定性的に矛盾しない。また、この特性の変化はイメージングにおけるボクセル毎のばらつきを考慮しても有意であることから、集団的に共有された傾向ある事も示された(図 7)。

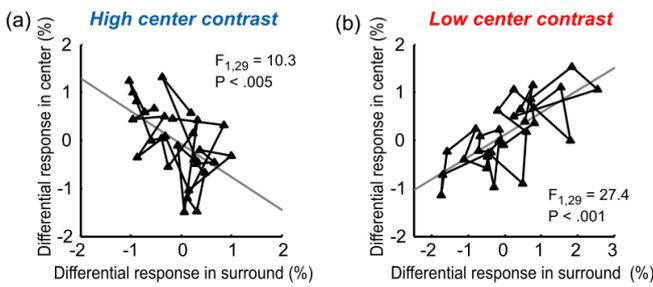


図 6 刺激統計量に依存した神経符号機構の解明

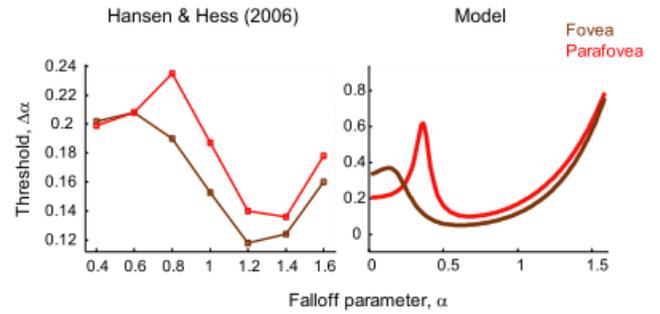


図 8 集団符号からのモデル推定と行動データの比較

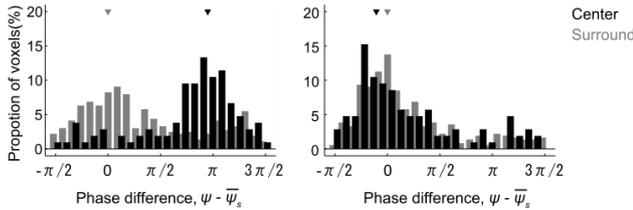


図 7 受容野の刺激依存性の集団的特性

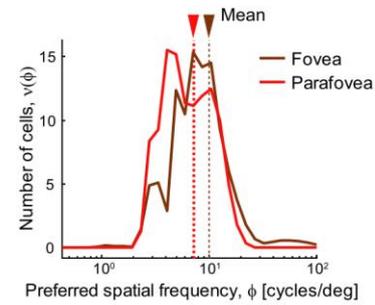


図 9 行動データから予測される神経集団の分布特性

一方、刺激の復号化の側面からの検証として、視覚刺激における統計的パラメータの知覚について、理論による予測を、先行研究で報告されている心理物理学的データと比較した。このとき、前章で行った定式化に基づいて、非一様な周波数選択性曲線を持つ細胞集団に関する画像統計量の復号化精度を理論的に導き、フィッシャー情報量を通して行動データとの関係を予測した。その結果、選好周波数についての細胞数分布を変化させることで、統計量の弁別精度の振る舞いに定性的な差異が生じることを明らかにした。これは報告されている中心視・傍中心視での心理物理実験データをよく説明できる(図 8)。この事は、従來說明されていなかった複雑な知覚特性が、本論文で提案する「刺激に関する最適な統計的モデル推定」の観点から自然に理解される例となっている。

さらに、実験で報告された弁別閾の特性から、逆に細胞数分布を推定することができる。この推定結果は、霊長類の生理実験により報告されている細胞分布の視野依存的特徴とも整合する結果となった(図 9)。これらの結果から、実際にヒトが自然画像統計量を認識する際、ここで仮定したような空間周波数選択的細胞の集団応答に基づいた推定を行っている可能性が示唆される。

4. モデル推定の動態

以上までの理論は、モデル自体の時間的変化を考慮したモデル推定問題(動的推定)に拡張できる。動的推定の第1のケースとして、統計モデル自体が時間的に変化

する場合を考える。第2のケースとして、**経験**に基づいて**観測モデル**を**学習**により獲得する場合を考える。

統計モデルが時間的に**変化する観測系**は、**隠れマルコフ過程**(図 10)として記述できる。前章までの議論では、モデル θ 自体に**対する**事前知識 $R(\theta)$ として、無情報事前分布を仮定してきた。一方、時間的な**連続性**がある場合、時刻 t における事前知識は、過去の**観測** $D_{t-1} \equiv \{r_{t-1}, r_{t-2}, \dots\}$ に基づき、 $R(\theta | D_{t-1})$ として**与えられる**。

時間的に変化するモデル推定の例として、**カテゴリカルな知覚**を考える。色覚や顔認知など、様々な知覚処理において、カテゴリが弁別や記憶などに影響することが知られている。本研究では、カテゴリを統計モデルの隠れたパラメータとみなし、隠れマルコフ過程に基づく推

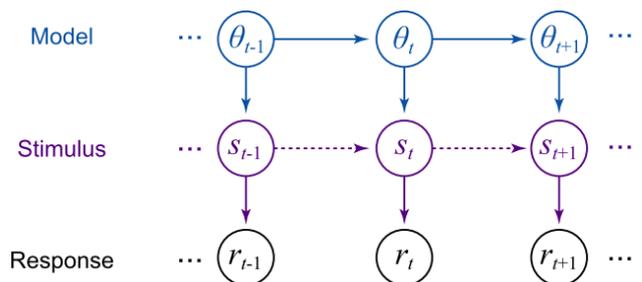


図 10 時間的変化を伴うモデル下での観測過程

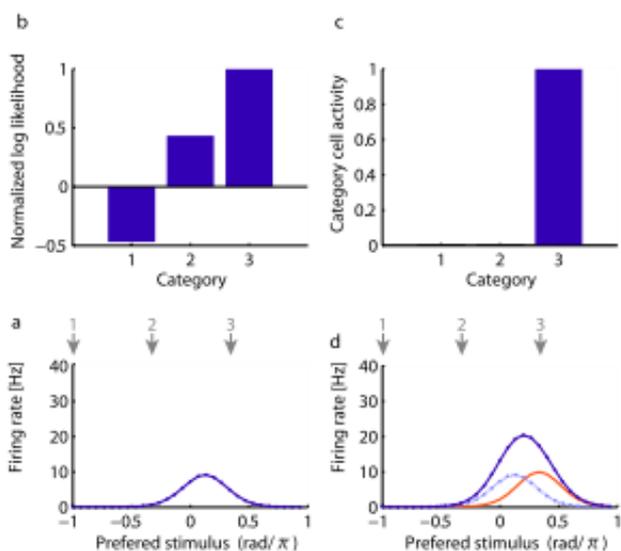


図 11 動的モデル推定によるカテゴリカル知覚

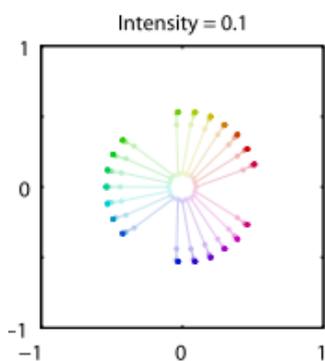


図 12 カテゴリカルな記憶におけるバイアスの再現

定問題として、推定アルゴリズムとその近似としての神経回路モデルを提案した(図 11)。その結果、カテゴリカルな影響下での知覚バイアス、記憶の時間的特性、神経表現のタスク依存性といった様々な現象が、統一的な理論的枠組みから理解できることを示した(図 12)。

最後に、経験に基づく観測モデルの獲得の問題を扱った。特に、生後発達における視覚刺激の複雑な生成モデル獲得について考察した。特に、現実の視環境では、非線形で複雑な光学過程の学習が必要となる。これは、第1段階に単純な刺激を学習した後に、第2段階でその観測モデルを学習する、という方略により効率よく学習できる(図 13)。これは、従来提案されてきた、線形なカルマンフィルタによる隠れマルコフ過程の推定モデルを、非線形な設定に拡張したものになっている。

非線形光学の例として、このモデルを遮蔽環境下での

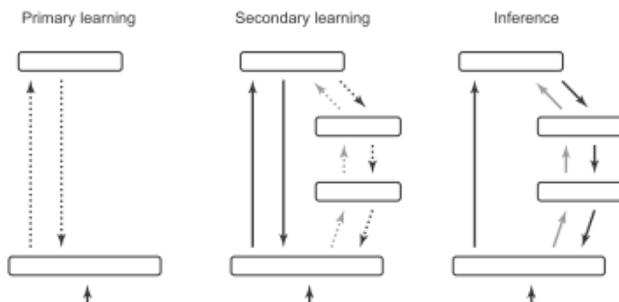


図 13 適応的な観測モデルの推定

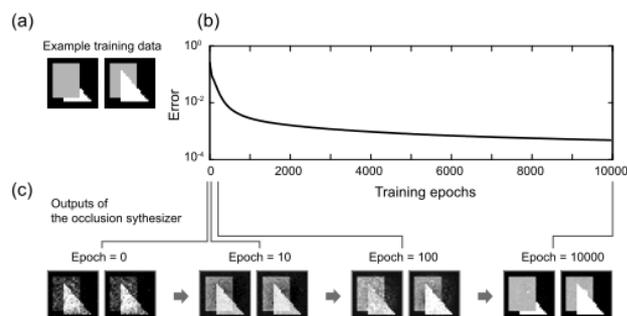


図 14 観測モデルの経験的獲得

観測モデルの獲得と物体認識に適用し、神経回路網による実装を行った(図 14)。これまで、高度な処理機能を獲得する際の乳児視覚の段階的発達に関して理論的モデルがほとんど提供されてこなかった。本論文では、視覚処理における後天的学習を「刺激の観測モデルに関する段階的な推定過程」ととらえることにより、視覚発達過程の計算論的解釈を与えた。

5. 結論

視覚処理の問題を経験ベイズ推定におけるモデル選択の視点から再解釈することで、広範な心理物理学および生理学的現象を統一的な理論的枠組みの中で説明した。また、これまで十分な理論的説明が与えられてこなかったいくつかの現象について、新たなモデルによる解釈を与えた。本論文の理論的、技術的成果は、視覚およびその他の感覚処理を理解する上で有用な視点を与えるだろう。