

論文内容の要旨

論文題目 Model and Applications of the Global Topographic Mapping Formation
(大域的トポグラフィックマッピング形成のモデルと応用)

氏名 松田源立

第1章：はじめに

脳内においては、各々のニューロンの選択性による情報表現（ある特定の傾きの直線にのみ反応するニューロン等）と、それらニューロン群の配置による情報表現（類似した傾きに反応するニューロン群がクラスタをなした方位コラム等）が存在しており、後者を「トポグラフィックマッピング」と呼ぶ。1970年代中期より、このマッピングを自動形成させる様々なモデルが提案されてきた。それら従来のモデルのほとんどは、生理学的な知見に基づき、ニューロン間の局所的相互作用（Kohonen 自己組織化マップの近接関数等）を利用している。この局所的相互作用の働きにより、近接したニューロンの対応する入力パターンが類似性を増す方向に適応が進み、結果として大域的なトポグラフィックマッピングが形成される。工学的な観点から見ると、これらのモデルは局所的な相互作用のみを利用しているので、要求される計算量が小さいことが大きな長所である。すなわち、これらのモデルは大規模データの次元縮約を効率的に行なう能力を持っている。そのため、このような局所的モデルは、数百万通のテキストデータの分析などに応用されており、成果をあげている。しかし、そこでは、データの元々持っている大域的な関係性は無視されているため、場合によっては不十分な分析しか出来ない場合がある。また、局所的モデルでは、局所最適解を避ける

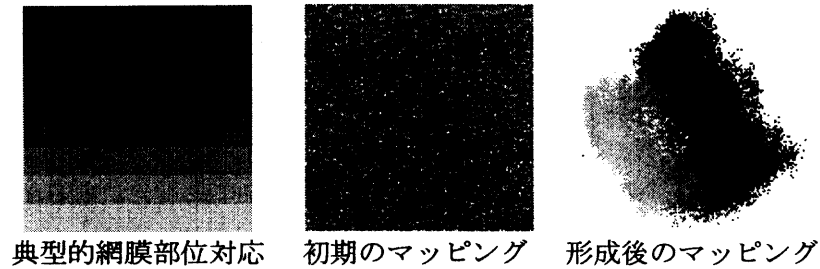


Figure 1: 標準 GMA モデルによる網膜部位対応の形成

のが困難である。

そこで、本論文では、元のデータの持つ大域的な関係性を全て考慮しながら、かつ効率的にマッピングを形成するモデル（大域的写像分析、GMA）を提案し、そのモデルの工学的応用を行なった。GMA の基本的なアイデアは、多次元尺度構成法（MDS）に確率勾配アルゴリズムを適用して、局所的要素と大域的要素を分離することである。それにより、GMA は計算量を増加させることなく、元のデータの全ての大域的關係性を出来るだけ保存するようなトポグラフィックマッピングを形成することが出来る。

第2章：標準 GMA モデル

標準 GMA モデルは、古典的多次元尺度構成法に確率勾配アルゴリズムを適用することによって構成されるモデルである。標準 GMA モデルは効率的に多次元尺度構成法を実行することが出来、かつ、ニューラルネットワークの形にインプリメントすることが出来る。標準 GMA モデルの大きな特徴は、古典的多次元尺度構成法に基づいているがゆえに、大域的最適性が保証されていることである。これは、多くの局所的モデルに比べて本モデルが大きく優位な点であると考えられる。標準 GMA モデルの有効性を検証するため、筆者は自然画像を入力とした数値実験を実行した。その結果、標準 GMA モデルは代表的なトポグラフィックマッピングの一種である網膜部位対応を形成することが可能であることが示された (図 1)。

第3章：離散的 GMA モデル

離散的 GMA モデルは、ニューロンが離散的な格子上に配置されている場合に有効なモデルである。筆者は焼きなまし法と離散的 GMA モデルを数値実験で比較し、ニューロンの数が膨大な場合には GMA が非常に有効であることを示した。さらに、さまざまな入力を離散的 GMA モデルに与えることで、興味深いトポグラフィックマッピングを形成させること

に成功した。例えば、両眼から視神経がオーバーラップして視覚野に投影されるマッピングや、相関のない異なる感覚系から投射された独立なマッピング群などを形成させることが出来た。これらの結果から、GMA はニューロンの数が膨大な状況下でもうまく作動してトポグラフィックマッピングを形成することが出来ることが示された。

第4章：GMA の拡張と応用

GMA は、標準 GMA モデル以外にも拡張可能であり、筆者は、SSTRESS と呼ばれる指標に基づく多次元尺度構成法に関する GMA モデル (GMA on SSTRESS モデル) を構築した。さらに、それらの GMA モデルを実際の大規模文書データの分析に応用した。具体的には、筆者は約 80000 通の Usenet の記事を収集して分析にかけた。各記事は単語に分解され、共通の単語を多く持つものをより近くに配置するようなマッピングを、標準 GMA モデル及び、GMA on SSTRESS モデルにより求めた。さらに比較のために、Kohonen の自己組織化マップ (SOM) による分析も同様に行なった。その結果、大域的な関係性の利用により、GMA は、SOM よりも明確に、データからクラスタを見つけ出すことが出来ることが分かった。また、標準 GMA モデルは、SOM のようにデータをゆがめる性質がないため、データからより有効な軸を抽出することが可能であることも分かった。これらの結果は、GMA が、大規模データの分析における応用において非常に有効である、という筆者の主張を実証するものである。

第5章：ツリー法による GMA の拡張

先に述べた通り、GMA は局所的要素と大域的要素の分離を利用した手法であり、分離不可能な場合は適用出来ない。しかし、ツリー法という計算天文学などでよく利用されるアルゴリズムを用いることで、その適用範囲を広げることが可能である。われわれは、LOGSTRESS という名の分離不可能な指標に基づく多次元尺度構成法を提案し、GMA とツリー法を組み合わせた解法を提案した。いくつかの簡単な数値実験により、ツリー法による拡張が有効であり、効率的アルゴリズムを構成することが可能であることが実証された。

第6章：おわりに

本論文では、元のデータの大域的な関係性を保存しつつ、かつ効率的にトポグラフィックマッピングを形成する手法、GMA を提案した。さらに様々な数値実験により GMA の有用性を実証した。今後は、GMA を、データマイニングで注目されている様々な問題に応用していく予定である。また、GMA の理論的側面を明らかにし、局所的モデルとの組合せを可能にしていきたいと考えている。さらに、トポグラフィックマッピングの形成を情報理論の観点から整理して原理を導き出し、それに対して GMA の適用を行ないたいと考えている。