

氏名	しもむらまさお 下村正夫		
授与学位	博士(工学)		
学位授与年月日	平成16年9月8日		
学位授与の根拠法規	学位規則第4条第1項		
研究科, 専攻の名称	東北大学大学院工学研究科(博士課程)電気・通信工学専攻		
学位論文題目	視覚情報処理系における階層的処理機構のモデル化に関する研究		
指導教官	東北大学教授 阿曾 弘具		
論文審査委員	主査	東北大学教授 阿曾 弘具	東北大学教授 矢野 雅文
		東北大学教授 堀口 剛	東北大学助教授 大町 真一郎
	(情報科学研究科)		

論文内容要旨

第1章 序論

ヒトを初めとする諸動物が行う知的情報処理として重要なものに、外界から得た情報を基にして外界の様相を知覚・認識するというパターン認識という行為が挙げられる。ヒトの機能を代替可能な人工物を作り上げる際に、パターン認識は欠かせないものといえる。そのため、パターン認識の能力を電子計算機に持たせようと幾多の研究が行われてきた。だが、研究の発展に伴い、諸動物が当たり前の事として行っているこの行為が見かけほど簡単には実装できないということも明らかになってきた。現在実用化されている工学的なアプローチによるパターン認識の機構は、研究者の認識対象に対する知識を想定した認識環境に合わせてこれをアルゴリズムに置き換え認識を実現しているため、限られた環境下でのみ有効に機能する。すなわち、環境や認識対象の変化に対して柔軟に対応可能な物とはなっていない。これに対して、ヒトを初めとする諸動物が行っているパターン認識の能力は非常に高度であり、かつ柔軟性に富む。このような高度な知的情報処理を実現するためには、「その基本原理を脳に求めることで処理機構を明らかにする必要がある」というアプローチが1940年代に登場した。現在では、神経生理学・情報数理学・心理学などの幅広い分野の研究者によって支持されており、学際的な研究が活発に行われている。このような計算論的神経科学のアプローチをパターン認識の研究に適用し、成功を収めた例が Fukushima によって提案されたネオコグニロンと呼ばれる数理モデルである。

ネオコグニロンという数理モデルを通じて脳における視覚情報処理の解明とその工学的応用を目指した研究は、過去にも様々な研究が行われてきた。しかし、過去の研究には視覚情報処理に欠かせない重要な機能の幾つかが欠落しているという問題点がある。例えば、視覚処理の情報伝搬経路の妥当性や注視機能の不完全性、数理モデルのパラメータ依存性などである。

本論文は、上記の考察で挙げたような既存のネオコグニロンおよびその発展モデルが抱える問

題点を解決し、より柔軟で使いやすく、生理学的にも妥当な視覚神経系の数理モデルを構築することを目的とする。

第2章 ネオコグニトロン

Fukushima によって提案されたネオコグニトロンは、視覚情報処理 (パターン認識) を目的とする数理モデルであり、Hubel-Wiesel の階層型仮説を基にして構築された。

ネオコグニトロンは、多層の階層構造を持つ神経回路網である。パターンを提示する入力層である U_0 層を先頭として、単純型細胞に対応する S 細胞で構成される U_S 層と、複雑型細胞に対応する C 細胞で構成される U_C 層とが交互に並んでいる。 U_S 層と U_C 層の組を細胞段と呼び、第 l 段目の細胞段を U_l と表記する。

各細胞層内では多数の細胞が 2 次元平面上に集まっており、これを細胞面と呼ぶ。 U_S 層内の細胞面を S 細胞面、 U_C 層内の細胞面を C 細胞面と呼ぶ。 1 つの細胞層内にはこのような細胞面が多数存在しており、同一細胞面内の細胞はいずれも同一形状の特徴に対して反応するという性質を持つ。

S 細胞は特徴抽出機能を持つ細胞で、入力結合として前段の C 細胞との間で可塑性のある興奮性結合を持つ。 C 細胞は位置ずれの許容を行う性質を持つ細胞で、前段の S 細胞との間で固定結合を持つ。 また、S 細胞・C 細胞以外に V 細胞と呼ばれるものも存在し、S 細胞との間で抑制性・可塑性の結合を持つ。 これら可塑性を持つ結合の強度は、学習時に自己組織的に決定される。

また、最終層の C 細胞の発火頻度は、 U_0 層に提示されたパターンの認識結果に対応している。例えば 0 から 9 までの数字を認識する場合、最終層である U_{C4} 層にはそれぞれの数字に対応した 10 個の C 細胞が配置され、最大発火している C 細胞に対応するクラスが認識結果を表す。

第3章 仮定と検証に基づく注視点移動によるパターン認識モデル

ヒトの視覚神経系では、視界内の処理対象となるパターンの位置は認識を通じて把握しており、視界内に混在した複数のパターン中で重要だと思われる位置へ視点を次々移すことで外界を認知している。この機構は注視 (パターンの切り出し) と呼ばれており、効率的な視覚情報処理を実現するために重要な役割を担っていることが知られている。

従来の視覚情報処理モデルでは、認識対象の視野内における位置がすでに決定されている状態、すなわち注視点移動後の処理のみを考慮したモデルが大半を占める。注視の機構を考慮した認識モデルもいくつか提案されているが、認識開始時にパターン全体が見えている場合にのみ有効に動作するものであった。しかし、実際のヒトの視覚情報処理系においては、パターンの一部のみが網膜に投影される場合が一般的に発生する。

本章では、認識開始時には網膜にパターン全体が射影されていない場合でも、視界外に及ぶ注視点移動を伴うことで、適切なパターンの切り出し・認識を実現するモデルを考案した。この状況で問題となるのは、注視点位置の計算とパターン認識処理とは互いの処理結果に依存しているため、両者を同時に、かつ正確に計算することは不可能である、という点である。本研究では、心理学的にヒトの認知活動で利用していることが指摘されている、「仮定と検証」という動作をモデルに組み込むことでこの問題の解決を図った。これは、同時に計算できない複数の事項を、反復演算によって徐々に解を得る方法である。

モデルは、網膜部・特徴抽出部・特徴種類比較部・仮定生成/検証部・注視点調整部の 5 つのモジュールで構成される。網膜部では、画像中から限られた注視点領域を切り出し特徴抽出部に送る。特徴抽出部はネオコグニトロンで構成され、網膜範囲内の画像中の特徴を抽出する。通常のパターン認識においては、特徴抽出で得られた特徴を全て用いて認識結果を即座に出力するが、本モ

デルでは抽出された特徴のうちでパターンの提示位置に依存する特徴と依存しない特徴とを分離して認識に使用する。まず、特徴種類比較部において、パターンの提示位置に依存しない特徴のみを用いて、現在の注視位置での仮の認識結果を出力する。仮認識結果は仮定生成/検証部に記憶される(仮定の生成)。続いて注視点調整部においてパターンの提示位置に依存する特徴を利用することで、仮定されたパターンが存在すべき位置を計算し、注視点の移動指示を網膜部に送る。注視点の移動後には再び仮認識結果を算出し、仮定生成/検証部において注視点移動前の仮定と相違が生じていないかを確認する(仮定の検証)。認識結果が同一の場合には現在の仮定が採択され、再び注視点調整部において注視点移動量の算出を行い、視点の移動を繰り返す。この手続きを注視点の移動が収束するまで繰り返すことで、最終的な注視位置に仮定されたパターンが存在するという最終認識結果を出力する。もし仮定の検証時に仮定とその時点での仮認識結果が一致しない場合には、現在検証中の仮定は棄却される(仮定の棄却)。その場合には、特徴種類比較部で次候補となった仮認識結果を新たな仮定として認識動作を続行する。棄却された仮定は仮定生成/検証部に記憶され、誤った仮定を生成し続けることを抑制する。

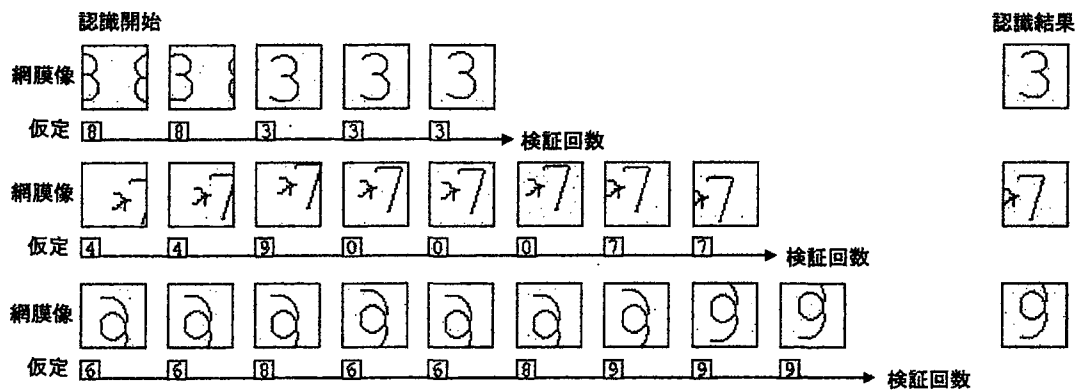


図 1: モデルの認識動作例

図 1 に数字を学習させた場合の本モデルの認識動作の一例を示す。認識開始時に断片的なパターンが網膜上に複数射影された場合や、ノイズが網膜上の大半を占めている場合など、既存のモデルでは認識が著しく困難であった状況下においても、仮定の検証と棄却を繰り返しながら適切な位置へと注視点を移動させ、正しくパターンを切り出し・認識することに成功した。

第 4 章 新しいネオコグニロン型ネットワークと ICA・PCA を用いた学習法

ネオコグニロンは、入力画像中の局所的情報を階層的に統合してゆくことで高い認識率を達成したモデルであり、このモデルの基本的な構造および動作は他の多くの研究にも採用されている。しかしながら、ネオコグニロンには認識率を左右するパラメータが数多く存在し、その調整はもっぱら作成者の経験に頼る方法が主流であった。この方法では特に、異なる認識対象に対して新たに適切なパラメータを設定する事が困難であり、実用上大きな問題となっていた。そこで本章では、ネオコグニロンの動作を数学的に解析することにより、ネオコグニロンの各階層で行われている処理が、基本的には高次元の入力データを低次元空間に射影するという処理であることを導いた。この知見を基に、高次元から低次元への変換を提供する手法である統計的次元圧縮法をネオコグニロン型ネットワークの学習方法として導入することで、パラメータ数を大幅に削減すると同時に、パラメータへの頑健性を高めた、ネオコグニロンを単純化した視覚認識モデルを提案

した。

導入した統計的次元圧縮法は、以下の3種類である。

- 低次元空間の基底間の相互情報量が最小となる変換を求める方法である独立成分分析 (ICA)
- 分散最大基準に基づいて低次元空間を構成する主成分分析 (PCA)
- パターンを識別するために有効な低次元空間を計算する統計的手法である部分空間法

モデルは3段構成のネットワークであり、1, 2段目にICA学習法及びPCA学習法を導入した。3段目には部分空間法を導入した。

手書き数字データベース、顔画像データベース及びテクスチャ画像データベースを用いた認識実験の結果から、1段目にICA学習法、2段目にPCA学習法を用いた場合、最もよい認識率を得られることがわかった。この組み合わせでは、ネオコグニトロンと同等の性能を保ちながら、パラメータ数を24から12に半減することに成功した。さらに、パラメータを変化させた場合でも、認識率の変化は微小な範囲に収まり、パラメータの変動に頑健であることも示された。また、ネオコグニトロンでは認識対象に応じてパラメータを調整する必要があったが、本モデルでは、手書き数字や顔画像といった認識対象の違いに関わらず、共通の計算方法および同じパラメータで高い認識率を達成しており、非常に汎用性の高い視覚認識モデルといえる。

なお、ICAは、脳において最も低次の視覚情報処理を行う、一次視覚野と呼ばれる部位での神経細胞の反応性を再現することが過去の研究から明らかになっており、この実験結果は未だ知見の得られていない高次の視覚情報処理機構が、PCAに基づくものである可能性がある事を示唆している。

第5章 結論

本研究は、計算論的神経科学のアプローチをパターン認識の研究に適用し、成功を収めた例としてネオコグニトロンと呼ばれる階層型の神経回路網に着目した。そして従来型のネオコグニトロンおよびその発展モデルが抱える問題点を解決した新たな視覚情報処理系における階層的処理機構のモデル構築を目的とした。

第3章では、物体形状の知覚に関わっている腹側経路および物体位置の知覚に関わっている背側経路という2つの情報処理経路を加味し、さらに上位の層からの情報のフィードバックを組み込むことで、認識開始時には網膜にパターン全体が射影されていない場合でも、視界外に及ぶ注視点移動を伴うことで、適切なパターンの切り出し・認識を実現するモデルを提案した。

第4章では、ネオコグニトロンの各階層で行われている特徴抽出処理が、基本的には高次元の入力データを低次元空間に射影するという処理であることに着目し、統計的な次元圧縮法である主成分分析 (PCA)、独立成分分析 (ICA) 及び部分空間法をネットワークの学習法として導入することで、パラメータ数の削減とパラメータ変動への頑健性向上を図った。また本章では、これらの統計的手法を導入するために、ネオコグニトロンの局所的・階層的な情報処理機構を継承しつつ信号伝搬式を線形化した、新しいネオコグニトロン型ネットワークを提案し、ネオコグニトロンとの定性的な比較を行った。

論文審査結果の要旨

ヒトの機能を代替できる知的情報処理系の構成には外界の様相を知覚・認識するパターン認識は欠かせない。パターン認識は、工学的見地からは対象パターンの特徴・性質を分析し、それを基に認識アルゴリズムを実現してきたが、ヒトの認識能力がもつ高度性、柔軟性はまだ得られていない。ヒトの視覚情報処理系は対象パターンについての特別な知識がないまま経験によって学習し、高度な認識能力を獲得していく。脳の視覚野の生理学的知見から神経細胞を階層的に構成し、学習により認識していくシステムとして、ネオコグニトロンが知られている。著者は、視覚情報処理系の基本原理の解明とその工学的応用を目標にして、ネオコグニトロンモデルを基礎にした階層的認識モデルについて研究を行い、認識と眼球運動とを連携させる注視機構を導入したパターン認識モデル、及び、パターン学習に統計的手法を導入した新しい階層的認識モデルを考案し、計算機に実装してその有効性を検証するとともに、視覚系の基本原理の一端を明らかにした。本論文はこれらの成果をとりまとめたもので、全編5章からなる。

第1章は序論である。

第2章では、提案モデルの基盤部分となるネオコグニトロンについて、その概要を説明している。

第3章では、パターンの形状と位置とが網膜から別々な経路で送られて処理されるという生理学的知見および注視点移動が高次から低次へのフィードバック経路により指令されるという知見をもとに、仮定と検証に基づく注視点移動によるパターン認識モデルを提案している。視野内に対象全体がはいっていない場合にも注視点を移動させてパターン全体を視野に入れて正しく認識する機能をもつことを計算機実験で検証した。また、極一部しか視野に入っていない場合には大きく注視点を移動させ、視野に入ってから固視によって詳細に認識を行うという認知現象が自然に起きることを明らかにした。これは興味深い成果である。

第4章では、3層構造で第1,2層の細胞への結線を、統計的次元圧縮法である独立成分分析(ICA)と主成分分析(PCA)を基に構成し、第3層の細胞への結線を部分空間法を基に実現した階層的認識モデルを提案している。認識対象画像を局所的部分画像の集まりと捉え、多数の訓練用画像の各部分画像にICA, PCAを適用して結線重みを学習させている。手書き数字、顔画像、テクスチャ画像という性質や特徴が異なる3種類の対象パターンに対して同一のパラメータ(細胞面数など)で学習させることができることを計算機実験により検証した。ネオコグニトロンの認識性能が多数のパラメータの値に敏感であったことに比べて、少数のパラメータでその値の変化に頑健な学習法であることを明らかにした。また、第1層にICA, 第2層にPCAを使った場合最も認識性能がよいことが明らかとなり、局所的特徴を抽出した後それらを統合した特徴が認識に有効であるという認識の基本原理の一端が明らかになった。これは重要な成果である。

第5章は結論である。

以上要するに本論文は、認識と注視点移動を仮定と検証で連携させる機構をもたせた階層的認識モデル及び統計的手法で結線重みを学習させる階層的認識モデルを提案し、計算機に実装してそれらの性能を検証するとともに視覚情報処理系の基本原理の一端を解明したもので、神経回路網工学及び情報通信工学の発展に寄与するところが少なくない。

よって、本論文は博士(工学)の学位論文として合格と認める。