

氏名	佐藤俊治
授与学位	博士（工学）
学位授与年月日	平成12年3月23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	階層的パターン統合処理に基づく視覚情報処理モデルに関する研究
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 星宮 望 東北大学教授 矢野 雅文 東北大学助教授 大町真一郎

論文内容要旨

第1章 序論

文字による情報伝達を行なうには記録，配布，認識の三段階を必要とする。記録に対しては活版印刷技術の向上，配布に対しては交通網の発展や情報ネットワークの技術の向上に伴ないそれぞれの時間的・金銭的コストが低下している。認識に関してはヒトの占有機能である時期が長かったが，電子計算機の登場により認識の研究が活発に行なわれるようになった。今日では近年の高度情報化に伴ない，文字認識技術の社会的重要性は一層高まっている。

文字認識を含むパターン認識活動はヒトが最も頻繁に行なっている知的情報処理であり，その柔軟性は非常に高い。例えばヒトはパターンの変形，大きさ，位置ずれ，ノイズ，回転に影響されずに瞬時に認識結果を得ることができる。パターン認識の研究はこれらパターンの多様性に影響されない認識機能を計算機で工学的に代用する試みから出発したが，一方で脳神経生理学や心理学等ではヒトの脳の認識機能の解明を目指した研究がなされてきた。これらの知見を活用することにより工学的実現可能なものを提案することは，ヒトの認識機能を解明するための証拠を示すことになる。

本研究は神経生理学や心理学的知見を導入して，パターンの多様性に影響されない視覚神経回路網モデルを構築することを目的としている。また最適なモデル構築のための処理時間の削減という工学的要求にも応えることも目的としている。

第2章 ネオコグニトロン

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た認識システムとしてネオコグニトロンが知られている。このネオコグニトロンに着目してその構造の概要を示すとともに，パターンの多様性に対する頑健性を調査した結果を述べている。ネオコグニトロンが変形，大きさ，位置ずれ，ノイズに対する頑健性は高いことを確認したが，回転に対してはほとんど頑健性がないことが明らかになった。

第3章 回転対応型ネオコグニトロン

本章ではパターンの回転にも頑健な認識システムを構築するために、パターン要素の方位を検出することができる細胞面群を導入した回転対応型ネオコグニトロンを提案している。細胞面群とは、従来型ネオコグニトロンを構成している細胞面に回転角の情報を表現する軸（回転軸）を付加したものである。

回転対応型ネオコグニトロンも従来型ネオコグニトロンと同様に階層型の神経回路網モデルであり、特徴検出細胞であるS細胞を含むS層、位置ずれ・回転ずれを許容するC細胞を含むC層が交互に並んでいる。各細胞は細胞面群内に3次的に配置されている。

S細胞はパターン要素の方位情報を検出する特徴検出を行ない、あるパターン要素が回転すると、そのパターン要素を検出しているS細胞の発火パターンは回転軸方向へ平行移動する。これは従来型ネオコグニトロンにおいてパターン要素が2次元平面内で平行移動すると、それに対応して発火するS細胞も細胞面内で平行移動する性質に対応している。細胞面からなる従来型ネオコグニトロンは、2次的にS細胞の発火パターンをぼかすことで変形や位置ずれに頑健な認識システムであったが、回転対応型ネオコグニトロンではC層においても細胞面群を導入し、回転軸方向にも発火パターンをぼかすことによって回転に頑健な認識システムとなっている。

回転対応型ネオコグニトロンの有効性を確認するために数値シミュレーションを行なった。回転対応型ネオコグニトロンに回転していない（正置した）手書き数字をサンプルとして与えて文字の学習を行なわせ、パターンの多様性にどの程度頑健性があるかを調査した。その結果、正置したサンプルの学習だけで任意の方位に回転した文字パターンを認識することを確認した。更に、文字の変形、大きさ、位置ずれ、ノイズに対する頑健性は従来型ネオコグニトロンよりも更に向上していることを確認した。文字の回転、変形、大きさ、位置ずれ、ノイズに対して頑健性を持つ単一の認識システムの実現は本研究が初めてである。

第4章 高速生成アルゴリズム

回転対応型ネオコグニトロンや従来型ネオコグニトロン（総称してネオコグニトロン型モデルと呼ぶ）には数多くのパラメータがあるが、その中でもS細胞のしきい値は認識性能に大きく影響する。また認識対象が異なれば最適なしきい値も異なるため、現在のところ学習前にその値を決定する方法は存在しない。よって最適なしきい値を決定するにはしきい値空間の全探索を行なう以外に方法はない。検索すべきしきい値空間は、例えばあるネオコグニトロン型モデルが4層のS層から構成されているとすると4次元空間となる。この空間内のある1点が一つのしきい値の組に対応する。最適なしきい値の組を求めるためには4次元空間内の各点に対応した大量のネオコグニトロン型モデルを学習によって構築し、その中で最も高い認識性能を持つネオコグニトロン型モデルを一つ選択するという操作を必要とする。1997年に行なわれたLovellらの研究によるとこのしきい値空間の全探索には数ヶ月の時間を要するため、解決すべき問題であることが指摘されている。

本章ではこの問題を解決するために細胞面群融合法を考案し、これを用いた高速生成アルゴリズムを提案している。高速生成アルゴリズムでは、実際に学習を行なうのは高しきい値を持つ1つのネオコグニトロン型モデルだけであり、その他の低しきい値を持つネオコグニトロン型モデルは学習を行なうことなく細胞面群融合法を用いて組織的に構成される。本章ではネオコグニトロン型モデルの学習過程の解析を行ない、高速生成アルゴリズムで使用される細胞面群の融合条件を導出している。

実際に数値シミュレーションを行ない有効性を確認した。すなわち本研究で使用した計算機を用いると、従来方法では計算時間に56日を要していたのに対して高速生成アルゴリズムでは約2時間に短縮されることを確認した。従って一般的には、最適なしきい値を求める計算時間が従来方法と比較して約680分の1に短縮できることが明らかになった。更に、高速生成アルゴリズムで構成されたネオコグニトロン型モデルを従来方法で構成されたネオコグニトロン型モデルと比較した結果、回路構成はほとんど同等であり、認識性能も変化がないことを確認している。

第5章 連想想起を伴う回転パターンの認識

ヒトは複雑な回転パターンを認識する際には瞬時にそのパターンを認識せず、心的回転によって与えられたパターンの回転角を補正し、そこで始めて認識結果を得ていることが心理実験により明らかにされている。本章ではこのようなヒトの認識機能と同等な数理モデルを考案し、その機能を実証している。

心理実験を行ない、どのようなパターン要素を用いて回転角の推定を行なっているかを調査した。その結果、ネオコグニトロン型モデルの第2細胞段で検出されるようなパターン要素を用いて回転角の推定をしていることが明らかになった。この結果を用いて、ボトムアップ型の神経回路網モデルとトップダウン型の連想想起モデルを結合した統合モデルを構築した。ボトムアップ型モデルの前層は回転対応型ネオコグニトロンで構成され、後層は従来型ネオコグニトロンで構成されている。

統合モデルの処理過程は以下の通りである。ボトムアップ型モデルは回転した入力パターンを直接認識できないため、仮の認識結果を生成する。仮の認識結果の生成はしきい値の低下、およびぼかしの範囲を広げることで実現される。連想想起モデルはこの仮の認識結果に対する正置パターンを連想し、第2細胞段における発火パターンを用いて回転角を推定する。推定された回転角で入力パターンを補正し、再びボトムアップ型モデルで認識を行なう。得られた認識結果が仮の認識結果と一致した場合には仮定が立証されたことになり、仮の認識結果が最終的な認識結果であるとして処理は停止する。一方、仮の認識結果と一致しない場合は仮定が否定されたことになり、新たな仮の認識結果を生成して再び検証を行なう。全ての仮の認識結果が否定された場合には、入力パターンは未知のパターンであるとして排除（リジェクト）される。

数値シミュレーションを行ない、実際に統合モデルが回転したパターンを認識できるかどうかを確認した。学習サンプルは正置したパターンのみで、評価に用いたパターンは変形、拡大縮小、位置ずれ、回転したパターンである。その結果、回転対応型ネオコグニトロンとほぼ同じ認識性能を示すことが明らかになった。ノイズに対する頑健性も確認している。すなわち、最初は鏡像でないものとして仮の認識結果を示すが、連想想起による検証の後に記憶しているパターンでないことを知り、このパターンを排除する。また、鏡像回転したパターンに対する統合モデルの動作が、心理学的事実と一致することも確認している。統合モデルは鏡像回転したパターンに対する動作について意図的に構成されていないにもかかわらずこのような結果が得られたことは、提案した統合モデルの妥当性を示すものと考えられる。

第3章で提案した回転対応型ネオコグニトロンは回転したパターンを直接認識できる場合のモデルであり、本章で提案した統合モデルは直接認識できない場合のモデルである。回転パターンを認識するという点で両モデルに違いはないが、視覚神経回路網モデルとしては心理実験との適合性から統合モデルの方が妥当であると言える。また工学的には回転対応型ネオコグニトロンの方が認識に要する時間が短く処理時間の点で優れているが、細胞数では統合モデルの方が少なく優位であるという特徴を持っている。

第6章 結論

回転したパターンを直接認識する回転対応型ネオコグニトロンと、仮定と検証に基づく心的回転を考慮した階層型視覚情報処理モデルを考案し、数値シミュレーションによりその有効性を確認した。また認識対象に応じて最適な認識性能を持つネオコグニトロンを実現するために、過去に指摘されていた計算時間に関する問題点を解決するアルゴリズムを考案しその理論的正当性を示すと共に実験的にその有効性を立証した。このアルゴリズムにより最適な回転対応型ネオコグニトロンを短時間で実現できることを明らかにした。

審査結果の要旨

工学的観点から考えたパターン認識技術は認識対象を限定することで成功していると考えられるが、ヒトの認識機能にはそのような限定はなく、より柔軟な認識機構の構築にはヒトの脳に学ぶことが必要である。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た認識システムとして、ネオコグニトロンが知られている。これは、回転した文字パターンが認識できない点で柔軟性に欠けていた。また、脳は回転したパターンを学習しなくても心的回転により認識できるが、その回路モデルは未知で認識システムとして構築されていない。著者は、ネオコグニトロンモデルを基礎にした視覚情報処理モデルについて研究を行ない、回転したパターンも認識できる数理モデルを考案し、その有効性を理論的・実験的に明らかにした。本論文はこれらの成果をとりまとめたもので、全編6章からなる。

第1章は序論である。

第2章では、ネオコグニトロンの構成・学習方法について定式化するとともに、回転パターンに頑健性がないこと、すなわち、大きく回転したパターンが認識できないことを実験により確認したことを述べている。

第3章では、パターン要素の方位を検出できる細胞面群を導入した回転対応型ネオコグニトロンを提案し、数値シミュレーションにより、その有効性を明らかにしている。サンプルとして与えた正置文字パターンの学習だけで、任意の方向に回転した文字パターンが認識でき、さらに、文字の変形、位置ずれ、拡大縮小、ノイズに対しても従来以上に頑健であることを実証している。これは文字パターンの考えられる多様な変形に対して耐性がある神経回路網による認識システムを初めて構築したもので、重要な成果である。

第4章では、従来型ネオコグニトロンや回転対応型ネオコグニトロンを認識対象に合わせて最適化するため、特定のしきい値での1回の学習により他のしきい値に対して学習することなく組織的に構成する細胞面群融合法を提案している。この手法により最適構成を求める計算時間を従来法に比べて約680分の1に短縮し、しかも認識性能に変化がないことを実証している。これは実用上極めて重要な成果である。

第5章では、回転した文字を必ずしも瞬時に認識せず心的回転により初めて認識するというヒトの認識機能を実現する数理モデルを提案している。これは、回転対応型認識モデルと連想想起モデルを有機的に結合したもので、入力パターンの仮の認識結果から回転角を自動検出し、回転を補正して認識結果の正しさを検証する機能を実現している。このモデルが心理学的事実と符合する機能をもつことなど興味深い知見を与えている。

第6章は結論である。

以上要するに本論文は、回転したパターンを直接認識する回転対応型ネオコグニトロンと心的回転を伴う階層型視覚情報処理モデルを提案し、その認識性能を明らかにするとともに、それらの最適構成の効率的な手法を理論的・実験的に明らかにし、その有効性を実証したもので、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。