

視野領域が機能分化した視覚系のモデル†

吉田千秋^{††} 豊田雅信^{††} 佐藤幸男^{††}

人間の視覚系は中心高解像・周辺低解像の視野を備えており、外界を認知する際には順次注視点を移動させながら個々の対象を認識し、またそれらの空間的な位置関係を理解している。本論文は人間の視覚系のこのような特徴を視野の機能分化とその協調という観点でとらえ、その意義を考察するとともに機能分化の概念を組み込んだ視覚系のモデルを提案する。これは認識前処理を行う周辺視および対象認識と注視点移動による空間関係理解を行う中心視という概念に基づいて、効率的な認識処理の実現を目指すものである。本論文では第一段階として単純な線図形を認識対象に設定し、機能分化した視覚系の基本的な認識原理を組み込んだモデルを計算機に構成した。この視覚系モデルは複数のニューラルネットワークモジュールから構成され、各モジュールが協調動作することによって認識動作を行う。またステップモータでカメラ視方向を自在に制御できるハードウェアシステムを試作し、周辺視と中心視とが協調し注視点移動しながら線図形を認識する動作の確認を行った。

1. はじめに

人間は情景を理解する際、視野内すべてに均等に注意を払うのではなく、興味のある対象には注視点を注ぎ認識を行うが興味のないものについては無視するという動作を無意識のうちに行っている。つまり情報の価値に応じて捨選択を行いつつ能動的に外界を認知しているのである。この認知に伴う注視点移動は機能的には眼球の自在な運動として実現されている。人間の視覚系は注視点すなわち視野の中心部では解像度が最も高く、周辺にいくほど解像度の低いものとなっている。しかも中心視と周辺視との差異は単に解像度の違いにとどまらず、機能的な違いにも関係すると考えられている^{1),2)}。

視野を中心視と周辺視とに分離して考えてみると、人間はまず周辺視で広い視野内を処理して興味対象あるいは探索対象の検出などを行い、狭い視野であるが高解像度の中心視を検出された対象に移動して対象の個別な認識を行う。周辺視によって特徴検出や群化などの低次の処理を実行し、認識など高次の処理が必要なものに対してだけ中心視を移動すればよいため、処理すべき情報量が削減され効率的な視覚情報処理が実現されていると考えられる。また個々の対象を順次注視することで対象間の空間関係を効率的に理解することができ、こうして得られた空間関係の理解を逐次的に組み立てることによって視野全体の認知を行うこと

ができると考えられる。このような人間の視覚系は能動的でトップダウン的な機能を有した処理系であるが、従来のコンピュータビジョンは解像度の均一な撮像系と機能的に一樣な処理系とによって構成されており受動的でボトムアップ的な処理を主としてきた。しかしながら外界を効率的に理解するためには視野内が機能分化し、しかも認識原理がそうした構成と適合した視覚システムの研究が必要になると考えられる。

従来から能動的な視覚システムに関する研究はいくつか行われている。効率化・知的化という観点から能動的な視覚系に対して取り組んでいるアクティブパーセプション³⁾、スマートセンシング⁴⁾など、またカメラに視線・視野制御機能を持たせ効率的な処理と認識対象の位置・大きさに対する制約の緩和を目指すもの⁵⁾、注視点移動とともに特徴抽出を行っていくことで知覚心像を作るもの⁶⁾などがある。しかしこれらの研究は視野が中心視と周辺視とに機能分化しながらも協調して動作するという概念に基づいているわけではない。人間が有する中心高解像・周辺低解像の機能分化した視野の概念を取り入れた視覚系のモデルとして、視野周辺部で粗いパターン情報を検出し、注視点移動により細かいパターン情報を分析して自己組織的に学習を進め外界の認識を行うもの⁷⁾もあるが、視野が機能分化した視覚系に関してその基本的な認識原理を追究するものではない。

本研究は視野が機能分化した視覚系に関するものであり、中心視と周辺視とを備えた視覚系に適合した認識原理を探り、注視点移動しながら効率的に外界の認識を行う視覚システムの構築を目指す。そのための第一段階として図形認識の基本的な原理を探るために単

† Vision System Model with Differentiated Visual Fields by CHIYAKI YOSHIDA, MASANOBU TOYODA and YUKIO SATO (Department of Electrical and Computer Engineering, Faculty of Engineering, Nagoya Institute of Technology).

†† 名古屋工業大学工学部電気情報工学科

純な線図形を対象を限定し、機能分化した視覚系のモデル化を行った。またカメラ視方向を自在に制御可能なハードウェアシステムを試作し、これを計算機上に構成された視覚系モデルと結合して認識動作の確認を行った。本論文においてはまず視野の機能分化と注視点移動の意味について考察し、本研究の意義を明確にする。そして新しい視覚系モデルを提案し、その設計方針と全体構造について述べる。さらに計算機上に構成したモデルの構成およびハードウェアシステムについて説明し、単純な図形を対象としたシステムの認識動作について述べる。

2. 視野領域の機能分化と注視点移動

2.1 視野領域の機能分化

機械的な視覚システムを構成するにあたり、人間の視覚系から学ぶことは多い。人間の視覚系での中心視と周辺視との差異は単に解像度の違いにとどまらず認識の深さなどの機能的な違いにも関係すると考えられている。例えば視野の各部位にアルファベットを瞬間提示した場合、中心視ではほぼ正確にアルファベットとして認識できるのに対し、周辺視ではその一部分しか知覚されなかったり、全形状を知覚できても文字として認識されないこともある。また周辺視に提示するパターンを拡大し、近似的に視力の差を補償しても同様の結果が得られることから、中心視と周辺視は異なる視野情報処理機能を持っていると考えられる⁹⁾。中心視では既得概念との照合が行われるのに対し、周辺視では主として特徴の抽出が行われている。しかし周辺視を覆い、視野を中心視のみに制限して人の顔の絵などを観察させると、顔の各部位は視点移動により認識できてもそれらの位置関係を把握することができないため、完全にその顔をイメージすることができないことが報告されている⁹⁾。また群化をとまなうゲシュタルト知覚も同様に中心視だけでは難しい²⁾。つまり周辺視は中心視に比べ認識能力が劣っているが、低次の特徴抽出による認識前処理や特徴の空間関係の把握、また群化による大局的知覚などは主として周辺視の機能であると考えられよう。

つまり視野を周辺視と中心視とに分離して考えると、人間は周辺視の低次の処理によって無意識的に情景内を体制化し、対象を整理・単純化した後に、中心視による詳細な認識を個々の対象に対して行っていることになる。

2.2 注視点移動の意義

中心視および注視点移動の役割は複数の対象が視野内に存在する状況で明確になる。例えば1冊の本の上にコーヒーカップが置いてある情景を考えると、人間は2つの対象をまとめて1つの物として認識するのではなく、本とコーヒーカップとを別々に認識し、その後それらの空間的な関係を理解すると考えられる。つまり逐次的に各対象を認識し、その後全体を理解するのである。このような逐次的な注意点の選択は物理的な注視点移動ではなく心理的な注意点の移動によっても可能ではあるが、一般には中心視を順次移動することによって実現されるであろう。もしこのような逐次的な認識ではなく、視野内のすべての対象を1つのまとまりとして同時に認識しようとしたら、認識すべき対象間の組み合わせの数は爆発的に増大することになる。例えば先の例で考えると“本”および“コーヒーカップ”というクラスのほかに“本の上にコーヒーカップ”をも1つのクラスとする概念を持っている必要があり、さらには遭遇しうる対象のすべての可能な組み合わせに対処しなければならない。逐次的な個別対象の認識と全体理解にはこのような組み合わせ数の爆発という問題を回避できる利点がある⁹⁾。

一般に、なじみのある対象すなわち過去に何度も見たことのある対象は同時・並列的に処理され、なじみのない対象に対しては逐次的に各特徴に対する処理が行われると考えられている⁹⁾。つまり初めて見る対象については視点移動により逐次的な特徴抽出が行われ位置の空間関係が理解される。しかし何度も見るに従って、目を動かさなくても周辺視によって同時・並列的に特徴の認識・空間関係の把握が行われていると思われる。実際に視点位置を観察すると、見慣れた対象ほど視点移動が少ないことが報告されている¹⁰⁾。またその学習過程における注目点の変化は、初めて見た時の視点移動を随所に現しながら次第に定まってくる¹⁰⁾。このように人間の視覚情報処理は基本的に特徴点群を逐次的に注目することでなされていると考えられ、脳内での知識表現にも大きく影響していると思われる。

われわれはこのような人間の視覚系に関する知見をもとにして、機械系の視覚のモデル化を目指す。しかし人間の視覚系は高度に発達しているため、生理学的に解明することは難しい。そのためここではパターン認識の立場からこの視覚系を考察しモデル化を試みる。その手法としては認識対象を限定し、その場合に

ついて注目すべき点(特徴点)を決定し、その時系列情報による認識手法を考案する。対象としては簡単のため三角形などの単純な線図形を考えることにする。この場合通常人間は注視点移動をほとんど行わなくとも図形を認識することが可能である。つまりこの場合は逐次的ではなくて並列的に認識処理を行っているといえる。しかし先天性白内障者が術後開眼した場合などには周辺視は見えているにもかかわらず、対象を一瞥しただけでは図形を判別することができず、注視点を移動させ頂点を逐次的に認識することによって初めて全体を三角形と判別することができることが知られている¹¹⁾。たとえ対象が簡単な多角形であっても未熟な視覚系にとってはなじみのない対象なのであり、注視点移動による逐次的な認識が不可欠であると考えられる。この場合は頂点が特徴点であり、全頂点を逐次注目することによって認識を行っているといえる。われわれは研究の第一段階としてこのような未熟な視覚系のモデル化を試みた。

このような未熟な視覚系が三角形を認識する過程と通常の視覚系が“本の上にコーヒーカップ”を認識する過程とは認識対象の複雑さに違いはあるものの、周辺視による特徴位置抽出と中心視による認識、そして注視点移動による空間関係理解という枠組みは共通している。つまり通常の視覚系は注視して“本”などを認識できるのに対して、未熟な視覚系は注視して頂点であるのか端点であるのかといった低いレベルの認識しかできないと考えられる。そして通常の視覚系では“本の上にコーヒーカップ”また未熟な視覚系では“頂点が3つ連結した閉図形”のように、注視点で認識できる対象間の位置関係を注視点移動で逐次的に理解すると考えられる。

本研究では視覚系の機能分化と協調という問題を研究していくうえで、第一段階として上に述べたような未熟な視覚機能の模擬を出発点とする。これは対象の複雑性に関わらず視覚系のモデルは低いレベルにおいては基本的に共通していると考えられ、対象を単純なものから始め段階的に複雑化していく過程において必要な機能等を検討していくことが望ましいと思われるからである。

3. 視覚系モデルの構成

ここでは視覚系モデルの設計方針、および第一段階として単純な直線図形を認識対象とした場合のモデルの全体構成について述べる。そして中心視および周辺

視がどのような処理を行い、またそれらの情報を用いてどのように注視点が移動され、またどのように認識が行われるのかについて説明する。

3.1 設計方針

われわれが提案する視覚モデルは人間の視覚系の優れた情報処理機構に学び、その機械系における実現を目指すものである。ここでは人間の効率的・能動的な情報処理がどのような機構で行われているかを工学的な視点から眺め、その処理体系を機能ごとに分類し、モジュール化することで検討した。

本視覚系モデルは中心視と周辺視とに機能分化していることを特徴とし、それぞれを異なった機能を備えたモジュールとして独立に構成する。また中心視の視方向を決定するためのモジュール、中心視および周辺視からの情報を基にして対象を認識するためのモジュール等を構成し、モデル全体を複数のモジュール構造とする。このように基本機能で各モジュールを規定し、大局的に各モジュール間の情報の流れを考察することによって、さらに具体的に各モジュールがどのような入力が必要とし、どのような出力を要求するのかを検討していく。

また以下に述べるような理由によって各モジュールは基本的にニューラルネットワークを用いて構成する。まず特徴抽出などのニューラルネットが得意とする処理はそれぞれの処理に適した形のネットワークに任せてしまうということである。本研究では各モジュール内の処理の具体的なアルゴリズムは重要ではなく、周辺視と中心視とがどのように協調動作するかが重要なのであり、その点ニューラルネットワークを用いれば具体的なアルゴリズムに関わることなく各モジュールの必要な入出力関係を得られる利点がある。また認識対象ごとに個別的な画像処理アルゴリズムを作成する心要がなくなる。さらに様々な機能を有した複数のモジュールを別個に設計してから結合すれば、1つのネットワーク構造では実現困難な高度な認識機能であっても視覚系モデルとして実現可能となる。人間の視覚系のように高度で大規模なネットワークを単一のネットワークで構成することは不可能であり、各機能ごとのモジュール構造にすることは意義深いと思われる。

3.2 モデルの全体構成

モデル全体は図1に示されるように複数個のモジュールが機能的に結合した構造を持つ。現時点では5個のモジュールが存在し、各々は基本的に3層パーセ

プトロン型のニューラルネットワークで構成されている。中心視は狭い視野を有し解像度が高く、周辺視は広い視野を有し解像度が低い。両者の間の機能的な違いをここでは対象を単純な直線図形とした場合について述べる。

対象が単一の線図形（三角形、四角形など）である場合は一般の情景内に存在する個々の興味対象を注視する場合と異なり、注視点の位置は興味対象というよりはむしろ図形中の特徴点の上にあると考えられる。実際に被験者に線図形を観察させ、そのときの注視点の動きを眼球運動測定器によって記録すると、線図形の頂点や端点などに注視点が停留することが知られている¹²⁾。したがって単純な線図形の場合には頂点や端点を特徴点とみなせると考えられる。

このような事実を鑑み、周辺視モジュールは図形全体を広い視野内におさめて同時・並列的に頂点や端点などの特徴点位置を検出する機能を備えることとする。この周辺視モジュールから供給される特徴点位置に関する情報を基にして、注視点移動モジュールが次の注視点位置を決定し、中心視を移動させる。中心視は高解像でありながらも狭い視野で特徴点をとらえることによって頂点と端点との区別・頂点の角度・線の延びている方向などの情報を抽出する機能を備える。さらに中心視の注視点移動によって得られた特徴点情報は注視点移動情報とともに符号化モジュールに送られ、認識に適した形に符号化される。そして認識モジュールは供給された符号の時系列に基づいて図形の認識を行う。

現時点では基本的なレベルとして交差しない直線図形を認識対象としている。このような図形に対する本視覚系モデルの認識動作は人間でいえば未熟な視覚系が図形を認識する場合に相当する。すなわち図形を一

瞥しただけでは特徴点の大まかな位置しか分からず、特徴点の形態やその連結関係などは理解できない場合である。しかし特徴点をそれぞれ注視することでその形態は判別でき、さらに線に沿って注視点を移動することで特徴点間の連結関係が理解できる。この動作を全特徴点について行えば、得られた情報から対象を認識することができる。このように基本的な認識動作を第一段階として実現し、次にこの動作を複数個組み合わせることで複雑な対象についても認識ができると考える。ここではまず基本的な認識動作の実現ということでモデルを構成した。以下に各モジュールのネットワーク構成と機能を示し、本視覚系モデルの認識動作について具体的に説明する。

3.3 各モジュールの処理

3.3.1 周辺視モジュール

周辺視モジュールは特徴点を検出するためのニューラルネットワークで構成される。このニューラルネットワークは低解像度の周辺視画像から特徴点（頂点と端点）の位置を検出するものである。この処理は同一のネットワークを2次元マトリクス状に並べることで実現されており、1つのネットワークは図2に示された構成になっている。入力層は2次元マトリクスであり、低解像度の周辺視画像の一部分を切り出す窓に相当する。そして入力層の各細胞の値は切り出された周辺視画像の各画素値に対応する。出力層は1つの細胞からなり、入力層2次元マトリクスの中心に特徴点があるときには1、それ以外のときには0の値をとるように教育される。この際の教育パターンは任意角度の1つの頂点（端点を含む）を入力層の任意の位置に配置したものである。この教育パターンをランダムに順

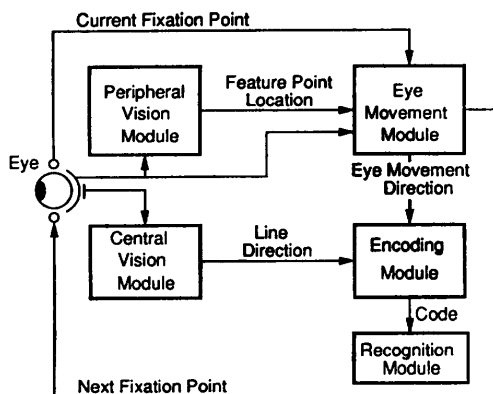


図1 視覚システム全体の構成
Fig. 1 Design of vision system.

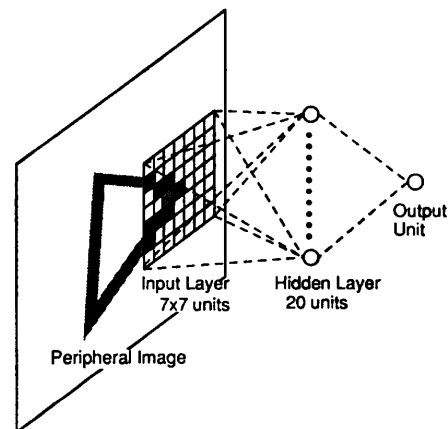


図2 特徴点を検出するネットワーク
Fig. 2 Feature detection network.

次生成し、かつ適切な教師信号をネットワークに与えることでバックプロパゲーションによる教育を行う。今回構成したネットワークでは入力層 7×7 であり、中間層の細胞数は 20 個である。

得られた特徴抽出用のネットワークを図 3 のように 2 次元マトリクス状に並べれば、周辺視画像の画素数と同数の出力を持った特徴抽出層が構成される。この特徴抽出層のなかで発火している細胞の位置が線図形中の特徴点の存在する場所を指し示すことになる。こうして得られた特徴点位置を示すマトリクス（ネットワークの出力）が注視点移動モジュールに送られる。

3.3.2 注視点移動モジュール

注視点移動モジュールは周辺視で位置検出された特徴点を中心視で順次処理していくために必要な注視点移動情報を供給するものである。

周辺視で検出された特徴点は複数個存在するので、まず最初にどの特徴点に注視点を移動するかを決めなければならない。また注視点を順次移動するときも同様であり、複数個の特徴点のなかから 1 つを選択する必要がある。ところで、中心視の視野は各特徴点を 1 つずつとらえるだけの広さしかないため、図形を認識するためには各特徴点間の連結関係を周辺視から検出する必要がある。そこで現時点では次の注視点位置を決定するために以下のような規則を設け、連結関係に関する情報は各特徴点間を注視点移動する順序で表現している。

具体的には現在ある特徴点を注視しているとして、必ずそれと連結している特徴点に注視点を移動することにする。こうすれば注視点移動によって得られた特徴点の系列は互いに連結している特徴点を順番に結んだものとなり、ここに含まれる情報から図形を認識す

ることができる。

3.3.3 中心視モジュール

中心視モジュールは中心視野に存在する頂点あるいは端点について、それを構成する線分の延びている方向を検出する。この方向検出ニューラルネットワークの構成を図 4 に示す。入力層は 2 次元マトリクスであり高解像度の中心視野に対応し、線図形画像の特徴点などの部分を狭い視野内にとらえる。現段階では出力層は 24 個の細胞からなり、各細胞は 0° ~ 360° の角度情報を 15° 間隔で離散化したものに相当する。例えば図 4 のような頂点が入力層の中心部に位置しているとき、出力層では 90° と 180° に対応する細胞が発火することになる。なお教育パターンは任意角度の 1 つの頂点あるいは端点であり、これを入力層の中心部に与えて教育を行う。今回構成したネットワークでは入力層 15×15 であり、中間層の細胞数は 20 個である。

周辺視画像と中心視画像との解像度は異なるため、周辺視画像から抽出された特徴点位置に中心視を移動したとき、入力層（中心視野）のちょうど中心に頂点が存在するとは限らない。しかし小規模のネットワークに対して頂点の位置に関係なく正しい線の方角を出力させることは困難である。そこでここで用いた方向検出ネットワークは中心に頂点が存在する場合だけを対象としている。そのかわりに周辺視と同様に特徴点（頂点・端点）に発火するように教育したネットワークを用いて中心視の位置の微調整を行うことにした。つまり周辺視で検出された特徴点位置にまず注視点を移動し、つぎに図 5 のネットワーク構成によってさらに特徴点の正確な位置を検出し、再度注視点移動を行って方向検出ネットワークのちょうど中心部に特徴点を位置させる。このようにして線の延びている方向が検出され、この情報が符号化モジュールに送られる。

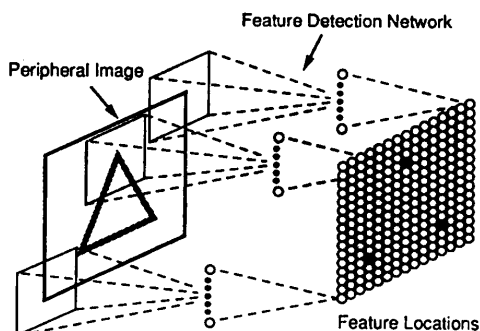


図 3 特徴点位置検出のために並列配置されたネットワーク

Fig. 3 Networks arranged to detect feature locations.

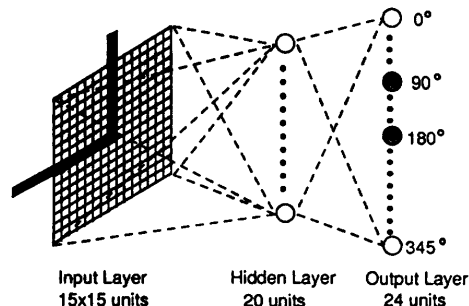


図 4 特徴点における線分方向検出のためのネットワーク

Fig. 4 Network for line direction detection.

3.3.4 符号化モジュール

符号化モジュールは中心視モジュールから得られた特徴点情報を認識モジュールが理解できるような符号に変換するニューラルネットワークである。

現段階では図形の認識方法として端点・凹・凸の3つの符号からなる時系列に基づいた判断を行っている。例えば、“凸・凸・凸”なら三角形，“端点・凸・凸・端点”であれば“コ”である。符号化ネットワークは、現在注視している特徴点の線の方向の情報（中心視モジュールから）と、現在位置に注視点移動してきた方向に関する情報（注視点移動モジュールから）を入力として、注視している特徴点が端点・凹・凸のいずれであるかを出力する（図6）。これは単純に考えれば、特徴点位置で左右どちらに曲がるかで凹・凸を判断していることになる。

当然ながら図形全体の情報から凹・凸は定まるもの

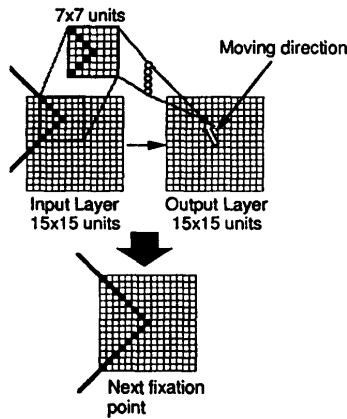


図5 視点位置微調整のためのネットワーク
Fig. 5 Network for fixation calibration.

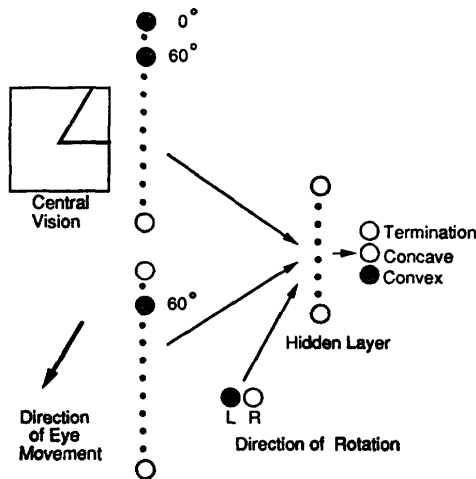


図6 符号化ネットワーク
Fig. 6 Encoding network.

であり、局所的な屈曲情報から定められるものではない。しかし三角形などの閉じた図形の場合には、左周りに注視点移動しているのが右周りに注視点移動しているのかが分かっていたら、凹・凸は判断できる。具体的には視野内の一番上の特徴点（必ず凸）を最初の注視点として、次の注視点に動くとき右周り・左周りどちらの方向に動いたのかを判断してこの情報を符号化ネットワークに入力している。また“コ”のように閉じていない図形の場合、凹・凸という概念は意味がなく、単に左右どちらに曲がったかを意味することになる。また閉じていない図形で最初の注視点端点でない場合、最終的に一方の端点に注視点到達するともう一方の端点などが未確認特徴点として残ってしまう。このときは現在の注視点位置から再度注視点移動を始め認識を行うことにした。

3.3.5 認識モジュール

認識モジュールは、符号化モジュールから送られた端点・凹・凸の符号系列から、その図形が何であるかを認識するネットワークである。用いたものは図7のようなリカレントネットワークであり、各細胞状態の更新は同期的に行われる。教育はこのリカレントネットワークと同一の動作をする多層フィードフォワードネットワークを対象としてバックプロパゲーションで行った¹³⁾。入力には端点・凹・凸のいずれかであり、例えば“凸凸凸”が入力された時点で三角形、もう1つ凸が入力された時点で四角形が出力される。また“凸凸凹凸”のような図形では凹が時系列上のどの位置にあっても同じ図形として認識されるように、起こりうる組み合わせ（この図形の場合は4つ）を同じカテゴリとして学習させる。どの時点での出力を認識結果とするかであるが、現段階では未確認の特徴点数を数えそれがゼロになったときの出力を認識結果としている。

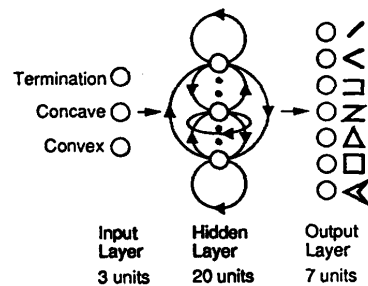


図7 認識ネットワーク
Fig. 7 Recognition network.

4. ハードウェアシステムの構成

注視点を移動させながら外界を認識する視覚システムの動作の検証のために、高解像の中心視および低解像の周辺視を備え、その注視点を移動させる視覚システムハードウェアを作成した(図8)。このシステムは周辺視用の広角レンズを装着した小型 CCD カメラおよび中心視用の望遠レンズを装着したもう一台の小型 CCD カメラとを備えている。そしてこれら2台のカメラを装着した台を水平回転用と垂直回転用の2台のステップモータで駆動して視方向を全方位自在に変化させる。2台のステップモータは直接にはパーソナルコンピュータによって制御され、ホストとなるワークステーションがシリアルラインを介してパーソナルコンピュータに命令を送ることによって間接的に全体を制御する。モータの駆動によるカメラの視方向移動速度は最大 288 (deg/sec) であり、中心視の視野角はズーム可変で $5.5 \times 4.2 \sim 32.2 \times 24.4$ (deg), 周辺視の視野角は 38.7×29.6 (deg) である。ワークステーション上には前章で述べた視覚系のモデルがソフトウェアとして構築されており、このモデルが視覚システムハードウェアを高速に制御しながら中心視および周辺視の2枚の画像をワークステーションに適時取り込み処理を行う。

本ハードウェアシステムは周辺視用のカメラと中心視用のカメラとが一体となって動く構成となっており、人間の視覚系と同様に中心視の視野は常に周辺視野のほぼ中心部に位置することになる。システムの構

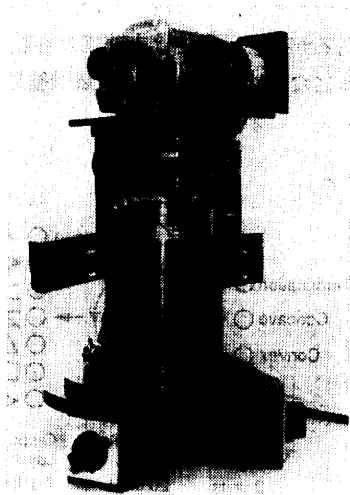


図8 視覚システム
Fig. 8 Vision system.

成としては周辺視用カメラは固定で中心視用カメラだけを駆動するものも考えられるが、人間の視覚系との相似性および周辺視も可動とすることによる潜在的な視野の拡大という点を鑑み本システムの構成とした。この場合外界の時間的変化がないなら必ずしも注視点移動をするごとに周辺視画像を取り込む必要はなく、認識対象の全体像が最初に周辺視用カメラで撮像されていれば認識過程においては中心視カメラからの画像入力を行うだけでよいことになる。

5. 視覚システムの動作

ワークステーション上に構成された視覚系モデルにハードウェア制御プログラムを結合し、実際にカメラの注視点を移動させながら線図形を認識する動作の確認を行った(図9)。今回認識対象とした線図形は単純なものであるため、周辺視用カメラ一台で実験を行った。最初にカメラから取り込んだ画像をフィルタリング・サブサンプリングして低解像の周辺視画像とし、注視点移動を伴う認識過程においてはカメラの中心部分の画像を適時取り込み高解像の中心視画像として処理した。

周辺視画像は最初にカメラから取り込まれた 512×512 の画像に 16×16 の平均値フィルタをかけ、その後 16 画素間隔で粗くサンプリングすることによって生成された 32×32 の画像である。この周辺視画像から視覚系モデルが特徴点位置を抽出し、選択された特徴点位置にカメラの中心を一致させるよう視方向を移動させる。カメラの中心を囲む 60×60 の画像を取り込み 4×4 の平均値フィルタリング・4 画素間隔のサブサンプリングを行った 15×15 の画像を中心視画像とした。この場合中心視画像は周辺視画像よりも4倍解像度が高いことになる。認識時間は周辺視の特徴点

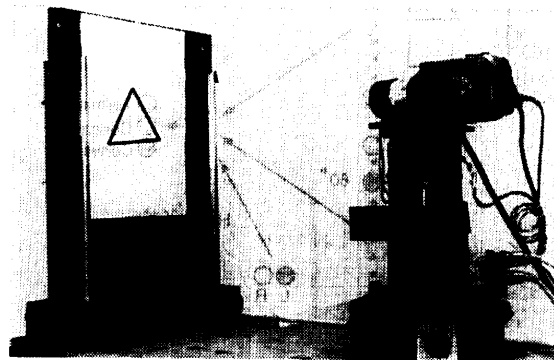


図9 実験風景
Fig. 9 Setting of experiment.

検出に 20 秒、一頂点あたりの視点移動・符号化処理に 5 秒程度である。例えば三角形の場合は 35 秒で認識が完了する。

現時点では本視覚システムの動作確認のため認識対象を 2~4 個の異なる符号系列で表される 7 つの線図形に限定した (図 10)。そしてそれらを適度に変形し、位置・大きさ・回転の変動を加えたものを対象に認識実験を行った。図 11 に認識の成功例と代表的な認識

失敗例とを示す。成功例から分かるように線図形の位置、大きさあるいは回転の変化に影響されずに、本視覚システムは正しい認識を行うことができる。認識に失敗する主な原因としては、周辺視で適切な特徴点位置を検出できなかった、頂点を端点と誤判断した、の 2 つが挙げられる。周辺視において特徴点位置の検出に失敗する主な原因は直線 (特徴点ではない) と鈍角 (特徴点である) との区別が難しいことにある。また頂点を端点と誤判断する原因は中心視モジュールの出力にある。出力は 15° 間隔の離散値であるから、これより鋭い角の頂点については端点との区別が困難である。出力の間隔を 15° よりさらに小さくすることはできるが、中心視の視野自体が離散的であるためこの問題を完全に解決することは難しい。



図 10 認識対象図形
Fig. 10 Line-drawings for recognition.

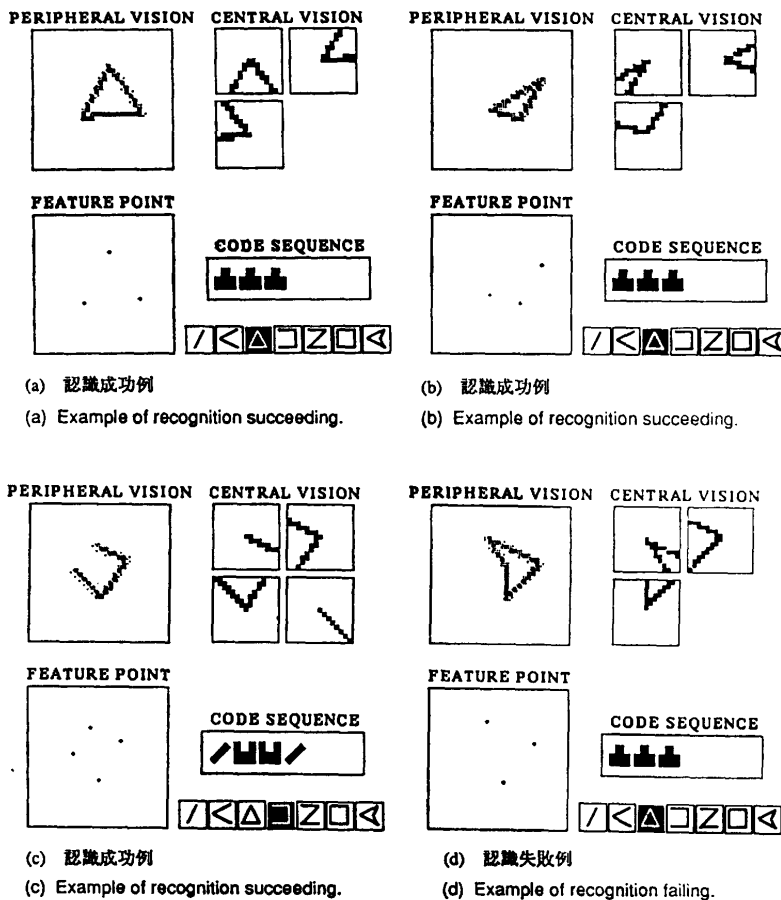


図 11 視覚システムの認識動作例
各動作例中、左上: 周辺視画像, 右上: 各注視点における中心視画像, 左下: 特徴点位置, 右中: 符号系列, 右下: 認識結果
Fig. 11 Demonstration examples.

Within each example, Top left: a peripheral vision image, top right: central vision images at each fixation point, bottom left: feature point locations, middle right: a code sequence, bottom right: the result of the recognition process.

6. むすび

視野領域が中心視と周辺視とに機能分化し、これらの協調作業によって注視点移動しながら外界を認識するという視覚システムのモデルを提案した。これは中心視と周辺視の機能分化と協調という概念および能動的な情報収集という視点に立つことによって、人間のような効率的で知的な視覚システムの構成を目指すものである。

現段階では認識対象は単純な直線図形に限定した。そしてニューラルネットワークからなる複数モジュール構造の視覚系モデルおよびステップモータでカメラ視方向を自在に制御できるハードウェアシステムを構成した。さらにこの視覚系モデルとハードウェアとを結合し、単純な線図形を対象として認識動作の確認を行った。この視覚系モデルは周辺視で特徴点位置を並列的に抽出し、中心視で各特徴を逐次的に処理する基本的な構成となっており、個別に機能定義・教育されたニューラルネッ

トワークモジュールの有機的結合体である。これによって位置・大きさ・回転に不変な認識などの高度な機能がニューラルネットワークだけを用いて実現されている。

われわれはこのシステムを一般三次元物体認識に応用する方向で考えている。そのためには例えばエッジ検出後に線図形の分割機能などを持たせ、図形を整理・単純化した後に注視点移動による逐次的な処理を適用していく必要があると考えられる。その際、各処理に適したニューラルネットワークがどのような形式のものなのかを検討していくことが重要であると考えられる。またさらには単純なフィードバック情報だけでなく、知識に基づいたトップダウン的な情報によって注視点位置を制御することも重要であると考えられる。

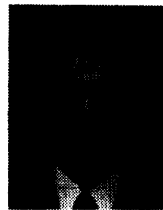
謝辞 本研究を進めるにあたり有益なご助言をいただいた日本工業大学渡部毅教授に感謝する。なお本研究の一部は文部省科学研究費補助金重点領域研究「脳の高次機能の計算論的および実験的研究」の補助を受けた。

参 考 文 献

- 1) 福田忠彦：図形知覚における中心視と周辺視との機能差，テレビ誌，Vol. 32, No. 6, pp. 492-498 (1978).
- 2) Watanabe, A. and Yoshida, T.: Roles of Central and Peripheral Vision in Pattern Perception, *NHK Technical Monograph*, Vol. 21, No. 3, pp. 23-31 (1973).
- 3) Bajcsy, R.: Active Perception, *Proc. IEEE*, Vol. 76, No. 8, pp. 996-1005 (1988).
- 4) Burt, P. J.: Smart Sensing within a Pyramid Vision Machine, *Proc. IEEE*, Vol. 76, No. 8, pp. 1006-1015 (1988).
- 5) 前川 仁，平岡茂彦，森田龍彌，藤井克彦，江見 勉：視線・視野制御機能をもつ画像処理システムの試作，信学論 (D)，Vol. J 63-D, No. 10, pp. 843-850 (1980).
- 6) 胡 金玲，倉田 是：回転可能な TV カメラによる直線複合画像の画像理解，信学技報，IE 89-11, pp. 49-56 (1989).
- 7) 大森隆司，中野 馨：自己組織化細胞を含む視覚系のモデル，計測自動制御学会論文集，Vol. 20, No. 1, pp. 56-63 (1984).
- 8) Ballard, D. H.: Eye Movements and Visual Cognition, *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 188-200 (1987).
- 9) Treisman, A.: Perceptual Grouping and Attention in Visual Search for Features and for Objects, *Journal of Experimental Psychology*, Vol. 8, No. 2, pp. 194-214 (1982).
- 10) Noton, D. and Stark, L.: Scanpaths in Saccadic Eye Movements While Viewing and Recognizing Patterns, *Vision Res.*, Vol. 11, pp. 929-942 (1971).
- 11) 鳥居修晃：視覚の心理学，サイエンス社 (1982).
- 12) Zavalishin, N. V.: Hypothesis Concerning the Distribution of Eye Fixation Points during the Examination of Pictures, *Autom. Remote Control*, Vol. 29, pp. 1944-1951 (1968).
- 13) Rumelhart, D. E., McClelland, J. L. and the PDP Research Group: *Parallel Distributed Processing*, pp. 354-361, The MIT Press, Massachusetts (1986).

(平成 3 年 11 月 1 日受付)

(平成 4 年 6 月 12 日採録)



吉田 千秋

1984 年慶應義塾大学工学部電気卒業。1986 年同大学院修士課程修了。同年名古屋工業大学工学部電気情報助手。1991 年ワシントン大学経営学大学院修士課程入学。名古屋工業大学在職中、視覚系の情報処理機能に関する研究に従事、現在に至る。



豊田 雅信

1990 年名古屋工業大学工学部電気情報卒業。1992 年同大学院博士前期課程修了。同年(株)C・T・I入社。在学中視覚情報処理、神経回路網などの研究に従事、現在に至る。



佐藤 幸男 (正会員)

1975 年慶應義塾大学工学部電気卒業。1980 年同大学院博士課程修了。同年東京農工大学工学部電子助手。1983 年名古屋工業大学工学部電気講師，1985 年より同大電気助教授。1986～87 年カリフォルニア大学サンタバーバラ校客員助教授。1987 年～88 年南カリフォルニア大学客員研究員。現在名古屋工業大学工学部電気情報助教授。3 次元物体の計測と認識，オンライン文字認識，画像解析とパターン情報処理の研究に従事。工学博士。電子情報通信学会，IEEE 各会員。