2. 画像処理の技術を拓く

2-1 人工神経回路

正会員 福 島 邦 彦[†]

1. ま え が き

1.1 人工神経回路とは

最近のコンピュータの発達はめざましく、種々の分野でのコンピュータ化が進んでいる。しかし、このようにコンピュータが発達し情報処理能力が高まったにもかかわらず、画像処理や視覚パターン認識に関しては、我々人間が日常簡単に行っているような処理すら満足に行えないのが現状である。これは、必ずしもコンピュータの演算速度や記憶容量の不足によるのではない。いったいどのような処理をコンピュータにやらせたらよいのかわからないことに原因がある。そこで、現にそのような能力を持っている生物の視覚神経系を調べて、それと同じような構造を持った人工神経回路を作ろうと考えるのはごく当然のことである。

人工神経回路とは、神経細胞の働きを単純化したような多入力1出力の演算素子(細胞と呼ぶ)を多数つなぎ合わせて作った回路網のことである。細胞間の結合の強さを結合係数あるいは重みと呼ぶこともある。電子回路にたとえてみれば、各細胞は、多入力1出力の演算増幅器のようなものである。結合係数は、各端子からの入力信号に対する増幅率に対応する。しかし演算増幅器に似ているとはいっても、その入出力特性は線形ではなく、2値の閾値特性を持った素子や、半波整流型特性を持った素子、あるいはS字型特性を持った素子など、目的に応じて種々の素子が用いられている。

1.2 パーセプトロン

このような人工神経回路を用いて視覚パターン認識を行おうとする試みは古くからあった。そのなかでも

有名なのは、1950年代の終わりに Rosenblatt が提唱した3層パーセプトロン¹⁾である。パーセプトロンは学習能力を持っていて、パターン認識能力を学習によって自分で身につけていくというので、多くの研究者の注目を集め、その後の神経回路の研究にも大きな影響を与えた。

3層パーセプトロンは、その名のとおりの3層の階層構造を持つ神経回路モデルである。第1層目は視細胞(光に反応する細胞)が並んだ入力層、第2層目はA細胞と呼ばれる細胞の層、第3層目はR細胞と呼ばれる細胞の層である。R細胞は、認識結果を出力する出力細胞である。そのなかで、第1層と第2層の間の結合(すなわち第2層のA細胞の入力結合)は変化しない固定結合であるが、最上位層のR細胞だけは、可変入力結合を持っている。

パーセプトロンに対しては種々の学習法が提案されているが、誤り訂正学習法が最も有名である。学習時には、パーセプトロンの外部にいる“教師”が、パーセプトロンに学習パターンを1つ呈示することにR細胞の出力を観測し、反応に誤りがあった場合にはR細胞に可変結合修正信号を送り、その反応が正しくなる方向に可変結合の強度を修正していく。すなわち、R細胞が出力を出すべきであるのに出し損なった場合には、R細胞に信号を送り込んでいた可変結合(つまりそのとき実際に出力を出していたA細胞からの結合)の強度を一定値だけ増加させて、R細胞が出力を出しやすくなるように修正する。逆にR細胞が誤って出力を出してしまった場合には、この逆方向に修正を加える。

パーセプトロンは当初、学習さえさせれば何でもできる万能のシステムのように思われていた。しかし研究が進むにつれて、その能力の限界が次第に明らかになってきた。一般に、階層構造を持つ多層神経回路の情報処理能力は、細胞層の数が増えるにつれて急激に増大する。3層パーセプトロンは、細胞層を3層しか

[†] 大阪大学 基礎工学部 生物工学科教授

“2. Enriching the Technology of Image Processing 2-1 Artificial Neural Networks” by Kunihiko Fukushima (Department of Biophysical Engineering, Faculty of Engineering Science, Osaka University, Osaka)

持たず、しかもその中で可変入力結合を持つのは最上位層の細胞だけであるので、その情報処理能力にもおのずから限界があったのである。しかし、多層の神経回路を学習させることのできる実用的な学習法は、なかなか見つからなかった。こうして、1970年代になると、神経回路の研究も下火になってしまった。

1.3 バックプロパゲーション

ところが1980年代後半になると、バックプロパゲーション (back propagation) と呼ばれる誤り訂正型学習法が、多層回路にも適用可能な学習手法として注目を集め、その有効性が、理論的な考察だけでなく、具体的な問題に対して実証されるようになってきた²⁾。これが1990年代のニューロコンピュータのブームのきっかけのひとつを作ることになったのである。もっとも、このバックプロパゲーションの考え方そのものは、Rosenblatt がすでに1962年に彼の著書¹⁾の中で述べていたが、まだ演算能力が低かった当時のコンピュータでは、その有効性を実験的に示すことはほとんど不可能に近く、当時はほとんど注目されなかった。

バックプロパゲーションでは、多層回路の最上位層の細胞の入力結合だけでなく、中間層の細胞の可変入力結合の値も、出力細胞の誤出力を減少させる方向に、最急降下法に従って少しずつ補正していく。つまり、中間層の細胞の入力結合に対しては、その値をどのように変化させれば最上位層の出力細胞の誤差が減るかを、その細胞から出力細胞に至る信号の流れを勘案して計算しながら、少しずつ補正していくのである。

バックプロパゲーションによって学習させると、多層神経回路は（ある程度の回路規模を持っていれば）、教えられたことは一応正しく答えるようになっていく。最近、人工神経回路にこれだけ多くの関心が寄せられているのは、このような学習能力に魅力があるからである。

1.4 汎化能力と回路構造

ところで、このような学習において重要なのは“汎化”の能力、すなわち“一を聞いて十を知る”能力である。例えば学習能力を持つ視覚パターン認識システムを作り、そのシステムに文字を教える場合を考えてみよう。もし、そのシステムが汎化の能力を持っていなければ、システムは学習時にあらかじめ教えられたパターンだけしか正しく認識することはできない。少しでも変形したパターンが与えられると、システムが正しく認識するという保証はまったくない。しかし手書きの文字は、書く人によっても、あるいは同じ人で

も書くたびごとに異なった形となる。将来現れる可能性のある変形パターンの数はほとんど無限に近いため、それらのすべてをあらかじめ教えておくなどということは不可能である。

しかし汎化能力があれば、未学習のパターンが与えられても、それが過去に学習したパターンと似ていさえすれば正しく認識できる。また、行うべき処理の手順を逐一指示されなくても、あいまいな指示に対しても相手の意志を汲んで正しく処理していくことができ、人間のように融通性のある情報処理が可能になる。

人工神経回路に生物の脳と同じような汎化能力を持たせるためには、脳が生まれながらにして(遺伝的に)持っているのと同じ構造上の制約を、人工神経回路にもあらかじめ持たせておくことが必要である。ところが最近では、「回路の構造をどのようにするかなどは何も考えないで、適当に回路を作って学習パターンを与えてやりさえすれば、あとは神経回路が自分で学習して、どのような仕事でもするようになっていく」と信じている人が多すぎるように思われる。たしかに、構造に制約のない回路の魅力は、何を学習させる場合にも同じ回路が使えるということである。しかし、どのようなものでも学習するということは、逆にいえば、学習結果に対する汎化の能力がほとんどないということにほかならない。

筆者らの研究の狙いは、生物の脳の神経回路の構造を解明し、人工神経回路の設計に取り入れていくことである³⁾。本稿では、生物の視覚神経系と同じような構造上の制約を持った回路の実現を目指して考えられたパターン認識システムの例として、筆者らが提唱した“ネオコグニトロン”と“選択的注意のモデル”について、最近の研究の発展を紹介する。

2. ネオコグニトロン

2.1 従来のネオコグニトロン

ネオコグニトロン⁴⁾は、視覚パターン認識能力を持つ多層の人工神経回路である。入力パターンが変形したり、大きさが変わったり、あるいは位置がずれたりしても、あまり影響を受けずに正しくパターンを認識する。

多層回路の最下位層は、光に反応する視細胞が並んだ入力層である。その後には、S細胞と名づけた細胞の層と、C細胞と名づけた細胞の層とが、交互に何段も結合されている。つまり、S細胞層とC細胞層とが組になって、1つのモジュールを形成し、そのようなモジュール(細胞段)が何段も積み重なって階層構造

をなしている。

S細胞は、入力パターンの特徴を抽出する細胞である。生理学という単純型細胞 (simple cell) に似ているので、その頭文字をとってS細胞と名づけられた。C細胞は、生理学という複雑型細胞 (complex cell) に似た性質を持っている。後述するように、このC細胞が、ネオコグニトロンに汎化の能力を持たせるうえで重要な役割を果たしている。

入力層に与えられた視覚パターンの情報は、このような回路の中で処理されながら上位の層に伝えられていく。認識結果を示すのは、最上位層のC細胞である。

ネオコグニトロンでは、個々の細胞は1つ前の層から信号を受け取っているが、前層のすべての細胞から信号を受け取っているわけではない。比較的狭い範囲にある細胞だけからしか結合を受け取っていない。しかも、信号を送り出している細胞と、信号を受け取っている細胞とは、それぞれの層の中でほぼ同じような位置を占めている。この点でも、ネオコグニトロンは生物の視覚系と似た構造を持っている。

2.2 特徴の抽出と統合

下位の層のS細胞は、入力層上の狭い範囲から信号を受け取って、入力パターンの局所的な特徴を抽出する。例えば、アルファベットのAを学習したネオコグニトロンの1番下の段 (U_{S1} 層) には、図1に示すように、Aという文字の種々の部分パターン、すなわち局所の特徴に反応するS細胞が並んでいる。

このような局所の特徴を抽出するS細胞が、どのような組合せで反応しているかを見ているのが、その次の段 (U_{S2} 層) のS細胞である。この細胞は、前の段 (U_{S1} 層) の細胞に比べると、もう少し大局的な特徴を見ている細胞であるということができる。

そのもう一段上には、この大局的な特徴を抽出する細胞の出力をたくさん集めて、もっと大きい特徴を見ている細胞がある。つまり、この細胞は、大局的な特徴を見る前段の細胞の反応を介して、間接的に入力層全体の情報を捕えていることになる。したがって、Aというパターンが与えられるとAに対応した細胞が

出力を出し、Bというパターンが与えられるとまた別の細胞が出力を出す。

2.3 変形に強いパターン認識

C細胞は、いくつかのS細胞から固定結合を受け取っていて、そのうちのどれか1つのS細胞が出力を出せば、C細胞も出力を出すようになっている。これらのS細胞はいずれも同じ形の特徴を抽出するが、特徴を抽出してくる場所はS細胞ごとに少しずつ異なっている。つまり、1個のC細胞は、わずかに場所が違ったところから、同じ特徴を抽出してくるような一群のS細胞の出力を受け取っている。したがって、入力パターンの位置が少し変わって、ある特徴の場所が入力層(網膜)の上で少しずれると、最初とは異なる別のS細胞がその特徴を捕えて反応するかもしれないが、いずれにしても、これらのS細胞の出力を全部まとめて受け取っているC細胞は、出力を出し続けることになる。つまりC細胞は、S細胞が抽出した特徴の、位置ずれの影響を吸収する性質を持っている。

多層回路内にはS細胞の層とC細胞の層とが交互に繰り返して階層的に結合されているので、入力パターンの情報は、多層回路内の各段でS細胞による特徴の抽出とC細胞による特徴の位置ずれの許容化を繰り返しながら上位段に送られていくことになる。この過程において、下位段で抽出された局所的な特徴は、次第に大局的な特徴に組み上げられていく。このとき、入力パターンの拡大や縮小をはじめとする種々の変形にともなう局所の特徴の相対的な位置ずれの影響は、C細胞の働きによって少しずつ吸収されるので、最終的には入力パターンのかなりの変形に対しても影響されない出力を得ることができるのである。

例えば図2(a)のように、3つの局所的な特徴を組合せた大局的特徴を抽出する中間段のS細胞に注目してみよう。このS細胞は、このような3つの特徴が、図のような位置関係を保って並んでいることを学習段階に覚えた細胞である。学習段階では、きれいな

図1 ネオコグニトロンにおけるパターン認識の原理

図2 変形に強いパターン認識の原理

A というパターンだけしか見せられていないのである。しかし、学習が終わったあとには、C 細胞の働きによって、これらの 3 個の局所の特徴の位置が多少ずれても、図に点線の円で示した程度の位置ずれであれば許容される。したがって、図 2(b)のように、大きい A というパターンを見せられても、小さいパターンを見せられても、3 つの特徴はそれぞれの点線の円内に入っているの、正しく認識する。このように、学習段階で与えられなかったパターンでも正しく認識するのである。

2.4 教師なし学習

ネオコグニトロンでは、S 細胞の入力結合を学習によって変えることによって、特徴抽出細胞を作り上げていく。学習には、“教師なし学習”と“教師あり学習”のいずれの手法を用いることも可能であるが、まず教師なし学習について説明する。

ネオコグニトロンに用いられている教師なし学習は、一種の競合学習である。学習パターンが与えられるごとに細胞の間で競合が起こり、そのとき一番大きな反応出力を出している細胞だけが入力結合を変えることができる。それでは、最大出力を出した細胞が具体的にどのように入力の結合を変えるかという、それは、この細胞に対して信号を送り出している細胞がどれくらい強く反応しているかによる。大きい出力を出している細胞からの結合は強められるが、出力が出していない細胞からの結合は変化しない。

このような競合は、神経回路内の各所で起こるが、競合の起こる範囲はあまり大きくない。つまり、ある狭い競合範囲ごとに最大出力を出した細胞が選ばれ、その最大出力細胞だけが入力の結合を変えることができるのである。

このようにすると、最大出力で反応した細胞の入力結合は、そのときの入力刺激にぴったりと一致したテンプレートのような形に成長することになる。例えば、A というパターンが与えられたときに、ある細胞がたまたま最大出力を出したとすると、その細胞の入力結合は A というパターンにぴったり一致するマスクのような形になる。つまり、この細胞は A に反応する細胞として成長していく。次に B というパターンが与えられると、先ほどの細胞はすでに A に対して反応するような性質を身につけてしまっているので、B には反応しない。したがって今度は、その周りにあるほかの細胞で、そのときたまたま B に対して大きい出力を出した細胞が B に対して反応するようになっていく。このようにして、細胞どうしの分業体制が自動的に形成されていく。したがって、同一のパ

ターンの特徴に反応する細胞が回路内に重複して作られることがなく、むだな回路の形成を防げる。

2.5 従来の教師あり学習法

前述の“教師なし学習”では、例えば A や B というパターンを見せても、これは A であるとか、これは B であるということは何も教えず、ただ A、B、C などのパターンを何回も繰り返して見せて学習させるのである。ネオコグニトロンは、自分自身で判断の基準を作り上げて、形の似ているものを 1 つのグループにして、これは同じグループのパターンだという見方をするようになっていく。

これに対して“教師あり学習”では、A を見せたらこれは A であると教え、B を見せたらこれは B であると教えながら学習をさせていく。どちらの学習法が良いかは目的によって異なる。文字読取りの場合に

図 3 従来の教師あり学習で学習したネオコグニトロンが正しく認識した英数字パターン例

は、形の違ったパターンでも同じ文字だと認識しなければならないことが多々ある。例えば‘4’と‘4’のようなパターンは、形は違っている、両方とも同じ文字であると認識しなければ困る。しかし教師なし学習では、おそらく別のパターンだと認識してしまう。このように、文字認識の場合には教師あり学習が重要になる。

ネオコグニトロンは、教師あり学習をさせることも可能である。教師あり学習では、学習パターンをあらかじめ上手に作成しておけば、かなり高い認識率が得られる。図3は、教師あり学習で英数字を教えたネオコグニトロンが、正しく認識した変形パターンの例を示している⁵⁾。かなり変形したパターンでも正しく認識していることがわかる。

しかし、ネオコグニトロンで従来用いられてきた教師あり学習では、パーセプトロンなどの教師あり学習とは異なり、教師は、学習パターンのカテゴリーだけでなく、学習パターンの中のどの場所に重要な特徴があるのかも、同時に教えていた。このため、高い認識能力を得ることはできるが、学習パターンの作成に、かなりの人手を必要とした。これに対して前述の教師なし学習では、学習にはほとんど人手を必要としないかわりに、認識率は教師あり学習に比べて多少低くな

っていた。

2.6 ネオコグニトロンの改良

そこで筆者らは、もっと手間のかからないパーセプトロン法を求めて研究を進めている。つまり、通常の教師あり学習の場合と同じように、学習パターンとそのカテゴリーだけを教えれば、あとは自動的に学習が進んでいき、しかも、従来の教師あり学習の場合と同じ程度の認識率が得られるような、新しい学習法の開発である。

この問題を筆者らは、2つの方向から攻めようとしている。ひとつは、回路の構造を、より人間の脳に近い構造に近づけていく方法である。つまり、脳が生まれながらに持っている構造をあらかじめ組み込んで、汎化の能力を高め、手間のかからない学習をさせるというものである。もうひとつは、新しい教師あり学習法そのものの研究である。すなわち、AならA、BならBという答えだけを教えれば、重要な特徴がパターンの中のどこにあるかは、いちいち教えなくてもよいような学習法の開発である。

まず最初に、脳が生まれながらにして持っているような回路をネオコグニトロンにも組み込んだ。この新しいネオコグニトロンは、図4の上半分に示すような5段の階層構造を持つ⁶⁾（下半分については後述す

図4 新しい回路構造を持ったネオコグニトロン

(a) 教師なし学習で教えた第1段階目の学習パターン

(b) 第1段階目の教師なし学習を終わったネオコグニトロンが正しく認識したパターンの例

(c) 第2段階目の教師あり学習を終わったネオコグニトロンが正しく認識したパターンの例

図5 新しい学習法で学習させたネオコグニトロン

る)。

第1段階目 (U_{S1} , U_{C1}) には、エッジ (白と黒の境界) を検出するモジュール (S細胞層とC細胞層の組) が作りつけられている。その次の段には、直線検出細胞が並んでいる。この回路では、まずエッジを検出し、その情報を用いて直線を検出するようになっているが、エッジを検出するS細胞の層と直線を検出するS細胞の層との間にも、C細胞の層がある。一般に、直線の両側の2つのエッジ間の間隔は、線の太さによって変わっていくが、C細胞の働きによってエッジの位置ずれの影響が許容されるので、線の太さによらずに直線を正しく検出することができる。

2段階目と3段階目のモジュールの間には、線の曲がりや端点を検出するモジュールが作りつけられている。ちなみに、このような細胞は生理学的な実験によって、ネコやサルなどの大脳でも見出されている。

3段階目以降の細胞は可変入力結合を持ち、その結合の値は学習によって変わっていく。従来の教師あり学習では、1つのS細胞では抽出しきれないほど変形した特徴が現れると、その特徴を別のS細胞に学習させて、それらのS細胞の出力を1個のC細胞に集めてくる (統合する) ことによって、問題を解決するようにしていた。従来はそのような統合を必要とする特徴は、教師がいちいち教えていたが、新しい学習法では、これを自動的に行えるようにした⁷⁾。

学習は2段階に分けて行う。まず最初に、学習パターンを各カテゴリーに対して1つずつ用意し、これを入力層に呈示して教師なし学習をさせて、可変結合の強度の第1近似値を求める。図5(a)は、教師なし学習で回路に教えた‘0’から‘9’までの数字パターンである。図5(b)は、この第1段階の教師なし学習を終わった回路が、正しく認識した変形パターンの例である。学習時には、図5(a)の10種類のパターンを見ただけで、変形したパターンは一切見せていない。しかし学習を終った後には、図5(b)に示すように、かなり変形したパターンでも正しく認識するようになった。このように、実際の生物の視覚系にあるのと同じような構造をあらかじめ回路に組み込んだうえで学習をさせると、教師なし学習だけでもかなり融通性の高いパターン認識ができるようになる。

しかし前述のように、文字認識という目的に対しては、いくら変形に強くても、形の類似性だけでは完全なパターン認識はできない。‘4’と‘4’のように形の違うパターンも同じ文字だということを教えてやるために“教師あり”の学習が必要になる。

ところで、ネオコグニトロンがパターンを認識するときには、回路の中には下から上に上がっていくフォワード (求心性) の信号だけが流れているのであるが、新しい“教師あり学習”をさせるために、今度は上から下に戻ってくるバックワード (遠心性) の信号の流れも取り入れたような回路を考えてみる。図4の下側の部分がそのバックワードの回路 (Wはバックワード回路の細胞層を表す) で、この部分は学習のためだけに用いられる (バックワードの回路は、後述する選択的注意のモデルの中のバックワードの回路と、ほとんど同じ構造をもっている)。図の上半分は、ネオコグニトロンの本来のフォワードの回路である。このフォワードの回路の可変結合は、前述の“教師なし学習”で一応でき上がっている。バックワードの回路内の細胞間の結合も、それと同じ値に合わせておく。

このフォワードの回路に種々の筆跡のパターンを呈示して、どのように認識するかを調べる。もし正しい

答えを出せば、もう何もする必要がないので、そのままにして次のパターンを与える。しかし、間違った答えを出したときには、正しい答え（正しい認識結果）をバックワードの回路の一番上の層の細胞に教えてやる。するとバックワードの回路には、下の段の細胞が本当はどのように反応すべきであったかという、希望出力が現れる。

そこで、フォワードの回路の各層の細胞に現れた実際の出力と、バックワードの回路の各層の細胞に現れた希望出力とを比較して、違いがある細胞だけに学習させるのである。つまり、両者の反応が違っている場所では、特徴を抽出しそこなったのだと考えて、新しい特徴を抽出する細胞をそこに作り上げていく。このようにすると、フォワードの回路でうまく抽出することができなかった特徴を抽出するような細胞が、自動的に作られていくことになる。

図5(c)は、このような2段階の学習を終わった回路が正しく認識した数字パターンの例である。かなり変形したパターンや、筆跡の違うパターンも正しく認識していることがわかる。このように、“教師あり学習”が、手間をかけなくてもできるようになった。

3. 選択的注意のモデル

3.1 選択的注意の基本モデル

上記のネオコグニトロンでは、階層構造を持つ多層細胞神経回路の中に、下から上に信号が流れていくフォワードの結合だけしか持っていなかった。しかし脳の中での情報処理には、フォワードの結合だけでなく、上位から下位に戻るバックワードの結合も重要な役割を果たしている。このバックワードの結合を取り入れて、もっと脳に近い回路に改善することが筆者らの望みである。

上から下に戻ってくるバックワードの信号といっても、いろいろな信号があるが、ここでは“選択的注意”という機能に注目してみよう。選択的注意とは、われわれがものを見るときに、ある部分だけに注意を集中して見るというような働きである。このような注意を向けるための信号として、バックワードの信号*を使ってみようというわけである。

そこで、選択的注意のモデル⁸⁾は、ネオコグニトロン型の階層型の多層回路に、バックワードの信号を伝える結合を付加した構造にした(図6)。バックワードの系は、フォワードの系を鏡に映したような構造を持

っている。両者の違いは信号の流れの方向だけである。フォワードの系では下から上へ信号が送られていくのに対して、バックワードの系では上から下に戻ってくる。

入力層に刺激パターンが与えられると、フォワードとバックワードの回路で形成されるフィードバックループの中を、信号がぐるぐる回りはじめる。しかし、信号は、フォワードの系とバックワードの系の中を、互いに独立に流れているわけではない。2つの系は、互いに相手側の信号の流れを強め合ったり弱め合ったりするような信号を授受し合っている。つまり、フォワードの系からバックワードの系に対しては、信号の流れの経路を制御する信号を送っている。このため、バックワードの信号は、フォワードの信号と同じ経路を逆向きにたどって流れていくようになる。また逆に、バックワードの系からは、そのバックワード信号が流れている経路に対応したフォワードの系の経路に向けて、フォワードの信号を流しやすくするような制御信号を送っている。これは、入力刺激の中の特定の部分だけに注意を集中する働きをしているということができる。したがって、回路の中を信号がぐるぐる回っているうちに、回路の反応は時間とともに次第に変わっていく。

フォワードの回路は、ネオコグニトロンとほぼ同じ構造を持っていて、パターン認識の処理を行っている。フォワードの系の一番下にあるのが入力層で、ここに刺激パターンが与えられる。一番上にあるのが認識細胞層である。

この認識細胞層の細胞の出力は、今度はバックワードの系を通して下に戻っていく。バックワードの系の一番下にあるのが、“連想出力層”である。この層の働きについては、いろいろな解釈ができるが、連想出力の現れる層、あるいはセグメンテーション（切り出し）の出力が現れる層であるということができる。

図6 フォワードとバックワードの信号の流れを持つ選択的注意のモデル

* 選択的注意のモデルや、2.6節のネオコグニトロンの学習に用いたバックワードの信号は、1.3節で述べたバックプロパゲーションとはまったく別のものである。

なぜセグメンテーション出力層と呼べるかという
と、例えば入力層に‘2’と‘3’という2つのパターン
が与えられたとすると、この回路はある瞬間には‘2’
の方に注意を向けて、これは‘2’であるという答えを
認識出力層に出すと同時に、「‘2’はこういう形をして
この場所にあった」といって、‘2’の部分だけを切り
出した結果をこの層に出すからである。

この同じ層を、連想出力層と呼ぶこともある。それ
は、入力層に不完全なパターン、すなわち、かすれた
パターンや汚れたパターンを見せても、回路の中を信
号が回っているうちに、かすれや汚れが修復された、
きれいなパターンが連想出力層に現れてくるからであ
る。

以上をまとめると、このシステムは図7に示すよう
に、2つ以上のパターンが同時に与えられても、個々
のパターンを順番に認識し、切り分けていくことがで
きる。不完全なパターンが与えられても、それを修復
した完全なパターンを連想することが出来る。入力に
含まれている汚れ(ノイズ)を除去する働きもある。し
かも、このような働きが、刺激パターンが学習パター
ンとは異なった位置に与えられたときにも、あるいは
形が歪んでいる場合にも、正しく働くというのが、こ
の選択的注意モデルが、従来のシステムに比べて大き
く優っている点なのである。

3.2 英文筆記体の続け文字の認識

この選択的注意のモデルの原理を用いると、英文筆
記体の続け文字を認識するシステムも実現することが
できる。続け書きされた英文筆記体の場合は、パター
ンのマスクのようなものを作っておいて、どのマスク

に一致するかを調べるというような方法、つまりパタ
ーンマッチング的な手法ではうまく認識できない。同
じ文字の系列であっても、書くたびごとに文字の間隔
は変わる。間隔が変われば当然、文字の形も変わっ
てくる。また、同じ文字でも、前後の文字の影響で形が
大きく変化する。それらを同じ文字として認識しなけ
ればならないため、通常のパターンマッチング的な手
法では、文字認識は困難なのである。

そこで筆者らは、選択的注意のモデルの考え方を
用いて英文筆記体の認識を試みた⁹⁾。しかし前述のモ
デルそのままでは、いくつかの問題があるため、多少手
直しをした。つまり、前述のように2, 3個のパター
ンが網膜に与えられた場合はよいが、もっと多くの文
字が同時に与えられると、互いに干渉しあって、正し
い出力が出てこないこともある。それを避けるため
に、与えられた文字列に対して、観測すべき領域を自
動的に制限する目的で下記のような機構を導入した。

ある文字列が与えられると、システムはまず観測す
べき領域を自分で定め、その領域を動かしながら文字
を1つずつ認識していく。この領域を探索領域と呼
ぶ。もし、この探索領域の位置や大きさを、読み取る
べき文字の位置や大きさにぴったりと合わせるように
厳密に制御しなければならないのだとすると、問題の
解決にはならない。しかし幸いなことに、このシステ
ムの基本になっている選択的注意のモデルは、それ自
体で、すでに文字の切り出しと認識の能力を持っている。
したがって、たとえ探索領域の中に文字が2, 3
個入ってきてもかまわない。ともかく、その中に少な
くとも1つ完全な文字が含まれていさえすれば、あと
は基本となる選択的注意のモデルそのものの働きによ
って、文字の切り出しと認識が進んでいく。したがっ
て、一度に処理する文字数を多少制限するというぐら
いの簡単な気持ちで探索領域を制御してやれば充分で
ある。具体的には、入力パターン全体の中で、線分の
密度が最も高い部分を探索領域の中心に持っていくと
いう簡単なコントロール機構を導入したが、このよう
な簡単なもので大丈夫である。

1個の文字の認識と切り出しが終わって回路の反応
がほぼ定常状態に落ち着くと、システムは自動的に注
意を切り換えて次の文字の認識に移る。このとき間違
って同じ文字を2回繰り返して読んでしまうことがな
いように、回路内の細胞に疲労のような効果を起こさ
せる機構を導入した。すでに1回見た文字に対して
は、疲労の効果によって反応が抑えられるのである。

このシステムの能力を調べるために行った予備実験
の結果を紹介しよう。図8の左上に示す *late* という

図7 選択的注意モデルの諸機能

文字列を与えると、システムは、まず最初に中央付近にサーチエリアを持っていて、 t という文字を認識し切り出してきた（図は連想出力層に現れた反応を時間を追って示している）。その反応がほぼ定常状態に落ち着いたことを検出すると、注意を切り換えてサーチエリアを動かし、左側の l を認識した。このよう

図 8 選択的注意のモデルを用いた英文筆記体続け文字の認識と切り出し。入力パターン（左上）と、連想出力層の出力の時間経過を示す。認識され、切り出された文字が連想出力層に現れる。

にして、 a 、 e を順次認識していった。読取っていく順序は、必ずしも左から順番にはなっていない。すべての文字を認識し終わると、もう新たに見るものはないと判断して、自動的に操作を止める。

ところで筆記体の e と l は、同じ形のパターンが変形しただけと考えることができる。選択的注意のモデルも、ネオコグニトロンと同じように変形に強いパターン認識の能力を持っているので、 e に対しても l に対しても、同じ認識細胞が出力を出す。そこで、認識され、切り出されてきたパターンの高さを比較することによって、両者を区別するようにしている*。

図 9 は、認識と切り出しが正しくできた単語の例を示す。このようにかすれた文字や、あるいは汚れた文字が並んでいても、正しく認識し、切り出すことができる。

3.3 漢字認識システム

同じような考え方を漢字の認識にも応用できる。漢字は文字の数が多いので認識が難しいといわれている。しかし幸いなことに漢字は、偏、旁、冠、足などに分解することができる。偏や旁などを、それぞれ独立したパターンとして認識することができれば、比較的小規模な回路で多数の漢字を判読することができると思われる。まだ小規模な予備実験ではあるが、筆者らは、選択的注意のモデルの考え方を、このような目的にもうまく用いることができることを確かめている¹⁰⁾。

4. む す び

このように、神経回路を設計するとき、生物の神経回路が生まれたときにすでに持っているような回路の構造を、あらかじめ作りつけておくと、パターンを認識するうえでの汎化の能力が非常に高まる。文字を

図 9 英文続け文字認識（認識と切り出しが正しくできた単語の例）

* 高さの比較を行う神経回路を作ることは容易であるが、このモデルには現在のところ作りつけてない。

読む機械は、人間と同じように読んでくれないと役に立たない。かりに汎化の能力があるとしても、もし機械が持っている汎化の能力と、人間が持っている汎化の能力が異質のものであったのでは、どうにもならない。人間が見れば非常に似ているのに、機械が見れば似ていないと判断しては困る。機械も人間と同じ判断基準、すなわち人間と同じ類似性の基準をもって判断しなければならない。このためには、生物の視覚系が生まれたときに持っているのと同じような回路構造を、機械にもあらかじめ作りつけておくことが不可欠である。

もちろん、ここで紹介したシステムは、まだ完全に生物の視覚系をまねしているとはいえない。今後、さらに生物の視覚系に近い回路構造を持たせるように改善していく必要がある。また、さらに優れた新しい学習方法を開発することも、これから取り組まなければならない重要な問題である。(1992年5月22日受付)

〔参 考 文 献〕

- 1) F. Rosenblatt: "Principles of Neurodynamics", Spartan Books, Washington, D. C. (1962)
- 2) D. E. Rumelhart, G. E. Hinton, R. J. Williams: "Learning Internal Representations by Error Propagation", Parallel Distributed Processing, Vol. 1, eds. D. E. Rumelhart, J. L. McClelland and the PDP Research Group, pp. 318-362, A

Bradford Book, MIT Press, Cambridge, MA (1986)

- 3) K. Fukushima: "Neural Networks for Visual Pattern Recognition", IEICE Trans. (信学英文誌), E74, 1, pp. 179-190 (Jan. 1991)
- 4) K. Fukushima: "Neocognitron: A hierarchical neural network capable of visual pattern recognition", Neural Networks, 1, 2, pp. 119-130 (1988)
- 5) K. Fukushima, N. Wake: "Handwritten Alphanumeric Character Recognition by the Neocognitron", IEEE Trans. Neural Networks, 2, 3, pp. 355-365 (May 1991)
- 6) K. Fukushima, N. Wake: "Improved Neocognitron with Bend-detecting Cells", IJCNN-92-Baltimore, to Appear (June 1992)
- 7) K. Fukushima, N. Wake: "An Improved Learning Algorithm for the Neocognitron", ICANN-92-Brighton, to appear (Sep. 1992)
- 8) K. Fukushima: "A Neural Network Model for Selective Attention in Visual Pattern Recognition and Associative Recall", Appl. Opt., 26, 23, pp. 4985-4992 (Dec. 1987)
- 9) 今川, 福島: "選択的注意機構による英字筆記体連結文字列認識", 信学論, J74-D-II, 12, pp. 1768-1775 (Dec. 1991)
- 10) 芦田, 福島: "選択的注意機構のモデルによる漢字認識", 信学技報, NC91-158 (Mar. 1992)

ふくしま くにひこ
福島 邦彦 昭和33年、京都大学工学部電子工学科卒業。同年、NHKに入局。放送科学基礎研究所、放送技術研究所などを経て、平成元年より、大阪大学基礎工学部生物工学科教授。神経回路モデル(視覚パターン認識、記憶、学習機構など)やニューロコンピュータの研究に従事。工学博士。正会員。

2-2 正則化——Regularization——

1. ま え が き

画像処理の研究分野の中で、コンピュータビジョンは、対象となるシーンをカメラで観測して得られた画像データからシーン内の実体を推定することが中心課題となる。それは、光学系を通して得られた投影像である画像から対象の幾何学的な性質や表面の物理的な性質を推定することであり、広い意味で、画像生成過程の逆問題を解くことにほかならない。例えば、2次元画像から3次元情報を復元する場合には、次のような逆問題を解かなければならない¹⁾ (図1参照)。

[†] 電子技術総合研究所 知能情報部

"2-2 Regularization in Computer Vision" by Naokazu Yokoya (Machine Understanding Division, Electrotechnical Laboratory, Tsukuba)

正会員 横 矢 直 和^{**}

- (1) 画像の明るさや色は、(i)対象の幾何学的構造、(ii)表面の物理的性質(表面反射率など)、(iii)光源の種類と位置、(iv)観察位置から決まるため、観測された明るさや色の情報から(ii)～(iv)の情報を分離し、(i)の情報を抽出す

図1 コンピュータビジョンにおける逆問題