

文字認識過程における競合学習

河合 克敏, 菅沼 義昇
名古屋大学

従来, 文字認識を行うとき, 前処理・特徴抽出の段階にまで上位の知識を用いてそれらを変更することは希であった. しかし, 文字認識に限らず, パターン認識においては, ボトムアップ処理だけでなくトップダウン処理が重要になる. そこで, ボトムアップとトップダウンを競合させることにより, 同じ入力パターンに対し様々な構造を生成できることを示す.

“Competitive Learning in Character Recognition”
by Katsutoshi KAWAI and Yoshinori SUGANUMA (Department
of Information Engineering, Faculty of Engineering,
Nagoya University, Furo-Cho, Chikusa-Ku, Nagoya City,
464-01, Japan)

Pattern recognition is influenced by context. It may be impossible to recognize characters correctly only by bottom-up methods. In this paper, this problem is treated. It is supposed that a structure is constructed by a given character and the structure is memorized as knowledge of the character. However, many structures can be constructed by a character and the different structures are recognized as different characters in certain circumstances. Therefore, in this paper, it is discussed how different structures are constructed by competition between bottom-up and top-down methods.

1. はじめに

人間は、どのようにして物事を学習し、認識できるのでしょうか。また、さまざまな環境に対し、どのようにして適応することができるようになるのでしょうか。比較的多くのことがわかっていると言われている視覚の分野においてもそれらは完全には解明されていない。三次元パターンは当然のこと、二次元パターン（文字パターンなど）ですらどのようにして認識しているのか明かにされていない。人間はパターンをどのように認識しているのか。この問題を解決するためには、まず「入力パターンからどんな特徴をどうやって抽出すべきか」という重大な問題を解決しなければならない。しかし、この問題に対する明確な解答は得られていない。そのため入力パターンは、さまざまな目的や対象に応じて認識するのに都合のいいように変換され、特徴抽出が行われている。変換方法は、挙げればきりがなほたくさんある。

学習に対する基本的な考え方は別の論文において詳細に述べるが、本論文に関係する部分について簡単に述べておく。『学習』といえは、多くの正の例の類似性を調べ、それらに共通する属性を見つけ出し、その属性を用いて概念を記述しているものが多かった。この場合負の例も多少は用いられてはいたが、あくまでも補助的なものにすぎなかった。しかし、対象物が何であることを認識する場合に必要な概念は、正の例だけによって本当につくられ得るものなのだろうか。正の例に共通な属性を抽出し、その属性だけを記述しておけば、認識できるものであろうか。もしそうだとすると、認識するのに必要かつ十分な属性は何を基準にどこでどうやって決めるのか。もっと言えば、人間は本当に正の例だけで物事を認識することができるようになるのであろうか。このように考えていくと、（当然正の例なしに学習はできないが）負の例との各属性の差異が物事を認識する上で非常に大切であると考えことは不自然な話ではないし、負の例との差異を有効に用いれば、正の例は少なくてもすむはずである。実際、人間においても、むやみやたらに多くの正の例だけを提示されただけで学習が行なわれるとは考え難い。人間は典型的な正の例をいくつか学習すれば多少変形したパターンでも認識できるようになるし、識別するとき、負の例との差異を有効に使っていると考えるのもごく当り前の話である。

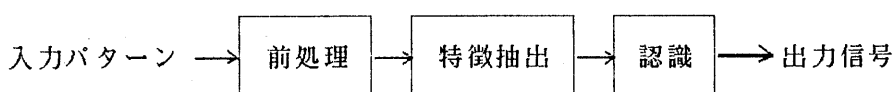
しかしながら、典型的な正の例といっても、常に全く同じ例を要求することには無理があるし、現実的ではない。そこで、多少の違いは前処理段階で吸収してくれることが望ましい。つまり、与えられた情報をそのまま使用するのではなく、事前に所有している知識によって、与えられた情報にある程度、一般化する必要性がある。さらに、この一般化は、トップダウン処理に対しても重要な役割を果たす。もしこのような一般化が行われずに、上位からの情報が視覚の各要素レベルに強く影響するような仮定を設ければ、その影響範囲は非常に狭められる¹⁾。

本論文においては、主として、この一般化過程、及び、その中にどの様な形

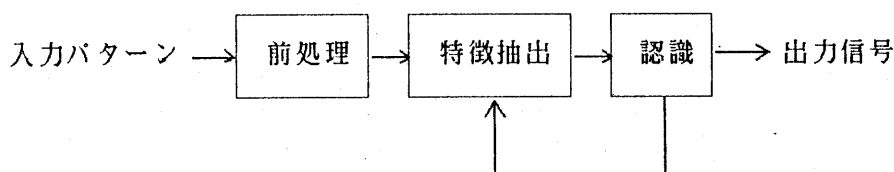
で上位からの情報が伝えられ、競合が行われるかについて論じる。

2. トップダウン処理の必要性

パターン認識システムにおいて、特徴を抽出する段階までの処理は入力パターンをシーケンシャルにボトムアップ処理するだけのものがほとんどである（図1.(a)参照）。特徴抽出を行う段階に上位からの情報を与えているシステムは少ない。すなわち、従来行われてきた方法の多くは「ボトムアップ処理に基づいている^{2) - 12)}」といえる。また、上位からの影響を考慮していたとしてもこの段階での処理は不自然なものが多い。



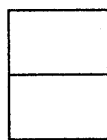
(a)



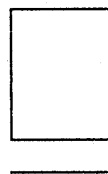
(b)

図 1

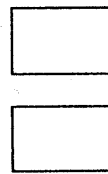
このようなボトムアップ処理だけで文字を認識するのが困難な場合が存在することは明かである。たとえば、『A'』の『'』の部分をノイズと見るか、あるいは意味のある記号とみるかは、この文字の現れた文脈の影響を強く受ける。単



(a)



(b)



(c)

図 2

純に、Grossberg の方法¹⁾のように、記憶されているパターンと一致する要素の数だけで決めてしまうわけにはいかない。また、図2の例においては、もし、

入力パターン『日』が与えられたとき、これを漢字と見なせば(b)の構造が望ましいし、もし、数字と見なせば(c)の構造が望ましい。

さらに、図3の例において、もし数字の『4』が3本の直線から構成されているような構造として記憶されていれば、入力パターン(a)が与えられたとき、それを、(b)のような構造として認識するよりも、(c)のような構造として認識する方が、『4』として認識され易いであろう。

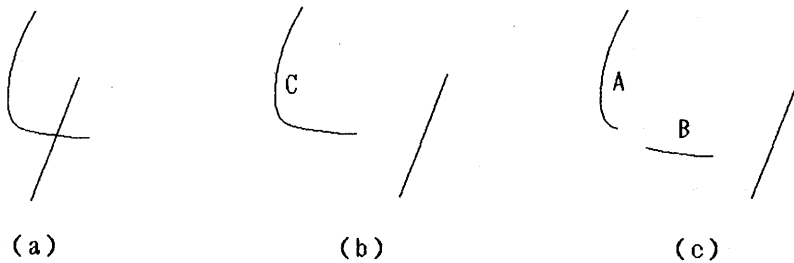


図3

そこで本論文では、文字認識を対象とし、文字がどのような構造になっているかを生成する段階で、ボトムアップ処理されたものと図1.(b)に見られるようなトップダウン処理によるものとを競合することによって、より適した文字構造を生成するシステムを提案する。我々が文字構造生成システムと呼ぶものは文字パターンを直線分あるいは曲線分に分け、それらからより一般的なグローバルノード（後述）やグローバルノード間の関係を生成するところまでを言う。すなわち、文字認識・学習システムの前処理・特徴抽出段階である。

3. 文字構造生成システムの概要

以下に文字構造生成システムの概要について記す。さきに述べたように、学習において、概念間の差異が非常に重要な役割を果たす。そこで、本論文において述べる一般化の過程においても、要素間、さらには、要素から生成されるノード間の関係に注目して入力パターンの階層的構造が生成される。

(1) 要素の方向特性

システムに与えられる入力パターンとしては、ドットパターンであり、その各要素は、上限及び下限が決められた刺激強度を持っているものとする。そこで、

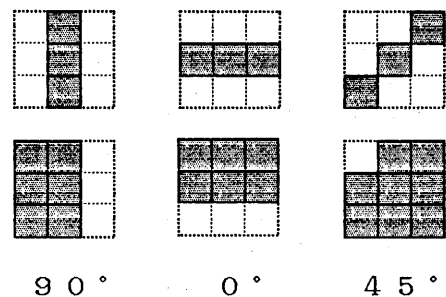


図4

それらの刺激強度及び要素位置の関係を利用して、各要素の方向特性を計算する。さらに、この方向特性は、それぞれ強度を所有しているものとする。例えば、図4の例においては、その中心要素の方向特性は上下とも同じであるが、強度は、上の図の方が大きくなる。

(2) 線分ノードの生成

同じ方向特性を持った要素がその方向特性の方向に並んでいれば、そこには直線が存在すると考えられる。また、方向特性がある一定の割合で変化していれば、そこには曲線が存在すると考えられる。このように、要素間の方向特性の差を利用して、直線または曲線に対応する線分ノードの検出を行なう。その一例を、図5に示す。この図において、各曲線で囲まれた部分が一つの線分ノードに対応し、図に描かれた細かな線分は、各要素の方向特性及びその強度を表す。

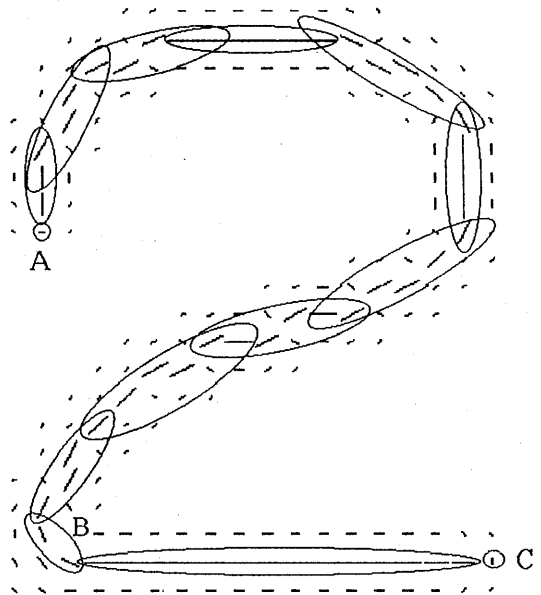


図5

(3) 線分ノードの連結

次に、上で生成された線分ノード間の関係を利用して、より一般的な線分ノードを生成する(図6)。

(4) グローバルノードの生成

同じ向きの曲率を持ち、それらを接続することによって、凸または凹の一つの線分を構成するような線分ノードを集めて新しい線分ノード(グローバルノード。例えば、図2.(b)及び(c)の長方形に対応するノード)を生成する。この処理によって、一つの閉曲線またはそれに近い線分を構成する線分ノードがまとめられる。

また、グローバルノードを生成する方法としては様々なものが考えら

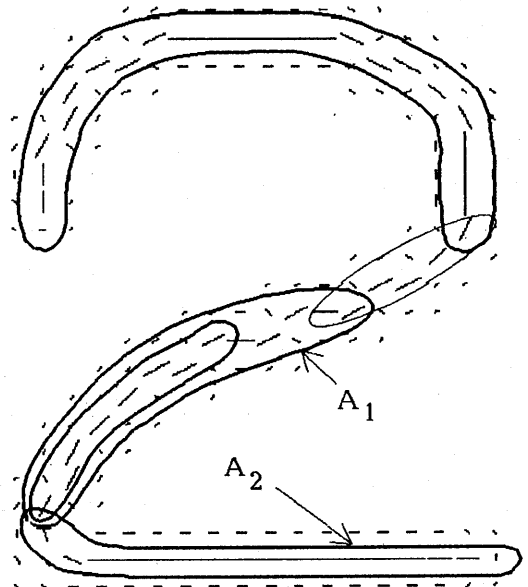


図6

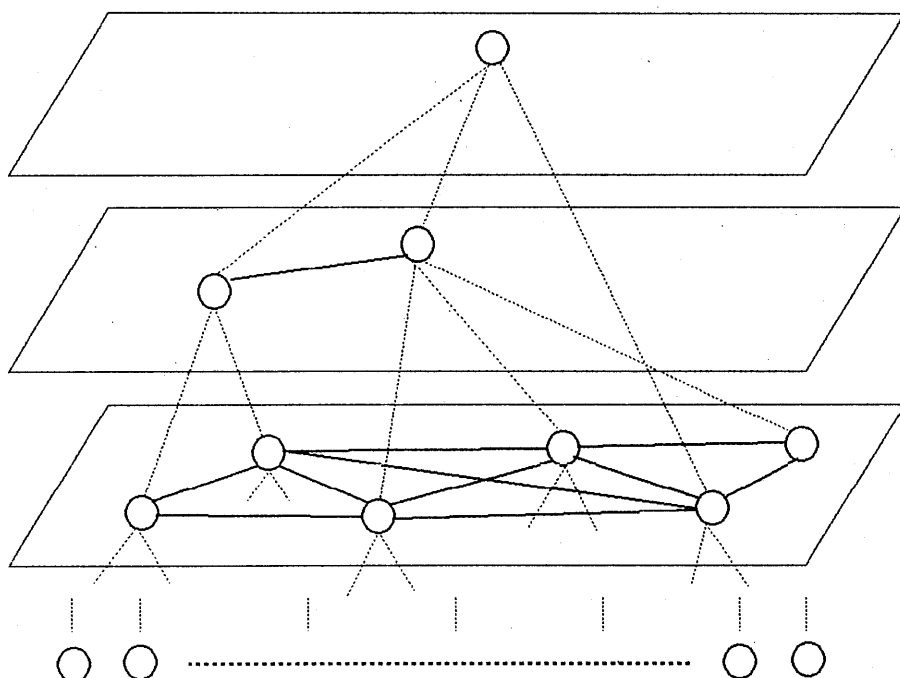


図 7

れるが、ここでは、注目するノード間の関係や属性を変化させる（重みを変化させる）ことにより、異なったグローバルノード、例えば、図 2.(b)や(c)を生成するものとする。

以上の処理によって、ある入力パターンが与えられると、図 7 に示すような階層的ネットワーク構造が生成される。この図において、最下層は各要素に対応し、最も上のレベルはグローバルノードに対応している。

4. トップダウン情報の利用

上に述べた処理の内、上位レベルからの情報を考えるにあたって、特に重要なのがグローバルノードの生成の段階である。この段階において誤った文字構造が生成されると、2章で述べた例のように、文字の認識がうまくいけなくなる可能性がある。そこで、本論文においては、文字構造生成のこれらの段階に上位レベルの知識が影響を与えると仮定し、競合によってグローバルノードの生成が行われるものとする。上位からの知識が影響した結果得られるグローバルノードがすべての場合に正しいければ、競合という手段をとらずに上位の知識に下位からのノードをあわせればよいのだが、いつでも正しいとは限らない。このことを考慮にいれるが為に競合という方法を用いることになった。

また、人間あるいはその他の動物においても、入り組んだ神経回路網の中で

このような処理が現実に行なわれている可能性は十分にあると思われる。では、競合といってもボトムアップ処理されたものとトップダウン処理されたものをどのような操作で扱うのが適当であろうか。少なくとも、パターンをそのまま記憶しておき、両者のマッチングをとり、マッチングした要素の数だけで処理することは、あまりにも不自然である。人間自体もそのような処理を行っているとは考えにくい。そこで本論文はもっと一般的なレベルで両者を競合させる一つの方法を提案する。その詳細は第5節で述べる。

5. グローバルノードと競合

グローバルノードを生成する際、注目するノード間の関係や属性を決定する重みの値としては、もしトップレベルから特別な要求がなければ、ある規準的な値が使用される。しかし、通常、トップレベルはある特定の仕事に関わっているとみるのが自然である。つまり、与えられた対象を、その仕事の範囲内で理解しようとするはずである。例えば、ディジタルの数字を読んでいるときであれば、図2の(a)を『8』と理解しようとするはずである。この時、『8』に対する記憶構造が2つの円からなっており、それを使用して図2の(c)を認識しようとするれば、図2の(c)のような構造がグローバルノードとして生成されることが望ましい。このように、先ほどの重みは、ボトムアップ的な方法だけでは決定することができず、トップレベルの影響を考慮して決めるべきである。つまり、ボトムアップ的な情報と、トップダウン的な情報の競合が起こることになる。競合に勝った重みを使用して、ある特定のグローバルノードが生成され、上位の認識部に渡され、適合度が計算される。その結果、適合度がある概念と判断するのに妥当なものであれば、与えられた入力パターンはその概念に相当すると判断され、認識が終わる。しかし、ある特定な概念と断定できない場合や該当する概念が存在しない場合は、認識が失敗に終わり、そのことを下位レベルに知らせ、新たな競合の基で、異なった重

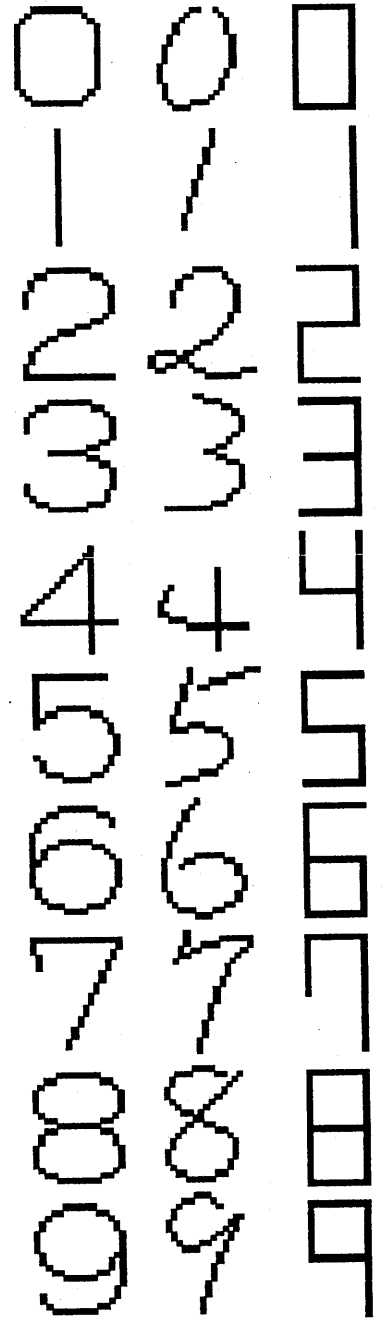


図8

みでグローバルノードをつくり直す。

6. 実行例

入力パターンとしては図8にあげたパターンを採用した。学習時には左の列のパターンしか与えていない。このことからこれらのパターンを学習パターンと呼ぶ。次に真ん中の列のパターンを手書きパターン、そして最後、右の列のパターンをデジタルパターンと呼ぶことにする。同じ行の各パターンは、同じ概念として、出力されることを期待するものである。グローバルノード生成の際に用いる重みを学習時のままで、学習パターンを認識させると（当り前の話であり、またそうでなければ困るが）、必ず成功する。次に手書きパターンを認識させる。この場合も、比較的癖のないパターンを与えたためであろうか、4を除いてすべて成功した。手書きパターンの4は、上部がつながっていないために学習パターンと同じものに認識されなかった。またこの場合の適合率は、みな一様に低かった。他の重みを用いてグローバルノードを作り替えてみようと試みたが、グローバルノード及びそれらの間の関係ともに変化しなかった（用意した重みの組は5組である）。

次に、デジタルパターンでグローバルノードを生成してみた。その結果、3と4と8が、このままでは明らかに認識が成功しないようなグローバルノードを生成した（実際成功しなかった）。まず3であるが、学習時と同じ重みでは、図9.(a)のように分割した。これでは、3と認識されないばかりでなく

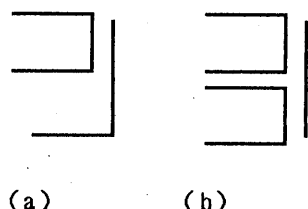
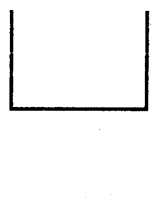


図9

他のどの概念とも一致しない。そこで重みを変えてみたら、(b)のようになった。この場合(b)のほうが勝ち上位に渡され、3と認識されて終わる。そして、デジタルパターン4に対しては、図10のようなグローバルノードを生成してしまい、いくら重みを変えてもこの形は変わらず、新しい概念と判定された。最後にデジタルパターン8に対しては、3とほぼ同じ結果が得られた（はじめは図11.(a)のようになり、つぎに(b)のようになったため、(b)が勝ち、8と認識された）。



7. 考察

手書きパターンの4や、デジタルパターンの4のようなものは学習パターン4とは別のものという結果が得られた。もし、学習時に学習パターンの4と手書きパターンの4をいっしょに学習すれば、4という概念は2つのうちどちらかに適合すれば、4であるというようになるであろう。また、こういった

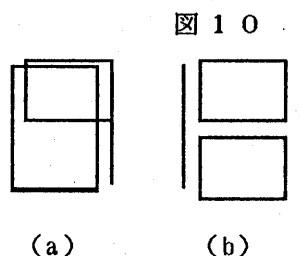


図11

知識構造を人間も持っていると考えるのは、全く当り前のことである。

単一パターンによる学習も4などの特殊な場合を除いて、十分有用である。また、競合させることにより、デジタルパターンの3や8のようなものまで認識可能となった。また、重みを変えらるゝとはこの場合、見方を変えることに他ならない。人間もいろいろな角度から物事を見、そして考えていることを思えば、この方法が意外に自然なものであるといえるのではないであろうか。

8. あとがき

トップダウン処理を考慮しなければならないということを考えだしてから、どういう方法で、そしてどこまで情報及び影響を与えれば十分かという問題がでてきた。われわれはこの問題に取り組んで日が浅く、本論文で提案した方法が十分であるとは思っていない。実際、線分ノードを連結する段階までトップダウンの情報を与えた方がいいのではないかと、もっと他にうまく競合が考えられないかなど、まだまだ検討・研究の余地を残している。

参考文献

- 1) S.Grossberg, Competitive Learning: From Interactive Activation to Adaptive Resonance, Cognitive Science, Vol.11, pp.23-63 (1987).
- 2) Fukushima,K., Miyake,S., and Ito,T. : Neocognitron: A Neural Network Model for a Mechanizm of Visual Pattern Recognition, IEEE Transactions on Systems, Man, and Cybernetics, Vol.SMC-13, No.5, pp.826-834 (1983).
- 3) Hull,J.J. : Hypothesis Generation in a Computational Model for Visual Word Recognition, IEEE Expert, Fall, pp.63-70 (1986).
- 4) 香田, 高木, ヌ木 : ニューラルネットによる手書き英数字認識—モデル規模及び学習データ数に関する評価検討—, 電子情報通信学会技術研究報告, Vol.88, No.177, pp.71-77 (1988).
- 5) Kurtzberg,J.M. : Feature Analysis for Symbol Recognition by Elastic Matching, IBM J. Res. Develop., Vol.31, No.1, pp.91-95 (1987).
- 6) 森, 横澤, 梅田 : P D Pモデルによる手書き漢字認識, 電子情報通信学会技術研究報告, Vol.87, No.428, pp.407-414 (1987).
- 7) 坂井, 麻田 : 手書き漢字の認識における分類—手書き漢字認識の手法とシステム—, 計測と制御, Vol.24, No.11, pp.1007-1012 (1985).
- 8) 塩野 : 多重辞書類似度法による手書き漢字識別の基礎実験, 情報処理学会論文誌, Vol.27, No.9, pp.853-859 (1986).
- 9) 鈴木, 有本 : 自己組織化学習機械の提案と相互干渉のある英字認識への応用, 情報処理学会論文誌, Vol.29, No.5, pp.472-479 (1988).
- 10) 津久井, 平井 : ニューロ・パターンマッチングによる文字認識, 電子情報

通信学会技術研究報告, Vol.88, No.177, pp.63-70 (1988).

- 11) Wang, q.r. and Suen, C.Y. : Large Tree Classifier with Heuristic Search and Global Training, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.PAMI-9, No.1, pp.91-102 (1987).
- 12) 山田, 上, 溝口, 天満 : ニューラルネットを用いた文字認識, 電子情報通信学会技術研究報告, Vol.88, No.177, pp.79-86 (1988).