生体の視覚系モデルによる文字パターン認識

正 員 梅 野 正 義 (名古屋工業大)

非会員 朱 小 燕 (名古屋工業大)

准 員 中 村 雅 司 (名古屋工業大)

1. はじめに

生体の視覚系は、一種の情報処理システムと考えたとき、電子計算機を初めとする現在の情報処理機械にはないような優れた機能を多数備えていることがわかる。例えば、手書き文字の読取り装置などはその良い例である。現在のところ、手書き文字を自由に読み取ることのできる文字読取り装置はまだ実現していないが、文字を読むということは我々人間にとってはそれほど困難もなく行える仕事である。神経系の持つこのような情報処理能力がいったいどのようなメカニズムに起因しているかを知ることは生理学の研究者だけでなく工学者にとっても大きな関心事となっている。

人間の脳には100億個以上のニューロンがあり、これらが複雑に影響し合って反応しているのであるから、その中の1個のニューロンの反応だけを観測して神経系全体の構造や機能を推定するのは決して容易ではない^{(1)~(3)}。そこで、神経系の働きを単純化あるいは抽象化したモデルを組み立ててみて、そのモデルをもとに神経系の働きを調べようとする試みがなされている⁽⁸⁾。モデルとしては、神経系の反応状態を数式的に記述した数学モデル、電子計算機でモデルの反応をシミュレートする電子計算機モデル、電子回路（ハードウェア）で組み立てた電子回路モデル、あるいはある種の化学反応が神経回路の反応に類似しているのに注目した化学モデルなど種々のモデルが考えられている。モデルを組み立てるときには通常、神経系の持っているすべての機能をそのまま模倣するのではない。神経系の示す特定の機能に注目し、その機能に対して

本質的な働きをしていると考えられる性質だけを選び出し、抽象化してモデルの機能として取り入れることになる。使用目的によって必ずしもすべての点で神経系に忠実なモデルが最も優れているとは言い切れないし、むしろ枝葉末節に至るような現象を取り除いて本質的な機能を浮かび上がらせるのがモデルの重要な役割の一つであると考えらるべきであろう。従ってモデルの導入は神経系の働きを単に解明し、理解するのに役立つだけでなく、工学者が神経系の優れた性質を情報処理装置の設計に取り入れようとする場合にもモデルの果たす役割は少なくない。

本研究の目的は、生体の視覚系の機能をシミュレートし、OCRのような文字認識装置とは異なる認識アルゴリズムについて提案することである。具体的には網膜における視細胞分布の違いによる中心視野と周辺視野の機能差から起こる眼球運動⁽⁴⁾を文字認識の分野に応用していくことである。

2. モデルの原理、構成

〈2・1〉 原 理 ヒトにあるパターンを示したとき、その注視点、つまりごく短時間眼球が停留して画像のある1部分に焦点が合う点は1点に固定されているわけではない。絶えずある間隔をおいてパターンを走査し、周辺視野で大きな情報を取り込み、中心視野によって細かい情報を取り込むことによって必要な情報を能率的に取り込んでいる。注視点が止まる場所は提示時間の長短に関係なく、主として白と黒との境界部、その中でも特に角の部分に集中している。

本モデルは特別な受容野を持つ細胞で入力パターンの特徴点——端点、交差点などを検出する。一番注目される特徴点から注視点を動かして各々の特徴点の特徴を判断しながらパターンを認識する。

この特徴抽出で注視点を検出するモデルのシステム図を図1に示す。注視点を検出する条件としては、網

Pattern Recognition by Biomedical Vision Model. By Masayoshi Umeno, Member, Xiaoyan Zhu, Non-member, & Masashi Nakamura, Associate (Department of Electrical and Computer Engineering, Nagoya Institute of Technology).

梅野正義：正 員，名古屋工業大学電気情報工学科
朱 小燕：非会員，名古屋工業大学電気情報工学科
中村雅司：准 員，名古屋工業大学電気情報工学科

図 1 注視点検出モデル

Fig. 1. Illustration of feature point detection.

図 2 受容野の概念

Fig. 2. Illustration of receptive field.

膜モデルを構成するニューロンモデルの信号値の中で最大値を出す成分の位置へ網膜モデルの中心を移動させるものとする。ただし、最大値をとる成分が複数存在するときは、入力パターンの情報を偏りなく検出するための注視点の位置から最も遠い成分を次の注視点とするのである。

〈2・2〉 受容野 人間の網膜は、5 段の構成層から構成されており、入力層にあたる視細胞層と出力層にあたる神経節細胞層の成分数の比は、約 150 対 1 で大脳視覚領へ情報が伝わるまでにある程度の情報処理が行われていることがわかる。特に、1 個の神経節細胞に影響を与える視細胞層の領域を受容野という (図 2)。受容野はいろいろあるが、本モデルは中心領域に光刺激を与えたとき正の信号を出す on 中心 off 周辺受容野と、逆に負の信号を出す off 中心 on 周辺型受容野を使い、受容野の荷重 $W(r)$ を以下の式で与える。

$$W_c(r) = K_c \cdot \exp(-\{r/r_c\}^2) \dots\dots\dots (1)$$

$$W_s(r) = K_s \cdot \exp(-\{r/r_s\}^2) \dots\dots\dots (2)$$

$$W(r) = W_c(r) + W_s(r) \dots\dots\dots (3)$$

ここで、 $K_c(>0)$ は正の中心荷重 W_c を決めるゲインで、 r_c は中心荷重の標準偏差である。同様に K_s は負の周辺荷重 W_s を決めるゲインで r_s は周辺荷重の標準偏差である。以上のように中心荷重 W_c と周辺荷重 W_s [(1) 式、(2) 式] を使って中心からの距離 r を決めることで、on 中心型受容野 $W(r)$ [(3) 式] を作った。

〈2・3〉 網膜モデル 網膜モデルは図 3 で示す工学的なニューロンモデルを 2 次元に配列して構成した。モデルの出力は (4) 式で表せる。

$$V = \sum X_i W_i - \theta \dots\dots\dots (4)$$

θ はしきい値である。 i は各ニューロンモデルの受容野を構成する成分である。 W_i はニューロンモデルの受容野荷重 (on 中心型と off 中心型の 2 種類ある) で受容野を構成する成分の座標に対応している。また、 X_i を入力パターン (0 と 1 の 2 値画像) の発火・不発火 (入力パターンの成分が 0 のとき不発火、1 のとき発

図 3 ニューロンモデルと受容野

Fig. 3. Neural model.

図 4 受容野のふるまい

Fig. 4. Responses of the neural models with different receptive fields.

火)を示す信号として、信号値 V を計算した⁽⁶⁾⁻⁽⁷⁾。

この on 中心型受容野と、off 中心型受容野を持つニューロンモデルは図4のような振舞いをする。(a)は棒状のパターンを on 中心型と off 中心型の受容野が矢印の方向に移動したときの信号値の方向を示している。(a)のようなパターンの場合、on 中心型受容野が、棒状パターンの端部で強い正の信号を出している。(b)は棒状パターンが交差している領域を on 中心型と off 中心型受容野が矢印の方向に移動したときの信号値の変化を示している。(b)のようなパターンの場合 off 中心型受容野が、棒状のパターンが交差している付近で強い正の信号を出している。

つまり、この網膜モデルは、さきに述べた2種類の受容野を持つニューロンモデルを2次元に配列した処理面で、入力されたパターンの端部や交差している部分を検出する特徴抽出を行っている⁽⁶⁾。

3. シミュレーション

さきに構成したモデルで、コンピュータシミュレーションを行った。モデルのパラメータは表1、表2に示されている。表中の r_c は on 反応の領域の広がり、 r_s は off 反応の領域の広がりを表している。図5のフローチャートのように、入力画面の網膜モデル上に対応している面(網膜画像)を作成する。この面は注視点が移動するごとに作成し直す。

次に網膜モデルを構成する成分(細胞)の種類によって(on 形受容野を持つか off 形受容野を持つかによって)特徴抽出し信号値 V が正になる入力画面上での成分座標を記録しておき、注視点を検出するときのデータとする。ただし、最大をとる成分が複数あるときは、前の注視点から最も遠い成分を次の注視点とし、その成分(注視点)の近傍(9×9の領域)内で縦(上)右上がり、横(右)、右下がり、縦(下)、左下が、横(左)、左上がりの中心視野モデルを構成する成分からみて、8方向の線分情報を検出した。検出の方法は方位性のある荷重を決めて処理を行っている。面積が大きすぎると計算量が多くなり小さければ検出が難しい。いろいろな面積の荷重で試した後、11×11の荷重にした。

ところが中心視野モデルは11×11の大きさがあるから、これを構成している成分について8方向の線分情報を検出すると、やはり情報にばらつきがある。線分によっては1つだけの線分情報を持つもの、つまり端点といった部分があれば、2つあるいは3つ以上の線分情報を持つものもある。多数の情報が検出されたときには最も多くの線分情報を持っている成分の情報

Table 1. The parameters of the computer simulation.

網膜モデルの大きさ	パターンの大きさ	ニューロンモデル (on, off types)					
		しきい値 θ		ゲイン K		r_c	r_s
		on	off	on	off	on	off
99×99	49×49	2.4	1.0	4.0	-4.0	0.5	-0.5

Table 2. The paramaters (W_i) of each receptive fields.

on-type				
	-0.17	-0.18	-0.17	
-0.17	-0.19	-0.06	-0.19	-0.17
-0.18	-0.06	3.08	-0.06	-0.18
-0.17	-0.19	-0.06	-0.19	-0.17
	-0.17	-0.18	-0.17	
off-type				
	0.17	0.18	0.17	
0.17	0.19	0.06	0.19	0.17
0.18	0.06	-3.08	0.06	0.18
0.17	0.19	0.06	0.19	0.17
	0.17	0.18	0.17

図5 シミュレーションフローチャート

Fig. 5. Flowchart of the program.

を検出し入力パターンの特徴とした。また複数の成分が同じ情報数で異なる情報を検出した場合にはそれらすべてを入力パターンの情報とした。

入力パターンを判断するため注視点の移動によりパターンの様々の特徴を検出する。検出された特徴は既

存のデータと照合して選択される可能性のあるパターンだけを認識の対象として選ぶ。途中もし選択可能なパターンが1つになったら認識を終了する。選択可能なパターンが1つに絞られないときには、数回の注視点移動を繰り返し選ばれた回数が一番多いパターンを入力パターンと判断する。アルファベット文字の場合 11 個の注視点があれば認識に十分なデータが得られる。またこれ以上は注視点の移動を繰り返しても新たな線分情報は得られないため、シミュレーションに於いては注視点の移動を 11 回にした。

入力の状態によっては文字の判別が難しい場合がある。例えば「C」のような入力パターンの場合、我々人間にとっても「C」なのか「O」なのか判断が付きにくい。それを考えて出力としては文字パターンの可能性を百分率で表示することにした。計算方法は次の式による。

$$X = C_i / C \dots\dots\dots (5)$$

X : 入力パターンが X である可能性, C_i : パターン X 内の既検出線分含有回数, C : 注視点移動回数 (11 回)

実際の結果はアルファベット 26 文字の基本パターンでシミュレートしたところすべての文字について認識が可能であった。

4. 結 び

本論文は、特徴抽出モデルを使ってパターンを正しく認識することを示した。認識のためのアルゴリズムには、特徴モデルのアルゴリズムを採用した。文字の部分部分の特徴だけを捉えて判断するのであるから、漢字など複雑な文字の認識を行ったり、柔軟な認識能

力を得るためには、まだまだ不十分であるが、周辺視野域の情報とした物体の大きな情報であるとか注視点の移動方向を情報、つまり特徴点間の位置関係のような別の情報を利用することで、かなり改善できるのではないと思われる。

本モデルの特長としては、特徴点で文字を判断するのでパターンの大きさは認識には全く影響しない。また回転しなければ位置ずれでも構わない。ところが認識のためのアルゴリズムは既存の 8 方向のデータを使っているので、あまり変形されたパターンは認識できない。そのためもっと高い柔軟性を持つ、しかも学習能力を得るためのアルゴリズムとして特徴点から出る線の本数で特徴を判断するモデルを今後の課題として考えている。また人工網膜のような 3 次元デバイスに適用できるように、特徴点を順次検出するよりもパターンのすべての特徴を一緒に検出し判断する並列処理システムへの応用も期待できる。

(昭和 62 年 12 月 24 日受付, 同 63 年 5 月 18 日再受付)

文 献

- (1) 永野:「視覚における周辺視野の一情報処理」, 信学論, J68-A, 46-53 (1985)
- (2) 福田:「運動知覚における中心視野と周辺視野の機能差」, テレビ学会誌, 33, 6, 39-44 (1979)
- (3) 稲波:「視野制限下における図形・文字の認識」, テレビ学会誌, 39, 5, 49-54 (1985)
- (4) 渡辺:「注視点と眼球運動」, 応用物理, 40, 3, 330-334 (1971)
- (5) 梅野・清水・野間・太田:「人工網膜のアプローチ」, 名工大学報, 35, 191-196 (1983)
- (6) 梅野・野間・清水:「人工網膜のシミュレーション」, 名工大学報, 36, 271-277 (1984)
- (7) 梅野・枇杷木:「文字パターン認識における自己組織化を利用した特徴量の検出」, 電子通信学会論文誌, J69D, 10, 1443-1449 (1986)
- (8) 福島: 視覚の生理とバイオニクス, 社団法人電子通信学会編