

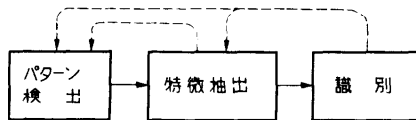
最近のパターン認識研究*

森 健一** 葛城 純夫**

1. はじめに

パターン認識の研究は1950年中頃から始まり、1960年代に入って急速に論文数が増加してきており現在でも最も活発な研究分野の一つである^{1,2,3)}。パターン認識研究の応用としては、文字読取装置、音声認識装置を始めとして指紋、声紋、顔や物体の輪廓線などの線図形、X線写真、航空写真などの平面パターン、3次元図形を認識する装置と非常に広範囲のものが考えられている。一方パターン認識の研究は研究者の数だけ方法論があるといわれるほど、研究者の考え方が色濃く反映される研究分野である。本文では、これらの数多くの研究の中から、最も有力と思われる2つの方法について、認識方法、具体的な成果の例について解説を試みたい。

一般にパターン認識の問題は図1に示すような3つのブロックに分けて考えると理解しやすい。第1のブ



第1図 パターン認識のブロック図

ロックのは対象としているパターンを見つけ出すことであり、文字読取であれば視野にあらわれた文字行の中から、1字分の文字像の位置を検出し光学像を電気信号に変換して装置に送り込むことである。パターン認識の応用例によって、パターンの抽出方法は異なるのが普通であるが、パターンの内容を知らせずに機械に自動的にパターン抽出を行わせることは考えるほどに容易なことではない。人間の場合には点線に示すように識別、あるいは特徴抽出の結果からのフィードバックがあり、パターンの内容を知ることと、パターンの抽出とが同時に行なわれていると考えられている。

現在までのところ図1のようなフィードバックを含ん

だ形式のパターン認識の研究については目立った成果は報告されていない。

第2のブロックは検出されたパターンから、そのパターンを特徴づける情報を抽出する過程である。簡単な文字読取装置などでは、パターンの検出と特徴抽出とが同時に行なわれることがある。多くの場合、パターン認識における実質的な中心問題はこの特徴抽出の方法にあり、入力のパターンに種々の変動(雑音信号)が加わっても影響をうけにくく、かつ異なるパターンを十分に区別できる特徴を抽出できる方法が見つかるならば、識別の過程は形式的なものにすぎなくなる。

これまではパターンの特徴抽出は対象とするパターンの性質によって異なるという理由で、すでに得られている特徴の情報をいかに取捨選択し、組み合わせるかどうかという第3のブロックについての研究が理論的な面から多くなされてきた。しかしながら、第1、第2のブロックの性質を無視しての理論は所詮根無し草で、論文の数ほどには具体的な成果は得られていない。パターンの性質に依存した特徴抽出方法、識別理論が最近になってようやく一つの流れにまとまりつつある。

2. 構造分析法によるパターン認識

パターン認識の方法論として有力なものの一つは、パターンの性質を分析する(Syntax Analysis^{4,7)}またはLinguistic Approach⁹⁾と呼ばれている方法である。この方法では、3段階に分けて特徴抽出と識別が行われる。まずパターンはいくつかの基本的な構成要素から成立していると考え、これに適当なシンボルを対応させる。シンボル化された基本要素は、対象とするパターンによってパターン全体、あるいは部分に付随した種々の変数のある量を意味することもありうる。次の段階で、得られたシンボル間の相互関係の構造を調べパターンを表現する記述(statement)を見出す。このパターンの構造をあらわした記述は、もとのパターンをシンボルによって表現したものであるから、逆に何らかの記述が与えられれば、この記述に属

* Recent Studies in Pattern Recognition, by Kenichi Mori, Sumio Katsuragi (Toshiba Research and Development Center)

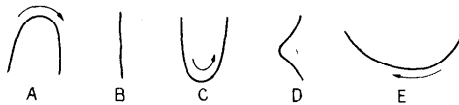
** 東芝総合研究所

する任意のパターンを発生することができる。この理由からパターンの記述はパターンの発生文法(Generating Grammar)とも呼ばれている。

パターンがいくつかの類に分けられるときには、各類型ごとに1個、必要であれば複数個の記述が用意される。そして、新たに入力パターンが与えられたときには、同様に基本要素に分解した上で、あらかじめ用意された記述のいずれと最もよくシンボルの構造が合致するかを調べ、識別が行なわれる。

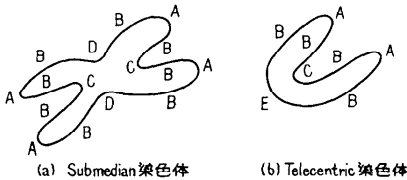
パターンの基本的な構成要素としては、文字の場合には文字線の断片(線素)の向きであるとか^{10,17)}、文字像の白黒の境界を追跡したときに得られる特定の形⁵⁾が考えられており、音声であれば周波数成分に分け、ある時間における出力がどの周波数帯に出現したか¹³⁾を基本要素に対応づけることが行なわれている。漢字の字画¹⁶⁾や、指紋の渦や流れの特異点¹¹⁾、染色体の特徴的な部分⁸⁾、泡箱写真の軌跡の部分図形など、応用例によって異なる基本要素が考えられるが¹⁵⁾、シンボル化された基本要素間の構造を問題にする点ではいずれも同一の考え方にもとづいている。

染色体の形、種類を認識する Ledley らの研究⁸⁾について構造分析法を紹介しよう。染色体の認識ではまずその外周を調べ、外周の曲率半径の変化から図2の



A(時計まわりのカーブ) B(逆時計まわりのカーブ)
C(直線) D(へこみ)
E(時計まわりの大きなカーブ)

第2図 染色体図形の基本的な構成要素



(a) Submedian染色体 (b) Telecentric染色体

第3図 染色体図形のシンボルづけ

ような5種類の特徴的な基本要素A~Eを抽出する。たとえば図3のような典型的な2種類の染色体はそれぞれ

BCBABDBABCABDBA
BCBABEBA

とシンボル列であらわされる。染色体の構造はいくつ

かの枝状の突起によって特徴づけられるが、その1つの枝 (arm) を記述するには、Aという基本要素の前後にいくつかの基本要素Bがついたものであると定義することができる。これを定義式として

$$\langle \text{arm} \rangle ::= B \langle \text{arm} \rangle | \langle \text{arm} \rangle B | A$$

と記述する。〈 〉 は5つの基本要素の適当な組合せによってつくられる部分図形、::=は定義式を、| はOR記号を意味している。このような記述式で2種類の染色体 Submedian Chromosome と Telecentric Chromosome を表現すると次のようになる。

$$\langle \text{arm} \rangle ::= B \langle \text{arm} \rangle | \langle \text{arm} \rangle B | A$$

$$\langle \text{side} \rangle ::= B \langle \text{side} \rangle | \langle \text{side} \rangle | B | B | D$$

$$\langle \text{bottom} \rangle ::= B \langle \text{bottom} \rangle | \langle \text{bottom} \rangle B | E$$

$$\langle \text{right part} \rangle ::= C \langle \text{arm} \rangle$$

$$\langle \text{left part} \rangle ::= \langle \text{arm} \rangle C$$

$$\langle \text{arm pair} \rangle ::$$

$$= \langle \text{side} \rangle \langle \text{arm pair} \rangle |$$

$$\langle \text{arm pair} \rangle | \langle \text{side} \rangle | \langle \text{arm} \rangle \langle \text{right part} \rangle$$

$$| \langle \text{left part} \rangle \langle \text{arm} \rangle$$

$$\langle \text{submedian chromosome} \rangle ::$$

$$= \langle \text{arm pair} \rangle \langle \text{arm pair} \rangle$$

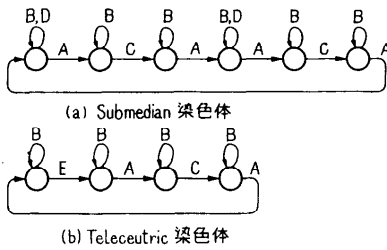
$$\langle \text{telecentric chromosome} \rangle ::$$

$$= \langle \text{bottom} \rangle \langle \text{arm pair} \rangle$$

任意の染色体から得られたシンボル列について、いずれの記述と合致するか調べて、2種類に分類することができる。

上の例ではシンボル間の順序関係をあらわすためのシンボルとして、OR記号|とAND記号(通常の表現法によって省略されている)とで充分であったが、一般的には演算記号は2以上あってもよく、漢字などでは5種の記号が考えられている¹⁶⁾。また記述の表現の仕方として、AND、ORの2種の記号を用いたシンボル列の記述には、デシジョン・グラフ¹⁸⁾を利用することもできる。記述式とデシジョングラフの関係は丁度オートマトン理論における正規表現と状態図との関係に相当するものといえよう。2種の染色体についてデシジョン・グラフで表現すると図4になる。デシジョン・グラフは、人間がパターンの構造を理解するのに適しており、手書き文字認識装置の設計に用いられて著しい効果があった¹⁷⁾。

構造分析法に用いられる記述は、リスト構造をなしているものが多く、言語の構造分析、グラフィック・ディスプレイにおけるファイルの処理¹²⁾などと多くの共通な問題を含んでいる。言語との間のアナロジーを



第4図 デンジョン・グラフで表現した染色図形の記述

図式化すると次のようになる⁹⁾。

パターン：特徴要素 \leftrightarrow 構造記述 \leftrightarrow 図形

言語：単語 \leftrightarrow 文法 \leftrightarrow 文章

構造分析法によるパターン認識を行なう装置は、現在の電子計算機に用いられている道具立てと同一のもので充分である点は実用化の面からは大きな利点である。ハードウェアを設計する場合には、パターンの記述をいかに簡潔なものにして記憶容量を少なくするか、記述の検索時間を短縮するにはどのような検索方式を採用するかに工夫がなされている¹⁰⁾。またパターンの構造を示す記述を、機械によって自動的に作成しようとする試みも行なわれている¹⁴⁾。

3. 位相空間を用いたパターン認識

構造分析法によるパターン認識は多次元のパターンを一度基本要素に分解し、再びパターンの構造を考慮しつつ再構成する過程で識別に必要な情報を得る方法であるが、その処理は常に1次元論理の連鎖によって行なわれている。パターン認識のもう一つの有力な方法は、パターンの構造を論理的な記述で表現するのではなく、多次元の関数によって表現する方法である。このためにパターンの集合全体は有限個の標本だけから必要な精度で推定できるようなある種の連続性が保存された位相空間を形成していると考えられる。パターンの表現形式を任意に座標変換してかえることができるためにアナログ量を基本に用いる。

文字パターンを領域 R の上で定義される \mathbf{x} の関数で表わされるものとして、入力パターンを $f(\mathbf{x})$ 、標準パターンを $f_0(\mathbf{x})$ とすると、実際に出現するパターン f は f_0 に雑音パターン $g_N(\mathbf{x})$ が加わって

$$f(\mathbf{x}) = f_0(\mathbf{x}) + g_N(\mathbf{x}) \quad (1)$$

となる。従来の多くのパターン認識の理論においては雑音パターン g_N は、 f_0 の関数に関係なくガウス型の雑音などが仮定され、その分散量を推定する問題な

どが取扱われてきた。この場合 \mathbf{x} の空間は全て一様な距離空間であると仮定され、パターンの類間を区切る線型判別関数を標本のデータから、計算あるいは漸近的な「学習法」によって求めることが中心問題であった。しかしながら g_N についての仮定は、実際のパターン認識の問題では充分満されている保障はなくよい成果が得られなかった。

最近になって、パターンの集合の空間を位相空間と考える理論が提案された^{20, 22, 23, 24)}。これらの理論においては g_N は f_0 に依存し、 f_0 の近傍のみ直交する関数によって f を展開することになる。飯島によって展開された一連の理論¹⁹⁾と、この理論にもとづいて試作された高性能文字認識装置 ASPET/70²¹⁾ に従って、位相空間を用いたパターン認識について概略を紹介しよう。

標準パターン $f_0(\mathbf{x})$ に対し、 $N-1$ 種類の微小変形を受けたパターンを考え

$$f(\mathbf{x}) = f_0(\mathbf{x}) + \sum_{n=1}^{N-1} \alpha_n g_n(\mathbf{x}) \quad (2)$$

なる形で表わす。いま $f_0(\mathbf{x}), g_1(\mathbf{x}), \dots, g_{N-1}(\mathbf{x})$ の N 種類のパターンから正規直交性を満す N 種類の関数 $\varphi_m (m=1, \dots, N)$ を求める。

$$\varphi_m(\mathbf{x}) = \sum_{n=0}^{N-1} \beta_{mn} g_n(\mathbf{x}) \quad (m=1, 2, \dots, N) \quad (3)$$

$$(\varphi_m, \varphi_{m'}) = \delta_{mm'} \quad (m, m'=1, 2, \dots, N) \quad (4)$$

ただし、 $g_0(\mathbf{x}) \equiv f_0(\mathbf{x})$

そうすると $f(\mathbf{x})$ は

$$f(\mathbf{x}) = \sum_{m=1}^N C_m \varphi_m(\mathbf{x}) \quad (m=1, 2, \dots, N) \quad (5)$$

ただし、 $C_m = (f, \varphi_m)$

の形に展開することができ、 $\{C_m\}$ はパラメータ $\alpha_1, \alpha_2, \dots, \alpha_{N-1}$ によって種々変化するが、 f_0 の近傍で変化する任意のパターン $f(\mathbf{x})$ に対しては常に

$$\|f\|^2 = \sum_{m=1}^N C_m^2 \quad (7)$$

なる関係が満足される。そこで任意の入力パターンを $h(\mathbf{x})$ としたとき、標準パターン $f_0(\mathbf{x})$ との間の類似度を

$$S[h, f_0] = \frac{1}{\|h\|} \sqrt{\sum_{m=1}^N (h, \varphi_m)^2} \quad (8)$$

と定義すれば、この値は0と1の間の実数値をとることになる。もし入力パターン $h(\mathbf{x})$ が(2)式の関係で標準パターン f_0 に属するパターンであれば(8)式の類似度は1になる。もし区別すべき標準パターンの種

類が K 種あるならば、それぞれの標準パターンの近傍で(2)~(6)式に従って展開を行ない $\{\phi_n\}$ を求、標準パターンの位相空間における構造を決定する。

任意の入力パターンが与えられたときは、各標準パターンごとに用意された展開関数との内積を求め、(8)式に従って標準パターンとの類似度を計算し、最も類似度の大きな標準パターンを答えとして識別する。この方法に従えば各標準パターンに個々の $N-1$ 種の雑音信号に対して極めて強い安定性を示すパターン認識を行なうことが可能となる。

一般にパターンの構造を反映した関数展開では、非線型関数になり回路化することが困難であるとされてきた。特にアナログ量を取扱う場合にはデジタル量の場合に較べて精度、速度、安定性、製造のしやすさなどの点に関して問題があった。しかしながら最近の集積回路技術はアナログ回路素子にも及び、演算増幅器など基本素子が非常に安価に生産され、精度、安定性、小型化などの面ではほとんど問題がなくなっている。将来アナログ量のメモリー、転送素子などが開発されるならば、デジタル素子と区別なく情報処理装置に用いることができるようになるであろう。

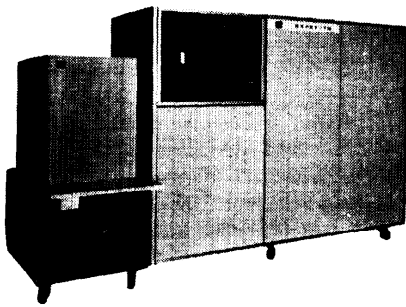
前記の理論を文字認識に応用して、英数字および記号の約 50 字を対象とした ASPET/70 (図 5) が開発された²⁾。このシステムは図 6 に示すような構成をし

ている。フライング・スポット走査装置により文字図形を 0.1 mm ごとのメッシュで走査し、白黒 2 値に量子化した上で 4 方向にシフトできる東西南北平面レジスタに貯えると同時に、文字位置と文字線の太さの正規化を行なう。ここまでデジタル回路で構成されているのは、パターンのメモリーおよびシフトがアナログ素子で現在得られないためである。標準化回路では 2 次元の量子化像から空間回路によるボケ演算と正規化操作により、デジタル量をアナログ量に変換するとともに、標本点の数を大幅に減らしている。識別回路では約 50 種の文字毎に入力パターンを(6)式に従って内積演算をした上で、(8)式によって 2 乗加算し平方根を求めるアナログの演算回路が用意されている。その結果得られる類似度 S_n の内、最も大きな値と次に大きな値を検出し、最大値が次大値より充分大きければ最大値をとる標準パターンを答えとし、最大値と次大値が接近している場合には認識拒否を答えとする。

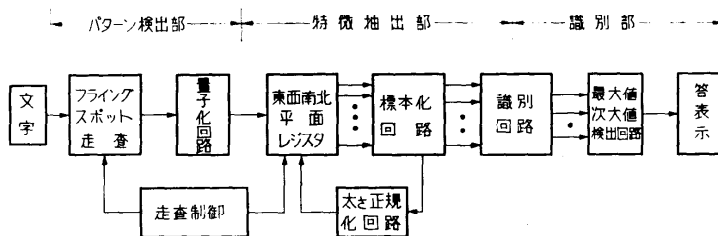
ASPET/70 が従来の OCR と性能の上で著しく異なる点は、文字の濃度が濃すぎてつぶれた図形になっても (図 7b) 正しく認識できる点である。この能力はパターン全体の構造を(5)式の関数展開が保存していることによって始めて可能となったもので、この位相空間を考慮した情報処理回路(空間回路網と呼ばれている)は、これからのパターン認識研究において種々の応用面を開拓してゆくであろう。

4. あとがき

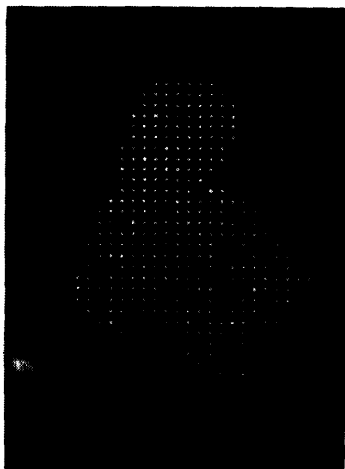
パターン認識研究は多くの研究者によって種々のアプローチの仕方で行われてきた。その中からパターンの構造を追求する 2 つの方法(構造分析法と空間回路網法)について紹介を行なった。前者は手書き文字読取、音声認識、指紋鑑定など構造を保存しつつ、大きな変形をうけるパターンを認識するのに力を発揮



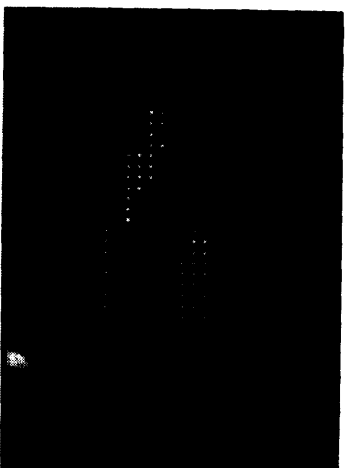
第 5 図 ASPET/70 の外観



第 6 図 ASPET/70 の構成図



第7図(a) つぶれた文字の例(数字4)



第7図(b) かすれた文字の例(数字4)

し、後者は活字読取、写真の自動解析など構造が一定しており、その上にパターン構造に依存した雑音が重畳するようなパターンを取扱うのに適している。いずれもパターンの構造を不変量として抽出している点では共通である。(図7参照)

今後パターンの構造を抽出する方法についての詳細な研究や、図1のパターン検出の研究および、特徴抽出、識別過程からのパターン検出へのフィードバックを含む総合的なパターン認識研究が必要であろう。

参考文献

1) G. Nagy, "State of the art in pattern recognition," Proc. IEEE, 56, pp. 836~862, May.

- 1968.
- 2) Y. C. Ho and A. K. Agrawala, "On pattern classification algorithms-Introduction and survey," Proc. IEEE, 56, pp. 2101~2114, Dec. 1968.
- 3) M. D. Levine, "Feature extraction; A survey," Proc. IEEE, 57, pp. 1391~1407, Aug. 1969.
- 4) R. Narasimhan, "Syntax-directed interpretation of classes of pictures," Comm. ACM, 9, pp. 166~173, March. 1966.
- 5) H. Freeman, "On the digital computer classification of geometric line patterns," 1962 Proc. National Electronics Conf. 18, pp. 312~324.
- 6) R. A. Kirsch, "Computer interpretation of English text and picture patterns", IEEE Trans. EC-13, pp. 363~376, Aug. 1964.
- 7) R. S. Ledley, "Thousand-gate computer simulation of a billion-gate computer," Computer and Information Sciences, J. T. Tou. and R. H. Wilcox, Eds. Washington D. C. Spartan Books, 1964, pp. 457~480.
- 8) R. S. Ledley, L. S. Rotolo, J. B. Wilson, M. Belson, T. J. Golab and J. Jacobsen, "Pattern recognition studies in the biomedical sciences," 1966, Proc. SJCC, pp. 411~430, 1966.
- 9) J. Feder, "Alinguristic approach to pattern analysis" Dept. of Elect. Eng., New York Univ. Tech. Rept. 400-133, Feb, 1966.
- 10) P. Mermelstein and M. Eden, "Experiments on computer recognition of connected handwritten words," Information and Control, 7, pp. 255~270, 1964.
- 11) W. J. Hankley and J. T. Tou, "Automatic fingerprint interpretation and classification via contextual analysis and topological coding," Pictorial Pattern Processing G. C. Cheng et al Eds. Washington D. C. Thompson Books, 1968, pp. 411~456.
- 12) I. E. Sutherland, "Sketchpad: A man-machine graphical communication system," Lincoln Laboratory, M. I. T. Tech. Rept. 296, Jan. 1963.
- 13) 樽松, 武田, 井上, "文法規則を用いたパターン認識についての考察" 電子通信会オートマトン研究会資料, A 69-70, 1970年1月.
- 14) 坂本, 堂下, 富田, 山口, "句構造サンプルパターンを識別するオートマトンの構造について", 電子通信学会オートマトン研究会資料, A 69-76, 1970年3月.
- 15) 坂井, 長尾, "電子計算機を用いた交通網地図の記憶と処理", 電気通信学会オートマトンと自動制御研究会資料, 1966年7月.
- 16) 藤村 靖, "文字認識に関する2,3の考察", 情

- 報と制御の研究, 昭和38年度, 第4号, pp. 2~7.
- 17) H. Genchi, K. Mori, S. Watanabe, and S. Katsuragi, "Recognition of handwritten numerical characters for automatic letter sorting," Proc. IEEE, **56**, pp. 1292~1300, Aug. 1968.
 - 18) S. Katsuragi, H. Genchi, K. Mori and S. Watanabe, "Recognition of handwritten numerals using a decision theory" Proc. International Joint Conference on Artificial Intelligence D. E. Walker. L. M. Norton Eds. Washington D. C. 1969.
 - 19) 飯島, "複合類似度法による文字識別理論", 昭和45年, 電気四学連合会予稿集, 2710, 1970.
 - 20) 飯島, "パターン認識", 日刊工業新聞社, 1969.
 - 21) 飯島, 森(俊), 野口, 森(健), 葛城, "文字読取装置 ASPET/70 の識別回路", 昭和45年電気四学連合大会予稿集, 2714, 1970.
 - 22) G. S. Sebestyen, "Decision-making processes in pattern recognition" The Macmillan Comp. New York, 1962.
 - 23) 上坂, "非線形判別関数の構成理論", 電子通信学会論文誌C, 51巻7号, pp. 295~302, 1968.
 - 24) H. Reitboeck et al. "A transformation with invariance under cyclic permutation for application in pattern recognition," Information and Control, **15**, pp. 130~154, Aug. 1969.
(昭和45年8月11日受付)