文字読取装置 ASPET/71

東京工業大学 飯 島 泰 蔵

文字読取装置は、電子計算機などへの情報入力機器として使用されるもので、最近とみに関心を集めつつあるものである。このたび従来の技術を飛躍的に凌駕する高性能文字読取装置 ASPET/71 が、わが国において研究開発された。本文はこの装置に採用された技術について、理論的側面から見た解説を行なったものである。

1. ま え が き

パターン認識についての研究が本格的に展開されるようになってから、はや十数年が経過してしまった。この研究は、当初考えられていたほど単純なものでなく、その難解さは、今日では定説となり始めているほどである。

文字読取装置は、電子計算機への情報入力機器のひとつとして考えられたものであり、当初は計算機で用いられている論理回路の応用技術によって、比較的安易に実用化が果せるものと思われていた。したがって、この装置の登場は、キーパンチ作業を介して行なわれてきた従来の情報入力手段にとって代わる機器の出現であるとして、大きな期待をもって迎えられたものである。

ところが文字読取装置は、パターン認識問題が持つ困難性を本質的にかかえているため、ある限度を越えて認識率の向上をはかることが、実は容易なことではなかったのである。一方、この装置が計算機の入力機器として実用に供されるためには、誤読が許されないという厳しい条件を満足しなければならなかったから、装置の認識性能が不十分であることから生ずる欠陥を回避するために、他の適当な手段を講ずる必要があった。文字フォントの規格統一や、用紙、印字品質等についての厳格な規制が始められるようになったのは、このような事情からである。

近年、電子計算機の発達はいよいよ著しく、超高速化・超大形化が進展するようになった。このため計算機システムにおいて取扱われる情報の量は激増の一途をたどり、キーパンチ作業による従来の情報入力手段によっては、システム内部での情報処理能力に追従して行けなくなる、という場面が、漸次表面化するようになってきた。このような事態は、単に省力化のためというよりも、システム全体の効率的運用をはかるために、入力機器の在

り方を革新しなければならない時期が迫りつつある、ということを示すものである。

情報技術の発達に支えられて進展を続けて行く情報化社会の中にあって、計算機システムの果す役割りは、ますます増大の一途をたどると推測される。そうすれば、人間から機械への情報伝達機能をつかさどる文字読取装置の果すべき任務は、いよいよ大きく広がって行くと考えねばならない。こうした前提に立って考えるならば、すでに人間社会の中で、広く使用されている文字に対して、過度の規制を加える方向は、決して望ましいものではない、ということが判明するであろう。

文字読取装置に対する大きな期待があったにもかかわらず、従来その使用分野が限られてきた裏には、上述のような理由がひそんでいたということができる。

昭和 41 年春、通産省では“超高性能電子計算機システムの研究開発”を目指す大形プロジェクトを開始した。電子技術総合研究所（当時の電気試験所）では、その一環として OCR（光学的文字読取装置）の研究開発を計画し、東京芝浦電気株式会社との間の緊密なる共同開発作業を展開することにより、昭和 45 年 2 月“ASPET/70”と称する認識装置試作機を発表するに至り、さらにこれを基として、昭和 46 年 11 月“ASPET/71”と銘打ったページ式 OCR を完成、発表した*。

前後 6 年間にわたった ASPET/71 の研究開発作業は、従来の OCR 技術が直面していた混迷の壁を打ち破り、印字品質に対する規制をほとんど必要としない新しい世代の OCR 技術を確立することに、その主眼が置かれた。ASPET/71 の技術が生み出された裏には、その根底を築くために十数年来、著者によって展開されてきた図形認識基礎理論^{1)~7)}の完成があったことを、記しておかねば

* ASPET の名称は、Analog Spatial Processor developed by Electrotechnical Laboratory and Toshiba の頭文字をとって命名されたものである。著者は、ASPET/71 の完成まで電総研に在籍していたが、1972 年 4 月、東工大へ移籍した。

“On the Optical Character Reader ASPET/71” by Taizo Iijima (Tokyo Institute of Technology, Tokyo)

ならない。実際には、これから導びかれた実用基礎理論が構成されるとともに、それに基づいた設計理論が確立された。また、この理論を受けた空間回路網技術が新たに開発され、従来の面目を一新する OCR 技術の確立に結びつくこととなった次第である。

本解説においては、ASPET/71に用いられている認識技術の本質に関する理論的背景を紹介することに重点を置き、ご参考にご供することとした。

2. 理論展開における基本的な立場

パターン認識問題は、一般に

(i) 与えられた個々のパターンからそれが担っている本質的な情報をいかにしてとらえるか、という問題と、

(ii) 認識の目的に応じていかに合理的な決定を行なうか、

という問題とに大別される。ところが、従来パターン認識に関する理論的研究では、もっぱら後者の問題に重点が置かれてきたのである。これは後者のほうが一般論の展開が容易である、という事情が手伝っていたものと思われる。後者の問題は、本来入力情報の中にあいまいさが含まれていたとき、統計的ないしは決定論的な意味で目的に応じた最も好ましい決定を行なうには、いかにすればよいか、という方法を追求する問題である。しかし、いかに巧妙な方法が確立されたとしても、入力情報を含むあいまいさがある限り、ある限界を超越した高い認識率に到達することは、原理的に不可能であることが知られている。

文字読取りの問題は、パターン認識問題における具体例のひとつであるから、当然上記の両者の問題を含んでいると考えられる。しかしながら、実際上からの要請を考えると、誤読を絶無にしなければならない、という厳しい条件に応えなければならないことが判明する。これは後者の問題に対する追求のみによっては解決できない問題であるといわねばならない。

従来、認識装置への入力情報にあいまいさが混入されていた理由を尋ねてみると、外界に提示されたパターンが十分な情報を保持しているにもかかわらず、これを観測し、その結果に前処理をほどこし、それから特徴を抽出してくる過程において、何らかの無理な処理が関与することによって、あいまいさが生じている場合が、きわめて多かった。これは明らかに前者に属する問題であるといわねばならない。したがって、文字読取りの問題を解決するには、前者の問題と正面から取り組む必要があったのである。

前者の問題が取扱い難い理由のひとつは、それが対象

とするパターンの性格に強く依存する問題である、という点である。そこで文字パターンは図形の一種であることに注目して、まず図形一般の持つ基本的性質を解明するとともに、図形の内部構成を明らかにする基礎理論の展開をはかった。次いでその基盤の上に立って、文字認識のための具体的な処理方法の確立を行なった。ここに展開された認識理論が従来のものと比較して際立って異なっている点は“理論が文字図形の持つ基本的性格を踏まえたものである”という点ばかりでなく“観測から決定に至る一連の認識処理過程が、一貫した理論に基づいて克明にとらえられている”という点に要約される。

以上が、ASPET/71における認識理論を展開するに当たって、とられたところの基本的な立場である。

3. 図形の基礎方程式

われわれは、外界にある図形をそのまま手に取るように眺めることができる。しかし、無意識のうちに行なわれていることであるが、光を用いた図形の観測操作は、実はそれほど単純な事柄ではない。外界に散在している図形情報の空間分布の模様を、眼の網膜上に再現するための光学的な図形伝送の方法は、本来随意であるはずの2次元図形を、ある限定された特性を持つ3次元図形像に変換してしまうことになるのである。以下にその事柄を解説しよう。

図形観測の基礎理論によれば、2次元平面 S_∞ 上で定義される任意の“図形”は、 $\mathbf{r} \in S_\infty$ を変数とする関数 $f(\mathbf{r})$ で表現され、これを観測することによって得られる図形像（これを“一般化図形”または広義の図形という） $f(\mathbf{r}, T(\tau))$ は

$$f(\mathbf{r}, T(\tau)) = \int_{S_\infty} G(\mathbf{r} - \mathbf{r}', T(\tau)) f(\mathbf{r}') d\mathbf{r}' \quad (1)$$

なる形式で表わされることになる。ここで $G(\mathbf{r}, T(\tau))$ はガウス型の関数で、

$$G(\mathbf{r}, T(\tau)) = \frac{1}{4\pi\sqrt{|T(\tau)|}} \exp\{-\mathbf{r} \cdot T(\tau)^{-1}\mathbf{r}\} \quad (2)$$

である。ただし $T(\tau)$ は2次元対称行列で、 $|T(\tau)|$ は τ の単調増加関数である。

(1) 式は、図形 $f(\mathbf{r})$ がパラメーター τ に応じた“ぼけ操作”を受けて $f(\mathbf{r}, T(\tau))$ に変換されることを表わしており、眼球内に作られる $f(\mathbf{r})$ の光学像は、一般に $f(\mathbf{r}, T(\tau))$ で表わされると考えられる。実際に網膜上に結像する図形像は、 τ をある値に選んだときの $f(\mathbf{r}, T(\tau))$ である。

正常な眼では、注目する図形は $\tau=0$ でとらえられるから、 $f(\mathbf{r}, T(0)) = f(\mathbf{r})$ となって、鮮明な図形像を得

るが、近視や遠視の眼では $\tau \neq 0$ であるため、図形像が外界の図形よりも、ぼけた状態でとらえられることになる。しかし正常な眼の場合でも、注目していない場所での図形の像は、一般に $\tau \neq 0$ なる状態で観測していると考えられる。

図形がぼけることは、図形を認識するうえで好ましくない、と考えられがちであるが、実は(1)式に示すような形式に基づいてぼけ操作が随意に行なわれるものであるということは、図形の本質に根ざしている事柄なのである。そしてこのぼけは、後述するように、認識を効率化するうえで、重要な働きをなすこととなるのである。

(1) 式を微分形式に書きかえると、

$$\nabla \cdot \left\{ \frac{d}{d\tau} T(\tau) \right\} \nabla f(\mathbf{r}, T(\tau)) = \frac{\partial}{\partial \tau} f(\mathbf{r}, T(\tau)) \quad (3)$$

なる放物形の2階偏微分方程式が導かれる。これは“図形の基礎方程式”と呼ばれるものであって、図形に対して本来許されている変換は、この基礎方程式の形式が不変に保たれるものに限られる、と考えることができる。実際の図形観測において、外界の図形から網膜上の図形像への変換は、よく知られているように、ぼけ操作ばかりでなく、“位置移動”、“伸縮”、“回転”、“傾斜”、“強度変化”、“一様な背景強度の無視”等を含むものである。ところがこのようなすべての変換は、いずれも(3)式の形式をまったく不変に保つ変換である、ということが厳密に証明されるのである。図形の変形を行なうに際して、上記の変換に基づくような変形をほどこすときに限って、その程度がかなり激しい場合でさえも、図形の本質がまったく損なわれないように見えるのは、このような理由からであると解釈される。

外界の図形と網膜上に写し出される図形像とは、一見したところでは、同じように見えるであろう。しかし、認識処理の本来の対象が外界の図形でなければならないにもかかわらず、実際上の処理における直接の対象が網膜上にある図形像である、という事実は、認識処理の在り方を本質的に限定することとなるのである。図形認識に関する基礎理論は、この単純な事実のみに立脚して築かれたものであった。この理論の立脚点となった(3)式が、図形基礎方程式と呼ばれることとなったのは、この理由からである。

4. 標準図形

2次元平面上に描かれた図形から情報を得ようとする場合、通常図形上のある点に視点を定め、それを中心とするある平面領域を視野の範囲として限定する操作が加えられるものである。この事柄を理論的に記述すれば次

図1 一般化図形における視点と視野の範囲

のようになる(図1)。

まず、パラメーター τ の中から $0 < \tau_1 < \tau_0$ なる2点を選んで、 $T_0 = T(\tau_0)$ 、 $T_1 = T(\tau_1)$ とおけば、(3)式の積分分解から、

$$f(\mathbf{a}, T_0) = \int_{S_{\infty}} G(\mathbf{a} - \mathbf{r}, T_0 - T_1) f(\mathbf{r}, T_1) d\mathbf{r} \quad (4)$$

なる関係式が求められる。 $f(\mathbf{a}, T_0)$ は、本来与えられた図形 $f(\mathbf{r})$ を T_0 だけばかすことによって得られる一般化図形の $\mathbf{r} = \mathbf{a}$ なる点での値を示すものであるが、(1)式の形式からわかるように、この値は、図形 $f(\mathbf{r})$ 上の $\mathbf{r} = \mathbf{a}$ なる点に視点を置き、その点を中心とする T_0 なる視野の範囲を眺めたときの、図形全体から得られる光エネルギーの総量を表わすものである、と解釈することもできる。ところが(4)式は、この値が $f(\mathbf{r})$ を T_1 だけばかした図形 $f(\mathbf{r}, T_1)$ に対して、視点を同じ $\mathbf{a}$ に置き、視野の範囲を $T_0 - T_1$ で置きかえたときの総エネルギーにも一致する、ということを示している。(4)式の右辺の値が、 T_1 の値に無関係に一定である、ということは重要である。

そこで任意に与えられる図形は、これをすべて $f(\mathbf{r}, T_1)$ なる図形に変換して考えるものとし、これを基準として考察を進めて行くことにする。任意に与えられる2つの図形間の“内積”を、

$$(f, f') = \int_{S_{\infty}} G(\mathbf{a} - \mathbf{r}, T_0 - T_1) f(\mathbf{r}, T_1) f'(\mathbf{r}, T_1) d\mathbf{r} \quad (5)$$

とし、各図形の“ノルム”を、

$$\|f\| = \sqrt{(f, f)} \quad (6)$$

で定義するものとする。この定義から明らかなように、元の図形が一定であっても、 $\mathbf{a}$ 、 T_0 、 T_1 の選び方によって上記の値は種々変化することになる。このことは、内積やノルムが、図形から情報を獲得しようとする側の主観的な図形の評価量を定量的に規定するものである、

ということを示している。 α や T_0 は、図形上のどの部分をいかなる範囲にわたって注視するかに関係しており、 $T_0 - T_1$ は、図形上の注視領域内からどの程度詳細な情報を得ようとしているかに、関係している。

いま α , T_0 , T_1 をある値に定めたとき、 α , T_0 の微小変動に対してそのノルムが極大値になっているような図形を想定して、これを“標準図形”と呼ぶことにしよう。本論文では、標準図形を特に $f_0(\mathbf{r}, T_1)$ なる記号で表わすものとする。そうすれば、

$$J = \|f_0\|^2 \quad (7)$$

なる汎関数の値は、定義によって極値をとることになるから、

$$\nabla_{\alpha} J = 0, \quad \frac{\partial}{\partial \tau_0} J = 0 \quad (8)$$

なる関係が成り立つことになる。そこで位置ベクトル $\mathbf{r}$ を直角座標成分に分けて、

$$\mathbf{r} = ix + jy \quad (9)$$

とおき、標準図形 $f_0(\mathbf{r}, T_1)$ およびその逐次導関数を、

$$\begin{cases} f_0 = f_0(\mathbf{r}, T_1), & f_1 = \frac{\partial}{\partial x} f_0, & f_2 = \frac{\partial}{\partial y} f_0 \\ f_{11} = \frac{\partial^2}{\partial x^2} f_0, & f_{12} = \frac{\partial^2}{\partial x \partial y} f_0, & f_{22} = \frac{\partial^2}{\partial y^2} f_0 \end{cases} \quad (10)$$

と略記するものとする。そうすれば (8) 式の要請から、

$$\begin{cases} (f_0, f_1) = 0, & (f_0, f_2) = 0, & (f_0, f_{11}) = -\|f_1\|^2 \\ (f_0, f_{12}) = -(f_1, f_2), & (f_0, f_{22}) = -\|f_2\|^2 \end{cases} \quad (11)$$

なる関係が導びかれる。この関係式は、標準図形 $f_0(\mathbf{r}, T_1)$ が持っている基本的な性質を表わしているものである。

5. 図形の正準化

図形 $f(\mathbf{r})$ は、本来光エネルギーの密度分布を表わす関数であるから、非負である。したがって (1), (2) 式から明らかなように、 τ のいかにかわらず、

$$f(\mathbf{r}, T(\tau)) \geq 0 \quad (12)$$

である。いま任意に与えられる一般化図形 $f(\mathbf{r}, T(\tau))$ に対して、これを、

$$g(\mathbf{r}, T(\tau)) = f(\mathbf{r}, T(\tau)) - f(\alpha, T_0) \quad (13)$$

なる関数に変換するものとする。そうすれば $g(\mathbf{r}, T(\tau))$ もまた (3) 式を満足するから、これも広義の意味で図形の一種であると考えられる。上記の変換をほどこすことを、図形の“正準化”という。

正準化された図形 $g(\mathbf{r}, T(\tau))$ はもはや非負ではないが、つねに、

$$(g, 1) = 0 \quad (14)$$

なる関係を満足している。正準化は単純な線形演算に過

ぎないが、後で示すようにこの演算は、実は本質的な寄与をなすこととなる。

なお、標準図形に対して得られた関係式 (11) は、正準化された標準図形に対しても、そのままの形式が成り立つということを注意しておくこととする。

6. 類似度によるパターンマッチング法

以下においては、すべて正準化された図形について議論を進める、ということをおこわておこう。

本章では、類似度を用いたパターンマッチング法の原理について、基礎理論的な考察を行なう。

1つの標準図形 $g_0(\mathbf{r}, T_1)$ と任意の図形 $g(\mathbf{r}, T_1)$ に関して、実数パラメーター λ を用いて次式を作るものとすれば、

$$\begin{aligned} \|\lambda g_0(\mathbf{r}, T_1) - g(\mathbf{r}, T_1)\|^2 &= \{\lambda \|g_0\| - \|g\|\}^2 \\ &+ 2\lambda \{\|g_0\| \|g\| - (g_0, g)\} \end{aligned} \quad (15)$$

なる関係式を得る。ここでもし、

$$g(\mathbf{r}, T_1) \equiv \lambda g_0(\mathbf{r}, T_1) \quad (16)$$

なる関係が成り立つと仮定すれば、

$$\|\lambda g_0 - g\| = 0, \quad \|g\| = \lambda \|g_0\|$$

となるから、(15)式によって、

$$\|g_0\| \|g\| = (g_0, g) \quad (17)$$

でなければならないこととなる。逆にもし (17) 式が成り立つと仮定すれば、 λ の値を、

$$\lambda = \|g\| / \|g_0\| \quad (18)$$

なる実数値に選ぶことによって、(15)式から、

$$\|\lambda g_0 - g\| = 0 \quad (19)$$

なる結果を得る。これはほとんどすべての点で (16) 式が成り立たねばならない、ということの意味している。

図形は、全体を一様に定数倍してもその本質が損なわれることはないから、(16) 式の関係は両図形 g , g_0 がまったく同等である、ということの意味している。故に両図形間の類似度を、

$$s(g_0, g) \equiv (g_0, g) / \|g_0\| \|g\| \quad (20)$$

で定義すれば、この値は本来、

$$-1 \leq s(g_0, g) \leq 1 \quad (21)$$

なる範囲の実数値をとるものであるが、特にその値が、 ± 1 となるか否かによって、 g が g_0 と同等な図形であるか否かを判定できることとなる。これが“類似度による”ところのパターンマッチング法の原理”である。

実際のパターンマッチング法では、両図形間の“相違度”を、

$$d(g_0, g) = \sqrt{1 - s^2(g_0, g)} \quad (22)$$

で定義したとき、この値がある閾値 θ よりも小さいならば、両図形は同等であるとみなす立場がとられている。そこで図形 $g(\mathbf{r}, T_1)$ を、

$$g(\mathbf{r}, T_1) = \|g\| \left\{ s(g_0, g) \frac{g_0(\mathbf{r}, T_1)}{\|g_0\|} + h(\mathbf{r}, T_1) \right\} \quad (23)$$

なる形式で表わすものとしよう。そうすれば、

$$(g_0, h) = 0, \quad \|h\| = d(g_0, g) \quad (24)$$

となるから、 $h(\mathbf{r}, T_1)$ は、 $g(\mathbf{r}, T_1)$ をそのノルムで正規化したものの $g_0(\mathbf{r}, T_1)$ に直交する図形成分を表わしていることになる。

任意に与えられた図形 $g(\mathbf{r}, T_1)$ を (23) 式を用いて $h(\mathbf{r}, T_1)$ に変換して考察するならば、 $h(\mathbf{r}, T_1)$ なる図形が超空間（図形空間）中の原点を中心とする半径 θ なる超球内にあるか否かによって、 g が g_0 と同等であるか否かを判定しようというのが、上記の方法の意味である。

7. 標準図形の微小変位

図形 $g(\mathbf{r}, T_1)$ が、特に標準図形 $g_0(\mathbf{r}, T_1)$ の位置が δ だけ変位した図形である場合を考察しよう。そうすれば、

$$g(\mathbf{r}, T_1) = g_0(\mathbf{r} + \delta, T_1) \quad (\delta = i\delta_1 + j\delta_2) \quad (25)$$

となる。いま変位量 δ の絶対値が充分小さいと仮定すれば、近似的に、

$$g(\mathbf{r}, T_1) = g_0 + g_1\delta_1 + g_2\delta_2 + \frac{1}{2}(g_{11}\delta_1^2 + 2g_{12}\delta_1\delta_2 + g_{22}\delta_2^2) \quad (26)$$

なる形で表わされる。ここで標準図形について定まる次の量：

$$\Delta_1 = \|g_0\|/\|g_1\|, \quad \Delta_2 = \|g_0\|/\|g_2\|, \quad \vartheta = (g_1, g_2)/\|g_1\|\|g_2\| \quad (27)$$

を定義すれば、

$$(g_0, g) = \|g_0\|^2 \sqrt{1 - \left\{ \left(\frac{\delta_1}{\Delta_1} \right)^2 + 2\vartheta \left(\frac{\delta_1}{\Delta_1} \right) \left(\frac{\delta_2}{\Delta_2} \right) + \left(\frac{\delta_2}{\Delta_2} \right)^2 \right\}} \quad (28)$$

$$\|g\| = \|g_0\| \quad (29)$$

なる結果を得る。したがって $g(\mathbf{r}, T_1)$ と $g_0(\mathbf{r}, T_1)$ との間の類似度は、

$$s(g_0, g) = \sqrt{1 - \left\{ \left(\frac{\delta_1}{\Delta_1} \right)^2 + 2\vartheta \left(\frac{\delta_1}{\Delta_1} \right) \left(\frac{\delta_2}{\Delta_2} \right) + \left(\frac{\delta_2}{\Delta_2} \right)^2 \right\}} \quad (30)$$

で与えられることとなる。

8. 標準図形のほけの微小変化

図形 $g(\mathbf{r}, T_1)$ が、特に標準図形 $g_0(\mathbf{r}, T_1)$ の T_1 が微小変化した図形である場合を考察しよう。そうすれば、

$$g(\mathbf{r}, T_1) = g_0(\mathbf{r}, T(\tau_1 + \varepsilon)) \quad (31)$$

であるから、 ε が充分小さいと仮定することによって近似的に、

$$g(\mathbf{r}, T_1) = g_0 + \varepsilon(\alpha g_{11} + 2\gamma g_{12} + \beta g_{22}) \quad (32)$$

となる。ここで α, β, γ は、

$$\frac{d}{d\tau_1} T(\tau_1) = \begin{pmatrix} \alpha & \gamma \\ \gamma & \beta \end{pmatrix} \quad (33)$$

で定まる定数である。そうすれば、

$$(g_0, g) = \|g_0\|^2 \left\{ 1 - \varepsilon \left(\frac{\alpha}{\Delta_1^2} + 2\vartheta \frac{\gamma}{\Delta_1 \Delta_2} + \frac{\beta}{\Delta_2^2} \right) \right\} \quad (34)$$

$$\|g\| = \|g_0\| \left\{ 1 - \varepsilon \left(\frac{\alpha}{\Delta_1^2} + 2\vartheta \frac{\gamma}{\Delta_1 \Delta_2} + \frac{\beta}{\Delta_2^2} \right) \right\} \quad (35)$$

なる結果を得る。したがって $g(\mathbf{r}, T_1)$ と $g_0(\mathbf{r}, T_1)$ との間の類似度を求めれば、

$$s(g_0, g) = 1 - O(\varepsilon^2) \quad (36)$$

なる結果が得られることとなる。

9. ほけ操作の積極的導入

入力図形 $g(\mathbf{r}, T_1)$ があらかじめ用意された既知図形 $g_0(\mathbf{r}, T_1)$ に類似した図形であるかどうかを見るために、類似度 $s(g_0, g)$ を調べるものとしよう。この際、入力図形 $g(\mathbf{r}, T_1)$ の位置を $g_0(\mathbf{r}, T_1)$ の位置に合わせることを、および T_1 なるほけの程度を合わせることを、あらかじめ前処理として行なっておく必要がある。ところが、この操作を精度よく行なうことが、実は実際上容易でないことが知られている。したがって、前処理が充分でない場合に算出される類似度は、一般には位置ずれ量 δ やほけの変化量 ε に比例した誤差を含むことになる、と考えられる。ところが $g_0(\mathbf{r}, T_1)$ が特に標準図形に選ばれているならば、(30)式や (36) 式に示されているように、これらの量の 2 乗に比例した誤差しか含まないことになるので、類似度の精度がこのことによって著しく高められる結果となる。

ところで (30) 式や (34), (35) 式をより詳細に眺めてみると、 δ や ε の係数の値は、 Δ_1 や Δ_2 の値に逆比例する関係になっている。しかるに、ほけの基準値 τ_1 を増大させて τ_0 に近づければ、

$$\lim_{\tau_1 \rightarrow \tau_0} \frac{\|g_1\|}{\|g_0\|} = \lim_{\tau_1 \rightarrow \tau_0} \frac{1}{\Delta_1} = 0, \quad \lim_{\tau_1 \rightarrow \tau_0} \frac{\|g_2\|}{\|g_0\|} = \lim_{\tau_1 \rightarrow \tau_0} \frac{1}{\Delta_2} = 0 \quad (37)$$

となることが証明されるから、類似度の精度をよりよく向上させるには、ほけの程度 τ_1 は高いほど有利である、という結論に到達する。

$g(\mathbf{r}, T_1)$ が $g_0(\mathbf{r}, T_1)$ の類似図形である場合、ほけを増大させて $\tau_1 \rightarrow \tau_0$ とすれば、類似度はその極限において $s(g_0, g) \rightarrow 1$ となる。しかし、この事実を逆に眺めれば、 $g(\mathbf{r}, T_1)$ が本来 $g_0(\mathbf{r}, T_1)$ とは異なる図形であった場合にも、 $s(g_0, g) \rightarrow 1$ となるのではないか、という疑問を生ずるであろう。実際、正準化を行なわない前の図形 $f(\mathbf{r}, T_1)$, $f_0(\mathbf{r}, T_1)$ に対しては、常にこの極限で

図 2 ぼけ操作による図形ベクトルの変化

$s(f_0, f) \rightarrow 1$ となることが証明されるのである。したがってこの疑問は、確かに的を射たもっともなものである、と考えてよい。ところが正準化操作の行なわれた図形間においては、実はこの関係が必ずしも成り立たないのである。この事実を理論的に示すには、図形の構造を詳細に論じなければならないので省略するが、以下に図形空間でのモデル的考察を通じて、そのことを解説しておくこととしよう。

図 2 は、図形空間中での図形ベクトル f_0, f を示したものであるが、これらの図形を正準化することによって得られる図形ベクトル g_0, g は、定数ベクトル OC に直交する超平面内にある。いま図形に対してぼけ操作を加えて行くと、図形ベクトル f_0, f の頂点は、それぞれ C 点に向かって移動して行くから、両ベクトル間の交角 φ は減少して 0 となり、 $s(f_0, f) = \cos \varphi \rightarrow 1$ なる結果を得る。ところが超平面内への射影成分である図形ベクトル g_0, g は、このときその絶対値を縮小させて行くことになるが、両ベクトル間の交角 θ はあまり変化を受けることなく θ_0 に収束する。したがって $s(g_0, g) = \cos \theta \rightarrow \cos \theta_0$ となって、類似度による評価量は、両図形ベクトル間の本質的な差異を、そのまま温存することとなるのである。

従来、ぼけ操作は、図形間の差異を消滅させてしまうという理由で、好ましくないと考えられてきた。ところが図形に対してあらかじめ正準化操作を行なっておきさえすれば、この心配は杞憂に過ぎない、ということが明らかになったのである。そうとすれば、その他の面においてぼけ操作のもつ有利な特性を積極的に利用する、という道が開かれることになる。

本来類似な関係にある図形間の類似度を見るために、ぼけ操作が有利であることはすでに述べたとおりであるが、ぼけ操作は、図形に含まれている高次空間高調波成分を抑制する効果がある、ということができる。この事柄は図形を標本化する際に、“ぼけの程度が高いほど少な

い標本点ですますことができる”という結果を生む。これは図形処理を行なう装置を電子回路によって構成しようとする場合、標本点が少ないほど装置が簡単化することになるので、経済性という観点からもぼけ操作が有利である、ということになるのである。

ASPET/71において、ぼけ操作が積極的に取り入れられた理由は、以上に述べたようなものであり、またこの新技術の重要な、特色のひとつとなっているものでもある。

10. 複合類似度法

既知図形を標準図形に選ぶとともに、正準化に関する前処理をほどこした入力図形に対して類似度に基づくパターンマッチング法を採用するならば、ぼけの積極的な導入をはかることによって、充分高い認識性能が得られることになろう、というのが、これまで述べてきた理論の概要である。

ところが、このような配慮にもかかわらず、実際の認識の場面における検討を進めてみると、まだ不充分であることが判明してくる。その原因をなしている事柄は、図形の位置を確定する操作が、図形の含む各種ノイズの影響にわざわざいされて、精度よくこれを行なうことが困難である、という点である。この事実、図形が実質的に含んでいるノイズ成分よりもはるかに大きな変形成分を、見かけ上図形が含んでいるという結果を生むことになり、そのことが認識性能を低下させる主要な原因を作るからである。しかしながら、この種の事柄は解析的に追求して行くことが可能であり、そのことを基にして理論的な立場から、対策を立案することができるようである、と考えられた。次に述べる複合類似度法は、このような観点から生み出されたものである。

いま標準図形 $g_0(\mathbf{r}, T_1)$ とその 1 次導関数とから、

$$\begin{cases} \varphi_0 \equiv \frac{g_0}{\|g_0\|}, \quad \varphi_1 \equiv \frac{1}{\sqrt{2(1+\vartheta)}} \left\{ \frac{g_1}{\|g_1\|} + \frac{g_2}{\|g_2\|} \right\} \\ \varphi_2 \equiv \frac{1}{\sqrt{2(1-\vartheta)}} \left\{ \frac{g_1}{\|g_1\|} - \frac{g_2}{\|g_2\|} \right\} \end{cases} \quad (38)$$

なる 3 種類の図形を定義するものとする。そうすればこれらは、

$$(\varphi_i, \varphi_j) = \delta_{ij} \quad (i, j=0, 1, 2) \quad (39)$$

なる正規直交系を作っていることが証明される。そこで入力図形 $g(\mathbf{r}, T_1)$ が、(25) 式で示されるような標準図形の変位した図形である場合を考察すれば、その近似展開式 (26) は、

$$g(\mathbf{r}, T_1) = \|g_0\| \left[\sqrt{1 - \left\{ \left(\frac{\delta_1}{A_1} \right)^2 + 2\vartheta \left(\frac{\delta_1}{A_1} \right) \left(\frac{\delta_2}{A_2} \right) + \left(\frac{\delta_2}{A_2} \right)^2 \right\}} \right] \varphi_0$$

$$\begin{aligned}
& + \sqrt{\frac{1+\vartheta}{2}} \left\{ \left(\frac{\delta_1}{\Delta_1} \right) + \left(\frac{\delta_2}{\Delta_2} \right) \right\} \varphi_1 \\
& + \sqrt{\frac{1-\vartheta}{2}} \left\{ \left(\frac{\delta_1}{\Delta_1} \right) - \left(\frac{\delta_2}{\Delta_2} \right) \right\} \varphi_2 \quad (40)
\end{aligned}$$

なる形式に書き換えられることが判明する。したがってこれから、

$$\|g\| = \|g_0\| = \sqrt{(\varphi_0, g)^2 + (\varphi_1, g)^2 + (\varphi_2, g)^2} \quad (41)$$

なる関係式を得る。

そこでいま、標準図形 $g_0(\mathbf{r}, T_1)$ に対して3種類の図形 $\varphi_0, \varphi_1, \varphi_2$ を用意するものとしよう。そうすれば任意の入力図形 $g(\mathbf{r}, T_1)$ に対して3種類の類似度 $s(\varphi_0, g)$, $s(\varphi_1, g)$, $s(\varphi_2, g)$ が算出できることになるが、ここでこれらを用いて作られる次の量：

$$s^*(g_0, g) = \sqrt{s^2(\varphi_0, g) + s^2(\varphi_1, g) + s^2(\varphi_2, g)} \quad (42)$$

を定義して、これを“複合類似度”と呼ぶことにする。そうすれば、入力図形 $g(\mathbf{r}, T_1)$ が (25) 式で表わされるような図形である限り、変位量 δ の大きさには無関係に、

$$s^*(g_0, g) = 1 \quad (43)$$

なる結果が得られることになる。

上記の結果は、近似式 (40) から導びかれたものであるから、変位量 δ が大きくなった場合には厳密でない。しかし類似度 $s(g_0, g)$ が図形の変化に対して特に敏感過ぎるという欠陥は、これを複合類似度 $s^*(g_0, g)$ で置きかえることによって、大幅に是正することが可能となるのである。

なお、ここに示した複合類似度は、図形の変化に影響され難い類似度を得るために、考えられたものであったが、その形式は同時に図形のぼけの基準 T_1 の微小変化に対しても、よりいっそう不変性を保たせるための方策になっている、ということが理論的に証明される。故に複合類似度を用いたパターンマッチング法を採用するならば、位置決めとぼけの程度を合わせるという前処理上最も困難な操作を、識別操作の段階において、比較的単純な方策によって補償できる、という結果になるのである。この巧妙な結果は、認識技術上きわめて有効な方式の確立を促すこととなった。

11. ASPET/71 の特長

ASPET/71 は、以上に述べたような理論を基盤として、設計製作されたものである(写真1)。ここに確立された認識方式の主要点は、

- 1° 正準化と称する前処理操作の導入
- 2° ぼけ操作の積極的な導入
- 3° 複合類似度法の採用

なる3点に要約される。これらの主要点は、いずれもま

写真1 ASPET/71の外観

ったく新しい独創的な考え方に基づくものであるが、こうした考えが図形の本質を究めようとして展開された基礎理論的研究の結果として生み出されたものであったという事実は、重要な意義を持っている。それはOCRの認識性能を向上させる方策が、図形本来の基本的性格に根ざす考え方によって初めて確立されることになるはずである、という著者の年来の思想が、このことによって立証されることとなったからである。これを契機としてOCRの技術が、従来のレベルを飛躍する新しい段階に突入するであろうことは、ほぼ間違いないことであると信ぜられる。

ASPET/71 は、印字を読取るページ式OCRとして設計された。したがって、文字図形の伸縮、回転、傾斜、等については、特に配慮されていない。この技術を手書き文字の場合にも適用しようとするならば、当然これらに対処する方策を加味する必要があるが、印字に限る場合はこれで充分である、ということが検証された。

実用上の立場から見た場合、ASPET/71の最も誇るべき特長は、従来ほとんど絶望的であると見られていた低品質文字が、驚くべき正確さで読取れるようになった、ということである。写真2は、低品質印字の例を示したものであって、(a)は多数枚のカーボンコピーによって印字がつぶれてしまった場合の例、(b)はタイプリボンを使い古して印字がかすれてしまった場合の例を示したものである。ASPET/71では、このいずれの場合も正確に読取れることが実証されている。

またASPET/71では、理論上示された図形処理の演算をそのまま忠実に実行させるため、アナログ量を並列的に演算する空間回路網が用いられた。この装置は紙送り機構や光電変換部の性能に押えられて、毎秒2,000字の読取速度を維持するよう設計されたが、認識部だけの処理速度は、実は毎秒20,000字程度の超高速性能であることが検証された。

ASPET/71では、認識対象となる字形から出発して、

写真 2 (a) つぶれた印字の例 (英字A)

写真 2 (b) かすれた印字の例 (英字A)

これを標準文字として整備し、これから複合類似度の演算に必要な回路諸定数を決定するまでの設計演算を理論どおりに行なった。したがってこの設計作業には、試行錯誤的な要素はまったく含まれていない。対象とする1つの字形を基にした設計演算であるにもかかわらず、この装置がその字形と類似した幅広い多数の字形に対処できる特性を得ているのは、字形の変化に対する理論的な考察が、奥深くなされているためである。このような

OCR の設計法は、従来の方法ではとうてい考えられなかったことであるが、このことによって使用者の要求に即座に応ずることのできる学習機能を持った OCR の実現をはかることが可能となるに至った。

12. む す び

本文は、最近われわれの手によって、研究開発された OCR “ASPET/71” の技術を、理論的側面から解説したものである。OCR に対する関心は、近年とみに高まりつつある現状であるが、すでに数年以上先行していた米国における OCR 技術に行きづまりが見え始めた今日の時点において、これをはるかに凌駕する高性能 OCR の新技術を開発したことに対して、世間の注目を集めることとなったということは、誠に幸運であったと考えている。本文の解説が、OCR だけにとどまらず、わが国におけるパターン認識技術の発展を望まれる読者の方々に対して、何らかの意味においてご参考になるならば、幸いである。

(昭和 47 年 8 月 31 日受付)

〔参 考 文 献〕

- 1) 飯島：図形の基礎方程式と観測変換，信学論，54-C，7 (1971) 641
 - 2) 飯島：図形空間の構成法に関する基礎理論，信学論，54-C，8 (1971) 736
 - 3) 飯島：図形の抑制核とその数学的諸特性，信学論，54-C，9 (1971) 835
 - 4) 飯島：図形の正規化に関する基礎理論，信学論，54-C，11 (1971) 1003
 - 5) 飯島：抽象図形空間の基本関数系，信学論，54-C，11 (1971) 1011
 - 6) 飯島：図形の特徴抽出に関する基礎理論，信学論，54-C，12 (1971) 1133
 - 7) 飯島：図形の構造認識に関する基礎理論，信学論，55-D，8 (1972) 499
 - 8) 飯島：パターン・マッチングによる図形識別法についての理論的考察，信学論，55-D，8 (1972) 507
- なお ASPET/70, 71 の紹介記事については、次のものが代表的である。
- 9) 飯島：新しい世代を開いた文字読取装置 ASPET/70，エレクトロニクス，15，5 (1970) 569-577
 - 10) 飯島：ASPET/70 から ASPET/71 へ，エレクトロニクス，17，2 (1972) 273-277
 - 11) 飯島泰蔵，森 健一：人間の識別能力に迫る OCR ASPET/71，日経エレクトロニクス，No. 30 (1972.5.22) 66-80