

手書き文字認識に関する二三の考察*

保原 信**

Abstract

Problems inherent in the automatic recognition of handwritten characters are discussed.

The author classifies the already attempted methods of automatic character recognition as "passive" in the sense that they utilize no knowledge other than given signals. In the present paper, the author emphasizes the significance of analysis by synthesis method by K. N. STEVENS as one of the "active" methods of handwritten character recognition, for it might give some further insights into the automatic recognition or graphological analysis of handwritten characters.

序

文字を機械で識別する方法については、従来多くの研究が報告されているが、十分な結果が得られたわけではない。とくに、手書き文字のようなより複雑な問題についても多くの文献が見られるが、いずれも多分に Heuristic な Level にとどまっている。

ここでは、少し観点を変えて、手書き文字認識の問題に含まれている本質的な点を二三論じ、とくに、手書き文字が具体的にどういう構造をもっているかということを中心として系統的に考察し、手書き文字識別に対する新しい接近の基礎を与えてみたい。この問題には後に述べるように、自動筆跡認定の問題が内包されており、この種の問題に対しても新しい定量的研究の糸口を与えるものと考えられる。

1.

手書き文字を認識するためには、まず、それがどういう形で（言語）情報を表わしているかを知らなければならぬ。そこで、手書き文字による情報伝達の過程を原理的に一般的な通信系と考えてみよう。通常の手書きによる通信では、情報源と考えられている物理的な空間図形としての文字そのものが、実はより根源的な各種の情報を人間の手書き過程という特殊な符号化過程によって符号化（物理的に実現化）した符号化

信号であり、したがって、最終的な情報の認識が、いわゆる手書き文字の再生という直接的な形ではなく、視覚の末梢における図形の検出から、大脳における最終的な情報の復号化までを含めたものである。しかも、その符号化の過程が、各人に固有な手書き過程によってのみ実現されている。

しかし、文字自体は特定の言語の使用者によって言語習慣として体得され、使用者間に共通に認められたある概念的図形であり、さらにまた、手書き文字は人間という共通の体系的構造をもった過程から生み出される信号である。

これらの意味で、手書き文字が具体的にどのように記述されているか、すなわち、どういう共通性と弁別性とで組み立てられているかをまず知る必要がある。

ところで、どの場合でも具体的な図形としての手書き文字は、つぎの条件で満たすものと考えてよい。

条件 I 一般の文字図形と同様に2次元空間におけるパターンである。

条件 II 筆順のほぼ一定した書字動作（筆点の変位 $x(t)$, $y(t)$ および筆圧 $p(t)$ ）の軌跡である。

このように、適当に定義された変域内の各位置に対して定められた2次元の空間関数として扱うことができるが、活字あるいはタイプライタによる印書とは、その生成の過程において本質的に異なっており、むしろ、同じ言語媒体である音声と似た面をもっている。すなわち、手書き文字はその生成過程では音声と同様に時間関数であるという意味では本質的には異なっていない。しかし、一度文字図形として記録されると様子

* Some remarks on the automatic recognition of handwritten characters, by Makoto Yasuhara (The University of Electro-Communications)

** 電気通信大学

が相当変わってくる*1。すなわち、音声では各人が各音素をその発声の手続きとして、動的・体験的印象としてしかとどめておくことができないが、文字などの2次元的空间図形は、いわゆる静的パターンとして、各人共通に定義することができる。しかし、空間の認知がまだ充分完成していない学習初期にある個体(新しい筆跡に出合った場合も含めて)、あるいは書かれた対象だけでなく、書字の動的特性が重要な手がかりとなる筆跡認定の立場から考えると、手書きの問題は音声のそれと全く異質のものではなく、むしろ、酷似しているものと考えるべきであろう。

2.

一般に“文字”は定まっているが、それを具体化した“文字パターン”は与えられていないという場合には、その文字体系の性質・構造がわからなくては、それによる制約を考慮することができない。そこで、個々の文字体系によらず、手書き文字一般について共通な制約が条件 I, II 以外に見い出されれば、それだけ手書き文字認識の設計方針は狭められることになる。ここに、人間の手書き過程の性質、構造を明確にしなければならぬ必然性が生まれてくる。

オランダの医者 D. van der Gon¹⁾ は人間の手書き過程の制約を、つぎの4つの仮定として要約した。

仮定 1 人間の手はある有限の質量をもった質点と等価である。したがって、手書き文字はこの質点が摩擦のある平面上を運動したときの軌跡である。

仮定 2 手の運動は互いに直交する方向に作用する独立した2つの筋肉群によって制御される。

仮定 3 Fast Cursive Handwriting の場合には、筆点の位置に関係する細かい帰還制御 (Visual Tactile, Proprioceptive な受容器からの Feedback Control) は無視できる。

仮定 4 書かれる図形は筋収縮の時間点で定まり、作用する筋力の大きさには関係しない。筋力の大きさは図形の全体的な大きさにのみ関係する。

以上のような仮定を導入することによって、手書き過程を具体的につぎのように定式化することが可能になる²⁾。

$$\left. \begin{aligned} \ddot{x} + k_x \dot{x} + l_x x &= F_x(t), \\ \ddot{y} + k_y \dot{y} + l_y y &= F_y(t), \end{aligned} \right\} \quad (1)$$

ただし、

$$k_x = (1/m) \left\{ \lambda_x + \mu_x \frac{\rho(t)}{V} \right\},$$

$$k_y = (1/m) \left\{ \lambda_y + \mu_y \frac{\rho(t)}{V} \right\},$$

$$l_x = \nu_x/m, \quad l_y = \nu_y/m,$$

$$F_x(t) = f_x(t)/m, \quad F_y(t) = f_y(t)/m,$$

ここに、

$\rho(t)$	筆圧
$V = (\dot{x}^2 + \dot{y}^2)^{1/2}$	筆速
k	正規化等価粘性率
l	筋肉の正規化弾性率
$F(t)$	正規化筋力
m	手の等価質量
λ	筋肉の粘性率
μ	筆点と平面間の摩擦係数
ν	筋肉の弾性率

このモデルの実験的検討については文献(2),(3)にゆずることとし、ここでは人間の手書き過程を1次元近似として(1)式のような簡単なモデルに置き替えられることを紹介しておくだけで充分である。

以上の結果を考慮して、手書き文字一般に共通な制約として条件 I, II の他につぎの条件を設定することができる。

条件 III 手書き文字(とくに Cursively Handwritten Character)は有限箇の画(Stroke)の連結である。

条件 IV 画は1回の筋収縮で生じる筆点の運動およびその空間的軌跡であって、つぎの筋収縮の生じる点までの部分である。したがって、 x 方向、 y 方向のそれぞれに対して独立に定義される。

3.

ここで新しく導入された画は、明らかに漢字のそれのように簡単な幾何学的線または点と異っており、あるいはまた M. Eden, Halle⁴⁾ らが試みた Cursive Handwriting における画の概念とも本質的に異なっている。いわゆる、人間の手書き運動に対する中枢からの制御指令(Motor Command)に呼応した形で定義されており、空間的形状には無関係である。以下、この考え方を実験的に検討しておこう。

*1 もちろん、音声でもテープレコーダを使うことによって、空間パターンとして直接記録することができるが、これを声帯模写することは本質的に不可能である。

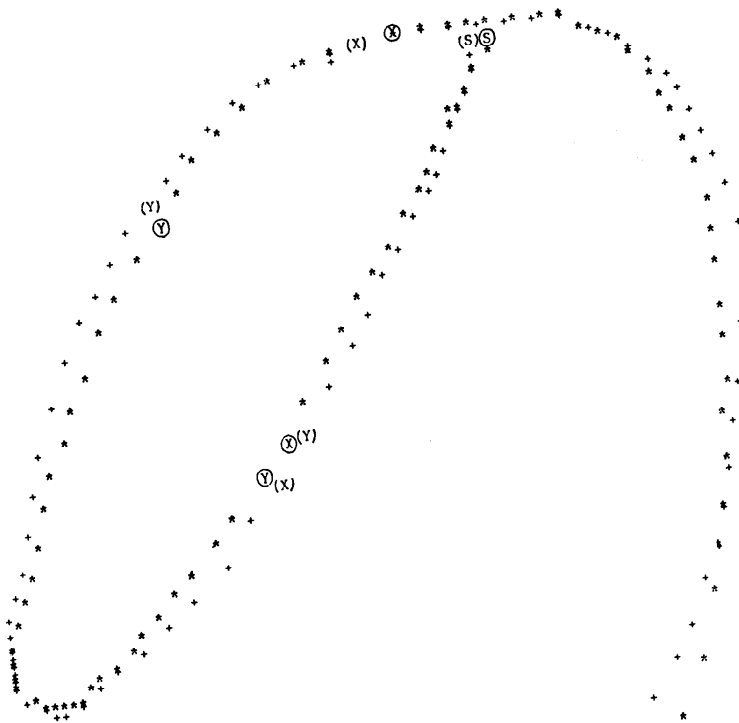


Fig. 1 Handwritten Japanese character “の” (*) and its synthesized pattern (+)

Fig. 1 は手書き分析装置*2を使って計算機に入力させた手書きのひらがな“の” (*印), および (1) 式のモデルを用いて, この原文字にできるだけ近い文字を合成したもの (+印) である. また, Fig. 2 は合成の際の力関数 $F_x^s(t)$, $F_y^s(t)$ である.*3

まず, 手書き分析装置により実際に手書きされた“の”について筆点の座標 $X(t)$, $Y(t)$ を測定し, これから速度 $\dot{X}(t)$, $\dot{Y}(t)$ および加速度 $\ddot{X}(t)$, $\ddot{Y}(t)$ を求める. つぎに

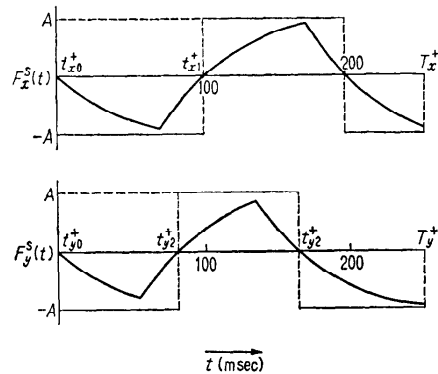


Fig. 2 The force functions which synthesize the pattern“の”

*2 手の運動を解析するために, 時間関数としての筆点の位置座標 $X(t)$, $Y(t)$ および筆圧 $F(t)$ を測定する装置*2.

*3 筆圧および筆速を一定と仮定し, 筋肉の弾性を無視して

$$\left. \begin{aligned} x + k\dot{x} &= F_x(t) \\ y + k\dot{y} &= F_y(t) \end{aligned} \right\}$$

のように (1) 式を単純化したモデルを採用し, k は各運筆者について, 別の実験から求める. 当被験者の場合 $k \approx 20(T^{-1})$ であった. また, 力関数 $F(t)$ については, 筋肉の活動が筋収縮開始の時間点と, その持続時間とで規定できるものと考え, その過渡特性を指数関数で近似した*1,3). Fig. 2 よりその様子がわかる.

$$\left. \begin{aligned} \ddot{X} + 20\dot{X} &= F_x(t) \\ \ddot{Y} + 20\dot{Y} &= F_y(t) \end{aligned} \right\} \quad (2)$$

に従って, $F_x(t)$, $F_y(t)$ を計算し, これが時間軸を切

る時点 T_D ($T_{D0}, T_{D1}, T_{D2}, \dots, D=X, Y$) を求める。

一方逆に, t_d ($t_{d0}, t_{d1}, t_{d2}, \dots, d=x, y$) を適当に与えて $F_x'(t), F_y'(t)$ を計算し

$$\left. \begin{aligned} \ddot{x} + 20\dot{x} &= F_x'(t) \\ \ddot{y} + 20\dot{y} &= F_y'(t) \end{aligned} \right\} \quad (3)$$

を解くことによって原文字 (実際の手書き文字) に最も近い文字を合成する*4。

この T_D, t_d に対応する点を Fig. 1 中に示したものが ⊗, ⊙ および (X), (Y) などの記号である。これら一連の実験から明らかかなことは

(1) (3) 式による合成モデルが実際の手書き文字のよい近似を与える。

(2) 合成モデルを逆に使って, 時間の消去された一般の手書き文字について, 新しくパラメータ τ (時間に対応するが, 時間であると考えない方がよい) を導入することによって, 最適な擬似変位 $\xi(\tau), \varphi(\tau)$ および擬似筋力 $\phi_\xi(\tau), \phi_\varphi(\tau)$ を zero 法によって計算で求めることが可能になる。ただし

$$\left. \begin{aligned} \frac{d^2\xi}{d\tau^2} + k_\xi \frac{d\xi}{d\tau} &= \phi_\xi(\tau), \\ \frac{d^2\varphi}{d\tau^2} + k_\varphi \frac{d\varphi}{d\tau} &= \phi_\varphi(\tau). \end{aligned} \right\} \quad (4)$$

なお, Fig. 1 の合成された“の”はこの考え方によっているが, τ の原点 (S点), およびその点における動きの方向はあらかじめ与えられた。

(3) 計算された $\phi_\xi(\tau), \phi_\varphi(\tau)$ を使って原文字の画抽出 (Stroke Extraction) が可能となる。Fig. 1 をみると, 実際の手書き文字を Stroke Segmentation した結果と, 合成文字に対するそれとが非常によく一致していることがわかる*5。

(4) 以上の手続きによりアナログ信号としての手書き文字を符号化することが可能となる。たとえば, Fig. 1 の“の”の場合は

$$\begin{pmatrix} 100^- & 96^+ & 54^- \\ 82^- & 82^+ & 86^- \end{pmatrix}$$

と表わされる。すなわち, 100, 96, 54 は x 方向のストローク持続時間 (あるいは筋肉収縮時間) に対応し, 82^- , 82^+ , 86^- は y 方向のそれに対応している。また肩の +, -, *などの記号は, 筋力作用の向きがそれぞれ正, 負および作用なしの状態を示している。また一方, この情報が与えられると, それに対応するパタ

ーンを一意的に合成することができることは自明であろう。

われわれはこれらの結果によって, 有効な情報を失なうことなく, 手書き文字の記述を物理的空間関数の領域から筋収縮に対応する画のレベルにおける記述へ可逆変換することができ, 同時に不必要な情報を落すことができる。

以上のように手書き文字の構造を研究するのに, 手書き文字を機械的にあるモデルから合成するための原理的法則を発見し, それを実際に適用して合成される文字図形が, われわれの概念的な手書き文字の特徴を十分に示すかどうかの判定を行ない, その結果によって用いた構造モデルの適否を判定したり, 逆に手書き文字の特徴を分析するという方法は, いわゆる Chomsky の Generative Grammar と同様であり, 音声における Stevens, Halle の Analysis by Synthesis とも似た原理をもつ。

4.

人間の音声処理系が Active であり, したがって, A-b-S 法が, 聴覚の問題を充分考慮していない点に疑問を残しているながらも, 音声分析法として有効であると考えられているが, 人間の文字識別系は一般には Active な系でないので, 手書き文字の分析に A-b-S の手法をそのまま適用することには問題がある。しかしながら, 活字やタイプライタによる印字, あるいは階書体などは問題外として, とくに, Cursively Handwritten Character の認識過程が全く Active でないと思うのには無理があるし, 筆跡認定を目的とする場合は, 必ず Active な要素が非常に強くなり, 筆速・筆圧などの動的特徴がより重要な役割をもつようになる。

手書き文字, 識別のモデルとして A-b-S 法を適用する場合, 陽に時間を考えることは間違いであり, 時間領域での比較が許されない点を注意すれば, A-b-S 法を手書き文字の分析に応用することができ, 実際有効であるという見とおしも得られた。とくに, 筆跡分析を目的とする場合には, 強力かつ重要な研究手段となる。またこの場合には, 文字の自動認識は不可欠の前提であり, この点からも可逆変換の可能な A-b-S 法の意味がはっきりとしてくる。

最後に, ここで結論としてまとめてみると

(1) 手書き文字認識のモデルとして, A-b-S 法を応用することには意味がある。

*4 実際には積分 $\int [X(t) - x(t)]^2 + [Y(t) - y(t)]^2 dt$ を最小とする

t_d を繰り返し法で求める。

*5 筋電流との対応もはっきりとしている*3。

(2) 筆跡認定を目的とする場合には、A-b-S 法は、強力かつ重要な研究手段である。

(3) 画の概念を改めて明確に定義し、A-b-S 法を適用することによって、手書き文字の画抽出を行なうことができ、人間の手書き過程の Hierarchy をさかのぼることができた。

とにかく、A-b-S 法は内部に Active な要素を含んでいる点で、従来の Passive な手法と本質的に異なっており、興味ある考え方である。

参 考 文 献

1) J. J. Denier van der Gon: A Handwriting

Simulator, Phys. Med. Biol. Vol. 6, pp. 407-414, 1962.

2) M. Yasuhara: Handwriting Analyzer and Analysis of Human Handwriting Movements, Jap. Psychol. Res. Vol. 11, No. 3, pp. 103-109, 1969.

3) 保原 信: 手書き文字の合成モデル, 情報処理 Vol. 11, No. 10, 1970.

4) M. Eden, M. Halle: The Characterization of Cursive Writing, Information Theory 4th London Symposium, pp. 287-299, 1960.

(昭和 45 年 8 月 17 日受付)