

手書き数字認識における特徴パラメータについて

青木正喜*・本山澄夫・水町守志

Abstract: It is devised and proposed here two recognition systems of handwritten numerals; these utilities are testified through computer simulation using 50×50 mosaic off-line data, obtained from handwritten numerals.

The first system is, principally based on the feature extraction of horizontal cross-section, the second on the topological information obtained through the curve-following processing. The optimization of processing parameters is sought by computer for these two systems.

1. ま え が き

パターン認識においてはパターンの観測、特徴パラメータ抽出と分類、および決定の三つが重要な点と考えられる。実際にはその中で特に特徴パラメータとして何をとりかというところが重要な問題である。これは考える対象によって全く変わってきてしまうため、場合に応じて決定せざるを得ない。ここでは扱う対象を手書き数字として特徴抽出に重点を置くこととし、パラメータ抽出に関して二つの提案と実験を行なって特徴抽出のための手段を見い出し、各種の基礎資料を得ることを目的としている。

特徴抽出に関しては、有効なパラメータを選定するための手段として計算機でのシミュレーションで行なった。このため、まず計算機入力データを得るための二次元二値化パターン読取装置を作製し、これによって得られたデータを使い、計算機の中でパターンを二値化マトリックスとして再生して各種パラメータに対するシミュレーションを試みた。

パラメータ抽出の第一の方法としては、対象を手書き数字に限ったことから簡単な方法の適用が考えられるので、横方向の投影を主体とした特徴抽出と、それを利用した認識についての実験を行なった。

第二の方法としては、字の持っている本質的な特徴と考えられるものは、字画の構造にあるとの考えに立って、数字以外にも一般に適用し得る方法として、字画の追跡による特徴抽出と認識についての実験を行なった。

2. 計算機入力データを得るための装置

手書き数字の特徴抽出に関する各種の実験を計算機により off-line で行なうに当って、計算機への入力データを得るために二値化された図形をマトリックスパターンのデータとして得るための装置を作製した。この装置の出力は、さん孔された紙テープであり、紙テープのデータにより計算機の中にマトリックスパターンとして手書きの数字が再生される。

* 東京大学大学院学生

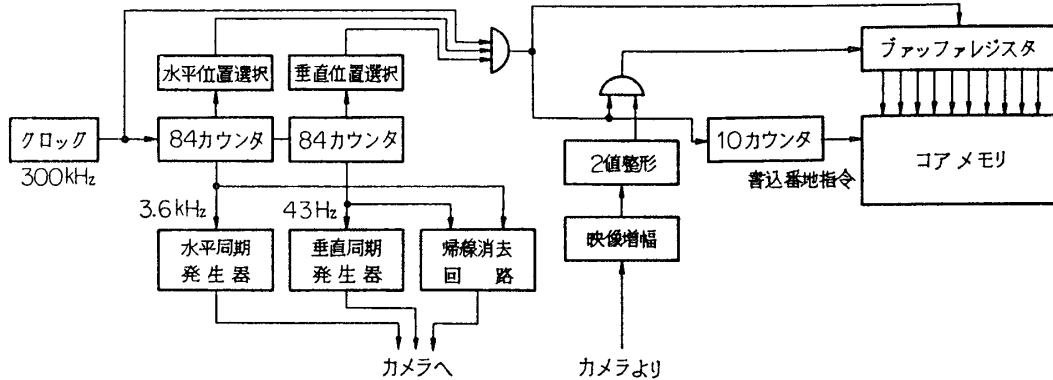


図1 書き込み部ブロック図

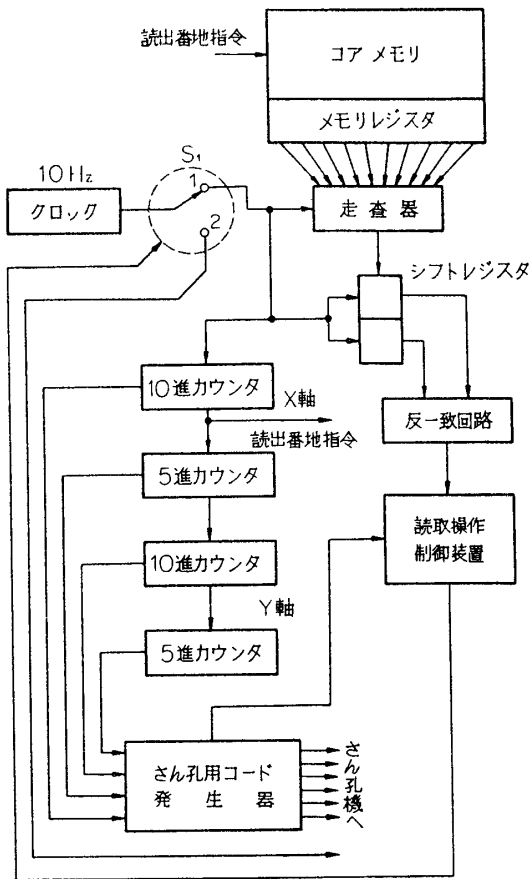


図2 読み出し部ブロック図

2.1 装置の構成

装置は図1および図2に示すとき書き込み部および読出し部さらにさん孔機より成る。読取り部分入力には低速走査 I. T. V. を用い信号振幅を二値レベル化し、画面が 50×50 の二次元マトリックスパターンとなるような標本化を行ない、さん孔のための buffer memory であるコアメモリに記憶する(書き込み部)。書き込み終了後、計算機入力として適当な内容に変換し、テープにさん孔する(読み出し部)ものである。以下書き込み部と読み出し部に分けて説明する。

2.2 書き込み部

ここに用いた I. T. V. は市販の標準方式のもので、その走査部を改良し低速走査とした(図1参照)。垂直走査線数 84 本(内帰線区間 4 本)、画面縦横比 1:1 としたためクロック周波数はおよそ 300kHz となった。このようにして得られる全画面(80×80メッシュ)のうちの任意の位置の

50×50 メッシュを選択して取り出す。順次走査して得られた映像信号を増幅し二値化整形を行なう。クロック周波数が約 300kHz であり直接テープのさん孔は行なえないため、二値化された信号は 10 ビット(コアメモリ 1 語分) ずつバッファレジスタにまとめられて、一たんコアメモリに書き込まれる。この動作は 1 フレーム分(250 語)の書き込みで終了する。

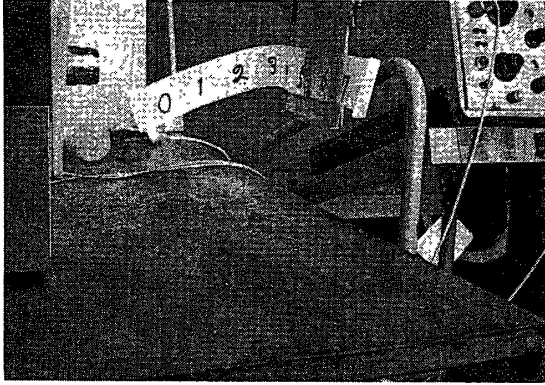


図 3 計算機入力装置読み取り部分

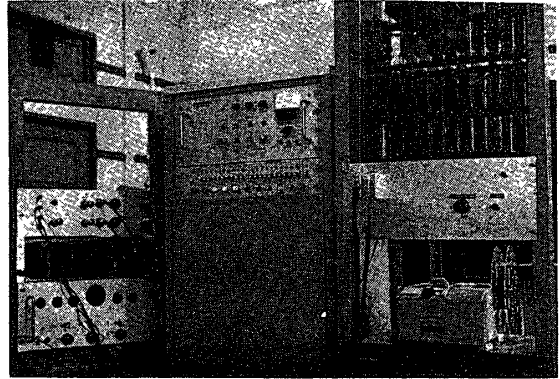


図 4 計算機入力装置制御部分

2.3 読み出し部

コアメモリに記憶された二値化パターンは、二値化パターンにおいて黒→白、白→黒の変化があった位置を (X, Y) 座標に変換してさん孔される (図 2 参照)。

コアメモリより 10 ビットずつまとめて読み出された二値化パターンは、1 ビットずつを順次に走査しシフトレジスタに送られる。前のビットとの変化があったとき反一致回路の出力があり、S1 が 1 から 2 に切り替えられ、その座標に対応したカウンタの内容が X 軸用 5 進カウンタ、10 進カウンタ、スペース、 Y 軸用 5 進カウンタ、10 進カウンタ、スペースの順に切り替えられてさん孔機に送られ、2 進化 10 進表示でさん孔される。さん孔が終了すると、S1 は 2 から 1 に切り替えられ走査がつづけられる。10 ビットごとに読み出し番地指令が出され次の語を読み出し、走査がつづけられて 2500 ビット (1 画面分) 走査し終ると読み出しは終了する。さん孔速度は 10 キャラクタ/秒である。また一画面のさん孔時間はおよそ 5 分である。

装置の写真を図 3, 4 に、二値化したパターンをブラウン管上に再現させたものを図 6 に、

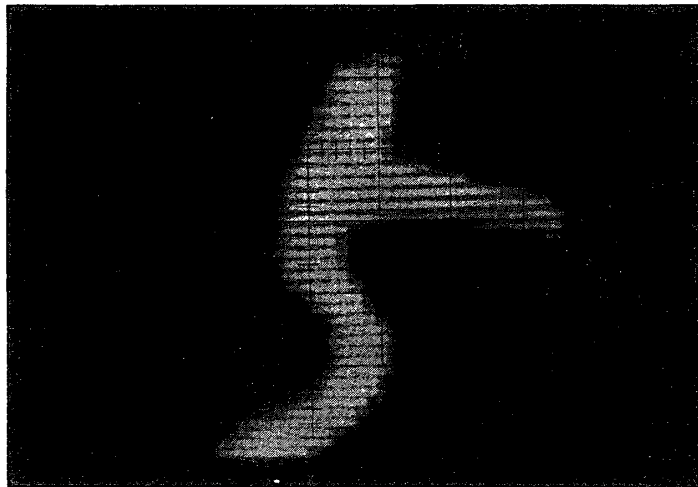


図 5 二値化されたパターン (CRT上)

```

0000
000000
00000100
0000000000
00000000100
000000000000
000000000000
00000000
0000000
00000000
0000000
00000000
0000000
000000
0000000
0000000
0000000
00000000
0000000
0000000
0000000
0000000
000000
000000
0000000
0000000
0000000
0000000
0000000
0000000
0000000
0000000
000000
000000
000000
000000
0000000
000000
0000000
0000000
0000000
0000000
000000
0000000
0000000
0000000
000000
000000
0000000
0000000
000000
0000000
0000000
0000000
000000
0000

```

図 6 計算機により
再現されたパ
ターン

得られたデータテープから計算機によって再現された数字を図 6 に示す。

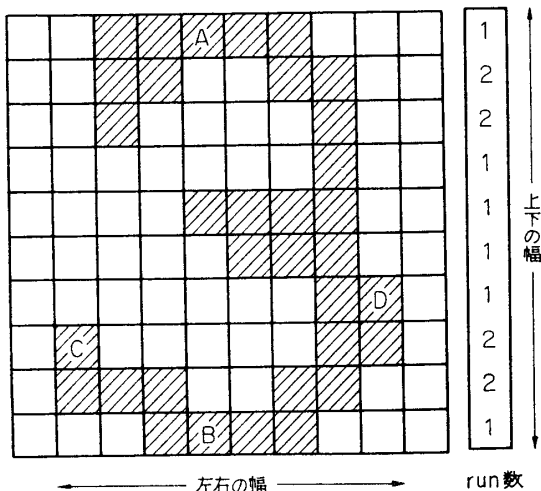
3. 横方向走査を主体とした特徴抽出

字を適当な投影軸に投影させて認識に応用することは、印刷文字に関してこれまでも行なわれてきている。この手法を手書き数字の特徴抽出に応用するべく、横方向と縦方向の両走査線のおのおのと手書き数字の字画との交差状態に着目した。走査線と字画との交差する数のみに注目すると、安定したパターンが得られ、特に横方向の走査線と字画との交差状態が安定しており、これによって手書き数字のパターン(字体の変化も考慮したもの)をグループ分けできることが判明した。そこでまず手書き数字をこのようなグループ分けを行ない、各グループの中では数字のパターンが限定されることを利用して、横方向の投影を主体とした微細構造による特徴を使用して個々の数字の識別を行うことを試みた。この方式では、字の大きさ位置の変化に対して強い認識機構ということがねらいである。

3.1 実験

前述の装置によって得られたデータを用いて、手書き数字を計算機の中に 50×50 の二値化マトリックスパターンとして再現する。この 50×50 のマトリックスの横方向の一行について、そこに連続した黒がいくつあるか (run 数と呼ぶ) を各行についてとる。た

だしそこで長さ 1 の run は雑音と見なして省略する。run と run の間長さ 1 の白である場合には、そこを黒に置き換えてやる。この 50 個の行の run 数を列方向に見た場合 (sequence と名付ける。図 7 参照) にこれがある特徴を持つので、これによって数字を大きくグループ分けする。手書き数字が各グループに分かれる様子を表 1 に示す。ここでは各行の run 数を列方向に見た sequence において、同一 run 数が 2 sequence 以上続かない場合には run 数の変化はないものと見なし、先行の run 数に置き替えている。表 1 に分類するに当っては、つながった同一の run 数の sequence を字の縦の大きさ (sequence の総和) で正規化したものを判定に使っている。なお、



実際は 50×50 のマトリックス斜線の部分
が黒をあらわす上下幅、左右幅を図のよう
にとり A, B, C, D は各々上下左右の頂

図 7 二値化したパターンの前処理

表 1 Sequence による分類

全部が1	1 1 2 3 3 5 5 7
長い2	0 2 3 4 5 6 9
下に長い1	1 2 3 4 7 9
上に長い1	2 3 4 6
その他	2 3 4 6 8

字の回転も考慮して表 1 は作ってある. 実際に各グループに含まれるパターン数は次のとおりである.

sequence によるグループ	含まれるパターン数
1 だけの sequence	17
長い 2 または上に 2 を含む	17
下に長い 1 を含む	16
上に長い 1 を含む	11
その他	14

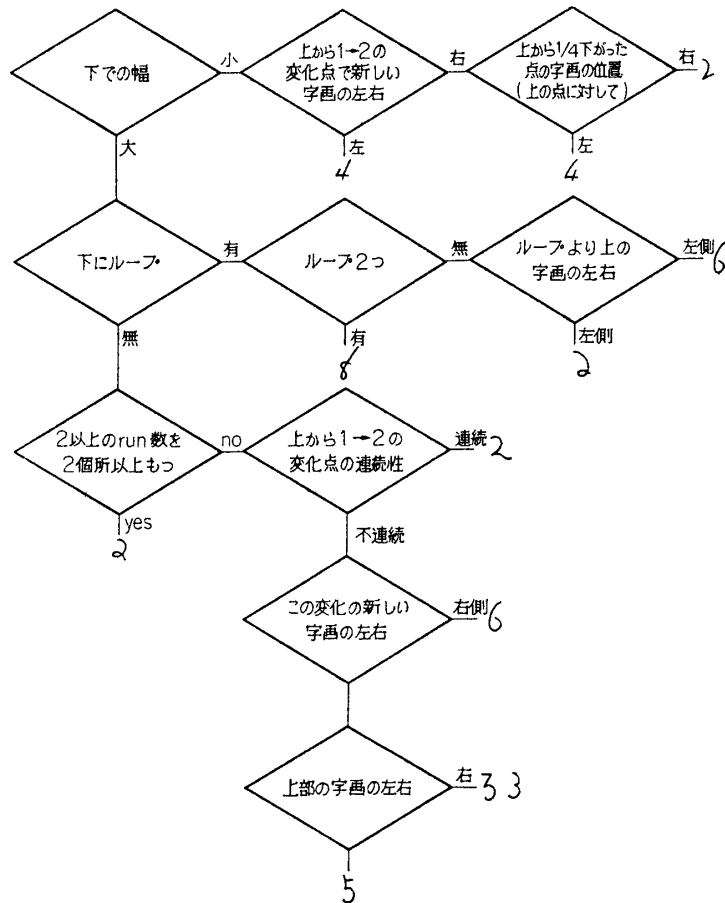


図 8 グループ内での個々の字の識別 (上に長い 1 をもつ Sequence)

各グループの中での個々の字の識別については、以下に述べるような特徴を用いて識別用の tree を cut and try で構成した。一つのグループについての例を図8に示す。

個々の字の識別に使用した特徴群

- ・上下左右の幅、頂点
- ・横方向に長い字面の有無
- ・sequence の変化点での字面の連続性
- ・ループの有無
- ・直線のあてはめと線の凹凸

3.2 結果と検討

手書き数字200字のデータについて計算機によるシミュレーションを行なった結果、認識率としては95%程度が得られた。ただし、この場合は rejection は許さないで判定を行なった。字の大きさ、位置など各種パラメータの変化を考えた場合には多少の変更は必要かと思われるが、手書きの数字と限った場合には、このような比較的簡単な認識方法でもかなりの字体の変化を許し、かつ認識率を高くすることが可能と思われる。

このシミュレーションにおいて、考察によりあらかじめ考えておいた数字のパターンは全部で75個であり、個々の字の認識までの tree における step 数の分布は次のとおりである。

step 数	3	4	5	6	7	8	9	計
パターン数	6	15	11	19	13	7	4	75

全パターンが等確率に起るものとすれば、平均 step 数は5.7となる。ただし、tree での分岐は必ずしも二者択一ではない。シミュレーションにおけるデータ数が少ないため、あらかじめ考えた全パターンは生起していない。

4. Curve Follow を用いた手書き数字の認識

手書き数字がその横方向への投影を主体とした特徴によって認識できることは判明したが、字体の多様性のために考えるべきパターンの数は多くなってしまった。特徴のうちでループを持つということは、字において大きな比重を持っており、前の方法でもあまり字体の変化に影響されずに検出が可能であった。しかし、その他の特徴は字体の少しの変化が投影を変化させるため投影から得ることは困難が多かった。これに対して字の持つ字面としての特徴は、字体が変化しても保たれると考えられる。そこで字体の変化にできるだけ影響されないような主として stroke としての特徴を見出すべく、文字の量子化パターンについて、それを字面のつながりと見て follow することによって特徴を抽出することを実験した。ここでは follow の行ない方が重要な factor であるので、考えられる二つの follow の方法を実験した。

第一の方法は、字の筆跡の中心を忠実に追うことを目的としたものであり、第二の方法は字面をその中に引くことのできる一番長い8方向の線分で折れ線近似するものである。識別にはこのような follow によって得られる主として stroke としての特徴、すなわち字面の端点、交差点、尖点ループ等を検出し、これらの特徴とその位置を使って個々の字を認識する tree を構成して計算機によるシミュレーションを行なった。

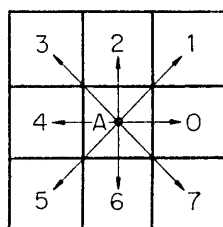


図 9

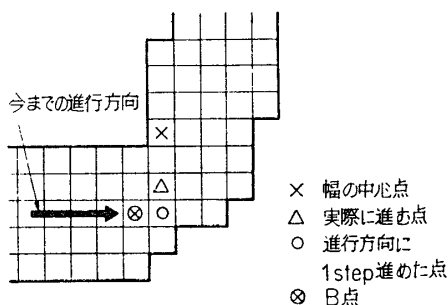


図 10

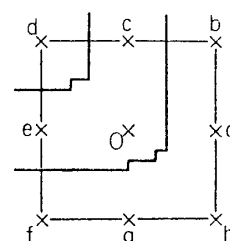


図 11

4.1 follow の方法 (その1)

ここでは手書き数字は、すでに前述データによって 50×50 のマトリックスに二値化されて計算機の中に再現されたパターンを follow するものとする。以下にその方法を列記する。

- 1) follow の 1 step で進み得る点は隣接する 8 点 (図 9 の A 点からは 0~7 の 8 点) のいずれかとする。(ただし方向は制限する。)
- 2) follow を開始し到達した点 B で次に進み得る点は、進行方向とその両側の 3 方向に対応した隣接した 3 点とする。
- 3) follow に normal mode と search mode の二つの mode を設ける。
- 4) normal mode: 進行方向に 1 step 進め、進行方向と直角な方向の幅をはかり、この幅が字の平均幅 (横方向の run length の平均値で代用する) を越えないならば許される三点のうち幅の中心の点に最も近い点へ進む。(図 10)
- 5) search mode: この mode は、次の三つの場合に適用される。(a) 出発点を検出する場合 (b) normal mode において進めてみた点が字画を飛び出した場合、(c) 進行方法に直角にとった幅が平均幅以上の場合。この mode に入ると与えられた点のまわりに正方形を作り、正方形の周と字画との交差の状態から進行方向を検出し (図 11), normal mode へうつる。この正方形の一辺の大きさとしては、字の平均幅のものとその二倍の二種類を用いている。

4.2 follow の方法 (その2)

字画の中に引くことのできる 8 方向 (図 9) の直線のうち、一番長いものを使って字を近似しながら follow を行なっていく。図 12 において説明すると、出発点 C で 8 方向に引ける直線の長さを調べ、その中で最も長い 4 の方向に進み D 点に行く。D 点では進行してきた方向へ戻らないよう制限を加えながら、最長の直線で近似できる 5 の方向を求め E 点に行く。このような操作を繰り返しながら、筆跡を折線近似しながら follow して行く。

この方法は、字の追跡の手続きが簡単であり、字が単

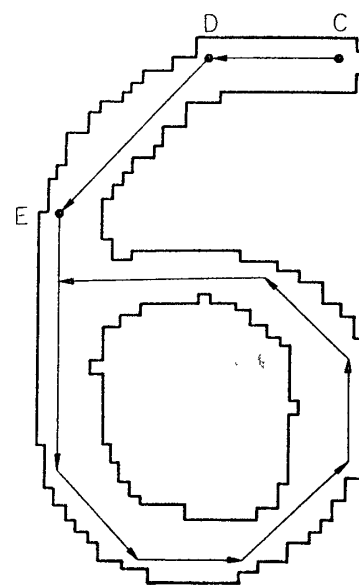


図 12

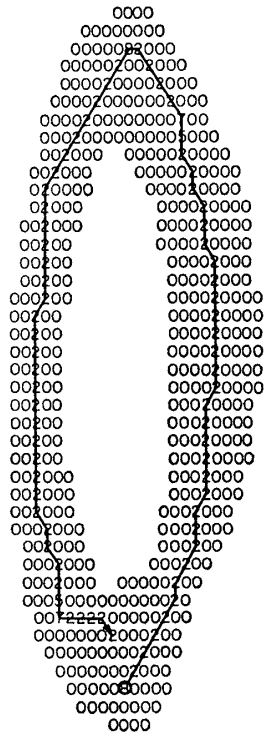


図 13 字画の中心の追跡

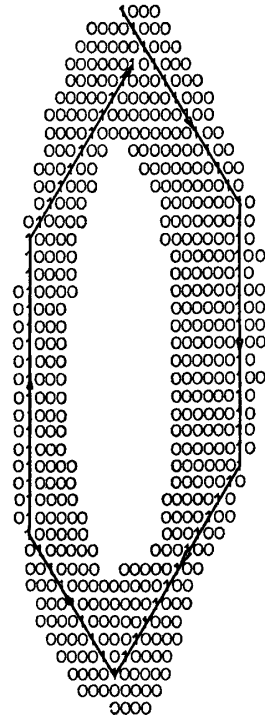


図 14 折れ線近似

純な形をしている場合には follow の誤りも少なく，特徴の抽出に関しても第1の方法と同様な結果を得ることができる．図 13, 14 にこれら二つの方法を対比させてある．なお表 2, 3 に第2の方法における折れ線の数とその平均の長さを示してある．

表 2 折れ線近似における線分数

データ No. 数字	1	2	3	4	5	6	7	8	平均
0	6	10	8	11	8	12	10	7	9.0
1	6	4	5	5	3	7	5	4	4.9
2	7	14	7	8	6	7	7	5	7.6
3	12	9	13	11	8	12	7	7	9.9
4	9	9	10	12	10	9	9	7	9.4
5	17	14	6	9	10	8	9	9	10.3
6	12	15	7	8	7	9	5	9	9.0
7	8	10	9	6	4	8	4	6	6.9
8	13	14	11	13	9	6	9	9	10.5
9	13	12	9	13	9	6	7	9	9.8

4.3 follow による特徴抽出と数字の認識

手書きの数字においては字体がさまざまに変化する．このような変化にかかわらず保たれ

表 3 折れ線近似の線分の長さ

データ No. 数字	1	2	3	4	5	6	7	8	平均
0	1.54	1.37	1.50	1.23	1.18	1.14	1.90	2.37	1.49
1	1.39	1.80	1.33	1.43	0.98	0.96	1.35	1.71	1.35
2	2.20	0.85	1.14	1.56	0.86	1.29	1.03	1.04	1.22
3	1.10	1.20	0.92	0.98	0.89	1.06	1.24	1.05	1.05
4	1.57	1.52	1.18	1.31	0.87	1.32	1.53	1.69	1.35
5	1.19	1.06	1.08	1.06	0.83	0.87	1.25	0.81	1.04
6	1.02	1.23	1.81	1.62	1.46	1.10	1.65	1.01	1.30
7	1.58	1.22	1.44	1.36	1.42	1.10	2.56	1.52	1.45
8	1.46	1.20	1.56	1.29	1.26	1.32	1.56	1.25	1.36
9	0.97	1.50	1.15	0.85	1.21	1.32	1.29	1.22	1.16

(注 平均 run length をと1してして normalize ある)

ている性質，すなわち我々が書き表わそうとしている字の性質というものは，トポロジカルな特徴の組合せであろうと考えられる。これらのことを考慮に入れて，数字の識別に有効であると思われる特徴を各数字に割りあてた。

数字のもつトポロジカルな性質

0 全体がループ。交差数 2

○→端点の近いものはループと見なす。

1 上下方向の長い直線。かざりのある場合がある。①，②→○の部分はかざりとみなす

2 下にループ，4 差，3 差または尖点があり，中央が右に凸。③：ループ，④：4 差，⑤：3 差，⑥：尖点

3 中央に 3 差，4 差，尖点があり，上下が右に凸。⑦：3 差，⑧：4 差，⑨：尖点

4 中央右の 3 差または 4 差から左へ直線があり，その端が 3 差または 4 差 ⑩

5 上部に水平方向の字画があり，その端点が右側 ⑪

6 下にループ最大交差数 3，または最大交差数 4 でその点から短い線。⑫：ループ ⑬：交差数 4

7 上下方向に長い線分と上部に水平方向の字画があり端点が左側。端点にかざりがある場合もある。

8 全体または下にループで最大交差数 4。⑭：全体のループ ⑮：下にループ

9 上にループで 4 でないもの ⑯ ⑰

ここで使用している特徴としては，端点，交差点，尖点，ループ，直線および線の凹凸などの全体に対する位置と大きさである。

follow を行ないながら，これらの特徴を検出していくわけであるが，follow の方法としては第 1 の方法を適用している。字画の端点，交差点，尖点などでは normal mode から search mode へうつるので，search mode で方向の検出を行ないどの場合に相当するかを決めて，その場所と共に記録しておく。follow を開始した点にもどった場合および既に follow

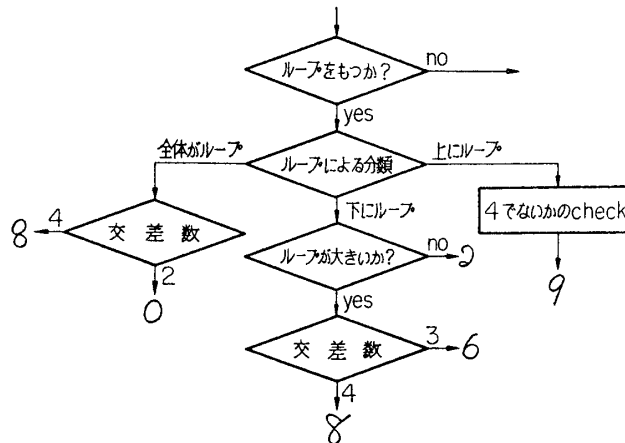


図 15 判定 tree の一部 (ループ)

した線にぶつかった場合には、ループであるかどうかを決め、さらにまだ follow していない部分がないかどうかを調べる。このようにして全体を follow することにより、特徴を抽出して記録を行なう。直線、線の凹凸に関しては follow を終わってから follow した点の中から適当な二点を取って直線のあてはめを行なうことにより得ている。

トポロジカルな特徴を用いて構成された判定 tree (その一部を図 15 に示す) によって、上に述べた特徴抽出を計算機の中に二次元二値化パターンとして再現された手書き数字に適用して識別の実験のシミュレーションを行なった。

認識率は 200 字のデータについて 90% であった。この場合には reject を許してある。これは判定の tree の不備によるものが多いと考えられ、実際にはもっと高い認識率を得ることが可能である。この方式では、字体の変化に影響されにくい本質的な特徴と考えられるものを使っているため、最終的に必要とされる特徴の数とその組合せも少なく、判定の tree も簡単なものとなっている (この場合には最終的に考えたパターンの数は 18)。反面、特徴抽出の手続きはかなり複雑なものとなることはやむを得ない。

5. 結 言

手書き数字の認識においては、横方向の投影を主体とした特徴が、かなりの字体の変化に対しても有効であることが判明した。これらの特徴は、読み取り時における走査の工夫により容易に得られ、かつ使用している特徴の数も比較的少ない。ただし字体の変化に伴って、これらの特徴も大きく変わることがあるので、あらかじめ考えておくパターンが、10 個の数字に対して 78 個とかなり多い。しかしさらにさまざまな字体の変化を考えても、考えるべきパターンは 100 内外で済むものと考えられる。そこでこの方法は、特徴のとり方が簡単でもあり、手書き数字の認識において一つの有効な手段であると思われる。

follow について実験した二つの方法は、細部を除いては同様な働きを示した。さらに両者の長所を組み合わせることにより、完全な follow を行なうことができると思われる。ただし、ここでは一度マトリックスとして二値化されたパターンについての follow を行なったわけ

で、他の条件のもとでは多少変えねばならない点もあると思われる。すでに書かれた字を follow することは、技術的にいくつかの問題を含んではいるが、字画の性質は字体の影響を受けにくいので、本質的に有効な手段であろう。手書き数字に限らずさらに複雑なパターンへの応用も可能であろう。シミュレーションにおいては、あらかじめ考えたパターンは上記の方法に比べて非常に少なく（この場合には 18 個）、さらに大きな字体の変化を考えても、この数はあまり増さないことが期待できる。このように follow を行なって、特徴を抽出することは、多少複雑にはなろうけれども、手書き数字の認識の根本的な解決に役立つものと考えられる。

1968 年 8 月 29 日 計測部