

知能作用過程のシュミレーション

その6 文字のパターン認識と「あいまいさ」の理論的考案

寺岡 宏 *矢吹 哲夫

Simulation of the Operating-Processes of the Human Intellect

Part 6. Pattern Recognition of letters and a theoretical study of the obscurity involved in the form of the letter

Hiroshi Teraoka and Tetuo Yabuki

In this paper, a model for pattern recognition by personal computer was proposed. The following three types of parameters were used in this model.

1. Five kinds of Space found in the figure of the letter
2. Four kinds of Lines enclosing the space
3. Two kinds of Symmetry in the figure of the letter

Using these parameters, capital letters were classed into three groups. V an U, H and X, Y and W showed the same pattern concerning space, line and symmetry in this model. To distinguish V and U, H and X, Y and W, one more parameter is required. Length of line was used for this purpose. The length of line shows some range of obscurity in pattern recognition. For the theoretical study of obscurity, membership-function and fuzzy-entropy were calculated for U and V. It was found that the letter V included more obscurity than the letter U.

この研究シリーズの前報¹においてA-Zの英語大文字22文字についてのパターン認識が扱われた。その方法としては、網膜をグラフィック画面としてモデル化し、その上に描かれた文字を49個の点に分解し、その情報がパターン認識の行なわれる場に伝達された。次に与えられた情報から垂直線、水平線、右下がりが斜線、左下がりが斜線、上下に閉じた曲

線、左右に閉じた曲線の6種類の線を識別し、それらの組み合わせと線が互いに交差する点の情報を総合して文字のパターン認識が行なわれた。今回の論文においては、前回用いられたパターン認識の方法を以下の観点について改め、パターン認識にともなう「あいまいさ」に対して理論的な考案を行なった。

文字のパターン認識においては、活字のように定型化されたものについては前報¹の方法が効果的に実行される。しかし画面上にマ

*北海道大学・理学部・物理学科・大学院博士課程

ウスによって手書きの文字を書く場合、最も単純な垂直線、水平線でも「ゆがみ」「ねじれ」が生じ、これを直線とするか、曲線とするかの判断は困難である。前報¹においては、文字の分解による情報を用いて線の形の判断が行われるためには、各タイプの線がもつ形の特徴を前もってコンピュータに記憶として与えておき、この記憶内容と伝達された情報との照合によって判断がなされるものとした。

「ゆがみ」「ねじれ」による曖昧さが生じる場合、それぞれのケースについての情報を前もって与えておくことは、プログラ的には複雑化をもたらすこととなる。特に曲線については形の「ゆがみ」が複雑、多様であり、取り扱いが困難となる。

今回の論文においては、人間のパターン認識の特徴に近づけることを目的として、以下の作業仮説に基づく、新しいパターン認識の方法を追求した。

1. 人間による文字のパターン認識においてはすべての視覚情報に対して、記憶との1対1の照合がなされるのではなく、文字の形に見られるいくつかの特徴のみを取り出し、これと記憶との照合がなされる。
2. 上記以外の形を構成する要素については、「あいまいさ」の対象とし、その判断においては前報²において扱われた所属度関数による取り扱いがなされるものとする。

文字のパターン認識の方法

1. グラフィック画面を網膜としてモデル化し、その上にかかれた文字をCOMMAND〈SYMBOL〉を用いて横7倍率、縦7倍率で所定の場を書く。この一文字を横20個、縦25個、合計500個の点に分解する。500個の分解点の占める面積は1個の文字面積の

24%に相当する。

2. 文字の形における特徴として空白の有無に着目し、空白を以下のA-Eの5つのタイプに区別した。

A: 図1に示されているように、文字の中心部分に横6点、縦6点の広がりをもつ空白

B: 図2に示されているように、文字の中央上側部分に横6点、縦4点の広がりをもつ空白

C: 図2に示されているように、文字の中央下側部分に横6点、縦4点の広がりをもつ空白

D: 図3に示されているように、文字の左側部分に横10点、縦15点の広がりをもつ空白

E: 図3に示されているように、文字の右側部分に横10点、縦15点に広がりをもつ空白

空白の検出にはCOMMAND〈GET@A〉を用いた。

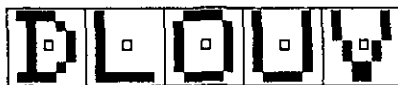


図1 空白Aの位置



図2 空白B、Cの位置

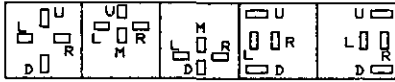


図3 空白D、Eの位置

3. 上記A-Eの空白部分を4方向から囲む線を以下の様に規定し、これを次の方法に従って調べた。

- L : 空白A-Eの左側に位置する線
- R : 空白A-Eの右側に位置する線
- D : 空白A、C、D、Eの下側に位置する線
- U : 空白A、B、D、Eの上側に位置する線
- M : 空白Bの下側、空白Cの上側に位置する線

上記L-Mの線の確認は図4に示される4角形を設定し、その内部が完全に空白のときは、線のない場合、4角形の内部になんらかの点が存在する場合はその場所に線があるものとした空白の検出にはCOMMAND<GET@A>を用いた。



空白: A B C D E

図4 空白A、B、C、D、Eを囲む線の位置

4. 各文字の分解によって得られた500個の点の情報を2次元配列に入れ、これを基に文字の形における左右対称性、上下対称性を求めた。

また配列の情報から上記3の空白を囲む線の長さをもとめた。

文字のパターン認識の結果

以上の方法に基づき文字のパターン認識の結果、英大文字は空白A、空白BC、空白DEを有する3つのグループに分けられる。表1、2、3に空白とそれを囲む線、線の長さ、対称性についての結果を示す。

表1 空白Aをもつ文字と空白を囲む線、線の長さ、対象性 (LRは左右対称性、UDは上下対称性)

Space and Line	Line Length	Symmetry	
		L.R	U.D
C A:LDU	L11 U10 D10	40%	100%
D A:LRDU	L25 R11 U14 D14	30%	100%
G A:LDU	L11 R7 U10 D10	0%	47%
J A:RDU	R21 U11 D11	0%	19%
L A:LD	L25 D21	50%	14%
O A:LRDU	L17 R17 U14 D14	100%	100%
Q A:LRDU	L17 R14 U14 D11	0%	47%
U A:LRD	L21 R21 D14	100%	0%
V A:LRD	L11 M7 R11 D7	100%	0%

表2 空白B、Cをもつ文字と空白を囲む線、線の長さ、対称性

Space and Line	Line Length	Symmetry	
		L.R	U.D
A B:LRMU C:LRM	L18 R18 U7 M21	100%	0%
B B:LRMU C:LRDM	L25 R7 U17 M14 D17	30%	100%
E B:LMU C:LDM	L25 U21 M14 D21	30%	100%
F B:LMU C:LM	L25 U21 M14	30%	14%
H B:LRM C:LRM	L25 R25 M21	100%	100%
M C:LRM	L25 M7 R25 M7	100%	28%
N C:LRM	L25 R25	30%	47%
P B:LRMU C:LM	L25 R7 U17 M17	60%	14%
R B:LRMU	L25 M7 R7 U17 M17	30%	14%
S B:LRMU C:LRDM	L7 R7 U14 M14 D14	60%	61%
W B:LRM	L25 M7 R25M7	100%	33%
X B:LRM C:LRM	L7 R7 M7	100%	100%
Y B:LRM	L11 M14 R11 M10	100%	33%
Z B:RMU C:LDM	L7 R7 U21 M7 D21	30%	38%

表3 空白D、Eをもつ文字と空白を囲む線、線の長さ、対称性

Space and Line	Line Length	Symmetry	
		L.R	U.D
I D:RDU E:LDU	M25 U10 D10	100%	100%
K E:LDU	L25 M10	0%	100%
T D:RU E:LU	M25 U18	100%	22%

以上の結果を空白の種類、囲む線、対称性の3点に着目するときUとV、XとH、YとWの文字については互いに同一の情報を示すことになる。これらの文字の区別には特定の線の長さを比較することが必要になる。

文字を構成する線は互いにある範囲の長さ

において、その文字の形の特徴を反映するものである。すなわち、その文字の固有の形と線の長さとの間にはある範囲の「あいまいさ」が許容され、文字のパターン認識には「あいまいさ」は不可欠な条件としてとりこまれている。例えばUとVの場合、底辺の線がある範囲より短い時はVとなり、反対にこの線がある範囲より長くなるときUとして認識される。

UとVにおける横線と縦線の比に基づく所属度関数 (Membership 関数)

上記表1の結果、文字Uでは、縦線：横線＝3：2の関係において活字の形を示している。しかし第1義的な形の特徴として中心の空白とこれを3本の線が囲み、100%の左右対称性を示す場合、縦線と横線の比はある範囲内において変化してもパターン認識には影響しない。以上の形に見られる「あいまいさ」は理論的には De Luca et al³によって提唱された、所属度の概念を用いて定量的に扱うことが可能とされる。

文字Uの所属度を縦線 (Y) と横線 (X) の長さの関数として次の式で規定した。

$$pi=3.1416$$

$$S(U)=\sin((X/7)*(7/10)*pi)^2*\sin((Y/5)*(5/14)*pi)^2\dots 1$$

1式においては縦線 (Y) と横線 (X) の長さの比が7：5のとき、S(U)=1 となる。所属度は $0 \leq S(U) \leq 1$ の範囲の値を示し、S(U)=1 のときは「不確かさ」は完全に無く文字Uの判断がなされる。S(U)=0 のときも同様に「不確かさ」は完全には無く、文字Uでは無いことの判断がなされる。S(U)=0.5のときは同様に文字Uであることと無いことの確率が等しくなり、判断は不可能とされる。このとき「あいまいさ」は最大を示す。

表4に1式の計算結果を示す。図5は横線

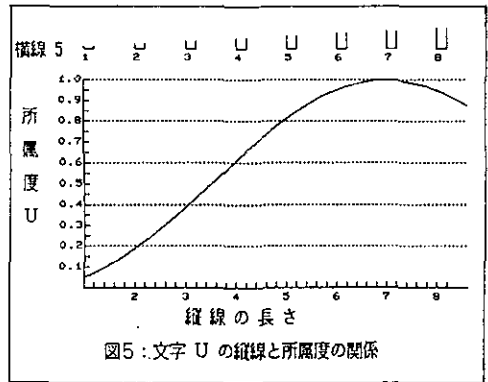


表4 文字Uの横線と縦線の長さとの所属度関数の値

縦線 \ 横線	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6
1	.004	.01	.017	.024	.032	.039	.044	.048	.049	.048	.044
1.5	.01	.022	.037	.054	.071	.086	.098	.106	.109	.106	.098
2	.017	.038	.065	.094	.123	.149	.17	.183	.188	.183	.17
2.5	.027	.058	.097	.141	.185	.224	.256	.276	.283	.276	.256
3	.037	.08	.134	.194	.254	.308	.351	.379	.388	.379	.351
3.5	.047	.103	.172	.249	.327	.396	.452	.487	.499	.487	.452
4	.058	.125	.211	.305	.4	.485	.552	.596	.611	.596	.552
4.5	.068	.147	.247	.358	.469	.569	.648	.699	.716	.699	.648
5	.077	.167	.28	.405	.531	.644	.734	.791	.811	.791	.734
5.5	.085	.183	.307	.445	.583	.707	.805	.869	.89	.869	.805
6	.09	.195	.328	.475	.622	.754	.859	.927	.95	.927	.859
6.5	.094	.203	.341	.493	.646	.783	.893	.963	.987	.963	.893
7	.095	.206	.345	.499	.654	.793	.904	.975	1	.975	.904

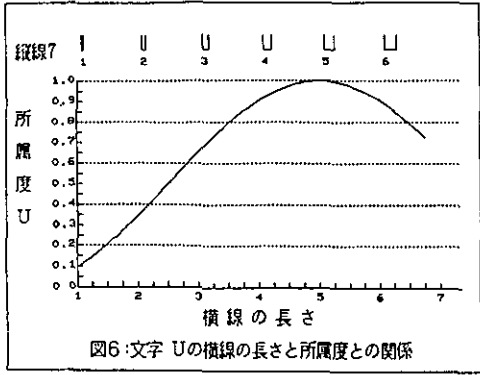


図6: 文字 U の横線の長さ と所属度 との関係

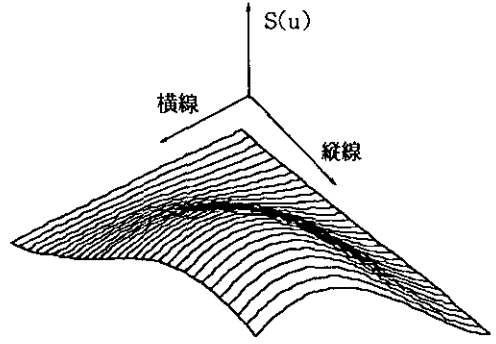


図7 文字Uの所属度関数の値

の長さを5と固定し縦線を変化させた場合の $S(U)$ の値、図6は縦線の長さを7と固定し横線を変化させた場合の $S(U)$ の値を示す。図5、図6の結果から文字Uのパターン認識にみられる「あいまいさ」が縦線と横線の長さの比として連続的に変化する様子が明らかにされる。図7に3次元グラフとして縦線(Y)と横線(X)の長さ と $S(U)$ 値の関係を示す。この曲面の形は文字の形の特徴を示すものと解釈される。文字Vにおいては図8に示されている縦線(Y)と横線(X)を用い、所属度関数として2式を設定した。

表5に2式の計算結果を示す。図9に3次元グラフとして縦線(Y)と横線(X)の長さ と $S(V)$ 値の関係を示す。曲面の形は文字Vの形の特徴として底辺の横線の長さが所属度に強い影響を及ぼすことを示している。

UとVにおける所属度関数に基づく「あいまいさ」とファジーエントロピー

所属度関数の値を用い、次の3式による値を計算した。3式は「あいまいさ」の指標としてのファジーエントロピー密度を示すものである³⁾。

$$S(V) = \sin(((3 - X)/7) * (7/6) * \pi) \\ \wedge 2 * \sin((7 - Y)/3) * (3/14) \\ * \pi) \wedge 2 \dots\dots\dots 2$$

$$M(A_i) = -S(A_i) * \log_2 S(A_i) - (1 - S(A_i)) * \\ \log_2 (1 - S(A_i)) \dots\dots\dots 3$$

表5 文字Vの横線と縦線の長さ と所属度関数の値

縦線 \ 横線	0	.5	1	1.5	2	2.5
0	1	.933	.749	.499	.249	.066
.5	.987	.921	.74	.493	.246	.066
1	.95	.886	.712	.475	.237	.063
1.5	.89	.831	.668	.445	.222	.059
2	.811	.757	.608	.405	.202	.054
2.5	.716	.668	.537	.358	.179	.048
3	.611	.57	.458	.305	.152	.04
3.5	.499	.466	.374	.249	.124	.033

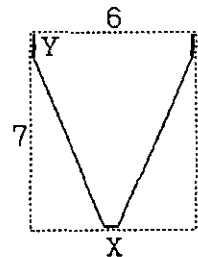


図8 文字Vの縦線と横線の位置

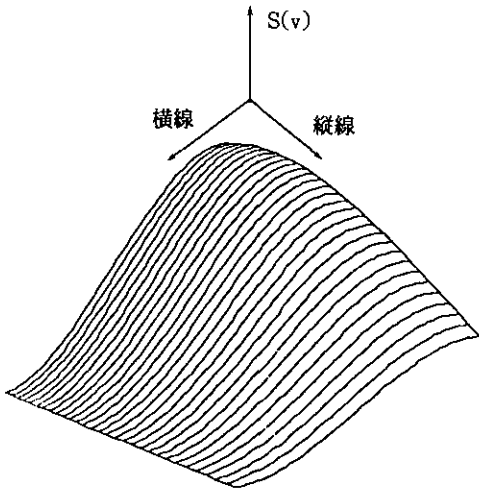


図9 文字Uの所属度関数の値

$S(A_i)$ は集合Aの1番目の要素の示す所属度関数の値で、今回のデータとしてはUは表1、Vは表2の値に相当する。 $M(A_i)$ の値は $0 \leq M(A_i) \leq 1$ の範囲の値を示し、 $S(A_i)=0.5$ のとき $M(A_i)=1$ となり、最大の「あいまいさ」を示す。 $S(A_i)=0$ または $S(A_i)=1$ のとき $M(A_i)=0$ として「あいまいさ」無しの状態を示す。図10に文字Uにおける縦線(Y)と横線(X)の長さ、図11に文字Vにおける縦線(Y)と横線(X)の

長さ、 $M(V_i)$ 値の関係を示す。図10、11の曲面は文字のパターン認識において縦線と横線によってもたらされる「あいまいさ」を数量化したものであり、各文字の形の特徴を「あいまいさ」を指標として図示するものである。図10、11の曲面の特徴を示すものとして前報²⁾において規定された「ファジーエントロピー」を用い、これを4式⁴⁾によって計算した。

ファジーエントロピー(A)

$$= 1/n * (M(A_1) + M(A_2) + M(A_3) + \dots + M(A_n)) \dots 4$$

ファジーエントロピーはファジー集合の各要素の示すファジーエントロピー密度の値を総計し、その平均値を求めたものであり、その集合の「あいまいさ」を指す係数としての意味を持つものである。

Uでは $1.0 \leq X \leq 6.0$ $1.0 \leq Y \leq 8.0$ の範囲 X、Yの各増分0.1

Vでは $0 \leq X \leq 2.5$ $0 \leq Y \leq 3.5$ の範囲 X、Yの各増分0.1

の条件でファジーエントロピーの計算がなさ

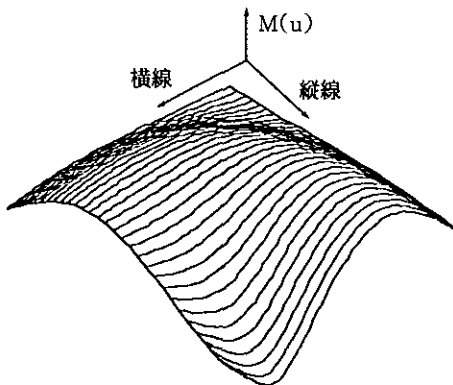


図10 文字Uのエントロピー密度と横線・縦線の関係

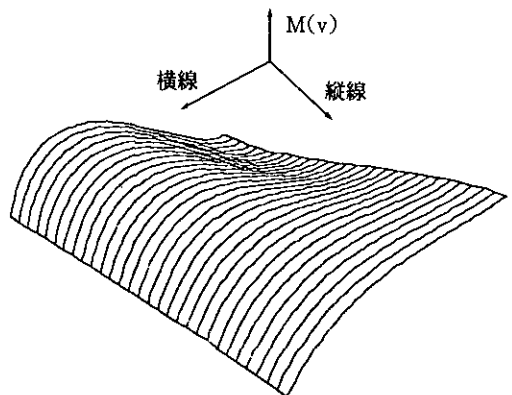


図11 文字Vのエントロピー密度と横線・縦線の関係

れた。

以上の計算結果

ファジーエントロピー(U)=0.657 (n=3621)

ファジーエントロピー(V)=0.742 (n=963)

の値が得られた。

横線と縦線の長さの違いによって構成されるUとVの集合においてUに比べVの方が全体的に見た場合、高い「あいまいさ」を有することが明らかにされた。このことは横線と縦線の長さの異なるUとVの文字が読まれる場合、Vの方がUよりもより高い不確かさを内包し、判断ミスを生じ易いことを示すものと推察される。

要 約

1. 文字のパターン認識の方法として文字の形に見られる空白とこれを囲む線、形に見られる対称性の3つのパラメーターが用いられた。その結果英大文字は3つのグループに分類されることが明らかにされた。
2. UとV、HとX、YとWの場合上記の3つのパラメーターでの比較においては同一の結果を示し、両者の区別のためには特定の位置の線の長さを用いる必要があることが明らかにされた。
3. UとVに関して縦線と横線の長さの違いによって構成される文字集合において、各要素の所属度の計算がなされた。縦線と横線によって構成される所属度の曲面を文字U, Vの形の特徴を示すものとして用いた。
4. 所属度に由来する、「あいまいさ」を示す関数として「ファジーエントロピー」が用いられ、UとVに関してその値が計算された。その結果UにくらべVの方がより高い「あいまいさ」を有する形であることが明らかにされた。

引用文献

- 1) 寺岡 宏、矢吹哲夫(1994) 知能作用課程のシュミレーション：その4 文字のパターン認識と記憶回路の形成 北星短大紀要 30:77-86
- 2) 寺岡 宏、矢吹哲夫(1994) 知能作用課程のシュミレーション：その5 文字情報の「あいまいさ」に関する理論的考察 北星短大紀要 30:87-94
- 3) De Luca, A and S. Termini(1972) A definition of nonprobabilistic entropy in the setting of fuzzy sets theory. Information and Control, 20: 301-312
- 4) 向殿政男、本多中二(1991) ファジー「あいまい」の科学. 35page. 岩波書店