

手書き文字のパターン認識に関する一手法

土橋 浩慶 原田 和郎 石原 好之 戸高 敏之

同志社大学 工学部

京都市上京区今出川通り烏丸東入

あらまし 本論文では、手書き文字のパターン認識に、構造解析法を用いた2段階の認識法を提案する。まず、第1段階で入力文字の全体的な構造特徴に近い登録文字を、近い順に20文字選び出し、第2段階で選び出した20文字の中から、分割された各部分の構造特徴により、入力文字に最も近い登録文字を選び出す方法である。この手法では、文字の個々の局所的な特徴を考慮することで、高い認識率を得ることができ、さらには、辞書の容量も非常に少なくすることができる。また、本手法による文字認識実験を、手書き文字142字(平仮名,片仮名)に対して行った結果、1文字約1秒の処理能力と認識率96.45%以上の良好な結果を得た。

和文キーワード 手書き文字認識、パターン認識、構造解析法、データ構造、パソコン処理

A Method of Pattern Recognition for Hand-Written Characters

Hironori Dobashi, Kazuou Harada, Yoshiyuki Ishihara and Toshiyuki Todaka

Faculty of Engineering, Doshisha University

Imadegawadori Karasuma Higashiiru, Kamigyo-ku, Kyoto 602, Japan

Abstract In this paper, a new method of pattern recognition for hand-written characters is proposed. This method is introduced a structurally analysis that has two stages. At the first stage twenty characters are selected by the feature of the whole structure of a inputted character. At the second stage a correct character is selected by the feature of the part structure in the selected twenty characters. As the proposed method uses the local feature of each character, it gives a high rate of recognition and reduces a size of dictionary. It is verified by the experiments.

Hand-Written Character Recognition, Pattern Recognition,

英文 key words Structurally Analysis Method, Data Structure, Personal-Computer Processing

1. まえがき

文字認識の一般的な手法は、パターン整合法と構造解析法とに大別できる。パターン整合法は、文字の大局的な特徴を把握するイメージ型の情報処理であり、入力文字パターンに加わる局所的な外乱には強いが、パターンの細かな違いが決定的な影響を与える類似文字の区別には弱い。一方、構造解析法は、文字を構成する線分間の局所的構造をとらえる論理型の情報処理で、パターンのわずかな違いも区別できる反面、文字のつぶれやノイズに弱く、多数の候補文字との比較のために莫大な計算量を必要とし、高精度であっても実用的ではない⁽¹⁾。

パターン整合法では、文字パターンを小領域に分割し、各領域のメッシュ特徴ベクトルを用いて類似度を求める方法⁽²⁾、構造解析法では、文字を全体的な構造情報とセグメント情報に分解して文字を認識する方法⁽³⁾、などが挙げられる。また最近では、ニューラルネットワークを用いた方法⁽⁴⁾⁽⁵⁾も検討されている。

これらは、いずれもそれぞれ優れた特徴をもっており有力な手法であるが、認識対象文字数、文字の種類、認識率、辞書の容量などの少なくともいずれかに問題点がある。

そこで本論文では、手書き文字のパターン認識に、構造解析法を2段階に分けて適用する認識法を提案する。文字の構造特徴を文字全体と分割された各領域に分けて入力文字と辞書との特徴量の比較を行い、入力文字を認識する。本手法では、辞書の容量が極端に減少でき、また、1文字当たりの認識も短時間でできる。さらに、本手法を用いて手書き文字142字(平仮名、片仮名)を対象に認識実験を行った結果を示す。

2. システムの構成および認識方法

2.1 システムの構成

文字の入力には、撮像装置として(株)東芝製のCCDカメラを用い、画像処理装置は(株)ソアテック社製の「VIDEO-98」を用いる。これによる入力文字は、16階調の濃淡画像として取り込まれている。また、CCDカメラにより入力された文字は画像データとして保存される。

その後、前処理(2値化,正規化,細線化)を

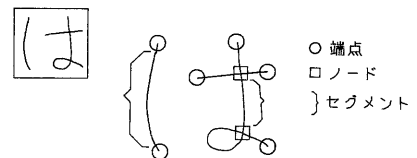
施し、特徴量の抽出を行い、辞書登録文字との2段階のマッチングにより、入力文字を認識する。これらの処理を、パーソナルコンピュータ(PC-9801RA)上で行う。またソフトウェアに用いた言語はCである。

2.2 認識方法

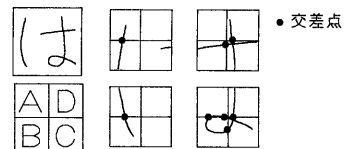
本論文では、構造解析法を2段階に分けて適用する認識法を提案する。まず、第1段階で入力文字の全体的な構造特徴に最も近い登録文字を近い順に20文字選び出し、さらに、第2段階で選び出した20文字の中から、分割された各領域の構造特徴により、入力文字に最も近い登録文字を選び出す方法である。この手法では、文字の個々の局所的な特徴を考慮することで、高い認識率を得ることが期待でき、さらには、辞書の容量も非常に少なくすることができる。

本論文で用いる文字の特徴量をFig.1に示す。文字全体の構造特徴量として、端点数、交点および分岐点数(以下、ノード⁽³⁾と呼ぶ)、セグメント数、半濁点数及び文字線の長さの5種類を用い、分割された各領域(Fig.1(b)のA, B, C, D)の構造特徴(以下、分割構造特徴と呼ぶ)として、端点数、ノード数、画素数、交差回数(縦横中央の2方向)の5種類を用いる。ここでセグメントとは、ノードとノード間、ノードと端点間および端点と端点間を結ぶ文字線とする。

個々の文字の特徴量分布図(相違図)の例をFig.2およびFig.3に示す。Fig.2から分かるよう



(a) 文字全体の構造特徴



(b) 分割構造特徴

Fig.1 特徴量の定義

に、例えば平仮名の「ね」と「れ」と「わ」は非常に似ており、全体的には端点数は変わらないが、ノード数は「ね」が他の二つに比べて一つ、セグメント数は各々の間に違い（後述する）がある。さらに部分的に見ると右下の横方向の交差回数に違いが生じる。

また、Fig. 3(a)～(d)を見ると平仮名の「い」「う」「こ」「り」では全体的な構造特徴はほぼ同一であるが、分割構造特徴の分割端点数や交差回数に違いが生じる。

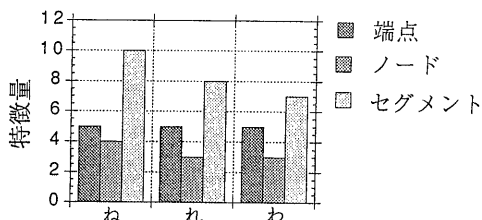


Fig. 2 全体の構造特数量による相違

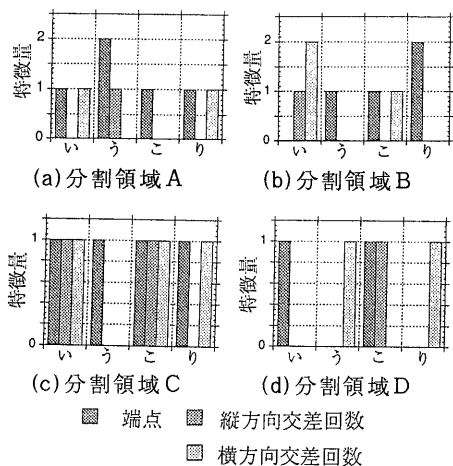


Fig. 3 分割構造特数量による相違

3. 文字の特数量抽出

文字の特数量抽出を行う前に、前処理として3種類の画像処理を行う。最初に、2値化を行い、大きさの異なる文字を処理できるように正規化を行う。次に、正確に特数量を抽出するために、細線化を行う。正規化は32×32ドットの大きさで行い、細線化は横井らの8連結の細線化手法⁽⁶⁾を用いて行う。

ここでは、細線化された文字から文字の構造特数量の抽出を行う。

3.1 文字全体の構造特徴

3.1.1 端点とノードの抽出

Fig. 4の3×3のマスクを用いて、文字全体の構造特徴としての端点とノードを抽出する。端点とノードかの決定は(1)式を用いて行う。

$$F(s) = \frac{1}{2} \sum_{i=1}^7 |s_i - s_{i+1}| \quad (1)$$

ここで、sは追跡中のセグメント上の注目の画素であり、s₀～s₇はその近傍の画素である。これらの画素が、黒画素の場合を値1、白画素の場合を値0とすると、

F(s) = 1のとき、sは端点

F(s) = 3または4のとき、sはノードとなる。また、それ以外はセグメントである。

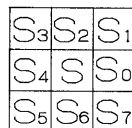


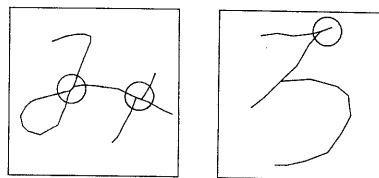
Fig. 4 3×3のマスク

3.1.2 セグメントの抽出

まず、文字画像のノードとノード間及びノードと端点間を結ぶセグメント、すなわち一つのノードに連結しているすべてのセグメントを抽出する。そして、ノードに連結しているすべてのセグメントを抽出した後、端点と端点を結ぶセグメント、および半濁点の有無の判定を行う。

3.1.3 ノードに連結したセグメントの除去

Fig. 5(a)のような変形交点及び同図(b)のようなヒゲは、変形セグメントによるものとして、これらを除去する。まず、抽出したセグメントの長さを求め、これが3以下のときはセグメント数と、ノード数（変形交点の場合）・端点数（ヒゲの場合）を除去する。



(a) 変形交点 (b) ヒゲ

Fig. 5 セグメントの除去

3.1.4 セグメントの分割

Fig. 6のように、抽出されたセグメントが曲がっている場合は、そこに屈折点が存在するとして、一つのセグメントを複数のセグメントに分割する。

まず、セグメントの長さを調べ、長さが4以上の場合、セグメント上の画素 S_i と S_{i+4} を結ぶ線分と S_i と S_{i-4} を結ぶ線分のなす角 θ を求める (Fig. 7)。

もし $\theta \leq \lambda$ ($\lambda = 90^\circ$ はセグメントを分割するときの角度のしきい値) ならば、画素 S_i の位置でセグメントを分割する。

セグメントの分割を行うと、片仮名の「フ」と「コ」と「ロ」などのよく似た文字に対して、文字の特徴をよく抽出し、誤認識が避けられる。

また、この屈折点によって分割されたセグメントは二つと数えるが、屈折点はノード数には数えない。

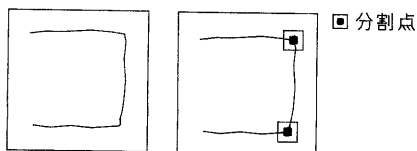


Fig. 6 セグメントの分割

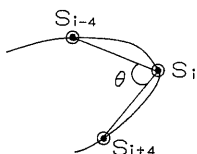


Fig. 7 セグメント分割の手法

3.2 文字の分割構造特徴

3.2.1 分割端点, ノード, 画素数の抽出

3.1.1で抽出された端点数とノード数をそれぞれの領域 (A, B, C, D) に分けておく。また、画素数の抽出も行う。画素数とは画素値1の数であり、端点数、ノード数および文字線の長さを加算したものである。

3.2.2 交差回数の算出

各部分の縦横中央に直線を引き、文字線との交差回数を算出する (Fig. 8)。また、文字線がこの直線と重なる場合は、交差回数を1回とする。

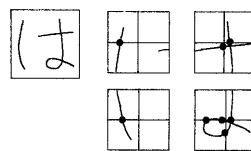


Fig. 8 交差回数の算出

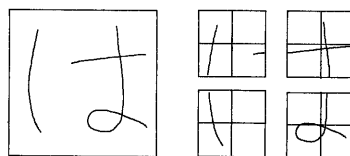
4. データ構造

前章で抽出された文字の構造特徴の探索が容易なように、文字全体の構造特徴と分割構造特徴を分割して登録することにする。

文字全体の構造特徴として、端点数H 1、ノード数H 2、セグメント数H 3、文字線の長さH 4、半濁点数H 5、および文字コードH 6を文字の全体構造特徴ファイルに登録する。ここで、文字線の長さとは、1文字のもつセグメント長を全部加算したもので画素数で表す。また、分割構造特徴として、分割端点数A 1、分割ノード数A 2、画素数A 3、および縦方向の交差回数A 4、横方向の交差回数A 5をそれぞれの領域 (A, B, C, D) に分けて分割構造特徴ファイルに登録する。

文字の全体構造特徴と分割構造特徴のデータファイルの一例をFig. 9に示す。同図の最上段(0)が全体構造特徴ファイルで、同図(1)(2)(3)(4)が分割構造特徴ファイルである。

理解を容易にするために、平仮名「は」の登録データの具体的な値を示す。全体構造特徴(0)では、「は」の端点は6個あるからH 1 = 6、



(b) 特徴量

(0) H 1, H 2, H 3, H 4, H 5, H 6

(1) A 1, A 2, A 3, A 4, A 5

(2) B 1, B 2, B 3, B 4, B 5

(3) C 1, C 2, C 3, C 4, C 5

(4) D 1, D 2, D 3, D 4, D 5

(a) データ構造

Fig. 9 特徴量とデータ構造

ノード数は $H_2 = 2$ 、セグメント数は $H_3 = 8$ 、半濁点数は $H_4 = 0$ 、さらに、文字線の長さが H_5 、および文字コードが H_6 に格納される。また、分割構造特徴(1) (分割領域A)では、分割端点数は $A_1 = 2$ 、分割ノード数は $A_2 = 0$ 、画素数 A_3 、縦方向の交差回数は $A_4 = 0$ および横方向の交差回数は $A_5 = 1$ が格納される。以下、各領域(B, C, D)の構造特徴が格納される。

したがって、辞書の容量は、端点数、ノード数、セグメント数、半濁点数、文字線の長さ、交差回数がそれぞれ1バイトで表せ、文字コードに2バイト要するが、142文字に対して必要なメモリ容量が約4kバイトと非常に少なくなる。また、処理時間については、辞書とのアクセスを1回にしているので非常に短くなる。

5. 判定式

前述したように、文字照合の処理量を減らすために、文字の認識は2段階に分けて行う。

5.1 第1段階

第1段階では、入力文字の全体的な構造特徴に近い登録文字を、近い順に登録ファイルから20文字選び出す。そのための基準を「全体構造特徴の近似度」 HS と呼ぶことにし、 HS を式(2)で定義する。

$$HS = \sum_{i=1}^3 \{ |T_i - t_i| \times \epsilon_1 + |L - l| + |M - m| \times \epsilon_2 \} \quad (2)$$

但し、 T_i と t_i は入力および登録文字のそれぞれの端点数、ノード数、セグメント数(順に $i=1, 2, 3$)であり、 L と l は文字線の長さ、 M と m は半濁点数である。

ここで、 ϵ_1 と ϵ_2 はそれぞれの特徴量の差をほぼ同じ大きさにするため決められた重みで、実験から $\epsilon_1 = 6$ 、 $\epsilon_2 = 50$ が適当であった。

また、 HS は同一文字なら最小となり、異なる文字なら大きな値となる。選出はこの値の少ない文字から選出される。

5.2 第2段階

第2段階では、第1段階で選び出した20文字の分割構造特徴を入力文字の分割構造特徴とマッチングさせる。マッチングの良さは次の基

準により決定する。

今、分割した一つの部分の構造特徴量それぞれに対して、式(3)を用いて得られる「分割構造特徴の近似度」 BS の値を求める。

$$BS = \sum_{i=1}^4 \{ |N_i - n_i| \times \epsilon_3 + |G - g| \} \quad (3)$$

但し、 N_i と n_i は入力および登録文字のそれぞれの端点数、ノード数、縦方向の交差回数、横方向の交差回数(順に $i=1, 2, 3, 4$)であり、 G と g は画素数である。また上述と同様に、 ϵ_3 は特徴量の差をほぼ同じ大きさにするため決められた重みで、実験から $\epsilon_3 = 4$ が適当であった。

さらに、 BS の値を各領域それぞれについて算出し、「分割構造特徴の近似度の和」 PS を式(4)で定義する。

$$PS = \sum_{i=1}^4 BS_i \quad (4)$$

認識には、全体構造特徴の近似度 HS と分割構造特徴の近似度 PS を用い、「入力文字の近似度」 SS を式(5)により決定する。

$$SS = HS + PS \quad (5)$$

式(5)は入力文字と登録文字の構造特徴がマッチングすると最小となる。これによって、入力文字に最も近い登録文字が選出できる。

6. 実験結果及び考察

本手法の有効性を確認するために、手書き文字142字(平仮名,片仮名)に対して、認識実験を行った。被験者は18歳から26歳までの20人(男10人,女10人)であり、データ収集時に特に使用目的等と言わなかった。

まず、先の20人の中から3人を選び出し、文字全体の構造特徴のみによる分類実験を行った(累積分類率)。累積分類率とは例えば20

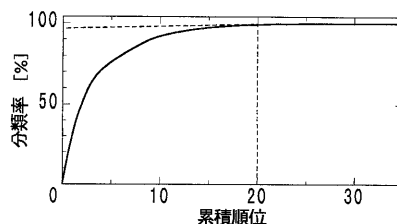


Fig. 10 累積分類率の変化

位の場合、文字の全体構造特徴を用いて、入力文字の構造特徴に近い20個の登録文字を選び出すとき、入力文字がこの20個の中に選ばれる率である。Fig. 10に累積順位による全文字に対する分類率の平均値の変化を示す。Fig. 10から本論文では20位以上をほぼ一定と見なし、第1段階で20個選び出すことにした。

次に、被験者20人に対して、本手法を用いた認識実験を行った。Table 1に20位累積分類率の結果を示す。20位までの累積分類率は、平仮名で97.89%、片仮名で98.67%、平均では98.28%となった。

Table 1 20位累積分類率 [%]

	男	女	平均
平仮名	97.89	97.89	97.89
片仮名	98.18	99.15	98.67
平均	98.04	98.52	98.28

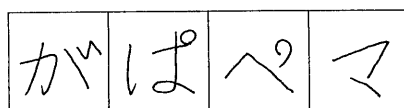
さらに、認識実験結果をTable 2に示す。3位累積認識率は、入力文字が認識結果の中で3個までに選ばれる率である。また、正読率は、入力文字と一致する登録文字が正しく選び出された率である。正読率は、平仮名で95.57%、片仮名で97.33%、平均では96.45%となった。

Table 2 認識実験結果 [%]

		男	女	平均
平仮名	3位累積認識率	97.18	97.05	97.11
	正読率	96.34	94.79	95.57
片仮名	3位累積認識率	97.89	98.03	97.96
	正読率	96.76	97.89	97.33
平均	3位累積認識率	97.54	97.54	97.54
	正読率	96.55	96.34	96.45

Fig. 11に誤認識の例を示す。誤認識は「が」や「ば」などの濁点によるもの、および「ば」などの半濁点によるものが多かった。濁点の場合は、セグメント除去により特徴量が変わってくることで、および濁点がつながっていることによるもので、半濁点の場合は、きちんと半濁点がかかれていないことによるものであった。また、特に片仮名の場合、「マ」や「ム」などで文字線が交差すべきところ(分岐点)で交差していない場合に誤認識が多少あった。しかし、本手法では大部分が認識できた。

したがって、本手法は手書き文字に対して有効であると言える。



(a) が (b) ば (c) ペ (d) マ

Fig. 11 誤認識の例

7. むすび

本論文では、パーソナルコンピュータ(PC-9801RA)を使って、手書き文字のパターン認識に構造解析法を用いた2段階の認識法を提案し、被験者20人に対して、認識実験を行った。本手法は142文字(平仮名、片仮名)に対して、平均96.45%の高い認識率が得られ、認識に要する時間も1文字当たり約1秒で達成された。よって、本手法は手書き文字に対して有効であると言える。

今後の課題は、本手法を漢字にも適応できるように改良すること、および文章から文字の読み取れるシステムを構築することである。

<参考文献>

- (1) 信学誌, Vol. 75, 6, pp557-690(1992).
- (2) 大谷, 塩: "情景画像からの文字パターン抽出と認識", 信学論(D), J71-D, 6, PP.1037-1047(1988).
- (3) 大和, 宮脇, 畑, 賀: "日本語印刷文書の読取りシステム", 信学論(D II), J75-D-II, 2, PP. 257-266(1992).
- (4) 朱, 山内, 神保, 梅野: "階層的ニューラルネットワークを用いた手書き文字認識", 信学論(D II), J73-D-II, 1, PP. 54-61(1990).
- (5) 森: "PDPモデルによる手書き漢字認識", 信学論(D II), J73-D-II, 8, PP. 1268-1274(1990).
- (6) 横井, 鳥脇, 福村: "標本化された2値図形のトポロジカルな性質について", 信学論(D), 56-D, 11, PP. 662-669(1973).