

P形フーリエ記述子を用いたオンライン手書き漢字認識

大友 照彦[†] 原 健一[†]

一筆書きされた手書き漢字をP形フーリエ記述子で表現すると、記述子の低域部分にはパターンクラスに関する情報が集約されているので、標準パターンとの重み付き類似度により手書き漢字の認識が可能である。本論文では、教育漢字543字種を認識対象として、3種類のP形フーリエ記述子に基づく認識法を提案している。手書き漢字を一筆書きの折線図形と考えると、(1)漢字全体、(2)水平方向、(3)垂直方向の筆点運動波形に対してP形フーリエ記述子を適用して得られる3種類のパワースペクトルを特徴として用いる。認識処理に先立ち、認識精度の向上と計算時間の短縮を目的として、正規化した手書き漢字の3種類の筆点運動波形の長さの分布、第1ストロークと第2ストロークの開始点間の方向と距離の分布および画数を調べて、参照する標準漢字パターンを制限する大分類法を導入している。543字種4887パターンの認識実験を行った結果、大分類を行った場合には平均の候補漢字数を6.4に制限することが可能で、クローズ認識率99.57%、オープン認識率98.67%が得られた。

On-Line Recognition of Hand-written Chinese Characters Using P-type Fourier Descriptor

TERUHIKO OHTOMO[†] and KEN-ICHI HARA[†]

This paper describes a method for the recognition of hand-written Chinese characters using three P-type Fourier descriptors. Three different waveforms are obtained by dividing each segment-line on the writing locus of each hand-written Chinese character into the vertical and horizontal components. Then, a total curvature function can be derived from each waveform. A P-type Fourier descriptor can be calculated by the Fourier transform of each total curvature function. In preprocessing each hand-written Chinese character, three features, such as the length of its writing locus, the direction from the starting point of the first stroke to that of the second stroke and the number of its strokes, are used for selecting some standard Chinese characters which should be compared to it. A total sum of the weighted difference of each Fourier coefficient between a character and a standard character is calculated. A hand-written Chinese character can be recognized by selecting the minimal sum out of the total sums for selected standard characters. A recognition rate of 98.67% was achieved for 4887 testing patterns.

1. はじめに

日本語情報処理の拡大に伴い、簡便な漢字入力の必要性が高まっている。このために、文字を手で書く過程で、ペンの運動にしたがって文字の軌跡情報を逐次取り込み、入力文字を認識するオンライン手書き文字認識の研究が活発に行われている¹⁾⁻⁹⁾。

オンライン手書き文字および手書き漢字認識の研究における基本的手法¹⁾は、(1)パターンマッチングによる方法、(2)構造解析の方法、(3)周波数領域での特徴抽出法、等に分類される。(1)はストロークの位置座標を特徴に用いて辞書とのマッチングをとるも

の^{2),3)}、(2)は入力文字の各ストロークをあらかじめ定めた複数の基本ストロークの1つで近似し、基本ストロークの系列を特徴として認識を行うもの⁴⁾⁻⁶⁾である。これらの方式では、基本ストロークの決定やストローク間の分離に問題があり、また、画数の大きい漢字に対しては、記憶容量や計算時間が共に増大する等の問題点がある。一方、(3)の周波数領域での特徴抽出法では、手書き文字の周辺分布のスペクトルを用いる方法⁷⁾や各ストロークごとの筆点座標の左右方向と上下方向の運動波形をフーリエ級数展開して得られる特徴を用いる方法⁸⁾が提案されている。さらに、後者の方法では特徴をファジー化データで表して認識をファジー推論により行う方法⁹⁾も検討されている。

上記の各認識手法では、いずれも基本ストロークの安定な分離、決定が必要であり、行書体や草書体で書

[†] 山形大学工学部共通講座
Department of Basic Technology, Faculty of
Engineering, Yamagata University

かれた漢字に対する対応が難しい。そこで、著者らはストロークの抽出が基本的には不必要で、崩し字などにも対応可能と考えられる認識手法として、上坂が提案したP形フーリエ記述子¹⁰⁾を用いた新しい手書き漢字認識法¹¹⁾を提案した。本手法では一筆書きされた漢字全体の曲線の曲がり方をP形フーリエ記述子で表現して特徴とするので個々のストロークの分離抽出が不必要であり、さらにフーリエ記述子の低域部分にパターンクラスに関する情報が集約されているので少ない情報量で認識が可能という特徴がある。ただし、前報では認識対象漢字数(36字種)が少なく、字種の増加に起因する問題をさらに検討する必要がある。

本論文では、教育漢字 543 字種を認識対象として、筆点座標から新たに水平方向と垂直方向の筆点運動波形を分離して、3種類の一筆書き図形のP形フーリエ記述子の特徴として用いることを提案する。さらに、手書き漢字認識の前処理段階として、計算時間の短縮を図るために、正規化した手書き漢字の一筆書き図形における3種類の線分の長さの分布、画数および第1ストロークと第2ストロークの開始点間の方向と距離の分布に基づいて、比較する標準漢字パターンの数を制限する大分類法を導入し、候補漢字数の低減効果と認識率に及ぼす影響を調べる。

2. P形フーリエ記述子と特徴パターン

平面上の曲線を周波数領域で記述する方法として、上坂は開曲線にも適用できるP形フーリエ記述子¹⁰⁾を提案した。これは、曲線の全曲率関数を指数関数の指数部に持つ複素数値関数をフーリエ展開して得られる記述子であり、任意の再生曲線の端点が原曲線のそれと常に一致するという優れた特徴を持つ。また、P形フーリエ記述子における低域の係数には、原曲線の概形を示す情報が集約されており、情報の圧縮も可能である。

前報¹¹⁾では、一筆書きされた手書き漢字の折れ線の偏角を用いた全曲率関数のP形フーリエ記述子から得られる低次数のパワースペクトルを手書き漢字の特徴として用いた。しかし、認識対象とする漢字の字種の増加に伴って認識率の低下が考えられる。

本論文で提案する認識法では、折れ線図形で表される手書き漢字の筆点座標系列を複素平面(x 軸が実軸、 y 軸が虚軸)上の点列

$$z(0), z(1), \dots, z(m)$$

で表して、これらの点列から得られる次の3種類の筆

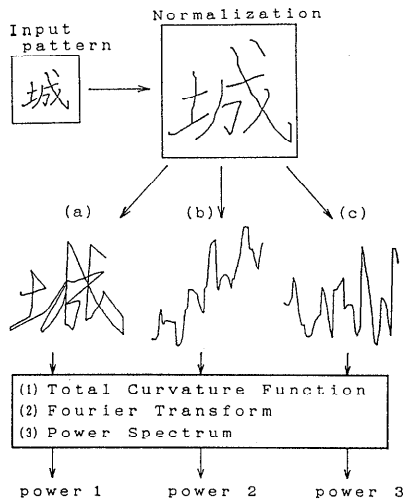


図1 特徴抽出の過程

Fig. 1 Process of feature extraction.

点運動波形のP形フーリエ記述子を手書き漢字の特徴として用いる。

- (1) 漢字全体の筆点運動波形
- (2) 水平方向の筆点運動波形
- (3) 垂直方向の筆点運動波形

図1は特徴抽出過程を示す。入力された手書き漢字の筆点座標の点列は、筆者により筆記速度、筆点の時間的停留などが異なるので、大きく変動する。そこで、前処理によって手書き漢字の大きさや位置を正規化した後、長さ一定($|z(j)-z(j-1)|=\delta$)の筆点座標系列に近似的に変換する。その筆点座標系列を、図1に示すように、手書き漢字を全体の筆点運動波形(a)と、水平方向(b)および垂直方向(c)の移動だけに着目した筆点運動波形に分解する。これらの3種類の波形において、2つのベクトル($z(j)-z(j-1)$)と($z(j+1)-z(j)$)とのなす偏角を $\rho(j)$ とする。ここで、手書き漢字の筆点運動波形におけるすべての偏角の累積関数である全曲率関数¹⁰⁾ $\theta(j)$ を次のように定義する。

$$\begin{cases} \theta(0) = \rho(0) \\ \theta(j) = \theta(j-1) + \rho(j) \quad (j=1, \dots, m-1) \end{cases}$$

この $\theta(j)$ ($j=0, \dots, m-1$)を用いて、複素数値関数 $w(j)$ ($j=0, \dots, m-1$)を

$$w(j) = \exp(i\theta(j))$$

と定義し、手書き漢字のP表現と呼ぶ。

$w(j)$ から定義される関数 $c(k)$

$$c(k) = \frac{1}{m} \sum_{j=0}^{m-1} w(j) \exp\left(-2\pi i \frac{jk}{m}\right) \quad (k=0, 1, \dots, m-1)$$

は $w(j)$ の離散フーリエ変換であり、 $c(k)$ をP形フーリエ記述子と呼ぶ。

3種類の波形のP形フーリエ記述子 $c_f(k)$ ($f=1, 2, 3$) の低域成分における正弦波成分 $a_f(k)$ と余弦波成分 $b_f(k)$ を用いて、次のような n 次のパワースペクトルを定義する。

$$\begin{cases} a_f(k) = 10 \cdot \log(\alpha_f^2(k) + \beta_f^2(k)) \\ \quad (k=0, \dots, n), (f=1, 2, 3) \\ b_f(k) = 10 \cdot \log(\alpha_f^2(m-k) + \beta_f^2(m-k)) \\ \quad (k=1, \dots, n), (f=1, 2, 3) \end{cases}$$

ただし、 $a_f(0)$ はパワースペクトルの直流成分を表す。このような3種類のパワースペクトルをそれぞれパワー1、パワー2、パワー3と略称する。

3. 手書き漢字の認識

3.1 候補漢字の選択

手書き漢字を標準漢字と比較照合する場合、認識対象とする漢字の数が多いほど比較に要する時間が増大し、誤認識も多くなる。そこで、比較照合する候補漢字数を制限する大分類が必要になる。ここでは、3種類のパラメータに基づく大分類法について述べる。

3.1.1 一筆書き漢字の線分の長さ

P形記述子を計算する過程で得られる手書き漢字の筆点運動波形の長さを特徴として、比較すべき標準漢字の数を制限することが考えられる。すなわち、正規化した一筆書き漢字の線分の長さの分布を調べて、認識すべき一筆書き漢字の線分の長さにより候補漢字を選択する。

手書き漢字の3種類のP形記述子を計算する場合、平面上に描かれる手書き漢字の筆点運動座標系において、隣合う筆点座標間の距離の総和、すなわち全長 L_f ($f=1, 2, 3$) を求める。ただし、 $z(k) = x(k) + iy(k)$

$$\begin{aligned} l_1(k) &= |Z(k+1) - Z(k)| \\ l_2(k) &= |x(k+1) - x(k)| \\ l_3(k) &= |y(k+1) - y(k)| \quad (k=0, \dots, m-1) \\ L_f &= l_f(0) + \dots + l_f(m-1) \quad (f=1, 2, 3) \end{aligned}$$

図2は、筆者9名が各々2回ずつ書いた手書き漢字の全長 L_f の分布を示す。実線は手書き漢字を一筆書きした時の筆点運動波形の長さ L_1 の分布、一点鎖線は水平方向の筆点の運動波形の長さ L_2 の分布、点線は垂直方向の筆点運動波形の長さ L_3 の分布を示す。長さ分布の変動幅が大きい漢字は個性が強く現れる漢字と考えられ、候補漢字を制限する効果は少ない。一方、長さ分布の変動幅が小さい漢字は筆者による形状

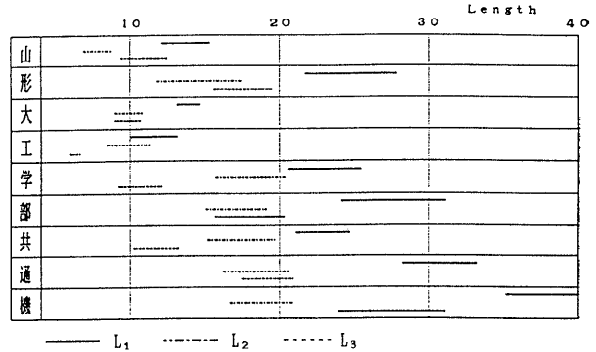


図2 一筆書き漢字の線分の長さの分布
Fig. 2 Distributions of lengths in writing locus of each Chinese character.

変化の少ない漢字であり、候補漢字を制限する効果が高い。図2から分かるように、一筆書き漢字の線分の長さの分布は、漢字の画数に依存して特定の領域に偏在している。例えば、 L_1 が20以上30以下に分布する漢字は形、学、部、共、通の5文字であり、さらに L_3 が13以下の条件を付加すると候補漢字は学、共の2文字に絞ることができる。このように一筆書き漢字の線分の長さの分布を用いると比較すべき候補漢字数を制限することが可能である。

すべての漢字 p ($p=1, \dots, 543$) に対して、それぞれ18パターンの手書き漢字の線分の長さ L_f' ($f=1, 2, 3$) の中で最大および最小の長さ $L_f'_{\max}(p)$, $L_f'_{\min}(p)$ を登録しておく。次に、認識すべき一筆書き漢字の線分の長さ L_f ($f=1, 2, 3$) を求めて、すべての f に対して次式を満たす候補漢字 p を選択する。

$$\begin{aligned} L_f'_{\min}(p) &= 1.25L_f'_{\min}(p) - 0.25L_f'_{\max}(p) \\ L_f'_{\max}(p) &= 1.25L_f'_{\max}(p) - 0.25L_f'_{\min}(p) \\ L_f'_{\min}(p) &< L_f < L_f'_{\max}(p) \quad (f=1, 2, 3) \end{aligned}$$

上式は、図2に示す一筆書き漢字の線分の長さ分布の変動を考慮して、それぞれ1.5倍にした領域を考えて、認識すべき一筆書き漢字の線分の長さ L_f を含む領域をもつ標準漢字をすべて比較の対象とする。

3.1.2 ストローク間の方向

漢字の一部の位相的特徴として、第1ストロークと第2ストロークの開始点間の方向が考えられる。図3(a)は『木』の手書き漢字の例を示す。すなわち、第1ストロークと第2ストロークの開始点間(点a-点b)の距離は D であり、第1ストロークの開始点から第2ストロークの開始点へ方向 H は、図3(b)より⊙の方向で与えられる。方向 H は、 D の大きい漢字群では安定で変動が少ない。しかし、 D が本来0

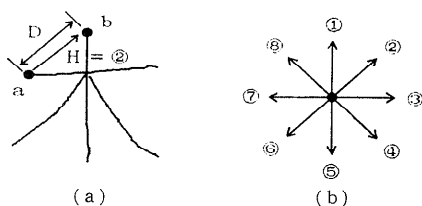


図3 第1ストロークと第2ストロークの開始点間の特徴
 (a) 距離と方向, (b) 八方向
 Fig. 3 Features on the starting points of first and second strokes.
 (a) Distance and direction,
 (b) Eight directions.

の漢字 (たとえば、『日』、『田』、『金』等) でも、手書き漢字では0にはならないことが多く、種々の方向を持つために位相的特徴として方向を利用できない。そこで、方向が利用できる漢字群を決定するために、手書き漢字の距離 D の分布を調べて、すべての手書き漢字を選別し標準辞書とした。

$$D \leq 20.0 \text{ のとき } DL=1$$

$$20.0 < D < 46.0 \text{ のとき } DL=2$$

$$D \geq 46.0 \text{ のとき } DL=3$$

ここで、20.0 および 46.0 は入力領域におけるドット数で、 D の分布から定めたしきい値である。 $DL=1$ は距離が短くて方向の情報を利用できない漢字群であり、 $DL=3$ は距離が長くて方向の情報を正しく利用できる漢字群である。 $DL=2$ は、 $DL=1$ あるいは $DL=3$ の漢字群と明瞭な区別がつかかねる漢字群である。

大分類では、認識しようとする手書き漢字が候補文字に必ず含まれるように制限規則を緩くする必要がある。そこで、認識する手書き漢字の距離 D と方向 H により、次のような標準辞書の利用規則を設けて候補漢字の制限を行う。

規則1 : $D \leq 30.0$ のとき

$DL \neq 3$ のすべての漢字を候補とする。

規則2 : $50.0 < D$ のとき

$DL \neq 1$ かつ $H \pm 1$ のすべての漢字を候補とする。

規則3 : $30.0 < D \leq 50.0$ のとき

すべての漢字を候補とする。

ここで、30.0 および 50.0 の数値は、標準辞書から候補漢字を選択するためのしきい値である。 $DL=3$ の漢字群に選別するときのしきい値は 46.0 であるが、その漢字群の変動を考慮してしきい値を 30.0 とした。また、 $DL=1$ の漢字群のしきい値は 20.0 であ

るが、 D の分布の検討によりその漢字群の変動が大きいと考えられることから安全性を考慮してしきい値を 20.0 から 50.0 に修正した。 $D > 50.0$ のときには、 $DL \neq 1$ の漢字群の中から方向の特徴を利用して候補漢字を選択する。このとき、認識する手書き漢字の方向 H が③の場合には、 $H \pm 1$ すなわち、②、③および④の方向を持つ漢字も候補にする。 D が2つのしきい値の間にあるときには、安全性を考慮してすべての漢字を候補漢字とする。

3.1.3 画数

漢字は、本来固有の画数を持っている。したがって、楷書で書くことを条件にする場合には、画数は候補を制限する重要なパラメータになる。画数によって候補漢字を制限するためには、認識する手書き漢字の画数を調べて、同じ画数を持つ標準漢字を候補漢字にすれば良い。一方、楷書で書くことが条件でない場合でも、行書体で書かれた文字では画数が極端に減少することは少ないと考えられる。そこで、正しい画数に制限した場合と、本来の画数より2画少ない漢字までを候補漢字とした場合について検討する。

3.2 重み付き類似度による認識アルゴリズム

前報¹¹⁾によれば、P形記述子を用いて手書き漢字を表すと、認識対象とする漢字は20~30次程度の再生曲線で我々人間は十分認識が可能であること、高次の係数ほど個性をより強く表していること、などが示された。

P形記述子による手書き漢字の認識において、認識すべき手書き漢字を、大分類によって選択された標準の候補漢字と比較するとき、両者の n 次のパワースペクトルの差の2乗和が最小の標準漢字を認識漢字とする方法が考えられる。

本認識システムでは、28次までのパワースペクトルを用いているが、10次以上のパワースペクトルによる差の2乗和が大きくなる。ところで、高次のP形記述子ほど漢字のクラスに関する情報量が少なくなると考えられるので、認識アルゴリズムを次のように定義する: まず、大分類により標準漢字の中から候補漢字 p を選ぶ。選ばれた候補漢字のP形記述子 $a_r'(k, p)$, $b_r'(k, p)$ と認識すべき手書き漢字のP形記述子 $a_r(k)$, $b_r(k)$ について、それぞれの差の重み付き2乗和を求めて、その和が最小である候補漢字を認識漢字とする。ただし、3種類のパワースペクトルを用いるので、次式に示すようにそれぞれの重み付き類似度の和を最小にする認識法を考える。

$$\min_p \sum_{f=1}^3 \sum_{k=0}^{28} W_k \cdot \{(a_f(k) - a_f'(k, p))^2 + (b_f(k) - b_f'(k, p))^2\}$$

ただし、添字の f はパワースペクトル ($f=1, 2, 3$) を表し、重み W_k は次式で表される。

$$W_k = \exp(-|k|/c_w)$$

3.3 認識システム

手書き漢字認識システムは、パーソナルコンピュータ PC 9801 RA と、タブレットとして機能するメディアグラフから構成される。メディアグラフは漢字入力の道具として利用する。

前処理段階では、位置や大きさの異なる手書き漢字を正規化する。手書き漢字は大きさが 200×200 ドットの正方形の領域内にスタイラスペンで入力するが、その入力パターンを文字の大きさが正方形の大ききの90% になるように縦、横各々異なる割合で拡大する。縦と横の比が0.3以下の文字に対しては縦も横も同一の拡大率で正方形の中央に位置するように拡大する。

手書き漢字認識システムの認識対象として、小学校4年生までに学習する教育漢字543字種を選んだ。手書き漢字は正しい筆順で書くという条件で被験者9名にそれぞれ日を異にして2回書いてもらった。1回目の手書き漢字(543(字種) \times 9=4887)を標準漢字の作成に使用する。さらに、これらの標準漢字はクローズ認識用データとしても使用する。また、2回目の手書き漢字(543(字種) \times 9=4887)はオープン認識用データとして使用する。

543個の標準漢字辞書として、手書き漢字について定義された3種類の長さの最小 $L_f'_{\min}(p)$ と最大 $L_f'_{\max}(p)$ 、画数、第1ストロークと第2ストロークの開始点間の距離 $D'(p)$ と方向 $H'(p)$ が登録される。さらに、入力された筆点座標系列から一筆書き漢字の線分の長さを求める。その系列は筆点間の長さが $\delta(=L/256)$ の筆点座標系列に近似変換されて、3種類のP形記述子が計算される。最後に、それぞれのパワースペクトルの平均 $a_f'(k, p)$ 、 $b_f'(k, p)$ が計算されて、登録される。

認識処理の手順のフローチャートを図4に示す。認識段階では、認識すべき手書き漢字の3種類の線分の長さ L_1, L_2, L_3 、画数および第1ストロークと第2ストローク間の距離と方向が計算されて、大分類による候補漢字の選択が行われる。次に、3種類のパワースペクトルが計算されて、選択された候補漢字の3種類の標準パワースペクトルとの重み付き2乗誤差を最

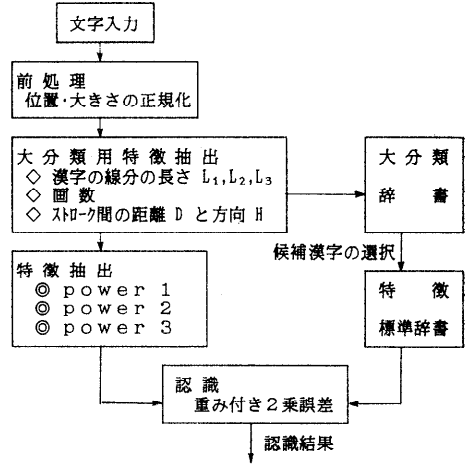


図4 認識処理の手順
Fig. 4 Procedure for recognition processing.

小にする基準により、認識漢字が決定される。

4. 認識実験

4.1 候補漢字の制限の効果

認識対象とする漢字数が多い場合、できるだけ候補漢字を制限することは、計算時間を短縮するために重要である。そこで、前章で述べた3種類の特徴による候補漢字の制限効果を調べる。

表1は、543字種のオープンデータにおいて、長さ L_f の分布により選択された候補漢字数の平均を示す。1種類の長さの分布を単独で用いると、543字種の漢字の1/2強から1/3強に候補漢字を制限することができる。特に、水平方向の筆点運動波形の長さ L_2 の分布を用いると、候補漢字の制限効果が高いことが分かる。さらに、3種類の長さの分布を用いると、約60個の候補漢字にまで制限することができる。

表1 大分類による平均候補漢字数
Table 1 Mean number of candidates of Kanji character due to classifications.
オープンデータ (543字種: 4887パターン)

分類方法	候補数	①共通候補数	②共通候補数	③共通候補数
L_1 分布	195.7	59.71/字	25.97/字	15.64/字
L_2 分布	207.1			
L_3 分布	165.9			
画数+2	126.8			
方向と長さ	298.1			
画数	44.3			

画数で候補漢字を制限すると、表1より543字種に対して候補漢字の平均は44.3個となり、制限効果は非常に高いが、筆記条件は厳しい。一方、行書体で書かれた文字のように認識すべき漢字の画数が2画少なくなっても、画数+2までを候補漢字にすると、候補漢字数の平均は126.8個となる。このような画数の制限をすると、筆記条件は厳しくなくて、候補漢字の制限効果も大きい。そこで、3種類の長さの分布と(画数+2)を併用すると、候補漢字数の平均を約26個まで絞り込むことができる。

第1ストロークの開始点から第2ストロークの開始点へ方向と距離により候補漢字数の制限を行うと、表1より543字種に対して候補漢字数の平均は298個となる。このように、単独では候補漢字の制限効果は小さいが、3種類の長さの分布と(画数+2)を併用すると、候補漢字数の平均は15.64個となり、制限効果は大きい。

4.2 最適重みの決定

表2は、3種類の長さの分布、(画数+2)およびス

表2 重み c_w の認識率に及ぼす影響
Table 2 Effect of weight coefficient c_w on recognition rate.

オープンデータ (4887パターン)

重み c_w	100.0	10.0	6.0	5.0	4.0
正解 (%)	4735 96.89%	4813 98.49%	4821 98.65%	4822 98.67%	4815 98.53%
2位以内 (%)	4824 98.71%	4866 99.57%	4873 99.71%	4874 99.73%	4874 99.73%
3位以内 (%)	4835 98.94%	4875 99.75%	4879 99.84%	4881 99.88%	4880 99.86%

トロークの開始点間の方向と距離により候補漢字を選択し、3種類のパワースペクトルを用いてオープン認識するとき、重み定数 c_w が認識率に及ぼす影響を比較したものである。ここで、正解は第1位で正しく認識される認識漢字数と認識率、第2候補および第3候補は、それぞれ認識漢字が第2位および第3位以上で正しく認識される漢字数とその百分率を表す。表2から、 $c_w=5$ のとき最大認識率 98.67% が得られる。 $c_w=100$ のときには、第28次の重み係数が0.75で、

表3 手書き漢字の認識結果
Table 3 Recognition results of hand-written Kanji patterns.

	大分類					特徴			クローズ認識	オープン認識			
	長さ L_1	長さ L_2	長さ L_3	画数+2	画数	方向	パワ1	パワ2		パワ3	正解(誤り)認識率(%)	第2候補認識率(%)	第3候補認識率(%)
(1)	×	×	×	×	×	×	○	×	×	4487(400) 91.82%	4336(751) 84.63%	4432(455) 90.69%	4541(346) 92.92%
(2)	×	×	×	×	×	×	×	○	×	3588(1299) 73.42%	2911(1976) 59.57%	3481(1406) 71.23%	3766(1233) 77.06%
(3)	×	×	×	×	×	×	×	×	○	3640(1247) 74.48%	3010(1877) 61.59%	3538(1349) 72.40%	3800(1087) 77.76%
(4)	○	×	×	×	×	×	○	×	×	4506(381) 92.20%	4166(721) 85.25%	4455(432) 91.16%	4565(322) 93.41%
(5)	○	○	○	×	×	×	○	×	×	4618(269) 94.50%	4369(518) 89.40%	4628(259) 94.70%	4711(176) 96.40%
(6)	×	×	×	×	×	×	○	○	○	4743(144) 97.05%	4513(374) 92.35%	4606(281) 94.25%	4621(266) 94.56%
(7)	○	○	○	×	×	×	○	○	○	4819(68) 98.61%	4682(205) 95.81%	4789(98) 97.99%	4815(72) 98.53%
(8)	○	○	○	○	×	×	○	○	○	4844(43) 99.12%	4737(150) 96.93%	4829(58) 98.81%	4847(40) 99.18%
(9)	○	○	○	○	×	○	○	○	○	4856(31) 99.37%	4769(118) 97.59%	4838(49) 99.00%	4850(37) 99.24%
(10)	○	○	○	×	○	○	○	×	×	4750(137) 97.20%	4596(291) 94.05%	4762(125) 97.44%	4813(74) 98.49%
(11)	○	○	○	×	○	○	○	○	○	4866(21) 99.57%	4822(65) 98.67%	4874(13) 99.73%	4881(6) 99.88%

重みの効果が小さいので、認識率は最大認識率より1.78%低下する。したがって、本認識実験では重み定数 c_w を5に固定する。

4.3 認識結果と考察

表3は、手書き漢字のクローズ認識およびオープン認識の結果を示す。クローズ認識とは、標準漢字の作成に使用した手書き漢字(543(字種) \times 9=4887)の認識であり、オープン認識とは標準漢字パターンの作成には使用していない手書き漢字(4887)の認識である。また、正解、第2候補および第3候補の意味は、表2の場合と同じである。

候補漢字の制限を行わないで、3種類のパワースペクトルをそれぞれ単独に使用したときのオープン認識率を調べると、パワー1では84.63%で最も高く、パワー2と3は約60%で低い。パワー2と3は筆点運動波形を水平と垂直の運動波形にそれぞれ分解したときのパワースペクトルであるから、単独ではパワー1より情報量が少なく認識率の低下は免れない。しかし、3種類のパワーを併用すると表3(6)のようにオープン認識率は92.35%まで改善される。

パワー1のみを用いて、大分類による認識効果を調べると次のようになる。認識率は、長さ分布 L_1 だけでは85.25%であるが、 L_1 、 L_2 、 L_3 を併用すると89.40%になる。さらに、ストロークの分離抽出を行って、画数と方向も併用すると表3(10)に示すように94.05%まで増加する。このように、大分類を行うことによって見かけ上オープン認識率を9.42%改善することができる。

3種類のパワースペクトルを併用した場合のオープン認識率は、大分類を行わない場合には92.35%であるが、3種類の一筆書き漢字の線分の長さによる大分類を用いると3.46%、また、(画数+2)を併用すると4.58%、さらに方向も加えると5.24%改善される。一方、(画数+2)を正しい画数に変えると6.32%上昇し、98.67%のオープン認識率が得られた。

このように、大分類による3種類の特徴の認識率に及ぼす効果は大きい。これは、P形フーリエ記述子のみでの認識では誤認識される類似漢字が、認識処理前の大分類によって候補漢字群から外れて、結果的に認識率の向上につながったと考えられる。特に、一筆書き漢字の線分の長さの分布による候補漢字数の制限では、一筆書き漢字の線分の長さにストロークとストロークの間の距離も加算されているので、認識すべき漢字が行書体や草書体で書かれた文字等のように各ス

トロークが正確に分離できなくてもある程度適用できると考えられ、その制限効果も大きい。表3(7)に示す認識率95.81%はストロークの分離抽出をしていない特徴データによるものであり、行書体程度の手書き漢字の認識率を見積もる指標になると考えられる。また、第1ストロークの開始点から第2ストロークの開始点への方向は、他の2つの特徴が手書き漢字全体の長さに関する特徴であるのに対して、ストロークの開始点間の方向という局所的な特徴のために、制限された候補漢字をさらに制限する効果がある。今後、グローバルな特徴を捉えるP形フーリエ記述子と共に、局所的な特徴を捉える特徴抽出法も検討していく必要がある。

5. おわりに

手書き漢字を一筆書きの折れ線図形と考えて、その筆点運動波形から得られる3種類のP形フーリエ記述子を用いたオンライン手書き漢字認識アルゴリズムを構成した。さらに、認識対象とする漢字の数の増加に伴う認識率の低下や計算時間の増加に対処するために、3種類の筆点運動波形の長さの分布、漢字の第1ストロークと第2ストロークの開始点間の方向と距離および画数などに基づく大分類法を導入した。認識実験では、543字種の手書き漢字4887パターンに対して、大分類法を用いると候補漢字数の平均は16.14で、クローズ認識では99.57%、オープン認識では98.67%の認識率が得られた。

今後は、JIS第1水準(3036字種)の手書き漢字を認識するシステムを構築する必要があるが、そのためにはグローバルな特徴だけでなく局所的な特徴抽出法も検討する必要がある。また、本手法は、一筆書きされた線図形の概形の特徴を抽出できるので、行書体や草書体で書かれた文字に対する認識にも適用する予定である。

参考文献

- 1) 住久洋子: くずし字など筆記制限を緩和する方向に進むオンライン手書き漢字認識, 日経エレクトロニクス, No. 331, pp. 115-133 (1983.12.5).
- 2) 小高和巳, 荒川弘熙, 増田 功: ストロークの点近似による手書き文字のオンライン認識, 信学論(D), Vol. J 63-D, No. 2, pp. 153-160 (1980).
- 3) 小高和巳, 若原 徹, 増田 功: 筆順に依存しないオンライン手書き文字認識アルゴリズム, 信学論(D), Vol. J 65-D, No. 6, pp. 679-686 (1982).
- 4) 寺井秀一, 中田和男: 手書き漢字・片仮名文字のオンライン実時間認識, 信学論(D), Vol. 56-D,

- No. 5, pp. 312-319 (1973).
- 5) 石井康雄: ストローク代表点に着目したオンライン手書き漢字認識, 信学論 (D), Vol. J 69-D, No. 6, pp. 940-948 (1986).
 - 6) 萬木正義, 永田静男, 小沼賢二, 久保田恵子: 階層分析法によるオンライン文字認識, 信学論 (D), Vol. J 68-D, No. 6, pp. 1320-1327 (1985).
 - 7) 中野康明, 中田和男, 中島 晃: 周辺分布とそのスペクトルによる漢字認識の改良, 信学論 (D), Vol. J 57-D, No. 1, pp. 15-22 (1974).
 - 8) 荒川弘熙, 増田 功: 手書き文字のオンライン認識, 信学論 (D), Vol. J 59-D, No. 11, pp. 809-816 (1980).
 - 9) 大森健児: フェジー推論による実時間手書き漢字認識, 信学論 (D-II), Vol. J 72-D-II, No. 3, pp. 369-379 (1989).
 - 10) 上坂吉則: 開曲線にも適用できる新しいフーリエ記述子, 信学論 (A), Vol. J 67-A, No. 3, pp. 166-173 (1984).
 - 11) 大友照彦, 原 健一: 線図形の曲がりの特徴としたオンライン手書き漢字認識, 信学論 (D-II), Vol. J 73-D-II, No. 4, pp. 519-525 (1990).

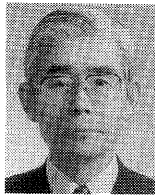
(平成3年11月8日受付)

(平成4年11月12日採録)



大友 照彦 (正会員)

昭和18年生。昭和46年山形大学大学院工学研究科修士課程修了。工学博士。山形大学工学部共通講座教授。手書き文字および連鎖母音認識、神経回路モデルを用いたパターン情報処理などに関する研究に従事。電子情報通信学会、神経回路学会、計測自動制御学会、日本音響学会各会員。



原 健一

昭和6年生。昭和39年東北大学大学院工学研究科博士課程修了。工学博士。山形大学工学部共通講座教授。脳の情報処理、パターン認識の研究に従事。電子情報通信学会、神経回路学会各会員。