

ファジー推論による手書き文字認識

5W-8

高橋敏明 大森健児

法政大学工学部

1. はじめに

手書き文字認識におけるシステムの評価基準は、認識速度と認識率である。本論文では、フーリエ変換を用いて文字のストローク毎の特徴(重心の位置, 始点と終点の離れ具合, 曲がり具合)を抽出し, ファジー推論を用いて手書き文字の曖昧さを吸収することにより, 少ない計算量で高い認識率を達成するオンライン手書き文字認識の方法を示す。

2. システムの構成

ハードウェアは, パーソナルコンピュータPC-9801VM と, タブレットとしてのメディアグラフで構成されている。OSとしては, PC-UNIXを使用し, この上に手書き文字認識のためのソフトウェアを作成した。

ソフトウェアの構成は, 図1に示すとおりであり, ①文字を入力する文字入力部, ②入力された文字を正規化する正規化部, ③正規化された文字データを, ストローク毎に解析し, X軸方向, Y軸方向各々の移動量に対してフーリエ変換を行い, 各周波数の強度(文字の特徴)を求めるフーリエ変換部, ④フーリエ変換により得られた各周波数の強度を曖昧なデータとして扱えるようにするためのファジー化部, ⑤基準文字のファジー化されたデータを記憶しておく基準文字部, ⑥基準文字部よりファジー化データを得て, 認識のためのプロダクションルールを生成するルール生成部, ⑦ファジー推論によって, 入力文字の認識を行なうファジー推論部, ⑧推論結果を出力する候補出力部からなる。

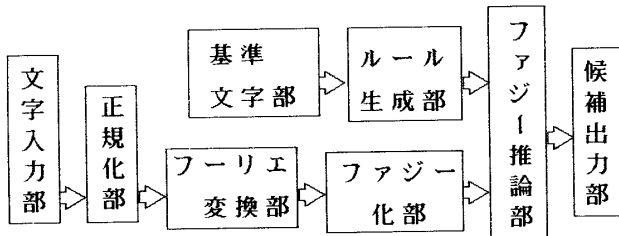


図1 手書き文字認識システム

3. 文字の入力

文字の入力はタブレットを使用して行なう。入力を行なう場合の条件は, 画数について正しく入力することだけであり, 大きさや書き順については制限をもうけていない。これは, 画数は, 「こぎとへん」や「しんによう」のように, 少数の間違いやすいものにさえ注意を与えておけば正しく入力できるのに対して, 書き順は, 書く人によってかなり変化に富んでおり, 制限をもうけると入力者に対して大変な負担となり, 本システムを実用的でないものになってしまうからである。

4. 正規化

入力された文字は, 大きさと筆速が異なる。そこで, 定められた大きさ(256×256)に正規化し, 単位時間に書かれる長さが一定となるように, ストロークの座標点列を修正する。特に, 大きさの正規化に関しては, 入力文字と基

準文字との縦横の長さの比のずれを吸収するために, 縦方向と横方向の拡大縮小を別々に行なう。

5. フーリエ変換

フーリエ変換は, ストローク毎に, X軸での移動量とY軸での移動量の各々について行なう。但し, 移動量そのままの波形に対してフーリエ変換を行なうと, 始点と終点の値が一致していないために, 非連続な波形についてのフーリエ級数を求めることになり, 収束率の悪いフーリエ級数になってしまう。そこで, 終点の位置で線対称に波形を折り返して波形を連続にし, それに対してフーリエ変換を行なう。

フーリエ変換により,

$$f(t) = a_0/2 + a_1 \cos t + b_1 \sin t + a_2 \cos 2t + b_2 \sin 2t + \dots$$

の各係数を得ることができる。数々のストロークにおける各係数を考察した結果, ストロークの特徴をよく示しているのは, $a_0/2, a_1, a_2$ であり, それぞれ, 重心の位置, 始点と終点の離れ具合, 曲がり具合を示しているといえる。ここでは, この3つの値とストローク長を用いて文字の認識を行なっている。但し, $a_0/2$ については絶対的な値ではなく, 前のストロークの $a_0/2$ からの相対的な値を用いている。これは, あるストロークの位置のずれが, それ以降のストロークの位置に影響するのを防ぐためである。

6. ファジー表現

ファジー表現とは, 私たちが日頃使っている, 「非常に長い」とか「極めて短い」というような曖昧な表現であり, その言葉の意味にある程度の幅をもっている。

手書き文字は, 同じ人が入力したとしても, ストロークの長さや周波数の強度が, 毎回異なっている。そこで, これらのデータを曖昧なものとして, 各々16のファジー値を用いて表わす。

ファジー理論では, ファジー値の中に含まれる可能性を, メンバシップ値という。また, ファジー値とメンバシップ値の関係は, メンバシップ関数で表わす。本システムでは, メンバシップ関数を, 図2に示す三角形で表わす。

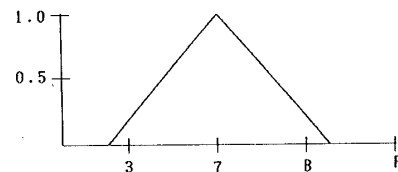


図2 メンバシップ関数

7. ルール生成

基準文字部よりファジー化データを受け取り, これにより, 各々の基準文字に対してのプロダクションルールを, ストローク毎に生成する。例として, 文字「疑」の第一ストロークについてのプロダクションルールを以下に示す。

- ルール疑1 第一ストロークにおいて, ストローク長が相当に短く,
 X軸の移動量で見たとき
 ストロークの重心が左端に相当に接近していて,
 終点が始点に対して右に相当に接近していて,
 曲がり具合は水平で,
 Y軸の移動量で見たとき
 ストロークの重心が上端に非常に接近していて,
 終点が始点に対して下に相当に接近していて,
 曲がり具合は垂直ならば,
 この文字は疑である。

8. ファジー推論

ファジー推論部では、ルール生成部により生成されたプロダクションルールに、入力文字のファジー化データをあてはめ、入力文字の基準文字に対する確信度（条件の満たされ具合）を計算する。ここでは、条件の確信度は基準文字のメンバーシップ関数と入力文字のメンバーシップ関数のmin-max、つまり、図3に示すように、2つのメンバーシップ関数が与えるメンバーシップ値のうち、小さい方を取り、その中で最大のものである。また、条件の確信度の論理積はminであるとする。また、結論の確信度は、条件の確信度とし、同じ結論が複数個存在する場合には、それらの平均ということにする。

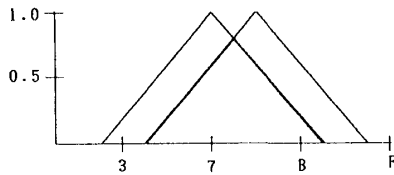


図3 メンバーシップ関数間の演算

9. DPマッチング

文字の入力の章で述べたように、文字の書き順は、書く人によってかなり変化に富んでいる。そればかりではなく、ストロークを書く方向が逆になっている場合もある。そこで、基準文字と入力文字を比較する場合、ストロークの順番の並べ替えが必要になる。また、ストロークを書く方向についても同様である。本システムでは、DP（ダイナミックプログラミング）を用いて、基準文字と入力文字のマッチングを行なっている。

10. 実験結果

認識実験は、3画、4画、5画、9画、14画の教育漢字全てについて、各々8種類の基準文字を用意し、①入力文字と基準文字の入力者が同一の場合、②入力文字と基準文字の入力者が異なる場合、③入力文字と基準文字の入力者が異なり、5種類の基準文字と比較した場合について、3人の入力データ（KN, TT, YS）に対して行なった。

その結果、最も認識率の低かったのが3画であり、画数が多くなるほど認識率が高くなっている。これは、ストロークの数が多いほど、文字の特徴が多いことによるものである。それ以外のことについては、全ての画数において、同様な認識結果が得られたので、ここでは、代表として3画（全24文字）についての認識結果を、以下の表とグラフ①～③に示す。但し、グラフについては、入力者TTの認識結果を代表として示す。表には、第1候補から第3候補までの累積認識率、正しい認識の確信度の平均と標準偏差、誤った認識の中で最大の確信度の平均と標準偏差を示す。グラフには、正しい認識の確信度の分布と、誤った認識の中で最大の確信度の分布を示す。

これらの結果を見ると、入力文字と基準文字の入力者が違う場合には、字体の違いによって認識率は低くなっている。しかし、5種類の基準文字と比較した場合には、入力文字と基準文字の入力者が違うにも関わらず、認識率のいちばん低かった3画でさえも、全て100%の認識率が得られている。

11. おわりに

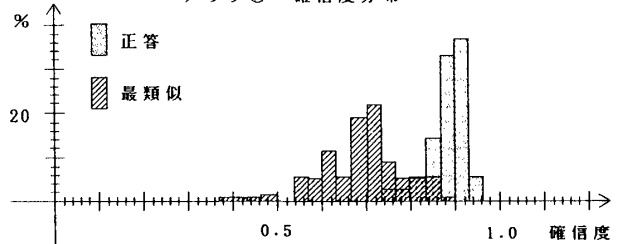
本論文では、フーリエ変換により文字の特徴を抽出し、ファジー推論を用いて手書き文字認識を行なう方法を示した。フーリエ変換により、文字の特徴を少数の値でうまく抽出できたこと、ファジー推論により、手書き文字の曖昧さを吸収できたことにより、少ない計算量で、高い認識率を達成できた。

①入力文字と基準文字の入力者が同一の場合の認識結果

表①

| 入力 デー タ | 累積認識率 | | | 正答の確信度 | | 最類似の確信度 | |
|---------------|-------|-------|-------|--------|-------|---------|-------|
| | 1位 | 2位 | 3位 | 平均値 | 標準偏差 | 平均値 | 標準偏差 |
| KN | 98.3 | 100.0 | 100.0 | 0.868 | 0.042 | 0.676 | 0.091 |
| TT | 100.0 | 100.0 | 100.0 | 0.869 | 0.042 | 0.672 | 0.093 |
| YS | 100.0 | 100.0 | 100.0 | 0.906 | 0.024 | 0.691 | 0.096 |

グラフ① 確信度分布

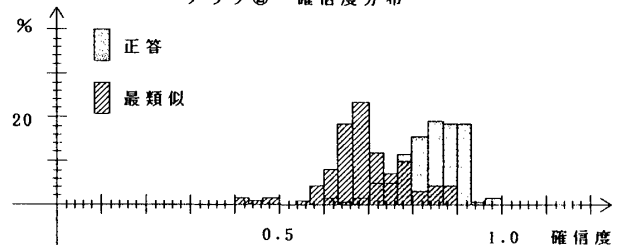


②入力文字と基準文字の入力者が異なる場合の認識結果

表②

| 入力 デー タ | 累積認識率 | | | 正答の確信度 | | 最類似の確信度 | |
|---------------|-------|-------|-------|--------|-------|---------|-------|
| | 1位 | 2位 | 3位 | 平均値 | 標準偏差 | 平均値 | 標準偏差 |
| KN | 93.3 | 94.1 | 96.6 | 0.804 | 0.080 | 0.667 | 0.096 |
| TT | 98.3 | 99.1 | 100.0 | 0.822 | 0.071 | 0.679 | 0.090 |
| YS | 96.7 | 100.0 | 100.0 | 0.845 | 0.057 | 0.681 | 0.100 |

グラフ② 確信度分布

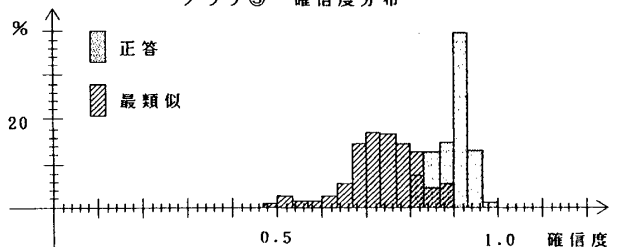


③入力文字と基準文字の入力者が異なり、5種類の基準文字と比較した場合の認識結果

表③

| 入力 デー タ | 累積認識率 | | | 正答の確信度 | | 最類似の確信度 | |
|---------------|-------|-------|-------|--------|-------|---------|-------|
| | 1位 | 2位 | 3位 | 平均値 | 標準偏差 | 平均値 | 標準偏差 |
| KN | 100.0 | 100.0 | 100.0 | 0.877 | 0.041 | 0.709 | 0.084 |
| TT | 100.0 | 100.0 | 100.0 | 0.881 | 0.038 | 0.720 | 0.083 |
| YS | 100.0 | 100.0 | 100.0 | 0.892 | 0.026 | 0.722 | 0.088 |

グラフ③ 確信度分布



(注. 図表①～③は、3画の教育漢字全24文字の認識結果である)