2N－3

トップダウン情報を用いた手書き漢字の部分構造への分割

白石知之　田中英彦
東京大学大学院工学系研究科

1　はじめに

文字認識の研究は古くから盛んに行われてきた分野である。特にわが国では漢字を対象とした研究を中心に行われてきた。

現在では非常に高精度なものが数多く作られているが、その多くは特徴整合法によるものである。これらは抽出した特徴を辞書内のものと比較し、最も近いものを解としているため、文字を理解しているとは言い難い。今後はそのような方向への質的な転換が必要であろう。

このためには漢字の構造を解析することが必須と考えられるが、英数字と比較して複雑な構造のため、そのままでは解析が困難である。

そこで我々は漢字の持つ階層的文字構造を利用して、トップダウンに構造情報を与えてやることで漢字を基本字根に分解し、構造解析を容易にする手法を提案する。

2　漢字の構造

漢字を分解していくと、一度に一つ一つのストロークではなく、偏や旁といったような、ある程度まとまったブロックに分解することができる。そしてそれらのブロックは、また更に細かいブロックへと分割されていく。

例えば「認」という文字を例に考えると、この文字はまず「言」と「忍」という部分に分割することができ、更にその後「忍」の部分は「刃」と「心」とに分割することができる。

このように漢字の構造は、基本構成要素の階層的な組合せによって成り立っていると見ることができる。

この要素の組合せは、以下に示す四種類に大別することができる。

上下型　要素が上下に並んで構成されているもの。
左右型　要素が左右に並んで構成されているもの。
外内型　一方の要素の内側に他方が入っているもの。
単体型　それ自身で構成されるもの。

したがって、単純な分割を再帰的に繰り返すことによって、基本構成要素に分割してやることか可能となる。

この性質を利用することで、分割を容易に行なう手法を提案する。

3　構造情報を利用した分割

漢字を部首程度の大きさのブロックに分割することで、それぞれの構造は全体と比較して単純なものとなるため、以降の解析を容易にすることが期待される。

しかし手書き漢字を対象とした場合、予期せぬ部分での接触や切断が考えられるため、字形情報を元にしたトポロジーの利用だけでは困難である。

そこで漢字の構造情報をトップダウンに与えることによって、トポロジーによらぬ分割を行う。

漢字の構造は前節で説明したような接続の階層的な組み合わせになっているため、分割の際はそれらの単純な接続を一つ一つ分割するだけで良い。

分割は以下の手順で行なう。

1. トップダウン情報から切断線を仮定する
2. 切断線に垂直な方向成分を入力図形から抽出する
3. 抽出した方向成分をストロークとみなし、中点ほど高く、端に近いほど低くなるように傾斜をかけたペナルティを設定する
4. 切断線に沿ってペナルティを加算し、ヒストグラムを作成する

ここで作成したペナルティの分布は、図1に示すようになる。例は漢字「認」のものである。

Segmentation of Handwritten Kanji Characters into Basic Parts.
Tomoyuki Shiraishi, Hidehiko Tanaka
Faculty of Engineering, University of Tokyo
{tshira,tanaka}@mtl.t.u-tokyo.ac.jp

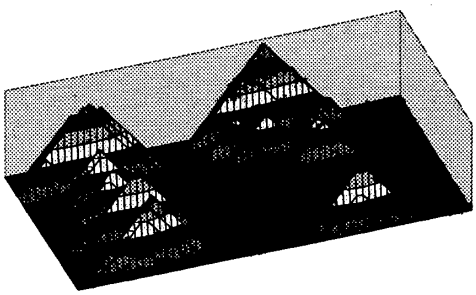

図 1: ペナルティの分布

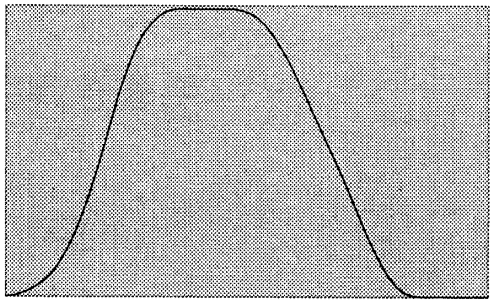

図 2: 判別基準値による分割

そこから作成したペナルティのヒストグラムから、判別基準法によって分割線の位置を決定する。

分割線によってヒストグラムが 2 つに分割されたとすると、それぞれの領域での生起確率 P_n、分散 σ_n、平均μ_n としたとき、

クラス内分散 $\sigma_W^2 = P_1\sigma_1^2 + P_2\sigma_2^2$
クラス間分散 $\sigma_B^2 = P_1(\mu_1 - \mu_T)^2 + P_2(\mu_2 - \mu_T)^2$

で定義される値を用いると、クラス間分散が大きく、クラス内分散が小さくなる位置が最良の分割線の位置であると考えることができる。

そこで全体の分散 σ_T が $\sigma_W^2 + \sigma_B^2 = \sigma_T^2$ なる性質を有していることを利用し、$\frac{\sigma_B^2}{\sigma_T^2}$ を最大にする位置を求める。

この解を解析的に求めるのは困難である。そこで対象となる全ての位置について計算を行ない、その最大値を取る位置を解とする。

このとき

$$P_1(k) = P_1(k-1) + \frac{n(k)}{N}, \quad P_2(k) = 1 - P_1(k) \quad (1)$$

$$\mu_1(k) = \frac{P_1(k-1)}{P_1(k)} \cdot \mu_1(k-1) + \frac{n(k)/N}{P_1(k)} \cdot k \quad (2)$$

$$\mu_2(k) = \frac{P_2(k-1)}{P_2(k)} \cdot \mu_2(k-1) - \frac{n(k)/N}{P_1(k)} \cdot k \quad (3)$$

といった漸化式が成り立つため、各位置について最初から計算することなく容易に計算を行うことが可能である。

図1のペナルティ分布から、分割線の位置による判別基準値を求めたものが図2である。

以上の式は 1 本の分割線によって 2 つの領域に分割する際であるが、3 つ以上の領域に分割する際も、クラス内分散、クラス間分散をそれぞれ

$$\sigma_W^2 = \sum_{k=1}^{n} P_k\sigma_k^2 \quad (4)$$

$$\sigma_B^2 = \sum_{k=1}^{n} P_k(\mu_k - \mu_T)^2 \quad (5)$$

と拡張することで、同様に分割位置を求めることが可能である。

4 おわりに

漢字の階層構造を利用して、入力図形に対してトップダウンに構造情報を与えることにより、手書き漢字における予期せぬ部分での接触や切断によらず領域を分割する手法を提案した。

この手法を用いることでそれぞれの領域内のパターンは単純なものとなるので、構造解析が容易に行えるようになることが期待される。

参考文献

[1] 白石知之, 田中英彦. 漢字の階層性に注目した文字認識手法. 情処第 49 回全大, Vol. 2, pp. 213–214, Sep. 1994.

[2] 白石知之, 田中英彦. 漢字の階層構造を利用した文字認識システム. 情処第 50 回全大, Vol. 2, pp. 57–58, Mar. 1995.

[3] 白石知之, 田中英彦. 漢字の階層的構造を用いた部分要素への分割による類似文 字弁別手法. 情処第 51 回全大, Vol. 2, pp. 173–174, Sep. 1995.