

デジタルペンを用いた文字認識の研究

～文字の大きさに左右されない一文字・一行の位置認識～

渡辺博芳研究室 毛塚遼太 田久幸輔

1.はじめに

近年、デジタルペンが様々な分野で活用されつつあるが、文字として認識されてパソコン画面上に映し出されているソフトはない。本研究では、デジタルペンで書かれた文字の認識を目指し、書かれた文字データがどの範囲で一文字、一行なのかを文字の大きさに依存することなく認識させることを目的とする。

2.デジタルペン(アノトペン)の概要

デジタルペンとは、デジタルペン対応用紙に書き込んだ情報を取得するペンを指す[1]。本研究で使用されているデジタルペンはアノト方式と呼ばれるデジタルペンである。記述時に筆圧センサーが作動し、カメラがデジタルペン対応用紙のドットパターン(以降「点」と表記する)を読み込むと、画像処理装置によってペン先の位置情報(座標データ(x:横方向,y:縦方向))と紙ごとに割り振られた固有のIDが取得される。ペンの軌跡以外に、書いた順序、筆圧、記入時刻なども取得することができる。取得した筆跡ストローク情報はメモリ内に蓄積され、ペンアクセスポイントへ挿し込むことでパソコンへ転送される。そのときXML形式とPNG形式のファイルが得られる。

使用例として、各種登録手続き、屋外調査、請求書発行、医療などがある。

3.一文字・一行の位置認識の定義

ここでの一文字位置認識とはデジタルペンで書かれ取得されたデータがどの文字を表しているのかを認識させるのではなく、パソコン画面上に再描画されている文字データの、どの範囲が一文字分になっているのかを認識させるものである。

一方、一行位置認識とは書かれた文字データがどの範囲で一行になっているのかを認識させるものである。

4.処理全体の流れ

デジタルペンで対応用紙に書かれた情報を取得、パソコンに転送し、そのデータ(XMLファイル)をJavaプログラムで読み込む。XMLファイルからストローク情報が座標データ(x,y)で得られ、読み込んだストローク情報から文字のサイズを予測する。予測された文字のサイズを基準に一文字・一行の位置認識をさせ、別窓で書かれた文字と認識された一文字・一行の範囲を表示させる[2]。

5.文字位置認識処理

5.1.文字のサイズの予測

書かれた文字の大きさに依存しない文字位置認識をするために、最初に文字サイズの予測をする。

最初に文字の区別なしで全てのストローク情報の座標データを取得する。その際、ストロークごとにx, yの最小値(x_min, y_min), 最大値(x_max, y_max)を記録し, x, yそれぞれmaxとminの差分(x_Sa, y_Sa)を計算する。計算結果を比較し, x_Sa, y_Saの最大値をその紙に書かれた最大の文字の大きさと予測し, それを一文字, 一行認識の範囲の基準とする。

5.2.一文字の位置認識

ストロークの最初の点から一文字が認識される処理の間、座標データ(x, y)と比較することで, x_minとx_maxを常に更新する。

ストロークごとに最初のxの値が $(x - x_{\min}) > x_{Sa}$ という条件と, $(x - x_{\min}) < x_{Sa}$ だが $(x - x_{\max}) > 5$ という条件, どちらかを満たした時にその前までのストロークが一文字だと認識させる。ここでの5はいろいろな人の文字データから割り出した平均的な文字と文字の幅である。

一文字だと認識されたら次のストロークの最初の点をx_min, x_maxに上書きし, また同じ処理を繰り返して一文字認識をしていく。

5.3. 一行の位置認識

一文字目の y_{\min} をその行の基準とし、ストロークごとに最初の点と y_{\min} が $(y - y_{\min}) > y_{\text{Sa}}$ という条件と、ストローク中の y の値と y_{\min} が $(y - y_{\min}) > y_{\text{Sa}}$ という条件、この 2 つを満たしたときにだけ、一行だと認識させている。

実行結果を図 1 に示す。

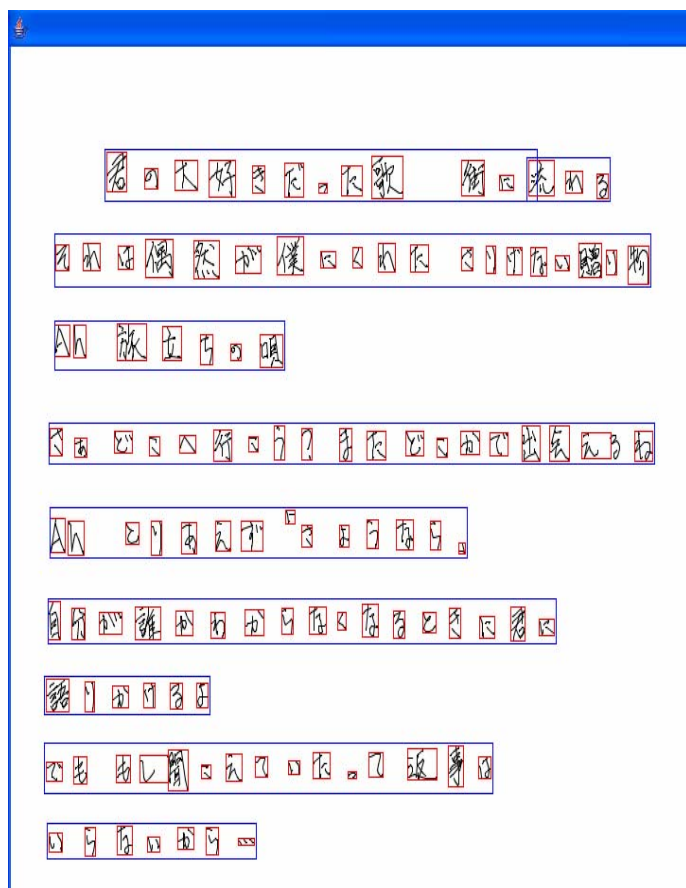


図 1 実行結果の画面

6. 認識率実験

6.1. 方法と結果

Mr.Children の旅立ちの唄という歌の歌詞の一部を抜粋し全 119 文字、7 行の平仮名、漢字、記号、英字で構成される文章を 18 人に横書きで書いてもらった。

集計方法は文字・行共に、正確に認識できた 18 人分の文字数(行数)を認識されるべき 18 人分の文字数 2142 文字、行数 126 行で割って認識率を出す。

その結果、文字認識は、1601 文字認識、認識率は 75%。

行認識率は、96 行認識、認識率 76%だった。

6.2. 考察

このプログラムは簡単な平仮名の文字データのみを使い作成したものであるため、漢字の認識率が少々悪い。

しかし、字のバランスが悪いと平仮名でも 2 文字と認識してしまう場合が多かった。例えば、濁点などが付いた文字の濁点と文字を離しすぎて書いてしまうと、濁点が別の文字として認識されてしまう。漢字の認識にも同じことが言える。

また、文字と文字の間が空いてないとプログラムの関係上認識できない。

行は、文字が大きく曲がって書かれていない限り正確に認識できていたが、図 1 の一行目を見て分かる通り、曲がっていないのに一行と認識されてしまう場合があった。この原因については、ストロークごとに行の判断をしているので同じ行でも下の方に書かれたストロークだと行認識の条件を満たしてしまっている。

7. おわりに

本研究で、文字の大きさに左右されずに一文字・一行の位置認識をするプログラムの作成を試みた。本プログラムでは、最大ストロークのサイズと文字間の距離を考慮して一文字を認識した。また、最大ストロークのサイズと現在のストロークとの距離を考慮して一行を認識した。

実験結果から一文字認識 75%、行認識 76%だった。

今後の課題は、現在出ている問題改善と認識率を上げるためにより良い認識方法を考えていき、検討することである。

参考文献

- [1] デジタルペン ソリューション：日立製作所，<http://www.hitachi.co.jp/Div/jkk/solution/tegaki/element.html>(2008 年 1 月 15 日アクセス)
- [2] 高田 美樹：改訂新版 Java スタートブック～基礎からしっかり徹底学習～，技術評論社，(2007)