

## (1) 手書き文字認識システムの開発

VBL非常勤研究員 水上嘉樹

### 1. はじめに

本研究では、手書き文字認識システムの開発を目的とする。システムの独自性は、研究担当者の以前の研究により有効性が明らかにされた変位抽出手法を用いるところにある。この技術により、手書き文字と計算機上に記憶されている標準文字の距離を測定する際に問題となる、文字形状の不安定性に対処できる。さらに、システムはGUI(グラフィカル・ユーザー・インターフェース)を装備することで、一般ユーザーに親しみやすい作業環境を提供する。新しい技術である変位抽出手法をGUIを備えたシステムに搭載することで、高精度かつ実用的なアプリケーションを構築する。

対象書類は予め文字枠が指定されていることとする。文字枠が指定された書類例として、銀行や郵便局での記載用紙、一般製品のユーザー登録用紙、各種アンケート用紙などが挙げられる。これらの書類は我々の日常生活に馴染みが深い。

システムは文字枠位置に書き込まれた文字部分を切り出した後に認識を行い、最終的にテキストファイルに変換する。このとき、文字枠位置に関する情報が自動的に得られることは重要である。何故なら、一般に書類上には多数の文字枠が存在し、これらの大きさ、位置を人手により測定することは容易な作業ではないからである。また、得られた認識結果を相互対話的に修正できるような環境を提供する必要がある。これは、変位抽出手法で許容できる形状変動を上回るような記載者独自の書き癖がある場合に、認識結果が必ず正しいことが保証できないことへの対処である。以下、2. で構築したシステムの原理、3. でシステムの操作方法、4. で研究を通して得られた成果を述べる。

### 2. システムの原理

開発したシステムは2つのソフトウェアから構成される。1つは未記入の用紙画像から文字枠の位置と面積(文字枠情報)を自動的に検出する「文字枠検出部」である。2つめは、記入済の原稿画像上の文字枠位置にある手書き文字を切り出し、これを認識する「手書き文字認識部」である。

#### 2.1 文字枠検出部

文字枠の位置と大きさには高い自由度がある。このため、位置および面積について可能性がある範囲内で文字枠を総当たりに探すことは現実的ではない。計算量がそれほど膨大でなく、文字枠を安定して検出できる方法が求められる。

本研究で設計した文字枠検出部は、罫線で囲まれた空白部分を文字枠候補とみなし、その中心位置から長方形領域を徐々に膨張させることで文字枠面積を決定する。これにより比較的低い計算量で、安定した文字枠の検出が期待できる。以下に、文字枠検出部の処理手順(図1)の詳細を述べる。

最初に、未記入の用紙画像をファイルシステムから主記憶上に読み込む(手順1)。続いて、白紙部分と罫線部分を分離するために閾値処理を行う(手順2)。一般に文字枠を指定する罫線は必ずしも実線ではない。つまり、点線などで文字枠が指定してある場合もこれを長方形の枠として安定に検出できなければならない。点線などの不明瞭な罫線を強調すると同時に、雑音的な罫線候補を除去するために、水平および垂直方向に画素明度の適切な領域で平均処理と閾値処理を行う(手順3)。ここでは、罫線は垂直または水平方向にある程度の密度で存在する黒画素で構成される、と仮定している。得られた結果は水平および垂直方向の罫線の所在を示す2枚の罫線画像として表現できる。なお、平均時の計算量を押さえるために簡単な工夫を行っている。水平方向の平均処理を行う際に、既に処理が終わっている横隣

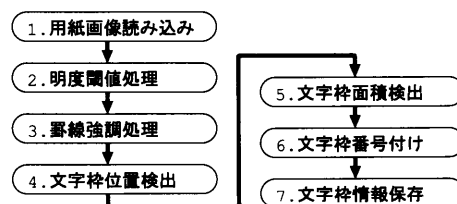


図1. 文字枠抽出部の処理手順



図2. 番号付けられた文字枠

の座標での平均値を継承し、平均対象領域の相違部分である画素明度の影響分だけを加味して注目座標での平均値を決定する。これにより、平均対象となる領域が大きくなっても計算量が変わらないという利点を得られる。ここでの計算量削減法は垂直方向の平均処理においても同様に用いる。

2枚の罫線画像から文字枠中心の候補位置を検出する(手順4)。文字枠の中心位置としての条件は、左右方向の等距離に垂直な罫線および上下方向の等距離に水平な罫線が存在することである。得られた中心位置から文字枠の面積を検出する(手順5)。ここでの考え方は、中心位置周囲の空白領域で最大の長方形領域を決定することである。最初に、中心位置に十分小さな長方形領域を設定する。なお、この時点で長方形領域に罫線が存在する場合は、文字枠の中心位置候補から削除される。次に、徐々に上下左右方向に長方形領域を拡大させる。上下方向の拡大では罫線の水平成分、左右方向の拡大では罫線の垂直成分に重ならない範囲でこの操作は行われる。但し、実際の計算時には安定な空白領域の決定のために、適切な定数に基づく閾値処理を行って長方形領域が罫線と重なっていないかを調べている。以上の操作により、原稿画像上の文字枠の位置および大きさが得られる。

最後に、得られた文字枠の番号付けを行う(手順6)。用紙が横書きで構成されているならば、同じ行内においては左側に位置する文字枠ほど若い番号を割り当て、違う行内では上側に位置する文字枠ほど若い番号を割り当てる必要がある。但し、行数の目安となる文字枠の垂直方向の位置は、実際には微妙に変動して検出されるために、縦方向の位置から文

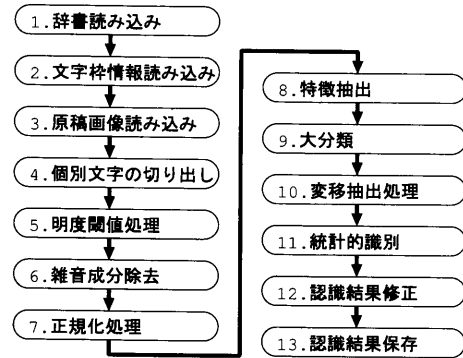


図3. 手書き文字認識部の処理手順

字枠が属する行数を直接的に決定することはできない。そこで、得られた全ての文字枠に対して変則的な並べ替えを用いて番号付けを行う。並べ替えは次の2つの規則に従う。[規則1]注目している2つの文字枠の垂直位置が適切な範囲より離れていれば、上部に位置する文字枠に若い番号を割り当てる。[規則2]垂直位置が適切な範囲より近ければ、左側に位置する文字枠に若い番号を割り当てる。得られた番号の順序に従い、文字枠情報をファイルシステム上に保存する(手順7)。なお、この保存時に手順2の閾値処理において背景として判断されなかった画素領域の位置と用いた閾値についても記録することとした。これらの位置情報は罫線やそれ以外の印刷部分と考えられ、後の手書き文字認識部において記入済の用紙原稿に文字枠位置を重ね合わせる際の相対位置を得るために有用である。以上の手法によって検出された文字枠の例を図2に示す。

## 2.2 手書き文字認識部

手書き文字認識部では、記入済の用紙画像から手書き文字を切り出した後に、これを認識してテキストファイルに変換する。切り出しでは、文字枠検出部で得られた文字枠情報が利用される。切り出し後の一連の認識処理は文献1において検討したものである。以下に、手書き文字認識部の処理手順(図2)の詳細を述べる。

最初に、文字形状の辞書である標準文字データをファイルシステムから主記憶上に読み込む(手順1)。標準文字データとは、事前に利用可能な大量の文字画像に対して特徴抽出処理を行った後に得られる各字種毎の平均と分散である。特徴抽出処理については後述する。続いて、未記入の用紙画像から取

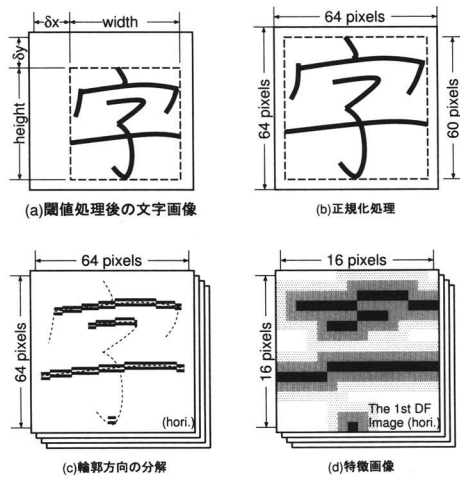


図4. 特徴画像の生成課程

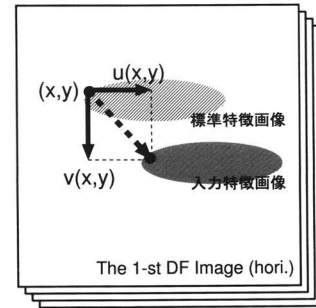
得した文字枠情報と印刷部分の位置情報（手順2）、および、記入済の原稿画像（手順3）を読み込む。

印刷部分の位置情報から文字枠情報に対する原稿画像の相対位置を決定する。すなわち、文字枠情報と原稿画像の間に並進的位置ずれがあることを仮定する。水平および垂直報告の位置ずれ量を調整しながら、印刷部分の位置情報に対応する原稿画像の明度を閾値処理する。最も高い相関を与える位置ずれ量を採用する。この後に、文字枠情報に従い原稿画像上の個々の手書き文字を切り出す（手順4）。

これ以降の処理は、切り出された個々の文字画像に対して独立に行われる。文字部分と背景領域を分離するために、文字画像に対して閾値処理を行う（手順5）。このとき、雑音的に文字部分と近い明度を持つ背景領域が、閾値処理によって誤って文字部分として分類されることがある。このような領域は面積が小さいことが予想されるので、文字部分として分離された領域の中で一定の面積を持たないものは除去する（手順6）。

画像上の文字パターンの大きさおよび位置の正規化処理を行う（手順7）。これにより、文字パターンは60×60画素の領域内で表現され、64×64画素の画像中心に配置される。文字パターンの外接枠を測定し、水平方向または垂直方向のより長い辺が60画素になるように、縦横比を保存した伸縮を施す（図4(a)(b)）。

正規化された文字画像から輪郭方向に着目した特徴を抽出する（手順8）。文字の輪郭を追跡して、水平、右下、垂直、左下の4つの成分に分解する（図4(c)）。これらは64×64画素の4枚の画像として表

図5. 変位関数 ( $u$ ,  $v$ )

現できる。続いて、個々の画像に対して平滑化処理を行い16×16画素の大きさに縮小する。このような次元数の削減は計算量を大幅に節約できるだけでなく、以下に述べる統計的識別器の利用に際しては必修である（文献2）。得られた4枚の平滑化画像を入力特徴画像とよぶ。これに対して、手順1で読み込まれた各字種毎の平均的な特徴画像を以下、標準特徴画像とする。標準特徴画像は手書き文字認識部に登録された字種の数だけ存在する。

既に述べたように、構築する手書き文字システムでは、入力特徴画像と標準特徴画像の間の変動を吸収するために変位抽出処理を導入している。但し、登録された全ての字種に対して変位抽出を行うことは、計算量が非常に大きくなってしまいう問題がある。そこで、変位抽出処理に先立ち、全ての字種の標準特徴画像に対してユークリッド距離を測定し、その上位を候補として選択する大分類処理を行う（手順9）。

変位抽出処理（手順10）は、以下の2つの制約についての最適化過程として考えることができる。[制約1] 標準特徴画像上の各座標において水平および垂直方向の変位関数 ( $u(i,j)$ ,  $v(i,j)$ )（図4）を考慮することで、入力特徴画像と標準特徴画像の間のユークリッド距離を小さくする。[制約2] 水平および垂直方向の変位は、その周辺領域で滑らかに変化する。これらの制約を最適化する変位関数 ( $u,v$ ) は、コスト汎関数の最小化問題から導かれるガウス-サイデル型漸化式（文献3）を用いて計算される。また、計算量の低減と局所解の回避のために事前に低解像度の画像に対して変位関数を求める粗密探索を採用している。

分散情報を用いて入力特徴画像の統計的識別を行う（手順11）。得られた変位を考慮して、入力特徴画



図 6. 認識結果の修正環境

像と候補字種の標準特徴画像の間の重み付けユークリッド距離を測定し、最も小さな値を与えた字種を認識結果とする。

1. で述べたように、認識結果の修正（手順12）を相互対話的に行い、完了後、結果をファイルシステムに保存する（処理13）。

### 3. 操作方法

GUIを採用したことにより、ほとんど全ての処理がマウス操作で指定できる。以下、文字枠検出部と手書き文字認識部の操作の概略について述べる。

手書き文字検出部は、仮想端末上にキーボードから layout [RETURN] と打ち込むことによって起動する。フレーム上部にファイル、文字枠検出、オプションと記述されたメニューが存在する。ファイルメニューを選択することで、用紙画像読み込み（手順1）、文字枠情報の保存（手順7）、処理全体の終了が指定できる。文字枠検出メニューを選択すると、文字枠の検出（手順2から6）が実行される。オプションメニューでは、文字枠検出に用いられる各種パラメータの変更が可能である。なお、メニュー以下のフレーム部には各処理の進行状況が視覚的に表示される。

手書き文字認識部は、ocr[RETURN]で起動する。辞書の読み込みは自動的に行われる（手順1）。フレーム上部にファイル、手書き文字認識、オプションのメニューが存在する。ファイルメニューを用いて、文字枠情報の読み込み（手順2）、原稿画像の読み込み（手順3）を行う。手書き文字認識メニューを選択すると、文字の切り出しから統計的識別にいたる一連の処理（手順4から11）が行われ、認識結果の修正環境に移行する（手順12）。メニュー以下のフレーム部には、元々

表 1. システムの実装環境

CPU	Alpha 600Hz
Memory	128MB
Operating System	Linux 2.0.35
Language	C
Graphic Library	Gtk+

の原稿上の位置配置を反映しながら、切り出された個々の文字画像と、認識結果が上下に隣接して表示されている（図6）。認識結果を変換したい場合は、認識結果の右隣に表示されるメニューをマウスで選択することで、縦方向に候補が表示される。この中から指定を行うことで認識結果を修正できる。また、計算機上の日本語入力システムを用いてキーボードから直接修正することも可能である。

### 4. まとめ

構築したシステムの特長は、以下の通りである。

(1) 認識精度の向上のために変位抽出手法を用いている。(2) GUIの実装。(3) 未記入の用紙画像から文字枠の位置と大きさを自動的に検出する。(4) 相互対話的な修正作業環境を提供する。

現時点では、平仮名および漢字が併せて3036字種が認識対象となっている。また、認識速度は約2.5文字/秒であった。システム構築に用いた環境を参考として表1に示す。認識辞書の作成は通産省電気総合研究所により提供されたデータベースETL-9Bを利用した。関係者各位に感謝する。

### 研究発表

- 1) R.March, Computation of stereo disparity using regularization. Pattern Recognition Letters, 8 (1988) 181-188, March 1998.
- 2) A. Jain et al., Dimensionality and sample size considerations in pattern recognition practice, Handbook of Statistics, 2, 1982.
- 3) Y. Mizukami, A handwritten Chinese character recognition system using hierarchical displacement extraction based on directional features, Pattern Recognition Letters, 12 (1998) 595-604.

### 連絡先

TEL/FAX 0836-35-9055

E-mail: mizu@is.csse.yamaguchi-u.ac.jp