

# 視覚障害者のための文庫本小説の 自動点字翻訳システムの構築

島田 恭宏\*・塩野 充\*\*

\*岡山理科大学大学院システム科学専攻

\*\*岡山理科大学工学部電子工学科

(1989年9月30日 受理)

## 1. まえがき

視覚障害者の文字情報獲得は、聴覚、触覚などの残存感覚で代行させている。しかし、これらの感覚による文字情報の獲得量は晴眼者のそれに比べると極めて少ない。これは日々出版される書物の量が膨大であるにも関わらず、これらを点字に翻訳したり、朗読し録音する作業はボランティア等による手作業が主流であるからである。従って、それらの作業量にはおのずと限界があり機械化が強く望まれている。活字で表現されている情報を視覚障害者が獲得できるような残存感覚に変換する装置として、文字認識装置(OCR)を利用した点字出力装置が考えられる。海外においてはOCRを用いたOptakonやKurzweil reading machine<sup>1)2)</sup>等が実際に市販され、使用されている。しかし日本語の場合、漢字認識の困難さ、言語処理の問題などがあり、これらに相当する市販品はいまだ開発されていない。

我が国における試みとしては、読書器の試作<sup>2)</sup>、漢字かな変換システムの試作<sup>3)</sup>等の研究成果があるが、これらは汎用コンピューターや専用OCRを用いているなど個人や小規模な施設等で購入、所有できる安価なシステムとは考えにくい。本研究では個々の視覚障害者の、いろんな本を自由に読みたいという多様な読書要望にきめ細かく対処しうるためにも、汎用コンピューター等を用いた集中型システムや専用OCRを用いた高価なシステムではなく、家庭や小規模な施設、図書館等でも所有できるようなパーソナルコンピューターレベルでの自動点字翻訳システムを目指している。そのような自動点字翻訳システムの実現のための第1段階として、パーソナルコンピューター(以下パソコンと略す)による印刷文字認識の実験を行った<sup>4)-10)</sup>。

これまでにパソコン上でのOCRの研究<sup>1)12)</sup>は、画像の入力に高価なドラムスキャナを用いたり漢字を対象としていない場合が多かった。本研究においては、使用装置としてパソコンと一般的なイメージスキャナ(以後スキャナと略す)を用いている。文庫本を入力対象としているがこの理由は、文庫本は種類が豊富で最も安価であり、小説類の書物の中では最も文字が小さく、かつ印字品質が悪い。従ってこれを対象にシステム構築を

行うことで、他の印字品質の良い大型本にも十分対応できるものと考えたからである。なお、本システムの使用形態としては晴眼者が操作を行い、視覚障害者にサービスを行うことを想定している。

本研究で想定しているシステムを図1に示す。まず文庫本をページ単位で入力し、傾き補正や文字切り出しの文書画像処理を行う。次に印刷文字認識を行い、その後、漢字かな変換等を施し、最終的な出力として点字出力、さらには音声出力を計画している。なお本論文では文字認識段階までの報告を行う。

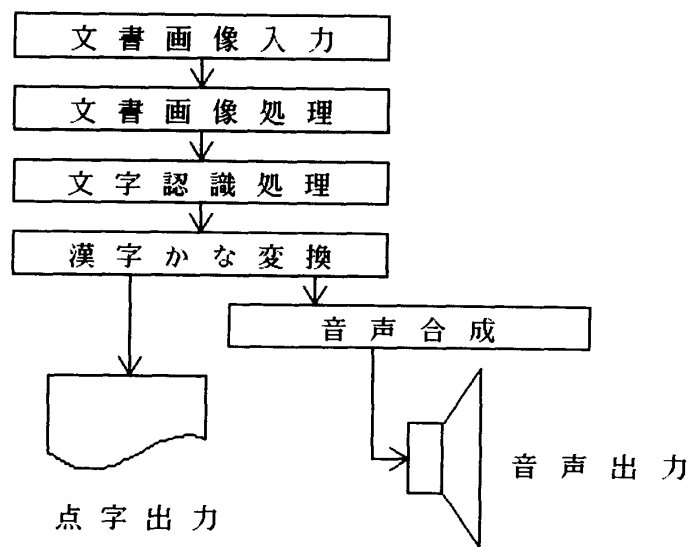


図1 本システムのブロック図

本システムのジェネラルフローを図2に示す。文書画像処理部は、基本的には縦、横方向の周辺分布をとることによって得られる情報を基にしている。傾き検出は、画像を横方向に等分し、分割された領域内において縦方向に周辺分布をとり、隣接する各領域の周辺分布の位相のずれから傾きを求める。各文字の切り出しは、縦方向の周辺分布の幅から行を抽出し、横方向の周辺分布の幅から各文字を切り出す。文字認識部における処理は大分類と個別認識の2段階で行う。大分類は、複雑指数<sup>13) 14)</sup>をユークリッド距離により候補を分類し、個別認識は単純類似度法を用いたパターン整合法により行う。各処理の詳細について以下の章で述べる。

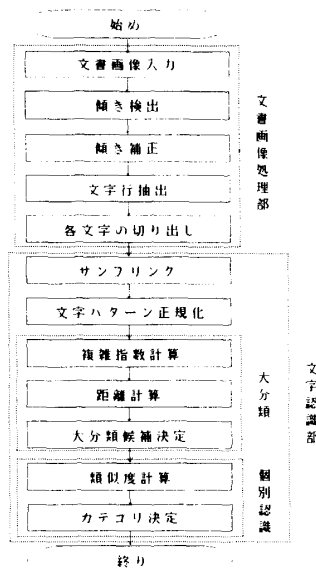


図2 本システムのジェネラルフローチャート

## 2. 文書画像処理部

通常のOCRでは、読み取る文字の大きさや位置を用紙上の絶対的な位置、或はガイドマークなどによって規定しておく必要がある、このことがOCRで取り扱い得る紙面の範囲に大きな制約を与えている。本研究では、文庫本のページという限定された紙面情報を与えているものの、印刷状況、入力方法等によっては全てのページにおいて同一とはならない場合が多い。特に入力時の画像の傾きは、文字認識に対しても重大な影響を与える。また文字認識を行う場合、1文字ずつの処理となるため文書画像からの個々の文字へと分離する必要があり、これらに対する処理を行う。

### 2.1 文書画像入力

認識対象である文庫本は、ページ単位で入力する。これは、後に計画している点字翻訳や音声出力において日本語としての文脈情報を用いたため、文章として処理データを扱う目的で処理をページ単位とし、入力操作は、1ページずつ入力する。

本研究では、パソコン用としては読み取り線密度の比較的高いスキャナを用いたため、パソコンのグラフィック画面2枚で文庫本の1ページが表示できる大きさとなる。スキャナの線密度は300dpi(dot per inch)であり、これは約12本/mmの解像度である。一方、文庫本の1ページの大きさは縦約120mm、横約80mmゆえ、1ページのドット数は縦1440ドット、横960ドットとなる。使用したパソコン(PC-98XA)のグラフィック画面は縦750ドット、横1120ドットゆえにパソコンのグラフィック画面1枚で文庫本の1ページの上半分または、下半分が表示できる。そこで全体の処理を1ページ単位で行うには、傾き検出の処理はスキャナから送られてくる生のデータを一旦ディスクに格納し、画像の傾き検出に必要な画像は、1/4に縮小して1ページ全体を表示する。画像の傾き補正以後は、ディスクに格納している画像データを読み出し、VRAM(ビデオRAM)1枚に文庫本のペー

ジ上半分を、他のVRAM 1枚にページ下半分を記憶させ、原情報本来の大きさと処理を行う。画像表示後、各処理を行う前に手動でマウスによるトリミングを行う。これは、処理範囲の決定を行うものである。

## 2.2 傾き検出、及び補正

本研究で用いたスキャナは原稿固定式のため、注意深く入力すると大きな傾きを生じることはない。よってマクロ的な傾きを検出し補正を行った後、ミクロ的な検出、補正を行うという一般的な方法ではなく、ミクロ的な検出方法のみで傾きを検出、補正を行う。本論文で使用した手法は、画像に対する処理量が少なく高精度に傾きを検出できる入力紙面の局所的周辺分布(LPP:Local Projection Profile)を用いる手法(以後LPP法と略す)<sup>15)</sup>である。この手法は、まず紙面を複数個の平行する帯状領域に分割し、それぞれの帯状領域内において周辺分布を求める。次に隣接する帯状領域の周辺分布の位相のずれを求めて傾きを検出する。図3に原画像を1/4に縮小した文庫本のページ画像に対し水平方向に6個の帯状領域に分割し、それぞれの領域内での周辺分布を求めたものを示す。図中 $\alpha_1$ は位相差であり、他の周辺分布についても同様である。そこで周辺分布 $P_1 \sim P_6$ を用い、お互いに隣接する周辺分布の相関値を計算して位相差 $\alpha_1 \sim \alpha_5$ を求めてこの平均値 $\hat{\alpha}$ から傾き $\theta$ を求める。位相差 $\alpha_1 \sim \alpha_5$ は式(1)を満足する値として求めることができる。また、傾き $\theta$ は式(2)より求めることができる。

$$\left. \begin{aligned} & \max_{-\beta \leq x \leq \beta} \left[ \sum_h P_k(h) P_{k+1}(h-x) \right] \\ & = \sum_h P_k(h) P_{k+1}(h - \alpha_k) \\ & \quad k = 1, 2, 3, 4, 5 \end{aligned} \right\} \quad (1)$$

$$\theta = \tan^{-1}(\hat{\alpha}/w) \quad (2)$$

ここで $w$ は帯状領域の幅、 $\beta$ は相関処理範囲である。本研究における精度に関する理論限界値は0.08度(=  $\tan^{-1}(1/750)$ )であり、計測の結果、 $\pm 1$ 度以内における傾き検出誤差の値は理論限界に達していた。

LPP法により求められた画像の傾き角度 $\theta$ を、アフィン変換の回転を用いて補正を行う。このような幾何学的変換は、原画像の各画素を変換式を用いて変換画像の画素に対応させるという方法をとると、変換画素の間に値が求まらない画素ができたり、2つ以上の画素が同じ画素に対応したりするので変換画像の各画素の値は最近傍法<sup>16)</sup>によって決定する。

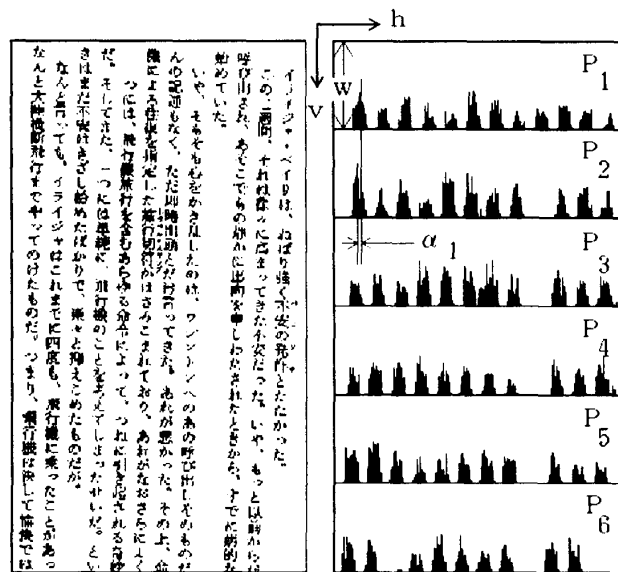


図3 紙面の部分的な周辺分布上に生ずる位相のずれ

### 2.3 行の抽出

文字行の抽出は、まず文書画像を文字の並びの方向に沿って（文庫本小説は縦書きゆえ垂直方向）に投影し周辺分布を求める。このとき各文字行の幅（文字行の周辺分布の幅）は検出できるので、入力された画像における全ての行の幅を検出し、その画像における文字行の幅の平均ピッチを求める。そして、各文字行に対する幅をこの平均ピッチと比較し、この平均ピッチと近いものはこれを文字行として切り出す。また、これよりかなり小さいものは、ふりがなの文字行と判断して切り出さない。この他、この平均ピッチよりかなり大きい場合はふりがな行と本文行の癒着と考えられ、その周辺分布の下の部分をある程度切り捨てて、これにより現れる各行の幅を、先述の処理を繰り返して行の抽出を行う（図4参照）。

### 2.4 各文字の切り出し

各文字の切り出しは、各文字行での周辺分布を幅方向にそれぞれ取り、基準となる線はその行の上部より最初に周辺分布が出現した点に引く。このとき、トリミング座標（トリミングを行う場合、文字により近い位置でその座標を決定しているため）よりも半文字分下であれば、その距離を上を持ちあげて基準線を引く（例：「，－，等）。日本語の個々の文字は、ほぼ正方形に近い矩形で囲むことができるという性質<sup>17)</sup>を利用し、その文字行の幅で候補点をあげていく。この候補点は、周辺分布より判断し、周辺分布が無い位置、またその周辺において、その値が最小の位置を仮切り出し位置としてあげる（図5参照）。一行の切り出しが終わった後、文字の平均切り出しピッチを計算し、この値を用いて先の切り出し法を再度行い、切り出し位置の決定を行う。全画面の切り出し終了後、切り出せなかった部分について、マウスを用いてユーザが手動で補正を行う。

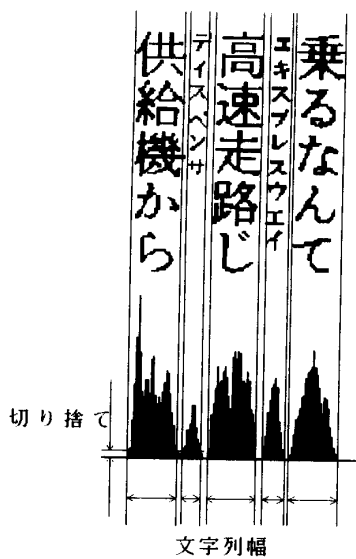


図4. 文字列抽出

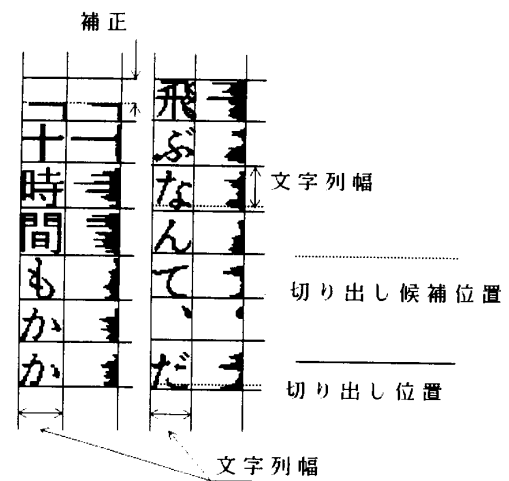


図5. 各文字の切り出し

### 3. 文字認識処理部

入力された文字を効率よく認識処理するために本研究では、大分類と個別認識で構成する。大分類には、複雑指数<sup>13) 14)</sup>を用い、入力された未知パターンと辞書パターンの複雑指数の距離を求め、100位までを候補として挙げる。個別認識には、単純類似度法を使用する。

#### 3.1 文字パターンの正規化

本研究で使用するスキャナの読み取り線密度は主、副走査共に300dpiであり文庫本の活字の大きさはほぼ8ポイント(約2.8ミリ角)である。これより抽出した文字パターンの一辺の大きさは、約33ドットとなる。しかし実際に文書画像から切り出された文字パターンの大きさは、切り出し位置のゆれがあるために正確に一定の位置に存在しているわけではない。従って、サンプリングしたパターンを直ちに一定の大きさの領域に入れることはできない。そこで、サンプリングしたパターンよりもかなり大きい作業領域を用意しておき、この作業領域にサンプリングパターンを入力する。文字パターンの正規化は、文字パターンの重心を求めそれを作業領域の中心に平行移動することによって重心位置の正規化を行う。この後、不要な枠の部分を切り捨て、パターンのサイズを白枠の部分を含め40×40画素とする。

#### 3.2 大分類

大分類に用いた文字パターンの特徴量は、漢字パターンの図形としての複雑さを示す指標として知られる複雑指数<sup>13) 14)</sup>を使用した。複雑指数は、式(3)により定義される。

$$C_x = l_y / \sigma_x, \quad C_y = l_x / \sigma_y \quad (3)$$

ここで $l_x$ および $l_y$ は、それぞれ横方向および縦方向の文字線の長さの和であり、次の方法で近似的に求める( $\sigma_x$ ,  $\sigma_y$ は後述)。図6においてN及びBはそれぞれ白地の内点、黒

字の内点であり，H，V，L，Tは白と黒の境界点である。ここでH及びVはこの点の近傍での原パターンの輪郭線方向が水平または垂直に近い場合，Lは斜線に近い場合，そしてTは細い斜線の両縁の場合に得られる量子化パターンである。文字パターンの輪郭線のどの部分もH，V，L，Tのいずれかに対応するので，輪郭線の長さはH，V，L，Tに適当な長さを割り当てることによって近似的に求めることができる。H，Vに単位1の長さを割り当てれば，L，Tに相当する長さはそれぞれ $1/\sqrt{2}$ ， $\sqrt{2}$ と考えてよい。図6の太線は各パターンに割り当てられた長さを示している。従ってパターンの輪郭線の近似値は，パターン全面についてH，V，L，Tの和 $\Sigma H$ ， $\Sigma V$ ， $\Sigma L$ ， $\Sigma T$ を計数し，それぞれに1，1， $1/\sqrt{2}$ ， $\sqrt{2}$ の重みをかけて加え合わせるによって求めることができる。すなわち輪郭線の長さを $\lambda$ とすると，

$$\lambda \doteq \Sigma H + \Sigma V + (1/\sqrt{2}) \Sigma L + \sqrt{2} \Sigma T \tag{4}$$

となる。そして輪郭線の長さ $\lambda$ は文字線分の長さの総和 $\ell$ の2倍に近似的に等しいので，

$$\ell \doteq (1/2) \{ \Sigma H + \Sigma V + (1/\sqrt{2}) \Sigma L + \sqrt{2} \Sigma T \} \tag{5}$$

となる。さらにLに割り当てられた長さの横方向の成分と縦方向の成分を考えると，それぞれ $1/2$ ， $1/2$ であり，Tの場合はそれぞれ1，1である。従って輪郭線の横方向成分 $\ell_x$ と縦方向成分 $\ell_y$ は式(6)で与えられる。

$$\left. \begin{aligned} \ell_x &\doteq (1/2) (\Sigma H + (1/2) \Sigma L + \Sigma T) \\ \ell_y &\doteq (1/2) (\Sigma V + (1/2) \Sigma L + \Sigma T) \end{aligned} \right\} \tag{6}$$

次に $\sigma_x$ および $\sigma_y$ は漢字パターンの横方向，縦方向への拡がり量を示す。この2つの量は，漢字パターンの重心回りの2次モーメントから与えられ，式(7)により求めることができる。

$$\sigma_x = \sqrt{M_{20}^*}, \quad \sigma_y = \sqrt{M_{02}^*} \tag{7}$$

ただし， $M_{20}^*$ ， $M_{02}^*$ の星印は， $M_{20}$ ， $M_{02}$ を濃度について正規化したものである。 $M_{20}^*$ ， $M_{02}^*$ は式(8)で与えられる。

$$\left. \begin{aligned} M_{20}^* &= \frac{1}{m_{00}} \left( m_{20} - \frac{m_{10}^2}{m_{00}} \right) \\ M_{02}^* &= \frac{1}{m_{00}} \left( m_{02} - \frac{m_{01}^2}{m_{00}} \right) \end{aligned} \right\} \tag{8}$$

ただし， $m_{00}$ ：0次モーメント  
 $m_{10}$ ， $m_{01}$ ：1次モーメント  
 $m_{20}$ ， $m_{02}$ ：軸まわりの慣性モーメント

$l_x, l_y$  および  $\sigma_x, \sigma_y$  は、いずれも横方向と縦方向に関して独立な分布をすること、位置と濃度に関する不変量であることが知られている。 $M_{20}^*$ ,  $M_{02}^*$  は [長さ×長さ],  $l$  は [長さ] であり漢字パターンの大きさに直接関係する量である。漢字パターンの大きさは、文字の本質的な情報ではないので、大分類のための量としては無次元量であることが望ましい。そこで、文字線分の長さの和である  $l_x$  ( $l_y$ ) を文字の拡がりの平方根  $\sqrt{M_{20}^*}$  ( $\sqrt{M_{02}^*}$ ) で除すれば無次元の量を得ることができる。この複雑指数は、単位当たりの広がりに対してどれだけの線分が含まれているかということ線分の密度を表している。大分類にはこれらの量を未知パターン、辞書パターンについて求め、ユークリッド距離を用い、距離の最小の候補から100位を取り大分類を行う。

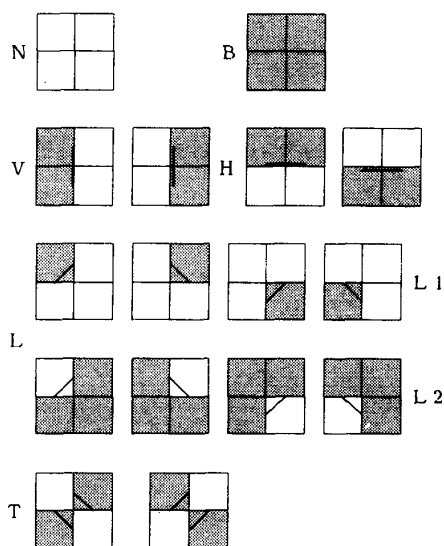


図6. 線長及び平均線幅を計算するための2×2メッシュ・パターン

### 3.5 個別認識

入力されたパターンPと第mカテゴリの辞書パターン $Q^m$ に対して類似度 $S^m(P)$ は、式(9)により与えられる。

$$S^m(P) = \frac{\sum_{i=1}^I \sum_{j=1}^J Q_{ij}^m \times P_{ij}}{\|Q^m\| \cdot \|P\|} \quad \left. \begin{array}{l} \text{ここで、} \|Q^m\| = \left\{ \sum_{i=1}^I \sum_{j=1}^J (Q_{ij}^m)^2 \right\}^{\frac{1}{2}} \\ \|P\| = \left\{ \sum_{i=1}^I \sum_{j=1}^J (P_{ij})^2 \right\}^{\frac{1}{2}} \end{array} \right\} (9)$$



実際にシステム内で類似度を計算する場合、C言語がビット操作、演算が簡単に行えること、また未知入力パターン、辞書パターンは2値パターンであることからサンプリングしたパターン、および辞書パターンを横方向に8画素毎にバイナリの8ビット（1バイト）情報として変換する。本論文では文字パターンは40×40画素で構成しており、この変換を行うことによって、パターンの情報量は1600バイトから200バイトへと圧縮できる。また、類似度を求める際、この圧縮パターンどおしの論理積をとり、ビット”1”の和をとる。

類似度法を用いた場合の認識規則は、原則として、未知入力パターンとM個のカテゴリの辞書パターンの各々との類似度 $S^m(P)$  ( $m=1, 2, 3, \dots, M$ )を計算した後、類似度を降順に並べなおし、次式により類似度最大となる辞書パターンのカテゴリ番号 $m_1$ を未知入力パターンの認識結果とすることで行う。

$$\left. \begin{aligned} \text{認識規則: } \max_{1 \leq m \leq M} \{S^m(P)\} &= S^{m_1}(P) \\ \Rightarrow \text{未知入力パターンのカテゴリ番号} &= m_1 \end{aligned} \right\} (10)$$

読み取り棄却機能を考慮した場合の認識規則<sup>18)</sup>を式(11)に示す。 $S^{m_1}(P)$ は第1候補の類似度、 $S^{m_2}(P)$ は第2候補の類似度である。

$$\left. \begin{aligned} \text{認識規則: } S^{m_1}(P) - S^{m_2}(P) &\geq \delta \\ \Rightarrow \text{未知入力パターンのカテゴリ番号} &= m_1 \\ S^{m_1}(P) - S^{m_2}(P) &< \delta \\ \Rightarrow \text{未知入力パターンの認識結果} &= \text{棄却} \end{aligned} \right\} (11)$$

ここで、 $\delta (> 0)$ は誤読、すなわち未知入力パターンの正しいカテゴリ番号が $m_1$ でない場合を救済するために導入されるしきい値である。

#### 4. 辞書の作成

認識に用いる辞書は、複雑指数、辞書パターン、ノルム、漢字コードからなり、カテゴリ数は本来ならば、小説一編は異なる字種が約2400種程度で構成されている<sup>19)</sup>のでこの程度のカテゴリ数が必要ではあるが、今回は基礎実験のためとりあえず1164カテゴリで構成した。単純類似度法を用いているために、辞書パターンとしては入力パターンと全く同じフォントを用いるべきであると考え、この本の9ページから60ページ間において1カテゴリに対し出現順に5枚までを重心位置の正規化を行った後に40×40画素でサンプリングする。大分類用辞書は、5枚のそれぞれのパターンの複雑指数 $C_x, C_y$ を算出し、この平均値を辞書値とした。個別認識に用いるパターン辞書は、この5枚を重ね合わせ、しきい値を1以上として2値化を行い、前述したようにバイナリ型のビットパターンに変換してファイルに格納している。なお、この場合のしきい値1は実験により決定したものである。

## 5. 認識実験

本研究の目的は、パソコンによる小説書物の点字翻訳システムの実現のための第一段階としての文字認識実験である。よって装置は容易に入手可能で、できるだけ安価な装置類を用いた。また入力データには安価で種類も豊富な文庫本を用いた。パソコンは、PC-98XA(CPUは8MHzの80286(80287付加))でプログラム記述にはC言語、スキャナは、主、副走査線密度共に300dpiのPIS-30(1・0データ機器社製)を用いた。入力データは、文庫本(早川書店、アイザック・アシモフ著、冬川亘訳、『はだかの太陽』、9～101ページ)で総文字数56,054文字を使用した。

認識実験は、現段階では処理時間がかかりかかるために人手の介入を必要とする処理と、人手を必要としない処理に分割した。これにより、まず外部記憶装置へ文庫本の各ページをスキャナから文書画像として入力、格納し、その後、バッチ処理により文字認識処理を行う。棄却等の処理は、後処理としてまとめて行い、認識処理中にはこれらの処理は行わない。認識処理に必要な辞書は、処理起動時に全てパソコンの主記憶上に格納しておく。棄却された文字については、個別認識結果の1位から10位までをディスプレイに表示し、ユーザが見て正解があればキーボードから候補番号をキー入力する。候補中に正解がない場合は日本語入力用ソフトを起動し、ユーザが正しい文字を入力する。

## 6. 実験結果と検討

分類処理における累積分類率を図7に示す。また表1に認識率および1文字当たりの平均認識時間を示した。大分類に用いた複雑指数は、文字パターンの特徴量として様々な望ましい特性を有しているが、雑音によるパターンのぼけに影響される<sup>14)</sup>。今回の実験でも複雑な文字に対して大分類の失敗が発生し、また簡単な文字でも横方向のストロークが多い場合に失敗している。これは入力文字の印字品質が悪く、文庫本の活字の字体が明朝体のため横線が細く、これに雑音が発生し大分類の失敗を招いていると思われる。表2に大分類の失敗文字例の一部を示す。

個別認識に用いた単純類似度法は、印刷文字認識では一般的であり計算方法も簡単である。しかし認識を失敗する場合があります、ことに類似文字の認識を失敗している。その例を以下に示す。

### I 濁点、半濁点などの形の良く似ているもの

[例] ぱ→ば, ぴ→び, ぷ→ぶ 等

### II 偏や傍の同じもの、または似ているもの

[例] 清→情, 諸→誰, 問→問 等

### III 全体的な形状が似ているもの

[例] 回→同, 人→入, 未→末 等

誤読の特徴として以上3つをあげたが、この原因としては大分類と同様、未知入力パターンに発生する雑音である。Ⅰの場合は、雑音によりその文字の特徴である濁点、半濁点があいまいな情報となって誤読につながっている。今回の実験の結果、濁点、半濁点（片仮名、平仮名の「は」行）を伴う文字の識別失敗が誤読の約26%を占めている。Ⅱ、Ⅲの場合は、文字の形状が類似しており、かつ雑音による文字パターンの歪、あるいは重心位置のずれにより誤読となっていると考えられる。

図8に読み取り棄却機能の設定を行うためのしきい値と正読率、棄却率、誤読率との関係を示した。一般に棄却判定のためのしきい値は、誤読率が棄却率の1/3から1/10程度になるように決定されている<sup>16)</sup>。本システムで扱うデータを考えた場合、小説での誤読を後処理で目視により捜すことは困難であり、極力誤読を抑える方向で処理を考える。そこで目標正読率を定め、図8より読み取り棄却を行うしきい値を決定することとした。今回は目標正読率を90%以上とすることで同図よりしきい値 $\delta$ を0.15、正読率を約90.99%とし、誤読率を約0.01%、棄却率を約9.00%とすることにした。このとき、未登録カテゴリの約99%が棄却となった。誤読率を完全に0にしようとする今回の結果では、正読率は13.39%と極めて低くなってしまいうために現実的ではないので、目標認識率に従いしきい値を決定した。

文書画像処理における各処理の成功率は、辞書作成において測定し、行抽出の成功率100%、各文字の切り出し成功率99.3%を得た。各文字の切り出しでの失敗は、文書形態の特殊な場合において発生している。例えば本文の文字長よりもその文字に対して振ってあるふりがなの文字長のほうが長い場合である。例えば「移動帯」のふりがなを「ムービング・ストリップ」としているような場合である。この場合、文字間をあけて印刷してあるために切り出しに失敗しており、また” “ ” (”等の場合、その開いた部分をつめて次の文字を印刷する場合があります。本研究で用いた「日本語の文字は正方形に近い矩形で囲むことができる」という規則がくずれているためである。LPP法は検出角度の範囲が狭いが高精度であり、1回の走査で処理が可能であり他の手法と比較して処理時間が短い。本研究においては、検出角度が1度以内であれば傾き検出誤差の理論限界値にほぼ等しい角度で検出できた。1度以内というのはかなりきびしい値と考えられるが、本論文で扱ったスキャナは原稿固定式のため注意して入力すれば傾きを抑えられるので1度程度の検出範囲でも実用上の支障はないと考えられる。なお処理の高速化、および1ページ単位での処理を前提にしたことで、縮小した文書画像に対して処理を行っているために検出範囲が狭くなっているものとする。補正に関しては、アフィン変換の回転を用い、変換座標の決定には最近傍法で近似的に行っているため誤差が生じている。

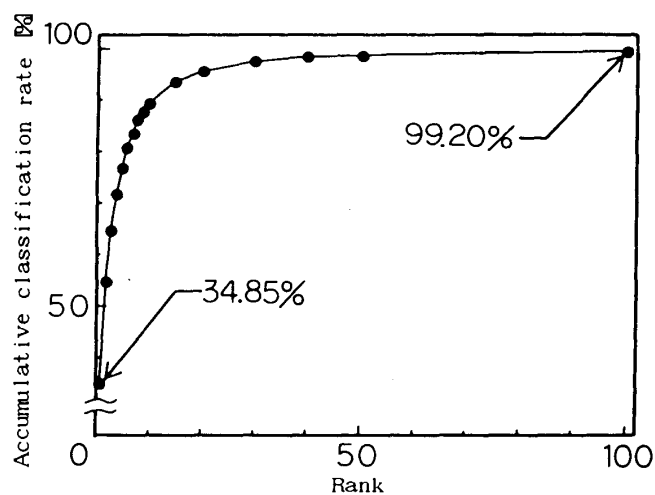


図7 累積分類率

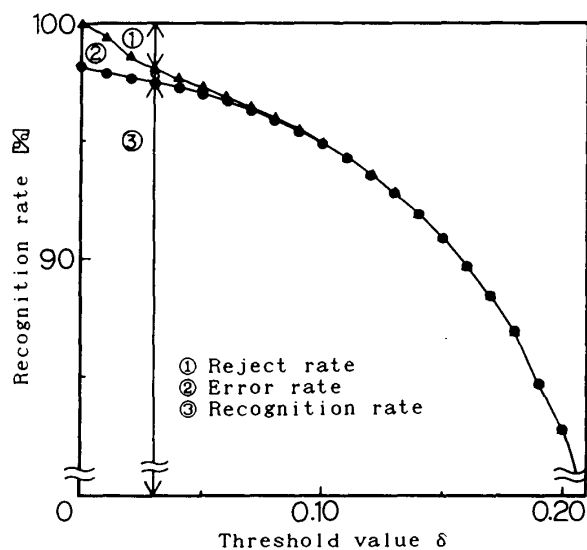
図8 認識率，誤読率，棄却率としきい値 $\delta$ との関係

表1. 認識結果

切り出し成功率	行の切り出し	100.0 %
	列の切り出し	99.3 %
大分類における正解存在率		99.20 %
平均認識率		98.15 %
1文字当たりの平均認識時間		約3.1 秒

表2. 大分類の失敗文字例

カテゴリ	出現回数	失敗回数	カテゴリ	出現回数	失敗回数
興	7	6	機	47	4
検	3	2	悪	24	2
根	8	5	面	26	2
慣	14	8	理	30	2
敵	4	2	当	31	2
急	4	2	無	53	3
膚	5	2	声	40	2
痛	5	2	感	42	2
車	20	6	全	43	2
離	24	7	手	113	5

## 7. むすび

文庫本点字翻訳のための第一段階として、パソコンによる印刷文字認識の実験を行った。印刷文字認識の手法としては特に目新しいものではないが、本研究の特徴は文字認識を行う場合、小説単行本としては文字が小さく印字状況が最も良くない文庫本を対象としていることと、パソコンベースのローコストなシステムを目指していることである。本論文で述べたシステムを全体的にみた場合、まず画像入力時は多少傾きに対し注意を払わなければならないが、ガイドマーク設定等、他の操作は必要とせず、また手動操作にマウスを用いることにより、コンピュータに不慣れな人でも操作が簡単なよう考慮した。文字認識における結果は、認識率のみで見た場合、かなり高い結果を得ているが、実用上、棄却処理は不可欠であり、この処理を入れた場合、正読率をある程度落とさなければ誤読率をおさえることができないが、これはトレードオフゆえやむをえないと考える。処理時間に関しては、入出力アクセス回数の減少、ビット演算による類似度計算等、高速化のための工夫は種々行っており、パソコン上のソフトウェアのみで構成した認識システムとしてはある程度評価できるのではないかと考える。しかし、実用化ということを考えた場合、まだ改善すべき項目が残っており、現在これらの改善を進めている。また同時に、漢字かな変換、および、かな点字変換、音声出力のための研究を進めている。

## 謝 辞

本研究に関し、熱心な御討論を頂いた本学大学院大倉 充氏、ならびに実験に御協力頂いた卒研究生有信進二氏（現・㈱岡山放送）、松岡孝則氏（現・㈱三菱電機東部コンピュータシステム）、大林一雄氏（現・㈱富士通エフ・アイ・ピー）、武中裕司氏（現・大学院）に厚く感謝する。

## 参考文献

- 1) 海保博之編：“漢字を科学する”，有斐閣，(1984)。
- 2) 篠原正美：“視覚障害者用読書器について”，システムと制御，Vol.29,no.1,pp.23-29(1985)。
- 3) 野村典正，森 健一：“漢字かな変換システムの試作”，信学論(D),Vol.J66-D,no.7,pp.789-795(1983)。
- 4) 島田恭宏，塩野 充：“パソコンによる文庫本の認識システム”，情処学中国四国支部研究会，(1987)。
- 5) 島田恭宏，有信進二，松岡孝則，塩野 充：“パソコンによる文庫本認識システムについて”，昭62中国連大，092107,(1987)。
- 6) 島田恭宏，塩野 充：“パソコンによる小説認識システムの一検討”，昭62信学情報システム部門全大，79,(1987)。
- 7) 島田恭宏，大林一雄，武中裕司，塩野 充：“パソコンによる文庫本点字翻訳システムについて”，昭62関西連大，G8-35,(1987)。
- 8) 島田恭宏，塩野 充：“パソコンによる小説書物の認識システム”，第18回画像工学コンファレンス，4-4,(1987)。
- 9) 島田恭宏，塩野 充：“文庫本点字訳のためのパソコンによる印刷文字認識”，信学技報，PRU87-92,(1988)。
- 10) 島田恭宏，塩野 充：“パーソナルコンピュータを用いた文庫本小説の印刷漢字認識実験”，テレビ誌，Vol.43,No.8,(1989)。
- 11) 上 博行，岡本 博，天満 勉，浅井 紘：“階層化判別法とその文字認識システムPC-OCR”，信学技報，PRU86-76,(1986)。
- 12) 亀沢 広，田所嘉昭：“印刷文字認識の簡単化アルゴリズムの検討”，信学技報，PRU84-79,(1984)。
- 13) 坂井邦夫，森 健一：“2000文字種を100文字／秒で読む印刷漢字OCRの開発”，日経エレクトロニクスno.172,pp.102-128,(1977)。
- 14) 坂井邦夫，森 健一：“漢字パターンの大分類”，信学技報，PRU73-17(1973)。
- 15) 秋山照雄，増田 功：“書式指定情報によらない紙面構成要素抽出法”，信学論(D),Vol.J66-D,no.1,pp.111-118,(1983)。
- 16) 長谷川純一，輿水大和，中山 晶，横井茂樹：“画像処理の基本技法”，技術評論社，(1986)。
- 17) 秋山照雄，増田 功：“印刷物の記事領域における文字の切り出し”，信学技報，PRU80-70(1981)。
- 18) 中田和雄編：“パターン認識とその応用”，コロナ社，(1987)。
- 19) 林 大監修：“図説日本語”，角川書店，(1982)。

# The Construction of Automatic Braille Translation System of Paperback Novel Books for Blind Persons.

Yasuhiro SHIMADA\* and Mitsuru SHIONO\*\*

*\*Graduate School, Okayama University of Science,*

*1-1 Ridaicho, Okayama 700 Japan*

*\*\*Faculty of Engineering, Okayama University of Science*

(Received September 30, 1989)

In this paper, a recognition system of novel books written in Japanese language for blind persons using a personal computer is proposed. This system consists of two units, a document image processing unit and a character recognition unit. In the former unit, normalization of the document image and extraction of the character patterns are carried out. In the latter unit, pre-classification and recognition of character patterns are carried out. As the result of the experiment with 56,054 characters of 93 pages in the novel books, average recognition rate of 98.15% was obtained. Average processing time was 3.1 sec/character.