候補文字ラティス法による オンライン手書き文字列の認識

Online Handwritten Character-string Recognition by Candidate Character Lattice Method

村瀬 洋　若原 徹　梅田 三千雄

Hiroshi MURASE , Toru WAKAHARA and Michio UMEDA

日本電信電話公社 武蔵野電気通信研究所

Musashino Electrical Communication Laboratory, N.T.T.

ABSTRACT: The segmentation of closely spaced handwritten characters is the major problem in online free-format text recogntion. This paper describes a method of combining segmentation with character classification by means of lattice description technique. The method accepts a string of connected and multi-sized characters since it can decompose the string into recognizable patterns by searching for the optimum path on the candidate character lattice. When applied to 105 handwritten strings of 9 to 11 characters, the method attained a recognition rate of 95.8%.

1．まえがき

近年の文書処理の発展に伴い漢字や図形情報の簡便な入力法に対する期待は高く、特に手書き漢字認識による円滑なマンマシンインターフェイスの実現が望まれている。これら実現のための基本技術として、低品質文字認識（文献(1)）、自由書式文字認識等、筆記時の制限の少ない認識技術の確立が必要となっている。本報告で述べるオンライン自由書式手書き文字列の認識は、即時性や対話性が良好であるというオンラインの特長に加え、各文字の枠内筆記の制限を解除するという特長を持つ。枠内筆記の制限が解除されれば、大量文書を入力する際の精神的・肉体的な負担は低減される。また、この認識技術は図面中の文字のような任意の場所に大小混在した文字や、罫線上に筆記された文字の認識等広い応用範囲を持つ。

従来、自由書式文字列を認識する手法は、文字と文字を分離するセグメンテーション段階と、個々の文字を認識する文字認識段階から構成されていた。セグメンテーション手法には、筆記者がセグメンテーションを意識して入力する手法として、文字と文字の間でピック等の指示を与える手法や、1文字を書き終えてから次の文字を書き初めるまでの時間間隔を利用する手法（文献(2)）があるが、いずれも筆記者の負担が大きくなる。一方、自動的にセグメンテーションを行う手法として、英文字列を対象にストロークの位置関係を利用する手法（文献(2)）、ストロークの横軸への投影を利用する手法（以下、単純投影法と呼ぶ。）等がある。しかし、漢字のように左右・上下に分離する可能性の高い対象では、ストロークの位置関係の利用だけで文字を確実にセグメンテーションすることは困難となっている。この他にOCR型文字列を対象とした報告（文献(3),(4),(10)）があるが、これらの手法をそのままオンライン型の文字列に適用することは不適当である。他方、文字認識には多数の手法（文献(5),(6),(7)）が提案されているが、文字が正しくセグメンテーションされなければ、これら文字認識技術は適用出来ない。

本報告では、タブレット上に手書きされた自由書式の文字列を認識するため、「候補文字ラティス法」を提案した。ここでは、自由書式文字列の一例として、セグメンテーション指示なく横方向に手書きされた仮名漢字文字列を扱う。「候補文字ラティス法」は筆者らの提案による線図形認識法「候補ラティス法」（文献(8)）を文字認識用に改良した手法であり、音声認識で知られる「音韻ラティス」（文献(9)）と同様の発想に基づく。「候補文字ラティス法」は各文字のセグメンテーションと各文字の認識を同

時に実現する手法であり、個々のストローク形状等の下位の特徴から、文字の包含関係等の上位の特徴までを有機的に利用することが可能なため、高精度に文字列を認識することが可能である。本手法の特徴を以下に述べる。

（1）手書き文字列の文字間隔の分散は大きく、文字が重なる場合から、文字間が1文字分以上離れる場合まで存在する。本方式では本質的に文字間隔（文字ピッチ）の情報を使用しないため、文字間隔の分散が問題にならない。

（2）漢字や平仮名には、"科","川","い" 等の分離文字が多数存在する。本手法では各文字の認識情報が自動的に文字のセグメンテーションに利用されるため、分離した部分図形が他の文字にならなければ、分離文字は正しく認識される。

（3）漢字には、"理","炉" 等のように左右に分離し、分離した各部分図形が他の文字になる場合がある。この場合には、文字の縦横比を利用した分離有意文字処理を導入している。

以下、2章で認識対象を、3章で手書き文字列の性質を、4章で認識アルゴリズムを、5章で認識実験結果を、6章で考察を述べる。

2. 認識対象

日本語文書には縦書きと横書きが存在するが、最近では横書きの文書の割合が極めて高く、技術図面中に出現する文字列にも横書きが多い。また横書き文字列の認識技術が確立されればこの技術を縦書き文字列に拡張することは容易であることから、認識対象として横書き文字列を扱う。その際、筆記者の負担を軽減するため、不定ピッチ、文字の大小混在、等の自由書式を許す。文字列は常用漢字1945字種と平仮名46字種の文字から構成されるものとする。

2. 1. 文字列データ

本報告で述べる認識アルゴリズムの効果を確認するため、認識実験用の文字列データを以下の条件で収集した。

(1)横罫線（1cm間隔）の用紙に、文字列を自由なピッチで筆記させた。
(2)各文字の筆順は自由であるが、画数は守らせた。
(3)文字列として以下の5文例を用意し、それを筆記させた。

① 地理的に離れた場所での
② 詳細な検討を続ける
③ 高騒音下での通話対策
④ 安定でしかも量産に向く
⑤ 利用するための手引

(4)文字ピッチを密・普通・粗の指示で、同一文例を3回づつ筆記させた。
(5)7名の筆記者に筆記させた。

データ収集数は（7名、各5文例、各3回）の105文例であり、文字総数は1050文字である。収集した文字列の一例を図1に示す。

図1　文字列の一例

2. 2. 辞書作成データ

文字認識に用いる標準パターンの作成には、常用漢字と平仮名のデータ（文献(7)）を利用した。文字の種類は常用漢字1945字種、平仮名46字種、画数変動文字204字種（例えば，" 子" は正しくは3画であるが2画で筆記されやすい。このような文字）の合計2195文字である。30名分の文字（65850文字）の内20名分（43900文字）の文字を用いて標準パターンを作成（4.6.(2)参照）した。また、残りの10名分の文字（21950文字）を予備実験の文字認識実験（4.6.(6)）に使用した。

3. 手書き文字列の性質

手書き文字列の性質を予備実験として調査する。

3. 1. 文字間隔の分布

手書き文字列の文字間隔の分散は印刷文字の場合と異なり極めて大きい。図2に、2章で述べた文字列データについて，X軸（横方向）に着目した文字間隔の分布を示す。但し、文字間隔の値は、文字列の縦幅を1.0とするように大きさを正規化（詳細は4．3．にて述べる。）した後の値である。負の文字間隔は文字が重なり合うことを意味する。図2より、手書きでは文字が10％程度重なる場合から、文字間が1文字分以上離れる場合まで存在することが確認された。この事実は、手書き文字列のセグメンテーションに単純投影法の適用が困難であることを示す。

3. 2. 分離文字の存在

漢字や平仮名には、" 科"，" 川"，" い" 等の

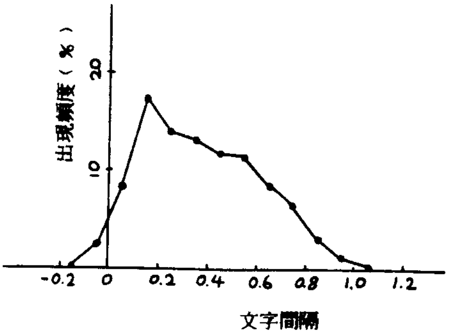


図２　文字間隔の分布

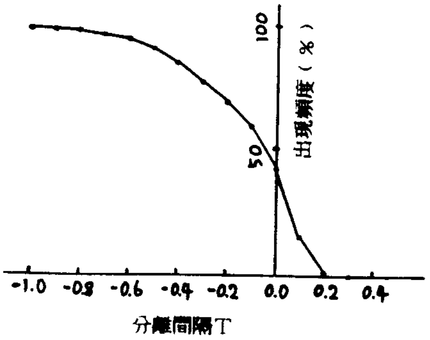


図３　分離間隔Tに対する分離文字の割合

ように、左右に分離する文字が多数存在する。それ以外にも、"口"の第１ストロークが第２、第３ストロークと分離するような、手書変形によりストロークの一部が左右に分離する場合なども存在する。そこで、文字の分離度を調査した。文字がMストロークから構成される場合に、第１ストロークから第ｉ（１≦ｉ＜Ｍ）ストロークまでの最大値をX_1とし、第ｉ＋１ストロークから第Mストロークまでの X座標の最小値をX_2とした際、$X_2 - X_1 > T$（Ｔは定数）となるｉが１個でも存在すれば、この文字は間隔Tで分離すると呼ぶことにする。図３に、間隔Tに対する分離文字の割合を示す。但し、対象は画数変動文字を含む常用漢字と平仮名２１９５字種を各５サンプルづつ、合計１０９７５文字である。各文字は前処理として大きさ1.0（縦幅と横幅の最大値を1.0とする。）に正規化する。図３よりX軸への投影が分離するような文字（Ｔ＞０）は４３％存在することが確認される。この事実は、文字のセグメンテーションに単純投影法を適用したならば、単独文字が複数の図形に分離される可能性が高いことを示す。

４．　認識アルゴリズム（候補文字ラティス法）

４．１．認識アルゴリズムの概要

本報告で提案する候補文字ラティス法は文字のセグメンテーションと認識を同時に実現することにより、セグメンテーション精度を高める手法である。図４に、処理の流れを示す。処理は、文字列から基本セグメントへの分解、基本セグメントから候補文字の生成、候補文字の認識と文字認識結果の候補文字ラティスへの登録、候補文字ラティスの探索、分離有意文字処理の各段階から構成される。以下、各処理を詳細に述べる。

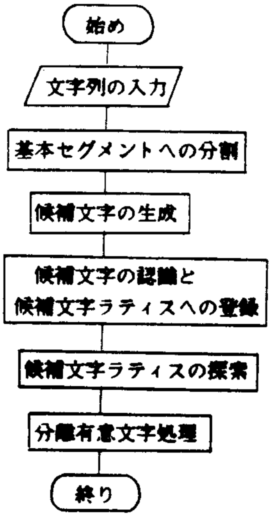


図４　処理の流れ

４．２．手書き文字列入力

手書き情報の入力にはデータタブレットを使用する。タブレットはペンのアップダウン情報とＸＹ座標の値を一定時間間隔でサンプリングする機能を有する。ペンのダウンからアップまでの座標値系列をストロークと呼ぶことにする。入力ストローク数がM，第ｊ（１≦ｊ≦Ｍ）ストロークの座標点数がN_jの場合の座標値系列を$\{(X_{ij}, Y_{ij})\ i=1,\ldots,N_j;\ j=1,\ldots,M\}$とする。

4．3．前処理

前処理として、文字列全体の大きさの正規化とストローク毎の座標点数の規格化を行う。大きさは、各座標点のY座標の2次モーメントの値が一定になるように正規化する。正規化後の座標値系列を｛（X'_{ij}，Y'_{ij}）$i=1,...,N_j; j=1,...,M$｝とすれば、X'_{ij}，Y'_{ij}は以下の式により与えられる。

$$\begin{cases} X'_{ij}=A \cdot X_{ij} \\ Y'_{ij}=A \cdot Y_{ij} \end{cases}$$

但し、

$$A=0.25/Y_m$$

$$Y_m=\frac{1}{P}\sum_{j=1}^{M}\sum_{i=1}^{N_j}(Y_{ij}-\bar{Y})$$

$$\bar{Y}=\frac{1}{P}\sum_{j=1}^{M}\sum_{i=1}^{N_j}Y_{ij}$$

$$P=\sum_{j=1}^{M}N_j$$

である。ここで、Pは総座標点数、$\bar{Y}$はY座標の平均値、Y_mはY座標に関する2次モーメントを表わす。この変換により、変換後のY座標の2次モーメントは0．25に正規化される。これは文字列の平均的な縦幅（高さ）を1．0とすることに相当する。

一方、入力された座標点は一定時間間隔でサンプリングされたものであるため、座標点間の間隔は筆記速度により変動する。そこで座標点間の間隔が一定値Wとなるように座標値系列を変換する（文献(1)）。Wは文字を表現するのに十分な値に設定した。変換後の座標値系列を｛（X''_{ij}，Y''_{ij}）$i=1,...,N'_j; j=1,...,M$｝とする。但し、N'_jは第jストロークの座標点数である。

4．4．基本セグメントへの分割

文字列を構成する入力ストローク列を複数の基本セグメントに分割する。方針は、基本セグメントの中に文字間の分離箇所が含まれない範囲で、なるべく大きなストロークの組が基本セグメントになるように、分割することにある。基本セグメントの決定パラメータをTh とする。第1ストロークから第k（$1 \leq k < M$）ストロークまでのX座標の最大値をX_1とし、第k＋1ストロークから第MストロークまでのX座標の最小値をX_2とし、（X_2-X_1）＞Thを満足するkをすべて求める。各kに対し第kストロークと第k＋1ストロークとの間で入力ストローク列を分割し、分割された各ストロークの組を基本セグメントとする。図5に、Thに対する基本セグメントの例を示す。Thを小さくすれば基本セグメントの単位が小さくなるが、Thを大きくしすぎれば基本セグメント内に文字間の分離箇所が含まれる。図2より、通常横方向に手書きされる文字列では文字間の重さなりは15％以下と考えられる。この結果より、Thとして－0．15を設定する。分割後の基本セグメントをe_1，・・・，e_t とする。但し、tは基本セグメント総数である。

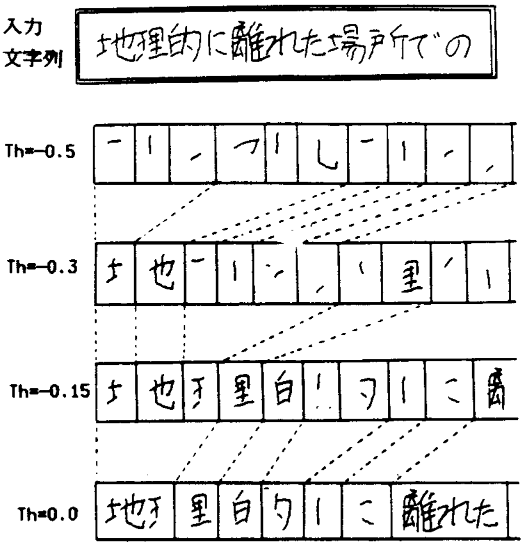


図5　Thに対する基本セグメントの例

4．5．候補文字の生成

文字列中の各文字は基本セグメント（e_1，・・・，e_t）を組み合わせることにより構成されるが、この段階では組合わせ方は不明である。そこで、基本セグメントを組合わせ、組合わせた図形が文字らしくなるすべての場合に対し、この図形を候補文字とする。基本セグメントを組合わせた図形が候補文字となるための条件は、

①図形は引き続く基本セグメントから構成される。

②図形の横幅は2．0以下（経験的に十分な値）である。

③図形を囲む長方形の長辺は0．2以上（経験的に十分な値）である。

④図形のストローク数は２３以下（常用漢字の最大画数は２３）である。

と設定した。図５で示した基本セグメント（Th=－０．１５）から生成した候補図形を図６に示す。候補文字数は実際の文字列に比較して一般に多数生成される。図６に示した例では、１１文字から構成される文字列に対し、２４の基本セグメント（$e_1, \cdots, e_{24}$）が抽出され、基本セグメントの組合わせにより９８個の候補文字が生成される。この場合、実際の文字数に比較し約１０倍の候補文字が生成されることになる。

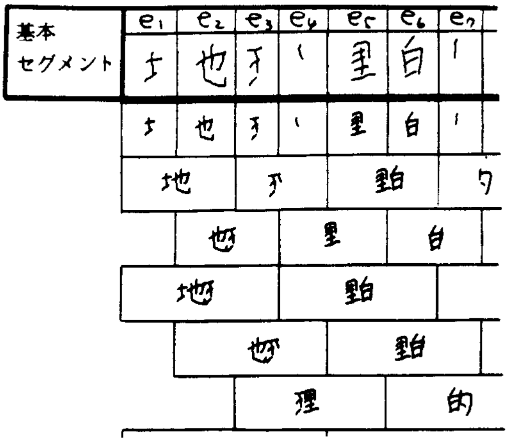


図６　候補文字の生成

４．６．候補文字の認識

前節で生成された各候補文字を認識することにより、各候補文字はどの文字らしいか（候補文字認識結果）、どの程度その文字らしいか（距離値）を求める。以下、候補文字の認識手順を述べる。

（１）前処理

前処理として、候補文字の座標値系列の重心と２次モーメントを利用して、候補文字の位置と大きさの正規化を行う。候補文字の座標値系列を｛（U_{ij}, V_{ij}）$i=1,\ldots,K_j; j=1,\ldots,L$｝（但し、Lはストローク数、$K_j$は第ｊストロークの座標点数）とし、正規化後の座標値系列を｛（U'_{ij}, V'_{ij}）$i=1,\ldots,K_j; j=1,\ldots,L$｝とした場合、$U'_{ij}$, V'_{ij}を

$$\begin{cases} U'_{ij} = B \cdot (U_{ij} - \bar{U}) \\ V'_{ij} = B \cdot (V_{ij} - \bar{V}) \end{cases}$$

で定義する。但し、

$$\bar{U} = \frac{1}{Q} \sum_{j=1}^{L} \sum_{i=1}^{K_j} U_{ij}$$

$$\bar{V} = \frac{1}{Q} \sum_{j=1}^{L} \sum_{i=1}^{K_j} V_{ij}$$

$$Q = \sum_{j=1}^{L} K_j$$

$$B = 0.5 / m$$

$$m = \frac{1}{Q} \sum_{j=1}^{L} \sum_{i=1}^{K_j} \{ (U_{ij} - \bar{U})^2 + (V_{ij} - \bar{V})^2 \}$$

である。ここで、Qは総座標点数、$\bar{U}$, $\bar{V}$は各U_{ij}, V_{ij}の平均値、mは（$\bar{U}$, $\bar{V}$）の回りの２次モーメント、Ｂは正規化係数である。

（２）候補文字認識用の辞書の作成

画数変動文字を含んだ常用漢字と平仮名２１５３カテゴリーについて辞書を作成する。辞書の作成手順を以下に示す。最初に、辞書作成用文字の各ストロークの座標系列を一定の点数（ここでは１５点）で近似し、前処理として上記（１）の処理を施す。次に各カテゴリーに対し２０個の文字を使用し、対応する座標点毎に座標値の平均値を求める。最後に、座標点間の距離が一定値になるように変換する（４．３．の後半の処理を参照）。この座標値系列を候補文字認識用の辞書とする。

（３）ストローク間の距離の定義（位置と方向のＤＰマッチング）

ストローク間の距離は（４）で述べる文字間の距離の計算に使用される。２本のストローク間の距離をストローク各点の位置（座標値）とその点での接線方向の値のＤＰマッチング（文献(8)）により定義する。ＤＰマッチングは本手法を画数変動を含む低品質文字認識に拡張する際に有効になる（文献(1)）。一方のストロークの座標値系列を｛（U_i, V_i）$i=1,\ldots,I$｝とし、他方を｛（U'_j, V'_j）$j=1,\ldots,J$｝とする場合、両者のストローク間距離ｄを、

$$d = \frac{1}{J} \min_{u} \{ \sum_{j=1}^{J} (g(j, u(j)) + \alpha \cdot h(j, u(j))) \}$$

で定義する。但し、Ｉ，Ｊはストロークの近似点数であり、ＩとＪとは一般に異なる。g（ｉ，ｊ）は（U_i, V_i）と（U'_j, V'_j）の間のユークリッド距離を表わす関数で、

$$g(i, j)=\sqrt{(U_i-U'_j)^2+(V_i-V'_j)^2}$$

となる。h(i, j)は、点(U_i, V_i)から点(U_{i+1}, V_{i+1})に向かうベクトルと、点(U'_j, V'_j)から点(U'_{j+1}, V'_{j+1})に向かうベクトルのなす角度で定義する。u(j)は座標点間の対応関係を表わす関数であるが、不適当な対応を防止するため制限条件として、

$$\begin{cases} u(1)=1 \\ u(J)=I \\ u(j)=i \quad \text{のとき} \\ \qquad u(j+1)=\{i \text{ or } i+1 \text{ or } i+2\} \end{cases}$$

を課した。αは接線方向の特徴の寄与度をあらわす定数であり、(6)で述べる予備実験で設定される。

（4）文字間距離の計算

候補文字は複数のストロークから構成される。まず、辞書の中から候補文字のストローク数（Lとする）と同じストローク数の標準文字を選択し、選択された標準文字のすべてと候補文字との間で文字間の距離を求める。文字間の距離は対応するストローク間の距離（（3）参照）の和で定義する。但し、候補文字の筆順は自由であるため、ストロークの対応は最初から決定できない。そこで、ストローク間距離行列（文献(7)）を用いて、対応するストロークを決定する。標準文字の第i（$1 \leq i \leq L$）ストロークと、候補文字の第j（$1 \leq j \leq L$）ストロークとのストローク間距離をd_{ij}とすれば、ストローク間距離行列は、

$$\begin{pmatrix} d_{11} & d_{21} & \cdot & \cdot & d_{L1} \\ d_{12} & d_{22} & \cdot & \cdot & d_{L2} \\ \cdot & \cdot & d_{ij} & & \cdot \\ \cdot & \cdot & & & \cdot \\ d_{1L} & d_{2L} & \cdot & \cdot & d_{LL} \end{pmatrix}$$

となる。この行列の各行から最小の値を選択し（対応するストロークの決定）、それらの和を文字間の距離とする。つまり、文字間の距離Dは、

$$D=\sum_{j=1}^{L}\{\min_{i=1,\cdots,L}(d_{ij})\}$$

で定義される。

（5）候補文字の認識

候補文字と、候補文字に等しいストローク数を持つすべての標準文字（辞書）との間で文字間距離を求め、最小の距離値を持つ標準文字のコードを認識結果とする。

（6）文字認識実験（予備実験）

（1）～（5）で述べた認識手法を評価するため、未学習文字21950文字（各カテゴリー10文字）を用いて認識実験を行った。図7にα（（3）参照）をパラメータとした認識率を示す。α＝0．3とすることにより、1～3画に対して98．8％、4～23画に対して99．9％、平均で99．8％の認識率を得ることを確認した。特に、接線方向特徴の利用は、平仮名等の画数の少ない文字に有効であることも確認した。

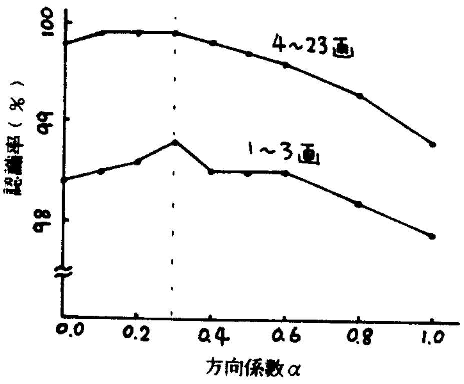


図7　方向係数αに対する文字認識率

4．7．候補文字ラティスの作成

候補文字ラティスは、入力文字列中の各基本セグメントがどの候補文字の一部であるかの情報、各候補文字の認識結果、及び認識時の距離値を、表形式に表現したものである。図6に示した表の各欄に候補文字の認識結果と距離値を代入することにより、候補文字ラティスが生成される。候補文字ラティスの例を図8に示す。

4．8．候補文字ラティスの探索

入力文字列中に含まれる各文字をセグメンテーションし、各文字を認識することは、候補文字ラティスの中からすべての基本セグメントを被う最適な文字系列を選択することにより実現できる。候補文字ラティスは2端子の重み付きグラフ（候補文字の認識結果をブランチに、距離値をその重みとするグラ

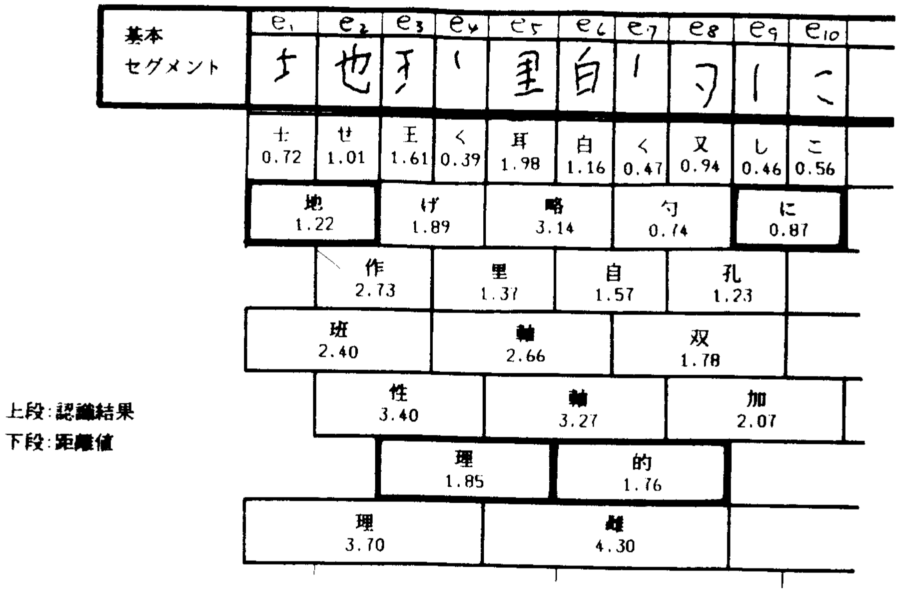


図8　候補文字ラティスの例

フ）により表現できる。図8に示す候補文字ラティスに対するグラフ表現の例を図9に示す。すべての基本セグメントを被う文字系列の選択は、このグラフの一方の端子から他方の端子に向かうパスの選択に対応する。このパスは一般に多数存在する。そこで、評価関数Sを導入し、Sを最小にするパスを探索する。パスの探索には動的計画法を採用した。Sは、

$$S=\sum_{\text{ブランチ}\in\text{パス}}(\text{ブランチの重み})$$

で定義する。つまり、ブランチの重みの和を最小にするパスを選択する。このパスに対応する文字系列を仮の認識結果とする。図8中に、太枠で仮の認識結果を示す。

4．9．分離有意文字処理

文字が手書き変形により左右分離した場合や、偏と旁により左右分離した場合でも、左右の部分図形の少なくとも一方が文字として意味を持たなければ、4.8.までの処理で文字列は正しく認識できる。しかし、文字が左右に分離し、且つ分離した各部分図形がどちらも文字として意味を有する場合（この文字を分離有意文字と呼ぶ）には、4.8.までの処理では必ずしも正しく認識されない。例えば、”説明”という文字列が、”言”，”兑”，”日”，”月”の4個の基本セグメントに分割された場合には、図13のような2端子グラフが生成される。この場合

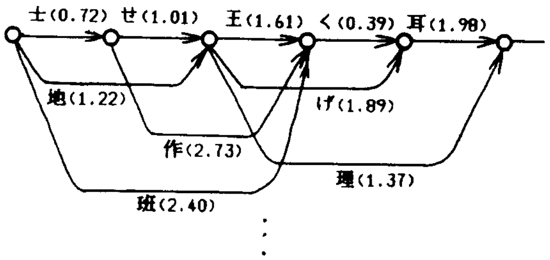


図9　候補文字ラティスのグラフ表現

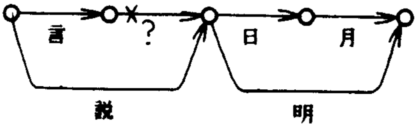


図10　候補文字ラティス（"説明"）のグラフ表現

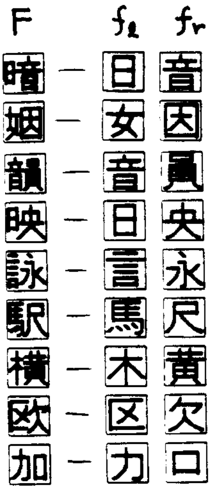


図１１　分離有意文字例（Ｆ＝ｆｌ＋ｆｒ）

には”説”は正しく、認識されるが、分離有意文字である”明”は、”明”と認識されるか”日”＋”月”と認識されるかは距離値情報だけでは決定できない。分離有意文字は常用漢字１９４５字種中２６６字種存在する。その一例を図１１に示す。

ある分離有意文字ＦをＦ＝ｆｌ＋ｆｒと表現する（例えば、Ｆ＝”明”，ｆｌ＝”日”，ｆｒ＝”月”）。仮の認識結果（４．８．参照）中にｆｌ，ｆｒなる文字列が存在し、且つ候補文字ラティス中にｆｌ，ｆｒに対応してＦなる候補文字が存在する場合、又は仮の認識結果中にＦなる文字が存在し、且つ候補文字ラティス中にＦに対応してｆｌ，ｆｒなる候補文字が存在する場合には、正しく認識されていない可能性がある。この場合には文字Ｆの縦横比と、文字ｆｌとｆｒの縦横比の平均値を比較し、その値が1.0に近い方の文字（文字列）に修正する。この処理を分離有意文字処理と呼ぶ。こうして選択された候補文字の認識結果の系列を、入力文字列の認識結果とする。

５．　認識実験結果

４．で述べた認識手法の効果を確認するため、１０５個の文字列データ（２．１．参照）を認識対象の一例として認識実験を行った。

５．１．分離有意文字処理を行わない場合

候補文字ラティス中から単純に距離値の和を最小とする文字系列を選択する手法（４．１．～４．８．参照）により文字列データを認識した結果、認識率は９４．９％であった。但し、認識率を、

認識率＝（正しく認識された文字数）／（総文字数）・１００　％

で定義する。表１に誤読５．１％の内訳を示す。

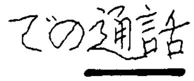
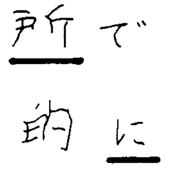
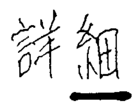
５．２．分離有意文字処理を行った場合

５．１．に示した誤読中、セグメンテーション誤りの一部は分離有意文字処理を導入することにより改善される。その結果、認識率は９５．８％に向上

表１　認識率と誤読内訳

		分離有意文字処理	
		無し	有り
認識率		９４．９％	９５．８％
セグメンテーション誤り	基本セグメント分割誤り	０．４	０．４
	ラティス探索誤り	３．７	２．８
文字認識誤り		１．０	１．０

表２　誤読例

	入力パターン例
（１）基本セグメント分割誤り 「通話」→「嚇」＋「舌」 「に検」→「論」 他	
（２）ラティス探索誤り 「所」→「戸」＋「斤」 「討」→「言」＋「寸」 「に」→「し」＋「こ」 「も」→「こ」＋「し」 他	
（３）文字認識誤り 「細」→「組」 「で」→「ぞ」 他	

した。表１に誤読４．２％の内訳を、表２に誤読例を示す。

６．　考察

６．１．本アルゴリズムの適用範囲

本アルゴリズムは文字のセグメンテーションに文字ピッチや文字の大きさ等の情報を使用していないため、以下のような文字列を認識することが可能である。

（１）文字の重なり

本手法では文字の大きさに比較して１５％までの文字間の重なりを許容する。Thを－0.15より更に小さい値に設定すれば、１５％以上の文字間の重なりをも許容することが可能になる（図１２（１）参照）。

（２）不定ピッチ

本手法では完全な不定ピッチを許容する（図１２（２）参照）。

（３）文字の大小混在

本手法では文字列における各文字の大きさの大小を許容する（図１２（３）参照）。

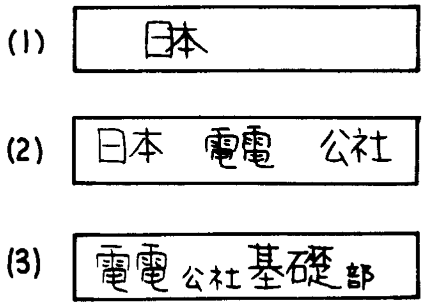


図１２　可読文字列例

６．２．計算量

本手法では、基本セグメントの組み合わせから生成されるすべての候補文字に認識処理を施して候補文字ラティスに登録する。これによりセグメンテーションは高精度化されるが、多くの計算量がかかる。例えば、図１に示した文字列の例（１１文字から構成）では９８回の文字認識処理がなされた。必要な文字認識回数は文字の重なりの許容度Thに依存する。長い文字列中の１文字に対し必要となる平均的な文字認識回数を文字列データで調査した。図１３に、その結果を認識率とともに示す。図１３より、文字の大きさに比較し１５％以下の重なりを持つ文

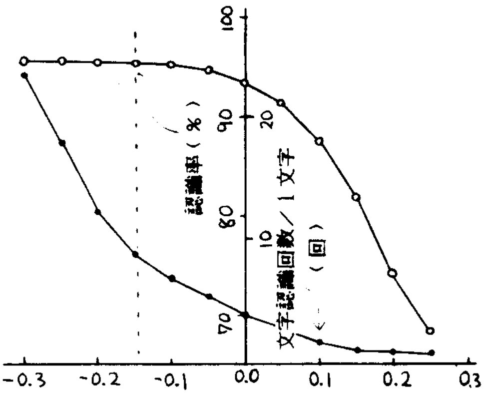


図１３　Thに対する文字認識回数と認識率

字列を認識するためには認識回数を文字数の約１０倍にすれば良いことがわかる。

６．３．他手法との比較

単純投影法により文字列から文字をセグメンテーションし、各文字を文字認識する手法と、本手法とを比較した。単純投影法は文字列を構成する各ストロークを横軸に投影し、その影の間隔がT（いき値）以上になる箇所で文字をセグメンテーションする手法である。文字列データ１０５文例に対し、Tをパラメータとしてセグメンテーションの実験を行った。セグメンテーションの成功率を図１４に示す。Tを小さな値に設定すれば１文字が複数の図形に分離されるため誤読は増加し、Tを大きな値に設定すれば

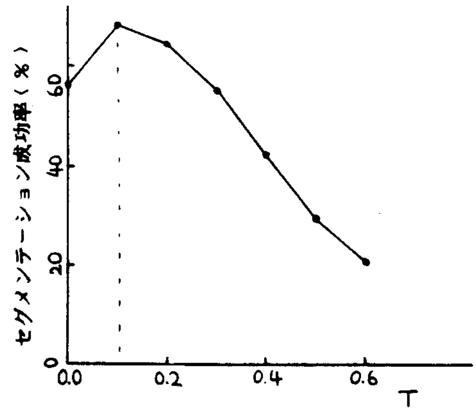


図１４　単純投影法におけるTに対する
セグメンテーション成功率

表2　文字列の認識率

単純投影法	候補文字ラティス法
６８．０％	９５．８％

文字間が連結するため誤読は増加する。Ｔを最適に設定することにより６８．５％のセグメンテーション成功率が得られる。これを文字認識したところ、文字列の認識率は６８．０％であった。表２に、本手法との認識率の比較を示す。

７．　むすび

本報告では、真に扱い易い入力形態を追求するため、セグメンテーション指示なく横方向に手書きされた文字列の認識手法として「候補文字ラティス法」を考案し、その性能を実験的に確認した。以下に、まとめる。

（１）文字列からの文字のセグメンテーションと各文字の認識を有機的に結合したことにより、文字のピッチ等の情報を全く使用することのない高精度な文字列認識法を実現した。その結果、完全な不定ピッチ文字列、文字間の重なりを許容する文字列、大小文字混在の文字列の認識が可能となった。文字列データ１０５文例に対し９５．８％の認識率を得た。

（２）文字の縦横比を利用した分離有意文字処理を導入したことにより、”理”，”炉”等の２文字が左右に配置されることにより構成される分離有意文字のセグメンテーション精度は向上し、導入しない際のセグメンテーション誤り４．１％を３．２％に低減することが可能になった。

本報告で提案した「候補文字ラティス法」はオンライン手書き文字列の認識に極めて有効である見通しを得た。今後の課題には、更に多数の文字列データで本手法の能力を定量的に分析すること、候補文字の生成処理を階層化し処理を効率化すること、分離有意文字処理に言語知識等の上位概念を導入し更に充実させること、等がある。

謝辞　日頃御指導頂く畔柳情報通信基礎研究部長、増田第三研究室長に深謝します。研究を進める上で有益な議論をしデータ収集に協力して頂いた第三研究室および視覚情報処理研究グループの諸氏に感謝致します。

文献

（１）若原，梅田："ストローク結合情報を用いたオンラインくずし字分類法"，信学技報，PRL83-30(1983)

（２）G.F.Groner:"Realtime recognition of handprinted text",AFIPS,FJCC,29,p591 (1966)

（３）馬場口，塚本，相原："手書き文字列における文字切り出しの基礎実験"，信学技報，PRL 83-58 (1983)

（４）T.Akiyama,S.Naito,I.Masuda:"A method of character extraction from format-unknown document images",ICIP,p85 (1983)

（５）吉田，迫江，"スタックＤＰマッチングによるオンライン手書き文字認識"，信学システム部門全大，90 (1983)

（６）M.Nakagawa,T.Manabe,K.Aoki,Y.Ikeda, N.Takahashi:"Online handwritten character Recognition as Japanese input method",ICTP (1983)

（７）小高，若原，増田："筆順に依存しないオンライン手書き文字認識アルゴリズム"，信学論(D)，J65-D，6，P679 (1982)

（８）村瀬，若原，梅田："候補ラティス法による手書きフローチャートのオンライン認識"，信学論(D)，J66-D，6 (1983)

（９）鹿野，好田："会話音声の機械認識における言語処理"，信学論(D)，J61-D，4 (1978)

（１０）R.G.Casey,G.Nagy:"Recursive Segmentation and Classification of Composite Character Patterns",Proc. 6-th ICPR