

オンライン手書き文字認識におけるパタンプリミティブの情報理論的決定とその辞書生成

相澤 正、中川正樹、高橋延匡
(東京農工大学 工学部 数理情報工学科)

1. はじめに

日本における情報処理システムを考える時、常に問題となるのは入力法である。今後情報処理システムがオフィスから家庭にまで広く日本の社会に浸透するためには、現在のようなキーボードを主体とした入力法でなく、素人でも容易に扱えることができる自然な入力法の確立が必要である。

現在、日本語入力法としては、現実的観点から仮名漢字変換方式が主流をなしているが、この方法には同音異義語選択の必要性、分ち書きの必要性、図表入力との親和性、等の問題があり、究極的な日本語入力法とは考えがたい。マンマシン・インターフェイスの点でも入力者に対する多大な負担、強制を黙殺している感がある。今後は人間が計算機に合わせるのではなく、計算機が人間に歩みよるといふマンマシン・インターフェイスの改善を志向すべきである。

このような観点到ち、本研究室においてはオンライン手書き文字認識方式を前提とした日本語入力システムの研究を行ってきた。この研究は、JOLIS (Japanese On-Line Input System) 計画と名づけられ、1981年度には実験システム JOLIS-1 の開発[1]を、また、1984年度にはその評価[2,3]を終了した。JOLIS-1 は平仮名、片仮名、教育漢字および句読点等の記号を構造解析的手法により、字種に依らない統一的な認識を行う。JOLIS-1 は、タブレットより入力される筆点位置座標点列に、前処理を施した後、8方向の直線セグメント系列に変換する。次に直線セグメント列として表現されたストロークは、方向変化とその前後の方向セグメント列として情報圧縮される。その後、ストローク辞書によって表現される29種類の基本ストロークの何れかに識別される。JOLIS-1 は、このように文字を基本ストロークの列で表現し、その列を標準字体の文字辞書と相違度マッチングをとることにより入力文字を認識するものである。ところが、この JOLIS-1 の評価により、低画数文字に対する認識の問題が明らかになった。それは要約すれば、JOLIS-1 においてはパタンプリミティブ(基本ストローク)の決定が適当でなく、ストロークあたりの特徴抽出量(情報量)が不足しているため、低画数文字の認識においては十分なカテゴリーの分離が行えないということである。これは構造解析的パタン認識におけるオープンプロブレム『いかにしてパタンプリミティブを決定するか』という問題に帰着する大きな課題である。JOLIS-1 の後継システム JOLIS-2 においてはこの課題の克服を目標としており、そのためのフィージビリティテストとして本研究は位置づけられる。JOLIS-2 は、英数字、ギリシャ文字、仮名、常用漢字、等の2265文字の認識を目標としている。このように、JOLIS-1 において問題となっていた低画数文字をあえて追加することによって、JOLIS-2 を特徴抽出方式研究における試金石と考えている。

以下の節では、まずパタンプリミティブ抽出の問題点を明らかにする。次に本研究の目標を述べた後、実際に行ったテスト内容について述べる。

2. パタンプリミティブ抽出の問題

現在主流をなしている統計的手法と異なり構造解析的手法は、対象をまずパタンプリミティブの集まりとみなし、その構成をもとに認識を行うものである。したがって、情景解析など対象が多くの要素よりなる問題に対して究極的な方法と考えられている。我々もこの手法を漢字という構造をもつ対象に応用しようと試みている。しかし、この手法にはいくつかの問題点が指摘されており、このため実際に構造解析的手法がパタン認識に応用された例は、2種の染色体の識別[4]、郵便番号読み取り装置[5]などごく少数にすぎない。

この構造解析的手法における問題点としては主に次の2点が挙げられている。

①パタンプリミティブの抽出に失敗した場合にいかに対処するか

②いかにしてパタンプリミティブ・セットを決定するか

まず①の問題に関して JOLIS では、文字認識処理に相違度の概念を導入することで対処した。これによって、抽出されたストローク列情報のささいなエラーについて、文字辞書の特別な拡張なしに対処できた。これは先に述べた郵便番号読み取り装置が、すべてのエラーを考慮した状態遷移を用意していたのに比べれば大きな差違である。

しかし、相違度の導入によってこの問題がすべて解消されたわけではなかった。すなわち、比較的高画数の文字にストローク誤識別が生じた場合には、この相違度の効果によって文字としては正認識に到達する可能性が高い。しかし、低画数の文字に同様なエラーが生じた場合には、この相違度の効果は現れず、最終的に誤認識またはリジェクトの結果を得ることが多かった。このような現象の原因は、JOLIS-1 の特徴抽出がストロークを主にしており、入力ストローク数がそのまま入力パタンより得られる情報量を決めてしまうということにある。すなわち、高画数文字パタンより得られる情報量は大きいため、一部にエラーが生じても全情報に対するその比率は微小である。しかし、低画数文字パタンでは同様なエラーが生じた場合、その比率は極めて高く許容できないことによる。また、この問題は文字辞書中の標準パタンの追加等によって対処できるケースは極めて少ない傾向を

もつ[3]。このことから、特に低画数文字からより多くの情報を抽出できるようなストローク識別方式が必要である。低画数文字からも高画数文字同様に多くの情報を抽出できれば、些細なエラーが生じた場合でも相違度の効果によって救済できることになる。このためには、JOLIS-1におけるストローク識別処理を抜本的に見直し、より高精度、安定的なものにする必要がある。

次の②の問題は、構造解析的手法において常に存在するオープンプロブレムと言えよう。現在、このパタンプリミティブ・セットをいかにして決定すれば最適なものとなりうるかという理論は存在しない[6]。それは、構造解析的パタン認識なるものが、その名のとおり構造解析とパタン認識の2面性をもつことによる。構造解析の立場によれば、入力された情報をいかに細分化し、その構造を明らかにするかという問題に帰着する。しかし一方、パタン認識の立場によれば入力された情報をいかに凝縮し、一つのカテゴリを決定するかという問題に帰着する。この両極の間において、パタンプリミティブの抽出という問題は入力パタンをいかに細分化し、いかに統合するかという相反する問題のトレードオフを考慮する必要がある。その上で最適なパタンプリミティブ・セットを決定しなければならない。これまで、構造解析的手法が応用された例の大部分はこの問題をおざなりにし、経験的直感に基づいてパタンプリミティブ・セットを決定してきた。その結果十分な特徴抽出が行われず、したがってパタンプリミティブ抽出の失敗に有効な対処が不可能であったという歴史が繰り返されてきた。

3. 本研究の目標

本研究の目標を次の5点と設定した。

- (1) ストロークの形状をより精密に表現可能な方法を導入する。
JOLIS-1 の一つの問題として、その表現能力の不足が挙げられる。ストロークを方向変化とその前後の方向セグメントによって表現した結果、同一基本ストローク列で記述されるカテゴリが多数存在し、その識別の分解能の低さが確認された。ストローク形状をより詳細化することによってこの問題に対処する。
- (2) ストローク修飾やループなどの変形に対し、可能なかぎりアルゴリズムにおいて吸収する。
ストロークの不安定要因として、ストローク修飾やループなどの変形の存在が確認された。これに対し JOLIS-1 では辞書学習という対策をとったが、十分な効果を得られなかった。本研究では、この問題に対しアルゴリズムにおいて対処する。
- (3) ストローク識別に際し、距離の概念を導入しセグメントの方向誤り等のエラーに対処する。
JOLIS-1 におけるストローク識別では、セグメントの方向誤り等の変形をストローク辞書によって吸収した。そのため、予期しない変形を許容できない性質をもっていた。この問題に対しストローク識別に距離の概念を導入することで対処する。
- (4) ストローク識別において、context を用いない方式での限界を追求する。
本研究のストローク識別においては、まずストローク形状情報のみでどの程度の識別能力が得られるかを考察する。したがって、ストロークの位置、大きさの情報は主として用いないものとする。まずストローク形状情報を主にしたストローク識別処理を構成し、“つ”、“っ”のように大きさの情報が必要とされる場合にはそれを付加的に利用するものとする。この方式の限界を考察した上で、ストロークの大きさ、位置情報、前後関係の導入に移行すべきであると考ええる。
- (5) 基本ストローク・セット、ストローク辞書の自動生成を目指す。
経験的に設定された基本ストローク・セット、辞書というものが曖昧、不統一なものとなりやすいことは JOLIS-1 において確認された。したがって本研究では基本ストローク・セット、ストローク辞書の可能なかぎりの自動生成を目標とする。

4. ストローク識別処理

圧縮表現を基に分類を行った後、より精密な識別が必要とされる場合にはさらに詳細識別を行うという階層的ストローク識別とした。処理の流れを図1、基本ストロークを図2、ストローク辞書を図3に示す。

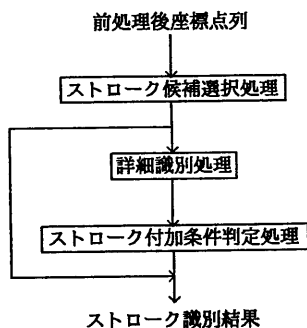


図1 ストローク識別処理の流れ



図2 基本ストローク・セット(134通り)

(*をつけたストロークは付加条件によって大小のストロークに分離する)

NO	転回数	標準圧縮表現	詳細識別標準パターン	形状
1	0	A 5点近似	→ 大
2	0	A	同上	→ 小
3	0	A 5点近似	→
:	:	:	:	:
8 6	2	A (LT) C (RT) H または H (LT) A (RT) G	ㄣ 15点近似	ㄣ
8 7	2	D (LT) A (RT) G または C (LL) C (RT) G	ㄨ 15点近似	ㄨ
:	:	:	:	:
1 3 4	5	F (LT) A (RT) F (LT) B (RT) A (LT) F	ㄨㄨ 30点近似	ㄨㄨ

図3 ストローク詳書

4.1 ストローク候補選択処理

このストローク候補選択処理は、JOLIS-1におけるストローク識別処理を基にしている。これは圧縮表現が比較的安定であることが、その評価を通じて明らかにされたことによる。このため本処理ではこの圧縮表現を利用し、次に示す処理を行う(図4参照)。

ステップ1:前処理後座標点を8方向セグメント量子化する。これは座標点列 P_1, P_2, \dots, P_n に対して $P_i P_{i+3}$

を抽出し、JOLIS-1と同様セグメント化処理フェーズ1、フェーズ2の処理を施すものである。 $P_i P_{i+3}$ は、セグメント方向の安定性を考慮し、実験的に設定したものである。

ステップ2:セグメント列より圧縮表現を得る。このときの転回種類はJOLIS-1と同様である。すなわち、1方向差以上5方向差以下の転回を右転回(RT)、左転回(LT)、6方向差以上の転回を右転回ループ(RL)、左転回ループ(LL)とする。

ステップ3:ストローク候補を次の規則によって複数個選択する。

- ①入力ストロークの圧縮表現に類似する圧縮表現をもつ候補をすべて選択する。この場合の類似とは転回種が同一であり、加えて単純ストロークでは方向差±1、複合ストロークでは始末端セグメントにおける方向差±2、中間セグメントにおける方向差±1の範囲中に存在するものと定義する。この類似性の許容範囲は実際の手書きパターン変形の分布を基に設定した。
- ②入力ストロークが複合ストロークである場合、始末端セグメントがそれぞれストローク修飾として疑わしいか否か判定する。この場合のストローク修飾の可能性の判定は、次の2条件を満たすか否かとする。

(i) 始末端セグメントの長さがストローク長に比較して20%以下であること。

(ii) 疑わしきセグメントを除去した場合、ストロークの転回数が減少すること。

もし、疑わしい場合には、該当する始端または終端セグメントを除去した圧縮表現についても類似する候補を選択する。

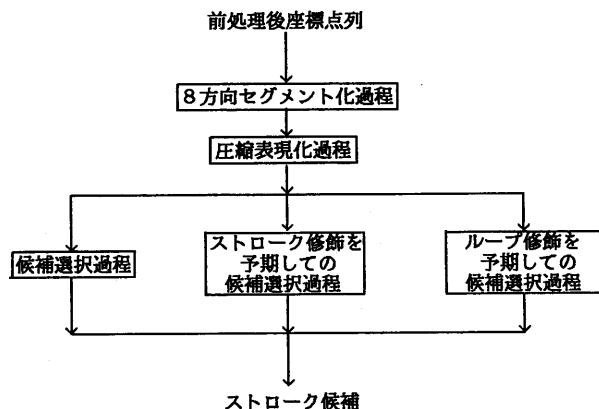


図4 ストローク候補選択処理の流れ

- ③入力ストロークがループを含む場合、そのループがターンの変形で入力された、有意ではないループ(ループ修飾)か否か判定する。座標点列 P_1, P_2, \dots, P_n より

$$P_i P_{i+1,0}$$

を抽出し、8方向セグメント化を行う。この処理によってループの消去が認められた場合、ループ修飾の疑いがあると判定する。 $P_i P_{i+1,0}$ は有意でないループの大きさの分布を基に、実験的に設定した。もし、ループ修飾と疑わしい場合にはそのループを除去した圧縮表現についても類似する候補を選択する。

4.2 詳細識別処理

前節のようなストローク候補選択処理を経た結果、ストロークが1つに特定できない場合に詳細識別を行う。ストローク候補選択処理が位相的特徴をもとに分類を行ったのに対し、この詳細識別では幾何学的特徴をもとに識別を行う。詳細識別処理は、図5に示す流れにそって行われる。まず、ストロークの大きさ、記入枠中の位置に対する正規化処理を行う。この後、座標点近似方式によって詳細識別を行いストロークを決定する。この結果が“つ”、“っ”などいくつかの大きさ情報を必要とするストロークであった場合には、先の大きさ正規化の時点で用いたパラメータによって大小を判定し、唯一の識別結果を得る。以下に各過程の内容を述べる。

4.2.1 正規化過程

まず、詳細識別にあたっては入力ストロークの大きさ、位置に関する正規化を行う。本研究におけるストローク識別では“つ”と“っ”のようにストロークの大きさまたは位置によって基本ストロークを分離しなければならない場合を除き、ストロークの大きさ、位置の情報をストローク識別処理において利用しない。これは形状情報による識別の限界を見極めるためである。

まず位置についてストロークの外接方形の中心を正規枠の中心点に合わせる。そして次に、大きさについてストロークの外接方形の縦、横の長さの大きい方を基準値に合わせるという方法をとる。これらはストロークの形状、特に縦横の比率を保存するといった配慮による。

4.2.2 詳細識別過程

ストロークは次のように表現される。

$$(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$$

x_i, y_i はそれぞれX、Y座標値である。この方式は、ストロークをいくつかの座標点で近似表現し、基本ストロークとの統計的識別を行うというものである。以下にこの流れを述べる。

(1) 点近似表現

この方式で問題とされるのは、主として近似点数である。小高らによれば、入力文字のストローク数によって3点(7画以上)または6点(6画以下)の近似を行っている[7]。また、北原らによれば、ストロークの入力点数によって4~9点の近似を行っている[8]。小高らの方式によれば、漢字のつづけ字などのストローク形状を抽出することは難しい。また、北原らの方式によれば、入力速度、ストローク長に表現が依存することになる。

このような点を考慮し、近似点数はストロークの複雑度によって可変とすべきであると考えた。そこで、このストロークの複雑度としてセグメントを圧縮表現した時の転回数をを用いることにした。これは、転回数がストローク修飾、ループ修飾という不安定要因を除けば比較的安定な特徴であることが立証されており、なおかつ、このことについてはストローク候補選択処理において考慮されていることによる。本研究では、転回数 t に対し近似点数を $5(t+1)$ 点とした。すなわち、転回数0の時5点近似、転回数1のとき10点近似、転回数2の時15点近似といったように転回数に比例して近似点数を増加させて扱った。これによれば先の小高、北原らの方式に比較して近似点数は多くなる傾向にあるが、ストローク候補選択処理によって詳細識別の対象となる基本ストロークは限定されており、冗長度は比較的悪くないと考えられる。

この点近似表現化に際しては、入力座標点列より始終端の点は必ず抽出するものとし、その他の点から残りの必要な数の点を抽出する。すなわち n 個の座標点列 P_1, P_2, \dots, P_n から m 点抽出する場合、次式によって抽出点を決定する。

$$\begin{cases} i=1 \text{ or } n \text{ or } \left\lceil \frac{n-2}{m-2} k \right\rceil + 1 & \text{ただし } k=1 \sim m-2 & \text{accept} \\ i = \text{otherwise} & & \text{not accept} \end{cases}$$

(2) ストローク間距離の算出

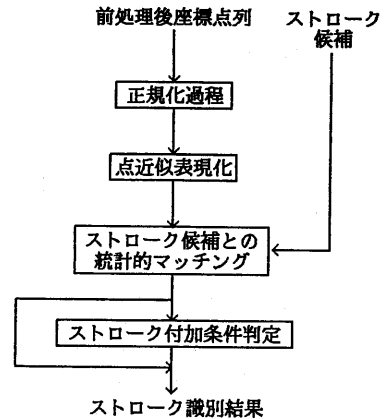


図5 詳細識別処理の流れ

ストローク間距離として、本研究ではシティ・ブロック距離を用いた。これは

$$\begin{aligned} \text{ストローク } S &= (x_1, y_1, x_2, y_2, \dots, x_n, y_n) \\ \text{ストローク } S' &= (x'_1, y'_1, x'_2, y'_2, \dots, x'_n, y'_n) \end{aligned}$$

のストローク間距離として、次式と定義する。

$$D_p = \sum_{i=1}^n (|x_i - x'_i| + |y_i - y'_i|)$$

このストローク間距離としては、マハラノビス距離を用いることでパタンの分布を考慮し、学習操作も容易になると考えられる。しかし、計算量の点でなお一考の余地があり本研究では採用しなかった。まず第一次近似として、シティ・ブロック距離を用い本方式の問題点、有効性を考察した後、マハラノビス距離などの導入を図るべきであると考ええる。

4.2.3 ストローク付加条件判定

以上のような詳細識別によって、ストローク識別結果が確定した上で、ストローク付加条件によって、さらに識別結果を分離する必要がある場合も存在する。これは例えば“つ”、“っ”などであり、ストロークの大きさ情報を用いなければ、異なるカテゴリが識別できない場合に付加的に行う。この条件にあてはまるストロークは調査の結果 JOLIS-2 認識対象2265カテゴリの内に14組存在した。詳細識別の結果がこの14組のストロークの何れかであった場合、正規化の処理において用いた縦、横の拡大縮小比率によって大小を判定し、その識別を行うものとする。

5. 基本ストローク・セットの決定およびストローク辞書の生成

5.1 基本ストローク・セット決定の条件

構造解析的手法におけるパタンプリミティブの決定方法に確立された理論は存在しない[6]。構造解析的手法を用いた他システムの例でも、大多数が経験的に定めたパタンプリミティブ・セットを用いている。JOLIS-1 においても経験的に設定した29通りの基本ストロークを用いたが、その結果、同一ストローク列で表現される文字が存在した。このため JOLIS-1 では付加条件を用いなければならなかったが、この条件は概して安定的ではなかった。

これに対し、本研究におけるパタンプリミティブ・セット決定の条件を次の2点と定義した。

- (1) パタンプリミティブ・セットは、認識対象を識別するに十分な集合である。
- (2) 上記(1)を満たすパタンプリミティブ・セットの中でも、よく誤識別される組みはまとめる。

この(1)は言い換えれば、パタンをパタンプリミティブ列として表現した場合、各カテゴリが一意に表現される必要があるということである。この条件なくしては、必ずパタンプリミティブ列だけでは区別のつかないカテゴリが存在することになり、JOLIS-1 においてその識別が問題となった。しかし、形状情報を基に基本ストロークを細分化してみても、認識対象文字を一意に表現することは困難である。したがって、本研究では文字を表現するに際して、ストローク間の位置関係情報として裏ストロークを導入する。この裏ストロークは、継続するストローク対の始終端点間を連結した、目に見えないペンの動きに対応するものである。これによってストローク間の位置関係のある程度表現することができ、文字に対する表現が緻密になる。

また、(2)は(1)の条件を満たすパタンプリミティブ・セットの中でも区別することの意味が少ないものはまとめて一つにすることを良しとする観点である。パタン認識においては、そのオーバーヘッドを考えると、パタンプリミティブ数が小さい方が有利である。この方針はすなわち、パタンを一意に記述するに理想として必要にして十分なパタンプリミティブの集合を求めようとするものである。

また、ストローク辞書の経験的決定には、パタン変形の分布が反映されていないため、ストローク辞書が不統一、曖昧となる傾向があった。このことより、本研究ではより自動的なストローク辞書の生成を行うものとする。また、完全な自動化が不可能な場合にはツールを用い、人間の簡単な操作によってストローク辞書の生成が行えるようにする。

5.2 基本ストローク・セットの決定および辞書生成の流れ

前節の条件に基づき、基本ストローク・セットの決定およびストローク辞書の生成を以下のように行った(図6参照)。また、これらのステップの副産物として、フラットな構造をもつ擬似的な文字辞書も生成することができる。

ステップ1: 圧縮表現の分類

まず第一に、ストローク辞書生成用パタン群1より、各文字パタンがどのようなストローク、裏ストローク列によって入力されるか調べる。この時ストロークは圧縮表現によって記述し、裏ストロークは8方向に量子化して記述する。そして、圧縮表現化されたストローク表現の分類を行う。各ストローク圧縮表現を仮の基本ストロークと定義する。これによって各文字パタンは、この(仮の基本ストローク+裏ストローク)列(以後、仮の文字表現と呼ぶ)によって表現されることになる。

ステップ2: 基本ストローク細分化の過程

次にステップ1で作成された仮の文字表現について、異なる文字が同一表現となっているか否かを

チェックする。これは、圧縮表現と裏ストローク列からでは表現が分離できないカテゴリの組を抽出することにある。この結果、異なるカテゴリが同一表現となっていた組に対しては、人間が適当なストロークを指示し仮の基本ストロークを追加することで対処する。すなわち、基本ストロークを細分化することで、各カテゴリの仮の文字表現が一意に表現されるよう修正を行う。また、このように修正された仮の基本ストロークについて詳細識別のためのストローク辞書を作成する。

ステップ3: 基本ストローク統合化の過程

最後にステップ1およびステップ2で作成した基本ストロークの候補選択辞書、詳細識別辞書を用いて、ストローク辞書生成用パターン群2のストローク識別処理を行う。ステップ1の結果、各カテゴリは仮の文字表現で記述されているが、これと本ステップにおいて実際に識別された結果の対応をとり、そのストローク誤識別確率を集計する。この集計データによって誤識別される確率の高い基本ストローク同士は統合することを検討する。この際、ストローク識別処理のモデルを情報路に求め、情報理論をこの統合化の判断基準に応用した。この基本ストロークの統合に対応して、ステップ1、ステップ2で作成した仮の文字表現、候補選択辞書、詳細識別辞書を修正する。これによって仮の文字表現は擬似的な文字辞書となり、候補選択辞書、詳細識別辞書は基本ストローク辞書となる。

次節よりこれらのステップについて、より詳細に述べる。

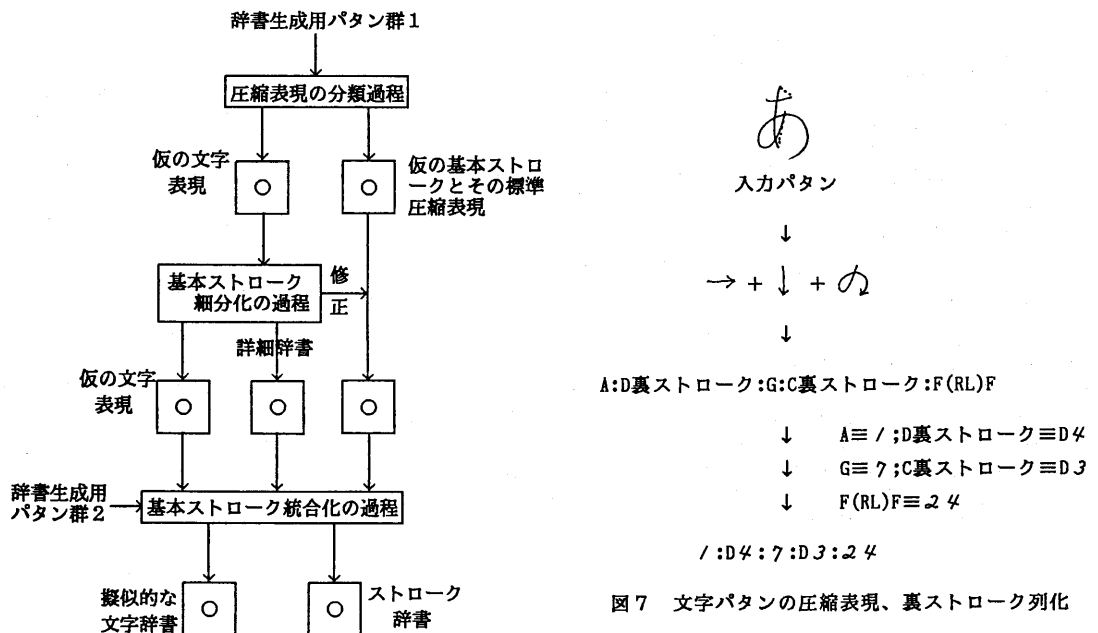


図6 ストローク辞書生成の流れ

5.3 圧縮表現の分類

各カテゴリについて文字データを収集し、ストローク辞書生成用パターン群1とする。この文字データに対し前処理[9]を施した後、第4.1節で述べたストローク候補選択処理と同様にストロークの圧縮表現化を行う。この際、生成する基本ストローク・セットの安定化を図るため、ストローク修飾、ループ修飾に関する判定は、本研究においては人間が行うものとした。これによって、ストローク辞書生成用パターン群1中の各文字を(圧縮表現+裏ストローク)列で表現する(図7参照)。この結果、どのような圧縮表現のストロークが生起するか明らかとなる。また、各カテゴリとこの(圧縮表現+裏ストローク)列との対応が得られ、後に述べる擬似的な文字辞書につながる。

次に、抽出したストロークの圧縮表現の一つ一つを仮に基本ストロークとして扱い、ストロークコードによって表現する。すなわち、さまざまな圧縮表現を各々基本ストロークと定義し、この時点では似かよった表現といえども統合化を行わない。図7の例ではカテゴリ“あ”が仮に、/:d4:7:d3:24と表現されたことを示している。このように各文字カテゴリを(仮の基本ストローク+裏ストローク)列によって表現する。

実際には JOLIS-2 の認識対象文字2265カテゴリの、平仮名、片仮名、英数字、ギリシャ文字(計319カテゴリ)については被験者3名より各1ボタン計3ボタン、漢字(1946カテゴリ)については被験者1名より1ボタン計1ボタンについて収集しストローク辞書生成用パターン群1とした。このパターンについて圧縮表現の分類を試みた結果、226通りの圧縮表現を得た。

5.4 基本ストローク細分化の過程

次に、前節で作成した各文字カテゴリの仮の表現について、異なるカテゴリが同一表現となされていないかチェックを行う。これは圧縮表現と8方向量子化裏ストローク列だけでは表現が分離できないカテゴリの粗を求め、その中で仮の基本ストロークを細分化することで各カテゴリを一意に表現する目的をもつ。

本過程ではパターンプリミティブ数を必要な数だけ増加させていく。すなわち、カテゴリ C_1, C_2, \dots, C_n がそれぞれ

$$C_i = s_{i1} d_{i1} s_{i2} d_{i2} \dots d_{i,m-1} s_{im} \quad (m: \text{ストローク数}, s: \text{ストローク圧縮表現}, d: \text{裏ストローク}, 1 \leq i \leq n)$$

と表現される場合を考えると、 $k \neq l$ の時 C_k, C_l の表現が一致してしまう場合に適当な s_{kx}, s_{lx} についてそれぞれ別の基本ストロークとし、基本ストローク・セットを拡大することにする。この結果、カテゴリ C_k, C_l が一意に表現されることになる。この s_{kx}, s_{lx} の選択についてはカテゴリ分析等の手法を用いて、より類似しないストロークを選択する必要があるが、現状ではこの点は人間の能力に依存した。

このような例としては“つ”と“フ”が挙げられる。すなわち“つ”でも“フ”でも圧縮表現では $\rightarrow(RT)\angle$ とされてしまう。そこで“つ”、“フ”のストロークをそれぞれ別の基本ストロークとして追加し、その識別を詳細識別で行うものとする。これは図8で示すように、ストロークの特徴空間内においてストロークの識別が十分になされていないために、文字パターン全体が識別できない場合、新たな軸を導入し(詳細識別を行い)このストロークを分離する。これによって文字特徴空間内において、十分な識別がなされることになる。しかし、図9のようにストローク特徴空間内においてストロークが十分に分離されてなくとも文字特徴空間内では、他のストロークの存在によって十分な識別が行える場合には、このストローク対を分離する必要はないであろう。このような処理によって各文字カテゴリを一意に表現するに必要な数の基本ストロークが設定されたことになる。

また分離される基本ストロークを含むカテゴリについては、ステップ1で作成した仮の文字表現を修正する必要がある。これについてもやはり人間がより類似する基本ストロークへと指示し、それにしたがって修正を行った。またここで、各基本ストロークと文字パターンデータとの対応より、ストローク詳細識別のための辞書を作成した。

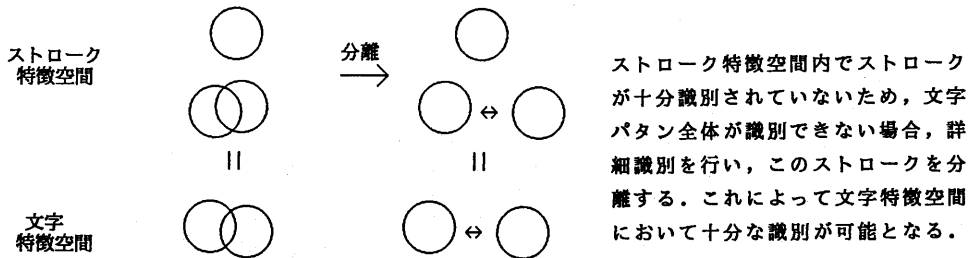


図8 基本ストローク追加のモデル

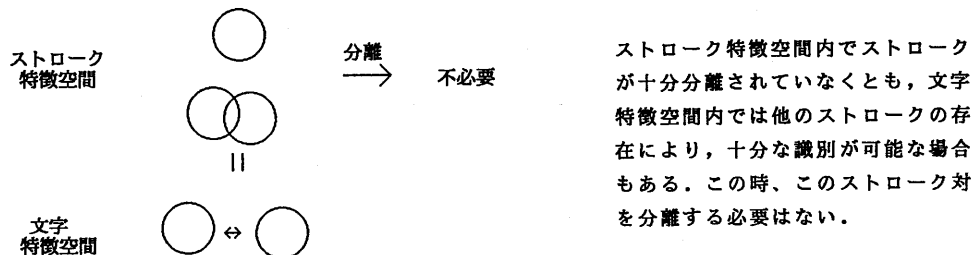


図9 基本ストロークを追加する必要のない場合

実際に前節の圧縮表現分類で述べた辞書生成用パターン群1について、本過程を行った結果、68組、161カテゴリについて表現が一致した。このため53通りの基本ストロークを増加させるに到った。例として、図10に示すよ

うに→(RT)↙なるストロークを含み、このチェックに挙げられたカテゴリは3組存在し、その結果9通りの基本ストロークへの細分化を行った。

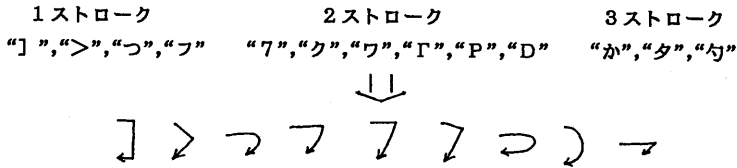


図10 →(RT)↙の細分化

5.5 基本ストローク統合化の過程

以上の過程によって各カテゴリを、仮のストロークと裏ストロークの系列によって一意に表現することができた。次に類似した基本ストロークを統合し一つの基本ストロークとする過程を考える。この過程を図8、図9と同様に示せば図11のように直感的に説明できよう。すなわち、ストローク特徴空間内において比較的接近しているストローク同士を統合しても、文字特徴空間内においてさほど影響のない場合、このような基本ストローク対については統合して一つの基本ストロークに再定義するというものである。

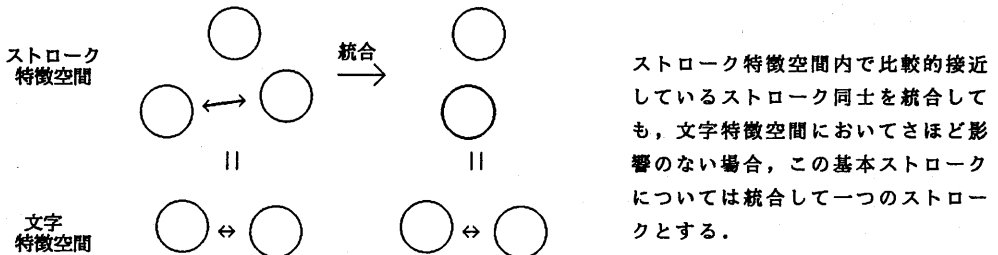


図11 基本ストローク統合のモデル

この過程では次のことが問題となる。

- ①どの基本ストローク対を統合するか
- ②どの程度基本ストロークを統合することが適当か

この問題に対しストローク識別処理のモデルを情報路に求め、情報理論をこの統合化の判断基準に応用した。これは図12のようにパタンの生起を情報源とし、ストローク識別処理を情報路とモデル化する。文字認識処理としては情報路(ストローク識別処理)より出力シンボル(ストローク識別結果)を受け取り、その情報によって入力パターンが何であったかを決定するわけである。ここで当然、文字認識処理より情報路(ストローク識別処理)に対する要求は、出力シンボル(ストローク識別結果)を1つ受け取ることにより平均して得られる情報量(相互情報量: mutual information)が十分大きいことである。

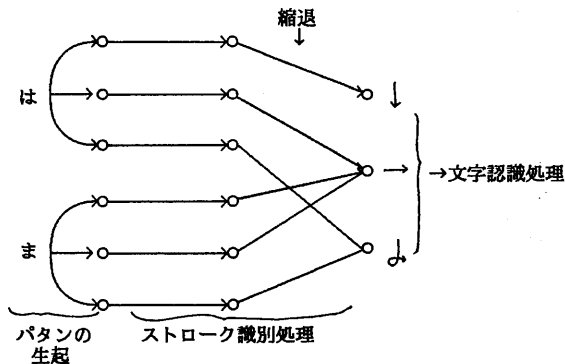


図12 ストローク識別処理の情報路モデル

これを定式化すれば、 r 個の入力シンボル a_i (この場合、仮の基本ストローク) と s 個の出力シンボル b_i (この場合、最終的な基本ストローク) に対し、

$$\begin{aligned} \text{相互情報量 } I(A; B) &= H(A) - H(A | B) \\ &= \sum_A P(a) \log \frac{1}{P(a)} - \sum_{A, B} P(a, b) \log \frac{1}{P(a|b)} \end{aligned}$$

A : 入力アルファベット、B : 出力アルファベット

P(a) : シンボル a の生起確率

P(a, b) : シンボル a とシンボル b の結合事象確率

P(a|b) : 出力 b を受信したという条件の下での入力 a の条件付き確率

H(A) : A の事前エントロピー

H(A | B) : A の B に対する事後エントロピー

を大きくするという条件である[10]。1つの入力シンボル a_i を指定するには平均として H(A) ビット必要であるに対し、その入力によってどのシンボルが出力されたか既知となった後では、1つの入力シンボルを指定するのに平均として H(A|B) ビットが必要となるだけである。すなわち、出力シンボルが既知となることによって、平均として [H(A) - H(A/B)] ビットの情報が得られると考えられる。この考えを構造解析的パタン認識に応用する。

しかし、図12に示すように十分類似した基本ストロークは統合し、一つの基本ストロークとすることがパタン認識の観点からは必要である。これは出力シンボル(基本ストローク)数を減少させることであり、情報路モデルにあっては縮退を意味する。これによって情報路の相互情報量は低下し、上述の要求と相反することになる。

そこで、構造解析的パタン認識におけるパタンプリミティブ決定の基準を、パタンプリミティブ数に比較した相互情報量に求めることを考えた。これは直感的には、適当なパタンプリミティブを統合しても、パタンプリミティブ抽出処理の性能(相互情報量)低下が、パタンプリミティブ数減少の効果より比較的小さい場合許容するというものである。本研究では、その具体的な評価関数として

$$\frac{\text{相互情報量}}{\text{パタンプリミティブ数}}$$

を用いた。したがって本研究におけるモデルでは、パタンプリミティブ決定の問題は生起確率が既知である入力シンボル(仮の基本ストローク)より、評価関数をより大きくする出力シンボル b_i の集合を決定する問題に帰着する。

また、相互情報量は入力シンボルの生起確率 P(a) に依存する。これは厳密には、日本語における各文字の出現確率の問題にまで発展し、容易に扱うことはできない。したがって本研究においては、近似的に各文字の生起確率は等しいものと仮定する。この仮定を認めても、なおその中で各基本ストロークの生起確率が異なるため、入力シンボルの生起確率 P(a) は異なる。

また評価関数の性質として、相互情報量はストローク識別方式に依存することがわかる。すなわち、ストローク識別方式によって、出力シンボルが既知となった後での入力シンボルの曖昧度 H(A|B) が異なるわけである。誤識別の多いストローク識別方式では曖昧度 H(A|B) が大きくなり、相互情報量 I(A; B) は低下する。逆に、理想的な100%のストローク識別性能が実現できた場合、相互情報量 I(A; B) は H(A) に等しくなりその時最大値をとる。

そして、最終的に基本ストロークが決定した時点において、入力シンボル a_i を文字辞書中に定義されている(入力しようと意図した)ストローク、出力シンボル b_i を実際のストローク識別結果とみなすと、相互情報量 I(A; B) はストローク識別処理が入力パタンから抽出した特徴の情報量とみなすことができる。これはストローク識別方式の性能評価の一つの尺度と考えられる。

実際には、第5.4節で述べた基本ストローク細分化の結果、279通りの仮の基本ストロークが得られていた。これに対し、ストローク辞書生成用パタン群2として、被験者1名より各カテゴリ(2265カテゴリ)1パタンずつ収集し、この基本ストローク統合化過程を行った。その結果、118通りの基本ストロークが得られた。そのうち、大小判定を必要とする16ストロークについては、基本ストロークを大小2通り用意するため、最終的には134通りの基本ストローク・セットを得た。これが図2に示したものである。

また、上述のように入力シンボルを文字辞書(本研究では擬似的な文字辞書)中に定義されたストローク、出力シンボルをストローク識別結果とすると、情報路モデルにおける相互情報量 I(A; B) はストローク識別処理によって入力パタンより抽出した情報量とみなすことができる。これを算出すると、本研究におけるストローク識別処理において6.7ビットという性能を得た。ただしこの時、つづけ字、筆順誤り文字は入力として除外して算出した。また、比較上 JOLIS-1 において同様の試算を行った結果、3.8ビット(経験的相違度、未学習の文字辞書を用いた)という性能を得ている。この数値を単純に比較しても大きな性能向上が図れたことがわかる。また、JOLIS-1 では1カテゴリあたり2.7表現であったのに対し、本研究では1カテゴリあたり約1.3表現である。本研究は基本ストロークの決定、ストローク辞書の生成方法の確立が主たる目的であるためだが、このことから JOLIS-2 が実現した時点ではより高い性能値が期待できる。加えて、本研究では裏ストローク情報も利用しており、その情報量も加算すると JOLIS-1 より格段良い特徴抽出方式の実現が図れたことがより明らかになるであろう。

6. おわりに

本研究における結論は次の通りである。

- ① ストローク識別処理のモデルを情報路に求め、情報理論をこのパタンプリミティブ決定の判断基準に応用した。
- ② 位相的特徴による大分類と、幾何学的特徴による詳細識別を融合させた階層的識別方式を実現した。
- ③ ストローク辞書、文字辞書の半自動生成を実現した。

しかし、今後の問題として、次のことが挙げられる。

- ① 本研究ではパタンプリミティブ決定の方法論の実践を怠りだため、用いた文字パタンデータが少ない。JOLIS-2 に組み込むには、文字パタンデータを十分に採集して、本研究の結果を精密化する必要がある。
- ② 本研究で決定した基本ストローク・セット、裏ストロークの系列によって、なお一意に表現できないカテゴリが存在するため、新たな特徴表現軸の導入または有効かつ安定的な付加条件の設定を検討する必要がある。
- ③ 極端に小さいストローク(例、“。”、“。”, etc.)を安定的に識別するために新たな処理フェーズの導入を検討する必要がある。
- ④ JOLIS-2 のストローク識別以後の処理によっては、上記の評価関数を修正する必要があるだろう。しかしそれでも、関数型が異なるだけで方法論に変更はない。

謝辞

本研究は、東京農工大学工学部数理工学情報工学科、阿刀田央一助教授、金子俊一助手を始めとする本学科の諸先生方に数々の御助言を頂いた成果である。また、JOLIS 計画に参加した五十嵐道弘氏(現・日本電気㈱)、木村慎一氏(現・キャノン㈱)、青木克郎氏(現・日本電気㈱)、真鍋俊彦氏(現・㈱東芝)、池田裕治氏(現・キャノン㈱)の各位に感謝する。

参考文献

- [1] 中川正樹, 他: “オンライン手書き文字認識システムJOLIS-1の設計と試作”, 情報処理学会日本文入力方式研究会資料3-1 (1982)
- [2] 中川正樹, 他: “オンライン手書き文字認識システムJOLIS-1の定量的評価”, 情報処理学会日本文入力方式研究会資料16-1 (1984)
- [3] 相澤正, 他: “オンライン手書き文字認識システムJOLIS-1の定量的評価(文字辞書の拡張とその効果)”, 情報処理学会第30回全国大会講演論文集4N-6, pp. 1279-1280 (1985)
- [4] 周藤安造: “統計的手法を導入した文法的パターン認識法”, 情報処理, Vol. 22, No. 8, pp. 797-806 (1981)
- [5] 久保光生, 他: “郵便番号自動読取区分機”, 東芝レビュー, Vol. 25, No. 1, pp. 11-18 (1970)
- [6] Fu, K. S.: “Syntactic Methods in Pattern Recognition”, Academic Press, New York (1974)
- [7] 小高和巳, 他: “オンライン手書き文字認識装置”, 電子通信学会論文誌, Vol. J65-D, No. 8, pp. 951-958 (1982)
- [8] 北原義典, 他: “方向余弦DPマッチングによる手書き「常用漢字」オンライン認識システム”, 電子通信学会論文誌, Vol. J64-D, No. 11, pp. 1013-1020 (1981)
- [9] 相澤正, 他: “オンライン手書き文字認識の前処理における信号処理的手法の応用”, 情報処理学会日本語文書処理研究会資料5-2 (1986)
- [10] アブラムソン: 情報理論入門, 好学社, 1969