U.D.C. 003.324.1.054:159.937.52]:681.327.5'21

電子計算機入力としての漢字認識

Chinese Character Recognition as Computer Input

中野康明* *Yasuaki Nakano*
山本真司** *Shinji Yamamoto*
中島　晃* *Akira Nakajima*
安田道夫* *Michio Yasuda*
中田和男** *Kazuo Nakata*

A research was conducted on the recognition of Chinese ideographs for enabling automatic inputting of Japanese sentenses into computers. An approach by means of pattern matching was adopted as the recognition method. To settle technical difficulties in this problem stemming from the fact that Chinese ideographs come in tremendous number of vocabularies and their patterns are highly complicated, the authors proposed (1) a method of compressing effectively the volume of Chinese ideograph information, (in which spectra of projection profiles are utilized,) and (2) a method of reducing applicant categories to be recognized (through hierarchical pattern matching.) Also, they developed a weighted correlation method which precisely differentiates between two similar ideographs. In computer simulation tests in which these methods were applied to 881 standard Chinese ideographs a recognition error of 10^{-6} and a rejection rate of 10^{-3} were recorded.

1 緒　言

電子計算機（以下，電算機と略す）の利用形態が高度化するに従って，常に数値計算だけを行なうにとどまらず文字や図形を入力することが要求され，さらに進んで「日本語文章情報」を取り扱うことが必要とされるようになった。日本語の特徴は漢字仮名交じり文で書かれることであり，将来においても漢字が全廃されるものとは考えられない。したがって，でき得れば漢字仮名交じり文のままで電算機に入出力することが望まれる。

日本語情報を電算機に入力する場合，最も問題となるのは漢字の入力である。この入力が必要となる背景を考えてみると，大きくいって二つの場合がある。その一つは，情報そのものが発生されるところでの入力であり，その場合，真に要求されるのは手書き漢字の認識である。たとえば，新聞記事の原稿の読取りとか，窓口業務での申請書などの処理である。現状では，すでに紙面に書かれてしまった手書き漢字の認識はきわめて困難である。一方，筆点運動の情報を利用するいわゆるオンライン リアルタイム認識の場合には，漢字を構成する字画の検出が比較的容易に行なえるので，教育漢字881字の認識がすでに可能となっている(1)。

漢字入力のもう一つは，いったん活字として印刷されて配布された情報の整理，記憶のための入力であり，この場合には印刷された漢字を読み取ることが必要とされる。政府発行の各種統計資料，白書類の記憶，新聞記事の選択記憶，特許公報の記憶，自動抄録などはその例である。この場合，将来そのような情報は印刷と同時に磁気テープにディジタル情報として記録され，配布されるようになるという見解もあるが，記憶しておきたい情報がすべてそのような形で入手できるとはかぎらないし，また必要なもののみを人間が選択して入力記憶するほうが便利で能率的なことも多い。

電算機ユーザーの立場から見ると，漢字入力装置としては漢字けん盤装置がほとんど唯一のものであり専門のオペレータを必要とする。したがって，漢字の入力を自動化したい希望は潜在的にかなり大きいものと思われる。

従来，印刷漢字に限定しても漢字認識はきわめて困難と考えられてきた。われわれは，将来漢字認識が必ず必要になるとの見通しのもとに，まず単一字体印刷漢字の認識を目標に研究を進めてきたが，かなりの成果を得ることができた。本稿は今までのわれわれの研究結果を集約したものである。

この間，一般社会においても漢字認識の必要性が認められ，通商産業省大型プロジェクト「パターン情報処理」にも一つの目標として取り上げられるに至っている。

2 印刷漢字認識の困難性

印刷漢字の認識を英数字のそれと比べると，その困難さは質的な問題というより量的な問題にあるといえる。

第一に認識すべき文字数（カテゴリー数）が非常に多いことである。当用漢字は1,850字であるが，通常の用途を考えてもほぼ2,000字が通常使用される字数である。この多数の文字を相互に区別するために文字パターンが複雑になっており，1文字を表現する情報量も大きくなっている。

たとえば，英数字36文字を認識する場合，1文字を表現するのに16×16メッシュでほぼ十分であり，全体で16×16×36≒10^4ビットの情報量となる。一方，漢字の場合は48×48メッシュ程度は必要であり，2,000字として全体で5×10^6ビットとなり，英数字の500倍にもなる。この比はそのまま標準パターン記憶容量，認識ハードウェアに効いてくる。

第二にいわゆる「外字」の問題がある。英数字の場合，その字数が明確に決まっており，用途によって増減するのは特殊記号などわずかにすぎない。これに対して漢字の場合は，当用漢字，教育漢字といった限定は一応の目安に過ぎず，用途によって文字の種類が変わり，しかも表外字が出現することを必ず想定せねばならないという困難がある。

3 印刷漢字認識の手法

3.1 印刷漢字認識へのアプローチ

単一字体印刷文字の認識手法としてパターン整合法は，安定な認識が行なえるものとして確立している。この方法は各文字に対して，各1個（複数個でもよい）の標準パターンを

* 日立製作所中央研究所　　** 日立製作所中央研究所　工学博士

設け，入力未知パターンと各標準パターンとの類似度を計算し，最大類似度を与える標準パターンのカテゴリーを認識結果とし出力するものである。この方法で，多数の対象を認識するために生ずる相互分離の低下を「対判定」によって解消し，相関手法によるかぎりでの最高の認識率を常に確保しようというのが後述する「対判定加重相関法」である[2]。

この方法を漢字に対して適用しようとすると，漢字の情報量の多いことが技術的な障壁となってくる。この難点を解決するためには次の三つの考え方がある。

(1)　漢字の情報量をなんらかの方法により有効に圧縮する。

(2)　漢字の情報量は圧縮せず，標準パターンの記憶容量の増大をがまんする。その代わり，なんらかの手法により漢字を分類し，認識の対象となる候補文字カテゴリーを減少させ，パターン整合の高速化を図る。

(3)　パターン整合法自体には手を加えず，メモリの低価格化ハードウェアの高速化に期待する。

このうち，(3)は一つの考え方ではあるが，技術突破のかぎが認識手法以外の点にあり，予測できない面がある。常識的に採用されるのは，(2)のアプローチであり，われわれもまず各種の分類法を検討した[3]。ここで検討したものは，(a)黒点数，(b)周長，(c)左側面プロファイル，(d)周辺分布の顕著なピーク，(e)ループ，枝などの位相幾何学的特徴，(f)へん，つくりなどの部分パターンの検出，(g)連長分布，(h)特殊マスクとの相関などを用いて分類できる。

これらの分類法に共通する欠点として，パターンの変形，雑音によって誤分類を生じやすい点があげられる。この欠点を避けるためには，漢字パターンの特徴を有効に圧縮して，次元数の下がったパターンを作り，このパターンについてのパターン整合により類似した文字を選び出すという方法が案出された。その考えを極限まで押し進めたものが階層的パターン整合法である。

また，圧縮した情報が有効な特徴ならば，分類にとどめず最終判定まで進めてしまってよい。この考えが周辺分布とそのスペクトルの利用である。

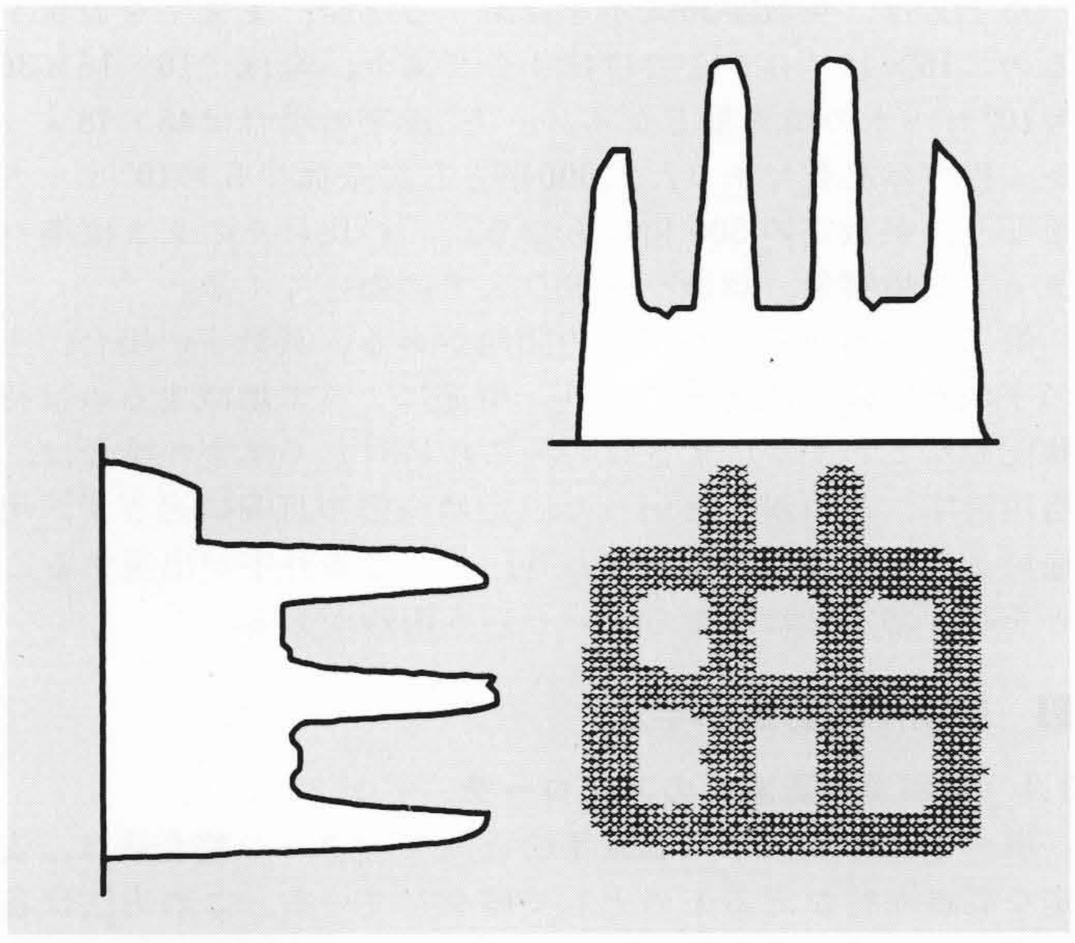

図 1　漢字の二値メッシュパターンとその周辺分布の例　　漢字
「曲」について，水平および垂直方向に投影して得た周辺分布を示す。

Fig. 1 An Example of Binary Mesh Pattern of a Chinese Character and its Projection Profiles

3.2　使用した文字サンプル

本研究で使用した文字サンプルは次のようなものである。

認識対象：教育漢字881字
字　　体：30ポイントゴシック体印刷活字(1 cm角)
観　　測：ビジコンにより光電変換，二値化した後50×50の分解能でサンプリング
二 値 化：5とおりに二値化レベルを変える

各文字を2回ずつ入力し，一方を標準パターン，他方をテスト パターン用として使用した。また，人工的なひずみパターンとして，

(1)　位置ずれパターン（上下，左右，斜め方向）

(2)　線幅変動パターン（太め・細め各1，2，3メッシュ）

(3)　線縁雑音パターン

を作成し，テスト パターン用入力として使用した。

また，一部の実験では4号タイプ文字サンプル(約4 mm角)も使用した。このサンプルでは，対象文字は9画以上の教育漢字500字である。各文字当たり4回の印字を使用し，各サンプルごとに4回二値化レベルを変えて電算機に取り込んだ。分解能は30ポイントと同様50×50であるが，実質的には1文字当たり42×42程度の分解能となっている。

電算機内に取り込んだパターン例は**図1**に示すとおりである。

3.3　各認識手法の説明

3.3.1　周辺分布とそのスペクトルの利用

漢字は主として垂直および水平の直線で構成されていることから，文字パターンの水平ならびに垂直方向への投影である周辺分布の利用が検討された[4]。

周辺分布の例は文字パターンとともに**図1**に示すとおりである。同図からも分かるように，周辺分布とは水平および垂直方向に文字パターンを積分したものである。周辺分布による情報圧縮効果は50×50メッシュの場合，ほぼ¼である。

周辺分布パターンを用いたパターン整合により，30ポイント ゴシック体印刷活字サンプルに対して99.4%以上の認識率が得られることが明らかとなった。周辺分布パターンは，投影方向の位置ずれに対してはきわめて強いが，投影方向と直交する方向の位置ずれには弱い欠点がある。このように，位置ずれに対して弱い点と，情報圧縮率がそれほど高くない点は改善されねばならない。

このような観点から，周辺分布の振幅スペクトルを検討した。周辺分布の振幅スペクトルは，周辺分布のフーリェ変換の絶対値として求めることができる。振幅スペクトルは，文字の位置ずれに対して不変であるという特長がある。また，周波数スペクトルとして分析されているため，認識に有効な成分だけを有効に選ぶことができ，雑音除去の効果とともに情報圧縮が期待できる。

予備的な検討により，振幅スペクトルの使用帯域として，水平，垂直それぞれ13チャネルを使用すればよいことが分かった。各成分を10ビットで表現するとすれば，全体で260ビットで済み，原パターンに比べてほぼ1/10に圧縮されている。認識率も，文字の二値化レベルが適当であるかぎり，99.9%以上を確保できることが示された。この認識率は周辺分布自体を使用したときよりかなり良いが，その理由として主帯域の利用による雑音の軽減と，位置ずれに不変なこと両者にあるものと考えられる。

周辺分布のスペクトルを用いる方法の欠点は，文字の線幅変動に弱いことである。この欠点を改良するために，スペクトルを補正する方法も考えられた。また，さらに認識率を高

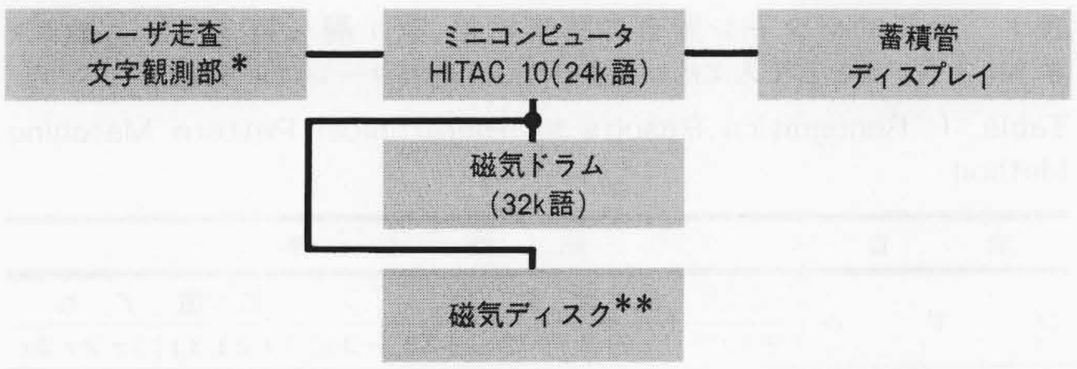

図2　漢字認識実験システム　　*は，HITAC-8959機構部，文字観測部を流用，**は，表示用パターンメモリとして使用した。

Fig. 2 An Experimental System of Chines Character Recogntion

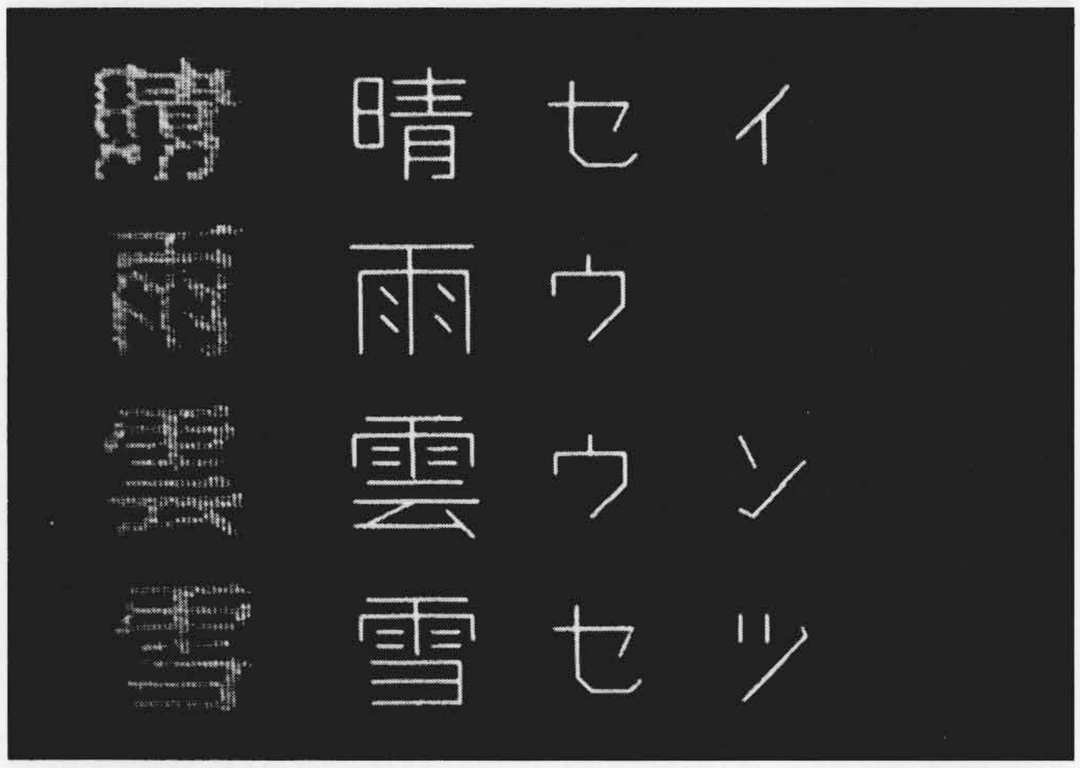

図3　認識結果の一例　　左端列は入力パターンを，第2列は認識結果の表示を，第3列は認識結果の音読み(確認用)を示す。

Fig. 3 An Example of the Recognized Output

める ために は周辺分布の投影方向を増加することも有効である。投影方向を無限に増加すると，結局二次元の振幅スペクトルに帰着する[5]。認識率と処理の単純さとのかねあいからは，水平，垂直のほかに±45度方向を加えた4方向の周辺分布を使用するのが最適と思われる。

この認識手法は処理が単純なこと，標準パターンのメモリ量が減少していることなどから，ミニコンピュータに磁気ドラムを付けた程度のシステムでもオンラインで実験が可能であり，われわれはHITAC 10に磁気ドラム，レーザ走査文字情報観測装置などを接続したシステムを実際に作製し，昭和48年10月に開催された日立技術展に出展した。システムのブロック図は**図2**に示すとおりである。このシステムは教育漢字，ひら仮名，かた仮名合わせて1,000字の読取りを行なうもので，字体としては4号明朝体タイプによるタイプオフセット印刷されたものを用いた。

図3は，この実験システムにおいて認識結果を表示したものである。左端は光電変換した文字面から1字分を切り出してそのまま表示したもの，第2行は認識結果の表示，3行め以降はふり仮名である。認識速度は1字当たり約2秒であった。このようにオンラインで1,000種にも上る漢字の認識実験を行なった例は世界でも初めての試みと思われる。

3.3.2　階層的パターン整合法

前述したように，漢字の情報量を圧縮しようとすると認識性能の低下は避けられないが，一方，漢字パターン全体の情報を使用するのでは処理速度が低下してしまう。この矛盾を無理なく解決したのが階層的パターン整合法である。

階層的パターン整合法では，前述した分類手法のうち後者のもの，すなわち，入力パターンに対して適応的に候補カテゴリーの類が形成される方法の発展とみることができる。この方法では，**図4**に示すように認識が多段に構成されており，層が進むに従って情報の精度は増加し，候補カテゴリー数は減少していくようになっている。

各層で用いるパターンは，初段のほうほどぼかされ，粗くサンプリングされており，後段のほうほど鮮鋭で細かくサンプリングされている。**図4**で示された例では第1層はメッシュ数8×8で各メッシュ点の濃度値は4ビットで表わされる。これを簡単に8×8×4と表わす。第2層でも8×8×4，第3層では16×16×2，第4層では32×32×2のパターンが使用される。

各層では各文字カテゴリーごとに標準パターンが用意される。ただし，第1層では後述するように分類用のパターンとなり，個々の文字とは違ったパターンが用いられる。

電算機内に取り込んだ漢字パターンを二次元的にぼかしたものをファックスに出力した例は**図5**に示すとおりである。このようなぼけパターンをさらに再サンプリングして使用する。

未知入力パターンに対する認識処理は，**図4**に示すように左から右へと進む。説明の都合上第2層から述べると，未知入力パターンをぼかして得た8×8×4のパターンを第2層の標準パターン(8×8×4)と順次比較し，距離の小さい順にいくつかのカテゴリーを候補として選出し第3層に送る。以下同様に続け，層が進むに従ってパターンの情報量は増加するが候補カテゴリーが減少するため，総合としての情報処理量を大幅に減らすことができる。

各層で標準パターンと入力パターンとの距離を計算するとき，二次元プロセッサ[7]を用いると高速に処理することができ，漢字認識を実用的な速度で行なうことが期待される。

次に第1層での分類手法を説明する。**3.1**で述べたように，入力文字と分類用パターンとの類似度を求め，類似度の大きな類を選び出す方法が雑音に影響されにくいと考えられる。この場合，問題となるのは分類用パターンの選び方である。この問題に対しては親近ペアを利用する方法を考案した。親近ペアとは互いに相手の文字から見て最も類似度の大きくなるような二つの文字の組である。親近ペアの平均パターンを分類用パターンとして用いる。この親近ペアの出現頻度は，パターン空間の中で分布の密なところに多く，疎なところで

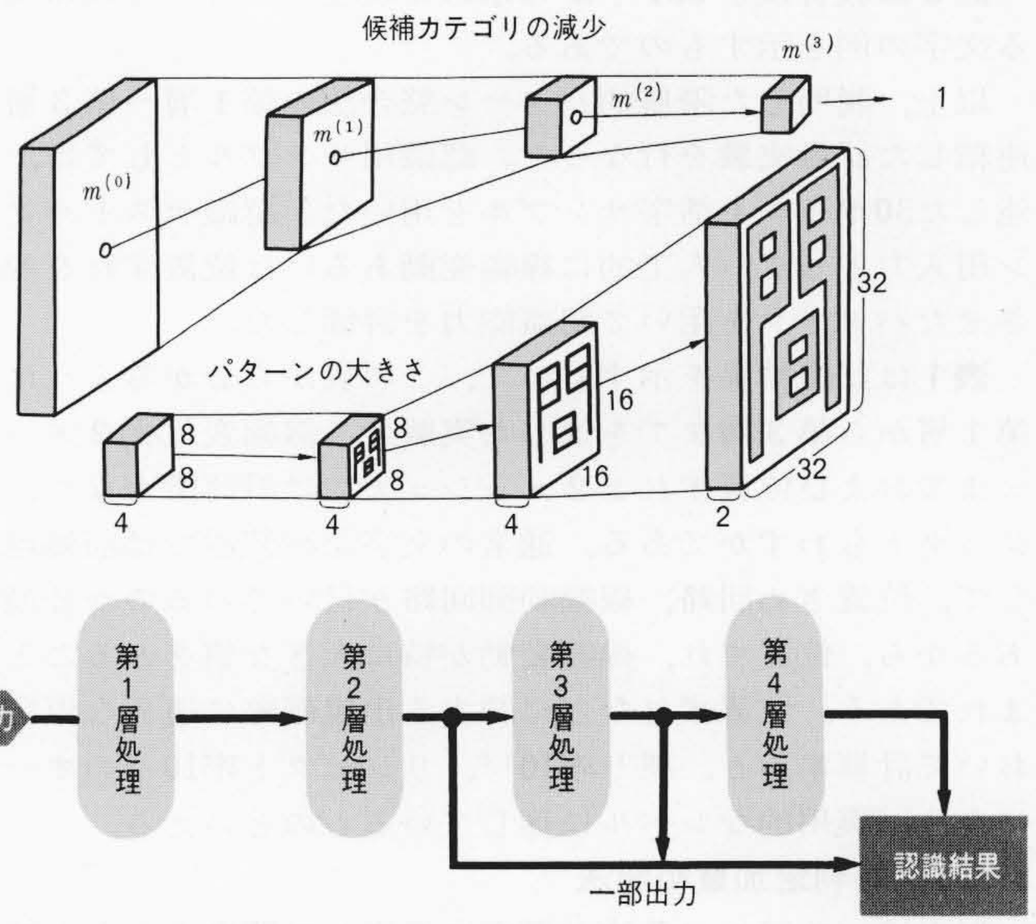

図4　階層的パターン整合法のブロック図　　粗い解像のパターンを用いて候補カテゴリーの選出を行ない，細かい解像のパターンにより最終判定を行なう。

Fig. 4 A Block Diagram of Hierarchical Pattern Matching Method

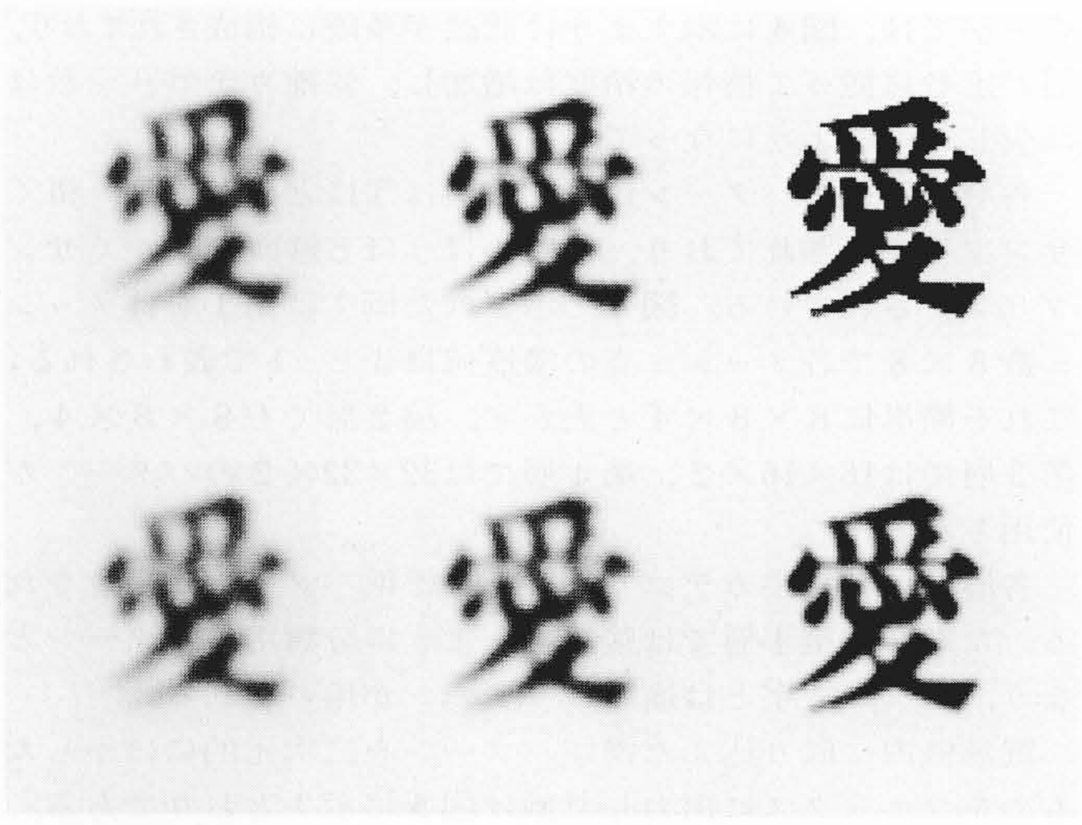

図5　ぼかしたパターンの例　　右上の標準パターンに，順次ぼかしを大きく加えて得たパターンを示した。

Fig. 5 An Example of Gradation Patterns

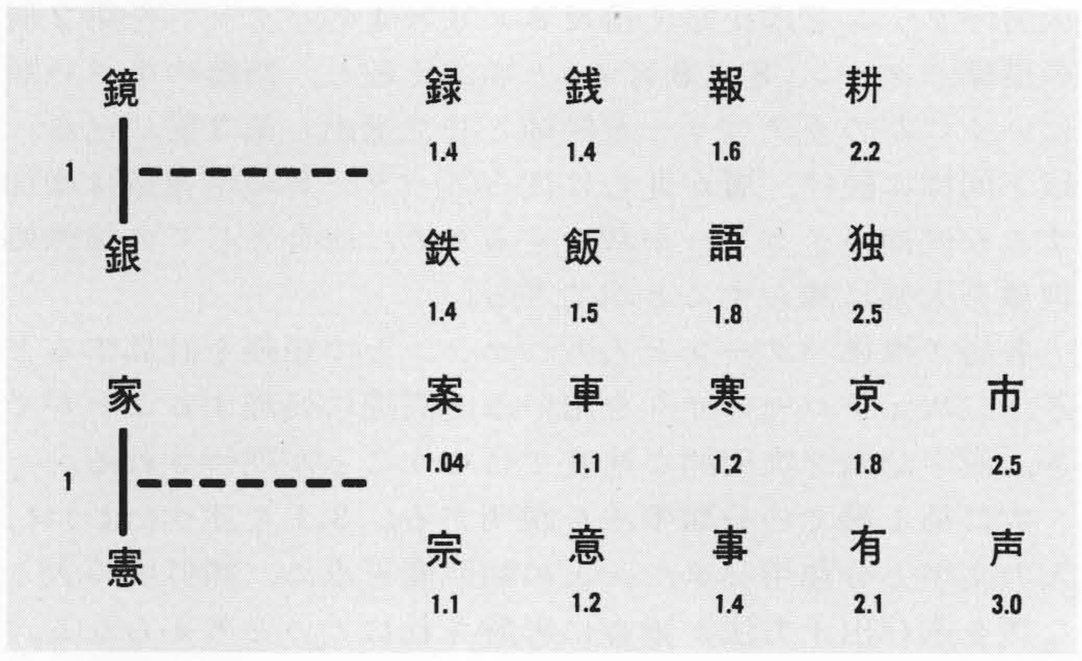

図6　親近ペアの例とそれに属する漢字の類　　各漢字の下の数字は親近ペアの距離に対する相対距離を示す。

Fig. 6 Examples of Closest Pairs and Clusters Belonging to Them

少ないという特性を持ち，パターンの出現分布に適応して分類用パターンがばらまかれるという特長を持っている。

図6は教育漢字881字から求めた親近ペアと，それに属する文字の例を示すものである。

以上，説明した階層的パターン整合法の第1層～第3層を連結した認識実験を行なった。認識用サンプルとしては，前述した30ポイント活字サンプルを用いた。認識テストパターン用入力として，人工的に線幅変動あるいは位置ずれを起こさせたパターンを用いて認識能力を評価した。

表1は認識結果を示すもので，この表から分かるように，第1層から第3層までを通した実験で，線幅変化±2メッシュまでおよび位置ずれ±2メッシュまでは誤認識がなく，リジェクトもわずかである。通常の文字認識装置では前処理として，位置ぎめ回路，線幅制御回路が付いているのが普通であるから，位置ずれ，線幅変動が特に大きな値をとることはまれである。位置ずれなどに関する出現確率に適当な仮定をおいて計算すると，誤り率10^{-6}，リジェクト率10^{-3}のオーダになり，実用的なレベルに達しているものといえる。

3.3.3　対判定加重相関法

以上述べた二つの手法が漢字を意識して開発されたものであるのに対し，対判定加重相関法[2]は印刷英数字に対して確立されたものであって，漢字認識に対しても本質的には同一の手法でよいということを主張するものである。したがって

表1　階層的パターン整合の認識結果(第1層～第3層)　教育漢字881字を対象として人工的に発生したひずみパターンにより実験。

Table 1 Recognition Results by Hierarchical Pattern Matching Method

項目	認識結果											
ひずみ	線幅変動						位置ずれ					
	$+1t$	$-1t$	$+2t$	$-2t$	$+3t$	$-3t$	$1t$	$2t$	$3t$	1τ	2τ	3τ
誤認識率(%)	0	0	0	0	0	0	0	0	0	0	0	1.1
リジェクト率(%)	0	0	0.1	0	6.1	47.0	0	0	3.2	0.5	1.9	43.6
平均誤認識率	0.0						1.0×10^{-6}					
平均リジェクト率	2.4×10^{-3}						2.0×10^{-3}					

注：t　メッシュ点間隔(サンプリングピッチ)，τ　斜め方向のメッシュ点間隔

この方法はむしろ文字認識全般を律する哲学的なものともいえ，パターン整合法によるアプローチ全体を支持するものである。

対判定法は，未知入力パターンが与えられたとき，それが二つのカテゴリーのどちらに属するかを決定し，その判定の組合せとして認識する。すなわち，未知文字がAであると結論するためには，(A，B)，(A，C)，……，(A，Z)のすべての対判定についてAに属することが言えないといけない。このとき，(A，B)の判定において，C～Zのカテゴリーが影響を与えないことが重要である。

このように，常に二つのカテゴリーのどちらであるかの判定を行なうので，対象が英数字であれ漢字であれ質的な困難さは変らず，量的な問題になる。対判定をさらに補強するものとして加重相関が考えられた。この方法は，たとえば，「問」という漢字と「間」という漢字とを区別する場合に「門がまえ」を除いた部分の重みを大きくして相関を計算することにより精密な判定を行なうものである。対判定に際し，対象の二つ以外のカテゴリーを考えなくてよいことが重要である。また，前例で未知の漢字がたとえば「向」であったとして，これが「問」と判定されるかも知れないが，「向」と「問」との対判定で「問」は否定されるのでなんら問題はない。

対判定加重相関法は，処理手順が一見複雑に見えるが，トーナメント法などの採用により，通常の相関法にわずかの処理を追加する程度で済む。**図7**は，トーナメント法により対判定を実行するプロセスを示すものである。

この方法を4号タイプ文字サンプルについて適用した。対象となる文字カテゴリーは9画以上の教育漢字500字をとった。各字16サンプル計8,000サンプルに対して，リジェクト率0.1

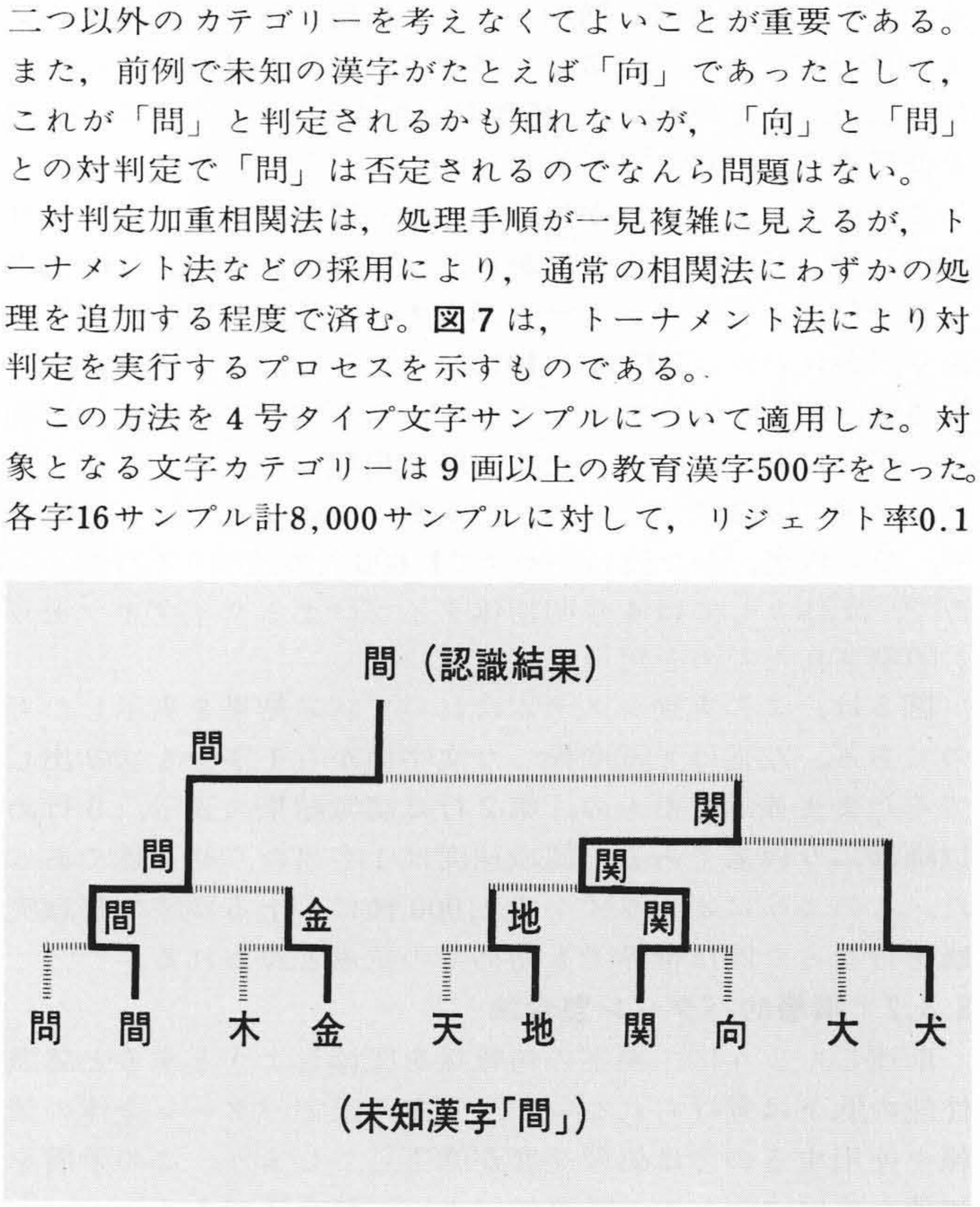

図7　トーナメント法で対判定を実行するときの過程（本図は説明用の例である）　　二つの漢字を対にし，未知漢字がどちらに属するかの判定を積み重ねて答えを出す。

Fig. 7 A Process of Pairwise Comparison by Tournament Method

%，誤り率0%の結果が得られた。この結果は，英数字と漢字とで困難の程度は同じであるという主張を裏付けた。

ここで使用した加重相関は二値パターンの相関に適用したものであるが，加重相関は前述したように一つの考え方であって，二次元のぼかしパターンにも，周辺分布のスペクトルにでも，あるいは広く一般のパターン整合全体にも応用できるものである。

経験によれば，対判定は全カテゴリーに厳密に適用しなくても，普通の相関法で少数個のカテゴリーを選んで，その中で対判定加重相関を使用すれば十分である。したがって，階層的パターン整合法の最終段の判定に適用して威力を発揮するものと考えられる。

4 結　言

3.で説明したいくつかの手法の優劣が実験結果を通して定量的に把握できるようになり，印刷漢字認識の最適手法が決定できる段階に達しつつある。その手法は，周辺分布スペクトル処理の単純さ，階層的パターン整合の全体としての認識性能の高さ，対判定加重相関法の最終判定としての有効さといったそれぞれの特長を統一した手法である。われわれは，すでにそのための一成案を得ており，単一字体印刷漢字については当用漢字を含む 2,000字程度を対象として，教育漢字について得た値以上の認識率を実験的に達成しつつある。

今後の問題として，

(1) 印刷漢字認識の最適手法の具体的な提案と，実験結果による裏付け

(2) 当用漢字を含む2,000字以上の認識実験

(3) 印字品質の劣化したサンプルについての認識性能の把握および印字品質の定量的な評価

(4) 具体的なハードウェアの開発

などが残されている。

終わりにご指導いただいた日立製作所中央研究所の渡辺所長，室井副所長，沼倉部長およびご討論いただいた漢字認識研究グループの各位ならびにプログラム作成を担当された内倉，古山，楠，奥村の各位に謝意を表わす次第である。

本研究の一部は通商産業省大型プロジェクトのパターン情報処理システムの研究開発の一環として行なったものであり，研究の機会を与えられたことおよび発表をお許しいただいたことを感謝する次第である。

参考文献

(1) 寺井，中田「手書き漢字・片仮名文字のオンライン実時間認識」信学論　56-D，312（昭48-5）

(2) 安田，門田，藤本，牧原，花野井「加重相関による単一字体印刷文字の認識」信学論　56-D，545（昭48-10）
加重相関法について詳述されており，印刷漢字に適用した例を併せ示す。

(3) 山本，中田「漢字パタンの二，三の性質」他　第12回情報処理学会全大，163（昭46-11）

(4) 中野，中田「周辺分布とそのスペクトルによる漢字の認識」信学論　56-D，146（昭48-3）
周辺分布のスペクトルを用いた認識手法について詳述されている。

(5) 中野，中田，中島「周辺分布とそのスペクトルによる漢字認識の改良」信学論　57-D，15（昭49-1）

(6) 山本，中田「階層的パターンマッチングによる漢字認識の基礎」信学論　56-D，365（昭48-6）

(7) 山本，中田，中島「階層的パターンマッチングによる漢字認識の実験」信学論　56-D，714（昭48-12）
階層的パターン整合法について総合的に報告されている。

手書き漢字・片仮名文字のオンライン実時間認識

日立製作所　寺井秀一・中田和男

電子通信学会論文誌D　56—D，5（昭48-5）

日本語情報処理において漢字入力の問題には切実なものがあり，従来の漢字テレタイプに代わる簡便な入力方式に対する要求が非常に大きい。現在まで漢字入力に関して様々な方式が発表されているが，それらを大別すると，(1)仮名・漢字変換方式，(2)けん盤入力方式，(3)パターン認識方式の三つに分類できる。(1)の方式は言語処理のソフトウェアによって行なわれる。この場合，同音異義の漢字をどのように扱うかが最も重要な問題となる。最適な漢字を一意的に指定することは非常に困難である。(2)の方式は操作に熟練を必要とするとか，多段階のけん盤操作による入力時間の問題などが指摘される。(3)のパターン認識方式は，これが実現できれば最も望ましいのであるが，現在，単一フォントの印刷漢字の認識が実験的に成功している段階であり，現実に需要が一番多い手書漢字の認識については当分実現しそうにない。

ここでは新しい漢字入力方式として，オンライン手書筆点座標入力装置（例えばランド タブレット）を使って手書きしている漢字を実時間で認識し，漢字入力を行なう方法について述べる。この方式では手書き過程を実時間で取り込むため，画の分離抽出，画数，筆順，各画の相互位置関係など認識に有効な情報が容易に得られ，その使用形態からみて，マン マシン システムに適した入力方式であるということができる。まず，手書過程において発生するペン先のup/down信号を利用して，文字を構成しているストロークを分離，抽出する。そしてこれらの各ストロークが，あらかじめ基本ストロークとして登録してあるもののいずれに相当するかを決定する。基本ストロークは「乚，冂，一，丿……」など11種類を用意している。次に入力文字のストローク数と，決定された基本ストロークのうちの特定のものとの組合せによって約 300カテゴリーの一つに大分類した後，基本ストロークのセット，シーケンス，相対位置関係「言，車，イ」などの部分パターンの有無など，個々の文字の特徴をある順序に配したDecision Tree をたどって最終文字を決定する。現在認識可能な文字は，教育漢字881字と濁点，半濁点を含む片仮名の合計952字である。入力装置のペンの使い方，文字の書き方などにわずかな教育を施すことにより90%以上の認識率が得られた。16k語のミニコンピュータ（HITAC 10）と65k語の磁気ドラムを用い，1文字の認識時間はドラムアクセスを含めて250msである。

オンライン手書き漢字認識方式は，文字を手書きするという性質上，大量一括の漢字入力には不向きであろう。しかし認識結果が直ちにフィードバックされるため，誤認識文字に対する訂正，再入力が可能であり，随時だれもが少量の文字を入力し，処理結果を得たいような場面に適している。例えば，情報検索のキー ワード入力，簡単な情報ファイルの作成時の入力手段，原稿校正システムへの適用などが考えられる。

論文抄録

階層的パターン マッチングによる漢字認識の基礎
——印刷漢字認識の研究——

日立製作所　山本真司・中田和男

電子通信学会誌　56—D，365（昭48-6）

印刷漢字を認識するアルゴリズムに関してはまだ本格的研究に乏しく，その実現の可能性は明らかでなかったが，階層構造パターン マッチング法と名付けられた新たなアルゴリズムによりその可能性を確認した。この考え方は，文字パターンの粗い解像によるマクロな観察から細かい解像によるミクロな観察への階層的パターン マッチング構造をとることにより，文字パターンを何段階かにぼかして，ぼけの多い階層からぼけの少ない階層へ，層を経るごとに文字の属するカテゴリーの候補を減らして認識する方式である。

教育漢字 881字を対象にした場合には，第1層～第4層の4層構造であれば十分であり，第1層では入力パターン及び比較すべき標準パターンを4×4絵素（1絵素4ビット）で表現し，以下第2層では8×8絵素，第3層では16×16絵素，第4層では32×32絵素で表現している。

この方式における処理の流れは次のとおりである。未知パターンXは初め第4層処理に対応する32×32絵素で光電変換，量子化されて一時記憶される（これを$X^{(4)}$と表現する)。この$X^{(4)}$に適当なぼかし操作を行なった後，等間隔サンプリングを行なって16×16絵素から成る第3層用パターン$X^{(3)}$を作る。以下同様にして第2層，第1層用のパターン$X^{(2)}$，$X^{(1)}$を作る。得られた$X^{(k)}$（$k=1\sim4$）を使って，今度は逆に第1層から第4層へと分類，識別処理が行なわれる。各層における標準パターンを$P_l{}^{(k)}$（$l=1\sim881$）とし，$X^{(k)}$，$P_l{}^{(k)}$の絵素成分をそれぞれ$x_{ij}{}^{(k)}$，$P_{ij}{}^{(k)}$（i，j：絵素番号）とすると，$X^{(k)}$，$P_l{}^{(k)}$間の距離$d^{(k,l)}$は，

$$d^{(k,l)}=\sum_i\sum_j(x_{ij}{}^{(k)}-P_{ij}{}^{(k,l)})^2\cdots\cdots(1)$$

で定義される。第1層においては，$P_l{}^{(1)}$のすべてのパターンと$X^{(1)}$の間で(1)式を計算し，$d^{(1,l)}$値の最小のものから$m^{(1)}$個のカテゴリー名を候補として抽出する。第2層においては，第1層によって選ばれた$m^{(1)}$個のカテゴリー名に対応する$P_l{}^{(2)}$のみを取り出し，これと$X^{(2)}$との間で(1)式を計算し，候補を更に$m^{(2)}$個に限定する（$m^{(2)}<m^{(1)}$)。以下同様にして，第4層においては$m^{(4)}=1$，すなわち認識した答を出す。

この方式の採用により，(1)漢字認識の処理が効率良く実行できること，(2)ノイズに対する信頼性が高いこと，(3)文字の特定の性質に依存しないアルゴリズムのため，字体，字種の変更に関係なく汎用的，統一的な装置設計が可能なことなどが明らかになり，漢字認識装置実現の可能性を確かなものとした。

論文抄録

周辺分布とそのスペクトルによる漢字認識の改良

日立製作所　中野康明・中田和男，他1名

電子通信学会論文誌　57-D，1，15（昭49-1）

周辺分布とそのスペクトルを利用した文字認識手法を前回提案し，印刷漢字の認識に適用して好結果を得た。この手法は，標準パターンメモリ容量が小さく，認識速度が高速で，位置ずれに強いという特徴を有するが，文字の線幅が変化すると認識率が低下する欠点があった。

本論文は，線幅変動の影響を補正する方法を中心として，周辺分布とそのスペクトルを使用する文字認識手法のその後の発展をまとめたものである。

文字パターン$P(i,j)$の周辺分布は，

$$X(j)=\sum_{j=0}^{N-1}P(i,j)$$

で与えられる。上式は水平軸上への投影であるが，同様に垂直軸上への投影，±45度方向への投影も定義される。

周辺分布のスペクトル

$A(k)$，$B(l)$，$C(m)$，$D(n)$，

$(k,l,m,n=1,\cdots\cdots,M-1)$

は，周辺分布のフーリエ変換の絶対値として定義される。ここでMはメッシュ数Nよりも大きな整数で，スペクトルのサンプリング間隔に関係する定数である。

本論文では，これらの振幅スペクトルの類似度により文字間の類似性を評価する。

認識実験に使用した漢字サンプルは，教育漢字 881字を対象とした30ポイントゴシック体印刷漢字サンプルである。これらのサンプルを1文字当たり50×50メッシュでサンプリングし，二値に量子化したメッシュパターンを標準パターンとした。この標準パターンから，人工的に位置ずれパターン，線幅変動パターンを作成した。

これらのサンプルを用いて，まず認識に使用する帯域の上下限周波数を系統的に変えた認識実験を行ない，最適帯域を定めた。

次に，線幅変動を補正したスペクトルを用いた認識実験を行なった。線幅変動の補正は周辺分布のスペクトルに，次のスペクトルを乗ずることにより行なわれる。

$$Q(\omega)=\exp(-\sigma^2\omega^2/2),$$

$$\sigma^2=(W_C^2-W_X^2)/10.0$$

ここで，W_Xは入力未知パターンの線幅で，W_Cは補正後の基準線幅である。この線幅変動の補正により，認識率が大きく改善されることが示された。

次に，二次元振幅スペクトルを利用する方法を実験した。周辺分布のスペクトルは二次元振幅スペクトルの一部分であることが示される。±45度方向の周辺分布のスペクトルを水平，垂直方向のそれに追加したものを簡易化二次元スペクトルと呼ぶ。

周辺分布のスペクトル，簡易化二次元スペクトル，二次元スペクトルの順に認識率が高くなることが判明した。簡易化二次元スペクトルは，計算速度と認識率とのかね合いで最適なものと考えられる。

以上は単一字体印刷文字の認識であるが，複数字体印刷文字の認識手法として，周辺分布の非線形伸縮整合を利用した準二次元非線形伸縮整合を提案し，異種字体印刷漢字の認識に適用した。