



印刷文字認識における位置ずれ補正の効果*

中野 康明** 安田 道夫** 中島 晃**

Abstract

Effective methods have been proposed for the recognition of the printed Chinese characters using pattern matching approaches. Through these works, the positioning errors have been proved to cause fatal deterioration of recognition.

The simplest and most powerful way to rectify the errors may be to seek the best fit between the input pattern and the standard of each category by translating one of them in horizontal and/or vertical directions.

We applied the technique to the recognition of printed Chinese characters whose vocabulary amounts to 1850. Two algorithms belonging to pattern matching methods were used. One of them is the hierarchical pattern matching method in frequency domain and the other is the simple binary pattern matching. Applying the technique of the correction, the error rates were reduced remarkably.

The effect of the correction resulted from the remarkable improvement of intraclass similarities and the slight rises of the interclass similarities.

1. ま え が き

計算機による日本語情報処理を行うにあたって、その最大のあい路は計算機への漢字情報の入力であることは広く指摘されている。漢字入力の理想的な形態は漢字認識装置であると考えられるが、認識対象を単一字体印刷漢字に限定しても、その実現はかなり困難であると考え意見がかなり根強いようである。

最近、印刷漢字を対象としたいくつかの認識手法が発表されている。その主なものは、中野らの周辺分布のスペクトルを利用する方法¹⁾、山本らの階層的パターン整合法²⁾、安田らの加重相関法³⁾、中野らの周波数領域での階層的パターン整合法⁴⁾などである。これらの結論を端的に言えば、「単一字体印刷漢字であれば、当用漢字を含む約2,000字程度に文字カテゴリー数が増したとしてもパターン整合法で認識できる」と

いうことである。

ところで、パターン整合法による文字認識においては、文字パターンの位置ずれによる影響が大きいことはよく知られている。上に述べたいくつかの手法においても、一応位置ずれに対する対策はなされている。また、印刷英数字に対して有効な認識手法として開発された複合類似度法⁵⁾は、位置ずれをはじめとする文字パターンの変動に対して同一カテゴリー内で安定な類似度が得られるように工夫したものであり、漢字パターンに対しても予備的な実験結果も報告されている⁶⁾。

これらの位置ずれ補正対策を検討するためには、次の二つの点を考える必要がある。一つは、入力パターンと同一カテゴリーの標準パターンとの類似度が位置ずれによって低下することに対する補正という点である。もう一つは、これらの補正によって異なるカテゴリー間の類似度が大きくなるかどうかという点である。前者については、ある程度解析的な取り扱いも可能であるが、後者については対象とするパターン自体の性質に大きく依存する性質のものであり、最終的に

* The effect of the correction of the positioning errors in printed character recognition by Yasuaki NAKANO, Michio YASUDA and Akira NAKAJIMA (Central Research Laboratory of Hitachi, Ltd.)

** (株)日立製作所中央研究所

は実験的な認識を必要とする問題である。

漢字のように文字カテゴリー数が多くなると、後者の問題が重要となってくる。すなわち、ある文字パターンに対して類似した文字パターンがあるかどうかは、かなり偶然的なものである。しかし、たとえば教育漢字の範囲では「遺」、「論」に類似した文字は存在しないが、当用漢字まで範囲を広げれば「遺一遺」、「論一論」のように類似した文字が存在するというように、ある文字パターンに対して類似した文字パターンが常に存在することを覚悟する必要がある。もちろん、教育漢字の範囲内でも、「休一休」、「王一玉」のように類似した文字パターンは存在し、英数字においてさえ「0一0」のような例もある。しかし、当用漢字 1,850 字を対象とすると、異なるカテゴリー間の類似性の問題が生ずる頻度はきわめて大きくなると予想される。

本論文は、以上の考察のもとに、きわめて多数の文字カテゴリーを対象とした場合の位置ずれ補正の問題を実験的に検討したものである。

2. 認識手法

5.1 位置ずれ補正手法の原理

本論文で検討する位置ずれ補正法としては、パターン整合を行う際、一方のパターンを他方に対して相対的に上下左右に移動させ、もっともよく整合する位置での類似度を両パターン間の類似度とする方法を採用した。この方法は、原理的にもっとも単純明快であり、つぎの理由によって認識能力もかなり高いことが期待される。すなわち、同一文字カテゴリー内においては、少なくともメッシュ間隔の整数倍の位置ずれに対しては完全な補正が期待できるので類似度の低下は少ないと考えられる。異なった文字カテゴリー間の類似度はあまり増えないと考えられるからである。

このような位置ずれ補正法は各種の認識手法に対して適用できると考えられるが、ここでは周波数領域での階層的パターン整合法⁴⁾と、基本的な二値パターンの重ね合せによる方法とについて実験した。

2.2 周波数領域での階層的パターン整合法

Fig. 1 に周波数領域での階層的パターン整合法の概念図を示す。この方法では、文字パターンを 2 次元フーリエ変換して、標準パターンの 2 次元フーリエ変換とのパターン整合を行う。整合は階層的に行われる。すなわち、第 1 層においては、2 次元フーリエ変換の低周波成分のみを用いて整合を高速に行う。これは、空間領域においてはぼかしたパターンを用いて整合を行うことと等価であるので、異なった文字カテゴリー間の類似度が大きくなる。したがって、第 1 層では候補カテゴリーの選出のみを行う。第 2 層以降では帯域を広げて等价的に密な解像のパターンを用いることにより精密性を上げて行く。したがって、この方法は第 1 層の高速性と最終層の精密性とを兼ね備えたものといえる。

周波数領域での階層的パターン整合法は、相関を周波数領域で行うということと、整合を階層的に行うという二つの特徴を有しているが、本論文ではとくに前者の特徴に着目し、その中で位置ずれ補正がどのように行われるかを論じている。

いま、文字パターンを次式で表す。

$$p(x, y), 0 \leq x, y \leq N-1. \quad (1)$$

ここで、 N は水平および垂直の標本化メッシュ数を表している。その 2 次元離散的フーリエ

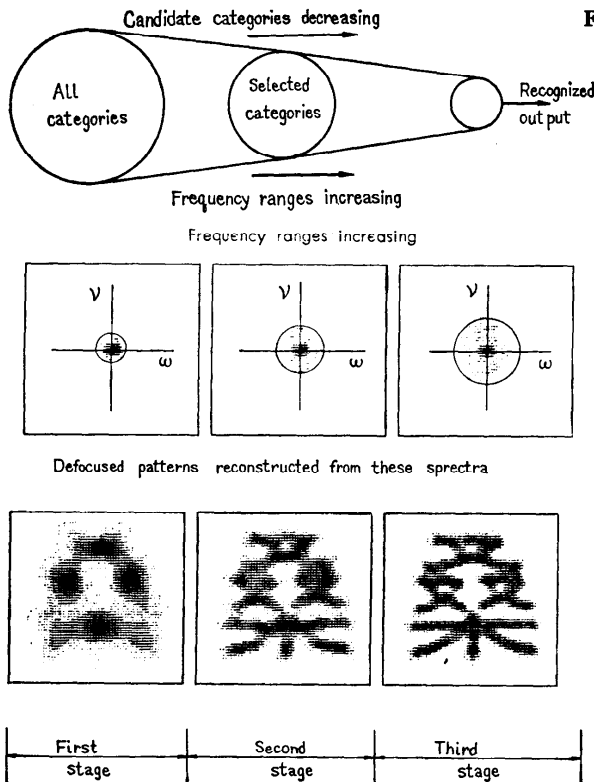


Fig. 1 Block diagram of hierarchical pattern matching in frequency domain (HMPF)

変換は

$$P(\omega, \nu) = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} p(x, y) \varepsilon^{x\omega + y\nu}, \quad (2)$$

$$\left(-\frac{M}{2} \leq \omega, \nu < \frac{M}{2}\right)$$

で与えられる。ここで、 ω, ν は整数であり、また

$$\varepsilon = \exp\left(-\frac{2\pi i}{M}\right), \quad i = \sqrt{-1}. \quad (3)$$

M は N よりも小さくない整数で、フーリエ変換を計算する周波数間隔に関するパラメータである。本論文では、 $N=50, M=64$ を使用している。

二つのパターン $p_1(x, y), p_2(x, y)$ が与えられたとき、その間の位置ずれを考えた相関（内積）は式(4)で与えられる。

$$c_{12}(\xi, \eta) = \sum_{x,y} p_1(x, y) p_2(x+\xi, y+\eta). \quad (4)$$

式(4)は周波数領域でも計算できる。

$$c_{12}(\xi, \eta) = N^{-2} \sum_{\omega, \nu} \varepsilon^{-(\xi\omega + \eta\nu)} P_1(\omega, \nu) P_2^*(\omega, \nu). \quad (5)$$

ここで、* は共役複素数を表す。

式(5)の右辺で、 $\varepsilon^{-(\xi\omega + \eta\nu)}$ は位置ずれ補正項であり、これを除いたものが周波数領域の相関である。第1層においては、式(5)の和はすべての ω, ν にわたるのではなく、低周波領域に制限される。したがって、位置ずれ補正項は比較的寄与が小さいので、省略しても実用上は差しつかえない。

二つのパターン $p_1(x, y), p_2(x, y)$ の間の類似度は、式(5)の内積を用いて定義される。

$$\rho_{12} = \frac{\max_{\xi, \eta} c_{12}(\xi, \eta)}{(c_{11}(0, 0) c_{22}(0, 0))^{1/2}} \quad (6)$$

未知入力パターンに対して、各カテゴリーの標準パターンとの類似度を求め、その値の大小によって、認識あるいは候補カテゴリーの選出を行う。

2.3 二値パターンの重ね合せ

式(4)の内積演算において、各メッシュ点での濃度は一般にはアナログ値(多値デジタル値)である。パターンのはかきを積極的にとり入れているいくつかの手法^{2,5)}では、各メッシュ点の濃度を多値で扱うことが重要な意義を有していると考えられる。

しかし、もともと文字が有している性質はその生い立ちから二値であると考えられ、十分な分解能と適当な二値化法とをもって観測すれば、二値メッシュパターンで十分表現できるはずである。その意味で、二値メッシュパターンの重ね合せは、もっとも基本的な手法といえる。

二値パターンを使用する場合の類似度は、式(4)および式(6)を用いて求められることは、一般の場合と異なる点はないが、内積を求めるときの演算が各メッシュ点ごとの論理積*を求めて、その中の「1」の個数を計数すればよいので、ハードウェアとして実現するときは比較的便利である。

3. 使用した文字サンプル

本実験で使用した文字サンプルは二通りの系統のものがある。

第一のものは、当用漢字1,850字を対象としたもので、4号明朝体タイプ活字を使用してワントゥタイムカーボンで印字されている。1文字あたり8回の印字を行ったものをそれぞれ光電変換、量子化、標準化を行い、磁気テープに収録した。タイプ活字の大きさは約4mmであり、標準化ピッチを0.1mmとしたので、1文字パターンあたりの実効的なメッシュ数はほぼ42×42程度である。量子化レベルは6である。

文字パターンの切り出しはつぎのようにした。文字パターン1個のみを含む十分広い領域をとり出し、上下左右からある閾値以上の黒レベルのある位置を探索し、文字パターンの外縁とする。左右および上下の縁の中間点を文字パターンの中心とし、この点を中心とする50×50の領域を切り出して使用する。

1文字について、8個のサンプルが使用できるが、そのうち4個のサンプルを各メッシュ点ごとに重ね合わせて平均したものを標準パターンとした。残り4サンプルを未知入力パターンとして使用した。なお、標準パターン、未知パターンともに中間的なレベルで二値化したものを使用した。Fig. 2(次頁参照)に、使用した文字サンプルの二値化メッシュパターンの例を示す。

第二のサンプル集合としては、電子技術総合研究所で集積し、公開している文字サンプルデータベースのうち、印刷漢字の部分**の一部を借用して使用した(以下、本論文ではこれを電総研データと略称する)。

電総研データで対象としている字種は、新聞印刷において高い頻度で使用される2,184字であり、漢字、仮名文字、英数字、記号などを含む。そのうち、第一

* 論理積の代りに排他的論理和を用いることも考えられるが、理論的にも実験的にも大差ないことが知られている。

** 電総研データのうち、印刷漢字については東京芝浦電気(株)総合研究所において作成され、電子技術総合研究所に納入されたものである。

頻度順位の 571 字については、各文字 10 サンプルが含まれているが、残りのものについては各文字 2 サンプルが利用できるだけであるので、本実験においては第一頻度順位の 571 字を使用した。

字体としては、新聞用 8 ポイント活字明朝体、ゴシック体および特許公報用 9 ポイント活字明朝体、ゴシック体の 4 種類が利用できるが、そのうちもっとも条件がきびしいと考えられる新聞用 8 ポイント活字明朝体のものを使用した。

文字パターンの分解能は、実効的な文字部分がほぼ 40×40 になるように標準化され、1 パターンの大きさは 60×60 である。本実験では、60×60 のエリアのうちの中央の 50×50 の部分を使用した。濃度の量子化レベルは 64 であるが、本実験ではほぼ中央値で二値化して使用した。

すでに述べたように、1 文字カテゴリについてそれぞれ 10 パターンが収録されているが、そのうちの 6 サンプルを各メッシュ点ごとに重ね合わせて平均したものを標準パターンとし、残り 4 サンプルを未知入力パターンとして使用した。

4. 実験結果

4.1 周波数領域での階層的パターン整合法による認識実験

周波数領域での階層的パターン整合法における位置ずれ補正効果を実験的に検討した。

階層としては 2 層構成をとり、第 1 層では全カテゴリ総当りにより候補カテゴリを 30 個選出し、第 2 層ではこの候補カテゴリに対して最終判定を行う。第 1 層および第 2 層で使用する周波数帯域はつぎのようにした。

$$\text{第 1 層: } 0.5 < \sqrt{\omega^2 + \nu^2} < 3.5, \quad (7)$$

$$\text{第 2 層: } 0.5 < \sqrt{\omega^2 + \nu^2} < 7.5. \quad (8)$$

ただし、式 (7)、(8) の表すリング状の領域のうち、対称性を利用して半分だけを用いればよい。

第 1 層では周波数帯域が十分低いので、位置ずれ補正は省略した。第 2 層では、左右、上下それぞれ 1 メッシュまでの位置ずれを補正した。いいかえれば、式 (5) における位置ずれ補正量 ξ, η として

$$(\xi, \eta) = (-1, -1), (-1, 0), (-1, 1), (0, -1), (0, 0), (0, 1), (1, -1), (1, 0), (1, 1),$$

の 9 通りをとっている。

* 最大類似度の次に大きな値を次大類似度と呼ぶ。



Fig. 2 Example of digitized patterns of printed Chinese characters

Table 1 Effect of the correction of positioning error in hierarchical pattern matching method (1)

Correction of positioning error	Threshold	0.	0.1	0.2	0.3	0.4
		Do	Total of input samples	7,400		
	Substitution error rate (%)	0.11	0.	0.	0.	0.
	Rejection rate (%)	0.	0.34	0.34	0.36	0.5
Undo	Total of input sample	7,400				
	Substitution error rate (%)	0.22	0.	0.	0.	0.
	Rejection rate (%)	0.	0.35	0.38	0.39	0.54

Pattern set: Typewritten Chinese characters
Vocabulary: 1850 Chinese characters for ordinary use

第 2 層における最大類似度、次大類似度*をそれぞれ、 ρ_m, ρ_n とするとき、これらが次の条件を満たすならばリジェクトする。

$$\rho_m - \rho_n < \delta \quad (9)$$

また、式 (9) の条件が成立しないとき、 ρ_m を与える文字カテゴリを認識結果として出力する。

以上の条件のもとで、タイプ文字データおよび電線研データに対して認識実験を行った。

Table 1 は、タイプ文字サンプル 1,850×4=7,400 パターンについて認識実験を行ったときの誤認識率およびリジェクト率を、位置ずれ補正を行わないときと行ったときの両者について結果を示す。これらの実験で、式(9)のパラメータ δ を変えたときの、誤認識率およびリジェクト率の変化が示されている。

Table 1 の結果から、位置ずれ補正の効果は明らかである。位置ずれ補正を行うことにより、誤認識は減少するが、同一のリジェクト閾値におけるリジェクト

Table 2 Effect of the correction of positioning error in hierarchical pattern matching method (2)

Correction of positioning error	Threshold	0.	0.1	0.2	0.3	0.4
Do	Total of input samples	2,284				
	Substitution error rate (%)	0.	0.	0.	0.	0.
	Rejection rate (%)	0.	0.	0.	0.	0.4
Undo	Total of input samples	2,284				
	Substitution error rate (%)	0.13	0.08	0.08	0.04	0.
	Rejection rate (%)	0.	0.08	0.08	0.13	0.21

Pattern set: ETL data base
Vocabulary: 571 characters including Chinese characters and katakana

率がほとんど同一であるので、位置ずれによる誤認識は位置ずれによって異なるカテゴリーに対する類似度(妨害類似度)が大きくなるために生ずるのではなく、同一カテゴリーに対する類似度が大きく低下するために生ずると考えることができる。

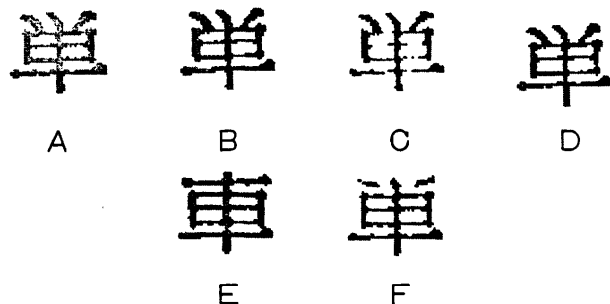
位置補正を行っても正しく認識できないものは、2メッシュ以上の位置ずれが生じているサンプルである。

Table 2 は、電総研データ 571×4=2,284 パターンに対する同様の実験結果を示している。

4.2 二値パターンの重ね合せによる認識実験

本節では、単純な二値パターンの重ね合せによる認識実験結果を述べる。この実験では電総研データを使用した。

認識実験の条件はつぎのようである。位置ずれ補正としては、式(4)のパラメータ ξ, η については $(\xi, \eta) = (-1, 0), (0, 1), (1, 0), (0, 0), (0, -1)$ の十文字型の5点をとる。階層性を導入していないので、未知入力パターンに対して全カテゴリー総当りで



A: input pattern D: Superposition of A and B (with correction)
B: standard pattern of A E: interfering pattern
C: Superposition of A and B (without correction)
F: superposition of A and E

Fig. 3 Effects of the correction of positioning errors.

Table 3 Effect of the correction of positioning error in binary pattern matching method

Correction of positioning error	Threshold	0.	0.1	0.2	0.3	0.4
Do	Total of input samples	2,284				
	Substitution error rate (%)	0.	0.	0.	0.	0.
	Rejection rate (%)	0.	0.	0.	0.05	0.07
Undo	Total of input samples	2,284				
	Substitution error rate (%)	0.35	0.26	0.21	0.17	0.13
	Rejection rate (%)	0.	0.13	0.35	0.61	0.70

Pattern set: ETL data base
Vocabulary: 571 characters

最大類似度、次大類似度を求める。リジェクトの規則は式(9)を使用する。

Table 3 に、位置ずれ補正を行ったときと行わないときの両者について、誤認識率とリジェクト率を示している。

Table 3 の結果に示されているように、位置ずれ補正を行わない場合には誤認識が生じており、しかもこの誤認識を完全にリジェクトするためには、リジェクトのパラメータ δ を十分大きくとらねばならないのでリジェクト率が1.4%とかなり大きくなってしまふ。これに対して、位置ずれ補正を行った場合には、誤認識は生じていないから、リジェクト率も0となる。

Fig. 3 に、位置ずれ補正を行わないとき誤認識となり、位置ずれ補正を行うことにより正認識となるようなパターンの例を示す。

Fig. 4 (次頁参照) に、位置ずれ補正を行わないとき、入力パターンと同一カテゴリーの標準パターンとの類似度の頻度分布および妨害カテゴリーの標準パターンとの類似度の頻度分布を示す。ここで、妨害カテゴリーとは入力パターンとは異なるカテゴリーで、類似度が最大となるものをいう。妨害カテゴリーの類似度が同一カテゴリーの類似度を上回るとき誤認識となる。

Fig. 5 (次頁参照) に、位置ずれ補正を行ったときの同様な結果を示している。

Fig. 4, Fig. 5 の各頻度分布について平均と標準偏差を求め、図中に示した。この値を見ると、同一カテゴリー内の類似度(ρ_0)、妨害類似度(ρ_i)の両者とも、平均としては同程度上昇する。しかし、標準偏差は ρ_0 の方がほぼ2分の1に減少しており、分布としての距離が開いている。

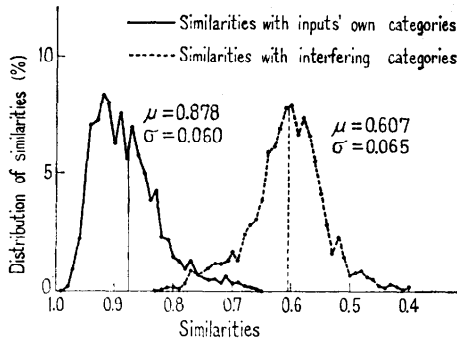
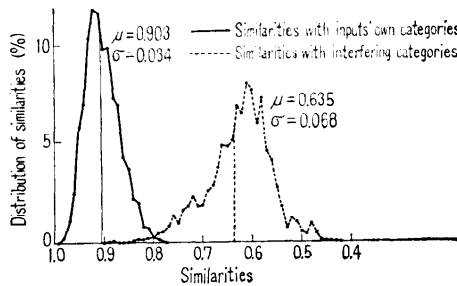


Fig. 4 Distribution of the similarities (without the correction of the positioning errors)



Similarity margin ($\rho_0 - \rho_i$)

ρ_0 : the similarity with the same category as input

ρ_i : the similarity with the interfering category

Fig. 5 Distribution of similarities (with the correction of positioning errors)

この点をさらに細く見るため、($\rho_0 - \rho_i$) の頻度分布を求めたものを Fig. 6 に示す。これらの図から明らかのように、位置ずれ補正の効果は同一カテゴリー内の類似度が大きく改善されることによってもたらされることがわかる。

Table 3 の結果は、位置ずれ補正を行うならば単純な二値パターンの重ね合せ法でも十分高い認識率が得られることを示した点で興味深い。とくに、比較的高品質の清刷りデータではあるが、8ポイント明朝体新聞活字という従来の常識では困難と思われてきたサンプルであっても、単純な重ね合せ法が適用できることは、実用上意義があると考えられる。

Table 2 と 3 とを比較すると、周波数領域での階層的パターン整合法は、位置補正を行わない場合には、単純な二値パターンの重ね合せよりもすぐれていることがわかる。この効果は、周波数領域での階層的パターン整合法においては高周波成分を切り落しているため、位置ずれに対して強くなっていることによると考えられる。

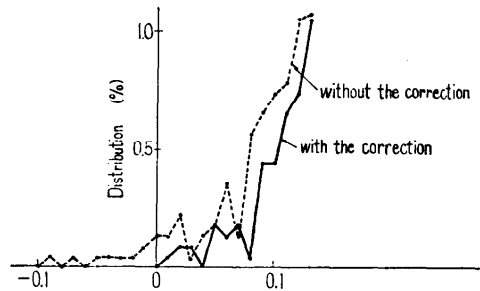


Fig. 6 Distribution of similarity margin

5. むすび

従来、漢字のように複雑かつ多字種のカテゴリーからなる対象にはパターン整合法が適用できないとする考え方もかなり根強く残っていたようである。その論拠の一つは類似した文字パターンが存在するので、誤認識が増加するのではないかという懸念にあったと思われる。本論文では、パターン整合法でとくに問題となる位置ずれについての実験的な検討を行うことにより、印刷漢字に対する適用可能性を検討した。その中で、上記の問題は信号対雑音比という観点、すなわち、位置ずれ補正による自分自身の標準パターンとの類似度の改善と、他のカテゴリーの標準パターンとの類似度（妨害類似度）との相対的な問題として把握すべきことを明確にした。

信号対雑音比という点から位置ずれ補正の効果を見ると、少なくとも実験的にはその有効性が確認されたことは、実用上有意義であると考えられる。この問題を理論的に考察することは今後の課題であるが、文字品質のモデル化⁷⁾のアプローチも一つの有望な方向であろう。

本研究で使用した電総研データベースの借用に際して、電子技術総合研究所森俊二室長および山田博三氏に多大の便宜をはかって頂いた。記して謝意を表す。

日頃御指導、御べんたつ頂く日立製作所中央研究所川崎部長、中田主管研究員、プログラム作成に御助力を頂いた古山嬢に謝意を表す。

本研究は通商産業省工業技術院大型プロジェクト「パターン情報処理システムの研究開発」の一環として行なわれたものである。

参考文献

- 1) 中野康明, 中田和男: 周辺分布とそのスペクトルによる漢字の認識, 電子通信学会論文誌, Vol. 56-D, No. 3, pp. 146~153 (1973).

- 2) 山本真司, 中田和男: 階層的パターンマッチングによる漢字認識の基礎, 電子通信学会論文誌, Vol. 56-D, No. 6, pp. 365~372 (1973).
- 3) 安田道夫, 門田彰三, 藤本好司, 牧原洋, 花野井歳弘: 加重相関による単一字体印刷文字の認識, 電子通信学会論文誌, Vol. 56-D, No. 10, pp. 545~552 (1973).
- 4) 中野康明, 安田道夫, 中田和男: 周波数領域での階層的パターン整合法による漢字認識, 電子通信学会論文誌, Vol. 58-D, No. 2, pp. 94~101 (1975).
- 5) 飯島泰蔵: マッチング法による文字認識の理論, 電子通信学会誌, Vol. 56, No. 5, pp. 668~677 (1973).
- 6) 平井彰一, 森健一: 漢字認識における単純類似度法と複合類似度法の比較評価, 昭和 49 年度電子通信学会全国大会, 1554 (1974).
- 7) 中野康明, 中田和男: 文字パターン品質の類似度に及ぼす影響, 情報処理学会創立 15 周年記念大会, 31 (1974)
(昭和 50 年 3 月 14 日受付)