

方向性特徴を用いない文字認識手法の検討

Study on Character Recognition without Directional Feature

北村浩治 二階堂真理恵 中島由美 安田道夫
Koji Kitamura Marie Nikaido Yumi Nakashima Michio Yasuda

要旨

文字図形からどのような特徴量を取り出すかは、OCR(文字認識装置)の性能を決定するひとつの要素である。文字図形から方向性成分を特徴量として取り出し、これを認識に利用することは、有効な手法のひとつとして考えられるが、方向性特徴を用いずに高い認識率を得る手法を探すため、本稿では方向性特徴を用いない(スカラー量による)文字認識手法の検討を行った。

具体的には、境界領域処理とぼけ処理を用いた場合の相関法による認識シミュレーションを行い、その結果について検討を行う。

1 はじめに

文字認識を含めたパターン認識の分野では、その認識対象となる図形から、何らかの特徴量を抽出し、利用することが多い。

相関法やsvm法などの認識手法だけでなく、認識対象からどのような特徴量を取り出し利用するかは、認識率を決定づける重要な事項である。

文字認識を行うために取り込まれた(オンラインでない)画像情報は、2次元平面上の2値あるいは多値のビットマップのような画素の集合(値はスカラー量)である。

方向性特徴を利用する文字認識では、このスカラー量から画素の連結方向などを特徴量として利用する。

認識のさいに方向成分を使用すれば、認識性能が向上することは、実験結果などから明らか[2, 9]であるが、方向成分に依らない方法をとっても99%以上の認識率を得ることができている。[3, 4]

本稿では、方向性特徴を用いない文字認識手法に対して、いくつかの処理を行うことによってさらなる改良が行えるかどうかを実際のシミュレーション実験によって検討する。なお、本稿で行う認識シミュレーション実験には、電子技術総合研究所(現、産業技術総合研究所)手

書き文字データベース ETL6 の数字部分を使用した。

2 文字認識のための相関法

2.1 相関法

相関法による文字認識手法は、直感的で計算も簡単な手法である。相関法では、各文字のカテゴリを代表する図形である関数(標準パターン)と認識対象となる文字図形を表現する関数(未知パターン)とを重ね合わせて、類似度(相関係数)を求める。

標準パターンと未知パターンをそれぞれ関数 f と g とすると、両者の類似度 S は、相関係数の式

$$S_{f,g} = \frac{(f,g)}{\|f\| \cdot \|g\|} \quad (1)$$

で計算される。

ここで、 $\|f\|^2 = \sum_{x,y} f(x,y)^2$ 、 $\|g\|^2 = \sum_{x,y} g(x,y)^2$ 、および $(f,g) = \sum_{x,y} f(x,y) \cdot g(x,y)$ である。

相関法による文字認識では、ひとつの未知パターンに対して、認識対象となるカテゴリを代表する標準パターンのすべてとの類似度を求める。そして、類似度を求めた結果、類似度が最大となった標準パターンのカテゴリをその文字として認識する。

たとえば、数字の場合であれば、0 から 9 までのカテゴリを代表する標準パターンがあり、類似度を求める計算は少なくとも 10 回は必要となる。

個別文字が決まった位置に、標準パターンと完全に同じものが表現されていれば、式 (1) の値は 1 となり 100% の認識ができる。

しかしながら、特に手書き文字などは、たとえ枠内にきちんと清書された文字図形であっても、位置、大きさ、傾きには、各文字ごとで多種多様であり、さらにノイズが含まれていることがあるため、ただ単に重ね合わせるだけの直接相関法では理想的な認識結果は得られそうもない。

ひとつのカテゴリに対してさまざまな形状をもつ手書き文字図形に対応できる手法としては、次の 2 つが考えられる。

- 必要となる文字を標準パターンとして記憶する
- 標準パターンまたは未知パターンの形を変えて対応する

前者の手法としては、サポートベクターマシン (svm) がある。svm では、サポートベクターと呼ばれる標準パターンにあたるデータをモデルとして採用 (記憶) することで、シミュレーション実験でも高い認識率を得ることができている。[1]

後者の手法としては、次に挙げる摂動相関法がある。

2.2 摂動相関法 (総当たり法)

先に述べた理由などから、ただ単に重ね合わせるだけでは、十分な認識結果を得ることができない。認識対象となる未知パターンとカテゴリを決定する標準パターンとの間にあるこのバラツキを補正するために、変形パラメータを少しずつ変えて (摂動) アフィン変換を行い、変形を行いながら、類似度を求める手法が摂動相関法である。

のちに述べるシミュレーション実験では、摂動パラメータは、軸の傾き (Angle), 伸縮 (Size), 並行移動 (Position) の 3 つとし、それぞれ、縦方向と横方向に対して単独に変えられるようにし、未知パターンの文字図形に対して変形を行う。摂動パラメータは次のとおりである。

- 座標軸の傾きの摂動

傾きの摂動は、直交座標 (左手系で縦軸を下向きとする) を基準として、縦軸と横軸をそれぞれ独立に一定の角度ずつ摂動させる。傾きの単位は 7.5° とする。

±3 単位の摂動で最大 45° の傾きをかけることができる。

- 座標軸の伸縮の摂動

伸縮の摂動は、 20×20 の作業領域に 14×14 の大きさで表現したものを基準として、縦方向と横方向に 1 画素単位で変わる。

±3 単位の摂動で、最小の大きさは 11×11 、最大の大きさは 17×17 となる。

- 平行移動の摂動

並行移動の摂動は、1 画素を単位として縦軸と横軸に摂動させる。形状は変わらない。

今回行う摂動相関法の処理の手順は、次の通りである。

1. 2 値化と切り出し

未知パターンとして、データベースから画像データを取り出す。この画像は大きさが 64×63 の多値モノクロ画像である。これに、簡単なノイズ除去を施し、2 値の画像にする。文字のサイズは一樣ではないので、まず縦方向、横方向にヒストグラムをとり、この情報を使って文字の部分だけを切り出す。

2. 傾きの摂動

摂動単位に従って縦軸と横軸を傾けて、アフィン変換による文字画像の変形を行う。

3. 大きさの伸縮による摂動

傾けた画像を、摂動単位に従って、画像の大きさを変える。次に自己相関値を求める。

4. 位置の摂動

大きさを変えた画像の位置を変える。

5. 相関法による類似度の算出

摂動させた未知パターンとあらかじめ用意しておいた各カテゴリを代表する標準パター

ンとの類似度を計算する．位置の摂動パラメータ全てについて計算が済んだら，大きさの摂動パラメータを1つ変えて，位置の摂動をしつつ類似度を求めていく．全ての大きさの摂動パラメータについても類似度の計算が済んだら傾きパラメータを1つ変えて，また同様の処理を行う．

6. 認識文字の決定

類似度が最大となった標準パターンのカテゴリが未知パターンのカテゴリであると認識する．

摂動相関法による処理は，摂動範囲の拡大に伴って計算量が膨大となる．直接相関を D，傾き (A)，大きさ (S)，位置 (P) で表し，乗数を摂動範囲として表したときの摂動範囲と計算量の関係は表 1 のとおりである．

摂動の範囲	変形範囲	相関計算量
D	± 0	1
ASP	± 1	729
$A^2 S^2 P^2$	± 2	15,625
$A^3 S^3 P^3$	± 3	117,649

2.3 適応摂動相関法 (山登り法)

摂動相関法で，摂動範囲を広げつつ，計算量を押えるためのひとつの改良方法として，適応摂動相関法 (山登り法) がある．

適応摂動相関法は，近傍の類似度を求める処理を逐次的に行う手法である．二次元の関数で考えると，まずは x が ± 1 の摂動範囲で y の値を求める． $x=-1, 0, 1$ のうち $x=0$ で y の値が最大となった場合は，そこで処理を終える． $x=1$ で y の値が最大となった場合は， $x=0, 1, 2$ で y の値を求め， $x=-1$ で y の値が最大となった場合は $x=-2, -1, 0$ で y の値を求める．すなわち，限定的な探索範囲で処理を行い，まわりよりも大きな値が中心にあればそこで終了，そうでない場合は，最大値を中心に据えた限定的な探索を行うという処理を繰り返す．

摂動範囲内で 2 次関数のような頂点があれば，探索の結果，その頂点に到達する．

3 境界領域処理とぼけ処理

3.1 境界領域処理

類似度の式 (1) で分数の分子にあたる部分は，標準パターンと未知パターンとの内積である．白い背景 (0) に黒 (1) で表現された文字の場合，2 次元平面上の画素同士の積をたし合わせたものとなり，標準パターンも未知パターンも黒の箇所は値をとり，それ以外ではゼロと

なる。

この式の分子だけを考えると、一方パターンで重ね合わさせられる部分が大きければ (黒の面積が広ければ)、字形が異なっても、偶然に重ね合わさってしまう機会が多くなり、結果として分子が比較的大きな値をとることが考えられる。

たとえば、文字の太さは、均一ではないので、細い線で構成された字形よりも、太い線で構成されたもののほうがこの値は大きくなりやすい。また、長さが異なる 2 つの数字の 1 どうしよりも、1 と 7 との方が偶然大きくなることなどが考えられる。

このように、本来は、字形による一致を試みるための相関計算であっても、図形の太さや濃さ、偶然の一致などにより、字形とは関係ない結果が得られることがある。

文字図形の形状情報だけを用いて、相関を行うのであれば、図形の輪郭であるその境界領域を取り出して利用することができるはずである。

2 値画像の場合、画像の各画素を注目画素に決め、次の手順をとることにより、境界領域を求めることができる。

1. 注目画素の評価

注目画素が白 (0) の場合は、白のままとする。注目画素が黒 (1) の場合は、次の周辺画素の評価を行う。

2. 周辺画素の評価

注目画素の周辺である上下左右の 4 点と、斜めの 4 点に白 (0) があることを確認する。この 8 点のいずれにも白 (0) がいない場合は、境界領域の画素ではないので、白 (0) にする。周辺に白 (0) がある場合は、次の孤立点の判定を行う。

3. 孤立点の評価

注目画素が黒 (0) で、その周辺画素に白 (0) があっても、周辺画素の全てが白 (0) の場合は、その注目画素は、孤立点である。通常、1 画素単独の点が文字の要素となることはないので、これはノイズとして採用せず白 (0) にする。そうでない場合は、境界領域なので、黒 (1) にする。

3.2 ぼけ (blur) 処理について

画像の小さい領域に対して、ガウス型の重みつき関数で畳み込み処理を行うことでぼけ処理を施した画像を得ることができる。

実際には、図 1 のようなマスクフィルタを用いて、注目画素とそのまわりの画素との計算を行い、適切な値で割って、その結果を注目画素の値にするという処理を画像全体に行うことで実現する。

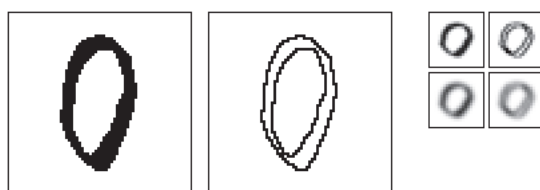
7	21	7
21	52	21
7	21	7

図 1 ぼけ処理のマスクフィルタ

フィルタは 3×3 の大きさなので、この、ぼけ処理をかけることにより、図形の大きさは上下左右斜めに対して 1 ずつ滑らかに広がった形になる。また、このぼけ処理を複数回行うことによって、より広い範囲に滑らかな裾野が広がる図形を得ることができる。

3.3 境界領域の抽出処理とぼけ処理の実施例

境界領域情報の抽出とぼけ処理の例を図 2 に示す。左側の大きめの図形は元画像で **ETL6** データベースの手書き数字、画像サイズは 63×64 (縦 \times 横) である。中央の画像は、この元図形から境界領域の情報を抽出した画像である。右端にあるの上 2 つの図形は、この 2 つの画像を 14×14 の大きさに縮小したものである。右端の下 2 つの図形は、上の 2 つの図形にぼけ処理を施したものである。



元画像 境界領域画像 縮小とぼけ処理

図 2 境界領域抽出処理とぼけ処理の例 (etl6 数字部分, 字種:0, No.3)

4 標準パターンの作成

4.1 単純平均による標準パターンの作成

標準パターンは、各カテゴリを代表するパターン図形である。単純平均による標準パターンを作成することは、母集団の平均図形を作成することである。

データベース **ETL6** の手書き数字は、1 カテゴリあたり 1,383 個の手書き数字画像がある。この中から、最後の 1 つを除く 1,382 個をサンプル番号が偶数番目のものと奇数番目のものに分けて、それぞれをたし合わせて、単純平均の標準パターンを作成する。ただし、大きさと位置のバラツキを無くすために、大きさを 14×14 にそろえ、 20×20 の領域の中心になるような処理を施す。

偶数番目のサンプルから作成した単純平均による標準パターンを図 3 に示す。



図 3 単純平均の標準パターン (偶数番目)

4.2 学習済み標準パターンの作成

各カテゴリの単純平均の標準パターンとその元データを摂動させて類似度が最大となる変形データを作成する。この変形データを再度、たし合わせて、学習済み標準パターンを作成する。

学習済み標準パターンは、単純平均の標準パターンを基準として整えられたパターン図形なので、単純平均よりも、より整った標準パターンとなる。

偶数番目のサンプルから作成した学習済み標準パターンを図 4 に示す。



図 4 学習済みの標準パターン (偶数番目)

4.3 境界領域情報による標準パターンの作成

単純平均の標準パターンまたは学習済み標準パターンを作成するときに、そのままの図形(画像)を用いずに、境界領域の情報を取り出し、この境界領域情報を用いて標準パターンの作成を行う。

今回は、元図形の境界領域情報を利用する。そのため、境界領域の情報を取り出すのは、大きさを揃える前に実施する。

単純平均による標準パターンを境界領域情報を使って作成したものを図 5 に示す。



図 5 境界領域情報を使った単純平均の標準パターン (偶数番目)

4.4 標準パターンのぼけ処理

作成した標準パターンのデータにぼけ処理を施して、ぼけのある標準パターンを作成する。ぼけ処理は、作成済みの標準パターンに対して行う(処理の最後に実施する)。

図 3 の標準パターンにぼけ処理を施したものを図 6 に示す。



図 6 ぼけをかけた単純平均の標準パターン (偶数番目)

4.5 認識シミュレーション実験

境界領域情報とぼけ処理を適用して行った相関法によるシミュレーション実験の結果を表 2 に示す。

境界領域情報とぼけ処理について適用する場合と適用しない場合の組合せで実験を行った。境界領域情報とぼけ処理は、標準パターンの作成と認識処理の両方に適用する。すなわち、ぼけ処理や境界領域の処理を標準パターンに適用した場合は、認識処理の中で未知パターンに対してもぼけ処理や境界領域の処理を適用する。

相関法は、総当たり法の場合と山登り法の場合両方について実験を行った。

実験では、すでに述べたようにデータを偶数と奇数の組に分けて行っている。たとえば、表中の“eo”は、標準パターンが偶数 (e) で、未知パターンが奇数 (o) であることを示している。

表 2 シミュレーション実験の結果 (誤読率:%)

境界領域	ぼけ処理	standard /unknown	単純平均		学習済み	
			山登り法	総当たり法	山登り法	総当たり法
なし	なし	ee	0.96	0.93	0.85	0.82
		eo	1.07	1.04	0.91	0.88
		oe	0.91	0.88	0.75	0.75
		oo	0.98	0.97	0.80	0.77
あり	なし	ee	1.11	0.10	0.93	0.90
		eo	1.29	1.24	0.98	0.93
		oe	0.98	0.96	0.93	0.90
		oo	1.11	1.04	0.94	0.85
なし	あり	ee	0.96	0.93	0.80	0.81
		eo	1.03	1.01	0.87	0.82
		oe	0.90	0.87	0.75	0.75
		oo	0.97	0.97	0.85	0.78
あり	あり	ee	1.11	1.07	0.88	0.88
		eo	1.33	1.26	1.01	0.93
		oe	1.04	1.03	0.87	0.85
		oo	1.19	1.13	1.01	0.90

5 ぼけ処理の効果について

上記のシミュレーション実験では、ぼけ処理は1回に限定して行ったが、さらに、ぼけの効果を見るため、ぼけ処理を複数回適用した場合のぼけの程度の違いによる認識率の変化をシミュレーション実験で確認した。ぼけ処理の程度(回数)は、最小は無しで、最大は5回実施する。

図7は、単純平均の標準パターンにぼけ処理をかけた例で、一番上の行の画像は、ぼけ処理なしの画像である。行が進むにしたがって、ぼけの程度を増やして、6列目は、ぼけ処理を5回行ったである。画像を見てわかるように、ぼけをかけるほど、画像としてもぼやけて見えるようになっている。

ぼけ無しの1画素は5回のぼけ処理で25×25画素の領域まで裾野が広がるが、ガウス型のマスク処理のため、最後まで元画像の形は見てとれる。

ぼけ無しで作成された画像は、この画像の濃度範囲である0から255の値で規格化されている。そのため、黒く濃い部分がある。(グレイ部分は、平均化のため)

一方、ぼけをかけた画像は、ぼけ処理のために、濃度のピークである黒く濃い部分は無くなり、少し霞んだような画像となっている。

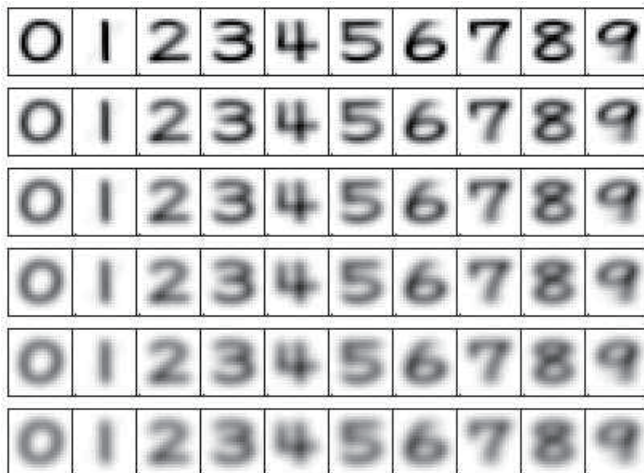


図7 ぼけをかけた単純平均による標準パターンの例(偶数番目のデータ)

ぼけ処理の程度と誤読率の関係をシミュレーション実験で確認した結果を図8で示す。

ぼけの範囲は、すでに述べた通りで、標準パターンは“単純平均”のものと“学習済み”のものを“画素データ”から作成した場合と“境界領域”から作成した場合とでそれぞれ作成した。

凡例は、“単純平均-画素データ”を“Avg”，“学習済み-画素データ”を“Lrn”，“単純平均-境界領域”を“Edge Avg”，“学習済み-境界領域”を“Edge Lrn”で表している。摂動法は、

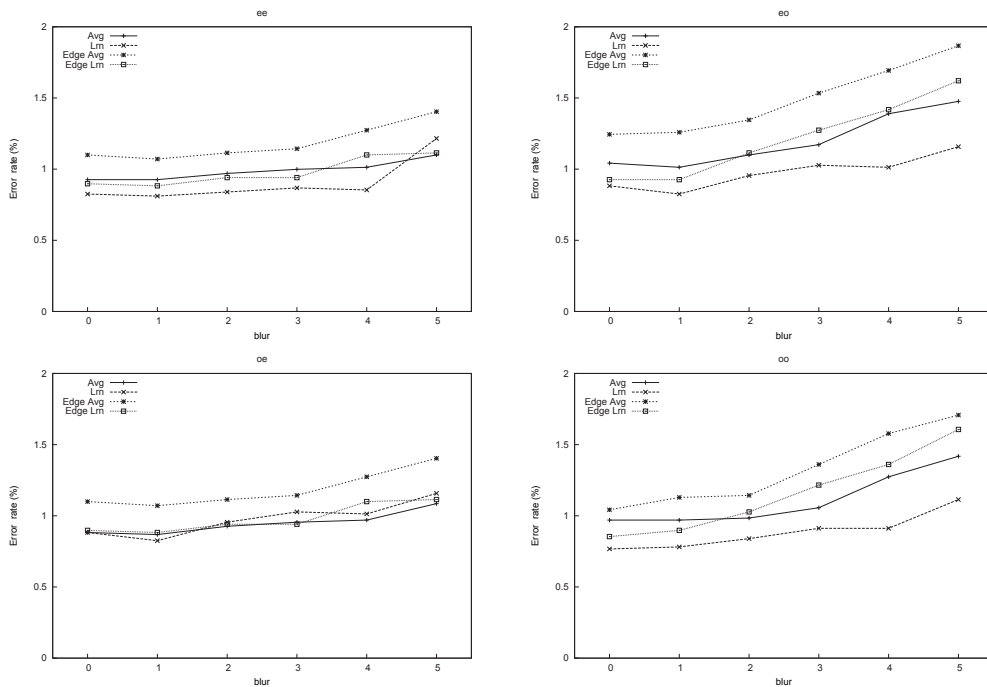


図 8 ぼけ処理の程度と誤読率の変化 (%)

すべて総当たり法で実施した。

認識シミュレーション実験で誤読となった一例を図 9 に示す。左端から見て 2 列目以降が、誤読となったサンプル文字図形である。サンプル文字図形の下に数字は、サンプル文字図形のデータ番号である。

6 検討

シミュレーション実験の結果，“学習済み標準パターン”を使った摂動相関法による文字認識は、表 2 で見ると誤読率は全て 1% 未満となり、認識率は 99% を越える認識結果が得られている。

しかし、従来法である画素情報による認識手法と比べて境界領域情報による認識手法は、残念ながら、認識率が低下する結果となった。

画素情報を利用する認識手法にぼけ処理を適用した場合は、ぼけ処理なしよりもぼけ処理ありの方が認識率が向上した。これは、ぼけ処理によってアフィン変換で生じる誤差が抑えられた効果であると考えられる。あるいは、離散的な刻みをもつ摂動法ではカバーできない変形パターンをある程度おぎなうことができた結果かもしれない。この効果の詳細な調査と実験は今

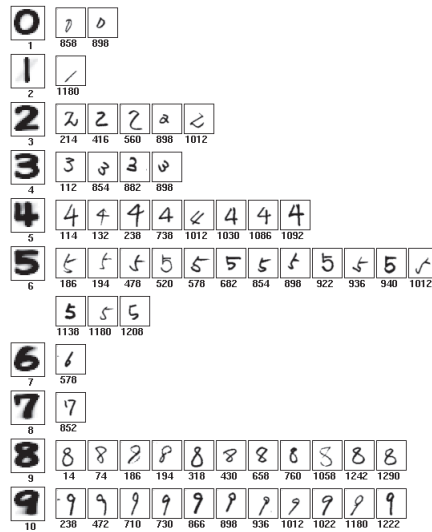


図 9 ぼけ処理 1 回 “学習済み-画素データ” oe の誤読データ

後の課題とする。

これに比べて、境界領域情報を利用する認識手法にぼけ処理を適用した場合は、ぼけ処理なしに比べぼけ処理があると認識率が低下した。この結果だけを見ると、ぼけ処理は、境界領域情報を利用する認識手法には適さないものに思えるが、認識率の低下は僅かであり、画素情報による方法では効果もみられることから、今後の工夫によっては改善の余地もあると考えられる。

総当たり法と山登り法とを比べると、ほとんどの場合において総当たり法の方が良好な結果となった。境界領域の情報とぼけ処理を使用した結果も総当たり法の方が良好な結果となったが、山登り法による大幅な認識率の低下が見られなかったことから、処理速度のことも考慮すれば、山登り法の有効性を確かめることができた。

ぼけの程度による文字認識の効果を見ると、画素情報を利用する認識手法において 1 回程度 (3×3 の範囲) のぼけでは、ぼけなしよりも良好な結果が得られ、ぼけの程度が大きくなると誤読率が高くなり、認識率が低下してしまう傾向が見られた。今回のシミュレーション実験では、1 回程度のぼけ処理が最適であることが分かった。

参考文献

- [1] 北村, 二階堂, 中島, 安田: “摂動相関法と svm 法”, 明星大学研究紀要-第 20 号, pp. 45-56, 2012 年
- [2] 北村, 二階堂, 中島, 安田: “類似度行列を用いる学習サンプルの選択法”, 明星大学研究紀要-第 20 号, pp. 23-32, 2012 年
- [3] 北村, 二階堂, 中島, 安田: “文字認識における標準パターンの学習効果”, 明星大学研究紀要-第 19 号, 2012 年
- [4] 北村, 二階堂, 中島, 安田: “摂動範囲の広域かとその効果の検証”, 信学技報, PRMU2004-163.
- [5] Yasuda, M., Yamamoto, K. and Yamada, H. : “Effect of the perturbed correlation method for optical character recognition”, Proc. 2nd ICDAR, 1993.
- [6] Ha, T. M. and Bunke, H. H. : “Off-line handwritten numerals recognition by perturbation method”, IEEE Trans, PAMI, Vol. 19, No. 5, pp. 535-539, May 1997.
- [7] Yasuda, M., Yamamoto, K. and Yamada, H. : “Effect of perturbed correlation method for optical character recognition”, J. of Pattern Recognition, Vol. 30, No. 8 pp. 1315-1320, Aug., 1997.
- [8] 二階堂, 北村, 中島, 安田: “擬似三次元特徴抽出の提案”, 信学技報, PRMU2004-164
- [9] 安田, 山本, 山田, 齊藤: “文字認識のための相関法の一改良-相補的特徴場の手書漢字への適用-”, 信学論, J68-D, 3, pp. 353-360, Mar. 1985
- [10] Watanabe, T., Kimura, Y. and Tomono, A. : “Affine-Invariant Recognition of Gray-Scale Characters Using Global Affine Transformation Correlation”, IEEE Trans, PAMI, Vol. 23, No. 4, pp. 384-395, Apr 2001.