

摂動法の検討

相関を用いた文字認識処理

Examination for the method of perturbation

character recognition using the correlation

二階堂真理恵
Marie Nikaido

北村浩治
Kouji Kitamura

中島由美
Yumi Nakashima

安田道夫
Michio Yasuda

要旨

周知の相関法を一から見直し、画素単位の摂動相関計算を行なうことで摂動範囲を細分化をはかる。本論文では、手書き文字データベースを用いたシミュレーション実験を行ない、その効果を検討する。実験には、電総研手書き文字データベース **ETL6** の手書き数字部分を使用した。また、標準パターンの学習やぼけ処理との組み合わせについても検討する。

1 はじめに

すでに報告した摂動相関法 [1],[2] から、文字認識処理との組み合わせによる効果が確認されている。相関を用いたパターン認識は、摂動処理と組み合わせることによりパターン間のズレを軽減させ、より大きな相関値を得る可能性がある。本論文では、摂動相関法の位置摂動に焦点を絞り、その摂動範囲を細分化することでより詳細な相関値を得られる可能性について検討をする。

2 特徴抽出

相関を算出するさいには、16 方向の特徴成分 [3] を用いた。文字認識における方向性ベクトルは、文字のストローク成分に沿ったものであるが、本論文における方向性ベクトルは、文字画素から背景画素への法線ベクトルとなっている。16 方向の特徴成分の詳細は、縦横斜めの 8 方向のベクトル成分を、相対成分の有無により二分したものである。図 1 に、特徴成分の実際の例を表す。図に用いた画像は、実験で使用している手書き数字『0』の平均標準パターンを用いたものである。

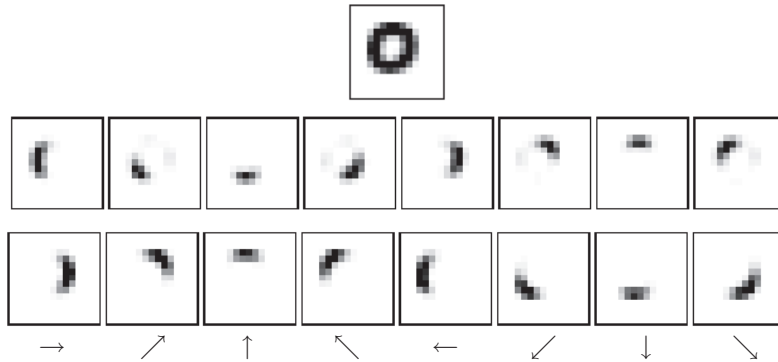


図 1：16 方向の特徴成分例

(上：全成分の総和，中：内側の成分，下：外側の成分)

3 標準パターンの学習

使用する標準パターンは，類似度を用いて使用するサンプルの選択を行なった [4]．同一パターン間の類似度を各カテゴリの各サンプルごとに算出し，その類似度が閾値を超えていない場合は，標準パターンとして採用する．また本論文では，選出したパターン以外に各カテゴリの平均をとったデータも標準パターンに含める．

また，閾値を超えていないサンプル以外にも学習処理を行なうことで，標準パターンに追加する．ここで，選出された標準パターンを学習パターンと称する．学習パターンと使用した全サンプルで認識処理を行ない，誤読が検出された場合，そのサンプルを学習パターンに追加する．これを誤読がなくなるまで繰り返すことで，学習パターンを決定する．

類似度算出に用いた式を，式 (1) に示す．

$$\mathcal{J} = \frac{(F \cdot G)}{\|F\| \cdot \|G\|} \quad (1)$$

式 (1) では，異なる 2 つのパターンを関数 F ，関数 G ，その類似度を \mathcal{J} としている．また， $\|F\|$ ， $\|G\|$ はそれぞれの関数の自己相関を表し， $\|F\| = \sqrt{F(k, j, i) \cdot F(k, j, i)}$ ， $\|G\| = \sqrt{G(k, j, i) \cdot G(k, j, i)}$ ，相互相関 ($F \cdot G$) は， $(F \cdot G) = \sqrt{F(k, j, i) \cdot G(k, j, i)}$ とする． k は特徴量， j, i はそれぞれ縦横方向の画素位置を表す．

4 摂動相関法

本論文で検討する摂動相関法は，位置に関する摂動である．これまで用いてきた位置摂動 [1],[2],[3] は，文字全体を水平および垂直に移動させるものであった．図 2 に 6×6 画素の簡単なモデルで表す．このモデルは数値『2』を表す文字で，左側をパターン A，右側をパターン B とする．また，摂動範囲は ± 1 画素とする．パターン A とパターン B の相関を算出する

さい、従来の位置摂動は図2の摂動パターンB(右下)のように画像全体を移動させる．摂動範囲 ± 1 画素の範囲で、摂動をしない場合も含めて全部で9つの摂動パターンをとる．この9種のパターンBとパターンAとの相関をとり、各摂動ごとに総和を求め、最大となるものを探索する．既知の位置摂動を用いて最大相関値 S を求める式を、式(2)に示す．関数 g, f はそれぞれ特徴成分を含める文字画像を、 $pmax$ は摂動量を表す．

$$S = \max \left(\sum_{dy=-pmax}^{pmax} \sum_{dx=-pmax}^{pmax} \left(\sum_{k=1}^{16} \sum_{j=1}^{jmax} \sum_{i=1}^{imax} g(k, j, i) \cdot f(k, j + dy, i + dx) \right) \right) \quad (2)$$

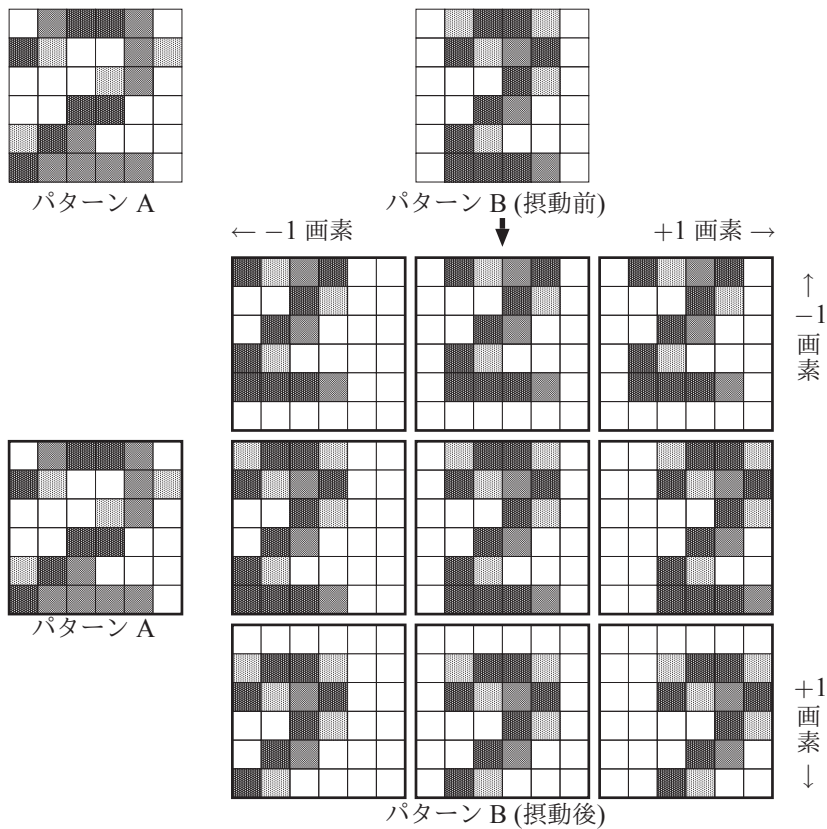


図2：位置の摂動相関法 (従来)

一方、本論文で検討する位置摂動は、画像全体ではなく1画素ごとに水平および垂直の摂動を行なう．1画素単位での相関をとることによって前述した手法よりも最大相関の選出範囲が細かくなることから、より詳細な相関値を得ることが期待できる．

説明のため、図 3 に図 2 と同様のパターンを示す．摂動範囲は、±1 画素とする．図をみても明らかなように、摂動処理を施す場合、対称画素を中心に摂動量分の隣接画素が必要となる．よって、実際に適用するさいは、文字画像の全範囲について相関計算を行なうため、摂動量分の作業領域を用意する必要がある．パターン A の対称画素 1 画素 (図 3：右、太枠部分) とそれに対応するパターン B の対称画素および隣接 8 画素の範囲 (図 3：左、1～9 にあたる) で相関計算を行ない、各画素ごとの総和を求める．摂動の結果、総和が最大となったもののみを採用する．位置摂動を用いて最大相関値 S を求める式を、式 (3) に示す．

$$S = \sum_{i=2}^{imax-1} \sum_{j=2}^{jmax-1} \max \left(\sum_{dx=-pmax}^{pmax} \sum_{dy=-pmax}^{pmax} \left(\sum_{k=1}^{16} g(k, j, i) \cdot g(k, j + dy, i + dx) \right) \right) \quad (3)$$

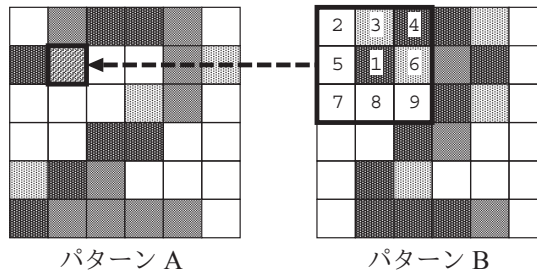


図 3：1 画素単位的位置摂動相関法

5 認識実験

実験に使用したデータは、電総研手書き文字データベース **ETL6** の手書き数字部分 (0～9 の 10 カテゴリ) である．データは 1 カテゴリあたり 1383 文字あり、これを奇数番目と偶数番目のデータに二分し、1 カテゴリあたり 961 文字の二種類のデータとして取扱う (以下、奇数番目のデータを『パターン o』、偶数番目のデータを『パターン e』と記す)．二分することでデータは各 691 文字となり、1383 番目の文字は実験では使用しない．元データは 63×64 画素のグレースケール、本実験では簡単なノイズ除去を施した後二値化をしたデータを用いている．認識実験は、学習パターンと未知パターンの相関を用いて正解カテゴリを検索する．2 種類のデータを、学習パターンと未知パターンにそれぞれ割り当て、4 通りの組み合わせで実験を行なう．表 1 に、各組み合わせを記す．

表 1：実験の組み合わせ

未知パターン \ 学習パターン	パターン o	パターン e
パターン o	oo	eo
パターン e	oe	ee

認識実験にかかる処理時間短縮のため、学習パターンに使用するデータは、あらかじめ元データから特徴量等を算出し、正規化を施したものを用意した。

正規化サイズは、文字領域が 8×8 画素、作業領域が 10×10 画素とし、作業領域の中心に文字領域を配置する。摂動相関法に適用した摂動範囲は、 ± 1 画素とした。

また本論文では、ぼけ処理を施した場合の実験も行なった。ぼけ処理に適用した重み係数を表 2 に、そして、実験に用いた環境を表 3 に記す。表 2 の $|h|$ は水平方向の、 $|v|$ は垂直方向の対称画素からの距離を表す。

表 2：重み係数

$ v \backslash h $	0	1
0	52	21
1	21	7

表 3：実験環境

MPU	Intel(R) Xeon(R) CPU 5160@ 3.00GHz
内部記憶装置用量	2GB
コンパイラ	Intel(R) C++ Compiler Version 10.0

6 実験結果

6.1 類似度

本論文で検討する摂動処理を取り入れた実験の結果、学習パターンと未知パターンの類似度が 100% を超えてしまう場合があった。図 4 に、摂動有り・学習パターン作成の閾値 95% 以上とした場合の、類似度のヒストグラムを表す。このグラフは、6910 文字全てのサンプルの正解カテゴリ類似度をカウントしたものである。実線が摂動処理のみの結果、破線が摂動処理の他にぼけ処理を施した結果となる。縦軸がサンプル数、横軸が類似度を表し、100% を 1000 とする。

図 4 のグラフから、oo, oe, eo, ee 全ての組み合わせにおいて、類似度が 1000 を超える場合があることが確認できる。ぼけ無しの場合、類似度は最大で 1400 近くになったものもあった。類似度のピークは、どの組み合わせも 1000 前後となった。

本来、類似度は同一パターンであると 100% になるというものであるが、先述のとおり本実験で得られた類似度からは 100% を超える場合が確認されている。これは、本論文で用いた位置摂動と正規化処理との組み合わせによる影響であると考えられる。

相関計算に使用している文字データは、 63×64 画素のデータから文字範囲を切り出し、その文字範囲の部分のみを 8×8 画素のデータへと正規化したものである。よって、正規化のさ

いの縮尺は、各サンプルごとに変化することになる。

また正規化処理のさい、元画像と正規化後の画像の画素同士が 1 対 1 で対応するということはほぼ起こらない。実際には、対応する座標位置の近接 4 画素に配分することとなる。近接 4 画素への配分の割合 (係数) は、その総和が各画素ごとに 1 となるように計算している。

本論文で用いた位置摂動は、1 画素ごとに対称画素と隣接画素をあわせた 9 画素との相関をとり、その値が最大となるものを採用している。9 画素のどれを採用するかは各画素の計算結果によって変化し、固定ではない。つまり、正規化のさいに最大で 1 となるよう割り当てられている係数の配分がずれてくることになる。他サンプルとの伸縮の比率もそろっていないため、異なるサンプル間では特に、正規化によるズレの影響を受けてしまう可能性がある。

この正規化によるズレを緩和する対策として、ぼけ処理を施した実験を行なった。図 4 から、ぼけ無しの場合と比べて類似度が全体的に減少していることからその効果が確認できる。ぼけ無しの場合の類似度は、最大で 1400 近くの値が確認されたが、ぼけ有りの場合は最大でも 1200 近くまでに減少している。

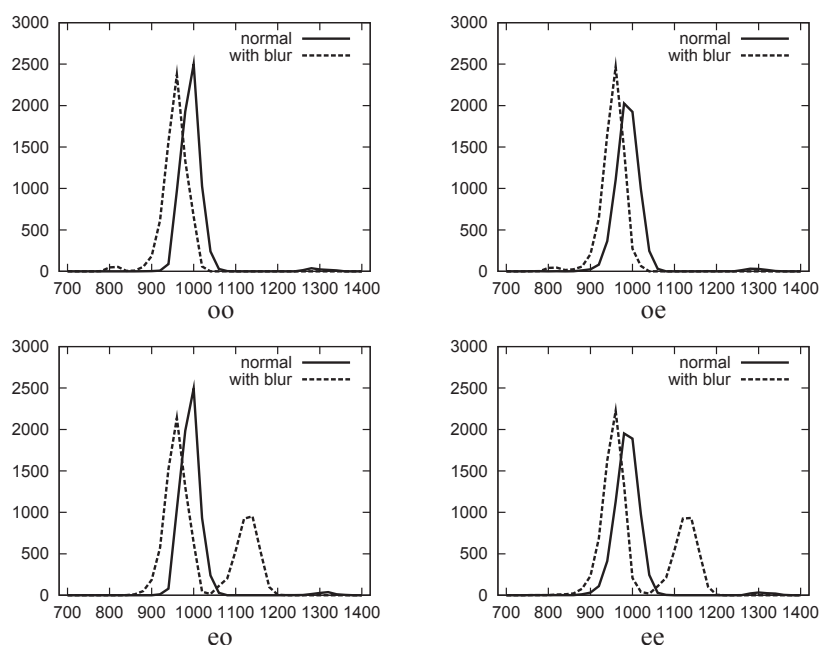


図 4：認識時の類似度ヒストグラム

図 5 に図 4 のヒストグラムをカテゴリごとに分けて表示したものを示す。グラフに使用したデータは、ぼけ処理無しで組み合わせが oo のものである。

グラフから、全カテゴリにおいて類似度 1000 前後にピークが確認できる。また、カテゴリ ” 1 ” の場合のみ類似度が 1300 ～ 1400 程度まで確認できる。これは、1 という文字画像の

正規化処理に関係していると思われる．具体的には，1 という文字は，他の文字と比較して縦長の形状をしており，正規化後も縦長の形状を保てるよう，水平方向の伸縮率を補整しているという点である．

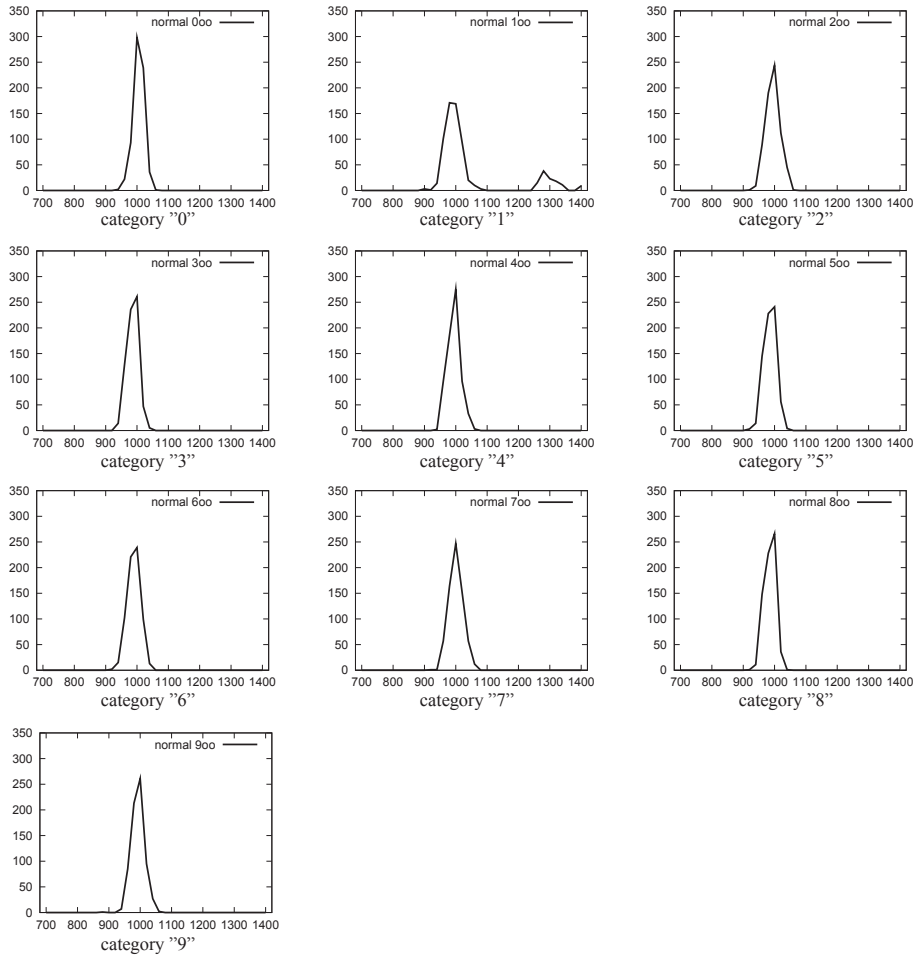


図 5：各カテゴリごとの認識時の類似度ヒストグラム (oo)

6.2 学習パターンのサンプル数と誤読文字数

使用する相関計算 (摂動処理を含む) と閾値 (3 節参照) により，学習パターンとして検出されるサンプル数は変動する．図 6 と表 4 に摂動相関を用いた学習パターンの選択数を示す．学習パターンに採用され得る全サンプル数は，平均パターンも合わせて 6920 文字である．グラフの縦軸は選択サンプル数，横軸は閾値を表す．

まずは，位置摂動有りの場合について述べる．図 6，表 4 から，偶数と奇数による違いは数

文字～十文字程度となり、ぼけ有り・無しの双方でパターンによる差に特定の傾向はみられなかった。ぼけ処理に関して注目してみると、閾値 97% を境にぼけ有りの場合の選択サンプル数から 1% ごとに 1000 文字程度の上昇がみられ、ぼけ無しの場合に比べ選択サンプル数の上昇が著しいことが確認できる。これは、先に述べた類似度のヒストグラムから説明できる。ぼけ無しに比べて、ぼけ有りの方が全体的に類似度が低い。よって、ぼけ有りの場合の方がリストに追加するさいに求められる閾値が低くなるということになる。

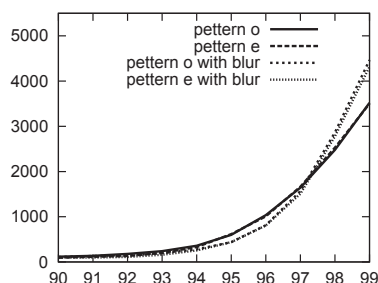


図 6：選択学習サンプル数
(摂動有り)

次に、摂動相関法を使用しなかった場合の学習パターンのサンプル数と比較をする。表 5 に、摂動無しの場合の学習パターンのサンプル数を示す。摂動相関を使用していないため最大の類似度は 100% となり、求められる閾値の値は大きく異なってくる。そのため、選択サンプル数が同程度の範囲であるデータを抜粋する。

表 5 から、パターン e と o の差が閾値 88% 以上で 100 文字以上検出されていることが確認できる。摂動有りの場合と比べて、全体的にパターン間の差が大きくなる傾向がみられる。

表 4：選択学習サンプル数 (摂動有り)

ぼけ 処理	閾値 パターン	90%	91%	92%	93%	94%	95%	96%	97%	98%	99%
無 し	o	116文字	136文字	177文字	239文字	358文字	600文字	1035文字	1643文字	2480文字	3520文字
	e	93	113	134	200	329	616	1010	1664	2530	3500
有 り	o	109	133	151	193	278	437	812	1567	2864	4465
	e	99	99	111	158	255	438	808	1509	2778	4319

表 5：選択学習サンプル数 (摂動無し)

閾値 パターン	80%	82%	84%	86%	88%	90%	92%
o	258文字	357文字	500文字	758文字	1275文字	2185文字	3602文字
e	223	303	443	680	1159	2028	3469

次に、認識シミュレーションの結果 (誤読文字数および正読率) についての結果を記す。図 7 に摂動相関法を用いた場合の誤読文字数のグラフを、表 6 に誤読文字数と正読率を示す。同じパターンの組み合わせ (oo, ee) については、学習サンプルを選択する時点で誤読が無くなるまで処理を繰り返しているため、結果を省略している。

閾値が大きいと使用する学習サンプル数が増えるため、当然認識の精度は上がる傾向にある。ただし、比較する対称が増えることでよい誤読を増やす可能性もでてくるため、必ずしも閾値が大きいほど認識率が良くなるわけではない(例として、ぼけ処理無し 組み合わせ oe の 94%～96% の場合等)。ぼけ処理の有無で比較をすると、図 7 からわかるように、ぼけ有りの場合の方が組み合わせによる誤読文字数の差が少ない傾向があることが確認できる。このことから、ぼけ処理は類似度の補整だけではなく、パターンの違いによる差を軽減させる効果もあると考えられる。

摂動相関法を用いた最良の結果としては、閾値 99% 組み合わせ oe の場合で、誤読文字数 10 文字(正読率 99.86%)となった。ぼけ処理有りの場合に関しても同様の結果であり、パターンによる差を軽減することはできても認識精度を上げる点では特筆すべき効果は得られなかった。

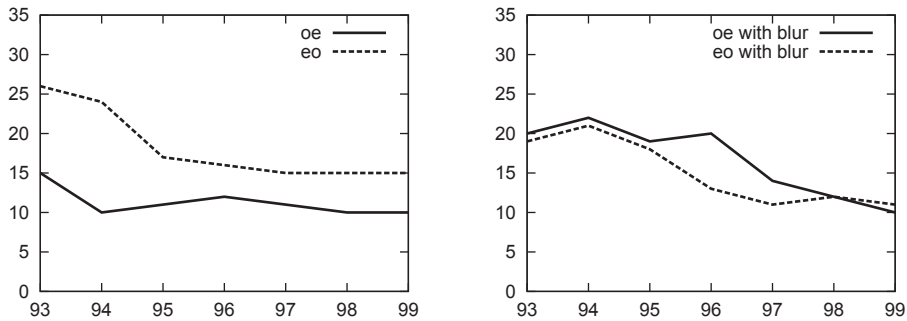


図 7：誤読文字数(左：ぼけ無し，右：ぼけ有り)

表 6：誤読文字数と正読率(摂動有り)

	ぼけ 処理	閾値 組み合わせ	93%	94%	95%	96%	97%	98%	99%
誤 読 文 字 数	無 し	oe	15文字	10文字	11文字	12文字	11文字	10文字	10文字
		eo	26	24	17	16	15	15	15
	有 り	oe	20	22	19	20	14	12	10
		eo	19	21	18	13	11	12	11
正 読 率	無 し	oe	99.78%	99.86%	99.84%	99.83%	99.84%	99.86%	99.86%
		eo	99.62	99.65	99.75	99.77	99.78	99.78	99.78
	有 り	oe	99.71	99.68	99.73	99.71	99.80	99.83	99.86
		eo	99.73	99.70	99.74	99.81	99.84	99.83	99.84

次に、摂動相関法を用いなかった場合の実験結果を表 7 に記す。先にも述べたように、閾値の範囲は摂動相関法を用いた場合と大きく異なるため、選択学習サンプル数の結果も考慮しながら比較する。

最良の結果で、閾値 92% 組み合わせ oe の場合の誤読文字数 5 文字 (正読率 99.93%) であった。選択学習サンプル数は 3602 文字 (パターン o) であり、摂動相関法を用いた場合で同等の選択学習サンプル数を持つデータは、閾値 99% の 3520 文字 (ぼけ無し パターン o) となる。この 2 つのデータを比較すると、誤読文字数 5 文字、正読率にして 0.07% 摂動無しの方が上まわっている。選択サンプル数を考慮して他の閾値の場合を比較しても、ほぼ摂動無しの場合の正読率の方が上まわる結果となった。

表 7：誤読文字数と正読率 (摂動無し)

	閾値 組み合わせ	80%	82%	84%	86%	88%	90%	92%
	誤読文字数	18文字	12文字	14文字	12文字	8文字	6文字	5文字
	正読率	99.74%	99.83%	99.80%	99.83%	99.88%	99.91%	99.93%
	組み合わせ	23	19	17	13	10	9	9
	誤読文字数	99.67	99.73	99.75	99.81	99.86	99.87	99.87
	組み合わせ	99.67	99.73	99.75	99.81	99.86	99.87	99.87

6.3 処理時間

最後に、本論文で行なった認識シミュレーション実験の処理時間について報告する。図 8 と表 8 に認識処理に要した処理時間を示す。グラフは摂動相関法を使用したもので、組み合わせは oe のデータである。横軸は閾値 (単位: %), 縦軸は処理時間 (単位: 秒) とする。また、実線はぼけ処理無しの場合、破線はぼけ処理有りの場合を表わす。

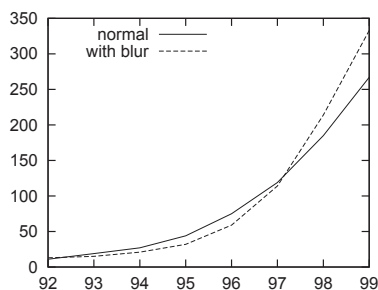


図 8：認識処理時間 (位置摂動相関使用, 組み合わせ oe)

表 8：認識処理時間 (位置摂動相関使用)

ぼけ処理	閾値 組み合わせ	92%	93%	94%	95%	96%	97%	98%	99%
	組み合わせ	11 ^S	19 ^S	27 ^S	44 ^S	75 ^S	119 ^S	185 ^S	267 ^S
無	oe	11	16	25	45	73	121	190	272
し	eo	13	15	21	32	59	114	214	333
有	oe	10	13	20	32	58	109	206	322
り	eo								

認識処理時間を検討する上で、重要になってくるのは学習パターンの文字数 (表 4 参照) である。認識処理では、選出された学習サンプル全てと未知文字パターンと相関をとるため、学習パターンが多い程処理時間を要することになる。

表 4 から閾値 97% 以下のデータでは、ぼけ処理無しの方が学習サンプル数が多い。この点を踏まえて図 8 や表 8 の処理時間を検討すると、ぼけ処理よりも主に学習パターンの文字数が処理時間に影響をおよぼしていることが確認できる。組み合わせ α 閾値 92% の場合では、ぼけ処理有りの方が 2 秒多くかかっているが、それ以外のデータでは、閾値 98% を超えるまで処理時間が逆転することは無い。このことから、認識処理においてぼけ処理による処理時間の影響は、ほぼ無いものと思われる。

7 まとめ

実験の結果から、考案する位置摂動相関法からは特筆すべき効果は確認できなかった。しかし、類似度の問題など本論文のシミュレーション実験ではまだ解決していない事項があることから、改良の余地はあると思われる。

今後の課題として、類似度の補整を含めた正規化処理の見直しや他の摂動処理との組み合わせ等、さらなる検討が必要である。

参考文献

- [1] 北村, 二階堂, 中島, 安田: ” 適応摂動相関法の提案”, 信学技報, PRMU2005-77
- [2] 北村, 二階堂, 中島, 安田: ” 文字認識のための適応摂動相関法の提案”, 明星大学研究紀要-情報学部-第 14 号, 平成 18 年 3 月
- [3] 二階堂, 北村, 中島, 安田: ” 線素特徴の拡張とその効果”, 明星大学研究紀要-情報学部-第 19 号 平成 23 年 3 月
- [4] 二階堂, 北村, 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-情報学部-第 20 号 3 月