

摂動相関法と svm 法

Perturbation and SVM for Character Recognition

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

現在、認識性能の優れた学習モデルのひとつとして知られているサポートベクトルマシン (support vector machine : svm) 法と、文字認識において、もっとも直観的かつシンプルな手法である重ね合わせ法 (相関法) とで比較と検証を行った。

相関法では、標準パターンと未知パターンとを使って認識を行う。一方、svm 法では、学習モデルと未知パターンを使って認識を行う。

同一データベースを使った相関法と svm 法との比較の他に、svm 法での学習モデルに採用されるサポートベクトルの数に着目し、同じ条件下での相関法との比較も行った。

具体的には、svm 法で採用するサポートベクトルの数が多くなった場合、やりかたによっては、相関法でも svm 法と遜色ない認識率を得られた。

1 はじめに

本稿では、パターン認識手法のうち、簡素な手法である摂動相関法と、優れた認識性能を持つ識別器として知られるサポートベクトルマシン (svm) 法との比較をシミュレーション実験で行う。筆者らが、従来から取り組んでいる摂動相関法は、アフィン変形による図形の摂動を行い、摂動した未知パターンと標準パターンの相関値を求めて、相関値が最大となるものをそのカテゴリとする、文字認識の手法である。(摂動させるのは、標準パターンでも構わない)

この摂動相関法単独での性能を評価するために、筆者らは方向性の特徴抽出などを使用しないスカラー単独での性能の評価を行っている [3, 4, 5, 6, 7].

一方、近年注目を浴びているサポートベクトル法による文字認識方法は、現在知られている多くの手法の中で認識性能の優れた方法のひとつ [2] とされている。

本稿では、摂動相関法、svm 法のシミュレーション実験を同じデータベースを使って比較を行った。また、この svm 法で得られた学習効果の一部 (サポートベクトル情報) を使って摂動相関法を行い、svm 法と比較を行った。シミュレーション実験の結果から、サポートベクトルとして採用されている情報を摂動相関法に適用すれば、直観的かつシンプルな摂動相関法でも、サポートベクトル法と遜色ない認識結果を得られた。

2 相関法

2.1 摂動相関法

文字認識のための相関法では、文字図形を表現する関数(未知パターン)と各文字のカテゴリを表現する関数(標準パターン)の類似度(相関係数)を求め、類似度が最大となる標準パターンが代表するカテゴリを未知パターンのカテゴリと認識する。

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。

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

アフィン変形の要素である、座標軸の傾き (Angle:A), 大きさの伸縮 (Size:S), 平行移動である偏移 (Position:P) の3つのパラメータを変えることによって摂動を行う。

アフィン変形は、標準パターンと未知パターンのどちらに対して行っても構わない。

摂動を伴わない直接相関を D とし、傾き (A), 大きさの伸縮 (S), 偏移 (P) の摂動をそれぞれの乗数として表した場合、摂動範囲を ± 3 まで拡張したときの相関計算量は表1のとおりである。

表1で見ると、摂動相関法を総当たりで行うと、その摂動範囲の拡大に伴って、相関計算量は膨大になる。実際、摂動範囲が最大となる ± 3 単位では、直接相関法と比較して、相関計算量は7の6乗倍となる(摂動範囲の6乗倍)。

表1 摂動範囲の拡大

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

2.2 適応摂動相関法

計算量が膨大になることを回避しつつ、摂動範囲を広げるための手法として、適応摂動相関法(山登り法)を提案している[6]。

この山登り法は、一言でいうと、局所的な相関値の最大値を探しながら、全体での相関値最大を探す手法である。

山登り法では、まず、摂動パラメータである傾き、伸縮、偏移のそれぞれについて独立に中心のまわり ± 1 単位までの摂動相関法を行う。ここで、中心で相関値が最大となるものが見つかった場合は、そこで計算を停止する。相関値最大となるものが中心以外で見つかった場合は、そこを再び、中心としてまわりのアフィン変形を行い相関量を求める。このとき、先に計算済みの箇所については、既に計算値が求まっているので改めて計算を行う必要はない。

図 1 は、アフィン変形量をマス目で表したものである。山登り法は、まず図 1 の (i) から始まる。中心 (アフィン変形量ゼロ) のまわり 8 点のアフィン変形を行って、中心を含めた 9 箇所について相関値最大のものを探査する。

探索の結果、中心が相関値最大となった場合は、(i) で計算を止める。上下左右のどれかの相関値が最大となった場合は、次は (ii-a) の状態となる。このとき、新しく中心となったまわりで、まだ計算がされていない 3 点についてのみ計算を行う。

中心から斜めのどれかの相関値が最大となった場合は、(ii-b) のようになる。この場合は、計算されていない周りの 5 点について計算を行う。

さらに外側に中心が見つかった場合は、(iii) のように再度、未計算の箇所についてアフィン変形と相関値の計算を行う。中心が端になった場合は、それ以上の変形ができないため、それ以上の計算を止める。

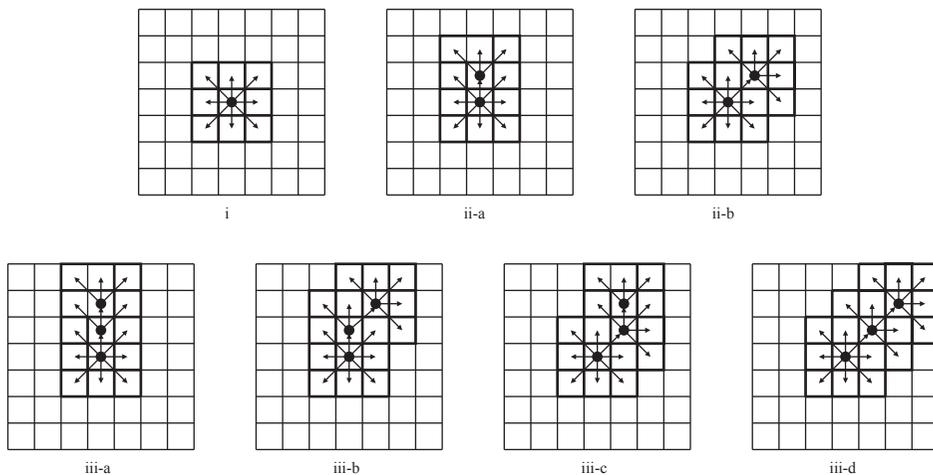


図 1 山登り法での探索方法

3 摂動相関法での標準パターン

3.1 認識実験用サンプルデータ

シミュレーション実験に用いた文字は、手書き文字データベース **etl6** の数字部分である。

1 画素 4 ビットで表現されたこのデータベースのデータに、簡単なノイズ除去を施して作成

した、 $64 \times 63 \times 1$ (縦 × 横 × ビット) の2値画素データにする。

この文字画像データの周囲に16個のゼロ画素を付加して(下の部分は17個)、 $96 \times 96 \times 1$ の2値画素データに変換する。

et16の数字データは0から9までの各カテゴリに1,383個のデータがある。実験では、各カテゴリの最後のひとつを除いた1,382個を使用して、サンプル番号の偶数番目と奇数番目の2つのグループに分ける。2つのグループの一方を標準パターン用にし、もう一方を未知パターン用とする。

3.2 標準パターンの作成

摂動相関法での標準パターンは、単純平均による標準パターン (average standard pattern) と学習済み標準パターン (learned standard pattern) とがある。

単純平均の標準パターンは、データを矩形に切り出したのち、 14×14 を基準サイズに揃えて足し合わせて単純平均をとっただけのものである。

標準パターンの作成で使ったデータをもう一度使い、摂動によって、各カテゴリの単純平均の標準パターンと類似度が最大となるようにアフィン変形させたものを再度、足し合わせたものを“学習済み標準パターン”と呼ぶことにする。

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るので、形状として、単純平均の標準パターンよりも、やや整った標準パターンを作成することができる。

なお、単純平均による標準パターンも学習済み標準パターンも1カテゴリにつき標準パターンは1つにしている。

作成した学習済み標準パターンを図2と図3に示す。

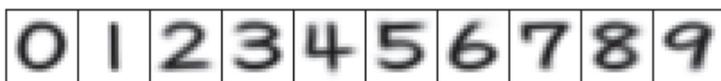


図2 学習済み標準パターン (奇数データから作成)

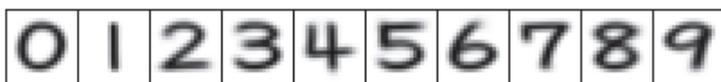


図3 学習済み標準パターン (偶数データから作成)

4 摂動相関法による文字認識シミュレーション実験

適応摂動相関法 (山登り法) を使って標準パターン、摂動範囲を変えた場合のシミュレーション実験結果を表2および図4に示す。表中のModeは、摂動範囲をあらわす。Standard/Unknownは、標準パターンと未知パターンの組合せを表し、“eo”は、標準パターン

が偶数 (e) で未知パターンが奇数 (o) であることを表している。認識に用いたデータ数は、6910(691/カテゴリ) で、エラー数から認識率は簡単に求めることができる。シミュレーション実験の結果から次のことがわかる。

- 摂動範囲を広げると認識精度があがる。
- 単純平均よりも学習済み標準パターンを用いた方が、認識精度があがる傾向がある。
- 学習済み標準パターンで摂動範囲を広げた場合は、99% 以上の認識率である。

表 2 山登り法によるシミュレーション実験の結果

Mode	Standard/ Unknown	Average standard pattern		Learned standard pattern	
		Number of Errors	User Times	Number of Errors	Uer Times
ASP	ee	118	20.661s	115	20.589s
	eo	128	20.425s	132	20.877s
	oe	115	20.493s	112	20.813s
	oo	121	20.457s	126	20.677s
A ² S ² P ²	ee	66	1m08.068s	62	1m10.152s
	eo	85	1m08.396s	75	1m10.380s
	oe	63	1m08.280s	64	1m10.380s
	oo	81	1m08.360s	68	1m10.272s
A ³ S ³ P ³	ee	63	1m57.791s	52	2m01.832s
	eo	78	1m56.487s	62	2m02.036s
	oe	63	1m56.295s	48	2m01.388s
	oo	76	1m56.171s	56	2m01.632s

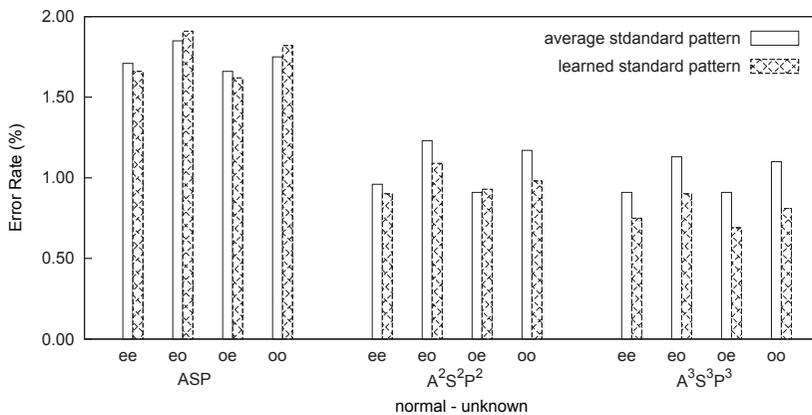


図 4 摂動範囲，標準パターンの違いによる誤読率 (山登り法)

5 svm 法による文字認識シミュレーション

サポートベクトル法を使った文字認識のシミュレーション実験を行うために、使用したソフトウェア LIBSVM[1] を紹介する。

LIBSVM は、国立台湾大学の Chih-Chung Chang と Chih-Jen Lin によって作成された svm のソフトウェアのセットであり、インターネットから入手することが可能である。

svm 法による文字認識シミュレーション実験は、この LIBSVM を使用した。

5.1 LIBSVM 用のデータ作成

摂動相関法で使用した `etl6` の手書き数字データをここでも使用する。

まず、摂動相関法と後で比較できるように、データを摂動相関法と同様に、偶数の組と奇数の組とに分けておく。

次に、簡単なノイズ除去をして、文字図形を矩形に切り出したものを 14×14 の大きさに揃えた後、 20×20 の領域の中心に設定する。この処理は、単純平均の標準パターンの作成処理の途中 (足し合わせる前) と同じである。

LIBSVM のデータフォーマットは `ascii` 形式で、個別文字認識で使用する場合、ファイルの 1 行に 1 文字分のデータを格納する。各行の先頭にはデータのラベルを設定し、データはその後に、“インデックス：値” の形式で格納する。ラベルとデータおよびデータとデータの間は、スペースで区切る。

ラベルとして数字データのカテゴリを設定する。インデックスは 20×20 の画像データの左上端を 1、そこから右方向に走査したものを 2, 3, ... とし、右下端を 400 の番号をあてる。値は、インデックスの位置にある画素の値を設定する。

svm 法では、入力の特徴ベクトルとして、任意のもの (例えば、方向性特徴など) を使うことができるが、ここでは、上記の摂動相関法との比較のため、スカラー量である各画素の画素値を特徴ベクトルとして使用する。

この方法で格納した、“インデックス：値” 形式の情報では、 20×20 の矩形の情報が残らない。また、この後にその矩形情報が必要となることもない。その理由は、サポートベクトル法では画像データを縦 \times 横の画像データとして扱わずに、400 次元のデータとして扱っているためである。

5.2 svm 法による文字認識シミュレーション実験

svm 法での文字認識処理は、学習によるモデルの作成と認識処理から成り立つ。また、認識処理の結果が学習モデルにフィードバックされることはないので、モデルの作成処理と認識処理とは完全に分けることができる。

データはモデル作成用のデータ (訓練用データまたは学習用データ) と認識処理用のデータ

とが必要となる。

あらかじめ、データは偶数と奇数に分けてあるので、一方を学習用に使い、そのモデルを作成する。そしてこの作成されたモデルを使い、認識処理用のデータを使って、svm 法による文字認識を行う。

svm 法では、通常、線形では分けることができないか、分けることが複雑なデータの境界をデータを高次元特徴空間に写像し、この特徴空間上で 2 つのクラスを分離する超平面を決定するという方式によってカテゴリを認識している。[1, 2, 8, 9]

しかも、実際には特徴空間の変数を陽に扱うことなく、カーネルと呼ばれる関数を使った計算だけで線形の識別を行うことができ、これは、カーネルトリックと呼ばれている。

カーネルのひとつである、ラジアル基底関数 (Radial Basis Function : RBF) カーネルは、次の式で与えられる。

$$K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2) \quad (2)$$

この式で表されているベクトル \mathbf{x} と \mathbf{x}' は、モデル作成の学習過程では、どちらも学習用のデータであり、モデル作成後の認識処理では、一方は学習用のデータで、もう一方は認識用のデータになる。

svm 法では、モデルの作成処理において、学習用のデータ全てが認識処理で用いられるわけではなく、マージン最大化 [2, 8, 9] という基準で、たくさんの学習用データの中の少数のデータだけが選択される。この少数の選択されたデータはサポートベクトルと呼ばれている。

学習に与えられない未知のデータに対して持つ認識能力のことを汎化能力と呼ぶが、svm 法を繰り返して学習することによって汎化能力を推定する交差検定法は、信頼性が高いが時間のかかる汎化能力の推定法である。

LIBSVM には、RBF カーネルを使って離散的な値を決めながら、もっとも汎化能力が高くなるパラメータ値を探索するグリッドサーチという手法が用意されており、シミュレーション実験ではこれを使用して学習モデルの作成を行った。

上記で述べたように、svm 法は、超平面によって分離するという、2 つのクラスの判定を行うものなので、多クラスへの拡張は、一意ではなく、いくつかの法がある。[9]

LIBSVM での多クラスへの拡張は 1 対 1 (one-against-one) SVM [1] という手法をとっている。この手法は、 n クラスの問題を $n(n-1)/2$ 個のクラス対からなる 2 クラス問題に変換する。そしてそれぞれのクラス対に対する超平面を決め、投票方式で分類するという手法である。[9]

10 クラス数字 (0-9) データを上記の手法に従って、モデルを作成して、svm による識別を行った結果を表 3 に示す。表中の“eo”は、学習モデルの作成には偶数データ (e) を使用して、未知データに奇数データ (o) を使用したことを表している。

図5と図6には認識の結果，誤読したデータを示す．ただし，この左端にある太字の図形は，相関法で用いた単純平均の標準パターンである．

誤読率は，eeとooでは0(%)，eoとoeでは0.5(%)以下という非常に認識性能のよい結果が得られている．

表3 svm法による文字認識シミュレーション実験結果

Learned / Unknown	Number of errors	User Times(s)
ee	0	4.272s
eo	23	4.272s
oe	16	4.540s
oo	0	4.540s

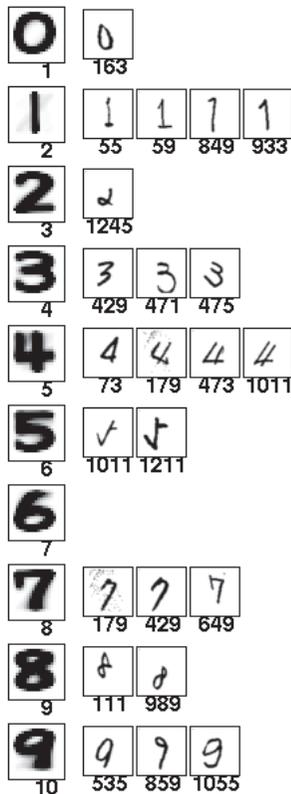


図5 eoの誤読データ

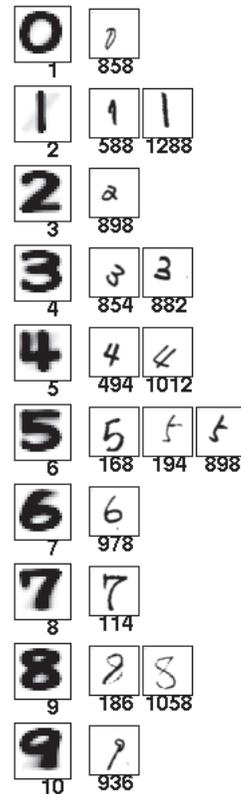


図6 oeの誤読データ

6 svm 法と相関法

筆者らのシミュレーション実験の相関法では、1 カテゴリ 1 標準パターンで認識を行っている。

一方、svm 法におけるモデルの作成では、多くの学習サンプルの中から、サポートベクトルとしてして少数のサンプルが選択されるとなっている。

偶数と奇数に分けたデータは 6910 データ (691 データ/カテゴリ) あり、そこから svm 法でサポートベクトルとして採用されるデータは、実際、どれくらいの数になるかをここで確認してみる。

先の実験で作成した学習モデルを使って調査したところ、偶数の学習モデルでは計 960 個が、奇数の学習モデルでは計 1,015 個がそれぞれサポートベクトルとして採用されていた。

偶数、奇数のそれぞれの学習モデルに採用されているサポートベクトルの数をカテゴリ単位に調べたものを表 4 に示す。そして、実際にどのような図形の手書き数字がサポートベクトルとして選ばれたのかを視覚的にわかるよう、偶数データを使った学習モデルで採用されていたサポートベクトルを図 7 に表す。

今回行った svm のシミュレーション実験では、モデルに採用されているサポートベクトルの数は、少ない偶数の場合でも、全学習データ数の 13% 以上になっている。

相関法で、すべての手書き数字のパターンを標準パターンとして採用することが、もしできるのであれば、認識率 100% の認識を行えることは容易にわかる。

しかし、実際には、すべての手書き数字のパターンを用意することは現実的ではない。

ただ、1 カテゴリあたりの標準パターン数を増やせば、認識率が良くなることも、仮定の話から想定できる。

ここでは、svm 法と相関法とを比較するため、svm 法で得られたサポートベクトルのデータを標準パターンとして、相関法のシミュレーション実験を行った。

すなわち、従来より 1 カテゴリあたり 1 標準パターンとしてきた相関法ではなく、サポートベクトルとして選ばれた、複数の標準パターンによる摂動相関法での認識実験を行った。

摂動範囲は、簡単のため直接相関 (D)、位置摂動のみ 1 単位 (P)、大きさ、伸縮、位置のそれぞれ 1 単位 (ASP) について行う。

また、これと比較のために、全データを標準パターンとして、そのまま使った直接相関 (D) および位置摂動のみ (P) の相関法による認識実験も行った。

表 4 サポートベクトルの数

class	0	1	2	3	4	5	6	7	8	9	total
even	55	36	88	90	104	162	83	78	150	116	960
odd	60	45	81	95	122	172	94	95	143	108	1015

実験の結果を表5および表6であらわす。

シミュレーション実験の結果から次のことがわかる。

- サポートベクトルを使った相関法は、今までの学習済み標準パターンより、良い認識結果となる。
- 標準パターンが複数でも、摂動範囲を広げると認識精度はあがる。
- 摂動範囲を±1単位まで広げたASPでは、僅かながらsvm法より良い認識率になった。
- 当然ながら、すべてを覚えてしまえば、認識率は100%となる。

表5 サポートベクトルを標準パターンとする認識結果

Mode	Standard/Unknown	Number of Errors	User Times
<i>D</i>	ee	1	2.780s
	eo	40	2.988s
	oe	29	3.316s
	oo	1	3.348s
<i>P</i>	ee	2	84.481s
	eo	26	84.433s
	oe	18	90.073s
	oo	3	89.737s
<i>ASP</i>	ee	0	6715.187s
	eo	15	6703.446s
	oe	7	7397.390s
	oo	0	7134.293s

表6 全データを標準パターンとする認識結果

Mode	Standard/Unknown	Number of Errors	User Times
<i>D</i>	ee	0	66.200s
	eo	31	66.400s
	oe	25	65.680s
	oo	0	69.232s
<i>P</i>	ee	0	590.996s
	eo	18	590.432s
	oe	12	591.080s
	oo	0	591.444

7 考察

シミュレーションによる実験結果から、svm法での文字認識は、認識性能も認識速度も優れていることが改めて確認できた。しかしながら、今回のシミュレーション実験で得られたsvm法での学習モデルのようにサポートベクトルの数が学習データの1割以上を占める場合、標準パターンとしてそれを採用した摂動相関法でも、svm法のそれと遜色のない認識性能が結果として得られた。処理時間について、svm法では学習モデルの作成には時間がかかったが、認識にかかる時間は摂動相関法と比べても、非常に早かった。一方、摂動相関法では、標準パターンの作成には時間がかからないが、認識には時間がかかった。

svm法のサポートベクトルに該当するような、相関法に適した複数の標準パターンを相関法に持ち込むことは、今後の課題とする。

さいごに、広島大学大学院の栗田多喜夫先生には、前回の紀要論文についてコメントをいただきました。そのなかで、svm法についての提案と、LIBSVMの紹介をしていただきました。この場を借りて栗田先生に感謝いたします。

参考文献

- [1] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [2] 栗田多喜夫：サポートベクターマシン入門
<http://home.hiroshima-u.ac.jp/tkurita/lecture/svm/index.html>
- [3] 下尾, 安田, : “文字認識のための摂動相関法とその深化”, 明星大学情報学部紀要, 10, pp. 159–176, 2002.
- [4] 北村, 二階堂, 中島, 安田 : “文字認識のための摂動相関法とその深化 (その2)”, 明星大学情報学部紀要, 12, pp. 1–12, 2004.
- [5] 北村, 二階堂, 中島, 安田 : “摂動範囲の広域化とその効果の検証”, 信学技報, PRMU2004-163
- [6] 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, 2006
- [7] 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-情報学部-第19号, 2011
- [8] C. M. Bishop : “Pattern recognition and machine learning” Springer-Verlag, 2006.
邦訳：元田浩, 栗田多喜夫, 樋口知之, 松本裕治, 村田昇監訳 : “パターン認識と機械学習 — ベイズ理論による統計的予測 (上下巻)” シュプリンガー・ジャパン, 2007–2009
- [9] 阿部重夫 : “パターン認識のためのサポートベクトルマシン入門” 森北出版株式会社, 2011