

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 22 日現在

機関番号：53203

研究種目：若手研究(B)

研究期間：2013～2014

課題番号：25871033

研究課題名(和文) カテゴリ追加を考慮した大規模データにおけるスパースサポートベクトルマシンの開発

研究課題名(英文) Sparse support vector machine for big problem which sequentially-add categories

## 研究代表者

北村 拓也 (Kitamura, Takuya)

富山高等専門学校・電気制御システム工学科・助教

研究者番号：40611918

交付決定額(研究期間全体)：(直接経費) 1,200,000円

研究成果の概要(和文)：本研究では、カテゴリ追加を考慮した大規模データにおけるスパースサポートベクトルマシンの開発を目的として行った。例として、顔認証において、あらかじめ全カテゴリ(登録する全ての人)を保持しているのではなく、逐一カテゴリが追加され、さらに大規模のデータを学習する問題を設定した上で、追加されるカテゴリと必要最低限の既存の情報を用いて学習できるスパースサポートベクトルマシンを開発・適用し、その性能の評価を行った。また、評価にはベンチマークデータセットを用いた計算機実験によっても行った。その結果、従来のシステムに比べて高い汎用性が確認できた。

研究成果の概要(英文)：In this research, I have developed sparse support vector machine for big problem which sequentially-add categories. For example, I apply this system to face identification which sequentially-add the categories (registrants). Then, this problem may be too big. However, this system uses only a new added category in training when this category is added. Using face identification problem and multi-category benchmark datasets, I evaluate the effectiveness of this system.

研究分野：機械学習

キーワード：パターン認識 機械学習 サポートベクトルマシン スパース オンライン学習

1. 研究開始当初の背景

近年、情報化社会の急速な発展と共に、個人認証や医療分野などにおける認識技術などのようなコンピュータに対象物のあるひとつのカテゴリに分類させる処理、すなわちパターン認識技術の必要性が非常に高まっている。パターン認識問題においてサポートベクトルマシンなどのような高い汎化能力を持つ識別器が開発されているが、個人認証技術などの逐次的に分類するカテゴリが追加される問題に適した識別器は现阶段では非常に少ない。

今後、このような逐次的に分類するカテゴリが追加される問題を取り扱うことが一般的になると想定され、本研究ではこの問題に有用な識別器の開発に取り組む。

2. 研究の目的

本研究の目的は、前述に示した通り、逐次的なカテゴリの追加を考慮した問題に適したパターン認識システムの開発である。ここで、逐次的にカテゴリが追加されることを前提にすると、最終的に学習するサンプルは無限大に増加することが容易に想定でき、追加される度それらのサンプルを取り扱うことは非現実的である。そのため、本研究で開発するシステムでは、分類するための判別関数の決定に新たに追加されたカテゴリにおけるサンプルと既存のサンプルの中のごく一部のサンプルのみを適宜選択し、可能な限りサイズの小さい最適化問題を解くシステムとすることが必要となる。

3. 研究の方法

カテゴリ追加を考慮した大規模データを取り扱うためには、学習の効率化とスパース化が重要となる。

そこで本研究では、標本特徴空間を用いたスパース最小自乗サポートベクトルマシン(以下、SLS-SVM)を改良し、学習の効率化とスパース化を同時に満たした。SLS-SVMで全教師サンプルから一部の教師サンプル(一次独立な教師サンプル)を用いて他の教師サンプルを表現し、判別関数決定のために用いる教師サンプルを削減し、解にスパース性を与える手法である。この手法をカテゴリ追加を考慮した大規模データに適用するために、カテゴリごとに一次独立なデータを選択し判別関数を決定するシステムに改良することにより、カテゴリが追加される度、このカテゴリの教師サンプルを学習する事のみで新たに判別関数を決定することができることから学習の効率化とスパース化が同時実現できる。

このシステムの評価を行うため、ベンチマークデータを用いた検証を行うと共に、多数の登録者を用いた顔認証により実問題への適用を行った。

4. 研究成果

本研究分野において広く用いられるベンチマークデータを用いた計算機実験において表 1, 2 に示す結果が得られた。ここで、

LS-SVM と SLS-SVM は従来手法を表し、FSL-SVM は提案方式を表している。

表 1. ベンチマークデータセットにおける提案方式の評価 (認識率 (%) とスパース性)

Data	LS-SVM		SLS-SVM		FSL-SVM		FSL-SVM (AAO)	
	Rec.	SVs	Rec.	SVs	Rec.	SVs	Rec.	SVs
Banana	89.5 ± 0.5	400	89.2 ± 0.5	44	89.0 ± 0.5	69	89.5 ± 0.4	276
B. cancer	73.6 ± 4.5	200	74.1 ± 4.5	52	73.3 ± 4.5	5	74.1 ± 4.4	105
Diabetes	77.0 ± 1.6	468	77.0 ± 1.7	165	76.9 ± 1.1	284	77.0 ± 1.7	324
German	76.2 ± 2.1	700	75.9 ± 2.1	189	75.8 ± 2.0	40	75.9 ± 2.0	268
Heart	84.2 ± 3.1	170	84.2 ± 3.3	126	84.1 ± 3.2	47	83.7 ± 3.1	75
Image	95.5 ± 0.7	1300	91.7 ± 1.2	279	91.0 ± 1.3	359	95.9 ± 0.5	1210
Ringnorm	96.3 ± 0.5	400	94.2 ± 3.0	22	94.3 ± 2.1	317	98.5 ± 0.1	400
F. solar	66.6 ± 1.6	666	66.6 ± 1.6	32	66.7 ± 1.6	104	66.6 ± 1.6	109
Splice	89.4 ± 0.7	1000	89.3 ± 0.7	977	86.8 ± 0.6	977	86.9 ± 0.6	974
Thyroid	93.8 ± 2.8	140	92.7 ± 2.8	29	92.6 ± 2.8	53	95.7 ± 2.1	108
Titanic	77.3 ± 1.2	150	77.2 ± 0.8	10	77.3 ± 1.1	6	77.2 ± 1.1	13
Twonorm	97.4 ± 0.2	400	97.5 ± 0.2	306	97.6 ± 0.1	40	97.6 ± 0.1	40
Waveform	90.3 ± 0.4	400	89.6 ± 0.6	393	90.2 ± 0.4	399	90.4 ± 0.4	400

表 2. ベンチマークデータセットにおける提案方式の評価 (計算時間 (s))

Data	LS-SVM	SLS-SVM	FSL-S	FSL-S (AAO)
Banana	0.219	0.211	*0.108	*0.052
B. cancer	0.008	0.023	0.025	*0.016
Diabetes	0.058	0.294	*0.111	*0.105
German	0.162	1.257	*0.439	*0.421
Heart	0.007	0.013	*0.005	*0.008
Image	0.803	7.597	*1.787	*4.121
Ringnorm	0.344	0.228	*0.174	*0.175
F. solar	0.091	0.883	*0.248	*0.263
Splice	0.731	6.076	*2.142	*2.150
Thyroid	0.004	0.011	*0.007	*0.008
Titanic	0.004	0.011	0.028	0.033
Twonorm	0.139	0.316	0.317	*0.305
Waveform	0.245	0.277	*0.235	*0.234

表 1 より、本研究で開発したシステムにおける識別能力は従来のシステムとほぼ同等であることが確認され、すなわち、カテゴリ追加を考慮することにより識別能力の劣化が生じないことを確認できた。

また実問題として、52 人の登録者が逐次的に追加されるような個人認証問題に適用した。ここで 1 名の登録につき、図 1 のように顔の傾きや輝度を変化させて行った。

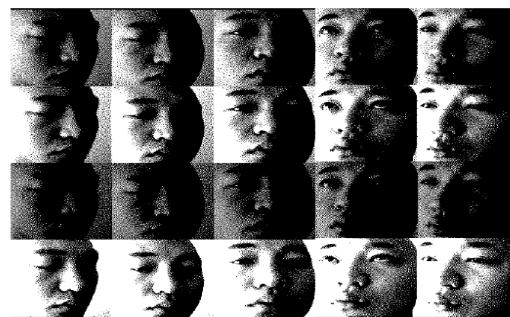


図 1. 個人認証における登録画像 (1 名分)

これにより、様々な環境下においてもシス

テムを適用できるかどうか検証できる。この問題における実験結果を表3, 4に示す。ここで, FLSL-SVMが付与される名称の識別器が提案方式でありその他が従来方式である。

表 3. 顔認証問題における提案方式の評価 (認識率 (%) とスパース性)

識別器	平均認識率 ± 標準偏差
KSM	96.14 ± 4.810
LS-SVM (OAA)	97.68 ± 3.462
LS-SVM (pairwise)	96.65 ± 6.541
SLS-SVM (OAA)	97.49 ± 2.973
SLS-SVM (pairwise)	97.61 ± 2.179
FSL-SVM (OAA)	86.55 ± 2.143
FSL-SVM (pairwise)	94.92 ± 5.482
部分空間法	96.90 ± 5.523
線形 LS-SVM (OAA)	86.54 ± 3.141
線形 LS-SVM (pairwise)	86.53 ± 3.274
線形 SLS-SVM (OAA)	83.01 ± 1.119
線形 SLS-SVM (pairwise)	83.01 ± 1.119
線形 FSL-SVM (OAA)	86.85 ± 2.047
線形 FSL-SVM (pairwise)	86.85 ± 2.047

表 4. 顔認証問題における提案方式の評価 (計算時間 (s))

識別器	学習時間
KSM	1.138
LS-SVM (OAA)	2.452
LS-SVM (pairwise)	1.781
SLS-SVM (OAA)	13.09
SLS-SVM (pairwise)	13.69
FSL-SVM (OAA)	0.9213
FSL-SVM (pairwise)	1.218
部分空間法	0.4352
線形 LS-SVM (OAA)	0.8788
線形 LS-SVM (pairwise)	0.8782
線形 SLS-SVM (OAA)	12.75
線形 SLS-SVM (pairwise)	14.02
線形 FSL-SVM (OAA)	0.4696
線形 FSL-SVM (pairwise)	0.5711

表3より, 認識率は従来手法より劣化していることが確認できるがその程度は微小であり, 表4より確認できる学習の高速化の観点を考慮すると非常に優れたシステムであることが明らかである。

これらのことより, 本研究によって従来のシステムに比べて非常に優れたシステムの開発が実現できたといえる。

##### 5. 主な発表論文等 〔雑誌論文〕(計2件)

[1] 北村拓也, 関根崇正, 塚越勇規, “カテゴリー別特徴抽出による高速スパース最小自

乗サポートベクトルマシン,” 情報処理学会論文誌 数理モデルと応用, Vol. 8, No. 1, pp. 7-17, 2015. (査読有)

[2] T. Kitamura and K. Ota, "Improved Multi-class Support Vector Machines Using Novel Methods of Model Selection and Feature Extraction," Proc. 20th International Conference on Neural Information Processing (ICONIP2013), Part , LNCS8227, pp. 299-306, 2013. (査読有)

〔学会発表〕(計9件)

[1] 江淵文人, 浅野孝平, 北村拓也, "目的関数値に基づいた特徴選択によるスパースLS-SVMの多クラス問題への拡張方式," 第42回知能システムシンポジウム講演, 北野プラザ六甲荘(兵庫県), 2015年3月17日.

[2] 浅野孝平, 北村拓也, "目的関数値に基づいた特徴選択によるスパースLS-SVM," 第42回知能システムシンポジウム, 北野プラザ六甲荘(兵庫県), 2015年3月17日.

[3] 浅野孝平, 北村拓也, "目的関数値に基づいた特徴選択によるスパースLS-SVM," 計測自動制御学会システム・情報部門学術講演会 (SSI2014), 岡山大学 (岡山県), 2014年11月21日.

[4] 北村拓也, 関根崇正, 塚越勇規, "カテゴリー標本特徴空間を用いた高速スパース最小乗サポートベクトルマシン," 第98回情報処理学会数理モデル化と問題解決研究会, 沖縄科学技術大学院大学 (沖縄県), 2014年6月26日.

[5] 舟木翔太, 北村拓也, "識別面を線形結合させた標本部分空間法に基づくサポートベクトルマシン," 第41回知能システムシン

ポジウム, 筑波大学東京キャンパス (東京都), 2014年3月14日.

[6] 塚越勇規, 舟木翔太, 北村拓也, "同時定式化方式を用いた高速スパース最小自乗サポートベクトルマシン," 第41回知能システムシンポジウム, 筑波大学東京キャンパス (東京都), 2014年3月14日.

[7] 舟木翔太, 北村拓也, "SS-SVMにおける識別面の線形結合による汎化能力向上方式," パターン認識・メディア理解研究会, 福岡大学 (福岡県), 2014年2月13日.

[8] 大林慎太郎, 塚越勇規, 舟木翔太, 北村拓也, "SVMを用いた顔認証実験の検証," パターン認識・メディア理解研究会, 福岡大学 (福岡県), 2014年2月13日.

[9] 北村拓也, 関根崇正, "多クラス問題におけるスパース最小自乗サポートベクトルマシンの高速学習法," 第23回インテリジェント・システム・シンポジウム (FAN2013), pp. 39-43, 九州大学 (福岡県), 2013年9月25日.

## 6. 研究組織

### (1) 研究代表者

北村 拓也 (Takuya, Kitamura)

富山高等専門学校・電気制御システム工学  
科・助教

研究者番号: 40611918