

研究種目：若手（B）
研究期間：2006～2008
課題番号：18700233
研究課題名（和文） カーネル特徴ベクトルを利用した多目的遺伝的アルゴリズムによる多クラス識別器の実現
研究課題名（英文） Multi-class Kernel-based Vector Machines using Multi-optimization Genetic Algorithm
研究代表者
渡邊 真也（WATANABE SHINYA）
室蘭工業大学・工学部・講師
研究者番号：30388136

研究成果の概要：

本研究における最大の成果は、「多目的最適化の枠組みを利用した多クラス識別器生成」の有用性を示したことである。具体的には、多クラス SVM に対する多目的最適化を用いたカーネルパラメータ最適化を行い評価基準の設定による識別性能への影響について調査を行い、複数の評価基準を併用する方法の優位性を示すことができた。一方、識別平面の明示化については残念ながら効果的な方法を実現できなかったものの、データ前処理による問題の容易化を実現する方法を考案し、低次元の問題に限定されているもののその有用性を示すことができた。

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	1,600,000	0	1,600,000
2007年度	1,100,000	0	1,100,000
2008年度	800,000	240,000	1,004,000
年度			
年度			
総計	3,500,000	240,000	3,740,000

研究分野：複合領域

科研費の分科・細目：「情報学」・「感性情報学・ソフトコンピューティング」

キーワード：カーネル法,多目的遺伝的アルゴリズム,多クラス識別器,サポートベクターマシン

1. 研究開始当初の背景

これまで多クラス識別問題においてニューラルネット（以下、NN）やサポートベクターマシン（以下、SVM）といった数多くの手法が提案され、多くの事例において良好な結果が報告されている。

しかしながら、従来の手法では得られた識別平面が明示的に示されない、識別平面を明示的に生成する場合では識別平面が線形あるいは準線形に限定されているという問題

点がある。識別平面はデータクラス間の差異を特徴的に表したものであり、その明示化は単に識別に有効な特徴ベクトルの抽出を可能にするだけでなく、そこから汎用的な知識を見いだすといった大きな有益性を有している。

そこで、本研究では識別平面を何らかの形で明示的に生成し、非線形識別も可能な仕組みを持つ識別器の開発を試みた。

2. 研究の目的

本研究の目的は、SVMにおけるカーネル法と多目的最適化の枠組みを利用することにより上述の問題点を解決した多クラス識別器の開発することである。多目的最適化を用いるのは、精度のみを評価基準として識別器を生成するのではなく、識別平面の形状の複雑さ等も考慮するためである。

以下、具体的な研究目的を幾つかの項目に分けて述べる。

(1) 多目的の枠組みで識別器設計を行う有用性の検証

本研究では、識別精度以外の評価項目（サポートベクターの数、識別平面の数など）を考慮した識別器設計の有用性について検証する。特に、多目的化による識別平面の複雑性緩和の効果について検証する必要がある。

(2) 識別平面の明示化について

識別平面の明示化およびそこからの特徴抽出の方法について検討を行う必要がある。本研究で前提としてカーネル関数を利用したSVMでは、カーネル関数により定義される特徴空間上で線形分離を行うため、実空間上での識別平面は明示的には得ることができない。そのため、特徴空間での線形識別平面が実空間上においてどのような識別平面となっているかを推測する必要がある。

3. 研究の方法

本研究では前述の「多目的の枠組みでの識別器設計の有用性」および「識別平面の明示化」について検証を行うために4つのアプローチから具体的な研究を進めた。以下、それぞれのアプローチごとに方法を述べる。

(1) 代表的なカーネル関数に基づくカーネルパラメータの最適化

RBFカーネル、線形カーネルといった代表的なカーネル関数を用いた多クラス識別器に対して、多目的遺伝的アルゴリズム（以下、多目的GA）を用いたカーネルパラメータ推定を行う。この取り組みにより、評価基準の設定と得られる精度の関係および多目的最適化の枠組みによるパラメータ推定の有用性について検証を行う。

(2) 複数のカーネル関数を組み合わせた場合における実験

より複雑な識別平面の問題にも対応するため、特徴の異なる複数のカーネル関数を組み合わせた場合についての実験を行う必要がある。これは複数のカーネルを組み合わせてもカーネル特性（半正定値性）が保たれるというカーネル関数の特徴を活かしたものである。

この実験では、問題に応じたカーネル関数の設計に関する指針を得ることを目指す。

(3) 得られた識別器からの識別平面の明示化についての検討

最適化により得られた多クラス識別器から識別平面を実空間上で明示的に抽出する方法について検討する。2次元の識別問題の場合には簡単に識別平面の視覚化が行えるが、3次元以上の場合には特徴空間における平面を実空間で表現するのは難しい。ここでは、何らかの局面近似の方法について検討を行う。

(4) 識別器ではなくデータ前処理による問題の容易化の検討

本研究では、識別器そのものの改良だけでなくデータ前処理による対象問題の容易化も研究対象として考える。難易度の高い識別問題は各クラスのデータが複雑に入り組んでいることが予想される。そのため、クラスのデータ分布を複雑化させている要因をデータ前処理により分析し、その要因の影響を軽減する方策について検討を行う。

4. 研究成果

得られた研究成果について、以下、幾つかの項目に分けて述べる。

(1) 代表的なカーネル関数に基づくカーネルパラメータの最適化に関する研究成果

多クラスSVMに対する多目的最適化を用いたカーネルパラメータ最適化を行い評価基準の設定による識別性能への影響について調査した。具体的には、クラス平均の精度(AUC(avg))、クラス間において最も悪い精度(AUC(min))、サポートベクターの数(#SV)を評価項目として設定し、どの評価項目の組み合わせが有効であるかについて実験を行った(表1参照)。

表1: 評価関数の組み合わせパターン

パターン	評価関数の設定
Pattern 1	AUC (avg)
Pattern 2	AUC (min)
Pattern 3	#SV
Pattern 4	AUC (avg), AUC (min)
Pattern 5	AUC (avg), #SV
Pattern 6	AUC (min), #SV
Pattern 7	AUC (min), AUC (min), #SV

UCI レポジトリから引用した問題に対して実験を行った結果、各評価項目を単独で用いるよりも3つ全てを評価項目として用いる場合が最も効果的であることが分かった。実験結果の一例を図1に示す。

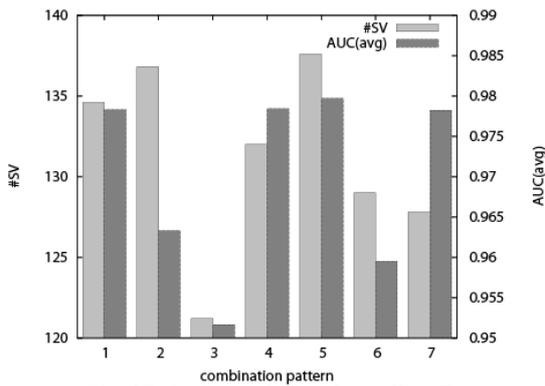


図1: 評価の組み合わせによる実験結果

このことより、識別問題の最適化では多目的の枠組みを利用することが効果的であることを明らかにすることができた。この研究成果については、後述の〔学会発表〕②、⑤、⑥にまとめられている。

(2) 複数のカーネル関数の組み合わせおよび識別平面の明示化について

代表的なカーネル関数を組み合わせて場合についても検討を行ったが良好な結果を得ることができなかった。当初、特徴の異なるカーネルを組み合わせることでそれぞれの特徴が機能を補完し合う効果を期待していたが、探索空間が広がり最適化の探索性能が悪化した一方、組み合わせによる補完の効果も期待したように得ることができなかった。

また、識別平面の明示化についても取り組んだが、特徴空間の識別平面を実次元空間で表現することは難しく、識別平面の特徴を何らかの形で抽出することもできなかった。

(3) データ前処理による問題の容易化の検討について

対象問題の容易化を目的とする新たなデータ前処理として、学習用データの分布より対象となるデータのクラスを分割し問題を単純化する方法を考案し、実験を行った。考案した多クラス化の概念図を図2に示す。

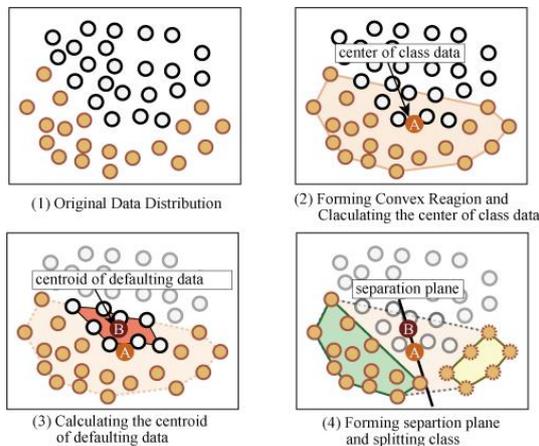


図2: 多クラス化に基づくデータ前処理の概念図

実験結果より、提案手法は2次元の複雑な形状を持つ問題に対して非常に効果的であることを示すことができた。実際が多クラス化の様子を図3に示す。

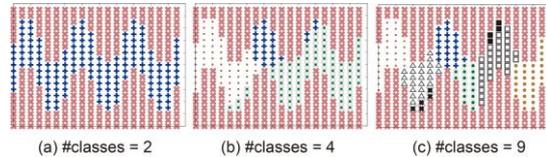


図3: 2次元データ分布に対する多目的化

一方、8次元を超える高次元の複雑なクラス分布形状を持つ問題に対しては良好な結果を得ることができなかった。これは、高次元空間におけるクラス分割が有効に機能していないためであり、今後、さらにこの問題に取り組んでいく予定である。この研究成果については、後述の〔学会発表〕①、④にまとめられている。

(4) その他の研究成果

上記以外の研究成果として「単目的最適化への多目的化の概念を利用したアプローチ」、 「多目的非劣解集合に対する設計支援」がある。

前者は、単目的最適化問題に対して、その目的関数の改善を促進するような別の目的を考慮して最適化する試みであり、探索促進のための触媒としての効果を期待するものである。なお、このアプローチで採用されている多目的化の概念は前述のデータ前処理につながるものである。本研究成果は、後述の〔雑誌論文〕③にまとめられている。

また、後者の非劣解集合からの設計支援は、高次元かつ多数の非劣解集合をどのように可視化し、内在する特徴を抽出するかという研究である。この研究では、カーネル次元削減による低次元化、代表解の抽出、グループ比較を通じた特徴抽出といった複数の手法を組み合わせた設計支援システムを提案し、ジェットエンジン最適化問題への適用を例にその有用性について示した。本研究成果は、後述の〔雑誌論文〕①にまとめられている。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計3件)

① 渡邊真也, 湊亮二郎, "多数非劣解集合からの設計支援手法の開発: ジェットエンジン最適化を通して", 人工知能学会論文誌, Vol. 24, No. 1, pp. 1-12, 2009 (査読有り)

② Shinya Watanabe, Hiroyasu Tomoyuki, Miki Mitsunori, "Effectiveness of an Evolutionary Algorithm for the

Multi-objective Rectangular Packing Problem”, Electronics and Communications in Japan, Part II., Volume 90, Issue 12, pp. 111-120, 2007 (査読有り)

③ 渡邊 真也, 榊原 一紀, “多目的化の概念を用いた単目的 Vehicle Routing 問題へのアプローチ”, 情報処理学会論文誌「数理モデル化と応用」, Vol. 48, No. SIG 2 (TOM 16), pp. 158-166, 2007 (査読有り)

[学会発表] (計 10 件)

① 渡邊真也, 木村幸代, 施建明, “境界領域の容易化を目的としたデータ前処理手法の提案”, 情報処理学会 研究報告「数理モデルと問題解決 (MPS)」 (No. 2009-MPS-073), pp. 133-136, 2009 (査読無し)

② Shinya Watanabe and Yukiyo Kimura, “A Methodology Using EMO for Parameter Estimation of SVM Kernel Function”, 2008 IEEE Conference on Soft Computing in Industrial Applications (SMCIA/08), pp. 211-216, 2008 (査読有り)

③ Shinya Watanabe and Ryojiro Minato, “Development of a Design Support System that can Efficiently Utilize Non-Dominated Solutions”, Joint 4th International Conference on Soft Computing and Intelligent Systems and 9th International Symposium on advanced Intelligent Systems (SCIS & ISIS 2008), CDROM-paper, 2008 (査読有り)

④ 木村 幸代, 渡邊 真也, “クラス分割に基づく多クラス識別器の開発”, 進化的計算シンポジウム 2008, 北海道 登別市, pp.173-180, 2008 (査読無し)

⑤ 木村 幸代, 渡邊 真也, “多クラス問題における SVM のカーネルパラメータ最適化”, 情報処理学会 情報システムと社会環境 研究会報告 No.105, pp.45-51, 2008 (査読無し)

⑥ 渡邊真也, 木村幸代, “多目的遺伝的アルゴリズムを用いた SVM カーネルパラメータ推定”, 進化計算シンポジウム 2007 講演論文集, pp.63-66, 2007 (査読なし)

⑦ S. Watanabe and K. Sakakibara, “A multiobjectivization approach for vehicle routing problems”, Evolutionary Multi-Criterion Optimization. Fourth International Conference (EMO 2007), Lecture Notes in Computer Science. Volume 4403, pp. 660-672, 2007 (査読有り)

⑧ 野一色 学, 榊原 一紀, 渡邊 真也, 玉置 久, 西川 郁子, “Pickup and Delivery 問題における探索空間の分割に基づく擬似焼き鈍し法の構成”, 第 50 回システム制御情報学会研究発表講演会, 469-470, 2006 (査読なし)

⑨ 渡邊 真也, 榊原一紀, “Vehicle Routing 問題への多目的最適化に基づくアプローチ”, 情

報処理学会 数理モデルと問題研究会報告 No.58, 兵庫県 豊岡市城崎町, pp.3-6, 2006 (査読なし)

⑩ 野一色 学, 榊原 一紀, 渡邊 真也, 玉置 久, 西川 郁子, “配送計画問題に対する分散型メタヒューリスティクスの構成”, 第 18 回自律分散システム・シンポジウム, 317-320, 2006 (査読なし)

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ

<http://is.csse.muroran-it.ac.jp/~sin/>

6. 研究組織

(1) 研究代表者

渡邊 真也 (Watanabe Shinya)

室蘭工業大学・工学部・講師

研究者番号 : 30388136

(2) 研究分担者

なし

(3) 連携研究者

なし