

研究種目：若手研究(B)

研究期間：2007～2010

課題番号：19700153

研究課題名(和文) カーネル部分空間法による学習理論の構築とその応用

研究課題名(英文) Kernel subspace method and its applications

研究代表者

鷲沢 嘉一 (WASHIZAWA YOSHIKAZU)

独立行政法人理化学研究所・脳信号処理研究チーム・研究員

研究者番号：10419880

研究分野：総合領域

科研費の分科・細目：情報学・知能情報学

キーワード：カーネル部分空間法, パターン識別, 機械学習

1. 研究計画の概要

初年度は、カーネル部分空間法に関する一般的な性能評価方法の開発と柔軟性のある信号処理による特徴抽出方法に関する研究を行う。一般的な性能評価方法の開発のために入力分布を仮定し、期待値を求める手法、任意の入力のもとで最悪の性能を求める手法などを検討する。柔軟性のある信号処理による特徴抽出方法では、脳信号のように異なる部位の信号が幾つもの同時に利用できる場合に対し、雑音を抑え、所望の特徴を多く含むような信号を抽出する手法を提案する。信号の抽出には、信号のスパース性や非負性あるいはその両方を利用する。さらに提案した手法を脳波信号などへ応用し性能を評価する。

次年度以降は、提案した信号抽出法とカーネル部分空間法の統合、BCIなどの脳信号処理への応用について検討を行う。

2. 研究の進捗状況

(1). 信号空間のスパース性に関する構造と部分空間との関連を利用し、スパース信号を抽出する手法を提案した。この手法は、従来法の信号の独立性を用いず、信号のスパース性のみを利用し、信号分離を行う手法であり、観測数が信号数よりも少ないときにも混合行列の推定が行える場合があるという特長を持つ。この結果は Y. Washizawa et al. "Sparse blind identification and separation by using adaptive K-orthodrome clustering," の題で論文雑誌に掲載されている。

(2). カーネル部分空間法の評価式を再定義し、制約付きの近似問題という枠組みで再

構成することにより、制約付き近似問題による特徴抽出法の族という新しい概念を開発した。信号の偏在する部分空間を求めるために、主成分分析やカルパーネンローブ展開がよく用いられる。これらは信号の存在する部分空間の次元を制限するという制約の下で、最良近似問題を解いていることに相当する。これにより特徴抽出が可能となるが、特徴抽出の本質は空間の次元を制約することではなく、変換行列そのものに対する制約であることに着目し、次元制約以外の制約方法を導入した。具体的にはフロベニウスノルム制約とトレースノルム制約を導入し、それぞれの性能と特長を検討した。さらにフロベニウスノルム制約を導入することで、雑音集合の特徴を抑制する手法を開発した。これらの結果は Y. Washizawa "Feature extraction using constrained approximation and suppression" の題で論文雑誌に掲載されている。

(3). カーネル部分空間法の評価式を制約付きの近似問題と再定義することで、最適化問題の定義が明確となり、部分カーネル主成分分析の概念を開発することができた。

カーネル部分空間法の1つであるカーネル主成分分析は、設計のために標本点数の固有値問題を解く必要があった。このためには標本数の二乗に比例する作業領域と標本数の三乗に比例する計算量が必要となる。しかしながら応用によっては数万から数百万の標本を用いることができる場合もあり、このような場合にはそのままカーネル主成分分析を適用することができなかつた。

カーネル主成分分析の問題を制約付きの近似問題と見直したことにより、計算量を増

大きせる原因が問題の空間の大きさに依ることが判明した。部分カーネル主成分分析では、問題の空間を標本の部分集合が張る空間に限定し、目的関数自身は全標本を用いることで効果的な近似問題を得ることができる。標準的な問題では、計算量を10分の1程度にしても誤差が1%程度しか変わらないことが確認できた。これらの結果については、Y. Washizawa, “Subset kernel principal component analysis” IEEE international workshop on machine learning for signal processingなどで発表している。

3. 現在までの達成度

①当初の計画以上に進展している。

拡張した概念から、新たな有効な手法など、当初の計画にはなかった事項まで研究が進んでおり、計画の内容についても概ね順調である。

4. 今後の研究の推進方策

(1). 脳信号の解析手法としてマルチウェイ解析が注目されている。従来の統計的データ解析手法はベクトルデータに対するものが主であった。画像などのベクトル以外のデータはベクトルに変換され、様々な解析手法が適用されることが多かった。しかしながらベクトル化により構造的な特徴が失われてしまったり、計算量、作業領域が大きくなってしまったりするなどの問題がある。例えば、脳波データは、被験者、トライアル、電極チャンネル、時間、周波数などのインデックスを持つ。これらをベクトル化すると巨大なベクトルになる上、構造的な性質が失われてしまう。

マルチウェイ解析では2つ以上のインデックスを持つデータ(テンソル)から直接特徴を抽出するものである。これまでの研究の成果をマルチウェイ解析に拡張することにより、より高度なデータ解析が期待できる。今後はこの拡張および、その応用について研究を行う予定である。

(2). 部分カーネル主成分分析の提案によ

り、大規模なデータに対する非線形解析が可能となった。例えば画像解析では、画像をパッチあるいはブロックと呼ばれる小領域に分割して解析を行う手法がよく用いられる。512x512の大きさの画像から5x5のパッチを取り出すとパッチの総数は26万程度となる。この大きさの標本は従来のカーネル部分空間法で扱うことができないが、提案法では問題の大きさを自由に設定できるため扱うことが可能となる。このような大規模な問題に対する応用も今後検討する予定である。

5. 代表的な研究成果

(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4件)

Y. Washizawa, “Feature extraction using constrained approximation and suppression,” IEEE trans. on neural networks, vol. 21, issue 2, pp. 201—210, 2010 査読あり

Y. Washizawa and A. Cichocki, “Sparse blind identification and separation by using adaptive K-orthodrome clustering,” Neurocomputing, vol. 71, no. 10-12, pp. 2321—2329, 2008.

[学会発表] (計17件)

Y. Washizawa, “Subset kernel principal component analysis,” IEEE international workshop on machine learning for signal processing, Sept. 2009, Grenoble, France.

Y. Washizawa, Y. Yamashita and A. Cichocki, “Blind source extraction using spatio-temporal inverse filter,” IEEE international symposium on circuits and systems, May 2009, Taipei, Taiwan.

[産業財産権]

○出願状況 (計3件)