

A Study of Efficient Kernel Adaptive Filtering Algorithm Based on Parallel Projection, Shrinkage Operator, and Model Learning

February 2021

TAKIZAWA, Masaaki

主 論 文 要 旨

No.1

報告番号	甲	第	号	氏 名	滝沢 雅明
主 論 文 題 名 :					
A Study of Efficient Kernel Adaptive Filtering Algorithm Based on Parallel Projection, Shrinkage Operator, and Model Learning (並列射影・軟閾値作用素・モデル学習に基づく 効率的な再生核適応アルゴリズムの研究)					
信号処理・機械学習の諸問題は、非線形関数推定問題として定式化される。近年、ビッグデータ解析をオンラインで実行できる適応学習アルゴリズムへの需要が高まっている。オンライン非線形関数推定への代表的なアプローチとして、ニューラルネットワーク・ボルテラフィルタ・カルマンフィルタ等が存在するが、それぞれ局所最適解・高演算量・モデル依存性の課題を抱えている。再生核に基づく適応アルゴリズムは、大域的最適性・低計算量・モデルへの非依存性などの利点がある一方、幾つかの課題が残されている。一つ目の課題は適応アルゴリズムの収束速度、二つ目は非線形関数を表現するための辞書サイズの増加、三つ目はカーネルパラメータの事前設計の困難性である。本研究では、これらの課題を解決する効率的な適応アルゴリズムを構築し、その性能を計算機実験により定量的に評価することを目的とした。 第1章は序論であり、本研究の背景と動機を述べ、先行研究を概説した。 第2章では、データ再利用による再生核適応アルゴリズムの高速化のための手法を提案した。提案アルゴリズムのポイントは、再生核ヒルベルト空間の幾何構造の利用と並列射影によるデータ再利用であり、これによって、少ない演算量で高速な収束、高精度な推定、ノイズに対する頑健性を同時に実現した。 第3章では辞書増加に伴うメモリ容量・計算量の増加および不要な辞書要素による性能劣化を防ぐために、軟閾値作用素を用いた適応的な辞書再構築法を提案した。ある条件下で、ℓ_1正則化付き二乗距離関数の最小解集合への単調接近性が保証されることを明らかにした。 第4章では、ガウス核のパラメータ（平均ベクトルと分散）を座標降下法で逐次的に学習する適応アルゴリズムを提案した。分散パラメータの初期値設定が学習結果に大きく影響することを実験的に明らかにするとともに、その影響を軽減することを目的として、分散パラメータの適切な初期値を選択するマルチスケールスクリーニング法を提案した。ガウス核のパラメータ学習は、数理モデルを適応的に学習することに対応しており、提案法は、これに加えてガウス核の係数ベクトルも同時に学習するものとなっている。 第5章は結論であり、各章で得られた内容をまとめ、本研究の成果を要約した。					

Thesis Abstract

No. _____

Registration Number	☒ "KOU" ☐ "OTSU" No. _____ *Office use only	Name	TAKIZAWA, Masaaki
Thesis Title A Study of Efficient Kernel Adaptive Filtering Algorithm Based on Parallel Projection, Shrinkage Operator, and Model Learning			
Various tasks in signal processing and machine learning can be cast as estimation of nonlinear functions. With the rise of big data, online/adaptive learning algorithms to solve nonlinear estimation problems in an online fashion have attracted particular attentions. Typical methods for online nonlinear estimation include neural network, Volterra filter, and Kalman filter, each of which has the following issues: local minima, high computational complexity, and model dependency. Although online/adaptive learning method with reproducing kernels, which is called the kernel adaptive filtering (KAF), overcome those issues, KAF suffers from the following efficiency issues: (i) convergence speed of algorithms, (ii) increase of the dictionary size, and (iii) the design of kernel parameters. This study aims to propose KAF algorithms which overcome those issues (i)-(iii) and verify the efficacy of the algorithm through computer simulations. In Chapter 1, the background, the objective of this study, and the previous studies were presented. In Chapter 2, an adaptive technique to enhance the convergence speed of KAF algorithms was presented. The proposed technique was based on the use of the geometric structure of reproducing kernel Hilbert space (RKHS)s as well as data-reusing implemented with the parallel projection. Thanks to the proposed technique, the proposed KAF algorithm had the advantages: high estimation-accuracy, low complexity, and fast convergence/tracking as well as noise robustness. In Chapter 3, an adaptive dictionary-reconstruction scheme based on a shrinkage operator was presented to avoid performance degradations of the algorithm caused by obsolete elements in the dictionary, as well as the increases of memory usages and computational complexities caused by the dictionary-size explosion. The proposed algorithm was shown to enjoy the monotone approximation property for a certain function. In Chapter 4, an adaptive algorithm which adapts the parameters of Gaussian kernels (scales and centers) iteratively by a coordinate descent-based algorithm was presented. We pointed out that the initial Gaussian scales largely affect the efficiency and the precision of the estimate. To alleviate the sensitivity to the initial scales, the multiscale screening method was presented which selects an adequate initial scale from multiple values efficiently. The proposed KAF algorithm included the adaptive learning of the coefficient vector of kernels as well as the kernel parameters, thereby learning mathematical models adaptively. In Chapter 5, the results of this study were summarized.			