

大学院 電気通信学 研究科 博士前期課程		情報通信工学 専攻	
氏 名	関口 涼平	学籍番号	0530026
論 文 題 目	カーネル主成分分析を用いた学習機械のパラメタ自動決定法		
<p>要 旨</p> <p>計算機による画像認識や音声認識，構文解析などは近年の情報化社会において，ますます重要性が増している．特に画像認識・動画認識については，生体認証や顔検出など実用化されてきているもの，様々な状況やポーズでも動画から識別するような高性能なものなど，多くの応用が期待されている分野である．</p> <p>しかし，技術の進歩から入力データが膨大な量に増えている今日では，入力データを直接学習させようとしても非効率であり，計算量の爆発的な増加や識別率が低下する事態も少なくない．入力される特徴から必要な特徴量だけを抽出し，学習に生かすことができなければ性能を向上させることは非常に難しい．</p> <p>本論文では，カーネル主成分分析という多変量解析法を利用し，入力データの特徴を効率的に抽出し入力データの規模を抑えることで，識別性能の向上と計算量の軽減について研究を行った．カーネル主成分分析は，Scholkopfにより提案され，画像処理などで応用が検討されている．</p> <p>主成分分析 (Principal Component Analysis: PCA) は，特徴量の相関関係を利用することによって，次元圧縮された部分空間への射影を求める方法であり，次元圧縮の手法としてパターン認識や信号処理の分野でよく利用されている．</p> <p>一方で PCA は線形変換であるため，分布の非線形性が強いデータに対しては，うまくデータの特徴量を抽出することができないという問題が存在する．</p> <p>Scholkopfによって提案されたカーネル主成分分析 (Kernel PCA: KPCA) は，カーネル写像によって高次元空間に非線形写像を行い，写像先で PCA を行うことにより，分布の非線形性の強いデータについて次元圧縮を行うことを可能にした．この手法は非線形座標軸上の成分を抽出できるため，近年画像認識の分野において注目されている．</p> <p>一般にカーネル法を識別問題に応用した場合，パラメタを持つカーネルではカーネルパラメタによって識別能力が大きく影響される．最適なパラメタを設定したとき，カーネル法の能力を最大限に引き出すことができるが，事前にこれを行う有効な手法は提案されていない．一般には事前実験をおこなって，良いパラメタを調べるが，多クラスの場合や複数のパラメタを持つカーネルを使用する場合，事前実験の規模が現実的ではなくなる．</p> <p>本研究では，KPCAと線形サポートベクトルマシン (Linear Support Vector Machine: LSVM) の結合による識別器を提案し，その性能を評価すると共に，主成分分析による次元削減の観点から最適なカーネルパラメタを決める効率的な方法を示す．この方法では事前実験を必要とせず，あらかじめ定めた条件と共に固有値分布に依存した評価関数を最適化することによって，最適パラメタを得ることができる．</p> <p>計算機実験の結果，提案手法によって構成された識別器は非線形サポートベクトルマシン (Non-Linear Support Vector Machine: NLSVM) の識別能力とほぼ同等の能力を得ることができた．さらに特徴量の効率的な抽出の結果，NLSVM よりも高速な実働計算量を実現することができた．</p>			