

图形、图像、模式识别

## 高维少样本数据的特征压缩

游文杰<sup>1, 2</sup>, 吉国力<sup>1</sup>, 袁明顺<sup>2</sup>

1.厦门大学 自动化系, 福建 厦门 361005

2.福建师范大学 福清分校, 福建 福清 350300

收稿日期 2009-8-24 修回日期 2009-10-9 网络版发布日期 2009-12-30 接受日期

**摘要** 针对一类高维少样本数据的特点, 给出了广义小样本概念, 对广义小样本进行信息特征压缩: 特征提取(降维)和特征选择(选维)。首先介绍基于主成分分析(PCA)的无监督与基于偏最小二乘(PLS)的有监督的特征提取方法; 其次通过分析第一成分结构, 提出基于PCA与PLS的新的全局特征选择方法, 并进一步提出基于PLS的递归特征排除法(PLS-RFE); 最后针对MIT AML/ALL的分类问题, 实现基于PCA与PLS的特征选择和特征提取, 以及PLS-RFE特征选择与比较, 达到广义小样本信息特征压缩的目的。

**关键词** [广义小样本](#) [主成分分析\(PCA\)](#) [偏最小二乘\(PLS\)](#) [特征提取](#) [特征选择](#)

分类号 [TP391](#)

## Feature reduction on high-dimensional small-sample data

YOU Wen-jie<sup>1, 2</sup>, JI Guo-li<sup>1</sup>, YUAN Ming-shun<sup>2</sup>

1.Department of Automation, Xiamen University, Xiamen, Fujian 361005, China

2.Fuqing Branch, Fujian Normal University, Fuqing, Fujian 350300, China

### Abstract

In view of the characteristics of small sample and high dimensional data, Generalized Small Samples (GSS) is defined. It reduces information feature of GSS: feature extraction (dimensionality extraction) and feature selection (dimensionality selection). Firstly, unsupervised feature extraction based on Principal Component Analysis (PCA) and supervised feature extraction based on Partial Least Squares (PLS) are introduced. Secondly, analyzing the structure of first PC, it presents new global PCA-based and PLS-based feature selection approaches, in addition recursive feature elimination on PLS (PLS-RFE) is realized. Finally, the approaches are applied to the classification of MIT AML/ALL, it performs feature extraction on PCA and PLS, and feature selection compared with PLS-RFE. The information compression of GSS is realized.

**Key words** [generalized small sample](#) [Principal Component Analysis \(PCA\)](#) [Partial Least Squares \(PLS\)](#) [feature extraction](#) [feature selection](#)

DOI: 10.3778/j.issn.1002-8331.2009.36.049

通讯作者 游文杰 [glji@xmu.edu.cn](mailto:glji@xmu.edu.cn)

### 扩展功能

#### 本文信息

▶ [Supporting info](#)

▶ [PDF\(909KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [参考文献](#)

#### 服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [复制索引](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

#### 相关信息

▶ [本刊中 包含“广义小样本”的 相关文章](#)

▶ [本文作者相关文章](#)

· [游文杰](#)

·

· [吉国力](#)

·

· [袁明顺](#)