基于智能计算的降维技术研究与应用

皋军 著

中国水利水电出版社
www.waterpub.com.cn

基于智能计算的降维技术研究与应用

皋军 著

中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书论述了智能计算中降维技术的建模思想和仿真方法，给出了运用最新的建模方法和理论对高维数据进行降维预处理的仿真计算的过程，向读者展示了在传统技术的基础上，如何通过结合其他智能计算的方法构造具有较强鲁棒性的特征降维方法，使读者了解特征降维技术在统计学意义下的一般演化规律。

本书以简洁易懂的语言描述了特征降维技术建模的理论基础和建模过程。本书可供智能控制、计算机等领域中的教师、研究生及其他相关人员参考。

图书在版编目（CIP）数据

基于智能计算的降维技术研究与应用 / 皋军著. -- 北京 : 中国水利水电出版社, 2013.11
ISBN 978-7-5170-1385-3

Ⅰ. ①基… Ⅱ. ①皋… Ⅲ. ①智能计算机－研究 Ⅳ. ①TP387

中国版本图书馆CIP数据核字(2013)第265714号

策划编辑：石永峰　责任编辑：杨元泓　加工编辑：孙　丹　封面设计：李　佳

书　　名	基于智能计算的降维技术研究与应用
作　　者	皋军　著
出版发行	中国水利水电出版社 （北京市海淀区玉渊潭南路1号D座　100038） 网址：www.waterpub.com.cn E-mail：mchannel@263.net（万水） sales@waterpub.com.cn 电话：（010）68367658（发行部）、82562819（万水）
经　　售	北京科水图书销售中心（零售） 电话：（010）88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排　　版	北京万水电子信息有限公司
印　　刷	三河市天润建兴印务有限公司
规　　格	170mm×240mm　16开本　11.75印张　217千字
版　　次	2013年11月第1版　2013年11月第1次印刷
定　　价	42.00元

前　言

随着社会信息化的发展，在具体的智能识别过程中需要处理的的数据越来越多地呈现出高维特征，比如图像处理、文本分类、视频检索、计算机视觉、微阵列数据基因选择和基于生物特征的身份识别等。造成这种现象的主要原因在于：在智能识别过程中，只有当样本已经包含了足够多的模式分类信息时，才能得到较好的智能识别效果。然而，如何确定特征中是否已经包含足够多的类别信息本身就是个很难解决的问题。因此为了提高模式识别效果，在通常情况下，人们通过采集尽可能多的特征去体现样本的类别信息，这导致原始样本空间的维数可能达到几千维甚至上万维，而如果在如此高维的原始空间直接使用模式识别方法，那么所得到的智能识别效果将受到较大的影响。这是因为在如此高维的特征中存在着大量冗余的特征，使得特征之间的相关性较强，从而增加了模式分类算法的负担，降低了算法的效率。同时，由于随着样本特征维数的增加，使得对样本的统计特性更加难以估计，从而会影响分类算法的泛化能力，呈现出所谓过学习的现象[1]。因此，在智能识别过程中必须要对高维数据进行相应的预处理，以达到在保持样本信息量的基础上尽可能降低特征的维数，提高模式分类的效果，而特征降维技术就是一种较为有效的数据预处理方法。

目前，特征降维技术作为一种关键的数据预处理技术被广泛地加以研究，并在不同的实际运用领域得到了较为成功的应用，但随着新理论和新技术的不断发展，特别是大量新兴的智能识别应用领域的需求，对特征降维技术提出了更高的要求，使得现有的特征降维技术面临更大的挑战。比如：如何提高基于支持向量机的特征选择方法的泛化能力和鲁棒性；如何更好地实现特征提取技术与模糊聚类技术的有机结合，以提高特征降维方法的鲁棒性；如何提高特征降维方法中的距离度量学习的有效性；如何将特征降维方法中的关键技术和理论运用到支持向量机中，以提高支持向量机的泛化能力和鲁棒性；如何结合张量理论提高特征降维的效果；如何在具有明显不同分布的源域和目标域实现提取技术等。

为此，本书从三个部分对上述问题进行描述和研究：

第一部分由第 2 章和第 3 章组成，这一部分分别讨论两种新颖的特征选择方法。具体来说，第 2 章主要针对势支持向量机 P-SVM 存在的泛化能力不强的问题，通过引入 Fisher 判别分析方法中的类内散度矩阵，重新构造 P-SVM 的目标函数，从而形成具有较强泛化能力的广义的势支持特征选择方法 GPSFM；第 3 章针对经典的模糊聚类方法 FCM 存在的对噪音数据和噪音特征敏感的问题，采用对样本点和样本特征同时加权的方式，重新构造 FCM 方法的目标函数，从而得到具有

特征排序功能的鲁棒性模糊聚类方法 FCA。

第二部分由第 4 章至第 7 章组成，这一部分分别讨论四种新颖的特征提取方法。具体来说，第 4 章针对线性拉普拉斯判别准则 LLD 方法存在的小样本问题以及如何确定原始样本空间类型的问题，通过引入语境距离度量并结合最大间距判别准则的基本原理，提出一种基于语境距离度量的拉普拉斯最大间距判别准则 CLMMC；第 5 章针对最大散度差判别准则的效果很大程度上依赖参数η的选取，并且该准则的划分属于硬划分，不能客观地反映现实世界的问题，通过引入模糊技术，重新构造一种新的模糊最大散度差判别准则，并根据这一新准则提出一种模糊聚类方法 FMSDC；第 6 章通过 FMSDC 方法并结合张量理论，提出一种矩阵模式的模糊最大间距判别准则 MFMMC，并在此基础上形成具有模糊聚类功能的双向二维无监督特征提取方法$(2D)^2$UFFCA；第 7 章主要讨论了迁移学习法，通过引入局部加权均值的方法和理论到 MMD 中，提出投影最大局部加权均值差异 PMLWD 度量，PMLWD 通过累积不同区域局部分块之间的局部分布差异来反映区域间的全局分布差异。在 PMLWD 的基础上，提出一种能实现迁移学习任务并具有一定局部学习能力的特征提取方法：最大局部加权均值差异嵌入 MWME。同时，在 PMLWD 的基础上，结合传统的学习理论，提出基于局部加权均值的领域适应学习框架 LDAF，在 LDAF 框架下衍生出两种领域适应学习方法：LDAF_MLC 和 LDAF_SVM。

第三部分由第 8 章和第 9 章组成，这一部分主要研究和讨论两种基于类内散度的支持向量机的方法。具体来说，第 8 章针对最小类内散度支持向量机 MCSVMs 面临的小样本问题，通过引入张量理论，重新构造 MCSVMs 支持向量机的目标函数，从而提出基于矩阵模式的最小类内散度支持向量机 $MCSVMs^{matrix}$ 及相应的非线性核方法 $Ker\text{-}MCSVM^{smatrix}$。$MCSVMs^{matrix}$ 方法不但克服了 MCSVMs 方法所面临的小样本问题，同时降低了算法本身具有的时间和空间复杂度。而且 $Ker\text{-}MCSVMs^{matrix}$ 方法首次实现了矩阵模式的非线性化；第 9 章针对经典 SVM 方法不能充分地反映样本内在几何结构及所蕴含的判别信息的问题，通过同时引入线性判别准则中的类内散度和局部保持投影 LPP 的基本原理，重新构造 SVM 的目标函数，提出基于全局和局部保持的半监督支持向量机 GLSSVM 及非线性核方法 Ker-GLSSVM。

本书由皋军独立完成。本书研究工作得到国家自然科学基金（NO：61375001，61272210）、江苏省自然科学基金（NO：BK2011417）、江苏省计算机信息处理技术重点实验室开放课题（NO：KJS1126）的共同资助。

由于时间仓促且作者水平有限，书中不当之处在所难免，敬请读者批评指正。

皋 军

2013 年 7 月

目　　录

第 1 章　绪论

1.1　课题研究背景

随着社会信息化的发展，在具体的智能识别过程中需要处理的数据越来越多地呈现出高维特征，比如图像处理、文本分类、视频检索、计算机视觉、微阵列数据基因选择和基于生物特征的身份识别等。造成这种现象的主要原因在于：在智能识别过程中，只有当样本已经包含了足够多的模式分类信息，才能得到较好的智能识别效果。然而，如何确定特征中是否已经包含足够多的类别信息本身就是个很难解决的问题。因此为了提高模式识别效果，在通常情况下，人们通过采集尽可能多的特征去体现样本的类别信息，这导致原始样本空间的维数可能达到几千维甚至上万维，而如果在如此高维的原始空间直接使用模式识别方法，那么所得到的智能识别效果将受到较大的影响。这是因为在如此高维的特征中存在着大量冗余的特征，使得特征之间的相关性较强，从而增加了模式分类算法的负担，降低算法的效率。同时，随着样本特征维数的增加，将使得对样本的统计特性更加难以估计，从而会影响分类算法的泛化能力，呈现出所谓过学习的现象[1]。因此，在智能识别过程中，必须要对高维数据进行相应的预处理，以达到在保持样本信息量的基础上尽可能降低特征的维数，提高模式分类的效果，而特征降维技术就是一种较为有效的数据预处理方法。

特征降维技术作为智能识别系统中非常关键的数据预处理方法，近年来在许多领域得到有效的运用。比如在数据可视化领域中，由于人类不能直接处理和感知真实世界的高维数据，而对低维数据的处理能力目前还强于机器智能。因此可以使用降维技术得到高维数据在低维空间上（二维或三维）的可视化投影，从而可以较为容易地发现隐藏在高维数据中的类别信息和空间分布结构；在图像识别领域，当要求处理的数据以矢量形式出现时，该数据就是一高维数据，而通常在处理如此高维数据时，为了在一定程度上防止过拟合现象的出现，一般要求具有足够大的训练集，这显然这是不可行的，因此必须对高维的数据进行有效降维，尽管在降维过程中，在一定程度上会丢失部分信息量，然而相对于智能识别的效率来说是合适的；在微阵列数据基因选择方面，DNA 微阵列数据具有明显的超高维超小样本的特点，正是由于这个特点造成了严重的维数灾难现象[1]。因此为了

有效地发现特定 DNA 数据的基因模式，找出最利于分类的基因个体，必须对此类型的基因数据进行特征降维处理。

特征降维方法在过去的几十年中被广泛地加以研究，但总体上可以将已有的方法分为两大类，即特征选择（Feature Selection）和特征提取（Feature Extraction）[2]。

1.1.1 特征选择技术

所谓特征选择，就是在原始的特征集中选取最有代表性的特征子集，重新构造一低维的样本空间。显而易见，最直观的特征选择方法就是枚举法，通过遍历原始的特征集，从所有的特征子集中寻找出最有利于智能识别的特征子集，得到全局最优解。从这一层面上讲，枚举法更适用于低维的原始样本空间，而在处理具有高维特征的数据时，枚举法将消耗大量的时间和空间资源，甚至在可计算状态下并不能获得全局最优解。因此，近年来具有时间和空间复杂度低、局部最优解或次优解特点的特征选择方法被大量地提出，比如基于支持向量机的特征选择方法[2-5]、基于概率密度估计的特征选择方法[6-7]、基于信息论的特征选择方法[8-10]和基于特征加权的特征选择方法[11-13]等，这些特征选择方法根据各自不同的评测标准来实现特征选择，而一般来说基于支持向量机、基于特征加权的特征选择方法相对于其他方法较为直观和简单。

基于支持向量机的特征选择方法一般依赖结构风险最小化原理，具有较强的泛化能力。因此，在特征选择问题上较之于基于经验风险最小的众多方法具有更好的鲁棒性。文献[2]提出的支持向量机的回归特征消除法（the SVM Recursive Feature Elimination，SVM-RFE），该方法首次使用留一交叉验证（Leave-One-Out Method）误差率作为信息准则（Information Criterion）来实现特征选择，然而该方法在评价不同特征子集的误差率时消耗了大量的额外计算，因此 SVM-RFE 方法的时间复杂度与样本特征数目成平方关系。为了在一定程度上降低 SVM-RFE 算法的时间复杂度，文献[3][4]则采用不同的特征选择判别信息准则来修正原始的 SVM-RFE 方法，从而使得改进的方法只具有线性关系的时间复杂度。然而，通过分析发现，以上几种方法在进行特征选择时还是通过选择支持向量来具体实现特征选择。文献[5]提出的势支持向量机（Potential Support Vector Machine，P-SVM）则是通过定义新的目标函数和相应的边界条件直接选取支持特征，从而提高了特征选择的效率。同时由于定义了新的边界条件，在一定程度上减少了边缘误差的传播。

而基于特征加权的特征选择方法出发点比较明确，就是通过对每一特征赋予相应的权值来表征不同特征对模式分类的贡献大小。文献[11]提出了加权 K-均值

类型聚类（Weighting in K-Means Type Clustering，W-K-Means），该方法通过无监督的模式分类方法（聚类）来得到每个特征所对应的权值，并对相应的权值进行排序，使用聚类的有效性来作为特征选择的标准。文献[12]提出了 RELIEF 特征选择方法，该方法根据识别相邻模式的区分能力来迭代产生相应特征的特征权值，算法简单有效。而文献[13]在充分分析了原始 RELIEF 理论上和算法构造上的不足后，依据最大期望原理重新构造迭代目标函数，提出了新的迭代 RELIEF 算法 I-RELIEF，该方法在一定程度上继承了原 RELIEF 算法的优点，同时可以实现多类模式分类的特征选择，提高了算法的适应性。

1.1.2 特征提取技术

不同于特征选择方法，特征提取则是对原始特征空间采用某种具体的变换映射操作，以获取低维的投影空间。总体上，特征提取可以分为线性方法和非线性方法。基于主成分析（the Principal Component Analysis，PCA）[14-15]、线性判别分析（the Linear Discriminant Analysis，LDA）[16]两种比较经典的线性特征提取方法被广泛地加以研究。PCA 方法作为一种无监督方法，是以方差大小来作为衡量信息量多少的标准，实现特征提取。然而，PCA 方法在计算相应的协方差时要做大量的浮点运算，因此相对时间复杂度较高，特别是在处理高维大数据集时尤为明显，同时 PCA 在处理一些特殊数据（比如图像数据）时，主成分在一定程度上会很难找到合适的物理解释等。和 PCA 不同，LDA 作为一种有监督的特征提取方法，在充分使用已知训练样本类别信息的前提下，通过构造所谓的类内散度和类间散度，并极大化类间散度与类内散度的广义 Rayleigh 商，以得到类间最大、类内最小的特征投影矢量，实现特征提取。该方法物理意义明确、几何含义直观，然而存在一较大问题——小样本问题，即在处理高维小样本数据时，类内散度矩阵容易发生奇异。鉴于此，近来研究者开发了很多能在一定程度上克服小样本问题的方法，比如 PCA+LDA [17]、RDA（Regularization LDA）[18]、OLDA（Orthogonal LDA）[19]、最大散度差判别准则（Maximum Scatter Difference Discriminant Criterion，MSD）[20]、2D-LDA（Two-dimensional LDA）[21]等。

诚然，PCA、LDA 方法比较适合处理线性可分的特征提取问题，而对于非线性的特征提取问题，上述两种线性方法一般并不能得到较好的效果。幸运的是，近十几年，核方法[22]被成功地运用到模式识别领域，核方法的本质就是通过定义一非线性影射将原始样本空间投影到高维特征空间（该空间一般为 Hilbert 再生核空间）[23]，使得原空间非线性不可分问题转化为在特征空间的线性可分问题，从而便于使用线性智能处理方法去处理非线性问题。此外，核方法能被广泛使用的另一个重要原因就是，在解决实际问题时并不真正需要构造相应的非线性函数去实现

将原始样本空间投影到高维特征空间，而是直接选择满足 Mercer 核容许条件[24]的核函数就可以实现。因此，将 PCA、LDA 分别同核方法相结合，就产生了能解决非线性特征提取问题的非线性核方法 KPCA（Kernel PCA）[25]、KLDA（Kernel LDA）[26]。区别于通过引进核技巧来解决非线性特征提取问题，近年来兴起的流形学习方法为实现非线性特征提取提供了另外一种思路。2000 年，Tenenbaum J B 和 Roweis S T 等人结合微分几何学中的黎曼流形理论和概念，分别提出了两种不同的流形学习方法：等距特征映射（Isometric Mapping，Isomap）[27]、局部线性嵌入（Locally Linear Embedding，LLE）[28]，从而开辟了基于流形学习的非线性特征提取方法的先河。随后，研究者在上述两种方法的基础上不断提出新的改进的算法[29-32]，然而以上各种流形学习方法虽然都能在一定程度上保持样本间内在的局部几何结构，体现内在蕴含的结构分类信息，从而发现样本内在的非线性流形结构，但都不同程度上存在所谓的 out-one-sample 问题[33]，即所对应的映射只能被定义在训练样本上，而在新的测试样本上很难获得低维的投影映射，同时还一定程度上表现为抗噪性不强。2003 年，He X F 等人提了一种新的线性流形学习方法——局部保持投影方法（Locally Preserving Projections，LPP）[33]，该方法作为拉普拉斯特征函数优化的线性近似方法，不但可以保持样本间的局部几何结构，而且又能在低维的表示空间提供显式的嵌入映射函数，同时容易被非线性嵌入，从而发现高维非线性流形结构。为了有效地提高 LPP 的泛化能力，文献[34][35]分别通过定义特殊的拉普拉斯矩阵和图嵌入（Graph Embedding）原则将 PCA、LDA 和 LPP 归结到同一个框架下，同时指出它们之间不可替代，LPP 更关注样本内在的局部几何结构，而 LDA 则更注重对样本全局信息的掌握。

除了上述的特征提取方法，近年来，为了有效处理具有张量模式的数据，大量基于张量模式（特别是基于矩阵模式）的特征提取方法被提出[36-44]，该类方法不但有利于保持张量数据固有的空间结构信息，同时还具有较低的空间和时间复杂度，而且在一定程度上可以提高特征提取效果的稳定性和鲁棒性，因此被广泛地用来解决图像（Second-order Tensor）识别和视频（Third-order Tensor）追踪等张量模式的智能识别问题。

综上所述，不管在传统的智能识别领域，还是具有广泛前景的新兴的应用模式识别领域，特征降维是数据预处理阶段不可或缺的智能处理技术，为后续的智能识别提供简单、有效的数据支撑。

1.2 特征降维技术面临的几个挑战

目前，特征降维技术作为一种关键的数据预处理技术被广泛地加以研究，并

在不同的实际运用领域得到较为成功的应用，但随着新理论和新技术的不断发展，特别是大量新兴的智能识别应用领域的需求，对特征降维技术提出了更高的要求，使得现有的特征降维技术面临更大的挑战。本课题主要关注以下几个关键挑战。

1. 如何提高基于支持向量机的特征选择方法的泛化能力和鲁棒性。传统的基于支持向量机的特征选择方法一般都是通过选取支持向量来实现特征降维，而文献[5]提出的势支持向量机（P-SVM）则通过定义新的目标函数和相应的边界条件，首次实现了对支持特征的直接选取。然而，通过结合 Fisher 判别分析[45]得知，P-SVM 方法的目标函数只是类内散度的特殊情况，因此，本课题要更加关注如何得到泛化性更强、特征冗余度低的势特征选择方法。

2. 如何更好地实现特征提取技术与模糊聚类技术[46]有机结合，以提高特征降维方法的鲁棒性。尽管目前已经有方法[11][47]可以在实现聚类的同时进行特征降维，但从总体上看，这些技术还不够成熟，这是因为在这些算法中存在着比较多的参数（比如：各种权值参数和调节参数等）需要人为地设定，而这些参数的设定往往会决定聚类效果的好坏和降维效果的稳定性。因此，本课题要对各种参数设定的合理性做较为详细的理论研究。

3. 如何提高特征降维方法中的距离度量学习的有效性。绝大部分特征降维方法都是假设原始样本空间的类型为欧氏空间，使用标准的欧氏距离度量。这样做在一定程度上是不合适的，这是因为并非所有的原始样本空间都是欧几里德空间，比如在计算机视觉中使用直方图表示的样本空间和非线性黎曼流形空间就属于非欧氏空间。因此，基于特征降维技术研究相对于原始样本空间类型依赖度较低的距离度量也是本课题重点研究的问题。

4. 如何将特征降维方法中的关键技术和理论运用到支持向量机中，以提高支持向量机的泛化能力和鲁棒性。经典的 SVM 方法作为一种有效的分类器，特别适合处理小样本问题，而且 SVM 依据结构风险最小化原则，构造相应的目标函数和边界条件，从而使得该方法具有一定分类精度和鲁棒性，但 SVM 比较关注于那些位于边界上所谓的特征向量，而并不能把握训练样本的分布信息。文献[48][49]分别将 LDA、LPP 的基本原理和方法引入到 SVM 中，分别提出了最小类内散度支持向量机（Minimum Within-class Scatter Support Vector Machines，MCSVMs）和拉普拉斯支持向量机（Laplacian Support Vector Machine，LapSVM），这些方法虽然在一定程度上反映了训练样本的分布信息，提高了支持向量机的分类精度和鲁棒性，但也存在着一些问题，比如 MCSVMs 存在着小样本问题，LapSVM 只关注训练样本内在的局部信息等，因此，研究将特征降维方法中的关键技术和理论应用到支持向量机中，提高分类的精度和鲁棒性也是本课题需要关注的问题。

5．由于在智能识别领域，越来越多的数据呈现张量模式[36-44]，如果还使用传统的子空间学习方法来处理这些具有高维特征的张量模式数据，在一定程度上会导致所谓的维数灾难[1]，同时会破坏原始数据内在的空间结构和相关性，故本课题关注于基于张量模式数据的子空间学习方法。

6．大多数的降维算法都是基于源域的训练样本和来自于目标域的测试样本具有相同分布或相同的特征空间，也就是说，训练样本和测试样本是独立同分布（Identically and Independently Distributed，I.I.D）的。但有时分布却不相同，为了解决这一非独立同分布学习问题，迁移学习（Transfer Learning，TL）方法被提出[14]。尽管目前已经有方法可以评判两个区域分布差异的度量，但从总体上看，这些技术还不够成熟。因此，本课题要对迁移学习做较为详细的理论研究。

1.3 课题的主要研究内容和组织安排

根据上述讨论的特征降维技术面临的挑战，本课题进行了相关的研究，具体内容分为七个章节，各章节概要如下：

1．第二章针对势支持向量机 P-SVM 的目标函数是 Fisher 判别分析[45]中类内散度的特殊情况，通过引入类内散度作为目标函数，提出广义的势支持特征选择方法（Generalized Potential Support Features Selection Method, GPSFM）。该方法不但继承了 P-SVMs 方法的优点，同时还具有特征冗余度低、选择速度快和分类精度高的特点，提高了特征选择的泛化能力和鲁棒性。

2．第三章通过对模糊聚类方法（Fuzzy C-Means，FCM）[46]的样本和特征同时加权的方式，提出具有特征排序功能的鲁棒性模糊聚类方法（Fuzzy Clustering Algorithm with Ranking Features and Identifying Noise Simultaneously，FCA）。该方法在聚类的同时，可以根据特征权值对特征进行排序，实现无监督的特征降维。而且在 FCA 中从权值的直观几何意义出发，合理地设定权值参数的取值范围，并从理论上加以证明，从而提高了 FCA 方法的聚类精度和特征降维的稳定性。

3．第四章针对线性拉普拉斯判别准则（Linear Laplacian Discrimination，LLD）[50]面临小样本以及如何确定原始样本空间类型的问题，通过引入语境距离度量（Contextual-distance Metric）[51]并结合最大间距判别准则的基本原理，提出一种基于语境距离度量的拉普拉斯最大间距判别准则（Contextual-distance Metric Based Laplacian Maximum Margin Criterion，CLMMC）。该准则不但在一定程度上避免了小样本问题，而且由于语境距离度量更关注输入样本簇内在的本质结构，而不是原始样本空间的类型，从而降低了该准则对特定样本空间的依赖程度。同时通过引入计算语境距离度量的新算法并结合 QR 分解的基本原理，使得 CLMMC

在处理高维矢量模式数据时更具适应性和效率。并从理论上讨论了 CLMMC 准则具有的基本性质以及与 LLD 准则的内在联系。

4．第五章基于最大散度差判别准则[20]提出了一种新的模糊最大散度差准则，并根据模糊最大散度差准则提出了一种基于模糊最大散度差判别准则的聚类方法（Fuzzy Maximum Scatter Difference Discriminant Criterion Based Clustering Algorithm，FMSDC），该方法在通过迭代优化方法实现聚类的同时，还可以得到最优鉴别矢量，实现特征降维。此方法在迭代过程中根据具体原则设定模糊最大散度差判别准则中的参数η，从而在一定程度上降低了由参数η引起的敏感性。因此 FMSDC 不但具有基本的聚类功能，同时还具有较好的鲁棒性和较强的特征降维能力。

5．第六章在最大间距判别准则（MMC）的基础上，结合模糊技术和张量理论，提出了一种矩阵模式的模糊最大间距判别准则（MFMMC）。在此基础上又提出了具有模糊聚类功能的双向二维无监督特征提取方法（$(2D)^2$UFFCA）。该方法不但能直接实现矩阵模式数据的模糊聚类，而且还可以对矩阵模式数据进行双向二维特征提取，实现特征降维。同时我们还从几何的直观含义出发，合理地设定矩阵模式的模糊最大间距判别准则中的调节参数γ，并从理论上证明其合理性。

6．第七章在最大均值差异（Maximum Mean Discrepancy，MMD）的基础上，引入局部加权均值的理论，提出投影最大局部加权均值差异（Projected Maximum Local Weighted Mean Discrepancy，PMLWD）度量，该度量能有效地度量源域和目标域中局部分块之间的分布和结构上的差异。同时，在 PMLWD 的基础上，结合基于局部加权均值的领域适应学习框架（Local Weighted Mean Based Domain Adaptation Learning Farmework，LDAF），衍生出两种领域适应学习方法：LDAF_MLC 和 LDAF_SVM。在 PMLWD 准则的基础上，提出一种能实现迁移学习任务并具有一定局部学习能力的特征提取方法——最大局部加权均值差异嵌入（Maximum Local Weighted Mean Discrepancy Embedding，MWME）。该方法不但能完成传统意义上的特征提取，同时还能完成在分布存在差异但相关的两个区域上实现领域适应学习，从而表明该特征提取方法具有较好的鲁棒性和适应性。

7．第八章针对最小类内散度支持向量机 MCSVMs 存在的小样本问题，通过引入矩阵模式的类内散度并结合矩阵模式的边界约束条件，提出基于矩阵模式的最小类内散度支持向量机（Matrix Pattern Based Minimum Within-class Scatter Support Vector Machines，$\text{MCSVMs}^{\text{matrix}}$），在此基础上运用 Mercer 核技巧首次提出真正意义上的非线性矩阵模式的支持向量机（Kernel-$\text{MCSVMs}^{\text{matrix}}$，Ker-$\text{MCSVMs}^{\text{matrix}}$）。该方法不但克服了 MCSVMs 方法的小样本问题，同时大大地降低了求解类内散度矩阵的时间和空间复杂度，特别是矩阵模式的非线性核方法的

提出更有利于求解非线性分类问题。

8．第九章针对 SVM 方法没有充分考虑训练样本局部和全局的分布信息，通过同时引入 Fisher 准则[45]的类内散度和 LPP[33]的基本原理，提出具有全局和局部保持的半监督支持向量机（Global and Local Preserving Based Semi-supervised Support Vector Machine，GLSSVM）。该方法在继承传统 SVM 方法的特点的基础上，充分考虑样本之间具有的全局和局部的几何结构，体现样本间所蕴含的局部和全局判别信息，同时满足作为半监督方法的必须依据的一致性假设，从而在一定程度上可以克服有监督方法训练不充分的不足，提高了分类精度。

本课题第 2 章到第 7 章研究和讨论特征降维方法，第 8 章和第 9 章研究和探讨特征降维方法的关键技术和理论在支持向量机中的运用，第 10 章对本课题进行总结和展望。