

SVM 和 RVM 对高光谱图像分类的应用潜能分析

张宇航, 张 晔

(哈尔滨工业大学 电子与信息工程学院, 150001 哈尔滨, zhyuhang88@hotmail.com)

摘要: 针对高光谱图像分类一直面临的小样本、非线性及高维数等问题, 分别从原理和实验两个方面分析比较了两种最新的核学习方法——支持向量机(SVM)和相关向量机(RVM)在高光谱图像分类中的异同点. 通过对稀疏性、运算时间及分类精度的实验仿真, 结果表明: 与 SVM 相比, RVM 模型更加稀疏, 从而测试时间更短, 更有利于大数据量在线测试; 然而, RVM 的缺点是分类精度略低于 SVM. 基于此, 本文利用 Fisher 线性鉴别分析(FLDA)技术, 在分类前对高光谱数据作可分性预处理, 一方面可以降低数据维数、减少计算量, 另一方面可以有效地提高小样本区域的分类精度, 进而提高 RVM 的总体分类精度, 使得 RVM 与 SVM 相比在高光谱图像精细分类方面更具优势.

关键词: 高光谱图像; 分类; 支持向量机; 相关向量机

中图分类号: TN911.73 **文献标志码:** A **文章编号:** 0367-6234(2012)03-0034-06

Potential analysis between SVM and RVM for hyperspectral imagery classification

ZHANG Yu-hang, ZHANG Ye

(School of Electronics and Information Engineering, Harbin Institute of Technology, 150001 Harbin, China, zhyuhang88@hotmail.com)

Abstract: To deal with the problems of limited samples, high dimension and non-linear in hyperspectral imagery classification, two new techniques, support vector machine (SVM) and relevance vector machine (RVM), are researched in this paper. Similarities and differences are compared and analyzed between SVM and RVM in hyperspectral imagery classification theoretically and experimentally. By simulations and experiments on classification accuracy, computational cost and sparsity, the results show that RVM model is sparser compared with SVM, which makes its test time much shorter, and thus more suitable for online testing of large amount of hyperspectral data. However, the main drawback of RVM is that its classification accuracy is slightly lower than that of SVM. To improve this performance, Fisher linear discriminant analysis (FLDA) is utilized before classification as a pre-processing to make transformation of hyperspectral data. In this way, not only the dimension of image is reduced, but also the classification accuracy of RVM is improved, especially in small land-cover patches, which makes the application of RVM more widely.

Key words: hyperspectral imagery; classification; support vector machine (SVM); relevance vector machine (RVM)

与多光谱遥感图像相比, 高光谱图像的突出

特点是光谱分辨率高, 因此可以实现多光谱不能实现的技术, 如图像的精细分类等. 然而, 高光谱图像分类一直面临着小样本、非线性及高数据维等问题, 传统遥感图像的处理模型和方法不再满足高光谱图像分类的需要, 这就给此类问题带来了困难与挑战.

收稿日期: 2011-01-25.

基金项目: 国家自然科学基金资助项目(60972143); 博士点基金资助项目(20092302110033).

作者简介: 张宇航(1988—), 女, 硕士研究生;
张 晔(1960—), 男, 教授, 博士生导师.

近年来,核方法的出现为高光谱图像分类带来了新的发展空间.其中,支持向量机(Support vector machine, SVM)和相关向量机(Relevance vector machine, RVM)是两种应用最为广泛的方法.这两种方法均是基于核映射技术,通过不同的核函数把原始低维非线性空间中的数据通过某种映射关系,映射到高维线性空间,再利用线性分类方法解决非线性问题,从而实现高光谱图像的精分类.

20世纪90年代,Vapnik等^[1]在结构风险最小化原则(Structural Risk Minimization, SRM)的基础上,提出了基于核函数的支持向量机方法.该方法具有一定稀疏性,在解决小样本、非线性及高维问题中表现出许多优势.然而SVM也存在一些缺点,主要表现为如下几个方面^[2]:

- 1) 虽然SVM模型相对稀疏,但是支持向量数目会随训练样本数的增加成线性增长;
- 2) 预测结果不具有统计意义;
- 3) 需要利用交叉验证等方法估计惩罚因子 C 的值,增加了模型训练的计算量;
- 4) 核函数必须满足Mercer条件,即核矩阵必须是连续对称正定阵.

2000年,针对以上问题,Tipping等^[2]在贝叶斯框架的基础上提出了一种与SVM相似的核学习方法——相关向量机,该方法可以有效弥补SVM的上述缺点.2007年,Demir等^[3]发表了将RVM应用于高光谱图像分类的文章,并取得了与理论相一致的预期结果.与SVM相比,RVM有更好的泛化能力,且模型更加稀疏,这就在一定程度上降低了运算的复杂度,更适用于小样本情况的应用.

目前,国内外对SVM和RVM在图像分类方面的研究非常活跃,这些研究要么是对模型算法进行的改进^[4-5],要么是研究核函数及其参数的选择问题^[6-7],而反映SVM和RVM的内在联系及差异性的文章较少.因此,本文分别从原理和实验两个方面详细比较了SVM与RVM在高光谱图像分类方面的异同点,目的是对当前SVM与RVM高光谱图像分类进行全面总结,同时也是对它们的进一步应用提供理论依据.在具体实验中,本文着重比较了两种方法的分类精度、运算时间及稀疏性,归纳总结两种方法的特点与局限性,在此基础上,分析其潜在的应用价值和发展可能.

1 原理及分析

本节主要对SVM和RVM分类相关的基本原

理进行阐述.首先,简述SVM和RVM分类的基本思想和方法,重点分析两者的共同点;之后,分别介绍两种方法各自的原理与算法,着重比较两者的差异性;最后,在二类分类的基础上介绍多类分类问题的解决办法.

1.1 分类问题原理概述

目前常用的数据分类方法中,根据在分类过程中是否需要人为指定学习样本的类别,可分为非监督分类和监督分类^[8].由于非监督分类没有使用任何关于数据的先验信息,所以分类效果通常较差,而且无法给出数据类别的含义.而在遥感数据处理中,研究者总是或多或少的掌握一些关于数据的先验信息,利用监督分类方法可以大幅度的提高分类的精度,因此本文采用监督分类方法实现高光谱图像的精分类.

监督学习方法是利用1个由输入向量 \mathbf{x} 组成的训练样本集 $\{\mathbf{x}_n\}_{n=1}^N$ 和对应的类别标号组成的标集 $\{t_n\}_{n=1}^N$,求解回归问题中的实数解或分类问题中的类别标号.希望通过学习得到输入向量 \mathbf{x} 与目标输出 t 之间的关系模型 $y(\mathbf{x})$,这样就可以由未知的输入 \mathbf{x} 得到 t 的预测值.通常情况下, $y(\mathbf{x})$ 可以通过定义在输入空间的 M 个基核函数:

$$\varphi(\mathbf{x}) = [\phi_1(\mathbf{x}) \quad \phi_2(\mathbf{x}) \quad \cdots \quad \phi_M(\mathbf{x})]^T$$
的线性组合来表示,即

$$y(\mathbf{x}; \mathbf{w}) = \sum_{i=1}^M w_i \phi_i(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}).$$

其中 $\mathbf{w} = (w_1, w_2, \dots, w_M)^T$ 为权向量^[3].

由此,监督学习过程可以概括为根据已知训练样本 $\{\mathbf{x}_n, t_n\}_{n=1}^N$,来估计权向量

$$\mathbf{w} = (w_1, w_2, \dots, w_M)^T$$

的合理取值,使 $y(\mathbf{x})$ 对于新的输入数据有较好的推广能力.

通常的SVM和RVM分类器,是用来处理两类分类问题的.分类的基本思想可以概括为:对于线性可分数据来说,训练过程的目的是寻找1个超平面,使得两类样本点能最大程度地被这个超平面区分开;对于线性不可分数据,可以借助线性分类思想,将输入空间的原始数据通过非线性映射变换到1个高维线性空间中,然后在这个新的线性空间中寻找最优分类面,使得样本的类间离散度最大,而类内离散度最小.通过SVM/RVM的训练过程,得到的权向量中只有少部分为非零值,这些非零权向量所对应的训练样本则被称为支持向量(SVs)/相关向量(RVs),而分类中实际用到的仅是这些支持向量/相关向量,这也就实现了模

型的稀疏化.

由于 SVM 和 RVM 均是基于核映射的学习方法,其高光谱图像分类过程均可由图 1 表示,分类过程主要分为离线训练和在线测试两大部分.而两种方法的差别仅在于训练过程中所用的核变换(映射)方法不同,关于 SVM 和 RVM 的核变换原理将在 1.2 及 1.3 节中详细阐述.

1.2 支持向量机原理

由上所述,SVM 分类问题可以看作是求解最优分类面的过程,因此,可以转化成带约束的优化问题,即在约束条件

$$y_i[\mathbf{w} \cdot \varphi(\mathbf{x}_i) + b] \geq 1 - \xi_i$$

下,求函数

$$\Psi(\mathbf{w}, \xi) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i$$

的最小值.其中 $\xi_i \geq 0$ 为松弛变量($i = 1, 2, \dots, N$),当类别划分出现错误时,相应的松弛变量大于 0; $C > 0$ 为惩罚因子,能够控制决策边界的形状.

在具体实现中,该带约束的优化问题可以用 Lagrange 函数来求解,从而得到的最优分类函数为

$$f(\mathbf{x}) = \text{sgn}\left\{ \sum_{i=1}^n a_i^* y_i K(\mathbf{x}_i, \mathbf{x}_j) + b^* \right\}.$$

其中: $K(\cdot, \cdot)$ 表示核函数,且必须满足 Mercer 条件.常用的核函数包括:线性核,多项式核,高斯径向基(RBF)核及 Sigmoid 核.有关 SVM 原理更详细的叙述,请参见相关的参考文献[1, 9-10].

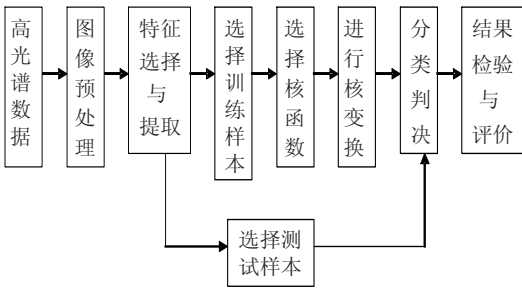


图 1 基于 SVM 和 RVM 的高光谱图像分类流程

1.3 相关向量机原理

Tipping 提出的 RVM 方法,是一种与 SVM 类似的核学习方法.不同的是,RVM 通过概率模型,引入了受超参数控制的高斯先验概率.每个权向量对应 1 个超参数,超参数的取值是通过最大化边缘似然函数来估计的.RVM 方法的最大特点是能够利用较少的训练样本,得到与 SVM 相近的分类效果.

对于两类分类问题,任何被分类目标都可以

归于其中一类,即 $t_n \in \{0, 1\}$. 因为 t 只有两种取值(0 或 1),似然函数 $p(t|\mathbf{w})$ 可以用伯努利分布来描述.在求解过程中,RVM 分类器利用 Logistic Sigmoid 函数 $\sigma(y) = 1/(1 + e^{-y})$,将线性组合 $y(\mathbf{x})$ 映射到(0,1) 区间内进行类别判定.根据伯努利分布的定义,似然函数可以表示为

$$p(t|\mathbf{w}) = \prod_{n=1}^N \sigma\{y(\mathbf{x}_n; \mathbf{w})\}^{t_n} [1 - \sigma\{y(\mathbf{x}_n; \mathbf{w})\}]^{1-t_n}.$$

如果不对权值进行约束,而直接求解似然函数的最大值,将产生严重的过拟合现象.为了提高模型的泛化能力,RVM 为每个权值 w_i 定义了高斯型先验概率分布,即

$$p(\mathbf{w}|\boldsymbol{\alpha}) = \prod_{i=1}^N \frac{\sqrt{\alpha_i}}{\sqrt{2\pi}} \exp\left(-\frac{\alpha_i w_i^2}{2}\right).$$

其中 $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_N)^T$ 为权向量

$$\mathbf{w} = (w_1, w_2, \dots, w_N)^T$$

对应的超参数.通过控制 $\boldsymbol{\alpha}$ 的大小可以改变先验概率的大小.

对于某个特定的 $\boldsymbol{\alpha}$ 值,后验概率分布可以根据贝叶斯准则获得,即

$$p(\mathbf{w} | t, \boldsymbol{\alpha}) = \frac{p(t | \mathbf{w}, \boldsymbol{\alpha}) p(\mathbf{w} | \boldsymbol{\alpha})}{p(t | \boldsymbol{\alpha})}.$$

其中: $p(t|\mathbf{w}, \boldsymbol{\alpha})$ 是似然函数; $p(\mathbf{w}|\boldsymbol{\alpha})$ 是先验概率; $p(t|\boldsymbol{\alpha})$ 是边缘似然函数.

由于权向量难以用一般的解析方法求解,因此,需要利用拉普拉斯逼近方法通过迭代来近似^[11].在迭代过程中,许多 α_i 取值趋于无穷,因此对应的权值 w_i 近似为 0 而被去除,进而剩下少量的非零权向量,这样就实现了对 RVM 模型的稀疏化.

1.4 SVM 与 RVM 多类分类问题

上述原理均是针对二类分类问题,实际应用中希望同时对多个类别数据进行分类.多类分类器可以借助二类分类器来实现.目前常用的两种方法为:一类对余类(One-against-Rest, 1-a-R)和一类对一类(One-against-One, 1-a-1)方法^[12].

1-a-R 方法的优点是简单、有效,训练时间较短,可用于大规模数据分类.但其缺点是当类别数较多,并且某一类的训练样本大大少于其他类的训练样本的总和时,这种训练样本的不均衡将对分类精度产生严重影响,且存在误分、拒分区域.

1-a-1 方法,由于每个 SVM 分类器只考虑两类样本,因此较容易训练.另外,虽然它的复杂度以类数按平方增长,但就分类速度来说,并不比通常的 1-a-R 方法慢,而且分类精度也较高.

因此,本文 SVM 及 RVM 多类分类问题将采

用 1-a-1 方法,其原理如下:在训练样本集 T (共有 k 个不同类别) 中寻找所有不同类别的两两组合 $T(i, j)$, 共有 $P = k(k-1)/2$ 个, 然后用求解两类分类问题的方法分别求出这 P 个判别函数, 如下:

$$f_{(i,j)}(\mathbf{x}) = \text{sgn}(g_{(i,j)}(\mathbf{x})).$$

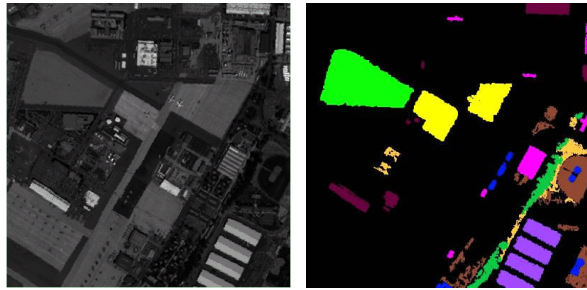
在判别时, 将输入信号 \mathbf{x} 分别输入到 P 个判别函数 $f_{(i,j)}(\mathbf{x})$ 中, 若 $f_{(i,j)}(\mathbf{x}) = 1$, 则判 \mathbf{x} 为 i 类, i 类获得一票, 否则判为 j 类, j 类获得一票. 分别统计 k 个类别在 P 个判别函数结果中的得票数, 得票数最多的类别就是最终判定类别.

2 仿真实验和结果分析

基于上述原理, 本节主要针对 SVM 和 RVM 的稀疏性、运算时间及分类精度进行实验仿真与比较.

2.1 实验数据简介

本文实验选用的高光谱图像为机载 AVIRIS 图像, 数据获取地点是美国 San Diego, 图像大小为 400×400 像素, 地面分辨率为 3.5 m. 图像经过大气辐射校正、几何校正等预处理, 并去除低信噪比波段后, 可用波段数为 126, 波段范围为 $0.4 \sim 1.8 \mu\text{m}$. 选择图像中有代表性的九类地物进行分类研究. 实验原始图像的第 25 波段示意图如图 2(a) 所示, 九类地物理想分类结果如图 2(b) 所示.



(a) 第 25 波段 (b) 所选九类地物理想分类结果

图 2 仿真实验中使用的高光谱图像

本文选择样本的策略是: 参考真实地物图, 人为地在待分类数据上选定感兴趣区域, 然后把这个区域的每个像素点提取出来. 在训练阶段, 每隔固定步长选取训练样本点, 本文分别每隔 30、100 个像素点选择 1 个像素点, 构成训练样本集 1 和训练样本集 2, 对样本进行训练; 在测试阶段, 选择全部数据作为测试样本集, 对其进行分类. 实验所用图像中各类样本数如表 1 所示. 实验中核函数选择目前应用最为广泛的 RBF 核, 其表达式如下:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2).$$

其中参数 $\gamma > 0$, 在 RBF 核中决定核的宽度.

由于参数值大小会在一定程度上影响高光谱图像的分类效果, 因此, 在 SVM 分类过程中, 需要对惩罚因子 C 和 RBF 核宽度 γ 进行合理取值. 本文根据分析和相应实验, C 取值为 2^{10} , γ 为 0.5; 而在 RVM 分类过程中, 由于不需设置参数 C 的值, 因此只需对 γ 进行合理取值, 本实验中 γ 取值为 2.5.

表 1 实验所用图像中各类样本数

| 类别标号 | 样本数 | | |
|-------|-------|-------|--------|
| | 训练集 1 | 训练集 2 | 测试集 |
| C_1 | 235 | 71 | 7 044 |
| C_2 | 158 | 47 | 4 721 |
| C_3 | 61 | 19 | 1 855 |
| C_4 | 162 | 48 | 4 851 |
| C_5 | 29 | 9 | 873 |
| C_6 | 59 | 17 | 1 748 |
| C_7 | 81 | 25 | 2 454 |
| C_8 | 72 | 21 | 2 135 |
| C_9 | 130 | 39 | 3 916 |
| Total | 987 | 296 | 29 597 |

2.2 稀疏性比较

稀疏性大小是衡量一种机器学习方法好坏的重要指标, 尤其是对于高维数、大数据量的高光谱图像, 稀疏性的大小决定了模型的复杂程度、运算效率以及分类效果, 因此, 实验比较了 SVM 与 RVM 的稀疏性, 实验结果如表 2 所示, 其中, SV_s/RV_s 表示 SVM/RVM 中所需的支持向量/相关向量数. 图 3 给出了 SV_s 和 RV_s 随训练样本数的变化规律, 图中横轴表示训练样本数占总样本数的比例, 由少到多分别为 $1/100 \sim 1/10$; 纵轴表示 SV_s/RV_s 数目.

表 2 SVM 与 RVM 稀疏性比较

| 训练集 | SV_s 数目 | RV_s 数目 |
|-----|-----------|-----------|
| 1 | 376 | 167 |
| 2 | 270 | 111 |

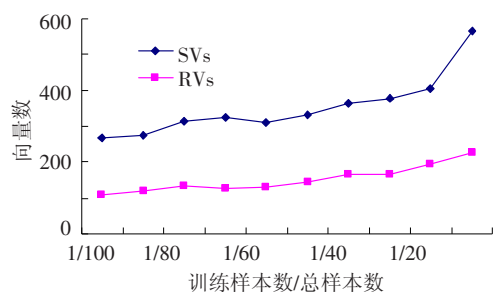


图 3 SV_s 和 RV_s 随训练样本的变化趋势

由以上结果可知,两种核学习方法均具有一定的稀疏性,且用到的向量数与训练样本数有关,当训练样本数增加时,SVs/RVs 数目会相应增加.然而,无论训练样本集如何变化,SVM 所需的支持向量数始终大于 RVM 所需的相关向量数,这就说明 RVM 具有更好的稀疏性.另外,由图 3 也可以看出,SVs 随训练样本数增加的变化率略大于 RVs,因此,在面临大量样本分类问题时,RVM 较 SVM 表现出更好的稀疏特性.

2.3 运算时间比较

利用 SVM 与 RVM 对高光谱图像进行分类所需的运算时间主要包括:对训练样本集的训练时间和对全部样本进行测试的测试时间.

表3 SVM 与 RVM 运算时间比较

| 方法 | 训练集 | 训练时间/s | 测试时间/s |
|-----|-----|--------|--------|
| SVM | 1 | 7.76 | 10.79 |
| | 2 | 6.05 | 7.88 |
| RVM | 1 | 90.76 | 3.20 |
| | 2 | 15.78 | 3.02 |

由表 3 结果可知,相比于 SVM,RVM 的训练时间较长,尤其是在训练样本数较多的情况下,训练时间会大量增加,这在某种程度上并不利于高光谱大数据量的处理.尽管 RVM 所需的训练时间较长,然而,实际应用中训练过程一般是离线进行的,对时间要求并不十分严格,而人们更关心的是样本的测试时间,由于 RVM 模型更加稀疏,因此,RVM 所需的测试时间更短,更适用于要求实时性的系统.

值得注意的是,此处所计算的运算时间,并不包含分类器参数的选择时间.由前面的原理可知,对于 SVM 分类器,需要确定惩罚因子 C 和 RBF 核宽度 γ 的取值,不同的参数取值会影响分类

结果.

然而,最优参数的选择并不是 1 个简单的工作,目前国际上常用的方法主要包括:网格搜索法和双线性法^[13].网格搜索法可以获得较高的预测精度,但计算量较大;双线性搜索法计算量较小,但预测精度略低,显然,无论哪种方法选择参数均需花费一定时间.而对于 RVM 来说,不需确定惩罚因子 C 的取值,这不仅省去繁杂的参数选择过程,也极大地缩短了整个分类过程的时间.

2.4 分类精度比较

分类问题中另 1 个非常重要的指标是分类器的分类精度,分类精度的大小直接影响着分类器的好坏.实验分别比较了 SVM 和 RVM 两种分类器的分类精度,结果如表 4 所示.

从结果可以看出,RVM 的分类精度略低于 SVM,这也是 RVM 应用于高光谱分类问题的 1 个显著缺陷.如何提高 RVM 的分类精度成为目前 1 个重要的研究方向.

在文献[14]中,作者提出一种改进方法来提高 RVM 的分类精度,即在分类之前,应用 Fisher 线性鉴别分析(FLDA)对数据进行特征提取,再应用 RVM 进行图像分类.利用该方法对实验图像进行分类,结果如表 4 中 3、4、7、8 行所示.从实验结果可以看出,该方法一方面可以有效降低高光谱的数据维数(如本实验中,数据维数由原始的 126 维降至 10 维),另一方面可以在一定程度上提高小样本区(如实验中 C_5 、 C_6 类)的分类精度,从而实现了高光谱图像的精细分类.特别是对于 RVM 分类方法,其分类精度可以提高到与 SVM 相近或是更高水平,这也就弥补了 RVM 分类精度略低于 SVM 的缺陷,使 RVM 与 SVM 相比在高光谱图像分类方面更具优势.

表4 SVM 与 RVM 分类精度比较

| 训练集 | 方法 | 分类精度/% | | | | | | | | | |
|-------|------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | | C_1 | C_2 | C_3 | C_4 | C_5 | C_6 | C_7 | C_8 | C_9 | 总体 |
| 训练集 1 | SVM | 99.90 | 97.73 | 98.33 | 94.54 | 87.97 | 92.22 | 96.74 | 96.67 | 99.26 | 97.19 |
| | RVM | 99.90 | 97.01 | 96.71 | 93.34 | 88.66 | 84.50 | 97.35 | 93.02 | 98.29 | 96.00 |
| | FLDA + SVM | 99.84 | 98.73 | 97.70 | 96.41 | 93.59 | 90.39 | 98.21 | 97.94 | 99.36 | 97.90 |
| | FLDA + RVM | 99.87 | 98.35 | 99.51 | 95.46 | 96.11 | 93.82 | 98.74 | 95.93 | 99.59 | 98.00 |
| 训练集 2 | SVM | 99.97 | 95.42 | 96.33 | 92.76 | 82.13 | 84.73 | 98.13 | 95.13 | 97.91 | 95.63 |
| | RVM | 99.99 | 94.15 | 97.95 | 88.89 | 80.64 | 80.84 | 96.37 | 88.10 | 96.73 | 93.82 |
| | FLDA + SVM | 100.00 | 96.06 | 98.49 | 92.74 | 90.16 | 86.84 | 98.21 | 97.70 | 97.50 | 96.39 |
| | FLDA + RVM | 100.00 | 97.12 | 99.03 | 93.49 | 89.35 | 82.44 | 96.98 | 97.80 | 96.76 | 96.22 |

3 结 论

本文分别从原理和实验两个方面比较了 SVM 和 RVM 在高光谱图像分类方面的异同点. SVM 和 RVM 均可有效解决高光谱图像带来的小样本、非线性及高维数等问题. 首先,从分类原理来看,两种方法的基本思想均是通过核函数将输入空间的非线性数据映射到 1 个高维线性可分空间中,然后在这个新的线性空间中寻找最优分类面,使得样本的类间离散度最大,而类内离散度最小. 不同点如下所示.

1) SVM 是基于结构风险最小化原则构建的学习机,而 RVM 是基于贝叶斯框架构建的学习机,因此,与 SVM 相比,RVM 不仅能获得二值输出,而且能获得概率输出;

2)在核函数的选择上,RVM 不受 Mercer 定理的限制,可以构建任意的核函数,且具有较好的泛化能力. 其次,从实验结果来看,两种方法均具有一定稀疏性,因此适用于高维、大数据量的高光谱图像分类问题. 而与 SVM 相比,RVM 模型更加稀疏,从而测试时间更短,更有利于在线测试;并且 RVM 不需对惩罚因子 C 做出设置,省去了大量参数选择时间. 然而,RVM 的主要缺点在于分类精度略低于 SVM,且训练时间较长,因此 RVM 更适用于要求实时性而对分类精度要求不高的系统. 文章最后应用 Fisher 线性鉴别分析技术对原始数据进行特征提取,不仅可以降低数据维数,而且能够有效提高 RVM 的分类精度,使得 RVM 的应用前景更加广泛. 如何继续提高 SVM 的稀疏性及 RVM 的分类精度是下一步研究重点.

参考文献:

- [1] VAPNIK V N. The nature of statistical learning theory [M]. New York: Springer-Verlag, 1995: 123 - 167.
- [2] TIPPING M E. Sparse Bayesian learning and the relevance vector machine [J]. Mach Learn Res, 2001, 1 (3): 211 - 244.
- [3] DEMIR B, ERTÜRK S. Hyperspectral image classification using relevance vector machines [J]. IEEE Geosci Remote Sens Letters, 2007, 4(4): 586 - 590.
- [4] DEMIR B, ERTÜRK S. Improving SVM classification accuracy using a hierarchical approach for hyperspectral Images [C]//Proceeding of 2009 IEEE International Conference on Image Processing (ICIP). Egypt: IEEE Signal Processing Society, 2009: 2849 - 2852.
- [5] CAMPS-VALLS G, RODRIGO-GONZALEZ A, MUÑOZ-MARI J, *et al.* Hyperspectral image classification with mahalanobis relevance vector machines [C]//Proceeding of 2007 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Barcelona: IEEE International, 2007: 3802 - 3805.
- [6] KAYA G T, ERSOY O K, KAMASAK M E. Support vector selection and adaptation for remote sensing classification [J]. IEEE Trans Geosci Remote Sens, 2011, 49(6): 2071 - 2079.
- [7] WU B, ZHANG W Q, CHEN L, *et al.* A GP-based kernel construction and optimization method for RVM [C]//Proceeding of 2010 International Conference on Computer and Automation Engineering (ICCAE). Singapore: IEEE Computer Society, 2010: 419 - 423.
- [8] MURAT D M, LANDGREBE D A. Toward an optimal supervised classifier for the analysis of hyperspectral data [J]. IEEE Trans Geosci Remote Sens, 2004, 42 (1): 271 - 277.
- [9] MULLER K R, MIKA S, RÄTSCHE G, *et al.* An introduction to kernel-based learning algorithms [J]. IEEE Trans. on Neural Networks, 2001, 12(2): 181 - 201.
- [10] RUIZ A, LÓPEZ-DE-TERUEL P E. Nonlinear kernel-based statistical pattern analysis [J]. IEEE Trans on Neural Networks, 2001, 12(1): 16 - 32.
- [11] MACKAY D J C. The evidence framework applied to classification networks [J]. Neural Comput, 1992, 5 (4): 720 - 736.
- [12] 薛明东, 郭立. 基于 SVM 算法的图像分类 [J]. 计算机工程与应用, 2004, 40(30): 230 - 232.
- [13] 李琳, 张晓龙. 基于 RBF 核的 SVM 学习算法的优化计算 [J]. 计算机工程与应用, 2006, 42(29): 109 - 204.
- [14] MIANJI F A, ZHANG Y. Improved hyperspectral land-cover analysis using relevance vector machine [C]//Proceeding of 2010 International Conference on Image Processing (ICIP). Piscataway: IEEE, 2010: 2281 - 2284. (编辑 张 宏)