

# 基于分类器集成的高光谱遥感图像分类方法

樊利恒<sup>1</sup> 吕俊伟<sup>1</sup> 邓江生<sup>2</sup>

(<sup>1</sup> 海军航空工程学院控制工程系, 山东 烟台 264001)  
(<sup>2</sup> 海军航空仪器计量站, 上海 200436)

**摘要** 高光谱遥感图像为地物的精确分类带来了机遇,但也面临着一些挑战,高光谱遥感图像分类中所面临的一个挑战是如何处理高的光谱维数和小的样本数目之间的矛盾,目前几乎全部采用降维方法来缓解这一矛盾。集成学习的出现和选择性集成概念的提出为解决这一问题提供了新的研究思路,基于这一思想提出了基于波段分组和分类器集成的方法。在高光谱遥感图像的原始光谱空间根据波段之间的相似性信息对光谱波段进行分类,从每类中随机抽取一个波段形成新的光谱组,并依靠限制不同光谱组中相同波段的数目增加不同光谱组之间的差异程度,将新的光谱组作为训练分类器的特征子集,在特征子集训练最大似然分类器,使用简单的多数投票法合成得到最终的集成分类器。实验结果表明,使用基于波段分组和分类器集成的方法可以得到更高的分类精度。

**关键词** 遥感;高光谱;分类器集成;波段分组;最大似然分类

中图分类号 TP751.1 文献标识码 A doi: 10.3788/AOS201434.0910002

## Classification of Hyperspectral Remote Sensing Images Based on Bands Grouping and Classification Ensembles

Fan Liheng<sup>1</sup> Lü Junwei<sup>1</sup> Deng Jiangsheng<sup>2</sup>

(<sup>1</sup> Department of Control Engineering, Naval Aeronautical Engineering Institute, Yantai, Shandong 264001, China)  
(<sup>2</sup> Naval Aeronautical Measure Station, Shanghai 200436, China)

**Abstract** The conflict of high dimensionality and the very limited number of available training samples is one of the problems in the classification of hyperspectral images. At the same time, the redundancy between different bands brings trouble to the classification. The ensemble learning provides a new way for solving the problem mentioned above. Based on the correlation between different bands, a band grouping is carried out. By selecting different bands from different groups new subsets of spectral bands is formed. The redundancy reduces bands in the new spectral band subsets are independent and used to train the maximum likelihood (ML) classifiers which can be used later for ensembling. The combining of classifiers is done by the simple majority voting and the ensemble classifier is formed. Experimental results of the hyperspectral remotely sensing image demonstrate that the method presented here has an excellent classification result and outperforms many other methods.

**Key words** remote sensing; hypersepctral; ensemble of classifiers; bands grouping; maximum likelihood classification

**OCIS codes** 100.2960; 100.4145; 100.5010

## 1 引 言

近年来,高光谱成像技术在国内外取得了重大

发展并在众多领域得到成功运用<sup>[1-2]</sup>。高光谱遥感将传统的光学成像技术与细分光谱技术结合在一

收稿日期: 2014-03-11; 收到修改稿日期: 2014-05-07

基金项目: 国家自然科学基金(61032001, 60801049)、国家 863 计划(2010AAJ140)

作者简介: 樊利恒(1986—),男,博士研究生,主要从事高光谱遥感图像分类与典型目标识别方面的研究。

E-mail: fan\_li\_heng@126.com

导师简介: 吕俊伟(1960—),男,博士,教授,博士生导师,主要从事机器视觉、目标识别与跟踪等方面的研究。

E-mail: ljwei369@163.com

起,在进行空间特征成像的同时,对目标的每个空间像元经过色散形成几十乃至几百个窄波段以进行连续的光谱覆盖,所采集的图像包含了丰富的空间、辐射和光谱三重信息,为地物的精确分类带来了机遇。然而高光谱遥感图像光谱分辨率的改善是以增加数据量和提高数据维数作为代价的,在提供丰富信息的同时波段之间具有较强的冗余性,为高光谱图像遥感地物分类和目标识别等实际应用带来了极大的困难,降低了数据处理的效率,甚至会影响分类的精度。此外,Hughes现象(即随着光谱维数的增加,分类精度先增加后降低的现象)反映了高的光谱维数和小的样本数目之间的矛盾,是高光谱遥感图像分类中所面临的一个难题<sup>[3-6]</sup>。

目前解决这一问题通常采用降维算法,将高维数据变换到有意义的低维表示。降维能够在一定程度上减轻Hughes现象,有效降低因较高的维数而导致的信息冗余,提高分类器性能。根据降维策略的不同,降维方法可以归纳为特征提取与波段选择<sup>[7]</sup>。特征提取如主成分分析、线性判别分析等,是通过数据变换的方式来实现降维的,它有严格的数学理论作为支撑,直接将数据从高维变换到低维,可以获得很好的降维效果,已经在高光谱分类中发挥了重要作用。然而该方法改变了高光谱数据原有的特性,会导致原始波段物理信息部分或全部丢失。波段选择如基于主成分变换的波段选择、基于聚类集成的波段选择等,是依据某个准则从原始波段中选择部分最能反映类别区分统计特性的相关波段。特征波段选择是直接从原始光谱空间的波段中选择用于分类的波段,波段的物理信息得以完整保留,有利于保持图像的原有特性。但特征选择算法也有其自身的局限性,由于运算量过大,使得算法需要较长的处理时间,在处理时的往往会设置一些强的假设和限制,一旦这些假设往往不成立,就会导致求解精度低和稳定性差。

集成学习作为机器学习的4个重大研究方向之一<sup>[8]</sup>,随着理论依据的完善,正引起越来越多研究者的关注。集成学习分类使用多个基分类器来解决同一分类问题,通过一些简单的分类算法,可以得到多个不同的基分类器,然后采用某种合成方式将这些基分类器组合成一个集成分类器。可以看出集成方法思路侧重于两个步骤:1)调用某些算法生成基分类器;2)选择特定的策略来组合基分类器。而在构建集成分类器时,有效地产生分类精度高、差异大的基分类器是关键。

集成学习的出现和选择性集成的提出为解决高的光谱维数与小的样本数目之间的矛盾和波段之间较强的相关性和冗余性这些问题提供了新的研究思路。基于此本文提出了一种将集成分类用于解决高光谱遥感图像分类中存在的高的数据维数和小的样本数目之间矛盾的方法。关于集成分类的具体研究思路为:1)有效地产生分类精度高、差异大的基分类器是集成分类的关键,而产生训练基分类器的特征子集的差异性在某种程度上决定了基分类器之间的差异性,这里提出波段分组的方法来产生具有差异性的特征子集,基分类器采用最大似然分类器,因为最大似然分类器不仅实现简单,而且在理论上有着最高的分类精度;2)在基分类器合成阶段,采用选择性集成方案,计算基分类器之间的相关性,选择两两之间负相关且满足精度要求的基分类器参与集成,基分类器合成采用简单多数投票策略。

## 2 基分类器的构建

### 2.1 波段分组

产生训练基分类器的特征子集的差异性在某种程度上决定了基分类器之间的差异性,提出了一种波段分组的方法来构造具有差异性的特征子集方法。将波段分组方法通过两个过程实现,分别为波段分类和重新分组。可以通过分析相邻波段间的信息量和相关程度来对原始波段空间中的波段进行分类,在分类过程完成后,从每个光谱类中选择一个波段组成一个新的光谱组,则这个光谱组包含了分类的基本光谱信息。

与普通分类相似的是,每一个波段类内的波段之间的相似度较高而不同类之间波段的相关度较低。定义波段间的相关程度,这里使用相邻波段之间的平均灰度差异、相邻波段之间的平均二次灰度差异和相邻波段的互信息这三种不同的参数来衡量相邻波段之间的相关程度。

假设 $X$ 为 $D$ 维高光谱遥感图像,像素数目为 $M \text{ pixel} \times N \text{ pixel}$ ,则波段数目为 $D$ 。 $X$ 可以表示为集合的形式,集合中的元素为单色图像即单一波段的图像,即 $X = [X_1, X_2, \dots, X_D]$ ,则 $X_i (i = 1, 2, \dots, D)$ 为大小为 $M \text{ pixel} \times N \text{ pixel}$ 的单色图像,是一个二维的矩阵。定义不同波段之间平均灰度差异 $\mu_{i,j}$ 为

$$\mu_{i,j} = \frac{\sum_{l=1}^M \sum_{k=1}^N |X_i(l,k) - X_j(l,k)|}{MN}, \quad (1)$$

式中  $X_i(l, k)$  表示  $X$  中第  $i$  波段中像素  $(l, k)$  的灰度。 $\mu_{i,j}$  越大则相关程度越低。

定义不同波段之间的平均二次差异程度  $\sigma_{i,j}$  为

$$\sigma_{i,j} = \frac{\sum_{l=1}^M \sum_{k=1}^N |X_i(l, k) - X_j(l, k)|^2}{MN}, \quad (2)$$

式中  $X_i(l, k)$  表示  $X$  中第  $i$  波段中像素  $(l, k)$  的灰度。 $\sigma_{i,j}$  越大则相关程度越低。

互信息是一种得到广泛应用的信号相关性度量,对变量的分布类型没有特殊要求,不仅能描述变量间的线性相关关系,也能描述变量间的非线性相关关系。对于高光谱遥感图像,可以用各波段间的互信息来衡量它们之间的信息依赖程度,较大的互信息往往意味着较高的相关程度<sup>[9-11]</sup>。下面介绍高光谱遥感图像波段间互信息的计算方法。

引入信息熵的概念,对于随机变量  $Y$ ,信息熵  $H$  定义为

$$H(Y) = - \sum_{y \in Y} p(y) \lg p(y), \quad (3)$$

式中  $p(y)$  是事件  $y$  的概率密度函数,求和对  $Y$  所有可能发生的事件进行。

考虑随机变量  $Y$  和  $Z$ ,互信息  $I$  定义为

$$I(Y, Z) = \sum_{y \in Y} \sum_{z \in Z} p(y, z) \lg \frac{p(y, z)}{p(y)p(z)}, \quad (4)$$

式中  $p(z)$  是事件  $z$  的概率密度函数,  $p(y, z)$  是事件  $y$  和事件  $z$  的联合概率密度函数,求和对  $Y$  和  $Z$  所有可能发生的事件进行。

互信息  $I$  也可以表达为信息熵的形式:

$$I(Y, Z) = H(Y) + H(Z) - H(Y, Z), \quad (5)$$

式中  $H(Y, Z)$  是联合熵,其定义为  $H(Y, Z) = - \sum_{y \in Y, z \in Z} p(y, z) \lg p(y, z)$ 。

观察高光谱数据的波段间的互信息。对于高光谱数据  $\mathbf{X} = [X_1, X_2, \dots, X_D]$ , 总共包含  $D$  个波段,像素数目为  $M \text{ pixel} \times N \text{ pixel}$ 。 $X_i$  和  $X_j$  分别代表了第  $i$  和第  $j$  个波段的图像信息,则可以定义  $X_i$  和  $X_j$  之间的互信息。

由此可以得到关于相邻波段之间冗余程度的参数  $\mu_{i,i+1}$ 、 $\sigma_{i,i+1}$  和  $I_{i,i+1}$ 。为了比较方便,接下来根据这 3 个参数构造一个总的指标  $\epsilon$ ,把  $\epsilon$  称为波段差异系数,简单的思路是将 3 个指标简单相加,由于 3 个指标的取值范围不同,因而直接相加导致取值范围较大的指标占主导地位,此外互信息和另外两个指标的性质也有不同,对互信息求倒数,再将 3 个指标线性映射到  $[0, 1]$  的范围内,最后进行线性相加,在不

引起混淆的情况下,仍用原来的符号表示 3 个参数,波段差异系数  $\epsilon$  的求取公式如下。

$$\epsilon_{i,j} = \mu_{i,j} + \sigma_{i,j} + \frac{1}{I_{i,j}}, \quad (i, j = 1, 2, \dots, D), \quad (6)$$

得到整个光谱波段的波段差异系数矩阵  $\boldsymbol{\epsilon}$ ,  $\epsilon_{i,j}$  越大则波段  $i$  和波段  $j$  之间的相关程度越低。

可以根据相邻波段的差异系数对整个光谱波段进行预处理,可以减小后续工作的计算复杂度,排除噪声波段的干扰。设定两个阈值噪声波段阈值 ( $t_{hh}$ ) 和相同波段阈值 ( $t_{hl}$ ),如果相邻波段间的差异系数  $\epsilon_{i,i+1}$  满足  $\epsilon_{i,i+1} \leq t_{hl}$ ,则认为波段  $i$  和  $i+1$  相同,随机删除其中一个波段;如果波段  $i$  与相邻波段之间的差异系数满足  $\epsilon_{i-1,i} \geq t_{hh} \& \epsilon_{i,i+1} \geq t_{hh}$ ,则认为波段  $i$  是噪声波段,予以删除。至此,预处理工作完成。

在得到相邻波段的差异系数之后,介绍采用的波段分组方法。分类原则是将相似程度较大的相邻光谱分为一类,用对整个原始波段求特征值的方法确定光谱分类的数目  $K$ ,具体为对整个波段求取特征值,对所有特征值排序,求和,如果前  $K$  个最大的特征值之和占有所有特征值总和的比例大于某一数值  $t$  (本文选的是 98%),则确定光谱分类数目是  $K$ ,在确定波段分组数目  $K$  之后将与前  $K$  值对应的波段作为分类界限,如  $\epsilon_{i,i+1}$  属于前  $K-1$  个最大的特征值,则认为波段  $i$  和波段  $i+1$  之间存在分类界限,波段  $i$  和波段  $i+1$  分属于不同的光谱类。进行重新分组,按照波段的编号对光谱类编号为  $1, \dots, K$ ,从第 1 光谱类中选择第 1 个波段,从第 2 光谱类中选择第 1 个波段,  $\dots$ , 从第  $K$  光谱类中选择第 1 个波段组成第 1 个光谱组;然后把第 1 个光谱组中来自第  $K$  类的波段换为第  $K$  类的第 2 个波段,直到把所有的光谱组合方式都完成,这样的方法简单,而且能够得到所有可能的光谱组合。如此方法得到的光谱分组中存在这样的情况,即两个光谱分组中有一半以上的波段相同,这会导致在这两个光谱分组上训练的分类器差异很小,对集成来说这是有害的,因此在上述过程完成后检测光谱组之间的相似性,如果有两个光谱分组的相同波段数目超过  $\lfloor K/2 \rfloor$  (其中  $\lfloor \cdot \rfloor$  表示向下取整),随机删除一个。

对波段分组步骤总结如下:

输入:原始高光谱遥感图像数据  $\mathbf{X} = [X_1, X_2, \dots, X_D]$ , 相同波段阈值  $t_{hl}$ , 噪声波段阈值  $t_{hh}$ , 特征根值比例  $t$ ;

输出:所有高光谱波段组的集合  $S = [S_1, S_2, \dots, S_L]$ ;

算法过程:

1) 预处理:计算波段之间的差异系数矩阵  $\boldsymbol{\varepsilon}$ , 并得到相邻波段之间的差异系数  $\varepsilon_{i,i+1}$

```
if  $\varepsilon_{i,i+1} < t_{hl}$ 
delete band  $i+1$  or band  $i$ 
else if  $\varepsilon_{i-1,i} > t_{hh}$  and  $\varepsilon_{i,i+1} > t_{hh}$ 
delete band  $i$ 
end
```

删除的波段数目为  $P$ , 则剩余的高光谱遥感图像数据为  $X' = [X'_1, X'_2, \dots, X'_{D-P}]$ 。

2) 对整个光谱特征波段求特征根, 得到特征根组成的向量  $\mathbf{e}_{ig}$ , 对  $\mathbf{e}_{ig}$  从大到小排序, 得到  $\mathbf{e}_{ig} = [e_{ig1}, e_{ig2}, \dots, e_{igD-P}]$ ;

$K \in [1, D-P]$ ; 如果  $e_{ig1} + \dots + e_{igK} > (e_{ig1} + \dots + e_{igD-P})t$  且  $e_{ig1} + \dots + e_{igK-1} < (e_{ig1} + \dots + e_{igD-P})t$ , 则  $K$  为光谱分类的数目;

3) 对  $\varepsilon_{i,i+1}$  排序, 取前  $K-1$  个最大的差异系数组成的向量为  $\boldsymbol{\varepsilon}' = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{K-1}]$ , 并记录  $\varepsilon_j$  ( $j = 1, \dots, K-1$ ) 对应的  $\varepsilon_{i,i+1}$ ;

$\varepsilon_j \in \boldsymbol{\varepsilon}'$ , 相应的差异系数为  $\varepsilon_{i,i+1}$ , 则波段  $i$  和之前波段分为一类, 波段  $i+1$  和之后的波段分为一类, 共得到  $K$  个光谱类, 每个光谱类的波段数为  $K_i$ , 则  $\sum K_i = D-P$ ;

5) 从每个光谱类中选择一个波段组成一个新的集合, 则新的光谱组合包含  $K$  个波段, 用遍历法得到所有可能的组合, 共  $\prod K_i$  个;

6) 比较任意两个光谱组合  $S'_i$  和  $S'_j$  之间相同光谱波段的数目  $N_{ij}$ , 则  $N_{ij} = N_{ji}$ , 如果  $N_{ij} > |K/2|$ , 删除  $S'_i$  和  $S'_j$  中任意一个;

7) 剩余的高光谱波段组共有  $(\prod K_i) |K/2| / K$  个, 作为训练基分类器的特征集。

## 2.2 基分类器的训练

近年来, 随着统计理论、模糊理论、模式识别、机器学习等方法的广泛研究, 多种高光谱图像分类算法相继发展起来。这些都为基分类器的构造提供了理论支撑, 同时也提供了多项选择。

最大似然(ML)分类方法是一种典型的基于统计分析的监督分类方法, 从理论上讲具有最小的出错率与最高的分类精度, 它利用遥感数据的统计特征, 假设各类样本数据服从高斯分布, 根据 ML 比贝叶斯判决准则建立非线性判别函数, 计算待测样本像素对不同类别的归属概率。Shafri 等<sup>[12]</sup>将 ML

分类方法与更为先进的人工神经网络(ANN)、光谱角制图(SAM)和决策树(DT)方法的分类效果进行对比, 发现 ML 的分类精度是最高的。这一结论显示了考虑数据和分类器之间相互关系对于高光谱遥感图像成功分类的重要性, 同时也说明了经典的 ML 分类方法仍然有机会比更先进的分类方法得到更高的分类精度。但将 ML 分类器作为基分类器参与集成一直没有得到学者应有的重视, 主要原因是 ML 作为一种统计分类方法不易构造具有差异性基分类器<sup>[13]</sup>, 而通过波段分组可以构造不同的特征子集, 从而产生具有差异的 ML 分类器, 使将 ML 分类器作为基分类器参与集成成为可能。

ML 分类算法由下式表示为

$$g_i(\mathbf{x}) = \ln p(\omega_i) - \ln |\boldsymbol{\Sigma}_i| - (\mathbf{x} - \mathbf{m}_i)^T \boldsymbol{\Sigma}_i^{-1} (\mathbf{x} - \mathbf{m}_i), \quad (7)$$

式中  $\mathbf{x}$  为像元光谱向量,  $\boldsymbol{\Sigma}_i$  为第  $i$  类的大小为  $K \times K$  的协方差矩阵,  $K$  为光谱波段数,  $\mathbf{m}_i$  为第  $i$  类均值向量,  $i = 1, 2, \dots, C$ ,  $C$  为类别数。ML 分类的实现需要从选择的训练样本中估计类均值向量和协方差矩阵。

判别准则为: 对于类别  $j$  ( $j = 1, 2, \dots, C; j \neq i$ )

$$\text{if } g_i(\mathbf{x}) > g_j(\mathbf{x}), \text{ then } x \in \omega_i, \quad (8)$$

式中  $\omega_i$  表示第  $i$  类别。

## 3 基于选择性集成的分类器合成

面对多个基分类器的分类结果, 需要寻求一种整合策略以获得最佳决策。分类器结果合成的思想来源分别是叠加法和元学习法。叠加法的思想是把基学习器分布在多个层次上, 用多个层次的学习器来完成学习任务。元学习算法的思想是训练一个元分类器对所有的基学习器的输出进行处理, 最终得到问题的输出<sup>[14]</sup>。根据是否所有基分类器都参与集成可以将集成策略分为以下两类:

1) 全部生成的分类器个体都参与集成的投票法

主要包括多数投票法和加权投票法。多数投票法的基本思想是多个基分类器进行分类预测, 通过某种投票原则进行投票表决, 是最简单, 也是最普遍的结论合成方法。加权投票法对每个分类器成员赋予一定的权重, 权重通过在训练集上测量每个成员分类器精度获得, 且权重与精度成正比, 即分类能力好的基分类器被赋予较大的权系数, 而分类能力相对差的基分类器赋予较小的权系数, 集成的结果取决于加权和。

2) 从集成中选择出部分个体的选择性集成学习算法

选择集成算法的基本思想就是利用对多个学习器进行适当的选择来剔除对学习系统有负作用的学习器,从集成中选择出部分个体参与集成,最后将所选择的部分个体进行结合从而得到比全部个体都参与集成性能更好的学习器。Zhou 等<sup>[15]</sup>提出了相应的选择性集成算法 GASEN,其理论分析和实验结果均表明,该算法性能优于 Boosting 和 Bagging 算法。

Brown 等<sup>[16-17]</sup>对基分类器的差异性和独立性进行了进一步的研究,并认为在多数投票集成时差异性有“好”有“坏”,分类器相互独立也并不是最有利于集成的情况。Kuncheva 等<sup>[18-20]</sup>将负相关学习引入差异性,认为负相关的分类器能更好地学习训练样本的不同方面,提高集成分类的性能和分类精度,并通过简单的公式和实验进行了证明。虽然分类器之间负相关与集成分类器的精度之间的具体关系还没被研究清楚,但他们的研究表明当分类器之间负相关时,总能保证得到优于单个分类器的分类结果。

可以用 Q 统计来计算两个基分类器之间的相关性。假设来自分类问题的数据集为  $Z = [z_1, \dots, z_N]$ ,

$z_k \in \mathcal{R}^N$ 。分类器  $f_i$  的分类结果表示为向量的形式  $y_i = [y_{1,i}, \dots, y_{N,i}]$ ,如果  $f_i$  对  $z_k$  正确分类,则  $y_{k,i} = 1$ ,否则  $y_{k,i} = 0$ 。则分类器  $f_i$  和  $f_j$  的 Q 统计为

$$Q_{ij} = \frac{N^{11}N^{00} - N^{01}N^{10}}{N^{11}N^{00} + N^{01}N^{10}}, \quad (9)$$

式中  $N^{ab}$  是数据集中  $y_{k,i} = a$  和  $y_{k,j} = b$  的个数。

虽然在生成训练基分类器的特征子集时考虑到了基分类器的差异性并对任意两个特征子集之间相同波段的个数进行了限制,但这些措施不一定能保证生成的基分类器之间的差异性,仍需要对所有的基分类器进行选择。选择基分类器不能单独根据单个分类器的分类精度,还应该考虑分类器之间的相关性。在训练基分类器之后,计算基分类器之间的相关性,然后选择分类精度高、负相关的那些基分类器。对选择出来的基分类器合成使用简单的多数投票法合成,得到最终的集成分类器。

设选择的部分负相关基分类器为  $f_i (i=1, \dots, S)$ ,  $S$  为选择的基分类器的数目,则简单多数投票法的最终分类结果  $f$  可以表示为

$$f(x, j) = \text{if} \left[ \sum_{i=1}^S f_i(x, j) > \frac{S}{2} \right], \quad (10)$$

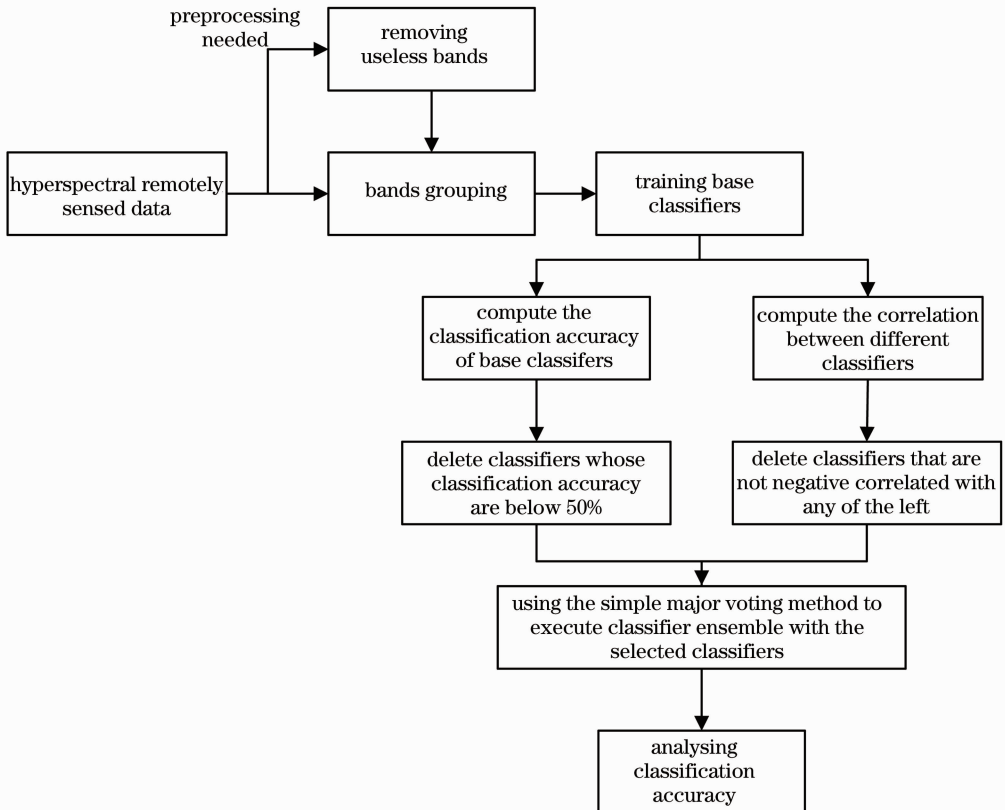


图 1 基于分类器集成的分类方法的流程图

Fig. 1 Flowchart of classification method based on ensemble of classifiers

式中  $f_i(x, j)$  表示分类器  $f_i$  对样本  $x$  的分类结果,  $j$  是类别标号, 如果  $x$  属于类别  $j$ , 则  $f_i(x, j) = 1$ , 否则  $f_i(x, j) = 0$ , 如果  $\sum_{i=1}^S f_i(x, j) > S/2$ , 则  $f(x, j) = 1$ , 说明样本  $x$  属于类别  $j$ 。

基于分类器集成的高光谱遥感图像分类方法流程如图 1 所示。

## 4 实验和结果分析

### 4.1 实验数据

#### 4.1.1 实验数据 1

实验所选用的数据是美国可见与红外成像光谱仪 (AVIRIS) 高光谱遥感数据 92AV3C。数据采集自美国印第安纳州西北部的一块印度松树测试地, 该高光谱遥感数据是公开的基准高光谱图像数据, 目前已有大量的国内外学者在高光谱图像处理的研究上采用了该

数据, 因此有利于重复实验和不同的算法之间的比较。该图像大小为  $145 \text{ pixel} \times 145 \text{ pixel}$ , 波长范围为  $0.4 \sim 2.5 \mu\text{m}$ , 原始波段数为 220 个, 移除水吸收波段(水吸收波段为第 104 波段到第 108 波段和第 150 波段到第 162 波段)后的波段数一共有 202 个, 图像的空间分辨率约为 20 m, 光谱分辨率小于 10 nm, 数据采集于 1992 年 6 月。将该数据的第 50 个波段作为红色, 第 27 个波段作为绿色, 第 17 个波段作为蓝色的 RGB 显示如图 2(a) 所示。可以看出图像的大约 2/3 被农作物覆盖, 大约 1/3 被森林和其他多年生植被覆盖, 由于拍摄的季节的原因, 耕地上的农作物没能较好的覆盖土地, 此外还有几条公路。其参考分类图像如图 2(b) 所示, 从图 2(b) 中看以看出该数据包括 16 个地物类别, 数据的地物类别情况如表 1 所示(高光谱数据图和地面观测数据图来源于下面的网站: <http://dynamo.ecn.purdue.edu/~biehl/MultiSpec>)。

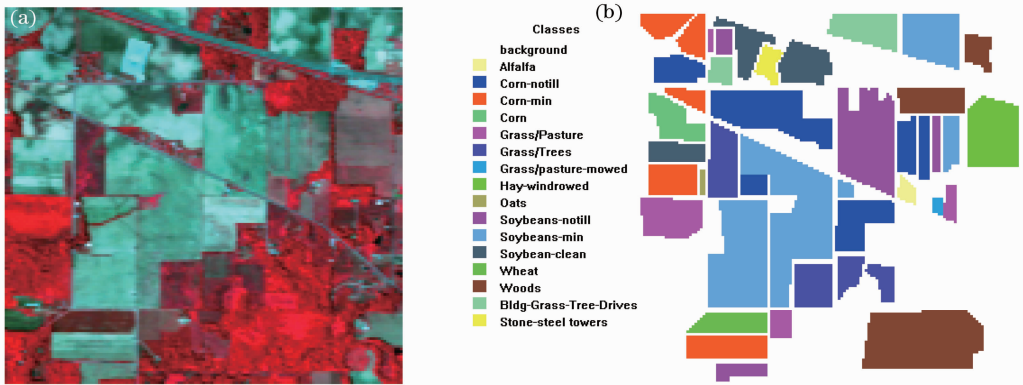


图 2 92AV3C 图像示意图。(a)假彩色合成图; (b)参考分类图

Fig. 2 Map about the hyperspectral dataset 92AV3C. (a) 92AV3C displayed in simulated color;

(b) reference map of the dataset

表 1 高光谱遥感数据的地物类别及样本数目 1

Table 1 Classes and number 1 of samples used in the experiments

No.	Landcover types	Number of samples
1	Alfalfa	54
2	Corn-no till	1434
3	Corn-min till	834
4	Corn	234
5	Grass/pasture	497
6	Grass/trees	747
7	Grass/pasture-mowed	26
8	Hay-windrowed	489
9	Oats	20
10	Soybean-no till	968
11	Soybean-min till	2468
12	Soybean-clean till	614
13	Wheat	212
14	Woods	1294
15	Bldg-grass-tree-drives	380
16	Stone-steel towers	95

由于类别 1、4、7、9、13、15 和 16 的样本数小于波段数的 2 倍, 有些文献只考虑了类别 2、3、5、6、8、10、11、12 和 14, 在从这 9 类的样本中随机选择一部分作为训练样本, 剩下的一部分则作为测试样本。

#### 4.1.2 实验数据 2

实验数据来自美国 AVIRIS 高光谱遥感数据 av920612\_NS\_line.lan, 如图 3(a) 所示(将该数据的第 50、27、17 个波段分别作为 RGB 合成显示)。该图像的空间分辨率约为 20m, 光谱分辨率小于 10 nm, 数据采集于 1992 年 6 月, 该高光谱遥感图像共有 220 个波段(包括噪声波段和水吸收波段), 图像大小为  $2478 \text{ pixel} \times 614 \text{ pixel}$ , 由图 3(b) 可以看出该图像一共可以分为 58 类。由于图像较大, 而且有较大一部分没有参考分类, 难以对分类结果进行评估和比较, 因此选择其中一部分[坐标范围为

[2389:2588,181:380]]进行实验,其原始图像和参考图像分别如图 4(a)和图 4(b)所示,该图像共有 23 个地物类别和 34225 个样本,包含样本最多的地物类为 Corn-Clean Till-EW,共有 7055 个样本,包含样本最少的地物类为 BareSoil,共有 24 个样本,所有 23 类地物共有 34225 个样本。由于一般认为

样本数目是波段数的 2~4 倍时分类效果最好<sup>[21]</sup>,而满足这一条件的地物共有 15 类。考虑到使用降维算法,有可能降维之后的波段数小于最少样本数的 1/2,所有这里保留了所有的地物类别,数据的地物类别情况如表 2 所示。

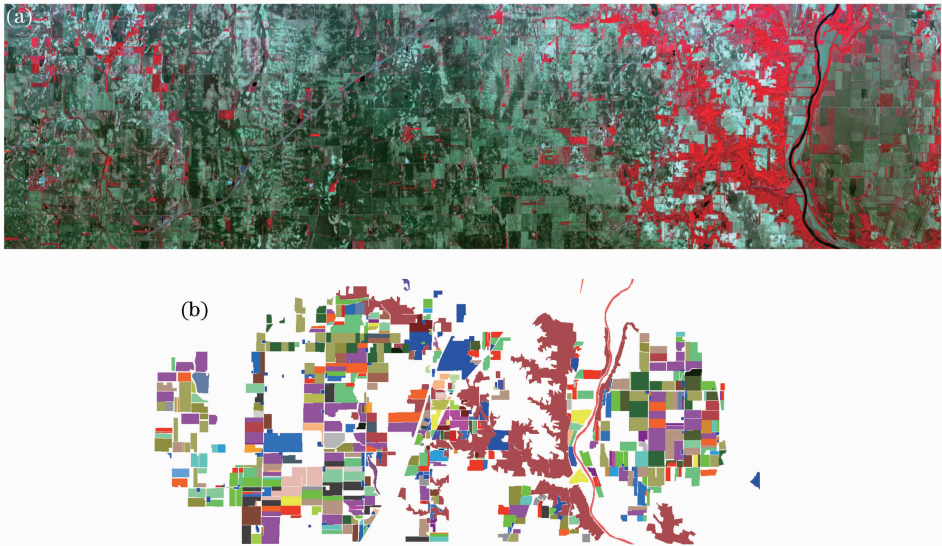


图 3 av920612\_NS\_line 高光谱遥感图像。(a)假彩色合成图像;(b)参考分类图像

Fig. 3 Map about the hyperspectral dataset av920612\_NS\_line. (a) av920612\_NS\_line displayed in simulated color; (b) reference map of the dataset

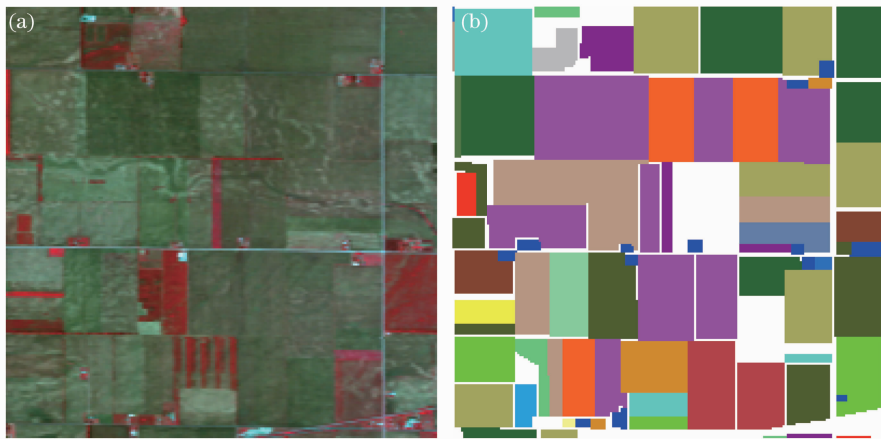


图 4 部分 av920612\_NS\_line 高光谱遥感图像。(a)假彩色合成图像;(b)参考分类图像

Fig. 4 Part of the hyperspectral dataset av920612\_NS\_line. (a) Part of av920612\_NS\_line; (b) reference map of the dataset

## 4.2 实验步骤

实验的目的是:1)验证本文算法的分类性能;2)将本文算法与当前相关文献所提出的算法进行比较。

作为对比,计算 ML 分类算法在整个波段范围的分类精度,再计算 ML 分类算法在各波段子集上的最优分类精度。

分类精度检验是分类技术中一个不可或缺的环节,精度检验一方面可以有效地对分类器进行评价从而改造分类器,另一方面也是对分类成果的最终评价,实际上对分类精度进行精确的分析是一件非常困难而又难以服众的事情。目前广泛使用的表示分类精度的方法是首先构造一个大小为  $K \times K$  混

淆矩阵或误差矩阵  $\mathbf{A}$  ( $K$  是类别数),  $a_{i,j}$  表示分类结果中第  $j$  类数据被分为第  $i$  类的的数据点数目, 然后依据混淆矩阵分别计算总体分类精度和 kappa 系数<sup>[22]</sup>。Montserud 等<sup>[23]</sup> 认为 kappa 系数的值大于 0.75 时表示分类器的分类性能良好, 而当 kappa 系数小于 0.4 时性能很差。

表 2 高光谱遥感数据的地物类别及样本数目 2

Table 2 Classes and number 2 of samples used in the experiments

No.	Landcover types	Number of samples
1	Baresoil	24
2	Buildings	559
3	Corn	382
4	Corn-NS	190
5	Corn-min till	1477
6	Corn-clean till-EW	4270
7	Corn-clean till-NS	7055
8	Corn-min till-EW	813
9	Corn-min till-NS	702
10	Corn-no till	308
11	Fescue	114
12	Hay	305
13	Not cropped	834
14	Oats	834
15	Pasture	55
16	Soybeans	196
17	Soybeans-clean till	1501
18	Soybeans-clean till-EW	4329
19	Soybeans-clean till-NS	2178
20	Soybeans-drilled	3247
21	Soybeans-min till-NS	1577
22	Soybeans-no till-NS	468
23	Wheat	2807

实验步骤:

- 1) 遥感图像预处理, 去除水吸收波段和噪声波段;
- 2) 波段分组, 确定训练基分类器的所有特征子集;
- 3) 在所有的特征子集上训练 ML 分类器, 得到基分类器;
- 4) 计算基分类器的分类精度和基分类器之间的相关性, 只保留两两负相关且分类精度大于 50% 的基分类器, 删除其他的基分类器;
- 5) 对选择的基分类器使用简单的多数投票法进行合成, 得到最终的集成分类器和分类结果;
- 6) 进行分类精度分析。

#### 4.2.1 实验数据 1

由于实验数据 1 是经过预处理的图像, 噪声波段已经去除, 因此不需要进行噪声波段检测和去除这一过程。

特征子集的生成。在波段分组步骤中, 令  $t_{hi}$  为 1, 如果  $\epsilon_{i-1,i} < 1$  并且  $\epsilon_{i,i+1} < 1$  则删除波段  $i$ , 满足这样条件的波段数为 37 个, 对剩余的 165 个波段进行分组, 对  $\epsilon$  进行降序排列, 选择前  $K$  个最大值所对应的波段作为分组的界限, 如果  $\epsilon_{i,i+1}$  满足属于前  $K$  个最大值这一要求, 则波段  $i$  和之前的波段分为同一组, 波段  $i+1$  和之后的若干波段分为另外的一组。当  $K$  确定后, 则  $K$  也是每个分组或训练基分类器的特征子集的容量(特征子集包含的波段数目), 可以得到每组中包含的波段数  $N_i$ , 则  $\sum_{i=1}^K N_i = 165$ , 则所

有可以得到的分组数目为  $\prod_{i=1}^K N_i$ , 这个数目可能会非常大, 而实际上为了取得较理想的集成效果并不需要非常多的基分类器<sup>[19]</sup>, 过多的基分类器反而会造成新的冗余和运算负担, 可以采取限制任意两个特征子集之间相同波段的数目, 得到最终的特征子集。通过实验的方法选择  $K$  的值。对原始数据求特征值, 然后进行排序, 发现前 5 个特征值占整个特征值总和的比例已经达到了 98.21%, 所以这里  $K$  取 5, 实验也证明了  $K$  取 5 时单个分类器的总体分类精度可以达到 50% 以上, 满足集成的条件。限制不同的特征子集之间的相同波段的数目为  $\lfloor 5/2 \rfloor = 2$  个, 得到的特征子集的数目为 570024。

在所有的特征子集上训练 ML 分类器, 随机选择 50% 样本作为训练样本, 剩下的 50% 的样本作为测试样本, 用来评价特征子集上训练的基分类器的分类精度, 该过程进行 10 次, 则可以得到 10 个不同的实验结果。得到的基分类器的数目为 570024。对于集成这些数目仍然显得非常多, 删除那些总体分类精度小于 55% 的分类器(实际上集成学习要求基分类器的分类精度大于 50% 即可, 但为了保证算法的稳定性和留有一定的误差空间, 这里选的精度稍微高于 50%), 对剩下的基分类器计算两两之间的相关性, 对两两负相关的分为一组。最后选择分类精度最高且两两之间负相关的一组(19 个)分类器进行集成。

#### 4.2.2 实验数据 2

实验数据 2 是包含了水吸收波段高光谱遥感数据, 需要进行水吸收波段检测和去除这一过程, 这里



水吸收波段可以通过观察其相关系数矩阵图像检测出来,实验数据 2 的相关系数矩阵如图 5 所示,纵向和横向各有两条黑色的条纹,它们相交的区域就是水吸收波段的范围,得到水吸收波段为第 104 波段到第 108 波段和第 150 波段至 163 波段,去除水吸收波段之后得到波段数为 201 个波段。其他实验步骤和实验设置与实验数据 1 完全相同。

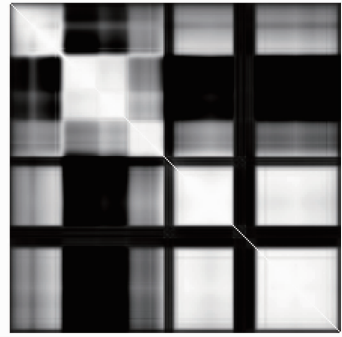


图 5 部分 av920612\_NS\_line 的相关系数矩阵  
Fig. 5 Correlation matrix of part of av920612\_NS\_line

### 4.3 结果与分析

对于实验数据 1,通过集成得到的最好结果用混淆矩阵表示,如表 3 所示。同时将分类结果以图像形式显示如图 6 所示。

表 3 混淆矩阵  
Table 3 Confusion matrix

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
C1	54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C2	0	1380	2	0	1	2	0	0	0	28	18	2	0	0	1	0
C3	0	4	827	0	0	0	0	0	0	0	0	3	0	0	0	0
C4	0	0	2	230	0	2	0	0	0	0	0	0	0	0	0	0
C5	0	0	0	0	483	4	0	0	0	0	10	0	0	0	0	0
C6	0	0	0	0	0	746	0	0	0	0	0	0	0	0	1	0
C7	0	0	0	0	0	0	26	0	0	0	0	0	0	0	0	0
C8	0	3	0	9	0	0	0	477	0	0	0	0	0	0	0	0
C9	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0
C10	0	3	0	0	2	2	0	0	0	957	4	0	0	0	0	0
C11	0	19	7	0	10	6	0	0	0	12	2401	11	0	0	2	0
C12	0	0	4	0	0	0	0	0	0	0	2	607	0	0	0	1
C13	0	0	0	0	0	0	0	0	0	0	0	0	211	1	0	0
C14	0	0	0	0	1	0	0	0	0	0	0	0	0	1274	19	0
C15	0	0	0	1	0	0	0	6	0	3	0	0	0	24	346	0
C16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	95

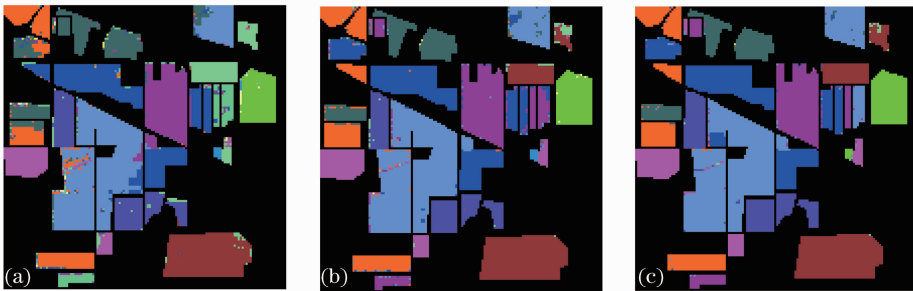


图 6 不同分类方法的分类结果。(a) ML 在特征子集上的分类结果; (b) ML 在整个波段的分类结果;  
(c) ML 集成的分类结果

Fig. 6 Classification maps with different classification methods. (a) Performance of ML classifier on the feature subset; (b) performance of ML classifier on the entire band set; (c) performance of ML ensembles

由表 3 可以计算得到总体分类精度为 97.76%, kappa 系数为 0.9745。而 ML 在所有的 202 个波段的分类方法记为 MLA, ML 在特征子集上的进行分类然后选择最好分类精度的分类方法记为 MLB,本

文的集成方法记为 MLC。不同方法的分类精度比较如表 4 所示。

对于实验数据 2,将最好的分类结果用混淆矩阵表示,并根据混淆矩阵计算总体分类精度和

kappa 系数,分别为 93.40% 和 0.9172。而 ML 在所有的 202 个波段的分类方法记为 MLA,ML 在特征子集上的进行分类然后选择最好分类精度的分类方法记为 MLB,本文的集成方法记为 MLC。不同方法的分类精度比较如表 5 所示。

表 4 不同 ML 分类方式的分类精度比较(实验数据 1)

Table 4 Comparison of different ML classification methods (dataset 1)

Sorting scheme	Overall sorting accuracy / %	Kappa coefficient
MLA	87.46	0.8489
MLB	84.67	0.8191
MLC	93.40	0.9172

表 5 不同 ML 分类方式的分类精度比较(实验数据 2)

Table 5 Comparison of different ML classification methods (dataset 2)

Sorting scheme	Overall sorting accuracy / %	Kappa coefficient
MLA	84.91	0.7952
MLB	83.08	0.7792
MLC	97.76	0.9745

由于实验数据 1 经常出现在目前的学术文章中,已有大量的国内外学者在高光谱图像处理的研究上采用了该数据,因此有利于不同的算法之间的比较,而实验数据 2 很少出现在其他文献中,由于实验条件的差别较大,为了比较结果的客观性,这里与其他分类方法进行对比时都是基于实验数据 1。Galal 等<sup>[24]</sup>使用多个光谱特征,并在整个波段和波段的某个子集分别进行了分类,得到的总体分类精度分别为 0.9020 和 0.9160。Forero 等<sup>[25]</sup>使用 AMD+PCA 的方法在 20 个波段上得到的分类精度和 kappa 系数分别为 94.64% 和 0.9253,使用的 AMD+TPCA 的方法在 15 波段上得到的分类精度和 kappa 系数分别为 94.51% 和 0.9301。Alajian 等<sup>[26]</sup>提出了基于监督-非监督的分类方法,得到了 0.9150 的总体分类精度。Landgrebe<sup>[27]</sup>提出了使用 DBFE 特征的 ECHO 分类器,可以得到 0.944 的总体分类精度。Valls 等<sup>[28]</sup>的用结合空间信息和光谱信息的分割方法得到了 91.8% 的分类精度和 0.91 的 kappa 系数。

将提出的方法应用到高光谱遥感图像上,取得了较好的分类结果。并与没有集成的 ML 分类方法做了比较,证明了集成的方法确实大大提高了分类精度和分类性能。同时本文算法与当前的其他一些高级和复杂的分类算法相比,分类精度和 kappa

系数也更高,说明算法具有较高的分类精度和更好的性能。

所提出的方法虽然较好的解决了高的数据维数和小的样本数目之间的矛盾,但也存在一些不足,主要是算法过程复杂,涉及到波段相关性信息的计算、波段分组时  $K$  值的确定、基分类器的训练、基分类器之间相关性的计算、对参与集成的基分类器选择和使用简单多数投票策略的集成,考虑到高光谱遥感图像巨大的数据量,在软件的环境下难以实现实时化。

## 5 结 论

采用集成的方法解决高光谱遥感图像的分类中高的光谱维数和小的样本数目之间的矛盾问题。根据波段之间的相似性信息对光谱波段分组,构造特征子集,然后在特征子集训练 ML 分类器,最后使用简单的多数投票法合成,得到最终的集成分类器和分类结果。结果显示可以得到比其他分类方法更好的分类结果,证明了集成方法可以应用到高光谱遥感图像的分类中,高的光谱维数反而成为一个优势,可以解决小的样本数目这一缺陷。值得注意的是集成方法并不总能保证比单个分类器更优秀的分类结果,但如果两个分类器之间的关系是负相关时几乎可以保证分类效果。本文算法难以实现对高光谱遥感数据的实时处理,对算法过程进行简化并采用优秀的近似计算是下一步将要展开的工作,此外还应开展与高光谱遥感信息流的结合研究。

## 参 考 文 献

- 1 Song Lin, Cheng Yongmei, Zhao Yongqiang. Hyper-spectrum classification based on sparse representation model and autoregressive model [J]. Acta Optica Sinica, 2012, 32(3): 0330003.  
宋琳,程咏梅,赵永强.基于稀疏表示模型和自回归模型的高光谱分类[J].光学学报,2012,32(3):0330003.
- 2 Wu Chao, Wu Yiquan. Target detection in hyperspectral image using projection pursuit based on chaotic particle swarm optimization [J]. Acta Optica Sinica, 2011, 31(12): 1211003.  
吴超,吴一全.基于混沌粒子群优化投影寻踪的高光谱图像目标检测[J].光学学报,2011,31(12):1211003.
- 3 Paul M Mather, Magaly Koch. Computer Processing of Remotely-Sensed Images; An Introduction [M]. New York: John Wiley & Sons, 2011. 229-285.
- 4 B M Shahshahani, D A Landgrebe. The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon [J]. IEEE Transactions on Geoscience and Remote Sensing, 1994, 32(5): 1087-1095.
- 5 Liu Xiaogang, Zhao Huijie, Li Na. Feature extraction based on multifraactal spectrum for hyperspectral data [J]. Acta Optica Sinica, 2009, 29(3): 844-847.  
刘小刚,赵慧洁,李娜.基于多重分形谱的高光谱数据特征提

- 取[J]. 光学学报, 2009, 29(3): 844-847.
- 6 Q Jackson, D A Landgrebe. An adaptive classifier design for high-dimensional data analysis with a limited training dataset [J]. IEEE Transactions on Geoscience and Remote Sensing, 2001, 39(12): 2664-2679.
- 7 L Zhang, B Du. Recent advances in hyperspectral image processing [J]. Geo-spatial Information Science, 2012, 15(3): 143-156.
- 8 T G Dietterich. Machine learning research: four current directions [J]. Ai Magazine, 1997, 18(4): 97-136.
- 9 Zhou Yang, Li Xiaorun, Zhao Liaoying. Modified linear-prediction based band selection for hyperspectral image [J]. Acta Optica Sinica, 2013, 33(8): 0828002.  
周 杨, 厉小润, 赵辽英. 改进的高光谱图像线性预测波段选择算法[J]. 光学学报, 2013, 33(8): 0828002.
- 10 Ge Liang, Wang Bin, Zhang Liming. Band selection based on band clustering for hyperspectral imagery [J]. Journal of Computer-Aided Design & Computer Graphics, 2012, 24(11): 1447-1454.  
葛 亮, 王 斌, 张立明. 基于波段聚类的高光谱图像波段选择[J]. 计算机辅助设计与图形学学报, 2012, 24(11): 1447-1454.
- 11 B Guo, S R Gunn, R I Damper, *et al.*. Band selection for hyperspectral image classification using mutual information [J]. Geoscience and Remote Sensing Letters, 2006, 3(4): 522-526.
- 12 H Z M Shafri, A Suhaili, S Mansor. The performance of maximum likelihood, spectral angle mapper, neural network and decision tree classifiers in hyperspectral image analysis [J]. Journal of Computer Science, 2007, 3(6): 419-423.
- 13 A Tsymbal, S Puuronen, D W Patterson. Ensemble feature selection with the simple bayesian classification [J]. Information Fusion, 2003, 4(2): 87-100.
- 14 Xie Yuancheng. Research on Classifier ensemble [D]. Nanjing: Nanjing University of Science and Technology, 2009.  
谢元澄. 分类器集成研究[D]. 南京: 南京理工大学, 2009.
- 15 Z H Zhou, J X Wu, W Tang. Ensembling neural networks: many could be better than all [J]. Artificial Intelligence, 2002, 137(1-2): 239-263.
- 16 G Brown, L I Kuncheva. "Good" and "Bad" Diversity in Majority Vote Ensembles[M]. Multiple Classifier Systems Berlin Heidelberg: Springer, 2010. 124-133.
- 17 L I Kuncheva, C J Whitaker, C A Shipp, *et al.*. Is independence good for combining classifiers? [C]. IEEE 2000. Proceedings. 15th International Conference on Pattern recognition, 2000, 2: 168-171.
- 18 L I Kuncheva, C J Whitaker, C A Shipp, *et al.*. Limits on the majority vote accuracy in classifier fusion [J]. Pattern Analysis and Applications, 2003, 6(1): 22-31.
- 19 Y Liu, X Yao. Ensemble learning via negative correlation [J]. Neural Networks, 1999, 12(10): 1399-1404.
- 20 Tang Yaohua, Gao Jinghui, Bao Qianzong. Novel selective support vector machine ensemble learning algorithm [J]. Journal of Xi'an Jiaotong University, 2008, 42(10): 1221-1225.  
唐耀华, 高静怀, 包乾宗. 一种新的选择性支持向量机集成学习算法[J]. 西安交通大学学报, 2008, 42(10): 1221-1225.
- 21 B Wu, C Chen, T M Kechadi, *et al.*. A comparative evaluation of filter-based feature selection methods for hyper-spectral band selection [J]. International Journal of Remote Sensing, 2013, 34(22): 7974-7990.
- 22 Liu Ying. The Study of Semisupervised Ensembled Support Vector Machines for Land Cover Classification [D]. Hefei: University of Chinese Academy of Sciences, 2013. 44-45.  
刘 颖. 基于半监督集成支持向量机的土地覆盖遥感分类方法研究[D]. 合肥: 中国科技大学, 2013. 44-45.
- 23 R A Montserud, R Leamans. Comparing global vegetation maps with the kappa statistic [J]. Ecological Modeling, 1992, 62(4): 275-293.
- 24 A Galal, H Hasan, I F Iman. Learnable hyperspectral measures [J]. Egyptian Informatics Journal, 2012, 13(2): 85-94.
- 25 S V Forero, J Angulo. Classification of hyperspectral images by tensor modeling and additive morphological decomposition [J]. Pattern Recognition, 2013, 46(2): 566-577.
- 26 N Alajian, Y Bazi, F Melgani, *et al.*. Fusion of supervised and unsupervised learning for improved classification of hyperspectral images [J]. Information Sciences, 2012, 217(6): 39-55.
- 27 L David. Hyperspectral image data analysis [J]. Signal Processing Magazine, 2002, 19(1): 17-28.
- 28 G Valls, D Tuia, L Bruzzone, *et al.*. Advances in hyperspectral image classification: earth monitoring with statistical learning methods [J]. Signal Processing Magazine, 2014, 31(1): 45-54.

栏目编辑: 张浩佳