基于ETAFSVM的高光谱遥感图像自动波段选择和分类

戴宏亮[1,2] 戴道清[1]

(中山大学数学与计算科学学院 广州 510275)[1] (广东商学院数学与计算科学学院 广州 510320)[2]

摘 要 提出了一种新型的具有良好特性的支持向量机——全间隔自适应模糊支持向量机(TAFSVM),并提出一种新的遗传算法——智能遗传算法(IGA)来设计一个TAFSVM分类器,称为ETAFSVM,同时优化高光谱遥感图像自动波段选择和TAFSVM参数集,并且结合5-fold交叉验证来确定其泛化能力,最后将ETAFSVM应用于高光谱遥感图像数据。通过先进行自适应波段选择后再用径向基神经网络分类器、K-最近邻分类器和标准支持向量机等3种方法进行全部分类精度比较,以及与这3种方法直接进行类别分类精度和平均分类精度比较,其结果表明运用ETAFSVM不仅可以自动进行波段选择,而且分类精度较高,对Hughes现象敏感性较低,是进行高光谱遥感图像分类的一种有效方法。

关键词 全间隔自适应模糊支持向量机,智能遗传算法,高光谱遥感图像,分类

中图法分类号 TP39 **文献标识码** A

Automatic Band Selection and Classification of Hyperspectral Remote Sensing Images Based on ETAFSVM

DAI Hong-liang[1,2] DAI Dao-qing[1]

(Department of Mathematics, Sun Yat-Sen (Zhongshan) University, Guangzhou 510275, China)[1]

(Department of Mathematics and Computational Science, Guangdong University of Business Studies, Guangzhou 510320, China)[2]

Abstract In this study, total margin-based adaptive fuzzy support vector machine(TAFSVM) which has good quality was proposed. Besides, this paper proposed an evolutionary approach to design a TAFSVM-based classifier(named ETAFSVM) by simultaneous optimization of automatic band selection and parameters tuning using an intelligent genetic algorithm(IGA), combined with 5-fold crossvalidation regarded as an estimator of generalization ability. Subsequently, the model of ETAFSVM was used to classify hyperspectral remote sensing images. Comparing with adaptive bands selecting firstly, then using radial basis functions neural network, K-nearest neighbors classifier and standard SVM to classify the test data for overall classification percentage accuracy, and then using the three classifiers to classify the test data for class and average percentage accuracy. The experimental results indicate that the proposed ETAFSVM model can achieve both higher classification accuracy and lower sensitivity to the Hughes phenomenon. Consequently, the ETAFSVM model provides a promising alternative for classification in hyspectral remote sensing images.

Keywords Total margin adaptive fuzzy support vector machine, Intelligent genetic algorithms, Hyperspectral remote sensing images, Classification

高光谱遥感又称为超谱遥感,它与多光谱遥感相比的突出特点是谱分辨力高,这对于利用遥感图像进行目标分类、识别与跟踪等具有重要的研究价值和应用意义。但是,高光谱遥感图像通常包括上百个连续分布的波段,使得传统的遥感处理技术不再适用,也就是通常所说的容易导致维数灾难,也称为"Hughes effect"[1]。高光谱遥感数据两个相邻波段之间一般相隔仅有10nm,相邻波段的图像空间、谱间相关性都非常高,并不是所有的波段都有同等的重要性。因此,通过特定的方法进行波段选择降维,形成新的高光谱图像空间,在不损失重要信息的条件下可以代表其它波段的信息。所以,进行适当的降维是必要的并且是可行的。

支持向量机(Support Vector Machine, SVM)是Vapnik等人根据统计学理论提出的一种新的机器学习方法,它是建立在统计学理论的VC维理论的结构风险最小原理基础上的,能很好地解决小样本、高维数问题[2,3]。近年来,SVM已经被广泛地应用于高光谱遥感图像分类。这些应用包括利用支持向量机进行有监督遥高光谱感图像分类[4-6]、半监督高光谱遥感图像分类[7]、利用原始支持向量机进行小样本高光谱遥感图像分类[8]等。一般情况下,进行特征选择可以有效地提升SVM的性能[9]。但是,SVM不能自动进行特征选择。在高光谱遥感图像中,SVM也就不能自动进行波段选择。因此,运用支持向量机对高光谱遥感数据进行最优分类面临着

到稿日期:2008-05-05 本文受国家自然科学基金(NSFC#60575004, NSFC#10771220),教育部高等学校博士点科研基金(SRFDP-20070558043)资助。

戴宏亮(1978-),男,博士研究生,讲师,研究方向为模式识别与知识发现、小波分析及应用,E-mail:daihongliang@tom.com;**戴道清**(1963-),男,教授,博士生导师,研究方向为复分析与小波分析、模式识别与知识发现、图像处理等。

3 个问题：(1)有效进行波段选择；(2)使用选择的波段对未知的样本进行精确和稳健的预测；(3)不使用先验知识自动进行分类器的有效设计。

为了解决上述 3 个问题，本文提出基于智能遗传算法(Intelligent Genetic Algorithms，IGA)和全间隔自适应模糊支持向量机(Total Margin-Based Adaptive Fuzzy Support Vector Machine，TAFSVM)的高光谱遥感图像分类方法。TAFSVM[10]是标准 SVM 的一种扩展，相对标准 SVM 而言，具有更加优良的特性。它可以利用惩罚的模糊性处理过拟合问题，也可以利用不同的损失函数调整由于不平衡数据集导致的最优分类面的偏斜问题。同时，还引进全间隔算法代替软间隔算法，使得全间隔自适应模糊支持向量机能够得到更低的泛化误差。它既可以应用于线性情形，也可以应用于非线性情形。但是，在运用 TAFSVM 之前，和标准 SVM 一样，其参数集需要预先确定。当 SVM 有许多参数时，与进化算法组合是可行的[11]。为了取得较好的分类性能，最好同时考虑特征选择和分类器设计[12]。但是，使用传统的遗传算法(genetic algorithm，GA)难以同时有效地优化波段选择和分类器中的参数集。本文提出智能遗传算法[13]同时优化 TAFSVM 参数集和波段选择中的大量参数，构成一种智能分类器，名为 ETAFSVM，进行高光谱遥感图像分类，不仅使得高光谱遥感图像分类能够达到较大的分类精度，而且可以有效进行高光谱遥感图像波段选择。为了验证 ETAFSVM 的有效性，运用实际高光谱遥感图像数据进行了实验。通过先进行自适应波段选择后再用径向基神经网络分类器 、K-最近邻分类器和标准支持向量机等 3 种方法进行全部分类精度比较，以及与这 3 种方法直接进行类别分类精度和平均分类精度比较，其结果表明运用 ETAFSVM 不仅可以自动进行波段选择，而且分类精度较高，对 Hughes 现象敏感性较低，是进行高光谱遥感图像分类的一种有效方法。

本文剩余部分结构如下：第 2 节简要介绍全间隔自适应模糊支持向量机的理论；第 3 节介绍智能遗传算法；第 4 节得到 ETAFSVM 模型和试验结果；最后是结论。

1 全间隔自适应模糊支持向量机

TAFSVM 相对于传统 SVM 具有 3 大优势：(1)传统 SVM 对所有样本点施加同样的惩罚，因此，传统 SVM 对样本中的离群点非常敏感。TAFSVM 通过对正负数据集施加不同的惩罚，并且每一个数据点都有相应的模糊隶属度，所以 TAFSVM 可以通过训练集的模糊性来增强泛化能力，处理过拟合问题。(2)对于不平衡训练集，TAFSVM 对正负数据采用不同的损失算法，可以提高正确分类率，因此 TAFSVM 对不平衡训练集具有自适应性。(3)TAFSVM 通过引进全间隔算法来代替软间隔算法，其不仅考虑误差，而且考虑构成最优超平面的正确分类点的信息，相对于传统 SVM，可以得到更低的泛化误差。

假设训练集为 $S=\{x_i,y_i\}_{i=1}^{L}$，在这里 $x_i\in R^n$ 是训练数据。对于分类的情形，y_i 是类标，取值 1 或者 -1；对于回归的情形，y_i 是实数。全间隔自适应模糊支持向量机可以归结为如下优化问题：

$$\text{Minimize}\ \frac{1}{2}\|\omega\|^2+C_1\sum\xi_i-C_2\sum\delta_i$$

$$y_i(\omega^T x_i+b)\geqslant 1-\xi_i+\delta_i$$

$$\xi_i\geqslant 0,\delta_i\geqslant 0,i=1,\cdots,L \tag{1}$$

其中 C_1 是对错分数据点，也就是松弛变量 ξ_i 进行加权；C_2 是对正确分类的数据点，也就是剩余变量 δ_i 进行加权。

1.1 线性不可分情形

线性不可分问题的原始问题为：

$$\text{Minimize}\ \frac{1}{2}\|\omega\|^2+C_1^+\sum_{S_f^+}\mu_i^+\xi_i^+ +C_1^-\sum_{S_f^-}\mu_i^-\xi_i^- - C_2^+\sum_{S_f^+}\mu_i^+\delta_i^+ -C_2^-\sum_{S_f^-}\mu_i^-\delta_i^- \tag{2}$$

使得

$$y_i(\omega^T x_i+b)-1+\xi_i^+-\delta_i^+\geqslant 0\quad \forall x_i\in S_f^+ \tag{3}$$

$$y_i(\omega^T x_i+b)-1+\xi_i^--\delta_i^-\geqslant 0\quad \forall x_i\in S_f^- \tag{4}$$

$$\xi_i^+\geqslant 0\quad \forall x_i\in S_f^+ \tag{5}$$

$$\xi_i^-\geqslant 0\quad \forall x_i\in S_f^- \tag{6}$$

$$\delta_i^+\geqslant 0\quad \forall x_i\in S_f^+ \tag{7}$$

$$\delta_i^-\geqslant 0\quad \forall x_i\in S_f^- \tag{8}$$

其中 C_1^+ 和 C_1^- 分别是对正负松弛变量加权，C_2^+ 和 C_2^- 分别是对正负剩余变量加权。直接解决此约束规划问题比较困难。和支持向量机类似，全间隔自适应模糊支持向量机也转化为对偶问题求解。运用拉格朗日乘数法，可以得到其对偶问题为

$$\text{Maxmize}\sum_{i=1}^{L}\alpha_i-\frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}\alpha_i\alpha_j y_i y_j x_i^T x_j \tag{9}$$

$$\mu_i^+C_2^+\leqslant\alpha_i^+\leqslant\mu_i^+C_1^+\quad \forall x_i\in S_f^+ \tag{10}$$

$$\mu_i^-C_2^-\leqslant\alpha_i^-\leqslant\mu_i^-C_1^-\quad \forall x_i\in S_f^- \tag{11}$$

其中，A_f^+ 和 A_f^- 分别是 S_f^+ 和 S_f^- 的子集：

$$S_f^+=\{x_i,y_i,\mu_i^+\}_{i=1,\cdots,L_p} \tag{12}$$

$$S_f^-=\{x_i,y_i,\mu_i^-\}_{i=1,\cdots,L_n} \tag{13}$$

$$A_f^+=\{x_i,i=1,\cdots,l_p\,|\,\mu_i^+C_2^+<\alpha_i^+<\mu_i^+C_1^+\} \tag{14}$$

$$A_f^-=\{x_i,i=1,\cdots,l_n\,|\,\mu_i^-C_2^-<\alpha_i^-<\mu_i^-C_1^-\} \tag{15}$$

其中隶属值 α_i 为拉格朗日乘子，$\mu_i^+\in[\varepsilon,1]$，$\mu_i^-\in[\varepsilon,1]$，$\varepsilon>0$，$L_p+L_n=L$。

1.2 非线性不可分情形

非线性不可分情形的对偶形式可以使用核函数 $k(x_i,x_k)=\Phi(x_i)\cdot\Phi(x_j)$。$\Phi(x):R^n\to R^m$，$m>n$ 是一个非线性映射，其目标函数为

$$\text{Maxmize}\sum_{i=1}^{L}\alpha_i-\frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}\alpha_i\alpha_j y_i y_j k(x_i,x_j) \tag{16}$$

约束条件和线性不可分情形一样。在优化问题中，KKT 条件起到很重要的作用。非线性 TAFSVM 的 KKT 条件为

$$\alpha_i^+[y_i(\omega^T\Phi(x_i)+b)-1+\xi_i^+-\delta_i^+\geqslant 0]=0\quad \forall x_i\in S_f^+ \tag{17}$$

$$\alpha_i^-[y_i(\omega^T\Phi(x_i)+b)-1+\xi_i^--\delta_i^-\geqslant 0]=0\quad \forall x_i\in S_f^- \tag{18}$$

$$\xi_i^+(\mu_i^+C_1^+-\alpha_i^+)=0\quad \forall x_i\in S_f^+ \tag{19}$$

$$\xi_i^-(\mu_i^-C_1^--\alpha_i^-)=0\quad \forall x_i\in S_f^- \tag{20}$$

$$\delta_i^+(-\mu_i^+C_2^++\alpha_i^+)=0\quad \forall x_i\in S_f^+ \tag{21}$$

$$\delta_i^-(-\mu_i^-C_2^-+\alpha_i^-)=0\quad \forall x_i\in S_f^- \tag{22}$$

b^* 的最佳值可以由下式计算：

$$b^*=\frac{\left(l_p\sum\limits_{x_i\in A_f^+}\sum\limits_{j=1}^{L}\alpha_j y_j k(x_i,x_j)+l_n\sum\limits_{x_i\in A_f^-}\sum\limits_{j=1}^{L}\alpha_j y_j k(x_i,x_j)\right)}{-l_p l_n} \tag{23}$$

对一个未知数据 x，它的决策函数为

$$D(x)=\sum_{i=1}^{L}a_i y_i k(x,x_i)+b^* \tag{24}$$

1.2.1 核函数

在机器学习理论中，流行的核函数有线性核、多项式核、径向基核等，其中径向基核已经被证明有好的泛化能力[14]。相应地，径向基核在本文中被作为核函数。径向基(RBF)核的表达式为：

$$K(x_i,x_j)=\exp(-\frac{x_i-x_j}{2\sigma^2}) \tag{25}$$

1.2.2 TAFSVM 参数

TAFSVM 对松弛变量惩罚的参数有两个 C_1^+,C_1^-，对剩余变量惩罚的参数也有两个 C_2^+,C_2^-。另外，每一个数据点都有相应的隶属度值，由文献[10]，可以按下式确定：

$$\mu_i^{+(-)}=\begin{cases}1-\dfrac{\|x_i-\bar{x}\|}{\max\limits_j\|x_j-\bar{x}\|}, & \text{如果 } \mu_i^+\geqslant\varepsilon_{1(2)}\\ \varepsilon_{1(2)}, & \text{其它}\end{cases}$$

$$x_i,x_j\in S_f^{+(-)} \tag{26}$$

其中 $\|\cdot\|$ 表示欧氏距离，$\varepsilon_1,\varepsilon_2$ 为非负实数，$\bar{x}$ 是 S_f^+ 或 S_f^- 中所有数据点的平均值。因此，TAFSVM 模型一共有 7 个参数：$C_1^+,C_1^-,C_2^+,C_2^-,\varepsilon_1^+,\varepsilon_1^-,\sigma$。

1.2.3 多类分类策略

在文献[18]中，作者比较了 3 种多类分类策略：one against one(OAO)，One against all(OAA)和 directed acyclic graph (DAG)。实验得出的结论是 OAO 和 DAG 更适合实际使用。文献[6]中从分类精度、计算时间和参数设定的稳定性等 3 个方面比较了 OAO、OAA 和两种分等级决策树、四种多类分类策略在高光谱遥感数据中的应用，得到的结论也是 OAO 比较适用。因此，本文选取 OAO 作为本文的分类策略。

2 智能遗传算法

本文提出 ETAFSVM 具有好的性能，一个主要原因就是能利用 IGA 同时进行高光谱遥感图像波段选择和 TAFSVM 参数优选。为了应用于大量参数集，IGA 采用一个有效的遗传算法染色体编码格式。然后通过分析被选择波段出现的频率，ETAFSVM 能够选择少量信息量大、与其它波段相关性小的波段，因此能够达到提高 TAFSVM 分类精度的目的。

IGA 良好的性能主要基于有效的遗传算法染色体编码和智能交叉操作。智能遗传操作建立在正交试验设计的基础上，使用“divide-and-conquer stategy”来解决具有大量参数难以处理的优化问题。而且，智能交叉采用“systematic reasoning method”而不是传统遗传算法的“generate-and-go method”来加速搜索。IGA 的详细介绍见文献[13]。

2.1 适应度函数和遗传算法染色体表示

适应度函数也称为评价函数，是根据目标函数确定的用于区分群体中个体好坏的标准，是算法演化过程的驱动力，也是进行自然选择的唯一依据。运用 IGA 设计 ETAFSVM 有两个目标：一个是使分类精度 C_a 达到最大；另外一个是使选择的波段数目 N_f 达到最少。适应度函数可以被定义为：

$$\max\ y(S)=C_a(S)-\omega_f N_f(S) \tag{27}$$

其中 S 表示需要被 IGA 优化的参数集，ω_f 是一个正的加权值。一般情况下，高的分类精度作为主要的目标。因此，ω_f 取一个较小的正数。

令 $S=\{t_i,g_i,C_1^+,C_1^-,C_2^+,C_2^-,\varepsilon_1^+,\varepsilon_1^-,\sigma\}$，遗传算法染色体编码如图 1 所示。

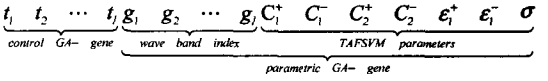

图 1 遗传算法染色体表示

图 1 中，$t_i\in\{0,1\}$ 用作选择有效的波段；$g_i\in[1,m]$ 是被选择波段的相应指标；m 是给定高光谱遥感数据的波段数；l 是提前设定的被选择波段的最大数目；

$C_1^+,C_1^-,C_2^+,C_2^-,\varepsilon_1^+,\varepsilon_1^-\sigma$ 是 TAFSVM 的参数。如果 $t_i=1$，表示具有指标 g_i 的波段被选择，否则，就表示排除在外。另外，为了减轻 IGA 的搜索负担，对 TAFSVM 参数设置整数个值构成搜索空间：

$C_1^+,C_1^-,C_2^+,C_2^-\in\{0.0001\times2^d,0.001\times2^d,0.01\times2^d,0.1\times2^d,1\times2^d,10\times2^d,100\times2^d\}$；$\varepsilon_1^+,\varepsilon_1^-,\sigma\in\{0.0001\times2^d,0.001\times2^d,0.01\times2^d,0.1\times2^d,1\times2^d\}$，$d=0,1,2,3$。

2.2 正交实验设计

正交实验设计综合正交列(orthogonal array，OA)和因子分析(factor analysis，FA)的优点。因子可以看作变量或者参数，一个集合中的一个元素可以看作因子中的一个阶层。一个完全的因子实验将考虑因子中所有的阶层组合。但是，在实际问题中，由于所有阶层组合数目太大，进行完全的因子实验是不可行的。因此，比较明智的做法是，考虑子集的阶层组合，进行部分因子实验。正交实验设计正是利用部分因子实验的良好特性来选取最优组合，达到解决问题的目的。下面对智能遗传算法中的二阶层正交阵列做一个简单的描述。令有 N 个因子，每个因子有两个阶层，一个完全的因子实验有 2^N 个组合。令 $M=2^{[\log_2(N+1)]}$，这里 $[M]$ 表示不大于 M 的最大整数。构造一个 M 行、$M-1$ 列的正交阵列 $L_M(2^{M-1})$，只用考虑前面的 N 列。正交阵列能大大减少因子分析中阶层组合的数目。正交阵列中需要分析的所有因子组合数目仅仅是 $M=O(N)$，在这里 $N+1\leqslant M\leqslant 2N$。再令 y_t 表示组合 t 的适应度函数值，$t=1,\cdots,M$。定义因子 i 阶层 k 的主要效应为 S_{ik}，$i=1,\cdots,N$

$$S_{ik}=\sum_{t=1}^{M}y_t W_t \tag{28}$$

在这里，如果因子 i 的阶层的组合 t 的值是 k，则 $W_t=1$，否则 $W_t=0$。当 $S_{i1}>S_{i2}$ 时，阶层 1 更好；当 $S_{i1}<S_{i2}$ 时，阶层 2 更好；当 $S_{i1}=S_{i2}$，阶层 1 和 2 有同样的贡献。最有效的因子 i 具有最大的主要效应差 $MED_i=|S_{i1}-S_{i2}|$。在每个因子两个阶层中更好的一个被确定好之后，即可以得到一个更有效的组合。

2.3 智能交叉操作

类似传统 GA，在一次交叉操作中，两个父代个体 P_1 和 P_2 产生两个子代个体 C_1 和 C_2。下面给出其具体操作步骤。

Step 1 使用一个正交阵列 $L_M(2^{M-1})$ 的前 N 列；

Step 2 令第 i 个因子的阶层 1 和阶层 2 分别表示父代个体 P_1 和 P_2 产生的第 i 个参数种群；

Step 3 计算第 t 次试验的适应度值 y_t，$t=2,\cdots,M$。y_1 是 P_1 的适应度值；

Step 4 计算主要效应差 S_{ik}，$i=1,\cdots,N,k=1,2$；

Step 5　决定每个因子中更好的一个阶层；

Step 6　从导出的相应父代个体中组合更好的 GA 基因，构成子代个体 C_1 的 GA 染色体；

Step 7　子代个体 C_2 的 GA 染色体构成和 C_1 类似，但是有一点不同的是具有最小主要效应差的因子选自其它阶层；

Step 8　在 P_1，P_2，C_1，C_2 和 $M-1$ 个正交阵列组合中的最好的两个个体作为最后的优秀子代个体 C_1 和 C_2。

一个智能交叉操作需要通过计算 $M+1$ 个适应度值来搜索 2^N 个组合，$N+1 \leqslant M \leqslant 2N$。

2.4　智能遗传算法

由上面的介绍可以得到具体的 IGA 算法如下：

Step 1　初始化。随机生成一个初始种群，该种群具有 N_{pop} 个可行个体，并且在一个遗传算法染色体中每一个基因 g_i 是唯一的。

Step 2　评价。评价种群中所有个体的适应度值。令 I_{best} 为种群中最优的个体。

Step 3　使用简单的截断选择，用最好的 $P_s \cdot N_{pop}$ 个个体来替换最坏的 $P_s \cdot N_{pop}$ 个个体，在这里 P_s 是一个选择概率。

Step 4　随机选择 $P_c \cdot N_{pop}$ 个个体(包括 I_{best})进行交叉操作，在这里 P_c 是一个交叉概率。

Step 5　使用变异概率 P_m 对种群进行变异操作。为了防止最优个体恶变，不对最优个体应用变异操作。

Step 6　终止测试。如果提前设定的终止条件满足，停止算法，否则转到步骤 2。

最后通过对遗传算法的最优个体解码，可以得到选择的波段子集和相应的 TAFSVM 参数集。

3　ETAFSVM 模型

图 2 是 ETAFSVM 流程图。传统 SVM 的参数必须提前确定，并且输出是一个 SVM 分类器。与传统 SVM 相比，ETAFSVM 通过 IGA 自动调节 TAFSVM 参数，最终输出是 TAFSVM 分类器和具有较大信息量，并且与其它波段相关性小的波段集。运用 IGA 优选 TAFSVM 参数和选择高光谱遥感波段集。为了避免过拟合，这里采用 5-fold 交叉验证技术[14]。ETAFSVM 具体步骤如下：

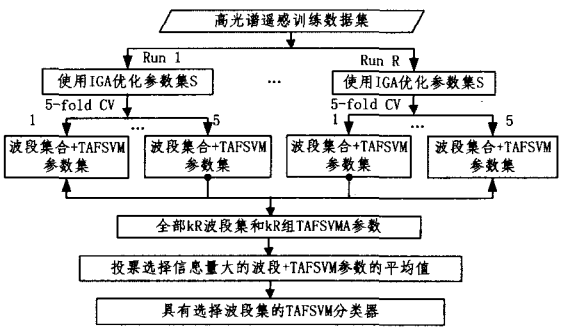

图 2　ETAFSVM 流程图

Step 1　进行 R 个独立运行。

Step 2　运用 IGA 调节 TAFSVM 参数，并且选择信息量丰富、与其它波段相关性小的波段集。结果得到 5 个波段集和 5 个 TAFSVM 参数集。

Step 3　根据波段集的出现频率，从 $5R$ 个波段集中选择 N_{top} 个信息量丰富、与其它波段相关性小的波段集。一般情况，$N_{top}=\lceil H \rceil$，H 是波段集的平均大小。

Step 4　对得到的 TAFSVM 参数集进行平均，把平均后的结果作为 TAFSVM 的最终参数。

Step 5　输出被选择的 N_{top} 个波段和具有参数平均值的 TAFSVM 分类器。

4　高光谱遥感数据分类

4.1　数据集

为了验证提出的 ETAFSVM 模型，本文以一个实际高光谱基准数据集作为训练样本和测试样本。该数据集表示 1992 年 6 月通过 AVIRIS 传感器拍摄的印第安纳州西北部印第安遥感试验区的一部分[15]。去掉噪声和水汽吸收较明显的谱段，选取 4757 个训练样本和 4588 个测试样本，如表 1 所列。

表 1　试验中训练样本数和测试样本数

类别	训练样本数	测试样本数
ω_1	742	692
ω_2	442	392
ω_3	260	237
ω_4	389	358
ω_5	236	253
ω_6	487	481
ω_7	1245	1223
ω_8	305	309
ω_9	651	643
总计	4757	4588

4.2　运用 ETAFSVM 模型的方法

IGA 参数设置如下：个体数目 $N_{pop}=40$，选择概率 $P_s=0.3$，交叉概率 $P_c=0.4$，变异概率 $P_m=0.1$，进化代数为 500。参数集 $S=\{t_i, g_i, C_1^+, C_1^-, C_2^+, C_2^-, \varepsilon_1^+, \varepsilon_1^-, \sigma | i=1,\cdots,l\}$ 中 $l=80$。适应度函数中加权值 $\omega_f=0.002$。$R=20$ 个独立运行被执行。

4.3　使用 ETAFSVM 进行波段选择及分析

为了评价本文提出的 ETAFSVM 模型的有效性，运用传统的自适应波段选择[16](Adaptive band selection，ABS)方法，再结合传统的径向基神经网络分类器[17](RBF)、K-最近邻分类器(K-nn)和支持向量机(SVM)方法对上述高光谱数据测试集先降维，然后分类，4 种方法进行比较，结果如图 3 和表 2 所列。图 3 横轴表示波段数目，从 20 到 200，步长为 10；纵轴表示分类精度。

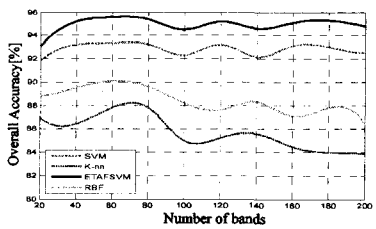

图 3　4 种分类器(ETAFSVM，RBF，K-nn 和 SVM)对高光谱遥感数据测试集先降维后分类性能对比

表 2　4 种不同方法全部平均分类精度和方差对比

方法	全部精度(%)	
	平均	方差
ETAFSVM	94.87	0.29
ABS+SVM	92.78	0.58
ABS+RBF	88.35	0.94
ABS+K-nn	85.72	2.27

由图 3 和表 2 可知，本文提出的 ETAFSVM 模型能够很好地进行波段选择，降维后的分类的性能明显高于另外 3 种方法。并且由表 2 的方差结果可以看出本文提出的 ETAFSVM 模型对 Hughes 现象敏感性最弱。表 3 是运用 ETAFSVM 方法根据波段的出现频率选择的 15 个信息量丰富并且和其它波段相关性小的波段。其中，指数由大到小排列。

表 3　指数及其对应的波段号(前 15 个)

序号	指数	波段号	序号	指数	波段号
1	1012.5	23	8	903.21	26
2	996.04	22	9	886.95	35
3	932.56	21	10	875.92	18
4	932.55	20	11	874.33	27
5	911.09	19	12	865.39	17
6	910.47	24	13	860.12	25
7	909.66	36	14	855.93	37

4.4　ETAFSVM 与其它常用分类器分类性能比较

运用 ETAFSVM 模型进行高光谱遥感数据分类，运用 IGA 得到 TAFSVM 参数的最优结果为：$C_1^+=800$，$C_1^-=40$，$C_2^+=200$，$C_2^-=10$，$\varepsilon_1^+=\varepsilon_1^-=0.4$，$\sigma=0.2$。另外，4 种不同分类器在高光谱遥感数据测试样本上的类别分类精度、平均分类精度如表 4 所列。

表 4　4 种不同分类器在测试样本上的类别、平均(AA)分类精度对比

方法	分类精度									
	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	AA
E	98.27	90.25	96.39	94.71	98.75	99.72	93.18	86.97	90.28	94.28
K	96.58	71.31	85.27	82.82	94.53	97.63	91.49	75.43	76.35	85.71
R	95.42	78.95	89.64	80.27	98.41	97.83	91.75	74.04	84.97	87.92
S	84.63	89.85	93.73	96.65	99.84	89.23	74.39	91.88	98.64	90.98

从表 4 的结果可以看出，本文提出的模型 ETAFSVM 无论是类别分类精度还是平均分类精度和全部分类精度都要好于另外 3 个常用分类器。

结束语　(1)本文提出的模型 ETAFSVM 运用于高光谱遥感图像，不仅可以进行分类，而且可以进行自动波段选择。与其它常用分类器相比，分类精度明显提高，是进行高光谱遥感图像分类的一种有效方法。(2)TAFSVM 是 SVM 的一种良好扩展，具备 SVM 所有优点。另外，TAFSVM 还可以通过训练集的模糊性来增强泛化能力，处理过拟合问题；对不平衡训练集具有自适应性，对正负数据采用不同的损失算法，这样也可以提高正确分类率；通过引进全间隔算法来代替软间隔算法，可以得到更低的泛化误差。TAFSVM 的良好特性符合高光谱遥感图像的内在规律。(3)TAFSVM 参数较多，根据经验或者运用交叉验证难度都比较大，但是运用智能遗传算法对参数进行 TAFSVM 参数优选，从结果看，该方法非常成功，可作为 TAFSVM 参数选择的一种有效方法。

参 考 文 献

[1] Hughes G F. On the mean accuracy of statistical pattern recognizers. IEEE Trans. Inf. Theory, 1968, IT-14 (1): 55-63

[2] Vapnik V N. Statistical Learning theory. NY: Wiley, 1998

[3] Vapnik V N. The Nature of Statistical Learning Theory. New York: Springer-Verlag, 1995

[4] Bazi Y, Melgani F. Toward an optimal SVM classification system for hyperspectral remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44 (11): 3374-3385

[5] Munoz M J, Bruzzone L, Camps V G. A support vector domain description approach to supervised classification of remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 2007, 45(8): 2683-2692

[6] Melgani F, Bruzzone L. Classificaiton of hyperspectral remote sensing images with support vector machines. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(8): 1778-1790

[7] Bruzzone L, Chi M, Marconcini M. Novel transductive SVM for semisupervised classification of remote sensing images

[8] Chi M, Feng R, Bruzzone L. Classification of hyperspectral remote sensing data with primal SVM for small-sized training dataset problem. Advances in Space Research, 2008, doi: 10.1016/j.asr.2008: 2-12

[9] Statnikov A, Aliferis C F, Tsamardinos I, et al. Acomprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. Bioinformatics, 2005, 21(5): 631-643

[10] Liu Y H, Chen Y T. Face Recognition Using Total Margin-Based Adaptive Fuzzy Support Vector Machines. IEEE Transactions on Neural Networks, 2007, 1(18): 178-192

[11] Mierswa I. Evolutionary learning with kernels: a generic solution for large margin problems // Proceedings of the GECCO'06. Washington, USA, 2006: 1553-1560

[12] Ho S Y, Liu C C, Liu S. Design of an optimal nearest neighbor classifier using an intelligent genetic algorithm. Pattern Recognition Letter, 2002, 23: 1459-1503

[13] Ho S Y, Shu L S, Chen J H. Intelligent evolutionary algorithms for parameter optimization problems. IEEE Trans. Evolutionary Comput, 2004, 8(6): 522-541

[14] Campbell C. Kernel methods: a survey of current techniques. Neurocomputing, 2002(48): 1-4, 63-84

[15] Aviris N W. Indiana's Indian Pines 1992 Data Set [Online]. ftp://ftp.ecn.purdue.edu/bieh l/MultiSpec /92AV3C.lan (original files), ftp://ftp.ecn.purdue.edu/biehl/ PC_MultiSpec/ThyFiles.zip (ground truth)

[16] 刘春红，赵春晖，张凌雁. 一种新的高光谱遥感图像降维方法[J]. 中国图像图形学报，2005，10(2)：218-222

[17] Bruzzone L, Prieto D F. A technique for the selection of kernel-function parameters in RBF neural networks for classification of remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 1999, 37 (2): 551-559

[18] Hsu C W, Lin C J. A Comparison of Methods for Multiclass Support Vector Machines. IEEE Transaction on Neural Networks, 2002, 13(2): 415-425