

高光谱图像分类的全面加权方法研究

王立国, 赵春晖, 乔玉龙, 陈万海

(哈尔滨工程大学信息与通信工程学院, 黑龙江 哈尔滨 150001)

摘要:像元分类是高光谱数据分析的最基本、最重要内容之一,而基于支持向量机(SVM)的分类方法以其高效性得以广泛使用。原始的SVM分类模型中并没有体现出样本、特征、类别对于分类或分析的不同重要性,从而影响了处理效果。为此,将各样本偏离其类中心的距离映射为样本加权系数;将类内散度矩阵应用于特征加权方法;将SVM方程系统中的单位矩阵对角元素加以调整来完成类别加权。不同加权方法既可以单独使用也可以联合使用。实验表明,所提出的加权方法有助于进一步提高高光谱图像的分类效果。

关键词:高光谱图像;分类;支持向量机;加权

中图分类号:TP75 **文献标识码:**A

RESEARCH ON ALL-AROUND WEIGHTING METHODS OF HYPERSPPECTRAL IMAGERY CLASSIFICATION

WANG Li-Guo, ZHAO Chun-Hui, QIAO Yu-Long, CHEN Wan-Hai

(College of Information and Communications Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: Pixel classification is one of the most basic and important contents of hyperspectral imagery (HSI) analysis, and support vector machine (SVM) based method is very popular in HSI classification for its high efficiency. The importance of samples, features, and classes, however, is not reflected in original SVM based classification model, and the classification effect is deteriorated consequently. In this study, the distance of each sample deviating from its class-center was mapped into the sample as weighting coefficient. And within-class scatter matrix was introduced into the feature weighting measure, and the diagonal elements in SVM equation system were adjusted for the purpose of class weighting. The weighted methods can be used solely or jointly. Experiments show that the proposed weighting methods are helpful to improve the effect of HSI classification.

Key words: hyperspectral imagery (HSI); classification; support vector machine (SVM); weighting

引言

随着遥感技术的发展,高光谱图像(HSI)得到了越来越广泛的应用^[1-3]。像元分类是高光谱数据分析的最基本、最重要内容之一。在众多的光谱图像分类方法中,支持向量机(SVM)^[4]以其优良的性能得到广泛好评和使用。SVM在发展过程中出现了许多结构类型。在这些结构类型中,最小二乘SVM(LS-SVM)^[5]受到了普遍欢迎,其主要原因在于,LS-SVM的代价函数是一个仅带有等式约束的优化问题,其求解可在线性系统中进行。虽然SVM在高光谱图像分类中表现出良好的性能,但如何进一步提高其分类性能仍然是一项值得研究的内容。在高光谱图像分类过程中,SVM的泛化性能对

于训练过程中的野值点和噪声干扰像元(统称为异常像元)较为敏感,而它们又常常不可避免地广泛存在于高光谱数据之中,影响了模型的准确性。SVM的建模方法过于依赖训练样本,对异常像元的存在很敏感,通常少量异常像元的引入就可能完全破坏模型的泛化性能。

文献[6]提出LS-SVM的加权方法,使得高光谱图像中受到噪声干扰严重的像元和野值点得到有效控制,从而获得了更加良好的鲁棒特性和推广能力。这种加权的方法包含一次完整的预备训练。而一次训练所需要的计算量一般较大,尤其是当训练样本较多时,该方法将变得极为耗时。由于这一原因,该方法并没有得到有效推广。

现有的高光谱图像分类加权方法一般都是针对

训练样本实施的,而对于如下两种情况却少有文献考虑。其一,高光谱图像不同的特征(或波段,谱段)对于类别可分性的影响是不同的,即它们对分类的作用是不同的,因此在分类器设计中不应等同对待;其二,在实际应用中,遥感数据类别众多,而不同类别对于高光谱数据分析的意义往往不同,或者说研究者对于他们所感兴趣的程度不同,因此也同样需要在分类器设计中加以考虑。为此,本文研究基于LS-SVM理论的高光谱图像分类问题中的多重加权方法,以进一步提高分类分析效果。

1 加权算法描述

本文的目的在于提供一种对像元、特征和类别进行多重加权处理的基于LS-SVM的高光谱图像复选性加权分类方法,为此需使得在建立分类模型时:异常程度不同的训练样本得到不同的控制;类别可分性不同的特征得到不同的体现;重要性不同的数据类别得到不同的对待,最终实现更好的数据分析效果。

1.1 样本加权方法

LS-SVM的数学模型是一个仅带有等式约束的误差代价函数平方和的优化问题,其求解可在线性系统中进行。其优化问题表达式为

$$\begin{aligned} \min_{w,b,e} J(w,e) &= \frac{1}{2} \|w\|^2 + \frac{\gamma}{2} \sum_{i=1}^n e_i^2 \\ \text{s.t. } y_i &= \langle w, \phi(x_i) \rangle + b + e_i, \\ i &= 1, 2, \Lambda, n, \quad \gamma > 0, \end{aligned} \quad (1)$$

其中, $x_i \in R^d$ 为样本数据, $y_i \in \{+1, -1\}$, e_i 分别为相应的类别标号和判别误差, $i = 1, \Lambda, n$ 。为了将异常程度不同的样本在分类模型中加以体现,需要将他们相应的分类误差在代价函数分配不同的权值,即获得LS-SVM的加权性训练模型。设 e_i 对应于权值 v_i , 式(1)变为

$$\begin{aligned} \min_{w,b,e} J(w,e) &= \frac{1}{2} \|w\|^2 + \frac{\gamma}{2} \sum_{i=1}^n (v_i e_i)^2 \\ \text{s.t. } y_i &= \langle w, \phi(x_i) \rangle + b + e_i, \\ i &= 1, 2, \dots, n, \quad \gamma > 0. \end{aligned} \quad (2)$$

如何合理地确定权值 v_i 便成为样本加权中的关键问题。由于训练样本中异常样本到其相应的类中心相对距离较远,因此可以通过距离尺度来度量其异常程度^[7]。这样,可以为异常程度较大的样本分配较小的权值来弱化其不良影响。另一方面,由于类内光谱的差异性,即使是纯样本也不可能集中在相应的类中心,而是存在一个相对较小的偏离。考虑

到这一点,在计算距离时,我们可以将前面所求得距离减去一个修正常数。为此,可首先确定以类中心为圆心,包含该类别规定比例样本点的最小半径,进而将该半径设为上述修正常数。

设样本 x_i 所对应的类中心为 x_{y_i} , 而以 x_{y_i} 为圆心半径为 r 的圆为包含该类指定比例样本的最小圆。用 $\hat{D}(x_i, x_{y_i})$ 表示样本 x_i 到 x_{y_i} 未经修正的距离, 则 $\hat{D}(x_i, x_{y_i})$ 的计算公式如下

$$\begin{aligned} \hat{D}(x_i, x_{y_i}) &= \|\phi(x_i) - \phi(x_{y_i})\| \\ &= (K(x_i, x_i) + K(x_{y_i}, x_{y_i}) - 2K(x_i, x_{y_i}))^{1/2}, \end{aligned} \quad (3)$$

从而可以规定 x_i 到其类中心 x_{y_i} 的修正距离 $D(x_i, x_{y_i})$ 为

$$D(x_i, x_{y_i}) = \hat{D}(x_i, x_{y_i}) - r, i = 1, 2, \dots, n. \quad (4)$$

记

$$\begin{aligned} D_{\max} &= \max_i (D(x_i, x_{y_i})) \\ D_{\min} &= \min_i (D(x_i, x_{y_i})) \end{aligned}, \quad (5)$$

并用 $ND(x_i, x_{y_i})$ 表示 $D(x_i, x_{y_i})$ 的正规化形式, 即

$$ND(x_i, x_{y_i}) = D(x_i, x_{y_i}) / D_{\max}, i = 1, 2, \dots, n. \quad (6)$$

则权值因子可以通过下面的公式来求得

$$\begin{aligned} v_i &= 1 - ND(x_i, x_{y_i})^2 + (D_{\min} / D_{\max})^2, \\ i &= 1, 2, \dots, n. \end{aligned} \quad (7)$$

容易验证 $0 < v_i \leq 1$ 。将式(1)中的误差项 $\{e_i\}_{i=1}^n$ 替换为其加权形式 $\{v_i e_i\}_{i=1}^n$ 便得到形如式(2)的样本加权型LS-SVM。

1.2 特征加权方法

特征加权关键是要找到一个合适的加权矩阵, 该矩阵可以加强有效的特征, 削弱类别可分性较差的特征。Fisher线性判别分析是一种广泛使用的分类技术, 在模式识别中得到了广泛的应用。其中的类内散度矩阵的逆矩阵可以很好地体现不同特征对于分类效果的不同贡献^[8], 这一效果已在光谱分离中得以验证^[9]。因此, 可将其应用于高光谱图像分类的特征加权之中, 具体方法如下。

设有 n 个训练样本向量用来分类, μ_j 是第 j 类样本的平均值 ($j = 1, 2, \dots, n$), 即

$$\mu_j = \frac{1}{n_j} \sum_{r_i \in C_j} r_i, \quad (8)$$

c_j, n_j 分别代表第 j 类样本集合及其样本数目, 据此可以定义类内散度矩阵 S_w 如下

$$S_w = \sum_{j=1}^p S_j, \quad (9)$$

这里

$$S_j = \sum_{r_i \in C_j} (r - \mu_j)(r - \mu_j)^T, \quad (10)$$

S_w 为实对称矩阵,从而存在正交矩阵 U 将其对角化为矩阵 B

$$U^T S_w U = B, \quad (11)$$

进一步可推知

$$S_w^{-1} = (UBU^T)^{-1} = (UB^{-1/2})(UB^{-1/2})^T, \quad (12)$$

记 $G = (UB^{-1/2})$, 则 G 可以用作分类问题中的特征加权矩阵.

1.3 类别加权方法

LS-SVM 优化表达式相应的对偶形式为

$$\begin{aligned} \min_{w,b,e,\alpha} L(w,b,e,\alpha) \\ = J(w,e) - \sum_{i=1}^n \alpha_i | \langle w, \phi(x_i) \rangle + b + e_i - y_i |, \end{aligned} \quad (13)$$

其最优 KKT 条件为

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^n \alpha_i \phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^n \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0 \rightarrow \alpha_i = \gamma e_i \\ \frac{\partial L}{\partial \alpha_i} = 0 \rightarrow \langle w, \phi(x_i) \rangle + b + e_i - y_i = 0 \\ i = 1, 2, \dots, n \end{cases} \quad (14)$$

利用消元法消去 w 和 e 后上式可进一步表示为

$$\begin{bmatrix} 0 & 1_v^T \\ 1_v & K + I/\gamma \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}, \quad (15)$$

其中, $y = [y_1, y_2, \dots, y_n]^T$, $1_v = [1, 1, \dots, 1]^T$, $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$, K 为训练样本的核函数矩阵, I 是一个 $n \times n$ 的单位矩阵, n 是训练样本的个数. 当 I 是单位矩阵时, 表示训练过程对每一训练样本等同考虑.

根据文献[5], 如果将不同类别的训练样本在优化模型中加以区别对待, 那么这将直接反映为在相应的线性方程组(15)中 I 的对角元素赋值的不同. 也就

是说, I 的对角元素值能够体现对各个训练样本的重视程度. I 的某一项权值相对越大, 则表示训练过程对相应的样本越不重视, 反之亦然. 类别加权指的就是通过重置 I 中某些类别样本的对应位置的对角元素值, 而不再是原始的等值设置, 以达到改变对各个类别的重视程度, 从而保护重要类别, 抑制非重要类别的目的. 因此, 把重要类别的训练样本对应的权值适当减小, 而把非重要类别的训练样本对应的权值适当增大, 能有效地提高重要类别的分类精度.

以上 3 种加权方法可以单独使用, 也可以以任何复选方式组合使用. 图 1 给出了高光谱图像复选性加权分类的操作界面示意图.

2 仿真实验

按照上述加权方法, 本文利用真实遥感图像进行性能评价. 遥感图像取自 1992 年 6 月拍摄的美国印第安纳州印第安农林高光谱遥感实验区的一部分, 图像大小为 144×144 , 220 特征. 原始监督数据将该图像界定为除背景外共 16 种地物类别.

第一组实验样本由图像中第 3、8 和 11 共三类地物(像元数目依次为 834、489 和 2468)的数据组合而成. 抽取部分像元的光谱特征作训练样本, 整类数据作为测试样本. 依次采用未加权、样本加权、特征加权、类别加权方式以及三重加权方式进行效果测试, 分类结果如图 2 及表 1 所示. 实验中, SVM 采用高斯核函数, 训练样本取自各个类别的前 100 个像元. 在类别加权实验中, 3 个类别的权值依次设置为 1、5 和 10, 即重点考虑类别 3 的分类效果. 在分类结果中, 以上 3 个类别依次被标记为不同颜色. 图像中分类错误的像元用白点显示出来. 表 1 中, 分类软误差指的是 SVM 判决结果未经二值量化前与监督分类数值之间的绝对误差均值统计, 这种方式较之判决结果量化后获取的硬分类精度更精确. 第二组实验选择第 2、6 和 10 共 3 类地物, 像元数目依次为 1434、747 和 968. 实验方式与第一组实验相同, 分类结果如图 3 及表 2 所示. 两组实验结果一致表明, 使用样本加权和特征加权的方法均能不同程度地降低整体分类误差, 如第一组实验中特征加权方法将未加权方法中 3 个类别的分类软误差分别降至原来的 85.3%、60.0% 和 71.6%. 在类别加权方法中可以看到被重点考虑的第 2、3 类的分类效果均好于无类别加权的同类分类效果, 尤其明显的是第 2 类的类别加权分类软误差低至其他加权分类方法的分类软误差的 80% 以下. 对于三重加权分类方法,

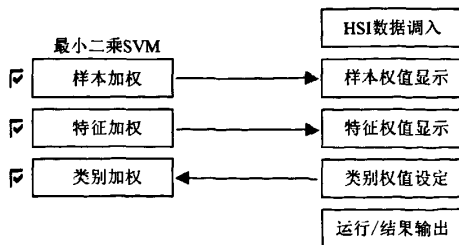


图 1 高光谱图像复选性加权分类的操作界面示意图
Fig. 1 The sketch map of operating interface of check weighting classification of HSI

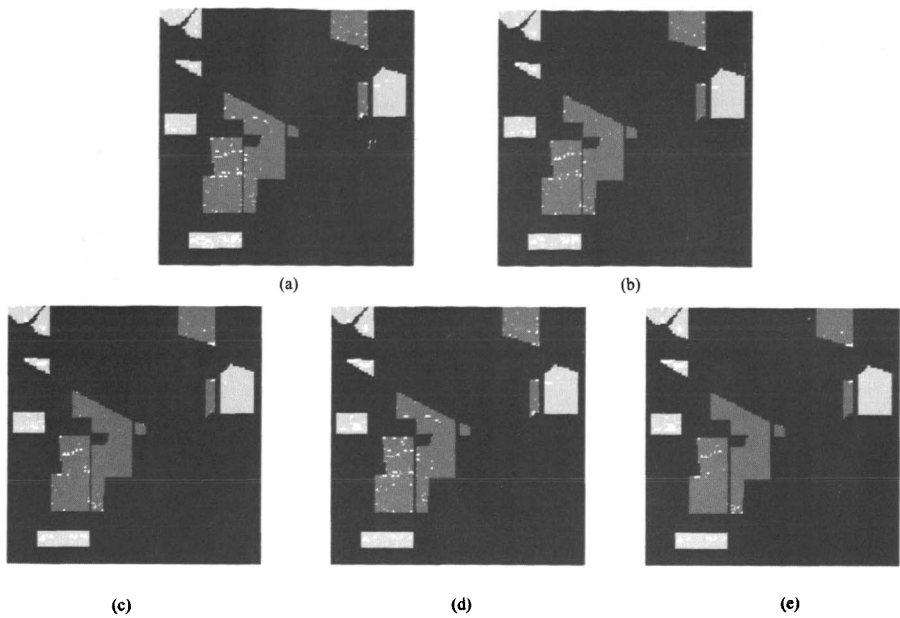


图 2 第一组分类实验中不同加权条件下的分类结果图 (a)未加权情形 (b)样本加权情形 (c)特征加权情形 (d)类别加权情形 (e)三重加权情形

Fig. 2 Classification result images under different weighting conditions in experiment group 1 (a)unweighted case (b)sample weighted case (c)feature weighted case (d)class weighted case (e)trily weighting case

万方数据

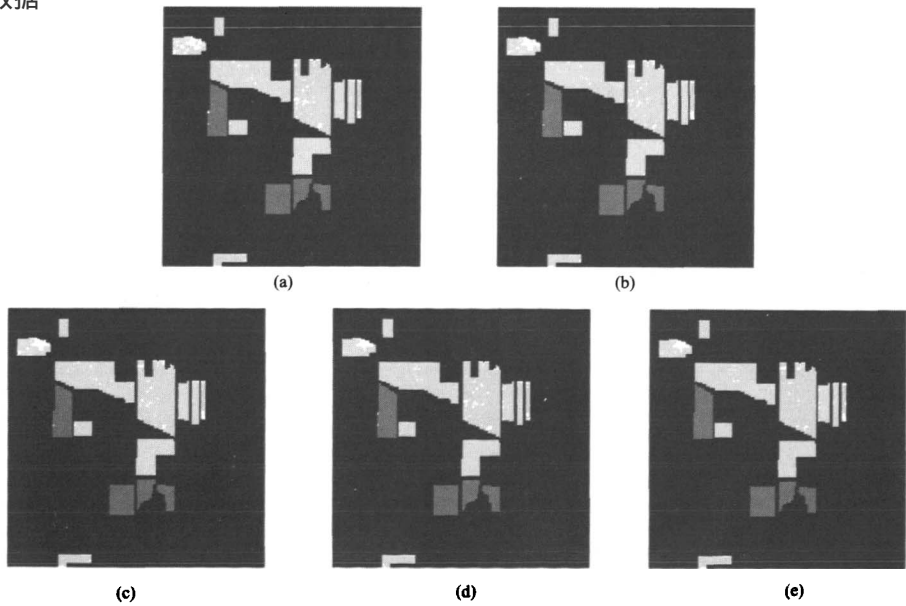


图 3 第二组分类实验中不同加权条件下的分类结果图 (a)未加权情形 (b)样本加权情形 (c)特征加权情形 (d)类别加权情形 (e)三重加权情形

Fig. 3 Classification result images under different weighting conditions in experiment group 2 (a)unweighted case (b)sample weighted case (c)feature weighted case (d)class weighted case (e)trily weighting case

虽然第 3 类和第 10 类的分类效果并未达到所有加权分类中的最优值,但从总体上统计,5 种分类方式

下错分像元总数依次为 513,428,394,460 和 352,显然三重加权的分析效果更佳.

表 1 第一组分类实验中的错分像元数目及分类软误差
Table 1 The number of misclassified pixels and the soft errors of classification in experiment group 1

地物类别	未加权	样本加权	特征加权	类别加权	三重加权
第 3 类	165/0.224	145/0.199	138/0.191	130/0.188	133/0.189
第 8 类	7/0.005	6/0.003	2/0.003	7/0.007	1/0.002
第 11 类	136/0.067	87/0.050	70/0.048	147/0.071	54/0.042

表 2 第二组分类实验中的错分像元数目及分类软误差
Table 2 The number of misclassified pixels and the soft errors of classification in experiment group 2

地物类别	未加权	样本加权	特征加权	类别加权	三重加权
第 3 类	165/0.224	145/0.199	138/0.191	130/0.188	133/0.189
第 8 类	7/0.005	6/0.003	2/0.003	7/0.007	1/0.002
第 11 类	136/0.067	87/0.050	70/0.048	147/0.071	54/0.042

3 结论

本文根据样本异常程度与样本偏离类中心距离之间的关系,将距离非线性映射为相应权值来完成样本加权;根据类内散度矩阵对线性光谱分离问题的加权特性,将其推广到 LS-SVM 分类问题中来完成特征加权;根据 LS-SVM 线性方程组中单位矩阵对角元素的特殊含义,将其设定为体现类别重要性的不同数值来完成类别加权. 在 3 种加权方法中,样本加权是专们对训练样本所施加的手段,特征加权是对全部数据进行的操作,而类别加权是在训练过程中对矩阵对角元素的重置. 3 种加权方法既可以单独使用,也可以以任何复选方式组合使用,在实际应用中可根据具体需求进行选择. 进一步的工作将对类别加权值的自适应确定、加权方法在其它分类问题和分类模型中的有效推广等内容加以研究.

REFERENCES

[1] LUO Xin, GUO Lei, LIU Zhen. Lossless compression of hyperspectral imagery by IPCT and 3-D TARP coder[J]. *J. Infrared Millim. Waves* (罗欣,郭雷,刘震. 基于 IPCT/3-D Tarp 的高光谱图像无损压缩. *红外与毫米波学报*), 2007, 26(6): 443—446.

[2] TAN Ke-Long, ZHOU Ri-Ping, WAN Yu-Qing, et al. Remote sensing monitoring method of hyperspectral and high-resolution for underground coal bed combustion[J]. *J. Infrared Millim. Waves* (谭克龙,周日平,万余庆,等. 地下煤层燃烧的高光谱及高分辨率遥感监测方法. *红外与毫米波学报*), 2007, 26(5): 349—352.

[3] YAN Fu-Li, WANG Shi-Xin, ZHOU Yi, et al. Monitoring the water quality of Tai-Hu Lake by using hyperion hyperspectral data[J]. *J. Infrared Millim. Waves* (阎福礼,王世新,周艺,等. 利用 Hyperion 星载高光谱传感器监测太湖水质的研究. *红外与毫米波学报*), 2006, 25(6): 460—464.

[4] Evgeniou T, Pontil M. *Support Vector Machines: Theory and Applications* [M]. Lecture Notes in Computer Science, Advanced Lectures. 2001, 2049, 249—257.

[5] Wu Hsu-Kun, Chen Pao-Jung, Hsieh Jer-Guang. Simple algorithms for least square support vector machines [C]. IEEE International Conference on Systems, Man and Cybernetics, 2006, 6: 5106—5111.

[6] Suykens J A K, Brabanter J D, Lukas L, et al. Weighted least squares support vector machines: robustness and sparse approximation[J]. *Neurocomputing*, 2002, 85—105.

[7] Song Q, Hu W J, Xie W F. Robust support vector machine with bullet hole image classification [J]. *IEEE Trans. on Systems, Man and Cybernetics, Part C*. 32, 2002, 440—448.

[8] Ji B, Chang C I, Jensen J O, et al. Unsupervised constrained linear Fisher's discriminant analysis for hyperspectral image classification[C]. In 49th Annu. Meeting, SPIE Int. Symp. Optical Science and Technology, Imaging Spectrometry IX (AM105), Denver, CO., 2004, 49: 2—4.

[9] Chang C I, Ji B H. Weighted abundance constrained linear spectral mixture analysis [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44, 378—388.