文章编号：0253-2239(2009)09-2607-08

基于非线性相关系数核方法的超谱数据分类

张 淼 沈 毅 王 强
(哈尔滨工业大学控制科学与工程系，黑龙江 哈尔滨 150001)

摘要 针对一对一策略的支持向量机算法进行了加权改造，提出一种新的基于非线性相关系数的核方法，在没有地物真实参考图的情况下进一步提高了超谱数据的分类精度。该方法考虑到遥感超谱数据信息依波段分布不均匀的特性，采用非线性相关系数对各波段数据在核函数内部进行加权，使得与参考图相关信息多的波段在分类器中发挥更为显著的作用。同时还提出一种基于非线性相关系数的参考图估计算法，解决了实际应用中真实参考图难以获取的问题。实验对比了采用径向基函数核的支持向量机分类器，结果显示在内部参数为典型值时，所提方法可在无需地物真实参考图的情况下将多分类平均精度和总体精度提高 2.90% 和 3.11%，且运算耗时无明显增加。

关键词 遥感；超谱数据分类；核方法；非线性相关系数；支持向量机；径向基函数
中图分类号 TP751.1 **文献标识码** A **doi**：10.3788/AOS20092909.2607

Nonlinear Correlation Coefficient Based Kernel Method for Hyperspectral Data Classification

Zhang Miao Shen Yi Wang Qiang
(*Department of Control Science and Engineering, Harbin Institute of Technology, Harbin, Heilongjiang* 150001, *China*)

Abstract Under the framework of support vector machines using one against one strategy, a novel kernel method based on nonlinear correlation coefficient is proposed to raise the classification accuracy under the most conditions of no ground truth reference map. This method takes into account the non-uniform information distribution of remote sensing hyperspectral data, and assigns nonlinear correlation coefficients as weights for the corresponding bands to make the band with greater correlation information play a more important role during the process of classification. Meanwhile a new estimated reference map based on nonlinear correlation coefficient is proposed to solve the realistic problem that the real one is usually unavailable. The experimental results show that for the support vector machines based on radial basis function, after adopting the proposed kernels, the average accuracy and the overall accuracy in multi-classification are increased by 2.90% and 3.11% with typical parameter configuration and no ground truth reference map, besides the computational time increment is unobvious.

Key words remote sensing; hyperspectral data classification; kernel method; nonlinear correlation coefficient; support vector machine; radial basis function

1 引 言

核方法(kernel method)在许多模式识别应用中都展现出了优良的性能，例如人们非常熟悉的支持向量机(SVM)分类器。SVM 通过最大化分类间隔来区分两个待分类的对象，并不依靠对训练样本概率密度函数的估计，因而该方法受输入空间高维

收稿日期：2008-12-29；**收到修改稿日期**：2009-04-13
基金项目：国家自然科学基金(60604021,60874054)资助课题。
作者简介：张 淼(1980—)，男，博士研究生，主要从事光谱分析、统计学习理论等方面的研究。
E-mail：ieeemiao@126.com
导师简介：沈 毅(1965—)，男，教授，博士生导师，主要从事超声成像技术、控制系统故障诊断等方面的研究。
E-mail：shen@hit.edu.cn

数的影响很小[1]。一些将 SVM 应用于超谱数据分类的文献都显示了其不逊于当前任何一种算法的优良性能[2,3]。但是,鲜有研究致力于扩展 SVM 方法使其更适合以超谱数据为对象的分类应用。

提出了一种新的基于非线性相关系数(NCC)[4]的波段加权核函数以更好地利用 SVM 分类器对超谱数据进行分类。分析了对于分类有用的相关信息沿波段分布的非均匀性。介绍了波段加权核函数,它考虑了每个波段对分类而言所具有的不同重要程度。波段加权核函数面临的突出问题是如何确定加权系数[5,6]。利用 NCC 对核函数中的各波段进行加权,使得有用信息多的波段在分类器中发挥更显著的分类作用。相对于增加数据滤波器的做法[7],该方法的优点是既不需要修改训练数据格式,也不需要对测试数据做相应的变换。考虑到分类任务往往是在没有地物真实参考图的情况下进行的,而真实参考图对于衡量各波段的有用信息又非常重要,为了解决这个困难,还提出了一种新的估计参考图解决方案,并在实验中取代真实参考图应用于所提出的核方法分类器中。

2 利用超谱数据特性进行波段加权

超谱传感器在广泛的光谱范围内采集信号,并以此作为精细区分各类地物的判定依据[8,9]。然而物质的光谱反射曲线通常是存在重叠的。许多类似的物质可能只在个别波段的光谱反射曲线上存在差异。可以预料,不同的波段对于鉴别出感兴趣的物质有着不相等的信息含量[10]。

图 1(a)举例说明了来自于 AVIRIS(Airborne Visible/Infrared Imaging Spectrometer)传感器的 92AV3C 数据集合中干草列(hay-windrowed)和小树林(woods)的光谱反射曲线。可以看到:同一类别地物的超光谱成像具有相似的曲线;不同类别地物的曲线之间虽有不同,但仍在多数波段上大致重叠。为了区分两类地物,还必须参考反射曲线在各波段的统计特征,例如均值和标准差。从图 1(c)来看,两类地物在波段 120-140 和 170-210 这两个子带上具有较小的标准差,再看图 1(b),两类地物的反射曲线在子带 120-140 上的差别更大,因而在该子带上两类地物的可分性要高于其余波段。就多分类而言,整个波段中的一些波段也会比其余波段包含更多对分类有帮助的信息,这就引出了一类改进办法,即通过波段处理手段来侧重于有效用的波段,从而提高分类精度。

在 SVM 分类的框架下,一种直接的方法就是定制专属核函数:根据各个波段所含有用信息的多少来给各波段分布不同的加权系数。令 $\boldsymbol{x}^i=[x_1^{(i)},x_2^{(i)},\cdots,x_N^{(i)}]$为 N 维超谱数据向量,代表超谱图像中某一点的各波段数据,$y_i\in\{-1,1\}$为分类目标,$\boldsymbol{\alpha}=[\alpha_1,\alpha_2,\cdots,\alpha_M]$为拉格朗日乘子,其中 M 是样本的数目。则 *SVM* 分类器可表示为

$$f(\boldsymbol{x})=\mathrm{sgn}\Big[\sum_{i=1}^{M}y_i\alpha_iK(\boldsymbol{x}_i,\boldsymbol{x})+b\Big],\qquad(1)$$

其中 $\boldsymbol{x}$ 是 N 维超谱数据输入向量,b 是阈值。此外,$K(\boldsymbol{x},\boldsymbol{x}')=\Phi(\boldsymbol{x})^{\mathrm{T}}\Phi(\boldsymbol{x}')$是一个具有对称内积的核函数。常用的核函数有径向基函数(RBF)核及多项式核分别为

$$K(\boldsymbol{x},\boldsymbol{x}')=\exp\Big(-\frac{\|\boldsymbol{x}-\boldsymbol{x}'\|^2}{2\sigma^2}\Big),\qquad(2)$$

$$K(\boldsymbol{x},\boldsymbol{x}')=(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}'+1)^d.\qquad(3)$$

标准的核函数,例如 RBF 及多项式核函数,对每个组成元素 $x_n^{(i)}$ 在特征空间的映射都是一致的[10]。文献[5]提出了一种更为有利的做法,即根据有用信息的多少来突出或减弱元素 $x_n^{(i)}$ 在核方法中

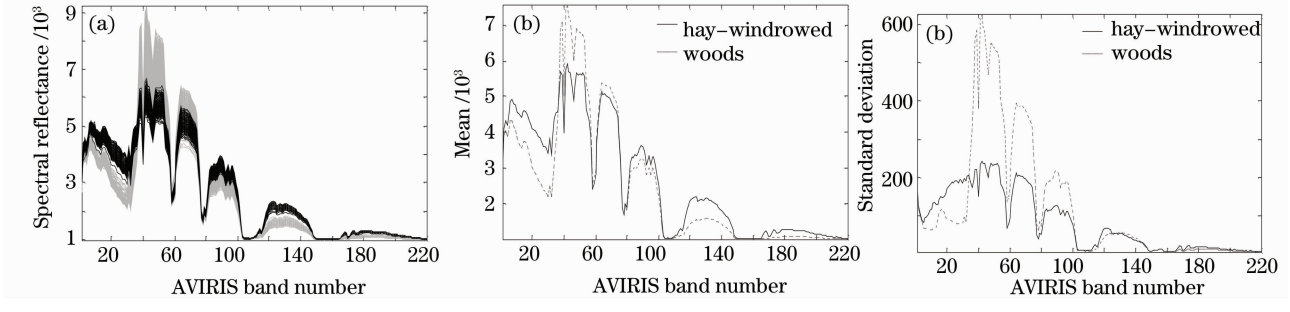

图 1 92AV3C 数据集合中两类地物各 150 个像素的光谱反射值(干草列为深色,小树林为浅色)(a)、相应均值(b)及标准差(c)

Fig. 1 Curves of 150 sampled spectral reflectances (hay-windrowed is in dark and woods in light)(a), corresponding means (b) and standard deviations (c) for two land-cover species in 92AV3C dataset

的作用，并通过将不同的加权系数分配给不同的波段来构造波段加权核函数。其做法是将一系列加权系数，$\boldsymbol{s}=[s_1,s_2,\cdots,s_N]$，依次分配给超谱数据向量中的每个元素 $x_n^{(i)}$，然后再将它们映射到特征空间中去。引入 $\boldsymbol{s}$ 后的波段加权 RBF 核和波段加权多项式核可表示为

$$K_{\mathrm{SW}}(\boldsymbol{x},\boldsymbol{x}')=\exp\left(-\frac{||\boldsymbol{S}(\boldsymbol{x}-\boldsymbol{x}')||^2}{2\sigma^2}\right), \quad (4)$$

$$K_{\mathrm{SW}}(\boldsymbol{x},\boldsymbol{x}')=(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{S}^{\mathrm{T}}\boldsymbol{S}\boldsymbol{x}'+1)^d, \quad (5)$$

其中 $\boldsymbol{S}$ 为对角矩阵，且 $\boldsymbol{S}=\mathrm{diag}(\boldsymbol{s})$。判断波段加权核函数是否是核函数的充分必要条件已由 Mercer 条件[10]给出，由于 $K_{\mathrm{SW}}(\boldsymbol{x},\boldsymbol{x}')=K(\boldsymbol{S}\boldsymbol{x},\boldsymbol{S}\boldsymbol{x}')$，因此易证 $K_{\mathrm{SW}}(\boldsymbol{x},\boldsymbol{x}')$ 是满足 Mercer 条件的合格核函数。为了使波段加权核函数能够达到提高分类精度的目的，关键是合理有效地估计加权系数，这也是本文重点研究的内容。

3　基于 NCC 的波段加权核函数

3.1　应用 NCC 估计加权系数

波段加权核函数中的加权系数直接决定了每个波段与核函数之间的关联程度，因而需要一种合理的方法来估算这些系数。可供选择的方法包括从原始数据中进行自学习或是通过先验知识进行预测。文献[5]提出了一种直接使用互信息(MI)来估算波段加权系数的直接方法。使用 MI 的优势是基于它与贝叶斯分类误差的密切联系以及其执行的高效性[11]。但是该方法有两方面的缺陷：1)由于 MI 不具有极值性，有用信息占很大比重的波段可能因信息总量不大而被淹没；2)所需的地物真实参考图往往难以及时绘制或根本无法绘制。

为了改进 MI 的不足，选取了具有极值性的 NCC[4]来进行加权系数的估计。考虑两个离散变量 X 和 Y，其元素个数均为 N，变量可取的状态数为 b。N 个元素中不同数值的元素个数需大于 b，否则会导致某些状态出现零个元素，从而产生奇异性的运算。状态的分布由以下的方式来确定：首先，将变量 X 和 Y 的元素分别按从小到大的顺序排列；然后，将最前面的 N/b 个值设为第一个状态，接下来的 N/b 个值设为第二个状态，依此类推，并称每个状态的最小值和最大值为状态阈值；最后，对于变量 X 和 Y，它们的元素对 $(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)$ 将根据前面所确定的状态阈值放入 $b\times b$ 的二维状态格中。

经过以上处理之后，变量 X 和 Y 的任意状态概率为 $p_i=1/b$，变量 X 和 Y 的联合概率为 $p_{ij}=n_{ij}/N$，其中 n_{ij} 是第 (i,j) 个二维状态格中元素对的个数。NCC 被定义为：

$$H_{\mathrm{NCC}}(X,Y)=H^r(X)+H^r(Y)-H^r(X,Y), \quad (6)$$

其中 $H^r(X,Y)$，$H^r(X)$ 及 $H^r(Y)$ 的定义如下：

$$H^r(X,Y)=-\sum_{i=1}^{b}\sum_{j=1}^{b}p_{ij}\log_b p_{ij}, \quad (7)$$

$$H^r(X)=H^r(Y)=-\sum_{i=1}^{b}p_i\log_b p_i, \quad (8)$$

注意到 $p_i=1/b$，则 H_{NCC} 可简化为：

$$H_{\mathrm{NCC}}(X,Y)=2+\sum_{i=1}^{b}\sum_{j=1}^{b}p_{ij}\log_b p_{ij}, \quad (9)$$

变量 X 和 Y 的 N 个元素对在 $b\times b$ 的二维状态格中的分配蕴含着两个变量间统计意义上的普遍相关性，因而能够衡量两个变量之间的非线性相关关系[12]。

3.2　应用 NCC 估计参考图

在利用波段加权核函数的时候，各加权系数决定了各波段数据对分类有用的信息含量。因而需要一个衡量有用信息所使用的参考标准。地物真实参考图是非常好的参考标准，但是由于需要人工辅助测绘往往难以及时提供或根本无法绘制，因而需要一种能够实时自动生成的估计参考图。文献[13]在分析了各波段的有用分类信息分布后，指定了一些相关信息趋势平缓的波段作为参考图通用的关键子带，但是该方法过于依赖专家知识，且对不同地域的超谱图像，事先指定的子带集合可能缺乏足够的适应性。

提出一种更为有利的做法，即利用相邻波段之间的相关程度从整个谱带中自动分离出紧密相关的子带作为对参考图进行估计的关键子带。MI 可以作为衡量相邻波段之间相关信息的度量，但是 MI 不具有极值性，因而在选取关键子带集合时往往难以确定选拔阈值，而 NCC 所具有的值域为[0,1]闭区间的极值性可以弥补这个不足，使得所选定的阈值具有普适性。

在进行参考图估计之前，首先要除去 220 个波段中受水汽吸收干扰的 20 个波段，这些波段的序列相对比较稳定，具体分别为第 104-108，第 150-163 以及第 220 波段[14]；然后对相邻波段求取 H_{NCC}(如图 2 所示)；最后，以 0.5 为阈值选取连续相关且波段数目大于等于 15 的子带作为关键子带集合。则可以得到 3 段符合要求的子带，分别为第 15-30，115-144 以及 170-218 波段。如图 3 所示：应用 H_{NCC} 计算出的关键子带(图中 * 号标识)亦符合专家认可的经验判断，即关键子带与地物真实参考图

的 MI 值较高且在一定长度上保持平稳[13]。最后，对该关键子带集合 T 的所有图像应用平均值法求出估计参考图 $\hat{R}$：

$$\hat{R}=\frac{1}{\text{length}(T)}\sum_{i\in T}M_i, \tag{10}$$

其中 length(T) 表示关键子带集合 T 的长度。

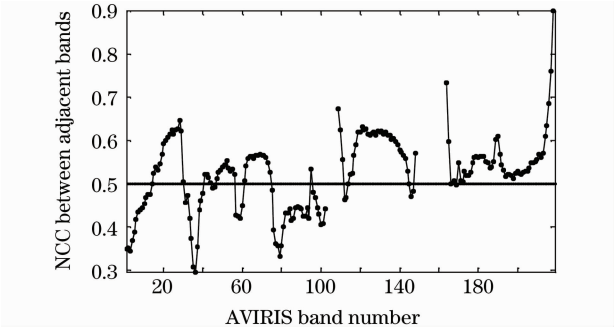

图 2 92AV3C 数据集合中相邻波段间的 NCC 值（不包括水汽吸收波段）

Fig. 2 The NCC values of 92AV3C dataset between adjacent bands (without water absorption bands)

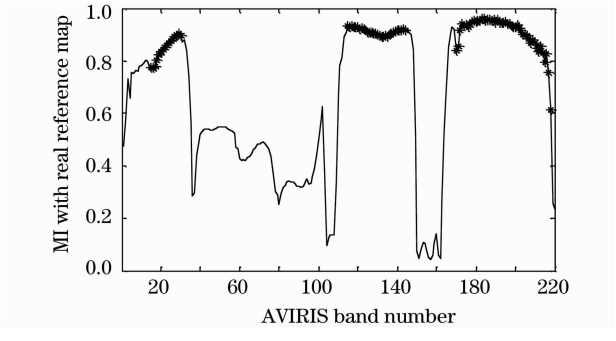

图 3 92AV3C 数据集合中各波段与地物真实参考图的 MI

Fig. 3 The MI values of 92AV3C dataset between each band and the ground truth reference map

在关键波段的选取上，只是通过简单的阈值划分及适当连续来达到筛选目的，但是该方法仍有待改进，例如，需要首先排除受水汽吸收影响的波段，否则这些受干扰的波段将使得关键子带集合的边界更加不确定。对此，可以通过粗糙集等理论来进一步研究集合对象之间的不可分辨性，从而得到更加适宜的估计参考图。

利用估计参考图 $\hat{R}$ 与各波段图像 $\boldsymbol{x}^{(i)}$ 进行计算所得到的 H_{NCC}，即可作为(4)式和(5)式中的各加权系数 s_i：

$$s_i=H_{\text{NCC}}(x^{(i)},\hat{R}), \tag{11}$$

4 数据描述与实验设计

选用 AVIRIS 传感器所采集的 92AV3C 超谱数据集合作为实验对象。该数据集合包含 224 个连续波段的 AVIRIS 图像，从 0.40 μm 到2.45 μm大约每隔 10 nm 一个波段，为美国印第安纳州西北部某农业地区[15]。采用该数据集合的优势是其附带了通过实地测绘而得到的地物真实参考图，利用该图可检验分类精度。去掉 4 个全为 0 值的波段以及 20 个受水汽吸收影响的波段（采用文献[14]所建议删除的第 104-108，150-163 以及 220 波段），实际采用的波段为 200 个。

针对文献[5]所提出的基于 MI 的波段加权核方法，文献[6]对其各个子分类器的分类性能进行了更为全面的实验分析，并同基于减小泛化误差(generalization error)而推导出的梯度下降法进行了对比。结果显示两者分类性能相近，但后者运算耗时过大，且每次迭代后的分类误差都无法准确估计，从而不能保证分类性能得到稳定提升，因此选取前者作为本文所提方法的比较对象。选取了 16 类地物中像素数最多的 7 类地物作为实验样本，这 7 类地物的像素总数占了所有 16 类地物像素总数的 80.64%，为进行多折交叉验证提供了充足的样本数量。实验采用 5 折交叉验证来计算其分类误差，即将各类地物的所有像素随机分为 5 份，每次取各类地物的 1 份用于训练，余下的 4 份用于分类（各类地物的具体像素数见表 1 所示）。

表 1 各类地物所对应的训练样本数和测试样本数

Table 1 Numbers of training and testing pixels in each class

	Class	Training pixels	Testing pixels
A	corn-no till	287	1147
B	corn-min till	167	667
C	grass/trees	149	598
D	soybeans-no till	194	774
E	soybean-min till	494	1974
F	soybean-clean till	122	492
G	woods	259	1035

本文所提出的核方法是基于 SVM 算法的框架，然而 SVM 本质上是两分类器，因此需要利用一系列该两分类器并辅以一定策略构造出多分类器。目前广泛应用的策略有 3 种，分别是决策树、一对多(OAA)和一对一(OAO)策略。决策树策略需要指定各分割面的分类对象，因而算法不具有通用性。OAA 策略是 SVM 早期最普遍的多分类方法，即每个 SVM 都要解决某一类对其余所有类的两分类问题，最后通过比较分类函数值的大小确定最终类别。该策略的主要问题是某一类对其它所有类的判别通

常会导致过于复杂的判别函数，此外每个 SVM 分类器都要解决相差很大的先验概率对训练所造成的困难。而 OAO 策略对任意两类都构造一个 SVM 二分类器，并将大量 SVM 并行运算，对结果投票选举确定最终类别。该策略使得各 SVM 判别容易，在训练时间上有着非常好的表现，因此采取 OAO 策略来进行所提核方法的多分类实验。

5　实验结果分析

分别采取了四种核方法对 92AV3C 数据集合进行 7 类地物的分类实验，它们是标准 SVM 方法，基于 MI 的波段加权核方法(SWKM)，本文改进的 MI-SWKM 以及本文提出的 NCC-SWKM。其中第一种方法无需参考图，第二种由文献[5]所提出的方法需要地物真实参考图，其余两种方法则自行估计参考图。4 种方法皆通过 OAO 策略来构造多分类器，针对 7 类地物的分类实验，每种方法都需要 C_7^2，即 21 个子分类器。实验中 NCC 算法的状态数选取 100(在 50 到 150 之间变化时分类效果几乎不变)，惩罚因子 C 以及 RBF 核函数的参数 σ 均依照文献[6]所选择的数值，即 $C=60$ 和 $\sigma=0.4$。为了准确衡量所提方法对分类性能所产生的影响，分别从单个子分类器及整体多分类器两个方面来分析实验结果。

在表 2 中，记录了各子分类器在 5 折交叉验证后的分类误差均值及标准差。由于被子分类器正确分类的样本仍有可能在最后的投票中被错误地标定为其它类别，而被子分类器错误分类的样本也有可能在最后的投票中被纠回到正确的分类，所以不同的统计方法往往得到不同的分类误差。鉴于实验的目的是为了反映各核方法相对于标准 SVM 对各子分类器的性能改善，因此采用各子分类器错误分类的样本数除以该子分类器所涉及到的样本总数来表示分类误差，并不计入由于其它子分类器的错误投票所导致的错误分类。

表 2　采用 RBF 核的各子分类器的分类误差及标准差比较，5 折交叉验证，$\sigma=0.4$，$C=60$

Table 2　Comparison of each sub-classifier's classification error and its standard deviation with RBF kernel, 5-fold cross validation, $\sigma=0.4$ and $C=60$

	Classification error (%) ± standard deviation (%)			
	Standard SVM	MI-SWKM	Improved MI-SWKM	Proposed NCC-SWKM
A|B	6.89±0.49	5.10±0.83	5.34±0.90	4.64±0.49
A|C	0.55±0.17	0.40±0.21	0.56±0.20	0.45±0.20
A|D	6.72±0.34	6.07±0.86	5.21±0.59	5.05±0.66
A|E	9.70±0.55	6.50±0.67	6.75±0.34	5.79±0.41
A|F	3.36±0.43	2.61±0.53	2.38±0.33	1.60±0.37
A|G	0.17±0.08	0.13±0.09	0.18±0.11	0.14±0.06
B|C	0.52±0.13	0.28±0.18	0.43±0.21	0.38±0.23
B|D	3.12±0.65	2.58±0.60	2.50±0.11	2.01±0.26
B|E	4.78±0.39	4.76±0.33	4.30±0.30	4.30±0.33
B|F	4.78±0.73	4.83±0.32	3.95±0.58	3.88±0.74
B|G	0.02±0.03	0.01±0.03	0.09±0.07	0.05±0.06
C|D	0.66±0.19	0.64±0.17	0.60±0.16	0.60±0.10
C|E	0.32±0.13	0.42±0.18	0.39±0.21	0.25±0.07
C|F	0.46±0.07	0.44±0.15	0.57±0.20	0.42±0.19
C|G	0.72±0.22	0.82±0.30	0.75±0.30	0.62±0.22
D|E	7.13±0.42	6.08±0.38	6.19±0.29	6.02±0.38
D|F	3.49±0.82	3.03±1.46	2.47±0.42	1.74±0.64
D|G	0.03±0.05	0.02±0.03	0.07±0.06	0.02±0.03
E|F	2.97±0.31	2.57±0.61	2.20±0.58	1.90±0.37
E|G	0.25±0.10	0.19±0.14	0.21±0.10	0.21±0.14
F|G	0.07±0.09	0.03±0.04	0.00±0.00	0.00±0.00

表 2 中各分类误差的标准差都比较小，这说明 5 折交叉验证所得分类精度的一致性较好，因此可直接用各子分类器的分类误差均值来评定分类效果。进一步分析可知：标准 SVM 在所有方法中效

果最差，其 21 个子分类器中有 13 个的分类误差比另外 3 种方法的都大；使用地物真实参考图的 MI-SWKM 略差于使用估计参考图的相同方法，前者有 11 个子分类器的误差大于后者，且在分类误差总和上前者也较后者多出 2.37%；同样使用估计参考图的 NCC-SWKM 在所有方法中效果最好，其 21 个子分类器中有 16 个的分类误差都是 4 种方法中最小的，分类误差总和上比前 3 种方法分别减少了 16.64%，7.44%和 5.07%。这说明针对各波段有用信息分布不均匀的特性所设计出的波段加权核函数使得 SVM 算法更好地适应了超谱数据的特点，从而使得大多数子分类器都有效地降低了分类误差。

在表 3 中，从平均精度、总体精度、预处理时间、训练时间、分类时间以及支持向量总数 7 个指标来分析 4 种方法的总体分类性能。其中平均精度定义为 7 类地物分类精度的均值，总体精度则定义为所有正确分类样本占总样本数的百分比。一般更侧重于平均精度指标，也有文献直接将平均精度称为总体精度[14]。预处理时间即波段加权系数的计算时间。运算量最大的训练部分由执行效率很高的 C 语言编写，时间统计上反而较短，另外两部分由 Matlab 语言编写，时间统计上则显得较长。

表 3 采用 RBF 核时的分类精度、时间消耗以及支持向量总数的比较，5 折交叉验证，$\sigma=0.4$，$C=60$

Table 3 Comparison of classification accuracy, time cost and the sum of support vectors with RBF kernel, 5-fold cross validation, $\sigma=0.4$ and $C=60$

Evaluation indexes	Standard SVM	MI-SWKM	Improved MI-SWKM	Proposed NCC-SWKM
Average accuracy/%	88.85	90.54	91.09	91.75
Overall accuracy/%	88.51	90.59	90.92	91.62
Preprocess time/s	0.00	23.00	22.38	69.01
Training time/s	1.64	1.53	1.56	1.42
Testing time/s	13.44	11.76	12.41	9.70
Support vector number	2703	2366	2357	1847

从表 3 可见，经过波段加权改造后，分类的平均精度和总体精度都得到一定提高，且基于 NCC 的核方法比基于 MI 的核方法在平均精度上高出了 1.21%。值得注意的是本文在对文献[5]方法进行改进后，分类效果略有提高，而且没有采用地物真实参考图，免去了由于绘制困难给应用造成的限制；此外，本文所提出 NCC-SWKM 也采用的是估计参考图，并带来了 3%左右的精度提升。造成这两种现象的原因可能来自于 3 个方面：1）是 MI 所用地物真实参考图中高达 50.7%的未标类像素对加权系数的估计造成了干扰；2）是 NCC 具有极值性，在加权过程中不容易淹没信息总量小的波段，造成有用分类信息的丢失；3）是 NCC 较 MI 能更好地度量非线性相关信息。在耗时方面，NCC-SWKM 较两种基于 MI 的核方法有着 3 倍左右的预处理时间，但由于预处理是由执行效率较低的 Matlab 语言编写，实际运算量远小于分类器的训练运算量。在训练时间上，4 种方法基本在同一水平上，相对来说 NCC-SWKM 能更快地满足训练终止条件。而在最后的分类计算中，运算量是固定的，它和支持向量数以及核函数复杂程度有关，由于该部分也是由 Matlab 语言编写所以在时间上也显得较长。综合 4 种方法在各阶段时间的表现，NCC-SWKM 在运算量最大且无法确定终止步数的训练阶段有着最短的耗时，所以在总体耗时上新方法并不会给应用造成困难。此外，通过表 3 中支持向量总数的统计，基于波段加权的核方法对支持向量还是有一定约减的，并直接导致了在分类时间上的优势。

本文所提出的基于 NCC 的波段加权算法可应用于不同的标准核函数（如 RBF 核、多项式核等），有必要对它们加权后的分类性能做出比较，因为不同的核函数有着不同的内部结构，而且即使应用同一种核函数，其内部参数的变化也会导致分类性能的差异。鉴于此，对采用 RBF 核和多项式核的 NCC 波段加权算法在不同参数下进行分类性能的对比，并选取平均精度（见表 4）为分类性能的评价指标。惩罚因子 C 的范围为 0.1 至 1000，RBF 核参数 σ 的范围为 0.2 至 1，多项式核参数 d 的范围为 1 至 11。同时对表 4 中分类精度较高（大于 88%）的分类器在训练时间和支持向量数方面做进一步统计（见表 5），以便于我们更好地分析参数变化对两种核函数加权算法的影响。

从表 4 来看，内部参数的变化对基于 RBF 核的 NCC-SWKM 的分类性能影响很大，但是最佳参数配置仍处在 $\sigma\in(0.2,1)$ 及 $C\in(10,100)$ 这一广泛认同的典型区间（进一步实验确认最佳参数为 $\sigma=0.4$ 及 $C=80$），所以并不会给应用带来困难。反观基于多项式核的 NCC-SWKM 则在分类精度上始

终存有 2%左右的差距。进一步分析表 5 所提供的信息，在分类效果可接受的参数配置中，基于多项式核的 NCC-SWKM 有着近乎相同的支持向量机数量，而基于 RBF 的核方法则相差很大。多项式核缺乏可变通的内部结构使得难以利用它发挥出更好的分类性能。在训练时间上，基于多项式核的 NCC-SWKM 平均为基于 RBF 核方法的 8 倍多，认为多项式核并不适合做基于 NCC 的加权改造。

表 4　NCC-SWKM 方法在不同参数下的平均精度(%)

Table 4　Average accuracies (%) of proposed NCC-SWKM in different parameters

Kernel type		C in Proposed NCC-SWKM				
		0.1	1	10	100	1000
RBF	$\sigma=0.2$	55.53	82.07	90.17	90.23	90.23
	$\sigma=0.4$	46.99	76.19	89.37	91.80	91.65
	$\sigma=0.6$	43.16	69.03	86.63	91.70	91.67
	$\sigma=0.8$	41.87	63.16	84.33	90.64	91.49
	$\sigma=1$	41.70	51.83	81.88	89.49	91.46
	$\sigma=2$	41.31	42.08	65.96	85.47	89.92
Poly.	$d=1$	41.77	48.63	75.98	85.40	87.15
	$d=3$	88.62	88.25	88.25	88.25	88.25
	$d=5$	88.84	88.84	88.84	88.84	88.84
	$d=7$	89.14	89.14	89.14	89.14	89.14
	$d=9$	89.30	89.30	89.30	89.30	89.30
	$d=11$	89.47	89.47	89.47	89.47	89.47

表 5　平均精度大于 88%的 NCC-SWKM 方法在训练时间和支持向量数上的比较

Table 5　Comparison of training time and support vector number of NCC-SWKM with average accuracy more than 88%

Evaluation indexes		NCC-SWKM	
		RBF	Poly.
Training time /s	Min	1.35	9.12
	Max	2.61	21.04
	Mean	1.85	15.88
Support vector number	Min	1170	1003
	Max	3505	1052
	Mean	1982	1012

6　结　　论

本文给出了一种较普通 SVM 有着更高分类精度的核方法。该方法根据分类对象(即超谱数据)所固有的信息分布不均匀的特点，利用 NCC 对各波段进行加权并将该过程封装在核函数内部，使得新的分类器无需对训练数据及测试数据做任何格式转换或滤波运算。NCC 可以用来衡量两波段之间的相关信息，同时具有极值性，用它作波段加权系数可有效避免有用信息比重大但总量不大的波段被淹没的现象，从而较基于 MI 的波段加权核方法进一步提升分类精度。通过对比基于 RBF 核函数的 4 种核方法在 AVIRIS 数据上的典型分类实验，本文所提 NCC-SWKM 算法无需地物真实参考图但却具有最佳的分类性能，运算时间增幅也不明显(计算量最大的训练时间反而有所下降)，而且其最佳参数配置位于广泛认同的 SVM 分类器的最佳参数选择区间。

同时，本文采用相邻波段间 NCC 大于阈值的子带集合来自动生成估计参考图，避免了对专家信息的过度依赖。鉴于 NCC 的极值性，该阈值在选取上相对比较稳定，也给参考图估计过程的在线实施带来便利。

最后，本文从参数选取及分类性能方面对比了基于 RBF 核和基于多项式核的 NCC-SWKM 算法。后者难以根据内部参数变化对支持向量作出调整，从而难以利用对象特性来进一步提高分类精度，而且在运算量最大的训练环节上耗时过大，因此多项式核并不适合做基于 NCC 的波段加权处理。

参　考　文　献

1 Zhang Rongxiang, Zheng Shijie, Xia Yanjun *et al*.. Application of support vector regression for reconstruction of non-uniform strain profile along the fiber grating[J]. *Acta Optica Sinica*, 2008, **28**(8): 1513～1517
张荣祥，郑世杰，夏彦君 等. 支持向量回归算法在光纤光栅非均匀应变重构中的应用[J]. 光学学报，2008，**28**(8)：1513～1517

2 M. Brown, H. G. Lewis, S. Gunn. Linear spectral mixture models and support vector machines for remote sensing[J]. *IEEE Trans. on Geoscience and Remote Sensing*, 2000, **38**(9): 2346～2360

3 G. Camps-Valls, L. Gomez-Chova, J. Calpe-Maravilla *et al*.. Robust support vector method for hyperspectral data classification and knowledge discovery[J]. *IEEE Trans. on Geoscience and*

Remote Sensing, 2004, **42**(7): 1530～1542

4 Wang Qiang, Shen Yi, Zhang Jianqiu. A nonlinear correlation measure for multivariable data set[J]. *Physica D: Nonlinear Phenomena*, 2005, **200**(3-4): 287～295

5 B. Guo, S. R. Gunn, R. I. Damper *et al*.. Hyperspectral image fusion using spectrally weighted kernels[C]. *Proc. Int. Conf. on Information Fusion*, 2005, **1**: 402～408

6 B. Guo, S. R. Gunn, R. I. Damper *et al*.. Customizing kernel functions for SVM-based hyperspectral image classification[J]. *IEEE Trans. on Image Processing*, 2008, **17**(4): 622～629

7 H. Kwon, N. M. Nasrabadi. Kernel spectral matched filter for hyperspectral imagery[J]. *International Journal of Computer Vision*, 2007, **71**(2): 127～141

8 Luo Qin, Tian Zheng, Zhao Zhixiang. Shrinkage-divergence-proximity locally linear embedding algorithm for dimensionality reduction of hyperspectral image[J]. *Chin. Opt. Lett.*, 2008, **6**(8): 558～560

9 Hou Ying, Liu Guizhong. Three-dimensional set partitioned embedded zero block coding alogrithm for hyperspectral image compression[J]. *Acta Optica Sinica*, 2008, **28**(1): 67～73

侯 颖，刘贵忠. 基于三维集合分裂嵌入式零块编码算法的超光谱图像压缩[J]. 光学学报，2008，**28**(1)：67～73

10 V. N. Vapnik. An overview of statistical learning theory[J]. *IEEE Trans. on Neural Networks*, 1999, **10**(5): 988～999

11 Jin Jing, Wang Qiang, Shen Yi. Registering multiple medical images using the shared chain mutual information[J]. *Chin. Opt. Lett.*, 2007, **5**(7): 389～392

12 Wang Qiang, Shen Yi. Performances evaluation of image fusion techniques based on nonlinear correlation measurement[J]. *Proc. IEEE Instrumentation and Measurement Technology Conf.*, 2004, **1**: 472～475

13 B. Guo, S. R. Gunn, R. I. Damper *et al*.. Band selection for hyperspectral image classification using mutual information[J]. *IEEE Geoscience and Remote Sensing Letter*, 2006, **3**(4): 522～526

14 F. Melgani, L. Bruzzone. Classification of hyperspectral remote sensing images with support vector machines[J]. *IEEE Trans. on Geoscience and Remote Sensing*, 2004, **42**(8): 1778～1790

15 Landgrebe D. Hyperspectral image data analysis[J]. *IEEE Signal Process Magazine*, 2002, **19**(1): 17～28