

DOI: 10.13203/j.whugis20130384

文章编号: 1671-8860(2015)05-0612-05

一种多特征转换的高光谱影像自适应分类方法

王 凯^{1,2} 舒 宁¹ 孔祥兵³ 李 亮¹

1 武汉大学遥感信息工程学院, 湖北 武汉, 430079

2 长江勘测规划设计研究院, 湖北 武汉, 430010

3 黄河水利科学研究院, 河南 郑州, 450003

摘要: 光谱相似性测度用来衡量像元光谱的相似程度, 是高光谱影像光谱匹配分类的重要工具之一, 一般通过设置阈值判断像元光谱和参考光谱是否相似来进行分类。在此基础上, 本文提出了一种多特征转换的高光谱影像自适应分类方法, 实现了各种光谱相似性特征和分类器相结合的一种自适应分类。实验结果表明, 本文提出的方法相比于传统的 SVM 方法, 分类的总体精度更高, 还可以避免部分传统光谱匹配分类方法中需要专家经验确定分类阈值的复杂过程。

关键词: 高光谱影像; 精细分类; 光谱相似性测度; 支持向量机

中图法分类号: P237.4; TP753 文献标志码: A

光谱匹配技术是高光谱影像地物识别的关键技术之一^[1], 常用于高光谱影像的精细分类。所谓光谱匹配分类, 就是利用光谱相似性测度值来判断地物光谱的类别属性。光谱相似性测度可用来衡量两条像元光谱曲线的相似程度。以光谱角匹配(spectral angel match, SAM)^[2]为例, 相似度域从 0(相似)到 1(不相似), 可以设定一个相似度静态阈值, 小于此阈值则表明两条光谱曲线相似。但是, 确定这个静态阈值往往需要大量的专家经验。近年来, 许多国内外学者针对传统的光谱相似性测度只考虑像元光谱单一特征的不足, 提出了一些新的测度, 取得了很好的效果^[3-4]。在这些研究的基础上, 本文提出了一种多特征转换的高光谱影像自适应分类方法, 通过这些光谱相似性测度值作为度量并引入分类器进行辅助, 用分类器自适应确定光谱相似性测度超平面, 既综合考虑了光谱的多种特征, 又避免了一些光谱匹配分类方法中阈值确定对分类结果造成的影响, 提高了高光谱影像分类的准确率。

1 光谱特征转换的自适应分类

1.1 常用的光谱相似性测度

光谱相似性测度通常可以分为确定性测度和

随机性测度两类^[3]。其中, 确定性测度主要包含光谱角匹配(SAM)、欧氏距离(ED)和光谱相关系数测度(SCM)^[2,5]等, 随机性测度包括光谱信息散度(SID)^[2]、正交投影散度(OPD)^[5]和光谱泛相似性测度(SPM)^[3]等。相关研究表明, 利用单一的光谱特征往往无法全面反映两条光谱曲线之间的相关性^[3-4], 因此, 本文综合考虑光谱矢量、形状、信息量和投影残差等特征, 将多种光谱测度应用于高光谱影像分类。

1.2 基于全光谱域的光谱特征自适应分类

基于全光谱域的光谱特征自适应分类(multi-feature conversion applied on full hyperspectral space, MFCFHS)方法通过以上多元光谱相似性测度和支持向量机(support vector machine, SVM)^[6]分类器的有效结合, 实现高光谱影像的精细分类, 主要分为光谱特征平滑、多特征光谱相似性测度矩阵构建、自适应超平面三个步骤。

首先进行光谱特征平滑。该步骤通过均值滤波最小化光谱特征矩阵的随机噪声, 将影像通过滤波器, 计算在特定正方形窗口下的均值作为窗口中心点的值, 则该中值滤波的计算公式为:

$$E_j = \frac{\sum_{i=j-m}^{j+m} S_i}{w}, j = 1, 2, \dots, n \quad (1)$$

收稿日期: 2013-08-06

项目来源: 国家 863 计划资助项目(2013AA102401); 中央高校基本科研业务费专项资金资助项目(201121302020007); 黄河水利科学研究院科技发展基金资助项目(黄科发 201301)。

第一作者: 王凯, 博士生, 主要从事高光谱影像解译研究。E-mail: wangk_wuhan@163.com

式中, S_j 表示真实的光谱; E_j 为经过平滑后的光谱; 设固定窗口的宽度为 ω , 则 $m = (\omega - 1)/2$, m 表示周围像元离窗口中心像元的距离; n 为影像的波段数。将原始影像进行固定窗口的光谱特征平滑, 可以有效去除影像中少数点由于某些因素造成的光谱值偏离正常值而产生的随机噪声。本文的固定窗口宽度为 3。

然后构建多特征光谱相似性测度矩阵。该部分主要分为计算类别均值向量、多特征光谱相似性测度计算、构建矩阵三个内容。

1) 计算类别均值向量。每一类的均值向量定义为:

$$\Phi_C = \left[\sum_{i=1}^k \frac{E_{1i}}{k} \quad \sum_{i=1}^k \frac{E_{2i}}{k} \quad \dots \quad \sum_{i=1}^k \frac{E_{ni}}{k} \right] \quad (2)$$

其中, E_{ji} 为训练样本集类别 C 中第 j 个波段经过中值滤波后的估计值; k 为类别 C 的样本数; n 为高光谱数据波段数。

2) 计算光谱相似性测度。通过构建光谱相似性测度模型, 计算影像每个像元与类别均值向量的光谱相似性测度。本文综合考虑光谱矢量、形状、信息量和投影残差等特征, 使用了 SAM、SCM、ED、SID、OPD 和 SPM 等 6 种光谱相似性测度。选择这些光谱相似性测度既能有效涵盖众多光谱特征, 又能避免由于选择了过多的光谱相似性测度而导致计算量过大。

3) 构建矩阵。根据以上 6 种光谱相似性测度计算光谱特征向量与类别均值向量的测度值, 得到的相似性测度值组成一组向量, 构建模型, 多特征光谱相似性测度矩阵定义为:

$$M = \begin{pmatrix} S_{11} & \cdots & S_{1q} \\ \vdots & & \vdots \\ S_{p1} & \cdots & S_{pq} \end{pmatrix} \quad (3)$$

式中, p 表示样本数; q 为类别数 \times 光谱相似性测度数。本文中含有 6 组光谱相似性测度、9 个类别, 则构建 6×9 即由 54 维的相似性测度值组成的矩阵。矩阵中的每一个列向量表示样本中所有像元与某一类别均值 C 在某一组光谱相似性测度下的测量值。

最后, 通过 SVM 进行自适应超平面。引入 SVM 分类器进行监督分类, 利用 SVM 分类器进行自适应相似度阈值选择。传统的 SVM 面对的

是二分类问题, 基于结构风险最小化原则, 通过间隔最大化分析, 在高维空间中求出一个最优超平面, 将数据一分为二。本文将 SVM 分类器作为自适应阈值的工具, 将多元高光谱测度进行自适应判定, 实现多种光谱特征融合的光谱匹配分类。

1.3 基于高光谱影像波段子空间的光谱特征自适应分类

基于高光谱影像波段子空间的光谱特征自适应分类方法 (multi-feature conversion applied on hyperspectral subspace, MFCHS) 将高光谱的整个波段分为不同的子空间分别进行处理, 这种处理被称为高光谱影像空间分解。将高光谱的光谱范围分为蓝波段 (400~499 nm)、绿波段 (500~570 nm)、红波段 (650~750 nm)、水汽吸收波段 (900~1 000 nm)、红外波段 (1 350~2 400 nm) 等几个空间。其余步骤与 MFCFHS 方法类似, 每一个类别 C 中的 p 个样本, 都有 270 个不同的相似性测度值 (9 个类别 \times 6 个不同的光谱测度 \times 5 个不同的光谱范围)。每一个训练样本和测试样本都将通过这种方式构建一组含有 270 维不同相似性测度值的向量, 建立分类模型, 利用 SVM 分类器进行自适应, 实现影像地物的分类。

2 实验与分析

2.1 AVIRIS 数据集

实验数据为 AVIRIS 传感器获取的美国印第安纳州西北部农田和森林混合地区的数据, 其拍摄于 1992 年, 包括 220 个波段, 去掉一些信噪比较低的波段 104~108 和 152~161 后, 保留了 205 个波段。

原始影像和实际地物分布图见图 1, 影像中共含有 17 个类别, 根据文献 [5] 的分析, 为了研究地物精细分类的需要, 本文从中选取了 9 类较难区分的地物进行精细分类。其中, 未耕玉米、疏耕玉米、未耕大豆、清除大豆的光谱特征用 SAM 或 SID 等光谱测度值衡量时非常相似, 湿草地与干草、苜蓿、树林的光谱曲线也难以区分, 影像中疏耕大豆区域为高度混合区域。因此, 选择的这些类别比较适合进行高光谱影像的精细分类实验。训练样本数和测试样本数分别为 1 938 和 3 359, 如表 1 所示。

表 1 样本选择

Tab. 1 Sample Selection

	未耕玉米	疏耕玉米	湿草地	苜蓿	干草	未耕大豆	疏耕大豆	清除大豆	树林	总计
训练	196	255	105	219	255	143	316	137	312	1 938
测试	256	368	266	366	370	471	484	329	449	3 359

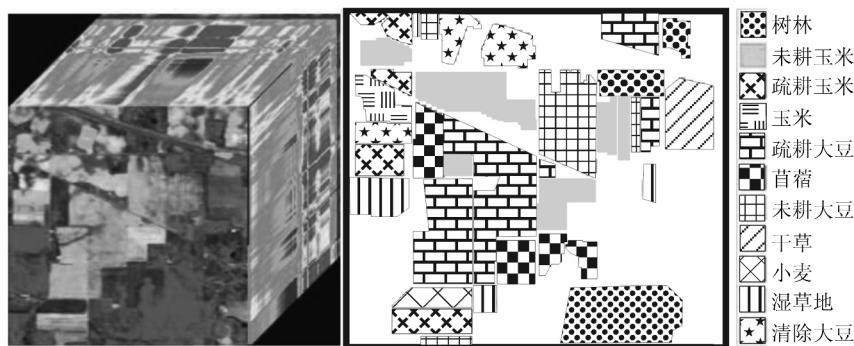


图 1 AVIRIS 影像和实际地物分布图

Fig. 1 AVIRIS Image and Classes Distribution

2.2 实验方法

实验方法设计为:①比较不同类型 SVM 分类器对本文方法实验结果的影响;②将本文方法与传统的 SVM 高光谱影像分类方法、最大似然分类法(maximum likelihood classification, MLC)和最小距离法(minimum distance, MD)进行比较分析。

本文提出了 MFCFHS 和 MFCHS 两种自适应光谱特征匹配分类方法。实验环境为 Matlab, 版本为 R2010b, 使用 LibSVM 工具箱^[6]的多类 SVM 分类器,SVM 分类器的核函数选择和参数设置参考相关文献^[7-9]。

2.3 结果分析

2.3.1 多类 SVM 分类器类型分析

表 2 为不同类型 SVM 分类器对分类结果的影响。LibSVM 算法进行多类分类时有两种策略:①一对多(one against all, OvA), 分类时依次把某个类别归为一类, 其他剩下的类别归为另一类, 这样, 含有 K 类地物的样本集就需要构建 K 个分类器;②一对一(one against one, OvO), 是将任意两类样本间构建一个分类器, 这样, 含有 K 类地物的样本集需要设计 $K(K-1)/2$ 个分类器。从表 2 中可以发现, OvO 的 SVM 分类结果优于 OvA 的 SVM 分类结果。这是由于 SVM 是一种二分类, OvA 是将某一类作为一类, 其他类作为一类进行二分类识别。这种分类方式不太适合精细分类, 当类别较多时, 有时无法区分一类和其他类的细微差别, 因而不能有效识别类别之间的细微差异, 造成 OvA 的分类结果较差。OvO 优于 OvA 是因为只采用两类的训练样本, 能较好地区分分类与类之间的差别, 因而具有较高的精细分类精度。

表 2 不同类型 SVM 分类器分类精度

Tab. 2 Different SVM Types Classification Accuracies

SVM 类型	训练样本	测试样本	分类精度/%	
			MFCFHS	MFCHS
OvA	1 938	3 359	76.18	78.48
OvO	1 938	3 359	78.27	80.74

2.3.2 不同分类方法结果和分析

分类结果通过总体精度、Kappa 系数、单类别使用者精度进行评价。表 3 为各分类方法的总体精度和 Kappa 系数, 可以看出, 传统的固定阈值光谱匹配分类方法(如最小距离法(MD))相比于 SVM 的分类精度而言, 下降了 20.37%。这是由于 MD 方法只考虑了光谱曲线的单一特征, 如果设置阈值则阈值需要人工调节, 这类算法虽然快速、简单, 但它对不同程度变化的光谱响应不敏感, 忽略了影像光谱中蕴含的其他有用信息, 因而分类结果在几种实验方法中精度最差。

表 3 各类方法的总体精度和 Kappa 系数

Tab. 3 Overall Accuracy and Kappa Value

	MD	SVM	MLC	MFCFHS	MFCHS
总体精度/%	51.65	72.02	76.93	78.27	80.74
Kappa 系数	0.5328	0.7141	0.7375	0.7789	0.8041

表 4 是单类别使用者精度。分析不同算法的单类别使用者精度, 可以发现 MLC 作为一种经典的分类方法, 对于某些类别(如苜蓿、干草、未耕玉米等)的使用者精度较高, 这是因为这些类别与其他类别光谱特征差异比较明显, 而对于湿草地、未耕大豆、清除大豆等光谱特征相似的类别, 最大似然法的分类精度较差。因此, 本文认为 MLC 不太适用于类别光谱差异较小的高光谱影像精细分类。

MFCFHS 方法和 MFCHS 方法相比于 SVM 方法在分类的总体精度上提高了 6.25% 和 8.72%, 因为它们通过使用 6 种不同的光谱相似性测度, 有效集合了光谱中各类有用信息精细分类, 各种光谱相似性测度组成的向量能通过 SVM

对难以区分的小类进行精细区分。MFCHS 方法在高光谱影像的精细分类中的精度最高,是因为其通过高光谱波段子空间分解和多元光谱相似性测度结合的方式,能最大程度上体现不同地物类别在光谱域的细微差别,增加了 SVM 对高度混合类别的区分能力。除此之外, MFCHS 方法的

分类精度优于 MFCFHS 方法,其主要原因是:MFCFHS 方法将光谱分为几个子空间,能最大化每个子空间内不同地物之间的信息差异,亦可最小化光谱范围过大而造成的对光谱信息统计独立性的影响,可以充分利用像元的光谱信息和波段间的细微差别。

表 4 单类别使用者精度/%

Tab. 4 User's Accuracy of Single Class/%

	未耕玉米	疏耕玉米	湿草地	苜蓿	干草	未耕大豆	疏耕大豆	清除大豆	树林
MD	51.17	32.88	25.57	46.17	60.27	31.42	77.06	36.60	81.73
SVM	53.13	63.04	77.07	59.84	68.92	51.80	100	69.00	92.87
MLC	92.35	80.43	45.49	99.73	100	32.27	98.55	38.91	99.78
MFCFHS	73.47	67.39	83.46	82.51	84.32	59.87	88.84	78.82	95.76
MFCHS	78.06	68.75	86.47	84.70	81.08	65.82	98.14	79.94	93.10

3 结语

本文设计了多特征转换的高光谱影像自适应分类方法,使用多种光谱相似性测度作为评价相似度的模型,引入一个分类器作为确定分类超平面的自适应工具。本文设计了两种实验方法,实验结果表明本文提出的方法相比于 SVM 方法在分类的总体精度上提高了 6.25% 和 8.72%,适应性更强。使用简单的可学习的光谱相似性测度进行高光谱影像精细分类,还可以有效避免一部分传统光谱匹配分类中需要专家经验确定分类阈值的复杂过程。

参 考 文 献

- [1] Zhang Liangpei, Zhang Lifu. Hyperspectral Remote Sensing [M]. Wuhan: Wuhan University Press, 2005(张良培,张立福. 高光谱遥感[M],武汉:武汉大学出版社,2005)
- [2] van der Meer F. The Effectiveness of Spectral Similarity Measures for the Analysis of Hyperspectral Imagery[J]. *Int J Appl Earth Obs*, 2006, 8(1):3-17
- [3] Kong Xiangbing, Shu Ning, Tao Jianbin, et al. A New Spectral Similarity Measure Based on Multiple Features Intergration[J]. *Spectroscopy and Spectral Analysis*, 2011, 31(8):2 166-2 170(孔祥兵,舒宁,陶建斌,等. 一种基于多特征融合的新型光谱相似性测度[J]. 光谱学与光谱分析, 2011, 31(8): 2 166-2 170)
- [4] Hosseini R S, Homayouni S, Safari R. Modified Algorithm Based on Support Vector Machines for Classification of Hyperspectral Images in a Similarity Space[J]. *Journal of Applied Remote Sensing*, 2012(6):3 115-3 123
- [5] Wu C, Chang C. Soft Decision-Made Hyperspectral Measures for Target Discrimination and Classification[J]. *Proc Soc Photo Opt Instrum Eng*, 2009(1):7 457
- [6] Chang C C, Lin C J. LIBSVM: A Library for Support Vector Machines, 2012 [EB/OL]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>, 2012
- [7] Shen Zhaoqing, Shu Ning, Tao Jianbin. An Algorithm of Weighted "1Vm" SVM Multi-classification for Hyperspectral Remote Sensing Image with NPA [J]. *Geomatics and Information Science of Wuhan University*, 2009, 34(12):1 444-1 447(沈照庆,舒宁,陶建斌. 一种基于 NPA 的加权 1VmSVM 高光谱影像分类算法[J]. 武汉大学学报·信息科学版, 2009, 34(12):1 444-1 447)
- [8] Tan Kun, Du Peijun, Wang Xiaomei. Multi-Class Support Vector Machine Classifier Based on Separability Measure for Hyperspectral Remote Sensing Image Classification[J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(2):171-175(谭琨,杜培军,王小美. 利用分离性测度多类支持向量机进行高光谱遥感影像分类[J]. 武汉大学学报·信息科学版, 2011, 36(2):171-175)
- [9] Shen Zhaoqing, Shu Ning, Tao Jianbin. On SVM Classification Algorithm of Hyperspectral RS Imagery Without Penalty Coefficient[J]. *Bulletin of Surveying and Mapping*, 2009(6):22-25 (沈照庆,舒宁,陶建斌. 无惩罚系数 SVM 高光谱影像分类算法研究[J]. 测绘通报, 2009(6):22-25)

A Multi-feature Conversion Adaptive Classification of Hyperspectral Image

WANG Kai^{1,2} SHU Ning¹ KONG Xiangbing³ LI Liang¹

1 School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

2 Changjiang Institute of Survey, Planning, Design and Research, Wuhan 430010, China

3 Yellow River Institute of Hydraulic Research, Zhengzhou 450003, China

Abstract: Spectral similarity measure is an important tool of hyperspectral remote sensing image classification. By setting the threshold to judge the pixel spectrum and the reference spectra is similar or dissimilar. To overcome this problem, this paper proposes a multi-feature conversion adaptive classification of hyperspectral image, this is done through using spectral similarity measure value as similarity patterns. Experimental results show that the proposed methods are, compared with the traditional SVM method in the overall accuracy of classification, increased by 6.25% and 8.72%, also it implies that using simple learnable measures outperforms complex and manually turned techniques used in traditional classification.

Key words: hyperspectral image; precise classification; spectral similarity measure; SVM

First author: WANG Kai, PhD candidate, specializes in hyperspectral imagery interpretation. E-mail: wangk_whu@163.com

Foundation support: The National 863 Program of China, No. 2013AA102401; Fundamental Research Funds for the Central Universities, No. 201121302020007; Development Funds for Yellow River Institute of Hydraulic Research, No. 201301.