

## 高光谱-LiDAR 融合的条件随机场分类方法

王雷光<sup>1,2</sup>, 耿若箏<sup>3</sup>, 代沁伶<sup>4</sup>, 王 军<sup>3</sup>, 郑 晨<sup>5\*</sup>, 付志涛<sup>6</sup>

- (1. 西南林业大学 大数据与人工智能研究院, 云南 昆明 650224;
2. 西南林业大学 森林生态大数据国家林业与草原局重点实验室, 云南 昆明 650224;
3. 西南林业大学 林学院, 云南 昆明 650224;
4. 西南林业大学 艺术与设计学院, 云南 昆明 650224;
5. 河南大学 数学与统计学院, 河南 开封 475004;
6. 昆明理工大学 国土资源工程学院, 云南 昆明 650093)

**摘要:** 为有效利用高光谱影像与 LiDAR 数据的互补性信息, 解决单一融合策略造成的场景解译地物边界不准确和分类精度低的问题, 提出了一种光谱-空间-高度特征融合、并顾及场景地物类别共生特性的条件随机场分类方法。首先, 对两种数据分别提取光谱及形态学特征, 对特征集采用图模型进行特征融合, 将特征输入概率支持向量机分类器, 得到初始分类结果。然后, 基于融合特征计算反映像素间类别本质差异的局部光谱-空间-高度协同的异质性值, 并统计类别间的空间共生关系。最后, 在条件随机场框架内, 整合初始分类结果、局部异质性信息及类别共生关系, 通过目标函数的迭代求解获得最终分类结果。通过将像素间的权重定义为对应像素位置融合特征的归一化欧式距离的单调减函数, 对标记不同但特征差异较大的类别间给予较小的权重, 以达到地物边界空间规整化的目的。通过对标记不同但共生概率较大的类别对给予较小的权重, 达到保留空间关系稳定的类别对的目的。采用城区场景的美国休斯顿地区数据集和林区场景的中国广西高峰林场两组数据集对提出方法进行了验证。实验结果表明: 休斯顿和高峰林场数据集精度分别达到 94.00% 和 92.84%, 分类结果的“胡椒盐”现象明显减少, 证明了该方法的有效性。

**关键词:** 高光谱; LiDAR; 条件随机场; 图模型特征融合; 局部异质性; 空间共生关系  
**中图分类号:** TP753 **文献标志码:** A **DOI:** 10.3788/IRLA20210112

## Conditional random field classification method based on hyperspectral-LiDAR fusion

Wang Leiguang<sup>1,2</sup>, Geng Ruozheng<sup>3</sup>, Dai Qinling<sup>4</sup>, Wang Jun<sup>3</sup>, Zheng Chen<sup>5\*</sup>, Fu Zhitao<sup>6</sup>

- (1. Institutes of Big Data and Artificial Intelligence, Southwest Forestry University, Kunming 650224, China;
2. Key Laboratory of National Forestry and Grassland Administration on Forestry and Ecological Big Data, Southwest Forestry University, Kunming 650224, China;
3. Forestry College, Southwest Forestry University, Kunming 650224, China;
4. College of Art and Design, Southwest Forestry University, Kunming 650224, China;
5. College of Mathematics and Statistic, Henan University, Kaifeng 475004, China;
6. Faculty of Land and Resources Engineering, Kunming University of Science and Technology, Kunming 650093, China)

收稿日期: 2021-02-17; 修订日期: 2021-06-07

基金项目: 国家自然科学基金 (31860182, 41771375, 41961053, 41571372); 云南省中青年学术和技术带头人后备人才项目 (2018HB026); 河南省重点研发与推广专项 (科技攻关)(192102210255); 河南省青年英才托举工程 (2020hytp013)

作者简介: 王雷光, 男, 教授, 博士生导师, 博士, 主要从事遥感图像分析与林业遥感应用方面的研究。

通讯作者: 郑晨, 男, 副教授, 博士生导师, 博士, 主要从事基于统计方法的遥感大数据智能解译研究。

**Abstract:** The interpretation of single remotely sensed data source may suffer from inaccurate boundary and low classification accuracy. The integration of hyperspectral and LiDAR data opens up the possibility to improve the classification performance. But, it is a challenge that how to appropriately integrate the considerable heterogeneity between the two types of data. In this paper, a conditional random field classification method was proposed to solve this problem by jointly taking both the heterogeneity of fused spectral-spatial-height features and co-occurrence of class labels into account. Firstly, the morphological features were extracted from two types of data respectively, and a graph model and training samples were jointly used to fuse the morphological features and spectral features. The obtained features were inputted into a support vector machine classifier to obtain the initial classification results with probabilistic outputs. Then, based on the fusion features, a local heterogeneity value was calculated to measure the essential difference of classes among pixels. Meanwhile, a class co-occurrence matrix, whose element calculated the spatial relationship between classes, was also obtained. Finally, a conditional random field framework was used to integrate the initial classification results, local heterogeneity information and the class co-occurrence matrix, and obtain the final classification results through inferencing two objective functions. In this process, by defining the weight between two neighboring pixel as a monotone decreasing function respect to the normalized Euclidean distance of the corresponding fused features, the object boundary could be regularized by giving a smaller weight to the class pairs with different labels and distinct features. Similarly, by giving a small weight to the class pairs with a strong spatial relationship, the purpose of maintaining the class pairs with stable spatial relations could be achieved. The method was validated with Houston and Gaofeng forest farm data sets. The overall accuracies of the proposed method reached to 94.00% and 92.84% respectively, and the "pepper and salt" phenomena of the initial classification results were significantly reduced. The result indicates the effectiveness of the proposed method.

**Key words:** hyperspectral; LiDAR; condition random field; feature fusion based on graph model; local heterogeneity; spatial co-occurrence

## 0 引言

遥感影像分类是遥感影像处理的研究热点和难点<sup>[1]</sup>。随着对地观测技术的不断发展,通过不同类型的传感器获取的遥感数据急剧增加。一方面,数据的增加对现有基于小样本的分类方法的效率和精度提出挑战;另一方面,使得利用不同类型、互补性强的观测数据进一步提升分类精度成为可能<sup>[2]</sup>。

高光谱影像具有光谱分辨率高、数据量大的特点,能精细的反映地物表面的光谱反射特性;LiDAR数据能够获得地物的立体结构信息。许多研究者<sup>[3]</sup>尝试利用两者的互补性,通过高度信息辅助光谱相似地物的识别,通过光谱信息辅助具有相同高度地物的识别,采用数据融合的方法解决单一数据源进行地物解译面临的挑战。

高光谱和 LiDAR 数据作为成像机理不同的异构数据,为实现高精度的场景解译,更适合采用特征级

和决策级的融合<sup>[4]</sup>。在特征提取阶段,由于高光谱影像存在数据的冗余,一般首先对其进行降维处理<sup>[5-6]</sup>,再从降维后的光学和 LiDAR 数据提取特征。然而,由于简单的特征叠加融合可能造成新的信息冗余,容易引发“维数灾难”,也难以反映不同特征对分类的贡献,融合的特征并不总是比使用单一特征源表现得更好。因此,通过特征选择或变换方法对特征集进行再次降维、利用对特征维度不敏感的分类器,都已成为提升分类精度的重要手段。文献 [7] 将高光谱影像和 LiDAR 数据提取的光谱、空间和高度特征通过拉普拉斯映射算法实现特征融合。文献 [8] 利用复核技术分别构造面向空间和光谱类特征的核函数,采用极限学习机实现分类。决策级融合通过将不同分类设置获得的分类结果加以综合实现。典型的决策融合方法有简单多数投票法<sup>[9]</sup>等。

此外,将初始分类结果与特征学习或区域分割算

法结合,以规整对象区域边界,并优化分类的思路,也获得了一定关注<sup>[7]</sup>。如参考文献<sup>[9]</sup>利用分割算法获得的区域对象对高光谱和 LiDAR 分类的结果进行后处理优化,以消除像素分类的“胡椒盐”噪声和规整地物对象边界。

条件随机场 (Conditional Random Field, CRF) 属于无向概率图模型。基于影像像素类别空间分布的局部平滑假设,CRF 可以有效平滑初始分类的噪声,也广泛应用于分类后处理中<sup>[9]</sup>。其技术关键在于有效保留地物的边界,防止分类结果过度平滑<sup>[9]</sup>。一般通过在空间能量项中引入边缘、局部光谱差异等度量指标<sup>[10-12]</sup>加以度量。然而,这类工作仍多针对多光谱或高光谱单一数据源,采用的是底层的统计特征,如何针对多源数据提出局部异质性指标,有效度量类别间的真实差异,仍有待深入研究。

上述研究表明,多源数据分类的精度提升和“胡椒盐”噪声的改善依赖于多阶段、多种处理策略的融合。有鉴于此,受上述多阶段融合思路的启发,同时为了有效表征场景中局部的异质性,保留分类优化结果中的对象边界,文中提出了一种顾及局部特征差异与全局类共生的 CRF 高光谱-LiDAR 融合分类方法。该方法的主要特点是:多源数据提取的光谱-空间-高度融合特征同时用于初始分类和描述地物局部空间异质性,并将全局类共生参数引入目标函数,综合特征融合和优化后处理技术,实现分类精度的提升和“胡椒盐”噪声的消除。

## 1 方法流程

图 1 为文中提出的 CRF 分类模型的流程。算法主要包含两个紧密联系的子模块。其一:基于光谱、空间、高度特征融合的初始分类。其目的是为了获得融合特征及初始概率分类结果。首先提取反映场景地物特点的光谱、空间及形态特征,形成高维特征向量,再基于训练样本和广义图模型求解映射函数,将特征映射到低维特征空间,将融合的低维特征输入概率支持向量机获得初始分类结果。其二:分类结果的 CRF 模型优化。映射获得的低维特征既用于初始分类和初始化特征场能量函数,又用于标记场能量函数中像素间的类别差异的度量。通过迭代求解特征场、

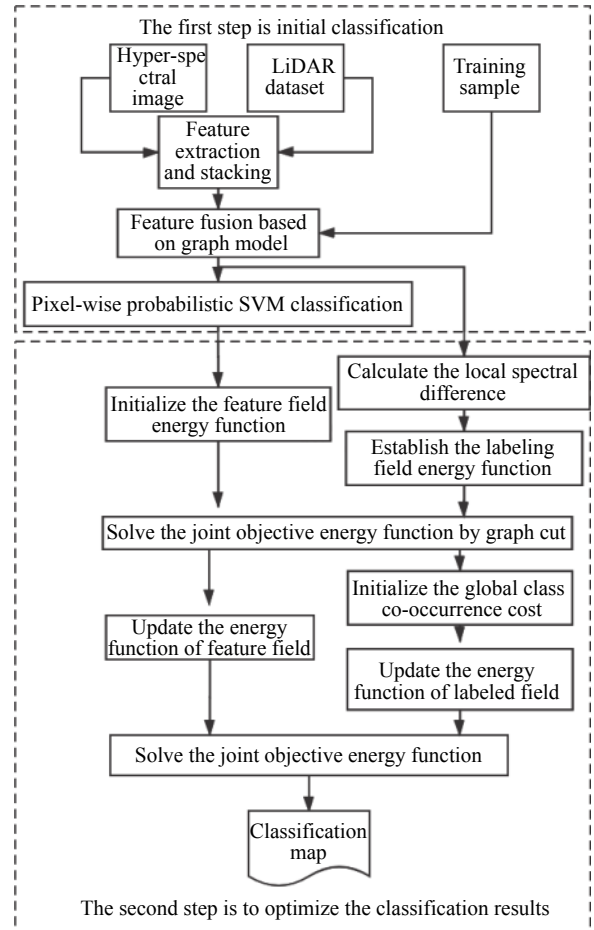


图 1 基于特征差异与类共生 CRF 方法的高光谱-LiDAR 分类流程  
Fig.1 Hyperspectral and LiDAR co-classification by CRF integrating feature dissimilarity and class co-occurrence

标记场联合目标函数获得初次优化结果及目标能量值;更新特征场能量函数,统计类别共生概率,更新标记场能量函数,对新的目标函数求解,获得最终分类结果。

### 1.1 标识说明及问题描述

文中:加粗的斜体大写字母代表集合或矩阵;斜体小写字母代表标量或列向量;字母的斜体下标代表集合或矩阵包含的列向量的索引;字母的正体上标单词缩写为不同类型的特征描述。

给定严格空间配准的高光谱影像和 LiDAR 生成的 DSM 栅格数据,即影像像素个数为  $n = n_1 \times n_2$ ,  $S$  表示影像所有  $n$  个位置的集合,  $\Omega = \{c_i\}_{i=1}^K$  表示类别标签集合,  $K$  为类别个数。高光谱和 LiDAR 的融合分类过程,即提取特征  $X = \{x_i\}_{i=1}^n$ ,并根据特征  $x_i$  赋予  $i$  像素为类别  $y_i \in \Omega$ ,并有分类结果  $Y = \{y_i\}_{i=1}^n$ 。

## 1.2 特征融合的初始分类

特征融合分类模块分为基于形态学的多源特征提取、基于图模型的特征融合<sup>[10]</sup>和基于概率支持向量机的分类三个主要步骤。

一般而言,高光谱影像可以提取光谱和空间特征,LiDAR 数据可以提取与高度相关的特征。为不失一般性,将获得的特征记为  $\mathbf{X}^{\text{Spe}} = \{x_i^{\text{Spe}}\}_{i=1}^n$ ,  $\mathbf{X}^{\text{Spa}} = \{x_i^{\text{Spa}}\}_{i=1}^n$  和  $\mathbf{X}^{\text{DSM}} = \{x_i^{\text{DSM}}\}_{i=1}^n$ 。其中,  $n$  为像素个数,  $x_i^{\text{Spe}}$ 、 $x_i^{\text{Spa}}$  和  $x_i^{\text{DSM}}$  分别代表第  $i$  个像素位置提取的光谱、空间和高程相关的三种特征列向量,其维数分别为  $d^{\text{Spe}}$ 、 $d^{\text{Spa}}$  和  $d^{\text{DSM}}$ 。

通过将三种特征进行堆叠,即可获得联合特征矩阵  $\mathbf{X} = [\mathbf{X}^{\text{Spe}}; \mathbf{X}^{\text{Spa}}; \mathbf{X}^{\text{DSM}}] = \{x_i\}_{i=1}^n \in \mathbb{R}^{D \times n}$ , 其中,特征总维数  $D = d^{\text{Spe}} + d^{\text{Spa}} + d^{\text{DSM}}$ , 并且  $x_i = [x_i^{\text{Spe}}; x_i^{\text{Spa}}; x_i^{\text{DSM}}] \in \mathbb{R}^{D \times 1}$ 。

特征矩阵  $\mathbf{X}$  仍具有较高维数,需要进一步约简。流形学习通过求出相应的嵌入映射,可以实现这一目的。拉普拉斯特征映射算法采用一个无向有权图描述一个流形,然后通过图的嵌入求解变换矩阵  $\mathbf{W}^*$ , 即

$$\mathbf{W}^* = \underset{\mathbf{W}}{\operatorname{argmin}} \left( \sum_{i,j=1}^n \| \mathbf{W}^T (x_i - x_j) \| C_{ij} \right) \quad (1)$$

式中:  $x_i \in \mathbf{X}$  代表图的节点;  $C_{ij}$  为节点邻接矩阵  $\mathbf{C}$  中  $x_i$  与  $x_j$  的邻接强度。采用与参考文献 [7,10] 类似的方式,  $C_{ij}$  也定义为二值矩阵。对  $x_i$  中的光谱、空间和高程三种子特征向量分别采用欧式距离寻找与其最近的  $k$  个样本,若  $j$  样本同时出现在三个样本集合中,则  $C_{ij} = 1$ , 否则  $C_{ij} = 0$ 。拉普拉斯特征映射算法具有计算复杂度低的特点,其详细求解过程可参考文献 [10]。

$$\mathbf{Z} = (\mathbf{W}^*)^T \mathbf{X} \quad (2)$$

利用公式 (2) 获得融合的特征后,采用概率支持向量机<sup>[13]</sup>即可获得所有像素的类别标记  $\mathbf{Y}^{\text{ini}} = \{y_i^{\text{ini}}\}_{i=1}^n$  和任意像素  $i$  属于给定类别  $y_i \in \Omega$  的概率  $l_{i,y_i}$ 。

## 1.3 基于条件随机场的特征融合与分类优化框架

上述特征融合分类中,分类器独立地预测每个像素的类别,所以无法避免结果中的“胡椒盐”噪声。条件随机场是给定特征变量条件下类别标记变量的马尔科夫随机场。经典条件随机场模型应用于后处理

优化<sup>[12]</sup>,通过类别标记变量的局部马氏性约束,可以有效地平滑局部的类别不一致现象,实现分类噪声的滤除。

像素分类结果的优化任务可以描述为:已知观测特征图像  $\mathbf{Z}$  和初始分类结果  $\mathbf{Y}^{\text{ini}}$ , 寻找一个地物类别标记随机场  $Y$  的最大后验估计  $\mathbf{Y}^*$  的问题,即

$$\mathbf{Y}^* = \operatorname{arg max}_Y (P(Y|\mathbf{Z})) \quad (3)$$

定义在以像元为节点  $V$ 、以像元 4-或 8-邻域为邻接关系的无向图  $G = (V, E)$  结构之上,如果给定观测特征  $\mathbf{Z}$  的观测随机场  $Y$  的条件概率分布  $P(Y|\mathbf{Z})$  为条件随机场,则对于任意一个节点  $i \in V$ , 随机变量  $y_i$  应满足:

$$P(y_i|\mathbf{Z}, y_j, i \neq j) = P(y_i|\mathbf{Z}, y_j, j \in N_i) \quad (4)$$

式中:  $i \neq j$  表示图  $G$  中除  $i$  外的所有节点;  $j \in N_i$  表示图  $G$  中与  $i$  邻接的节点集合,文中定义为  $i$  的 8 邻域位置。同时,根据 Hammersley-Clifford 定理,  $P(Y|\mathbf{Z})$  可以表示为一个 Gibbs 分布,即

$$P(Y|\mathbf{Z}) = \frac{1}{M(\mathbf{Z})} \exp \left( - \sum_{c \in C_G} \varphi_c(y_c|\mathbf{Z}) \right) \quad (5)$$

式中:  $C_G$  定义为所有一阶和二阶势团的集合;  $\varphi_c$  为团  $c$  的对应势函数; 规范化因子  $M(\mathbf{Z})$  保证  $P(Y|\mathbf{Z})$  构成概率分布。将公式 (5) 代入公式 (3), 并定义  $E(Y|\mathbf{Z}) = \sum_{c \in C_G} \varphi_c(y_c|\mathbf{Z})$  为 Gibbs 能量,有:

$$\mathbf{Y}^* = \underset{Y}{\operatorname{arg min}} \left( \sum_{c \in C_G} \varphi_c(y_c|\mathbf{Z}) \right) \quad (6)$$

即求解 Gibbs 能量  $\sum_{c \in C_G} \varphi_c(y_c|\mathbf{Z})$  最小时对应的像素类别标记序列。由于同一特征在所有像素位置都有定义,  $E(Y|\mathbf{Z})$  又可以重写成一阶势能函数  $E^{\text{Fea}}$  和二阶势能函数  $E^{\text{Lab}}$  的和,即

$$E(Y|\mathbf{Z}) = E^{\text{Fea}} + E^{\text{Lab}} \quad (7)$$

$$E^{\text{Fea}} = \sum_{i \in S} \varphi_m(y_i, z_i) \quad (8)$$

$$E^{\text{Lab}} = \beta \sum_{i \in V} \sum_{j \in N_i} \varphi_n(y_i, y_j, \mathbf{Z}) \quad (9)$$

式中:一元势函数  $\varphi_m(y_i, z_i)$  描述了给定像素位置  $i \in S$  特征  $z_i$  时,  $i$  属于  $y$  类别的概率;二元势函数  $\varphi_n(y_i, y_j, \mathbf{Z})$  描述了给定特征  $\mathbf{Z}$  时,像素  $i$  与邻域像素  $j$  的类别标记关系;超参数  $\beta$  用于调节二阶势能函数的贡献,  $\beta$  取值越大,分类图的平滑效果越强。

为获得目标函数(公式(6))可求解的数学形式,首先将公式(8)中一元势函数定为:

$$\varphi_m(y_i, z_i) = -\log(l_{i,y_i}) \forall i \in S \quad (10)$$

$l_{i,y_i}$ 类别概率计算时融合了光谱、空间和高程多种特征,因此相比仅采用光谱特征获得的分类结果将更为精确<sup>[12]</sup>。

在经典 CRF 模型中,公式(9)中二元势函数  $\varphi_n(y_i, y_j, \mathbf{Z})$  定义为:

$$\varphi_n(y_i, y_j, \mathbf{Z}) = 1 - \delta(y_i, y_j) \quad \forall i \in S, j \in N_i \quad (11)$$

式中:当  $y_i = y_j$  时,  $\delta(y_i, y_j) = 1$ , 否则  $\delta(y_i, y_j) = 0$ 。由于目标函数(公式(6))计算整体能量的极小值,因此二元势函数(公式(11))倾向于为中心像素  $i$  和邻域像素  $j$  赋予相同的标识。

文中在公式(11)中引入  $f(\Delta_{i,j})$ , 并将二元势函数  $\varphi_n(y_i, y_j, \mathbf{Z})$  重新定义为:

$$\varphi_n(y_i, y_j, \mathbf{Z}) = f(\Delta_{i,j}) [1 - \delta(y_i, y_j)] \quad \forall i \in S, j \in N_i \quad (12)$$

式中:  $\Delta_{i,j}$  度量了中心像素  $i$  和邻域像素  $j$  的特征差异性,  $f(\Delta_{i,j})$  为  $\Delta_{i,j}$  的单调递减函数。公式(12)表明:当局部像素的异质性指标  $\Delta_{i,j}$  较小时,即相邻的像素具有类似的特征时,像素有更大的可能性属于同类;反之,像素具有较大的可能性不属于同类。文中,  $f(\Delta_{i,j})$  定义为归一化欧式距离  $\Delta_{i,j}$  的指数函数,即

$$f(\Delta_{i,j}) = \exp(-|\Delta_{i,j}|) \quad (13)$$

$$\Delta_{i,j} = \sqrt{\sum_b \left( \frac{z_{i,b}}{z_b} - \frac{z_{j,b}}{z_b} \right)^2} \quad (14)$$

式中:  $z_{i,b}$  和  $z_b$  分别代表融合特征矩阵  $\mathbf{Z}$  中  $i$  位置上第  $b$  维的特征值及第  $b$  维特征的平均值。将公式(7)~(14)整合,代入公式(6),有目标函数:

$$Y^* = \operatorname{argmin}_Y \sum_{i \in S} \left[ -\log(l_{i,y_i}) + \beta \sum_{j \in N_i} \exp \left( -\sqrt{\sum_b \left( \frac{z_{i,b}}{z_b} - \frac{z_{j,b}}{z_b} \right)^2} \right) [1 - \delta(y_i, y_j)] \right] \quad (15)$$

相比采用原始数据的像素梯度作为局部异质性度量指标的方法<sup>[11]</sup>,文中采用的特征是空间、光谱和高程特征的融合。由于场景中对象的真正边界往往是多种局部视觉线索的综合,因此融合特征可以更有

效地描述不同地物对象的差异。采用欧式距离作为特征差异的度量,主要考虑到采用的基于流形学习的特征融合和降维后,特征空间是一种局部近似为欧式空间的特殊拓扑空间。另外,各特征维逐一归一化的处理方式有效避免了不同特征类型可能存在的量纲不一致的问题。

另一方面,目标函数(公式(15))在二元势函数中仅考虑局部的特征差异性,没有考虑类别间的空间依赖特性,往往会在对象边界上出现细碎颗粒状噪声<sup>[12]</sup>。因此,在以上获得初次优化结果  $Y^*$  的基础上,考虑将类别的共生约束引入二元势函数,对初次优化结果进行再次优化,即定义目标函数为:

$$Y^* = \operatorname{argmin}_Y \sum_{i \in S} \left[ -\log(l_{i,y_i}) + \beta \sum_{d \in \{1,2,\dots,8\}} [1 - g_d(y_i, y_j)] \times [1 - \delta(y_i, y_j)] \right] \quad (16)$$

式中:  $d \in \{1,2,\dots,8\}$  代表中心像素与 8-邻域像素的 8 种位置关系;  $y_j$  为像素位置  $i$  在  $d$  方向上的邻域像素的标记;  $g_d(y_i, y_j)$  为  $y_i$  与  $y_j$  类别在  $d$  方向上的共生概率。公式(16)表明:  $g_d(y_i, y_j)$  为各项异性的函数,不同方向上相同类别对的共生情况也有所不同,目标函数倾向于保留具有更强空间联系的标记对。

最后,通过顺序求解目标函数(公式(15)和(16)),顺次对初始结果进行两次优化,即可获得最终的优化分类结果。与参考文献[12]类似,首先采用 Graph Cut 算法求解目标函数(公式(15)),然后基于优化结果统计初始化类别共生概率的  $g_d(y_i, y_j)$ ,再借助 ICM 算法迭代优化(公式(16)),获得最终分类结果。

## 2 数据与实验设置

### 2.1 实验数据

为验证所提出的融合分类算法的有效性,采用城区和林区不同场景的数据进行验证实验。

#### 2.1.1 城区数据集

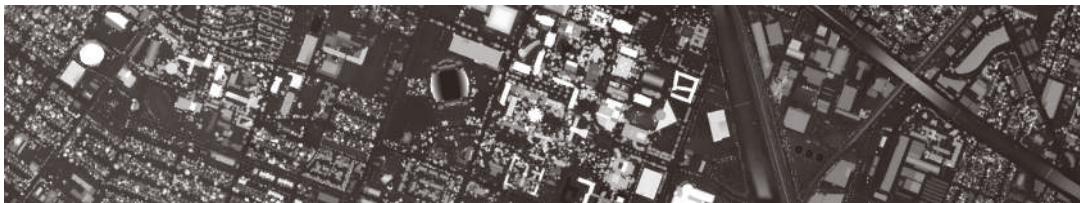
该数据集(下文简称 Houston 数据集)<sup>[10]</sup>,在休斯顿大学校园和邻近城市地区获取,包括 1 景高光光谱影像和 1 景由 LiDAR 点云生成的数字表面模型(DSM)栅格数据。其中,高光光谱影像的光谱范围为 380~

1050 nm, 包含 144 个波段, 空间大小为 349×1 905, 空间分辨率为 2.5 m, 共包含 15 个地类。图 2 为高光谱影像(真彩色模式)、DSM, 训练和测试样本的数量见表 1(a)。在高光谱影像的获取过程中, 由于云层的

遮挡导致部分研究区域位于阴影之下, 在该区域也没有选择训练样本, 但部分测试样本位于阴影区域, 增大了实验的挑战性。



(a) 高光谱真彩色影像  
(a) Hyperspectral true color image



(b) LiDAR 生成的 DSM 影像  
(b) DSM images generated by LiDAR

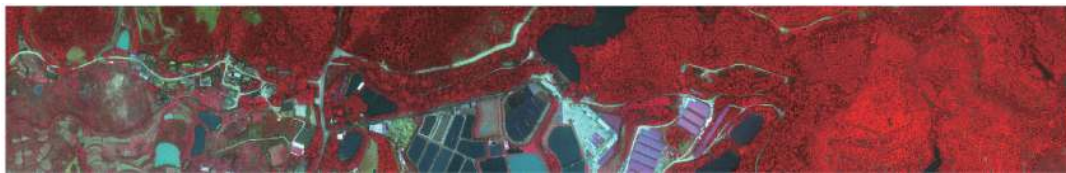
图 2 休斯顿数据集

Fig.2 Houston data set

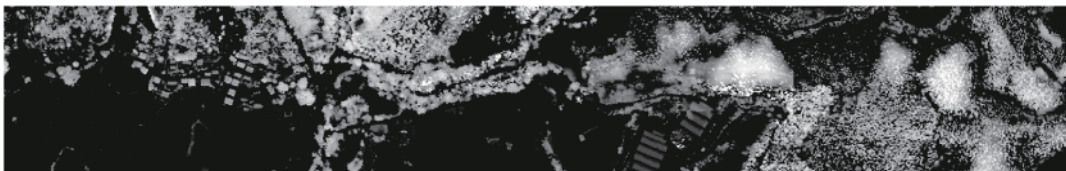
### 2.1.2 林区数据集

该实验数据集在广西壮族自治区南宁市高峰林场地区获取<sup>[4]</sup>, 经过几何校正、拼接、配准等预处理后形成的数据集(下文简称高峰林场数据集)也包括 1 景高光谱和 1 景 DSM 栅格数据。其中, 高光谱影像

包含 125 个波段, 空间大小为 351×2281, 空间分辨率为 1 m, 共 15 个类别。结合野外调查, 获得该场景的训练和测试样本, 两者没有交叉。图 3 为该高光谱影像和 DSM, 训练和测试样本的数量和图例信息见表 1(b)。



(a) 高光谱假彩色影像  
(a) Hyperspectral false color image

















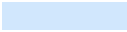















(b) LiDAR 生成的 DSM 影像  
(b) DSM images generated by LiDAR

图 3 高峰林场数据集

Fig.3 Gaofeng forest farm data set

表 1 实验样本类别及数量

Tab.1 Class names and their numbers

(a) 休斯顿 (a) Houston			(b) 高峰林场 (b) Gaofeng forest farm		
Class name	Number of training/testing samples/pixel	Sample color	Class name	Number of training/testing samples/pixel	Sample color
Healthy grass	198/1 053		Eucalyptus	193/315	
Stressed grass	190/1 064		Road	74/106	
Synthetic grass	192/505		Tilia tuan	40/52	
Trees	188/1 056		Cultivated land	95/127	
Soil	186/1 056		Acacia crassicarpa benth	208/308	
Water	182/143		Wasteland	16/20	
Residential	196/1 072		Michelia macclurei dandy	69/95	
Commercial	191/1 053		Building	165/251	
Road	193/1 059		Other broad leaved forests	184/275	
Highway	191/1 036		Pinus massoniana lamb	214/300	
Railway	181/1 054		Cunninghamia lanceolata	34/47	
Parking Lot 1	192/1 041		Water	390/562	
Parking Lot 2	184/285		Mixed shrub forest	53/84	
Tennis court	181/247		Bamboo	21/34	
Running track	187/473		Grassland	23/20	

2.2 实验设置

拓展形态剖面特征 (EMP) 通过采用不同形状、依次增大的结构元素对波段进行序列开闭运算,可以有效提取场景中的对象大小和形状空间结构信息,广泛用于空间信息的提取。考虑到研究场景中,道路、房屋、草地、树林、树冠等往往具有规整的边界或规则的形状,基于序列尺寸的圆形和线形结构元素,对 DSM 和高光谱波段提取多尺度特征,具体参数设置参考文献 [7]。

首先,采用序列尺寸的圆形和线形结构元素,分别对 DSM 波段和高光谱影像提取形态剖面特征,然后,对于共计 210 维的光谱、空间和高度特征进行基于图模型的特征融合(见 1.2 节)后,通过不同维数融合特征的精度比较(见 3.1 节)确定最终保留的特征维数。融合后的特征基于径向基函数,采用概率支持向量机<sup>[13]</sup>进行分类。其中的模型参数由 10-折交叉验证获取。

在分类结果优化阶段,涉及目标函数(公式 (15) 和 (16))中两个超参数  $\beta$  的选择问题,默认设置两个超参数取值一致,并通过实验确定较优的参数值。结果评价采用定量精度评价和目视解译结合的方式。分

类结果的定量评价采用基于混淆矩阵的整体精度 (OA)、Kappa 系数以及相关的类别生产者精度和用户精度等指标。

3 实验结果与分析

基于两个数据集,从特征提取阶段中保留特征维数的影响、超参数  $\beta$  的取值、初始分类方法的影响、与主流方法的对比等方面评估所提出的方法的优劣。

3.1 特征维数的影响

首先对图模型融合后 Houston 数据集整体精度 OA 随保留特征维数 (1~32) 变化的情况进行分析。实验结果表明:随着特征维数的增加,GGF 和 GGF-CRF 的分类精度都在逐渐提升;且对于保留不同维数获得的初始分类结果采用后处理算法后精度均有所提升。高峰林场数据集也有相似的表现。当保留特征维数超过 24 维后,分类精度开始逐渐下降。保留特征维数在 20~32 之间时,OA 均超过 90%;且当融合特征维数取 24 维时精度最高。因此,后续的保留特征维数均设置为 24 维。

3.2 超参数  $\beta$  取值的影响

表 2 为不同  $\beta$  取值下, Houston 总体分类精度

(OA)、Kappa 系数及平均精度 (AA) 的值。当  $\beta$  在一定范围内取值时, 3 个指标均比较稳定, 表现出算法对于  $\beta$  取值的鲁棒性。可以认为  $\beta$  取值在 [0.5, 4.5] 范围时, 算法对于超参数的取值不敏感。这一结论也

被其他研究<sup>[9]</sup>和文中的高峰林场数据所证实。因此, 后续实验中, Houston 数据集  $\beta$  取值固定为 1, 高峰林场数据集  $\beta$  取值固定为 2.5。

表 2  $\beta$  取值对最终分类精度的影响 (Houston)

Tab.2 Influence of different  $\beta$  values on the final classification accuracy (Houston)

Precision	$\beta$								
	0.5	1	1.5	2	2.5	3	3.5	4	4.5
OA	93.99%	<b>94.00%</b>	93.93%	93.88%	93.89%	93.86%	93.85	93.84%	93.83%
Kappa	0.935	<b>0.935</b>	0.934	0.933	0.934	0.933	0.933	0.933	0.933
AA	93.47%	<b>93.42%</b>	93.19%	93.07%	91.30%	93.06%	93.04%	93.04%	93.03%

### 3.3 Houston 数据集

#### 3.3.1 初始分类及不同优化方法对比

为全面评价提出方法的优势和不足, 设置不同特征及组合、不同特征融合方法以及不同 CRF 后处理方法, 共 7 种方法进行对比分析。具体而言:

- (1)  $F^{Spe}$ : 基于光谱特征的 SVM 分类;
- (2)  $F^{DSM}$ : 基于 DSM 提取的形态学的 SVM 分类;
- (3)  $F^{Spe}+F^{Spa}$ : 基于形态学空间特征及光谱主成分结合的分类;
- (4)  $F^{Spe}+F^{DSM}$ : 光谱-高度特征叠加融合分类;
- (5) GGF: 空间-光谱-高度特征经过图模型融合后进行分类, 即文中采用的初始分类方法;
- (6) GGF\_CRF1: 对于 GGF 的输出结果仅考虑融合特征局部差异性的 CRF 后处理方法;
- (7) GGF-CRF: 文中提出的方法。

上述方法中, (1)~(5) 属于像素级分类方法 (第一组); (6)~(7) 属于基于 CRF 的分类优化方法 (第二组)。前者的对比可以评价各类特征的重要性及特征融合方法的有效性; 后者的对比可以评价提出优化模型的有效性。

表 3(a) 展示了两组七种方法在处理 Houston 数据集的生产者精度对比。整体而言, 文中提出的方法 (GGF-CRF) 在大部分类别上获得了最优或次优的结果, 且 OA、AA 和 Kappa 均为最高。合成草地、裸土、网球场及跑道的分类精度达到 100%, 且大部分类别的分类精度均在 90% 以上。

在像素级分类方法中, 单一特征的分类结果  $F^{Spe}$  和  $F^{DSM}$  的整体精度均较低, 分别为 81.98% 和

60.48, 体现出光谱特征对于遥感分类的基础性作用, 也体现出高度特征在区分高度差异不大的地物上的局限性。但阴影区内外地物光谱特征的差异和阴影区训练样本的缺失导致  $F^{Spe}$  在阴影区的表现 (图 4(c)) 不佳。

光谱-空间联合特征 ( $F^{Spe}+F^{Spa}$ ) 和光谱-高度联合特征 ( $F^{Spe}+F^{DSM}$ ) 的应用将分类精度分别提升到 85.12% 和 85.14%, 表明了光谱与空间、高度特征的互补性。而在  $F^{Spe}+F^{Spa}$  简单叠加的特征融合方式建立的特征空间中, 阴影区域内外相同地类的样本差异仍然较大, 导致该部分分类结果不佳 (图 4(e))。应用流形学习的广义图融合算法 GGF 后, 由于引入了训练样本在联合特征空间的相似性信息, 分类精度进一步提升到了 93.34%, 表明该算法不仅能降低特征的维数, 还能有效去除波段间的冗余, 使得样本间的特征差异更适合采用欧式距离度量。如图 4(g) 所示, 阴影区的分类结果明显改善, 特别是跨越阴影和非阴影区域的高速公路 (高架桥) 分类最为完整。

但如图 5(a) 所示, 由于上述 5 种分类过程均以像素为单位, 特征提取即使考虑了局部邻域的信息, 分类结果中仍然存在不同程度的“胡椒盐”噪声, 地物边界不清晰, 仍需要借助后处理手段进行平滑。

在 GGF 方法分类的基础上, 两种后处理算法结果各类别和整体精度均略有提高, GGF-CRF 的 OA、AA 和 Kappa 系数均最高, 但数值上 GGF-CRF1 和 GGF-CRF 方法精度差异不大。图 5 为 GGF 分类及两种优化方法的结果。直观上, 图 5(b) 分类图的视觉差异要远大于精度统计反映的差异, 造成这一现象的原



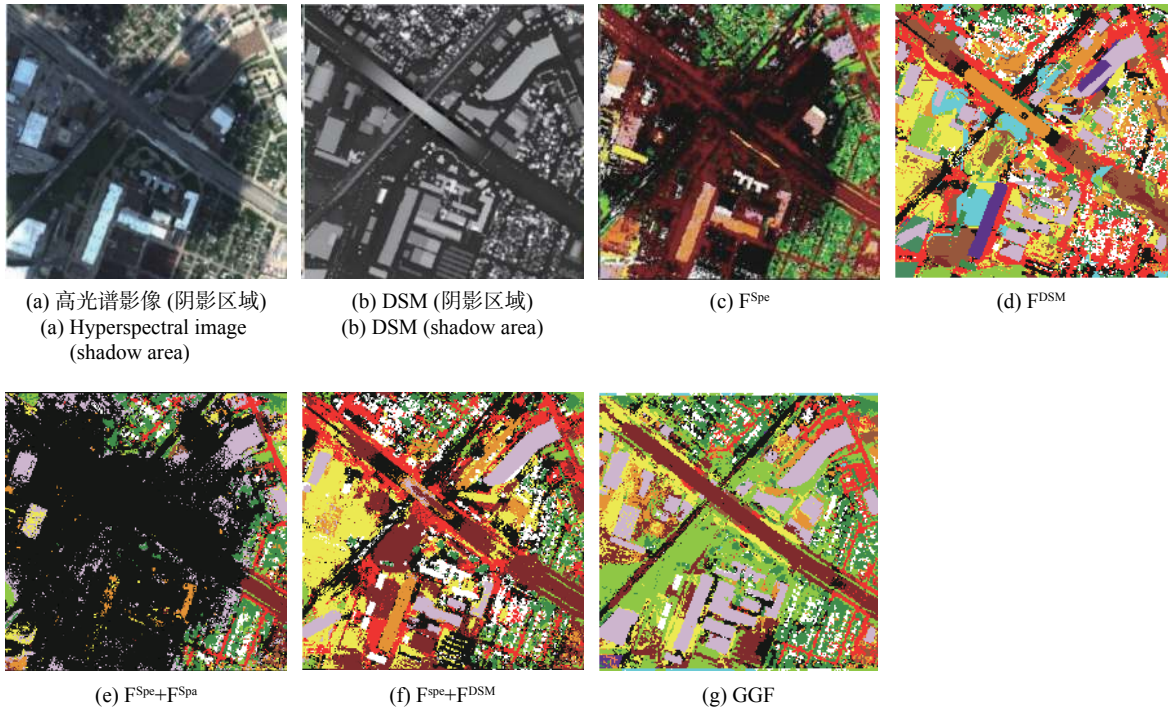


图 4 局部阴影区域不同特征融合设置的分类结果对比 (Houston 数据集)

Fig.4 Classification results obtained from different feature fusion settings in the shaded area (Houston data set)

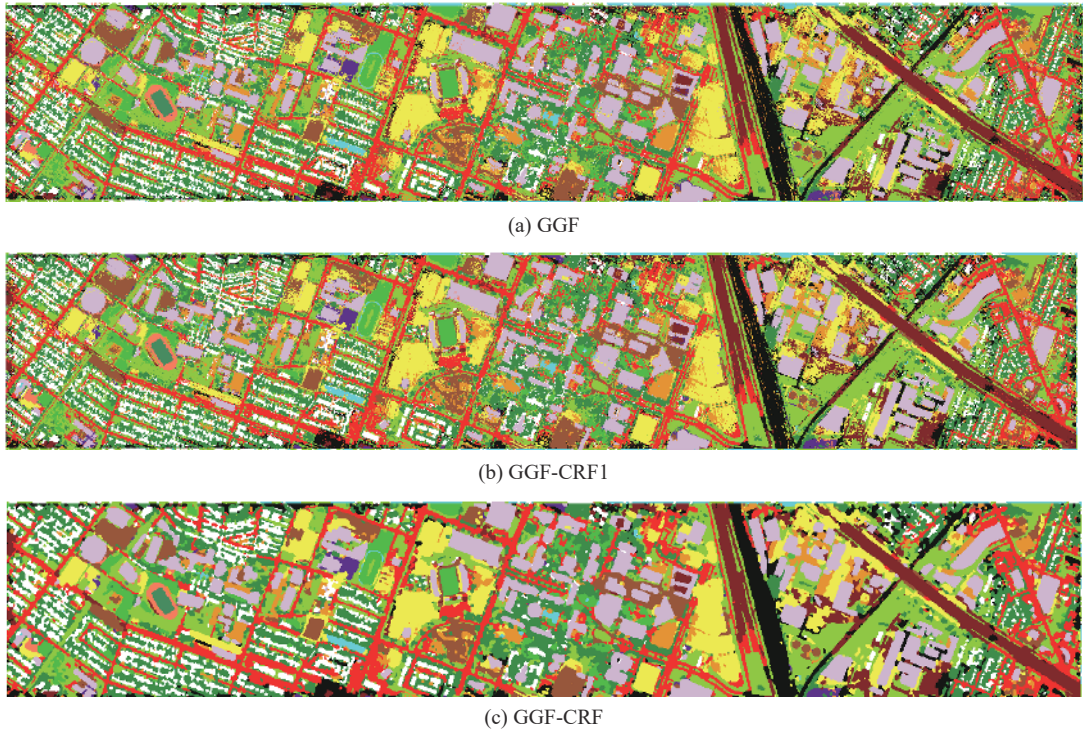


图 5 初始分类和采用不同 CRF 方法优化的结果 (Houston 数据集)

Fig.5 Initial classification map and results optimized by different CRF methods (Houston data set)

因是缺少位于不同对象邻接区域的测试样本。从图 5(b)~(c)中还可以看出:两种优化方法在不同程度上减少了 GGF 中图 5(a)的“胡椒盐”噪声。GGF-

CRF1 方法考虑了相邻像素间融合特征异质性值,不仅能准确反映不同地物类别之间的边界,并且对于较小的地物类别,在分类结果中也能被保留下来。

GGF-CRF 在 GGF-CRF1 的结果上进行再优化, 通过考虑类别间的空间共生关系, 能有效改善 GGF-CRF1 方法中的噪声。

根据公式 (16), 当相邻像素类别相同时,  $1-\delta(y_i, y_j)$  取 0; 当相邻像素类别不同时, 较强空间关系对

应的  $1-g_d(y_i, y_j)$  较小。求解公式 (16) 的极小值, 倾向于保留邻域空间关系更稳固的类别对。因此, 如表 3 所示, 从 GGF-CRF1 到 GGF-CRF, 大多数类别的精度均有所提升, 特别是居民区、商业区、高速公路等人工地物类别。

表 3 七种方法用于不同数据集分类结果的生产者精度对比

Tab.3 Producer's accuracy comparison of seven classification methods for different data sets

Category	Pixel level classification method				CRF classification optimization method		
	F <sup>Spe</sup>	F <sup>DSM</sup>	F <sup>Spe</sup> +F <sup>Spa</sup>	F <sup>Spe</sup> +F <sup>DSM</sup>	GGF	GGF_CRF1	GGF-CRF
	(a) 休斯顿数据集 (a) Houston data set						
Healthy grass	82.34	24.88	55.69	55.89	81.67	82.43	<b>83.1</b>
Stressed grass	83.36	55.92	84.40	84.49	99.34	99.62	<b>99.81</b>
Synthetic grass	<b>100</b>	91.88	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>
Trees	93.37	67.23	91.57	98.11	99.24	99.24	<b>99.62</b>
Soil	98.30	76.80	<b>100</b>	99.15	<b>100</b>	<b>100</b>	<b>100</b>
Water	91.61	80.42	<b>99.30</b>	96.50	95.10	95.10	94.41
Residential	76.59	71.74	82.84	91.32	92.35	92.26	<b>93.47</b>
Commercial	56.51	61.92	53.09	52.42	94.59	94.78	<b>95.73</b>
Road	66.57	51.37	79.04	83.95	<b>86.02</b>	85.93	85.74
Highway	72.39	53.86	68.15	79.92	93.24	93.63	<b>94.98</b>
Railway	92.88	83.97	<b>97.34</b>	87.76	90.70	90.80	90.61
Parking Lot 1	78.58	60.71	<b>97.70</b>	79.63	94.24	94.43	<b>97.41</b>
Parking Lot 2	72.98	57.19	<b>81.05</b>	<b>74.04</b>	72.28	71.93	66.67
Tennis Court	98.79	97.17	<b>100</b>	98.79	<b>100</b>	<b>100</b>	<b>100</b>
Running Track	98.31	28.96	98.52	97.67	99.37	99.37	<b>99.79</b>
OA	81.98%	60.48%	85.12%	85.14%	93.34%	93.47%	<b>94.00%</b>
AA	84.17%	64.27%	85.91%	85.31%	93.21%	93.30%	<b>93.42%</b>
Kappa	0.805	0.597	0.839	0.839	0.928	0.929	<b>0.935</b>
(b) 高峰林场数据集 (b) Gaofeng forest farm data set							
Category	Pixel level classification method				CRF classification optimization method		
	F <sup>Spe</sup>	F <sup>DSM</sup>	F <sup>Spe</sup> +F <sup>Spa</sup>	F <sup>Spe</sup> +F <sup>DSM</sup>	GGF	GGF_CRF1	GGF-CRF
Eucalyptus	73.65	60.63	90.79	77.46	86.67	96.82	<b>97.14</b>
Road	48.11	52.83	<b>90.57</b>	66.98	74.53	73.50	<b>73.58</b>
Tilia tuan	5.77	46.15	<b>59.62</b>	53.85	25	32.69	32.69
Cultivated land	83.46	98.43	<b>100</b>	96.85	<b>100</b>	<b>100</b>	<b>100</b>
Acacia crassicarpa benth	71.75	88.31	97.08	87.66	90.91	<b>97.73</b>	<b>97.73</b>
Wasteland	80	55	95	90	95	<b>100</b>	<b>100</b>
Michelia macclurei dandy	31.58	55.79	75.79	70.53	67.37	83.16	<b>84.24</b>
Building	83.27	84.06	96.41	92.83	<b>98.01</b>	97.21	97.21
Other broad leaved forests	70.91	66.55	<b>96.36</b>	65.82	83.27	85.45	85.82
Pinus massoniana lamb	73.67	92.67	92.00	85.67	89.00	96.67	<b>97.00</b>
Cunninghamia lanceolata	12.77	68.09	<b>95.74</b>	65.96	78.72	<b>95.74</b>	<b>95.74</b>
Water	99.82	98.22	<b>100</b>	99.64	<b>100</b>	<b>100</b>	<b>100</b>
Mixed shrub forest	2.38	51.19	73.81	67.86	73.81	<b>88.1</b>	<b>88.1</b>
Bamboo	0	2.94	17.65	<b>17.65</b>	0	0	0
Grassland	10.00	40.00	55.00	85.00	95.00	100	<b>100</b>
OA	71.46%	78.58%	92.41%	83.32%	87.71%	92.37%	<b>92.84%</b>
AA	49.81%	64.06%	82.39%	74.92%	77.15%	83.14%	<b>83.28%</b>
Kappa	0.674	0.756	0.914	0.811	0.860	0.913	<b>0.919</b>

3.3.2 与其他融合分类方法的对比

表 4 所示为文中提出的方法与主流融合方法的结果对比。其中, Deep fusion<sup>[15]</sup> 首先使用卷积神经网络方法将光谱数据和激光雷达数据映射为成高度抽象的特征向量, 再使用全连接神经网络来融合光谱和激光雷达数据特征, 通过网络来训练学习这两种特征的非线性组合。HyMCKs<sup>[8]</sup> 利用复合核来实现异构数

据地融合分类。多级融合<sup>[4]</sup> 方法针对特定地物类型的特点设计融合准则, 使用投票方法对特征级融合结果进行决策级融合, 然后再利用条件随机场对分类结果进行后处理。EC-CRF<sup>[11]</sup> 方法利用 Sobel 算子分别对高光谱和 LiDAR 进行边缘提取, 并将两种边缘进行组合, 根据梯度自适应地控制一个可变加权系数并输入 CRF 模型中。

表 4 不同融合方法的分类精度对比 (Houston 数据集)

Tab.4 Comparison of classification accuracy of different methods on Houston data set

Precision	Deep fusion <sup>[15]</sup>	HyMCKs <sup>[8]</sup>	Multi level fusion method <sup>[4]</sup>	EC-CRF <sup>[11]</sup>	GGF-CRF
OA	91.32%	90.33%	93.22%	91.70%	<b>94.00%</b>
Kappa	0.9057	0.8949	0.930	0.907	<b>0.935</b>

从表 4 中可以看出, 文中的方法在总体精度和 Kappa 值上都处于较高的水平。相比于其他融合算法, 该算法能在高光谱和 LiDAR 数据融合分类上获得较好的效果。

3.4 高峰林场数据集

表 3(b) 展示了上述两组七种方法在高峰林场数据集分类结果的生产者精度对比。整体而言, 文中提出的方法 (GGF-CRF) 在大部分类别上获得了最优或次优的结果, 且 OA、AA 和 Kappa 均为最高。耕地、荒地、水体及草地的分类精度达到 100%, 大部分类别的分类精度均在 80% 以上。同时, 目视对比, 该数据集的分类结果也具有与图 5 类似的表现。

因此, 与 Houston 数据集类似的实验结果, 实验证实了该算法在提升分类精度和改善“胡椒盐”现象的有效性。

3.5 算法效率分析

在 CPU Intel i7 2.50 GHz、内存 64 GB、Windows 7 操作系统、Matlab 2019b 环境下测试了算法耗时。在第一特征融合及初始分类步骤中, 由于初始分类耗时因特征参数、样本数量和分类器设置而异。统计未包含形态特征提取和 SVM 交叉验证步骤的时间。

Houston 整体处理时间为 220.05 s, 高峰林场整体处理时间为 240.45 s, 第二优化步骤比第一特征融合及初始分类步骤需要更多的计算时间。其中, 目标函数 (公式 (15) 和 (16)) 的求解占整个处理时间的约 70%。第一步处理中, 形态特征提取、拉普拉斯特征映射和 SVM 参数交叉寻优占用超过 80% 的处理时间。

4 结 论

为了充分利用高光谱影像与 LiDAR 数据的互补性信息, 达到提升分类精度和保留地类边界的目的, 文中提出了一种特征融合与条件随机场模型耦合的分类方法。该方法对于 Houston 高光谱-LiDAR 实验数据集取得了良好的分类效果。

研究结论主要有以下三点: (1) 拓展形态滤波特征对于高分辨率影像中的空间信息具有良好的描述能力, 结合流形学习的降维方法能获得较高的分类精度。但因其仍然属于像素级的分类方法, 无法从根本上避免分类结果中的细碎噪声。(2) 以融合的特征描述条件随机场中二阶势团中的邻域像素的异质性, 有效避免了传统随机场模型的标记过度平滑问题; 类共生代价函数的引入进一步提升了分类精度。(3) 采用融合的特征及概率分类结果初始化条件随机场模型特征场能量, 同时将融合特征引入标记场建模, 通过特征、决策的协同融合, 可以实现地物对象的高精度分类和分类噪声的改善。

下一步工作中, 拟将改进该方法在林区场景中树种分类上的应用, 以及深度特征与随机场模型的结合。

**致谢** 感谢国家重点研发项目“人工林资源监测关键技术研究 (2017YFD0600900)”提供文中使用的高峰林场数据集; 感谢休斯顿大学高光谱图像分析组和国家航空激光测绘中心 (NCALM) 提供文中使用的 Houston 数据集。

## 参考文献:

- [1] Ghamisi P, Rasti B, Yokoya N, et al. Multisource and multitemporal data fusion in remote sensing: A comprehensive review of the state of the art [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2019, 7(1): 6-39.
- [2] Muram D, Prasad S, Pacific F, et al. Challenges and opportunities of multimodality and data fusion in remote sensing [J]. *Proceedings of the IEEE*, 2015, 103(9): 1585-1601.
- [3] Rasti B, Ghamisi P, Gloaguen R. Hyperspectral and LiDAR fusion using extinction profiles and total variation component analysis [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(7): 3997-4007.
- [4] Cao Qiong, Ma Ailong, Zhong Yanfei, et al. Hyperspectral-LiDAR multi-level fusion urban land cover classification [J]. *National Remote Sensing Bulletin*, 2019, 23(5): 892-903. (in Chinese)
- [5] Shi Guojun. Infrared image target recognition method based on joint characterization of depth feature [J]. *Infrared and Laser Engineering*, 2021, 50(3): 20200399. (in Chinese)
- [6] Hou Banghuan, Yao Minli, Jia Weimin, et al. Hyperspectral image classification based on spatial structure preserving [J]. *Infrared and Laser Engineering*, 2017, 46(12): 1228001. (in Chinese)
- [7] Liao W, Pižurica A, Bellens R, et al. Generalized graph-based fusion of hyperspectral and lidar data using morphological features [J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(3): 552-556.
- [8] Ghamisi P, Rasti B, Benediktsson J A. Multisensor composite kernels based on extreme learning machines [J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(2): 196-200.
- [9] Huang X, Lu Q, Zhang L, et al. New postprocessing methods for remote sensing image classification: A systematic study [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52(11): 7140-7159.
- [10] Debes C, Merentitis A, Heremans R, et al. Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2014, 7(6): 2405-2418.
- [11] Ni L, Gao L, Li S, et al. Edge-constrained Markov random field classification by integrating hyperspectral image with LiDAR data over urban areas [J]. *Journal of Applied Remote Sensing*, 2014, 8(1): 085089.
- [12] Wang L, Huang X, Zheng C, et al. A Markov random field integrating spectral dissimilarity and class co-occurrence dependency for remote sensing image classification optimization [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2017, 128: 223-239.
- [13] Chang C C, Lin C J. LIBSVM: A library for support vector machines [J]. *ACM Transactions on Intelligent Systems and Technology*, 2011, 2(3): 27.
- [14] Feng B, Zhang C, Zhang W, et al. Analyzing the role of spatial features when cooperating hyperspectral and LiDAR data for the tree species classification in a subtropical plantation forest area [J]. *Journal of Applied Remote Sensing*, 2020, 14(2): 022213.
- [15] Cheng Y, Li C, Ghamisp P, et al. Deep fusion of remote sensing data for accurate classification [J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(8): 1253-1257.