

# 基于 Point-Net 的多源融合点云地物分类方法

释小松\*, 程英蕾, 薛豆豆, 秦先祥

空军工程大学信息与导航学院, 陕西 西安 710077

**摘要** 为了提高城区机载激光雷达点云数据地物分类的分类精度,提出了一种基于 Point-Net 网络的多源融合点云地物分类方法。点云在地物三维特征表示上具有优势,而遥感影像包含丰富的光谱信息,因此设计了一种点云与遥感影像的配准融合方法,综合利用两种数据的优势。针对 Point-Net 网络存在缺少邻域信息的问题,提出一种针对融合点云数据的多尺度 Point-Net 分类模型,实现对融合点云数据的有效分类。利用城区点云数据验证本文算法,通过分析分类精度和分类时间对分类效果进行评价。结果证明:相比其他算法,本文算法有效提高了点云数据分类效果,实现了对城区点云数据的有效分类。

**关键词** 图像处理; 点云数据; 遥感影像; 数据融合; Point-Net; 地物分类

中图分类号 P237

文献标志码 A

doi: 10.3788/LOP57.081019

## Object Classification Method for Multi-Source Fusion Point Clouds Based on Point-Net

Shi Xiaosong\*, Cheng Yinglei, Xue Doudou, Qin Xianxiang

Information and Navigation College, Air Force Engineering University, Xi'an, Shaanxi 710077, China

**Abstract** In order to improve the accuracy of object classification of point cloud data from airborne LiDAR, an object classification method for multi-source fusion point cloud data based on Point-Net is proposed. Point clouds can effectively represent three-dimensional features of objects, and remote-sensing images contain detailed spectral information. Therefore, a registration and fusion method for point cloud data and remote sensing images is designed to comprehensively utilize their advantages. Meanwhile, considering the lack of neighborhood information in Point-Net, a multi-scale Point-Net classification model for fusion point clouds is also proposed to realize effective classification of fusion point cloud data. The proposed algorithm is verified with point cloud data from urban regions and the classification effect is evaluated by analyzing the classification accuracy and time. Results show that, compared with other methods, the proposed method can effectively improve the classification accuracy of point cloud data, and achieve effective classification of point cloud data in urban areas.

**Key words** image processing; point cloud data; remote-sensing image; data fusion; Point-Net; object classification

**OCIS codes** 100.4996; 280.3420; 280.3640

## 1 引言

一直以来,遥感影像数据作为地球空间信息和遥感测绘等领域的主要数据源被国内外研究者深入研究,其中基于遥感影像数据的城市地物识别分类研究已取得了很多成果<sup>[1]</sup>。遥感影像具有高分辨率特性,能够提供丰富的光谱、纹理等信息<sup>[2]</sup>,因此得到广泛应用。高分辨率特性保证了遥感影像的连续

性,能够精确表征地物轮廓,从而保证了地物轮廓提取的有效性<sup>[3]</sup>;而光谱信息能够表征不同地物的不同光学特性,从丰富的光谱信息中能够获得多种特征(如归一化植被指数<sup>[4]</sup>)用于地物分类和三维模型重构,光谱信息在植被分类方面作用明显<sup>[5]</sup>。但是,利用遥感影像进行地物分类和三维建模也存在几个问题:1)使用遥感影像重建三维模型时从二维空间映射到三维空间的过程会丢失大量信息,而且三维

收稿日期: 2019-09-02; 修回日期: 2019-09-09; 录用日期: 2019-09-16

基金项目: 国家自然科学基金(41601436)

\* E-mail: shixiaosong321@126.com

模型需要多幅二维影像匹配生成,费时费力;2)城区环境复杂,某些地物具有相似的光学特性,使用遥感影像进行地物分类时会产生误分类现象;3)传统影像数据获取周期长、成本高,不满足很多实际应用的需求<sup>[6]</sup>。激光雷达(LiDAR)技术的出现和发展为地球空间信息的获取提供了全新技术手段,被誉为继全球定位系统以来遥感测绘领域的又一场技术革命<sup>[7]</sup>。机载激光雷达克服了传统遥感影像的缺点,数据采集自动化程度高,能够快速、精确地获得城区地物的三维点云数据;随着机载激光雷达的不断发展,现有激光雷达也可以携带遥感影像传感器,其采集遥感影像的周期也大大缩短。点云数据能够提供地物的空间三维信息,尤其是高精度的高程信息,弥补了遥感图像无法有效表示建筑物侧面墙体、植被高度等立面信息的缺陷<sup>[8]</sup>,而且点云数据可以直接应用于三维模型重构。但是点云本身也存在缺陷,其点集离散分布,无法有效表示地物的轮廓线,并且缺少光谱、纹理等信息。遥感影像和点云各有利弊,对两种数据进行融合处理能够有效利用彼此的优势弥补自身缺陷,因此近年来越来越多研究者将点云与遥感影像融合后的新数据应用于地物分类提取和

三维模型重构等方面的研究。

本文基于改进 Point-Net 网络提出了一种融合点云与遥感影像的地物分类算法,利用 Qi 等<sup>[9]</sup>在 2017 年提出的 Point-Net 网络作为分类器进行城区地物分类,针对 Point-Net 网络忽略局部特征的问题,设计多尺度邻域点云特征提取网络弥补原始 Point-Net 网络的缺陷,使用遥感影像和点云融合后的多源融合点云作为新数据源用于分类器的训练分类。本文方法首先对点云与遥感影像进行配准,配准之后将点云每个点对应遥感影像的光谱信息加入该点特征,生成融合遥感影像的点云数据。之后对融合数据进行预处理,预处理后的点云数据作为训练数据和测试数据,训练 Point-Net 分类器,用训练好的分类器对测试点云数据进行分类处理并检验算法的精度和效率,通过与其他算法进行对比分析,验证本文方法的有效性。

## 2 方法具体流程

本节将介绍本文方法的具体流程,分为点云与遥感影像融合和多尺度 Point-Net 模型地物分类两部分,流程图如图 1 所示。

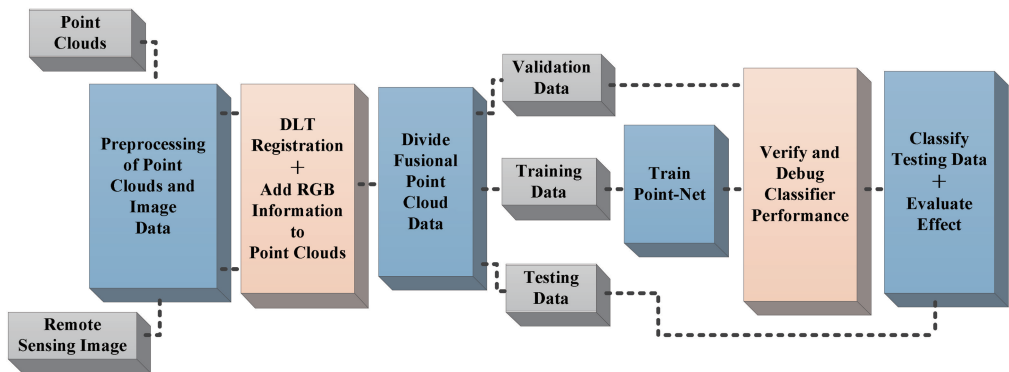


图 1 方法流程图

Fig. 1 Flow chart of proposed method

第一部分为点云与遥感影像融合。本文基于直接线性变换(DLT)法<sup>[10]</sup>和 SIFT 算法<sup>[11]</sup>设计了一种点云和遥感图像的配准融合方法。首先要进行 LiDAR 点云数据与遥感影像数据的配准,即通过配准模型推算两种数据的变换关系,确定点云中各点在相应遥感图像上的像素坐标<sup>[12]</sup>,这是数据融合的前提。配准步骤如下:1) LiDAR 点云预处理,先将点云转换为深度图像,通过多阈值滤波、聚类分析和形态学平滑,提取建筑物二值图像;2) 遥感影像预处理,对遥感图像进行 RGB 聚类、阈值滤波和形态学平滑,提取建筑物二值图像;3) SIFT 特征提取和配

准模型构建,用 SIFT 算法提取两幅二值图像的 SIFT 点特征,用最小二乘法匹配确立同名点,以同名点为配准控制点建立 DLT 配准模型,进行点云和遥感影像配准。配准完成后进行点云数据和遥感图像的融合,利用配准模型找到点云中点在遥感影像中对应的像素点,将像素点的 RGB 光谱特征添加到对应点的特征向量中,得到融合多源点云数据。

第二部分为多尺度 Point-Net 模型地物分类。本文方法在原始 Point-Net 模型的基础上加入多尺度局部邻域提取网络,构建多尺度 Point-Net 模型作为 LiDAR 点云地物分类的分类器。首先将点云

样本数据分为训练数据集、验证数据集和测试数据集三部分。之后训练多尺度 Point-Net 分类器,并调整模型参数、训练拟合分类器模型和对模型分类能力进行初步评估调整。训练集和验证集用于分类器的训练过程,训练集用于参数调试和分类器训练,验证集用于检验训练过程中分类模型的性能是否达到预定标准,防止过拟合。最后测试本文算法的地物分类效果,使用训练完成的多尺度 Point-Net 分类器对测试集进行分类,利用相关指标对分类结果进行评估。

### 3 点云与遥感影像融合

本章主要阐述 LiDAR 点云与遥感影像融合的

相关原理和具体方法。由于本研究主要针对城区 LiDAR 点云数据进行地物分类,为后续将点云数据应用于建筑物三维模型重构、生成数字高程模型 (DEM) 等用途打下基础,因此本章的融合处理旨在将遥感影像的光谱数据作为补充信息加入点云数据特征空间中。融合遥感影像的 LiDAR 点云数据的优势包括:1)把光谱等信息作为点云的新特征,增加了单个点的特征向量维度;2)多传感器融合后的数据能够更全面地描述被测对象,增加影像光谱等信息能够提升点云对植被、建筑物等地物的精细分类能力;3)多传感器并行工作能够有效提高效率,在相同时间内获取更多信息。LiDAR 点云与遥感影像融合流程图如图 2 所示。

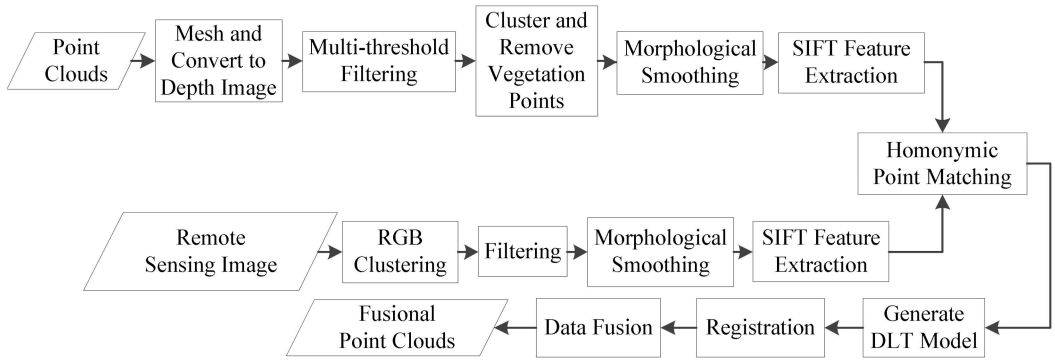


图 2 配准融合流程

Fig. 2 Process of registration and fusion

#### 3.1 点云与遥感影像配准

目前的 LiDAR 点云与遥感影像配准主要有两种配准方法:一种是基于光束法平差<sup>[13]</sup>的点云与序列影像的配准,该模型对内外方位元素初始值要求严格,参数复杂,配准结果精确;另一种是基于 DLT 模型的点云与单张影像的配准,该模型简化了参数,直接建立物点与像点的线性关系,无需推算内外方位元素的初始值,是目前主流的配准变换模型。本文使用 DLT 模型进行配准,对点云与单张遥感影像进行融合,DLT 模型可表示为

$$\begin{cases} x + \frac{l_1 X + l_2 Y + l_3 Z + l_4}{l_9 X + l_{10} Y + l_{11} Z + 1} = 0 \\ y + \frac{l_5 X + l_6 Y + l_7 Z + l_8}{l_9 X + l_{10} Y + l_{11} Z + 1} = 0 \end{cases}, \quad (1)$$

式中: $(X, Y, Z)$ 为点云数据物点的物方坐标; $(x, y)$ 为与物点对应的遥感影像像素点的像平面坐标; $l_i (i=1, 2, \dots, 11)$ 为配准模型的待定参数,表示点云空间坐标系和影像平面坐标系之间的转换关系。

配准前,对选取的小范围对应区域点云和遥感图像数据提取建筑物二值图像,旨在利用两种数据

的建筑物特征相似性构建配准模型。由于点云数据的无序性,本文采用网格法<sup>[14]</sup>将点云数据规范化,网格宽度由点云平均点距确定,即

$$d = \sqrt{\frac{S}{N}}, \quad (2)$$

式中: $d$ 为点云的平均点距; $S$ 为点云区域在水平面上的最小外界多边形面积; $N$ 为点云中点的个数。点云通过网格规则化后,每个网格单位都为深度图像的一个像素点,每个像素网格的灰度值由该网格内所有点的高程值 $z_j (j=1, 2, \dots, n)$ 确定,灰度值计算公式为

$$I = \frac{\sum_{j=1}^n z_j}{n \cdot z_{\max}} \times 255, \quad (3)$$

式中: $I$ 为当前像素网格的灰度值; $z_{\max}$ 为点云所有点高程最大值; $n$ 为当前网格中点的个数。遍历所有网格并计算灰度值,生成深度图像。

生成深度图像后,依据图像中建筑物的灰度值分布范围并设定多个范围阈值进行多阈值滤波,符

合阈值范围的像素点灰度值设为 255, 剩余像素点灰度值设为 0, 综合分析不同阈值的提取效果并通过加权得到最终提取阈值, 滤波后得到初始二值图像。由于初始二值图像除了提取出建筑物以外还有少量植被等小型地物, 接下来通过聚类分析剔除植被等地物。通过聚类分析首先找到多个中心点  $(x_k, y_k) (k=1, 2, \dots, m)$  进行八邻域分析, 将满足生长准则的中心点归为同一类别, 继续进行邻域聚类, 本文中设置生长准则为两点间的距离和灰度差均小于一定阈值, 通过迭代聚类提取多个区域, 计算每个区域的面积, 把面积小于阈值的区域认定为非建筑物区域并将其删除。最后利用形态学开运算对建筑物边界进行平滑, 得到最终的建筑物二值图像。

针对遥感图像, 首先进行 RGB 聚类, 找到初始像素中心点, 计算相邻像素点之间的 R、G、B 三个通道的差值, 当三者的差值均小于给定阈值时, 判定两像素点为一类, 直到遍历所有像素点, 提取多个区域。之后依据 RGB 值和区域面积设定阈值并进行滤波, 当区域的 RGB 值均在阈值范围内并且区域面积大于给定阈值时判定为建筑物区域, 提取所有建筑物区域并生成初始二值图像, 利用形态学开运算对建筑物边界进行平滑, 得到最终的建筑物二值图像。

得到二值图像后, 首先提取 SIFT 特征进行配准, 用 SIFT 算法提取特征点并将其匹配为同名点对, 点云深度影像的 SIFT 点特征向量  $\mathbf{A}_p = (a_{p1}, a_{p2}, \dots, a_{p128})$ , 遥感影像的 SIFT 点特征向量  $\mathbf{B}_q = (b_{q1}, b_{q2}, \dots, b_{q128})$ , 用最小二乘法进行匹配, 计算两组点特征之间的欧氏距离作为相似性度量, 当点云深度影像与遥感影像中最近的 SIFT 点特征距离  $D(\mathbf{A}_p, \mathbf{B}_q)_{\min}$  和次近的 SIFT 点特征距离  $\bar{D}(\mathbf{A}_p, \mathbf{B}_q)_{\min}$  之比小于给定阈值时, 认为两者是同名点, 该过程可表示为

$$\begin{cases} D(\mathbf{A}_p, \mathbf{B}_q) = \sqrt{\sum_{r=1}^{128} (a_{pr} - b_{qr})^2} \\ \frac{D(\mathbf{A}_p, \mathbf{B}_q)_{\min}}{\bar{D}(\mathbf{A}_p, \mathbf{B}_q)_{\min}} < d \end{cases}, \quad (4)$$

式中:  $d$  为设定的距离比阈值;  $a_{pr}$  为第  $p$  个 SIFT 点特征向量  $\mathbf{A}_p$  中的第  $r$  个子特征;  $b_{qr}$  为第  $q$  个 SIFT 点特征向量  $\mathbf{B}_q$  中的第  $r$  个子特征。由于直接线性变换模型有 11 个位置参数, 因此要提取 6 对以上的同名点作为配准控制点进行参数估计, 求解所有参数后得到点云和遥感图像之间的 DLT 配准模型, 根据配准模型对待配准图像进行重采样, 得到配

准后的遥感图像。

### 3.2 点云与遥感影像融合

完成 LiDAR 点云与遥感影像的配准后, 就能确定点云中物点坐标  $(X, Y, Z)$  所对应的遥感图像像素点坐标  $(x, y)$ 。点云原有的特征向量包含三维坐标  $(X, Y, Z)$  和回波强度  $P_I$ , 需要将遥感图像的 RGB 信息加入点云特征向量, 依据点云和遥感影像的对应关系进行融合, 具体步骤如下:

1) 依次选取点云数据中的所有点, 对于当前点  $(X', Y', Z')$ , 依据 DLT 配准模型找到其在遥感影像中对应的像素点  $(x', y')$ ;

2) 提取对应像素点  $(x', y')$  的 R、G、B 三个通道值  $(V_r, V_g, V_b)$ , 加入当前点  $(X', Y', Z')$  的特征向量, 得到新的特征向量  $(X, Y, Z, P_I, V_r, V_g, V_b)$ ;

3) 遍历所有点, 完成点云与遥感影像的融合, 得到多源融合点云数据。

## 4 构建多尺度 Point-Net 分类模型

与二维图像不同, LiDAR 三维点云是一种具有旋转不定性且排列无序的数据形式, 而之前的深度学习模型需要对输入的数据进行有序排列, 因此之前的研究中会将点云体素化或转换成深度图片的形式输入深度学习模型, 这些操作既增加模型复杂度同时又损失了点云的原始信息, 而且使得融合点云数据的特征维度增加, 只有改进现有神经网络结构才能增大深度学习模型处理数据特征的维度。

Point-Net 神经网络引入输入变换矩阵 T-Net 和最大池化层: T-Net 通过初始点云位置生成旋转矩阵, 依据损失函数返回值不断迭代调整 T-Net, 最终把输入点云旋转到分类的最优角度; 最大池化层是通过降采样的方式, 降低输出维度, 保留特征中的重要信息<sup>[15]</sup>用于提取点云的全局特征。Point-Net 神经网络通过引入以上两种结构, 克服了神经网络训练点云数据时的无序性<sup>[16]</sup>和旋转性问题, 可以直接把点云数据作为输入进行处理, 无需再进行体素化或二维压缩处理, 能够提取点云数据的单点特征和全局特征并在后续神经网络模型中完成点云分类, 文献[9]测试了不同点云深度学习分类模型对场景点云的分类精度, 其中 Point-Net 分类模型精度最高, 平均精度达到 83.7%, 原始 Point-Net 网络结构如图 3 所示。

原始 Point-Net 模型存在两个问题: 1) 融合点云数据包含三维坐标、回波强度、RGB 值等特征, 原始 Point-Net 只有三个数据输入通道; 2) 原始 Point-Net 未提取局部邻域特征信息, 造成分类器在某些细节

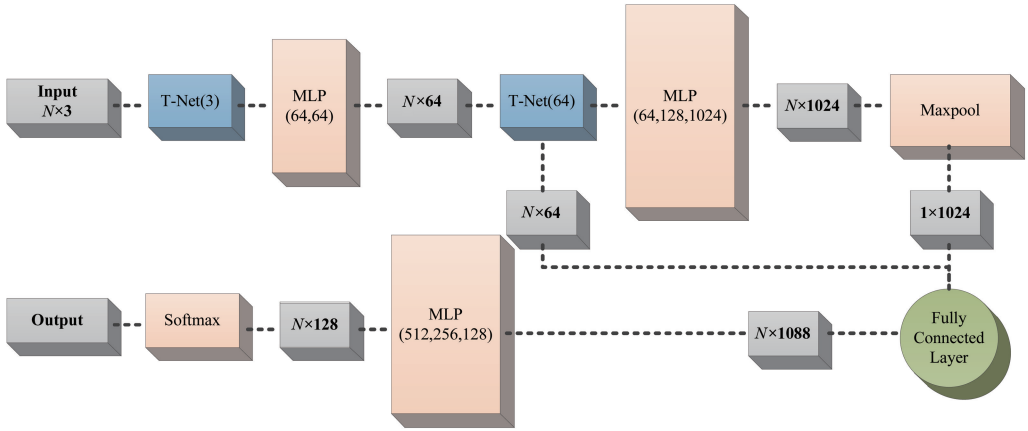


图 3 Point-Net 网络结构

Fig. 3 Architecture of Point-Net

上出现分类错误。本节在 Point-Net 网络结构的基础上构建多尺度 Point-Net 分类模型,提升模型的输入特征维度,并且提取多尺度邻域点云特征用于点云地物分类。具体改进思路如下:针对融合点云特征空间维度增加的问题,通过调整输入变换矩阵维度增加通道数,使矩阵输入从原来的三维特征向量变为融合后的七维特征向量;针对融合点云数据扩展特征空间带来

的数据量增加,通过加深 MLP(multilayer perceptron)层的网络层数充分提取点云的深度特征;针对缺少局部邻域特征的问题,在原有 Point-Net 神经网络提取单点特征和全局特征的基础上,搭建两个局部特征提取网络,提取了不同尺度的局部点云数据,用局部特征提取网络提取点云局部特征用于分类模型训练。改进后多尺度 Point-Net 网络结构框图如图 4 所示。

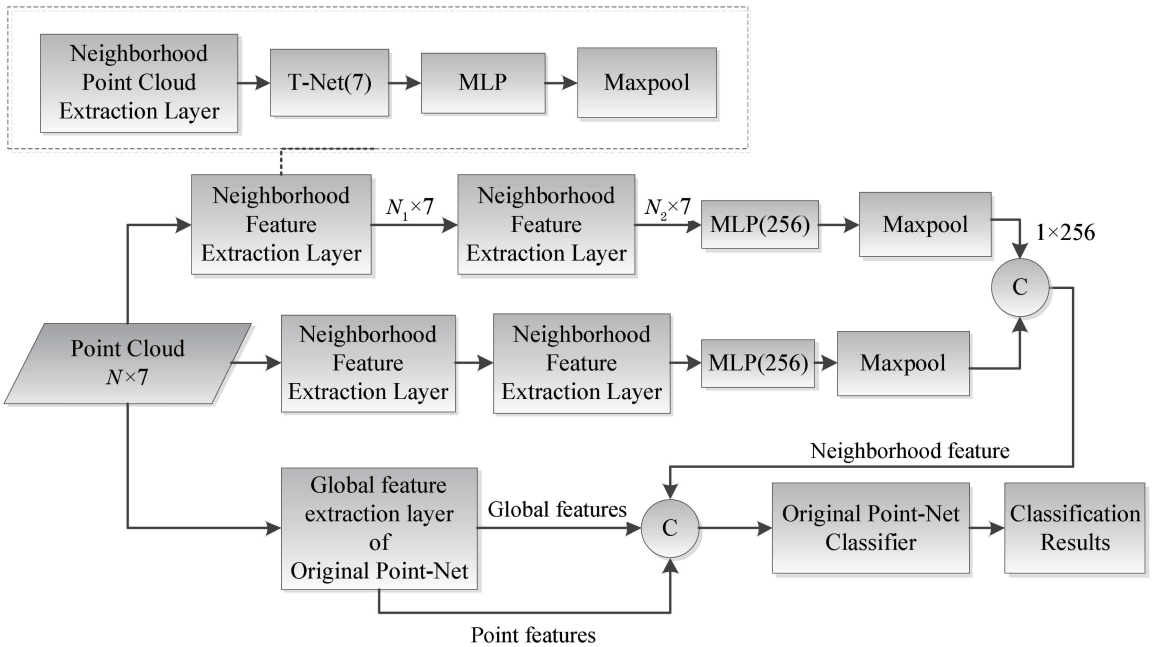


图 4 多尺度 Point-Net 网络结构

Fig. 4 Architecture of multi-scale Point-Net

按照改进思路,构建多尺度 Point-Net 网络的具体措施有:1) 将原来的输入变换矩阵 T-Net(3) 替换为 T-Net(7),替换后的网络能够进行七维融合点云数据的空间变换,在局部特征提取网络和原始网络全局特征提取部分的 MLP 层增加了两个卷积

层,以增加特征提取深度;2) 在原始 Point-Net 网络基础上加入两个不同尺度的局部特征提取网络,每个网络都包含一个中心点采样层、一个邻域聚类层和一个 Point-Net 网络,中心点采样层通过最远点采样法提取中心点,邻域聚类层提取离中心点的欧

氏距离  $d$  小于阈值  $d_0$  的所有点, 对其进行聚类形成邻域点云子集, 每个邻域点云子集子点集之间可以存在重叠, 并且每个子点集的点数不一定相等, 网络结构如图 4 中虚线框内所示。网络的邻域特征提取的步骤为: 1) 利用采样层采样选取中心点, 利用邻域聚类层提取所有中心点不同尺度邻域点云子集; 2) 提取多尺度邻域点云后, 将大尺度和小尺度的邻域点云分别输入一个局部特征提取网络用于提取特征, 通过 T-Net(7) 校正数据, 之后通过 MLP 网络和 Maxpool 结构得到一个 7 维特征向量作为此子集的中心点的输出特征; 3) 所有  $N_1$  个输出特征点又可以组成新的特征点集, 再次提取若干邻域特征点块, 将其输入相同结构的 Point-Net 网络进行特征提取。算法中每个尺度都设置了三层小型 MLP 网络和 Maxpool 结构, 随着提取次数增加, 输出特征点集中的点越来越少, 每个特征点包含的局部邻域信息越来越多, 第三层网络通过池化得到一个 256 维特征向量作为局部点云特征。两种尺度邻域特征与原始 Point-Net 提取的单点特征和全局特征通过全连接层连接, 得到多尺度邻域特征向量。改进后的多尺度 Point-Net 神经网络提取了点云的

局部特征, 克服了原始 Point-Net 神经网络缺少邻域特征的问题, 邻域特征更好地描述了点云在局部区域的细节。

## 5 仿真与结果分析

### 5.1 仿真

仿真使用国际摄影测量与遥感学会提供的 Vaihingen 城区 LiDAR 点云数据集测试分类算法的有效性, 数据集包含各种复杂地物场景, 使仿真结果具有普适性。点云的平均点距为 0.66 m, 分类算法将点云数据分为地面、建筑物、植被和人造物体四类。本文数据集共有 573026 个离散点, 将数据集划分为训练集(90%, 其中包含 20% 的验证集)和测试集(10%), 所有样本数据集中点云类别已经过人工标记。本文选取面积为 22995 m<sup>2</sup> 的点云数据(包含 57302 个离散点)作为测试集, 将其作为验证本文点云算法模型分类精度的依据。用于仿真的点云部分数据如图 5(a) 所示, 对应的遥感影像如图 5(b) 所示, 点云与遥感影像的重叠影像如图 5(c) 所示, 由图可知两图之间相同地物位置差别较大。

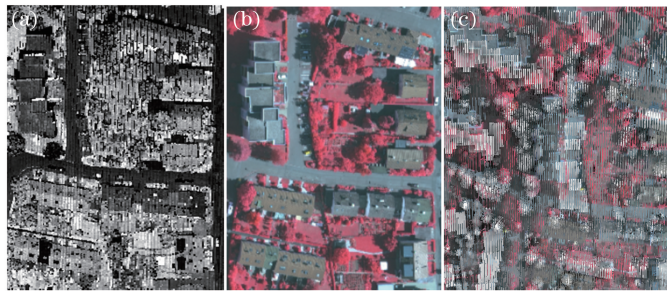


图 5 仿真数据。(a) 点云数据; (b) 遥感影像数据; (c) 重叠影像

Fig. 5 Simulation data. (a) Point cloud data; (b) remote-sensing image data; (c) overlapping image

点云与遥感影像配准首先利用部分点云数据确定 DLT 配准模型, 之后完成整体点云数据的配准。仿真采用的点云数据平均点距为 0.4 m, 遥感影像分辨率为 0.08 m, 分辨率之差在合理范围内, 满足一个像素的配准精度。通过 2.1 节中的预处理操作提取 LiDAR 点云数据和遥感图像的建筑物二值图像, 点云和遥感影像生成的建筑物二值图像结果如图 6(a) 和 6(b) 所示, 未配准图像之间在位置上存在一定的差异。由图可知, 预处理操作有效滤除了其他地物, 并且得到了完整的建筑物轮廓。之后在提取建筑物的基础上提取二值图像的 SIFT 点特征, 找到两幅图像的同名点, 推算出 DLT 模型的所有参数并最终确定配准模型, 根据配准模型对待配准

图像进行重采样, 得到配准后的遥感图像与点云数据, 配准结果如图 6(c) 所示。从图中可知, 配准后点云与遥感影像重合度高, 配准效果良好。配准后, 将遥感影像的 RGB 信息加入点云对应点的特征向量, 依据点云和遥感影像的对应关系进行融合, 融合结果如图 6(d) 所示, 图中点云数据按 RGB 信息显示。

生成多源融合点云数据后, 用多尺度 Point-Net 网络进行点云地物分类。首先用训练数据训练多尺度 Point-Net 分类器模型, 将融合点云分别输入原始 Point-Net 网络和两个局部特征提取网络, 通过原始 Point-Net 网络得到点云的单点特征和全局特征, 通过局部特征提取网络得到局部邻域特征, 用全

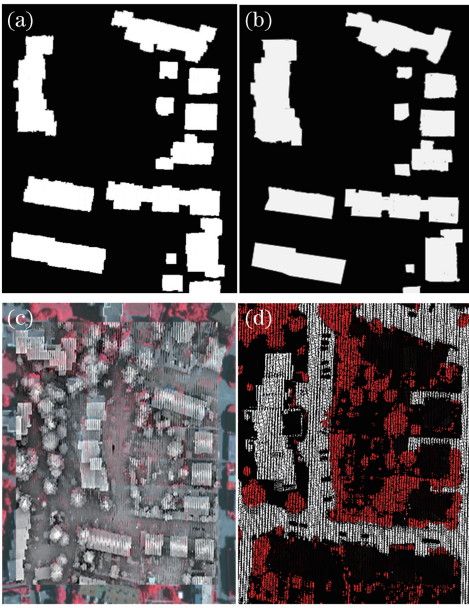


图 6 点云、遥感影像二值图像以及配准、融合点云数据。  
(a)点云二值图像;(b)遥感影像二值图像;(c)配准结果;(d)融合点云数据

Fig. 6 Binary images of point clouds and remote-sensing image, results of registration, and fusion point clouds. (a) Binary image of point clouds; (b) binary image of remote-sensing image; (c) registration result; (d) fusion point cloud data

连接层连接所有特征, Point-Net 网络通过 MLP 层对所有特征进行多层提取, 得到 128 维特征向量作为点云分类依据对训练集进行分类并利用损失函数计算残差, 调整分类器参数, 用验证集对分类模型的性能进行交叉验证。提取局部特征时, 第一层

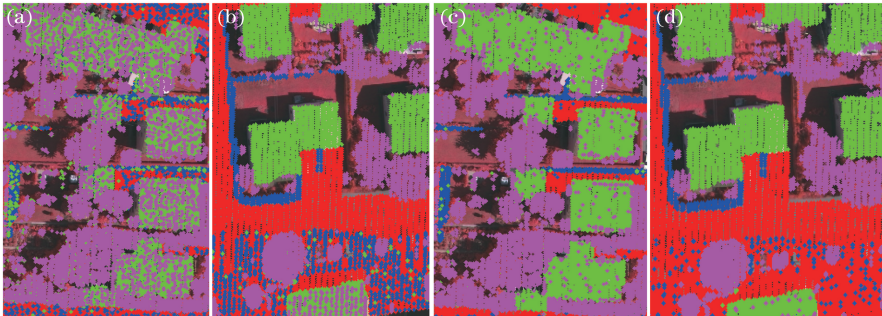


图 8 分类结果。(a)(b)未加入 RGB 点云;(c)(d)加入 RGB 点云

Fig. 8 Classification results. (a)(b) Without RGB point clouds; (c)(d) with RGB point clouds

### 5.2 结果分析

为了定量检测模型的分类效果, 本文采用  $F_1$  值[由准确率( $P$ )和召回率( $R$ )求得]来评估模型对测试集的分类精度, 采用分类时间(time)评估模型训练与分类的速度, 准确率、召回率和  $F_1$  值计算公式分别为

$$\begin{cases} P = \frac{T_P}{T_P + F_P} \\ R = \frac{T_P}{T_P + F_N} \\ F_1 = \frac{2PR}{P + R} \end{cases} \quad (5)$$

特征提取网络设置 512 个中心点, 阈值半径  $d_0$ 。经过调试分别设置为 1.5 和 4; 第二层特征提取网络设置 64 个中心点, 阈值半径  $d_0$ 。设置为 0.1。Point-Net 网络训练时, 最大训练轮次(epoch)设置为 100, 初始学习率(learning rate)为 0.01, 学习率衰减率(decay rate)为 0.7, 批处理尺寸(batch size)为 32。网络调用 Adam 优化算法<sup>[17]</sup>拟合分类模型, 依据分类结果重新调整模型参数, 该算法基于训练数据迭代更新神经网络权重, 比一阶随机梯度下降算法性能拟合效果更优, 训练误差与训练轮次的关系曲线如图 7 中含圆圈曲线所示。从图中可知, 随着训练轮次增加, 模型的训练误差逐渐减小, 当训练轮次达 90 次以上时, 训练误差减小到最小值附近并且趋于稳定。利用训练好的多尺度 Point-Net 网络模型对测试数据集进行分类, 多尺度 Point-Net 网络对测试集分类的结果与原始 Point-Net 网络的分类结果如图 8 所示。

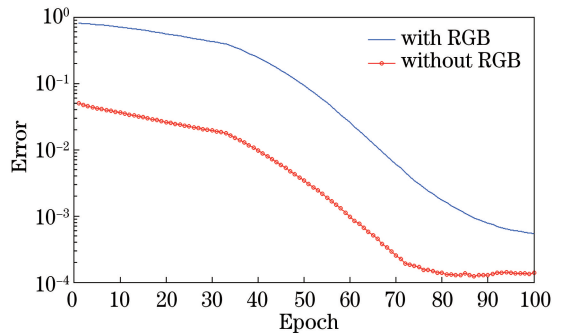


图 7 训练误差曲线

Fig. 7 Training error curve

式中:  $T_p$  为判定为该类的样本中判断正确的点数;  $F_p$  为判定为该类的样本中判断错误的点数;  $F_N$  为判定为其他类的样本中判断错误的点数。  $F_1$  值越高, 精度越高。相对于准确率,  $F_1$  值能够更全面表示分类精度, 分类精度和分类时间的评估结果如表 1 和表 2 所示。

由表 1 可知, 多尺度 Point-Net 网络在对点云数据的各类地物  $F_1$  值和整体分类精度方面高于支持向量机和原始 Point-Net 网络, 分类效果更优, 这说明加入多尺度邻域点云特征的确能改善分类器对细节的分类。从图 8 可知, 建筑物边角、植被与建筑物交汇处等细节得到了有效分类。而将遥感影像与点云融合作为输入数据并通过分类得到的  $F_1$  值和整体分类精度则比单一点云作为输入数据要高, 特别是在地面、植被和建筑物方面分类效果得到明显提高。从图 8 可知, 加入 RGB 通道值的融合点云数据能够通过光谱差异有效区分植被和起伏较大的建筑物顶部, 克服单一点云数据无法解决的植被与建筑物误分类问题, 同时针对平坦地面和低矮人造物体误分类问题, 融合点云也能够通过两者光谱信息上的差异进行有效区分。由图 7 可知, 加入 RGB 通道值的融合点云数据的训练误差始终低于单一点云数据。由表 2 可知, 原始 Point-Net 网络的训练时间

表 1 不同方法分类精度

Algorithm	Category	$F_1/\%$	Overall accuracy /%
SVM	Ground	79.2	78.7
	Vegetation	81.2	
	Building	78.5	
Point-Net	Artificiality	45.1	84.7
	Ground	89.4	
	Vegetation	87.4	
Multi-scale Point-Net without RGB	Building	78.2	92.5
	Artificiality	24.2	
	Ground	92.7	
Multi-scale Point-Net with RGB	Vegetation	91.6	96.8
	Building	93.3	
	Artificiality	57.1	
Point-Net with RGB	Ground	98.5	96.8
	Vegetation	97.6	
	Building	96.1	
	Artificiality	76.9	

最短, 但是与未加入融合点云数据的多尺度 Point-Net 网络的训练时间相差不大, 这是由于多尺度特征提取网络与原始特征提取网络并行处理, 节约了训练时间; 而加入融合点云数据的多尺度 Point-Net 网络由于需要配准融合并且点云特征维度增加, 因此训练时间增加。

表 2 不同方法分类时间

Table 2 Classification time of different methods

Method	Point-Net	Multi-scale Point-Net without RGB	Multi-scale Point-Net with RGB
Training time	4375	4649	5568
Test time	204.9	209.8	204.9

## 6 结 论

提出了一种融合点云与遥感影像的城区地物分类方法用于城区点云数据的分类, 该方法将 LiDAR 点云数据与遥感图像融合生成融合点云数据, 以 Point-Net 为基础, 提出多尺度 Point-Net 分类模型, 改进后的分类模型结构能够处理融合点云数据。对相应理论和具体流程进行了详细描述, 通过仿真分析检验该方法的有效性。研究表明: 融合遥感影像的 LiDAR 点云数据包含了丰富的光谱信息, 克服了原始点云数据缺少光谱特征的缺点, 为点云分类增加了更多信息, 改进后的 Point-Net 分类模型不仅能够提取点云的单点特征和全局特征, 而且能够提取点云的局部特征, 改善了原始 Point-Net 网络对于局部细节的误分类问题, 有效提高了地物分类精度, 所提方法自动化程度较高。后续工作可以

在现有点云深度学习模型基础上进行深入研究, 在特征提取、分类效率等方面对现有模型进行改良, 提出更加实用高效的分类模型。

## 参 考 文 献

- [1] Guo Y, Bennamoun M, Soheli F, et al. 3D object recognition in cluttered scenes with local surface features: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(11): 2270-2287.
- [2] Wu Z H, Gao Y M, Li L, et al. Fully convolutional network method of semantic segmentation of class imbalance remote sensing images [J]. Acta Optica Sinica, 2019, 39(4): 0428004.  
吴止锲, 高永明, 李磊, 等. 类别非均衡遥感图像语义分割的全卷积网络方法 [J]. 光学学报, 2019, 39(4): 0428004.



- [3] Du S J, Zou Z R, Zhang Y S, et al. A building extraction method via graph cuts algorithm by fusion of LiDAR point cloud and orthoimage [J]. *Acta Geodaetica et Cartographica Sinica*, 2018, 47(4): 519-527.  
杜守基, 邹峥嵘, 张云生, 等. 融合 LiDAR 点云与正射影像的建筑物图割优化提取方法 [J]. *测绘学报*, 2018, 47(4): 519-527.
- [4] Alonzo M, Bookhagen B, Roberts D A. Urban tree species mapping using hyperspectral and lidar data fusion [J]. *Remote Sensing of Environment*, 2014, 148: 70-83.
- [5] Gao S, Shen X, Dai J S, et al. Tree species classification in urban forests based on LiDAR point cloud segmentation and hyperspectral metrics extraction [J]. *Remote Sensing Technology and Application*, 2018, 33(6): 1073-1083.  
皋厦, 申鑫, 代劲松, 等. 结合 LiDAR 单木分割和高光谱特征提取的城市森林树种分类 [J]. *遥感技术与应用*, 2018, 33(6): 1073-1083.
- [6] Chen J H. Data fusion of aerial image and point cloud derived from LiDAR system [D]. Xi'an: Chang'an University, 2013.  
陈军红. 机载 LiDAR 点云数据与影像数据融合处理 [D]. 西安: 长安大学, 2013.
- [7] Yang S J, Zhang K S, Shao Y S. Classification of airborne LiDAR point cloud data based on multiscale adaptive features [J]. *Acta Optica Sinica*, 2019, 39(2): 0228001.  
杨书娟, 张珂殊, 邵永社. 基于多尺度自适应特征的机载 LiDAR 点云分类 [J]. *光学学报*, 2019, 39(2): 0228001.
- [8] Guan H Y, Deng F, Zhang J Q, et al. Object-based fusion and classification of airborne laser scanning data and aerial images [J]. *Geomatics and Information Science of Wuhan University*, 2009, 34(7): 830-833.  
管海燕, 邓非, 张剑清, 等. 面向对象的航空影像与 LiDAR 数据融合分类 [J]. *武汉大学学报(信息科学版)*, 2009, 34(7): 830-833.
- [9] Qi C R, Su H, Mo K, et al. Point-Net: deep learning on point sets for 3D classification and segmentation [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 21-26, 2017, Honolulu, Hawaii, USA. New York: IEEE, 2017: 652-660.
- [10] Wei Q C, Zhou G Q, Zhu Q, et al. On DLT method for CCD camera calibration [C] // *Proceedings of Third International Conference on Signal Processing (ICSP'96)*, October 18-18, 1996, Beijing, China. New York: IEEE, 1996: 883-885.
- [11] Stamos I, Liu L, Chen C, et al. Integrating automated range registration with multiview geometry for the photorealistic modeling of large-scale scenes [J]. *International Journal of Computer Vision*, 2008, 78(2/3): 237-260.
- [12] Gong X, Wu L, Xie Z, et al. Classification method of high-resolution remote sensing scenes based on fusion of global and local deep features [J]. *Acta Optica Sinica*, 2019, 39(3): 0301002.  
龚希, 吴亮, 谢忠, 等. 融合全局和局部深度特征的高分辨率遥感影像场景分类方法 [J]. *光学学报*, 2019, 39(3): 0301002.
- [13] Zheng S, Huang R, Zhou Y. Registration of optical images with LiDAR data and its accuracy assessment [J]. *Photogrammetric Engineering & Remote Sensing*, 2013, 79(8): 731-741.
- [14] Hovad J, Komarkova J, Sedlak P. Slope based grid creation using interpolation of LIDAR data sets [C] // *ICSOFT*, July 29-31, 2013, Reykjavik, Iceland. [S. l.: s.n.], 2013: 227-232.
- [15] Hackel T, Savinov N, Ladicky L, et al. Semantic3D. Net: a new large-scale point cloud classification benchmark [J/OL]. (2017-04-12) [2019-09-05]. <https://arxiv.xilesou.top/abs/1704.03847>.
- [16] Schmidhuber J. Deep learning in neural networks: an overview [J]. *Neural Networks*, 2015, 61: 85-117.
- [17] Kingma D P, Ba J. Adam: a method for stochastic optimization [J/OL]. (2017-01-30) [2019-09-05]. <https://arxiv.xilesou.top/abs/1412.6980>.