DOI: 10.19783/j.cnki.pspc.201461

基于 PSO-KNN 的变电站设备三维点云识别方法研究

李 科

（国网四川省电力公司信通公司（省数据中心），四川 成都 610072）

摘要：针对传统的三维重建方法因数据缺失而造成的精度差、效率低等问题，在三维激光扫描点云的基础上，提出了一种将粒子群优化算法和 *k*-近邻分类算法相结合的变电站设备三维点云识别方法。通过粒子群优化算法对各子空间特征的系数权重进行优化，*k*-近邻分类算法完成分类。通过实验分析点云子空间的大小和丢失率对识别效果的影响，并与改进的迭代最近点算法进行比较，验证该方法的优越性和准确性。实验结果表明，该方法具有较好的识别准确率和效率，识别准确率达到 95%以上，平均识别时间为 0.19 s，具有一定的应用价值。

关键词：变电站设备；粒子群算法；*k*-近邻分类算法；三维点云识别；三维重建

3D point cloud research on an identification method based on PSO-KNN substation equipment

LI Ke

(State Grid Sichuan Information & Telecommunication Company (Provincial Data Center), Chengdu 610072, China)

Abstract: There are problems of poor accuracy and low efficiency caused by the lack of data in traditional three-dimensional reconstruction methods. Thus, based on a three-dimensional laser scanning point cloud, a recognition method for substation equipment is proposed. This combines the particle swarm optimization algorithm and the *k*-nearest neighbor classification algorithm. The particle population optimization algorithm is used to optimize the coefficient weight of each subspace feature, and the *k*-nearest neighbor classification method is used to classify the equipment. The influence of the size and loss rate of the point cloud subspace on the recognition effect is analyzed through experiment. It is compared with the improved iterative closest point algorithm to verify the superiority and accuracy of the method. Experimental results show that this method can effectively improve recognition accuracy and efficiency. The recognition accuracy can reach more than 95%, and the average recognition time is 0.19 seconds, which has application value.

This work is supported by the Science and Technology Project of State Grid Corporation of China (No. 5216A0182020R).

Key words: substation equipment; particle swarm optimization; *k*-nearest neighbor classification algorithm; 3D point cloud recognition; 3D reconstruction

0 引言

随着地理信息系统向三维领域的不断发展，变电站三维数字化的研究越来越受到重视，并逐渐成为研究的热点[1]。作为一种新的测量技术，3D 激光扫描可以从目标对象的表面采集大量的点云数据。具有非接触、高精度的特性，它为重构物体的数字 3D 模型提供了新的思路[2]。 由于数据采集环境的复杂性和可变性，在实际场景中采集的数据可能包括噪声、缺失等，甚至同一类型设备的点云数据也可能有不同程度的差异。 因此，研究智能变电站设备的三维点云识别方法在现实中具有十分重要的意义。

目前，国内外学者对变电站设备的三维识别技术进行了大量的研究。文献[3]提出了一种采用三维激光获取设备的点云数据的算法，使用由支持向量机改进的 AdaBoost 算法进行识别，通过仿真验证了该算法的优越性。结果表明，训练精度和测试精度均超过 98%，具有良好的识别精度。文献[4]提出了一种霍夫投票算法用于设备点云数据的识别，并比较模型库中的距离直方图以计算目标的相似度。根据相似度阈值获得初始识别结果。通过霍夫投票进

基金项目：国家电网公司科技项目资助(5216A0182020R)

行最后识别，识别准确率为 90.1%，单个设备的平均识别时间为 15.6 s。文献[5]提出了一种通过全景摄像机和三维扫描系统获取室外设备和变电站的多角度场景照片和点云数据的方法。根据指定的建模标准，使用 Autodesk 3d-Max 工具在 3D 场景中对变电站设备、建筑物和周围的地形环境进行建模，获得了与场景匹配的 3D 模型。文献[6]提出了一种将激光扫描和无人机相结合的预警系统，智能检测变电站周围环境。测试结果表明，该系统可以构建周围三维模型。但是，以上方法计算量大、费时、不能满足实际场景识别的要求。这些研究为变电站设备三维点云识别方法的研究提供了理论依据。

在此基础上，提出了一种基于粒子群优化算法(Particle Swarm Optimization, PSO)和 *k*-近邻分类算法(*k*-NearestNeighbor, kNN)的变电站设备三维点云识别方法。 PSO 算法用于优化每个子空间特征的系数权重，而 kNN 算法则用于对设备进行分类。通过实验分析了该算法的性能，验证了该方法的优越性。

1　识别整体流程

文中主要研究的是通过设备点云数据识别设备类型、位置、方向和角度来加速变电站三维仿真模型的重建[7]。 整个过程从简化三维坐标数据开始，并着重于点云数据的目标识别[8]。点云数据处理可以分为三个阶段：

(1) 数据预处理。由于点云数据多且存在噪声点，需要先进行数据预处理。

(2) 点云数据特征提取。首先，根据点的特征将提取过程分为两个步骤。划分子空间和提取子空间中特征形成特征向量[9]。

(3) 目标识别。通过 PSO 算法和 kNN 算法相结合的方式完成三维点云识别。图 1 为识别系统的整体流程。

2　变电站设备的三维识别

变电站设备的三维点云识别，根据点的特征划分子空间并提取子空间的特征形成特征向量，选取整个点云的质心与每个子空间点云的质心连线相对于 *Z* 轴正方向夹角的余弦值作为点云的子空间特征。使用 kNN 获得分类结果，并使用 PSO 算法对每个子空间特征的系数权重进行优化。

2.1　*k*-近邻分类算法

kNN 算法的原理：空间中最近邻有 *k* 个样本属于一个类别，则该样本也属于该类别[10]。在 kNN 算法中，通过特征值距离对选定的相邻 kNN 进行分类。

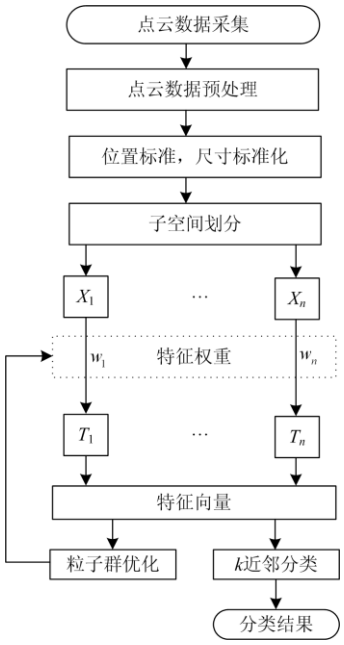


图 1　识别系统的整体流程

Fig. 1 Identifying the overall process

k 值影响着 kNN 算法的分类效果。如果 *k* 很小，则测试样本将在较小范围内对类别进行计数，容易出现分类错误[11]。如果 *k* 较大，则测试样本将在较大范围对类别数进行计数，具有较多的统计信息，可以避免分类错误，但是分类器效率较差，*k* 值一般采用交叉检验确定[12]。*k* 近邻决策图如图 2 所示。

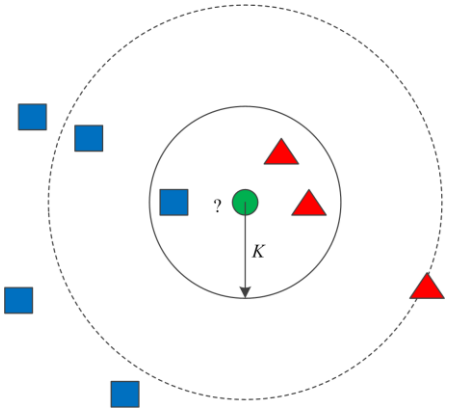


图 2　*k* 近邻决策图

Fig. 2 *k*-nearest neighbor decision graph

如果已知模板和标签中的数据，输入测试数据，计算测试数据的特征与模板特征之间的距离，找出最相近的 *k* 个数据[13]。测试数据的类别就是 *k* 值最多的类，算法实现步骤如下：

1) 计算测试数据 *T* 和模板数据 *MT* 之间的距离 *d*，并选择欧氏距离，如式(1)所示[14]。

$$d(T,MT)=\sqrt{\sum_{i=1}^{n}(T_i-MT_i)^2} \quad (1)$$

2) 通过距离排序，选择距离最小的前 k 个值，k 值设置为 3[15]。

3) 对 k 个样本类别出现频率进行计算。

4) k 个样本属于测试数据预测分类中出现最高的一类。如果频率相等，距离最小的类别为预测类别。

2.2 PSO 优化子空间特征权重

PSO 算法是基于群体智能的启发式算法，是一种全局进化算法，用于求解多目标、非线性和多变量[16]。PSO 算法迭代初始粒子，对粒子的位置和速度进行更新，并跟踪空间中的最优粒子以找到最优解[17]。停止迭代的条件通常为是否达到目标迭代次数。PSO 具有易于编程、收敛速度快和精度高的优点，因此被广泛使用[18]。文中使用 PSO 算法优化各子空间中特征的权重。流程图如图 3 所示。

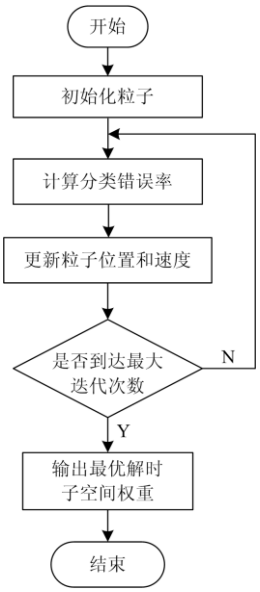


图 3 粒子群优化算法流程

Fig. 3 Particle swarm optimization algorithm flow

通过对各子空间特征的系数权重进行调整，提高了特征识别和分类的准确性。使用 PSO 优化特征权重时，将分类误差率定义为适应度函数，粒子的最终位置表示对应特征所占权重。算法实现过程如下：

1) 对粒子的位置和速度进行初始化。

2) 根据式(2)和式(3)对粒子的位置和速度进行更新[19]。

$$V_{id}^{k+1}=\omega V_{id}^{k}+c_1r_1(pbest_{id}-x_{id}^{k})+c_2r_2(gbest-x_{id}^{k}) \quad (2)$$

$$x_{id}^{k+1}=x_{id}^{k}+V_{id}^{k+1} \quad (3)$$

式中：c_1 和 c_2 分别为自学习和社会学习系数，$c_1=c_2=1.5$；r_1 和 r_2 为[0,1]之间的随机数；x_{id}^{k} 和 V_{id}^{k} 分别第 k 次迭代的位置和速度；$pbest_{id}$ 和 $gbest$ 分别表示种群进化过程中的个体最优和全局最优。

3) 根据式(4)计算适应度函数[20]。

$$fitness=1-\frac{m}{N} \quad (4)$$

式中：函数表示误识错误率；N 为测试数据的总数；m 为被正确认识数量。

4) 更新个体最优和全局最优。

5) 达到迭代次数，停止并输出参数[21]。如果尚未达到迭代次数，重复步骤 2)—4)，直到达到最大迭代次数。

3 实验结果分析

3.1 实验参数

在建立模板特征库之后，为了提高分类识别的精度，PSO 算法用于优化每个子空间特征的系数的权重，而 kNN 算法则用于对设备进行分类。计算机为具有 8 GB 内存的 CPU i5-2450m。图像处理软件为 Matlab 2014[22]。

设置 PSO 算法最大迭代次数为 100。100 个测试数据。表 1 为变电站设备尺寸统计数据，包括断路器(CB)、电流互感器(CT)、隔离开关(DS)、接地开关(ES)、避雷器(LA)、电压互感器(PT)、陷波器(LT)等[23]。

表 1 设备尺寸

Table 1 Equipment size

设备类型	长/m	宽/m	高/m
500 kV_CB	1.475 0	5.031 0	6.475 0
500 kV_CT	1.325 0	1.369 0	6.150 0
500 kV_DS	1.439 0	13.452 0	8.225 0
500 kV_ES	1.597 0	3.896 0	5.602 0
500 kV_LA	1.812 0	1.821 0	6.539 0
500 kV_PT	0.999 0	1.003 0	6.076 0
500 kV_LT	2.217 0	2.211 0	6.824 0

为了验证算法的优越性，从以下三个方面进行分析：(1) 分析点云子空间的大小对识别结果的影响；(2) 分析点云缺失对识别结果的影响；(3) 与改进的 ICP 算法进行比较。

3.2 子空间大小对识别结果的影响

理论上，点云的子空间越小，子空间的边长越小，点云的子空间数目越多，特征向量的维数就越大。与标准设备相比，分类和识别结果更加准确，反之则越不准确[24]。

边长对不同子空间的影响如下所述。在子空间分割之前，将点云数据标准化，并根据 3×3×6 立方体缩放所有点云。 然后，如果选择不同的子空间边长，则子空间的数目将不同，点云特征向量的长度将不同，并且分类效果也将不同。 表 2 所示边长分别为 1.0、0.8、0.6、0.5 和 0.4 的子空间的分类结果。

表 2 分类效果(子空间大小不同)

Table 2 Classification effect (different subspace sizes)

边长	个数	优化前		PSO 优化后	
		分类准确率/%	用时/s	分类准确率/%	用时/s
1.0	54	70.01	0.023	83.05	33.012
0.8	128	79.03	0.025	88.08	45.101
0.6	250	79.03	0.026	88.08	62.087
0.5	432	79.03	0.028	88.08	91.860
0.4	960	88.05	0.033	95.33	120.187

表 2 计算了 PSO 优化前后不同子空间大小的分类准确度和时间。为了更形象地展示子空间大小对分类结果的影响，图 4 为针对不同子空间边缘长度 PSO 优化前后的分类准确率。

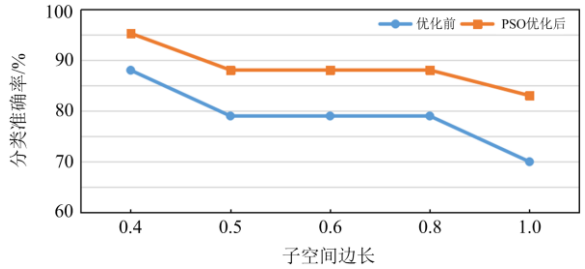


图 4 优化前后不同边长分类准确率

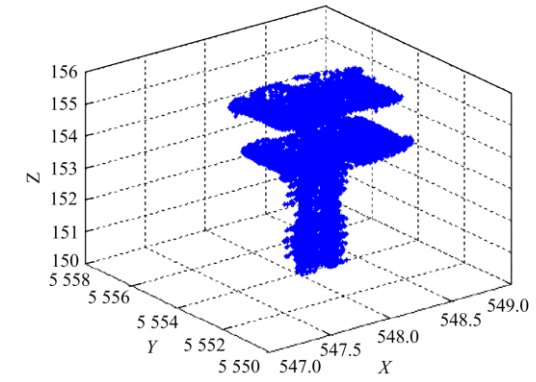
Fig. 4 Classification accuracy of different length before and after optimization

由表 2 和图 4 可知，PSO 优化后的分类精度得到了显著提高。由于 PSO 算法迭代过程需要时间，但时间在合理范围内，因此可以利用 PSO 算法优化子空间特征权重来提高分类精度。边长越短，子空间的总数越多，特征向量的维数越多，使其更容易发生“维数灾难”。

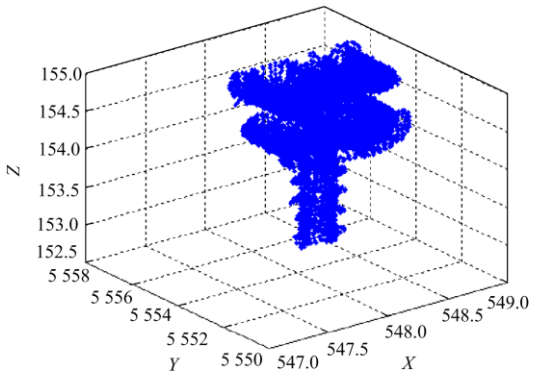
在满足分类准确率的基础上，尽量提高分类速度。

3.3 点云缺失对识别结果的影响

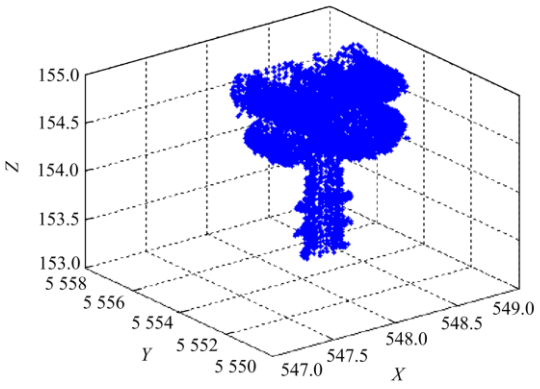
对于文中的识别方法，通过点云缺失分析其性能[25]。由于变电站设备获取的都是理想点云数据。文中以断路器为例，随机丢失一些点云数据，10%、20%、30%的数据丢失。通过文中的识别方法对上述三种情况进行识别、分类，如图 5 所示。在缺失 10%和 20%数据时可以正确识别，缺失 30%数据时识别错误。



(a) 缺失10%三维视图



(b) 缺失20%三维视图



(c) 缺失30%三维视图

图 5 点云缺失三维图

Fig. 5 Point cloud missing 3D map

采用该方法对所有测试数据进行缺失试验，可识别的平均缺失率为 13%，点云缺失太多，识别效果不理想。所有测试数据可识别的最小缺失为 10%，大部分集中在 10%~13%。因此，文中提出的识别方法可以应用于点云缺失较低的场合。

3.4 算法对比

为了进一步说明该算法的有效性，将文中算法与改进的迭代最近点(Iterative Closet Points, ICP)算法进行了比较。改进的 ICP 算法基于传统 ICP 算法

较高的配准精度，使用投影轮廓预选择模型，大大缩短了识别时间，是当前相对理想的识别方法之一。100 个待识别的点云设备，子空间的边长是 0.4，识别结果见表 3。

表 3 不同算法识别对比

Table 3 Identification comparison of different algorithms

设备名称	改进 ICP		文中方法	
	是否正确识别	用时/s	是否正确识别	用时/s
500 kV_CB	是	82.01	是	0.13
500 kV_CT	是	56.71	是	0.15
500 kV_DS	是	150.85	是	0.23
500 kV_ES	是	130.50	是	0.20
500 kV_LA	是	92.56	是	0.18
500 kV_PT	是	70.51	是	0.20
500 kV_LT	是	125.83	是	0.19
…	…	…	…	…
	准确率	平均	准确率	平均
	98.00%	85.81	97.00%	0.19

对于 100 个待识别的点云设备，改进的 ICP 识别方法正确识别了 98 个，仅 2 个识别错误。文中方法正确识别了 97 个，错误 3 个，识别准确率为 97.00%。就算法的识别时间而言，改进后的 ICP 为 1 min 以上，文中方法在 1 s 内完成。文中方法的识别速度较为理想，也有较好的准确率，实际应用较为理想。

4 结论

文中提出了一种结合 PSO 算法和 kNN 算法的变电站设备三维云识别方法。使用 PSO 算法优化子空间特征权重，并用 kNN 算法进行识别分类。通过实验分析了这种方法的性能。实验结果表明，该方法的识别精度为 96.67%，平均识别时间为 0.19 s。下一个工作重点是不断改进和完善变电站设备的三维云识别方法。

参考文献

[1] 吕哲，王增平. 基于暂态波形特征的输电线路雷击干扰与故障识别方法[J]. 电力系统保护与控制，2020, 48(6): 41-48.
LÜ Zhe, WANG Zengping. Lightning interference and fault identification method of transmission line based on transient waveform characteristics[J]. Power System Protection and Control, 2020, 48(6): 41-48.

[2] 张婕，曾国辉，赵晋斌，等. 基于改进冒泡排序的模块化多电平换流器电容电压均衡策略[J]. 电力系统保护与控制, 2020, 48(6): 92-99.
ZHANG Jie, ZENG Guohui, ZHAO Jinbin, et al. Capacitor voltage equalization strategy for modular multilevel converter based on improved bubble sorting[J]. Power System Protection and Control, 2020, 48(6): 92-99.

[3] 王菲，王球，任佳依，等. 三维激光云扫描技术在变电站电气设备识别中的应用[J]. 佳木斯大学学报：自然科学版, 2020, 38(4): 124-128.
WANG Fei, WANG Qiu, REN Jiayi, et al. Application of 3D laser cloud scanning technology in substation electrical equipment identification[J]. Journal of Jiamusi University: Natural Science Edition, 2020, 38(4): 124-128.

[4] 纪勇，刘丹丹，罗勇，等. 基于霍夫投票的变电站设备三维点云识别算法[J]. 郑州大学学报：工学版，2019, 40(3): 1-6, 12.
JI Yong, LIU Dandan, LUO Yong, et al. 3D point cloud recognition algorithm for substation equipment based on Hoff voting[J]. Journal of Zhengzhou University: Engineering Edition, 2019, 40(3): 1-6, 12.

[5] 田小壮，赵丰刚，刘海影，等. 基于点云数据的变电站三维仿真模型的实现及展望[J]. 四川电力技术，2018, 41(4): 32-36.
TIAN Xiaozhuang, ZHAO Fenggang, LIU Haiying, et al. Realization and prospect of substation 3D simulation model based on point cloud data[J]. Sichuan Electric Power Technology, 2018, 41(4): 32-36.

[6] 雷亚兰，杜治兴，方艳斌，等. 无人机载激光扫描成像在变电站设备运维管理中的应用[J]. 电子设计工程，2019, 27(14): 185-193.
LEI Yalan, DU Zhixing, FANG Yanbin, et al. Application of UAV laser scanning imaging in substation equipment operation and maintenance management[J]. Electronic Design Engineering, 2019, 27(14): 185-193.

[7] 戴志辉，黄敏，苏怀波. 基于 MMC 的环状直流配网在不同接地方式下的故障特性分析[J]. 电力系统保护与控制, 2019, 47(1): 1-10.
DAI Zhihui, HUANG Min, SU Huaibo. Fault characteristics analysis of circular DC distribution network under different grounding modes based on MMC[J]. Power System Protection and Control, 2019, 47(1): 1-10.

[8] 黄文婧，李华强，杨植雅，等. 基于模糊聚类排序及状态均匀性的电网安全性风险评估[J]. 电测与仪表，2018, 55(9): 21-26.
HUANG Wenjing, LI Huaqiang, YANG Zhiya, et al. Power grid security risk assessment based on fuzzy clustering ranking and state uniformity[J]. Electrical Measurement & Instrumentation, 2018, 55(9): 21-26.

[9] 叶波. 基于负载均衡度的云计算任务调度算法[J]. 东北电力大学学报, 2019, 39(1): 88-95.
YE Bo. Cloud computing task scheduling algorithm

based on load balancing degree[J]. Journal of Northeast Electric Power University, 2019, 39(1): 88-95.

[10] 谢黎, 周华良, 于同伟, 等. 一种智能变电站新型双网冗余设备及实现[J]. 电力系统保护与控制, 2019, 47(11): 151-156.
XIE Li, ZHOU Hualiang, YU Tongwei, et al. A new dual network redundant equipment and its implementation for intelligent substation[J]. Power System Protection and Control, 2019, 47(11): 151-156.

[11] NAHMAN J, SALAMON D. Mutual interference of neighboring grounding systems and approximate formulation[J]. Electric Power Systems Research, 2017, 41(12): 151-156.

[12] MOSES V, HATHERLEY R, BISHOP Ö T. Bioinformatic characterization of type-specific sequence and structural features in auxiliary activity family 9 proteins[J]. Biotechnology for Biofuels, 2016, 9(1): 88-94.

[13] HOU Kaiyuan, SHAO Guanghui, WANG Haiming, et al. Research on practical power system stability analysis algorithm based on modified SVM[J]. Protection and Control of Modern Power Systems, 2018, 3(2): 119-125. DOI: 10.1186/s41601-018-0086-0.

[14] PATEL G K, DABHI V K, PRAJAPATI H B. Clustering using a combination of particle swarm optimization and K-means[J]. Journal of Intelligent Systems, 2017, 12(3): 457-469.

[15] GAUTAM J V, PRAJAPATI H B, DABHI V K, et al. Empirical study of job scheduling algorithms in Hadoop MapReduce[J]. Cybernetics and Information Technologies, 2017, 21(1): 146-163.

[16] CAETANO C E F, LIMA A B, PAULINO J O S, et al. A conductor arrangement that overcomes the effective length issue in transmission line grounding[J]. Electric Power Systems Research, 2018, 46(5): 159-162.

[17] HU Jianjiang, FICHTNER M, BARICCO M. Preparation of Li-Mg-N-H hydrogen storage materials for an auxiliary power unit[J]. International Journal of Hydrogen Energy, 2017, 42(27): 17144-17148.

[18] JIA Zhiwei, WANG Lijun, ZHANG Jinchuan, et al. High efficiency, low power-consumption DFB quantum cascade lasers without lateral regrowth[J]. Nanoscale Research Letters, 2017, 12(1): 88-95.

[19] 孙杰, 李亮曹, 兴砚. 基于三维激光点云的矿区建筑物形变特征提取方法[J]. 激光杂志, 2019, 36(4): 36-40.
SUN Jie, LI Liangcao, XING Yan. A method for extracting deformation features of mining buildings based on three-dimensional laser point cloud[J]. Laser Magazine, 2019, 36(4): 36-40.

[20] 吴凯峰, 刘万涛, 李彦虎, 等. 基于云计算的电力大数据分析技术与应用[J]. 中国电力, 2015, 48(2): 111-116.
WU Kaifeng, LIU Wantao, LI Yanhu, et al. Cloud computing-based large power data analysis technology and application[J]. Electric Power, 2015, 48(2): 111-116.

[21] 王方建, 习晓环, 万怡平. 大型建筑物数字化及三维建模关键技术分析[J]. 遥感技术与应用, 2014, 29(1): 144-150.
WANG Fangjian, XI Xiaohuan, WAN Yiping. Key technologies analysis of digital and three-dimensional modeling of large buildings[J]. Remote Sensing Technology and Applications, 2014, 29(1): 144-150.

[22] 陈朋, 谭晔汶, 李亮. 地面三维激光扫描建筑物点云特征线提取[J]. 激光杂志, 2016, 37(3): 9-11.
CHEN Peng, TAN Yewen, LI Liang. Extraction of point cloud feature lines from ground three-dimensional laser scanning buildings[J]. Laser Magazine, 2016, 37(3): 9-11.

[23] 何原荣, 郑渊茂, 潘火平. 基于点云数据的复杂建筑体真三维建模与应用[J].遥感技术与应用, 2016, 31(5): 1091-1099.
HE Yuanrong, ZHENG Yuanmao, PAN Huoping. True three-dimensional modeling and application of complex buildings based on point cloud data[J]. Remote Sensing Technology and Application, 2016, 31(5): 1091-1099.

[24] 贾雪, 刘超, 徐炜, 等. 海量点云数据的建筑物三维模型重建[J]. 测绘科学, 2019, 44(4): 124-129, 181.
JIA Xue, LIU Chao, XU Wei, et al. 3-D model reconstruction of buildings based on massive point cloud data[J]. Surveying and Mapping Science, 2019, 44(4): 124-129, 181.

[25] 李小梅, 李国柱. 基于移动激光扫描数据的建筑物精细化建模研究[J]. 软件, 2019, 40(8): 163-167, 180.
LI Xiaomei, LI Guozhu. Fine building modeling based on mobile laser scanning data[J]. Software, 2019, 40(8): 163-167, 180.

收稿日期：2020-11-25；　**修回日期：**2021-01-19

作者简介：

李　科(1988—)，男，工程师，主要从事电力设备运维检修、电力数据挖掘分析工作。Email: like20201123@163.com

（编辑　张爱琴）