

危化品仓储中激光测距的自适应分层聚类去噪算法研究

刘学君¹ 魏宇晨¹ 袁碧贤^{1,2} 卢浩^{1,2} 戴波¹ 李翠清³

(北京石油化工学院信息工程学院 北京 102617)¹ (北京化工大学信息科学与技术学院 北京 100029)²
(北京石油化工学院化学工程学院 北京 102617)³

摘要 为实现危化品仓储安全预警,采用激光测距、编码器等进行库内五距的监测。针对测距数据的噪声问题,设计了一种分层去噪、反馈补偿的自适应调节阈值的聚类去噪算法。针对噪声特点及距离的不同,以被测实物为中心将其由远及近划分为 3 类,第一层、第二层使用峰值去噪,第三层使用夹角分段拟合,同时进行层间的反馈纠错实现闭环去噪。实验表明,与拟合去噪算法相比,所提方法的方差降低了 0.83;与差值去噪算法相比,所提方法解决了小部分噪点集中出现而不能去除所带来的灾难性影响的问题,方差值降低了 1.93。该算法可准确去除噪点并还原监测物体位置。

关键词 激光测距,点云,聚类,去噪算法,危化品仓储

中图分类号 TP311.5 **文献标识码** A

Study on Adaptive Hierarchical Clustering De-noising Algorithm of Laser Ranging in Storage of Dangerous Chemicals

LIU Xue-jun¹ WEI Yu-chen¹ YUAN Bi-xian^{1,2} LU Hao^{1,2} DAI Bo¹ LI Cui-qing³

(College of Information Engineering, Beijing Institute of Petrochemical Technology, Beijing 102617, China)¹

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)²

(College of Chemistry Engineering, Beijing Institute of Petrochemical Technology, Beijing 102617, China)³

Abstract In order to realize the safety early warning of dangerous chemical products, this paper used laser ranging and encoder to monitor the five distance in the warehouse. In order to solve the noise problem of ranging data, a new algorithm was designed for reducing noise and feedback compensation. According to the characteristic of the noise and the distance, taking the tested object as the center, objects are divided into three categories from far to near. The first layer and the second layer use peak denoising, the third layer uses piecewise fitting angle, simultaneously interlayer feedback error correction is used to realize the closed loop denoising. Experiment results show that the variance is reduced by 0.83 compared with the denoising algorithm. Compared with the comparison algorithm of difference value, this algorithm can removed catastrophe caused by a small amount of noise concentrated and the variance value is decreased by 1.93. The algorithm can remove the noise and restore the position of the object.

Keywords Laser ranging, Point cloud, Clustering, Denoising algorithm, Storage of dangerous chemicals

1 引言

随着危化品仓储安全监测方面相关研究的不断发展,以及大数据、机械化、智能化的到来,借助于传感器获取、计算机分析等方式,对仓储监测等问题的管理呈现出更高的要求,研究者对其重视程度也明显增加^[1-5]。其中激光扫描所得的相关参数信息具有数字化、清晰化的特点,适用于三维重建等技术,逐渐成为研究热点^[6]。而激光扫描所获得的点云,不可避免会含有噪声点,其原因有很多,诸如物体表面粗糙、波纹、平滑;测量系统本身的设备缺陷;实验中步进电机的抖动、固定架的松动以及光线影响等^[7]。因此,设计一种自适应分层聚类去噪算法,并总结分析提取公式以使其适用于其他传感器所得数据,诸如超声波探测数据、相机标定离散数据等,同时针对不同场景不同情况完成阈值的自动计算与调整,由易到难逐层聚类去噪^[8-11]。

1.1 激光扫描测距简介

危化品仓储激光扫描模块采用激光测距、旋转云台、编码器装置组合而成,其中激光仪采用 650 nm 激光光源,测距精度为 1 mm,测距范围为 70 m,输出功率小于 1 mW,采样频率为 15 Hz。把多个激光测距仪有机组合,形成整体激光扫描阵列^[6-7],如图 1 所示。每个激光测距仪旋转扫描一个平面,激光遇到物体时返回并测得距离,同时编码器记录角度值^[8]。

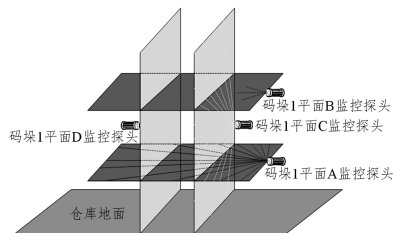


图 1 激光扫描阵列示意图

本文受国家重点研发计划项目(2016YFC0801500),北京市教育委员会市属高校创新能力提升计划项目(2016014222000041)资助。

刘学君(1977—),男,博士,副教授,主要研究方向为光通信、嵌入式开发;魏宇晨(1995—),男,硕士生,主要研究方向为软件嵌入式开发、图像处理, E-mail:1017689090@qq.com(通信作者)。

1.2 研究方案

传感器测得的数据会受光线、测量面粗糙程度等因素影响,在传输过程中时常包含噪点。而由于无法确定阈值,人们经常使用人机交互的方式去除噪点,大大降低了监测的时间,且不适用于实时性的传输环境^[8-11]。因此,根据传感器测距特点(如图 2 所示),设计了一种依据不同情境自主调节阈值的算法。

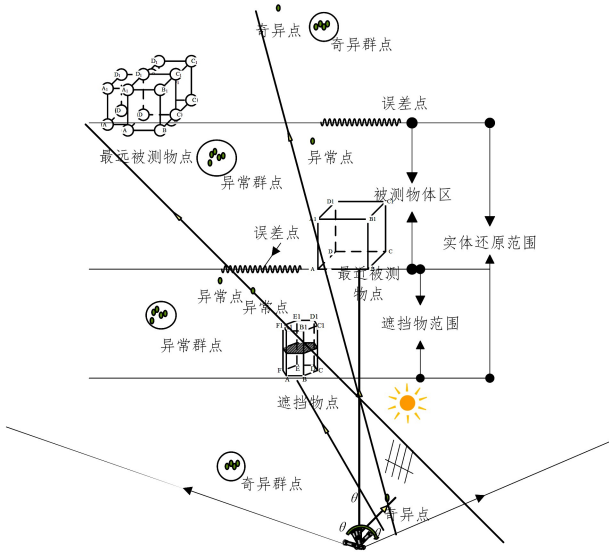


图 2 传感器测距原理示意图

另外,去噪过程中由于噪点位置的不同,不能基于一个阈值整体去噪,故将对噪点的出现情况分层讨论,每一层对初始阈值进行相应的自适应调节,以此聚类样本去除噪声点。为适用于普遍传感器及噪声特点,将层数分为 3 层,第 2 节将详细介绍。由于去噪过程产生了大量计算,算法优化极为重要,因此需要计算机分析当前所处环境从而挑选适合的算法进行处理。为达到最优效果,规定后面的去噪算法层会反馈调节前面的去噪算法,以减小去噪时带来的误差。

2 自适应分层聚类去噪算法

2.1 分层依据及算法概况

基于传感器所得数据中的噪点情况,将单独出现的噪点称为孤立点,小范围集中出现的噪点称为孤立群点,如图 3 所示^[19]。按噪点与正常点距离的远近,可将噪点中的孤立点划分为极大值奇异点、误差点以及极小值奇异点;可将孤立群点划分为极大值奇异群点、误差群点以及极小值奇异群点,如图 4 所示。依据噪声特点以及分布位置的不同,将噪声分为 3 层。首先调节阈值大小,并按层次将第一层与第二层归为一类,使用峰值去噪算法迭代去除。将交织在正常点云附近的第三层噪点单独归为一类,使用夹角分段最小二乘拟合算法,逐段将交织的噪点上下平移为一条直线。

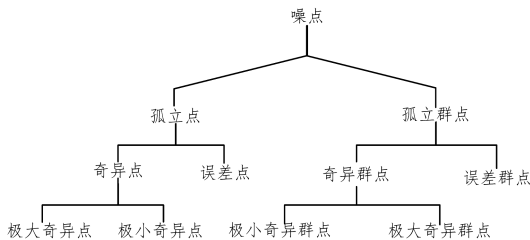


图 3 噪点分类示意图

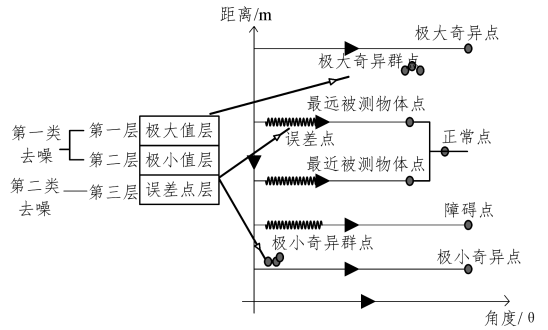


图 4 噪点分层示意图

2.2 初始阈值设定

设置扫描异常点的初始局部阈值,以此进行第一层去噪以及相应阈值调节,完成后续层去噪^[12-18]。根据激光扫描特点可知测量同一水平面激光的最远距离为 L_2 ,本实验激光仪测距最远为 30m,最近距离为 L_1 ,实验编码器的分辨率为 n ,本实验中 n 为 16384,故在步进电机的控制下激光扫描相邻两点的角度值相差公式如式(1)所示。本实验中 $\theta=0.02^\circ$ 。

$$\theta = 360^\circ / n \tag{1}$$

由于微分变化量极小可将弧长近似看成相邻两点的相差距离,即编码器每转动一次最大两点相差距离公式如式(2)所示。本实验中 $d_{max}=0.01047$,考虑到奇异点可能小部分集中出现形成群点,故为每 1° 划分一个范围,求解阈值大小 N 的公式如式(3)所示。

$$d_{max} = (L_2 * 2 * 2 * PI) / 360 * \theta \tag{2}$$

$$N = (L_2 * 2 * PI) / 360 * \theta * (\alpha / \theta) \tag{3}$$

本实验中初始阈值大小 $N=0.523$,针对不同传感器及不同情境可相应调节 α 的角度以改变权值,本实验中 $\alpha=1^\circ$ 。求得初始阈值后进行第一类噪点去除^[6-8]。

2.3 第一类噪声

第一类噪声包括两层,一般由步进电机的抖动、固定架的松动等器材原因以及光线影响所产生,其数值特点极为鲜明,一般远远偏离被测样本方差且出现量极少,将它称为噪声中的奇异(群)点,如图 5 所示。按位于实体还原范围的远近,将噪声分为极大值奇异(群)点和极小值奇异(群)点。针对这类噪声的特点,我们将极大值奇异(群)点放置于第一层首先处理,极小值奇异(群)点放置于第二层进行处理。

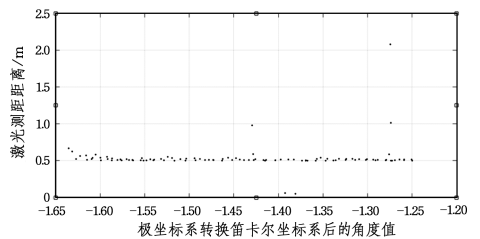


图 5 奇异点分布示意图

首先进行第一层去噪,从样本中选取峰值作为参考点,设 X_1 为参考点横坐标值, Y_1 为参考点纵坐标值, X_2 为参考点左侧相邻横坐标值, Y_2 为参考点左侧相邻纵坐标值,通过两点距离公式(4)计算左侧相邻两点相差距离,右侧相邻两点相差距离的计算同理^[6-8]。

$$\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} \tag{4}$$

若左侧比较距离大于初始阈值则观察右侧情况,若右侧相邻两点比较距离同样也大于初始阈值则标记为噪声进行去除,反之若右侧相邻两点比较后小于阈值范围则将两点归为一类,如图6所示。使用迭代算法将相邻点作为参考点继续重复右侧比较操作,当比较距离大于阈值时将其归于一类的点作为孤立群点去除,反之若右侧相邻两点比较到末尾仍未发现有大于阈值的比较点,则将其归于一类,这些点即为正常点。

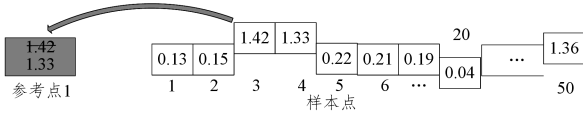


图6 参考点选取示意图

若左侧相邻两点距离比较后小于初始阈值则将两点归为一类,使用迭代算法将相邻点作为参考点继续重复左侧比较操作,若相邻两点比较到末尾仍未发现有大于阈值的比较点,无需比较右侧即为正常点,反之出现大于阈值的比较点,则归入参考点进行右侧比较,方法同理。第一层去噪原理图如图7所示。考虑到算法计算量的问题,选取最值时使用冒泡排序

算法,即在角度 θ 升序矩阵
$$\begin{pmatrix} F_{1,1} & \cdots & F_{1,50} \\ \vdots & \ddots & \vdots \\ F_{n,1} & \cdots & F_{n,50} \end{pmatrix}$$
 中以50个点作为一维数据,求各维数据 $F_{i,50} (\max |\nabla \Delta\theta| (i=1, \dots, n))$ 作为峰值。

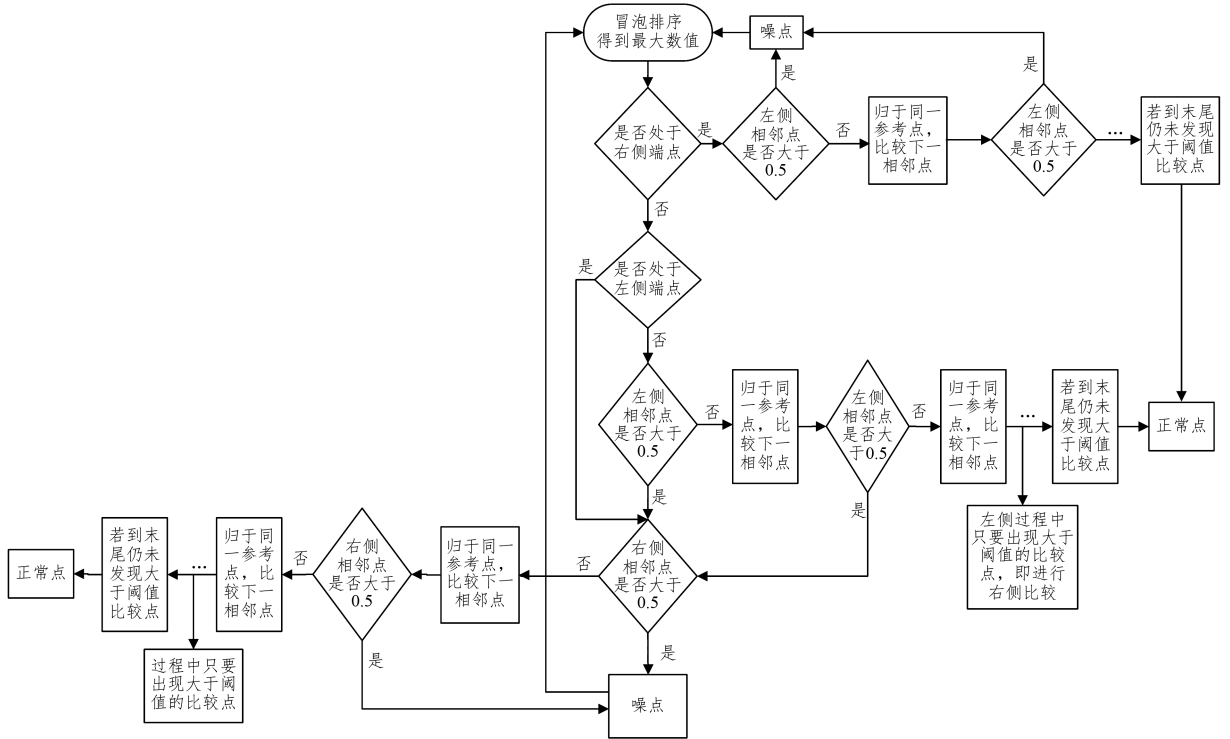


图7 第一类去噪算法流程图

图8为监测不同位置时,每转动 1° 所得的50点数据,经过反复迭代后的方差对比图,其中A处为情境最近监测点,B处为情境最远监测点。可知以选取最值为参考点,依据阈值范围进行的差值迭代比较可明显去除第一类噪声中位于第一层的噪点,即极大值奇异点、极大值奇异群点。

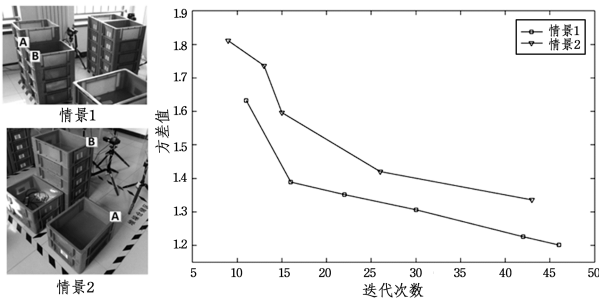


图8 第一层迭代去噪方差结果图

数量少等特点,依旧沿用式(2)、式(3),同时使用上述所证算法,但本层选取的最小值为参考点,依据阈值范围进行差值迭代去除噪点^[19]。由上述算法可知,在选取峰值作为参考点进行迭代比较时,考虑了噪点集中的问题。图9给出传统差值去噪算法情况。

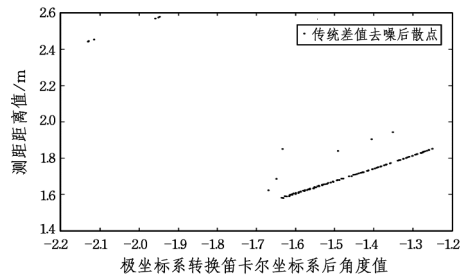


图9 传统差值去噪情况示意图

其次,完成第二层去噪算法,此层噪声特点与第一层噪声特点相似,基本都是由于强紫外线等光线影响以及器械抖动等原因产生噪点,使得激光在未达到被测物体或遮挡物时就折射回来,称其为极小值奇异(群)点。根据本类噪声数值小、

图10为改良后双峰法去除第一类噪点的情况。可以明显看出异常点小部分集中出现时,传统差值不能有效去除噪点,尽管有后续层的进一步补救措施,但也极大影响了图像还原的准确度。

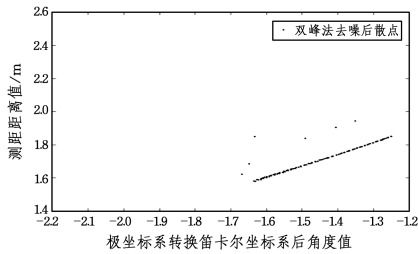


图 10 双峰法去噪情况示意图

2.4 第二类噪声

去除完第一类噪声后,剩下样本点包含于正常点及夹杂于正常点周围的误差点,其属于第二类噪声,主要产生原因分为两类:外部原因和内部原因^[20]。这类噪声点的去除比第一类更为复杂。

内部噪声可能由于电子元件的光和电的基本性质所引起,如电子或空穴粒子的随机运动形成的散粒噪声或导体中自由电子的无规则热运动所形成的热噪声;也可能由于器材材料本身引起的噪声,如正片和负片的表面颗粒性和磁带磁盘表面缺陷所产生的噪声。

外部噪声即指系统外部干扰以电磁波或经电源串进系统内部而引起的噪声,如电气设备、天体放电现象等引起,而这种噪声可能就是高斯噪声、脉冲噪声等多个噪声合成累计的。综上所述虽然其形式复杂,但仍具有规律特点,即含噪数据由原始正常数据点在某段中附加上随机噪声所形成,用数学公式可以描述为:

$$G(x, y) = F(x, y) + V(x, y) \quad (5)$$

其中, $F(x, y)$ 表示正常数据点, $V(x, y)$ 为噪声点,某段含噪数据记为 $G(x, y)$ 。基于本类噪声特点,第三层去噪使用分部拟合去噪算法,即逐段分类拟合。最小二乘的原理是在研究两个变量 (x, y) 之间的相互关系时,得到一系列成对的数据 $(x_i, y_i), i=1, 2, 3, \dots, n$ 。设经验方程为 $Y(x, y)$, 其中包含待定系数 a_n , 将 (x_i, y_i) 代入方程求差 $y_i - Y(x)$, 并以此取其平方和防止正负差值抵消,如式(6)所示:

$$E = \sum (y_i - Y(x))^2 \quad (6)$$

借助上述式子通过求 e 的极小值可以求出 a_n , 从而求出该组数据的最佳拟合函数,该函数使得误差平方和最小。假设经验方程是线性的,那么利用最小二乘法就可得到某一区间内的线性回归,利用此点性质结合上文去噪所得数据加以讨论,使其适用于传感器测距图像还原等领域。

沿用 2.1 节设置阈值的方法,但为适应本层去噪特点对其进行改进。首先仍将每 1° 划分为一个范围,即扫描 50 个点为一个范围,所以将 n 个有序序列分为 k 组:

$$N_1 = (x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_m, y_m)$$

$$N_2 = (x_{m+1}, y_{m+1}), (x_{m+2}, y_{m+2}), \dots, (x_n, y_n)$$

⋮

$$N_n = (x_{n+1}, y_{n+1}), \dots, (x_{n+n}, y_{n+n})$$

其次寻找这 50 个点的最大值即为点 $V_{\max}(x_i, y_j)$ 和最小值即为点 $V_{\min}(x_i, y_j)$, 设定起始端点为 $B(x_i, y_j)$, 末尾结点为 $E(x_i, y_j)$, 如图 11 所示^[21]。

直线 L_1 的一般式方程为 $Ax + Bx + C = 0$, 法向量 $\vec{n} = (A, B)$; 直线 L_2 的一般式方程为 $A'x + B'y + C' = 0$, 法向量 $\vec{n}_2 = (A', B')$ 。可得 $\cos \alpha = |\cos \langle \vec{n}_1, \vec{n}_2 \rangle| = \left| \frac{\vec{n}_1 * \vec{n}_2}{|\vec{n}_1| * |\vec{n}_2|} \right|$, 由此分别计算 $\angle BV_{\max}E$ 和 $\angle BV_{\min}E$, 如式(7)所示:

$$\alpha = \arccos \left| \frac{AA' + BB'}{\sqrt{A^2 + A'^2} \sqrt{B^2 + B'^2}} \right| \quad (7)$$

若只有 $\angle BV_{\max}E$ 的夹角大于 90° , 则记录最大值位置, 并且分部拟合范围阈值从起始端点到末尾结点变为起始端点到最大值点和最大值点到末尾结点两部分; 若 $\angle BV_{\min}E$ 大于 90° 同理; 若 $\angle BV_{\max}E$ 和 $\angle BV_{\min}E$ 都大于 90° , 则选取位置点较小的作为参考点以此分为两段。图 12 是分层去噪与单独使用分部拟合去噪的对比图。

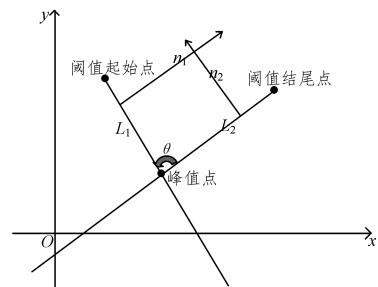


图 11 分段拟合夹角示意图

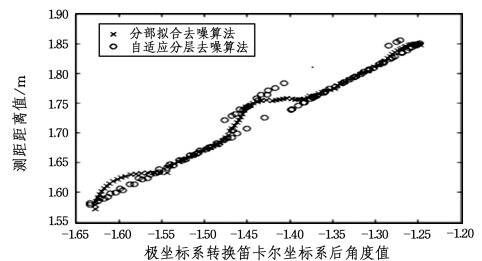


图 12 夹角分段拟合与传统拟合的比较示意图

3 实验结果

对于第一类噪声的去除,在选取峰值时使用了以下 3 种算法进行对比,其中冒泡法遍历峰值的平均时间复杂度为 $O(n^2)$ 。加权枚举法即首先随机选取 $F(x)$ 一点与其他点 $Y(x_i) (i=1, \dots, n)$ 进行比较, 选取出比 $F(x)$ 大的点作为加权点 $Z(x_i) (i=1, \dots, n-m)$, 以此循环反复求解峰值。分治算法遍历是传统冒泡遍历的改良版, 将数据一分为二同时冒泡遍历两块, 求取两块最大峰值 $F_1(x)$, 故另一点即为 $Y(x)$, 其后遍历 $F_1(x)$ 一侧数据与 $Y(x)$ 进行对比, 较大者即为 $F_2(x)$, 随后以此类推。

表 1 遍历峰值数据时间统计

遍历峰值 算法	找寻一次 峰值时间	找寻两次 峰值时间	找寻 3 次 峰值时间	找寻 10 次 以上峰值时间
冒泡算法遍历	0.021	0.033	0.069	0.322
加权枚举法遍历	0.040	0.064	0.071	0.216
分治算法遍历	0.025	0.031	0.054	0.082

(单位: s)

- [13] WEIJER J, GEVERS T. Boosting color saliency in image feature [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(1): 150-156.
- [14] 徐天芝, 张贵仓, 贾园. 基于形态学梯度的分水岭彩色图像分割 [J]. 计算机工程与应用, 2016, 52(11): 200-203, 208.
- [15] NITHYA A, KAYALVIZHI R. Medical Image Segmentation using Modified Morphological Reconstruction [J]. International Journal of Computer Applications, 2014, 86(2): 20-26.
- [16] 黄展鹏, 张琦, 赵洁. 基于自适应标记分水岭算法的肝脏 CT 图

像自动分割 [J]. 北京生物医学工程, 2017, 36(4): 378-382.

- [17] 汪梅, 李琳, 王斌, 等. 分水岭算法应用于主动轮廓模型能量分割算法的研究 [J]. 计算机科学, 2017, 44(5): 314-319.
- [18] 缪慧司, 梁光明, 刘任任, 等. 结合距离变换与边缘梯度的分水岭血细胞分割 [J]. 中国图象图形学报, 2016, 21(2): 192-198.
- [19] 余焯, 李冰飞, 张小魏, 等. 面向 RGBD 图像的标记分水岭分割 [J]. 中国图象图形学报, 2016, 21(2): 145-154.
- [20] 张海涛, 李雅男. 阈值标记的分水岭彩色图像分割 [J]. 中国图象图形学报, 2015, 20(12): 1602-1611.

(上接第 211 页)

由表 1 可以看出, 当寻找第一类样本噪点的次数小于 10 时, 可选用普通的冒泡算法, 大于 10 次可使用改良后的冒泡算法, 即分治算法遍历。但一般我们无法预知样本噪点情况, 所以在样本点不多的情况下使用第一种或第二种均可。表 2 为自适应分层去噪算法与传统差值算法及传统分段拟合算法的数据对比, 可知此类算法能有效保留数据点且恢复被测物体数据。

表 2 去噪算法效果对比统计

数据点类别	长方体障碍物	圆柱体障碍物	锥体障碍物
分层聚类去噪算法总数据数	103	103	103
传统差值去噪算法总数据数	103	103	103
传统分段拟合算法总数据数	103	103	103
分层聚类去噪后数据点数	103	103	103
传统差值去噪后数据点数	102	99	101
传统分段拟合去噪后数据点数	103	103	103
分层聚类去除噪点数	12	15	6
传统差值去除噪点数	5	7	4
传统分段拟合去除噪点数	14	20	16
分层聚类去噪后方差值	0.83	0.61	0.57
传统差值去噪后方差值	2.76	0.85	0.72
传统分段拟合去除后方差值	1.21	1.03	0.22

结束语 激光扫描等传感器获得的大量数据, 除了正常数据点外, 还包含了本文所述的个别奇异点以及大量误差点, 这些噪点分布的位置不同且特点也不同, 因此采用了自适应分层聚类去噪算法。由实验结果可知, 该算法可准确去除噪点并还原监测物体位置。与分部拟合算法相比, 其细化了阈值的设定并且在拟合前滤除了奇异点等大噪声点的干扰; 与差值去噪算法相比, 其考虑了噪点集中出现等问题且可由后一层的去噪弥补多去除的正常点和少去除的噪声点。最后做了算法量化比较, 使得该算法在成功还原数据的同时极大降低了计算量。

后续应进一步对自适应改变阈值的方法加以讨论和优化, 并且需进一步开发人机交互界面, 使得在一次输入传感器规格数值后, 通过读取捕获数据完成自动化去噪、传输等后续操作。

参 考 文 献

- [1] 潘爱华. 贯彻落实《中国制造 2025》纲要精神以信息化智能化促进危化品行业安全发展 [J]. 中国石油和化工经济分析, 2016, 1(1): 5-7.
- [2] 裘静. 探讨港口危险品的仓储安全管理问题思考 [J]. 中国金属通报, 2016(6): 61-62.

- [3] 师立晨, 王如君, 多英全. 我国危险化学品重大危险源安全监管存在问题及建议 [J]. 中国安全生产科学技术, 2014, 30(12): 161-166.
- [4] 天津港“8·12”特别重大火灾爆炸事故调查报告公布 [J]. 消防界(电子版), 2016(2): 35-40.
- [5] 中国安科院参与的国家 863 重点课题“危化品生产储存使用安全监测系统开发及应用示范”试点项目开始现场施工工作. 中国安全生产科学技术 [J]. 中国安全生产科学技术, 2008, 15(8): 94.
- [6] 刘学君, 江帆, 戴波, 等. 基于激光测距仪的危化品仓库堆垛安全距离的网格化监测方案 [C] // 中国自动化大会, 2015: 21-50.
- [7] 刘学君, 江帆, 戴波, 等. 基于 ARM 的危化品仓库堆垛安全距离监测装置的研究与开发 [J]. 制造业自动化, 2016, 38(4): 11-14, 25.
- [8] 刘学君, 江帆, 戴波, 等. 危化品仓库堆垛安全距离的精确扫描定位监控算法的研究 [J]. 计算机测量与控制, 2016, 24(11): 73-76.
- [9] 戴波, 吕昕, 刘学君, 等. 基于改进的超宽带危化品仓储堆垛定位技术 [J]. 上海交通大学学报, 2016, 50(6): 963-967.
- [10] 戴波, 吕昕, 刘学君, 等. 基于 UWB 四参考点矢量补偿的危化品仓储堆垛货物定位方法 [J]. 化工学报, 2016, 67(3): 871-877.
- [11] 戴波, 李志超, 刘学君, 等. 基于泰森多边形的 UWB 危化品堆垛仓储货物定位技术 [J]. 化工学报, 2016, 67(3): 878-884.
- [12] 韦航, 卢必松, 陆斌. 物证检验照相中噪点的产生原因及解决方案 [J]. 职业圈, 2007, 15(1): 144-145.
- [13] 贺峰, 刘承兰, 朱国富, 等. 基于距离和差值序列的复杂运动目标微动特征提取 [J]. 电子学报, 2012, 40(3): 422-428.
- [14] 张宏伟, 赖百炼. 三维激光扫描技术特点及其应用前景 [J]. 测绘通报, 2012, 10(10): 320-322, 337.
- [15] 周军, 朱博楠, 杨圣强, 等. 基于动态差值法的直流系统绝缘监测技术 [J]. 电工技术学报, 2015, 30(1): 235-241.
- [16] 邓文君, 叶景杨, 张铁. 面向机器人磨抛的激光点云获取及去噪算法 [J]. 光学学报, 2016, 17(30): 180-188.
- [17] 李明磊, 李广云, 宗文鹏. 激光扫描点云准确快速去噪方法 [J]. 测绘通报, 2015, 25(12): 27-29.
- [18] 陈晓曦, 王延杰, 刘恋. 小波阈值去噪法的深入研究 [J]. 激光与红外, 2012, 20(1): 105-110.
- [19] 徐鹏, 惠振阳. 基于 MATLAB 实现点云噪声剔除算法研究 [C] // 贵州省岩石力学与工程学会学术年会, 2014.
- [20] 蔺素珍, 韩泽. 基于深度堆叠卷积神经网络的图像融合 [J]. 计算机学报, 2017, 15(11): 2506-2518.
- [21] 田垅, 刘宗田. 最小二乘法分段直线拟合 [J]. 计算机科学, 2012, 15(6): 482-484.