

基于加权 FCM 和法向离群因子文物点云去噪

宋凯瑞¹ 汤 慧¹
SONG Kairui TANG Hui

摘 要

针对已有点云去噪算法在处理文物碎片点云时存在几何特征保持性较差、去噪精度较低且时间复杂度较高等问题,提出一种基于模糊加权 C-均值聚类(FCM)和法向离群因子双边滤波的点云去噪算法。首先,利用模糊加权 FCM 算法对大尺度噪声进行初步去除,确保对文物整体形状的良好保持。然后,在双边滤波算法中引入法向离群因子,并采用该算法进行小尺度噪声去除,实现文物点云的精细去噪。最后,采用含噪的公共点云模型 Bunny 和 Dragon 以及四组含噪文物碎片点云模型进行实验。实验结果表明,提出的算法对比文献算法去噪精确度提高了约 15%,耗时降低了约 20%,有效保持了文物碎片点云尖锐特征,为数字化文物保护和复原工作提供了重要的技术支持。

关键词

点云去噪;特征保持;模糊加权;法向离群因子;双边滤波

doi: 10.3969/j.issn.1672-9528.2024.09.012

0 引言

三维激光扫描技术是获取物体空间点云数据信息的重要手段之一,目前已在逆向工程、虚拟现实、数字城市建设、无人驾驶车辆技术以及生物医学研究等领域发挥了重要作用^[1-5]。然而,在实际的三维激光扫描过程中,点云数据的质量会受到多种因素的影响,包括外部环境的不稳定性和扫描设备的固有局限性。受这些因素的影响,采集到的点云数据含有大量噪声,给后续的数据处理(如精确特征提取、建模、虚拟重建等应用)带来诸多不便^[6]。在文物复原过程中,三维点云噪声的去除更是必不可少的工作。因此,需要选择合适的算法对点云噪声进行处理。

近年来,国内外学者对点云去噪算法进行了大量研究。例如,廖中平等^[7]利用异向法矢平滑处理修正法向量的双边滤波算法对噪声点进行去除,相比于传统双边滤波算法提高了特征保持性,但对于离群噪声点去除效果不理想;苏本跃等^[8]利用 K-means 聚类算法去除点云噪声,该算法运行速度明显提高,但去噪精度不够且效果难以掌控;Fleishman 等^[9]将双边滤波的方法应用到三维网格的去噪上,对原始点云数据进行调整以实现去噪;Duan Yao 等^[10]通过提取主成分将三维数据降维生成二维数据,在有效去除噪声的同时保留目标特征的细节信息,但存在去噪精度不够的问题;李仁忠等^[11]利用统计滤波、半径滤波和双边滤波等方法库算法对点云噪声进行去除,但其阈值设置

会影响算法的运行效率;张宏伟^[12]利用动态半径滤波和改进双边滤波算法去噪;焦亚男等^[13]利用两种结构结合的双边滤波算法进行去噪,效果较好,但只适用于有序点云,现实中不易实现;任彬等^[14]将点云噪声划分为远信号噪声点和近信号噪声点,分别利用非线性函数的阈值自适应算法和基于曲率的算法进行去噪。

以上算法在点云特征保持性上取得了一定成效,但在文物点云去噪精度和时间效率上还有待提高。鉴于此,本文提出一种基于加权 FCM 和法向离群因子的点云去噪算法,通过将点云噪声分为大尺度噪声和小尺度噪声分别进行处理。采用引入模糊加权系数的 FCM 算法去除大尺度噪声,采用法向离群因子双边滤波算法去除小尺度噪声。该算法不仅在点云去噪精度和时间效率上有所提高,还能够较好保持文物碎片点云模型原有的几何特征。

1 基于加权 FCM 的大尺度去噪

大尺度噪声点往往在数据中形成较为明显的簇状结构或分布区域,与周围的正常数据点在特征上有显著差异。通过引入模糊加权系数,基于加权 FCM 算法使得点云每个数据点的隶属度不仅考虑了点与聚类中心之间的相似性,还通过调整模糊加权系数,使得大尺度噪声点在隶属度上表现出显著的差异,能够更有效地去除大尺度噪声。

1.1 FCM 算法

模糊 C-均值聚类(fuzzy c-means algorithm, FCM)是一种基于目标函数的模糊聚类算法,它通过迭代减小目标函数误差来得出最终的聚类结果^[15]。

1. 西安财经大学信息学院 陕西西安 710100

设待去噪点云为 $P = \{p_i\}$ ($i = 1, 2, \dots, n$), n 为点云 P 中点的个数, 对其采用 FCM 算法进行去除, 将任意点集合 $\{p_i\}$ 分为 m 个聚类组 $V = (v_1, v_2, \dots, v_m)$ 。

假设每组的聚类中心为 v_j , 则目标函数为:

$$J_c = \sum_{i=1}^n \sum_{j=1}^m \mu_{ij}^c \|p_i - v_j\|^2 \quad (1)$$

式中: n 为点云 P 中点的个数; m 为聚类组的个数; c 为加权指数; $\|p_i - v_j\|$ 表示第 i 个点云点 p_i 与第 j 个聚类中心 v_j 之间的欧式距离; μ_{ij} 表示样本 x_i 对聚类中心 v_j 的隶属度。 μ_{ij} 满足以下条件:

$$\begin{cases} 0 \leq \mu_{ij} \leq 1, \forall i, j \\ \sum_{i=1}^n \mu_{ij} = 1, \forall j \\ 0 < \sum_{j=1}^m \mu_{ij} < m, \forall i \end{cases} \quad (2)$$

1.2 加权 FCM 算法

该算法通过在 FCM 中引入模糊加权系数, 使得欧氏空间中数据点的特征程度不同, 距离远的变得更远, 距离近的变得更近^[16]。模糊加权系数定义为:

$$W_j = \mu_{ij}, i=0, 1, \dots, n \quad (3)$$

$$W_{ij} = \frac{\mu_{ij}}{W_j} \quad (4)$$

则模糊加权距离为 $\left\| \frac{1}{W_{ij}} (p_i - v_j) \right\|^2$ 。

将模糊加权距离引入 FCM 算法中, 目标函数 (1) 转化为:

$$J_c' = \sum_{i=1}^n \sum_{j=1}^m \mu_{ij}^c \left\| \frac{1}{W_{ij}} (p_i - v_j) \right\|^2 \quad (5)$$

鉴于较大的数据集可能导致聚类复杂性增加, 该算法只考虑在以数据点 p_i 为球心的包围球内进行聚类。如果在包围球内的数据点数量低于设定的阈值 ε , 那么这些数据点 p_i 被识别为噪声点并会被直接移除。相反, 如果包围球内的数据点数量高于阈值, 则它们被认为是正常点, 并会经过改进的 FCM 聚类算法进行调整, 以接近聚类中心。

基于加权系数的 FCM 算法去噪具体步骤总结如下。

- (1) 设包围球内数据点个数阈值 ε , 以及包围球半径 r 。
- (2) 计算以数据点 p_i 为中心, r 为半径的包围球包含的点数 N_r 。
- (3) 对包围盒包含的点数 N_r 以及阈值 ε 进行比较, 若 $N_r < \varepsilon$, 则认为该数据点 p_i 为噪声点, 直接剔除。
- (4) 若 $N_r \geq \varepsilon$, 则认为该数据点为正常点, 对该点所在包围球内所有数据点进行模糊加权聚类。
- (5) 计算出聚类中心, 用其代替该包围球内所有数据点。
- (6) 重复步骤 2 ~ 5, 对所有数据点进行处理, 最终得

到去除大尺度噪声后的点云 P' 。

采用以上引入模糊加权系数的 FCM 算法, 可以有效去除点云大尺度噪声。

2 双边滤波算法

点云去除大尺度噪声后, 对于其散布在数据中, 与周围正常数据点的特征差异较小的噪声点, 引入一种基于法向离群因子的双边滤波算法。该算法考虑了点云中每个点的法向信息, 通过结合空间距离和法向相似性, 能够更有效地区分小尺度噪声和正常数据点。

2.1 双边滤波算法

传统双边滤波算法是一种非线性滤波方法, 它利用邻近数据点的加权平均来调整当前数据点, 以改善滤波结果。对点云中的每个点进行加权平均, 可以实现去除噪声的效果, 同时保持点云的细节和边缘信息。

双边滤波表达式为:

$$p_i' = p_i + \alpha \cdot n_i \quad (6)$$

式中: p_i' 为双边滤波后的点云数据点, p_i 为除去噪后的点云数据点, α 为双边滤波权重因子, n_i 为点 p_i 的法矢。

$$\alpha = \frac{\sum_{p_j' \in N(p_i')} w_c(\|p_j' - p_i'\|) w_s(\langle n_i, p_j' - p_i' \rangle) \langle n_i, p_j' - p_i' \rangle}{\sum_{p_j' \in N(p_i')} w_c(\|p_j' - p_i'\|) w_s(\langle n_i, q_j - q_i \rangle)} \quad (7)$$

式中: $N(p_i')$ 为数据点 p_i' 的邻域点集合, n_i 为数据点 p_i' 的法向量。

定义光顺平滑权重函数 $w_c(x)$ 为:

$$w_c(x) = \exp[-x/2\sigma_c^2] \quad (8)$$

式中: σ_c 为数据点 p_i' 到邻近点距离对 p_i' 的影响因子, σ_c 的取值大小会对滤波结果产生影响, 较大的取值将会有更好的去噪效果; x 为数据点 p_i' 与邻域点 p_j' 之间的距离, 其表达式为:

$$x = \|p_j' - p_i'\| \quad (9)$$

定义特征保持权重函数 $w_s(y)$ 为:

$$w_s(y) = \exp[-y/2\sigma_s^2] \quad (10)$$

式中: σ_s 为数据点 p_i' 到邻近点距离在其法向上的投影对数据点 p_i' 的影响因子, σ_s 的取值大小会对点云数据的几何特征信息保持程度产生影响, 较大的取值将会更好地保持几何特征; y 为数据点 p_i' 的法向量和邻域点 p_j' 的法向量的内积, 其表达式为:

$$y = \|n_i - n_j\| \quad (11)$$

式中: n_j 为数据点 p_j' 的法向量。

由于数据点的邻域特征对原始双边滤波因子有一定的影响, 因此对其进行改进, 以提高其稳健性和保持特征性^[17]。改进后的双边滤波因子为:

$$\alpha' = \frac{\sum_{p'_j \in N(p'_i)} w_c(\|p'_j - p'_i\|_2) w_s(\langle n_i, p'_j - p'_i \rangle) \langle n_i, p'_j - p'_i \rangle}{\sum_{p'_j \in N(p'_i)} w_c(\|p'_j - p'_i\|_2) w_s(\langle n_i, n_j \rangle - 1)} \quad (12)$$

2.2 引入法向离群因子

双边滤波算法通常将邻域中的所有点放在一起考虑,这可能会导致两个问题:首先,它们对 k 邻域的选择非常敏感;其次,杂散噪声可能会错误定位正确的点。法线方向上的投影距离可用于测量邻域与测量点之间的相似性,计算这些投影距离的平均值 μ 和标准差 σ ,可以用来检测法向方向上的异常点。在有限的邻域范围内,当投影距离大于 σ 时,将该邻域点视为法向离群点, u 为离群倍数,用于控制去除异常值的灵敏度,引入法向离群因子^[18];当邻域点为法向离群点时, d 为0,反之,则 d 为1。

引进法向离群因子后的双边滤波因子为:

$$\alpha'' = \frac{\sum_{p'_j \in N(p'_i)} d \cdot w_c(\|p'_j - p'_i\|_2) w_s(\langle n_i, p'_j - p'_i \rangle) \langle n_i, p'_j - p'_i \rangle}{\sum_{p'_j \in N(p'_i)} d \cdot w_c(\|p'_j - p'_i\|_2) w_s(\langle n_i, n_j \rangle - 1)} \quad (13)$$

式中: d 为法向离群因子。

高斯核函数相关计算式为:

$$\omega = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (14)$$

$$\sigma^2 = \frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \dots + (x_n - \mu)^2}{n-1} \quad (15)$$

式中: μ 是相应样本数据平均值; σ^2 是样本数据无偏估计方差。

在点云双边滤波因子权函数计算式中,可依式(14)和式(15),利用各个量值与其均值的差异性,通过计算数据点在邻域内与各点对应的范数标量,采用各个数据邻域点对影响因子 σ_c 和 σ_s 进行无偏估计,构造自适应影响因子 σ_c^2 和 σ_s^2 ,其表达式分别为:

$$\sigma_c^2 = \frac{\|p'_1 - p'_i\|_2^2 + \|p'_2 - p'_i\|_2^2 + \dots + \|p'_n - p'_i\|_2^2}{n-1} \quad (16)$$

$$\sigma_s^2 = \frac{|\langle n_i, n_1 \rangle - 1|^2 + |\langle n_i, n_2 \rangle - 1|^2 + \dots + |\langle n_i, n_n \rangle - 1|^2}{n-1} \quad (17)$$

得到相应的光顺平滑权函数 w_c 和特征保持权函数 w_s ,其表达式分别为:

$$w_c = \exp[-\|p'_j - p'_i\|_2^2 / 2\sigma_c^2] \quad (18)$$

$$w_s = \exp[-|\langle n_i, n_j \rangle - 1|^2 / 2\sigma_s^2] \quad (19)$$

基于法向离群因子双边滤波去噪具体步骤总结如下。

(1) 计算每个点云数据点 p'_i 的 k 个最近邻域点集合 $N(p'_i)$ 。

(2) 计算数据点 p'_i 与邻域点 p'_j 之间的距离 x 和数据点 p'_i 的法矢和邻域点 p'_j 的法矢的内积 y 。

(3) 计算自适应参数 σ_c^2 和 σ_s^2 。

(4) 将计算得到的参数 x 、 y 、 σ_c^2 和 σ_s^2 代入光顺平滑权函数 w_c 和保持特征权函数 w_s 中,求出函数。

(5) 将 w_c 和 w_s 代入引进法向离群因子后的双边滤波因子中,计算双边滤波因子。

(6) 利用双边滤波表达式计算新数据点 p'_i 的位置,建立新的点云位置。

(7) 将每个点云数据都进行上述操作,即得到去除小尺度噪声的点云 p'' ,算法终止。

3 评价指标

3.1 倒角距离

倒角距离(chamfer distance, CD)在三维点云中常用于三维重建网络损失函数的设计,其定义为:

$$c(P'', Q) = \frac{1}{N} \sum_{p'' \in P''} \min_{q_i \in Q} \|p'' - q_i\|_2 + \frac{1}{M} \sum_{q_i \in Q} \min_{p'' \in P''} \|q_i - p''\|_2 \quad (20)$$

式中: P'' 为去噪后的点云, Q 为原始无噪声点云, N 和 M 分别是两个点云的个数。

$c(P'', Q)$ 越小, P'' 与 Q 的分布就越相似,这意味着它更接近真实表面,也就是能更有效地去除噪声。

3.2 均方误差

均方误差(mean squared error, MSE)是去噪后点云各数据点偏离原始点云数据点的平方和的平均数,作为去噪后效果的客观衡量指标,定义为:

$$E_{MSE} = \frac{1}{M} \sum_{i=1}^M (P''_i - Q_i)^2 \quad (21)$$

式中: M 表示 Q 中所有点的个数。

MSE用于计算 Q 距离的平均值,该平均值越小时, P'' 与 Q 越接近,去噪效果越好。

4 实验与结果分析

实验先后采用公共点云模型和文物点云模型验证所提算法。首先进行大尺度点云噪声去除,再进行小尺度噪声去除。首先以Bunny和Dragon模型为例,去噪结果如图1和图2。



(a) 含噪 Bunny 模型 (b) 大尺度去噪 (c) 小尺度去噪

图1 Bunny模型去噪



(a) 含噪 Dragon 模型 (b) 大尺度去噪 (c) 小尺度去噪

图2 Dragon模型去噪

为验证本文算法的性能,对 Bunny 和 Dragon 含噪点云模型分别采用文献[19]算法、文献[20]算法对其进行去噪,去噪最终结果对比分别如图3~图4和见表1。

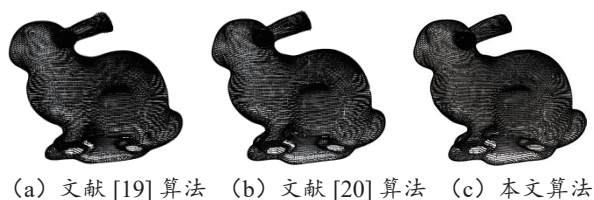


图3 Bunny去噪算法对比



图4 Dragon去噪算法对比

表1 不同算法对公共点云去噪结果对比

公共点云	去噪算法	耗时/s	倒角距离/ $\times 10^{-8}$	均方误差/ $\times 10^{-8}$
Bunny	文献[19]算法	22.1	1.519 9	1.897 6
	文献[20]算法	20.8	1.354 7	1.646 8
	本文算法	13.5	1.080 2	1.339 4
Dragon	文献[19]算法	24.9	3.220 7	1.175 1
	文献[20]算法	22.5	3.020 1	0.927 0
	本文算法	16.4	2.794 3	0.691 3

从公共点云数据模型的去噪结果可以看出,对比文献[19]算法和文献[20]算法,本文算法的去噪结果耗时最短,倒角距离和均方误差相比其他算法更小,去噪效果良好,同时保持了点云清晰的几何特征。

对文物碎片点云模型进行实验,以四组含噪文物碎片点云模型为例,如图5。

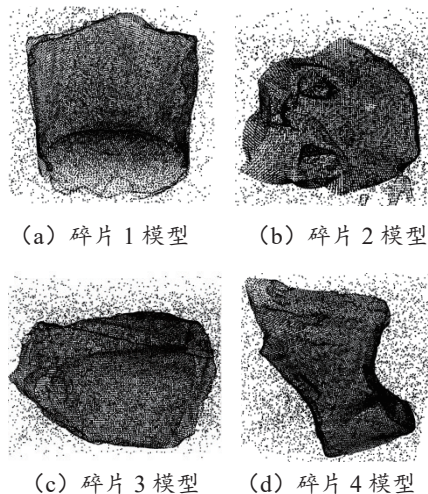


图5 含噪文物碎片模型

对以上四组含噪文物碎片点云模型,采用文献[19]算法、文献[20]算法和本文算法进行去噪,最终去噪结果如图6~图9和见表2。

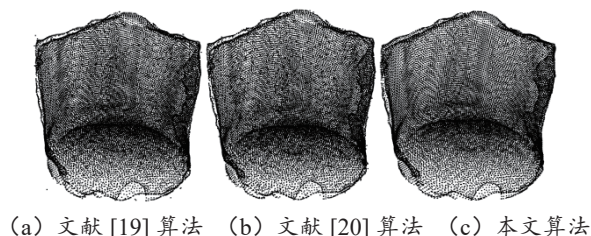


图6 碎片1去噪结果

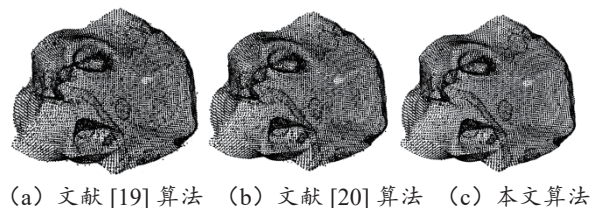


图7 碎片2去噪结果

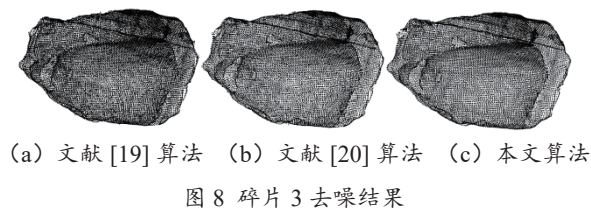


图8 碎片3去噪结果

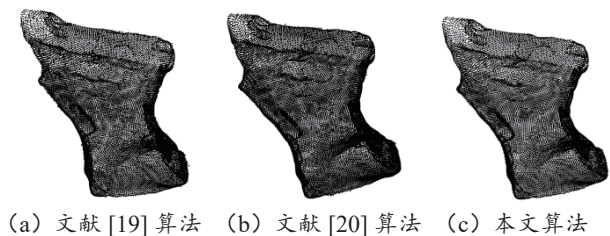


图9 碎片4去噪结果

表2 不同算法对文物点云去噪结果对比

文物碎片	去噪算法	耗时/s	倒角距离/ $\times 10^{-8}$	均方误差/ $\times 10^{-8}$
碎片1	文献[19]算法	24.1	1.710 4	2.005 1
	文献[20]算法	22.5	1.527 1	1.753 3
	本文算法	15.4	1.269 2	1.430 6
碎片2	文献[19]算法	25.7	1.861 1	2.161 7
	文献[20]算法	24.2	1.679 0	1.903 2
	本文算法	16.9	1.420 5	1.579 2
碎片3	文献[19]算法	23.5	1.762 8	1.963 9
	文献[20]算法	22.1	1.594 3	1.795 2
	本文算法	16.2	1.306 1	1.482 9
碎片4	文献[19]算法	22.8	1.702 2	1.849 2
	文献[20]算法	20.7	1.573 2	1.628 4
	本文算法	13.5	1.283 9	1.328 4

从文物碎片点云模型的去噪结果表明,相比文献[19]算法和文献[20]算法,本文算法去噪后的碎片模型不仅除去了大尺度离群噪声点,还更好地去除了小尺度噪声点,可以在保持点云尖锐几何特征的基础上达到最佳的去噪效果,是一种精确快速的点云模型去噪算法。这是由于文献[19]算法是采用引入采样点曲率改进双边滤波因子,对非离群点噪声进行去噪,该算法对非离群点噪声具有良好的去噪效果,但是对离群点噪声的识别和去噪效果不佳;文献[20]算法是一种统计滤波和空间密度聚类相结合的去噪算法,先后利用统计滤波和空间密度聚类来去除离群点,该算法对于高噪声点可以有效去除,但对低噪声点的去除不够显著;本文所提去噪算法采用加权FCM算法实现大尺度噪声删除,采用法向离群因子双边滤波算法实现小尺度噪声删除,可以有效识别两种不同类型的噪声,达到较高的去噪精度,较好保持文物碎片点云模型原有的几何特征。

5 总结

在三维文物点云预处理中,有效的点云去噪算法不仅保留原始数据的特征信息,还有较高的去噪精度和速度。本文提出了一种分层次去噪的算法,分别采用加权FCM算法和法向离群因子双边滤波算法对不同尺度噪声点进行去除。利用公共和文物点云数据模型进行实验,结果表明,本文算法有效识别了不同尺度的噪声,相比其他两种算法去噪精确度提高了约15%,耗时降低了约20%,且较好地保持了文物点云模型原有几何特征。该算法在设定阈值和计算邻域信息时还有一定的主观性,如何设置最优阈值与计算更好的邻域信息是下一步研究的重点。

参考文献:

- [1]XIAO J, QU W, JIANG H, et al.Three-dimensional fractal charaterization of concrete surface subjected to sulfuric acid at tacks[J].Journal of nondestructive evaluation, 2020,39(3): 39-57.
- [2]XU S, JIA X, CHOI Y, et al.Three-dimensional scanning technique in the congenital microtia reconstruction with tissue expander[J].Chinese medical journal,2021,134:842-844.
- [3]王春阳,李国瑞,刘雪莲,等.基于IVCCS的三维点云配准算法[J].红外与激光工程,2022,51(6):469-480.
- [4]宋丽梅,张继鹏,李云鹏,等.基于多视角红外传感器的三维重建方法[J].液晶与显示,2023,38(6):759-769.
- [5]周静,胡怡宇,胡成玉,等.基于点云补全和多分辨Transformer的弱感知目标检测方法[J].计算机应用, 2023, 43(7): 2155-2165.
- [6]陈建文,赵丽丽,任蓝草,等.深度学习点云质量增强方法综述[J].中国图象图形学报,2023,28(11):3295-3319.
- [7]廖中平,白慧鹏,陈立.基于双边滤波改进的点云平滑算法[J].计算机技术与发展,2019,29(11):42-46.
- [8]苏本跃,马金宇,彭玉升,等.基于K-means聚类的RGBD点云去噪和精简算法[J].系统仿真学报,2016,28(10):2329-2341.
- [9]FLEISHMAN S, DRORI I, COHEN-OR D.Bilateral mesh denoising[J].ACM transactions on graphics, 2003, 22(3): 950-953.
- [10]DUAN Y, YANG C, CHEN H, et al.Low-complexity point cloud denoising for LiDAR by PCA-based dimension reduction[J]. Optices communication, 2021,482(1):126567.
- [11]李仁忠,杨曼,冉媛,等.基于方法库的点云去噪与精简算法[J].激光与光电子学进展,2018,55(1):251-257.
- [12]张宏伟.基于双尺度算法的激光雷达点云去噪[J].激光与红外,2023,53(2):169-175.
- [13]焦亚男,马杰,钟斌斌.一种基于尺度变化的点云并行去噪方法[J].武汉大学学报(工学版),2021,54(3):277-282.
- [14]任彬,崔健源,李刚,等.基于自适应阈值的三维点云分段式去噪方法[J].光子学报,2022,51(2):319-332.
- [15]孙科学,渠吉庆.基于线性优化模糊C均值算法和人工神经网络的光照传感器布局方法[J].电子与信息学报, 2023, 45(5):1766-1773.
- [16]崔鑫,闫秀天,李世鹏.保持特征的散乱点云数据去噪[J].光学精密工程,2017,25(12):3169-3178.
- [17]曲金博,王岩,赵琪.DBSCAN聚类和改进的双边滤波算法在点云去噪中的应用[J].测绘通报,2019,512(11):89-92.
- [18]陈亚超,樊彦国,禹定峰,等.考虑法向离群的自适应双边滤波点云平滑及IMLS评价方法[J].图学学报,2023, 44(1): 131-138.
- [19]刘永生,蔡世阳,陈一馨,等.混合滤波与改进双边滤波的点云去噪算法[J].东北大学学报(自然科学版),2023, 44(5): 682-688.
- [20]赵涵,刘永生,赵德鹏,等.基于混合滤波的点云去噪算法研究[J].应用激光,2023,43(9):147-155.

【作者简介】

宋凯瑞(1999—),女,陕西榆林人,硕士研究生,研究方向:图形图像处理、可视化、大数据。

汤慧(1982—),女,山东聊城人,硕士生导师,研究方向:计算机图形图像处理、虚拟仿真、可视化。

(收稿日期:2024-05-31)