Filtering Processing of Lidar Point Cloud Data

Yaxiong Jin1,a, Xitun Yuan1,b, Zhe Wang1,c, Boqiang Zhai1,d

1Department of Surveying and Mapping Engineering, Xi'an University of Science and Technology, Xi'an City, Shaanxi Province, 710000, China
aemail: 1848719171@qq.com, bemail: 110523602@qq.com, cemail: 729659365@qq.com, dmail: 904256266@qq.com
*Corresponding author· s e-mail: cehuiyjs@126.com

Abstract. During the acquisition of point cloud data by lidar, due to system and environment reasons, there will be a large number of discrete points and noise points in the obtained point cloud data, which will affect the extraction of target ground object information. In view of the noise introduced by the system itself and the environment where the target object is when the lidar is scanning, a de-noising method combining statistical filtering and bilateral filtering is proposed in this paper.

1. Introduction
This article uses the statistical filtering and combination of bilateral filtering denoising method to deal with point cloud, including statistical filtering is used to remove the laser point cloud is a lot of noise from the group, under the premise of bilateral filtering in keeping the characteristics of point cloud to smooth denoising laser point cloud [1], and for the large number of point cloud voxel grid filter, while maintaining the characteristics of point cloud of point cloud amount to streamline, speed up the point cloud filtering and improve the efficiency of the laser point cloud data processing.

2. Point cloud under-sampling
For more number of lidar point cloud data, requires the sampling point cloud to streamline, through the use of voxel grid method to streamline of point cloud [2], under the point cloud sampling before as shown in fig.1, downsizing as shown in fig.2 after here, voxel three-dimensional grid for the side length is equal to 1 cm cubes, the number of sampling points before 460400, the number of points after filtering for 67194, the number of point cloud to reduce greatly, at the same time, the sampling point cloud contour shape before and after the basic remains the same, for the subsequent operations to save a large amount of computing resources, as shown in table 1 it can be seen through the study of the streamline of point cloud, The time of point cloud filtering is greatly reduced. The simplification of laser point cloud plays an important role in realizing the real-time performance of point cloud filtering.

| Data processing          | Grid side length (cm) | Table point cloud filtering time (s) |
|--------------------------|-----------------------|-------------------------------------|
| Point cloud before       |                       | 27.32                               |
| streamlining              | 0.2                   | 16.81                               |
| Point cloud after         |                       | 7.54                                |
| streamlining              | 0.5                   | 2.98                                |
|                          | 1                     |                                     |
3. Point cloud statistical filtering algorithm

Statistical filtering algorithm is a very common filtering algorithm. By calculating the distance distribution between query points and adjacent points in the input data, the average distance between each point and all adjacent points is calculated. Then the points whose average distance is outside the standard range can be defined as outliers and removed from the data.

Suppose a point cloud data set with discrete points $X = \{X_i, i = 1, 2, \ldots, n\}$, point cloud data after filtering discrete points $X' = \{X_i, i = 1, 2, \ldots, m\}$, for any point $X_i$, let $S_i$ represent its average distance to $k$ points in the neighboring region. The statistical filtering algorithm assumes that the average distance $S_i$ of all points in $X'$ to their respective neighborhoods obeys a Gaussian distribution, and the Gaussian distribution has mean value and standard deviation [3], which is calculated from the global average distance. If $\mu$ is its mean value and $\sigma$ is its standard deviation, then $S(\mu, \sigma)$ is denoted, and the formula is as follows:

$$
\mu = \frac{\sum_{i=1}^{n} S_i}{n} \quad (1)
$$

$$
\sigma = \frac{\sum_{i=1}^{n} (S_i - \mu)^2}{n} \quad (2)
$$

Thus, the threshold range based on Gaussian distribution $(\mu - \text{std} \cdot \sigma, \mu + \text{std} \cdot \sigma)$ is obtained according to mean $\mu$ and standard deviation $\sigma$, and all points in the point cloud outside the range of $(\mu - \text{std} \cdot \sigma, \mu + \text{std} \cdot \sigma)$ from their neighborhood are regarded as outliers. In this algorithm, $K$ value and standard deviation multiple $\text{std}$ can be set. The former controls the size of the neighborhood, and the latter controls the range of conditions, that is, the severity of screening. Only when the two parameters are set properly can high quality noise reduction results be obtained.

The statistical filtering denoising process is as follows:

1. Input the point cloud requiring denoising into the statistical filter.
2. Set neighborhood search points $K$ for the statistical filter, and calculate the average distance $S_i$ from each query point to $K$ neighborhood points for denoising.
3. Set standard deviation multiple $\text{std}$, calculate the mean value $\mu$ and standard deviation $\sigma$ of the mean distance of all data points, obtain the screening range of discrete points $(\mu - \text{std} \cdot \sigma, \mu + \text{std} \cdot \sigma)$, provide conditions for the screening of discrete points.
4. Screen the discrete points, save the point cloud and noise points after denoising, use the visualization module to visualize and observe the effect of denoising. If the denoising effect is poor, then adjust the parameters and adopt statistical filtering operation again.

Fig.1 Data Diagram of Original Point Cloud        Fig.2 Data diagram after point cloud simplification
The effect diagram of table point cloud after statistical filtering of point cloud is shown in Fig.4 (a) and (b). It can be seen that most separation group noise points circled in Fig.2 have been filtered out.

For table point clouds, within the range of reasonable parameters, 1.8%-3.2% of point clouds are defined as outliers removed; for corridor lidar point clouds, 2.6%-6.3% of outliers in point clouds are removed.
Table 2  Number of point clouds before and after filtering

| Data processing | The number of neighborhood points K | Ratio Std | Number of table point clouds |
|-----------------|------------------------------------|-----------|------------------------------|
| Before filtering | 30                                 | 1         | 65016                        |
|                  | 30                                 | 2         | 65909                        |
| After filtering  | 50                                 | 1         | 65063                        |
|                  | 50                                 | 2         | 65935                        |

4. Bilateral filtering algorithm

Outliers can be well removed by point cloud statistical filtering, but the "noise" mixed with useful information of point cloud cannot be removed by statistical filtering. Therefore, the weighted average method is adopted by using bilateral filtering to denoise the internal point cloud while maintaining the edge\(^4\). The procedure flow chart of the algorithm is shown in Figure 5. First, the K neighborhood of the current point is found by searching the KdTree neighborhood, then the point cloud normal vector is estimated by PCA method, and the bilateral filtering factor is calculated by calculating the \(W_C\) and \(W_S\) kernel functions\(^5\). Finally, the data points move along the point cloud normal vector normal direction, so as to preserve the characteristics and filter the noise. The specific steps are as follows:

1) Build the point cloud Kd-Tree topological relationship, and then calculate the set \(N(P_i)\) of K neighborhood of each data point \(P_i\).

2) The least square method is used to fit the point cloud \(N\) in the K neighborhood of \(P_i\) point into a plane. Then, the normal vector \(N_i\) of \(P_i\) point is estimated, and the covariance matrix \(C\) is constructed. The eigenvector \(V\) corresponding to the minimum eigenroot \(\lambda\) is used as the approximate estimation of the normal vector of \(P_i\) point.

3) each near point of \(P_i\) calculation \(W_c\) and \(W_s\) parameter \(| | P_j - P_i | | \) and \(| | < n_j, n_i > | | 1.\)

4) Calculate the Gaussian kernel functions \(W_C(x)\) and \(W_S(y)\).

5) Calculate the bilateral filtering factor \(\lambda\).

6) Calculate the filtered data points until all the points are updated.

Fig. 5  Flow chart of bilateral filtering

Fig. 6 below shows the point cloud data of a person: (a) is the original point cloud image of the filtered person, (b) is the statistical filtering effect picture, and (c) is the effect picture after the statistical bilateral filtering.
By comparing the Fig.6 (a) - (c) original point cloud image characters, statistical filtering point cloud and statistics after bilateral filtering after the point cloud, clear, statistical filtering noise in point cloud outliers in better removal, statistics bilateral filtering is not only to remove outliers, and in keeping the edge for the point cloud internal smoothing denoising, achieve better denoising effect.

5. Conclusion
The voxel raster method is used to subsample the data with a large number of point clouds, which speeds up the operation efficiency of the algorithm and reduces the running time. For lidar point cloud data is proposed and bilateral filtering method based on statistical filtering to remove noise of point cloud, statistical filtering can effectively remove outliers laser point cloud data, can keep the laser point cloud point cloud bilateral filtering characteristics under the condition of smoothing denoising, and proved through the experiment of laser point cloud data has better denoising effect, and provides a good data for subsequent processing of the laser point cloud.

References
[1] Yang,X.(2020) Filtering and segmentation of lidar point cloud data.D. Harbin Institute of Technology.
[2] Li,R.Z.,Yang,M.,Rang,Y.(2018) Point cloud denoising and simplification algorithm based on method library.J. Laser and Optoelectronics Progress.,(1):251-257.
[3] Xu, Z.Y., Yan, Y.J. (2018) Research on filtering algorithm based on lidar point cloud. J. Information and communication., 000(002):80-82.

[4] Du, X.Y., Jiang, X.F., Hao, C.G., et al. (2010) Bilateral filtering denoising algorithm for point cloud model. J. Computer applications and software.,(07):245-246.

[5] Yu, B., Guo, L., Qian, X.L., et al. (2012) Bilateral filtering algorithm of controllable kernel function. J. Journal of Jilin University (Engineering and Technology Edition).,(06):1553-1557.