Indexed in
Scopus Compendex and Geobase Elsevier,
Geo-Ref Information Services-USA, List B of Scientific
Journals, Poland, Directory of Research Journals

International Journal of Earth Sciences and Engineering

www.cafetinnova.org

ISSN 0974-5904, Volume 09, No. 05

October 2016, P.P.2273-2277

Research and Implementation for Scattered Point Cloud Data Denoising Method

GUIGANG SHI[1], XINGHUA DANG[2] AND FENGYU YE[1]
[1]School of Civil Engineering, Anhui Jianzhu University, Hefei, CHINA
[2]Technical Service Department of Library, Anhui Jianzhu University, Hefei, CHINA
Email: snowpeace@126.com

Abstract: In the science and technology of Surveying and Mapping, 3D laser scanning technology is an important means of Surveying and mapping and geographic information data acquisition. But there is noise and redundant data in the original point cloud data collecting by 3D laser scanning technology. The paper analyzed systematically point cloud data of the different types, the noise causes and consequence, and according to the categories summarized various denoising methods. On the basis of studying the law of scattered point cloud data, the denoising algorithms were analyzed to Laplacian, Mean Curvature Flow filtering, Mean Shift, Bilateral Filtering and so on. And their denoising ideas were illuminated, the program has implemented algorithm by Matlab software compilation, the feasibility of these denoising algorithms were verified by the cases. The research results provide a reference for point cloud data optimization, and create a good foundation for the point cloud data processing and 3D model construction, which is better serve the construction of digital city, smart city and other fields.

Keywords: *Scattered Point Cloud Data, Mean Shift Denoising, Bilateral Filtering Denoising, Case verification.*

1. Introduction

Three dimensional laser scanning technology is a breakthrough in surveying and Mapping Science and technology. It can automatically, quickly measure object space point information. These massive amounts of information provide a new technical means for the rapid establishment of the object three dimensional image models. The technology plays an important role in digital city, intelligent city, reverse engineering, virtual simulation and other fields.

However, with the further application of 3D laser scanning technology, the drawbacks of the technology are becoming more and more serious. The point cloud data is a kind of spatial data collection which acquired by the 3D laser scanning technology. The data points are dense redundant, discrete and scattered. At the same time, the point cloud is a collection of massive data, up to hundreds of thousands or even millions of data points, which needs a huge amount of storage. And because the influence comes from the laser scanning equipment, measurement surface reflection, sensor, artificial disturbance, measurement environment, processing error and so on, point cloud data contains a large number of useless data called the noise point.

The noise has a serious impact on the smoothness of curved surface, and even point cloud data cannot achieve model reconstruction because of its influence. If do not handle it directly and effectively, it will reduce the efficiency of the geometric model reconstruction, occupy a large amount of calculation and processing time, consume a large amount of time and computer resources at storage, display and construction of surface model and so on. These factors desperately need to remove the noise of point cloud data, in order to facilitate point cloud data processing and deep processing. The result will provide a good data source for the surface and 3D model construction, which is the target of this article to research denoising method of point cloud. This paper mainly researches the algorithm and principle of various denoising for scattered point cloud, which is realized and verified by the program and the case. So that the engineers and researchers can deepen and apply them according to their own needs.

2. Point cloud data classification and noise processing method

Because the structure of Terrestrial 3D Laser Scanner (TLS) and the principle of collecting point cloud data are different, the processing method of noise point is also different. At present, the point cloud data array form mainly has the following categories [1]:

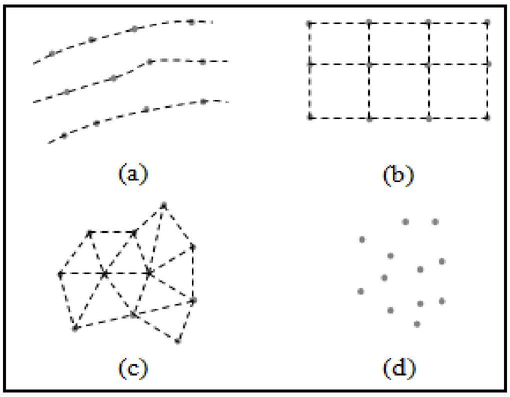

Figure 1: *Different representation of point cloud data*

2.1. Scanline-based point cloud data

It is scattered point cloud data in a particular direction, as shown in Figure 1 (a). The noise of this kind of point cloud can be removed by Minimum Distance Method, Uniform Sampling Method, String Value Method, Angle Deviation Method, Angle Chord Method and Chord Method, and so on.

2.2. Array point cloud data

It is an ordered point cloud data in a certain order arranged, as shown in Figure 1 (b). The noise of this kind of point cloud can be removed by Rate Reduction Method, Equal Space Reduction Method, Chord Method, Equivalent Reduction Method, etc.

2.3. Grid point cloud data

It is also an ordered point cloud data by triangular network interconnection as shown in Figure 1 (c). Minimum Bounding Area Method, Equal Distribution Density Method and others are applied to weaken denoising of this point cloud data.

2.4. Scattered point cloud data

Data distribution has no chapter to follow, completely scattered, as shown in Figure 1 (d). They can weaken denoising by the methods of Clustering, Iteration, Particle Simulation, Uniform Grid Method, Bounding Box Method, Random Sampling and Curvature Sampling, etc.

The front three kinds belong to the ordered or partial ordered point cloud data, which have the topological relation between point and point. This denoising method is relatively simple, and effective methods are also many. It is difficult to process the noise of the last kind of scattered point cloud data. Its points and points is out of order. The ordered point cloud data denoising method cannot be applied directly here, otherwise the effect is also very poor. At present, for scattered point cloud data, there is not a fast, concise denoising method.

Therefore, This paper researches the algorithm and principle of various denoising for scattered point cloud, especially, focus on the research of Laplacian Algorithm, Mean Curvature Flow, Mean Shift and Bilateral Filter, which is realized by the program, achieved good results. Thus, it provides a good reference for engineers and researchers. They can deepen and apply them based on their needs.

3. Scattered point cloud denoising methods

3.1. Laplacian algorithm

It is one of the most common and the simplest smoothing methods for Laplacian Operator to remove noise. Its basic principle is to apply the Laplacian Operator to each vertex of the model. Laplacian Operator is:

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \qquad (1)$$

Figure 2 described the process of point cloud by Laplacian Operator, and the geometric meaning is very intuitive.

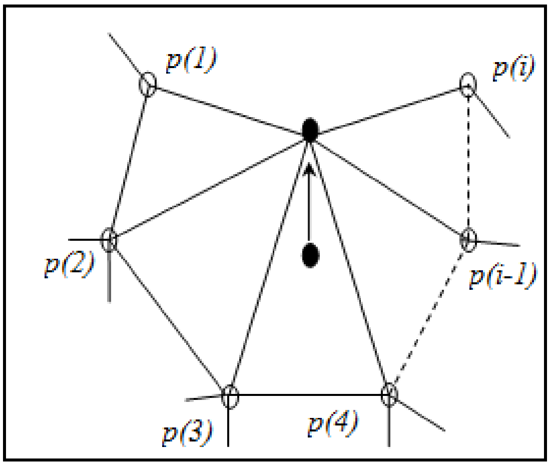

Figure 2: *Laplacian operator denoising smoothing process*

In the model of point cloud, set up $p_i = (x_i, y_i, z_i)$, Laplacian Operator can be dispersed by the neighborhood points of pi , then:

$$\delta_i = L(p_i) = p_i - \frac{1}{d_i} \sum_{j \in N(i)} p_j \qquad (2)$$

Vertex is moved according to the result of Formula 2. The denoising of the point cloud data can be regarded as a diffusion process on, then:

$$\frac{\partial p_i}{\partial t} = \lambda L(p_i) \qquad (3)$$

Laplacian denoising method achieved the smoothing purpose by spreading the high frequency geometric noise. Although the algorithm is simple, with the increase in the number of iterations, it is easy to make the convex and concave feature of model fuzzy and vertex drift.

3.2. Mean Curvature Flow filtering

Mean Curvature Flow Filtering method, also known as Desbrun Filtering [2]-[3], is to use the curvature flow to guide the smoothing of mesh model, and moves the vertex along the direction of the normal vector with the rate of the average curvature.

$$\frac{\partial L}{\partial t} = -G(L)n(L) \qquad (4)$$

In Formula 4, $G(L)$ is the average curvature, n(L) is outside normal direction of point. Solving the above equations, we can get:

$$L^{t+1} = L^t + \lambda G(L)n(L) \qquad (5)$$

Making the vertex V_i, the vertex can be updated as follows:

$$V_i^{t+1} = V_i^t + \lambda G(V_i^t) n(V_i^t) \qquad (6)$$

Figure 3.7 describes the process of smoothing based on the mean curvature flow, which can get a better smoothing effect, but it makes the sampling rate of grid points worse.

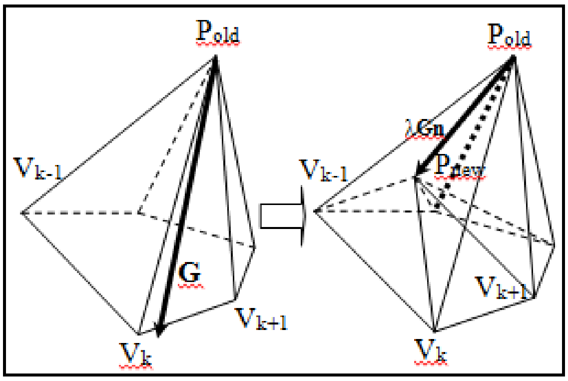

Figure 3: *Desbrun Filtering principle of smooth operation*

3.3. Bilateral Filtering algorithm

Jones et al. [4] put forward a non-iterative Bilateral Filtering algorithm under the condition of keeping the grid model feature which can be realized to control the size of the vertex neighborhood. This may be applied on point cloud data and non-manifold surfaces. Though, the algorithm drawbacks lie in searching adjacent regions of the vertex, when the neighborhood is too large, the processing time will need a large amount of computation. When the neighborhood is too small, it is not effective to deal with some of the slightly larger noise, on the contrary, even to enhance the noise. Fleishman et al. [5] proposed a similar method based on the idea of Image Bilateral Filter, which is different with Jones. It moves the position of the sampling points along the normal direction of the vertex, which makes the smooth order of the surface rapid increase, and achieves the effect of smoothing. However, the algorithm can cause excessive smoothing to lead to abnormally maintain the fine characteristics of the model when it processes slightly larger noise. Although these methods applied on the mesh model surface in papers, they are also suitable for point model surfaces.

3.3.1. Bilateral Filtering principle of point cloud data

Image Bilateral Filtering substitutes the weighted average value of the surroundings point gray for the current point. The weight factor is not only related to the geometric distance between the current point and the surrounding points, but also the difference of the gray value. The Bilateral Filtering method is used to 3D point cloud data filter, which is mainly to deal with the noise in the point cloud data, and there are many similarities with the two dimensional image smoothing. In 3D point cloud data filtering, It can define:

$$p' = p + \lambda n \qquad (7)$$

In the above formula, p' is the new point after the data point p filtering, and the λ is the bilateral filter weight factor, and the n is the normal vector of the data point p. The λ is defined as follows:

$$\lambda = \frac{\sum_{k_{ij} \in N(p_i)} H_C(\| p_i - k_{ij} \|) H_S(\langle p_i - k_{ij} \rangle)(n_i, q_i - k_{ij})}{\sum_{k_{ij} \in N(p_i)} H_C(\| p_i - k_{ij} \|) H_S(\langle p_i - k_{ij} \rangle)} \qquad (8)$$

In the Formula 8, $N(p_i)$ is the neighbor points of the data point p_i, smoothing filter is the standard Gauss Filter, which can be expressed as:

$$H_C(x) = e^{-x^2/2\sigma_C^2} \qquad (9)$$

The weight function keeping the characteristics is similar to the smoothing filter, can be defined as:

$$H_S(y) = e^{-y^2/2\sigma_S^2} \qquad (10)$$

In the Formula 10, the parameter σ_c is impact factor on the distance between the data point p_i and the neighbor points; the parameter σ_s is the impact factor about distance vector projection on the data point normal n_i to data point p_i , the distance is between the data point p_i and the neighbor points. They are represented respectively the data point Gauss filter constant coefficients on tangent plane direction and normal direction, which reflects the impact of the bilateral filtering operation of the tangential and normal direction on any data point p.

The Bilateral Filtering method is used to filter the noise of the point cloud data, the method is simple and effective, and the operation speed is fast. It can remove the noise while maintaining the characteristic. But it cannot deal with a large range of noise, especially when the iteration number is more it is easy to produce the problems such as excessive smoothing, distortion and so on.

3.3.2. Design and software implementation of Bilateral Filtering algorithm for point cloud data

Software Matlab programmed point cloud data bilateral filtering algorithm, the main process is as follows:

(1) The point cloud data is initialized, the filter is set to 5, the standard deviation is *sigma* = [3 0.1].
(2) Calculated H_s = *exp*(-(*X*.^2+*Y*.^2)/(2**sigma*^2)).
(3) Calculated λ = *exp*(-(*dL*.^2+*da*.^2+*db*.^2)/(2**sigma*^2)).
(4) Calculated normal vector *n*, the point cloud obtained after filtering is obtained by the Formula 7.

Figure 4 is the original point cloud data, and Figure 5 is point cloud data after bilateral filtering.

3.4. Mean Shift Method

3.4.1. Mean Shift Algorithm Principle

Mean Shift is a density gradient estimation of certain space position, around which the point distribution can be counted in small sub area [6]. Mean Shift moves continuously the space points along the gradient direction until the gradient is zero. Given the d dimensional Euclidean space R^d, for the point data set $P=\{x_i, i=1,2,\cdots,n\}$, with the kernel function $K(x)$ and the nuclear window range h the multivariate kernel density estimation function is:

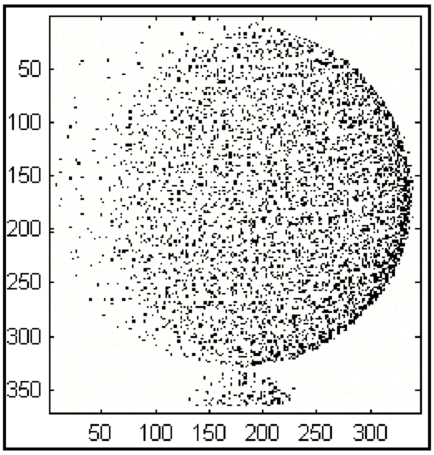

Figure 4: *The original point cloud data*

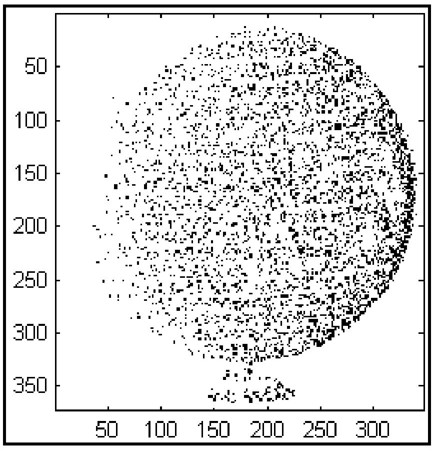

Figure 5: *Point cloud data after bilateral filtering*

$$f(x)=\frac{1}{nh^d}\sum_{i=1}^{n}K\left(\frac{x-x_i}{h}\right) \quad (11)$$

Where h is called bandwidth, it shows x neighborhood range to estimate the density of the x point. $K(x)$ is called the density kernel function:

$$K(x)=c_{k,d}k(\| x\|^2) \quad (12)$$

In the formula $c_{k,d}$ is the normalized constant, and the $K(x)$ integral is 1. Differentiated the Formula 11, the gradient can be gotten at x, and the mean shift iteration vectors are also obtained:

$$M_s(x)=\frac{\sum_{i=1}^{n}\frac{x_i}{h^{d+2}}g\left(\left\|\frac{x-xi}{h}\right\|^2\right)}{\sum_{i=1}^{n}\frac{1}{h^{d+2}}g\left(\left\|\frac{x-xi}{h}\right\|^2\right)}-x, \quad g(x)=-K'(x) \quad (13)$$

The above formula shows along the $M_S(x)$ direction it is the fastest to move x to the maximum density place in the range of h bandwidth. This iterative process will eventually converge, that is, the $M_S(x)$ is finally 0 which is called as Mean Shift, and the process of repeated movement is called the mean shift algorithm.

3.4.2. Point cloud data denoising by Mean Shift method

Suppose the point model is $P\in R^3$ in the point cloud data, which consists of two parts: the spatial position information $v_i=(x_i,y_i,z_i)$ and the normal vector n_i, that is, $p_i=(v_i,n_i), i=1,2,\cdots,n$, n is the number of point set$\{p\}$. Ordering k nearest neighbor points of the general point p_i as $N(p_i)=\{q_{i,1},q_{i,2},\cdots,q_{i,k}\}$, the Mean Shift vector:

$$M_s(p_i)=\frac{\sum_{j=1}^{k}g(\|n_i-q_{ij}^h\|)(q_{ij}-M(p_i))}{\sum_{j=1}^{k}g(\|n_i-q_{ij}^h\|)} \quad (14)$$

Where $g(x)$ is generally Gaussian kernel function, n_i is the general point of the normal, q_{ij}^h is the neighbor point feature information, $M_S(p_i)$ is called as the mean shift point of p_i. In addition, p_i is $M(p_i)$ initial value, $M_S(p_i)$ is the Mean Shift vector of $M(p_i)$. It can be known that the mean shift process is the gradual moving process from the vertex to the sample mean point, there is the following formula:

$$M(p_i)=M(p_i)+M_s(p_i) \quad (15)$$

From the above iterative process, each point will converge to a stable point, called as the mode point. That is, mean shift process is repeat iterative, the convergence is end until each point can find the mode point. In the realization, the convergence condition is that $M_s(p_i)$ is less than a certain positive value ε. Through Mean Shift iterative, each sampling point would move to the maximum probability place on the surface of point model, which is also the denoising process of point model.

3.4.3. Design and software implementation of Mean Shift algorithm for point cloud data

According to the principle of Mean Shift algorithm, it was achieved to the point cloud data smoothing by software Matlab, the main process is as follows:

(1) Inputed initialization point cloud data, set hr<=0, *hs* <=0, andε=2**hs*.
(2) Calculated $G=exp(-G/hs/hs)$.
(3) Iterated calculation Formula 13.
(4) Calculated the Formula 14, the point cloud data can be smoothed.

Figure 6 is the original point cloud data, and Figure 7 is point cloud data after smoothing by Mean Shift method.

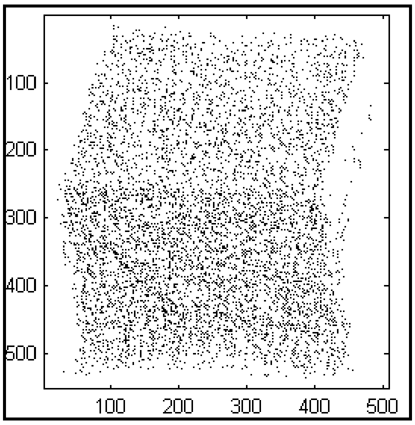

Figure 6: *The original point cloud data*

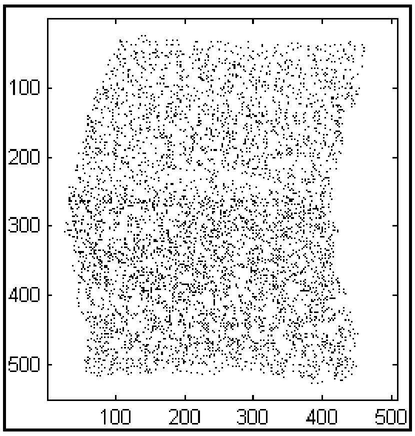

Figure 7: *Point cloud data after Mean Shift smoothing*

4. Conclusion

Point cloud data is one of the basic data sources of Digital City, Smart City, Reverse Engineering and so on. Point cloud data optimization and compression is the key technology of point cloud data processing. The noise and redundant data of point cloud must be deleted which facilitates the subsequent three-dimensional model reconstruction, thus, provide good model data for the construction of Digital City, Smart City, Reverse Engineering and so on. This paper focuses on the algorithm analysis of Laplacian Algorithm, Mean Curvature Flow, Mean Shift and Bilateral Filter, and illustrates the core idea of the algorithm which is verified by cases, and the result and denoising function are better. This further illustrates the feasibility of above all kinds of denoising methods in the point cloud data, and provides reference for point cloud denoising workers. Next, after point cloud denoising this paper will further research on the point cloud compression and modeling algorithm.

5. Acknowledgements

This work was supported by Science and Technology Project of Anhui Province (1501041132) and 12th Five-Year National Science and Technology Support Program (2012BAJ08B01, 2012BAJ08B03).

References

[1] Dai JL. Research on preprocessing algorithm of massive point cloud, Master Dissertation of Zhejiang University, 2006, 34-38 (in Chinese), 2006.

[2] Mathieu Desbrun, Mark Meyer, Peter Schro der, Alan Bats. Implicit fairing of irregular meshes using diffusion and curvature flow, Proceeding of SIGGRAPH, 1999,45-52, L.A., CA.

[3] Vollmer J, Mencl R, and Muller H. Improved Laplacian smoothing of noisy surface meshes, Proceedings of EUROGRAPHICS'99, 1999:131-138.

[4] Jones T. Feature preserving smoothing of 3D surface scans. Department of Electrical Engineering and Computer Science, 33, 89-95, MIT, 2003.

[5] Fleishman S, Drori I, and Cohen-Or D. Bilateral mesh denoising, Proceedings of ACM SIGGRAPH' 03, 2003, 950-953.

[6] Bogdan Georgescu, Ilan Shimshoni, Peter Meer. Mean Shift based clustering in high dimensions a texture classification example, Ninth IEEE International Conference on Computer Vision, 1, 13-16, 2003.