

Density-Based Denoising of Point Cloud

Faisal Zaman, Ya Ping Wong and Boon Yian Ng

Abstract Point cloud source data for surface reconstruction is usually contaminated with noise and outliers. To overcome this deficiency, a density-based point cloud denoising method is presented to remove outliers and noisy points. First, particle-swarm optimization technique is employed for automatically approximating optimal bandwidth of multivariate kernel density estimation to ensure the robust performance of density estimation. Then, mean-shift based clustering technique is used to remove outliers through a thresholding scheme. After removing outliers from the point cloud, bilateral mesh filtering is applied to smooth the remaining points. The experimental results show that this approach, comparably, is robust and efficient.

Keywords Point cloud · Denoising · Optimal bandwidth · Particle swarm optimization · Bilateral filter

1 Introduction

In the past few decades, surface reconstruction has gained significant attention due to the availability of commodity structured light sensors such as Microsoft Kinect. Point cloud data acquired from these devices is usually contaminated with noise and outliers [1], both of which could be caused by the lighting or reflective nature of the surface or artifact in the scene. In recent years, a number of point cloud denoising methods have been proposed [1, 2]. Most of the methods work only on certain type of noise or outliers. Denoising algorithms such as data clustering [14, 15] are robust to outliers but require prior knowledge about the input objects. Majority voting method

F. Zaman (✉) · Y.P. Wong (✉) · B.Y. Ng (✉)
Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia
e-mail: talk.console@gmail.com

Y.P. Wong
e-mail: ypwong@mmu.edu.my

B.Y. Ng
e-mail: ngyian@mmu.edu.my

[16] can detect several types of outliers cluster, but high computation time makes it infeasible for large datasets. Smoothing algorithms such as moving least square [7] and mean curvature flow [5] are also used to remove noise and outliers. These methods treat outliers as point with large noise and project the noisy points to an estimated surface. However, these methods are sensitive against large number of outliers and oversmooth the data points.

In this paper, a point cloud denoising method to remove noise and outliers from the data is proposed. This method consists of the following steps. First, the density of the input data point using kernel density estimation (KDE) technique is evaluated. The performance of KDE is highly influenced by the choice of smoothing parameter, also known as bandwidth. Most of the bandwidth selection methods such as [6, 9, 11, 17] involved brute-force or exhaustive search strategies have expensive computation. A particle swarm optimization (PSO) (proposed by Kennedy and Eberhart 1995) aided bandwidth selection criterion using leave-one-out cross-validation (LOOCV) for multivariate kernel density estimation is proposed. Because of the simple implementation and quick convergence rates of PSO, it has been adapted in many practical applications [3]. The optimal bandwidth selection method ensures the robust performance of density estimation. After estimating the density, the mean-shift clustering algorithm is applied to detect the local maxima of the constructed density estimation. One of the main advantages of using mean-shift algorithm is its ability to localize cluster modes automatically without any prior knowledge of the number of clusters. Once clustering is done, outliers can be removed by the proposed thresholding scheme. After removing the outliers, bilateral mesh filtering is applied to smooth out the rest of the points. Experiments on various point cloud datasets show that the method used in this research gives better result than that of other approaches.

2 Our Method

In this section, an overview of the optimal bandwidth selection framework is provided, starting from a review of the classical bandwidth selection methods.

2.1 Bandwidth in Kernel Density Estimation

Kernel density estimation (KDE), also known as Parzen window estimation, is a non-parametric way of estimating the underlying distribution or probability density function for a data set. Given n data points X_i , $i = 1, \dots, n$, the kernel density estimate obtained with multivariate kernel function $K_H(x)$ and bandwidth or smoothing parameter H , which is a symmetric positive definite matrix, computed at the point x is defined as:

$$\hat{f}_H(x) = \frac{1}{n} \sum_{i=1}^n K_H(x - X_i) \quad (1)$$

where $x = (x_1, x_2, x_3)^T$ is the 3-dimension vector, $K_H(x) = |H|^{-1/2}K(H^{-1/2}x)$. The resulting density through KDE depends more heavily on the choice of bandwidth than that of the kernel. For the same data, different bandwidths can produce different results. Optimal bandwidth matrix—which minimizes the error between the estimated density $\hat{f}_H(x)$ and the true density $f(x)$ is estimated. The performance of $\hat{f}_H(x)$ can be determined by the risk function or integrated mean squared error (IMSE) such as,

$$R(\hat{f}_H(x), f(x)) = E(L(\hat{f}_H(x), f(x))) \tag{2}$$

where, L is the loss function and $L(\hat{f}_H(x), f(x)) = \int [\hat{f}_H(x) - f(x)]^2 dx$. The optimal bandwidth can be obtained by minimizing the risk function, but since the true density $f(x)$ is unknown leave-one-out cross-validation is used to estimate the risk function. Therefore, loss function can be expressed as a function of H .

$$L(H) = \frac{1}{n} \sum_{i=1}^n \log \hat{f}_{H,i}(x_i) \tag{3}$$

where, $\hat{f}_{H,i}(x_i)$ is the leave-one-out estimator, $\hat{f}_{-i}(x_i) = \frac{1}{(n-1)} \sum_{j \neq i}^n K_H(x_i - x_j)$ denotes the density estimation of x_i by using the other $(n - 1)$ observed data points. Then the optimal bandwidth matrix H^* can be estimated by,

$$H^* = \arg \max_H \frac{1}{n} L(H) \tag{4}$$

Several classes of parameterizations of the bandwidth matrix are available in the literature. Some popular choices include full bandwidth matrix, a diagonal matrix with positive elements, or using a single bandwidth. In the application, the diagonal bandwidth matrix, $H = \text{diag}(h_1^2, h_2^2, h_3^2)$, is employed since it is computationally more efficient than the full bandwidth matrix.

2.2 PSO for Optimal Bandwidth Selection

The PSO technique is proposed to carry out the bandwidth selection for better approximation of kernel density estimation. PSO is a nature-inspired meta-heuristic non-linear stochastic optimization technique. PSO is popular for its ability to optimize complex non-linear functions and for its simple implementation. PSO consists of a swarm of particles that fly through hyperspace of the of objective function’s landscape. The pipeline of the optimal bandwidth selection method shown in Fig. 1. The optimization problem is solved from Eq. 4 using PSO, where $H = [h_1, h_2, h_3]^T$ is a 3-dimensional vector to be optimized and $L(\bullet)$ is the cost function. Here are the steps of PSO algorithm to find the optimal bandwidth; where H_i^t is the potential solution of the i th particle at iteration t .

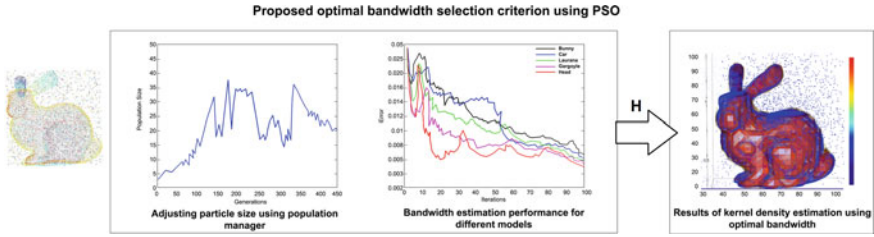


Fig. 1 The pipeline of the optimal bandwidth selection method, using PSO

1. **Swarm Initialization:** Set the iteration index $t = 0$. S is the total number of particles, where $\{H_i^t\}_{i=1}^S$ randomly generated particles in solution space.
2. **Swarm Evaluation:** Each particle denoted as p_i^t remembers its best position visited so far, which provides the cognitive information. Every particle denoted as p_g^t also knows the best position visited so far among the entire swarm, which provides the social information. The cognitive information p_i^t and the social information p_g^t are updated at each iteration as follows:

```

if  $L(H_i^t) < L(p_i^{t-1})$  then
     $p_i^t \leftarrow H_i^t$ 
end if
if  $L(p_i^t) < L(p_g^{t-1})$  then
     $p_g^t \leftarrow p_i^t$ 
end if
    
```

3. **Swarm Update:** Each particle H_i^t has a velocity, denoted as v_i^t . The velocity and position of the i th particle are updated in each iteration according to:

$$v_i^t = \omega v_i^{t-1} + C_c U[0, 1](H_i^{t-1} - p_i^{t-1}) + C_s U[0, 1](H_i^{t-1} - p_g^{t-1}) \quad (5)$$

$$H_i^t = H_i^{t-1} + v_i^t \quad (6)$$

Here $U[0, 1]$ uniform random number between 0 and 1 and inertia weight (proposed by Shi and Eberhart [13]) sets to $\omega = U[0, 1]$. C_c and C_s are constants and called cognitive learning rate and social learning rate, respectively. It controls a particles learning behavior. Large value of C_c causes particles scatter around and slow convergence. On the other hand, large value of C_s causes fast convergence, and sometimes it can cause the swarm to converge to local optima. These parameters are chosen carefully so that global optima can be attained without compromising convergence speed. In this application, the parameters C_c reduced from 2.5 to 0.5 and C_s varied between 0.5 and 2.5 according to $C_c = 2.5 - 2t/T$ and $C_s = 0.5 + 2t/T$ [12].

The global best position p_g of all particles during the previous three steps is defined as $p_g = \arg \max_{p_i} [L(H_i)]$, where $p_i = H_i$. Once the maximum number of iterations reached algorithm stopped and return the solution p_g .

Searching the solution depends highly on the number of selected particles. The number of particles increases the chances of finding global optimal solution, but it also increases the computation expenses significantly. To identify the number of particles suitable for solving the current problem, a method also known as population manager [10] is adopted. The idea behind the population manager is to adjust the total number of particles in PSO according to the solution-search status. The population manager determines the number of suitable particles in the following ways: Assuming k is population-manager activating threshold. For k consecutive generations, if there is no update in p_g and current population size does not exceed the maximum population size, then a new particle will be added to the swarm. This new particle will be added by combining the information from randomly selected two particles in previous best solutions. If population size is equal to the maximum population and there is no update in p_g , then remove a particle with poor performance to make space for new potential particle. On the other hand, in k consecutive generations, if p_g update multiple times, then remove a particle with poor performance. Figure 2a demonstrates how population manager dynamically adjust the particle in order to facilitate the searching process.

The initial population size in this application is set as one, and the maximum population size is set as 45. To evaluate the PSO based bandwidth selection method, a method to different types of point cloud models is applied (Fig. 2b). As the number of iterations increases, the mean square error (MSE) gradually converges to zero and becomes steady once the optimal bandwidth is found.

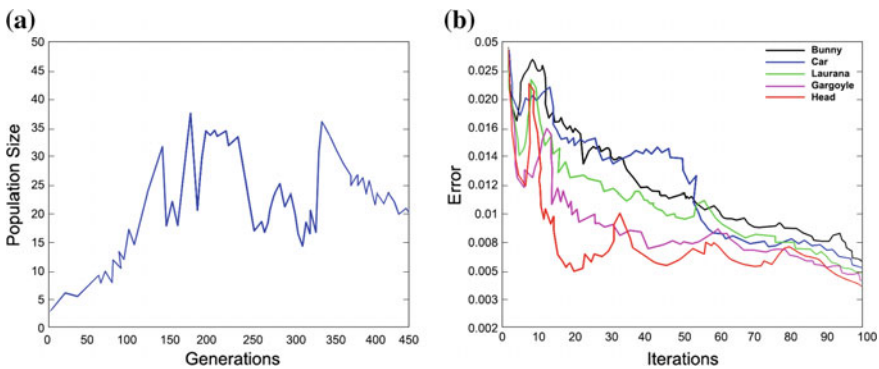


Fig. 2 **a** Change in the number of particles using Population manager. **b** Estimation performance of our optimal bandwidth selection method

3 Noise and Outliers Removal

3.1 Algorithm for Outlier Removal

Mean-shift-based clustering algorithm [4] is a non-parametric unsupervised clustering technique that does not require any prior information about the number of clusters. It is applied to the kernel density estimate to identify each local maxima or mode which represents one cluster. The mode of the density function are located at the zeros of the gradient function $\nabla f(x) = 0$. For each data point, say x , the mean-shift procedure generates a sequence of points $\{y_j\}_{j=1,2,\dots}$ with

$$y_{j+1} = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{y_j - x_i}{h_i}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{y_j - x_i}{h_i}\right\|^2\right)} \quad j = 1, 2, \dots \quad (7)$$

and $y_1 = x$. This sequence converges to a mode of density. The point x belongs to the cluster corresponding to this mode.

After the clustering is done, the outliers need to be determined. For each point p in a cluster, where x_p is the original location of p , the average distance (\bar{x}) of all its neighboring points using k-nearest neighbor (kNN) is calculated. Then the point p is shifted iteratively to a new position \bar{x} . After that, the distance between two positions, $|\bar{x} - x_p|$, is calculated. If the difference between $|\bar{x} - x_p|$ and the average distance between its shifted neighbors below a certain threshold, then point p is considered outlier (Fig. 3b). The outliers are discarded, leaving behind a new set of points, which have better representation of the true shape of the cluster. Figure 4 demonstrates the results of the outliers removal algorithm in different models.

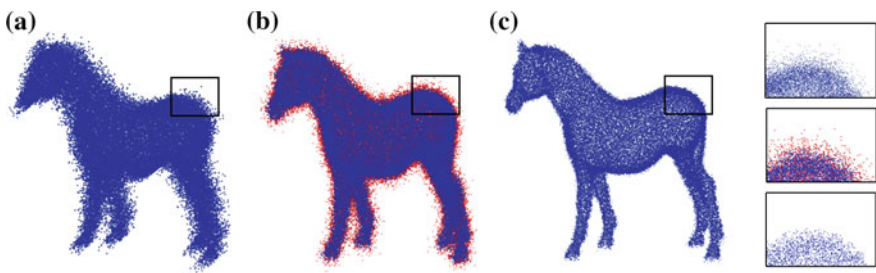


Fig. 3 Results of outliers removal on horse model. **a** The original point cloud with outliers. **b** The outliers detected by our method demonstrated in red color. **c** The resultant point cloud after removing outliers (see the back side of the horse in zoom-views)

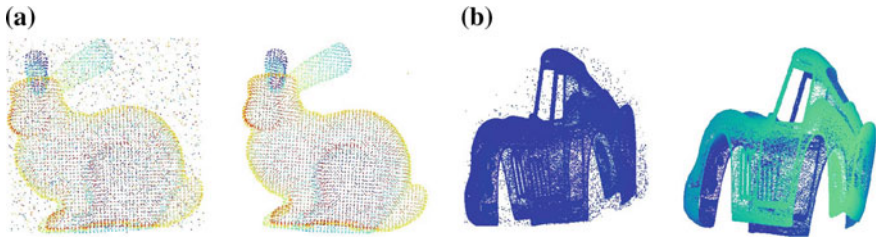


Fig. 4 Results of the outliers removing algorithm: **a** *Left* Bunny point cloud with uniformly distributed noise, *Right* Corresponding point cloud after using the method used in this research, **b** Point cloud model of a car with outliers data, *Right* Car model after removing outliers

Table 1 Computation time of the optimal bandwidth selection and outliers removal method for various models

Data	Input points	Filtered points	T_h	T_f
Bunny	330 K	293 K	29 s	30 s
Car	720 K	600 K	1 m 45 s	1 m 25 s
Horse	364 K	214 K	44.43 s	18.45 s
Gargoyle	2.1 M	796 K	3m 2s	4 m 44 s
Head	1.9 M	1.2 M	2 m 5 s	3 m 25 s

3.2 Experimental Results and Analysis

In order to validate the performance and effectiveness of the proposed method, the estimated performance in both synthetic and real world dataset is tested. This program is written in Python, and the data obtained from the research is from the tests running on Intel Core-i7 CPU 860 at 2.80 GHz plus 4 GB RAM PC. Figure 4 demonstrates some of the results of this method (standford bunny and car models). To evaluate the applicability of this method, both raw and synthetic noise was tested. Table 1 summarizes timings and the parameters used to generate the results. T_h and T_f is the computation time for the optimal bandwidth selection using PSO and outliers removal using thresholding scheme respectively.

Results of the denoising approach on scanned data show that it has good performance on different types of scanned data. Experiments illustrate the capability of this method in removing outliers.

3.3 Algorithm for Noise Removal

After removing outliers, triangular mesh for the rest of the points is calculated. After that, the noise from the mesh by applying bilateral mesh filter [8] is removed.

Bilateral mesh filter is popular for its ability to remove noise and preserve features at the same time. For denoising a point p , other points around its neighborhood $\{q_i\}$ is identified; where, $\|p - q_i\| < \rho = |2\sigma_c|$, here σ_c is the radius of the neighborhood of point p . To determine a desire value for σ_c , the average distance of all adjacent triangles in an input mesh is used. The bilateral filters in this method also have two parameters, σ_c and σ_s . The σ_s is the standard deviation of the offsets in the selected neighborhood. The result of bilateral mesh filtering operation on the input model is illustrated in Fig. 5.

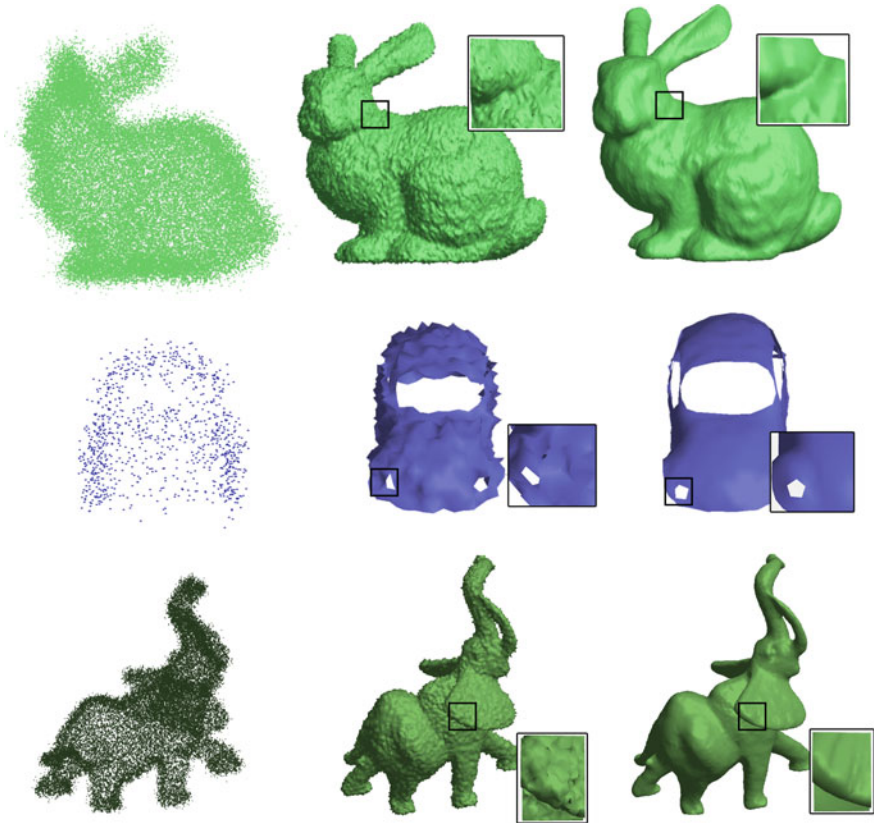


Fig. 5 Point cloud denoising results: The *first column* is the point cloud models after removing outliers; *second column* represents the triangular mesh representation of the point cloud and *third column* is the models after applying bilateral mesh denoising

4 Conclusion

In this paper, a density-based approach for denoising point cloud data is proposed. In particular, particle swarm optimization technique is used for estimating the optimal bandwidth matrix of kernel density estimation, and mean-shift clustering technique to remove outliers. To remove noise from remaining points, bilateral mesh denoising method is applied. Our method is very robust with highly noisy dataset, as can be seen from the examples shown. While this research illustrates only a few applications of the method within the domain of point cloud data, it is believed that its simplicity and effectiveness will lead to its application in other domain.

References

1. Berger M, Levine JA, Nonato LG, Taubin G, Silva CT (2013) A benchmark for surface reconstruction. *ACM Trans Graph (TOG)* 32(2):20
2. Berger M, Tagliasacchi A, Seversky L, Alliez P, Levine J, Sharf A, Silva C (2014) State of the art in surface reconstruction from point clouds. *EUROGRAPHICS star reports*, vol 1, pp 161–185
3. Blum C, Gro R (2015) Swarm intelligence in optimization and robotics. In: *Springer handbook of computational intelligence*. Springer, pp 1291–1309
4. Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. *IEEE Trans Pattern Anal Mach Intell* 24(5):603–619
5. Desbrun M, Meyer M, Schroder P, Barr AH (2000) Anisotropic feature-preserving denoising of height fields and bivariate data. In: *Graphics interface*, vol 11, pp 145–152. Citeseer
6. Duong T, Hazelton ML (2005) Convergence rates for unconstrained bandwidth matrix selectors in multivariate kernel density estimation. *J Multivariate Anal* 93(2):417–433
7. Fleishman S, Cohen-Or D, Silva CT (2005) Robust moving least-squares fitting with sharp features. In: *ACM transactions on graphics (TOG)*, vol 24. ACM, pp 544–552
8. Fleishman S, Drori I, Cohen-Or D (2003) Bilateral mesh denoising. In: *ACM transactions on graphics (TOG)*, vol 22. ACM, pp 950–953
9. Guidoum AC (2013) Kernel estimator and bandwidth selection for density and its derivatives
10. Hsieh ST, Sun TY, Liu CC, Tsai SJ (2009) Efficient population utilization strategy for particle swarm optimizer. *IEEE Trans Syst Man Cybern Part B: Cybern* 39(2):444–456
11. Hyndman RL, Zhang X, King ML (2004) Bandwidth selection for multivariate kernel density estimation using MCMC. In: *Econometric Society 2004 Australasian meetings*. No. 120, Econometric Society
12. Ratnaweera A, Halgamuge SK, Watson HC (2004) Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. *IEEE Trans Evol Comput* 8(3):240–255
13. Shi Y, Eberhart R (1998) A modified particle swarm optimizer. In: *The 1998 IEEE International Conference on evolutionary computation proceedings, IEEE world congress on computational intelligence 1998*. IEEE, pp 69–73
14. Song Y (2010) Boundary fitting for 2d curve reconstruction. *Vis Comput* 26(3):187–204
15. Sotoodeh S (2006) Outlier detection in laser scanner point clouds. *Int Arch Photogram Remote Sens Spat Inf Sci* 36(5):297–302
16. Wang Y, Feng HY (2015) Outlier detection for scanned point clouds using majority voting. *Comput-Aid Des* 62:31–43
17. Zhang X, King ML, Hyndman RJ (2006) A Bayesian approach to bandwidth selection for multivariate kernel density estimation. *Comput Stat Data Anal* 50(11):3009–3031