Journal of Engineering and Sustainable Development
http://jeasd.uomustansiriyah.edu.iq/index.php
https://doi.org/10.31272/jeasd.25.2.1

JEASD

Vol. 25, No. 02, March 2021
ISSN 2520-0917

POINTS CLOUD PRE-PROCESSING AND SAMPLING BASED ON DISTANCE ALGORITHM TECHNIQUE

***Ali M. Al-Bdairy[1]** **Ahmed A. A. Al-Duroobi[2]** **Maan A. Tawfiq[1]**

1) Production Engineering and Metallurgy Dept. University of Technology Baghdad, Iraq.
2) Prosthetics and Orthotics Engineering Dept. Al-Nahrain University Baghdad, Iraq.

Received 30/6/2020 ***Accepted in revised form 20/9/2020*** ***Published 1/3/2021***

Abstract: Although the rapid development of reverse engineering techniques such as a modern 3D laser scanners, but can't use this techniques immediately to generate a perfect surface model for the scanned parts, due to the huge data, the noisy data which associated to the scanning process, and the accuracy limitation of some scanning devices, so, the present paper present a points cloud pre-processing and sampling algorithms have been proposed based on distance calculations and statistical considerations to simplify the row points cloud which obtained using MATTER and FORM 3D laser scanner as a manner to obtain the required geometrical features and mathematical representation from the row points cloud of the scanned object through detection, isolating, and deleting the noised points. A MATLAB program has been constructed for executing the proposed algorithms implemented using a suggested case study with non-uniform shape. The results were proved the validity of the introduced distance algorithms for pre-processing and sampling process where the proficiency percent for pre-processing was (18.65%) with a single attempt, and the counted deviation value rang with the sampling process was (0.0002-0.3497mm).

Keywords: *Point Cloud, Pre-Processing, Sampling, Distance Technique.*

1. Introduction

The 3D laser scanner can be considered as one of the modern technologies which used for reverse engineering applications, but due to accuracy limitation, environmental noise sources such as lighting and contrast, the complexity and reflectivity property of the scanned objects, there are some of the problems which associated to acquired points cloud from the 3D scanning process such as a huge number of points and appearance of unnecessary data points, so the attention of researcher increased to treat with the same problems to simplification the row points clouds and produce a points clouds without noisy points with a minable number of points. D. Xia et al performed numerical and analytic investigations to reduce the noise data properties for regions-of-interest ROI to reconstruct the images based on chord image reconstructions using different scanning configurations [1].

M. Wand et al present an algorithm for out of core outlier removal and geometrical filtering tool for multi-resolution data structure visualization, interactive selection, transformation, painting, of huge scanner data (point cloud) sets [2].

Z. Min was present an algorithm to surface reconstruction, depending on determining the primitives of the surfaces and construct blends surfaces between them, the normal vectors were used to classify the surface primitives into triangular patches as a manner to get the

**Corresponding Author:* 70120@uotechnology.edu.iq

boundary representations and model reconstruction [3].

J. Liu et al proposed an effective and simple reconstruction method based on interpolation of the B-spline surfaces from section curves for blade surface with little measured points, to improve the aerodynamics performance of blade surfaces. The little measured points of section curves were acquired from the coordinate measuring machine [4].

B. Cyganek et al proposed a multidimensional data classification approach to processing complex structures. This approach depends on a tensor kernel applied to train of the Support Vector Machines (SVM) and Sequential Minimal Optimization (SMO) with a chord distance kernel [5].

K. W. Lee and P. Bo proposed a reconstruction method for feature curves based on the intersections of the developable strip pairs which represent the regions besides the features to getting the smooth feature curves from the point cloud of model [6].

J. Kisztner et al introduce an algorithm based on image processing and remote sensing to generate procedure and software of cluster and spectral behavior analysis for general use to vegetation separation based on the 3D cloud which acquired by Terrestrial Laser Scanning [7].

N. Leal et al presents a linear model to simplify the point cloud, taking the density assessment of the point cloud, to determining and ignoring the outlier and noise points, then use the curvature to clustering the point cloud set and detecting the noise points which have a high curvature to decrease the point cloud sets [8].

S. Gauthier et al proposed an analyze and segment method based on digitized 3D surface mesh curvature histogram taking a real object through identifying the valley and peak primitives, also suggests using the analysis of curvature histogram to determine the quality of mesh for varied reverse engineering steps, which can be used to digitize a CAD model from a digitized one. [9].

C. Mineo et al, introduces a boundary point detecting algorithm and spatial filtering approach based on Fast Fourier Transform (FFT), the algorithm gives the tessellated surfaces directly with less noise from the data point cloud, through identifying the points of sharp edges and creases which are not based on pre-defined threshold values [10].

In the present paper a points cloud data pre-processing and sampling algorithms have been proposed, based on the 3D distance calculations and some of the statistical considerations, the acquired points cloud was obtained using the 3D laser scanner as a modern device for reverse engineering domain. As a result to implement the proposed algorithms the row points cloud will be as possible as free of noisy points with reasonable total points number in processed points cloud, while preserving the geometrical topology of the studied case study. The presented paper has been organized as follows: in section 2 the proposed pre-processing algorithm based on the distance calculation will be demonstrate, then the methodology of this algorithm and the adopted program that built using MATLAB software will be illustrated in sub sections I and II respectively. In section 3 the suggested case study will be show and scan using MATTER and FORM 3D laser scanner device to digitize the studied case study as points cloud, also the results of the implementation of the adopted pre-processing for the intended case study has been presented. The proposed sampling algorithm and it's methodology have been presented in section 4, while sections 5 will be illustrated and record the

results, and discussion for the studied case study. Finally, in sections 6 the conclusion will be recorded and illustrate.

2. Distance Pre-Processing Algorithm

Such a method measures the distance between the adjacent data points as a basis to determine whether to delete the point or saved. The distance algorithm can be considered as relatively simple method with regarding of the mathematical expression, and produce little change in curvature of the data point cloud. The algorithm takes into consideration a constant distance between the adjacent points in the point cloud which acquired from a 3D laser scanner device. According to that, whenever the distance becomes more than the specific value, that indication to presence a noise point need to be processed.

I. Distance Algorithm Methodology

The distance algorithm is mathematically simplest, fastest, easy to calculate, and directly treat with the original points clouds. The distance algorithm has better effectiveness to detect the noisy points with small curvature changes. While it is ineffective with points that have high curvature changes [11].

The algorithm of minimum distance is based on setting a minimum value of distance (d_{min}), then comparing this value with the distance value for each points pair along the direction of the scanning line ordered (d) using the distance equation, which is commonly used to compute the distance as squared of the distance between two points $X = (x_0, x_1)$, $Y = (y_0, y_1)$ and $Z = (z_0, z_1)$ [12].

$$\text{Distance}^2(X,Y,Z) = (x_0-x_1)^2 + (y_0-y_1)^2 + (z_0-z_1)^2 \quad (1)$$

Fig. 1 show the arbitrary set of points from the point cloud, and descript the basic distance between these points, it can notice the effect of the noised points to measure the distance (d_3) which is a greater than the others measured distances such as (d_1 and d_2).

If recorded $d_{min} \geq d$, after comparison for points, that referred to the scanning points is in uniform order. Otherwise, the other points need to be remove according to the state of the assessment. The methodology of distance algorithm could be performing with the basic steps using a MATLAB program software:

step1: Input the row point cloud,

step2: set the iteration i for the points in point cloud,

step3: compute the minimum distance d_{min} ,

step4: compute the distance d between each pair of adjacent points in the point cloud, using "eq. (1)".

step5: compare the distance d with the minimum distance d_{min} and take one of the decisions,

a: if $d \leq d_{min}$ that indicate, the point with iteration i is the point belonged to the surface of the scanned object, so it will be save.

b: if $d > d_{min}$ that indicate, the point with iteration i is the noised point need to be delete, and set i=i+1.

step6: repeat steps 4 and 5 to perform the calculation and assessment for all points in the point cloud.

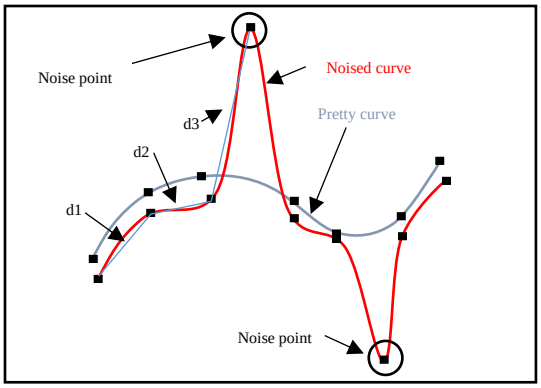


Figure1. Demonstrate the Distance Measuring Sketch.

II. Proposed Distance Algorithm Program

A general MATLAB program is created to execute the adopted distance algorithm procedure. Fig. 2 show the flowchart of the basic steps for the building pre-processing program. The constructed program begins with receiving the measured data as the points cloud. Then compute the mean distance value between the overall points of the points cloud. After that, the distance for each two-adjacent points will be computed and compared with the mean value of the distance continuously and individually. Finally, the decision will be made according to the comparison assessment.

3. The Suggested Case Study

To ensure the validity of the proposed algorithm to simplify the row points clouds, a case study with non-uniform shape has been selected which showed in Fig. 3.

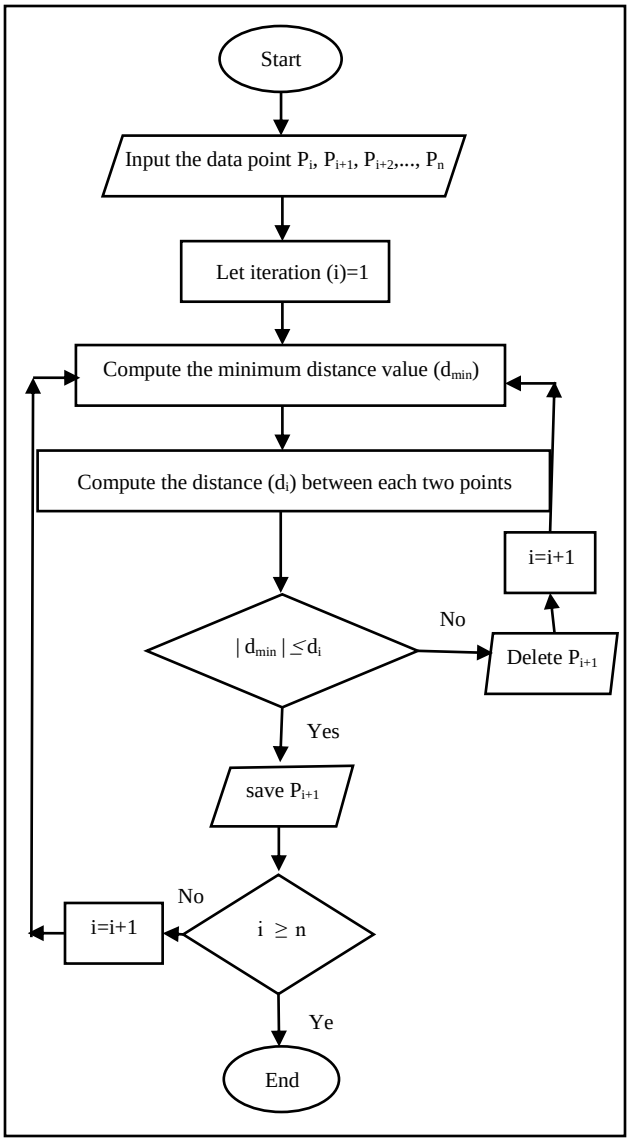


Figure 2. The Proposed MATLAB Program of Distance Pre-Processing Algorithm.

Figure 3. The Proposed Case Study.

The studied case study has been 3D scanned using a 3D laser scanner device named ``MATTER and FORM``, the output of the scanning process was a data file in form of points cloud with x,y,z coordinate axis that achieved by the assisted software for the scanner device. Then the data file is imported in Excel software to rearrange, prepare the points cloud and save it in (.xlsx) format to process with the built MATLAB program of the adopted distance pre-processing algorithm as mentioned in subsection II from section 2. The scanned points cloud was contained (1133648) points and as a result to implement the proposed pre-processing algorithm for a single attempt the points cloud become contain (922194), throughout detection and delete the noisy points. the resulted point cloud has been saved with data format ('file name.dat.'), to preparing for the surface reconstruction process.

4. The Proposed Statistical Mean Sampling Algorithm

The statistical mean algorithm data point sampling is simple to apply and proved a noticeable sampling process, it's based on a uniform steps between the sequential points in the points cloud to compute the average for adjacent points, it can use the statistical mean to obtain the average according to the proposed formula:

$$x_{new}(i)=(x(i)) +x(i+1) +\dots. +xn). /n \quad (2)$$

The deviation value was computed for all points along the streamline of the point cloud. That depending on computing the distance between the opposite points for pre-processed and sampled points using the proposed "equation (3)" for measuring the deviation distance.

$$d=\sqrt{((x_{new}(i+1)-x(i))^2+(y_{new}(i+1)-y(i))^2)} \quad (3)$$

After complete the data pre-processing for acquisition point cloud using the proposed algorithm, it can simplify the resulted point cloud furthermore using the adopted sampling process, which will be based on the jumping of the sequence in the points cloud orientation. After that, the mean and minimum deviation will be computed between the resulted points clouds of the pre-processing algorithm and the points cloud of the sampling process. The jumping process means make the steps between the adjacent points. It can be used the equation (2) to compute the mean value for the point between the adjacent points. The deviation value was computed for all points along the streamline of the point cloud, that depended on computing the distance between the opposite points using the distance equation (3).

The adopted sampling process was developed using a building MATLAB software, and implemented to a proposed case study for explain the effect of the sampling process to enhance the curves rather than the surface of the object.

5. Result and Discussion

As a result to apply the adopted pre-processing algorithm which was based on the distance calculations to simplify the points cloud, the produced cloud of the suggested case study reduced to (507826) points. The noisy points percentage was (18.65%), which was ignored as undesired points with a single attempt for executing the proposed distance algorithm. The indication that the adopted distance algorithm gives acceptable results for simplifying the data clouds through determining and removing the noisy points. Fig. 4 shows the resulted points cloud for the studied case study, where it is illustrated the points cloud without noisy point faraway of the cloud surface, and this cloud can be processed to produce a pretty surface model of the object.

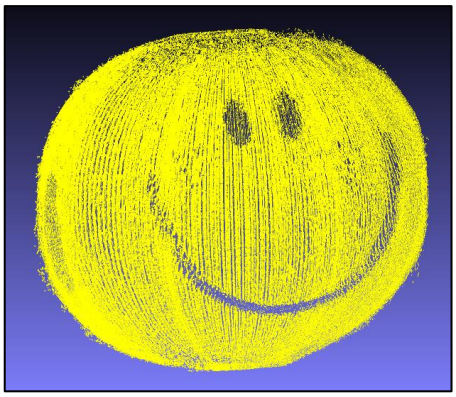


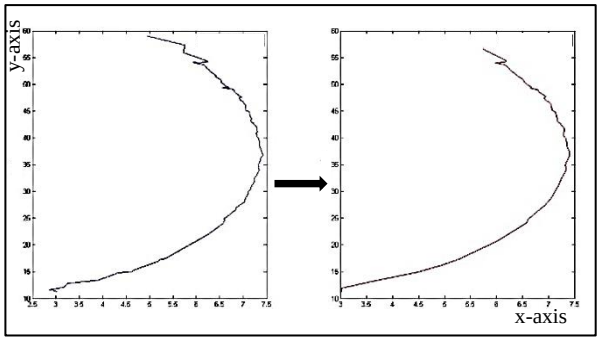
Figure 4. Resulted Point Cloud of The Studied Case Study.

Table 1 presents the result of the sampling process with different percentages. Shows the mean and minimum deviation values that occur at each percentage value of sampling. Where, it is illustrated that, the deviation value is about (0.0002-0.3497) mm, that can be produced with the 80% sampling, while the deviation value is range (0.0002-0.1769) mm at 50% sampling. Therefore, the selected sampling percent will be depended on the required deviation range.

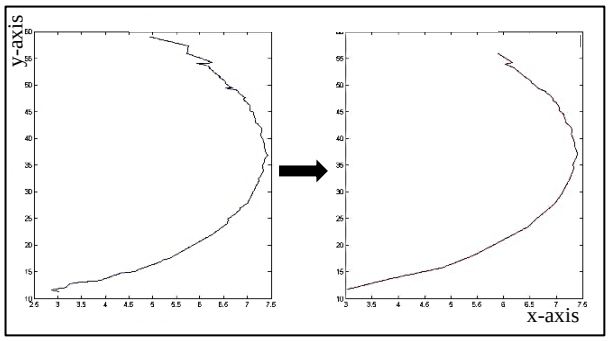
Table 1. Sampling Results for the Studied Case Study.

Percent of sampling%	Number of resulted points	Mean deviation	Minimum deviation
80	113915	0.3497	0.0002
75	142394	0.3250	0.0004
66	189858	0.2815	0.0002
50	284788	0.1769	0.0002

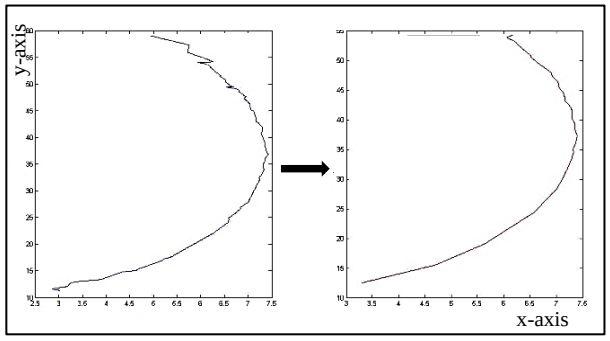
The result of the sampling process is shown in Fig.5 for the segment curve of the studied case study. Where, the enhancement of the resulting shape of the case study will be achieved with increasing the percentage of sampling, which leads to a decrease in the fluctuating of the beaks.



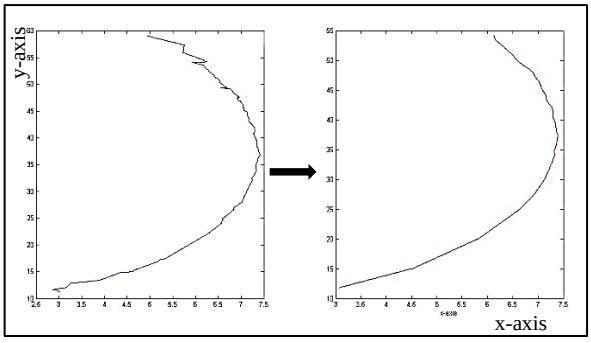
a. 50% sampling



b. 66% sampling



c. 75% sampling



d. 80% sampling

Figure 5. Sampling Results of Curve Segment for the Studied Case Study.

The surface reconstruction process was done using the facilities of Meshlab software for the proposed case study as shown in Fig.6.

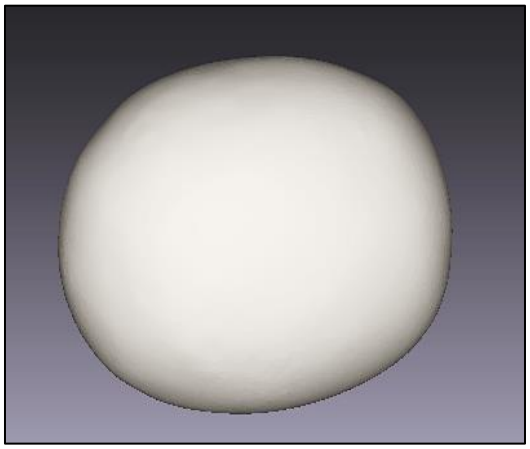


Figure 6. Resulted Surface Model for The Studied Case Study.

The surface fitting process was conducted for the resulted points cloud that pre-processed and sampled according to the developed algorithms. The surface fitting process achieved using the facility of the MATLAB software to generate a mathematical expression for the studied case study.

The result of the fitting process is a linear polynomial with third degree for x and y directions.

$$F(x,y)= n_{00} + n_{10}*x + n_{01}*y + n_{20}*x^2 + n_{11}*x*y + n_{02}*y^2 + n_{30}*x^3 + n_{21}*x^2*y + n_{12}*x*y^2 + n_{03}*y^3 \quad (4)$$

Table 2 present the Coefficients of the produced polynomial for the studied case study.

6. Conclusion

In the present paper, two simplification algorithms have been proposed to treat with the acquired points cloud of the scanned objects. The results of the pre-processing algorithm proved the proficiency of the adopted algorithm to detect, isolate and delete the noisy points in the points cloud, that based on the instantaneous distance calculations between every adjacent pair of points. The achieved percent of deleting the noisy points was (18.65%) from the total number of points for a single attempt of application the adapted algorithm. Therefore, the large number of points in the point could of the scanned objects can be treated with the adopted distance algorithm. The results of the proposed sampling algorithm showed the validity of this algorithm to further sampling process of the points cloud based on the deviation and the statistical mean calculations preserving the geometrical surface topology of the scanned objects. The proposed sampling algorithm was conducted with different percentages of sampling varying from 50% to 80%, the deviation value is about (0.0002-0.3497) mm with the 80% sampling, while the deviation value is range (0.0002-0.1769) mm at 50% sampling. Also, it can be noticed decreasing the fluctuating with a higher percent of sampling. Finally, the selected sampling percent will be depended on the required deviation range.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Table 2. Polynomial Coefficients of the Proposed Case Study.

Coefficient	n_{00}	n_{10}	n_{01}	n_{20}	n_{11}	n_{02}	n_{30}	n_{21}	n_{12}	n_{03}
Value	50.84	0.1455	-0.03677	-0.01994	-0.0007274	-0.0186	-0.0001763	-0.0002367	0.0003491	0.0004609

7. References

1. D. Xia et al, ``Noise Properties of Chord-Image Reconstruction``, IEEE Transactions on Medical Imaging, vol. 26, No. 10, 2007.
2. M. Wand et al, ``Processing and interactive editing of huge point clouds from 3D scanners``, Elsevier, Computers and Graphics, vol. 32, pp.204–220, 2008.
3. ZHOU Min ''A New Approach of Composite Surface Reconstruction Based on Reverse Engineering'', Elsevier, Procedia Engineering, Vol. 23, pp. 594 – 599, 2011.
4. J. Liu, J. Zhao, X. Yang, J. Liu, X. Qu, and X. Wang, "A Reconstruction Algorithm for Blade Surface Based on Less Measured Points", Hindawi Publishing Corporation, IJAE, International Journal of Aerospace Engineering, Vol. 2015, Article ID 431824, 2015.
5. B. Cyganek, B. Krawczyk, and M. Woźniak, "Multidimensional Data Classification with Chordal Distance Based Kernel and Support Vector Machines", Elsevier, EAAI, Engineering Applications of Artificial Intelligence, Vol. 46 pp. 10–22, 2015.
6. K. W. Lee, and P. Bo, ``Feature Curve Extraction from Point Clouds via Developable Strip Intersection``, Elsevier, journal of Computational Design and Engineering, Vol. 3, pp. 102–111, 2016.
7. J. Kisztnera, et al, ``3D Documentation of Outcrop by Laser Scanner — Filtration of Vegetation``, Elsevier, Journal of Perspectives in Science, Vol. 7, pp. 161—165, 2016
8. N. Leal, E. Leal, and S.T. German, ``A Linear Programming Approach for 3D Point Cloud Simplification``, IAENG International Journal of Computer Science, Vol. 44:1, 2017.
9. S. Gauthier, W. Puech, R. Bénière, and G. Subsol, "Analysis of Digitized 3D Mesh Curvature Histograms for Reverse Engineering", Elsevier, Computers in Industry, Vol. 82, pp. 67–83, 2017.
10. C. Mineo, S. G. Pierce, and R. Summan,'' Novel Algorithms for 3D Surface Point Cloud Boundary Detection and Edge Reconstruction'', CDE, journal of Computational Design and Engineering, Vol. 6, pp. 81–91, 2019.
11. S. Wang et al, ``Measured Scan Data Reduction Combination Algorithm Based on Scan Line``, Applied Mechanics and Materials, Trans Tech Publications, Switzerland, vols. 651-653, pp 2335-2338, 2014.
12. P. J. Schneider and D. H. Eberly, ``Geometric Tools for Computer Graphics``, Elsevier Science, Morgan Kaufmann Publishers, 2003.