Point Data Pre-Processing Based on Fuzzy Logic for Reverse Engineering Modelling

Igor Budak1,∗ - Mirko Soković2 - Janez Kopač2 - Janko Hodolić1
1 Faculty of Technical Sciences, Novi Sad, Serbia
2 Faculty of Mechanical Engineering, Ljubljana, Slovenia

State of the art 3D-digitization systems designed for reverse engineering (RE) modelling are characterized by increased scanning speed and also by the possibility to generate a large number of points in a short time. Generally, this improves the efficiency of the RE-modelling process. In practice however, a significant number of points in the stage of the CAD model generation may become an issue. In addition, the 3D-digitization process is often connected with measuring errors because of the properties of measuring systems and/or the object. Therefore, nowadays considerable attention is devoted to the problem of pre-processing the results of 3D-digitization, often by means of tools of artificial intelligence.

This paper introduces a pre-processing system based on fuzzy logic, along with the results of its practical application.

Keywords: reverse engineering, pre-processing, 3D-digitization, point data reduction

0 INTRODUCTION

State of the art 3D-digitization systems which are applied in reverse engineering modelling are characterized by increased scanning speed as well as by the possibility to generate a large number of points in a short time. Actual laser 3D-scanners are able to generate over 40,000 points per second. Generally, this improves the quality and efficiency of the RE-modelling process. In practice, however, later in the stage of CAD model generation, a huge number of points which are generated in the stage of 3D-digitization may become a serious problem. Also, the 3D-digitization process is often connected with measuring errors because of the nature of measuring systems themselves, the properties of digitized objects and the operator’s subjective mistakes.

Considering the above mentioned, the stage of pre-processing the results of 3D-digitization, which includes error filtering, data smoothing and the most sophisticated process of data-point reduction, becomes inevitable in almost any RE system. The fact that pre-processing the results of 3D-digitization has become a frequent subject of recent research with the central role of reducing the number of points, confirms the previous statement.

Therefore, based on different approaches, a large number of systems is developed for pre-processing the results of 3D-digitization, which depend on the RE methodology above all, i.e. on the 3D-digitization system for which they are designed. This applies to all the three pre-processing segments: filtering, smoothing and reduction. However, it should be noted that the reduction depends on RE methodology and to a larger degree on the technique of 3D-digitization. In other words, most of the filtering and smoothing approaches may be successful in RE systems based on different methodologies, while in the case of reduction it is very rare.

Among the many data-point reduction approaches, it is possible to integrate the methods of artificial intelligence into the process of pre-processing, i.e. into point data reduction above all, in order to achieve better quality and process efficiency.

Point reduction at the "cross-sectional" RE methodology is based on the so called sampling methods, among which the uniform (factor),
spatial, chordal and angular are the most known [5] and [6]. All these methods are based on different parameters (based on which they decide on point reduction). Each of the methods have certain advantages/drawbacks on different (scanned curves in) point clouds. In other words, each method is more or less efficient (yields in larger level of reduction with smaller error-deviation).

This paper introduces a system for integrated pre-processing the results of 3D-digitization which is designated primarily for RE modelling systems based on the "cross-sectional" approach where special attention was paid to point data reduction module based on fuzzy logic.

1 SYSTEM FOR POINT DATA PRE-PROCESSING

The system developed for point data pre-processing (the working title of the SiPreF) which is primarily intended for the needs of reverse engineering modelling on the bases of the "cross-sectional" methodology is designed on the modular principle (Fig. 1) and consists of:

1) Module for 3D filtering of points (errors);
2) Module for extraction of point data in cross-sections;
3) Module for filtering and smoothing of point data in cross-sections;
4) Module for reduction of point-data based on fuzzy logic;
5) Module for formatting of output data.

The system for traditional data-point pre-processing was the starting point for developing the SiPreF system, which is presented in more details in [7]. The following are the main improvements of the SiPreF as compared to the above mentioned basic system:

- development and integration of decision-making procedures on point reduction based on fuzzy logic in the spatial method, chordal method and the method of straightness.

Fig. 1. General model of the SiPreF system

In the realization of the SiPreF point data pre-processor, conventional programming tools as well as artificial intelligence (fuzzy logic) tools were used. SiPreF was developed in Matlab (MathWorks) 7.2 (Release 2006 a).

The development of the modules, i.e. the integrated software tools was based on a wider analysis of the most often used pre-processing tools. SiPreF integrates 17 tools grouped in five previously mentioned modules:

1) Import and (re)orienting of point data in 3D space,
2) Volumetric filtering,
3) Data filtering by segmented line,
4) Elimination of individual points,
5) Selection and sorting of cross-sections data,
6) Change of resolution,
7) Data filtering by the angle method,
8) Filtering /smoothing of data by the median method,
9) Data smoothing by the mean method,
10) Analysis of reduction-induced deviations,
11) Reduction of data by the chordal method with an analysis of deviations,
12) Reduction of data by the straightness method with an analysis of deviations.
13) Reduction of data by the spatial method with analysis of deviations,
14) Fuzzy-based reduction of data by the chordal method,
15) Fuzzy-based reduction of data by the straightness method,
16) Fuzzy-based reduction of data by the spatial method,
17) Generation of output format.

2 MODULE FOR POINT DATA REDUCTION BASED ON FUZZY LOGIC

In the course of SiPreF development major emphasis was placed on the module for point data reduction, which was designed according to a novel approach with integrated deviation analysis and fuzzy logic reasoning. This constitutes the main difference in comparison to the systems developed so far.

The algorithm presentation of the developed software solution of this module is given in Fig. 2.

Building on the weak spots and deficiencies of current approaches to reduction of point data by sampling methods - i.e. a lack of information on the level of deviation in reduced point clouds and the necessity to employ parameters which are abstract to the user - a novel approach was developed for an analysis of the level of deviation of the reduced point cloud in comparison with the initial point cloud. This novel approach introduces an additional parameter termed maximum allowed reduction error (MARE), in the reduction-related decision process with the sampling-based methods. The approach was tested with three different methods of sampling: chord, straightness and spatial (more detail in [7] and [8]) showing a significant improvement of the maximum deviation/level of the reduction ratio, i.e. an improvement of the average deviation and the level of reduction.

In order to improve the process of reduction, the novel approach was enhanced by implementing fuzzy logic in the process of reduction-related decision making. Besides the additional improvement of the deviation/mean level of reduction ratio, implementation fuzzy logic allowed a more user-friendly and intuitive application. The reduction process is controlled by simply entering the deviation tolerance, allowing the user to gain better insight into the quality of the reduced point cloud.

![Algorithm presentation of the module for data-point reduction](Image)

---

**Fig. 2. Algorithm presentation of the module for data-point reduction**

### 2.1 The Novel Approach for Point Data Reduction

The process of point data reduction is, without a question, associated with the deviations of the resulting cross-sectional curves (generated with reduced point data), relative to initial (original) cross-sectional curves (generated with original point cloud). The basic question here knows the tolerable level of deviations for the given level of reduction. This very problem renders the process of point data reduction complex and sophisticated.

One of the largest deficiencies of the actual sampling-based methods for the reduction of point data is the lack of feedback information on the level of deviation of the resulting cross-sectional curve, which is the result of elimination of a point or a set of points. By the same token, these methods also lack decision-making mechanisms, which would allow them to select...
the points for reduction using resulting deviation as a criterion.

The key feature of the novel approach to the sampling-based point data reduction, proposed here, is the procedure for an analysis of deviations of cross-sectional curves, which are the result of point data reduction.

A practical realization of the procedure for the analysis of deviation is based on computation and analysis of maximum deviation of the resulting cross-sectional curve relative to the original cross-sectional curve generated from initial point cloud. Least squares method, modified to meet specific requirements of the problem in hand, was used to compute maximum deviation - designated MRE (Maximum Reduction Error) in this dissertation.

Least squares method, widely used in analysis of errors of curve approximation, evaluates deviation of approximated curve relative to original points (Fig. 3). The method considers the curve for which the sum of squares of deviations $\varepsilon_i(x_i, y_i)$ is minimal as best solution [9]:

$$ S = \varepsilon_1^2 + \varepsilon_2^2 + ... + \varepsilon_n^2. \quad (1) $$

Deviations $\varepsilon_i(x_i, y_i)$ are calculated at points defined by resolution $Q$ (Fig. 4) which can be varied to suit the length of the scanned curve, i.e. the density of scanned points within array.

Besides the parameter MRE, the procedure also employs the ARE (Average Reduction Error) - an additional parameter which allows the assessment of deviation of the resulting cross-sectional curve. ARE is a mean sum of deviations computed at points defined by resolution $Q$ (Fig. 4) and can be expressed as:

$$ ARE = \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i. \quad (3) $$

The methods for reduction of point data by sampling were improved by implementing fuzzy logic into the procedure for decision-making on which the elimination of points is based.

To eliminate the problems which stem from the specific values of decision-critical input parameters entered by the user, and create a more user-friendly system, a new, synthetic parameter was introduced under the name reduction coefficient (RC), and its maximum allowed value was defined as the maximum allowed reduction coefficient (MARC). For all three methods the RC parameter was derived based on method-specific parameters, and an additional input parameter maximum reduction error (MRE), i.e. the maximum allowed reduction error (MARE). MARE was introduced to allow the maximum reduction error to be controlled.

The details of the novel approach for point data reduction are presented here for the case of the spatial method.
The spatial method for point data reduction is based on the parameter of spatial (Euclidean) distance \( (d_E) \) [7] which, together with MRE, was used as input parameter for the fuzzy-logic-based decision-making system.

The structure of this fuzzy system is shown in Fig. 5.

**Fig. 5. Structure of fuzzy system for point data reduction using spatial method**

The input is formed by two state variables – \( d_E \) and MRE, whose values are fuzzified into fuzzy sets with appropriate input spaces, while the fuzzy sets of input values are defined by their membership functions [10] and [11].

Due to its simplicity, triangular membership function was chosen for all state variables.

Figs. 6 a and b show graphs of membership functions for input state variables \( d_E \) and MRE. It should be noted that the input space for \( d_E \) (0 to 2 mm) was defined based on practical experience, while for the state variable MRE, this input space was defined on the basis of real-application experience with \( MARE = 0.05 \) mm as the pivotal parameter. The graphs show that the input space was segmented in the following way - for \( d_E \) it was sectioned into three segments with appropriate linguistic terms (shorter, medium, and longer), while for the MRE it was segmented into three fuzzy sets (slight, moderate and significant).

The output from this fuzzy system is variable \( RC \) (a non-dimensional value) which, for the sake of simplicity, was allotted an output space from 0 to 100 (Fig. 6 c) and the parameter \( MARC \) (whose function was previously explained) must fall within that space. To allow finer control of \( RC \) parameter, the input space was segmented with finer resolution, resulting in a total of nine fuzzy sets denoted with linguistic qualifiers - minor, very low, low, medium-low, medium, medium-high, high, very high, and huge.

Using the defined fuzzy variables (\( d_E \), MRE, and RC) and their belonging fuzzy subsets with their membership functions, fuzzy control rules which represent the knowledge base of the proposed fuzzy system were defined. A total of nine fuzzy rules were defined which are presented in Table 1.

**Fig. 6. Graphs of membership functions for the state variables**

The system was modelled using reference values of output variable, according to which the membership functions of input variables were adjusted. Table 2 gives the reference values of output variable \( RC \), depending on input variable values: \( d_E \) and MRE, which were fundamental to this system. Table 3 reviews the final values of \( RC \) variable (for the same combinations of input
variable values) after the membership functions were adjusted to input variables, and the output variable.

Table 1. Control fuzzy rules of the fuzzy system for point data reduction using spatial method

<table>
<thead>
<tr>
<th>Rule</th>
<th>Condition</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$(d_E$ is Shorter) and (MRE is Slight)</td>
<td>(CR is Huge)</td>
</tr>
<tr>
<td>2.</td>
<td>$(d_E$ is Medium) and (MRE is Slight)</td>
<td>(CR is Very-high)</td>
</tr>
<tr>
<td>3.</td>
<td>$(d_E$ is Longer) and (MRE is Slight)</td>
<td>(CR is High)</td>
</tr>
<tr>
<td>4.</td>
<td>$(d_E$ is Shorter) and (MRE is Moderate)</td>
<td>(CR is Mid-high)</td>
</tr>
<tr>
<td>5.</td>
<td>$(d_E$ is Medium) and (MRE is Moderate)</td>
<td>(CR is Mid)</td>
</tr>
<tr>
<td>6.</td>
<td>$(d_E$ is Longer) and (MRE is Moderate)</td>
<td>(CR is Mid-low)</td>
</tr>
<tr>
<td>7.</td>
<td>$(d_E$ is Shorter) and (MRE is Significant)</td>
<td>(CR is Low)</td>
</tr>
<tr>
<td>8.</td>
<td>$(d_E$ is Medium) and (MRE is Significant)</td>
<td>(CR is Very-low)</td>
</tr>
<tr>
<td>9.</td>
<td>$(d_E$ is Longer) and (MRE is Significant)</td>
<td>(CR is Minor)</td>
</tr>
</tbody>
</table>

Table 2. Reference values for output variable RC depending on the values of input variables $d_E$ and MRE

<table>
<thead>
<tr>
<th>$d_E$ [mm]</th>
<th>MRE [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>95 82.5 72.5 65 35 25</td>
</tr>
<tr>
<td>0.5</td>
<td>90 77.5 70 62.5 30 20</td>
</tr>
<tr>
<td>1</td>
<td>85 75 65 60 25 15</td>
</tr>
<tr>
<td>1.5</td>
<td>80 70 60 55 22.5 10</td>
</tr>
<tr>
<td>2</td>
<td>75 65 55 50 20 5</td>
</tr>
</tbody>
</table>

Table 3. Reference values for output variable RC depending on the values of input variables $d_E$ and MRE after the adjustment of the membership functions

<table>
<thead>
<tr>
<th>$d_E$ [mm]</th>
<th>MRE [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>95.9 81.4 72.5 67 35.3 23.4</td>
</tr>
<tr>
<td>0.01</td>
<td>89.2 76.6 69.5 63.6 31.4 18.8</td>
</tr>
<tr>
<td>0.02</td>
<td>87.5 77.1 67.3 60 25.1 16.7</td>
</tr>
<tr>
<td>0.03</td>
<td>81.2 69.7 62 55.5 24.5 12.5</td>
</tr>
<tr>
<td>0.04</td>
<td>75 64.6 54.8 51.1 22.9 6.14</td>
</tr>
</tbody>
</table>

As the criterion for adjustment of the membership functions, mean square deviation was adopted [9]:

$$E = \sqrt{\frac{e_1^2 + e_2^2 + \ldots + e_n^2}{n}} \quad (4)$$

using the three sigma rule:

<table>
<thead>
<tr>
<th>$e_{\text{max}}$</th>
<th>Acceptable level of adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt; 3E$</td>
<td>Unacceptable level of adjustment</td>
</tr>
</tbody>
</table>

In the first case, considering that $E = 1.5618$ and $|e_{\text{max}}| = 2.9$, the three sigma rule is satisfied: $2.9 < 4.685$ according to which this level of adjustment is considered acceptable.

The mechanism of fuzzy decision-making is based on the Mamdani method. This method uses the minimum of operation, i.e. the minimum of intersection, to form the fuzzy implication function [10] and [11].

The procedure of fuzzy reduction involves the following steps (Fig. 7):

- MARE is set and then used to compute MARC;
- a set of points is selected (two sequential points – i.e. the first and the second point in the first step ) to compute parameters;
- $d_E$ is computed;
- preliminary elimination of the second point in the selected array is performed, in order to compute the MRE parameter;
- using the computed parameters and the preset fuzzy rules, RC is computed;
- based on the comparison of the computed RC against the preset value of MARC, a decision is made whether to eliminate or retain the analyzed point;
- in case the second point in the point array has been eliminated, the new set of points is composed of the first point from the previous set, and the next point in the scanned array, while in the case when no elimination occurred, the new set is composed of the second point from the previous set, and the next point in the scanned array;
- after the last point in the scanned array has been selected, and the first run of the reduction procedure is finished, a change of direction in the selected point set is performed which means that the last point in the scanned array of the last run becomes the first point in the next run, which makes the process of reduction more even;
the process is re-iterated on a reduced scanned point array until it is possible to eliminate at least one point in a single run.

3 PRACTICAL APPLICATION OF SiPreF

Functionality and efficiency of the developed system (SiPreF) are verified through its practical application. Here the results of SiPreF's application on two case studies are presented.

3.1 Case Study 1

The first case study used to test SiPreF involved a sports glasses lens (Fig. 8). The choice of this part, which due to ergonomic intent is of a relatively simple geometry, was motivated by the complexity of digitized data (Fig. 9) which requires adequate fixture and locating. 3D-digitization was performed by a contact system Cyclon II-Renishaw, resulting in a total of 412,111 points, of which a large number represent error-points, which actually belong to the fixture and measuring table.

![Fig. 8. Sports glasses and a lens](image)

![Fig. 9. The results of 3D-digitization of a sports glasses pane](image)

The pre-processing included 3D-filtering (volumetric filtering, filtering by segmented line, and elimination of individual points), cross-sectional filtering/smoothing of point data (elimination of end points), a change of resolution and a reduction of point data.

The point cloud subject to reduction contains 109,528 points in 214 cross-sections. Fuzzy-chordal reduction method was chosen, with $MAD = 0.03$ mm. The results of reduction are presented in Table 4 and in Fig. 10.

<table>
<thead>
<tr>
<th>Table 4. Results of point data reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD [mm]</td>
</tr>
<tr>
<td>Maximum error [mm]</td>
</tr>
<tr>
<td>Average error [mm]</td>
</tr>
<tr>
<td>No. of eliminated points</td>
</tr>
<tr>
<td>Reduction level [%]</td>
</tr>
<tr>
<td>No. resulting points</td>
</tr>
</tbody>
</table>
Surface model was generated in Pro/ENGINEER Wildfire 4 by automated generation of cross-sectional curves on the bases of reduced point cloud in IBL format (Fig. 11).

Verification of the generated surface model from “reduced” point cloud was conducted through a comparison of the deviation of the “reduced” and “original” surface models from the “original” point cloud (used as input in reduction process). Deviations were analysed by the application of CAD-inspection technique. The results are shown in Fig. 12. Numerical values for maximum positive and negative deviations are given in Table 5.

Judging by the graphical output from CAD-inspection and the maximum values, one can conclude that the level of deviation of the “reduced” surface model is very close to that of the “original” surface model. More regions with deviations within the “reduced” surface model are direct consequences of the approximate surface generation on the bases of reduced point data.

3.2 Case Study 2

The second verification of SiPreF was performed on artistic plaquette (Fig. 13).
Compared to the previous example, this object has much more complex geometry, but is also easier to set up in the fixture, considering its flat bottom surface. In that respect, this case study focuses on the reduction aspects of SiPreF application because the preparation of 3D-digitization results (Fig. 14) is much simpler. The result of 3D-digitization comprises a total of 2,970,761 points.

Before the reduction phase, pre-processing included 3D-filtering (volumetric filtering), cross-sectional filtering/smoothing of point data (elimination of end points) and change of resolution.

Prepared for reduction, the point cloud contains 1,117,774 points in 763 cross-sections. Reduction in this verification was performed using the fuzzy-spatial method, for MAD = 0.03 mm (Table 6, Fig. 15).

As in the case study 1, surface model was generated in Pro/ENGINEER Wildfire 4 by automated generation of cross-sectional curves on the bases of reduced point cloud in IBL format (Fig. 16).

The results of CAD-inspection of deviation of the “original” and “reduced” surface models in comparison with “original” point cloud are shown in Fig. 17. Numerical values of maximum positive and negative deviations are given in Table 7.

A graphical representation of CAD-inspection results and extreme deviations indicates very similar deviations for both surface models. Furthermore, the deviation of “reduced” model is even slightly smaller than that of the “original”, which can be attributed to the more efficient surface approximation due to smaller number of points.

To verify the functionality of the generated model, in this case study a physical model was manufactured using the machining technology (Fig. 18).
To provide additional verification, comparison was made between deviations of the machined physical model and the “reduced” surface model. The procedure included 3D-digitization of the physical model by Cyclon II – Renishaw system, followed by CAD-inspection of deviation of the resulting point cloud in comparison with the “reduced” surface model. The analyses were performed in Pro/ENGINEER module Pro/VERIFY (Fig. 19, Table 8).

Table 8. Extreme deviations obtained by CAD-inspection

<table>
<thead>
<tr>
<th>Deviation [mm]</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.1706</td>
<td>0.1889</td>
<td></td>
</tr>
</tbody>
</table>

Considering the complexity of surface, deviations of “reduced” surface model, the accuracy of the machining system, and the accuracy of the scanning system, the results obtained by verification of the machined physical model confirm practical applicability of the SiPreF.

4 CONCLUSION

Within the paper a system for preprocessing the results of 3D-digitization, designed for use in systems for RE-modelling based on cross-sectional methodology, has been presented.
Special attention was paid to a module for point data reduction, which is the central and most complex part of the system. Building on the weak spots and deficiencies of current approaches to reduction of point data by sampling methods - i.e. the lack of information on the level of deviation in reduced point clouds and necessity to employ parameters which are abstract to user – a novel approach was developed for analysis of the level of deviation of the reduced in comparison with the initial point cloud. The novel approach was enhanced by implementing fuzzy logic in the process of reduction-related decision making. Beside the additional improvement of the deviation/mean level of reduction ratio, implementation fuzzy logic allowed a more user-friendly and intuitive application.

This paper also provides practical results of the application of the introduced system. Although still in the experimental stage, the results show that it is satisfying.

Future research will be directed towards analyzing parameter relations, i.e. adequate functions of affiliation in fuzzy procedures with the goal of fine-tuning the performance of the methods themselves as well as adjusting the level of reduction coefficients among the methods.

5 REFERENCES


