Computer-Aided Design 62 (2015) 31–43

Contents lists available at ScienceDirect

Computer-Aided Design

journal homepage: www.elsevier.com/locate/cad

Outlier detection for scanned point clouds using majority voting[☆]

Yutao Wang, Hsi-Yung Feng [*]

Department of Mechanical Engineering, The University of British Columbia, Vancouver, BC, Canada V6T 1Z4

HIGHLIGHTS

- A robust method to detect and remove all types of outliers in scanned point clouds.
- Ability to preserve valid point clusters of small size.
- Effectiveness validated with a variety of scanned point clouds.

ARTICLE INFO

Article history:
Received 26 April 2014
Accepted 10 November 2014

Keywords:
3D Laser scanning
Point cloud
Outlier detection
Non-isolated outliers
Majority voting
Feature preservation

ABSTRACT

When scanning an object using a 3D laser scanner, the collected scanned point cloud is usually contaminated by numerous measurement outliers. These outliers can be sparse outliers, isolated or non-isolated outlier clusters. The non-isolated outlier clusters pose a great challenge to the development of an automatic outlier detection method since such outliers are attached to the scanned data points from the object surface and difficult to be distinguished from these valid surface measurement points. This paper presents an effective outlier detection method based on the principle of majority voting. The method is able to detect non-isolated outlier clusters as well as the other types of outliers in a scanned point cloud. The key component is a majority voting scheme that can cut the connection between non-isolated outlier clusters and the scanned surface so that non-isolated outliers become isolated. An expandable boundary criterion is also proposed to remove isolated outliers and preserve valid point clusters more reliably than a simple cluster size threshold. The effectiveness of the proposed method has been validated by comparing with several existing methods using a variety of scanned point clouds.

1. Introduction

Measurement outliers are inevitable by-products of 3D scanning. When scanning an object with complex geometry and varying surface reflectiveness, the collected scanned point cloud may contain extensive outliers, which are false measurement points not belonging to the scanned surface. Such erroneous data points pose some problematic issues to the applications of the scanned point cloud. In the context of quality inspection, outliers indicate abrupt large deviations from the reference design model and this may lead to false inspection result. In reverse engineering, since outliers can form dense clusters, they can lead to erroneous surface patches, known as ghost surfaces, which do not exist on the scanned object [1]. Therefore, it is critical to detect and remove these measurement outliers as a preprocessing step. In comparison to the common tedious manual removal process, which is time consuming and relies on the operator's experience, it is highly desirable to develop an automatic outlier detection and removal method. However, automatic and effective removal of measurement outliers is challenging since the scanned object surface model is unavailable and the estimate of the object surface shape would be inaccurate in the presence of extensive outliers.

From a local point density perspective, outliers in a scanned point cloud can be sparsely distributed or densely clustered. The dense outlier clusters can be further classified into isolated clusters and non-isolated clusters. The non-isolated outlier clusters seamlessly connect to the scanned object surface and cannot be separated using a simple distance criterion [1]. Fig. 1 shows the three types of outliers present in a scanned point cloud from a gear. Because the surface is reflective and contains sharp features and concave geometry, extensive outliers are formed around these features due to the undesirable specular reflections, multipath reflections and occlusions in the scanning process [2]. Unlike scanning noise, which is commonly assumed to be randomly distributed, these outliers are systematic and can form ghost surfaces not part of the scanned object. To alleviate the effect of surface

☆ This paper has been recommended for acceptance by Vadim Shapiro.

* Corresponding author. Tel.: +1 604 822 1366; fax: +1 604 822 2403.
E-mail address: feng@mech.ubc.ca (H.-Y. Feng).

http://dx.doi.org/10.1016/j.cad.2014.11.004
0010-4485/

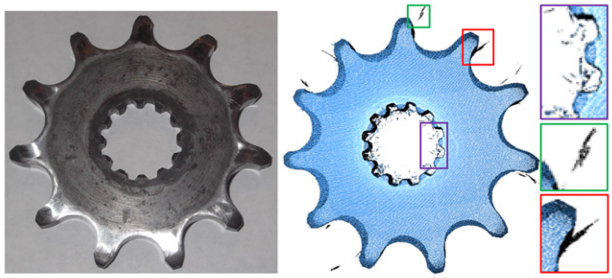

Fig. 1. A scanned point cloud from a gear showing the sparse outliers, isolated and non-isolated outlier clusters.

reflectiveness, it is common to spray the object with a thin layer of gray powder so that the resulting outliers would be less. However, such a treatment may introduce the issues of accuracy reduction and part contamination, and is inapplicable when scanning, for example, final machined parts, medical components, clean room parts, polished parts [1]. Moreover, some industrial applications, such as in-process inspection, require parts being scanned in the mist of their manufacturing operations. In such situations, the extra process of spraying would not be ideal for productivity concern. Therefore, an effective method that can reliably detect outliers is definitely in great demand.

Many outlier removal methods have been proposed in the literature but they generally focus on a certain type of outliers and are inapplicable to other types of outliers. For example, methods to remove sparse outliers [3,4] detect outliers based on low local point density and rely on the condition that the local neighborhood of each point contains sufficient data points. Such methods are unable to detect dense outlier clusters since the local neighborhood of each point in an outlier cluster still contain a good number of data points. To detect outlier clusters, existing methods [5,6] segment a scanned point cloud into many clusters and then set small clusters as outliers, which will mistakenly remove good clusters with smaller numbers of data points. When non-isolated outlier clusters exist in the point cloud, these outliers will be indistinguishable and blend into the scanned object surface. To remove these non-isolated outliers, Shen et al. [1] proposed a surface propagation method based on the assumption that outliers were randomly distributed and should not form smooth surface patches. However, such an assumption is not always valid in practice since non-isolated outlier clusters are in fact systematic and can actually form planar surface patches. Thus, non-isolated outliers remain undetected.

In this work, a comprehensive outlier detection method is developed to effectively identify sparse outliers, isolated outlier clusters, and non-isolated outlier clusters in scanned point clouds. Among the three types of outliers, non-isolated outlier clusters are the most challenging to detect since they are connected to the scanned surface and can even form planar patches. Common clustering methods will always end up with clusters in which non-isolated outliers are mixed together with good surface points. To address this issue, a majority voting based approach is proposed to reliably detect and transform non-isolated outlier clusters into isolated clusters. Then, an expandable boundary criterion is employed to remove all the sparse outliers and the isolated outlier clusters, as well as to preserve critical geometric features such as sharp edges.

2. Existing methods

Many methods have been proposed in the literature to detect sparse outliers, isolated outlier clusters, and non-isolated outlier clusters. Sparse outliers are erroneous measurement points with low local point density. Based on the assumption that the local point density of a scanned point cloud follows the Gaussian distribution, Rusu et al. [3] proposed an efficient approach to detect sparse outliers, which correspond to low point densities. In practice, however, the scanned point cloud density for good surface points can be non-uniform and incomplete. For example, small subtle geometric features with color and reflectiveness variations are often sparsely sampled and would be falsely categorized as outliers. Sotoodeh [2] proposed a density-based algorithm, which also encounters the similar issue since point density is insufficient to distinguish sparse outliers from sparse good points. Weyrich et al. [4] applied three criteria to detect outliers in the neighborhood of each data point using user-specified parameters. For a point in a dense outlier cluster, its neighborhood does not contain sufficient good points and this makes this approach ineffective for outlier clusters.

For isolated outlier clusters, which have high point density and are relatively separated from the scanned surface, clustering techniques, such as spatial graphs [6] and region growing [5], have been adopted to classify the whole scanned point cloud into many clusters. Then, the small clusters are treated as outliers and removed. However, small good clusters due to insufficient sampling are removed as well. Moreover, a point cloud may have non-isolated outlier clusters that are not separable using a distance criterion. Therefore, clustering approaches will end up with some clusters where good points and outliers coexist. Recently, Wang et al. [7] utilized a distance-based deviation factor to detect sparse outliers and then detected small outlier clusters using region growing. Unfortunately, non-isolated outlier clusters were not considered.

To detect non-isolated outliers, Shen et al. [1] proposed an iterative surface propagation scheme based on the assumption that all outliers are randomly distributed so that they are hardly on the same plane. However, such an assumption is not always valid. The outliers caused by undesirable specular reflections are systematic and can form local planar patches with an offset from the scanned surface. As a result, these outliers will be mistakenly treated as regular points under the surface propagation scheme. Köhler et al. [8] developed a special structured light scanning system with an aim to detect and remove outliers during object scanning through additional sensory information. The involved data processing algorithms are thus only applicable to the particular scanning system.

Many point cloud smoothing algorithms can be used to reduce outliers. In such a context, outliers are simply treated as points with large noise. For example, common moving least-squares methods [9–11] can robustly project noisy points to the estimated surfaces in the presence of sparse outliers. However, such locally estimated surfaces can be biased towards outliers if the local neighborhood contains a large number of outliers from dense outlier clusters. Other smoothing techniques, such as Laplacian [12], mean curvature flow [13,14], bilateral filter [15,16], anisotropic diffusion [17], non-local denoising [18] and statistical methods [19,20] are all fundamentally based on the local neighborhood concept and may not work reliably among dense outlier clusters.

3. Non-isolated outlier detection

Detecting the non-isolated outlier clusters is challenging since these outliers are attached to the scanned surface and cannot be easily separated. Some non-isolated outlier clusters can even form smooth patches like good surface patches. To address this issue, a majority-voting scheme is proposed to cut the connections between non-isolated outlier clusters and the scanned surface so that the non-isolated outlier clusters become isolated.