euspen's 23rd International Conference & Exhibition, Copenhagen, DK, June 2023
www.euspen.eu

A new data fusion algorithm for point cloud registration

Zhongyi Michael Zhang, Sofia Catalucci, Adam Thompson, Samanta Piano

Manufacturing Metrology Team, Faculty of Engineering, University of Nottingham, UK

zhongyi.zhang@nottingham.ac.uk

Abstract

In recent years, numerous methods based on Gaussian process algorithms, weighted-least-square algorithms and machine learning algorithms have been proposed for three-dimensional point cloud registration. However, these algorithms have often been tested on point clouds of similar sizes and point densities, frequently sharing similar initial alignments and orientations. In this paper, we propose a new algorithmic pipeline for registering two point clouds of different sizes and point densities, that do not share initial alignment. This algorithm is used to register one dataset that is dense and small with one that is sparse and large; the former representing a region in the latter dataset. Our algorithm, firstly, sets the large point cloud as a reference and segments it into subsections ("sub-clouds") of the exact size of the small point cloud. Then, the algorithm compares the geometrical similarities between each sub-cloud and the small point cloud: both are further partitioned into layers along an arbitrary axis, with each layer again being partitioned into identical voxels. The number of points contained in each voxel is divided by the total number of points in each point cloud (i.e. converted to a percentage of points). If the percentages of points in a pair of corresponding voxels in both point clouds are similar, the pair is considered to be matched. Then, if > 90 % of the total number of voxels in this layer are matched, this pair of layers are labelled as matched. If > 90 % of the total number of layers are matched, this sub-cloud is regarded as successfully matched to the small point cloud. Finally, the small point cloud is registered to the location of the large point cloud in the reference coordinate system (the coordinate system of the big point cloud). Our algorithm has been tested with synthetic datasets, showing initial success. In future work, the pipeline will be tested on real measurement data.

Point cloud registration, metrology, data fusion, machine learning, geometric comparison

1. Introduction

Data fusion is a popular subject in metrology research, because of its advantages in improving measurement accuracy and expanding measurement coverage [1]. Existing data fusion methods in optical coordinate metrology (for example, point cloud registration) can be classified into three categories according to their mathematical principles: Gaussian processes algorithms [2], weighted least square algorithms [3] and machine learning algorithms [4,5]. These algorithms have shown success in fusion of point clouds of similar size, with initial coarse alignment and small numbers of points. However, the point cloud registration scenario where the two point clouds have disparity in size and arbitrary initial orientations has been rarely investigated. In this paper, we propose a new data fusion pipeline designed for this point cloud registration scenario.

2. Methodology

The proposed data fusion pipeline assumes that the point clouds are collected by two individual optical measurement systems with separate frames of reference, i.e. the two point clouds do not share any initial alignment. Additionally, it is assumed that the two point clouds are of different sizes, for example, one point cloud could represent the three-dimensional (3D) geometry of an engineered part (a 'large-and-sparse' point cloud), while the other point cloud could be the surface geometry of a tiny portion of that part (a 'small-and-dense' point cloud).
The simplest method to determine the correct location and orientation for registration is geometrical similarity comparison, i.e. employment of an algorithm to analyse the geometrical characteristics of every portion of the large-and-sparse point cloud, then the portion showing the most similar geometrical characteristics to the small-and-dense point cloud will be the area on which the small point cloud is registered. The complete registration pipeline has the steps shown in Figure 1.

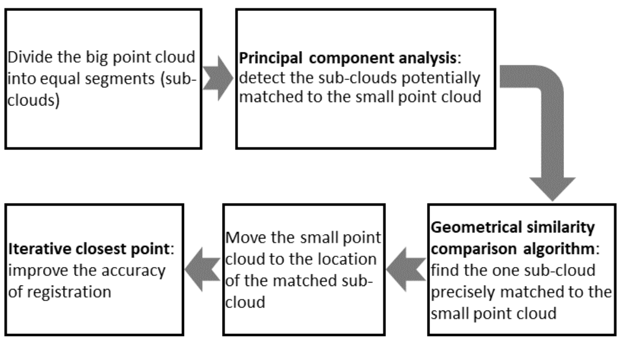

Figure 1 The point cloud registration pipeline.

2.1. Geometrical similarity comparison

We propose a method of comparing cross-sectional shapes along same axis in a global coordinate frame (denoted as the *z*-axis), to compare the geometrical similarity between two point clouds in 3D. To apply this comparison algorithm, the large-and-sparse point cloud is at first equally divided into multiple segments, which we call "sub-clouds". Each sub-cloud has the exact same size as the small-and-dense point cloud. The sub-cloud and the small-and-dense point cloud are then voxelised (i.e. decomposition into a regular grid of cubic cells in 3D space).

The percentage of points falling into each voxel relative to the total number of the points in each point cloud is calculated. Then, starting from the first layer of voxels along *z*-axis for both point clouds, the cross-sectional shapes are compared: if two corresponding voxels for two point clouds contain the same percentages of points, then these two voxels are considered as matched. Then, if in each layer of voxels, there is more than a certain number of matched voxels, this pair of layers is considered as a matched pair. Finally, if there is more than a certain number of matched voxels layers, the two point clouds are considered to be matched. The small point cloud will finally be registered onto the location of the matched sub-cloud in the frame of the big point cloud. This comparison algorithm is shown in Figure 2.

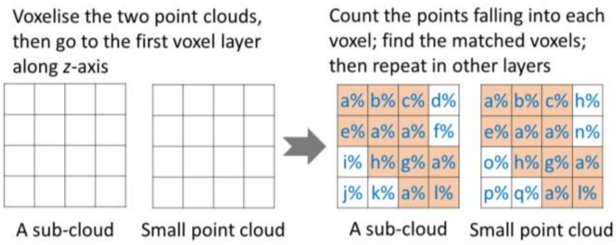

Figure 2 After voxeling the two point clouds, the algorithm calculates the percentage of points calling into each voxel, then counts the matched voxels in each layer along *z*-axis.

2.2. Sub-cloud pre-selection via principal component analysis

A problem of the proposed geometrical similarity comparison algorithm is the necessity of the pre-alignment along *z*-axis of the two datasets, leading to large computational costs. As such, we propose an additional step based on principal component analysis (PCA) to pre-select a number of sub-clouds before applying the comparison algorithm introduced in section 2.1.

The pre-selection algorithm finds the 1st and 2nd PCA axes of each point cloud at first; the two PCA axes form a PCA plane for each point cloud. Then, the histogram of the distances between the points and the PCA plane is plotted for each point cloud. If the histogram of a sub-cloud shows similar pattern to that of the small point cloud, this sub-cloud is considered as potentially matched with the small point cloud. With a number of candidates selected, the geometrical similarity comparison algorithm will determine which sub-cloud most matches the small point cloud. The pre-selection is shown in Figure 3.

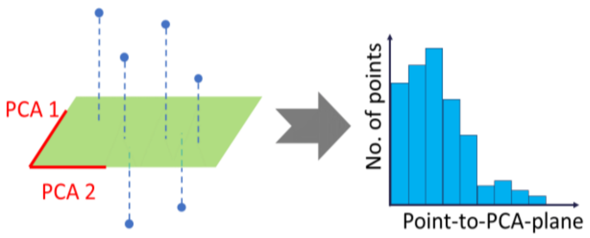

Figure 3 The histogram indicates the distribution of point-to-PCA-plane distances of a point cloud. Similar point clouds have similar histograms.

3. Tests and results

3.1. Registration of simple point clouds

Before applying the proposed pipeline to any point clouds representing engineered parts, all point clouds in the following tests were generated from 3D models with simple geometries. In the test shown in Figure 4, we used the pipeline to register the point cloud of a cylinder (red) into a large point cloud (green) containing this cylinder feature. In this test, we defined the matching threshold as follows: if > 90 % voxels in a layer along the *z*-axis of the sub-cloud are matched with the corresponding layer of the small point cloud, these two layers are determined to be matched; then if > 90 % voxel layers are matched, this sub-cloud is regarded matched with the small point cloud and the latter is registered into the location of the former in the coordinate frame of the big point cloud.

There is a mismatch between the small point cloud and the target area as shown in Figure 4(c), mainly due to the segmention of the big point cloud to reduce the number of sub-clouds being compared with the small point cloud, as restricted by the computing power. In this test, the target feature in the big point cloud is initially aligned along *z*-axis, so detecting the correct orientation of registration is easy and fast compared to industrial cases.

The point-to-point distances of the registration in this test are displayed in Figure 5, where the distances from the points in the red point cloud to the closest points in the big point cloud are represented by various colours. The statistics of the point-to-point distances in this test are visualised with the histogram in Figure 6. The accuracy of the registration decreases with an increase in the point-to-point distance of a specific point. According to these statistics, after the small point cloud is registered into the location of the matched sub-cloud, 24 % of the points have point-to-point distances < 0.33 ('small' distances), 60 % have point-to-point distances between 0.33 and 1.80 ('medium' distances) and 16 % have point-to-point distances larger than 1.80 ('large' distances). The maximum point-to-point distance is 2.30. This statistics are summarised in Table 1, and all units are arbitrary.

3.2. Fast detection of potentially similar point clouds via point-to-PCA-plane distances

As mentioned in section 2.2, we proposed the point-to-PCA-plane distance histogram to detect the sub-clouds potentially matched with the small point cloud when their orientations are unknown. The results of this test indicate that the distribution of the point-to-PCA-plane distances can be used to estimate the geometrical similarity between two point clouds, as seen in Figure 7.

The total number of points in all three test point clouds is 500. For point clouds with similar geometries but different orientations, the distributions of point-to-PCA-plane distances have similar patterns (Figure 7 (a) and (b)), i.e. the maximum bins show similar boundaries and heights; the boundaries of the top five bins are similar too. This initial result displays that the point-to-PCA-plane distance is applicable to detecting potentially matched sub-clouds without being disturbed by orientations of the point clouds.

Table 1 The percentages of points in the small point cloud have different point-to-point distances.

Point-to-point distances	Short (< 0.33)	Medium (0.33 – 1.80)	Long (> 1.80)
Number of points (% relative to the total number of points in the small point cloud)	24 %	60 %	16 %

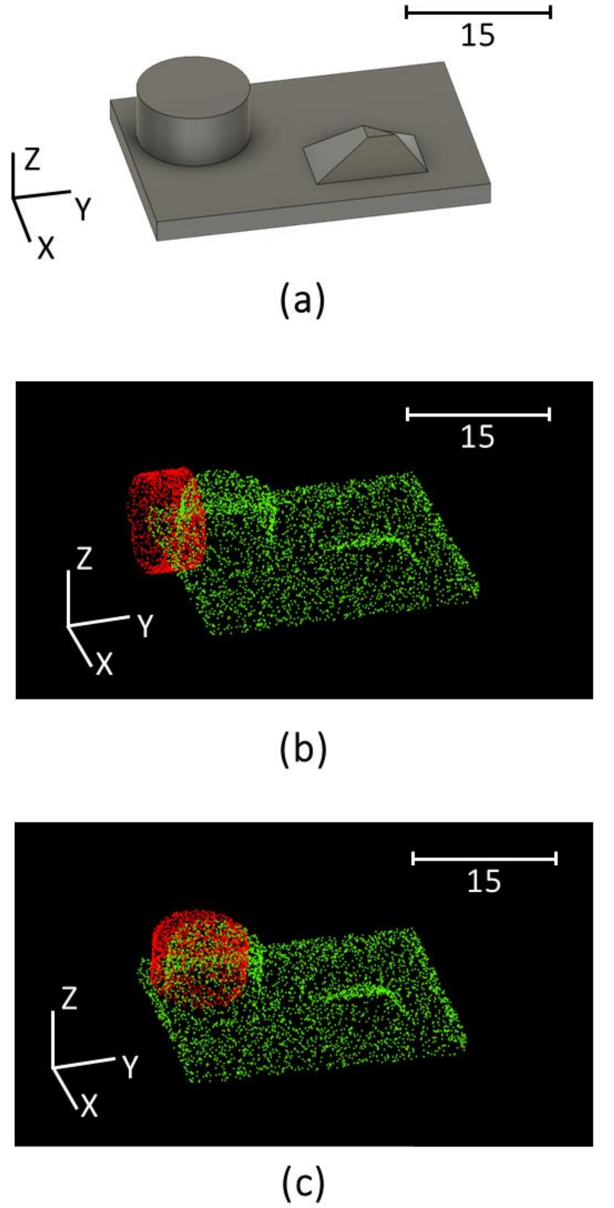

Figure 4 The registration of two point clouds having simple geometries. The red point cloud consists of 2,000 points and the green one consists of 4,000 points. (a) the CAD model for generating point clouds; (b) the red cylinder point cloud will be registered into the green point cloud; (c) the registration result given by our pipeline.

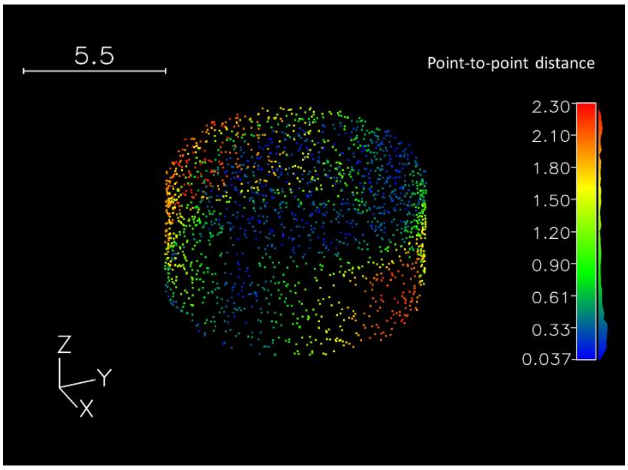

Figure 5 The distances from each point in the red point cloud to its nearest point in the matched area in the green point cloud.

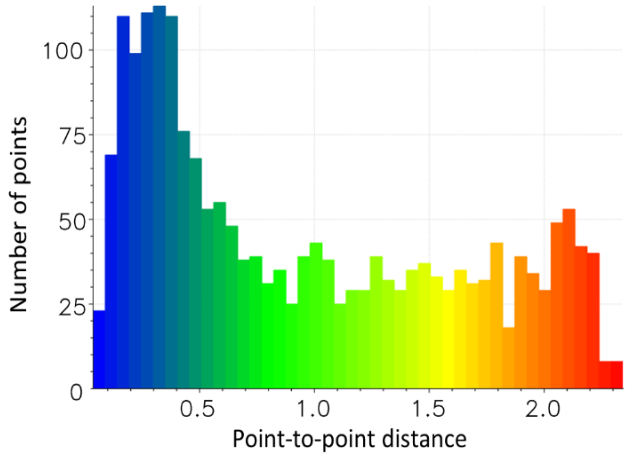

Figure 6 The histogram of the point-to-point distances. The colour scale is in correspondence to the one in Figure 5.

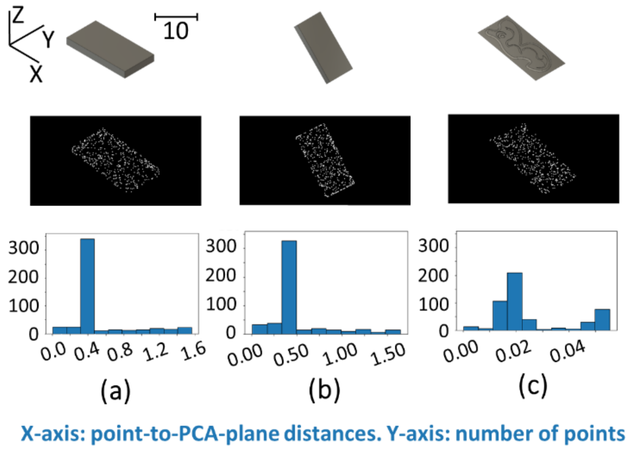

Figure 7 Top: CAD models used for generating the point clouds, shown in the middle row. Bottom: histograms of the point-to-PCA-plane distances. (a) benchmark; (b) object with same shape but different pose; and (c) object with same pose but different geometrical features on top.

4. Discussion

The test of the geometric similarity comparison algorithm with simple synthetic point clouds indicates that the algorithm can detect the location and orientation of one point cloud, relative to another, for registration. In terms of the accuracy of the registration using the similarity comparison algorithm, more than 80 % of the points in the small point cloud have point-to-point distances which are less than a quarter of the size of the small point cloud itself (the diameter of the cylinder model is 8 and its height is 4). The level of error can be further reduced by decreasing the step size of segmentation (creating more sub-clouds) and running the similarity comparison algorithm on a better computational hardware.

To reduce the number of sub-clouds that have to be processed by the similarity comparison algorithm, we proposed an initial selection stage based on the point-to-PCA-plane distances. Without aligning the orientation of the small point cloud with each sub-cloud to make a comparison of geometric similarity, all sub-clouds are scanned at this stage and the candidates, which are potentially similar to the small point clouds, are selected. Our test results indicate that the point-to-PCA-plane distance is an effective index for detecting geometrically similar, but misaligned, point clouds.

In the test shown in section 3.1, the small point cloud and the big point cloud are assumed to be initially aligned on the *x*-*y* plane, i.e. along the *z*-axis. In reality, however, the initial relative orientations of the small point cloud and the big point cloud can make both algorithms miss the target area (the area in the big point cloud where the small point cloud should be registered into). As described in section 2, the big point cloud is segmented with the dimensions of the small point cloud. As such, in the segmentation process, if the initial orientation of the target area is drastically different from the one of the small point clouds, the target area might be divided into multiple sub-clouds. Hence, none of these sub-clouds will show similar geometric characteristics in the similarity comparison process or the initial selection process with point-to-PCA-plane distances. This probable failure is visualised in Figure 8 and indicates that a boundary of application scenarios should be defined for using this data fusion pipeline.

Apart from the failure caused by initial alignments, noise in datasets can also reduce the quality of any point cloud registration pipelines. In the tests displayed in this research, all point clouds were generated from synthetic 3D models designed with CAD software. As such, the datasets used in this research have no noise or outliers. Examination of the robustness against noise of our pipeline represents an interesting area of research for further work on this topic.

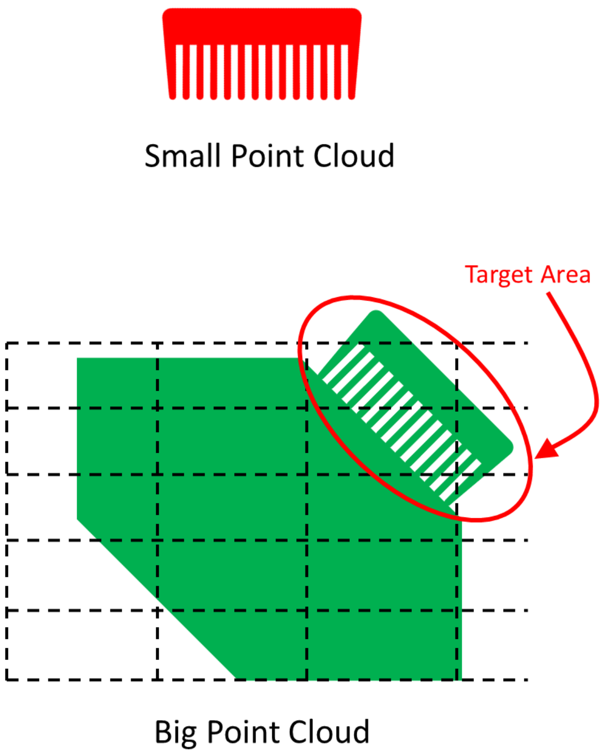

Figure 8 If the initial orientations of the small point cloud and the target area in the big point cloud are drastically different (as in the figure demonstrates), the segmentation, which is based on the dimensions of the small point cloud, can disassemble the target area into several different sub-clouds. As such, neither the similarity comparison algorithm or the initial selection based on point-to-PCA-plane distances can detect the target area.

5. Conclusion and future work

The presented tests indicate that the proposed point cloud registration pipeline shows good potential in registering 3D datasets with simple geometries. As for the next step, the combination of sub-cloud pre-selection with point-to-PCA-plane strategy into the registration pipeline will be proposed. To avoid missing the target area in the big point cloud due to the difference of initial orientations, the range of difference of initial orientations between the two point clouds has to be defined. In addition, the further application of the pipeline to point clouds representing real complex parts will be investigated.

Acknowledgements

The authors would like to thank the UKRI Research England Development (RED) Fund for funding this work via the Midlands Centre for Data-Driven Metrology.

References

[1] Llinas J, Hall D L 1998 Introduction to multi-sensor data fusion *Proc. - IEEE Int. Symp. Circuits Syst, Monterey, USA* 537–40
[2] Ren M J, Cheung C F, Xiao G B 2018 Gaussian process based bayesian inference system for intelligent surface measurement *Sensors* **18** 4069
[3] Forbes A B 2012 Weighting observations from multi-sensor coordinate measuring systems *Meas. Sci. Technol.* **23** 025004
[4] Eastwood J, Sims-Waterhouse D, Piano S 2020 *Machine learning approaches* In: Leach R K *Advances in Optical Form and Coordinate Metrology* (IOP Publishing) 6-1-6–20
[5] Zhang Z M, Catalucci S, Thompson A, Leach R, Piano S 2022 Applications of data fusion in optical coordinate metrology: a review *Int. J. Adv. Manuf. Technol.* **124** 1341–56