A Method of Position Precision Investigation Based on Large Pointcloud Data in Topography Survey

Dafu Zhang, Jing Zhang, Jili Yao, Keli Liu and Tongle Xu

ABSTRACT

Due to the disordered irregularity of the 3D laser scanning point cloud, the homonymy points of the checkpoint cannot be determined, so the repeated measurement comparison method and the comparison method of different mapping methods cannot be used for position precision investigation of 3D laser scanning topography survey. Therefore, a proposition in this paper is that the proximal-point method is used to detect the position precision of 3D laser scanning topography survey, which realizesthat the point of the same name is replaced by the point closest to the checkpoint in the dense point cloud, and calculate the position precision of the plane and elevation firstly; in addition, find the homonymy points from the super-large point cloud with more than 100 million points. Experiments show that the proposed method can detect the position precision of 3D laser scanning topography survey in ordinary computers.1

KEYWORDS

Large Pointcloud; Proximal-Point Method; Position Precision; Precision Detection.

1Dafu Zhang, Jili Yao, Keli Liu School of Civil Engineering and Architecture, Shandong University of Technology, Zibo, China
Jing Zhang, Tongle Xu School of Mechanical Engineering, Shandong University of Technology, Zibo, China
INTRODUCTION

Topography survey methods include classic topographic mapping with plane table, topographic mapping with transit and modern total station mapping, photogrammetry, RTK mapping, etc. [1-2]. With the development of science and technology, the advantage of laser scanners is that it can acquire high-precision three-dimensional coordinates of the object under test at high speed without contact, so it is widely used in surveying and mapping. Laser scanning and ground data for topography survey have also become one of the hot topics of many scholars. Some documents use the relationship of point cloud data to calculate the elevation accuracy of the topographic map (DEM). For example, when Andrej Kobler established DTM in laser scanning data, he calculated DTM elevation accuracy by using the fitting parameters in the iterative interpolation method of REIN (repetitive interpolation) [3-9]. In Document 10, a digital elevation model established by using a three-dimensional point cloud is proposed to perform quality inspection on its mathematical precision. There are also some literatures on how to improve the accuracy of point cloud conversion. For example, Nora Csanyi and Charles K. Toth use the ground-specific double concentric circular which controls target to improve the precision of the lidar point cloud data, and perform data calculation to make the accuracy of the laser radar-based data mapping to the centimeter level [10], George Vosselman proposed to improve and evaluate the plane measurement accuracy of point clouds by extracting and comparing the offset values of the overlapping region feature lines in point cloud data by aeronautical laser scanning [11]. In the topography survey using 3D laser scanning technology, the number of point cloud scanning points in a region after multi-station scanning and splicing may exceed 10 million. The existing software cannot directly find the same name of thousands of checkpoints. A method called proximal-point method is proposed in this paper, which is based on detecting three-dimensional accuracy of topography survey position of large point cloud.

THE TRADITIONAL TOPOGRAPHIC MAP POSITION PRECISION DETECTION METHOD

Position Precision Tolerance

The traditional topographical mathematical precision is measured by positional accuracy, while the positional accuracy includes horizontal accuracy and elevation accuracy. There are two methods for positional accuracy detection: one is repeated measurement comparison, and the other is the comparison method of different mapping methods. The accuracy requirements specification has clear provisions [10]. The repeated measurement comparison uses the same measurement method to perform two or more measurements, and compares its coordinates and elevation results, then calculates the position error and calculates the horizontal position errors.
and elevation error; the comparison method of different mapping methods uses
different measurement methods to measure the same target, and uses position
difference to calculate position precision, such as GPS and total station measurement
results comparison, leveling and trigonometric leveling comparison.

The specification stipulates that the point error on the map of the planimetric
point and the nearest field control point shall not be greater than that specified in
Table I. The error in elevation of elevation point with notes, contour and the nearest
field control point shall not be greater than that specified in Table II.

| Table I. THE TOLERANCE OF THE POINT ON THE MAP OF
| PLANIMETRIC POINT AND THE NEAREST FIELD CONTROL POINT. |
|-------------------------------------------------------|
| Measuring scale | Flatground,Hill | Mountain | Alpine |
|                 | country(mm)    | land(mm) |
| 1:500~1:2000    | 0.6            | 0.8      |

| Table II. THE TOLERANCE OF THE ERROR IN ELEVATION OF ELEVATION
| POINT WITH NOTES,CONTOUR AND THE NEAREST FIELD CONTROL POINT. |
|---------------------------------------------------------------|
| Flat ground        | Hillcountry        | Mountainland | Alpine land |
| (mm)               | (mm)               | (mm)         | (mm)        |
| 1:500 Note with point | 0.4                 | 0.4          | 0.5          | 0.7          |
| Contour line       | 0.5                 | 0.5          | 0.7          | 1.0          |
| 1:1000 Note with point | 0.5                 | 0.5          | 0.7          | 1.5          |

**Horizontal Accuracy Detection**

The traditional detection method is characterized in that the points on the map
can be found in the field, and the coordinates on the map can be in one-to-one
correspondence with the measured coordinates.

The homonymic checkpoints measured by the same method or different methods
should be evenly distributed in the middle and the edge of the work area, and the
number of checkpoints should be higher than 10%.

For the planimetric position of the checkpoint, the point error of checkpoint \( m_{ci} \)
should be calculated according to formula (1).

\[
m_{ci} = \pm \sqrt{\frac{\sum \Delta S_{ci}^2}{2n_c}}
\]  

(1)

Where \( \Delta S_{ci} \) is the difference in the planimetric position of the checkpoint and \( n_c \)
is the number of points in the checkpoint.
Elevation Accuracy Detection

The elevation accuracy check of the topographic map data still uses the same method or different methods to measure the elevation of the same-named point, and error in elevation of the checkpoint \( m_{ch} \) is calculated according to formula (2).

\[
m_{ch} = \pm \sqrt{\frac{\sum \Delta H_{RSi}^2}{2n_c}}
\]  

(2)

Where \( \Delta H_{RSi} \) is the difference in the elevation of the checkpoint and \( n_c \) is the number of checkpoints.

PROXIMAL-POINT TERRAIN ACCURACY DETECTION METHOD

In this paper, proximal-point terrain accuracy detection method is used for position precision detection of 3D laser scanning topographic points of large point clouds. With the advantage of the scanning speed of 3D laser scanners, when using the laser scanning technology for topographic measurement, the amount of data generated is often large. After a region is spliced, the number of point cloud scanning spot may exceed 100 million, for thousands of checkpoints, the existing software cannot find homonymy points corresponding to the checkpoint. The feature of the proximal-point method proposed in this paper is to dynamically search for the "homonymy points" and position accuracy detection for large point clouds. The homonymy point refers to a certain scanning spot in the point cloud corresponding to the GNSS RTK (or total station) measurement point.

The way to implement this method is to (1) divide the point cloud into several blocks, each point cloud does not exceed 5 million points; (2) determine the location of the homonymy points for the GNSS RTK measurement checkpoint (or total station measurement point), which point cloud it belongs to; (3) searching for the closest point of the checkpoint in the determined block as the homonymy points; (4) calculate the root mean square error of the planimetric position, point with notes and error in elevation of contour line. In order to find the closest point in the point cloud and speed up the processing speed of the computer without affecting the accuracy of the scanning topographic measurement position, this paper firstly dynamically partitions the point cloud data, and then judges the location which the GNSS RTK measurement checkpoint or the total station measurement point is, and then finds the nearest point from the partition, which greatly reduces the number of calculation points.
Point Cloud Data Processing

The number of point clouds is too large and dense, so directly reading may cause the computer to crash due to insufficient memory. In order to reduce the amount of data, the point cloud data is denoised and simplified firstly[12-13], after the software automatically denoises, there will be residual noise points, which needs to be manually deleted [14], then the data of the scanned topographic map is synthesized by multi-station scan data, and each station is an independent coordinate system, so the data of each station must be converted to a unified [15-16] in the coordinate system.

(1) Point cloud simplification

In order to make the massive data obtained meet the engineering needs, and the software can be used to manage and optimize the data well, the data needs to be simplified and compressed. The sampled data set obtained by laser scanning measurement is generally densely redundant and scattered, it will be difficult to construct a mesh point by point, the data set must be filtered and compressed. The general screening method is to set a minimum distance, and then compare the distance between adjacent two points in the direction of the scanning line, and the point smaller than the threshold point is deleted. You can also connect the first two points of the scan line, select the point closest to the two points from the middle point as the break point, and then repeat the above process between the first point and the middle point, between the middle point and the last point, until the distance sought is less than the predefined threshold, the point is aborted.

(2) Point cloud denoising

In the ground laser scanning, noise is generated mainly due to the reflection of the laser by the flying object or the large particle dust, and the strong reflection of the smooth surface of the ground. The denoising method is combined with automatic denoising by computer and manual denoising.

(3) Ground point cloud acquisition

The ground point cloud acquisition is also called point cloud filtering. The purpose is to delete the point cloud on the tree and the building, and obtain the point cloud data reflecting the ground shape, which can be obtained by the filtering method. The slope-based filtering method was originally proposed by Vosselman [17], the maximum slope of the two laser foot points around it is used as the processing criterion for extracting the ground point. The premise of this assumption is that the gradient of the slope of the ground point is very different from the gradient of the ground point (trees, houses, etc.). If the maximum slope value of the currently processed laser foot point and the neighborhood laser foot point is less than a predefined threshold, then this point is judged as a ground point, and vice versa. It can be seen that the slope threshold plays a crucial role in the accuracy of the filtering result. The smaller the threshold is set, the more laser foot points in the original point cloud data will be identified as ground points. This threshold setting can be derived from repeated experiments in the training data set. In general, slope-
based filtering methods are used to achieve better filtering effects in flat terrain areas. The accuracy of filtering in areas such as cliffs or steep slopes where terrain features change is often unsatisfactory. In order to break through this limitation, Sithole proposed an improved slope-based filtering method, which makes the slope threshold change with the topographical features, and the filtering accuracy is greatly improved, especially in the more complex terrain environment [13].

**Terrain Position Accuracy Detection Method**

(1) Planimetric position accuracy detection

Based on the measurement of the location accuracy of the super-large point cloud, it is impossible to find the same name of the checkpoint in the huge data like the traditional detection method, but because the point cloud data is very dense, it can be considered as the checkpoint or total station with GNSS RTK measurement. The closest point of the measurement point is the point of the same name. In this way, the central and edge of the work area selects the checkpoint, and the GNSS RTK measurement or total station measurement is used for repeated measurement. The number of checkpoints should be higher than 10%. For the planimetric position of the checkpoint, the point error of checkpoint \( \Delta s_m \) should be calculated according to formula (1). Where \( \Delta s_m \) is the difference of planimetric position between the GNSS RTK measurement checkpoint or the total station measurement point and the nearest point in the point cloud, \( n_c \) is the number of points in the checkpoint.

In the horizontal position errors, there is a terrain representative error (the GPS measurement point and the neighboring point in the point cloud are not strictly the same point), so the point cloud density should not be too small. In this paper, the point cloud resolution is 0.1m.

(2) Elevation accuracy detection

In the elevation accuracy check of the topographic map data, the error in elevation of the checkpoint \( \Delta h_m \) is still calculated by the GNSS RTK measurement checkpoint or the coordinates of the total station and the closest point in the point cloud according to formula (2), where \( \Delta H_{RSI} \) is the elevation difference between the total station/GNSS RTK measurement detection point elevation and the nearest point of the scanning point cloud, and \( n_c \) is the number of checkpoints. In the elevation comparison error, it also contains the terrain representative error, and sometimes there will be errors. For example, in the place where there is a steep ridge, there will be errors with the same name on the ridge and the ridge, so the difference is above 1m. In the statistical elevation accuracy, terrain representative errors do not participate in the calculation.

(3) Super large point cloud processing method
When calling a large point cloud, because the number of points is too large, directly reading will cause insufficient memory and crash, so the processing of the large point cloud must be read in blocks, and then subsequent operations. The number of the first partition area of the super large point cloud should not be too large, preferably no more than 100. The partition principle is shown in Table III.

If the number of points in each zone is still more than 5 million, the dynamic quadtree principle is used to divide into four smaller cells, as shown in Figure 1, until the number of points in each cell does not exceed 5 million.

Of course, any square partition can also be used, and the side length of each area is arbitrarily set to $a$, the X and Y coordinates of the lower left corner of the super large point cloud are calculated, and then the partition is performed in the direction in which X and Y are increased. The area number starts from the bottom left corner partition 1 and proceeds to the right upwards, as shown in Figure 2. After the partition is completed, for each point or total station measurement point measured by GNSS RTK, first determine the area where the point is located, and then find the nearest point of the point in the area. If the point is less than 5 cm from the edge of the area, you should find the closest point in this area and adjacent areas, which greatly reduces the amount of calculation. After finding the closest point, calculate the positional accuracy between the GNSS RTK measurement point and the nearest scanning point to generate a report file in EXCEL format.

| Serialnumber | scale   | zone length (m) |
|--------------|---------|-----------------|
| 1            | 1:5000  | 2500            |
| 2            | 1:2000  | 1000            |
| 3            | 1:1000  | 500             |

Figure 1. Dynamic quadtree.  
Figure 2. Segmentation at any scale.
EXPERIMENT AND ANALYSIS

Data Introduction

In the Qingdao Development Zone, a total of 49 stations were scanned with the STONEX X9 laser scanner, and 30.19 GB of point cloud data containing 983873940 points was obtained. The area is about 1,900 meters long from east to west and about 1,500 meters wide from north to south, about 3 square kilometers. There is a certain ups and downs in the terrain. After vegetative filtering and simplifying the point cloud data according to the spatial resolution $0.2m \times 0.2m \times 0.1m$, the remaining points are 285375632. In order to study the method of scanning topographic measurement position accuracy detection, 1198 GNSS RTK detection points were measured. In order to prevent errors, two groups were selected and measurements were taken separately, then compare the results, which eliminates the possibility of GNSS RTK measurement detection point errors.

Data Processing

IDL is a application development language of new generation with interactive, cross-platform and object-oriented feature. It is powerful. The data processing software in this paper is an EEXLT (Map Feature Extraction System) system independently developed based on the IDL (Interactive Data Language) programming language. Position accuracy detection is a function in this system.

The purpose of data processing is to find the closest point to the 1198 GNSS RTK detection points from 285,375,632 points, comparing the planimetric position and elevation of the two points and calculating the root mean square error.

The data processing process is to first read the super large point cloud in physical order, and read 5 million points per segment and partition the point cloud. The total number of points with less than 5 million points is 261, and then calculate the area which each GNSS RTK detection point locates in, and then search for the nearest point from the point cloud data in the area to the checkpoint as the homonymy points of the checkpoint, and finally calculate difference in the planimetric position of homonymy points, the planimetric point error, the elevation difference and error in elevation. The program runs on computers with different configurations, and there is a difference in the running time.

| Plane position accuracy error interval (m) | Number of points | Height accuracy error interval (m) | Number of points |
|--------------------------------------------|-----------------|----------------------------------|-----------------|
| 0.00~0.30                                  | 1001            | 0.00~0.20                        | 1062            |
| 0.30~0.40                                  | 192             | 0.20~0.25                        | 101             |
| Above 0.40                                 | 5               | 0.25~0.35                        | 24              |
|                                           |                 | Above 0.35                       | 11              |
Analysis of Results

For the 1:500 large-scale topographic map, as for the planimetric accuracy of the planimetric point relative to the point error of the nearest field control point, the maximum error in the flat or hilly land is 0.3 meters, and the maximum error in the mountain or mountain is 0.4 meters; as for elevation point with notes, contour line relative to error in elevation of the nearest field control point in the elevation accuracy, the maximum error in the flat land and hilly land is 0.2 meters, the maximum error in the mountain is 0.25 meters, and the maximum error in the mountainous area is 0.35 meters. The planimetric position accuracy interval statistics table and the elevation accuracy interval statistics table are listed in Table IV.

In this experimental plane, there were 5 points larger than 0.4 meters, and only 11 points with elevations greater than 0.35 meters. The reason is that after the point cloud data is simplified, the GNSS RTK measurement detection point does not coincide with the laser scanning point. It is also possible that the point is far away from the scanning station and related to the parameters set by the scanner during scanning. The entire point cloud data is partitioned by 1 meter resolution, and each partition only retains the lowest point, and the DEM is generated. These points are placed together with the GNSS RTK measurement detection points in the southern cass software for verification, which proves that the method is reliable.

CONCLUSIONS AND PROSPECTS

Through the experiment, the following conclusions can be drawn:

For large point clouds, if you use the point-by-point search method from 200 million scan points to find the detection points and the homonymy points of GNSS RTK measurements which number is 1198, you need to calculate 230 billion times for distance, and the computer memory cannot store 200 million points. At this time, the software automatic detection is a feasible method. The data processing process of this experiment proves that the algorithm for processing the segmentation reading and dynamic partitioning (point cloud block) storage of the large laser point cloud data is advanced, as long as the computer hard disk allow, no matter how big the point cloud data is, the computers of various configurations can process them, and the speed is faster, which provides technical guarantee for the realization of the proximal-point method.

In the topographic measurement of super-large point cloud which is based on laser scanning technology, for the topographic position accuracy detection, point cloud data and position difference of the homonymy points obtained by GNSS RTK ground repeat measurement can be used to reflect the measurement accuracy. There is no special requirement for the selection of repeated measurement points, and there is no need to mark in the field. The ground repeat measurement work is simple and convenient to implement, and the proximal-point method is practical and feasible.
The proximal-point method has representative error, and it is related to the point cloud density in data acquisition or the spatial resolution setting during data reading. The higher the density of acquisition is or the higher the spatial resolution setting is during reading, the more accurate the similarity of the proximal points is and it is much closer to the homonymy point, how to adopt the appropriate density and spatial resolution to achieve terrain position accuracy detection, further research is needed.

REFERENCES

1. Dongliang Wu, Yingjuan Shang, Dahua Gu, Dezhang Zhong. 2009. Research on the application of RTK system, total station and CASS in the mapping of terrain[J]. Science of Surveying and Mapping. 34(2):194-196.
2. Zhihao Wang, Ping Liu. 2011. Analysis of large-scale mapping experiment using UAV aero photographic system[J]. Bulletin of Surveying and Mapping. 7:18-20.
3. Xuefeng Guan, Huai Wu. 2010. Leveraging the power of multi-core platforms for large-scale geospatial data processing: Exemplified by generating DEM from massive LiDAR point clouds. Computers & Geosciences. 36 (2010) 1276–1282.
4. Giuseppe Casula, Paolo Mora et al. 2010. Detection of Terrain Morphologic Features Using GPS, TLS, and Land Surveys: “Tana della Volpe” Blind Valley Case Study[J]. Journal of Surveying Engineering. 136:132-138.
5. Andrej Kobler, Norbert Pfeifer, Peter O.et al. 2007. Repetitive interpolation: A robust algorithm for DTM generation from Aerial Laser Scanner Data in forested terrain. Remote Sensing of Environment. 108 (2007) 9-23.
6. Jinjun Xu, Minwei Zhang. 2007. Status and development of terrestrial laser scanner[J]. Science of Surveying and Mapping. 1:47-50.
7. Jinjun Xu, Minghui Yu, Yanbing Zheng. 2008. Overview of the application of Terrestrial 3D laser scanner[J]. Geotechnical Investigation & Surveying. 12:31-33.
8. Yongbo Wang, Yehua Sheng. 2011. Patrol inspection of EHV electronic power transmission line using 3D laser measuring technology[J]. Science of Surveying and Mapping. 36(5):60-63.
9. Zhengjun Liu, Jianguo Qian, Zheng peng Zhang, Jian Wang. 2006. Experiment on high resolution DTM acquisition by 3D terrestrial laser scanner[J]. Science of Surveying and Mapping. 31(4):72-74.
10. Nora Csanyi and Charles K. April 2007. Toth. Improvement of Lidar Data Accuracy Using Lidar-Specific Ground Targets. Photogrammetric Engineering & Remote Sensing. Vol. 73, No. 4, pp. 385-396.
11. George Vosselman. 2012. Automated planimetric quality control in high accuracy airborne laser scanning surveys. ISPRS Journal of Photogrammetry and Remote Sensing. Vol. 74, pp.90-100.
12. Ruwen Schnabel, Sebastian Moser, Reinhard Klein. Fast vector quantization for efficient rendering of compressed point-clouds. Computers & Graphics. Computers & Graphics. 32 (2008) 246-259.
13. George Sithole, George Vosselman. 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. ISPRS Journal of Photogrammetry & Remote Sensing. 59(2004)85-101.
14. Flood, M. 2001. LIDAR activities and research priorities in the commercial sector. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XXXIV (Pt. 3/W4), 3-8.
15. JiliYao. 2005. Sarc model of 3D coordinate transformation[J]. Geomatics and Information Science of Wuhan University. 30(9):825-827.
16. Jili Yao. 2006. Rigorous formula for direct calculating parameter in 3D transformation[J]. Science of Surveying and Mapping. 5:7-9.

17. George Vosselman, Patrice Kessels, Ben Gorte. 2005. The utilisation of airborne laser scanning for mapping[J]. International Journal of Applied Earth Observation and Geoinformation. 6:177-186.