

DATA PROCESSING METHOD AND DATA PROCESSING SYSTEM

BACKGROUND

Technical Field

The present disclosure relates to a data processing method and a data processing system.

5 Description of Related Art

The statements in this section merely provide background information related to the present disclosure and do not necessarily constitute prior art.

LiDAR bundle adjustment (LBA), used to simultaneously calibrate sensor pose and point cloud accuracy, is a fundamental task in photogrammetry and robotics. LBA is a necessary
10 pre-processing step to reduce point cloud errors and improve consistency. It plays a key role in subsequent applications using pre-mapped images, such as forest inventory, infrastructure maintenance, and robotic delivery.

LBA has been extensively studied over the past few decades and has been applied in traditional mobile mapping systems, including aerial laser scanning and vehicle-borne laser
15 scanning. With the rapid adoption of small, low-cost 3D sensing systems and the increasing demand for 3D mapping in complex environments without global navigation satellite system (GNSS) signals, existing LBA methods face limitations in stability, efficiency, and computational overhead, making them difficult to directly apply to low-cost platforms.

Therefore, the data processing method and data processing system designed for small-
20 scale sensing systems in large-scale complex scenarios is one of the existing problems that need to be solved urgently.

SUMMARY

In order to solve the above-mentioned problems, the present disclosure provides a stable, efficient, and scalable data processing method and data processing system.

25 The data processing method includes steps of: receiving a plurality of first scanned data and

a plurality of second scanned data corresponding to a preset environment, wherein the first scanned data correspond to a first position range in the preset environment, and the second scanned data correspond to a second position range in the preset environment; sparsifying the first scanned data and the second scanned data according to a relation graph to generate a plurality of first sparsified data and a plurality of second sparsified data; uniformly sampling the first sparsified data and the second sparsified data according to a first voxel size to generate a plurality of first sampled data and a plurality of second sampled data; generating a plurality of smoothed data according to the first sampled data and the second sampled data; dividing the smoothed data into a plurality of data clusters; optimizing relative relations of the data clusters and obtaining a state matrix; optimizing relative relations of a plurality of nodes in each data cluster according to the state matrix; generating point cloud map information according to the data clusters.

In some embodiments, “receiving a plurality of first scanned data and a plurality of second scanned data corresponding to a preset environment” includes steps of: obtaining the first scanned data of the preset environment in the first position range by a first LiDAR device; obtaining the second scanned data of the preset environment in the second position range by a second LiDAR device.

In some embodiments, “receiving a plurality of first scanned data and a plurality of second scanned data corresponding to a preset environment” includes steps of: obtaining the first scanned data of the preset environment in the first position range by a LiDAR device; receiving the second scanned data from a storage device.

In some embodiments, “generating point cloud map information according to the data clusters” includes steps of: determining whether the point cloud map information is complete; outputting the point cloud map information if the point cloud map information is complete.

In some embodiments, the data processing method further includes steps of: uniformly sampling the first sparsified data and the second sparsified data according to a second voxel size to generate a plurality of third sampled data and a plurality of fourth sampled data if the point cloud map information is not complete; generating the smoothed data according to the third sampled data and the fourth sampled data; wherein the second voxel size is less than the first voxel size.

In some embodiments, “determining whether the point cloud map information is complete” includes a step of: determining whether the point cloud map information is complete according to an absolute position error.

In some embodiments, “generating a plurality of smoothed data according to the first sampled data and the second sampled data” includes steps of: obtaining a transformation matrix according to an adjacent position relation of the first sampled data; transforming the second sampled data from a first coordinate system to a second coordinate system according to the transformation matrix; generating the smoothed data in the second coordinate system according to the first sampled data and the second sampled data.

In some embodiments, “dividing the smoothed data into a plurality of data clusters” includes steps of: dividing the smoothed data into a plurality of sub-clusters that are mutually independent; randomly merging the sub-clusters according to a modular gain and a probability distribution; stopping merging the sub-clusters and generating the data clusters when a preset condition is met.

In some embodiments, the preset condition includes that the number of the nodes in each data cluster reaches a preset number.

In some embodiments, “optimizing relative relations of the data clusters and obtaining a state matrix” includes steps of: selecting a central node from the nodes of each data cluster;

associating position relations of the central nodes of the data clusters to obtain the state matrix.

In some embodiments, “optimizing relative relations of a plurality of nodes in each data cluster according to the state matrix” includes a step of: associating position relations between the central node and other nodes of each data cluster.

5 In some embodiments, the relation graph includes a plurality of node information and a plurality of edge information connecting the node information; the step of “sparsifying the first scanned data and the second scanned data according to a relation graph” includes steps of: selecting the edge information according to a preset parameter, wherein the preset parameter includes a preset number and a preset ratio; sparsifying the first scanned data and the second
10 scanned data according to the edge information of the preset parameter.

The data processing system includes a data input unit, a first processing unit, a second processing unit, a third processing unit, a fourth processing unit, and a data output unit. The data input unit receives a plurality of first scanned data and a plurality of second scanned data corresponding to a preset environment, wherein the first scanned data correspond to a first
15 position range in the preset environment, and the second scanned data correspond to a second position range in the preset environment. The first processing unit sparsifies the first scanned data and the second scanned data according to a relation graph to generate a plurality of first sparsified data and a plurality of second sparsified data. The second processing unit uniformly samples the first sparsified data and the second sparsified data according to a first voxel size to
20 generate a plurality of first sampled data and a plurality of second sampled data, and generate a plurality of smoothed data according to the first sampled data and the second sampled data. The third processing unit divides the smoothed data into a plurality of data clusters. The fourth processing unit optimizes relative relations of the data clusters and obtain a state matrix, optimize relative relations of a plurality of nodes in each data cluster according to the state

matrix, and generate a point cloud map information according to the data clusters. The data output unit outputs the point cloud map information.

In some embodiments, the data processing system further includes a fifth processing unit. The fifth processing unit determines whether the point cloud map information is complete, and
5 if the point cloud map information is complete, the data output unit outputs the point cloud map information.

In some embodiments, if the point cloud map information is not complete, the second processing unit uniformly samples the first sparsified data and the second sparsified data according to a second voxel size to generate a plurality of third sampled data and a plurality of
10 fourth sampled data, and generates the smoothed data according to the third sampled data and the fourth sampled data, wherein the second voxel size is less than the first voxel size.

In some embodiments, the fifth processing unit determines whether the point cloud map information is complete according to an absolute position error.

In some embodiments, the third processing unit divides the smoothed data into a plurality
15 of sub-clusters that are mutually independent, randomly merges the sub-clusters according to a modular gain and a probability distribution, and stops merging the sub-clusters and generates the data clusters when a preset condition is met, wherein the preset condition includes that the number of the nodes in each data cluster reaches a preset number.

In some embodiments, the fourth processing unit selects a central node from the nodes of
20 each data cluster, associates position relations of the central nodes of the data clusters to obtain the state matrix, and associates position relations between the central node and other nodes of each data cluster.

In some embodiments, the first processing unit, the second processing unit, the third processing unit, and the fourth processing unit are integrated into a processor.

In some embodiments, the relation graph includes a plurality of node information and a plurality of edge information connecting the node information; the first processing unit selects the edge information according to a preset parameter, and sparsifies the first scanned data and the second scanned data according to the edge information of the preset parameter, wherein the
5 preset parameter includes a preset number and a preset ratio.

As described above, the data processing method of this embodiment processes multiple scanned data obtained in different ranges under a predetermined environment (e.g., scanned data obtained by different LiDAR devices or stored scanned data). The method first sparsifies the scanned data according to a relation graph and then samples the sparsified data at a
10 predetermined voxel size for smoothing. Afterward, the smoothed data is divided into multiple data clusters. After optimizing these data clusters, a state matrix is obtained. According to the state matrix, the relative relations between the multiple nodes in each data cluster are further optimized to generate point cloud map information. Therefore, the data processing method of this embodiment efficiently and accurately processes scanned data (e.g.,
15 point cloud data) through sparsifying, then increases the accuracy of the sparsified data through smoothing. Afterward, the data is divided into multiple data clusters, and the state matrices of these data clusters are obtained for clustered data processing. This reduces the burden on system computing performance and expands the scale of data that can be processed.

It is to be understood that both the foregoing general description and the following detailed
20 description are exemplary, and are intended to provide further explanation of the present disclosure as claimed. Other advantages and features of the present disclosure will be apparent from the following description, drawings, and claims.

BRIEF DESCRIPTION OF DRAWINGS

The present disclosure can be more fully understood by reading the following detailed

description of the embodiment, with reference made to the accompanying drawing as follows:

FIG. 1 is a flowchart of a data processing method according to a first embodiment of the present disclosure.

FIG. 2A is a schematic diagram of scanned data according to the present disclosure.

5 FIG. 2B is a schematic diagram of sparsified data according to the present disclosure.

FIG. 3A is a schematic diagram of sampling data according to the present disclosure.

FIG. 3B is a schematic diagram of smoothed data according to the present disclosure.

FIG. 4A is a schematic diagram of smoothed data according to the present disclosure.

FIG. 4B is a schematic diagram of a data cluster according to the present disclosure.

10 FIG. 5A and FIG. 5B are schematic diagrams of data clusters and state matrices according to the present disclosure.

FIG. 6A and FIG. 6B are flowcharts of the data processing method according to a second embodiment of the present disclosure.

15 FIG. 7 is a block diagram of a data processing system according to a first embodiment of the present disclosure.

FIG. 8 is a block diagram of the data processing system according to a second embodiment of the present disclosure.

FIG. 9 is a schematic diagram of the data processing system according to the present disclosure.

20 DETAILED DESCRIPTION

Reference will now be made to the drawing figures to describe the present disclosure in detail. It will be understood that the drawing figures and exemplified embodiments of present disclosure are not limited to the details thereof.

25 As used herein, terms such as “first”, “second”, “third”, “fourth”, and “fifth” describe various elements, components, regions, layers, and/or portions, which should not be limited by these terms. These terms may only be used to distinguish one element, component, region,

layer, or portion from another. Terms such as “first”, “second”, “third”, “fourth”, and “fifth” as used herein do not imply a sequence or order unless the context clearly indicates otherwise.

Please refer to FIG. 1, which shows a flowchart of a data processing method according to a first embodiment of the present disclosure. As shown in FIG. 1, the data processing method disclosed herein includes steps S01 to S08. The step S01: receiving a plurality of first scanned data and a plurality of second scanned data corresponding to a preset environment, wherein the first scanned data correspond to a first position range in the preset environment, and the second scanned data correspond to a second position range in the preset environment. The step S02: sparsifying the first scanned data and the second scanned data according to a relation graph to generate a plurality of first sparsified data and a plurality of second sparsified data. The step S03: uniformly sampling the first sparsified data and the second sparsified data according to a first voxel size to generate a plurality of first sampled data and a plurality of second sampled data. The step S04: generating a plurality of smoothed data according to the first sampled data and the second sampled data. The step S05: dividing the smoothed data into a plurality of data clusters. The step S06: optimizing relative relations of the data clusters and obtaining a state matrix. The step S07: optimizing relative relations of a plurality of nodes in each data cluster according to the state matrix. The step S08: generating a point cloud map information according to the data clusters.

Please refer to FIG. 2A, which shows a schematic diagram of scanned data according to the present disclosure. As shown in FIG. 1A and FIG. 2A, in the step S01, a plurality of first scanned data S1 and a plurality of second scanned data S2 corresponding to a preset environment can be received. The preset environment may be, for example, an airport, a factory, building, an urban environment, a natural environment, or other environments, which are not limited here. The first scanned data S1 correspond to a first position range R1 in the preset environment, and the second scanned data S2 correspond to a second position range R2 in the preset environment. The first position range R1 and the second position range R2 may

overlap completely or partially, that is, the first position range R1 and the second position range R2 may overlap at least part of the area in the preset environment. Therefore, a relation graph between the first scanned data S1 and the second scanned data S2 may be established. In addition, the scanned data may be, for example, point cloud data.

5 In some embodiments, “receiving a plurality of first scanned data S1 and a plurality of second scanned data S2 corresponding to a preset environment” includes steps of: obtaining the first scanned data S1 of the preset environment in the first position range R1 by a first LiDAR device, and obtaining the second scanned data S2 of the preset environment in the second position range R2 by a second LiDAR device. In other words, The first scanned data S1 and
10 the second scanned data S2 may be scanned data obtained at the same or different time points through the same or different LiDAR devices (e.g., drones, robots, wearable equipment, etc.). Therefore, the first scanned data S1 and the second scanned data S2 may have no sequence relationship in time, but may be scanned data of different ranges or the same range obtained by different means in the same environment.

15 In another embodiment, “receiving a plurality of first scanned data S1 and a plurality of second scanned data S2 corresponding to a preset environment” includes steps of: obtaining the first scanned data S1 of the preset environment in the first position range R1 by a LiDAR device, and receiving the second scanned data S2 from a storage device. In other words, the first scanned data S1 may be real-time data obtained through the LiDAR device, while the second
20 scanned data S2 is data stored in the storage device. Similarly, the first scanned data S1 and the second scanned data S2 may have no sequential relationship in time, but may be scanned data of different ranges or the same range obtained by different means under the same environment.

Please refer to FIG. 2B, which shows a schematic diagram of sparsified data according to
25 the present disclosure. As shown in FIG. 1 and FIG. 2, in the step S02, the first scanned data and the second scanned data are sparsified according to the relation graph to generate a plurality

of first sparsified data P1 and a plurality of second sparsified data P2. The relation graph is established according to the first scanned data and the second scanned data, and is usually expressed as $G=(X,E)$, where G is the relation graph, X is the node set, and E is the edge combination. That is, the relation graph includes multiple node information and multiple edge information connecting the node information. Therefore, for example, the relation graph may be optimized by numerical calculation to reduce the node information and edge information, and then the first scanned data and the second scanned data are sparsified to generate the first sparsified data P1 and the second sparsified data P2. In particular, “optimization” referred to in the present disclosure means finding a set of variables by numerical methods to make a function reach the minimum (or maximum) value, that is, to obtain the optimized numerical result.

For example, in some embodiments, “sparsifying the first scanned data and the second scanned data according to a relation graph” includes steps of: selecting the edge information according to a preset parameter, wherein the preset parameter includes a preset number and a preset ratio, and sparsifying the first scanned data and the second scanned data according to the edge information of the preset parameter. Specifically, for example, the relation diagram may be optimized by optimizing the following equation (1).

$$\begin{aligned} \operatorname{argmax} F_{set}(\mathbf{E}^o) \text{ s.t. } \mathbf{E}^o \subset \mathbf{E}, |\mathbf{E}^o| \leqslant \mathbf{N}^o, \\ F_{set}(\circ) = \lambda_{min}(\mathbf{A}(\circ)), \end{aligned} \quad \dots \text{ equation (1)}$$

$\mathbf{N}^0$ is a preset parameter, which may include a preset number and a preset ratio. $\mathbf{E}^0$ is the selected edge information. For example, $\mathbf{N}^0$ may be set to 20% of the edge information, thereby obtaining the optimized $G^0=(X,E^0)$, and then obtaining the representative first sparsified data P1 and second sparsified data P2.

Please refer to FIG. 3A, which shows a schematic diagram of sampling data according to the present disclosure. As shown in FIG. 1 and FIG. 3A, in the step S03, the first sparsified data and the second sparsified data are uniformly sampled according to a first voxel size to generate a plurality of first sampled data and a plurality of second sampled data. For example,

the first sparsified data and the second sparsified data are sampled according to the first voxel size (for example, set to γ), and the first sampled data $\hat{p}_i^W$ and the second sampled data $\hat{p}_j^W$ can be obtained, where $\{\hat{p}_i^W, i \in \Psi\}$ and $\{\hat{p}_j^W, j \in \Phi_i\}$. The second sampled data $\hat{p}_j^W$ is a node within the first voxel size γ . In FIG. 3A, the first sampled data $\hat{p}_i^W$ and the second sampled data $\hat{p}_j^W$ are, for example, represented under a first coordinate system F^W , and the first coordinate system F^W is, for example, a global reference coordinate system. In particular, the first sampled data $\hat{p}_i^W$ and the second sampled data $\hat{p}_j^W$ are derived from the first scanned data and the second scanned data, respectively, and therefore the first sampled data $\hat{p}_i^W$ and the second sampled data $\hat{p}_j^W$ also represent data of different or same ranges obtained by different means under the same environment.

Please refer to FIG. 3B, which shows a schematic diagram of smoothed data according to the present disclosure. As shown in FIG. 1 and FIG. 3B, in the step S04, a plurality of smoothed data $\hat{p}_j^S$ are generated according to the first sampled data and the second sampled data. Specifically, the first sampled data may be used as a smoothing kernel, and the second sampled data may be projected into the tangent space of the smoothing kernel to generate smoothed data $\hat{p}_j^S$.

In some embodiments, “generating a plurality of smoothed data $\hat{p}_j^S$ according to the first sampled data and the second sampled data” includes steps of: obtaining a transformation matrix according to an adjacent position relation of the first sampled data; transforming the second sampled data from a first coordinate system to a second coordinate system according to the transformation matrix; generating the smoothed data in the second coordinate system according to the first sampled data and the second sampled data. The method of obtaining the transformation matrix is described as follows: first, for example, by the following equation (2), a normal vector $\hat{n}_i$ of the first sample data as the smoothing kernel is obtained.

$$\begin{aligned} \underset{\hat{\mathbf{n}}_i, |\hat{\mathbf{n}}_i|=1}{\operatorname{argmin}} \mathbf{G}(\hat{\mathbf{n}}_i) &= 1 - \hat{\mathbf{n}}_i^\top \check{\mathbf{n}}_i + \mu |\mathbf{D}(\hat{\mathbf{n}}_i)|_0, \\ \mathbf{D}(\hat{\mathbf{n}}_i)_j &= 1 - \hat{\mathbf{n}}_i^\top \check{\mathbf{n}}_j, \end{aligned} \quad \dots \text{ equation (2)}$$

When the normal vector $\hat{\mathbf{n}}_i$ is obtained, three coordinate axes of the second coordinate system F^S of the tangent space of the smoothing kernel may be established by the following equation (3). The second coordinate system F^S is, for example, the kernel space reference coordinate system.

$$\hat{\mathbf{n}}_2^i = \hat{\mathbf{n}}_i, \hat{\mathbf{n}}_1^i = [\hat{n}_{i,1}, -\hat{n}_{i,0}, 0]^\top, \hat{\mathbf{n}}_0^i = \hat{\mathbf{n}}_1^i \times \hat{\mathbf{n}}_i \quad \dots \text{ equation (3)}$$

The transformation matrix may be expressed as $\hat{\mathbf{M}}_i = [\hat{\mathbf{n}}_0^i, \hat{\mathbf{n}}_1^i, \hat{\mathbf{n}}_2^i]^\top$. After obtaining the transformation matrix by using the first sampled data, the second sampled data can be transformed from the first coordinate system to the second coordinate system F^S by using the transformation matrix, thereby generating smoothed data $\hat{p}_j^S$ in the second coordinate system F^S . Specifically, it can be expressed as equation (4).

$$[\hat{x}_j^S, \hat{y}_j^S, \hat{z}_j^S]^\top = \hat{\mathbf{p}}_j^S = \hat{\mathbf{M}}_i (\hat{\mathbf{p}}_j^W - \hat{\mathbf{p}}_i^W) \quad \dots \text{ equation (4)}$$

Furthermore, the smoothed data $\hat{p}_j^S$ in the second coordinate system F^S may be expressed by using equation (5).

$$f_i(\hat{x}_j^S, \hat{y}_j^S) = \boldsymbol{\alpha}^i{}^\top \left[(\hat{x}_j^S)^2, (\hat{y}_j^S)^2, \hat{x}_j^S \hat{y}_j^S, \hat{x}_j^S, \hat{y}_j^S \right]^\top \quad \dots \text{ equation (5)}$$

$\boldsymbol{\alpha}^i$ is a parameter of the first sample data $\hat{p}_i^W$ as the smoothing kernel, for example, $\boldsymbol{\alpha}^i = [\alpha_0^i, \alpha_1^i, \dots, \alpha_5^i]^\top$. Furthermore, the smoothed data $\hat{p}_j^S$ can be used to obtain a fitted polynomial surface S by using equation (6).

$$\begin{aligned} \underset{\boldsymbol{\alpha}^i}{\operatorname{argmin}} \sum_{j \in \Phi_i} \left| (f_i(\hat{x}_j^S, \hat{y}_j^S) - \hat{z}_j^S) w(|\hat{\mathbf{p}}_j^S|) \right|^2 \\ w(d) = e^{-d^2/\gamma^2}. \end{aligned} \quad \dots \text{ equation (6)}$$

Please refer to FIG. 4A, which shows a schematic diagram of smoothed data according to

the present disclosure. As shown in FIG. 4A, after processing all smoothed data $\hat{p}_j^S$, a fitted polynomial surface S for the preset environment can be obtained. Please refer to FIG. 4B, which shows a schematic diagram of a data cluster according to the present disclosure. As shown in FIG. 1 and FIG. 4B, in the step S05, the smoothed data is divided into multiple data
5 clusters C. Therefore, by further dividing the nodes in the smoothed data into small, independent multiple data clusters C and ensuring that the nodes within each data cluster C are closely related, data can be processed more efficiently.

In some embodiments, “dividing the smoothed data into a plurality of data clusters C” includes steps of: dividing the smoothed data into a plurality of sub-clusters that are mutually
10 independent; randomly merging the sub-clusters according to a modular gain and a probability distribution; stopping merging the sub-clusters and generating the data clusters when a preset condition is met. Specifically, the weight of the edge information in the smoothed data can be measured, for example, by calculating the minimum eigenvalue according to the relation graph. Initially, each node in the smoothed data can be treated as a single sub-cluster, and then the sub-
15 clusters are randomly merged according to the modular gain and the probability distribution. The modular gain Q may be expressed, for example, by equation (7), and by maximizing Q, the optimal cluster can be obtained.

$$Q = \frac{1}{2s} \sum_{\mathbf{E}_k \in \mathbf{E}} \delta(\mathbf{E}_k^0, \mathbf{E}_k^1) \left(\lambda_k^{\min} - \frac{w(\mathbf{E}_k^0)w(\mathbf{E}_k^1)}{2s} \right) \dots \text{equation (7)}$$

where, $s = \sum_{\mathbf{E}_k \in \mathbf{E}} \lambda_k^{\min}$, $w(E_k^0)$ is the total weight of the edges connected to node E_k^0 ,
20 $w(E_k^1)$ is the total weight of the edges connected to node E_k^1 . When nodes E_k^0 and E_k^1 belong to the same cluster, $\delta(\mathbf{E}_k^0, \mathbf{E}_k^1)=1$; otherwise, $\delta(\mathbf{E}_k^0, \mathbf{E}_k^1)=0$.

Furthermore, when a preset condition is met, the merging of the sub-clusters may be stopped and the data cluster C is generated. The preset condition may include, but is not limited to, when the number of nodes in each data cluster reaches a preset number. For

example, when the number of nodes in each cluster reaches 300, the merging of the sub-clusters may be stopped and the data cluster C can be generated. In particular, the preset number may be set differently depending on the requirements.

Please refer to FIG. 5A and FIG. 5B, which show schematic diagrams of data clusters and
5 state matrices according to the present disclosure. As shown in FIG. 5A and FIG. 1, in the step S06, the relative relations of the data clusters C are optimized to obtain the state matrix M. Specifically, each data cluster C may be considered a rigid body, and therefore by optimizing the relative relations of the data clusters C, the state matrix M can be obtained.

In some embodiments, “optimizing relative relations of the data clusters C and obtaining
10 a state matrix M” includes steps of: selecting a central node CN from the nodes N of each data cluster C; associating position relations of the central nodes CN of the data clusters C to obtain the state matrix M. Therefore, each data cluster C can be considered a rigid body. By associating the positions of the central nodes CN of each data cluster C, the state matrix M can be obtained. In other words, by converting the three-dimensional data cluster C into the two-
15 dimensional state matrix M, the relative position relations are preserved and computational complexity is reduced.

As shown in FIG. 5B and FIG. 1, in the step S07, the relative relations of the plurality of nodes N of each data cluster C is optimized according to the state matrix M. Specifically, after optimizing the position relations of each data cluster C, the relative relations of the plurality of
20 nodes N of each data cluster C may be further optimized according to the state matrix M.

In some embodiments, “optimizing relative relations of a plurality of nodes N in each data cluster C according to the state matrix M” includes a step of: associating position relations between the central node CN and other nodes C of each data cluster C. Therefore, since each data cluster C may be considered a rigid body, the relative relations between these data clusters
25 C can be calculated via the central node CN. By associating the position relations between the central node CN and other nodes C in each data cluster C, the relative relations between all

nodes C can be further calculated. Finally, in the step S08, the point cloud map information is generated according to these data clusters. Therefore, after calculating the relative relations between all nodes in these data clusters, the point cloud map information can be generated.

As described above, the data processing method of this embodiment processes multiple
5 scanned data obtained in different ranges under a predetermined environment (e.g., scanned data obtained by different LiDAR devices or stored scanned data). The method first sparsifies the scanned data according to a relation graph and then samples the sparsified data at a predetermined voxel size for smoothing. Afterward, the smoothed data is divided into multiple data clusters. After optimizing these data clusters, a state matrix is obtained.
10 According to the state matrix, the relative relations between the multiple nodes in each data cluster are further optimized to generate point cloud map information. Therefore, the data processing method of this embodiment efficiently and accurately processes scanned data (e.g., point cloud data) through sparsifying, then increases the accuracy of the sparsified data through smoothing. Afterward, the data is divided into multiple data clusters, and the state matrices
15 of these data clusters are obtained for clustered data processing. This reduces the burden on system computing performance and expands the scale of data that can be processed.

Please refer to FIG. 6A and FIG. 6B, which show flowcharts of the data processing method according to a second embodiment of the present disclosure. As shown in FIG. 6A and FIG. 6B, the difference between the data processing method of the second embodiment of the present
20 disclosure and the data processing method of the first embodiment is that the step S08 includes steps S081 to S084. The step S081: determining whether the point cloud map information is complete. If the point cloud map information is complete, the step S082: outputting the point cloud map information. If the point cloud map information is not complete, the step S083: uniformly sampling the first sparsified data and the second sparsified data according to a second
25 voxel size to generate a plurality of third sampled data and a plurality of fourth sampled data, wherein the second voxel size is less than the first voxel size. The step S084: generating the

smoothed data according to the third sampled data and the fourth sampled data.

In some embodiments, the completeness of the point cloud map information can be further determined. In some embodiments, “determining whether the point cloud map information is complete” includes a step of: determining whether the point cloud map information is complete
5 according to an absolute position error. The absolute position error (APE) is the difference between each estimated node and its corresponding real node. Therefore, for example, the APE can be used to determine the completeness of the point cloud map information generated by the present disclosure. If the determination result is “Yes”, the point cloud map information can be directly output. In particular, the determination criteria may be, for example, that the
10 point cloud map information is considered complete when the APE is within a predetermined range. Of course, the APE may have different predetermined ranges for different preset environments.

On the other hand, if the determination result is “No”, the first sparsified data and the second sparsified data are re-sampled according to a second voxel size to generate a plurality
15 of third sampled data and a plurality of fourth sampled data. In particular, the second voxel size is less than the first voxel size. For example, if the first voxel size is set to γ , the second voxel size may be set to γ/T_D . In other words, by reducing the voxel size, the size of the sampled data can be reduced, thereby reducing the smoothed data (i.e., the smoothing kernel) generated from the sampled data, thereby obtaining more accurate point cloud map information.

Please refer to FIG. 7, which shows a block diagram of a data processing system according
20 to a first embodiment of the present disclosure. As shown in FIG. 7, the data processing system 10 according to the first embodiment of the present disclosure includes a data input unit 101, a first processing unit 102, a second processing unit 103, a third processing unit 104, a fourth processing unit 105, and a data output unit 106. The data input unit 101 receives a
25 plurality of first scanned data S1 and a plurality of second scanned data S2 corresponding to a preset environment. In particular, the first scanned data S1 correspond to a first position range

in the preset environment, and the second scanned data S2 correspond to a second position range in the preset environment. The first processing unit 102 sparsifies the first scanned data S1 and the second scanned data S2 according to a relation graph to generate a plurality of first sparsified data P1 and a plurality of second sparsified data P2. The second processing unit 103 uniformly samples the first sparsified data P1 and the second sparsified data P2 according to a first voxel size to generate a plurality of first sampled data $\hat{p}_i^W$ and a plurality of second sampled data $\hat{p}_j^W$, and generate a plurality of smoothed data $\hat{p}_j^S$ according to the first sampled data $\hat{p}_i^W$ and the second sampled data $\hat{p}_j^W$. The third processing unit 104 divides the smoothed data $\hat{p}_j^S$ into a plurality of data clusters C. The fourth processing unit 105 optimizes relative relations of the data clusters C and obtain a state matrix M, optimizes relative relations of a plurality of nodes in each data cluster C according to the state matrix M, and generates a point cloud map information PM according to the data clusters C. The data output unit 106 outputs the point cloud map information PM.

In some embodiments, the relation graph includes a plurality of node information and a plurality of edge information connecting the node information. The first processing unit 102 selects the edge information according to a preset parameter, and sparsifies the first scanned data S1 and the second scanned data S2 according to the edge information of the preset parameter.

In some embodiments, the third processing unit 104 divides the smoothed data $\hat{p}_j^S$ into a plurality of sub-clusters that are mutually independent, randomly merge the sub-clusters according to a modular gain and a probability distribution, and stop merging the sub-clusters and generate the data clusters C when a preset condition is met. In particular, the preset condition comprises that the number of the nodes in each data cluster C reaches a preset number.

In some embodiments, the fourth processing unit 105 selects a central node from the nodes of each data cluster C, associates position relations of the central nodes of the data clusters C to obtain the state matrix M, and associates position relations between the central node and

other nodes of each data cluster C.

The specific implementation of the data processing system 10 has been described in detail in the above embodiments and will not be repeated here.

Please refer to FIG. 8, which shows a block diagram of the data processing system according to a second embodiment of the present disclosure. As shown in FIG. 8, the data processing system 10A of the second embodiment of the present disclosure differs from the data processing system 10 of the first embodiment in that the data processing system 10A further includes a fifth processing unit 107. The fifth processing unit 107 determines whether the point cloud map information PM is complete. If the point cloud map information PM is complete as determined by the fifth processing unit 107, the data output unit 106 outputs the point cloud map information PM. If the point cloud map information PM is not complete as determined by the fifth processing unit 107, the second processing unit 103 uniformly samples the first sparsified data P1 and the second sparsified data P2 according to a second voxel size to generate a plurality of third sampled data $\hat{p}_i^W$ and a plurality of fourth sampled data $\hat{p}_j^W$, and generates the smoothed data $\hat{p}_j^S$ according to the third sampled data $\hat{p}_i^W$ and the fourth sampled data $\hat{p}_j^W$. In particular, the second voxel size is less than the first voxel size. In some embodiments, the fifth processing unit 107 determines whether the point cloud map information PM is complete.

The specific implementation of the data processing system 10A has been described in detail in the above embodiment and will not be repeated here.

Please refer to FIG. 9, which shows a schematic diagram of the data processing system according to the present disclosure. The data processing system 1 is, for example, a computing device including a processor 1005 and a non-transitory, computer-readable recording medium 1001. The recording medium 1001 stores a program 1002 (i.e., a set of executable instructions) and other components. The processor 1005 at least partially executes the program 1002 to implement part or all of the data processing methods described in one or more embodiments of

the present disclosure. In some embodiments, the first processing unit 102, the second processing unit 103, the third processing unit 104, the fourth processing unit 105, and/or fifth processing unit 107 shown in FIG. 7 and FIG. 8 may be partially or fully integrated into an independent processor 1005, but this is not intended to be limited.

5 The processor 1005 is electrically coupled to the recording medium 1001 via a bus 1003. The processor 1005 is also electrically coupled to an input/output (I/O) interface 1004 via the bus 1003. The network interface 1006 is also electrically connected to the processor 1005 via the bus 1003. The processor 1005 executes the program 1002 encoded in the recording medium 1001 to enable the data processing system 1 to implement part or all of the data
10 processing methods described herein. In some embodiments, the processor 1005 is a central processing unit (CPU), a multiprocessor, a distributed processing system, an application-specific integrated circuit (ASIC), and/or a suitable processing unit.

In some embodiments, the recording medium 1001 is an electronic, magnetic, optical, electromagnetic, infrared, and/or semiconductor system (or device or apparatus). For example,
15 the recording medium 1001 includes semiconductor or solid-state memory, magnetic tape, removable computer disk, random access memory (RAM), read-only memory (ROM), hard disk, and/or optical disk. In one or more embodiments using optical disks, the recording medium 1001 includes compact disk-read read-only memory (CD-ROM), compact disk-read/write (CD-R/W), and/or digital video disc (DVD).

20 The I/O interface 1004 is coupled to an external circuitry. In some embodiments, the I/O interface 1004 includes a keyboard, keypad, mouse, trackball, trackpad, touchscreen, and/or cursor keys for communicating information and commands to the processor 1005.

The network interface 1006 enables an assembly process production system 1 to communicate with a network, which can be connected to one or more other computer systems.
25 The network interface 1006 includes a wireless network interface, such as Bluetooth, wireless fidelity (WIFI), worldwide interoperability for microwave access (WIMAX), general packet

radio service (GPRS), or wideband code division multiple access (WCDMA); or a wired network interface, such as Ethernet, universal serial bus (USB), or Institute of Electrical and Electronic Engineers-1364 (IEEE-1364).

As described above, the data processing method of this embodiment processes multiple
5 scanned data obtained in different ranges under a predetermined environment (e.g., scanned data obtained by different LiDAR devices or stored scanned data). The method first sparsifies the scanned data according to a relation graph and then samples the sparsified data at a predetermined voxel size for smoothing. Afterward, the smoothed data is divided into multiple data clusters. After optimizing these data clusters, a state matrix is obtained.
10 According to the state matrix, the relative relations between the multiple nodes in each data cluster are further optimized to generate point cloud map information. Therefore, the data processing method of this embodiment efficiently and accurately processes scanned data (e.g., point cloud data) through sparsifying, then increases the accuracy of the sparsified data through smoothing. Afterward, the data is divided into multiple data clusters, and the state matrices
15 of these data clusters are obtained for clustered data processing. This reduces the burden on system computing performance and expands the scale of data that can be processed.

The components of several embodiments are summarized above so that a person having ordinary knowledge in the technical field to which the present disclosure belongs can better understand the concepts of the embodiments of the present disclosure. Although the present
20 disclosure has been described with reference to the preferred embodiment thereof, it will be understood that the present disclosure is not limited to the details thereof. Various substitutions and modifications have been suggested in the foregoing description, and others will occur to those of ordinary skill in the art. Therefore, all such substitutions and modifications are intended to be embraced within the scope of the present disclosure as defined
25 in the appended claims.

WHAT IS CLAIMED IS:

1. A data processing method, comprising steps of:

receiving a plurality of first scanned data and a plurality of second scanned data corresponding to a preset environment, wherein the first scanned data correspond to a first position range in the preset environment, and the second scanned data correspond to a second position range in the preset environment,

sparsifying the first scanned data and the second scanned data according to a relation graph to generate a plurality of first sparsified data and a plurality of second sparsified data,

uniformly sampling the first sparsified data and the second sparsified data according to a first voxel size to generate a plurality of first sampled data and a plurality of second sampled data,

generating a plurality of smoothed data according to the first sampled data and the second sampled data,

dividing the smoothed data into a plurality of data clusters,

optimizing relative relations of the data clusters and obtaining a state matrix,

optimizing relative relations of a plurality of nodes in each data cluster according to the state matrix, and

generating point cloud map information according to the data clusters.

2. The data processing method as claimed in claim 1, wherein the step of “receiving the plurality of first scanned data and the plurality of second scanned data corresponding to the preset environment” comprises steps of:

obtaining the first scanned data of the preset environment in the first position range by a first LiDAR device, and

obtaining the second scanned data of the preset environment in the second position range by a second LiDAR device.

3. The data processing method as claimed in claim 1, wherein the step of “receiving the

plurality of first scanned data and the plurality of second scanned data corresponding to the preset environment” comprises steps of:

obtaining the first scanned data of the preset environment in the first position range by a LiDAR device, and

5 receiving the second scanned data from a storage device.

4. The data processing method as claimed in claim 1, wherein the step of “generating point cloud map information according to the data clusters” comprises steps of:

determining whether the point cloud map information is complete, and

outputting the point cloud map if the point cloud map information is complete.

10 5. The data processing method as claimed in claim 4, further comprising steps of:

uniformly sampling the first sparsified data and the second sparsified data according to a second voxel size to generate a plurality of third sampled data and a plurality of fourth sampled data if the point cloud map information is not complete, and

generating the smoothed data according to the third sampled data and the fourth sampled data,
15 data,

wherein the second voxel size is less than the first voxel size.

6. The data processing method as claimed in claim 4, wherein the step of “determining whether the point cloud map information is complete” comprises a step of:

determining whether the point cloud map information is complete according to an absolute position error.
20

7. The data processing method as claimed in claim 1, wherein the step of “generating the plurality of smoothed data according to the first sampled data and the second sampled data” comprises steps of:

obtaining a transformation matrix according to an adjacent position relation of the first sampled data,
25

transforming the second sampled data from a first coordinate system to a second

coordinate system according to the transformation matrix, and

generating the smoothed data in the second coordinate system according to the first sampled data and the second sampled data.

8. The data processing method as claimed in claim 1, wherein the step of “dividing the
5 smoothed data into the plurality of data clusters” comprises steps of:

dividing the smoothed data into a plurality of sub-clusters that are mutually independent,
randomly merging the sub-clusters according to a modular gain and a probability
distribution, and

10 stopping merging the sub-clusters and generating the data clusters when a preset condition
is met.

9. The data processing method as claimed in claim 8, wherein the preset condition
comprises that the number of the nodes in each data cluster reaches a preset number.

10. The data processing method as claimed in claim 1, wherein the step of “optimizing
relative relations of the data clusters and obtaining the state matrix” comprises steps of:

15 selecting a central node from the nodes of each data cluster, and
associating position relations of the central nodes of the data clusters to obtain the state
matrix.

11. The data processing method as claimed in claim 10, wherein the step of “optimizing
relative relations of the plurality of nodes in each data cluster according to the state matrix”
20 comprises a step of:

associating position relations between the central node and other nodes of each data cluster.

12. The data processing method as claimed in claim 1, wherein the relation graph
comprises a plurality of node information and a plurality of edge information connecting the
node information,

25 wherein the step of “sparsifying the first scanned data and the second scanned data
according to the relation graph” comprises steps of:

selecting the edge information according to a preset parameter, wherein the preset parameter comprises a preset number and a preset ratio, and

sparsifying the first scanned data and the second scanned data according to the edge information of the preset parameter.

5 13. A data processing system, comprising:

 a data input unit configured to receive a plurality of first scanned data and a plurality of second scanned data corresponding to a preset environment, wherein the first scanned data correspond to a first position range in the preset environment, and the second scanned data correspond to a second position range in the preset environment,

10 a first processing unit configured to sparsify the first scanned data and the second scanned data according to a relation graph to generate a plurality of first sparsified data and a plurality of second sparsified data,

 a second processing unit configured to uniformly sample the first sparsified data and the second sparsified data according to a first voxel size to generate a plurality of first sampled data
15 and a plurality of second sampled data, and generate a plurality of smoothed data according to the first sampled data and the second sampled data,

 a third processing unit configured to divide the smoothed data into a plurality of data clusters,

 a fourth processing unit configured to optimize relative relations of the data clusters and
20 obtain a state matrix, optimize relative relations of a plurality of nodes in each data cluster according to the state matrix, and generate point cloud map information according to the data clusters, and

 a data output unit configured to output the point cloud map information.

 14. The data processing system as claimed in claim 13, further comprising:

25 a fifth processing unit configured to determine whether the point cloud map information is complete, and

the data output unit is further configured to output the point cloud map information if the point cloud map information is complete.

15. The data processing system as claimed in claim 14, wherein if the point cloud map information is not complete, the second processing unit is further configured to uniformly
5 sample the first sparsified data and the second sparsified data according to a second voxel size to generate a plurality of third sampled data and a plurality of fourth sampled data, and generate the smoothed data according to the third sampled data and the fourth sampled data, wherein the second voxel size is less than the first voxel size.

16. The data processing system as claimed in claim 14, wherein the fifth processing unit
10 is configured to determine whether the point cloud map information is complete according to an absolute position error.

17. The data processing system as claimed in claim 13, wherein the third processing unit is configured to divide the smoothed data into a plurality of sub-clusters that are mutually independent, randomly merge the sub-clusters according to a modular gain and a probability
15 distribution, and stop merging the sub-clusters and generate the data clusters when a preset condition is met, wherein the preset condition comprises that the number of the nodes in each data cluster reaches a preset number.

18. The data processing system as claimed in claim 13, wherein the fourth processing unit is configured to select a central node from the nodes of each data cluster, associate position
20 relations of the central nodes of the data clusters to obtain the state matrix, and associate position relations between the central node and other nodes of each data cluster.

19. The data processing system as claimed in claim 13, wherein the first processing unit, the second processing unit, the third processing unit, and the fourth processing unit are integrated into a processor.

25 20. The data processing system as claimed in claim 13, wherein the relation graph comprises a plurality of node information and a plurality of edge information connecting the

node information,

wherein the first processing unit is configured to select the edge information according to a preset parameter, and sparsify the first scanned data and the second scanned data according to the edge information of the preset parameter, wherein the preset parameter comprises a preset

5 number and a preset ratio.