



A Volumetric Fusing Method for TLS and SFM Point Clouds

Wei Li, Cheng Wang , Senior Member, IEEE, Dawei Zai, Pengdi Huang, Weiquan Liu, Chenglu Wen, Senior Member, IEEE, and Jonathan Li , Senior Member, IEEE

Abstract—A terrestrial laser scanning (TLS) point cloud acquired from a given ground view is incomplete because of severe occlusion and self-occlusion. The models reconstructed by aligning the cross-source point clouds [TLS and structure-from-motion (SFM) point clouds] provide a more complete large-scale outdoor scene. However, because of differences in nonrigid deformation, stratified redundancy of alignment is inevitable and ubiquitous. Therefore, this paper presents a volumetric fusing method for cross-source three-dimensional reconstructions. To eliminate the stratification of aligned cross-source point clouds, we propose a graph-cuts method with boundary constraints for blending the two cross-source point clouds. Then, to reduce the gaps that exist in the blending results, we develop a progressive migration method combined with the local average direction of normal vectors to smooth the unconnected boundary. Finally, experimental results demonstrate the effectiveness of eliminating stratification with the proposed blending algorithm, and the progressive migration method achieves a smooth connection in the boundary of the blended point clouds.

Index Terms—Boundary constraints, graph cuts (GC), progressive migration, volumetric fusion.

I. INTRODUCTION

WITH the rapid development of laser scanning technology, multisource point clouds [Kinect, LiDAR, range cameras, structure-from-motion (SFM), and simultaneous location and mapping (SLAM)] are acquired more conveniently. There are many works on information extraction and analysis of three-dimensional (3D) point cloud, such as automatic road vector extraction [1], object detection [2], [3], and line segment extraction [4], [5]. However, most of those works are based on a single data source. Because of the differences of

scanning precision, range, and view in multisource point clouds, every 3D scanning method or means has its own characteristics and applicable environment. For example, the terrestrial laser scanning (TLS) [6] system provides an array of capabilities in terms of instrument range, scan speed, field of view, size, and portability. However, because of self-occlusion and occlusion, the collected point cloud is often incomplete. Furthermore, it is difficult to fully collect the surface point cloud of an outdoor scene because, in particular, the roof point cloud of a building is often missing. To avoid some self-occlusion and occlusion, a mobile laser scanning (MLS) [7] system (a type of mobile mapping system) collects high-precision point clouds rapidly from both sides of a road. Moreover, also because roof point clouds of buildings on both sides of a road are missing, the collected point cloud, limited by scanning perspective, still is incomplete. SFM [8] and multiview stereo (MVS) [9] methods automatically produce large-scale urban models from airborne imagery; the multiview images are obtained easily by aerial photography from an unmanned aerial vehicle (UAV). However, the SFM point cloud from UAV usually contains the sparse point cloud of a large-scale scanned scene, but omits the details of the ground. For complete 3D reconstruction, using the algorithm directly is the preferred method to repair the missing parts of the point cloud. For example, to handle the occlusion problem occurring during data acquisition, Cai *et al.* [10] presented an indispensable process for completing the point cloud. However, this method is invalid when a large range of 3D information is missing.

Reconstructed methods by fusing multisource point clouds achieve complementary advantages and are becoming a new trend. However, because of differences in the cross-source 3D point clouds (e.g., point cloud density, distribution uniformity, accuracy, scene size, and occlusion), it is inevitable that much stratified, redundant noise exists in the aligned point clouds. Furthermore, the stratified, redundant information leads directly to a decline in the quality of the reconstructed 3D model. As seen from Fig. 1(a), the SFM point cloud is nonrigid and more seriously deformed than the TLS point clouds. Thus, as shown in Fig. 1(b), when a SFM point cloud is aligned to a corresponding TLS point cloud (yellow), there are many redundancy such as stratified in the reconstruction. Therefore, to eliminate this redundancy, we propose a volumetric method based on a graph-cuts (GC) model to fuse the aligned cross-source point clouds (i.e., SFM and TLS point clouds). First, with a commercial software (RiSCAN software from RIEGL), we eliminate

Manuscript received September 15, 2017; revised February 26, 2018, May 14, 2018, and June 18, 2018; accepted July 11, 2018. This work was supported in part by the general grant from the Natural Science Foundation of China under Grant U1605254, Grant 61771413, and Grant 41471379. (Corresponding author: Cheng Wang.)

W. Li, C. Wang, D. Zai, P. Huang, W. Liu, and C. Wen are with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Information Science and Engineering, Xiamen University, Xiamen 361005, China (e-mail: weili_xs@163.com; cwang@xmu.edu.cn; david102812@gmail.com; alualu628628@163.com; 472483520@qq.com).

J. Li is with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Information Science and Engineering, Xiamen University, Xiamen 361005, China, and also with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: junli@uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2018.2856900

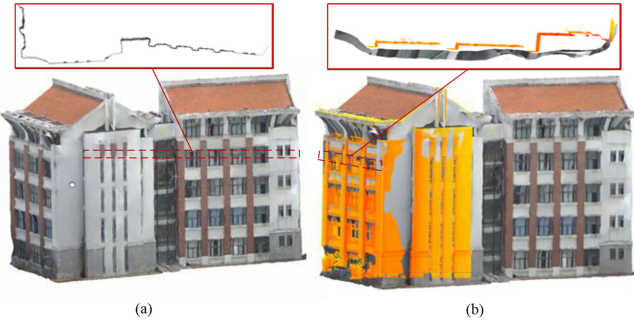


Fig. 1. (a) SFM point cloud, which samples densely from 3D mesh models, is nonrigid. (b) Coarse reconstructed 3D point cloud.

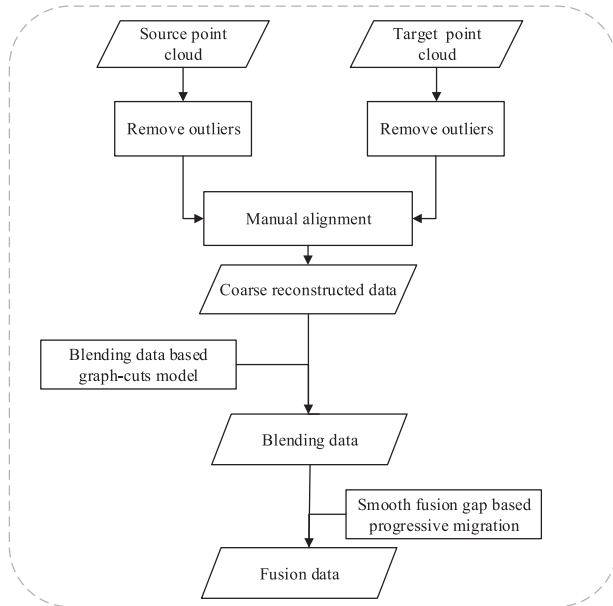


Fig. 2. Flowchart of the proposed method.

some outliers and redundant targets. Second, inspired by [11], to eliminate the stratified point cloud, we propose an improved GC model to eliminate the deformed point cloud (SFM point cloud), which overlaps the target 3D point cloud (TLS point cloud). Third, because there are many gaps as well as fractured areas, to ease and smooth the gaps, we propose a method of progressive migration perception based on Euclidean distance and normal vector of local region. The specific operational framework is shown in Fig. 2.

The contributions of this paper are as follows: (1) to eliminate stratification in the reconstructions of the cross-source 3D point clouds, we propose a new volumetric blending method based on a GC model to remove the deformed point clouds (SFM point clouds) in the overlapped region; (2) to smooth the gaps that exist in the result of blending, we propose a method of progressive migration perception.

II. RELATED WORKS

Recently, urban 3D reconstruction has been applied widely to high-precision mapping and augmented reality. 3D

reconstructed methods from multisource point clouds are becoming popular. Many scholars have recognized the benefits of urban 3D reconstruction and modeling from cross-source 3D point clouds [12]–[16].

Many methods have been proposed for 3D reconstruction from a series of different point clouds, the following methods are representative. Hornung *et al.* [17] proposed a reconstructed method of watertight 3D models; however, their method still requires localizing the acquisition vehicle and reconstructing the 3D models based on the street views that lack the roof parts. Richard *et al.* [18] presented a system based on Kinect for accurate real-time mapping of complex and arbitrary indoor scenes. However, because of the limit of sensor scanning distance, their system was unsuitable for reconstructing outdoor scenes. Cheng *et al.* [19], using 3D road networks and 3D building contours, proposed a hierarchical method for the automatic registration of airborne laser scanning (ALS) and MLS 3D point clouds. However, the alignment algorithm depends on the driving trajectory, and does not consider the fusion of aligned results. Peng *et al.* [20], using classical iterative closest points (ICP), registered two cross-source point clouds (SFM and LiDAR point clouds). But, many assumptions are made, including removing outliers and manually selecting objects. Huang *et al.* [21] proposed a systematic method for registering cross-source point clouds. They proposed a scale normalization method to eliminate the scale problem and used a new graph construction method to combine the two structures. However, they do not deal with the post-registration redundancy.

The volumetric fusion methods of the aligned 3D point clouds (cross-source point clouds) have received much attention. Some works combine street-side and aerial data for mesh reconstruction [22]–[24]. Fiocco *et al.* [11], using a distance field over an octree and an out-of-core dual contouring approach, proposed a volumetric fusion method as a segment of the deformed point cloud integrated over 200 000 points from a tripod-mounted ground LiDAR and an aerial digital surface model (DSM). Despite the complex algorithm, the results are noisy and contained many holes. Forkuo *et al.* [25] performed the automatic fusing of photogrammetric imagery and laser scanning point clouds. However, their fusion focuses on TLS 3D point clouds and digital images. Fruh *et al.* [26] constructed meshes over street-side LiDAR range maps and generated a larger scale DSM. However, they reconstructed separately a facade and an airborne mesh without topological fusion. Consequently, the simulation is rough, and there still existed some holes in the 3D reconstruction data. Frueh *et al.* [27], using planar or horizontal interpolation, generated textured 3D building facade meshes from laser scanning point clouds and camera images, which mainly filled holes in the background layer, caused by occlusion from foreground layer objects. However, due to the characteristics of the TLS systems, point clouds of the roofs are still missing. Shan *et al.* [23] applied the method of Poisson surface reconstruction directly to join point clouds computed by patch-based MVS [28], where there is not a cross-consistency check between the airborne and street-side data. Recently, Bdis-Szomor *et al.* [29] proposed an efficient view-driven meshing approach for street-side images and large-scale height maps. Then, they

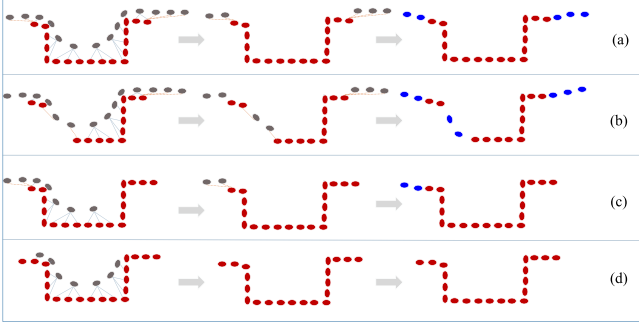


Fig. 3. Illustration of volumetric fusion in 2D. (a) and (b) Target point set (red) completely overlap with source point set (grey), and the size of the overlapping area is equal to the target point set. (c) Target point set (red) and source point set (grey) partially overlap each other. (d) Source point set (grey) completely overlap with target point set (red), and the size of the overlapping area is equal to the source point set.

superimposed a 3DT on the MVS points and classified the inside/outside of the tetrahedra while enforcing line-of-sight and photo-consistency constraints [13]. They blended two barely different 3D point clouds using a GC model; therefore, many gaps exist in the 3D point cloud. Additionally, as an SFM-based algorithm, the method is time consuming and of lower precision.

Compared with [29], we propose an improved GC model to blend the cross-source point clouds (TLS and SFM point clouds), and, to smooth the gaps of boundary, develops a progressive migration method of SFM points near the blending gap.

III. METHODOLOGY

As shown in Fig. 2, to reconstruct a more complete 3D model using a SFM point cloud, we designed a point cloud fusion framework. Given two cross-source 3D point clouds: source and target point clouds, first, using industrial software (RiSCAN software of RIEGL), we performed some preprocessing that included removing the outliers from each point cloud and manually aligning source point clouds with the target point clouds. Then, we proposed a new volumetric fusion method based on the GC model in [30] and [31], where two 3D cross-source point clouds, as well as TLS and SFM point clouds, are applied directly. Thus, the deformation point cloud overlapping the reference data was eliminated. In the smoothing step, progressive migration of the local normal vector is designed to smooth the fusion gap.

Fig. 3 gives an illustration of volumetric fusion in 2D. The point distribution of two point clouds in one dimension are summarized as four cases [see Fig. 3(a)–(d)]. The charts in the first column represent four cases as well as the aligned cross-source point clouds. The charts in the middle column represent the blending of the indicated results by the proposed GC model. The charts in the third column show the indicated results of smooth acquired with the method of progressive migration. In Fig. 3, the red points represent source point set (i.e., SFM point clouds), and the grey points represent target point set (i.e., TLS point clouds). Points connected by light blue lines are assigned to overlapping regions. The other points, connected by orange

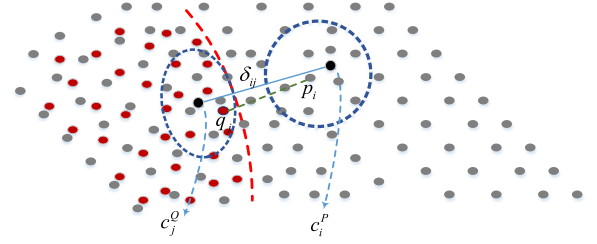


Fig. 4. Illustration of boundary constraints. The grey points represent source point set (i.e., SFM point cloud), and the red points represent the target point set (i.e., TLS point cloud). The grey points in the dashed box is the K_1 nearest points of point p_i , wherein the centroid is denoted by c_i^P . The red points in the dashed box is the K_2 nearest points of point q_j , wherein the centroid is denoted by c_j^Q . q_j is in the target point set, which represents the nearest point of p_i . δ_{ij} represents the distance between the centroid c_j^Q and the centroid c_i^P .

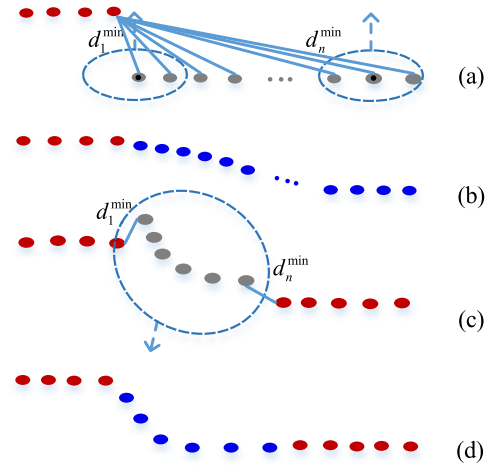


Fig. 5. Illustration of smooth fusion in 2D. The red points represent the target point set (i.e., TLS point cloud), the grey points represent source point set (i.e., SFM point cloud) before processing, and the blue points represent source point set (i.e., SFM point cloud) after processing. The two pairs figures [(a), (b) and (c), (d)] represent two different cases of smooth. The blue arrow indicates the direction of migration points.

dotted lines are assigned to nonoverlapping regions. The details are described as follows.

A. Preprocessing for Cross-Source Point Clouds

Laser scanners typically collect point clouds with varying point densities. Additionally, reflective materials such as glass and dust in the air exist in the acquired point cloud which results in sparse outliers. Also, redundant noise, due to calibration errors in the camera, exists in the 3D data generated by the SFM or MVS algorithms. Therefore, before fusing the two cross-source point clouds, we first eliminate the redundant noise and sparse outliers. Then, using industrial software for raw source and target point clouds, we remove the redundant targets and outliers as much as possible. However, noise and many outliers still remain. Therefore, we make use of the method [32], where some of the outliers are filtered further by performing a statistical analysis on the neighborhood of each point and trimming those that do not meet a certain criteria. Coarse alignment is completed first

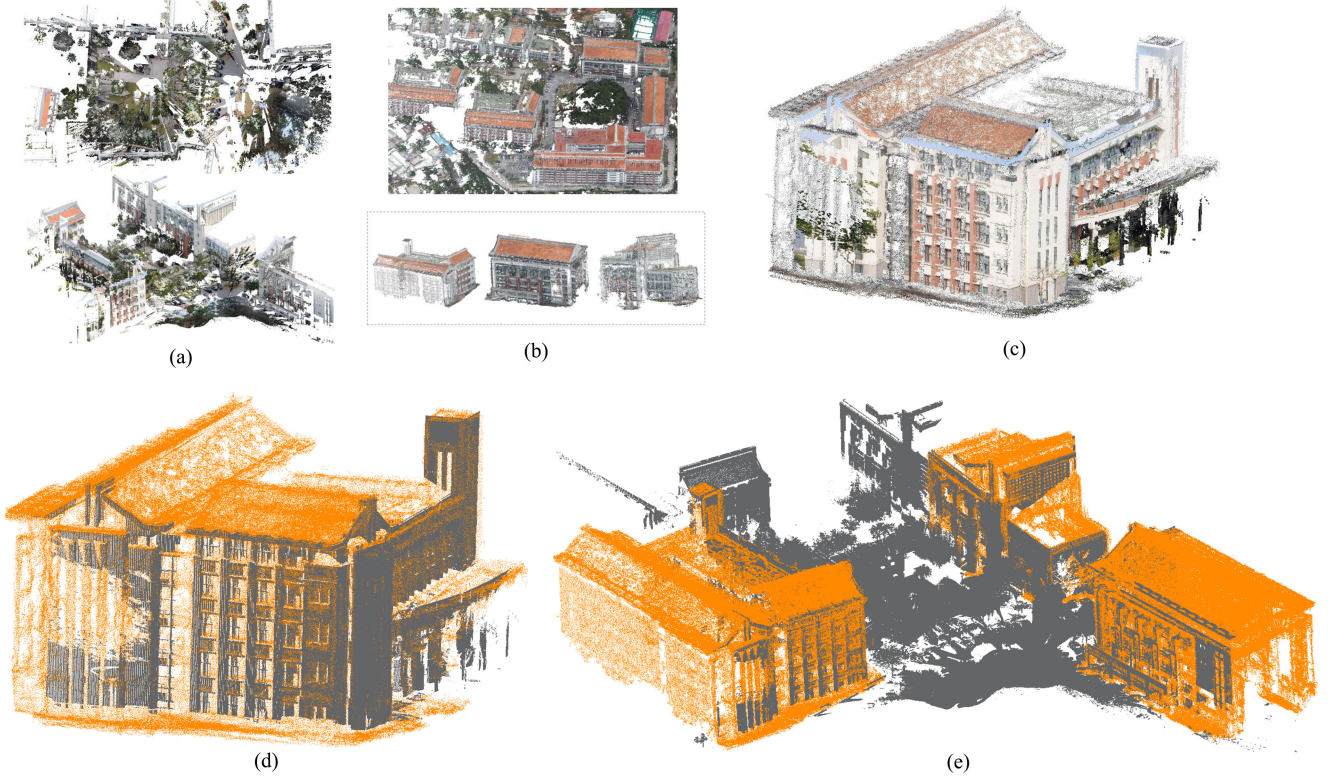


Fig. 6. Aligned results of the two different point clouds. (a) Registered result of 3D point clouds from the scanning mutations. (b) Several 3D point clouds from SFM raw data. (c) Registration result of the above two point clouds. (d) Fusing result of a pair of the cross-source point clouds. (e) Fusing results of multiobjects point clouds.

by the RiSCAN software and then fine registration is achieved by the ICP algorithm.

B. Point Cloud Blending

Given two aligned cross-source point clouds (i.e., SFM point cloud P and TLS point cloud Q), as shown in the first column of Fig. 3, the point $p_i \in P$ is deformed and is inaccurate, compared with its corresponding point $q_j \in Q$, which generates hierarchical redundancy. To eliminate the redundancy, we formulated volumetric fusion as a segmentation issue for the source point cloud P . We assign a binary label $l_i \in \{0, 1\}$ to each point $p_i \in P$, which means the point marked by 0 is removed and the point marked by 1 is retained.

To measure the substitutability of source point p_i , we first compute the angle between the normal vectors of the source point p_i and the nearest point q_j in the target point cloud Q . Then, the nearest distance d_{ij} is computed by the two coordinates of points p_i and q_j . Thus, the likelihood for a source point p_i to have a substitute is formulated as follows:

$$\varphi_i = e^{(-d_{ij}^2)/(2\sigma^2)} \bullet \max \{0, \cos \theta_{ij}\} \quad (1)$$

where $\cos \theta_{ij} = \mathbf{n}(p_i)^T \mathbf{n}(q_j)$; $\mathbf{n}(p_i)$ represents the normal vector of the point p_i , $\mathbf{n}(q_j)$ represents the normal vector of the point q_j ; σ represents a blending parameter to control our notion of vicinity, which involves the deviations of P and Q due to the errors of alignment. Thus, the likelihood φ_i ranges from zero (no substitute) to one (perfect substitute). Additionally, normal

vectors are computed in P and Q by k-nearest neighbor (K-NN) and least squares plane fitting, respectively, and by flipping the normal according to the relative perspective.

For the segmentation of the points near the boundary of the fusing point cloud, it is necessary to enlarge the corresponding distance such that it becomes more impossible to substitute the points. We develop a method of boundary constraints to reduce the gaps in the fusing boundary. As Fig. 4 shows, the grey points (SFM) and the red points (TLS) partially overlap. Because of the characteristics of the point cloud boundary, the nearest neighbor of p_i is the point q_j in the point cloud Q . To enlarge the distance between p_i and q_j , the likelihood function is modified as follows:

$$\varphi_i = e^{(-\omega_{ij} d_{ij}^2)/(2\sigma^2)} \bullet \max \{0, \cos \theta_{ij}\}. \quad (2)$$

The weight factor ω_{ij} is formulated as

$$\omega_{ij} = \begin{cases} 1 & \delta_{ij} < \delta_0 \\ M & \text{others} \end{cases} \quad (3)$$

where the distance δ_{ij} between two centroids represents the nearest distance between a point in cloud point P and a point cloud Q . δ_0 ($\delta_0 = 0.05$ in our experiments) is a constant, that is somewhat larger than the resolution factor \bar{r} ($\bar{r} = 0.03$ in our experiments) for point cloud Q . M ($M = 10\delta_0$ in our experiments) is a penalty threshold of the boundary to reduce the gap of blending boundary of cross-source point clouds.

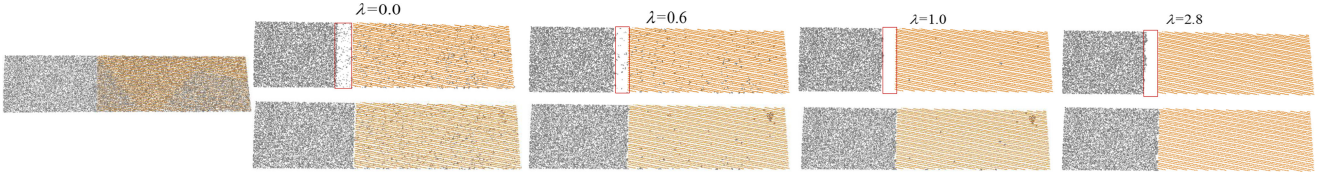


Fig. 7. Comparative effectiveness of blending. The aligned cross-source 3D point clouds are shown at the first column, and the following columns represent the blending results according to different regularization parameters λ . The top row shows the result of the fusion method based on GC, and labeling overlapped point cloud with the orange data, the second row shows the results of the proposed algorithm. The red box represents fracture area.

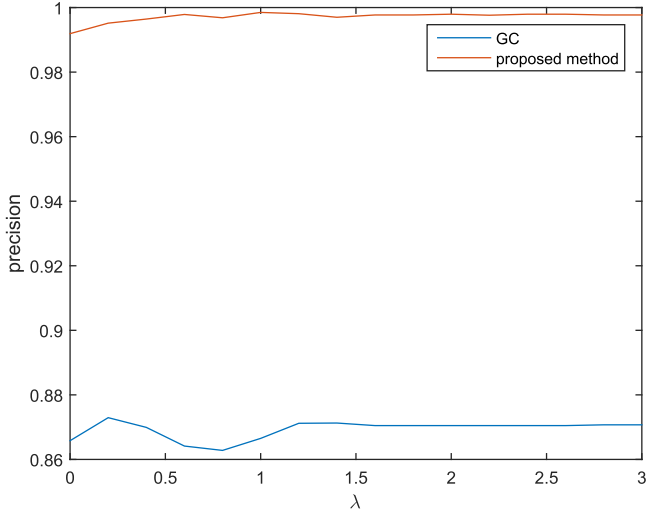


Fig. 8. Comparison of accuracy. The figure shows different distribution of accuracy respectively by GC and proposed method.

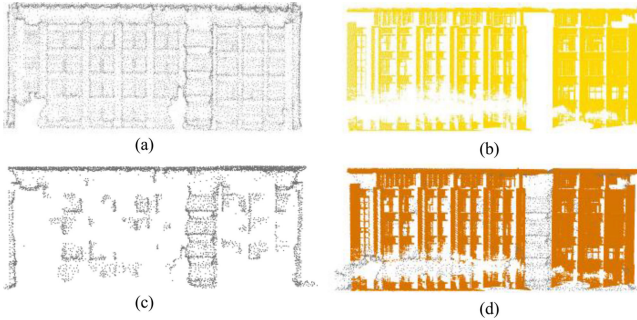


Fig. 9. Output of our blending. (a) and (b) 3D point clouds from SFM and REIGL VZ1000, and had been aligned. (c) The segmented result by using GC. (d) The final fusion result.

For a smooth segmentation, we define the influence between adjacent points of the K-NN graph over P as follows:

$$\psi(p_i, p_{i_N}) = e^{-d_{i,i_N} / \text{medd}_{i,i_N}} \quad (4)$$

where d_{i,i_N} represents the distance between any two adjacent points $p_i \in P$ and $p_{i_N} \in P$; p_{i_N} represents one of the K-NN points to p_i , where $i_N \in \{1, 2, \dots, K\}$ is an index of p_{i_N} , K is the number of K-NN points; and medd_{i,i_N} represents the median of all K-NN distances in P .

As in the method of Kolmogorov *et al.* [31], by solving energy minimization problems, we seek the binary labels over source

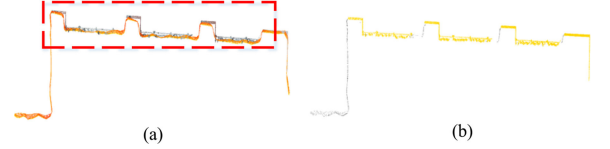


Fig. 10. Elimination of point cloud stratification. (a) Stratified demonstration of point cloud registration. (b) Result of fusion.

point cloud P . The energy function is defined as follows:

$$E(L) = \sum_{i:p_i \in P} E_i(l_i) + \lambda \sum_{i:i_N} \psi(p_i, p_{i_N}) \cdot \mathbf{I}[l_i \neq l_{i_N}] \quad (5)$$

where l_{i_N} represents the label of the point adjacent to point p_i . \mathbf{I} represents an indicator function that $\mathbf{I} = 1$ if $l_i \neq l_{i_N}$, otherwise $\mathbf{I} = 0$; λ is a regularization factor; $L = \{l_1, l_2, \dots, l_n\}$ represents a complete labeling. The unary penalties $E_i(l_i)$ for point p_i to obtain label l are defined as follows:

$$E_i(l_i) = \begin{cases} 1 - \varphi_i & l_i = 0 \\ \varphi_i & l_i = 1 \end{cases} \quad (6)$$

where φ_i is computed by (2).

C. Smoothing Fusing Gaps

The source point cloud P overlapped with the target point cloud Q , and was eliminated in the blending step. Because of registered errors and the nonrigid deformation of data, the fusing point cloud is discontinuous and the grating is inconsistent. The remaining 3D point clouds are denoted as the blended point clouds, which have no overlapping point clouds in 3D space. However, many fractures exist at the intersection of the fusion. To eliminate the gap in the fusing point cloud, we must obtain the distance and direction that each point would move, and then move the points. Therefore, we propose a progressive migration method with a local average direction of normal vectors.

Given a blended point cloud, we first compute the minimum Euclidean distance (denoted by d_i^{\min} , $i = 1, \dots, n$) between a point p_i and a target point cloud Q . Then, the mutual nearest neighbor distance of the cross-source point clouds near the gap is denoted as d_{nn} . Therefore, given a distance threshold T , for each boundary point $p_i \in P$, the migration distance near the gap is designed to be proportional to the Euclidean distance d_i , denoted as follows:

$$d_i = (1 - (d_i^{\min} - d_{nn})/T)d_i^{\min} \cos \theta_i \quad (7)$$



Fig. 11. Smoothing fusion based on progressive migration at two views as well as (a) and (b). The three rows in (a) and (b) represent the aligning, blending, and smoothing results of cross-source validated point clouds. 3D data of the orange yellow color represents TLS point cloud, the gray color represents SFM point cloud.

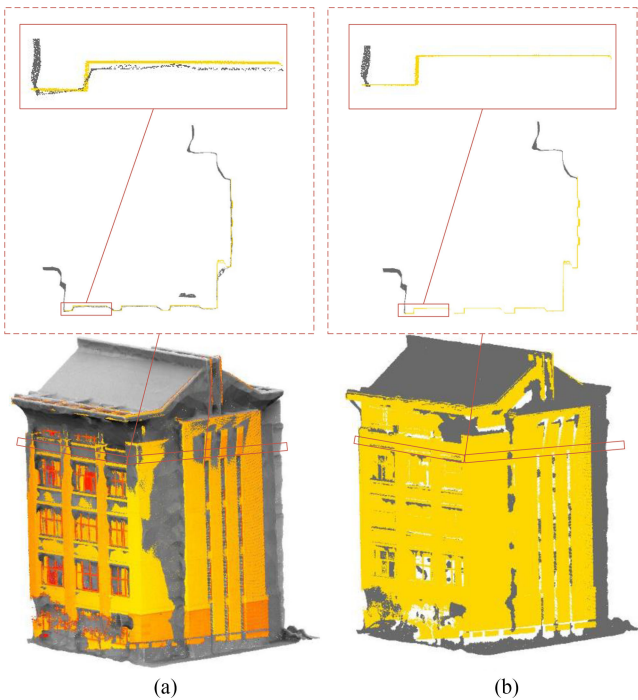


Fig. 12. Two cross-source point clouds are sampled densely from their corresponding triangular mesh models. (a) Aligned result of two cross-source point clouds, on which exists the stratification redundancy. (b) Corresponding fusion result.

where θ_i represents the angle between the normal vector of the searching point p_i and the vector generated by the point p_i and its nearest point $q_j \in Q$.

The migration direction is determined by the local average normal vector and relative distance d_i . Each point p_i is shifted by distance d_i , along its given direction so that it has a smooth effect. Thus, we define the shifting vector as follows:

$$\mathbf{m}_i = \begin{cases} p_i + d_i \cdot \mathbf{n}(p_i) & d_i < T \\ \mathbf{0} & d_i \geq T \end{cases} \quad (8)$$

where $\mathbf{n}(p_i)$ represents the normal vector of point p_i . The migration distance threshold T is set so that the points away from

the gap remain stationary, which greatly improves the efficiency of the calculation.

As Fig. 5 shows, the red point set (TLS) and the blue point set (SFM) are cross-source point clouds. The blending step has been completed. Then, it is necessary to smooth the blending results with loopholes. Fig. 5(a) shows how each point of source data migrates near the intersection of the cross-source point cloud along a given direction. Fig. 5(b) indicates the effect of fusion. From (7), it can be seen that, when the nearest Euclidean distance is larger than T ($T = 10\delta_0$ in our experiments), the points do not migrate, but remain stationary. Fig. 5(c) and (d) show the smoothing of the gap of a small area within a loophole.

D. Evaluation Method

To evaluate the results from fusing cross-source point clouds, we first formulate the validation of the blending operation as the size of the regional gaps. The area of a gap region is measured by projecting the area of the selected point cloud like a spatial plane in the normal vector direction. The segmentation precision of the overlapping area is computed as

$$\text{precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (9)$$

where TP is the number of removed true positive points. FP is the number of removed false positive points.

Second, we compute the segmentation precision with the increased value of the regularization factor. Finally, the effect of gap smoothing is compared with related algorithms.

IV. EXPERIMENT

Our method relies on the Point Cloud Library (PCL) [33] and the libLAS library [34] for processing of the 3D point clouds, and the GCOptimization library [30] for the GC in C++. We could not find any publicly available dataset with both terrestrial LiDAR 3D and SFM point cloud for the same geographic location. Therefore, we performed the experiments using our datasets: Haiyun Campus [$200 \times 200 \text{ m}^2$, Fig. 6(a), (b)] captured in Xiamen, China. The airborne 3D data is generated by VisualSFM [35], which automatically generates 3D point clouds. The terrestrial LiDAR 3D point clouds are acquired via REIGL VZ-1000.

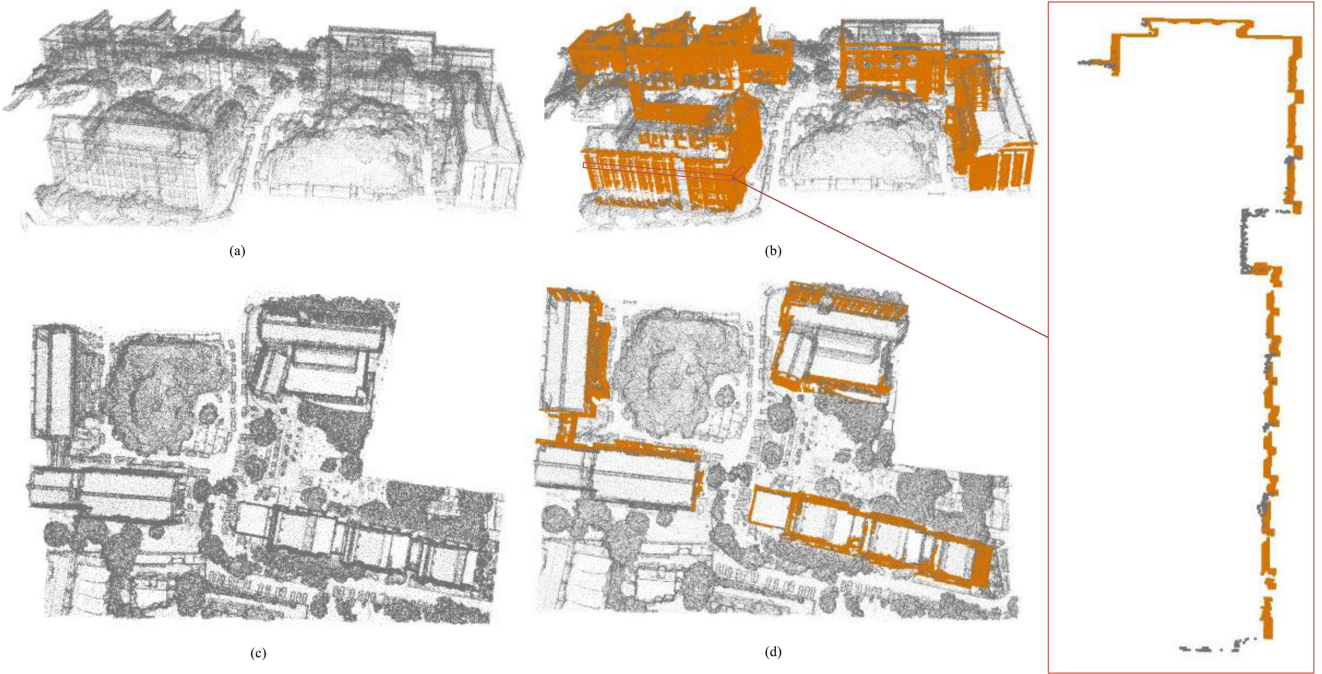


Fig. 13. Final output of our blending at two views. The left figures (a) and (c) show that a SFM point cloud is on two different views, the right figures (b) and (d) show the corresponding fusion results.

Terrestrial laser scanners, such as REIGL VZ 1000 (selected), Z+F Imager 5010, etc. collect highly dense and accurate 3D point cloud data by mutistation scanning based on a ground view. In regard to SFM/MVS point clouds, although the Terrestrial 3D LiDAR data has many advantages, it does not cover the roof of buildings; occlusion problems are always present. To compensate for the roof and some of the occlusion in Terrestrial LiDAR point clouds, the scanned objects such as buildings from an SFM point cloud were selected and segmented manually. Then, we eliminated outliers and aligned them with corresponding targets in the Terrestrial LiDAR point clouds. Fig. 6(c) shows the results of one of the alignments. However, hierarchical phenomena and redundant noise exist [See Fig. 6(a)]. To implement fusion of two different point clouds, we must eliminate stratification. In this section, we first compare our proposed method with previous methods, and then show the results of fusion. Fig. 6(d) and (e) show the fusing results of single and multiple pairs of cross-source point clouds respectively.

A. Comparative Experiments

To evaluate the results, we define the accuracy of the blending step as the number of segmented out-points divided by the actual number of out-points. As shown in Fig. 7, we fix our parameters as $\sigma_b = 0.2$, $\delta_0 = 0.05$, $M = 100$; the related volumetric regularization parameters λ are 0.0, 0.6, 1.0, 2.8. First, the SFM cloud in the overlap becomes cleaner as lambda increases. The second row in Fig. 7 shows that a boundary gap exists in the fusion results. However, compared with the algorithm of [13] as well as GC, our method reduces the gap between the two cross-source point clouds. Clearly, there is no gap with our proposed method. Our experiments were based on the labeling data. As

shown in Fig. 8, the accuracy with the GC algorithm is less than that of our method and about 20% less than the accuracy of our proposed algorithm. Therefore, the corresponding quantitative results show that our method more effectively eliminates the boundary gap of the point cloud fusion.

B. Fusing Experiments

To include the roofs of buildings in the TLS point cloud and improve relative accuracy, using the proposed method, we first eliminate the airborne 3D (SFM) data overlapped with the TLS point cloud (specifically, the SFM point cloud that is spare and deformed) and minimize the energy function through GC with boundary constraints.

The airborne 3D point cloud [See Fig. 9(a)], which is generated by SFM technology, but not yet made into a triangle mesh model, is spare and deformed. We use our proposed volumetric fusion procedure to remove the airborne 3D point clouds that overlap with TLS point clouds [See Fig. 9(b)]. We combine the distance between corresponding points and their difference of normal vectors and optimize the energy function by GC. In our experiments, the regularization parameter is set as $\lambda = 3.0$. The results of blending are shown in Fig. 9(c). Our results of fusion are shown in Fig. 9(d).

Fig. 10 shows the results of local fusion. From Fig. 10(a), it is seen that stratification exists on the two registered point clouds. However, after applying our method, the stratification is eliminated and fused. Fig. 6 shows the final fusion results.

It is important to smooth the faults because doing so reduces point cloud discontinuity and grating inconsistencies. After the registration of the cloud data, data stratification is a common phenomenon. To validate the effect of smooth fusion (Fig. 11),

a group of cross-source 3D data is used to verify the effect of the gap buffer. We first blend the two different point clouds [see the second row of Fig. 11(a) and (b)]. There are faults near the fusing gap; therefore, we must reduce the gap. Because of the deformation and low relative accuracy of SFM point clouds, we smooth the SFM point cloud near the gap using our proposed method of progressive migration perception. As shown in the third row of Fig. 11, the proposed method repairs the gap and eases the faults. To validate our method and better show the effect of volumetric fusion, the two SFM point clouds are sampled densely from their corresponding triangular mesh models (See Fig. 12). We eliminate the stratification and achieve the fusion of cross-source point clouds. Therefore, the effect of volumetric fusion shows that the proposed method is effective. However, the aligned error, which has different normal directors and corresponding distances, generates the volumetric gap at the bottom of Fig. 12(b).

Moreover, our method also applies to the complementary of the SFM large-scale outdoor point cloud which enhances the details of the street-view 3D point clouds. As seen in Fig. 13, airborne acquisition and ground-view 3D point clouds provide complementary 3D information on a city scale. A relative complete SFM point cloud lacks ground-view details, while the TLS point cloud is incomplete for higher floors and severe occlusion. Thus, we also strengthen the data types by fusing them as a segmented problem of a point cloud. The results on the right show that, using our proposed volumetric fusing method, the cross-source point clouds are blended and smoothed; therefore, finer 3D models can be generated.

V. CONCLUSION

We proposed a new volumetric fusion method for TLS and SFM point clouds. In the blending step, a GC-based segmentation method with boundary constraints was developed for eliminating the redundant fusion information. The fusion results, using our method and evaluated with the designed dataset, demonstrate that our method is more effective than previously proposed methods. In addition, the progressive migration method was applied to achieve smooth connection in the boundary of the blending point cloud. Thus, we combined the strengths of the data types to reconstruct 3D point cloud models that enhance the details. Our experiments showed good fusion quality. Part of our future work will be to apply our method on the campus, even on a city-wide scale via cross-source 3D data.

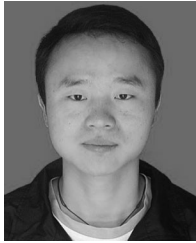
ACKNOWLEDGMENT

The authors would like to acknowledge the anonymous reviewers for their valuable comments.

REFERENCES

- [1] W. Cheng, T. Hassan, N. El-Sheimy, and M. Lavigne, "Automatic road vector extraction for mobile mapping systems," *Int. Arch. Photogramm. Remote Sens.*, vol. XXXVII, no. Part B3b, pp. 515–521, 2008.
- [2] H. Wang, C. Wang, H. Luo, P. Li, M. Cheng, C. Wen, and J. Li, "Object detection in terrestrial laser scanning point clouds based on hough forest," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1807–1811, Oct. 2014.
- [3] D. Zai, J. Li, Y. Guo, M. Cheng, Y. Lin, H. Luo, and C. Wang, "3-d road boundary extraction from mobile laser scanning data via supervoxels and graph cuts," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 802–813, Mar. 2017.
- [4] Y. Lin, C. Wang, J. Cheng, B. Chen, F. Jia, Z. Chen, and J. Li, "Line segment extraction for large scale unorganized point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 102, pp. 172–183, 2015.
- [5] Y. Lin, C. Wang, B. Chen, D. Zai, and J. Li, "Facet segmentation-based line segment extraction for large-scale point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 4839–4854, Sep. 2017.
- [6] M. Lemmens, "Terrestrial laser scanning," in *Geo-information*, M. Lemmens, Ed., Netherlands: Springer, 2011, pp. 101–121.
- [7] S. Mendenhall, "Mobile laser scanning," *CE News*, vol. 23, no. 5, 2011.
- [8] J. Heinly, J. L. Schonberger, E. Dunn, and J.-M. Frahm, "Reconstructing the world in six days (as captured by the Yahoo 100 million image dataset)," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2015, pp. 3287–3295.
- [9] E. Tola, C. Strecha, and P. Fua, "Efficient large-scale multi-view stereo for ultra high-resolution image sets," *Mach. Vision Appl.*, vol. 23, no. 5, pp. 903–920, 2012.
- [10] Z. Cai, C. Wang, C. Wen, and J. Li, "3d-patchmatch: An optimization algorithm for point cloud completion," in *2nd IEEE Int. Conf. Spatial Data Mining Geographical Knowl. Serv.*, July 2015, pp. 157–161.
- [11] M. Fiocco, G. Bostrom, J. G. Gonçalves, and V. Sequeira, "Multisensor fusion for volumetric reconstruction of large outdoor areas," in *5th Int. Conf. 3-D Digital Imag. Modeling*, 2005, pp. 47–54.
- [12] R. Cabezas, O. Freifeld, G. Rosman, and J. W. Fisher, "Aerial reconstructions via probabilistic data fusion," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2014, pp. 4010–4017.
- [13] A. Bódis-Szomorú, H. Riemenschneider, and L. Van Gool, "Efficient volumetric fusion of airborne and street-side data for urban reconstruction," in *23rd Int. Conf. Pattern Recognit.*, 2016, pp. 3204–3209.
- [14] C. Strecha, T. Pylvänäinen, and P. Fua, "Dynamic and scalable large scale image reconstruction," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2010, pp. 406–413.
- [15] P. Musialski, P. Wonka, D. G. Aliaga, M. Wimmer, L. v. Gool, and W. Purgathofer, "A survey of urban reconstruction," *Computer Graphics Forum*, vol. 32, no. 6., 2013, pp. 146–177.
- [16] B. Curless and M. Levoy, "A volumetric method for building complex models from range images," in *Proc. 23rd Annu. Conf. Comput. Graph. Interactive Techn.*, 1996, pp. 303–312.
- [17] A. Hornung and L. Kobbelt, "Robust reconstruction of watertight 3-d models from non-uniformly sampled point clouds without normal information," in *Symp. Geom. Process.*, 2006, pp. 41–50.
- [18] R. A. Newcombe *et al.*, "Kinectfusion: Real-time dense surface mapping and tracking," in *10th IEEE Int. Symp. Mixed Augmented Reality*, Oct. 2011, pp. 127–136.
- [19] L. Cheng, Y. Wu, L. Tong, Y. Chen, and M. Li, "Hierarchical registration method for airborne and vehicle lidar point cloud," *Remote Sens.*, vol. 7, no. 10, pp. 13 921–13 944, 2015.
- [20] F. Peng *et al.*, "Street view cross-sourced point cloud matching and registration," in *IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 2026–2030.
- [21] X. Huang, J. Zhang, Q. Wu, L. Fan, and C. Yuan, "A coarse-to-fine algorithm for registration in 3d street-view cross-source point clouds," in *Int. Conf. Digit. Image Comput.: Techn. Appl.*, Nov. 2016, pp. 1–6.
- [22] Q.-Y. Zhou and U. Neumann, "2.5 d dual contouring: A robust approach to creating building models from aerial lidar point clouds," *Proc. 11th Eur. Conf. Comput. Vision*, 2010, pp. 115–128.
- [23] Q. Shan, R. Adams, B. Curless, Y. Furukawa, and S. M. Seitz, "The visual turing test for scene reconstruction," in *Int. Conf. 3D Vision*, 2013, pp. 25–32.
- [24] K. Vanhoey *et al.*, "VarCity-the video: the struggles and triumphs of leveraging fundamental research results in a graphics video production," in *Proc. ACM SIGGRAPH Talks*, 2017, p. 48.
- [25] E. K. Forkuo and B. King, "Automatic fusion of photogrammetric imagery and laser scanner point clouds," *Int. Arch. Photogrammetry Remote Sens.*, vol. 35, no. 2004, pp. 921–926, 2004.
- [26] C. Früh and A. Zakhor, "Constructing 3d city models by merging aerial and ground views," *IEEE Comput. Graph. Appl.*, vol. 23, no. 6, pp. 52–61, Nov./Dec. 2003.
- [27] C. Frueh, S. Jain, and A. Zakhor, "Data processing algorithms for generating textured 3d building facade meshes from laser scans and camera images," *Int. J. Comput. Vision*, vol. 61, no. 2, pp. 159–184, 2005.
- [28] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1362–1376, Aug. 2010.

- [29] A. Bódis-Szomorú, H. Riemenschneider, and L. Van Gool, "Efficient edge-aware surface mesh reconstruction for urban scenes," *Comput. Vision Image Understanding*, vol. 157, pp. 3–24, 2017.
- [30] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [31] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, Feb. 2004.
- [32] A. Nurunnabi, G. West, and D. Belton, "Outlier detection and robust normal-curvature estimation in mobile laser scanning 3d point cloud data," *Pattern Recognit.*, vol. 48, no. 4, pp. 1404–1419, 2015.
- [33] R. B. Rusu and S. Cousins, "3d is here: Point Cloud Library (PCL)," in *proc. IEEE Int. Conf. Robot. Autom.*, Karlsruhe, Germany, vol. 47, pp. 1–4.
- [34] H. Butler, M. Loskot, P. Vachon, M. Vachon, and M. Vales, "libLAS: ASPRS LAS LiDAR data toolkit," vol. 15, 2011.
- [35] C. Wu, "Towards linear-time incremental structure from motion," in *Proc. Int. Conf. 3D Vision*, 2013, pp. 127–134.



Wei Li received the M.Sc. degree in applied mathematics from the College of Science, Jimei University, Xiamen, China, in 2013. He is currently working toward the Ph.D. degree in information and communication engineering with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen, China.

His current research interests include computer vision, 3D point cloud processing, registration of 3D data, object detection and extraction, and mobile laser scanning data processing.



Cheng Wang (M'07–SM'16) received the Ph.D. degrees in Communication and Signal Processing from the National University of Defense Technology (NUDT), Changsha, China, in 2002.

He is currently a Professor in the School of Information Science and Engineering, and Executive Director of Fujian Key Laboratory of Sensing and Computing for Smart City, Xiamen University. He is also the Chair of Working Group I/6 on Multi-sensor Integration and Fusion, in the International Society of Remote Sensing (ISPRS). He has coauthored more

than 150 papers in referred journals and top conferences including IEEE-TGRS, PR, IEEE-TITS, CVPR, AACL, and ISPRS-JPRS. His current research interests include point cloud analysis, multi-sensor fusion, mobile mapping and geospatial bigdata. He is a Fellow of IET.



Dawei Zai received the B.Sc. degree in aircraft design and engineering from Xian Jiaotong University, Xian, China, in 2012. He is currently working toward the Ph.D. degree in information and communication engineering with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen, China.

His current research interests include computer vision, machine learning, and mobile laser scanning point cloud data processing.



Pengdi Huang received the Ph.D. degree in information and communication engineering from the School of Information Science and Engineering, Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen, China, and the Department of Communication Engineering, Xiamen University, in 2017.

He is currently working with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Information Science and Engineering, Xiamen University. His current research interests include computer vision, 3D point cloud processing, object detection and extraction, and mobile laser scanning data processing.



Weiwan Liu received the B.S. and M.S. degrees in applied mathematics from the College of Science, Jimei University, Xiamen, China, in 2016. He is currently working toward the Ph.D. degree at the Department of Computer Science, Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen, China.

His current research interests include computer vision, machine learning, mobile laser scanning point cloud data processing, and Augmented Reality.



Chenglu Wen (M'14–SM'17) received the Ph.D. degree in mechanical engineering from China Agricultural University, Beijing, China, in 2009.

She is currently an Associate Professor with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Information Science and Engineering, Xiamen University, Xiamen, China. Her current research interests include 3D point cloud processing, machine learning, and robot vision.

Dr. Wen has authored or coauthored more than 40 research papers published in refereed journals and proceedings. She is the Secretary of the ISPRS WG I/6 on Multisensor Data Fusion (2016–2020), and the Associate Editor for IEEE-GRSL.



Jonathan Li (M'00–SM'11) received the Ph.D. degree in geomatics engineering from the University of Cape Town, Cape Town, South Africa, in 2000.

He is currently a Professor with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, School of Information Science and Engineering, Xiamen University, Xiamen, China. He is also a Professor and the Head of the WatMos Lab, Faculty of Environment, University of Waterloo, Waterloo, ON, Canada. His current research interests include information extraction from LiDAR point clouds and

from earth observation images.

Dr. Li has authored or coauthored more than 300 publications, over 130 of which were published in refereed journals, including IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, *ISPRS Journal of Photogrammetry and Remote Sensing*, *International Journal of Reliability and Safety*, *Photogrammetric Engineering and Remote Sensing*, and *Remote Sensing of Environment*. He is the Chair of the ISPRS Working Group I/2 on LiDAR-, Air-, and Spaceborne Optical Sensing (2016C2020), the Chair of the ICA Commission on Sensor-driven Mapping (2015C2019), and an Associate Editor for IEEE-TITS and IEEE-JSTARS.