AN EFFICIENTLY VOLUMETRIC FUSING METHOD FOR STRUCTURE-FROME-MOTION AND TERRESTRIAL POINT CLOUD

Wei Li[1], Cheng Wang[1], Dawei Zai[1], Pengdi Huang[1], Weiquan Liu[1], Jonathan Li[12]

[1]Fujian Key Laboratory of Sensing and Computing for Smart City, School of Information Science and Engineering, Xiamen University, Xiamen, Fujian361005,China
[2]GeoSTARS Lab, Department of Geography and Environmental Management, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

ABSTRACT

Airborne acquisition and ground-view 3D point cloud provide complementary 3D information at city scale. A complete but lacks ground-view details, while the latter is incomplete for higher floors and severe occlusion. In this paper, First, a volumetric fusion method based on graph cuts were applied for fusing of airborne and terrestrial 3D LiDAR data. Second, we propose a method of constraints based on the local centroid of point cloud to eliminate the gap of fusion boundary. Finally, the experiments show that the improved fusion algorithm implement blending effectively.

Index Terms—graph cut; volumetric fusion; boundary constraints

1. INTRODUCTION

With the rapid development of laser scanning technology, LiDAR 3D point clouds are becoming convenience. There has been a lot of research work based on the data , for instance, Automatic road vector extraction[1], Object detection[2], and Line segment extraction[3]. However, previous researches focus on information processing for the single data sources. And recently, structure-from-motion (SFM)[4] or Multi-View Stereo (MVS) [5] method have enabled the automated production of large-scale unban models from airborne imagery, and also have been applied to street-side mapping. However, such large-scale point cloud still often sparse especially on the ground view. Therefore, the generated 3D models still lack façade ground-level details due to occlusions and shadows. Some scholars proposed two methods mainly include SFM and mobile mapping of SFM. However, first, the acquired efficiency of SFM point cloud is lower than LiDAR systems. the density of point cloud is smaller than LiDAR 3D point cloud, and the precision of point cloud is lower than LiDAR 3D point cloud. Second, constrained by the reason that many places do not have roads, it is difficult for us to cover all the perspective of the 3D large-scale objects only base on street-side. 3D data based on LiDAR is density, high precision, and have no roofs, higher floors, and there is no traffic-free area such as courtyards for Terrestrial laser scanner. Therefore, Attempts to model and abstract buildings from such data to provide more detailed 3D information for large-scale 3D scene. It has practical application value that terrestrial LiDAR 3D data is prepared to provide complementary solution for both MVS and LiDAR moblile mapping methods. Therefore, both airborne and terrestrial LiDAR 3D data needs to be exploited in order to produce the next generation of large-city models. The results of method should be complete and more detailed.

Many approaches have been proposed to combine street-side and aerial data for joint mesh reconstruction. For examples, Fruh and Zakhor [6] construct meshes over street-side LiDAR range maps and over a larger-scale Digital Surface Model(DSM). However, they reconstruct a façade and an airborne mesh separately without topological fusion. There are rough and some gap at the results of fusion. Bódis-Szomorú A and Riemenschneider H etc.[7] build a 3DT on top of MVS points and do inside/outside classication of tetrahedra while enforcing line-of-sight and photo-consistency constraints. And then blending the two different 3D point clouds using graph cut. They had obtained some meanful results in their experiments. However, because of affecting of SFM algorithm, the method is Time-consuming and lower precision.

In this paper, we propose an improved solution to fuse airborne point cloud of large coverage but possibly low detail and a detailed but incomplete street-side point cloud in a volumetric fusion. In the alignment, the terrestrial LiDAR 3D data are aligned one-to-one to the objects SFM point cloud, which has corresponding geographical location, by using industrial software. In the blending step, inspired by reference[8], we also formulate this as a segment over the airborne point cloud and terrestrial LiDAR 3D data. constrained by using of normal vector and Euclidean distance, there are many wide gaps at the boundaries of the fusion results. Therefore, we propose a volumetric fusion based on boundary constraints to eliminate the gap.

The contributions of this paper are as follows: 1) we applied a volumetric fusion method based on graph cuts to eliminate the stratification of two kinds of point cloud aligned manually; 2) a modified volumetric fusion method based on boundary constraints was used to eliminate the gap, which existed at the result of fusion.

2. POINT CLOUD BLENDING

We formulate volumetric fusion as a segmentation over the airborne point cloud P, which assigns a binary label $l_i \in \{0,1\}$ to each point $p_i \in P$, and the points, being marked by 0, would be removed.

As the reference[8], we seek the binary labeling over P that minimizes

$$E^b(L) = \sum_{i:p_i \in P} E_i^b(l_i) + \lambda_b \sum_{ii_N} \psi(p_i, p_{i_N}) \cdot 1[l_i \neq l_{i_N}] \quad (1)$$

Where l_{i_N} is the $i\text{-}th$ adjacent point of labeling point p_i, 1 is indicator function, λ_b is regularization parameter. The unary penalties $E_i^b(0)$ for point p_i to obtain label l are defined as

$$E_i^b(l) = \begin{cases} 1-\phi_i & l=0 \\ \phi_i & l=1 \end{cases} \quad (2)$$

Given SFM point cloud P and the street-side point cloud Q, an airborne point p_i has street-side substitute $q_j \in Q$ which is supposedly better quality if q_j is the nearest neighbor of p_i in Q. The nearest substitute was assigned that if both the Euclidean distance d_{ij} between them and the angle θ_{ij} between their normal $n(p_i)$ and $n(q_j)$ are small. Constrained by Euclidean distance, the substitute when points $p_i \in P$ is near the outer boundary of street-side point cloud Q was inaccurately, as figure 3(b). Therefore, it is necessary to enlarge the corresponding distance such that the substitute become more impossible on the boundary of point cloud$P$. As fig.1, according to the characteristics of point cloud boundary, First, the nearest neighbor of p_i in the point cloud Q is searched and named q_j. And then, we compute the local centroid of p_i in the point cloud P, and local centroid of p_i in the point cloud Q, so the distanceδ_{ij}between two centroids is used to represent the distance between a point of cloud point P and the boundary of point cloud Q. At the end, we construct the weight factor ω_{ij} as formula (4), δ_0 is constant, which more than resolution factor of point cloud Q. Therefore, the likelihood for an airborne point to have a substitute can be formulate as

$$\varphi_i = \varphi(d_{ij}, \theta_{ij}) = e^{-\omega_{ij} d_{ij}^2/(2\sigma_b^2)} \bullet \max\{0, \cos\theta_{ij}\} \quad (3)$$

$$\omega_{ij} = \begin{cases} 1 & \delta_{ij} < \delta_0 \\ M & others \end{cases} \quad (4)$$

Where $\cos\theta_{ij} = n(p_i)^T n(q_j)$, and φ ranges from 0(no substitute) to 1(perfect substitute). σ_b is a blending parameter to control our notion of vicinity, which should incorporate deviations of P from Q due to alignment and reconstruction errors. Normal vectors are computed in P and Q separately by *K-NN* and least squares plane fitting, and by flipping normal according to relative perspective.

For a smooth segmentation, we define the influence between adjacent nodes of the *K-NN* graph over P as

$$\psi(p_i, p_{i_N}) = e^{(-d_{ii_N}/medd_{ii_N})} \quad (5)$$

Where d_{ii_N} is the distance between any two adjacent points p_i and p_{i_N}, and $medd_{ii_N}$ is the median of all *K-NN* distances in P.

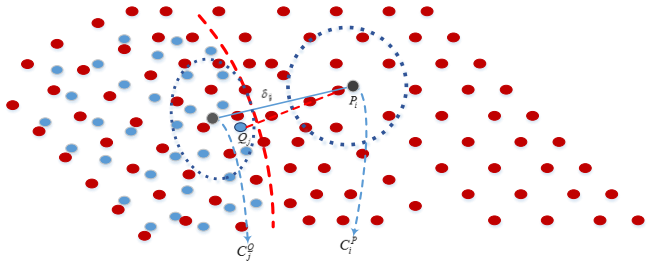

Fig.1. 2D illustration of boundary constraints. The blue points in the dashed box inside the ellipse is the K nearest points of point p_i, q_j is the nearest point of p_i, C_j^Q and C_i^P are the corresponding centroids.

3. EXPERIMENT

Our method relies on PCL library[12], liblas library[13] and the GCOptimization library[14] for graph-cuts in C++. We could not find any publicly available dataset with both airborne and terrestrial LiDAR 3D point cloud for the same geographic location. Therefore, We show experiments on our datasets, Haiyun Campus ($100 \times 210\, m^2$, Fig. 6(a).) captured in Xiamen, China. The airborne 3D data is generated by industrial software Acute3D, which is able to automatic generate 3D point cloud. The terrestrial LiDAR 3D point cloud is acquired via REIGL VZ1000.

3.1 PRE-PROCESSING AND ALIGNMENT

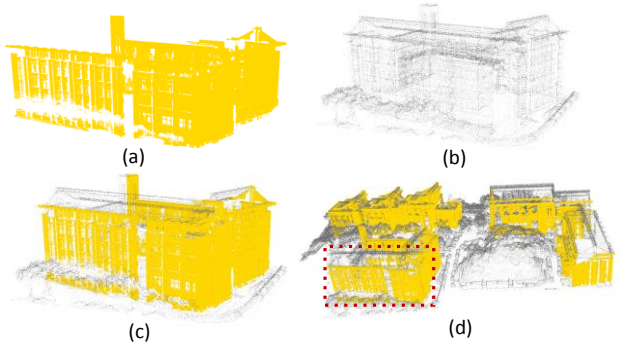

Fig.2. aligned results of the two different point clouds. (a) is registered result of 3D point clouds from the scanning muti-stations, (b) is the 3D point cloud from SFM, and (c) is registration result of the above two point clouds, (d) is the whole result of muti-buildings point clouds and airborne data.

Ground-view 3D point cloud can be got by variety of equipment and technology, Terrestrial laser scanner as REIGL VZ 1000, Z+F Imager 5010 etc. compared to SFM/MVS point cloud, Terrestrial 3D LiDAR data has higher precision and can be more efficiently accessed by

scanning. In our experiments, Terrestrial 3D LiDAR data was acquired by REIGL VZ-1000. The scanning objects such as buildings, were selected and segmented manually. And then, we eliminated outliers and align them each other manually, which has overlapped part, by industrial software. We got airborne data from industrial software smart 3D capture. The airborne data was used as baseline and registered it with terrestrial 3D LiDAR data. Fig.2 show the aligned results that every target building point cloud is aligned to the SFM point cloud. However, it exist stratification of point clouds (as show fig.5(a)) after aligning all the targets to airborne 3D point cloud, we have to eliminate stratification to implement fusion of two different point clouds.

3.2 BLENDING OF POINT CLOUD

We evaluate the effect of our modified fusion algorithm with boundary constraint compared to the algorithm of reference [8]. Our experiments were based on the labeling data. As Fig.3 (a), the two different point clouds has been registered, the size of grey one is 31423, the size of orange one is 8494, and the size of labeling set that should be substituted is 18830. We applied our method to fusing the two point clouds, and fix our parameters as $\sigma_b = 0.2$, $\delta_0 = 0.05$, $M = 100$ volumetric regularization parameter as $\lambda = 3.0$ Fig.3 (b) show that the boundary gap exists at the fusion results, however, our result Fig.3(c) eliminate the gap between the two point clouds. The corresponding quantitative results in table 1 show that our method eliminate the boundary gap of point cloud fusion.

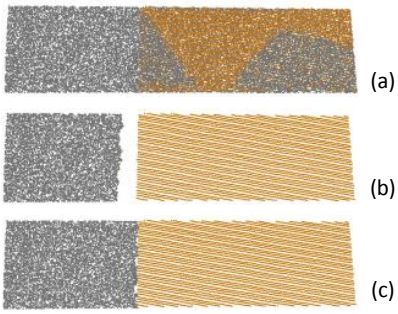

Fig.3.The elimination of the boundary gap. The top (a) show two registration 3D point cloud, and labeling overlapped point cloud with the orange data, the mid (b) is the result of reference[8], the below result is ours.

Table.1 statistics of eliminating overlapped points

	Labeled points	Filtered points	error
Ref[8]'method	18830	20438	1608
Our method	18830	18805	25

In order to enhance the details of buildings and improve its relative accuracy, we eliminate the airborne 3D data which is spare and deformation by constructing formulation, and minimize the energy function though graph cut.

As Fig.4, it obviously show that the airborne 3D data at Fig.4 (a) is spare and deformation. We take use of modified volumetric fusion procedure to segment the airborne 3D data which overlap with terrestrial 3D LiDAR data. We fix our regularization parameter as $\lambda = 3.0$, and minimize the energy function by graph cut. The segmented result is show at Fig.4 (b).and Fig.4(c) show our result of fusion.

Fig.5 shows the local result of fusion. From Fig. 5(a), we can observe that it exists stratification on the two registered point clouds. And after applying our method, the stratification is eliminated and fused each other. And Fig.6 shows the final fusion results.

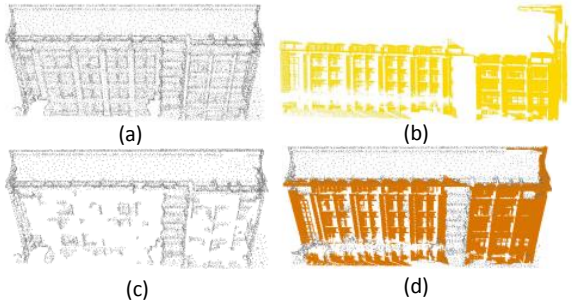

Fig.4. Output of our blending. (a) and (b) are 3D point clouds from SFM and REIGL VZ1000, and had been aligned, (c) is the segmented result by using graph cuts,(d) is the final fusion result.

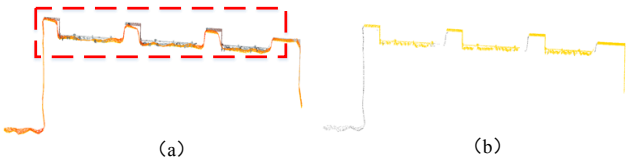

Fig.5. Elimination of point cloud stratification. (a) is the stratified demonstration of point cloud registration, (b) is the result of fusion.

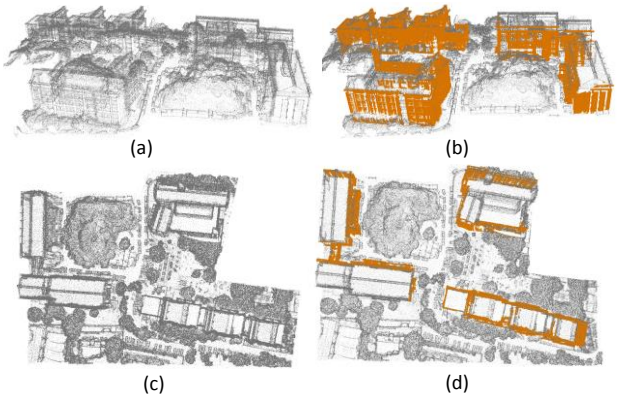

Fig.6. Final output of our blending at two views. The left column are same SFM point cloud, the right show fusion results from two different views.

4. CONCLUSION

The paper proposes an efficient method for fusion a 3D point cloud from a detailed, high precision but incomplete ground view and a complete but low-detail airborne point cloud. Our method provide a semi-automatic method to efficiently implement alignment and evaluate their accuracy. Thus, we also joins strengths of the data types by fusing them as a segmented problem of point cloud. Our detailed experimentation show good fusion quality and consuming

reasonable time. Applying the method at campus even city scale via terrestrial reconstructing data is part of our future work.

5. ACKNOWLEDGMENT

This work was support by grants form Natural Science Foundation of China (No.61371144 and No.U1605254), and Foundations of Fujian Higher Education (JA14292, JA15461).

6. REFERENCES

[1] Cheng W, Hassan T, El-Sheimy N, et al. Automatic road vector extraction for mobile mapping systems[J]. Int Arch Photogramm Remote Sens, 2008, 37(Part B3b): 515-521.

[2] Wang H, Wang C, Luo H, et al. Object detection in terrestrial laser scanning point clouds based on Hough forest[J]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(10): 1807-1811.

[3] Lin Y, Wang C, Cheng J, et al. Line segment extraction for large scale unorganized point clouds[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 102: 172-183.

[4] Heinly J, Schonberger J L, Dunn E, et al. Reconstructing the World* in Six Days*(As Captured by the Yahoo 100 Million Image Dataset)[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3287-3295.

[5] E. Tola, C. Strecha, and P. Fua. Efficient large scale multiview stereo for ultra high resolution image sets. *Machine Vision and Applications*, 23(5):903–920, 2011.

[6] Fruh C, Zakhor A. Constructing 3D city models by merging aerial and ground views[J]. IEEE computer graphics and applications, 2003, 23(6): 52-61.

[7] Bódis-Szomorú A, Riemenschneider H, Van Gool L. Efficient Volumetric Fusion of Airborne and Street-Side Data for Urban Reconstruction[J]. arXiv preprint arXiv:1609.01345, 2016.

[8] Fiocco M, Bostrom G, Gonçalves J G M, et al. Multisensor fusion for volumetric reconstruction of large outdoor areas[C]//Fifth International Conference on 3-D Digital Imaging and Modeling (3DIM'05). IEEE, 2005: 47-54.

[9] Zhou Q Y, Neumann U. 2.5 d dual contouring: A robust approach to creating building models from aerial lidar point clouds[C]//European conference on computer vision. Springer Berlin Heidelberg, 2010: 115-128.

[10] Musialski P, Wonka P, Aliaga D G, et al. A survey of urban reconstruction[C]//Computer graphics forum. 2013, 32(6): 146-177.

[11] Zhou Q Y, Neumann U. 2.5 d dual contouring: A robust approach to creating building models from aerial lidar point clouds[C]//European conference on computer vision. Springer Berlin Heidelberg, 2010: 115-128.

[12] http://pointclouds.org/.

[13] http://www.liblas.org/

[14] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. PAMI, 23(11):1222–1239, Nov 2001.