INTEGRATING SMARTPHONE IMAGES AND AIRBORNE LIDAR DATA FOR COMPLETE URBAN BUILDING MODELLING

Shenman Zhang [a], Jie Shan [b*], Zhichao Zhang [a*], Jixing Yan [a], Yaolin Hou [a]

[a] School of Remote Sensing and Information Engineering, Wuhan University, 430079, Wuhan, China –
(smzhang, zhichao, yanjx, houyaolin)@whu.edu.cn
[b] Lyles School of Civil Engineering, Purdue University, 47907, West Lafayette, IN, USA – jshan@purdue.edu

Commission V, WG V/4

KEY WORDS: Smartphone Images, Airborne LiDAR, Building Reconstruction, Registration

ABSTRACT:

A complete building model reconstruction needs data collected from both air and ground. The former often has sparse coverage on building façades, while the latter usually is unable to observe the building rooftops. Attempting to solve the missing data issues in building reconstruction from single data source, we describe an approach for complete building reconstruction that integrates airborne LiDAR data and ground smartphone imagery. First, by taking advantages of GPS and digital compass information embedded in the image metadata of smartphones, we are able to find airborne LiDAR point clouds for the corresponding buildings in the images. In the next step, Structure-from-Motion and dense multi-view stereo algorithms are applied to generate building point cloud from multiple ground images. The third step extracts building outlines respectively from the LiDAR point cloud and the ground image point cloud. An automated correspondence between these two sets of building outlines allows us to achieve a precise registration and combination of the two point clouds, which ultimately results in a complete and full resolution building model. The developed approach overcomes the problem of sparse points on building façades in airborne LiDAR and the deficiency of rooftops in ground images such that the merits of both datasets are utilized.

1. INTRODUCTION

Building reconstruction is a popular topic not only actively discussed in the research community (Brenner, 2005; Tack et al., 2012; Vosselman, 1999; Vosselman and Dijkman, 2001) but also in industry (Terrasolid, Smart3D Capture). As a technology capable of collecting 3D dense and accurate point clouds, airborne LiDAR has been widely used for building reconstruction in recent years. Although building rooftops extracted from airborne LiDAR point clouds are meeting the needs of more urban applications (Chen et al., 2012; Galvanin and Poz, 2012; Mongus et al., 2014), façade reconstruction remains to be a problem due to its sparse coverage or even complete occlusion. In fact, building façades are important and informative components for a complete building model, especially in virtual reality and navigation applications.

Recently, smartphones with built-in high quality cameras and other sensors, e.g., GPS and digital compass have become ever prevalent. Such images along with their sensor information can be acquired from image-based social media websites, such as Flickr and Instagram. These photo collections downloadable from the Internet are abundant, being cost free, up-to-date, and high resolution (Avrithis et al., 2010; Goesele et al., 2007; Snavely et al., 2010). State-of-the-art Structure-from-Motion (SfM) and Multi-View Stereo (MVS) reconstruction techniques (Agarwal et al., 2011; Crandall et al., 2013; Fonstad et al., 2013) allow us to process these images to reconstruct building façades in fine detail and precision. In contrast to airborne LiDAR point clouds, the building rooftops produced from ground images are coarse or even incomplete. As a result, the two sources are complementary to each other and their successful integration can lead to a complete building reconstruction.

This paper presents a solution to complete building reconstruction of both rooftops and façades by combining airborne LiDAR point clouds and ground image point clouds generated from smartphone images. To start the procedure, images are first clustered by geo-tags obtained from their metadata and then processed by SfM and MVS algorithms. LiDAR point cloud corresponding to the target building captured in the image set is found by using the GPS and compass info in the metadata. The two point clouds, one from airborne LiDAR data and one from ground smartphone images, are registered and combined based on the target building outline points visible to both data sets. A number of tests demonstrate that this approach can achieve a satisfactory reconstruction of building models with finer detail and higher completeness than using one dataset alone.

2. RELATED WORK

The proposed method consists of following major steps: 1. Generating ground image point clouds using SfM and MVS algorithms; 2. Corresponding building extraction from the airborne LiDAR point clouds based on the ground image point clouds; 3. Geo-registration of the ground image point clouds; 4. Matching building outlines respectively extracted from the LiDAR point clouds and ground image point clouds. A wealth of relevant researches for some of the above individual steps have been conducted. Below is a brief summary.

2.1 Structure-from-Motion and Dense Multi-View Stereo

Structure-from-Motion (Snavely et al., 2006) is a range imaging technique that recover camera parameters, pose estimates and sparse 3D scene geometry from 2D image sequences. It falls into two categories: incremental SfM and global SfM. Incremental SfM methods reconstruct target objects by recursively adding

---

* Corresponding authors: jshan@purdue.edu, zhichao@whu.edu.cn.

images to the point cloud generated by previous images (Agarwal et al., 2011; Goesele et al., 2007; Irschara et al., 2012). Global SfM approaches reconstruct the scene from all input images by estimating a global transform of all cameras so that the time complexity is smaller (Crandall et al., 2012). State of the art method is able to reconstruct cities consisting of 150K images in less than a day on a cluster with fine detail (Agarwal et al., 2011). MVS algorithms generate much denser point cloud of the object from the SfM results. A lot of dense multi-view stereo methods focus on the reconstruction of small objects. In the meantime, it has been proven that dense multi-view stereo methods are more adapted to larger scenes reconstruction, e.g., outdoor buildings with an initial sparse point cloud from SfM (Strecha et al., 2006; Gargallo and Sturm, 2005; Kolmogorov and Zabih, 2002). Furukawa and Ponce (2010) proposed an impressively accurate algorithm that generates dense set of reconstruction patches.

2.2 Building Extraction from Airborne Lidar Point Clouds

In classical methods, building extraction from airborne LiDAR data has been resolved by classifying the LiDAR points to different object types such as terrain, buildings, vegetation or others. Rottensteiner and Briese (2002) proposed a method to remove terrain point using a skew error distribution function and then separate building points from other points by analysing height differences. Sampath and Shan (2007) clustered the building cloud into planar segments, followed by their topologic reconstruction. Dorninger and Pfeifer (2008) proposed a comprehensive method for extracting rooftops from LiDAR point clouds. A target based graph matching approach was used to handle both complete and incomplete laser data (Elberink and Vosselman, 2009). Rau (2012) proposed a TIN-Merging and Reshaping roof model reconstruction algorithm. Sohn et al. (2012) generalized noisy polylines comprising a rooftop model by maximizing a shape regularity.

2.3 Geo-registration of Ground Image Point Clouds

Registering ground image point clouds to a map was often completed manually when geo-tags and camera pointing information were not available (Snavely et al., 2006). Today, the emergence and prevalent use of mobile cameras enable us to easily acquire images with embedded geo-tags and pointing information. Geo-tags and text labels are commonly used for geo-registration of models generated from ground images (Zamir and Shah, 2014; Zamir, 2014; Grzeszczuk et al., 2009), usually followed by an iterative refinement of aligning the models to building footprints. Furthermore, 3D models from Google Earth and building footprints from OpenStreetMap may also be used to improve geo-registration (Untzelmann et al., 2013; Wang et al., 2013).

2.4 Point Set Registration

Point set registration, aiming to find a spatial transform that aligns two point sets, is a key component in many computer vision and pattern recognition tasks. Because of its simplicity and low computational complexity, the Iterative Closest Point (ICP) algorithm is one of the most popular methods for rigid point set registration (Besl and Mckay, 1992). It gets globally optimal solution based on the criterion of closest distance and least-squares. However, the registration result of ICP depends on the initial values of the transform. Various probabilistic methods were developed to overcome this limitation of ICP (Luo and Hancock, 2001; Rangarajan et al., 1996). The Coherent Point Drift (CPD) algorithm (Myronenko and Song, 2010) considers the registration of two point sets as a probability density

estimation problem: one point set represents the Gaussian Mixture Model centroids and the other represents the data points. CPD can get accurate results with a decent performance.

3. METHODOLOGY

The proposed approach for both building rooftops and façades reconstruction consists of the following steps, as shown in Figure 1. First, the smartphone images are clustered by GPS data obtained from image metadata. Every building is reconstructed by SfM and MVS algorithms from a clustered smartphone image set. The corresponding LiDAR point cloud is extracted by using the GPS data and the image pointing information obtained from the clustered image set. Following this stage, two forms of top view 2D outlines of target building are respectively extracted from LiDAR point cloud and ground image point cloud. The two point clouds are then matched by using the building outlines. The target building is finally completely reconstructed in both rooftops and façades through the combined point cloud. This process can be repeated for every building within the study area to create a virtual scene.
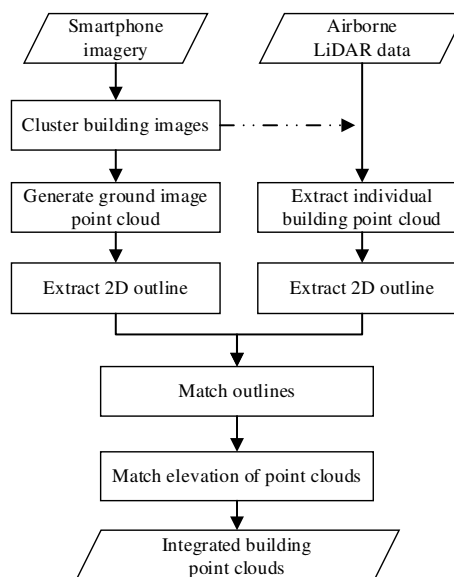
Figure 1. Workflow for the integration of airborne LiDAR data and ground smartphone imagery.

3.1 Building Reconstruction from Smartphone Imagery

Collection of smartphone imagery can be approximately clustered for individual buildings using geo-tags and image pointing information obtained from the Exchangeable Image File (Exif) format data. After that, we only need to handle each clustered image set for generating individual building point cloud in the proposed method.

Images captured by smartphones are lack of camera intrinsic parameters, pose and sensor response characteristics that are usually provided by calibrated cameras. Therefore, the image qualities are relatively poor and the distortions are significant. Geometric and radiometric calibrations of smartphone images are necessary before using them to reconstruct buildings. For smartphone images, camera models and approximate lens data are available from the Exif tags of JPG image files. Adobe Camera Raw, a commercial software tool, is utilized to correct geometric distortion using camera models and lens data (Adobe Camera Raw, 2015). Then, linear stretching is used to minimize colour difference between neighbouring images. This step is

necessary to assure both the reliability and precision of building reconstruction. Once much of the geometric and radiometric deformation of images have been calibrated, we proceed to full 3D reconstruction of the clustered smartphone imagery using the SfM algorithm, which is open source and developed by Agarwal et al. (2011). Sparse 3D point clouds and camera matrix are generated after the SfM process. Then a patch-based multi-view stereo algorithm (Furukawa and Ponce, 2010) is utilized to create dense ground image point clouds.

3.2 Building Point Cloud Extraction from Airborne LiDAR

The method presented in this paper reconstructs individual buildings one at a time. Once ground image point clouds have been generated, each corresponding airborne LiDAR point cloud of the target building is found and extracted by using GPS and imaging pointing information stored in the Exif tags when the images were taken. Although the precision of GPS readings in smartphones are generally about 10 meters in urban areas, it is sufficient for locating a candidate section of the LiDAR building point cloud.

For each image with a GPS coordinate $G_i(x_i, y_i)$, a circular buffer zone B_{G_i} can be created at G_i:

$$B_{G_i} = \{P_i| \parallel P_i - G_i \parallel \leq R\} \qquad (1)$$

Possible location where the image actually captured is within zone B_{G_i}. Extending the buffer zone B_{G_i} in a straight line L_i (Figure 2 (a)) for a certain distance D and adding all buffer zones, the point cloud for the target building is highly possible located in the buffer zone collection B_{L_i}, as grey area indicated in Figure 2 (b).

$$B_{L_i} = \{P_i| \parallel P_i - L_i \parallel \leq R\} \qquad (2)$$

Potential area of target building point cloud section has been found in buffer zone B_{L_i}, the exact building point cloud is extracted by classifying the point cloud in buffer zone B_{L_i}. Ground and vegetation points are eliminated by classifying the point cloud to building and non-building points. Building LiDAR points are segmented to possibly different individual buildings after processing by a region-growing algorithm on the building point class. The major part of building points occupies the largest area can be considered as point cloud of the target building, illustrated in Figure 2 (c), while the others may be points of adjacent ones.
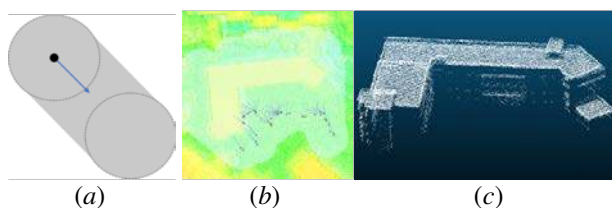

(a)   (b)   (c)

Figure 2. Extracting LiDAR building point clouds corresponding to ground image point cloud. (a) buffer zone of possible building LiDAR points from GPS position of a camera. (b) top view of elevation coloured LiDAR point cloud in the background. Grey translucent region around the target building (yellow region at the centre) shows the overlay buffer zones of all cameras. (c) extracted LiDAR building point cloud.

3.3 2D Building Outlines Extraction

Aligning Point Cloud to Up-right direction. The LiDAR point cloud and the ground image point cloud have common 2D building rooftop outlines which can be used for their registration. To generate a precise up-right 2D building outlines, ground image point cloud must be aligned to the up-right vector at first. Let N_{p_i} represents the normal of point p_i in façade ground image point cloud $\{P\}$. Similar to the method proposed by Untzelmann et al. (2013), we obtain the initial up-right direction N_0 by calculating the normal of a plane fitted from camera positions. After removing noise points p_i where the angle between its normal and N_0 is larger than a certain threshold, we apply a RANSAC-based approach to refine the up-right vector N_{up} by iteratively selecting two points p_i and p_j from the façade points and estimating the cross product of N_{p_i} and N_{p_j}:

$$N_{up} = RANSAC(N_{p_i} \times N_{p_j}) \quad (p_i, p_j \in P, i \neq j) \qquad (3)$$

The 2D building outlines can be acquired by projecting the point cloud to the horizontal plane after aligning ground image point cloud to up-right direction. The 2D outlines of LiDAR point cloud is extracted in two steps. First, rooftops and façades are clustered by analysing the normal vectors. Second, 2D building outlines corresponding to the one extracted from ground image point cloud are extracted by projecting the façades to the horizontal plane, as shown in Figure 3.

Ground image point cloud   *up-right 2D outlines*

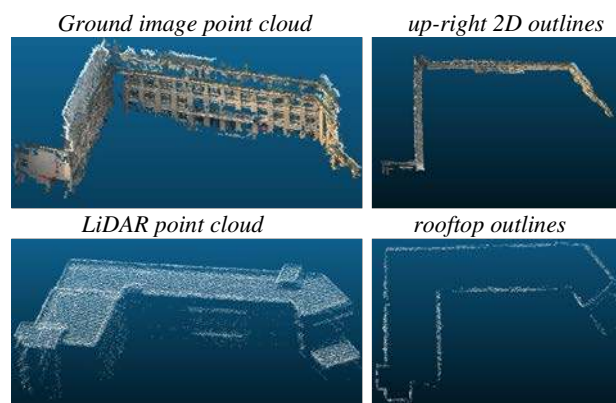*LiDAR point cloud*   *rooftop outlines*

Figure 3. Extracting building outlines of ground image point cloud (top) and LiDAR point cloud (bottom).

Extracting Visible Part of 2D Building Outlines. Ground image point cloud covers only the visible part of building façades from camera positions. Mismatch of 2D building outlines extracted from the two point clouds may occur due to repetitive structures of the building if the entire 2D building outlines from the LiDAR point cloud are involved in the matching process.

In order to work around this problem, image pointing information, normal of façade points and Field of View (FOV) are utilized to extract the visible part of 2D building outlines from LiDAR data. First, potentially visible points in outlines are extracted by searching the area in the FOV from individual camera position. Then, the visibility of each façade point is checked by using the angle between the normal vector and the image pointing vector, illustrated in Figure 4.
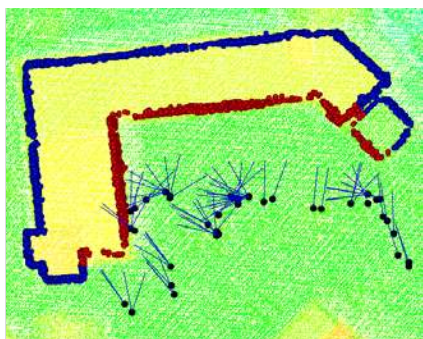
Figure 4. Extraction of visible part of the building outlines by using image pointing information, normal and FOV. Red points show the visible part corresponding to cameras and blue points show the invisible part.

3.4 Integration of LiDAR and Ground Image Point Clouds

Once the building outlines of the two point clouds are matched, LiDAR point cloud and ground image point cloud can be matched by the same transform. In order to precisely register the two 2D building outlines generated in different coordinates systems, we propose a two-step solution.

Coarse Integration by Matching Camera Positions. Two sets of camera position data have been acquired: A. GPS data of camera positions from Exif tags; B. Camera positions from SfM result. Although the former is erroneous about 10 meters, it is sufficient for coarse registration of the two point sets. The combination can be combined by a 7-parameter similarity transform $\mathbf{Trans} = (R, S, T)$. An iterative procedure is adopted to minimize the errors between the two point sets:

$$e = \sum_{i=1}^{n} \omega \parallel S(Rp_i + T) - q_i \parallel$$

$$\omega = 1 - \frac{|angle(D_{p_i} - D_{q_i})|}{\pi} \tag{4}$$

p_i : camera position from the GPS data set;
D_{p_i} : pointing vector of camera p_i;
q_i : camera position from the SfM result;
D_{q_i} : pointing vector of camera q_i.

The scale of ground image point cloud is changed to be similar to that of the LiDAR point cloud after coarse integration (shown in Figure 5) so that the accurate registration in the next step can converge.
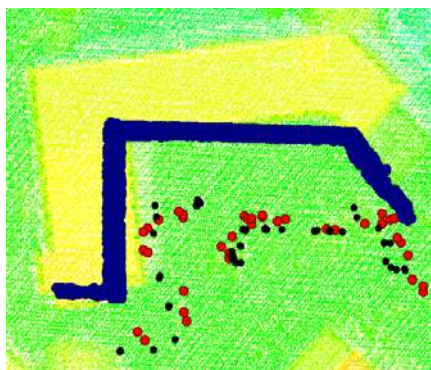

Figure 5. Coarse integration of LiDAR point cloud (elevation coloured in the background) and ground image point cloud (blue dots) by matching camera positions respectively from SfM (black dots) result and GPS data (red dots).

Accurate Integration by Registering Building Outlines. In the previous step, the ground image point cloud has been approximately matched with the LiDAR point cloud and converted into the world coordinate system by registering the geo-tags of images (global coordination) and the camera positions from SfM (local coordination). But due to deficient accuracy of GPS from smartphones, the matching is not accurate enough to meet the need of building reconstruction.

In order to accurately integrate the LiDAR point cloud and ground image point cloud, we adopt the Coherent Drift Point (CPD) algorithm (Myronenko and Song, 2010) to match building outlines extracted respectively from LiDAR point cloud and ground image point cloud. CPD represents the centroids of one point set using a Gaussian mixture model and aligns them with another set of data points. It allows points to displace at all degrees of freedom, and makes the displacement vectors point to the similar directions by regularizing high frequency components of the displacement field.

The two point clouds are accurately registered in 2D space after matching 2D outlines using the CPD algorithm, however, elevation is missing during this process. In order to accurately register LiDAR point cloud and ground image point cloud in 3D space, the median point of building height is calculated respectively from LiDAR point cloud and ground image point cloud. Then, the two point clouds are matched on the z axis by aligning the median points. Due to the incompleteness of the point clouds, errors may exist in the calculated median points and previous matching on the z axis might not be ignorable. When this happens, we manually move the ground image point cloud on the z axis at a distance to make sure the two point clouds integrated at a proper height. After that, the LiDAR point cloud and ground image point cloud are finally integrated, as illustrated in Figure 6.

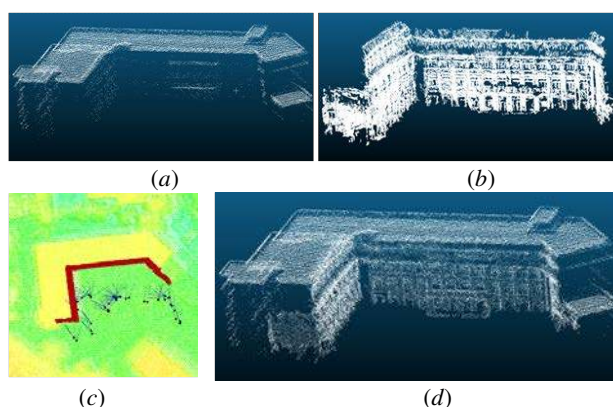

(a)   (b)

(c)   (d)

Figure 6. Accurate integration of LiDAR and ground image point clouds by matching building outlines. (a) LiDAR point cloud. (b) ground image point cloud. (c) top view of integrated result. (d) forward view of integrated result.

4. EXPERIMENTS AND RESULTS

The test datasets were processed with the methods described in the previous section. Several different buildings on Wuhan University campus are reconstructed. Integrated results from two of the buildings will be presented. Provided with global coordinate information, the original airborne LiDAR data cover about 4 square kilometres of the campus with a point density of 10 pts/m^2. A total of 582 smartphone images having GPS coordinate and digital compass data are captured in the test area. Through clustering analysis, 78 are clustered for a student dorm and 93 for an instruction building. Ground image point clouds are

generated by the clustered image sets after geometric and radiometric calibrations, as shown in Figure 7.

93 images

78 images
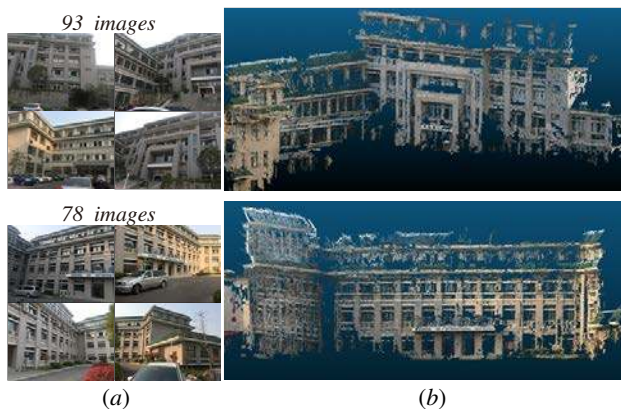
        (*a*)   (*b*)

Figure 7. Clustered building images and ground image point clouds for a student dorm (top, 93 images) and an instruction building (bottom, 78 images). (a) clustered building images; (b) ground image point clouds.

A RANSAC-based method is utilized to acquire the up-right direction vector of the ground image point cloud by using the normal of façade points. Then, 2D building outlines can be obtained by projecting the adjusted point cloud to the horizontal plane. By using the image geo-tags, individual building LiDAR point clouds of the student dorm and the instruction building can be extracted by the method described in section 3.2, as illuminated in Figure 8. In the next step, the building LiDAR point cloud is classified to rooftops and façades. Visible 2D outlines of buildings are extracted by projecting the façade point cloud to the horizontal plane and check the visibility of façade points using the camera pointing data and camera FOV information, shown in Figure 9.

.
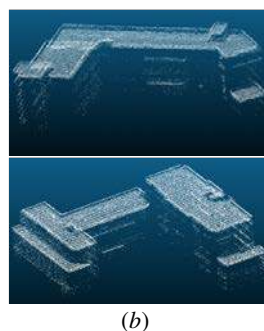

        (*a*)   (*b*)

Figure 8. Airborne LiDAR data of the test area. (a) is part of the study area. (b) is the point clouds for two buildings. Top: the student dorm, bottom: the instruction building.

In order to solve the problem of different scales in the two types of datasets, a 7-parameter similarity transform is utilized to approximately convert them into the same coordinate system by matching camera positions from geo-tags and SfM results. Next, outlines are accurately registered using the CPD algorithm. Using the transform matrix generated by the CPD algorithm in the previous step, LiDAR point cloud and ground image point cloud can be matched in 2D space.

After matching the median height point of the two point clouds and manually moving ground image point cloud along z axis to a proper height, point clouds are finally integrated in 3D space. Integrated results of the test buildings are shown in Figure 10.
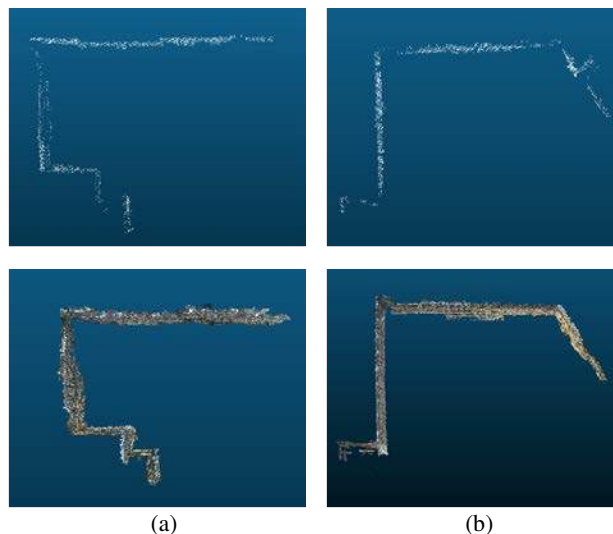

        (a)   (b)

Figure 9. Two visible building outlines extracted from airborne LiDAR (top) and ground image point clouds (bottom).
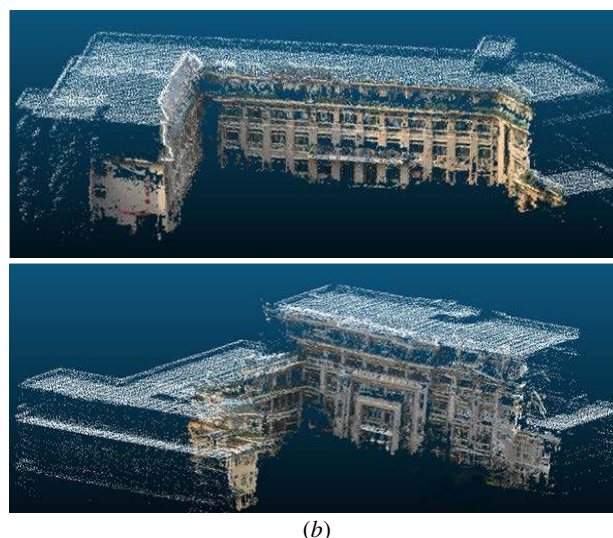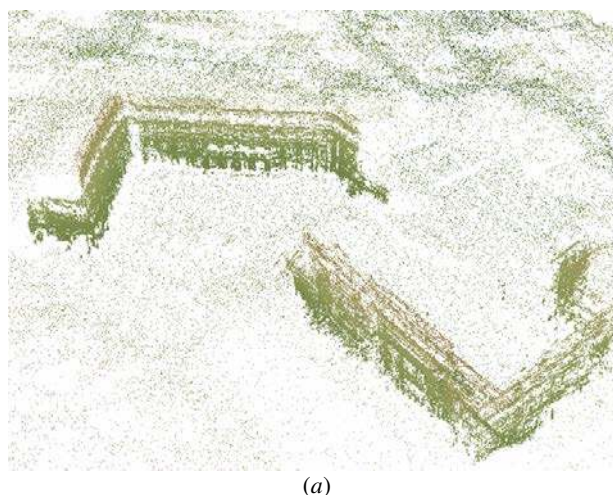

(*a*)

(*b*)

Figure 10. Integrated point clouds. (a) the two buildings within airborne LiDAR point cloud. (b) the student dorm (top) and instruction building (bottom).

5. CONCLUSIONS

This paper has presented a novel method for complete reconstruction of both façades and rooftops of buildings by integrating ground smartphone images and airborne LiDAR data. Building outlines respectively from the LiDAR point cloud and the ground image point cloud are used to match the two point clouds, leading to a complete and full resolution building model. Taking full advantage of geo-tags and digital compass information obtained from smartphone images, the developed approach, which is highly efficient and low-cost, overcomes the issue of sparsity of points on façades in airborne LiDAR data and the deficiency of rooftops in ground images. Test buildings reconstructed by our method are much higher in completeness and finer in detail than using airborne LiDAR data alone. Most steps in the proposed method can be implemented automatically. Since buildings are reconstructed one by one in this method, building reconstruction for a large area can be achieved by performing a similar process to all individual buildings. In future work, we intend to acquire more images of buildings in the test area and achieve large scale reconstruction.

Our methods achieve desired results on buildings with complex outlines. But it may not be effective for compact buildings with a matchbox shape outlines due to lack of junction points for outlines matching. Due to the lack of building ground truth, no effective method is found to accurately evaluate the integrated results at this time. Most steps in our methods are automated except for the step of recovering elevation of the buildings. For future work, we plan to automatically recover elevation by matching line features extracted respectively from LiDAR data and images. Building model reconstruction from the integrated point cloud is also the issue that should be addressed in the next step.

ACKNOWLEDGEMENTS

REFERENCES

Adobe Camera Raw. *https://helpx.adobe.com/camera-raw*.

Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S.M. and Szeliski, R., 2011. Building rome in a day. *Communications of the ACM*, *54*(10), pp.105-112.

Avrithis, Y., Kalantidis, Y., Tolias, G. and Spyrou, E., 2010. Retrieving landmark and non-landmark images from community photo collections. In *Proceedings of 18th the ACM international conference on Multimedia*, pp. 153-162.

Besl, P.J. and McKay, N.D., 1992. Method for registration of 3-D shapes. In *Robotics-DL tentative* (pp. 586-606). International Society for Optics and Photonics.

Brenner, C., 2005. Building reconstruction from images and laser scanning. *International Journal of Applied Earth Observation and Geoinformation*, *6*(3), pp. 187-198.

Chen, D., Zhang, L., Li, J. and Liu, R., 2012. Urban building roof segmentation from airborne lidar point clouds. *International journal of remote sensing*, *33*(20), pp. 6497-6515.

Crandall, D., Owens, A., Snavely, N. and Huttenlocher, D., 2011. Discrete-continuous optimization for large-scale structure from motion. In *Computer Vision and Pattern Recognition, IEEE Conference on*, pp. 3001-3008.

Dorninger, P. and Pfeifer, N., 2008. A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. *Sensors*, *8*(11), pp. 7323-7343.

Elberink, S.O. and Vosselman, G., 2009. Building reconstruction by target based graph matching on incomplete laser data: Analysis and limitations. *Sensors*, 9(8), pp. 6101-6118.

Fonstad, M.A., Dietrich, J.T., Courville, B.C., Jensen, J.L. and Carbonneau, P.E., 2013. Topographic structure from motion: a new development in photogrammetric measurement. *Earth Surface Processes & Landforms*, 38(4), pp. 421-430.

Furukawa, Y. and Ponce, J., 2010. Dense 3D motion capture from synchronized video streams. In *Image and Geometry Processing for 3-D Cinematography*, pp. 193-211.

Galvanin, E.A.D.S. and Poz, A.P.D., 2012. Extraction of building roof contours from LiDAR data using a Markov-random-field-based approach. *Geoscience and Remote Sensing*, IEEE Transactions on, 50(3), pp. 981-987.

Gargallo, P. and Peter S., 2005. Bayesian 3D modelling from images using multiple depth maps. In *Computer Vision and Pattern Recognition, IEEE Conference on*, pp. 885-891.

Goesele, M., Noah, S., Brian, C., Hugues, H., and Steven, M., 2007. Multi-view stereo for community photo collections. In *Computer Vision, ICCV. IEEE 11th International Conference on*, pp. 1-8.

Grzeszczuk, R., Košecka, J., Vedantham, R. and Hile, H., 2009. Creating compact architectural models by geo-registering image collections. In *Computer Vision Workshops (ICCV Workshops), IEEE 12th International Conference on*, pp. 1718-1725.

Irschara, A., Christopher, Z., Manfred, K., and Horst, B., 2012. Large-scale, dense city reconstruction from user-contributed photos. *Computer Vision and Image Understanding*, 116(1), pp. 2-15.

Kolmogorov, V. and Ramin, Z., 2002. Multi-camera scene reconstruction via graph cuts." In *Computer Vision, ECCV*, pp. 82-96.

Luo, B. and Edwin, R. H., 2001. Structural graph matching using the EM algorithm and singular value decomposition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(10), pp. 1120-1136.

Mongus, D., Lukač, N. and Žalik, B., 2014. Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, pp. 145-156.

Myronenko, A. and Song, X., 2010. Point set registration: Coherent point drift. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(12), pp. 2262--2275.

Rangarajan, A., Haili, C., Eric, M., Suguna, P., Lila, D., Patricia, G., and James, D., 1997. A robust point-matching algorithm for

autoradiograph alignment. *Medical Image Analysis* 1(4), pp. 379-398.

Rau, J., 2012. A line-based 3D roof modell reconstruction algorithm: TIN-merging and reshaping (TMR). In *Proceedings of the ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences Conference*, 25, pp. 287-292.

Rottensteiner, F. and Briese, C., 2002. A new method for building extraction in urban areas from high-resolution LIDAR data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(3/A), pp. 295-301.

Sampath, A. and Shan, J., 2007. Building boundary tracing and regularization from airborne LiDAR point clouds. *Photogrammetric Engineering & Remote Sensing*, 73(7), pp. 805-812.

Smart 3D Capture. *http://www.acute3d.com*.

Snavely, N., Steven, M., and Richard, S., 2006. Photo tourism: exploring photo collections in 3D. In *ACM transactions on graphics (TOG)*, 25(3), pp. 835-846

Snavely, N., Ian S., Michael, G., Richard, S., and Steven, M., 2010. Scene reconstruction and visualization from community photo collections. *Proceedings of the IEEE*, 98(8), pp. 1370-1390.

Sohn, G., Jwa, Y., Kim, H.B. and Jung, J., 2012. An implicit regularization for 3D building rooftop modeling using airborne LIDAR Data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1(3), pp. 305-310.

Strecha, C., Rik, F., and Luc V.G, 2006. Combined depth and outlier estimation in multi-view stereo. In *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2, pp. 2394-2401.

Tack, F., Buyuksalih, G. and Goossens, R., 2012. 3D building reconstruction based on given ground plan information and surface models extracted from spaceborne imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, pp. 52-64.

Terrasolid. *http://www.terrasolid.com*.

Untzelmann, O., Sattler, T., Middelberg, S. and Kobbelt, L., 2013. A scalable collaborative online system for city reconstruction. *Computer Vision Workshops (ICCVW), IEEE International Conference on*, pp. 644-651.

Vosselman, G., 1999. Building reconstruction using planar faces in very high density height data. *International Archives of Photogrammetry and Remote Sensing*, 32(3), pp. 87-94.

Vosselman, G. and Dijkman, S., 2001. 3D building model reconstruction from point clouds and ground plans. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(3), pp. 37-44.

Wang, C., Keith, W., and Noah, S., 2013. Accurate geo-registration of point clouds using geographic data. In *3D Vision-3DV, International Conference on*, pp. 33-40.

Zamir, A.R., 2014. *Visual Geo-localization and Location-aware Image Understanding* (Doctoral dissertation, University of Central Florida Orlando, Florida).

Zamir, A.R. and Mubarak, S., 2014. Image Geo-Localization Based on Multiple Nearest Neighbour Feature Matching Using Generalized Graphs. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(8), pp. 1546-1558.