





3D Facade Reconstruction Using the Fusion of Images and LiDAR: A Review

Haotian Xu^(✉)  and Chia-Yen Chen 

Department of Computer Science, The University of Auckland,
Auckland 1010, New Zealand
hxu454@aucklanduni.ac.nz

Abstract. Three-dimensional (3D) urban reconstruction becomes increasingly crucial in many application areas, such as entertainment, urban planning, digital mapping. To achieve photorealistic 3D urban reconstruction, the detailed reconstruction of building facades is the key. Light Detection and Ranging (LiDAR) point clouds and images are the two most important data types for 3D urban reconstruction, which are complementary regarding data characteristic. LiDAR scans are sparse and noisy but contain the precise depth data, whereas images can offer the color and high-resolution data but no depth information. In recent years, an increasing number of studies show that the fusion of LiDAR point clouds and images can attain better 3D reconstruction results than a single data type. In this paper, we aim to provide a systematic review of the research in the area of the 3D facade reconstruction based on the fusion of LiDAR and images. The reviewed studies are classified by the different usage of images in the reconstruction process. We hope that this research could help future researchers have a more clear understanding of how existing studies leverage the data in LiDAR scans and images and promote more innovations in this area.

Keywords: Data fusion · Point cloud · Optical image · LiDAR · Urban reconstruction · Facade reconstruction · 3D reconstruction · Registration · Sensor fusion · Laser scanning

1 Introduction

Three-dimensional (3D) urban reconstruction is a significant topic with commercial and intellectual values [13], which has a great diversity of applications, such as traffic planning, visualization for navigation, virtual tours, utility management, civil engineering, and crisis management [2]. Therefore, in recent years, there is an increasing demand for the 3D reconstruction of photorealistic urban building models [21]. As building facades are the essential part of urban buildings, the detailed facade reconstruction is especially crucial for photorealistic 3D urban reconstruction.

Light Detection and Ranging (LiDAR) point clouds and images are the two main kinds of data used for 3D reconstruction. Many studies which aim at the automatic generation of 3D models based on LiDAR and photographs have been conducted in computer vision, photogrammetry, and computer graphics communities [20]. Nevertheless, images have a long history and are usually captured by different kinds of cameras, while the LiDAR point cloud is a new 3D data type [25]. LiDAR devices acquire the range data of target objects through the time-of-flight of lasers [4]. In general, LiDAR devices can be classified by resolution level. High-resolution LiDAR can generate dense point clouds but are slow and have small working volumes, while the low-resolution LiDAR is fast and easy to use but usually generate noisy and sparse data points [9]. In the last two decades, most of the studies for facade reconstruction are based on either of these two data types, including image-based methods [12, 22, 23], and LiDAR-based methods [5, 15, 18]. Typically, images have high resolution and color information but lack 3D data, while LiDAR point clouds have the demerits like noise, sparsity, and the lack of color information but naturally contain precise 3D data [20]. Thus, the characteristics of LiDAR point clouds and images are complementary. Moreover, an increasing number of researchers recently reported that the fusion of LiDAR scans and photographs could have a better performance in many different kinds of applications than a single data type [25].

There have already been a few surveys related to 3D urban reconstruction [3, 7, 13, 19, 20]. However, the systematic review for the studies on the facade reconstruction using the fusion of LiDAR data and images is still rare. Therefore, this paper aims to fill this research gap. In all the studies that we reviewed, we found that LiDAR point clouds are only used to reconstruct facade structures, but images can be used for different purposes. Therefore, we classify these studies into two groups by the different application purposes of images. The first category only uses images to texture the 3D model generated from LiDAR point clouds, while the second category uses images not only for texturing 3D facade models but also for assisting the reconstruction of facade structures in point clouds. In Tables 1 and 2, we provide a quick reference to the relevant studies of the two categories. In the rest of this paper, we will introduce the two classes of methods respectively.

Table 1. Methods only using 2D images for texturing.

Study	Reconstruction scale	Features used for 2D-3D registration
[16]	Individual building	Sets of 2D and 3D linear features
[17]	Individual building	Clusters of 3D and 2D linear features
[10]	Individual building	3D rectangular parallelepipeds and 2D rectangles
[11]	Individual building	2D and 3D linear features
[26]	City	Geo-referenced information
[6]	City	Pre-calibration of 2D and 3D sensors

Table 2. Using 2D images for texturing based on 2D-3D registration.

Study	Reconstruction scale	Approach to assisting the reconstruction of 3D facade structures
[24]	Individual building	Combining 2D and 3D facade linear features in 3D space
[14]	Individual building	Combining 2D and 3D facade linear features in 2D space and projecting the refined 2D linear feature back to 3D space
[9]	Individual building	Fusing 2D and 3D data to detect the main planes with different depth of facades
[1]	Individual building	Using 2D data to reconstruct 3D window crossbars
[8]	City	Fusing 2D and 3D data to determine the shape grammar of facades

2 Using Images Only for Texturing Process

In this group of studies, usually 3D facade structures are first generated from only the point cloud captured by LiDAR, and then the texture is mainly produced by the registration of 2D images and 3D point clouds. Usually, this kind of approaches utilizes high-resolution LiDAR to collect the depth data as LiDAR is the only data source used for reconstructing facade structures. This group of methods can be divided into two classes. The first category has no requirement that the relative position and orientation of the camera and LiDAR should be fixed, while the second category has such a requirement.

Both classes of methods have advantages and disadvantages. The first category of methods has the complete flexibility for capturing 2D and 3D data, which makes the captured data more complete in comparison with the second category of methods when the placement of 2D and 3D sensors are constrained by geographical conditions. However, this kind of methods may require different 2D-3D registration methods depending on different facade structure, which makes it relatively difficult to be applied to large-scale urban reconstruction. Therefore, the first kind of methods usually focuses on the reconstruction of individual buildings [10, 11, 16, 17]. In contrast, since the relative position and orientation of 2D and 3D sensors are fixed, and 2D and 3D sensors are pre-calibrated, the registration of 2D and 3D data are quite easy for the second kind of methods. Thus, the second one is proper for large-scale urban reconstruction [6, 26]. Nevertheless, for this kind of methods, the flexibility of the data capturing process and the completeness of data are sacrificed in some particular situations [11]. In the rest of this section, we will introduce the two classes of methods respectively.

2.1 Methods Using LiDAR and Cameras with Unfixed Relative Position and Orientation

[16] first proposes a method for the photorealistic reconstruction of urban buildings using unfixed LiDAR and cameras. The method mainly utilizes the corresponding linear features detected in both 2D and 3D data for 2D-3D registration. 3D linear features are extracted from the intersection of the planar regions segmented from point clouds, and 2D linear features are extracted by edge detection in images. Based on the registration result, 3D building models are textured by using 2D images. The authors then proposed another slightly different registration approach based on the clusters of 3D and 2D lines instead of the sets of 3D and 2D lines [17].

Based on the previous work [17], some updates were then made in [10]. One of the critical updates is that the clusters of the higher-level 3D and 2D features, i.e., the vertical or horizontal 3D rectangular parallelepipeds extracted from 3D point clouds and the 2D rectangles acquired from 2D images, are used for 2D-3D registration. The authors stated that the use of such higher-level features is because of the large search space which makes the matching of 3D individual lines and 2D individual lines almost impossible, and the inexistence of the corresponding 2D lines of some 3D lines in 2D data or the inexistence of some corresponding 3D lines of 2D lines in 3D data. However, the authors then proposed a new method for 2D-3D registration which utilizes only linear features instead of clusters of significantly grouped linear features [11]. This approach employs a more efficient algorithm for achieving the faster matching process of linear features.

2.2 Methods Using LiDAR and Cameras with Fixed Relative Position and Orientation

There are relatively few papers using rigidly mounted cameras and LiDAR for urban reconstruction. This kind of methods often is used for large-scale urban reconstruction. Generally, the methods use a car with rigidly mounted LiDAR and cameras to collect a large number of 2D and 3D data of urban environment. 3D facade models usually are first reconstructed by 3D point clouds. Then, the 3D models are textured by using geo-referenced information [26] or the pre-calibration of 2D and 3D sensors [6].

3 Using 2D Images for both Texturing Process and Assisting the Reconstruction of 3D Facade Structures

As mentioned before, the point clouds produced by LiDAR generally have problems including sparsity, noisiness, and missing data. Therefore, some other papers about the building facade reconstruction based on the fusion of LiDAR and images utilize 2D images to enhance 3D point clouds. Accurate facade features, like linear features, can be extracted from images and then used to consolidate

the structure of 3D facade models [9,14,24]. Besides, images can provide the detailed information of facade elements which LiDAR can hardly capture, such as the crossbar of windows [1]. In addition, 2D images can also be used for texturing 3D facade models.

Linear features are the most significant component in the facade structure of many different kinds of buildings and can be relatively easily extracted from 2D images. [24] proposes a 3D reconstruction method for the building facade whose structure is mainly composed of straight lines. First, the pre-processing of the 3D point cloud and 2D image are executed for filtering the noise and outliers of the 3D data points, detecting the target building, and registering 3D point clouds and 2D images. Then, straight lines existing in facade structures are extracted from the 2D space of photographs and projected to the 3D space of LiDAR point clouds. Finally, these projected 3D lines are employed for consolidating the corresponding feature lines extracted in point clouds.

In [14], a similar method which also employs the linear features extracted from 2D images to refine the 3D facade model produced from LiDAR data is introduced. The main difference regarding the approach to 2D-3D fusion between the paper and [24] is the space used for matching and enhancing the linear features of facade structures. In [24], 2D linear features are projected to the 3D space to directly enhance the 3D linear features of the point cloud, whereas this approach projects 3D linear features to the 2D space for the matching and consolidation process. Thus, once the projected linear features are improved in the 2D space, they will be projected back to the 3D space for completing the 3D model.



Fig. 1. The generated depth-layers of a building facade (differently colored) [9].

2D images can enhance not only the linear features of 3D point clouds but also planar features. An approach to reconstructing the building facade with large-scale repetitions is introduced in [9]. The decomposition of the planes with different depths (depth-layers) of building facades in 2D images (Fig. 1) is the core of this method. This is achieved by assigning the depth values obtained from each part of facades in 3D point clouds to the corresponding part of facades in

images. Once the depth-layers in 2D images are extracted, the self-symmetries in facade structures can be recognized and used for model texturing and handling the missing data in point clouds.

Furthermore, 2D images can be used to capture elements which may be missed out by the LiDAR since images usually have higher resolutions. In [1], terrestrial LiDAR scans and photographs are used to reconstruct the different levels of details of building facades. Since it is hard to capture the accurately detailed structure inside windows by using LiDAR, images are used for reconstructing the small structures inside windows like windows frames and windows crossbars.

Moreover, the fusion of images and LiDAR can be used for assisting the determination of a specific kind of facade style. This is another way to use images for assisting the reconstruction of facade structures. In [8], a workflow used for the automatic reconstruction of the 80% of buildings in the city of Graz, Austria is introduced. As Graz has plenty of different kinds of complex building styles, many grammar templates of building styles are pre-generated for guiding the feature detection. First, the fusion of images and LiDAR data are used to generate the grammar representation of facades. Then, the corresponding grammar template of a facade is found by matching its grammar representation against all the templates. In this research, the key to reconstructing building facades is to get the corresponding shape grammars by processing the combination of the detected features from orthophotos, the segmented plane regions from depth images, and the corresponding shape grammar template.

4 Conclusion

This paper presents a comprehensive systematic review of the research on the 3D facade reconstruction based on the fusion of LiDAR and images. It can be seen from our review that the fusion 2D and 3D sensors is able to reconstruct high-quality textured 3D building facade models. Also, most of the studies in the early stage of this area only utilize images for texturing purpose. However, most of the subsequent studies focus on using images for both the refinement of facade structures in 3D point clouds and the texturing process. We believe that this trend is reasonable and promising for the photorealistic 3D reconstruction of building facades.

Currently, most of the studies in this areas aim to reconstruct the building facades with regular or straightforward structures, such as the one mainly composed of straight line and planes. However, if the building facade which needs to be reconstructed contains more complicated structures, like the highly decorated neo-classical facades in [8], such direct reconstruction based on refinement and texturing would be quite challenging. Shape grammar is a potential solution for this kind of situation. However, this method is not efficient and generic, especially in the case that there are a large number of various elaborate building facades to be reconstructed. Hence, the primary challenge for this research area is how to leverage the rich color information in 2D images and the precise depth

information in 3D LiDAR point clouds for achieving the balance between the quality and the efficiency of 3D facade reconstruction.

We hope that this paper can boost the future research on 3D facade reconstruction from different communities including remote sensing, computer vision, and computer graphics. Most of the challenges in this research area would be resolved by the improvement of both algorithms and hardware. Finally, with the increasing number of the applications of 3D urban reconstruction, we believe that this area will be increasingly crucial.

References

1. Becker, S., Haala, N.: Refinement of building facades by integrated processing of LIDAR and image data. *Int. Arch. Photogrammetry Remote Sens. Spat. Inf. Sci.* **36**, 7–12 (2007)
2. Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., Çöltekin, A.: Applications of 3D city models: state of the art review. *ISPRS Int. J. Geo-Inf.* **4**(4), 2842–2889 (2015). <https://doi.org/10.3390/ijgi4042842>
3. Brenner, C.: Building reconstruction from images and laser scanning. *Int. J. Appl. Earth Obs. Geoinf.* **6**(3), 187–198 (2005). <https://doi.org/10.1016/j.jag.2004.10.006>
4. Campbell, J.B., Wynne, R.H.: *Introduction to Remote Sensing*. Guilford Press, New York (2011)
5. Edum-Fotwe, K., Shepherd, P., Brown, M., Harper, D., Dinnis, R.: Fast, accurate and sparse, automatic facade reconstruction from unstructured ground laser-scans. In: *ACM SIGGRAPH 2016 Posters, SIGGRAPH 2016*, pp. 45:1–45:2. ACM, New York (2016). <https://doi.org/10.1145/2945078.2945123>
6. Fruh, C., Zakhor, A.: 3D model generation for cities using aerial photographs and ground level laser scans. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001*, vol. 2, p. II, December 2001. <https://doi.org/10.1109/CVPR.2001.990921>
7. Haala, N., Kada, M.: An update on automatic 3D building reconstruction. *ISPRS J. Photogrammetry Remote Sens.* **65**(6), 570–580 (2010). <https://doi.org/10.1016/j.isprsjprs.2010.09.006>
8. Hohmann, B., Krispel, U., Havemann, S., Fellner, D.: Cityfit-high-quality urban reconstructions by fitting shape grammars to images and derived textured point clouds. In: *Proceedings of the 3rd ISPRS International Workshop 3D-ARCH*, vol. 2009, p. 3D (2009)
9. Li, Y., Zheng, Q., Sharf, A., Cohen-Or, D., Chen, B., Mitra, N.J.: 2D–3D fusion for layer decomposition of urban facades. In: *2011 International Conference on Computer Vision*, pp. 882–889, November 2011. <https://doi.org/10.1109/ICCV.2011.6126329>
10. Liu, L., Stamos, I.: Automatic 3D to 2D registration for the photorealistic rendering of urban scenes. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, vol. 2, pp. 137–143, June 2005. <https://doi.org/10.1109/CVPR.2005.80>
11. Liu, L., Stamos, I.: A systematic approach for 2D-image to 3D-range registration in urban environments. In: *2007 IEEE 11th International Conference on Computer Vision*, pp. 1–8, October 2007. <https://doi.org/10.1109/ICCV.2007.4409215>

12. Müller, P., Zeng, G., Wonka, P., Van Gool, L.: Image-based procedural modeling of facades. In: ACM SIGGRAPH 2007 Papers, SIGGRAPH 2007. ACM, New York (2007). <https://doi.org/10.1145/1275808.1276484>
13. Musialski, P., Wonka, P., Aliaga, D.G., Wimmer, M., Gool, L., Purgathofer, W.: A survey of urban reconstruction. *Comput. Graph. Forum* **32**(6), 146–177 (2013). <https://doi.org/10.1111/cgf.12077>
14. Pu, S., Vosselman, G.: Building facade reconstruction by fusing terrestrial laser points and images. *Sensors* **9**(6), 4525–4542 (2009). <https://doi.org/10.3390/s90604525>
15. Sadeghi, F., Arefi, H., Fallah, A., Hahn, M.: 3D building Façade reconstruction using handheld laser scanning data. *Int. Arch. Photogrammetry Remote Sens. Spat. Inf. Sci.* **40** (2015). <https://doi.org/10.5194/isprsarchives-XL-1-W5-625-2015>
16. Stamos, I., Allen, P.K.: 3-D model construction using range and image data. In: CVPR, p. 1531. IEEE (2000). <https://doi.org/10.1109/CVPR.2000.855865>
17. Stamos, I., Allen, P.K.: Automatic registration of 2-D with 3-D imagery in urban environments. In: Proceedings Eighth IEEE International Conference on Computer Vision, ICCV 2001, vol. 2, pp. 731–736, July 2001. <https://doi.org/10.1109/ICCV.2001.937699>
18. Wang, J., et al.: Automatic modeling of urban facades from raw lidar point data. *Comput. Graph. Forum* **35**(7), 269–278 (2016). <https://doi.org/10.1111/cgf.13024>
19. Wang, R., Peethambaran, J., Chen, D.: Lidar point clouds to 3-D urban models: a review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **11**(2), 606–627 (2018). <https://doi.org/10.1109/JSTARS.2017.2781132>
20. Wang, R.: 3D building modeling using images and lidar: a review. *Int. J. Image Data Fusion* **4**(4), 273–292 (2013). <https://doi.org/10.1080/19479832.2013.811124>
21. Wang, R., Bach, J., Ferrie, F.P.: Window detection from mobile LIDAR data. In: 2011 IEEE Workshop on Applications of Computer Vision (WACV), pp. 58–65. IEEE (2011)
22. Xiao, J., Fang, T., Tan, P., Zhao, P., Ofek, E., Quan, L.: Image-based Façade modeling. *ACM Trans. Graph.* **27**(5), 161:1–161:10 (2008). <https://doi.org/10.1145/1457515.1409114>
23. Xiao, J., Fang, T., Zhao, P., Lhuillier, M., Quan, L.: Image-based street-side city modeling. *ACM Trans. Graph.* **28**(5), 114:1–114:12 (2009). <https://doi.org/10.1145/1661412.1618460>
24. Yang, L., Sheng, Y., Wang, B.: 3D reconstruction of building facade with fused data of terrestrial LIDAR data and optical image. *Optik - Int. J. Light Electron Opt.* **127**(4), 2165–2168 (2016). <https://doi.org/10.1016/j.ijleo.2015.11.147>
25. Zhang, J., Lin, X.: Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing. *Int. J. Image Data Fusion* **8**(1), 1–31 (2017). <https://doi.org/10.1080/19479832.2016.1160960>
26. Zhao, H., Shibasaki, R.: Reconstructing a textured cad model of an urban environment using vehicle-borne laser range scanners and line cameras. *Mach. Vis. Appl.* **14**(1), 35–41 (2003). <https://doi.org/10.1007/s00138-002-0099-5>