ISPRS Journal of Photogrammetry and Remote Sensing 195 (2023) 220–221

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Editorial of theme issue "3D Modeling of Indoor Environments"

This theme issue is dedicated to the topic of "3D Modeling of Indoor Environments". 3D modelling of indoor environment has emerged as a significant and challenging task in terms of the provision of a semantically rich and geometrically accurate indoor model. Recently, there has been an increasing amount of research related to 3D modeling of indoor environment for many applications such as navigation guidance, emergency management, building maintenance and renovation planning, and a range of indoor location-based services (e.g. way-finding, contextualized content delivery, etc.).

The sixteen contributions for this theme issue can be roughly divided into two groups: seven papers deal with various aspects of acquiring the 3D geometry of indoor environments, whereas the other nine manuscripts focus on geometric, semantics and topological perspectives of indoor models.

1 Data acquisition of the 3D indoor environments

In general, a diversity of techniques can be applied for a highly accurate acquisition of the 3D geometry of indoor environments. For example, Simultaneous Localization and Mapping (SLAM) is the online process of capturing and recovering the geometrical structures of the whole scene by using either active (e.g., laser sensor) or passive (e.g., monocular and binocular stereo) sensing techniques. Wen et al. proposed a dense map construction algorithm based on a light-and-fast stereo visual-inertial navigation system (VINS). Unlike the original VINS-Mono, their proposed method estimated the scale factor using stereo matching instead of visual-inertial alignment. In addition, a dense map was constructed for mobile vehicles. Images can be used also to reconstruct advanced depth maps. Zioulis et al. focused on exploiting the weak layout cues for the indoor scene depth estimation with a single omnidirectional camera. They explicitly coupled the layout and depth estimation tasks to offer increased depth estimation performance. As the rapid development of laser scanner, Light Detection and Ranging (LiDAR)-based SLAM systems become an efficient way to quickly acquire the 3D indoor environment information. For the indoor scene, it is characterized by few features, similar components and large scales, which seriously influences data association and cumulative error elimination. As a result, the quality of SLAM-based mapping is degraded. Jia et al. presented a cross-correction LiDAR SLAM method for constructing high-accuracy 2D maps. They addressed problematic scenario by taking the cross correction between pose and map into account to enhance the data association and cumulative error elimination capacities. Mobile laser scans usually show the data characteristics of the sparsity and inhomogeneity. Cong et al. investigated the precision of 3D mapping by a LiDAR-based SLAM system, instead of the efficiency and robustness. In their developed method, structural feature extraction and directional constraint sphere are used to improve the performance of laser odometry. A non-rigid registration based on the spline motion further refined the mapping results. They concluded that the problems of motion distortion and degeneration are jointly optimized. To address the inhomogeneous density of laser scans, Chen et al. introduced a feature equalizer that seeks to balance the inhomogeneous feature points of buffered scans in a sliding window and made them uniformly distributed to the best possible extent. Moreover, a self-supervised prior module was designed, which would also provide an initial guess for the registration process. With the process of indoor mobile mapping systems, new opportunities for surveying and mapping complex sites have emerged, which makes scan planning very necessary for maximizing data completeness and minimizing data repetitiveness. Frías et al. developed a novel optimal scan planning method with static and mobile mapping systems to achieve specific mobile laser scanning systems constraints (such as maximum acquisition time or closed-loops requirement). Besides, the advancement of smartphones also offers great opportunities for crowdsourcing-based indoor mapping, which is one of the most promising applications due to its low cost and flexibility. Zhou et al. provided a comprehensive review of crowdsourcing-based indoor mapping solutions using smartphones.

2 Geometric, semantics and topological perspectives of indoor models

Vectorized indoor reconstruction has attracted increasing attention in recent years due to its high regularity and low memory consumption. Wall structures are always the focus of vectorized indoor reconstruction due to their complexity, and they are usually represented by floorplans. The floorplan is considered as a planar graph where each simple cycle represents the polygonal boundary of a room. Hübner et al. presented an automatic method for room partitioning from triangle meshes. Their proposed approach is applicable to a range of challenging scenarios encompassing indoor environments with curved walls and complex room layouts extending vertically over multiple storeys. Fang et al. partitioned the indoor scenes into a set of polygonal spaces and selected edges belonging to wall structures for floorplan generation. The potential of their methods was proved from simple to complex indoor scenes, compared with existing methods in terms of geometric accuracy and output simplicity. Currently, the generation of 3D building models has many far-reaching applications in the architectural engineering and construction community. Han et al. developed a multistep pipeline to not only produce floorplans but also further compact polygonal models from point clouds obtained by LiDAR or Multi-View-Stereo. Their produced models confirm to CityGML LoD2 without overhang structures. For generating the vectorized models closer to the real scenes (e.g. LoD3), Li et al. focused on reconstructing small regular objects in indoor scenes from RGB-D data by combining data and model-driven method.

https://doi.org/10.1016/j.isprsjprs.2022.11.019
Received 22 November 2022; Accepted 22 November 2022
Available online 7 December 2022

Instead of directly generating compact polygon representations, Fang et al. proposed a structure-aware indoor scene reconstruction method by reducing the number of facets of dense meshes. To this end, two meshes with different levels of abstraction were produced, which would be used for different indoor applications according to user preferences.

Efficient semantic expression of indoor environments is of great significance to many applications. Huan et al. developed an encoder-decoder framework to learn and fuse multi-scale features for indoor semantic segmentation, where it propagates inter-scale contextual information and aggregates the hierarchical features for producing the discriminative representations. Their qualitative and quantitative results suggest that the developed model is capable of scale-robust indoor semantic segmentation and avoids the strong preference to large-scale instances in the prediction results. For most hierarchical feature approaches, their architectures are generally manually-designed, which heavily relies on human experience and massive trials. Instead of manually defining a multi-level feature fusion network, Lin et al. presented a neural architecture search method from point cloud semantic segmentation. In their approach, a feature pyramid network was automatically found for integrating the feature representations at different levels. This automation on the architecture design saves massive trials and human errors. The objective of a single task usually fails to generate enough information for robustly achieving the ideal requirements. Multi-task collaborative optimization offers redundant and complementary information from different perspectives to assist task completion. In addition to the hierarchical feature aggregation, Wang et al. proposed a densely connected graph convolutional network through jointly semantic and instance segmentation. Their framework combined graph convolutional network, attention pooling and multi-level feature aggregation for achieving the indoor semantic modeling.

Unlike outdoor counterparts, where a specific location is defined and represented by coordinate reference systems, the indoor positions are usually marked using a cell identifier or a symbolic code due to the building structures (e.g. walls, doors). The spatial relationship reconstruction plays a significant role in indoor applications. Yang et al. proposed a semantics-guided method for indoor navigation element reconstruction from colorized point clouds. Moreover, the reconstructed relationship of indoor navigation elements was encoded by pre-defined XML Schemas and therefore can be used as the base for indoor navigation applications regardless of the platforms. Their experimental results demonstrate the capability of the automatic reconstruction of indoor navigation elements.

Here, we would like to thank all the reviewers for their time and dedication to the review process. Moreover, the Guest Editors are also very grateful to Dr. Clément Mallet, Editor-in-Chief of the ISPRS Journal of Photogrammetry and Remote Sensing and Prof. Derek Lichti, former Editor-in-Chief of the ISPRS Journal of Photogrammetry and Remote Sensing, for their great collaboration and assistance in preparing this collection of scientific publications.

Zhizhong Kang[a], Cheng Wang[b], Kourosh Khoshelham[c], Ville Lehtola[d]

[a] *School of Land Science and Technology, China University of Geosciences (Beijing), China*

[b] *School of Informatics, Xiamen University, China*

[c] *Department of Infrastructure Engineering, University of Melbourne, Australia*

[d] *Faculty of Geo-Information Science and Earth Observation, University of Twente, the Netherlands*

E-mail addresses: zzkang@cugb.edu.cn (Z. Kang), cwang@xmu.edu.cn (C. Wang), k.khoshelham@unimelb.edu.au (K. Khoshelham), v.v.lehtola@utwente.nl (V. Lehtola).