Zhilu YUAN, You LI, Shengjun TANG, Ming LI, Renzhong GUO, Weixi WANG, 2021. A survey on indoor 3D modeling and applications via RGB-D devices. *Frontiers of Information Technology & Electronic Engineering*, 22(6):815-826. https://doi.org/10.1631/FITEE.2000097

A survey on indoor 3D modeling and applications via RGB-D devices

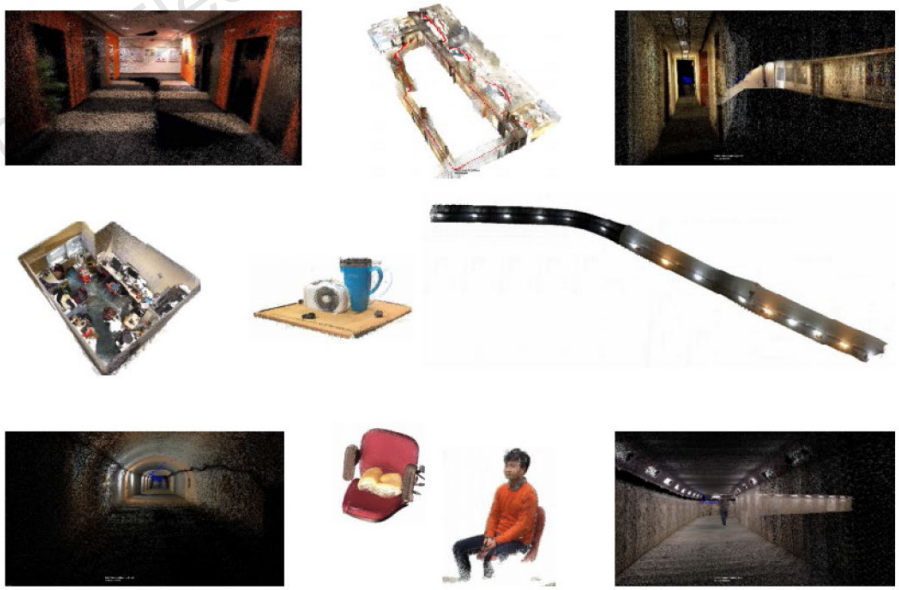
Key words: 3D indoor mapping; RGB-D; Indoor localization; Construction monitoring; Emergency evacuation

Shengjun Tang
E-mail: shengjuntang@szu.edu.cn
ORCID:https://orcid.org/0000-0002-8262-7397

Researches on RGB-D devices

- Generation of detailed 3D maps for indoor environments is essential for many mobile robot applications, including indoor navigation, facility management, virtual reality, and building information models (BIMs) (He and Habib, 2018).

Researches
For the modeling part, we have witnessed the development of 3D dense mapping and simultaneous localization and mapping (SLAM) pipelines. Those that depend on only the RGB-D devices are now an important part.
Since commodity RGB-D sensors were introduced, the 3D geometry capture technology has blossomed rapidly. Applications based on these technologies are expanding over a broad field, such as semantic understanding of scenes, indoor localization, and BIM reconstruction.



Commercial RGB-D devices

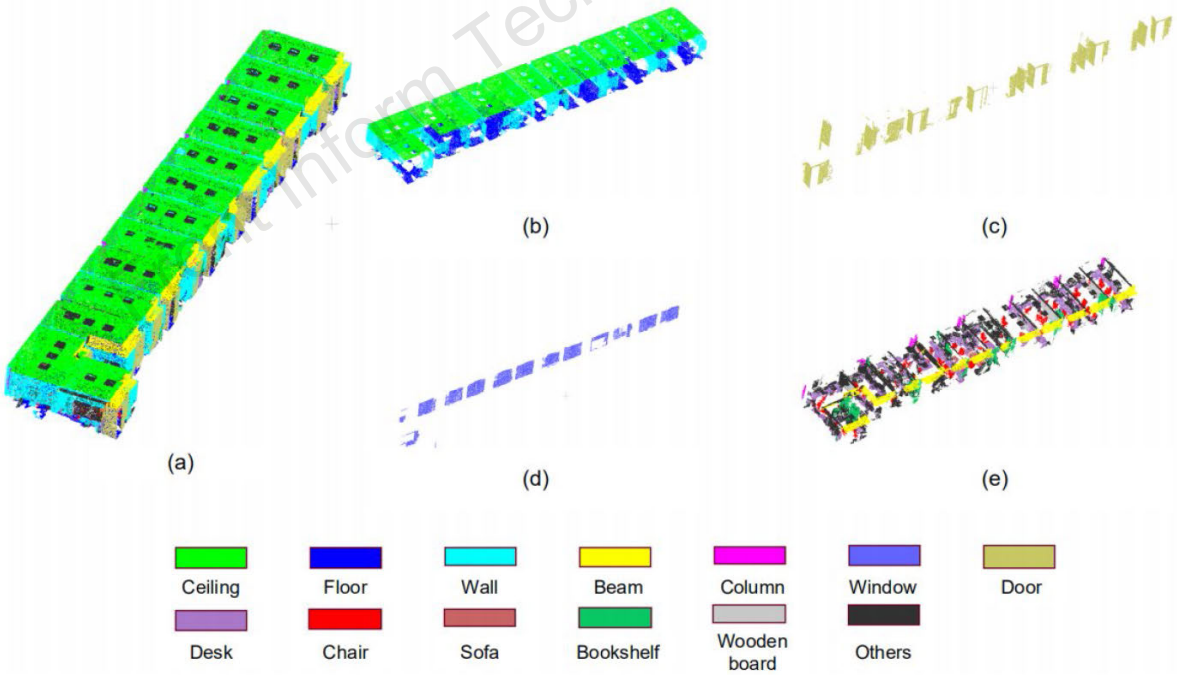
- In recent years, RGB-D devices, such as Kinect and Structure Sensors, **have gained wide acceptance for SLAM and indoor-mapping applications**. There are typical RGB-D sensors, like Kinect 1, Kinect 2, Azure Kinect DK, and Structure Sensor.

Device	Size (height×width×length)	Depth image resolution	Data acquisition frequency (frame/s)	Battery usage	Weight (g)
Kinect 1	64 mm×76 mm ×305 mm	320×240	<30	No battery	1360
Kinect 2	76 mm×165 mm ×350 mm	512×424	<30	No battery	1225
Azure Kinect DK	39 mm×103 mm ×126 mm	640×576, 320×288, 512×512, or 1024×1024	<30	No battery	400
Structure sensor	29 mm×28 mm ×119 mm	640×480	<30	3–4 h continuous mapping and 1000+ h standby	95

Device	Maximum effective distance (m)	View angle	Inertial measurement unit	Interface	Multi-device synchronization
Kinect 1	4.5	57°×43°	×	USB2/ USB3	√
Kinect 2	4.5	70°×60°	×	USB3	×
Azure Kinect DK	4	75°×65°, or 120°×120°	√	USB3	√
Structure sensor	3.5	58°×45°	×	Lightning	×

Datasets and modeling techniques

- Public datasets: common datasets and explicit evaluation indices can assist in the improvement of most advanced technologies. This view has been proved by several successful examples in the computer vision area.
  - SUN RGB-D (Song et al., 2015)
  - ICL-NUIM (Chen K et al., 2014)
  - TUM RGB-D datasets (Sturm et al., 2012)
  - Washington RGB-D scenes dataset (Lai et al., 2014)



Datasets and modeling techniques

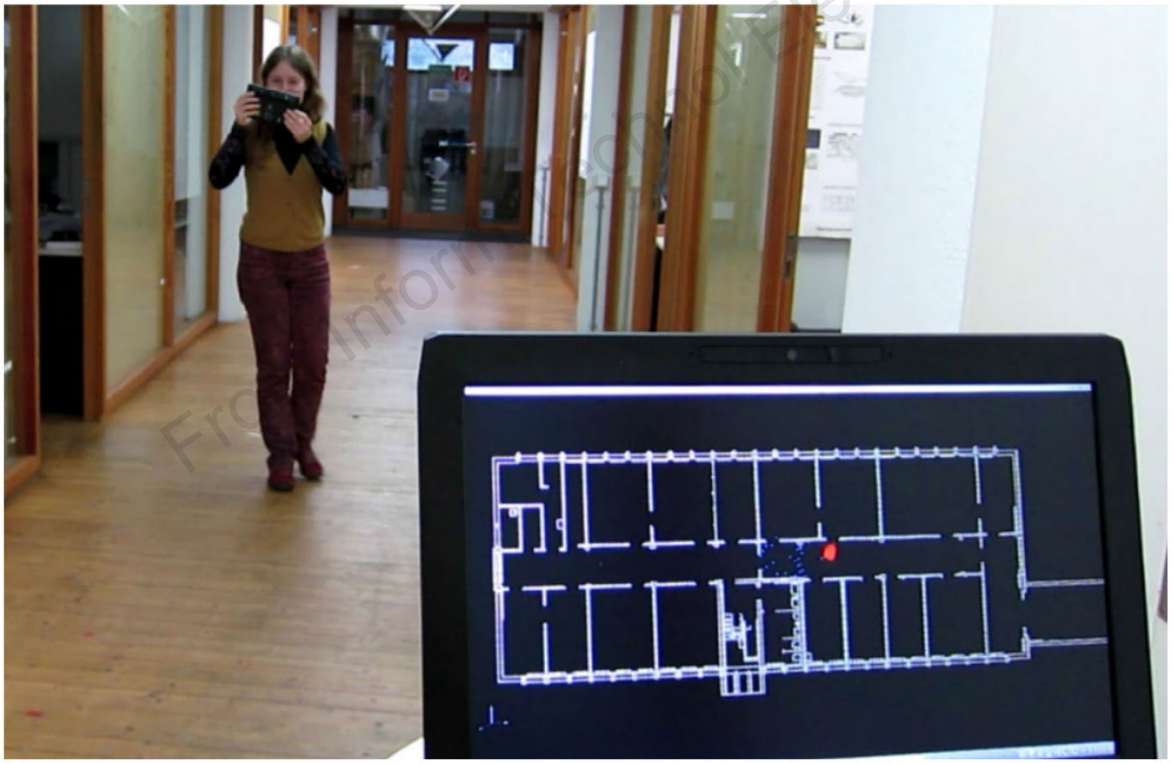
- Many 3D techniques that use RGB-D devices and laser systems have been developed in robotics and computer vision research areas.

  - The feature-based SLAM method is commonly used in the visual SLAM system. The core concept of this method is to achieve camera tracking using the detected feature points, and pose updating is conducted by minimizing the distances of features.

  - For the dense system, the ICP algorithm and relative variants are commonly used techniques. The ICP algorithm conducts pose updating by minimizing the whole distances between two sets of point clouds.

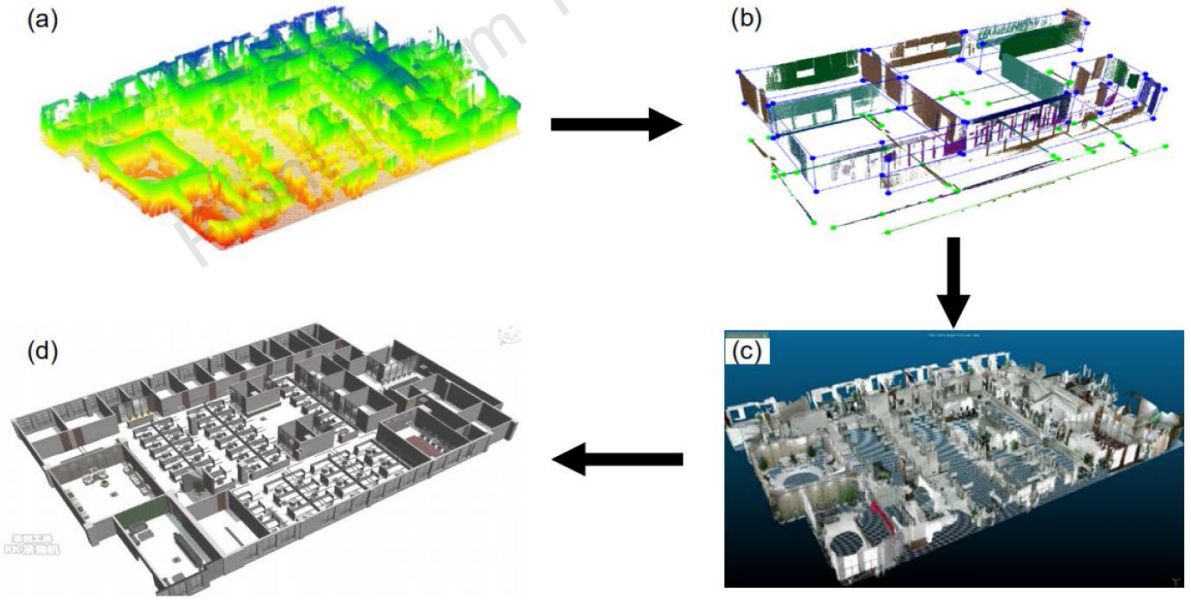
Applications: indoor localization

- It is universally accepted that locating people precisely in indoor conditions is key to services requiring location perception.
- More and more research has been carried out on RGB-D devices, which can calculate the positions of smartphones or tablets with **six degrees of freedom** (Winterhalter et al., 2015).



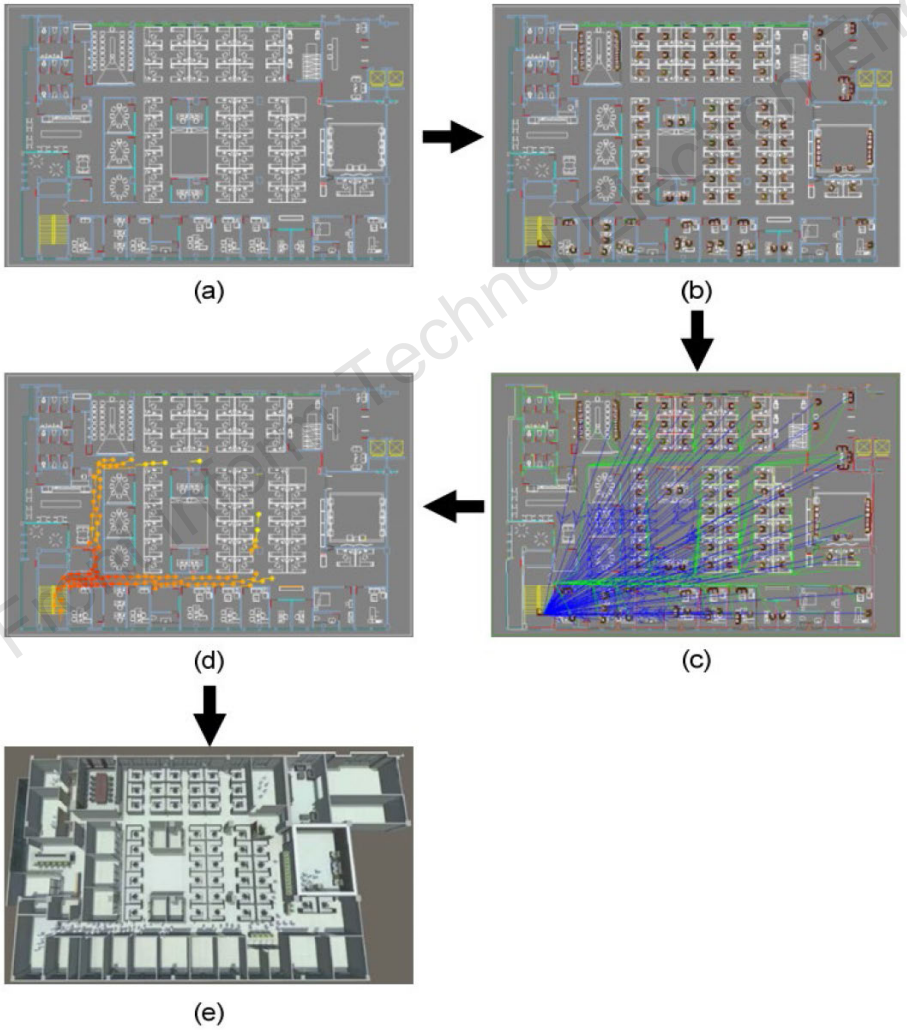
Emergency evacuation

- Emergencies such as fires, earthquakes, and terrorist attacks occur frequently. **How to ensure safe and efficient evacuation** of people in the case of emergency is a key point in the field of emergency evacuation management (Yuan et al., 2019).
- The research hotspot of indoor emergency evacuation is to **use simulation technology to restore and quantify the layout of indoor space facilities**, the **topological structure** of escape routes, and the evacuation management strategy, to verify the rationality of the design and management scheme for indoor public areas (Tang et al., 2019a).



Emergency evacuation

- After 3D reconstruction by RGB-D devices, the simulation of the pedestrian movement process was realized by combining the micro simulation models of pedestrian movement such as social force.



Conclusions

Outlook
• In this study, we presented an extensive survey on indoor scene modeling and applications with RGB-D data.
• We briefly introduced some public RGB-D datasets and modeling technologies, and divided the technologies for 3D modeling using an RGB-D system into two categories: feature-based and dense styles.
• After that, two typical applications including indoor localization and emergency evacuation based on RGB-D devices were introduced. By summarizing a broad spectrum of literature related to RGB-D devices, we hope this work gives some insights into this important topic.

References

[1] He FN, Habib A, 2018. Three-point-based solution for automated motion parameter estimation of a multi-camera indoor mapping system with planar motion constraint. ISPRS J Photogr Remot Sens, 142:278-291. https://doi.org/10.1016/j.isprsjprs.2018.06.011

[2] Song S, Lichtenberg SP, Xiao J, 2015. SUN RGB-D: a RGB-D scene understanding benchmark suite. IEEE Conf on Computer Vision and Pattern Recognition, p.567-576. https://doi.org/10.1109/CVPR.2015.7298655

[3] Chen K, Lai YK, Wu YX, et al., 2014. Automatic semantic modeling of indoor scenes from low-quality RGB-D data using contextual information. ACM Trans Graph, 33(6):208. https://doi.org/10.1145/2661229.2661239

[4] Sturm J, Engelhard N, Endres F, et al., 2012. A benchmark for the evaluation of RGB-D SLAM systems. IEEE/RSJ Int Conf on Intelligent Robots and Systems, p.573-580. https://doi.org/10.1109/IROS.2012.6385773

[5] Winterhalter W, Fleckenstein F, Steder B, et al., 2015. Accurate indoor localization for RGB-D smartphones and tablets given 2D floor plans. IEEE/RSJ Int Conf on Intelligent Robots and Systems, p.3138-3143. https://doi.org/10.1109/IROS.2015.7353811

[6] Yuan ZL, Guo RZ, Tang SJ, et al., 2019. Simulation of the separating crowd behavior in a T-shaped channel based on the social force model. IEEE Access, 7:13668-13682. https://doi.org/10.1109/ACCESS.2019.2894345

[7] Tang SJ, Li Y, Yuan ZL, et al., 2019a. A vertex-to-edge weighted closed-form method for dense RGB-D indoor SLAM. IEEE Access, 7:32019-32029. https://doi.org/10.1109/ACCESS.2019.2900990