A VR-Oriented Efficient 3D Scene Synthesis System

Yi-ke Li
Tsinghua University
lyk20@mails.tsinghua.edu.cn

Hou Tam
Tsinghua University
th21@mails.tsinghua.edu.cn

Jia-Hong Liu
Tsinghua University
liujiaho19@mails.tsinghua.edu.cn

Shao-Kui Zhang
Tsinghua University
zhangsk18@mails.tsinghua.edu.cn

Song-Hai Zhang
Tsinghua University & BNRist
shz@tsinghua.edu.cn

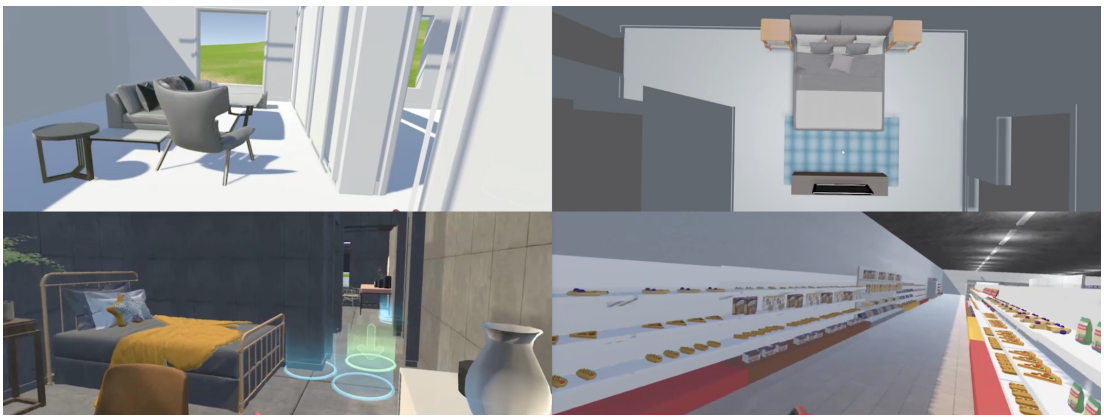

Fig. 1: Several typical works of our system: Real-time recommendations for function area of the indoor scene to meet user preferences (top-left). Real-time recommendations of objects in the indoor scene to meet user preferences (top-right). Automatic generation of indoor viewpoints based on photographic rules (bottom-left). Automatic layout of commercial scenes (bottom-right).

***Abstract*—With the rapid development of virtual reality (VR), the demand for 3D scenes has been boosted dramatically. Massive algorithms of automatic arrangement for objects have been proposed to solve this problem. However, few of these techniques are applied in VR. To make the automatic 3D scene synthesizing more suitable for VR, we put forward a series of innovative methods, considering the interaction and presentation features in VR. Our methods, including automatic scene synthesis and intelligent scene editing, contribute to the efficiency and effectiveness of VR scene synthesis. We present a multi-person, multi-terminal remote 3D scene collaborative layout system with the above technical sustain. These attributes, consequently, promote the applicability and convenience of our system.**

***Index Terms*—Scene Synthesis, Intelligent Scene Editing**

I. PROBLEM STATEMENT

Indoor scene synthesis triggers discussion due to the difficulty of selecting object categories reasonably and arranging objects plausibly. [11] proposes four criteria for the synthesis techniques and objectively compares the advantages and drawbacks between them. In [4], [5], choosing and arranging objects are respectively formulated as a supervised generative model and a Bayesian model, which are solved by MCMC.

One-to-one motion tracking has been widely applied in VR interaction technologies, but it contains drawbacks due to physical limitations. [6] proposes a redirected walking (RDW) method by reducing the reset times and distance of the point of interest (POI) when reset. [9] proposes a RDW algorithm to divide the physical boundary into a set of endpoints with a reset direction vector. In [10], the heated map is calculated, which helps to identify the most approximate position for a reset. [1] trains a classifier using the eye head-coordinated information sequences to detect whether the motion of the head is relevant to the scene.

Other methods based on scans, videos, cameras, and point of interest can be applied in various situations related to VR scenes. [2] provides a system that guarantees users to avoid obstacles in the physical environment. [8] proposes a scale-aware technique to insert virtual objects of suitable size into monocular video. In [7], a home fitness monitoring system is used to detect whether the user's fitness action is standard by the monocular camera. In [3], the living convenience index of residential buildings in Beijing's urban areas is investigated.

II. OUR NOVELTIES

Increasing the Efficiency of VR-Oriented Scene Synthesis. The manual layout of VR scenes consumes a lot of time and energy. By comparison, the techniques of both automatic synthesis and interactive synthesis in our system substantially increase the layout's efficiency.

Developing an Integrated System. We develop an open-source system for manipulating 3D scenes such as exploring and modifying. Our system supports the multi-user, multi-terminal remote collaborative operation in a VR environment, which provides a vivid and immersive experience for the user.

Improving the Interaction of Intelligent Editing in VR. The basic operations of editing 3D scenes, such as

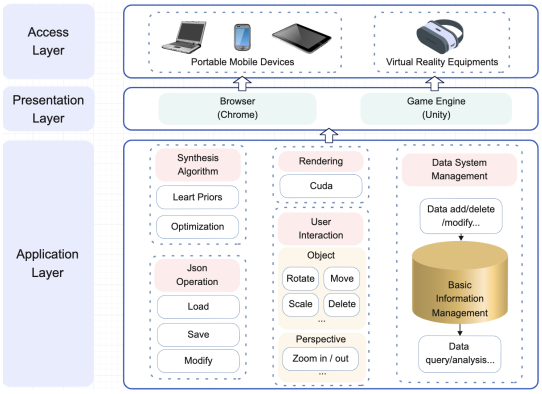

Fig. 2: VR-oriented 3D scene synthesis system structure.

inserting, deleting, moving or rotating the object, can be completed through the VR controllers. Besides, some superior interaction is also employed in our system. Interactions with minimum efforts, such as hand gestures in VR are used for intelligent editing. For each scene, our method generates a set of viewpoints. When user teleport at these places by VR controllers, their field of vision of HMD can be adjusted to the most aesthetic one.

III. TECHNICAL HIGHLIGHTS

The system architecture is demonstrated in Figure 2.

Automatic VR-Oriented Scene Synthesis. [12] proposes a fast indoor scene synthesis method by inputting the room shape and an object list with the learned priors. By tests for complete spatial randomness (CSR), we obtain the intensity of spatial relations. Furthermore, we learn the discrete priors from samples that accurately represent layout patterns. Using the Hausdorff metric, the layouts come out in seconds and are more reasonable than other methods. In [15], we first learn hyper-relations among three or more objects, which requires merely *O(1)* time to sample layout. Then, we propose a geometry-based framework for indoor furniture layout generation, considering various details. We also propose an automatic synthesizes method in commercial places. We decompose the traffic flow into computational patterns. Then, we synthesize these patterns by applying loss functions guiding the optimization. The loss functions ensure a plausible and practicable result considering components like space utilization, elasticity, roundness, etc.

Interactive Scene Synthesis in VR. An impressive framework [13] is proposed to meet both user satisfaction and arrangement efficiency demand. While the user is moving and clicking the cursor in the scene, the framework automatically recommends suitable objects in real time. This technique is based on priors learned from the dataset. The object is placed according to the current scene context. Besides this single object inserting method, we also proposed a synthesis tool to preview the potential group of objects by hand gestures in VR. This tool is achieved by the novel idea of "coherent group sets", a concept of layout strategies.

Advanced Scene Interaction in VR. Since our system is multi-user and multi-channel, users can cooperate with the same scene via a mobile terminal, PC terminal and VR equipment. The Socket-IO protocol is used to connect different terminals. When editing the same scene, any terminal's changes can be reflected in other devices in real time. Particularly in VR, the user can teleport to a specific spot and acquire aesthetic views recommended by our system. This method is achieved by our automatic view selections algorithm in [14], seeking a better view according to interior photography.

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation of China (Project Number 62132012) and the Tsinghua-Tencent Joint Laboratory for Internet Innovation Technology.

REFERENCES

[1] X. Fan, Y. Cai, Y. Yang, T. Xu, Y. Li, S. Zhang, and F. Zhang. Detection of scene-irrelevant head movements via eye-head coordination information. *Virtual Reality & Intelligent Hardware*, 3(6):501–514, 2021.

[2] Y. He, Y. Liu, Y. Jin, S.-H. Zhang, Y.-K. Lai, and S.-M. Hu. Context-consistent generation of indoor virtual environments based on geometry constraints. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[3] J. Li and S. Zhang. Research on beijing residential convenience index based on point of interest. *Journal of Computer-Aided Design & Computer Graphics*, 33:609–615, 04 2021. doi: 10.3724/SP.J.1089.2021.18539

[4] Y. Liang, S.-H. Zhang, and R. R. Martin. Automatic data-driven room design generation. In *International Workshop on Next Generation Computer Animation Techniques*, pp. 133–148. Springer, 2017.

[5] Y. Liang, S.-H. Zhang, and R. R. Martin. Learning guidelines for automatic indoor scene design. *Multimedia Tools and Applications*, 78(4):5003–5023, 2019.

[6] S.-Z. Xu, T.-Q. Liu, J.-H. Liu, S. Zollmann, and S.-H. Zhang. Making re-sets away from targets: Poi aware redirected walking. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[7] P. YU, L. LIU, Y. CAI, Y. HE, and S. ZHANG. Home fitness monitoring system based on monocular camera. *Journal of ZheJiang University (Science Edition)*, 48(5):521–530, 2021.

[8] S. Zhang, X. Li, Y. Liu, and H. Fu. Scale-aware insertion of virtual objects in monocular videos. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 36–44. IEEE, 2020.

[9] S.-H. Zhang, C.-H. Chen, Z. Fu, Y. Yang, and S.-M. Hu. Adaptive optimization algorithm for resetting techniques in obstacle-ridden environments. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[10] S.-H. Zhang, C.-H. Chen, and S. Zollmann. One-step out-of-place resetting for redirected walking in vr. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[11] S.-H. Zhang, S.-K. Zhang, Y. Liang, and P. Hall. A survey of 3d indoor scene synthesis. *Journal of Computer Science and Technology*, 34(3):594–608, 2019.

[12] S.-H. Zhang, S.-K. Zhang, W.-Y. Xie, C.-Y. Luo, Y.-L. Yang, and H. Fu. Fast 3d indoor scene synthesis by learning spatial relation priors of objects. *IEEE Transactions on Visualization and Computer Graphics*, 28(9):3082–3092, 2021.

[13] S.-K. Zhang, Y.-X. Li, Y. He, Y.-L. Yang, and S.-H. Zhang. Mageadd: Real-time interaction simulation for scene synthesis. In *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 965–973, 2021.

[14] S.-K. Zhang, H. Tam, Y.-X. Li, T.-J. Mu, and S.-H. Zhang. Sceneviewer: Automating residential photography in virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–14, 2022. doi: 10.1109/TVCG.2022.3214836

[15] S.-K. Zhang, W.-Y. Xie, and S.-H. Zhang. Geometry-based layout generation with hyper-relations among objects. *Graphical Models*, 116:101104, 2021.