2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)

Interactive Panoramic Ray Tracing for Mixed 360° RGBD Videos

Jian Wu*
State key laboratory of virtual reality technology and systems, Beihang University, Beijing, China

Lili Wang†
State key laboratory of virtual reality technology and systems, Beihang University, Beijing, China
Peng Cheng Laboratory, Shenzhen, China

Wei Ke‡
Macao Polytechnic University

2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 979-8-3503-4839-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/VRW58643.2023.00231

Abstract

This paper introduces an interactive panoramic ray tracing method for rendering real-time photo-realistic illumination and shadow effects when inserting virtual objects into 360° RGBD videos. First, we approximate the geometry of a real scene with a panoramic depth buffer and a screen space depth buffer. Then, a sparse sampling ray generation method is proposed to accelerate the tracing process by reducing the number of rays that need to be emitted in ray tracing. After this, an irradiance estimation pass is introduced to generate a noisy Monte-Carlo image. At last, the final result is smoothed by interpolation, spatial-temporal filtering, and differential rendering. We tested our method in some natural and synthetic scenes and compared the results of our approach with the ground truth and that of the Image-Based Lighting method. The results show that our method can generate visually photo-realistic frames for virtual objects in 360° RGBD videos in real-time, making the rendering results more natural and credible.

Index Terms: Mixed reality—360° RGBD video—Real-time rendering—Panoramic ray tracing

1 Introduction

In mixed reality applications, photo-realistically inserting virtual objects into the natural world environment is challenging. One of the problems is that it is difficult to capture the entire environment's lighting, geometry, and material information. The missing natural environment information often leads to inconsistent final images due to some artificial lighting and geometry hypothesis. Another problem is that when simulating the interaction between virtual objects and natural environment lighting, some approximate processing is often done to meet the high requirements for the frame rate of mixed reality applications. However, many details were lost, such as surface shadow, spatially-varying illumination, etc.

This paper provides an interactive panoramic ray tracing method to facilitate real-time realistic virtual object insertion and rendering for 360° RGBD videos in mixed reality. We first approximate the geometry of a real scene with a panoramic and a screen space depth buffer. Then, a sparse tracing acceleration is proposed to accelerate the tracing process by reducing the number of rays emitted during ray tracing. After this, an irradiance estimation pass is introduced to generate a noisy Monte-Carlo image. At last, the final result is smoothed and synthesized by interpolation, spatial-temporal filtering, and differential rendering. Our method supports spatially varying global illumination for multiple virtual objects materials, including diffuse, glossy, and specular, allowing dynamic virtual objects and changing real-world geometry. We tested our method in natural and synthetic scenes and compared the result of our method with the ground truth and that of the IBL method. The results show that our method can generate visually photo-realistic frames for inserting virtual objects into 360° RGBD videos in real-time.

*e-mail: lanayawj@buaa.edu.cn
†e-mail: Wanglily@buaa.edu.cn, Lili Wang is also the corresponding author
‡e-mail: wke@mpu.edu.mo

2 Interactive Panoramic ray tracing Method

According to the depth of the environment panorama, we can know precisely from where light is emitted. Then, assuming all materials in the natural world environment are diffuse, any position inside the real world space can be seen as light by the panorama pixels since diffuse material objects reflect light uniformly in any direction. By calculating the appearance of every pixel on the virtual object surface according to their position and environment lighting, we can simulate a spatially varying virtual-real merging result.

Inspired by the ideas of fast spatial-temporal filtering algorithm [6] and the image space ray tracing approach [3], we propose an interactive panoramic ray tracing method for mixed 360° RGBD videos. Our method can render frames with virtual objects naturally merged in the real-world scene with realistic illumination, shadows, and reflections in real-time.

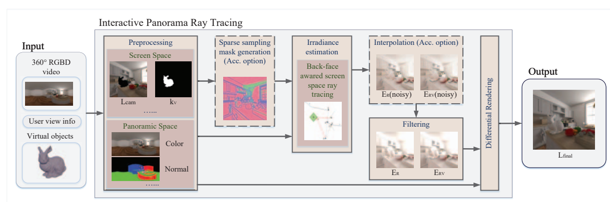

Figure 1: Our method merges virtual objects into a 360 ° RGBD frame and keeps the lighting, reflections, and shadows consistent. The pipeline consists of six passes.

Fig. 1 gives an overview of our IPRT pipeline for an RGBD panoramic frame based on a differential rendering framework [1]. The pipeline consists of six passes. The first pass is the preprocessing pass. In this pass, we generate G-buffers of the merging scene, which will be used in the subsequent passes.

The second sparse sampling mask generation pass and the fourth interpolation pass are designed to control the sampling rate of the rays traced from the user's view to satisfy the high frame rate requirement of many mixed reality applications. They are packaged as acceleration options, meaning the system can normally run without them.

The third pass is an irradiance estimation pass. In this pass, we first calculate only the rays reaching the real surface and update the color of the real surface. Then, we calculate the rays reaching the virtual object's surface. We already know that with Monte-Carlo estimation, one can calculate the irradiance of a pixel E with:

$$E = \frac{1}{N}\sum_{i=1}^{N} L_{\omega_i} \cdot f_r \cdot cos\theta / p_{f_r}(\omega_i) \quad (1)$$

Where, L_{ω_i} is the radiance of the sampled ray from direction ω_i, f_r is the BRDF, $cos\theta$ is the cosine term and $p_{f_r}(\omega_i)$ is the probability

979-8-3503-4839-2/23/$31.00 ©2023 IEEE
DOI 10.1109/VRW58643.2023.00231

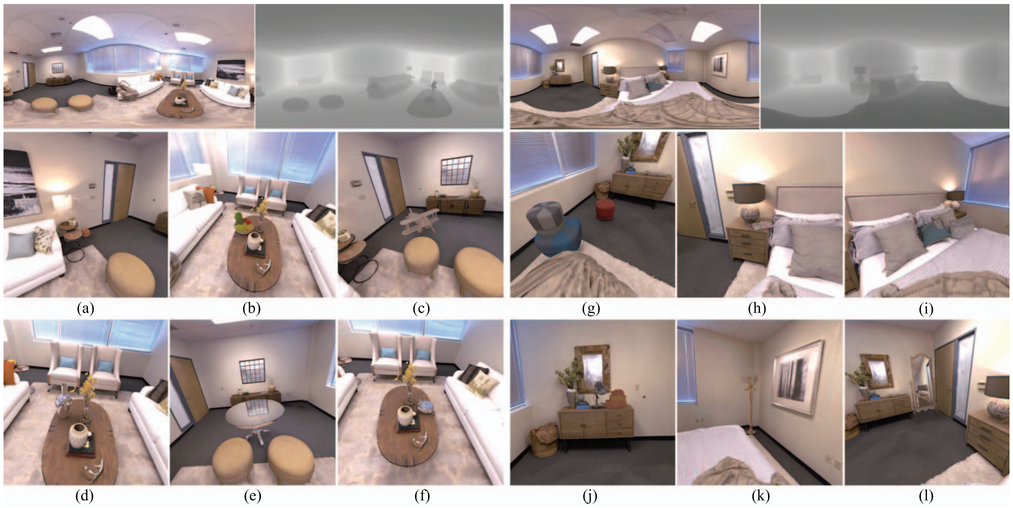

Figure 2: The augmented images of scenes *Apartment* (left half) and *Bedroom* (right half).

density function (PDF) of the sampled ray distributed with BRDF. To run the system in real-time, we only sample at most one Monte-Carlo ray for each pixel in a frame ($N = 1$). And for every frame, we get noisy irradiance images.

After the above passes, we insert a filtering pass to accurately reconstruct the noisy Monte-Carlo irradiance images. The reconstruction performs two main steps. In the first step, we use the screen space G-buffer to drive a 65×65 joint bilateral spatial filter [2] to get an initial irradiance estimation. The filter is implemented in a 5-level *à*-trous manner for efficiency [6]. In the second step, we temporally filter the result images of the spatial filtering process using motion vectors. We also performed a clamping operation to reduce ghosting during temporal accumulation [7].

In the final differential rendering pass, we calculate the final result of the merging scene. As illustrated in the differential rendering framework [4], the final rendering image of the virtual-real hybrid scene can be calculated as follow:

$$L_{final} = (1-M) \cdot k_V E_{RV} + M \cdot (\frac{E_{RV}}{E_R} L_{cam}) \quad (2)$$

Where, L_{final} is the final rendered image of the hybrid scene, M is the fraction of the pixel covered by the real-world environment, k_V is the virtual objects albedo texture, E_{RV} is the pixel's irradiance considering both real and virtual objects, E_R is the pixel's irradiance considering only real objects and L_{cam} is the original real-world G-buffer radiance image (color image). We can easily compute k_V, L_{cam}, and M in the preprocessing pass. With the equation, we can synthesize the merging result image (L_{final}) of the virtual objects and the real-world 360 $^\circ$ RGBD frame.

3 Results and Discussion

Fig. 2 shows the rendering result of our IPRT method for two 360° RGBD frame scene: *Apartment* and *Bedroom* [5]. We insert 12 virtual models into the two scenes, with different diffuse, glossy, and specular materials. Our method can generate realistic global illumination effects on these virtual objects. For example, the lighting effect on the wing of the wooden plane (c), and the pillow (i) that changes with the environment, the reflections of the yellow leaves on the surfaces of the bottle (d) and the teapot (f), and also the images of the windows, the lamp and the cupboard on the surface of the tabletop (e) and the mirror (l). In addition, our method can also generate shadows on the real surface caused by these inserted virtual objects, such as the shadows next to the basketball, the plane, the soft stools, the hanger, etc. Furthermore, we can see the spatially varying effects when the virtual object moves around in the scene; these global illumination effects change coherently of the inserted virtual models (please also reference our accompanying video).

4 Conclusion

We have presented a real-time rendering framework based on our interactive panoramic ray tracing method to facilitate photo-realistic global illumination in MR for 360 $^\circ$ RGBD video. We give a panoramic ray tracing method to calculate the interaction between rays and the surrounding environment. We also provided a sparse tracing acceleration strategy to accelerate the frame rates of our framework further, taking into account the different performances of various graphics hardware. We compared our method's results with the ground truth rendered with the path tracing method in two synthetic scenes. The results showed that our method could produce photo-realistic images at pretty high frame rates.

Acknowledgments

This work was supported by National Key R&D plan 2019YFC1521102, by the National Natural Science Foundation of China through Projects 61932003, by Beijing Science and Technology Plan Project Z221100007722004.

References

[1] P. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *ACM SIGGRAPH 2008 classes*, pp. 1–10. 2008.

[2] J. Kopf, M. F. Cohen, D. Lischinski, and M. Uyttendaele. Joint bilateral upsampling. *ACM Transactions on Graphics (ToG)*, 26(3):96–es, 2007.

[3] M. McGuire and M. Mara. Efficient gpu screen-space ray tracing. *Journal of Computer Graphics Techniques (JCGT)*, 3(4):73–85, 2014.

[4] S. U. Mehta, K. Kim, D. Pajak, K. Pulli, J. Kautz, and R. Ramamoorthi. Filtering environment illumination for interactive physically-based rendering in mixed reality. In *EGSR (EI&I)*, pp. 107–118, 2015.

[5] M. Rey-Area, M. Yuan, and C. Richardt. 360MonoDepth: High-resolution 360deg*monoculardepthestimation.InCVPR*,2022.

[6] C. Schied, A. Kaplanyan, C. Wyman, A. Patney, C. R. A. Chaitanya, J. Burgess, S. Liu, C. Dachsbacher, A. Lefohn, and M. Salvi. Spatiotemporal variance-guided filtering: real-time reconstruction for path-traced global illumination. In *Proceedings of High Performance Graphics*, pp. 1–12. 2017.

[7] Z. Zeng, S. Liu, J. Yang, L. Wang, and L.-Q. Yan. Temporally reliable motion vectors for real-time ray tracing. In *Computer Graphics Forum*, vol. 40, pp. 79–90. Wiley Online Library, 2021.