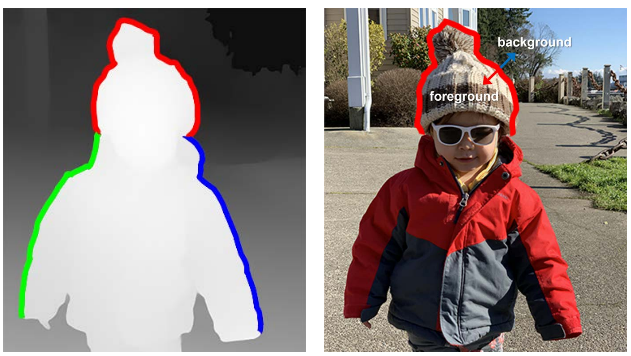
Diffuse3D: Wide-Angle 3D Photography via Bilateral Diffusion – Supplementary Material –

Yutao Jiang[1,4]*, Yang Zhou[1,4]*, Yuan Liang[1,4], Wenxi Liu[2], Jianbo Jiao[3], Yuhui Quan[1], and Shengfeng He[4]†
[1]School of Computer Science and Engineering, South China University of Technology
[2]College of Computer and Data Science, Fuzhou University
[3]School of Computer Science, University of Birmingham
[4]School of Computing and Information Systems, Singapore Management University



(a) Depth edges (b) Subproblem

Figure 1: Demonstration for depth edges and a subproblem. Different colors of edges represent different subproblems.

(a) Synthesis region (b) Local context (c) Global context

Figure 2: Demonstration for synthesis and context regions.

This supplementary material provides additional details on how to select subproblems for initial LDI and clarifies some key terms used in the process. We also provide a video demonstration for a better understanding.

*Both authors contributed equally to this research.
†Corresponding author: shengfenghe@smu.edu.sg.

1. Selecting Subproblem

To inpaint the occlusion areas in the initial LDI, we iteratively select a subproblem from the initial LDI. An iteration process can be divided into the following steps:

1. **Selecting depth edges.** The depth edges refer to the discontinuities between foreground and background. Following [1], we first sharpen the input depth to reduce blurring in the depth map. Next, we identify raw discontinuities based on the depth difference threshold between adjacent pixels. These discontinue pixels are further merged into linked segments by connected component analysis. Finally, we discard any segments that contain less than 10 pixels, resulting in the final depth edges. When the camera moves to a novel view, the occluded areas behind the foreground become visible. This happens along the depth edges. Therefore, we choose the depth edge as the basic unit of the subproblem. For instance, Figure 1a shows 3 depth edges, resulting in 3 subproblems. As a demonstration, we choose one edge for the next step (Figure 1b).

2. **Deciding synthesis areas.** To determine the region for inpainting, i.e., the area of occlusion behind the foreground, we must first identify the specific depth edge. To accomplish this, we utilize a flood-fill algorithm that starts at the edge and incrementally expands the region in four directions (left, right, up, and down) over 40 iterations. However, we must take care to prevent overlapping with the background, so we strictly constrain the expansion to the same side of the depth edge. In Figure 2a, we highlight the synthesis region corresponds to the depth edge in Figure 1b.

3. **Deciding global and local context.** Just as we determine the synthesis region, we use flood-fill expansion to identify the local context region on the opposite side

of the depth edge. The local context (Figure 2b) is the reference area situated close to the synthesis region. Following [1], we run the flood-fill algorithm for 100 iterations. On the other hand, the larger depth areas beyond the synthesis region comprise the global context (Figure 2c). Together, these context regions provide the necessary information for the inpainting process.

Once the previous steps have been completed, we employ the proposed bilateral diffusion method to perform inpainting on the synthesis region. The resulting image is then merged back into the initial LDI. We repeat this process iteratively by selecting subproblems until all depth edges have been processed. This enables us to generate a final LDI, from which we can create new viewpoints.

2. Video Demonstration

We have provided additional qualitative results in the attached video demonstration for a better understanding of our approach and its capabilities (see d3d-video.mp4) . The codec for generating video is H.264.

References

[1] Meng-Li Shih, Shih-Yang Su, Johannes Kopf, and Jia-Bin Huang. 3d photography using context-aware layered depth inpainting. In *CVPR*, pages 8028–8038, 2020. 1, 2