

# Large-scale and Drift-Free Surface Reconstruction Using Online Subvolume Registration *Supplementary Material*

## 1 Introduction

This document provides additional results to highlight the strengths and weaknesses of our approach. First, we provide high resolution images dealing with the *stonewall* sequence. We illustrate the merits of the three key steps of our proposal, namely *Subvolume Creation* (Sec. 2), *Subvolume Registration* (Sec. 3) and *Volume Blending* (Sec. 4). Then, in Sec. 5 we compare the reconstruction quality of our method to the moving volume approach [4, 15, 16, 21] and Zhou and Koltun [23], while Sec. 6 focuses on the reconstruction of the room in complete darkness, where all other state-of-the-art global optimization methods [21, 23, 24] would fail.

The reconstructions obtained by our method and the moving volume approach rely on a TSDF with voxel size 0.96cm. Zhou and Koltun [23] do not provide a voxel resolution for their available meshes.

## 2 Subvolume Creation

Sec. 4 of the main paper describes the *erosion* process, whereby the active volume in GPU memory always represents the fusion of the last  $K$  frames. Then, after every  $K$  tracked frames, a subvolume is copied back onto host main memory. In Figs 1, 2 and 3 below we show the surfaces extracted from 9 of the 55 subvolumes created by our algorithm on the *stonewall* sequence [23] using  $K = 50$ . The visualized meshes highlight how subvolumes are low-drift models of relatively large portions of the workspace.



Figure 1: Leftmost column in Fig. 6

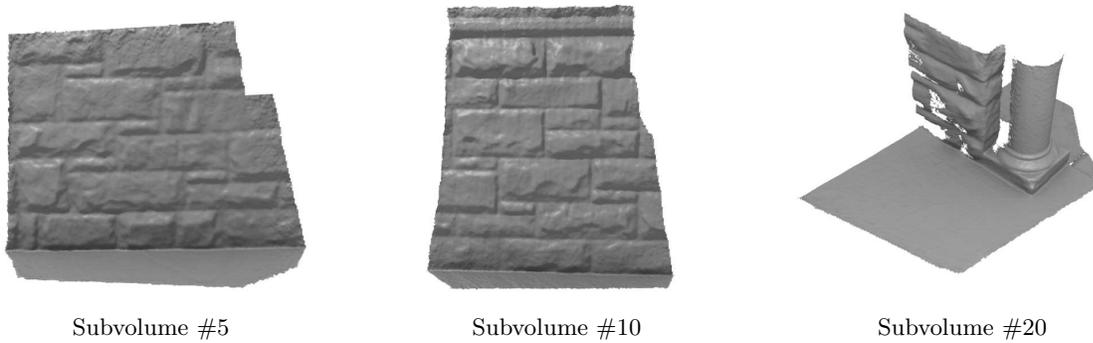


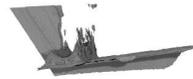
Figure 2: Front wall and rightmost column in Fig. 6



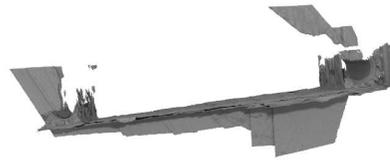
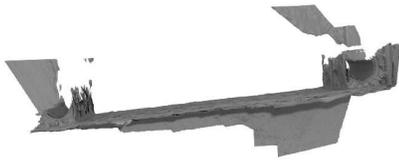
Figure 3: Leftmost column in Fig. 6 at loop closure.

### 3 Subvolume Registration

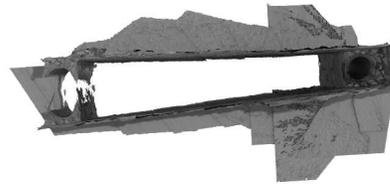
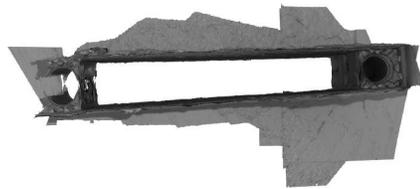
Every time a new subvolume is created, a pose optimization is run (Sec. 5 of the paper). Here we compare the surfaces extracted from subvolumes at successive time instants with (left) and without (right) the proposed pose optimization algorithm. As made clear by Figs 4 and 5, our approach keeps the drift error low even when a loop closure has not been reached yet. Please note that to highlight the specific merit of the optimization step, subvolumes have *not* been blended here.



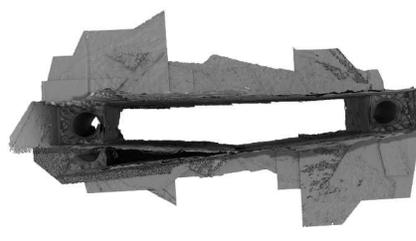
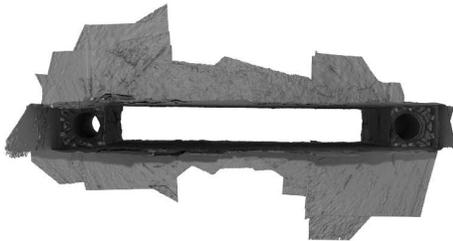
(a) First 5 subvolumes.



(b) First 20 subvolumes.



(c) First 36 subvolumes.



(d) All the 55 subvolumes.

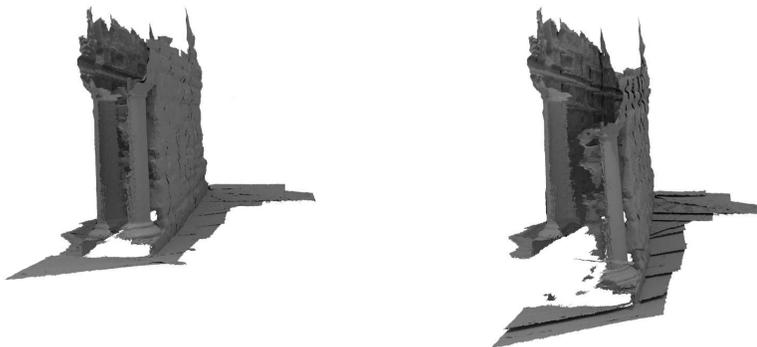
Figure 4: *Stonewall* sequence, top view. Left: with optimization; right: without optimization.



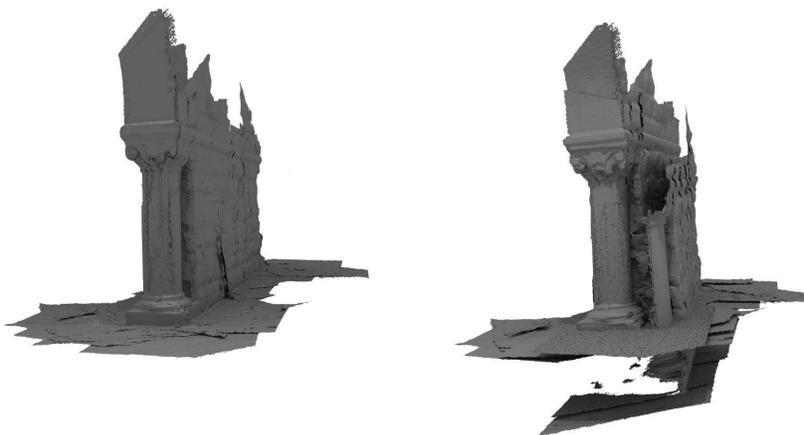
(a) First 5 subvolumes.



(b) First 20 subvolumes.



(c) First 36 subvolumes.



(d) All the 55 subvolumes.

Figure 5: *Stonewall* sequence, leftmost column. Left: with optimization; right: without optimization.

## 4 Volume Blending

Compared to the standard KinectFusion approach [14], our method initially fuses into a volumetric representation only a limited amount of depth measurements, which results in noisier surfaces (see Fig. 6, right column). The volume blending step, however, fixes this issue as can be seen in the left column of Fig. 6.

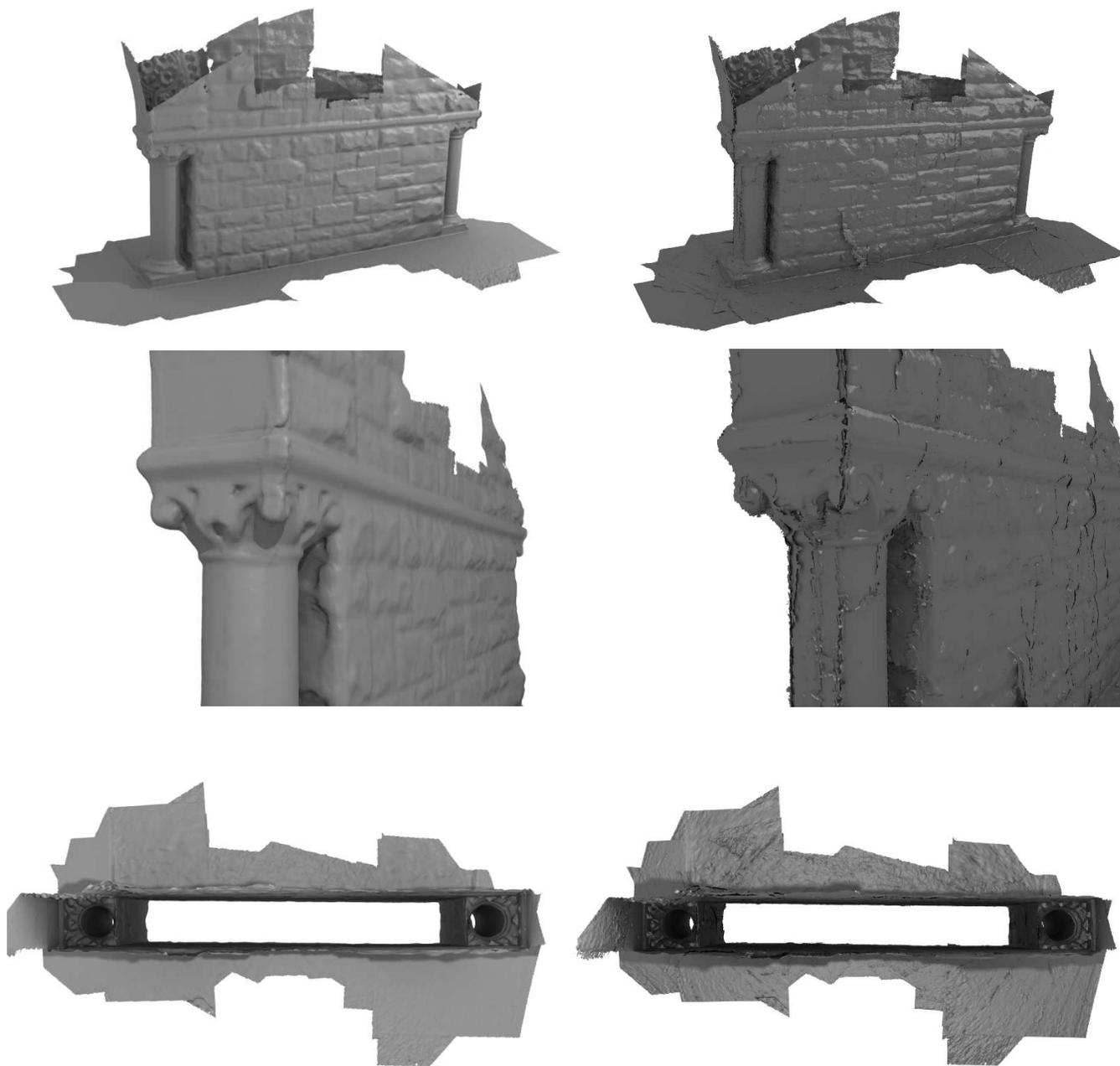


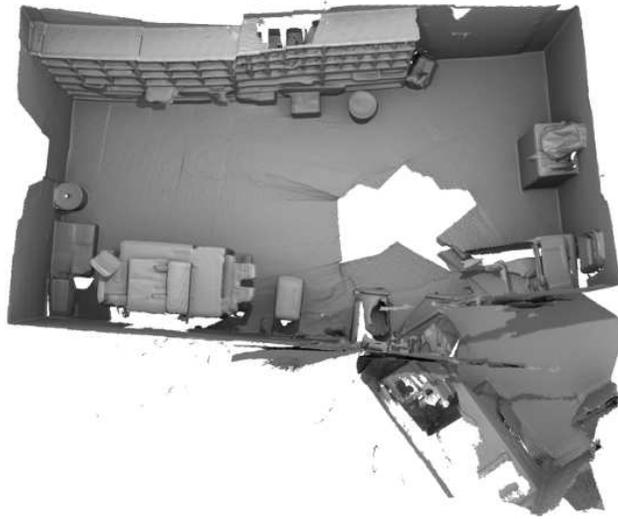
Figure 6: Reconstruction from *Stonewall* sequence. Left: with volume blending; right: without volume blending.

## 5 Comparison to other methods

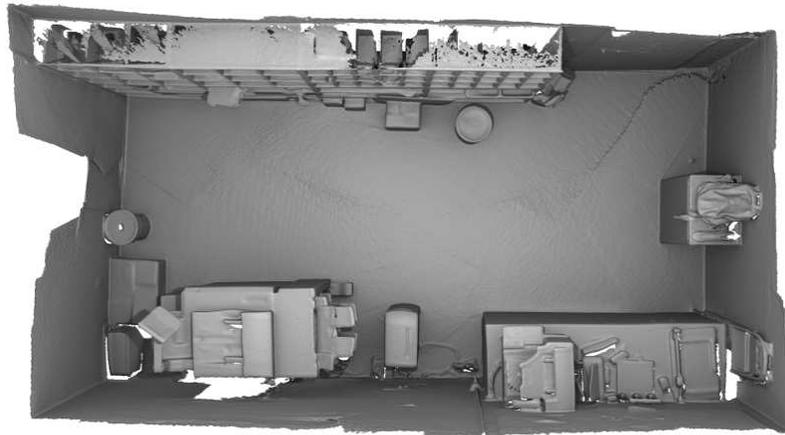
Sec. 7 of the paper reports qualitative comparative results obtained with our method, a moving volume approach [4, 15, 16, 21] and Zhou and Koltun [23] on four sequences from the dataset introduced in [23]. Here, we provide additional and/or larger images.



(a) Our approach.



(b) Moving volume approach.

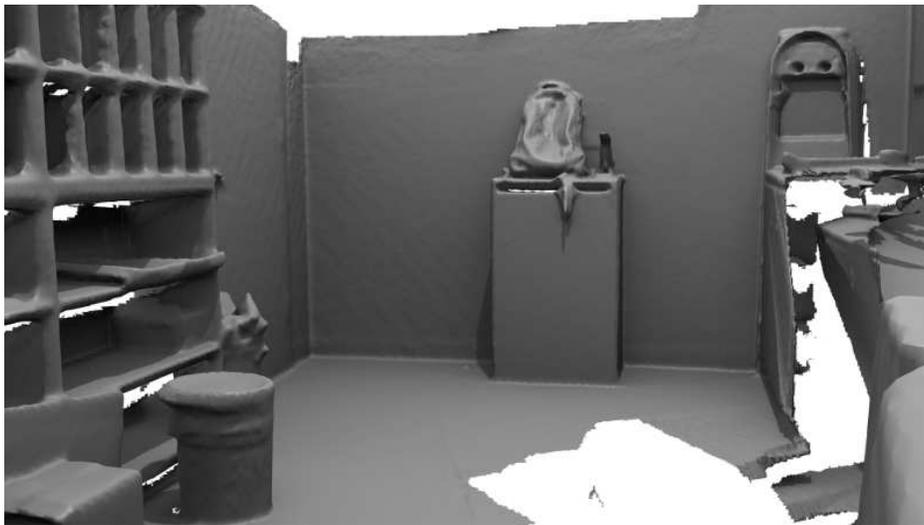


(c) Zhou and Koltun [23].

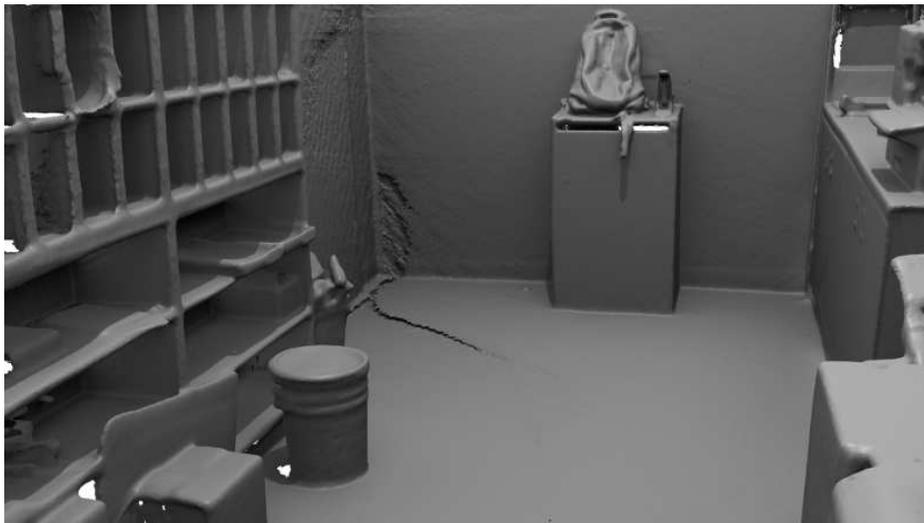
Figure 7: *Copyroom* sequence, top view. Note gross errors in moving volume and artifacts on floor for Zhou and Koltun [23].



(a) Our approach.



(b) Moving volume approach.



(c) Zhou and Koltun [23].

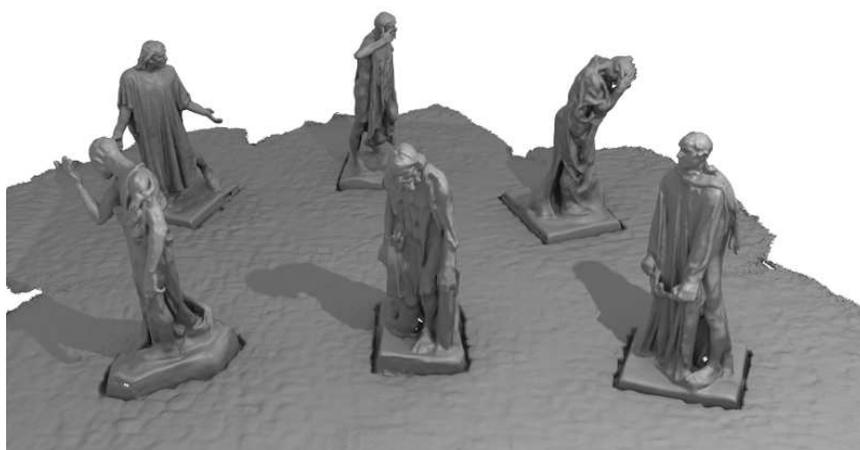
Figure 8: *Copyroom* sequence, the corner.



(a) Our approach.



(b) Moving volume approach.

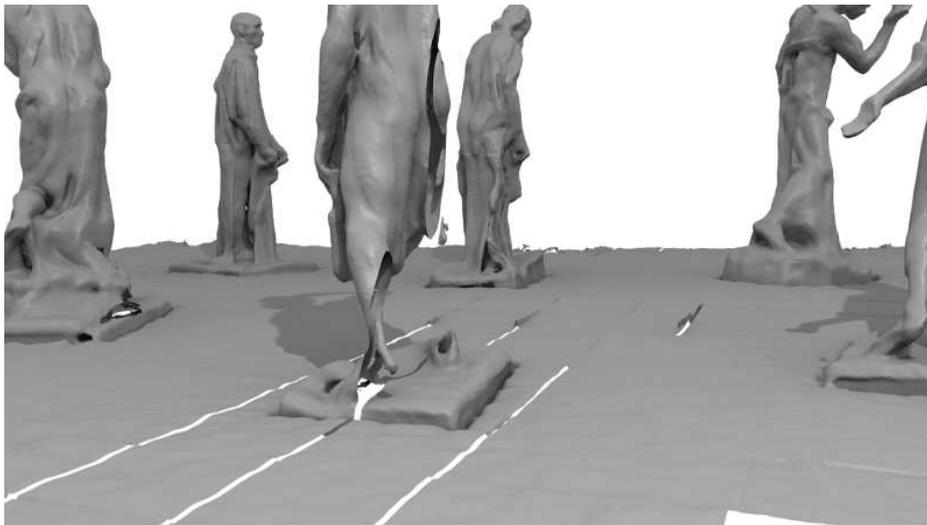


(c) Zhou and Koltun [23].

Figure 9: *Burghers* sequence, front view.



(a) Our approach.



(b) Moving volume approach.



(c) Zhou and Koltun [23].

Figure 10: *Burghers* sequence, rear view.

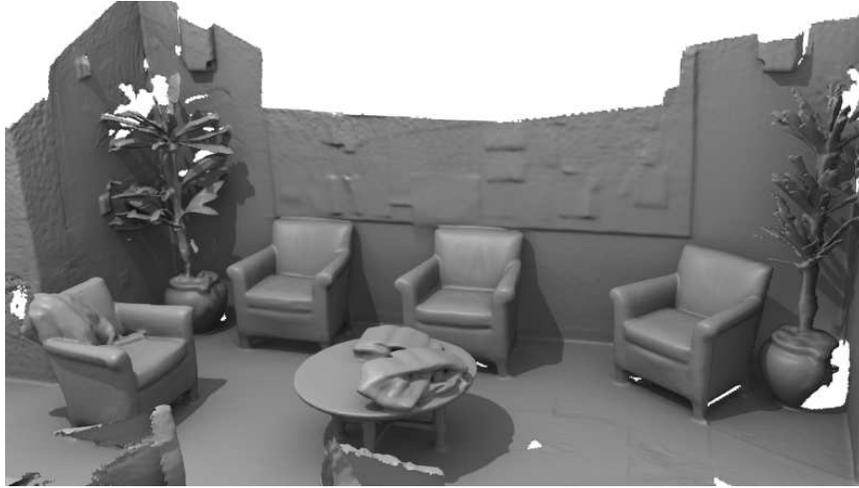


Figure 11: Our approach.

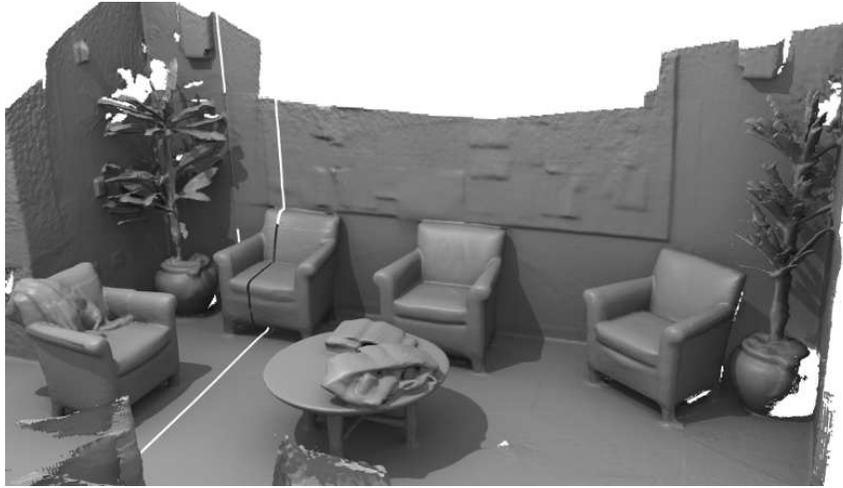


Figure 12: Moving volume approach.



Figure 13: Zhou and Koltun [23].

Figure 14: *Lounge* sequence, final reconstruction.

## 6 Reconstruction “in the dark”

Sec. 7 of the paper describes the reconstruction of a room in complete darkness. It is worth pointing out that existing RGB-D SLAM algorithms (e.g. [20, 23, 24]) include a pose optimization step that needs color data together with depth measurements, and hence would not be able to process this sequence. Here, we show larger images than those in Fig. 8 of the paper and more details from the final reconstruction.

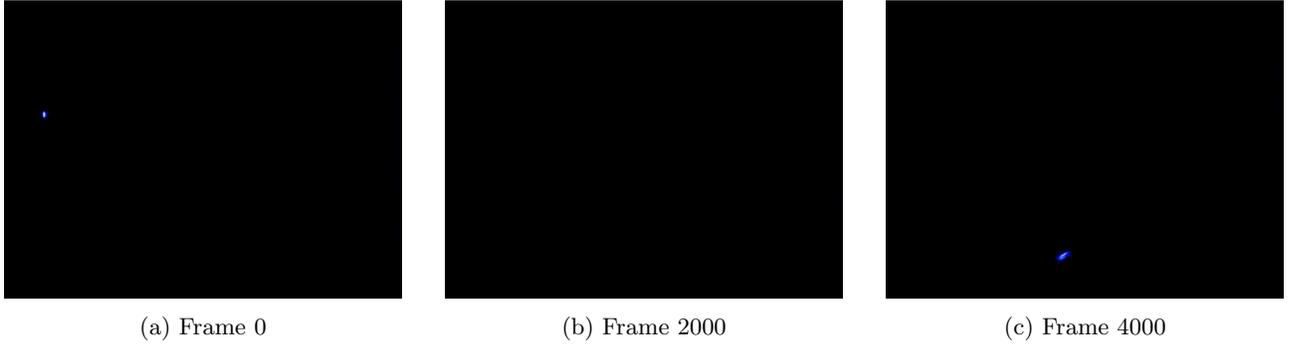


Figure 15: *Dark room* sequence: RGB frames from sequence show recording in complete darkness. All RGB-D SLAM systems would fail in this case.

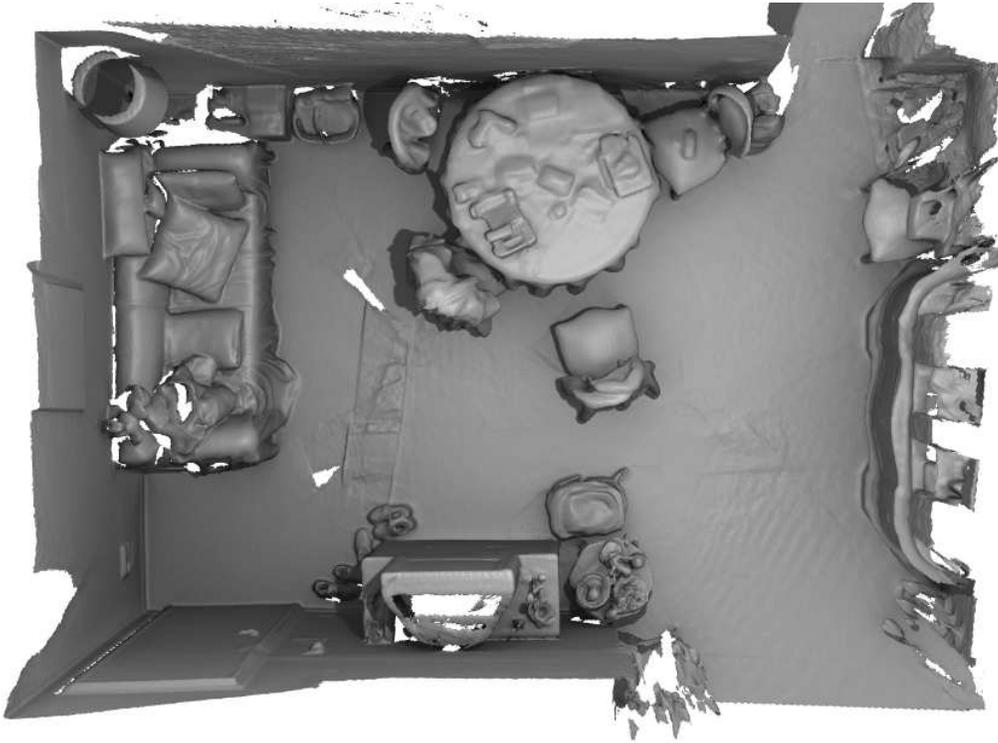


Figure 16: *Dark room* sequence, larger top view.



Figure 17: *Dark room* sequence, larger detail of the couch.



Figure 18: *Dark room* sequence, additional details.



Figure 19: *Dark room* sequence, additional details.

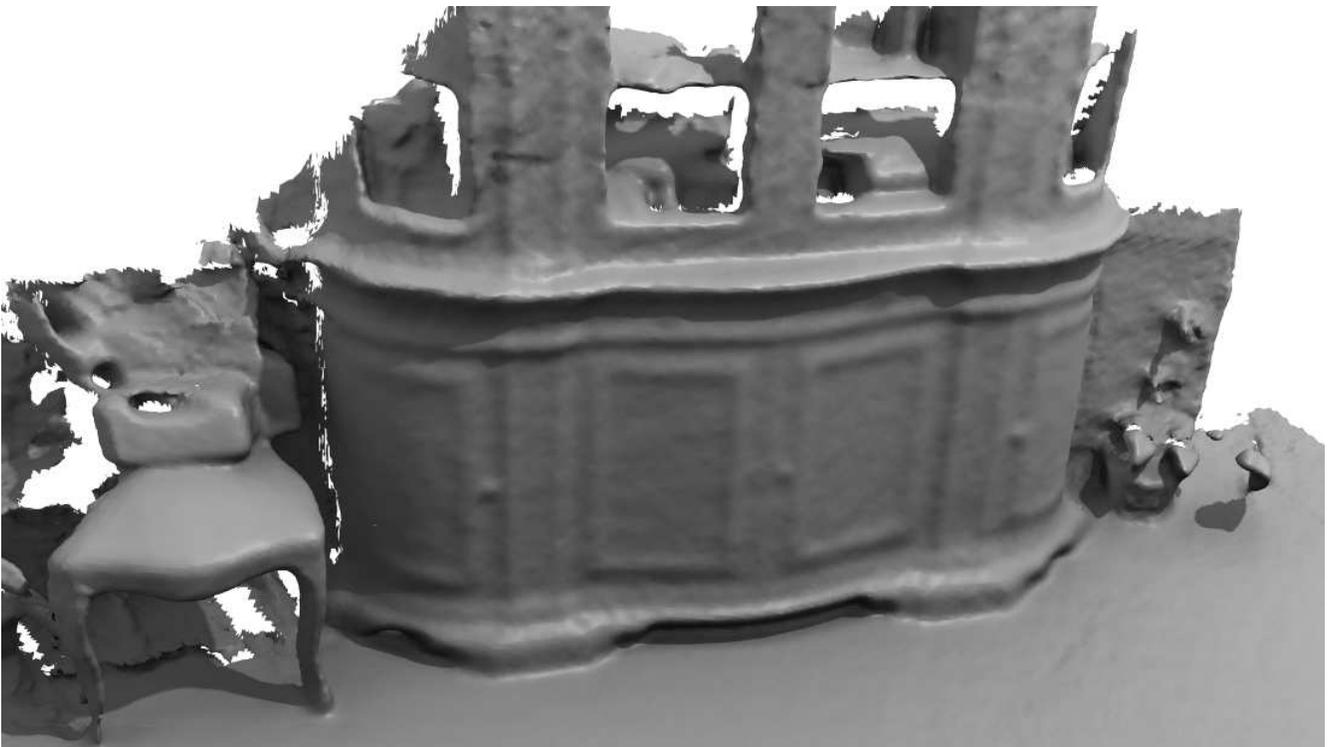


Figure 20: *Dark room* sequence, additional details.