

Here's a curiosity driven sneak peak of a real world application that this project can do with minor modifications.

The implemented Depth-Regressor does not accurately predict real world depth, as it was trained exclusively on synthetic data and used a simple U-Net architecture. Instead I hacked together the depth outputs from (1):

<https://github.com/sniklaus/3d-ken-burns> (1)

These depthmaps are voxelized instead of the ones predicted by our Unet. After calibrating my phone camera (Mi-i9) and taking several pictures of my bedroom (at high resolution) and then jointly resizing intrinsics matrix and pictures, I attained for the following image and corresponding results:

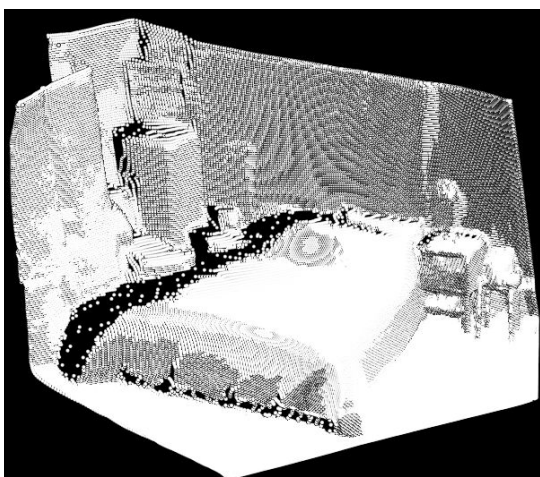
Resized to 320x240 (WxH)



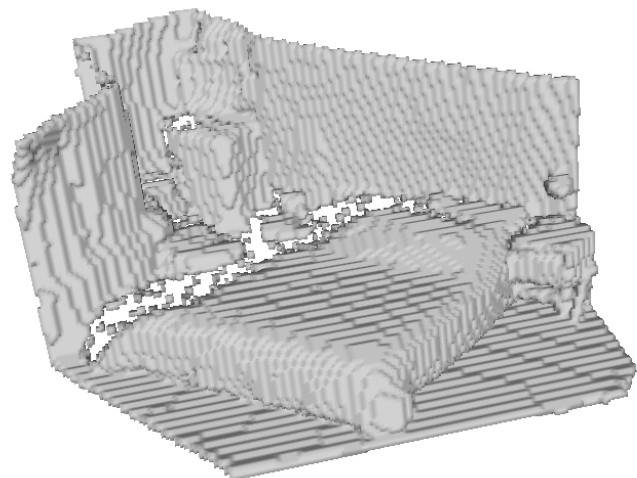
Depth (via (1))



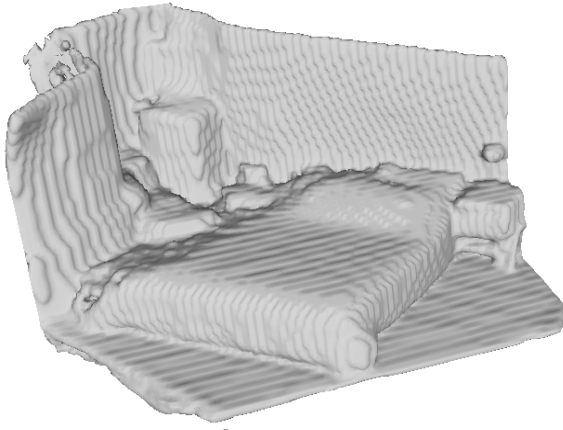
Projection with intrinsics



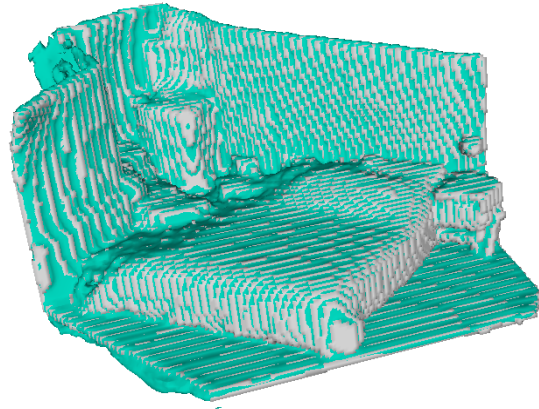
differentially voxelized



Prediction



Prediction (teal) & Voxelized Pointcloud



Note that the If-Net here has not been jointly trained with the depth regressor of (1), and could therefore in theory achieve much better results. Furthermore, we plan on completing the voxel grid via a scene-completion network before feeding it into the If-Net for refined results.