ETH zürich IIII Google Microsoft

SceneFun3D: Fine-Grained Functionality and Affordance Understanding in 3D Scenes

Introduction

"Open the drawer"



"Turn off the floor lamp"



Motivation

- coarse object level.



lement annotations



We introduce **SceneFun3D**, the first large-scale dataset with **highly accurate** interaction annotations in 3D real-world indoor environments

Building the SceneFun3D dataset



Alexandros Delitzas¹, Ayça Takmaz¹, Federico Tombari^{2,3}, Robert Sumner¹, Marc Pollefeys^{1,4}, Francis Engelmann^{1,2}

¹ETH Zürich ²Google ³TUM ⁴Microsoft

Existing methods are heavily focused on understanding 3D scenes on a

In real-world applications, agents need to interact with the functional interactive elements (e.g., handles, knobs, buttons) and reason about their purpose in the scene context to successfully complete tasks.

The SceneFun3D dataset





Registration and	d alignment	Re



Hires RGB-D 3D iPad camera \pose refinement Reconstruction



Global reg.







