Multi-user, Scalable 3D Object Detection in AR Cloud

Siddharth Choudhary, Nitesh Sekhar, Siddharth Mahendran, Prateek Singhal
Magic Leap, Inc
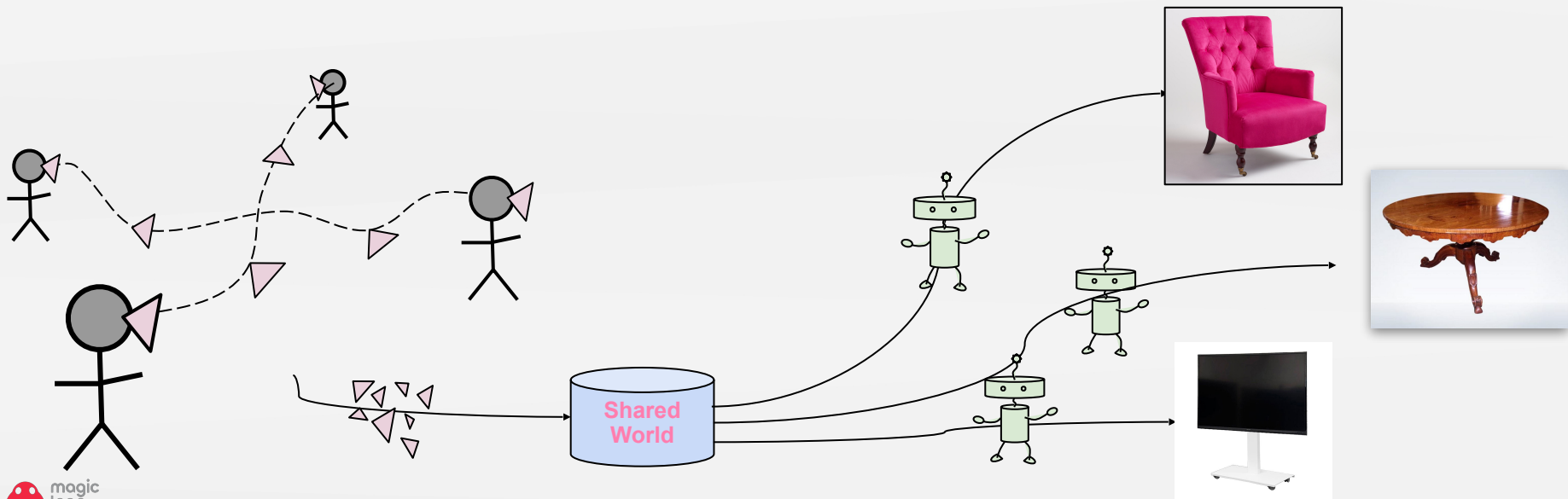
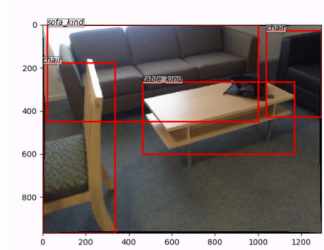# Multi User, Across Time, Environment Generation

Approach

RGB frame

Depth frame

Voxelize detection masks

Load object visible in
the current frustum

Associate and Fuse

Approach

RGB frame

Depth frame

Voxelize detection masks

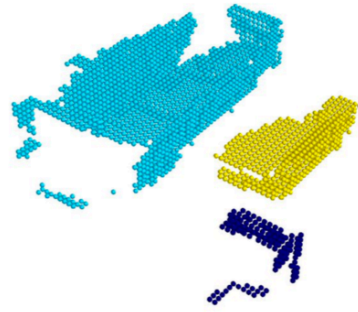Load object visible in
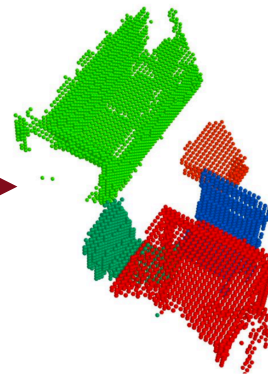the current frustum

Associate and Fuse
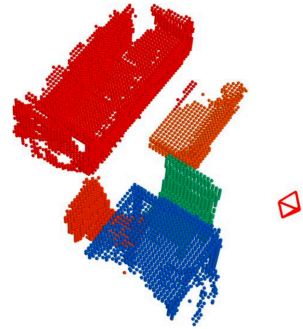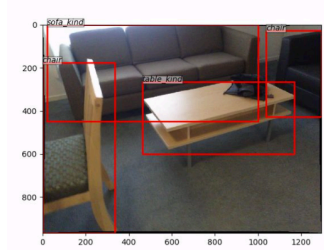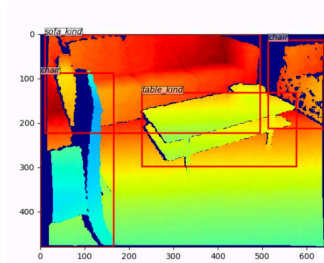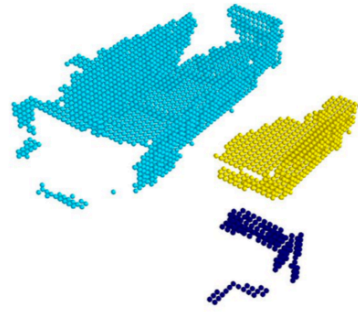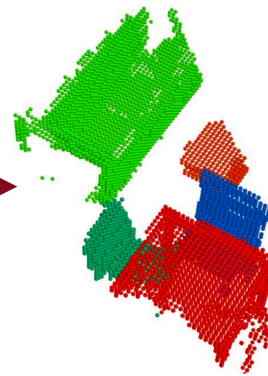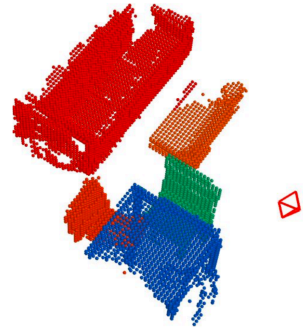
Approach

RGB frame

Depth frame

Voxelize detection masks
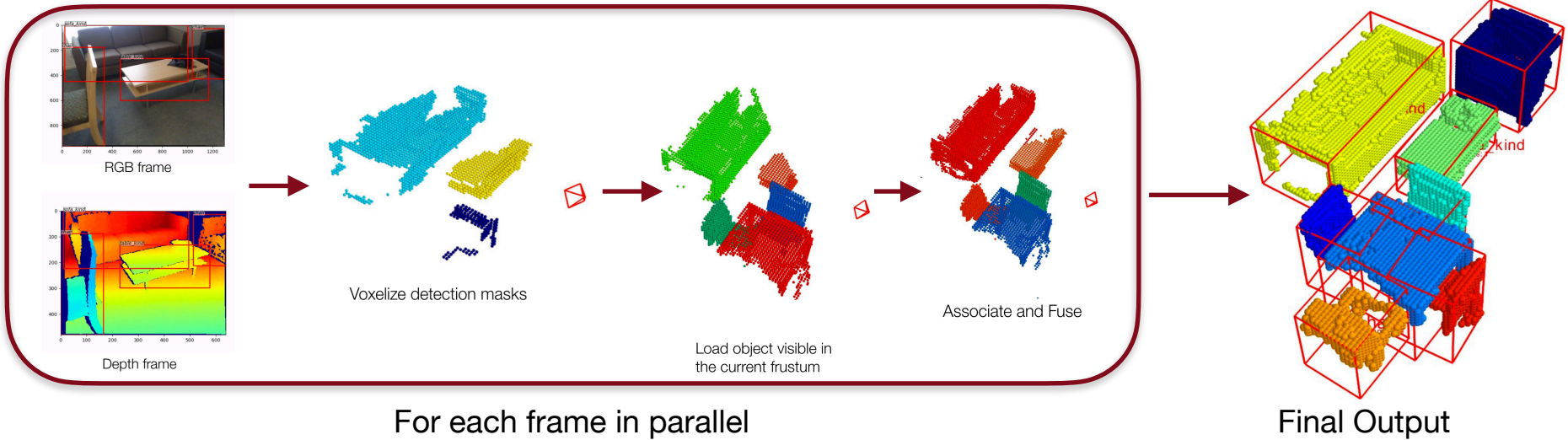
Load object visible in
the current frustum

Associate and Fuse

For each frame in parallel

Final Output

Note: We only store the object voxels, not the background mesh

Why is it scalable?

RGB-D Frames

Shared World

Multi-User

Each frame can be processed
on an independent node

Each object can be updated
asynchronously without locking
all the objects in the map

Why is it scalable?

RGB-D Frames

Multi-User

Shared World

Scalable in number of users

Scalable in number of objects
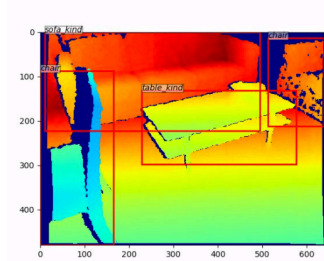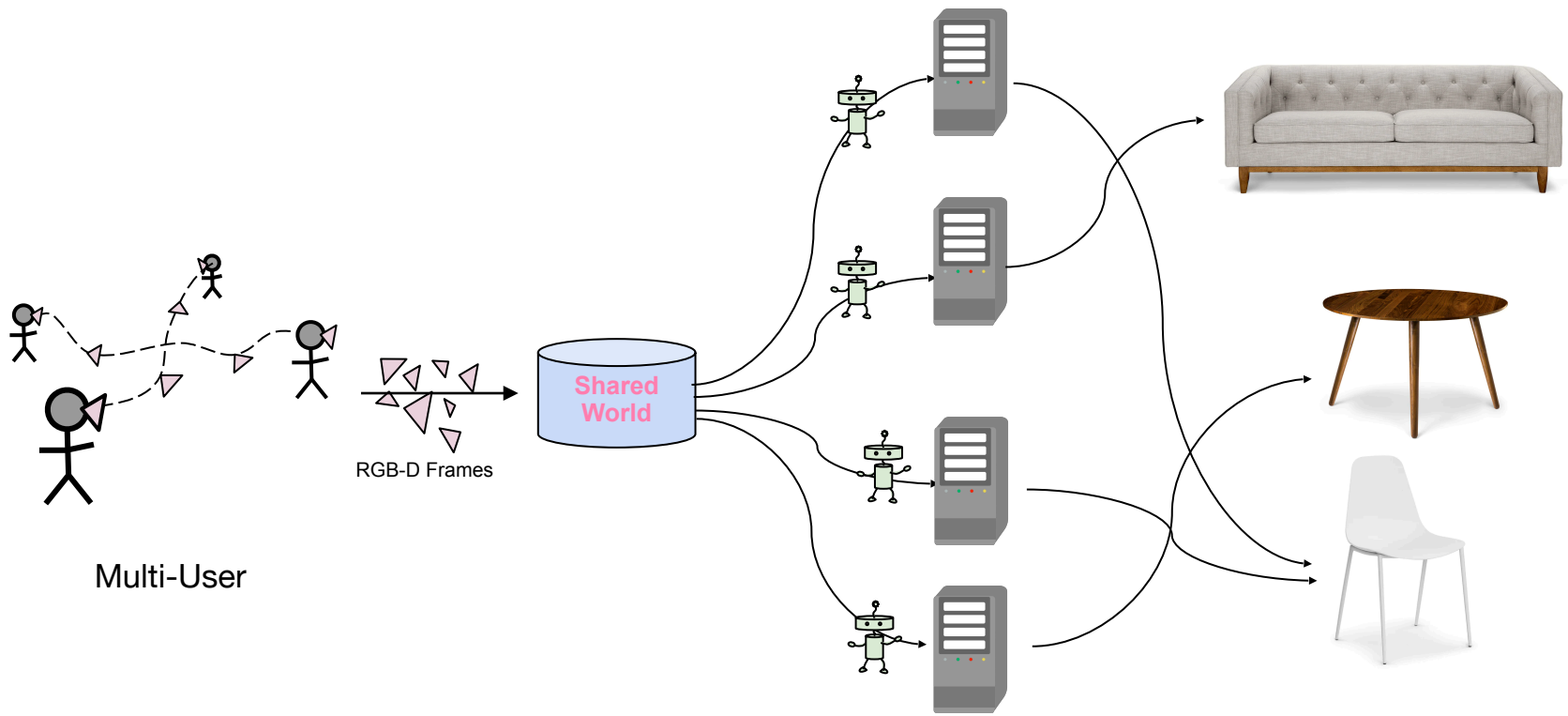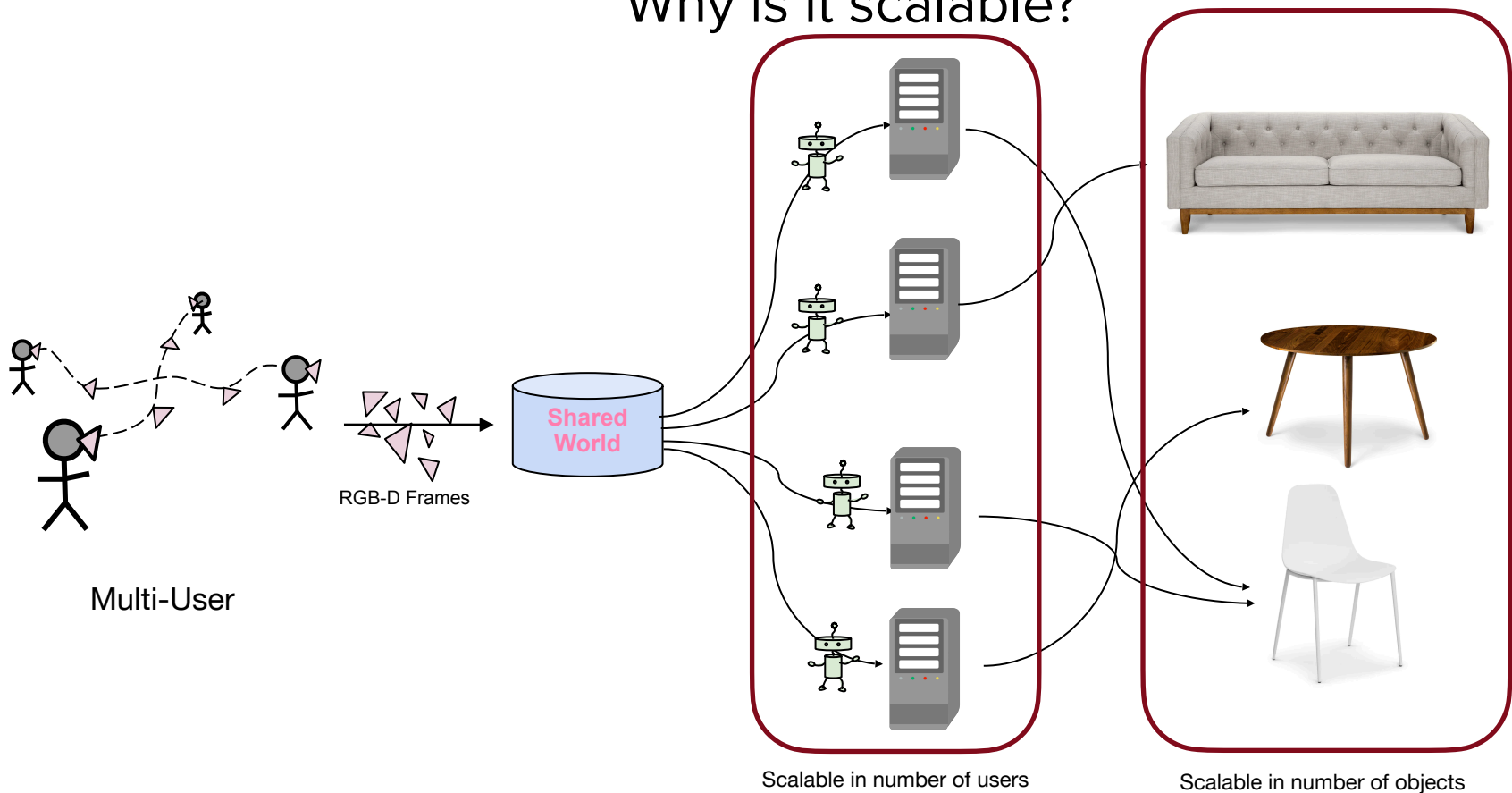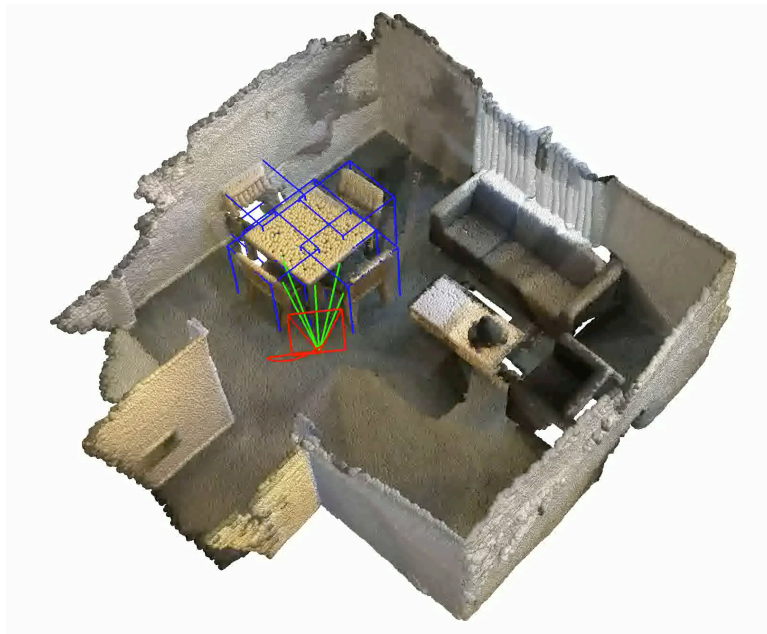
Dataset

Single-user dataset
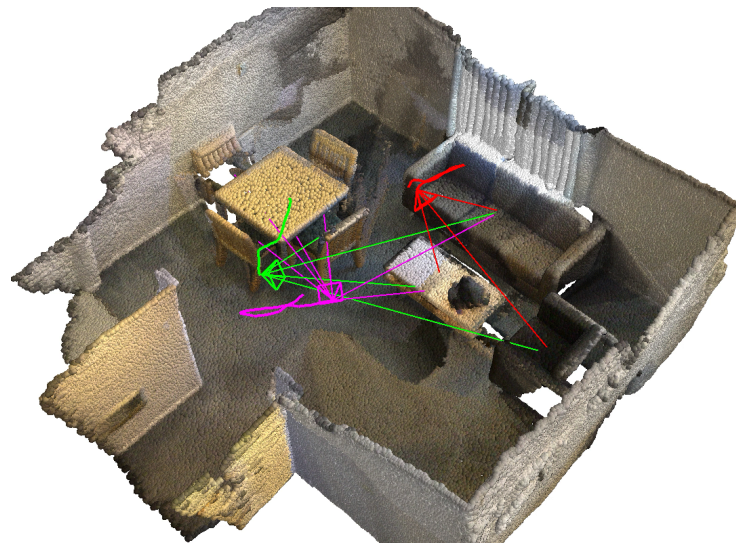
Multi-user dataset

Results for a single scene

1 user

10 users

50 users

Red: Estimated Bounding Box
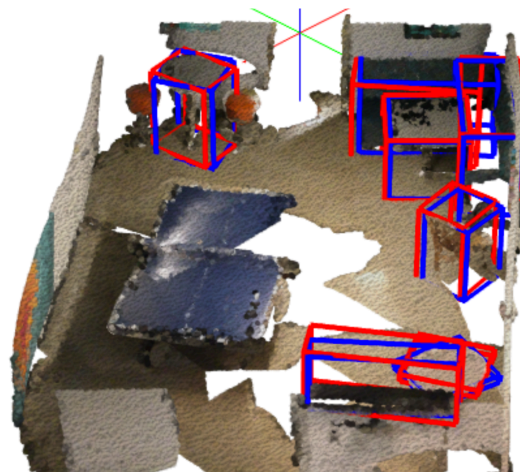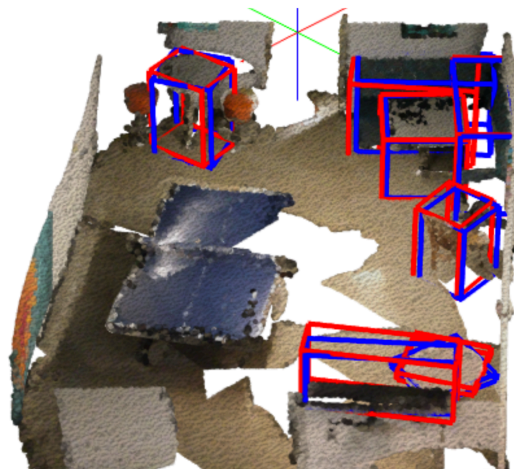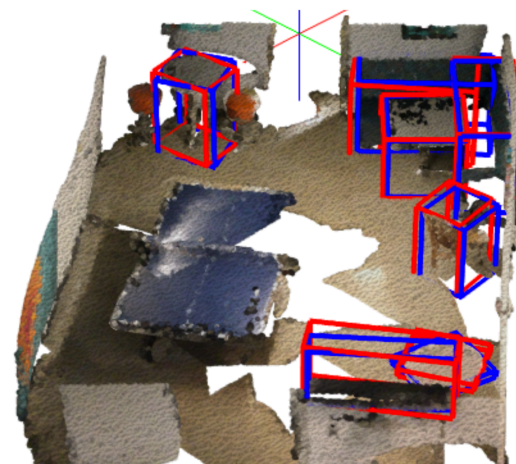**Blue**: Groundtruth Bounding Box

Mean Average Precision at IoU=0.25
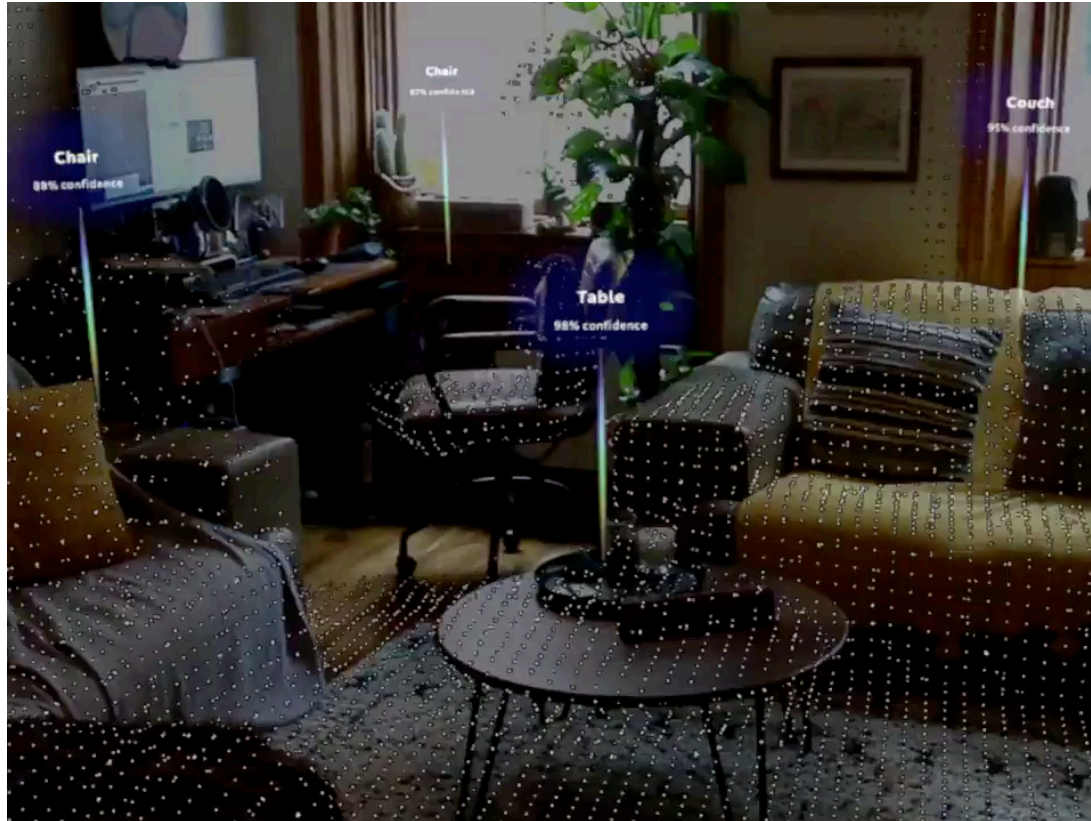
Num of Users	mAP@0.25	Variance
1	0.699	0
10	0.646	0.0027
50	0.672	0.0025
100	0.664	0.0044

Metrics are computed over 40 ScanNet scenes
Variance is computed across 10 Monte-Carlo runs

Memory Requirement

Metric	Dense Mesh	Object Level Map	Ratio
Mean	201.3	12.001	16.77
Median	198.16	11.63	17.07

Object level map also includes the memory required to store the sparse map

Shipped with latest update to Magic Leap 1

Thank you!