

MOT & Reconstruction W13

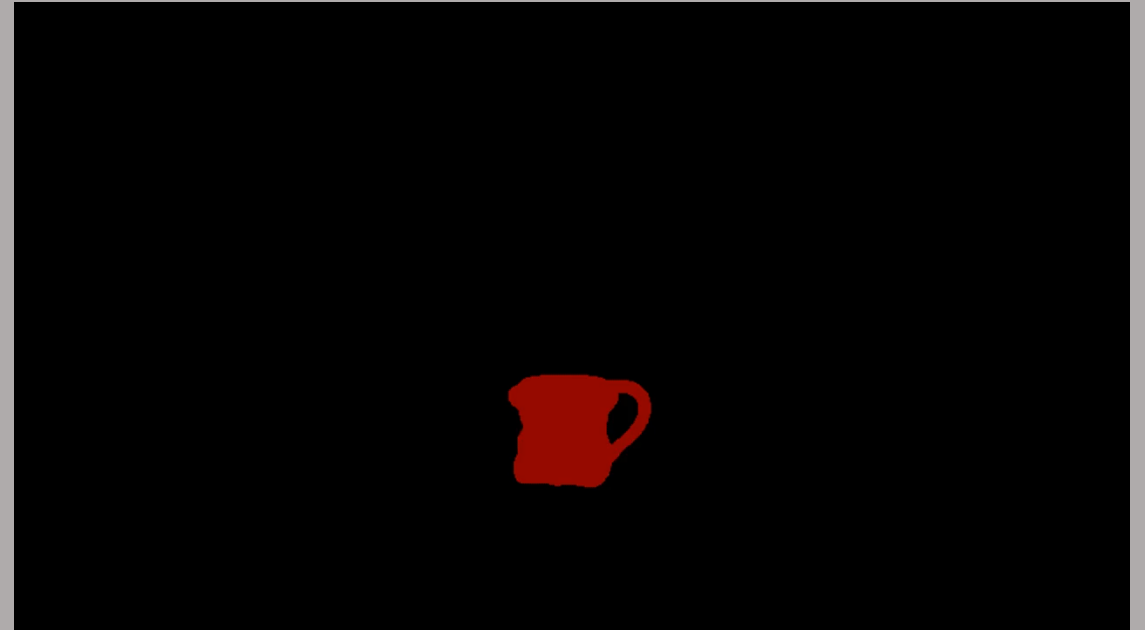
- Recap week 12 progresses
- Poses and reconstructions scores algorithms
 - ADD Score
 - CD Score
- New papers

MOT & Reconstruction W12

- Recorded 21 new videos
 - RGBD using the Kinect Azure
 - 500 frames each
 - 1280x720
 - 5 objects:
 - “Teddy mug”
 - Spray-paint can
 - Metal mug
 - Glass mug
 - CD-ROM

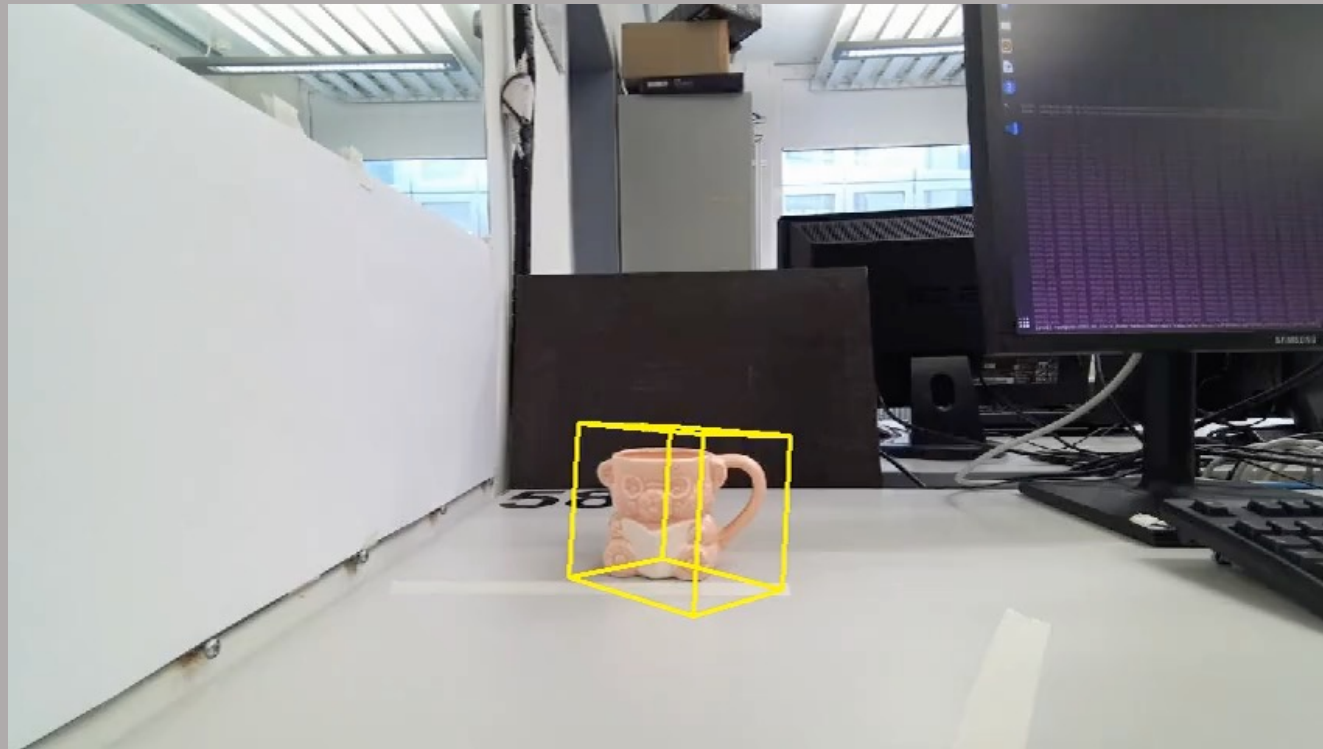
MOT & Reconstruction W12

Teddy mug



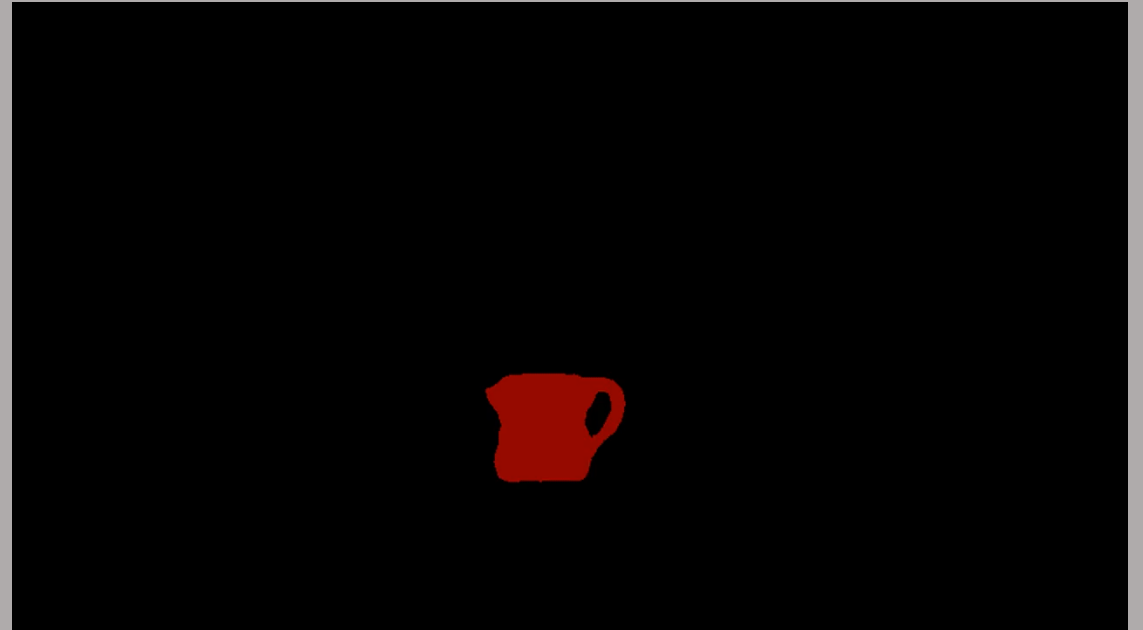
MOT & Reconstruction W12

Teddy mug



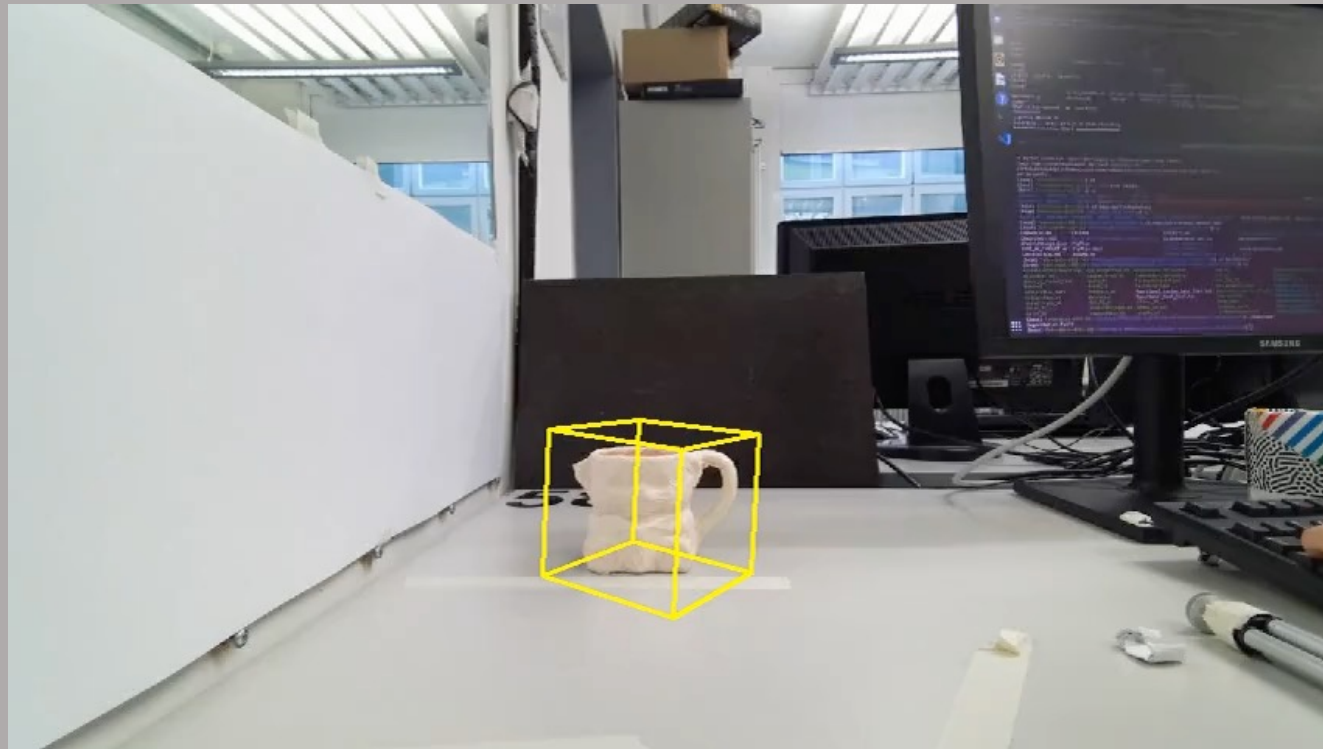
MOT & Reconstruction W12

Teddy mug



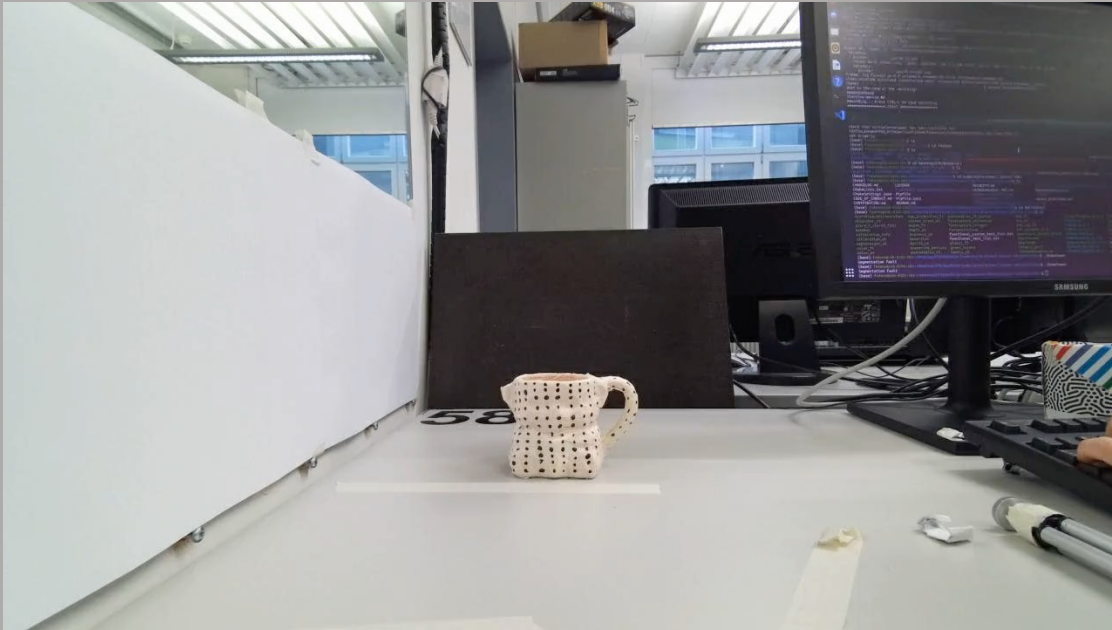
MOT & Reconstruction W12

Teddy mug



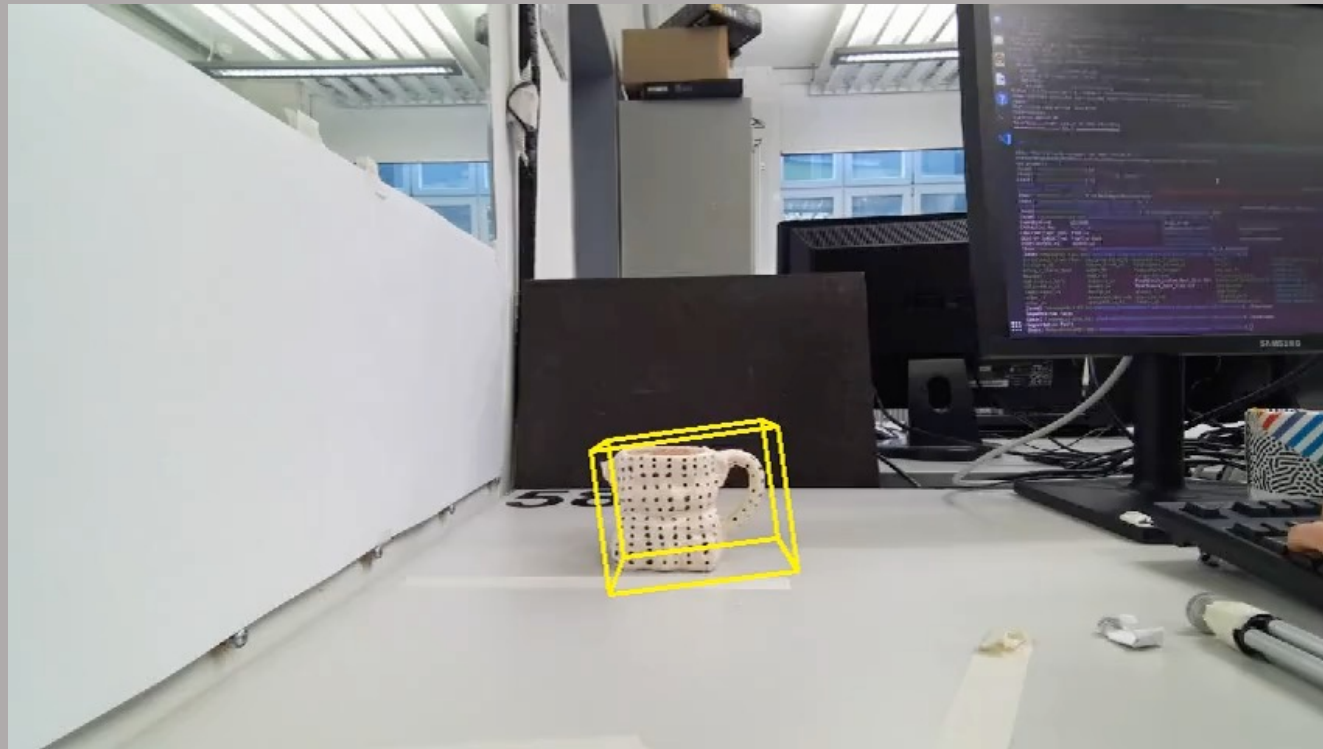
MOT & Reconstruction W12

Teddy mug



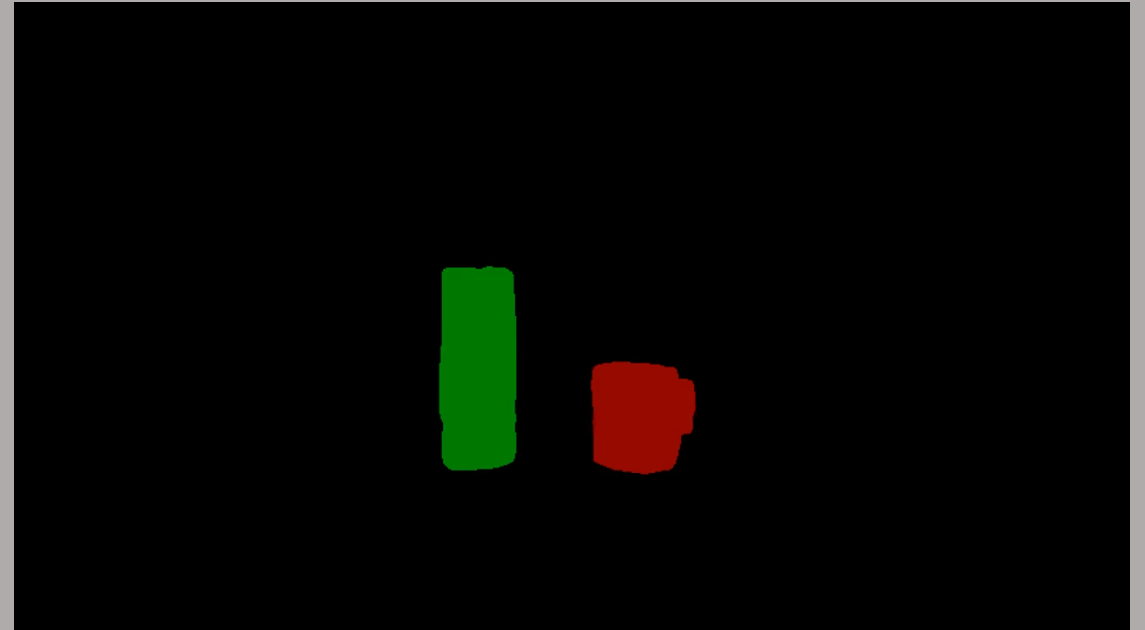
MOT & Reconstruction W12

Teddy mug



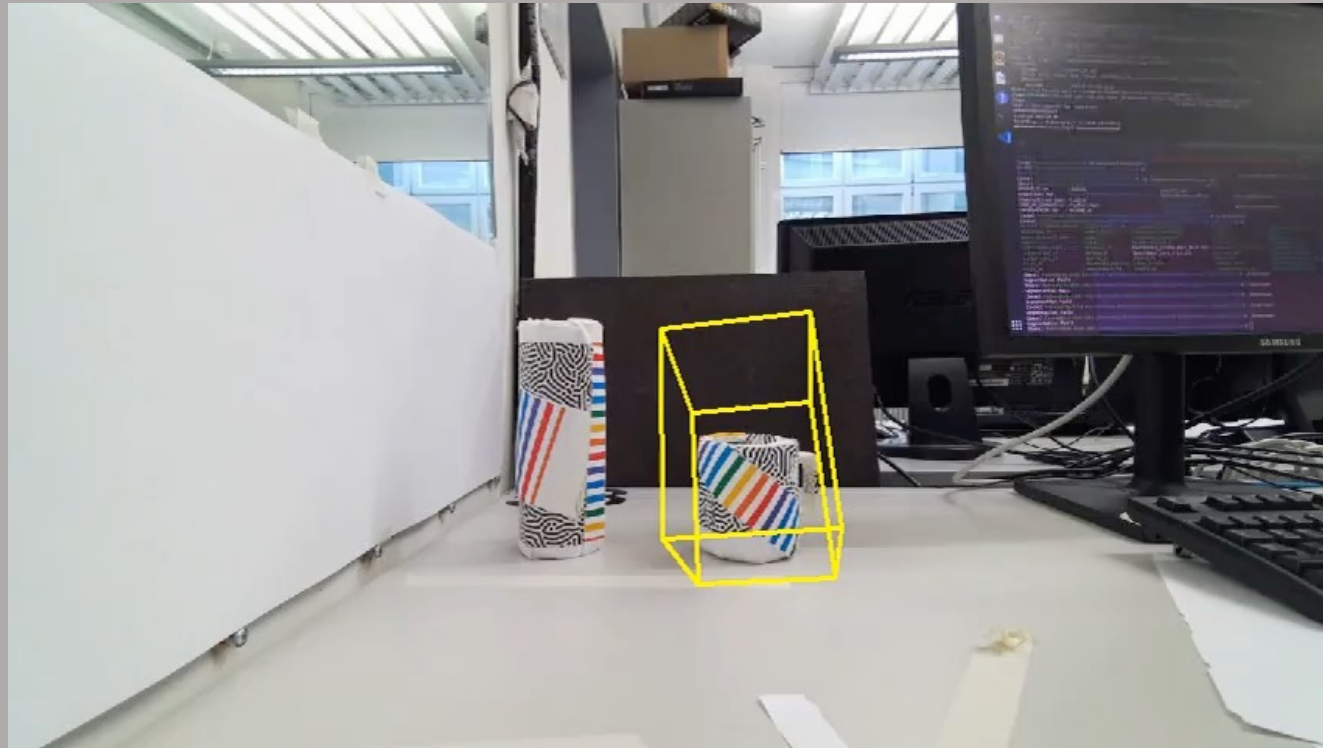
MOT & Reconstruction W12

Spray-paint with mug



MOT & Reconstruction W12

Spray-paint with mug



Pose Score

- BundleSDF uses ADD-S and ADD scores metrics during the evaluation
 - Average distance 3D, point to point
 - Initial standard pose in BundleSDF

```
local > home > fedona > Desktop > ETH > Bachelor > Results > LinemodTest > 000004 > bundlesdfLM_4 > {} scene_prediction.json > [ ] 138 > {} 0 > [ ] cam_R_m2c > #
1 {
2   "20": [{"cam_R_m2c": [1, -0, 0, -0, 1, -0, 0, -0, 1], "cam_t_m2c": [0.05610579997, -0.07472649962, 0.718210876],
3   "27": [{"cam_R_m2c": [0.9093547463, 0.2137267739, -0.3569227159, -0.3390604258, 0.877884686, -0.3381656706, 0.241
4   "33": [{"cam_R_m2c": [0.8994082808, 0.1897131652, -0.39379403, -0.220684886, 0.9747365713, -0.0344481878, 0.37731
5   "35": [{"cam_R_m2c": [0.8812286854, 0.2069754899, -0.4249669909, -0.2400559634, 0.9704332352, -0.02515137568, 0.4
6   "44": [{"cam_R_m2c": [0.9093547463, 0.2137267739, -0.3569227159, -0.3390604258, 0.877884686, -0.3381656706, 0.241

local > home > fedona > Desktop > ETH > Bachelor > Datasets > LinemodTest > 000004 > {} scene_gt.json > [ ] 33 > {} 0 > [ ] cam_R_m2c > # 1
1 {
2   "20": [{"cam_R_m2c": [-0.37412, 0.904486, -0.20479301, 0.83160299, 0.229463, -0.50575, -0.41045099, -0.35951701
3   "27": [{"cam_R_m2c": [0.0124993, 0.99848402, 0.0536055, 0.99687397, -0.00826068, -0.0785758, -0.0780139, 0.0544
4   "33": [{"cam_R_m2c": [0.044085, 0.99476701, 0.0921647, 0.88371199, 0.0041978, -0.46801299, -0.465951, 0.102079,
5   "35": [{"cam_R_m2c": [0.0831454, 0.98718798, 0.136188, 0.87734503, -0.00770351, -0.47979701, -0.472601, 0.15937
6   "44": [{"cam_R_m2c": [-0.35177401, 0.90803897, -0.227421, 0.833709, 0.193441, -0.517214, -0.425657, -0.371546,
7   "47": [{"cam_R_m2c": [-0.374679, 0.92411, -0.0750765, 0.86028302, 0.31631801, -0.39982, -0.34573001, -0.2143909
```

Pose Score

- BundleSDF uses ADD-S and ADD scores metrics during the evaluation
 - Average distance 3D, point to point
 - Initial standard deviation

ADD score always = 0

```
local > home > fedona > Desktop > ETH > { } scene_gt.json > [ ]33 > { }0 > [ ]cam_R_m2c > #1
1 {
2   "20": [{"cam_R_m2c": [0.72649962, 0.718210876],
3   "27": [{"cam_R_m2c": [0.86, -0.3381656706, 0.241
4   "33": [{"cam_R_m2c": [0.23, -0.0344481878, 0.37731
5   "35": [{"cam_R_m2c": [0.8812286854, 0.2069754899, -0.4249669909, -0.2400559634, 0.9704332352, -0.02515137568, 0.4
6   "44": [{"cam_R_m2c": [0.0001514778, 0.02101513501, 0.1452556252, 0.02476740383, 0.0005237725, 0.01705740042, 0

local > home > fedona > Desktop > ETH > Bachelor > Datasets > LinemodTest > 000004 > { } scene_gt.json > [ ]33 > { }0 > [ ]cam_R_m2c > #1
1 {
2   "20": [{"cam_R_m2c": [-0.37412, 0.904486, -0.20479301, 0.83160299, 0.229463, -0.50575, -0.41045099, -0.35951701
3   "27": [{"cam_R_m2c": [0.0124993, 0.99848402, 0.0536055, 0.99687397, -0.00826068, -0.0785758, -0.0780139, 0.0544
4   "33": [{"cam_R_m2c": [0.044085, 0.99476701, 0.0921647, 0.88371199, 0.0041978, -0.46801299, -0.465951, 0.102079,
5   "35": [{"cam_R_m2c": [0.0831454, 0.98718798, 0.136188, 0.87734503, -0.00770351, -0.47979701, -0.472601, 0.15937
6   "44": [{"cam_R_m2c": [-0.35177401, 0.90803897, -0.227421, 0.833709, 0.193441, -0.517214, -0.425657, -0.371546,
7   "47": [{"cam_R_m2c": [-0.374679, 0.92411, -0.0750765, 0.86028302, 0.31631801, -0.39982, -0.34573001, -0.21439009
```

Pose Score

- BundleSDF uses ADD-S and ADD scores metrics during the evaluation
 - Average distance
 - Initial standard deviation

Should I align the first frame predicted pose with the first ground truth pose?

```
local > home > fedona > Desktop > ETH
1 {
2   "20": [{"cam_R_m2c": [0.72649962, 0.718210876],
3   "27": [{"cam_R_m2c": [0.086, -0.3381656706, 0.241
4   "33": [{"cam_R_m2c": [0.0, -0.0344481878, 0.37731
5   "35": [{"cam_R_m2c": [0.2352, -0.02515137568, 0.4
6   "44": [{"cam_R_m2c": [0.37735, -0.01795740042, 0.

local > home > fedona > Desktop > ETH
1 {
2   "20": [{"cam_R_m2c": [-0.37412, 0.904486, -0.20479301, 0.83160299, 0.229463, -0.50575, -0.41045099, -0.35951701]
3   "27": [{"cam_R_m2c": [0.0124993, 0.99848402, 0.0536055, 0.99687397, -0.00826068, -0.0785758, -0.0780139, 0.0544
4   "33": [{"cam_R_m2c": [0.044085, 0.99476701, 0.0921647, 0.88371199, 0.0041978, -0.46801299, -0.465951, 0.102079,
5   "35": [{"cam_R_m2c": [0.0831454, 0.98718798, 0.136188, 0.87734503, -0.00770351, -0.47979701, -0.472601, 0.15937
6   "44": [{"cam_R_m2c": [-0.35177401, 0.90803897, -0.227421, 0.833709, 0.193441, -0.517214, -0.425657, -0.371546,
7   "47": [{"cam_R_m2c": [-0.374679, 0.92411, -0.0750765, 0.86028302, 0.31631801, -0.39982, -0.34573001, -0.21439009]
```

Reconstruction Score

- BundleSDF uses Chamfer Distance

- $CD(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} ||a - b|| + \frac{1}{|B|} \sum_{b \in B} \min_{a \in A} ||b - a||$

Reconstruction Score

- BundleSDF uses Chamfer Distance
- $CD(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} ||a - b|| + \frac{1}{|B|} \sum_{b \in B} \min_{a \in A} ||b - a||$
 - What if a mesh is not “centered”?
 - Compute centroid fails with artifacts...

BundleSDF fail

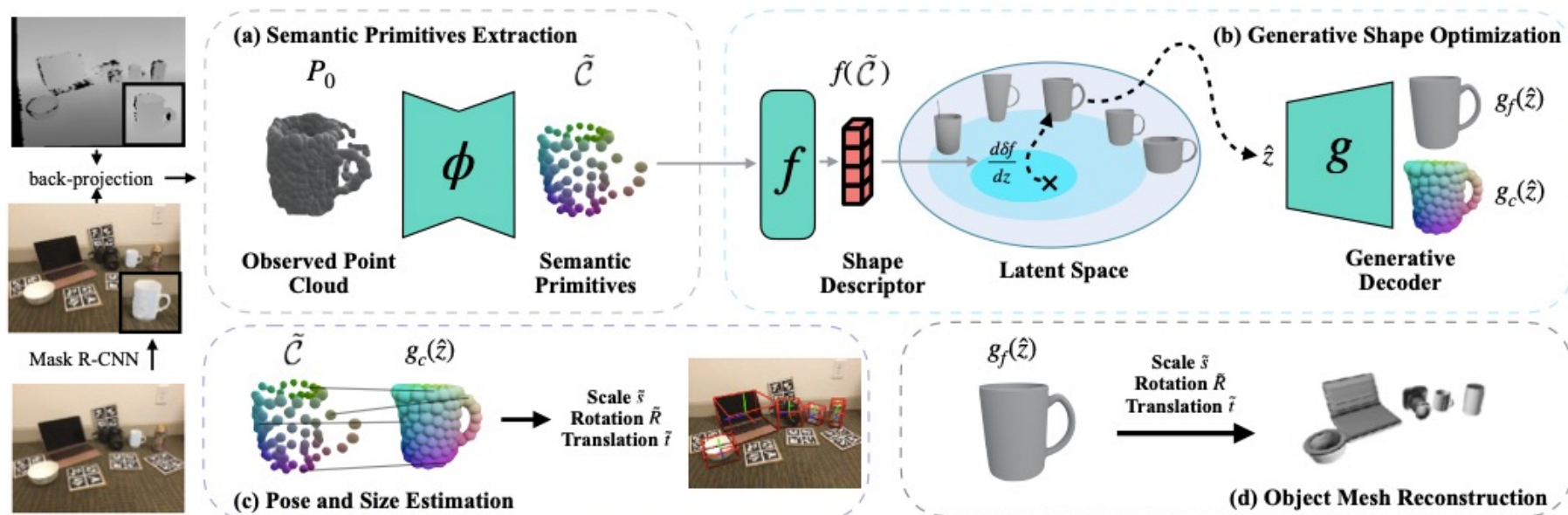
- The code did fail sometimes to reconstruct the objects
 - "bug fixed" with new version of the code (Nov 4th)
 - Much slower to run
 - Every time it has too few matches it interrupts and asks if should continue
 - On Linemod Dataset keeps interrupting
 - Still have to try modify the code and re-run on the videos

New papers?

- Hard to find something similar to BundleSDF
 - There would be BundleTrack (-> fails to build)
- Simpler to find codes for either pose estimation or 3D reconstruction
 - Pose estimation:
 - Gasp: Generative Category-Level Shape and Pose Estimation with Semantic Primitives
 - DOPE: Deep Object Pose Estimation - ROS Inference
 - HybridPose: 6D Object Pose Estimation under Hybrid Representations

Gasp

- Shape Estimation is focused on Object semantics (recognize an object between a fixed amount of categories)

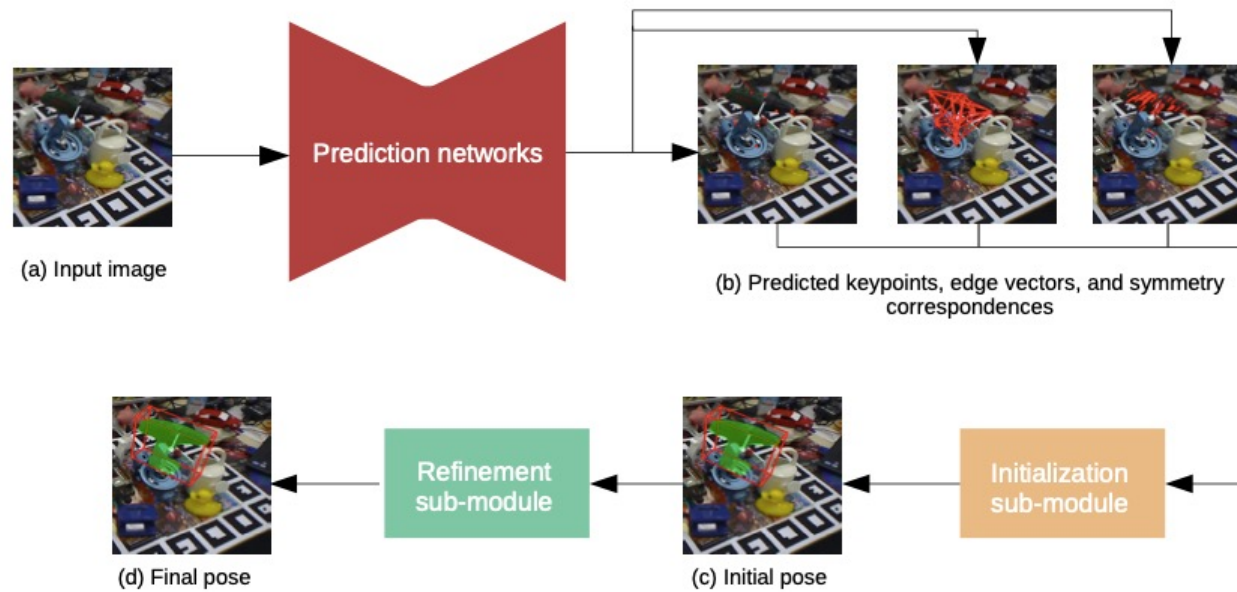


DOPE

- Nvidia Project, 2018
- RGB

Hybrid Pose

- 2020
- Tested on Linemod
- Real Time, SOTA performance



Other Papers?

- KinectFusion
 - No official code available?
- DPOD: Dense Pose Object Detector