

# T-LESS: An RGB-D Dataset for 6D Pose Estimation of Texture-less Objects

**Tomáš Hodaň<sup>1</sup>, Pavel Haluza<sup>1</sup>, Štěpán Obdržálek<sup>1</sup>,  
Jiří Matas<sup>1</sup>, Manolis Lourakis<sup>2</sup>, Xenophon Zabulis<sup>2</sup>**



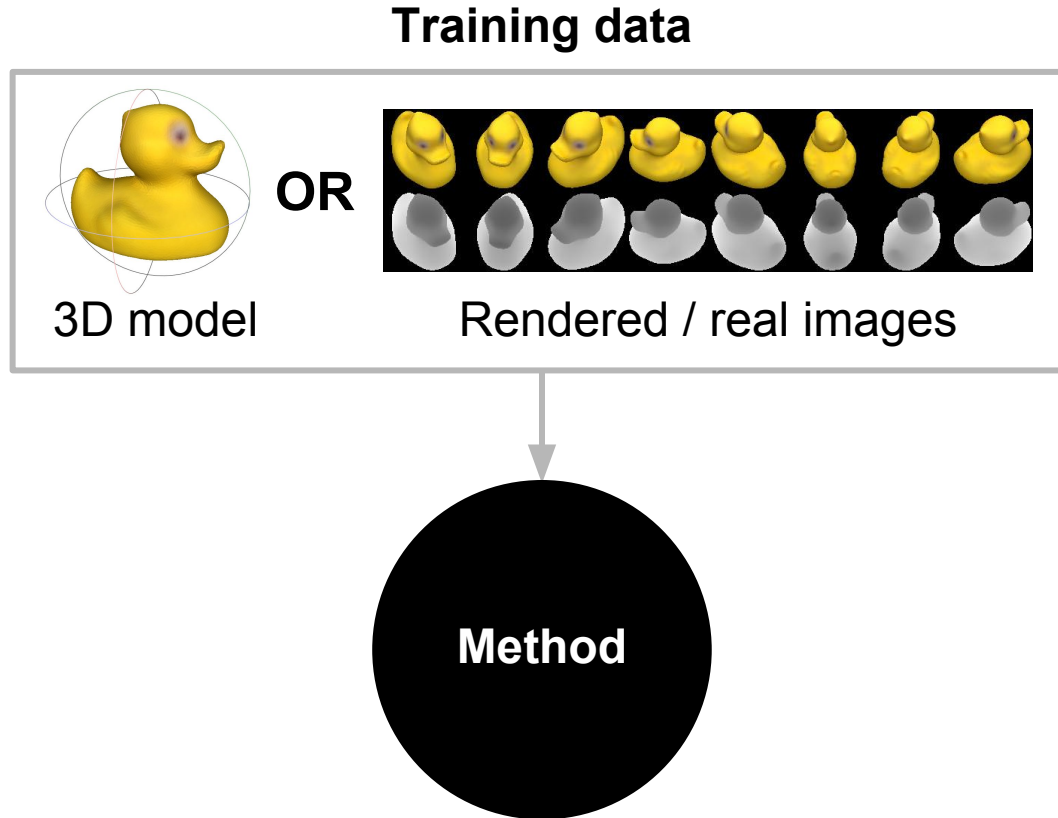
<sup>1</sup> Center for Machine Perception, CTU in Prague, CZ



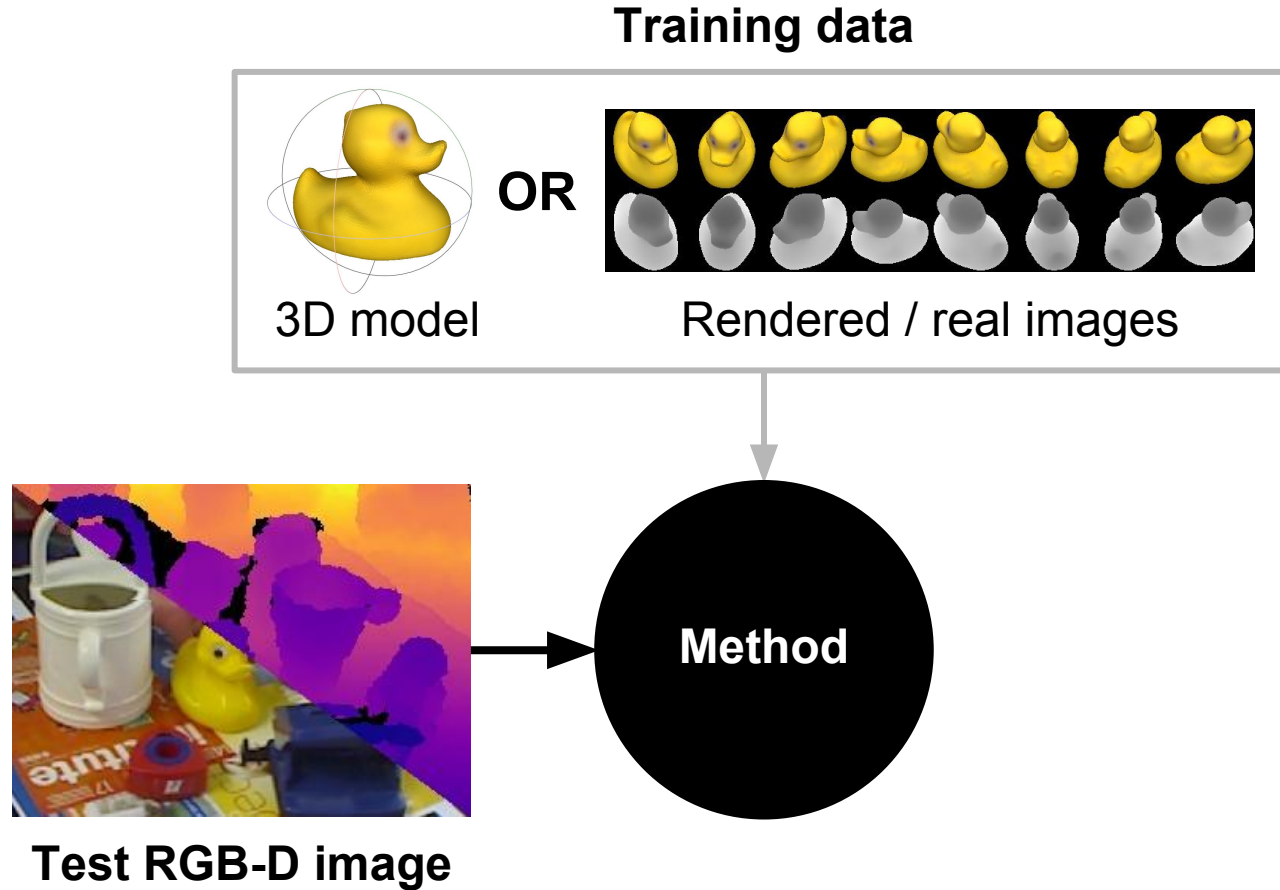
<sup>2</sup> Institute of Computer Science, FORTH, Heraklion, GR

WACV 2017, 28th March 2017, Santa Rosa

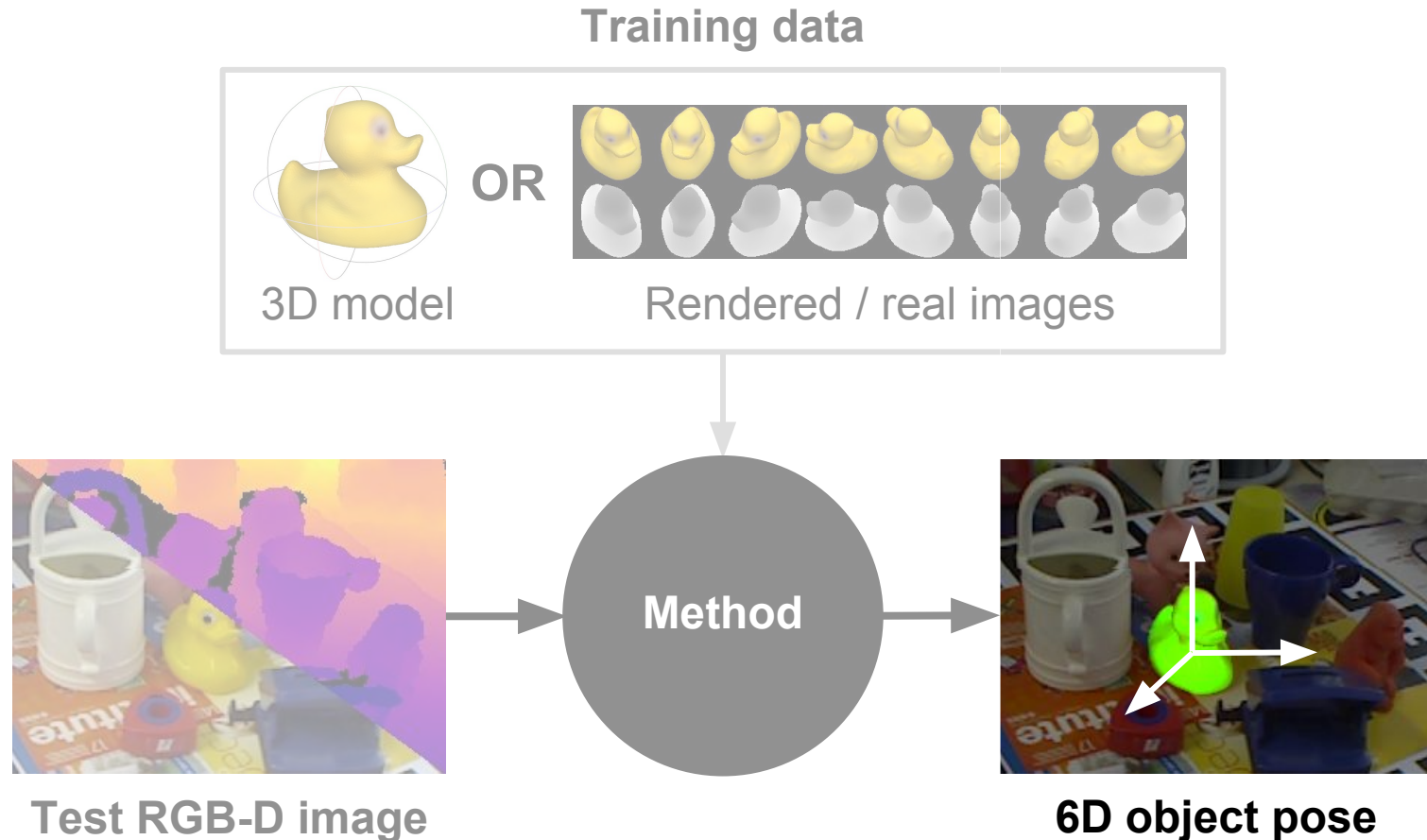
# 6D Object Pose Estimation



# 6D Object Pose Estimation

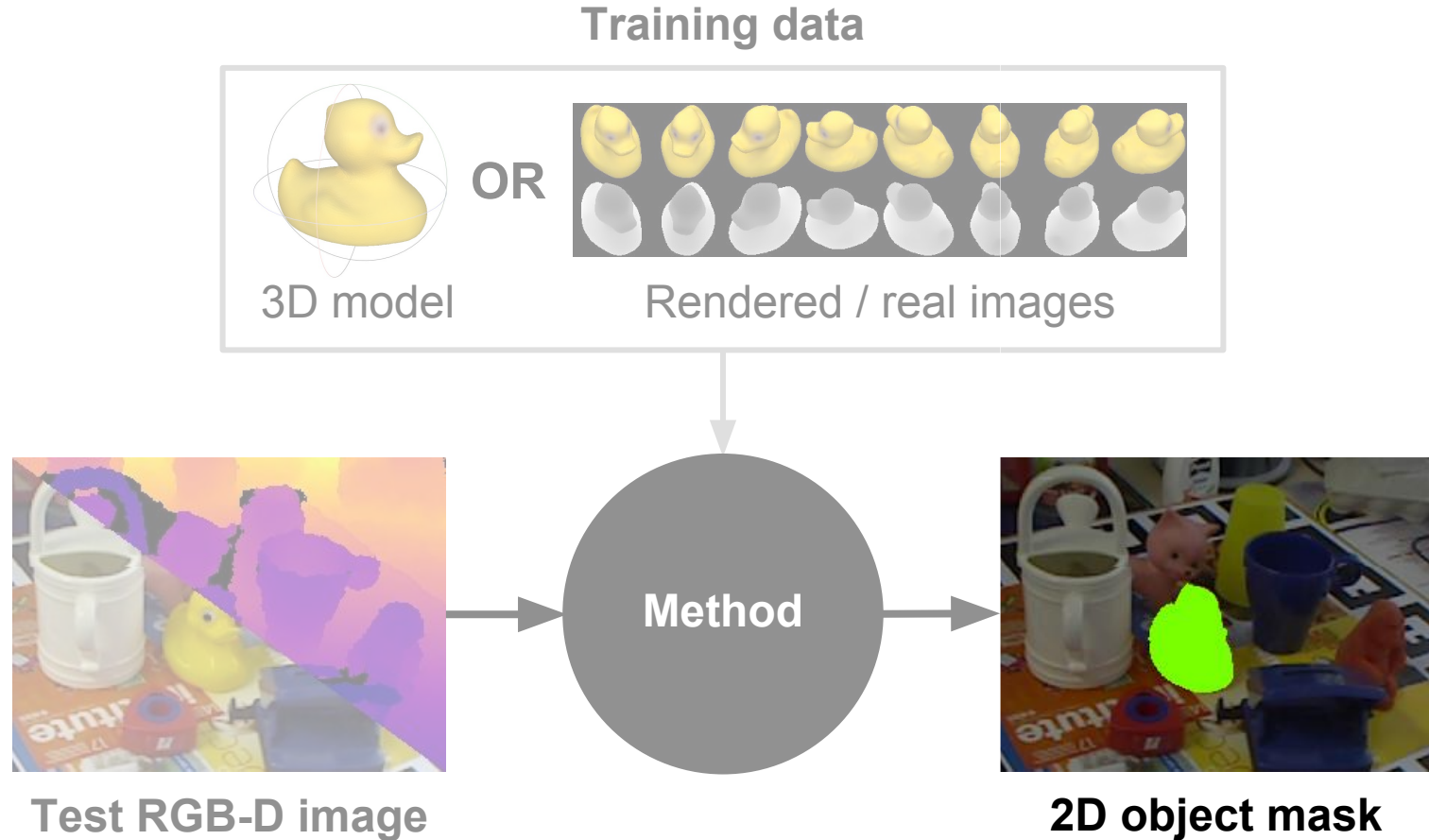


# 6D Object Pose Estimation



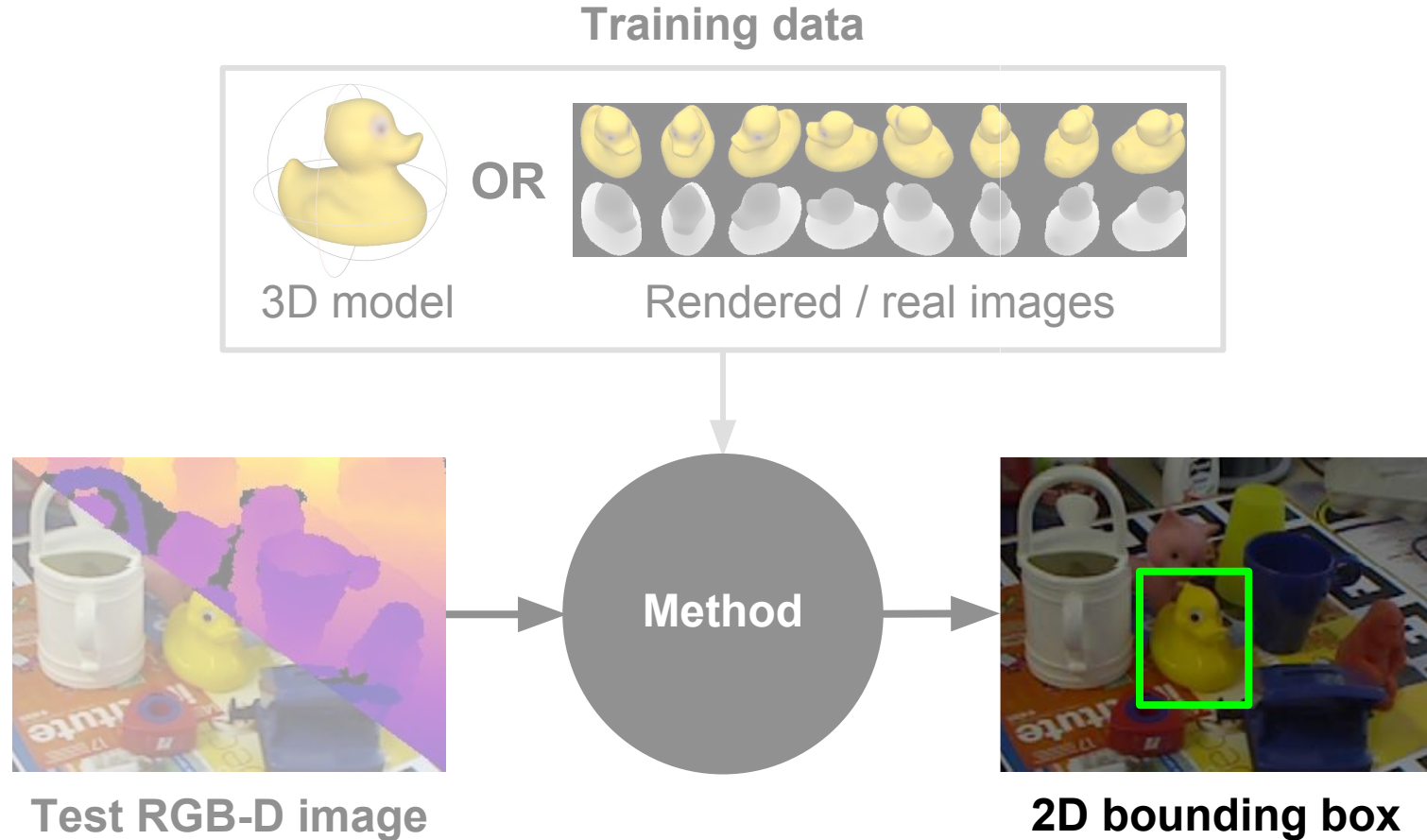
1. **Hinterstoisser et al.** "Model based training, detection and pose estimation of texture-less 3d objects in heavily cluttered scenes." ACCV'12.
2. **Hodaň et al.** "Detection and fine 3D pose estimation of texture-less objects in RGB-D images." IROS'15.
3. **Brachmann et al.** "Uncertainty-driven 6d pose estimation of objects and scenes from a single rgb image." CVPR'16.

# 2D Object Segmentation



1. **Lai et al.** "A large-scale hierarchical multi-view rgb-d object dataset." ICRA'11.
2. **Richtsfeld et al.** "Segmentation of unknown objects in indoor environments." IROS'12.
3. **Silberman et al.** "Indoor segmentation and support inference from rgb-d images." ECCV'12.
4. **Georgakis et al.** "Multiview RGB-D Dataset for Object Instance Detection." 3DV'16.

# 2D Object Detection

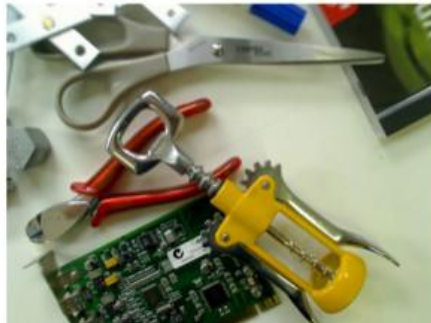


1. **Damen et al.** "Real-time Learning and Detection of 3D Texture-less Objects: A Scalable Approach." BMVC'12.
2. **Tombari et al.** "BOLD features to detect texture-less objects." ICCV'13.
3. **Rios-Cabrera and Tuytelaars.** "Discriminatively trained templates for 3d object detection: A real time scalable approach." ICCV'13.



# Texture-less Objects

Detection and accurate localization of texture-less objects is often required in **robotics** and **augmented reality**.



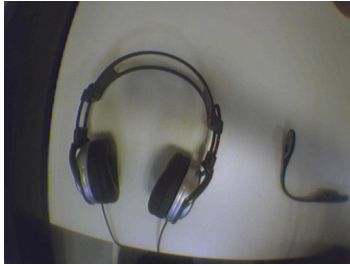
# Existing Datasets with Texture-less Objects



## RGB datasets:



Cai et al.



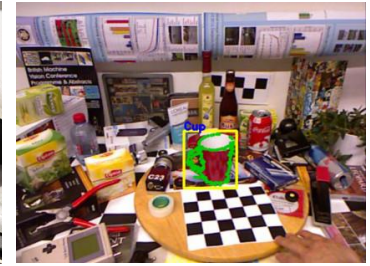
Damen et al.



Tombari et al.



Hsiao et al.

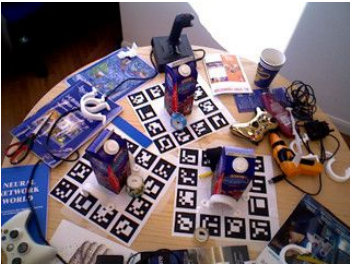


Rios-Cabrera et al.

## RGB-D datasets:



Hinterstoisser et al.



Tejani et al.



Doumanoglou et al.



Walas et al.



Michel et al.



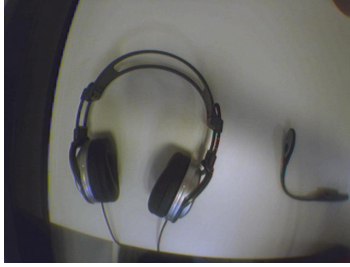
# Existing Datasets with Texture-less Objects



## RGB datasets:



Cai et al.



Damen et al.



Tombari et al.



Hsiao et al.

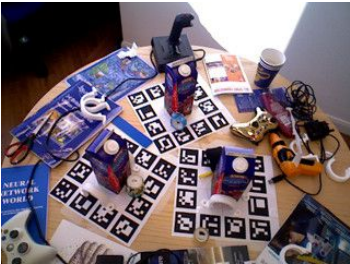


Rios-Cabrera et al.

## RGB-D datasets:



Hinterstoisser et al.



Tejani et al.



Doumanoglou et al.



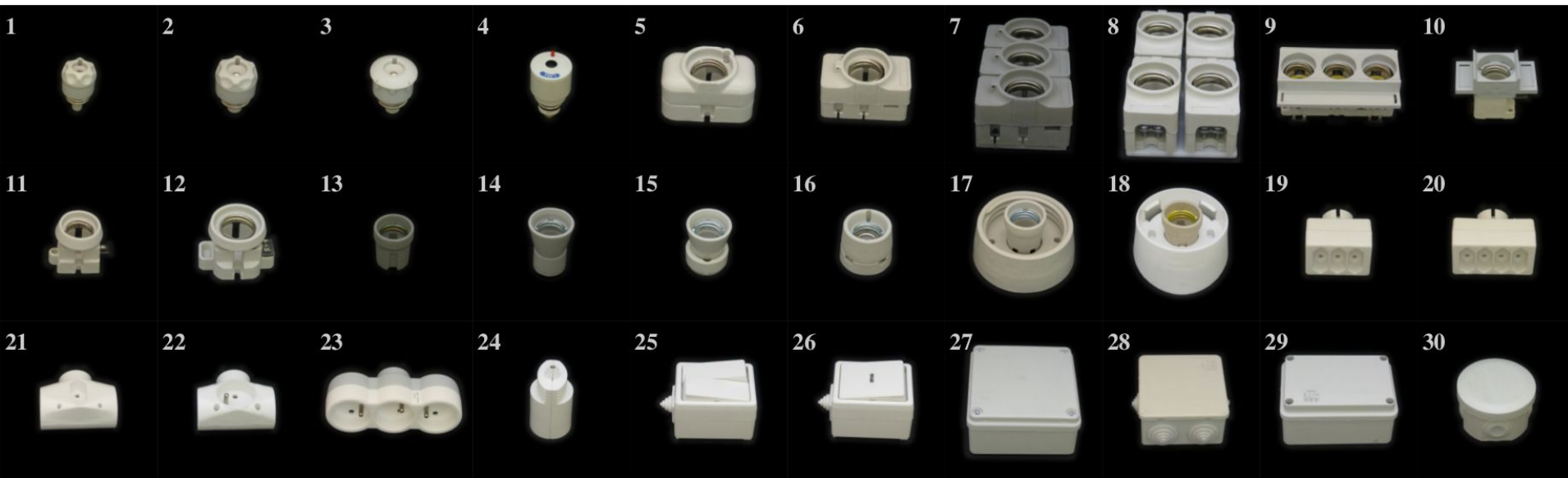
Walas et al.



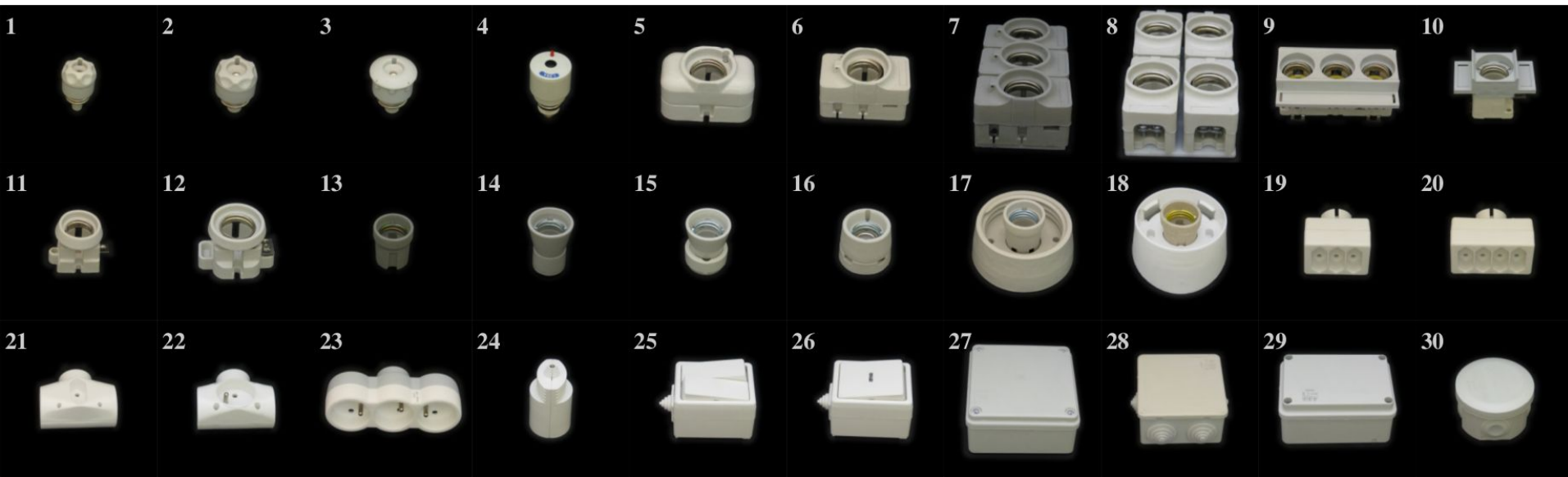
Michel et al.

Common aspect: **objects with discriminative size, shape or color**

# 30 Industry-relevant Objects in T-LESS

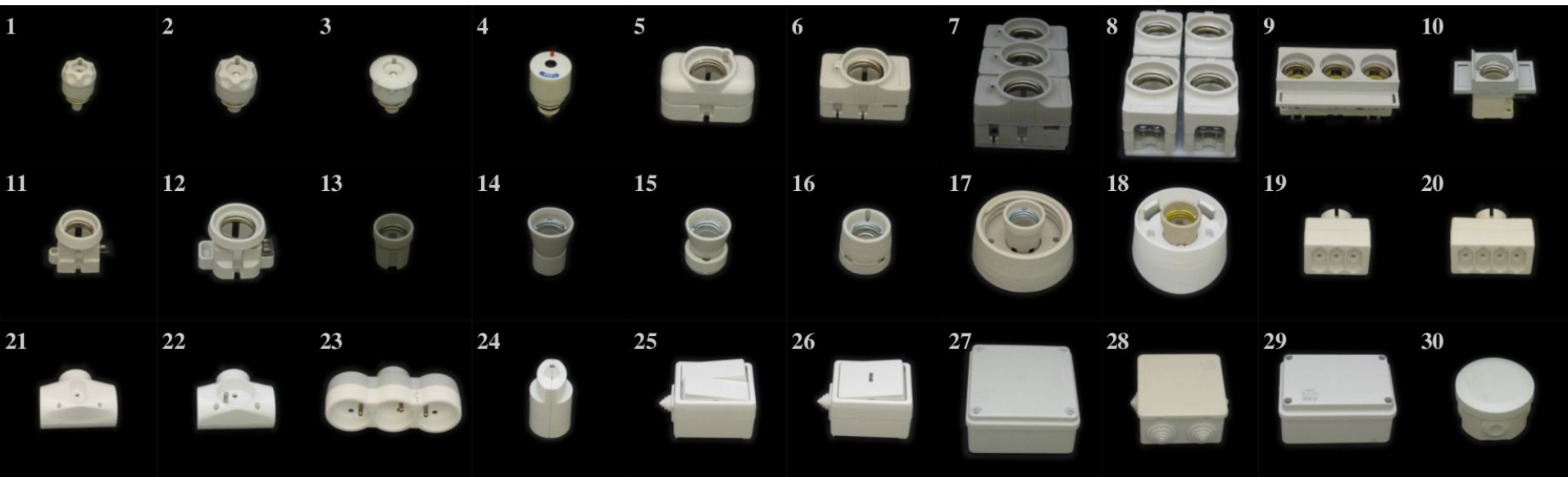


# 30 Industry-relevant Objects in T-LESS



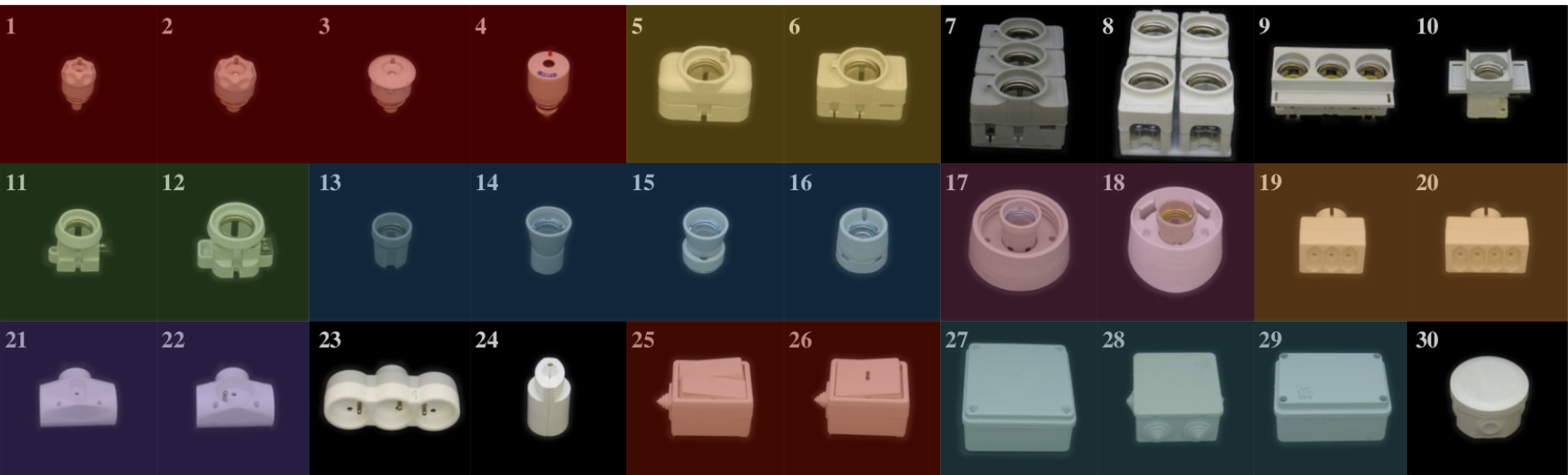
- No significant texture

# 30 Industry-relevant Objects in T-LESS



- No significant texture
- No discriminative reflectance properties

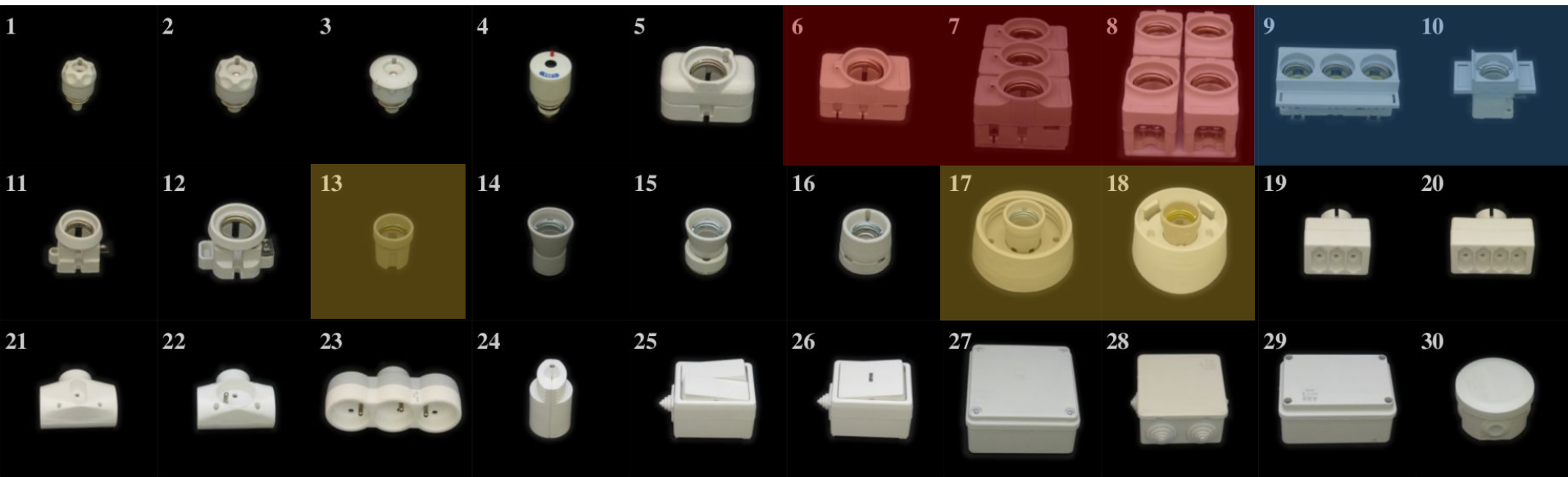
# 30 Industry-relevant Objects in T-LESS



- No significant texture
- No discriminative reflectance properties
- Symmetries and mutual similarities in shape or size



# 30 Industry-relevant Objects in T-LESS



- No significant texture
- No discriminative reflectance properties
- Symmetries and mutual similarities in shape or size
- Some objects are parts of others

# Three Synchronized Sensors



Carmine 1.09

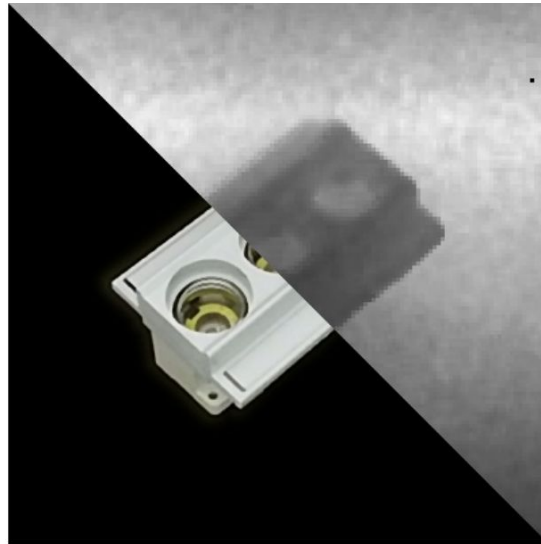
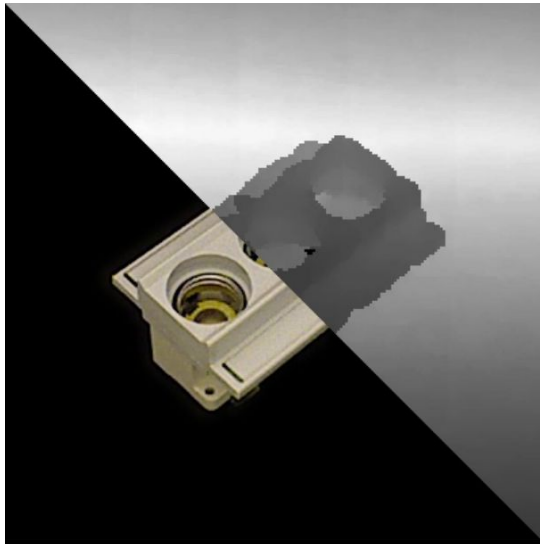


Kinect v2

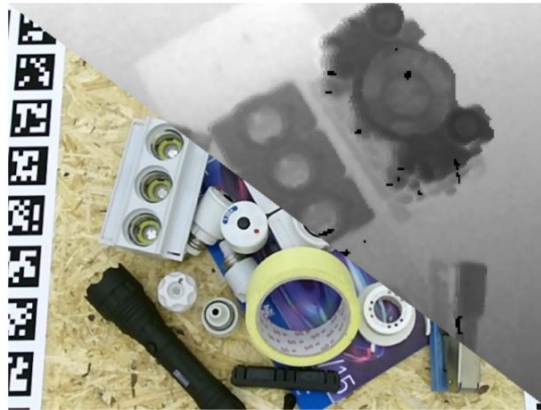
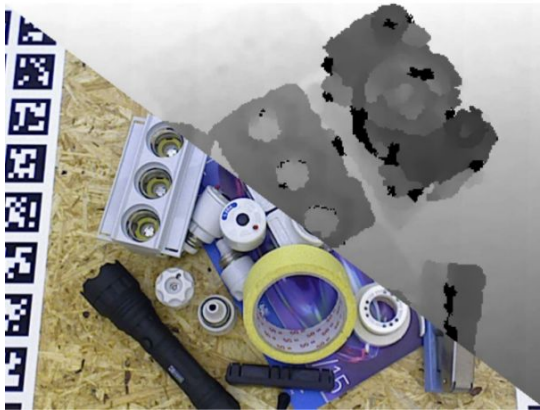


Canon IXUS 950 IS

Training images



Test images

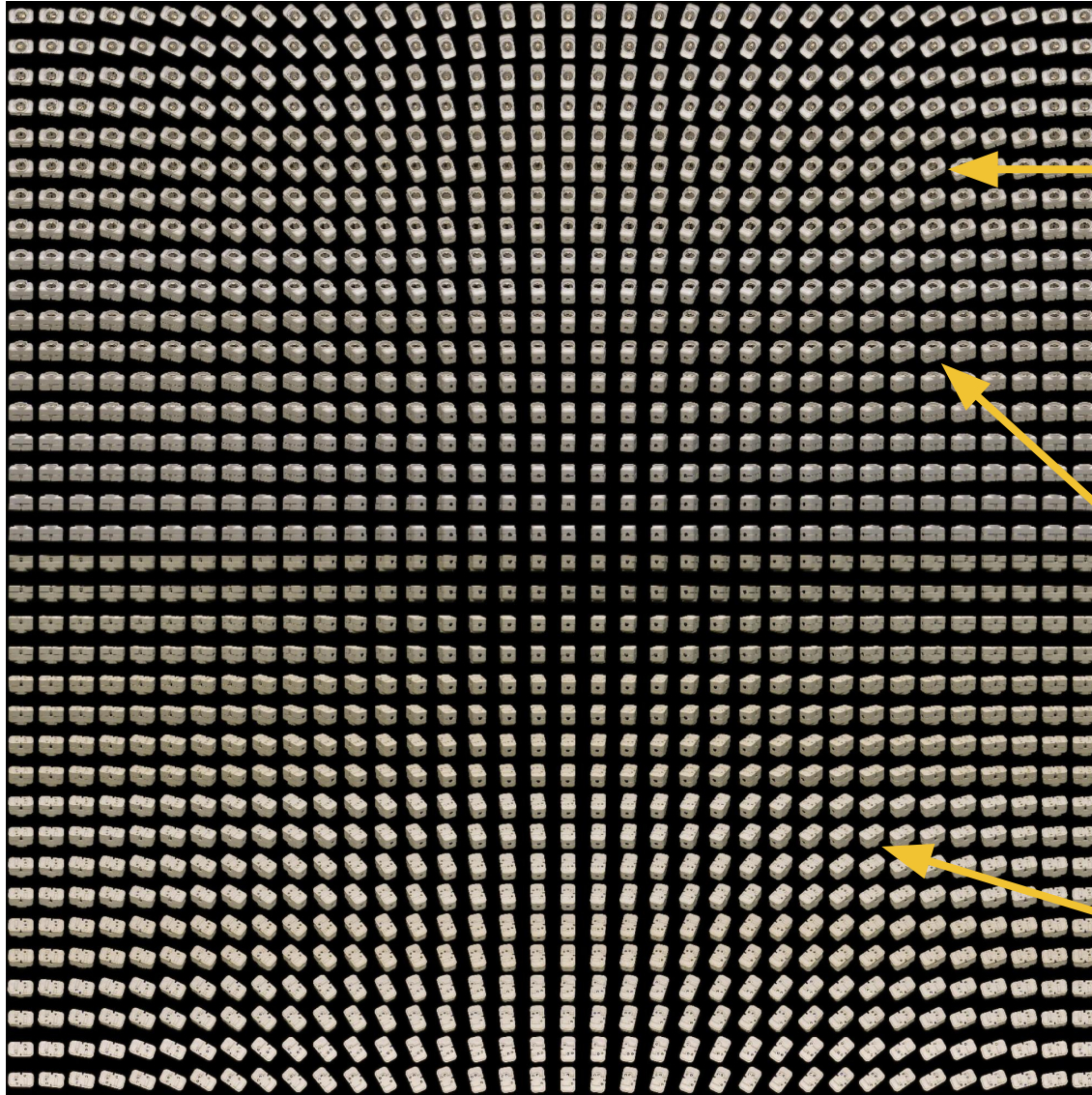




# Training Images, 39K per Sensor



- Depict individual objects against a **black background**
- Captured from a **systematically sampled view sphere**





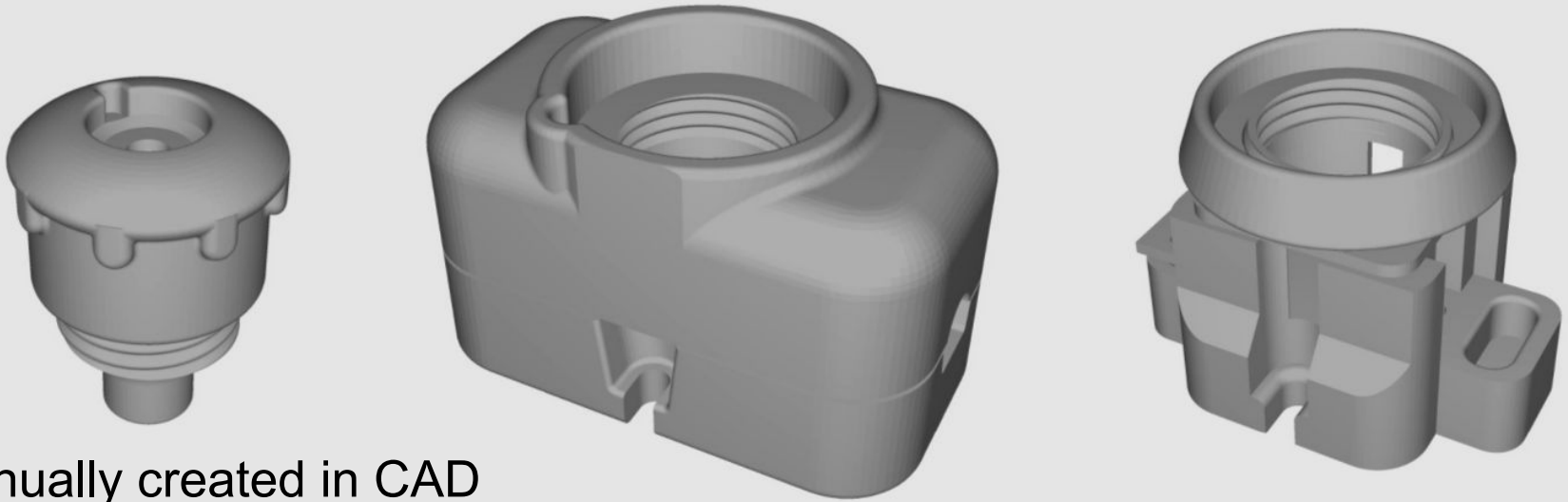
# Test Images from 20 Scenes, 10K per Sensor



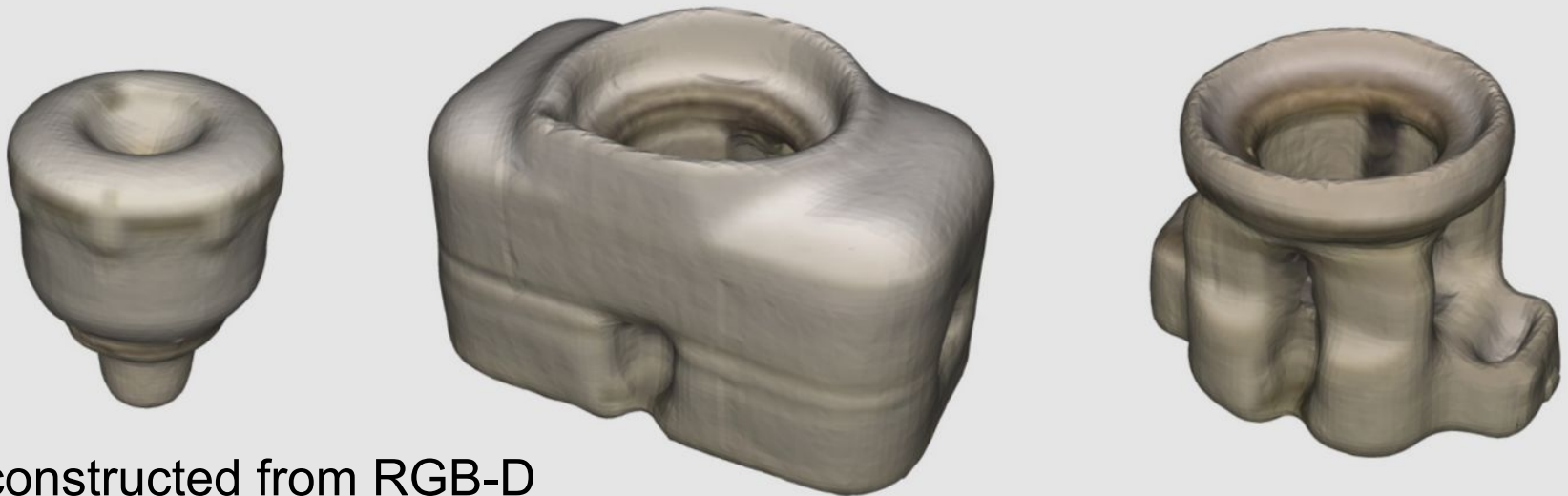
- From simple scenes with **several isolated objects** to very challenging ones with **multiple objects** and a high amount of **clutter and occlusion**
- Captured from a **systematically sampled upper view hemisphere**



# 3D Object Models



Manually created in CAD



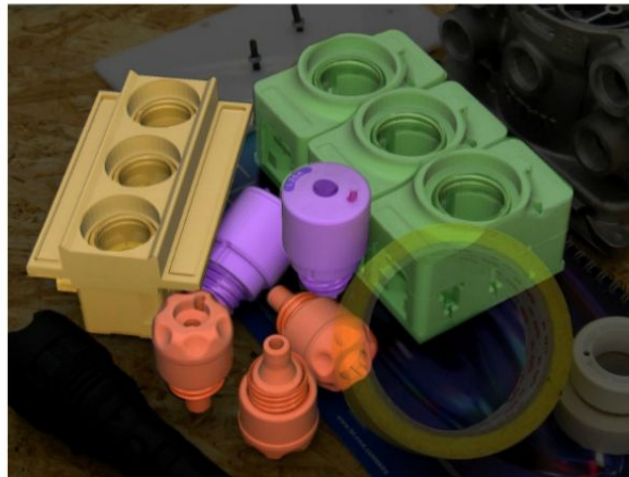
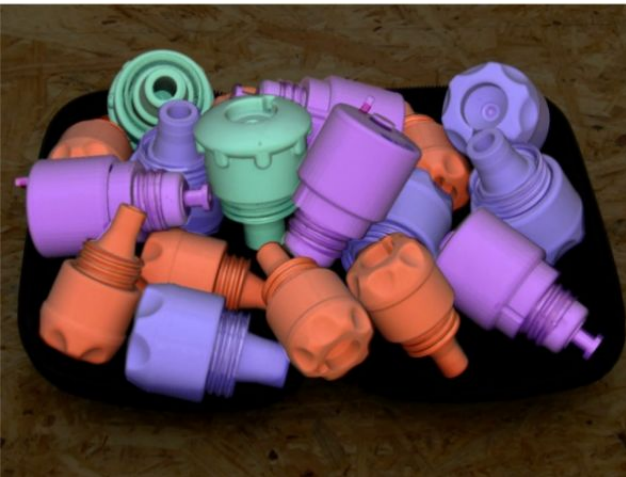
Reconstructed from RGB-D



# Accurate Ground Truth 6D Poses



Fully annotated - GT 6D pose for each instance of the modeled objects



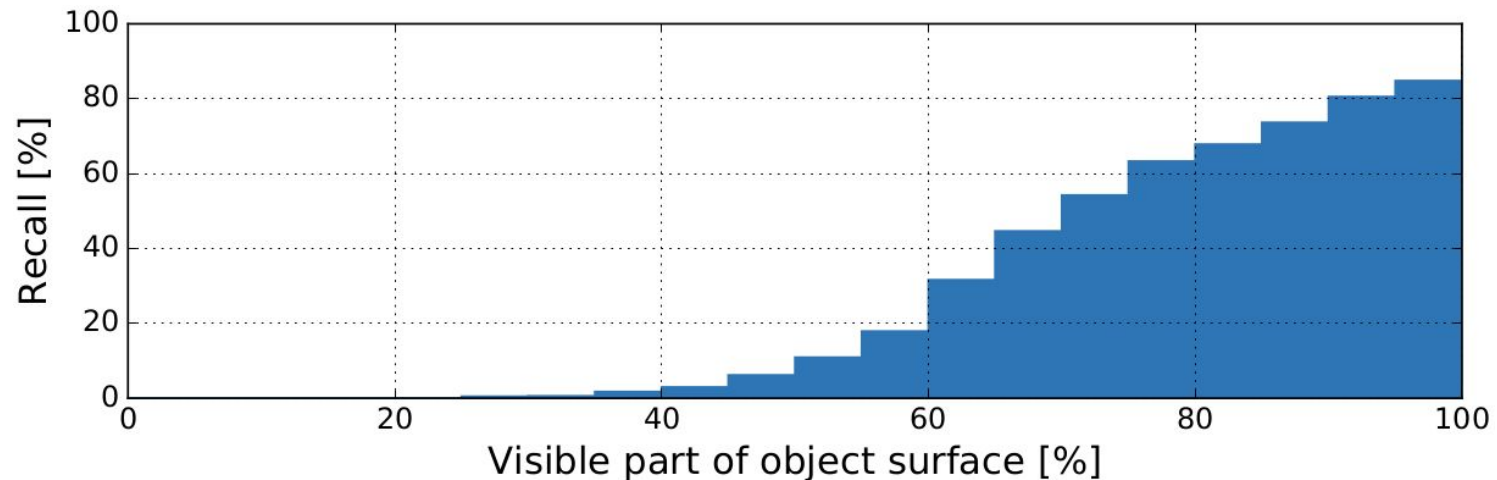
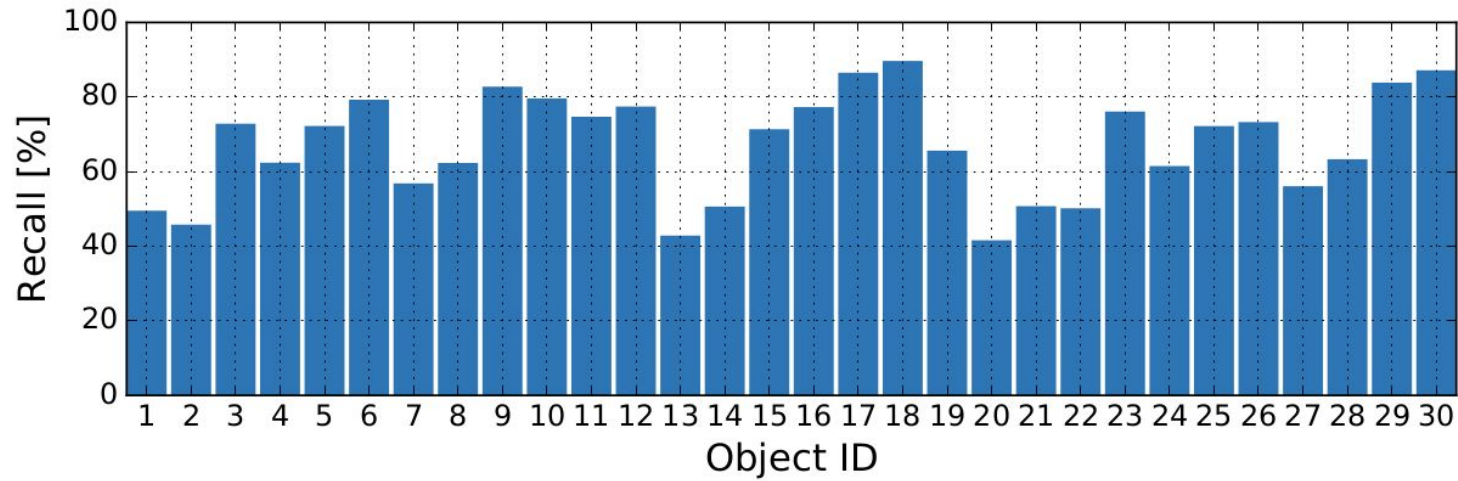
# 6D Localization - Hodañ et al. (IROS'15)



**Input:** A test image + IDs of object instances that are present in the image

**Output:** 6D pose estimates of the object instances

Success rate: **67.2%** → **There is ample room for improvement!**





# SIXD Challenge at ICCV 2017 Workshop



- At the **3rd Workshop on Recovering 6D Object Pose**
- **To establish SOTA** in 6D object pose estimation
- Will be **announced in April**
- 6 selected datasets converted to a **unified format**:



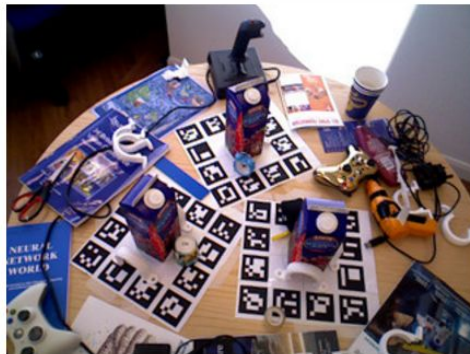
Hinterstoisser et al. [3] with extra ground truth from Brachmann et al. [4]  
[DOWNLOAD](#)



T-LESS [2]  
[DOWNLOAD](#)



TUD Light  
[DOWNLOAD](#)



Tejani et al. [5] - reduced version  
[DOWNLOAD](#)



Doumanoglou et al. [6] - scenario 2 coming soon



Rutgers APC [7] - reduced version  
[DOWNLOAD](#)

### An RGB-D dataset and evaluation methodology for detection and 6D pose estimation of texture-less objects

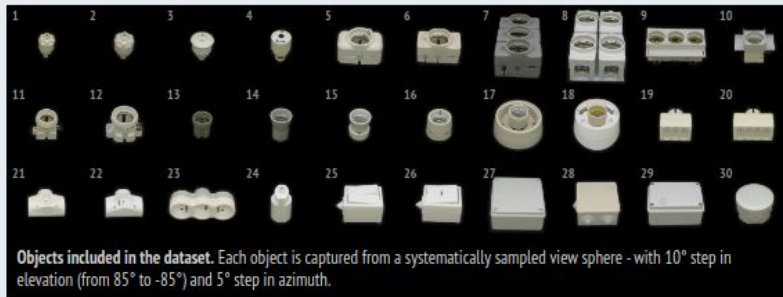
- **30 industry-relevant objects:** no discriminative color, no texture, often similar in shape, some objects are parts of others.
- **Three synchronized sensors** used to capture the training and test images: Primesense CARMINE 1.09 (a structured-light RGB-D sensor), Microsoft Kinect v2 (a time-of-flight RGB-D sensor), and Canon IXUS 950 IS (a high-resolution RGB camera).
- **Training images (39K from each sensor)** depict individual objects against a black background.
- **Test images (10K from each sensor)** originate from 20 test scenes. The scene complexity varies from simple scenes with several isolated objects to very challenging ones with multiple object instances and a high amount of clutter and occlusion.
- **Two types of 3D models for each object:** a manually created CAD model and a semi-automatically reconstructed one.
- **A new evaluation methodology** which deals with pose ambiguity that can be caused by object symmetries and occlusions.

Please cite the following paper if you use the dataset. This work is licensed under [Creative Commons Attribution-ShareAlike 4.0](#).

T. Hodaň, P. Haluza, Š. Obdržálek, J. Matas, M. Lourakis, X. Zabulis,  
**T-LESS: An RGB-D Dataset for 6D Pose Estimation of Texture-less Objects,**  
 IEEE Winter Conference on Applications of Computer Vision (WACV), 2017, Santa Rosa, USA

[\[PDF\]](#), [\[BIB\]](#)

- **19/01/2017** - WACV 2017 paper about T-LESS is available on [arXiv](#).
- **23/09/2016** - The first complete version (v2) of the dataset is released.
- **16/03/2015** - A preview version (v1) of the dataset is available.



**Sample test images.** The images are overlaid with colored 3D object models at the ground truth poses. Each test scene is captured from a systematically sampled view hemisphere - with 10° step in elevation (from 75° to 15°) and 5° step in azimuth.

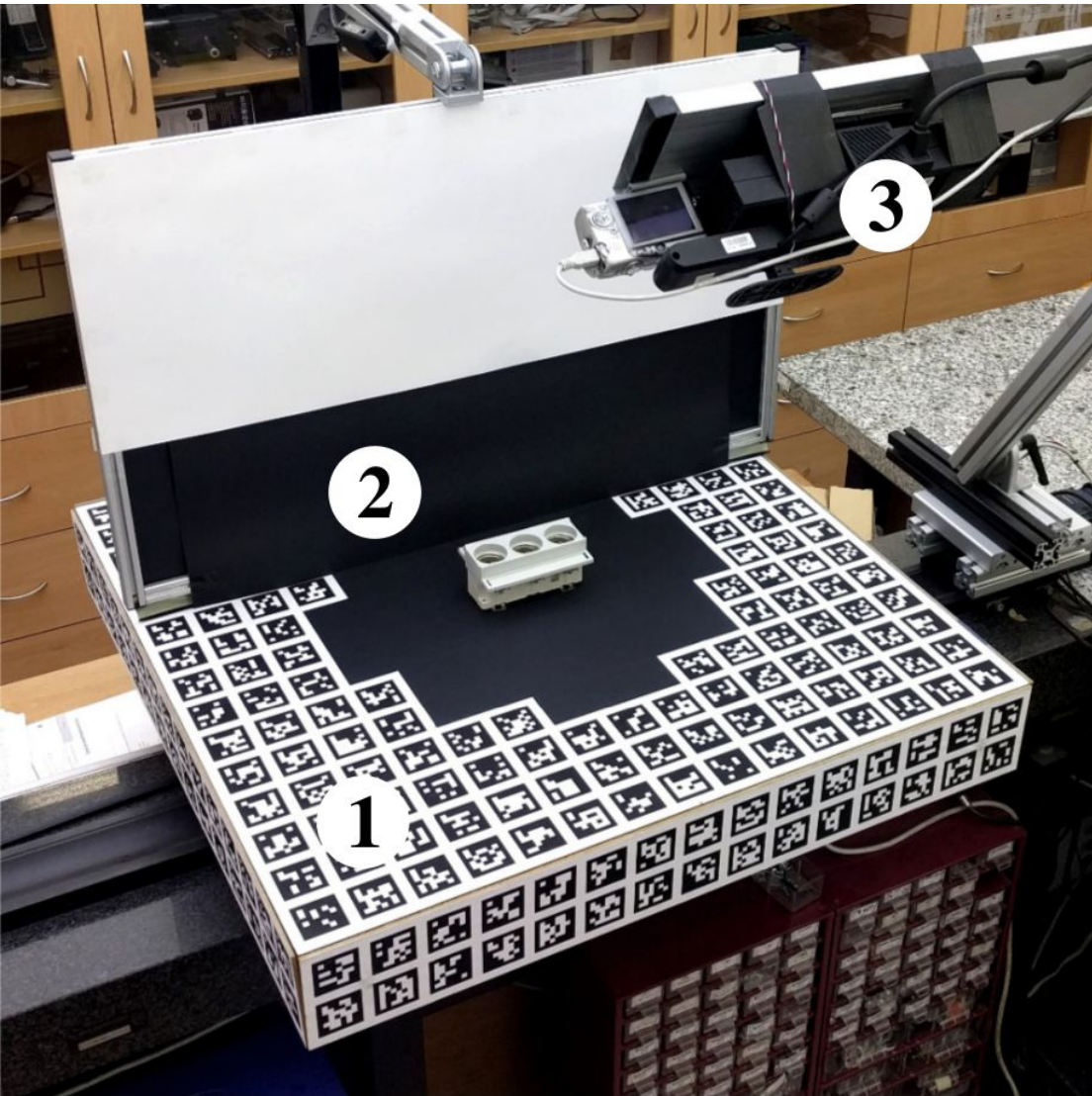
T-LESS available online:  
[cmp.felk.cvut.cz/t-less](http://cmp.felk.cvut.cz/t-less)

Poster 185





# Acquisition Setup



1) **Turntable** with marker field.

2) **Screen** ensuring a black background for training images, removed when capturing test images.

3) **Triplet of sensors** attached to a jig with adjustable tilt.

# Estimation of Ground Truth 6D Poses



1. Manually align the object models to the scene model to get **initial object poses**.
2. **Render the object models** at the current poses into several scene images.
3. **Identify misalignments** and **manually refine** the object poses accordingly.
4. **Go to step 2** until a **satisfactory alignment** of the renderings with the images.

