From Domestic Service Robots to Industrial Mobile Manipulators

Sven Behnke

Computer Science Institute VI
Autonomous Intelligent Systems

universität bonn
ais

Robot Competitions

- Provide common test bed for benchmarking
- Foster robotics research

DARPA Urban Challenge

RoboCup Humanoid Soccer

DLR SpaceBot Cup

DARPA Robotics Challenge

RoboCup@Home

- Since 2006
- Focus on applications in domestic environments and on human-robot interaction
- **Goal:** Develop robots that support humans in everyday tasks
- **Competition:**
  - Predefined tests
    - Follow a person
    - Find and put away objects
    - Fetch drinks
    - Understand complex speech commands
  - Open demonstrations
- Bar is raised every year

Requirements for Domestic Service Robots

- **Robust navigation** in indoor environments
  - Map of the home
  - Navigational sensors
  - Not too large mobile base (narrow passages)

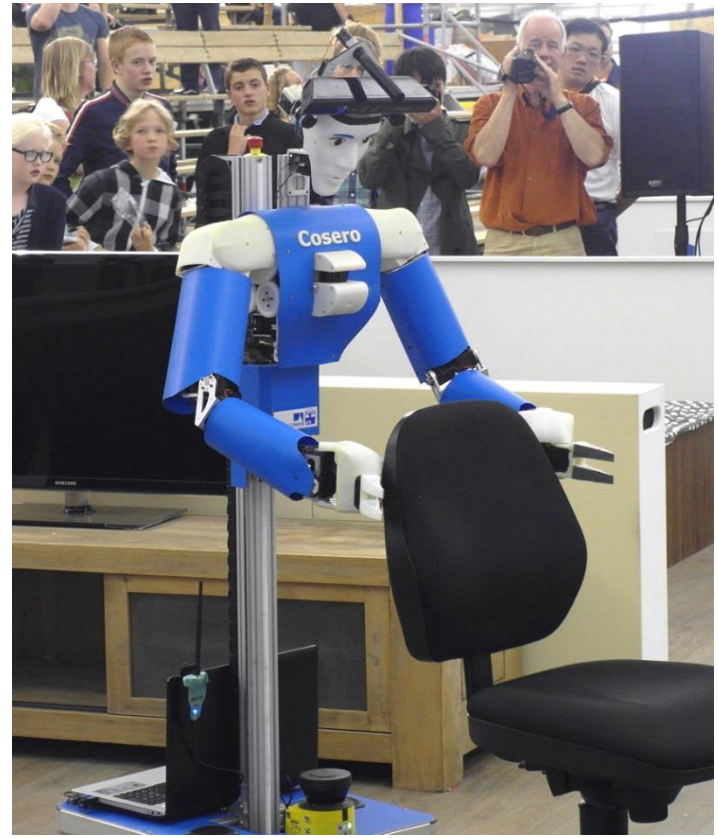
- **Object manipulation**
  - Dexterous arm(s)
  - Grippers that can handle the payload of common household objects
  - Sensors to detect and recognize objects

- **Intuitive communication** with the users
  - Appropriate communication height for face-to-face interaction
  - Multiple modalities, such as speech, gestures, mimics, and body language

When we started, most available domestic robot systems supported only one or two of the above skills.

Our Domestic Service Robots

Dynamaid



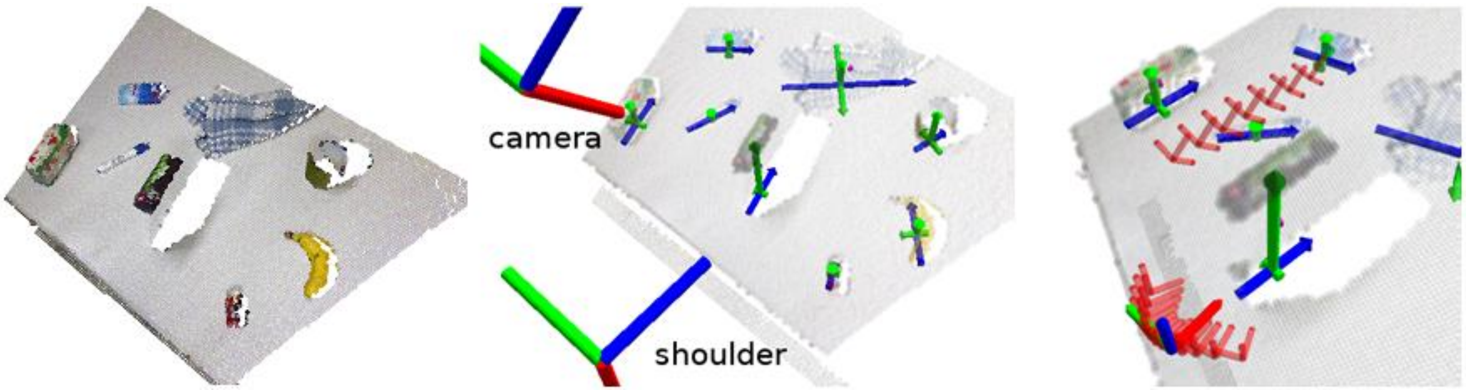
Cosero

- Size: 100-180 cm, weight: 30-35 kg
- 36 articulated joints
- PC, laser scanners, Kinect, microphone, ...

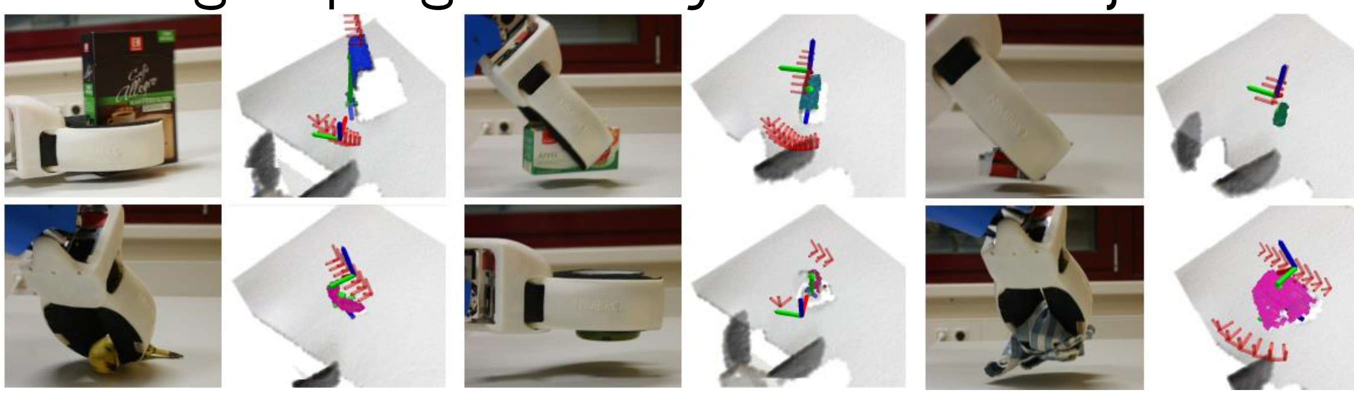
RoboCup 2013 Eindhoven

Analysis of Table-top Scenes and Grasp Planning

- Detection of clusters above horizontal plane
- Two grasps (top, side)

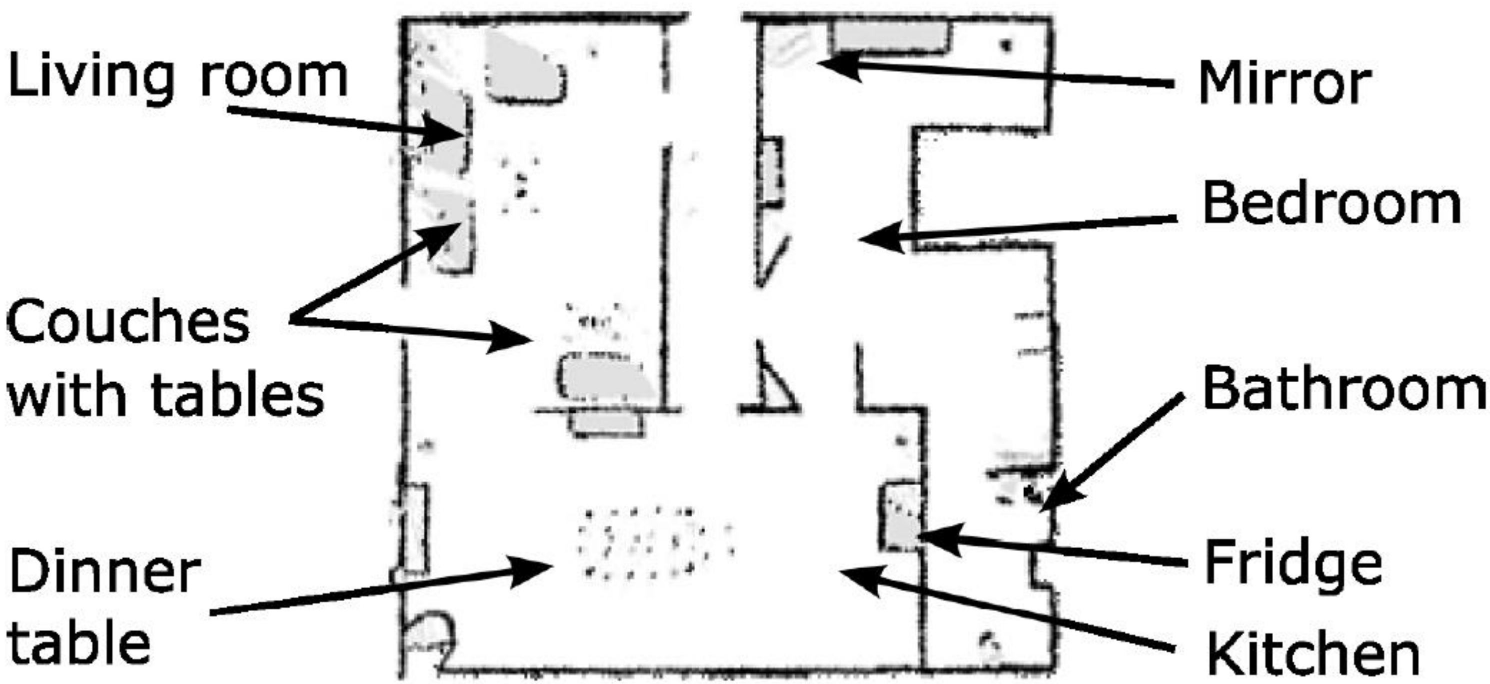


- Flexible grasping of many unknown objects



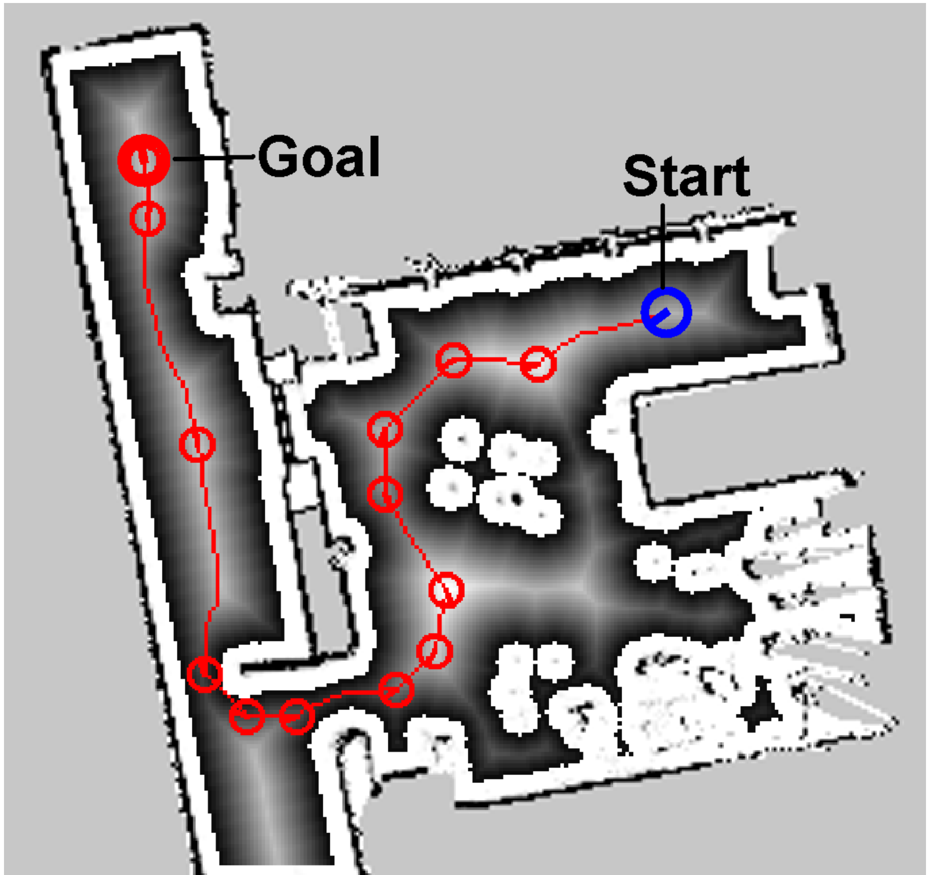
[Stückler, Steffens, Holz, Behnke, Robotics and Autonomous Systems 2012]

Mapping the Environment



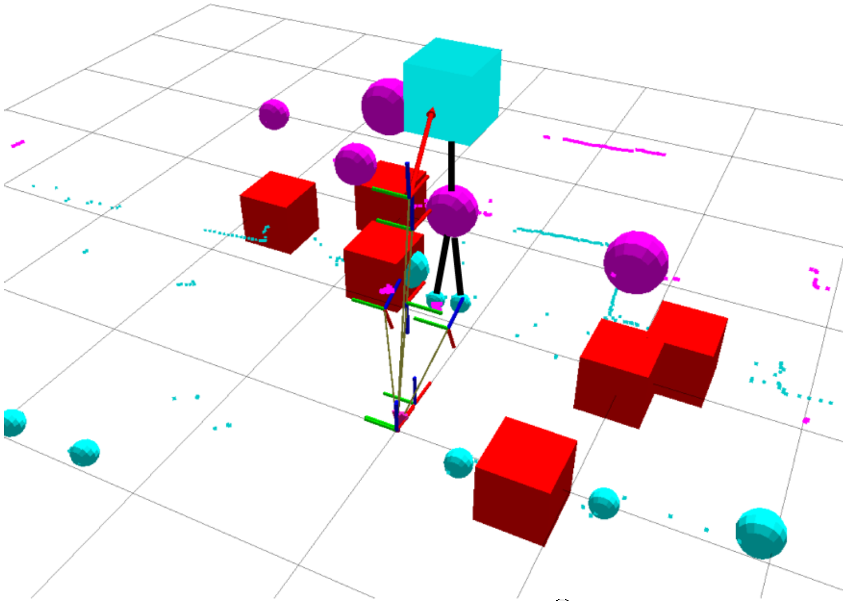
Path Planning

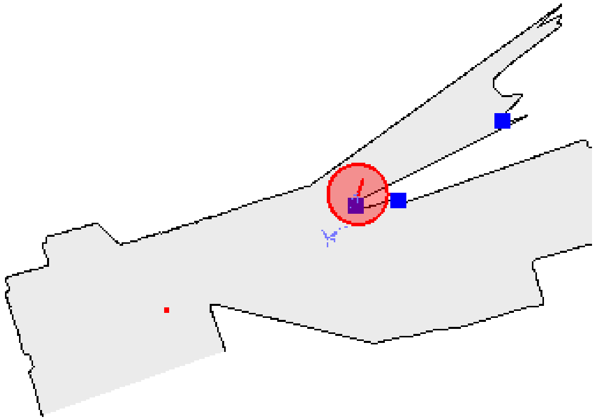
- Global planning tries to keep away from obstacles
- Obstacle avoidance using two lasers, ultrasound
- Robot alignment in narrow passages
- Plan revision when path blocked



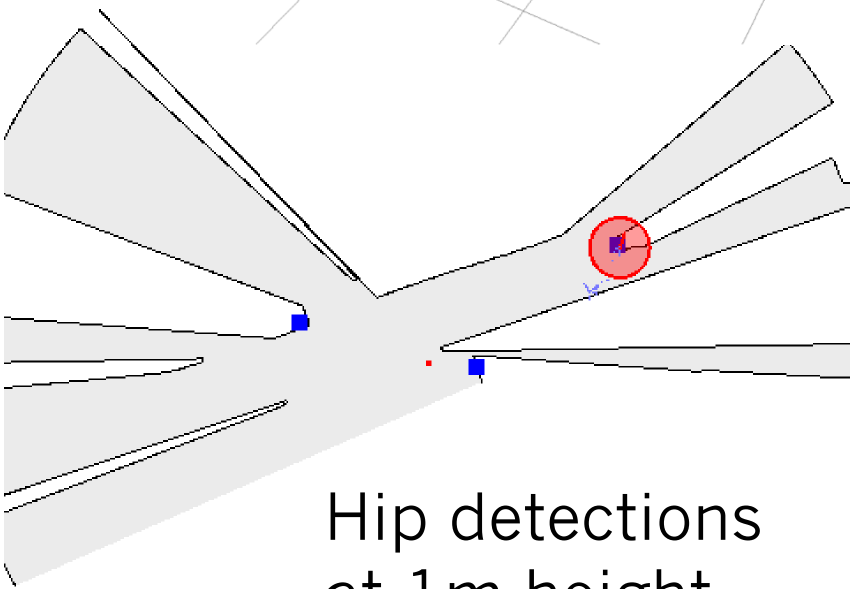
Continuous Person Awareness

- Detect persons using LRFs at two heights
- Visual person verification and identification
- Natural gaze strategies
- Face recognition





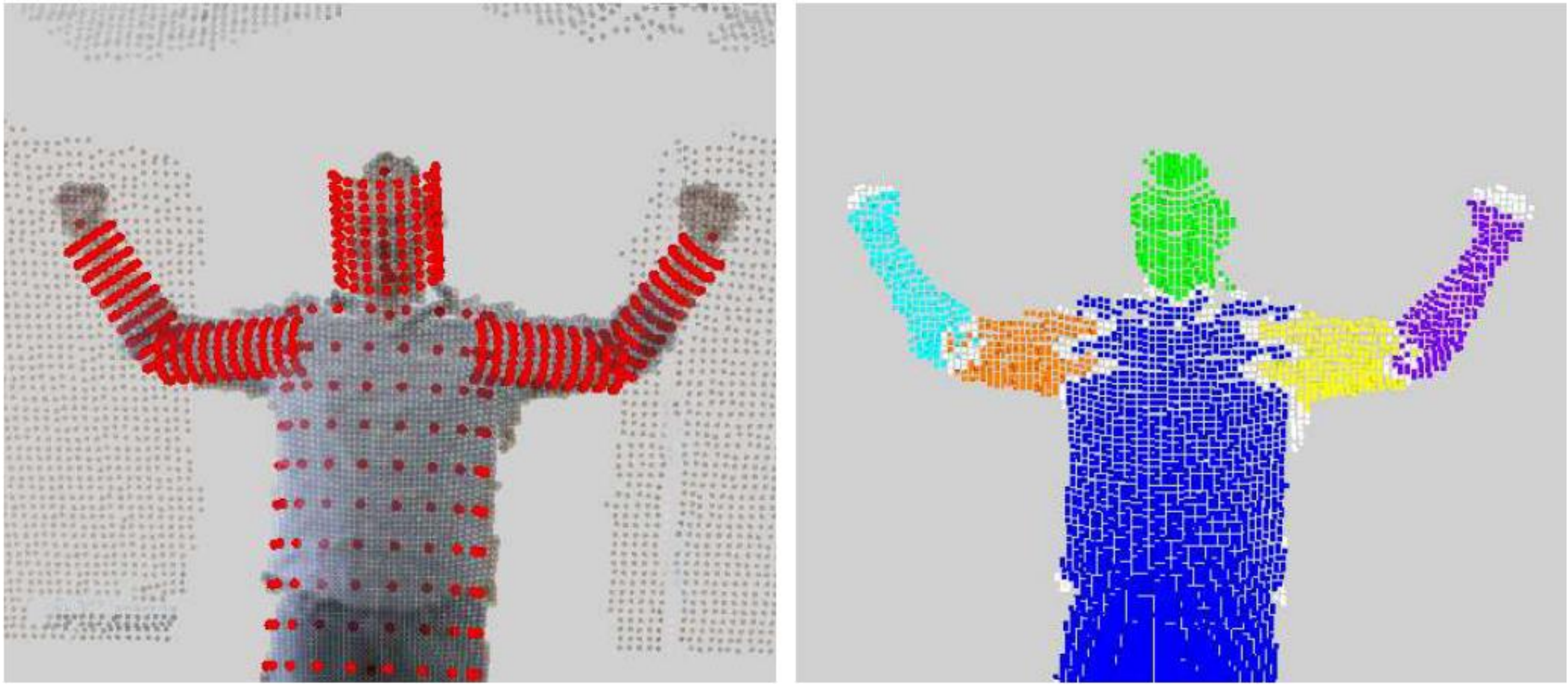
Leg detections at 30cm height



Hip detections at 1m height

Adaptive Person Model

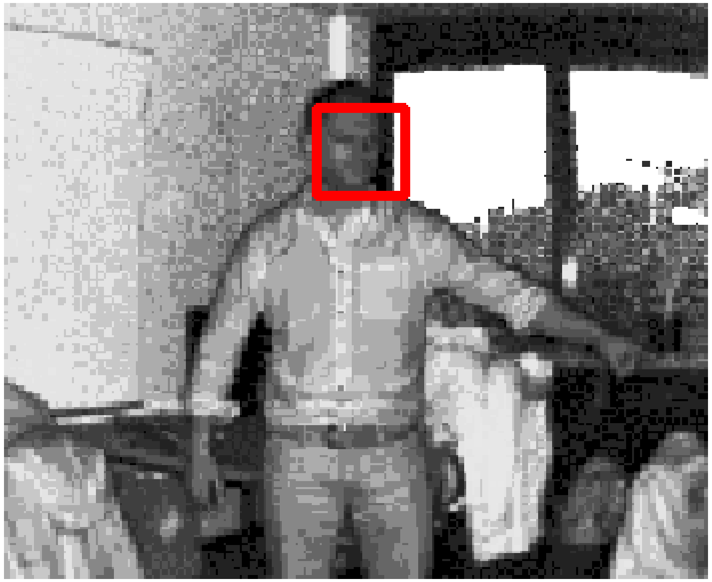
- Model: geometric primitives, connected by joints
- Registration through articulated ICP
- Adaptation of primitive parameters to body proportions

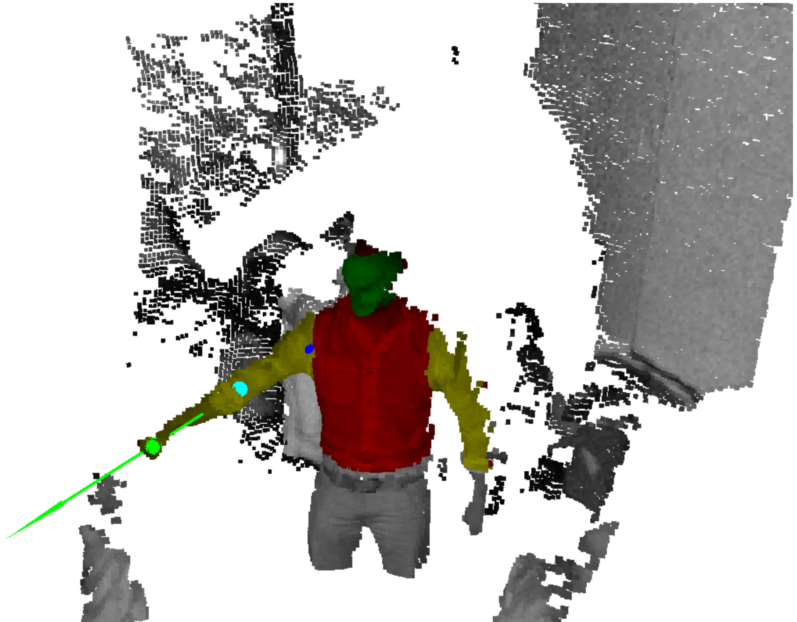


[Droeschel, Behnke: ICIRA 2011]

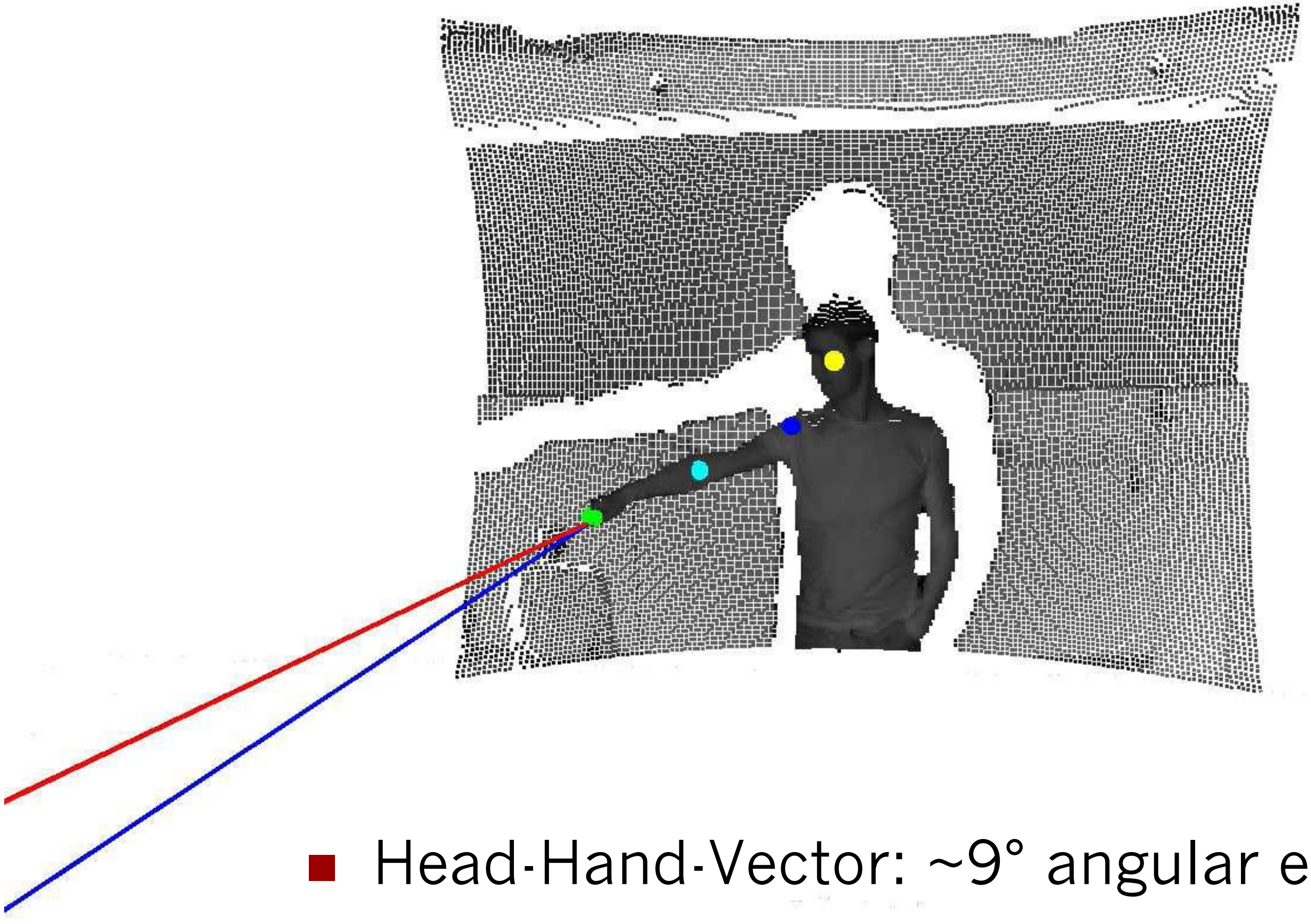
Gesture Recognition using a ToF Camera

- Find and track face in amplitude image
- Find body by region growing
- Segment torso and arms
- Identify elbow and hand





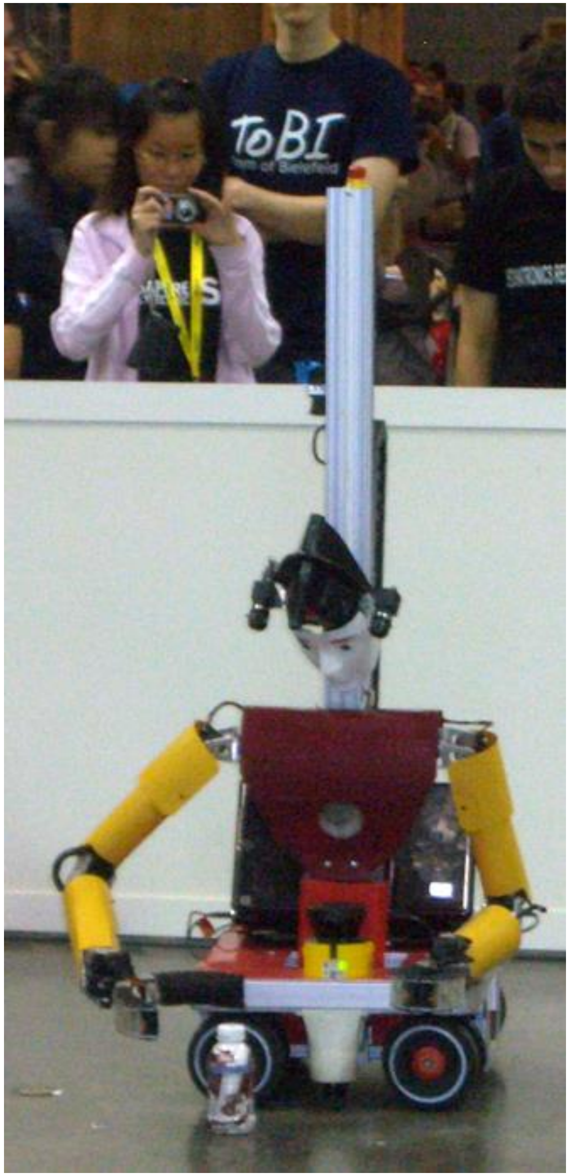
Estimating Pointing Direction

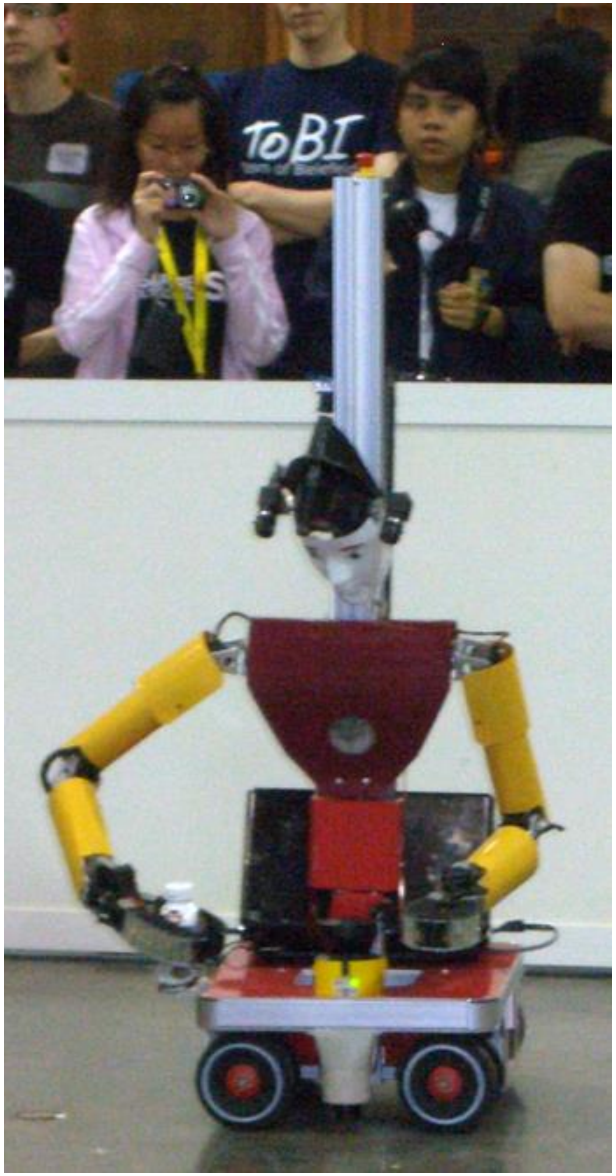


- Head-Hand-Vector: ~9° angular error
- GPR function approximator: ~3° error

[Dröschel, Stückler, Behnke: HRI’11]

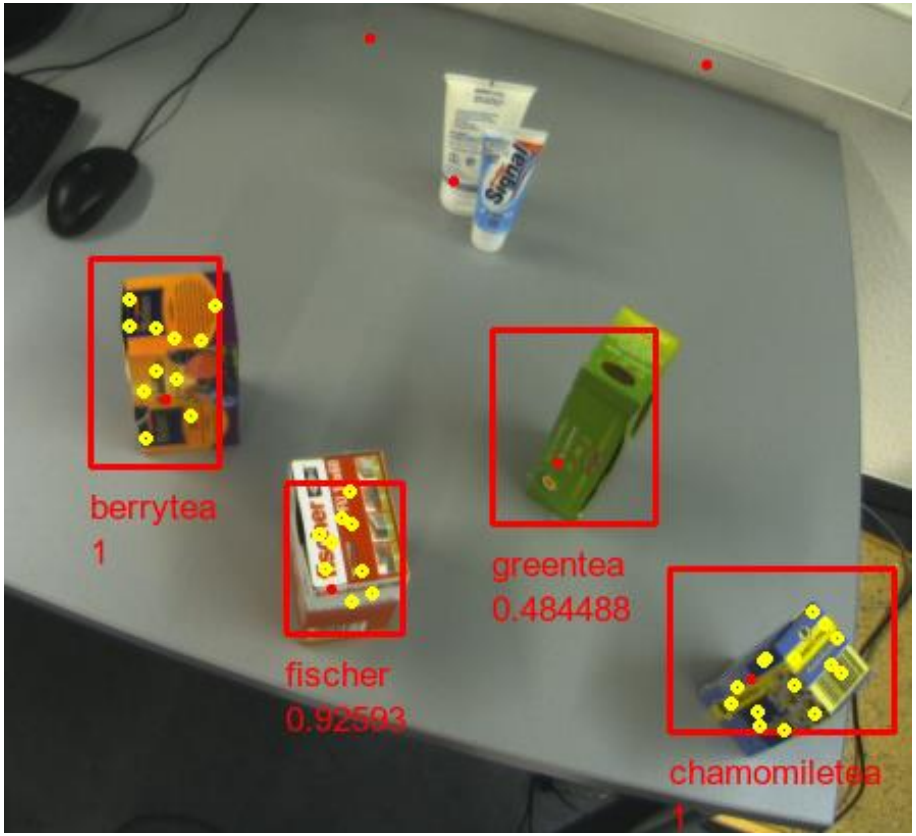
Picking-up Objects from the Floor

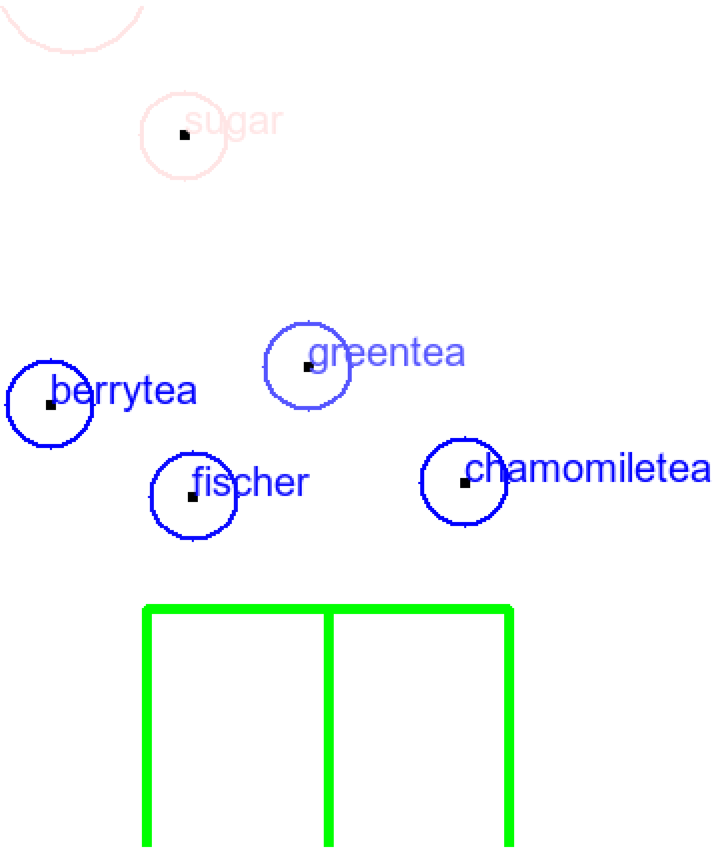




Showing Gesture

Visual Object Recognition

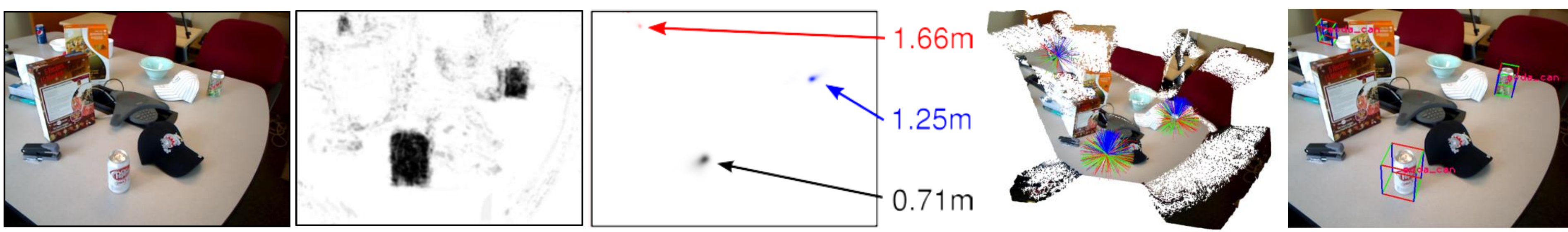




- Detect objects with laser
- Map them to image plane
- Recognize them using color and SURF features
- Kalman filter for objects

Object Class Detection in RGB-D

- Hough forests make not only object class decision, but describe object center
- RGB-D objects data set
- Color and depth features
- Training with rendered scenes
- Detection of object position and orientation



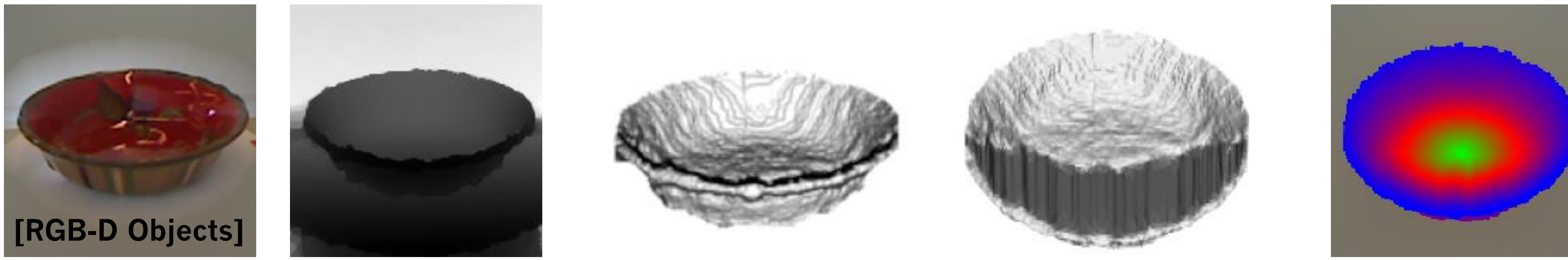
Scene Class prob. Object centers Orientation Detected objects

- Depth helps a lot

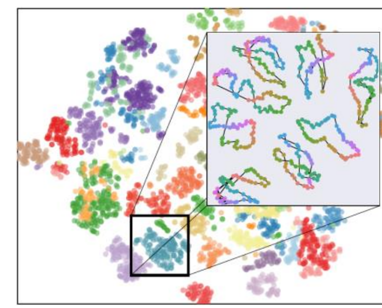
[Badami, Stückler, Behnke: SPME 2013]

Object Recognition and Pose Estimation

- Rendering canonical views

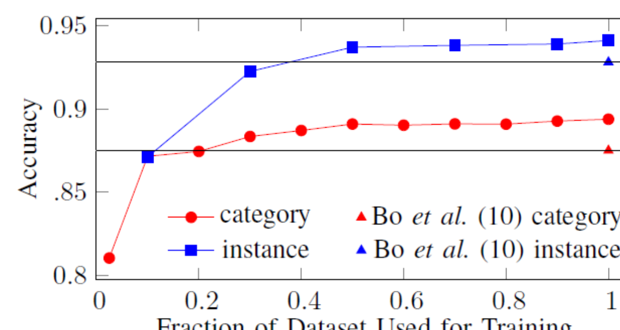


- Pretrained convolutional neural network



- Linear SVM / SV regression

Method	Category Accuracy (%) RGB	Category Accuracy (%) RGB-D	Instance Accuracy (%) RGB	Instance Accuracy (%) RGB-D
Lai *et al.* (8)	74.3 ± 3.3	81.9 ± 2.8	59.3	73.9
Bo *et al.* (10)	82.4 ± 3.1	87.5 ± 2.9	**92.1**	92.8
Ours	$\mathbf{83.1 \pm 2.0}$	$\mathbf{89.4 \pm 1.3}$	92.0	**94.1**

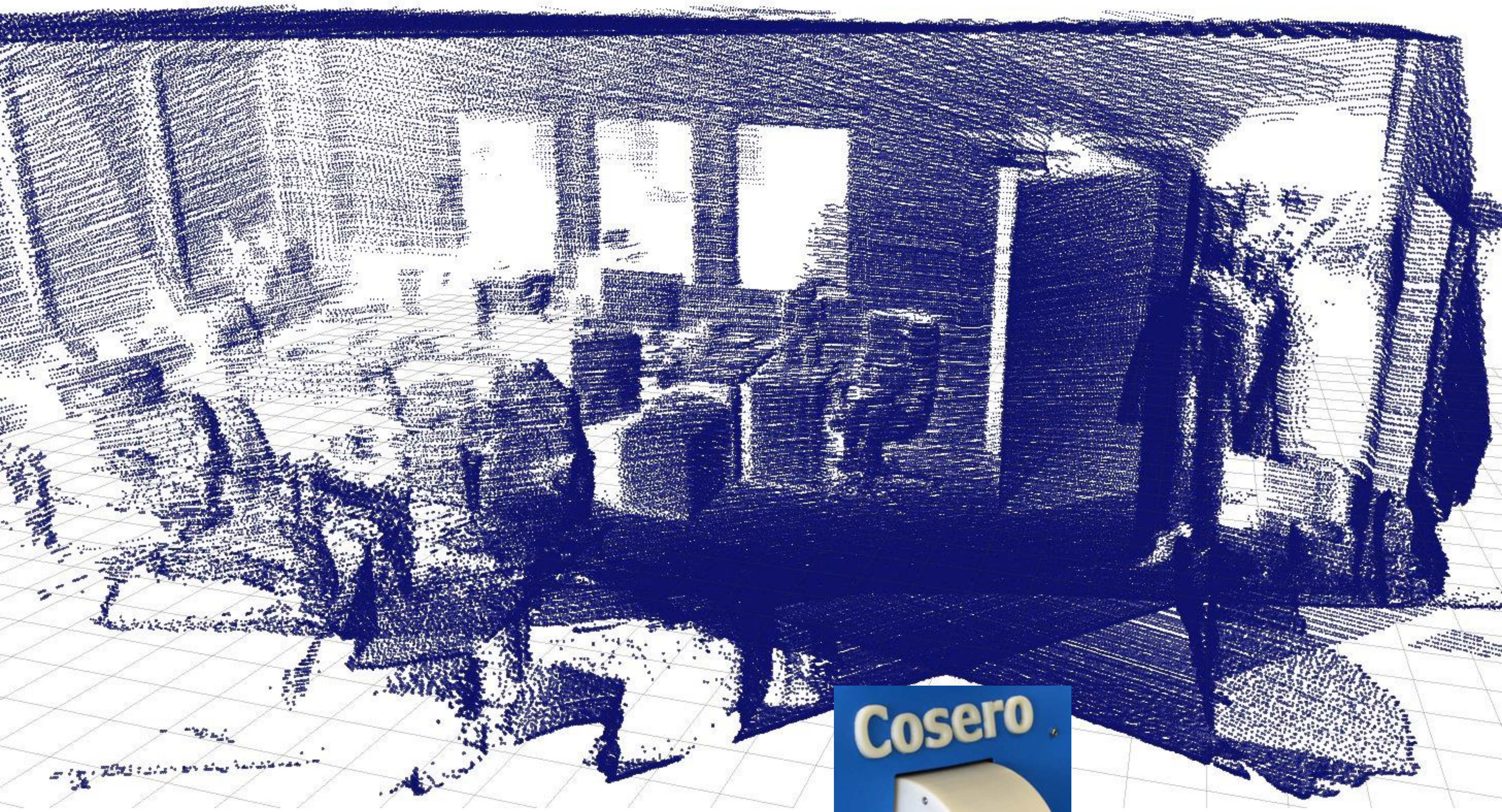


Work	MedPose	MedPose(C)	MedPose(I)	AvePose	AvePose(C)	AvePose(I)
Lai *et al.* (9)	62.6	51.5	30.2	83.7	77.7	57.1
Bo *et al.* (10)	20.0	**18.7**	**18.0**	53.6	47.5	44.8
Ours – instance level pose regression	20.4	20.4	18.7	51.0	50.4	**42.8**
Ours – category level pose regression	**19.2**	19.1	18.9	**45.0**	**44.5**	43.7

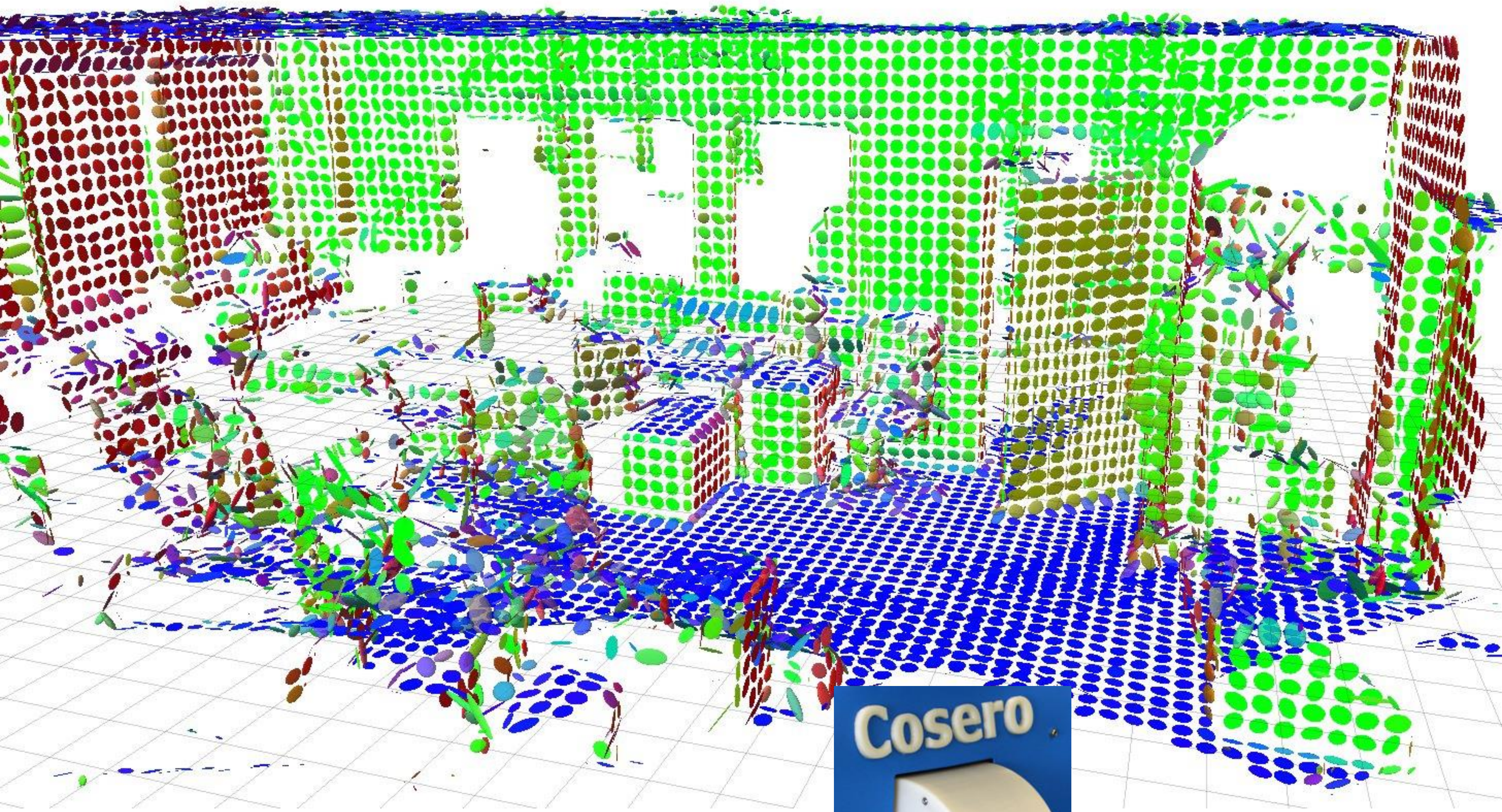
Angular error in °

[Schwarz, Schulz, Behnke, ICRA 2015]

3D-Mapping with Surfels

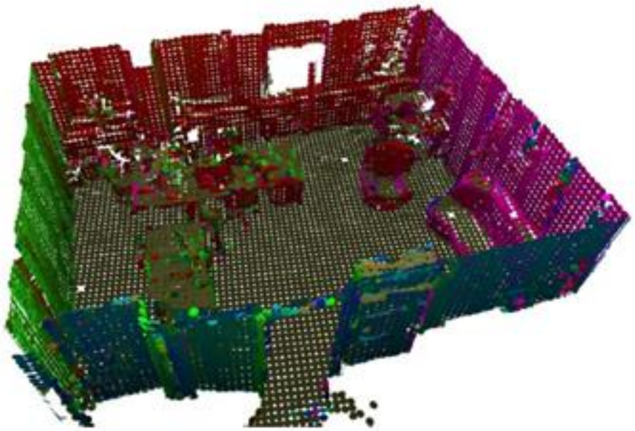
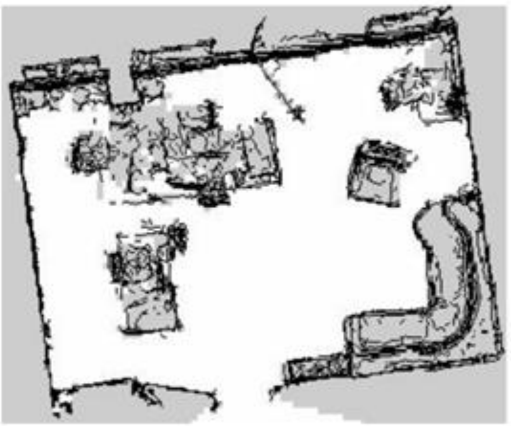


3D-Mapping with Surfels



3D-Mapping and Localization

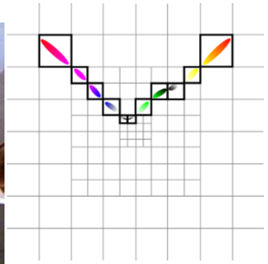
- Registration of 3D laser scans
- Representation of point distributions in voxels
- Drivability assessment trough region growing
- Robust localization using 2D laser scans

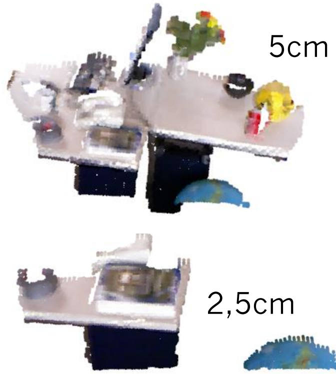




[Kläß, Stückler, Behnke: Robotik 2012]

3D Mapping by RGB-D SLAM

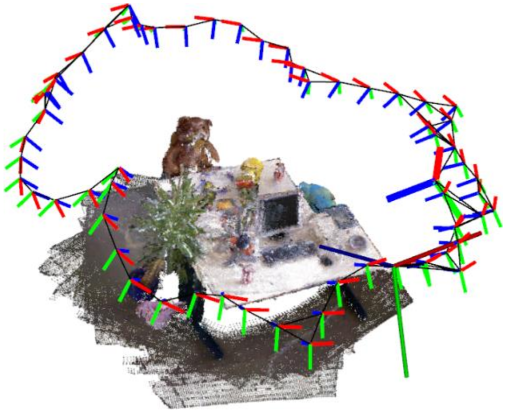
- Modelling of shape and color distributions in voxels
- Local multiresolution
- Efficient registration of views on CPU

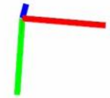




[Stückler, Behnke:
Journal of Visual Communication
and Image Representation 2013]

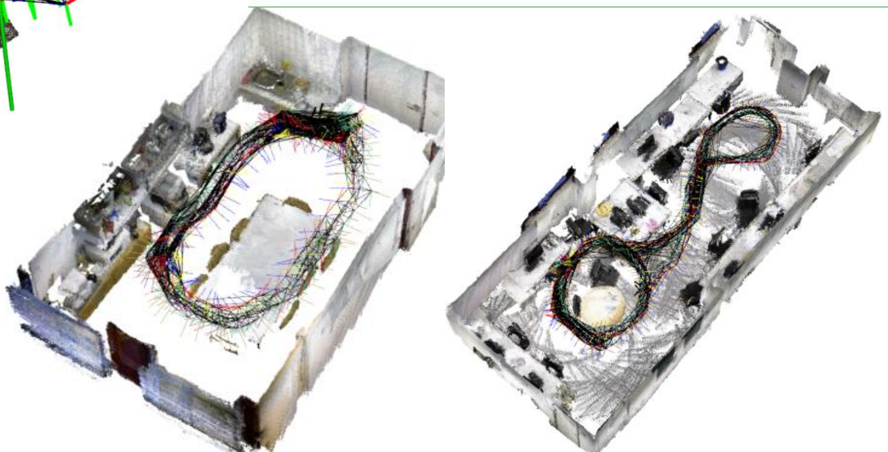
- Global optimization





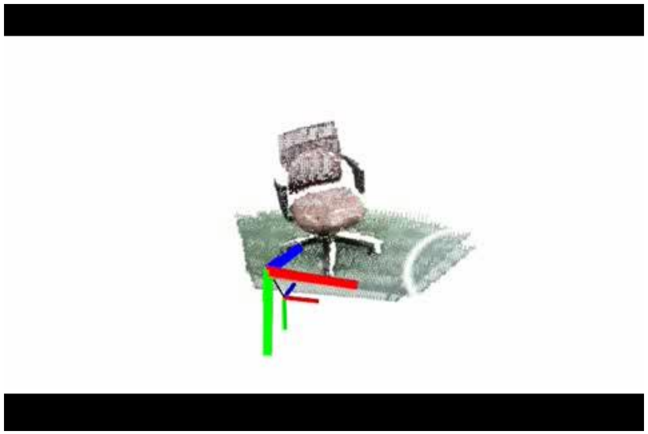
- Multi-camera SLAM

[Stoucken, Diplomarbeit 2013]

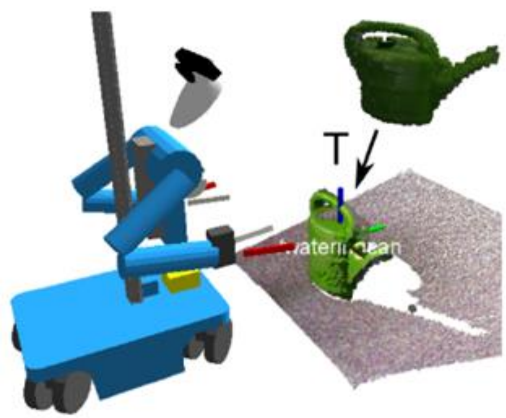


Learning and Tracking Object Models

- Modeling of objects by RGB-D-SLAM

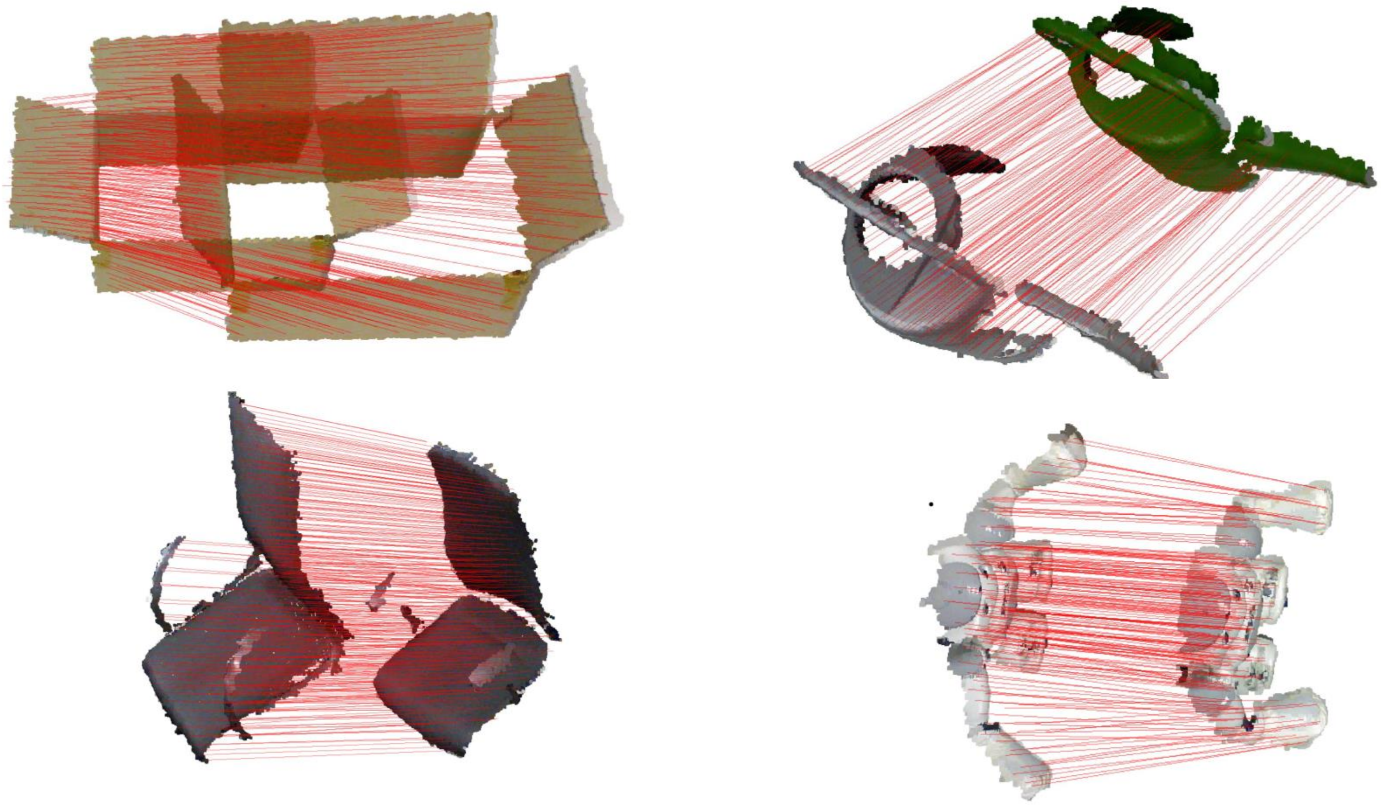


- Real-time registration with current RGB-D image



Deformable RGB-D-Registration

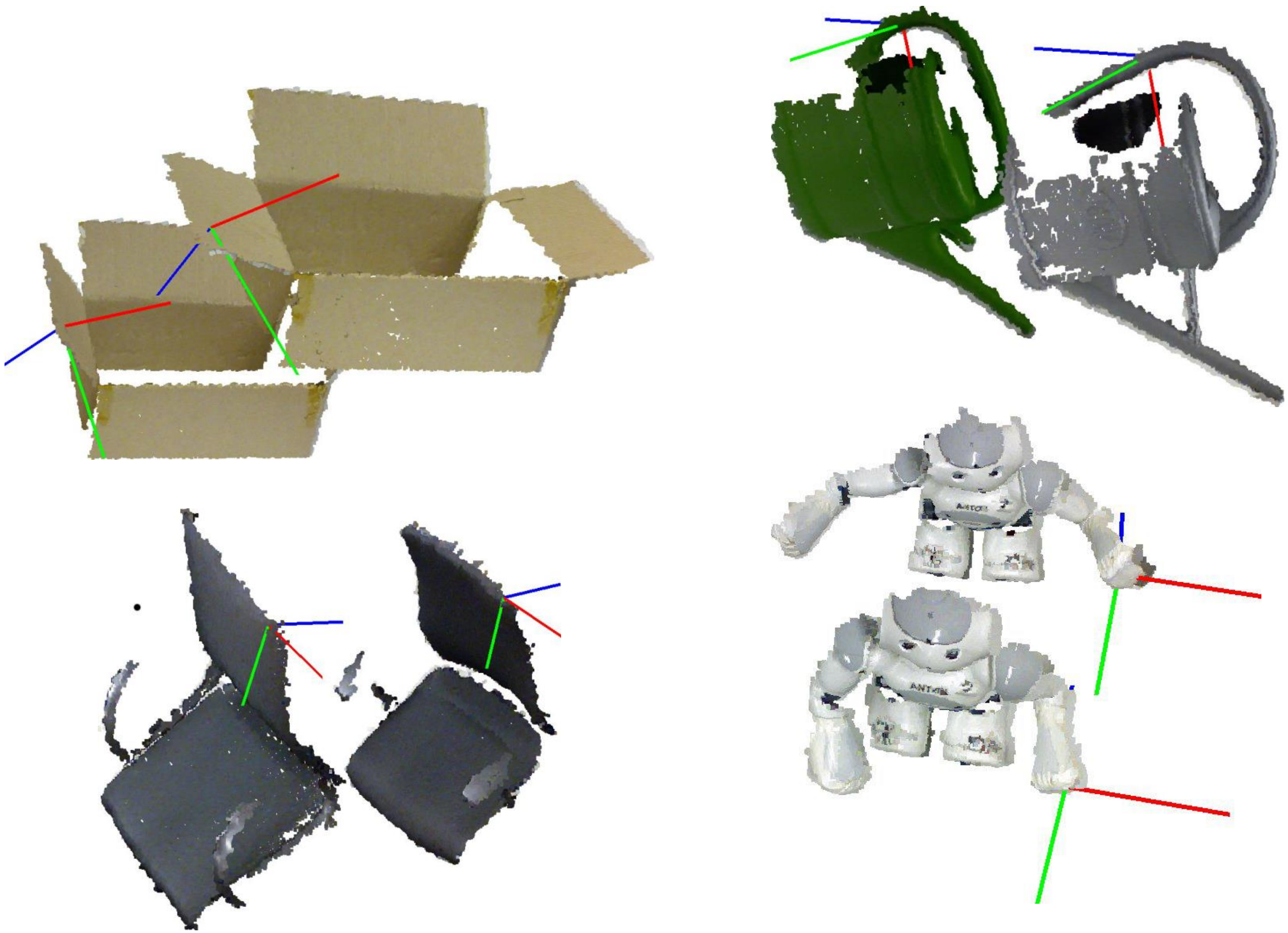
- Based on Coherent Point Drift method [Myronenko & Song, PAMI 2010]
- Multiresolution Surfel Map allows real-time registration



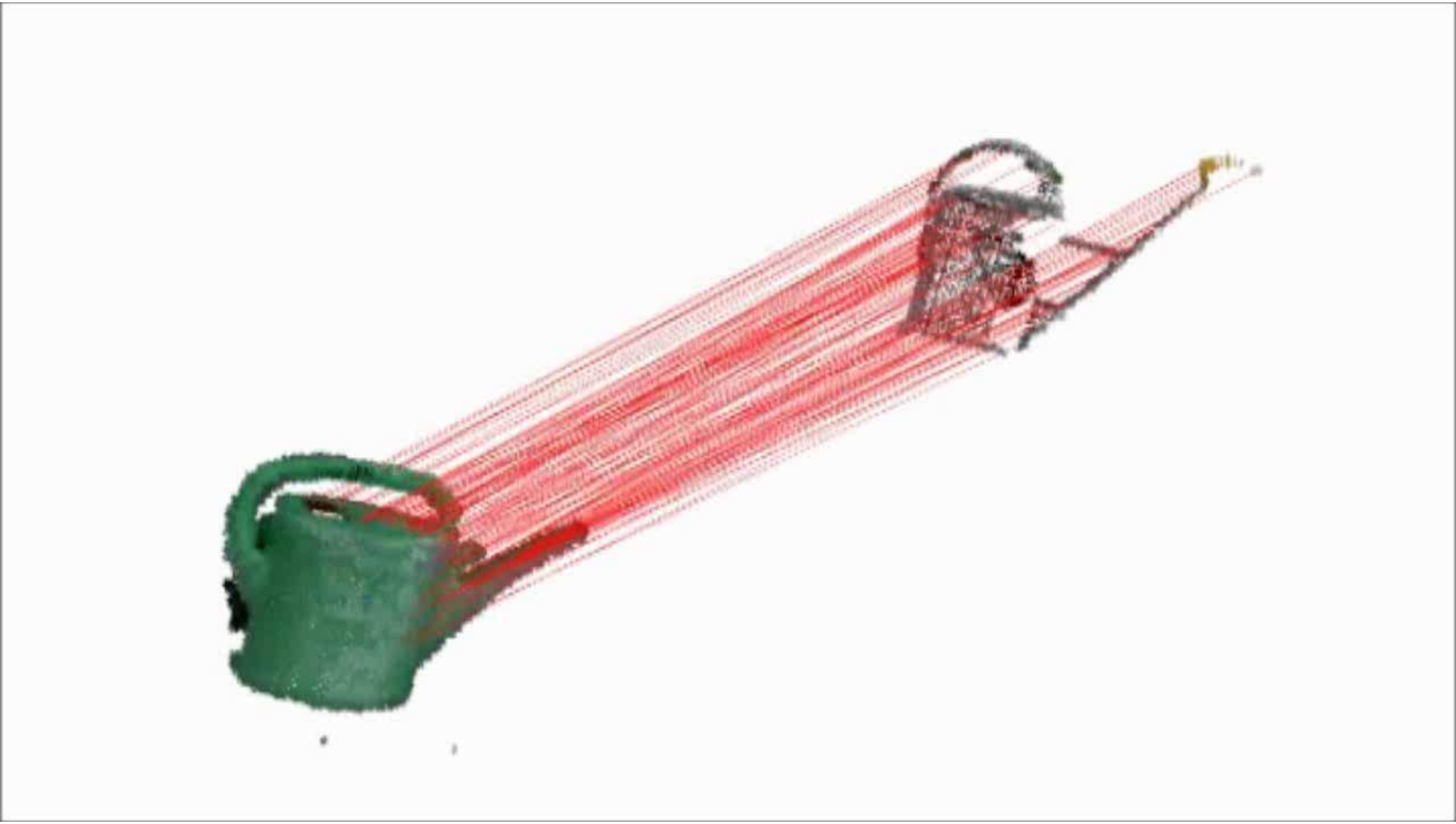
[Stückler, Behnke, ICRA2014]

Transformation of Poses on Object

- Derived from the deformation field



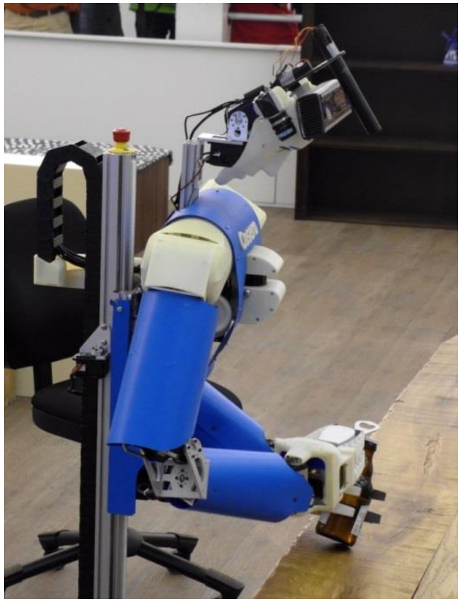
Grasp & Motion Skill Transfer

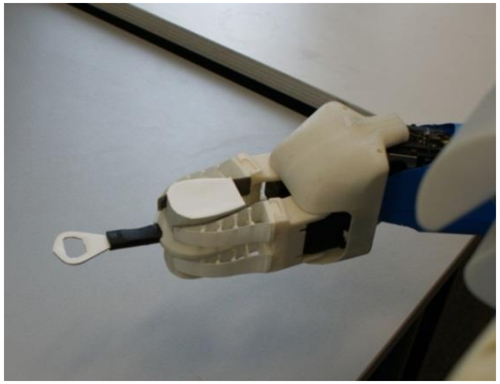


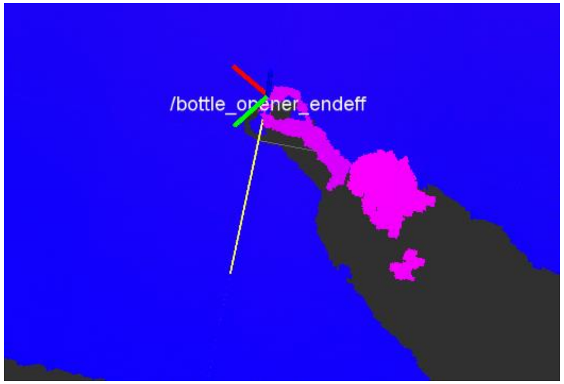
- Demonstration at RoboCup 2013

Tool use: Bottle Opener

- Tool tip perception
- Extension of arm kinematics
- Perception of crown cap
- Motion adaptation

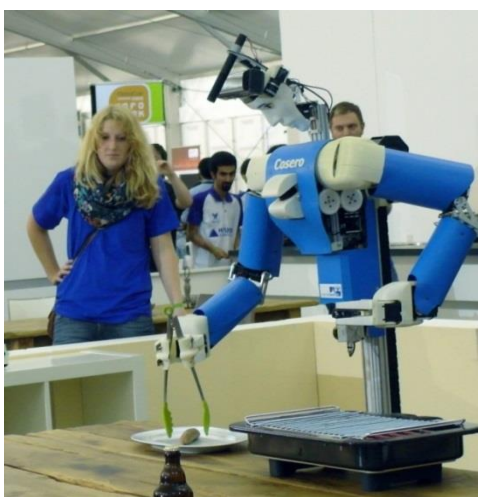






Picking Sausage, Bimanual Transport

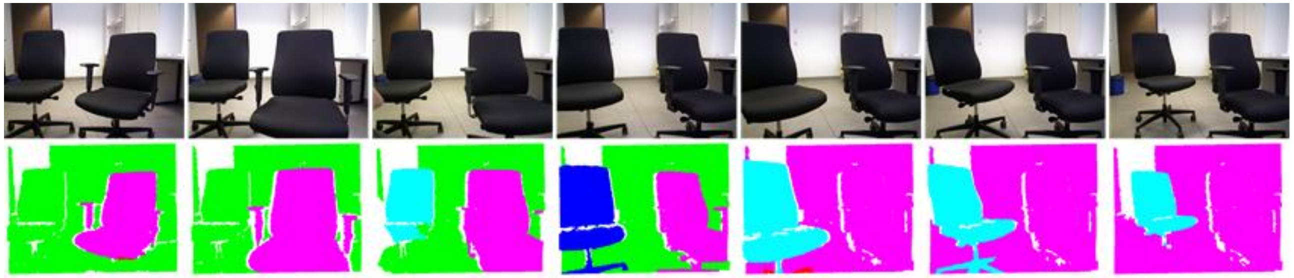
- Perception of tool tip and sausage
- Alignment with main axis of sausage



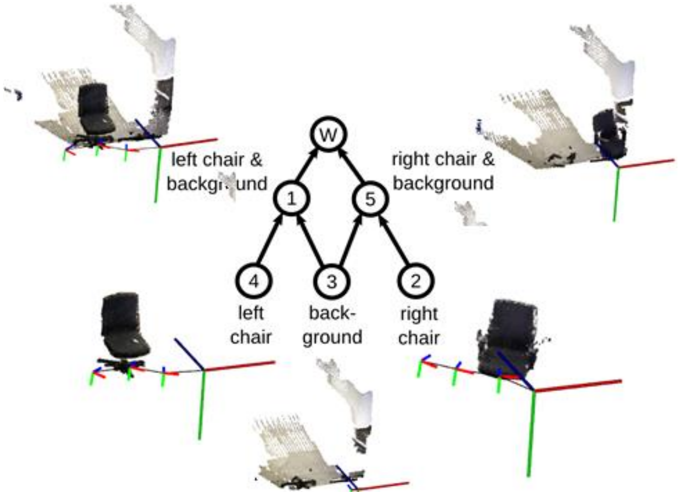
- Our team NimbRo won the RoboCup@Home League in three consecutive years

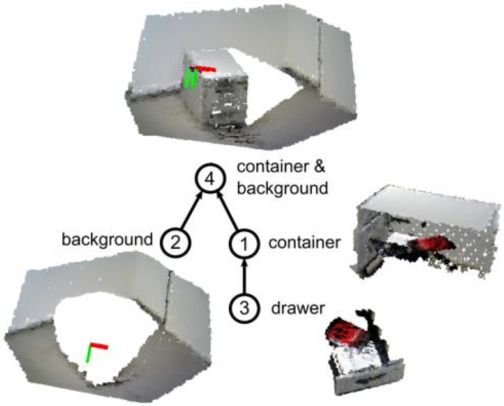
Hierarchical Object Discovery trough Motion Segmentation

- Motion is strong segmentation cue
- Both camera and object motion
- Segment-wise registration of a sequence



- Inference of a segment hierarchy

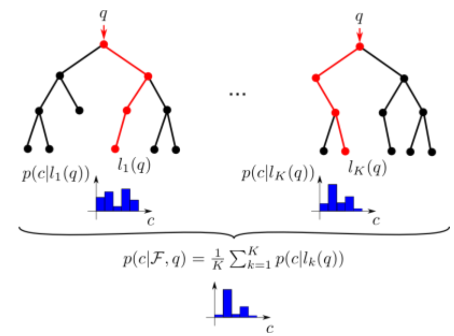


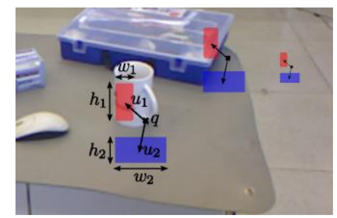


[Stückler, Behnke: IJCAI 2013]

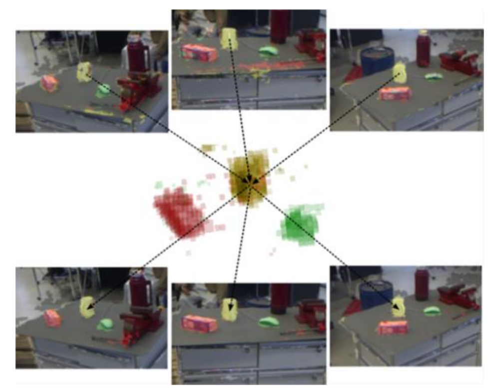
Semantic Mapping

- Pixel-wise classification of RGB-D images by random forests
- Inner nodes compare color / depth of regions
- Size normalization
- Training and recall on GPU
- 3D fusion through RGB-D SLAM
- Evaluation on own data set and NYU depth v2





[Stückler, Biresev, Behnke: IROS 2012]



Ground truth

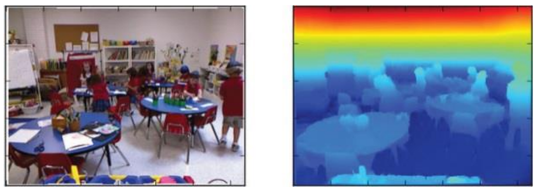
Segmentation

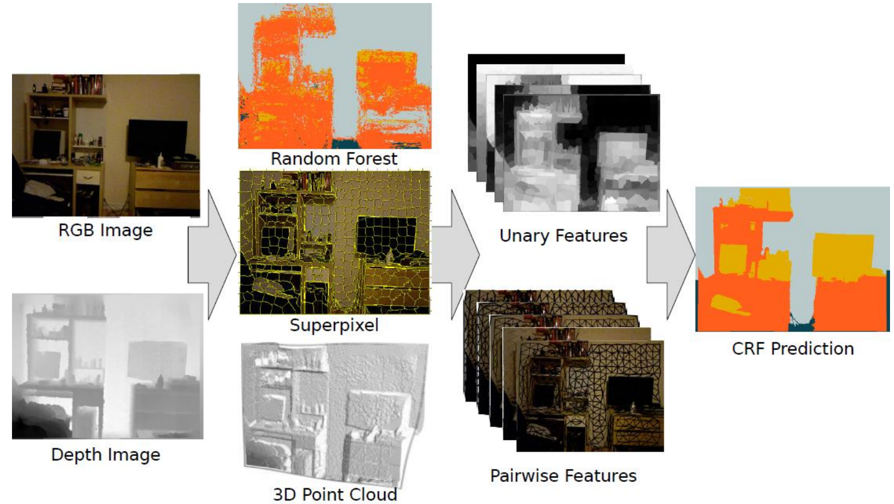
Accuracy in %	Ø Classes	Ø Pixels
Silberman et al. 2012	59,6	58,6
Couprie et al. 2013	63,5	64,5
Random forest	65,0	68,1
3D-Fusion	**66,8**	**70,6**

[Stückler et al., Journal of Real-Time Image Processing 2014]

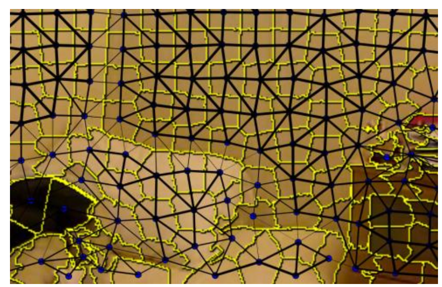
Learning Depth-sensitive CRFs

- SLIC+depth super pixels
- Unary features: random forest
- Height feature





- Pairwise features
  - Color contrast
  - Vertical alignment
  - Depth difference
  - Normal differences



similarity between superpixel normals

- Results:

Random forest

CRF prediction

Ground truth

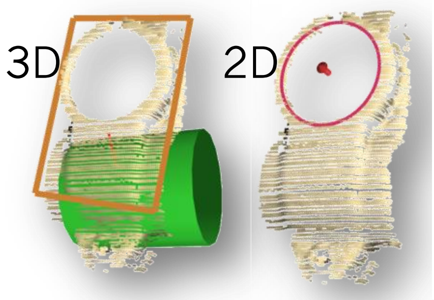
	class average	pixel average
RF	65.0	68.3
RF + SP	65.7	70.1
RF + SP + SVM	70.4	70.3
RF + SP + CRF	**71.9**	**72.3**
Silberman *et al.*	59.6	58.6
Couprie *et al.*	63.5	64.5

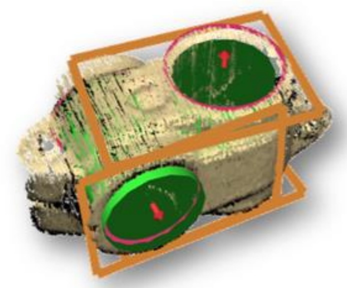
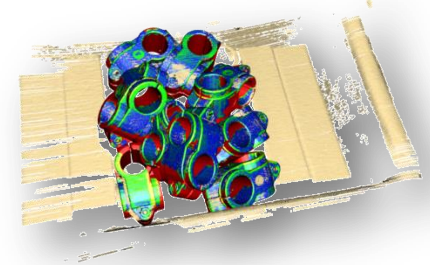
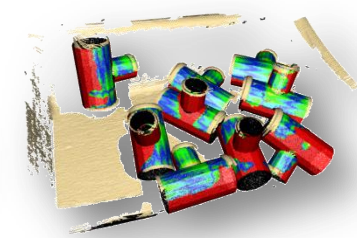
[Müller and Behnke, ICRA 2014]

Bin Picking

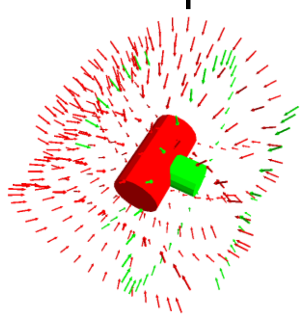
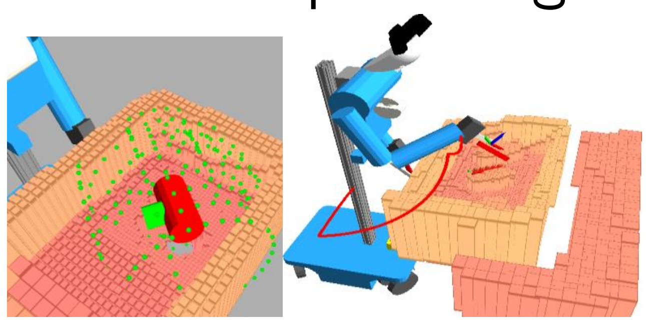
- Known objects in transport box


- Matching of graphs of 2D and 3D shape primitives



- Grasp and motion planning

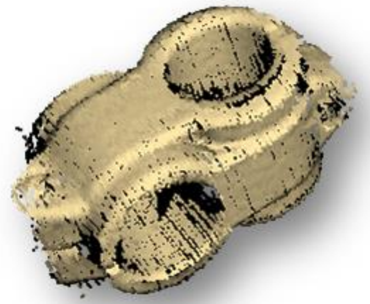



Offline Online

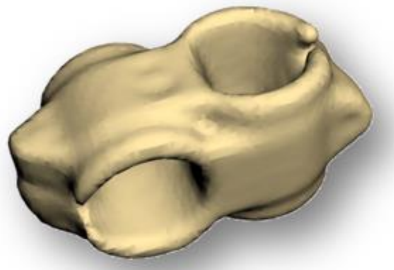
[Nieuwenhuisen et al.: ICRA 2013]

Learning of Object Models

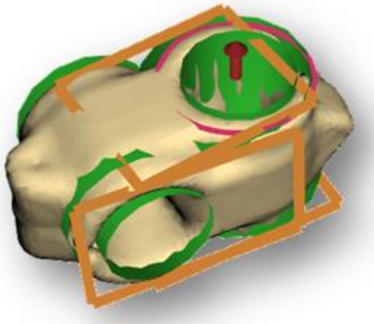
- Scan multiple objects in different poses
- Find support plane and remove it
- Segment views
- Register views using ICP
- Recognize geometric primitives



Registered views

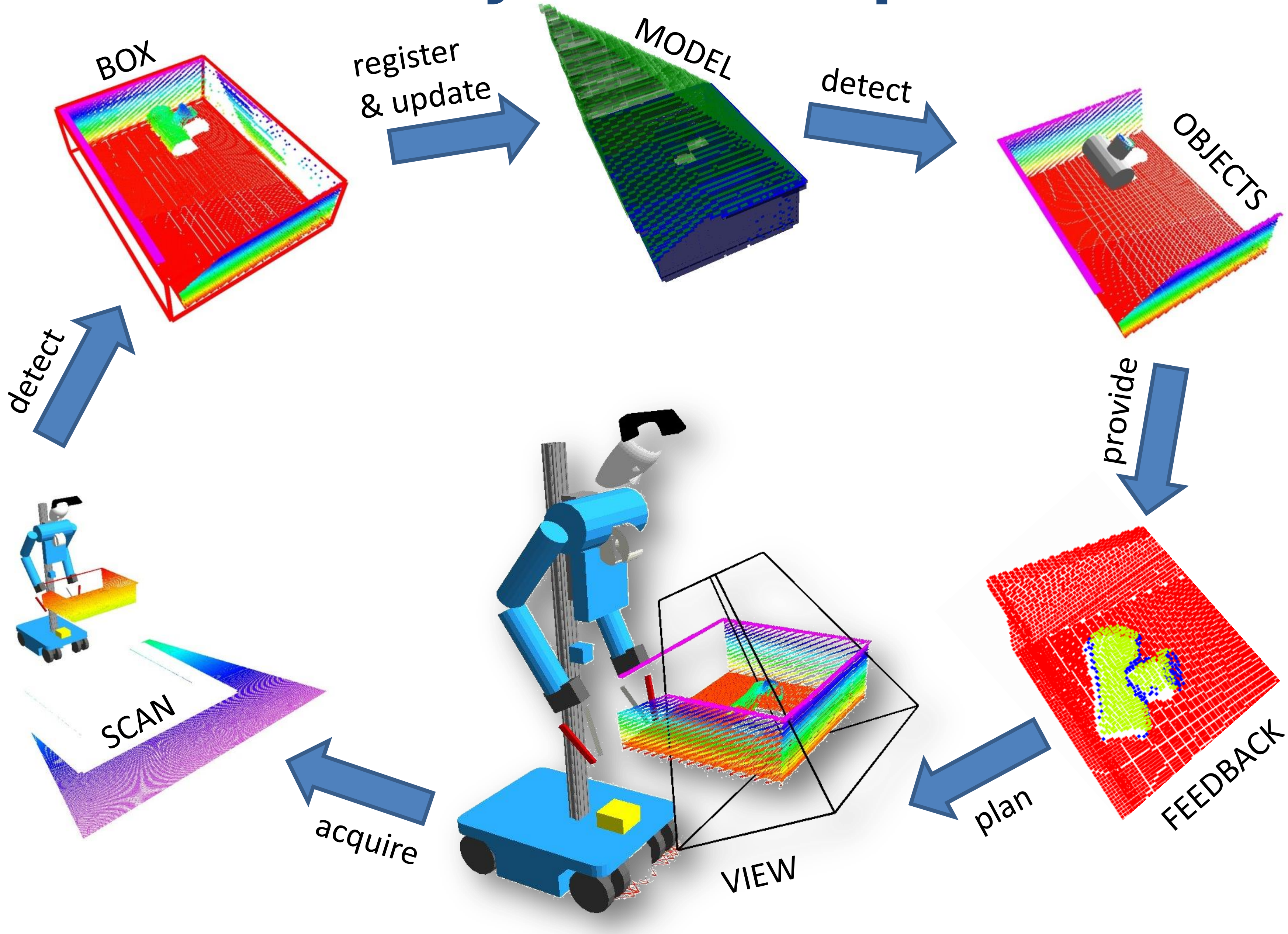


Surface reconstruction

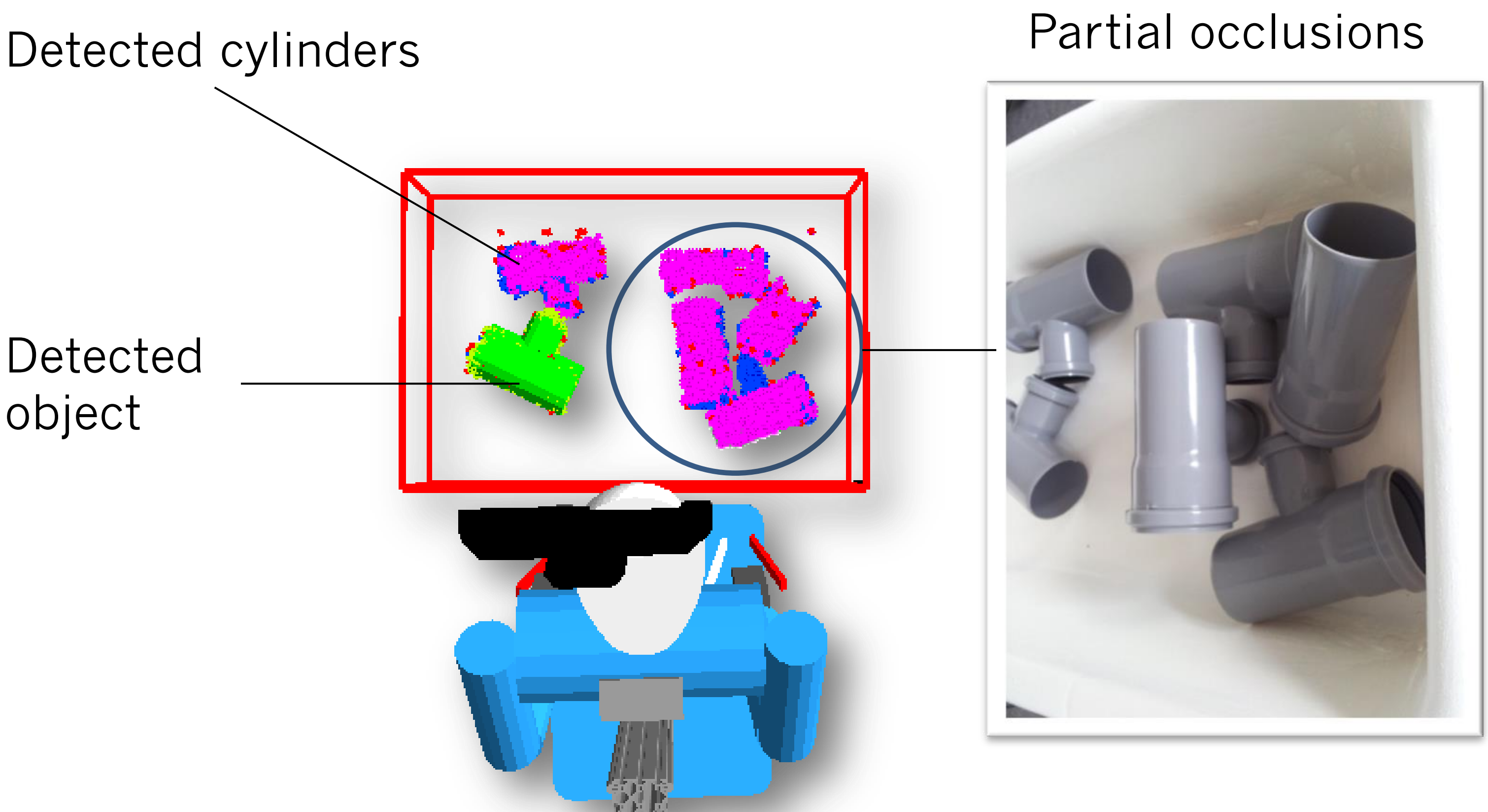


Detected primitives

Active Object Perception

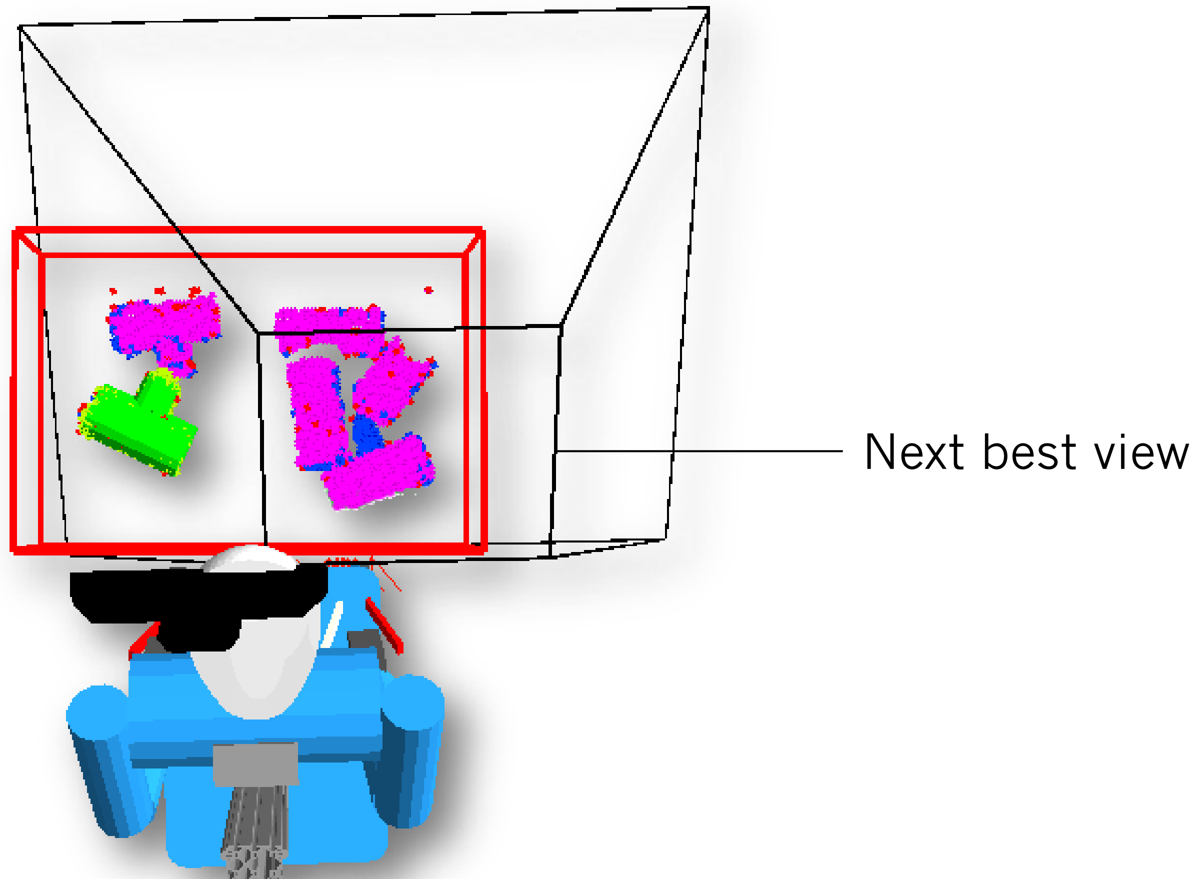


Active Object Perception



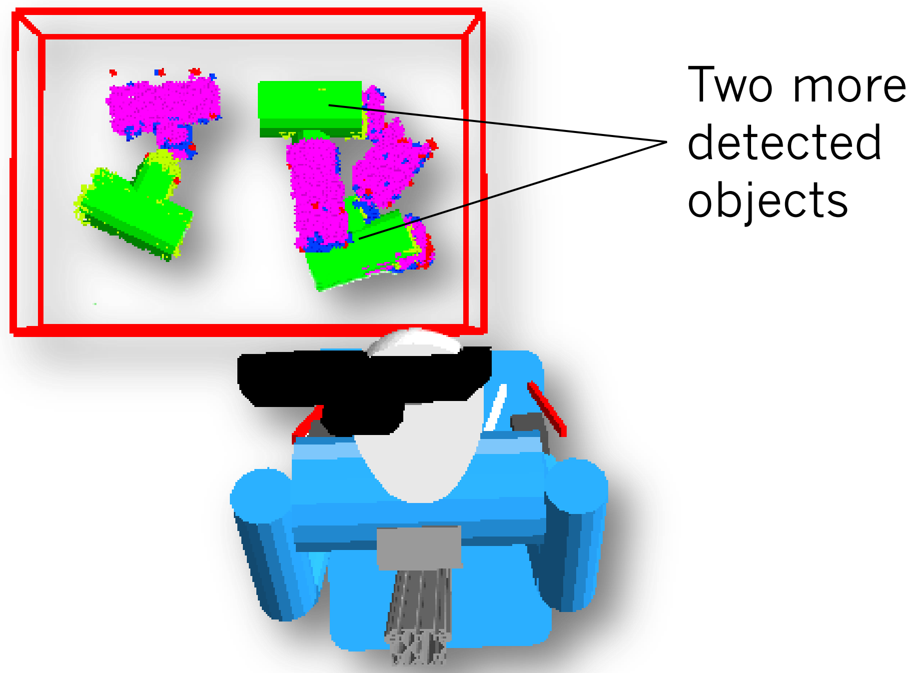
- Efficient exploration of the part arrangement in the transport boxes to handle occlusions

Active Object Perception



- Efficient exploration of the part arrangement in the transport boxes to handle occlusions

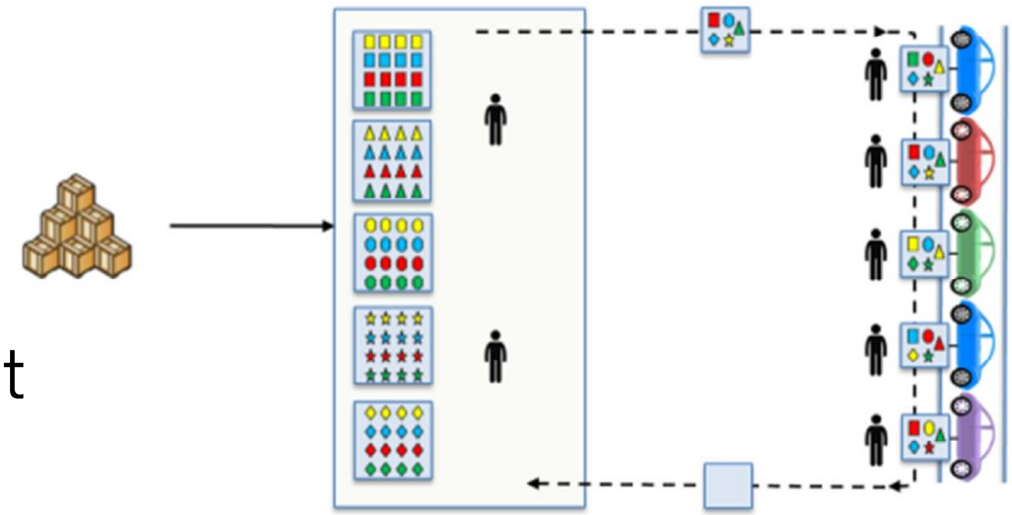
Active Object Perception



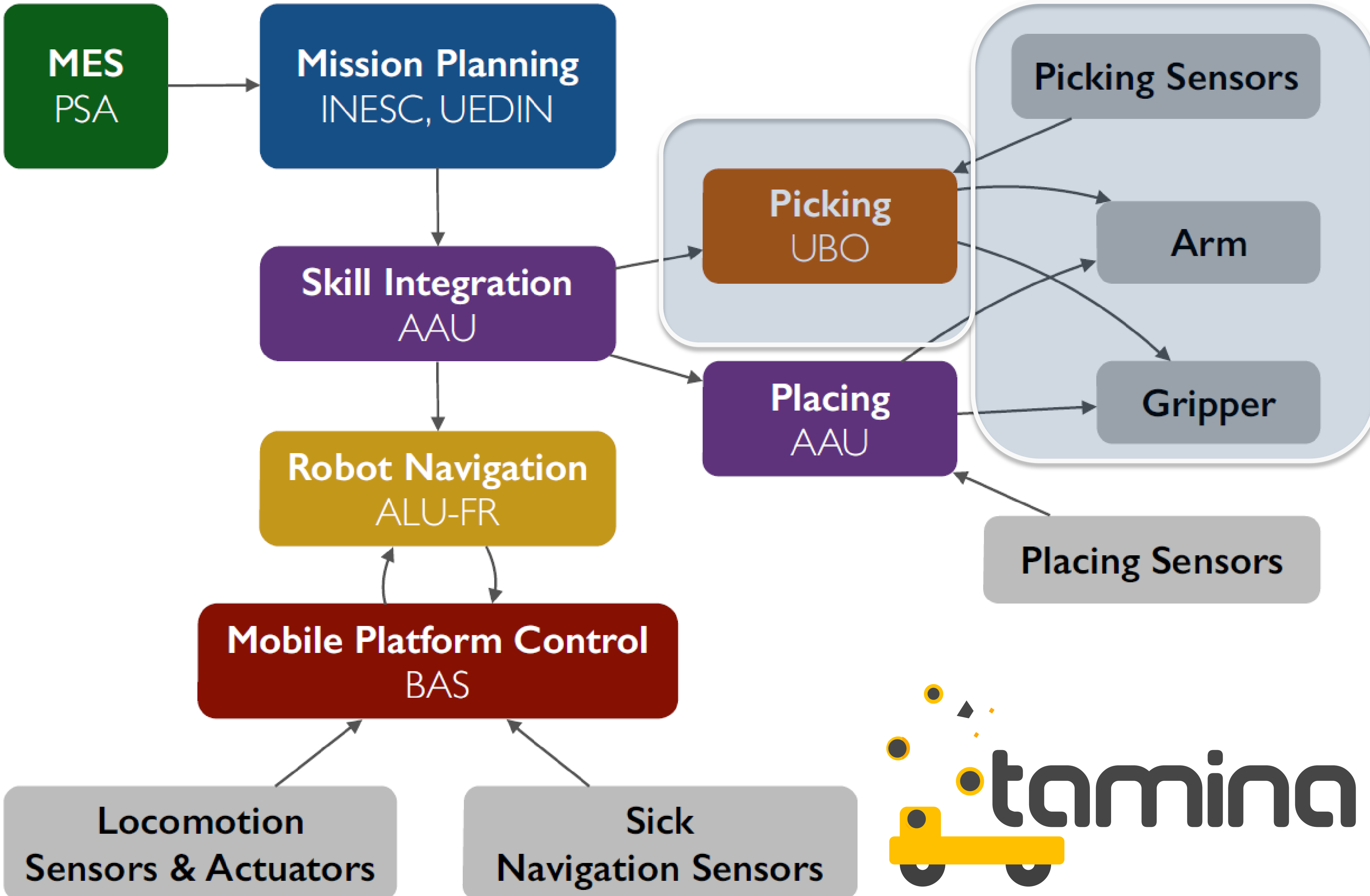
- Efficient exploration of the part arrangement in the transport boxes to handle occlusions

Use-case: Kitting of Automotive Parts

- Many car variants
- Collect the parts needed for the assembly of a particular car in a kit
- Parts in different variants are available in a supermarket
- Robot needs to
  - navigate to the transport boxes,
  - grasp the parts, and
  - place them in the kit

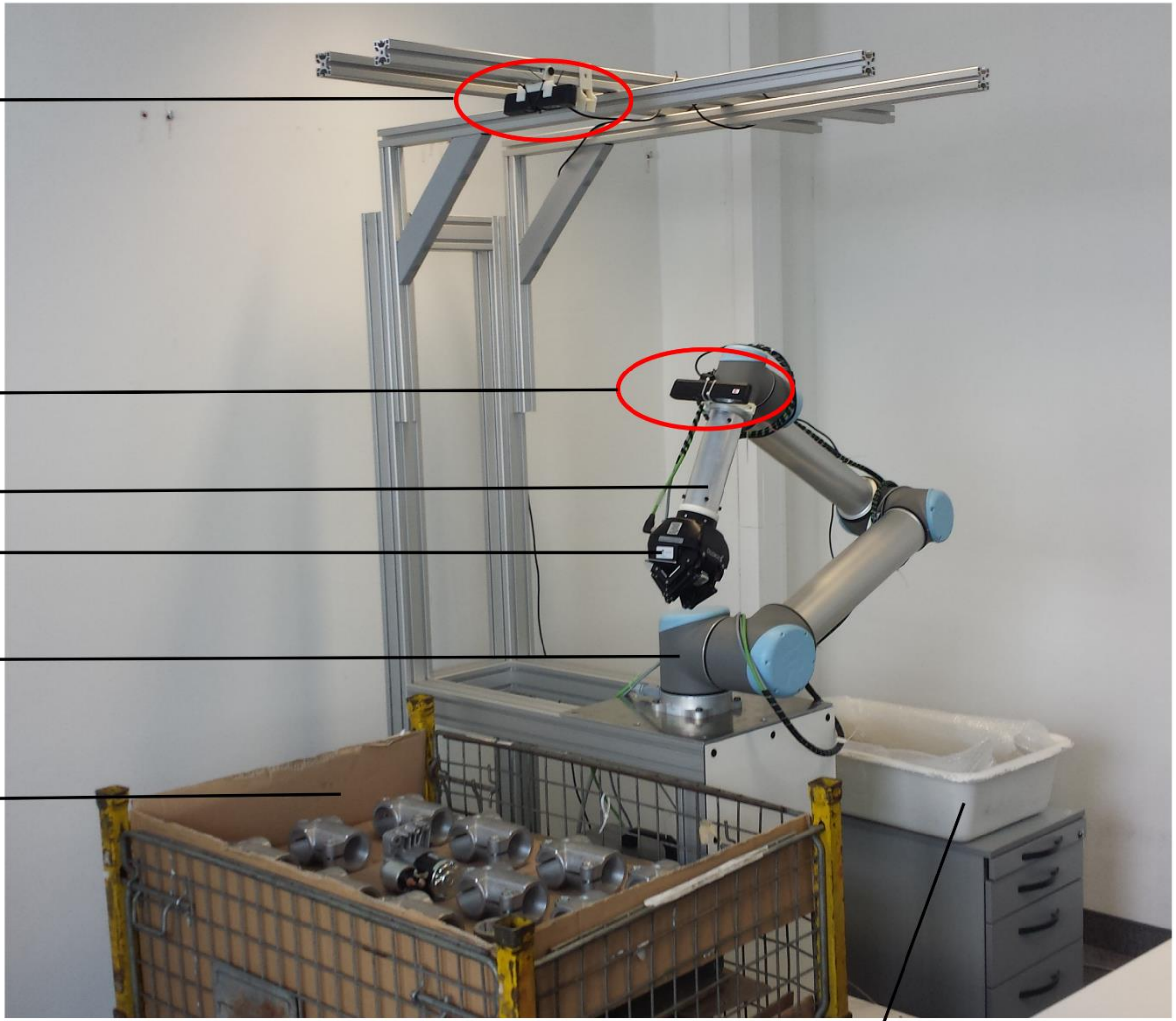


Sustainable and Reliable Robotics for Part Handling in Manufacturing Automation



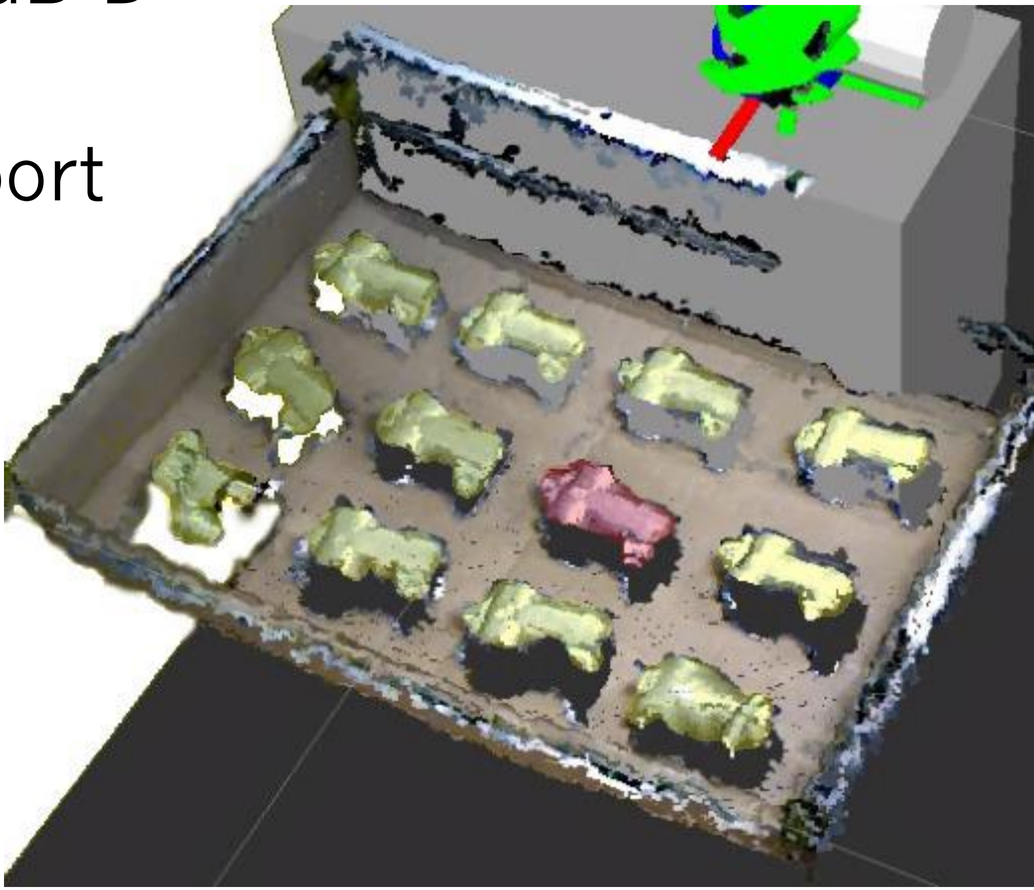
tamina

Realized Lab Demonstrator



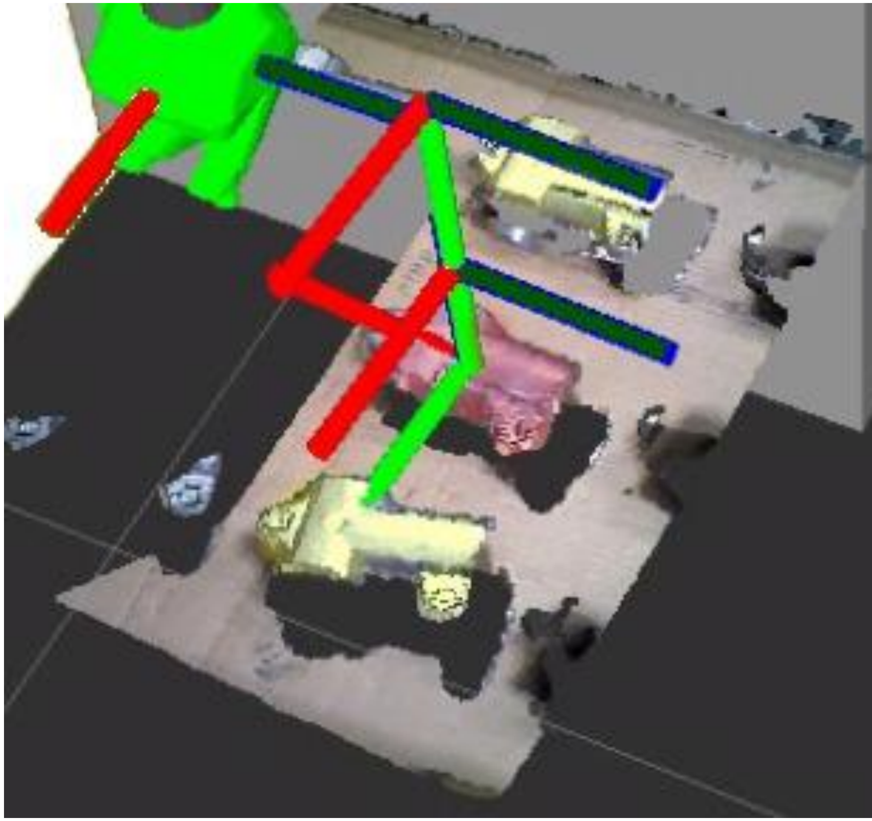
Object Candidate Detection

- Using work space RGB-D camera
- Initial pose of transport box roughly known
- Detect dominant horizontal plane above ground
- Cluster points above support plane
- Estimate main axes



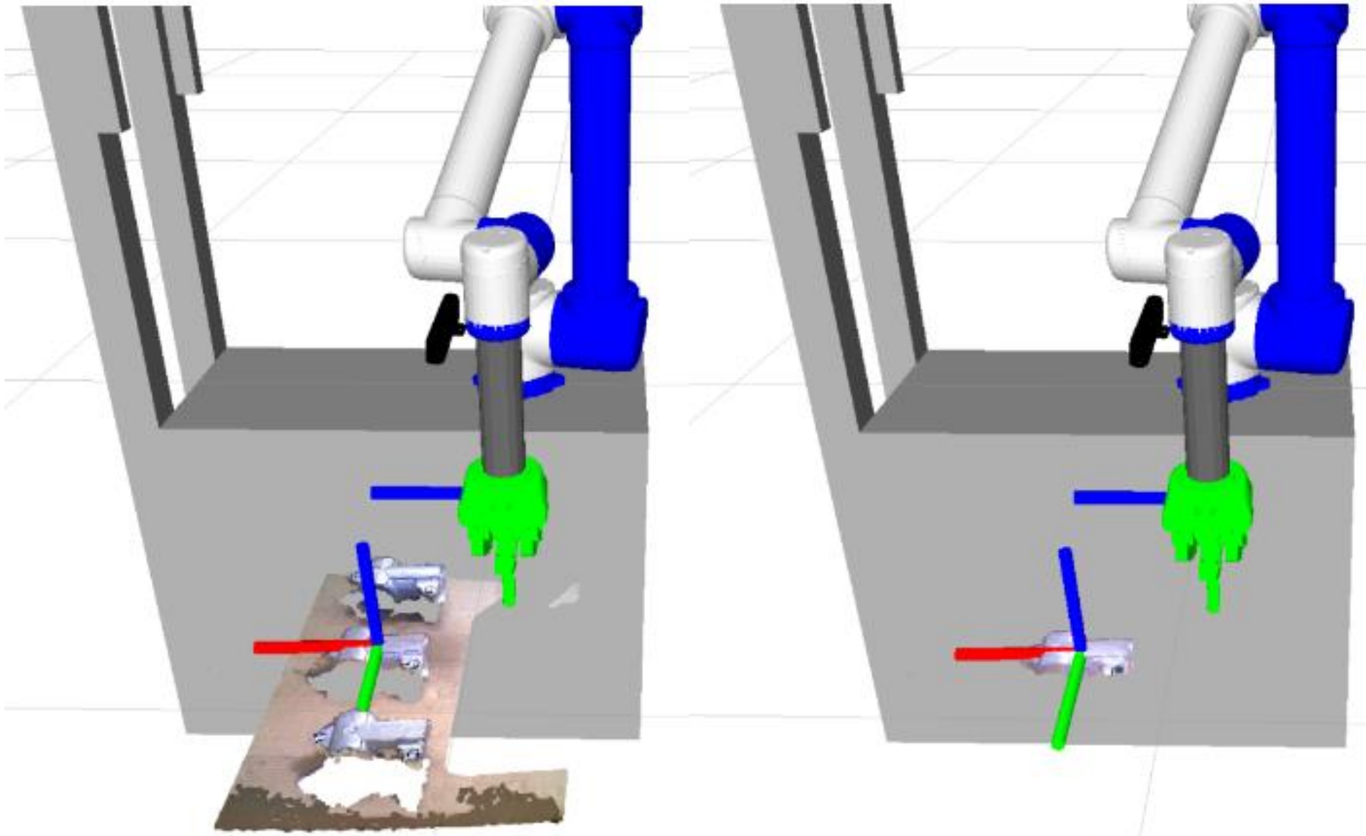
Object View Registration

- Wrist RGB-D camera moved above innermost object candidate
- Object views are represented as Multiresolution Surfel Map
- Registration of object view with current measurements using soft assignments
- Verification based on registration quality



Registered Object Model

- Registration yields the object pose

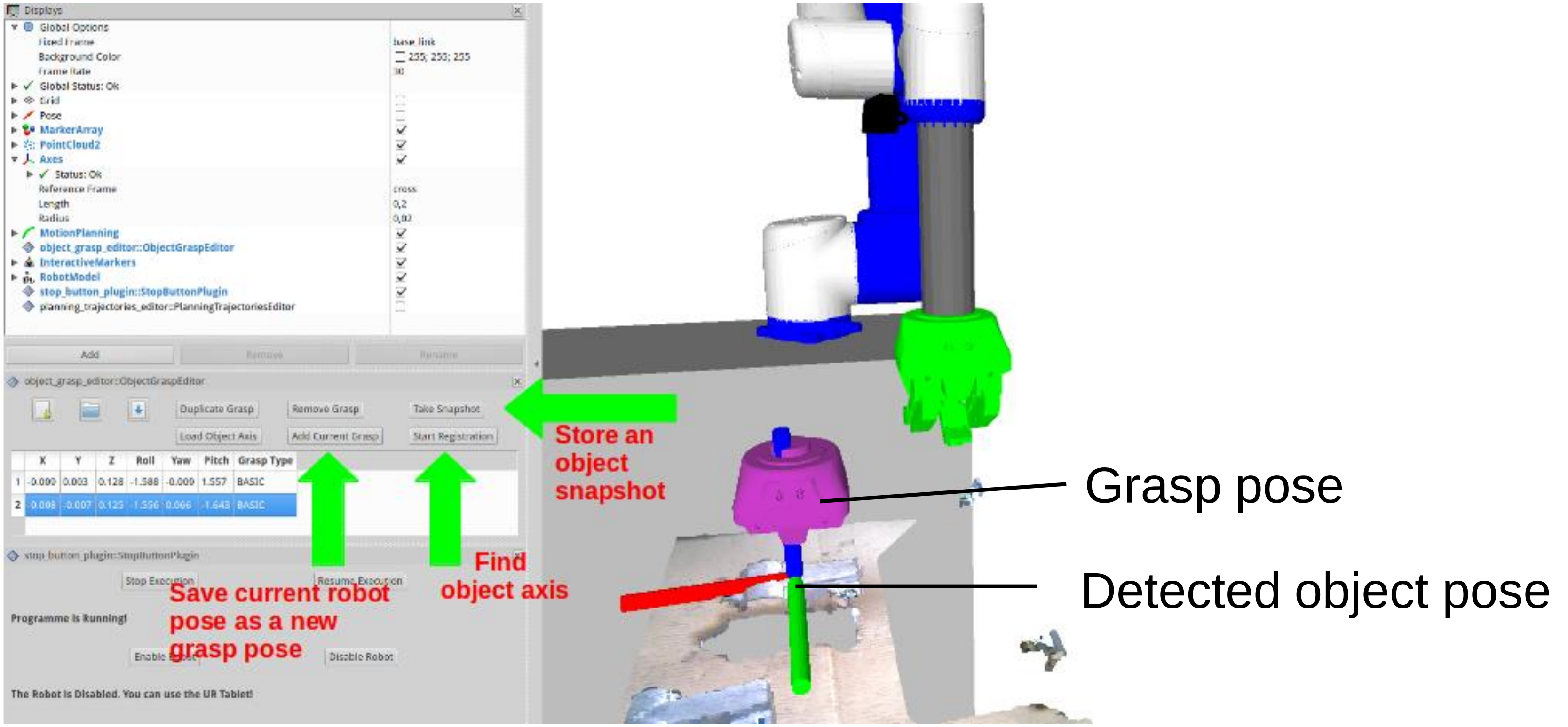


Detected Object Pose

Object Model and Pose

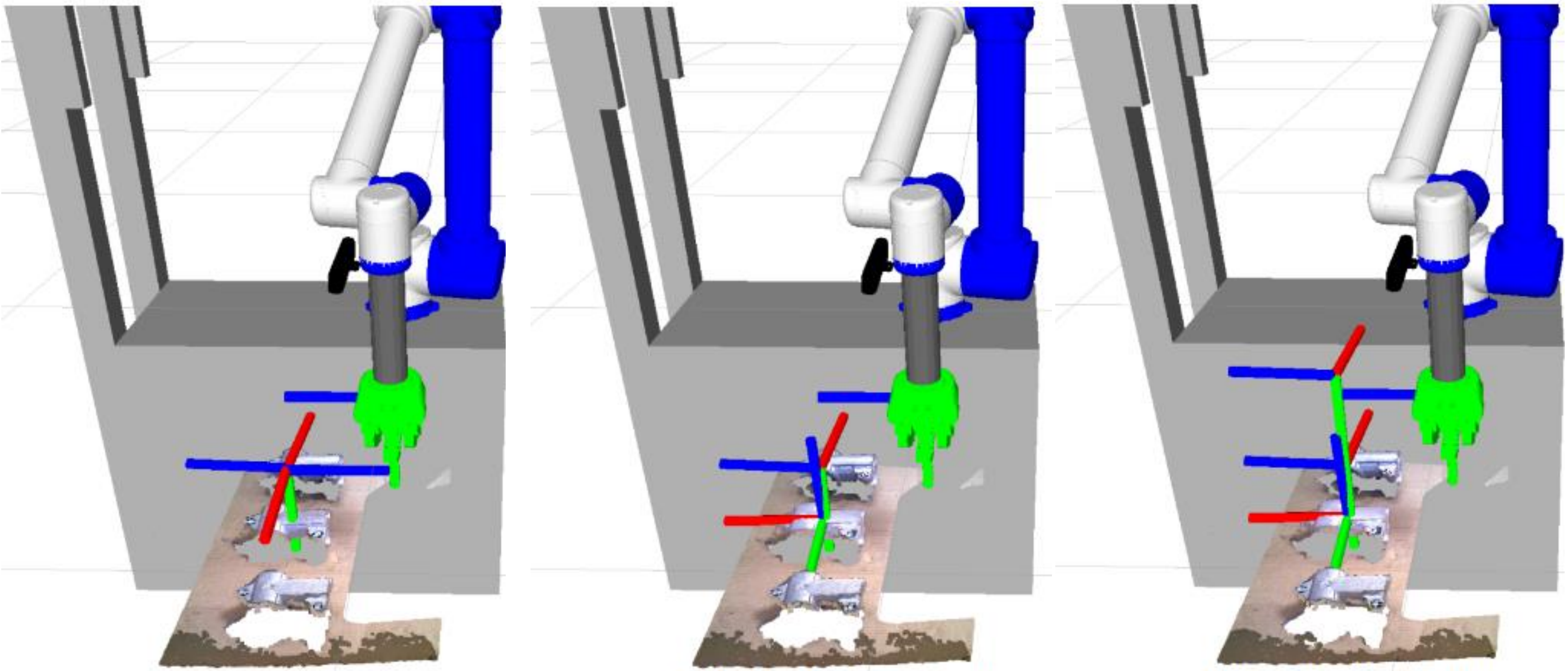
Grasp Definition

- GUI for object model acquisition and grasp definition, relative to object model



Grasp Selection

- Grasps are selected according to object pose

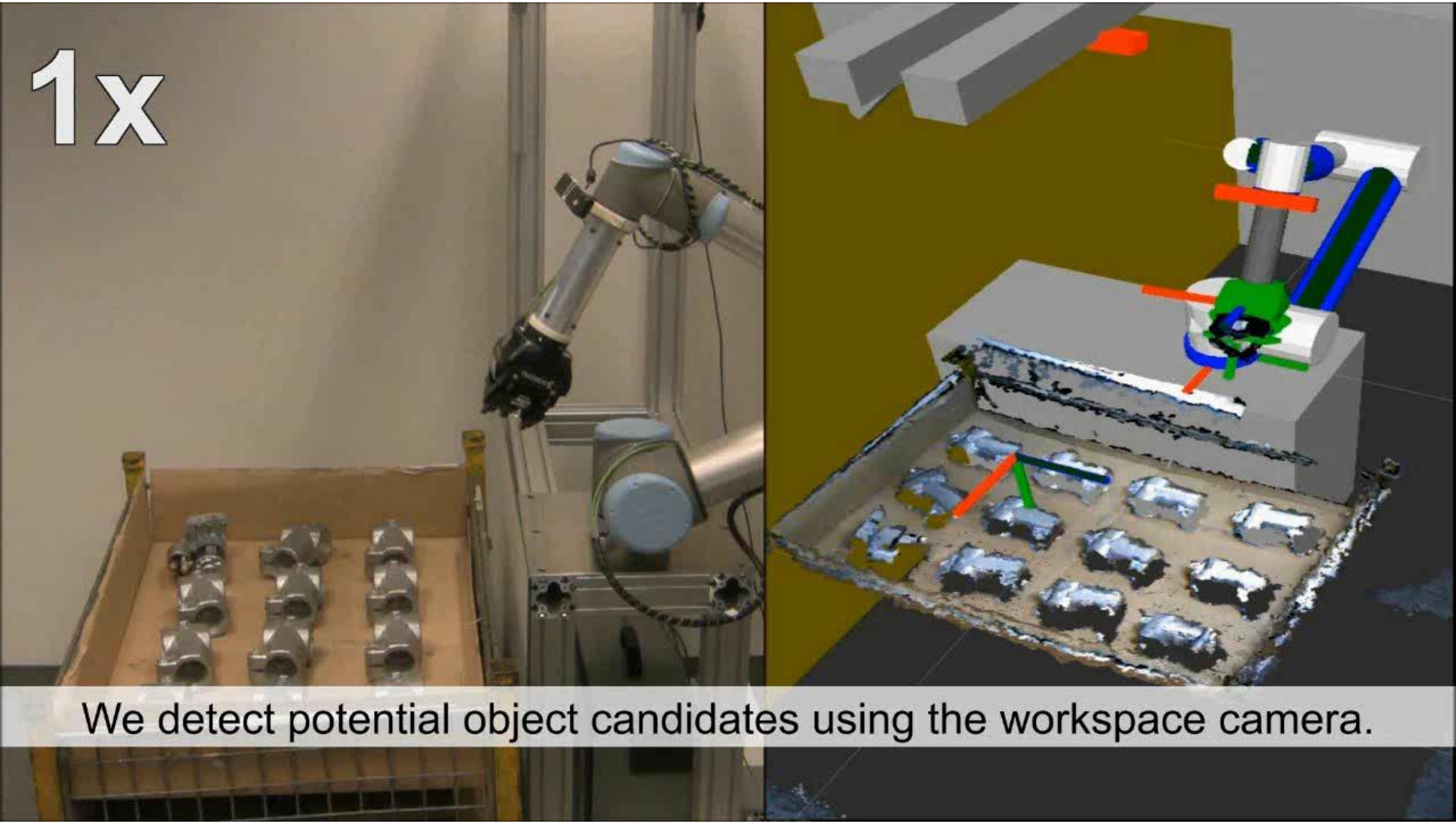


Grasp Candidates

Selected Grasp and Object Pose

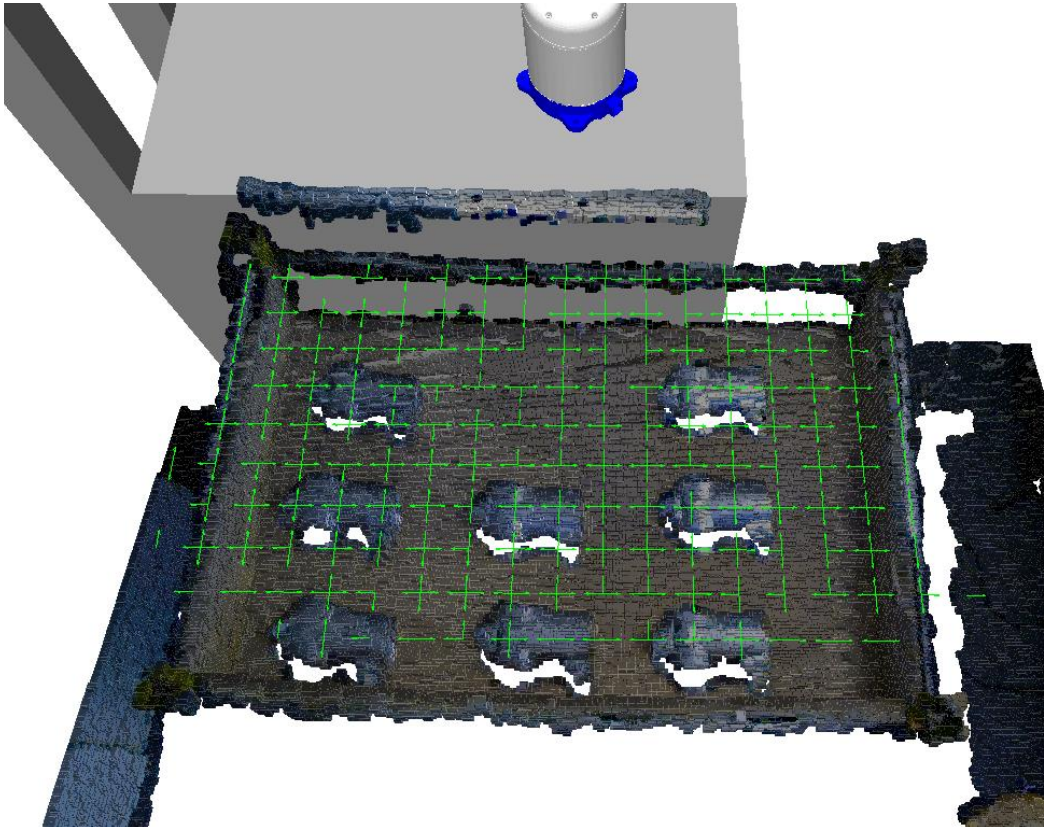
Pre-Grasp, Grasp and Object Pose

Part Detection and Grasping



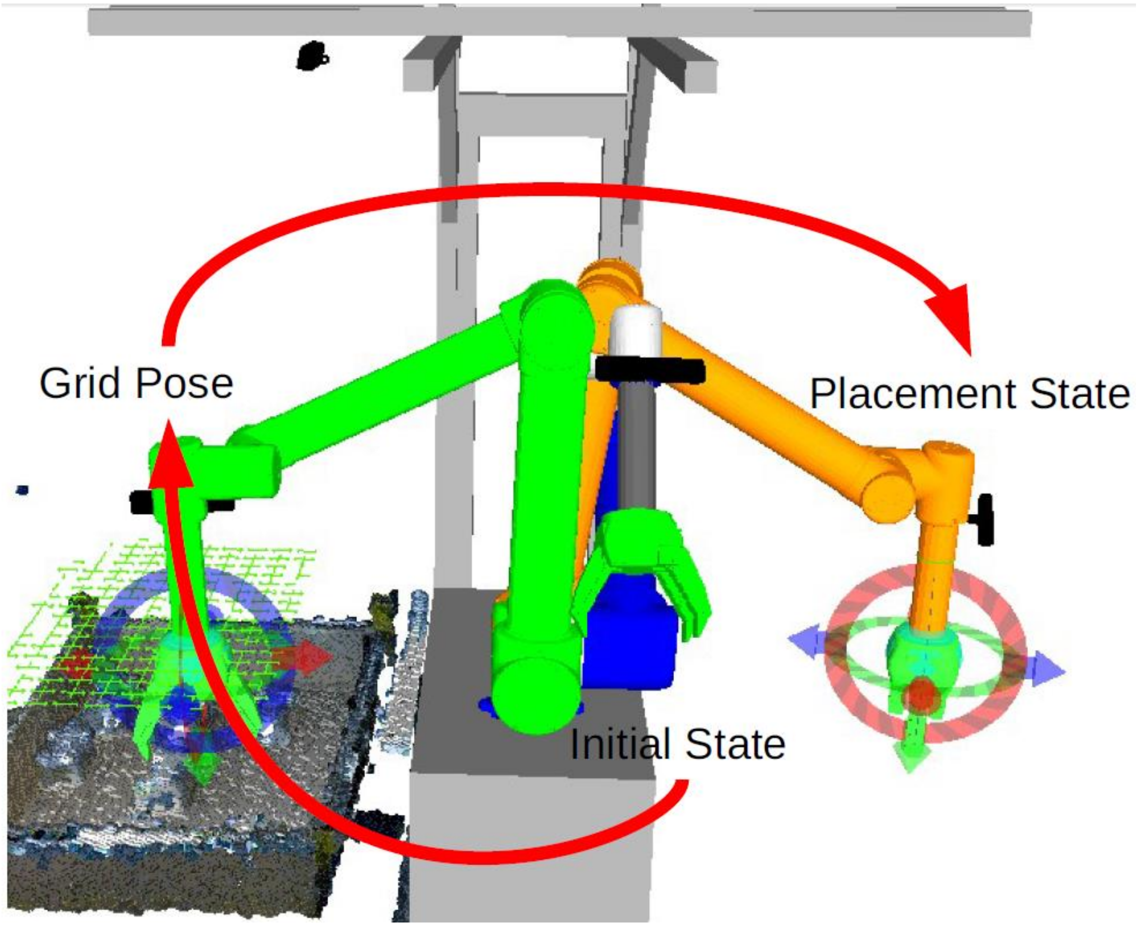
Motion Planning

- Use ROS MoveIt for motion planning and execution
- Predefined poses (initial, placement) and grid of poses above the objects
- Preplanned paths
- Only short trajectories must be planned online

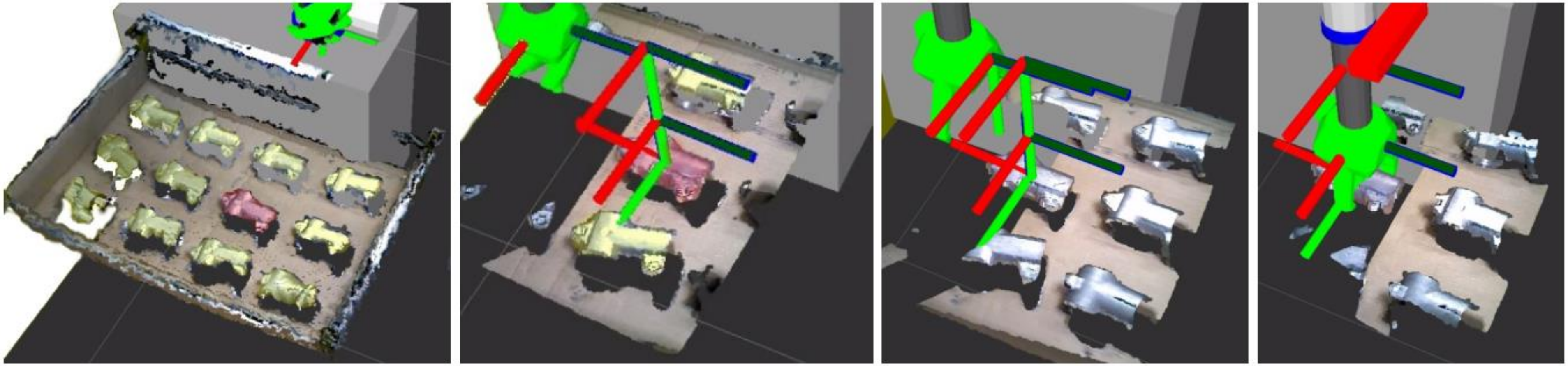


Concatenation of Motion Segments

- Interpolation for smooth segment transitions



First Test Sprint Results

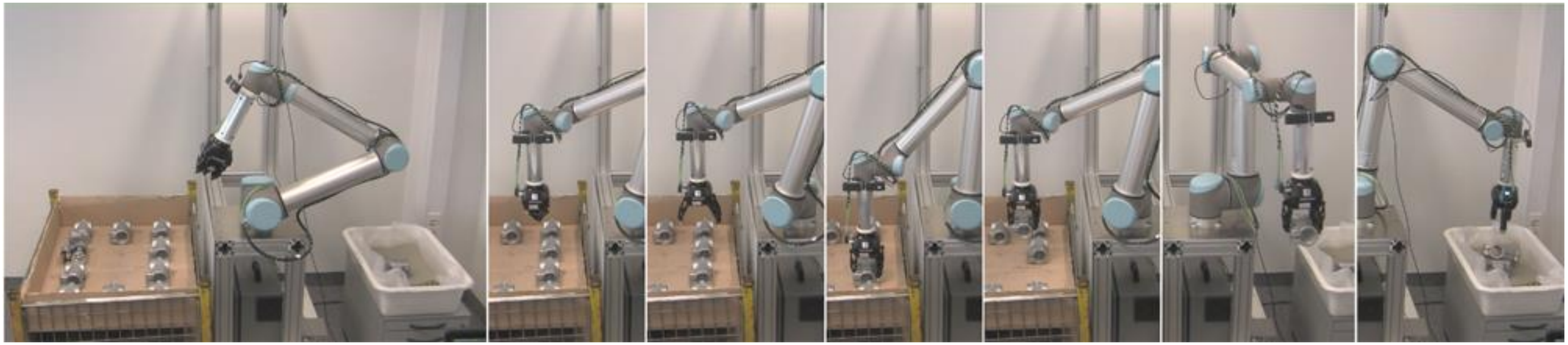


Fast Pipeline for Object Detection, Localization & Verification

1. Initial segmentation of the pallet’s support surface and object candidates (in real-time, <30 ms) using the workspace camera
2. Moving the arm to the object candidate closest to the pallet center
3. Registration using multi-resolution surfel mapping against a known object model (roughly 500 ms) using the wrist camera

Depalletizing of Starters

Depalletizing Results: 10 Runs



- Total time

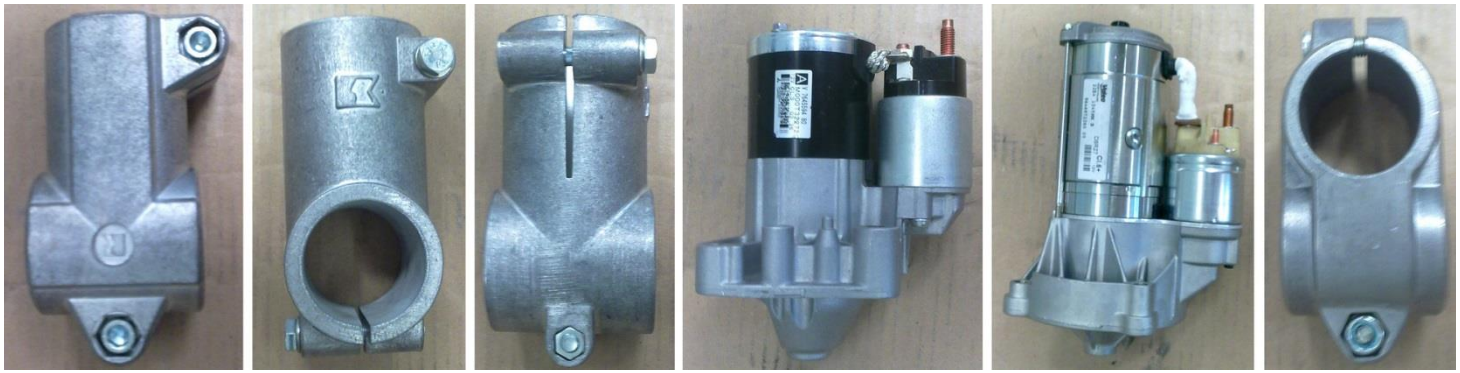
Component	Mean	Std	Min	Max
Object detection and grasping	13.84 s	1.89 s	10.42 s	23.81 s
Full cycle (incl. release and returning to initial pose)	34.57 s	3.01 s	29.53 s	49.52 s

- Component times and success rates

Component	Mean	Std	Min	Max	Success Rate
Initial object detection	26.3 ms	10.3 ms	0.02 ms	38.5 ms	100 %
Detecting that the pallet is empty					100 %
Object localization & verification	532.7 ms	98.2 ms	297.0 ms	800.1 ms	100 %
Identifying wrong objects					100 %
Grasping a found object	7.80 s	0.56 s	6.90 s	10.12 s	99 %

Part Verification Results

- Parts used for verification



- Detection confidences

Object	Mean	Std	Min	Max
Correct object (“cross clamp”)	0.901	0.024	0.853	0.951
Similar cross clamp (pose 1)	0	0	0	0
Similar cross clamp (pose 2)	0.407	0.034	0.299	0.452
Small starter	0	0	0	0
Large starter	0.505	0.055	0.398	0.581
Smaller cross clamp	0	0	0	0

Different Lighting Conditions

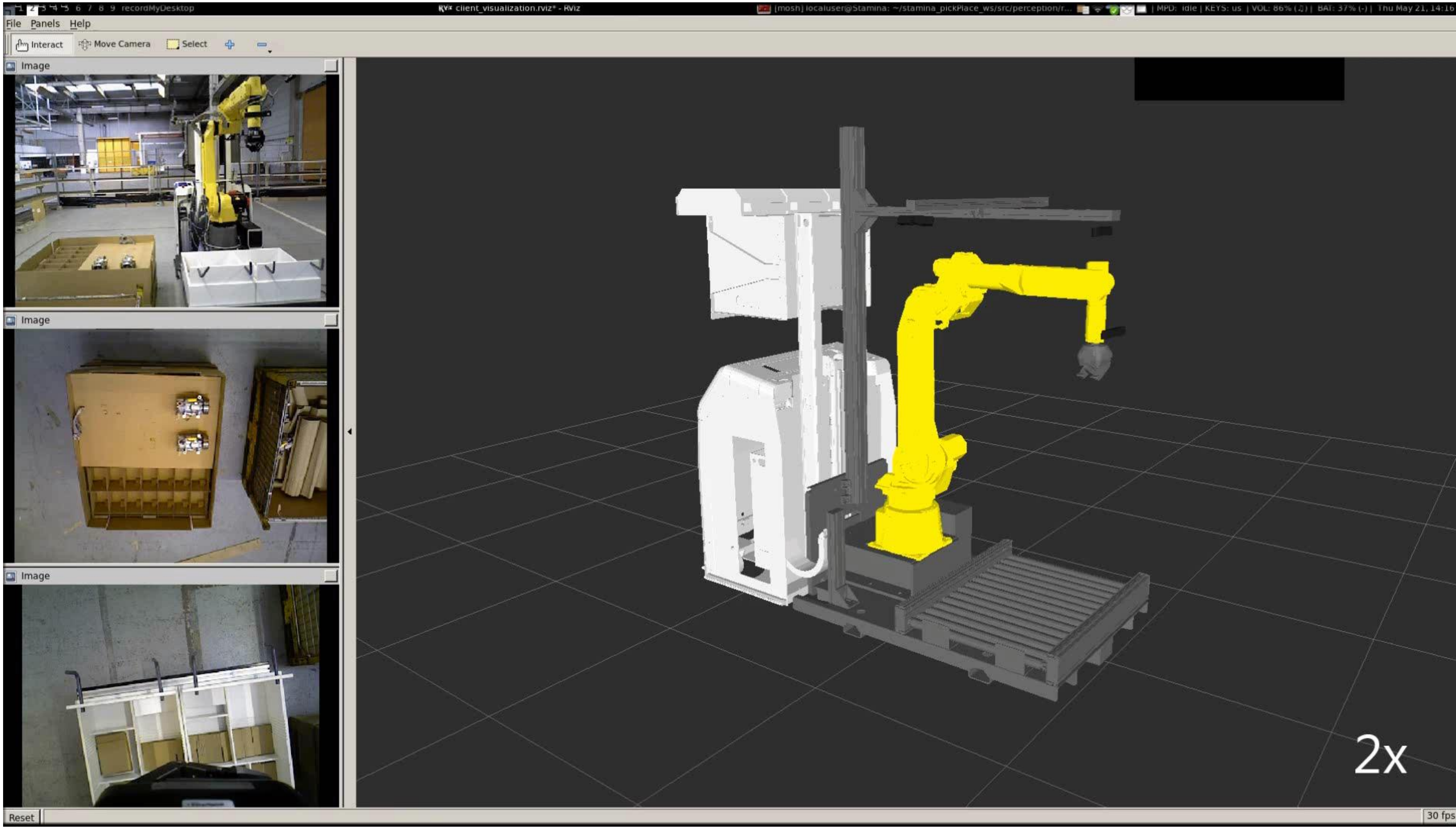
Artificial light and day light

Only daylight

Low light

- In all cases, the palette was successfully cleared.

Integrated Mobile Manipulation Robot



- 2nd Test Sprint, 05/2015

EuRoC Challenge 2: Shop-floor logistics and manipulation

- Addresses scenarios Logistics and Robotic Co-Workers
- Mobile manipulators as logistic carriers and dexterous manipulators
- Platform: Kuka miiwa
- Common benchmarking tasks
  - Pick and place, simple pre-assembly
  - Limited workspace, fast adaptation
- KittingBot project: NimbRo Logistics + PSA Peugeot Citroën
  - Freestyle: Depalletizing
  - Show-case: Sequencing
  - Pilot: Collaborative kitting

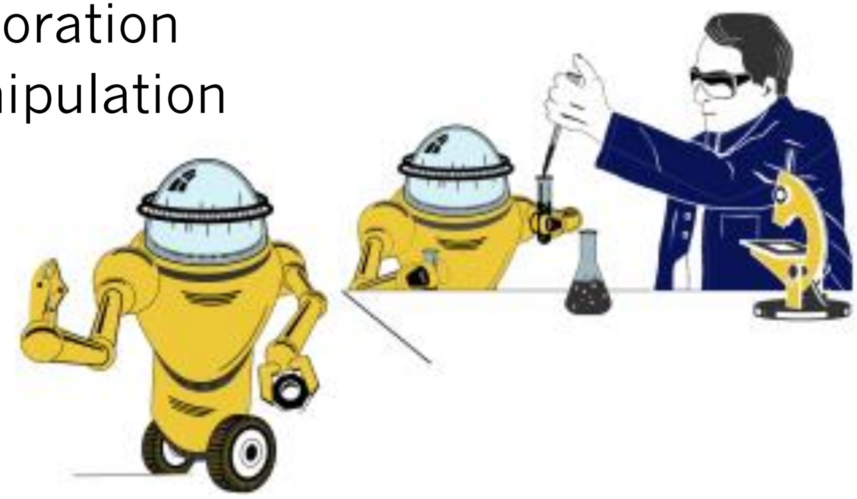

[EuRoC Challenge 2]


[Kuka miiwa]


[PSA Supermarket]

Conclusion

- Developed methods for environment perception, navigation, manipulation, and human-robot interaction in domestic service domain
- Methods applicable to industrial settings
- Challenges
  - Autonomy in semi-structured environments
  - Reliability of operation
  - Safety in human-robot collaboration
  - Speed of navigation and manipulation
- Need for further research
  - Robot construction
  - Perception
  - Behavior control
  - Learning



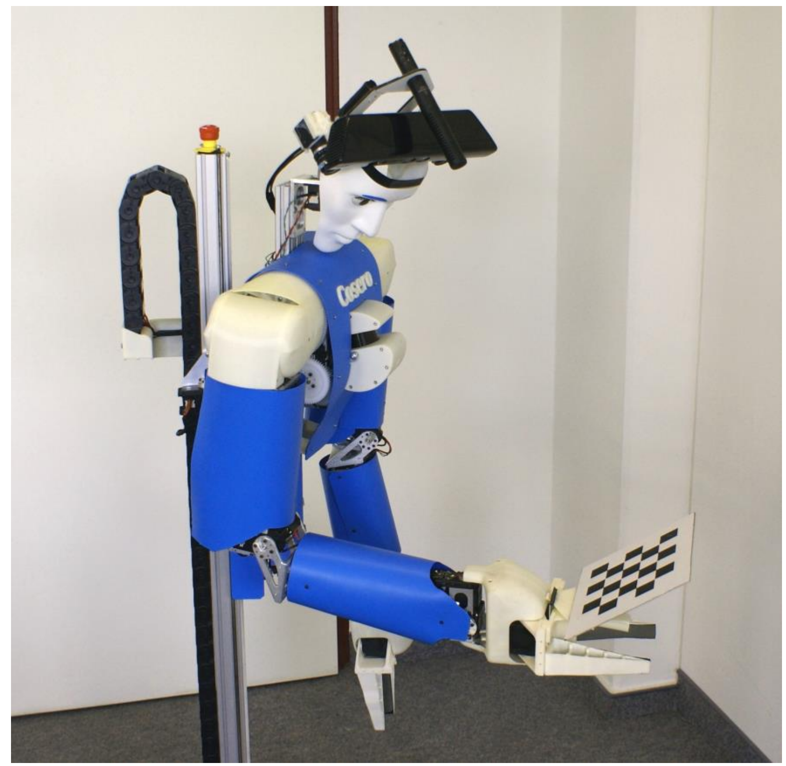
- Easier: mobile manipulators in smart environments

Thanks for your attention!

Questions?

Calibration of the Hand-Eye Kinematics

- Kinematic model
- Sample observations in various poses
- Estimate model parameters



[Hubert et al., Humanoids 2012]