Perception and Planning for Autonomous Mobile Robots

Sven Behnke

University of Bonn
Computer Science Institute VI
Autonomous Intelligent Systems

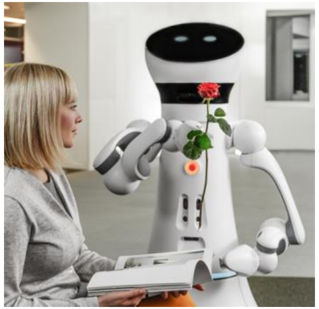
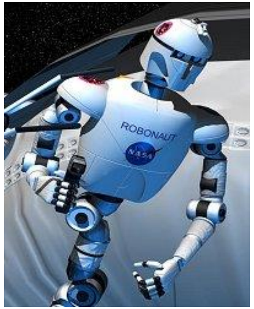
Many New Application Areas for Robots

- Self-driving cars
- Logistics
- Agriculture, mining
- Collaborative automation
- Personal assistance
- Space, search & rescue
- Healthcare
- Toys

Need more cognitive abilities!

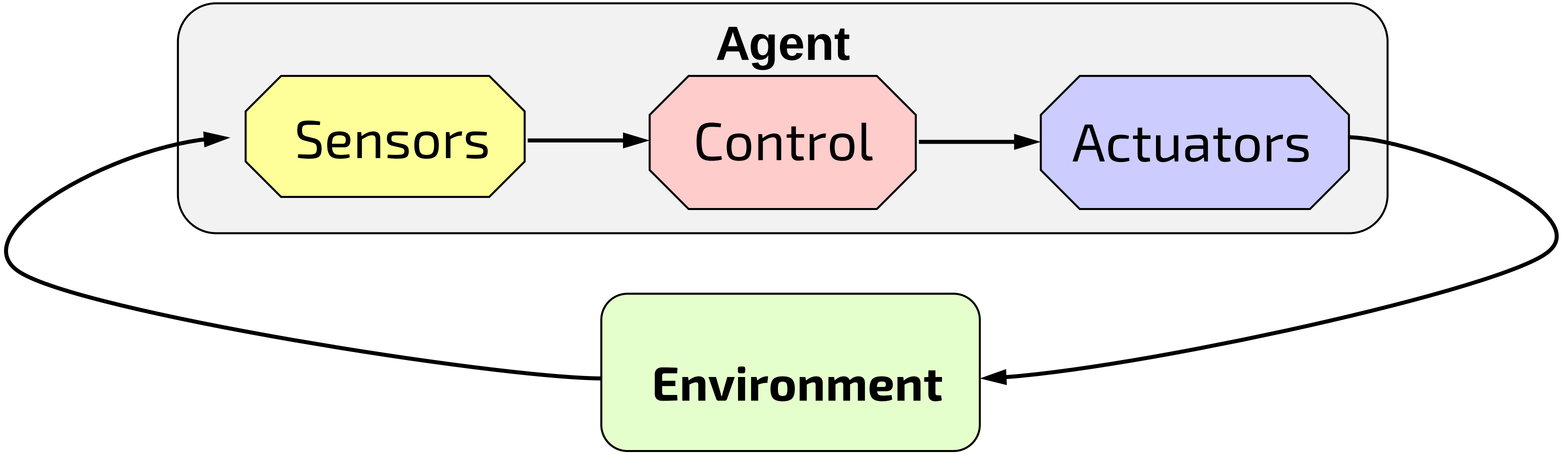





Sub-problems

- Environment perception
- Behavior planning
- Action generation



Some of our Cognitive Robots

- Equipped with numerous sensors and actuators
- Complex demonstration scenarios

Soccer

Domestic service

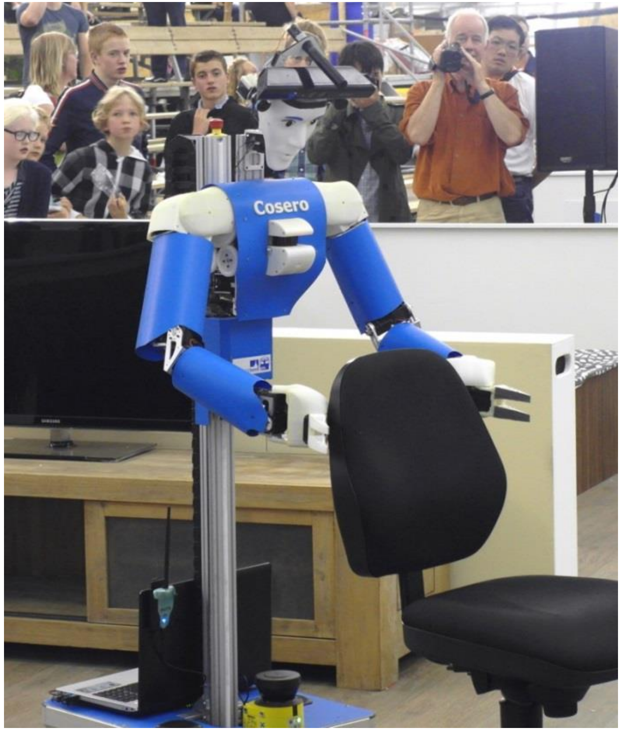
Mobile manipulation

Bin picking

Aerial inspection

Our Domestic Service Robots

Dynamaid



Cosero

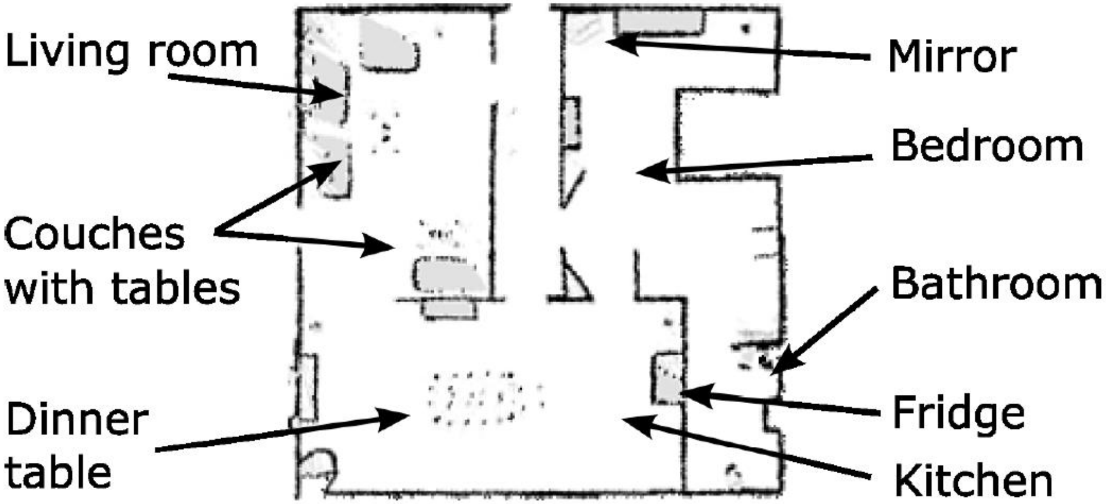
- Size: 100-180 cm, weight: 30-35 kg
- 36 articulated joints
- PC, laser scanners, Kinect, microphone, …

[Stückler et al.: Frontiers in Robotics and AI 2016]

Cognitive Service Robot Cosero

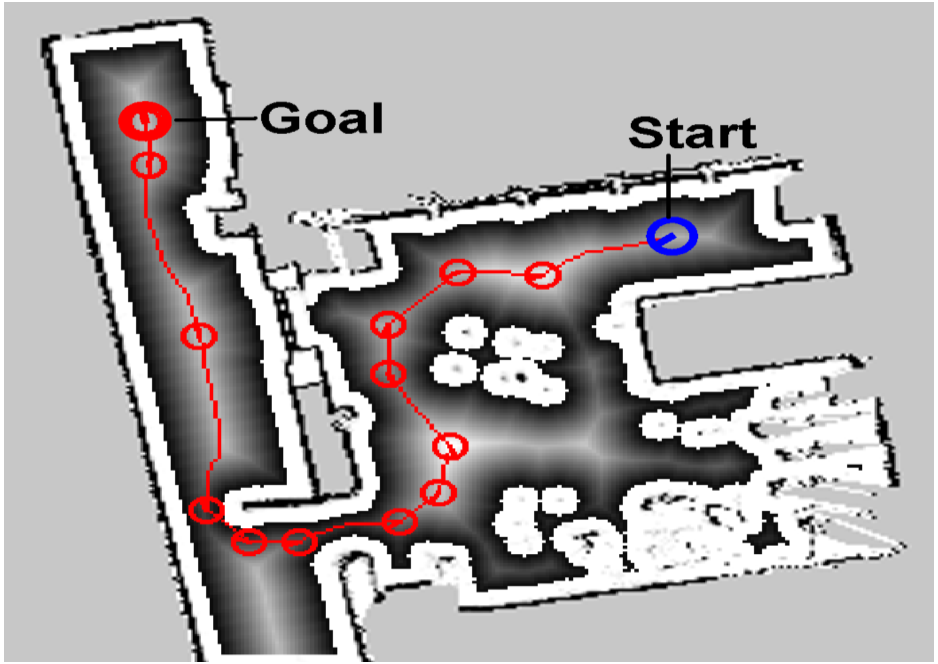
UNIVERSITÄT BONN AIS

Mapping the Environment



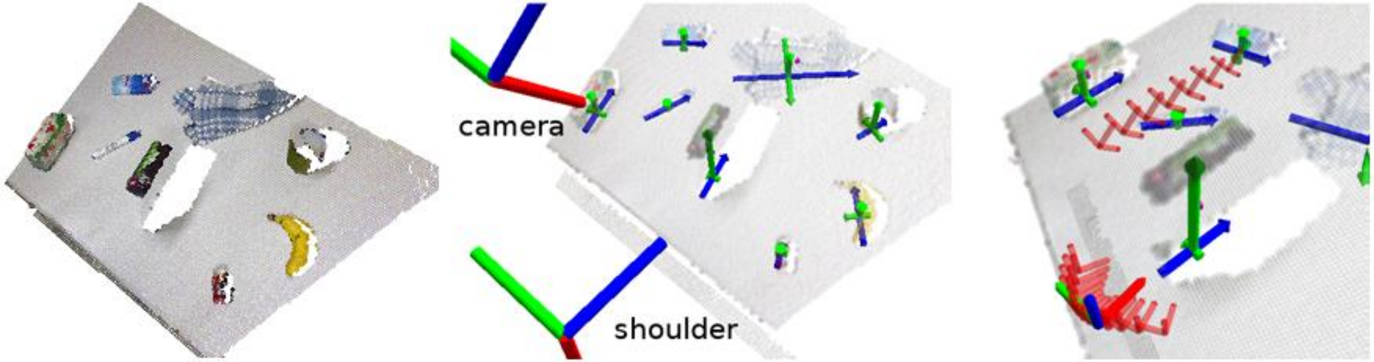
Path Planning

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

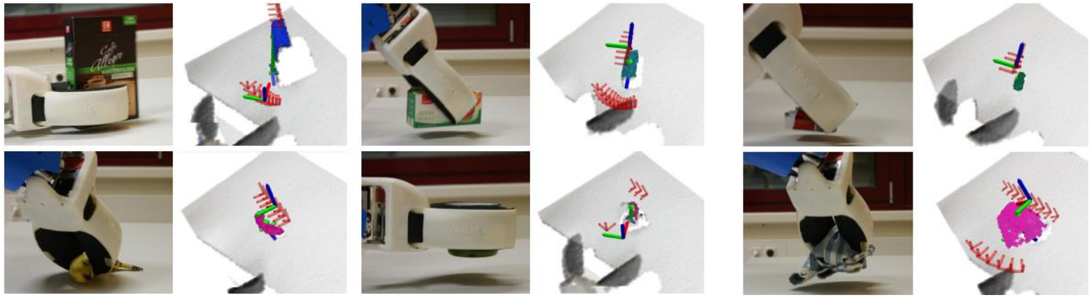


Object Perception and Grasp Planning

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

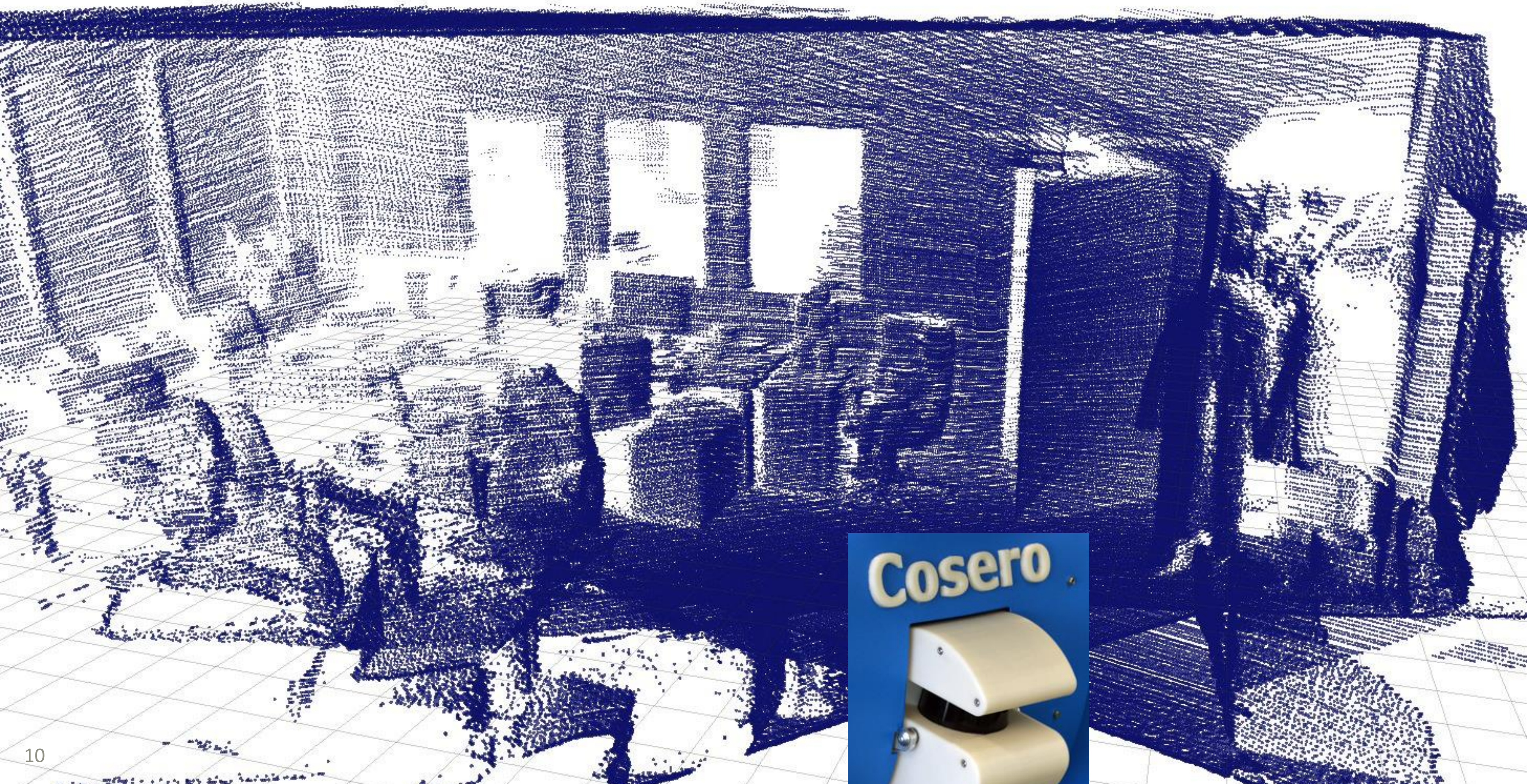


- Flexible grasping of many unknown objects

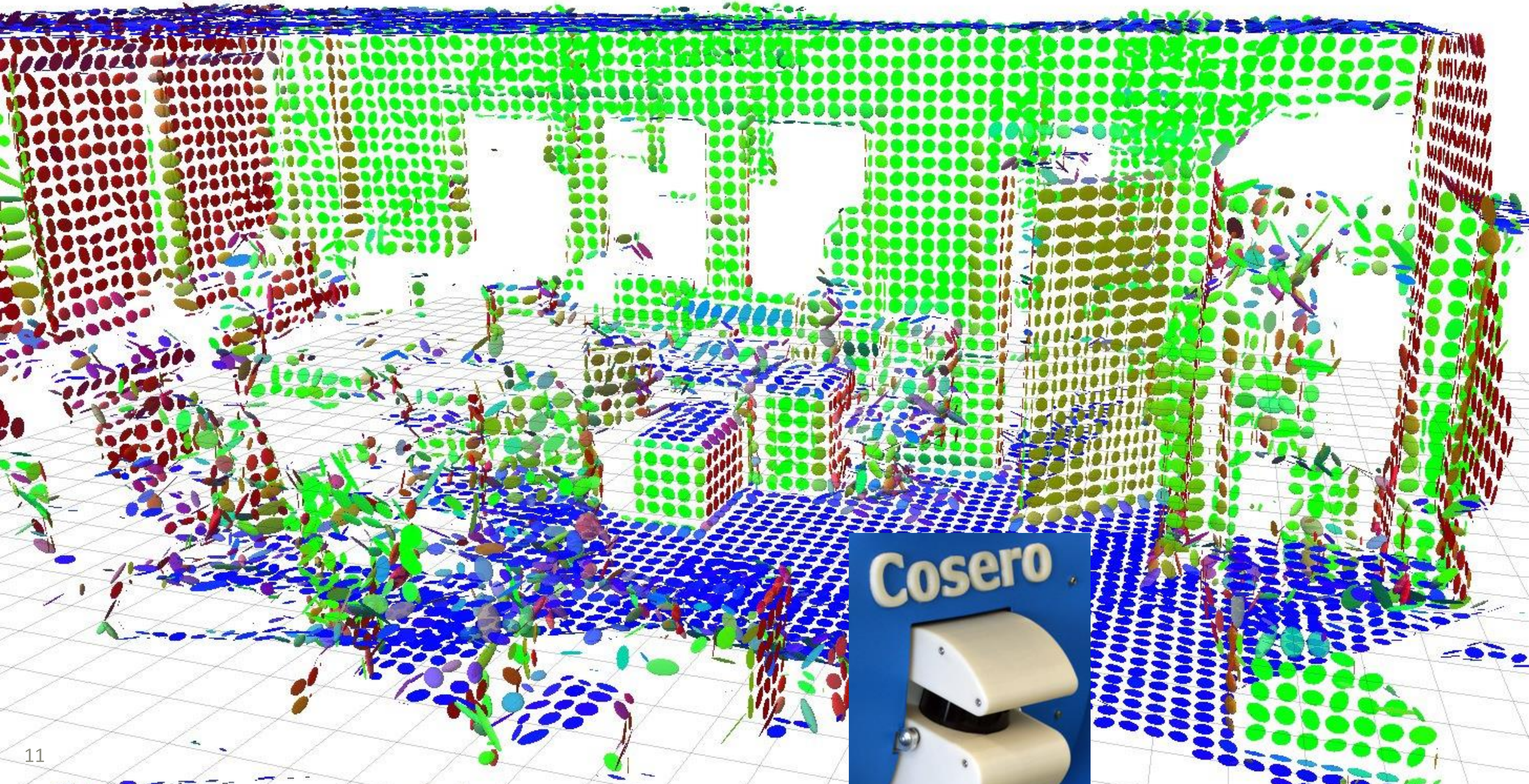


[Stückler et al, Robotics and Autonomous Systems, 2013]

3D-Mapping with Surfels

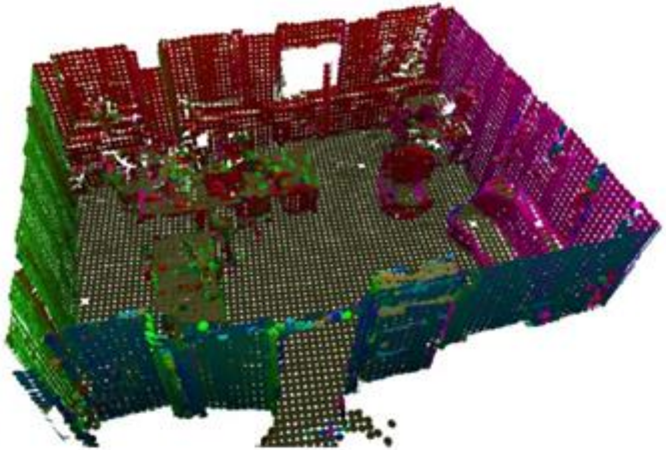
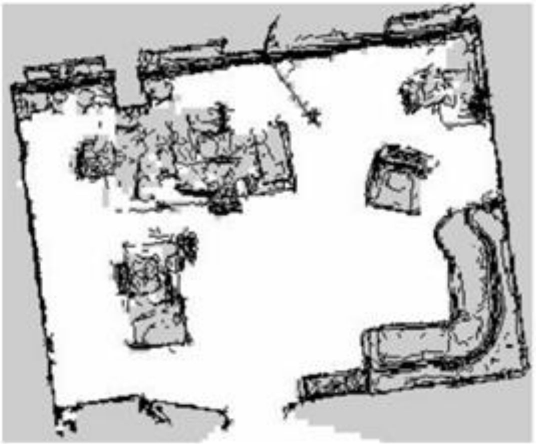


3D-Mapping with Surfels



3D-Mapping and Localization

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

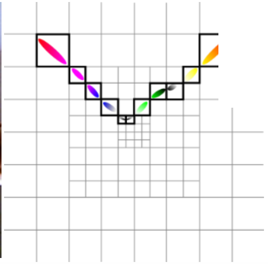




[Kläß, Stückler, Behnke: Robotik 2012]

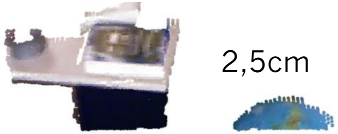
3D Mapping by RGB-D SLAM

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

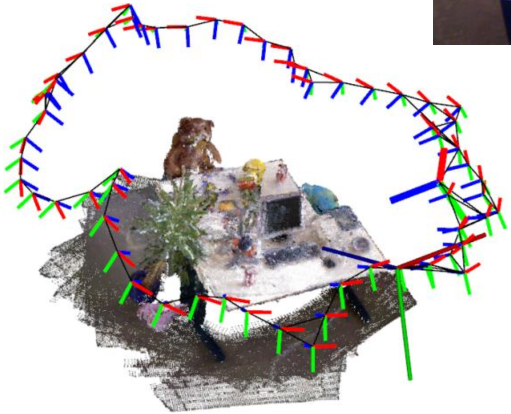
- Modelling of shape and color distributions in voxels
- Local multiresolution
- Efficient registration
  of views on CPU
- Global
  optimization

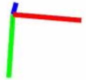


5cm

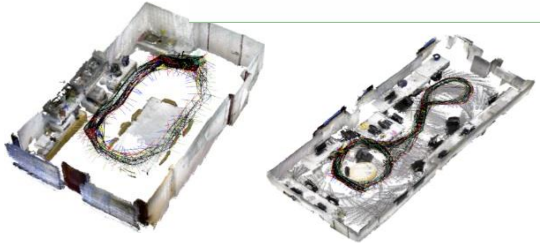


2,5cm





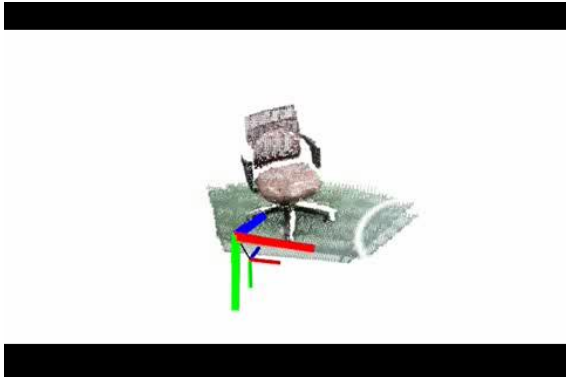
- Multi-camera SLAM



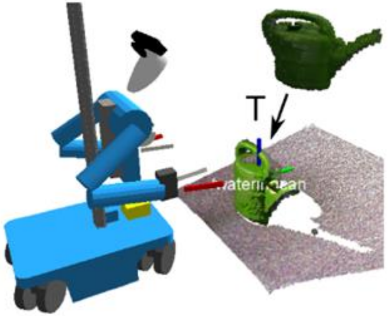
[Stoucken]

Learning and Tracking Object Models

- Modeling of objects by RGB-D-SLAM

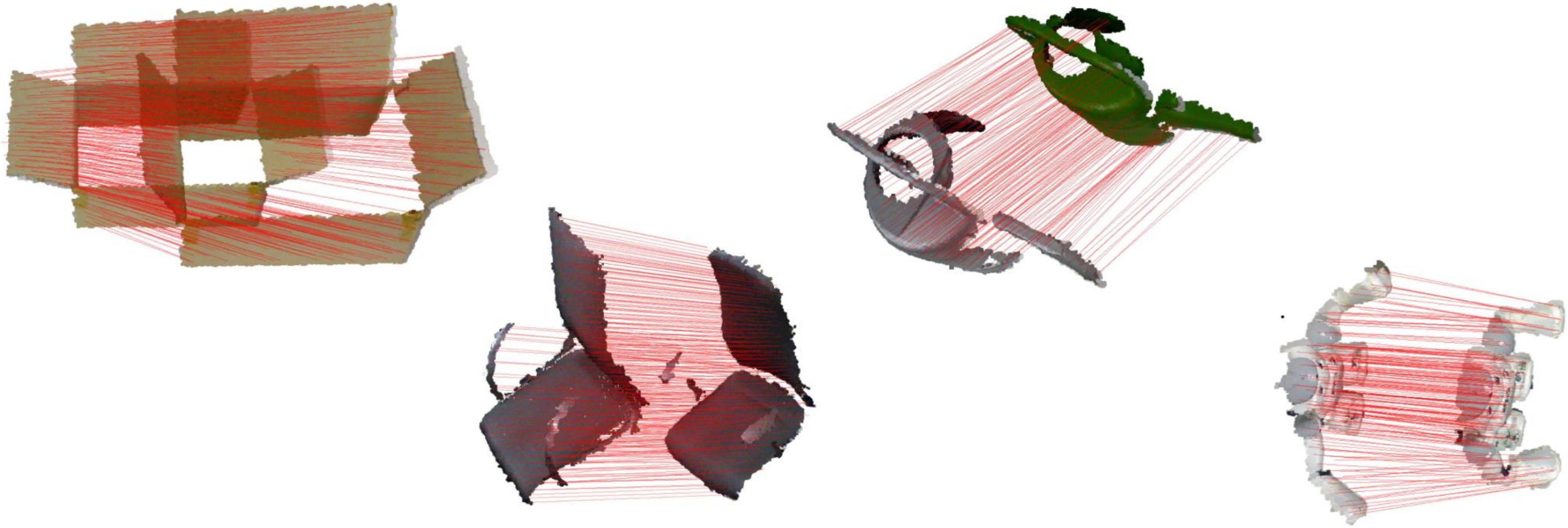



- Real-time registration with current RGB-D frame

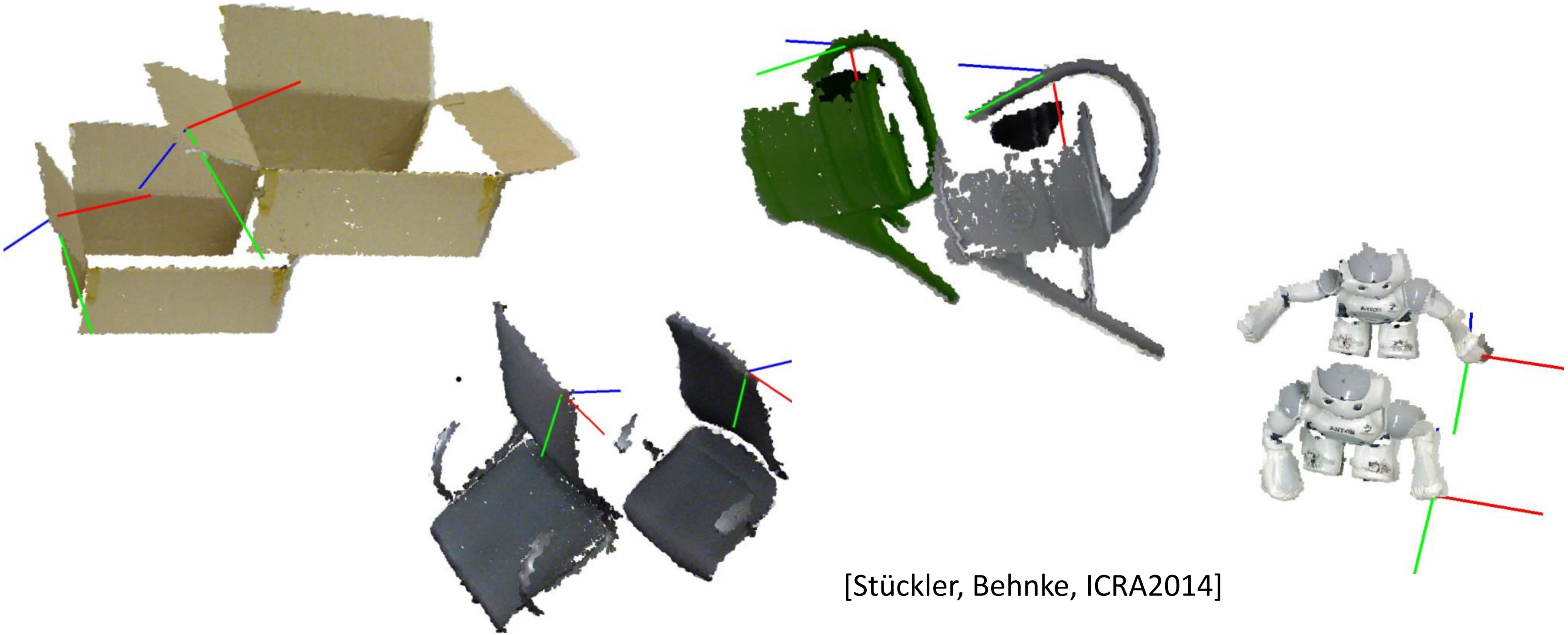
Deformable RGB-D-Registration

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



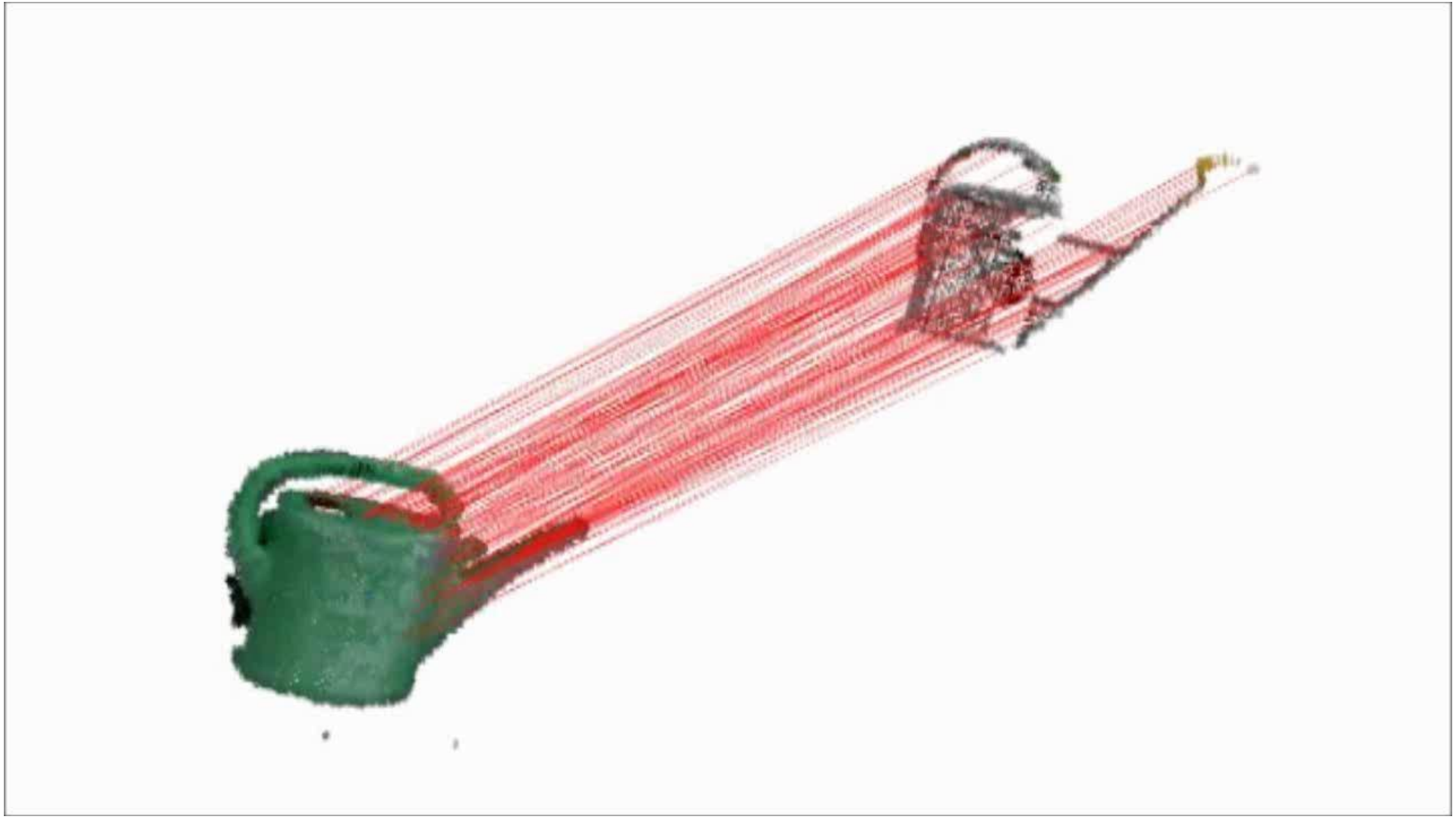
Transformation of Poses on Object

- Derived from the deformation field



[Stückler, Behnke, ICRA2014]

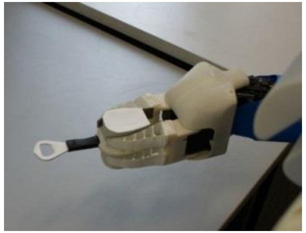
Grasp & Motion Skill Transfer

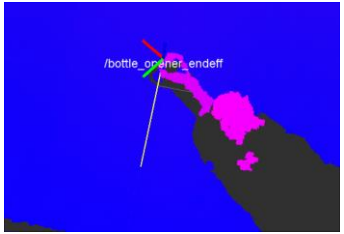


[Stückler, Behnke, ICRA2014]

Tool use: Bottle Opener

- Tool tip perception



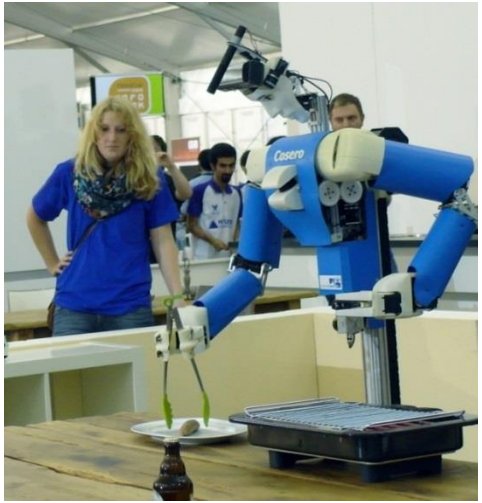


- Extension of arm kinematics
- Perception of crown cap
- Motion adaptation

[Stückler, Behnke, Humanoids 2014]

Picking Sausage, Bimanual Transport

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

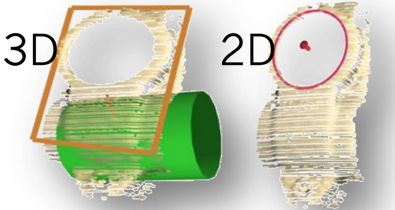


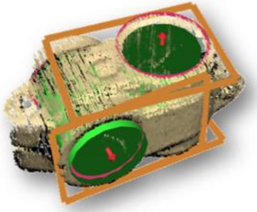
- Our team NimbRo won the RoboCup@Home League in three consecutive years

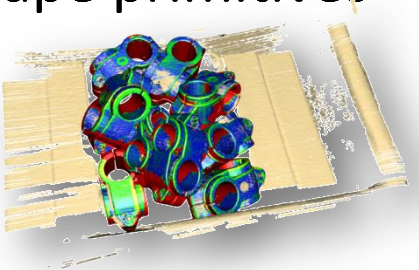
Bin Picking

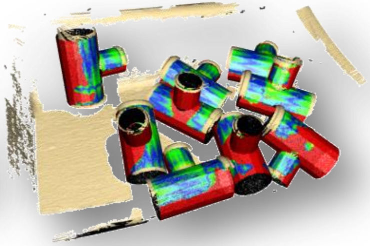
- Known objects in transport box

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

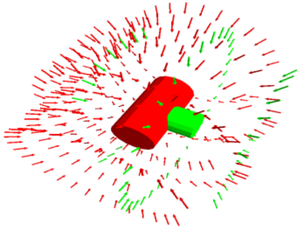


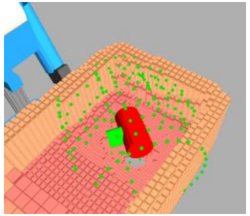


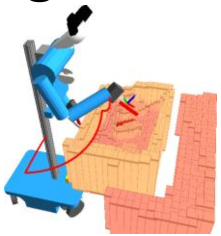




- Grasp and motion planning





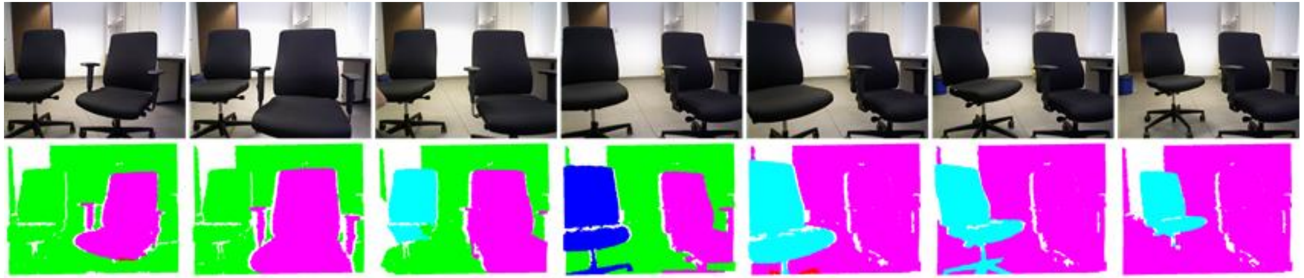


Offline Online

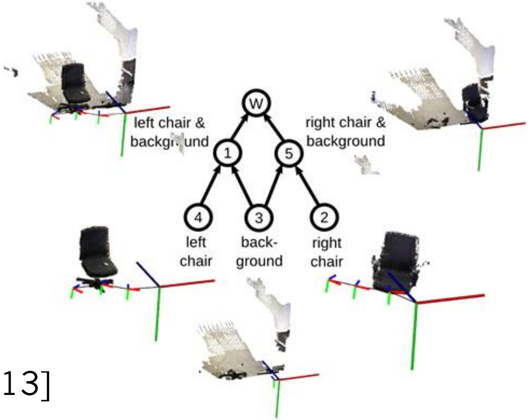
[Nieuwenhuisen et al.: ICRA 2013]

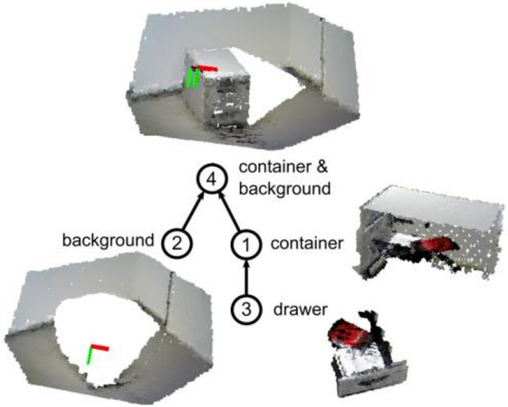
Hierarchical Object Discovery trough Motion Segmentation

- Simultaneous object modeling and motion segmentation



- Inference of a segment hierarchy

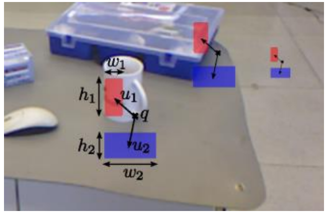


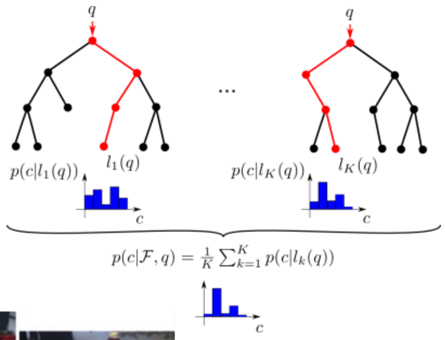


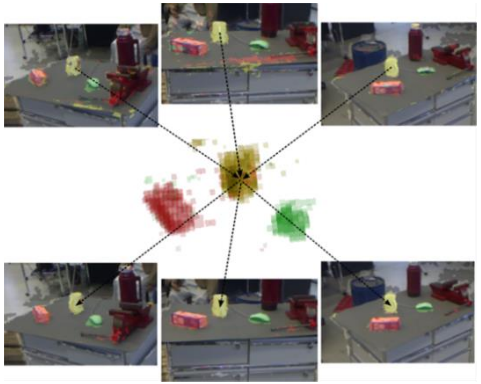
[Stückler, Behnke: IJCAI 2013]

Semantic Mapping

- Pixel-wise classification of RGB-D images by random forests
- Compare color / depth of regions
- Size normalization
- 3D fusion through RGB-D SLAM
- Evaluation on NYU depth v2







[Stückler, Biresev, Behnke: IROS 2012]

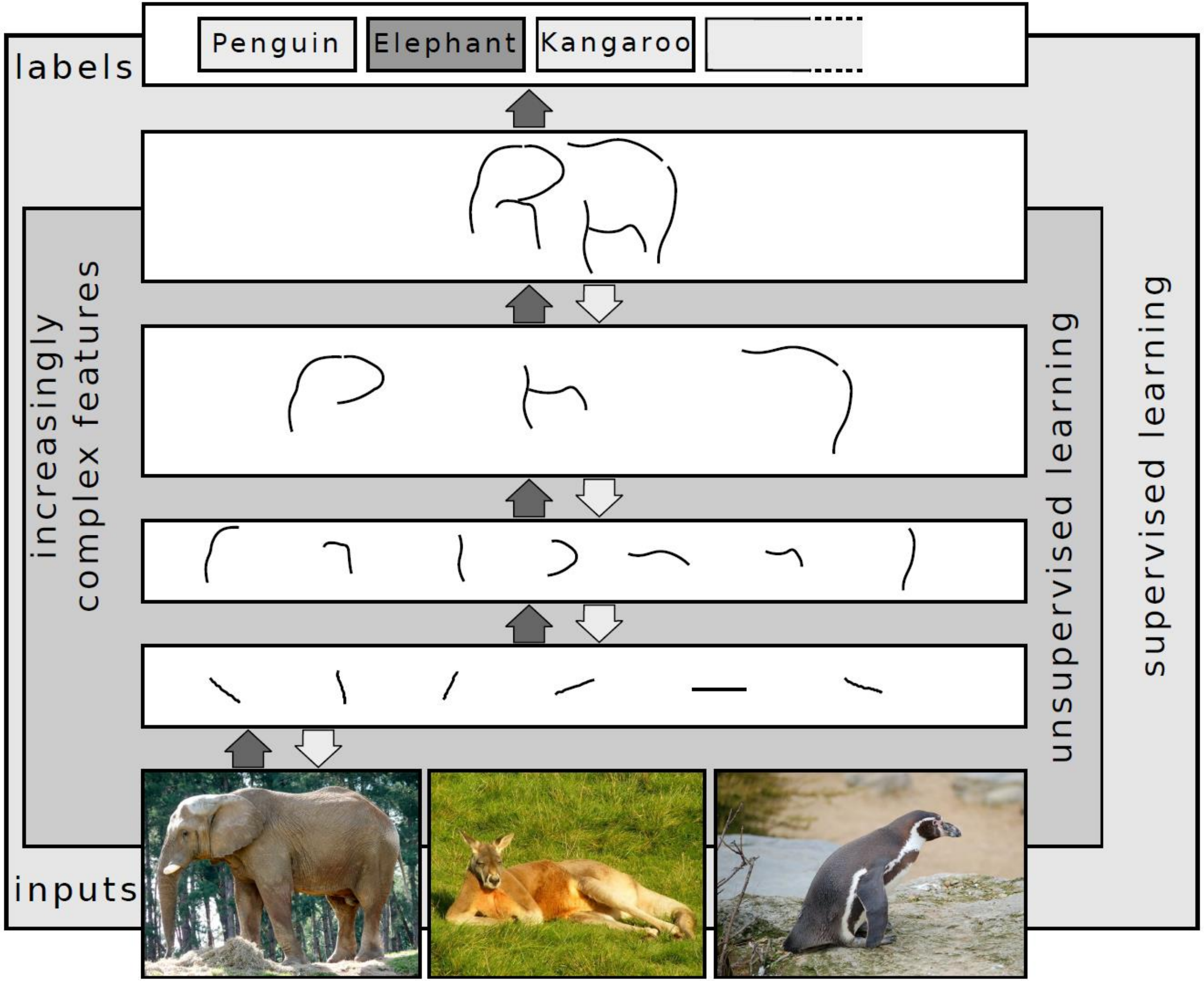
Ground truth

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**	

Deep Learning

- Learning layered represen- tations

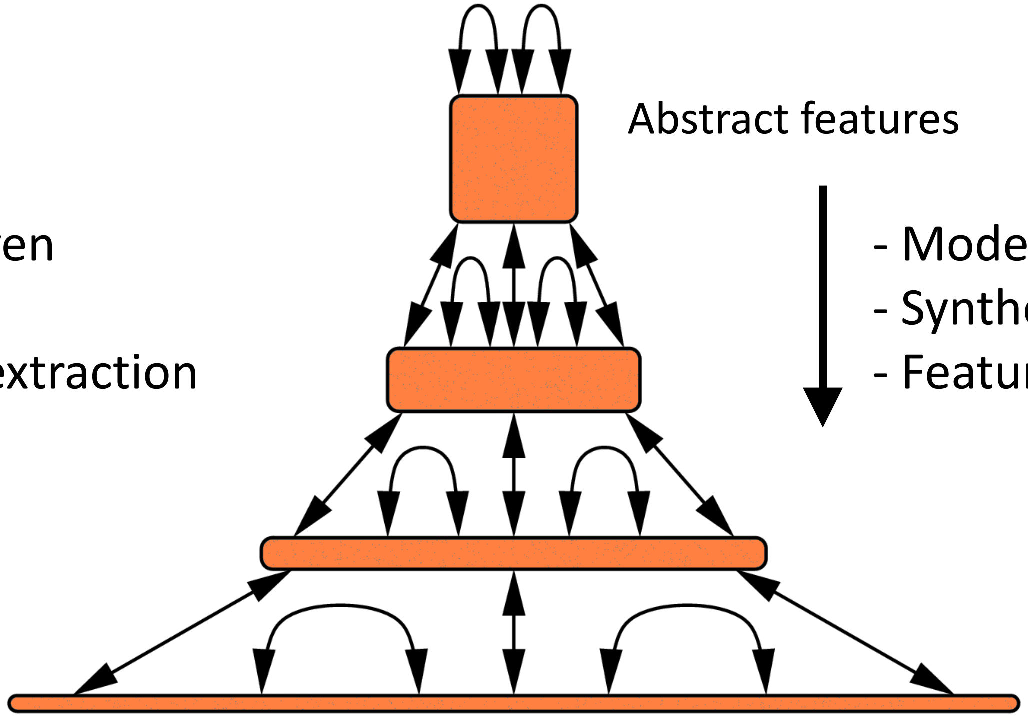


[Schulz; Behnke, KI 2012]

Neural Abstraction Pyramid

[Behnke, Rojas, IJCNN 1998]
[Behnke, LNCS 2766, 2003]

Abstract features



- Data-driven
- Analysis
- Feature extraction

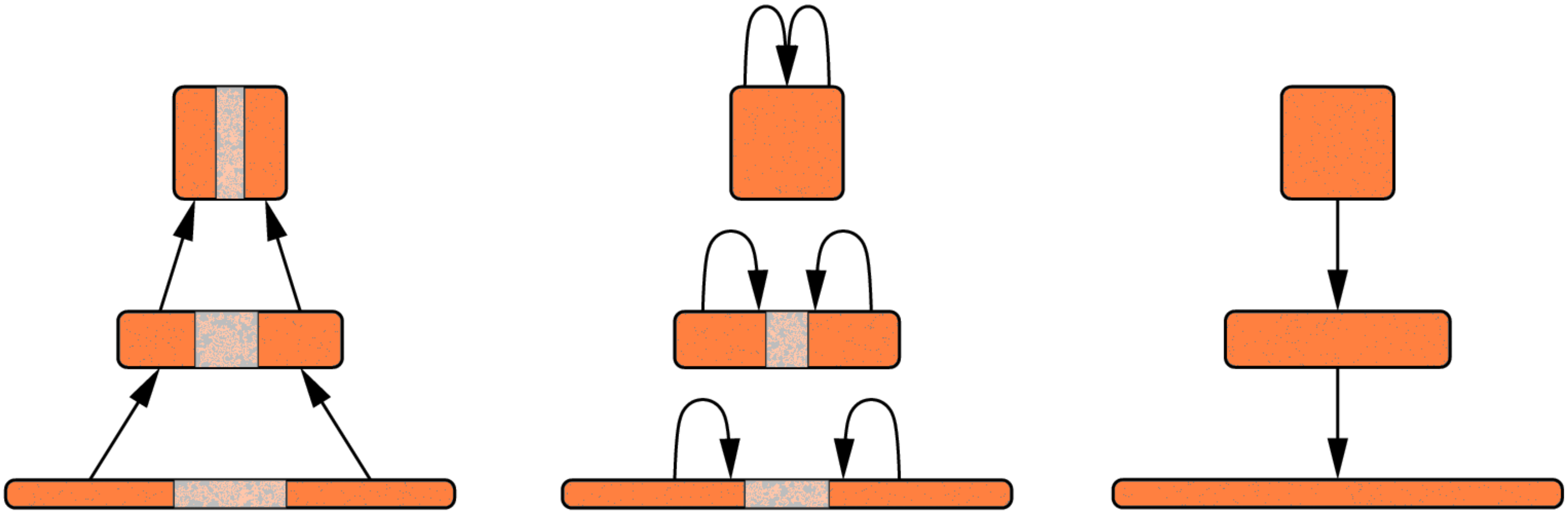
- Model-driven
- Synthesis
- Feature expansion

Signals

- Grouping - Competition - Completion

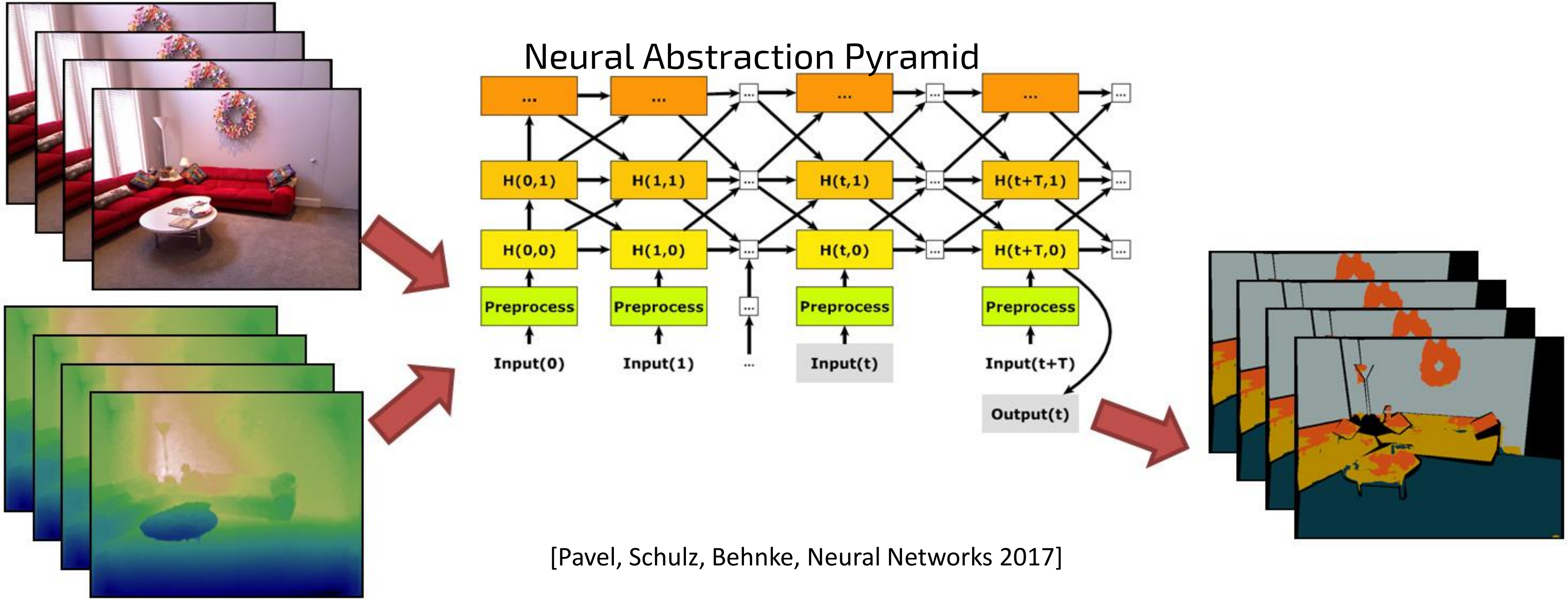
Iterative Image Interpretation

- Interpret most obvious parts first
- Use partial interpretation as context to resolve local ambiguities



Neural Abstraction Pyramid for RGB-D Video Object-class Segmentation

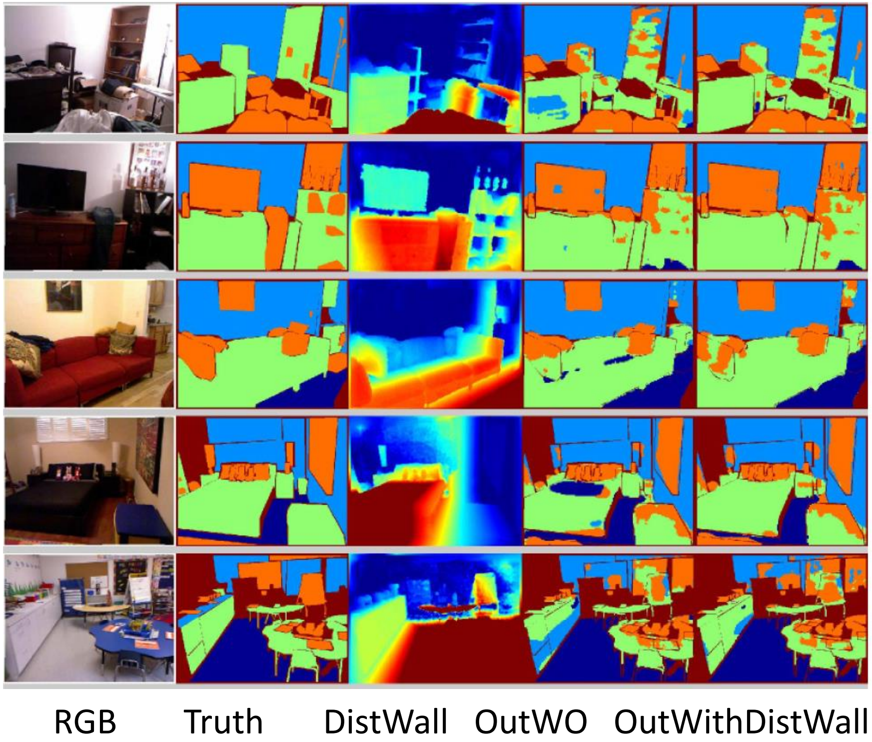
- Recursive computation is efficient for temporal integration



[Pavel, Schulz, Behnke, Neural Networks 2017]

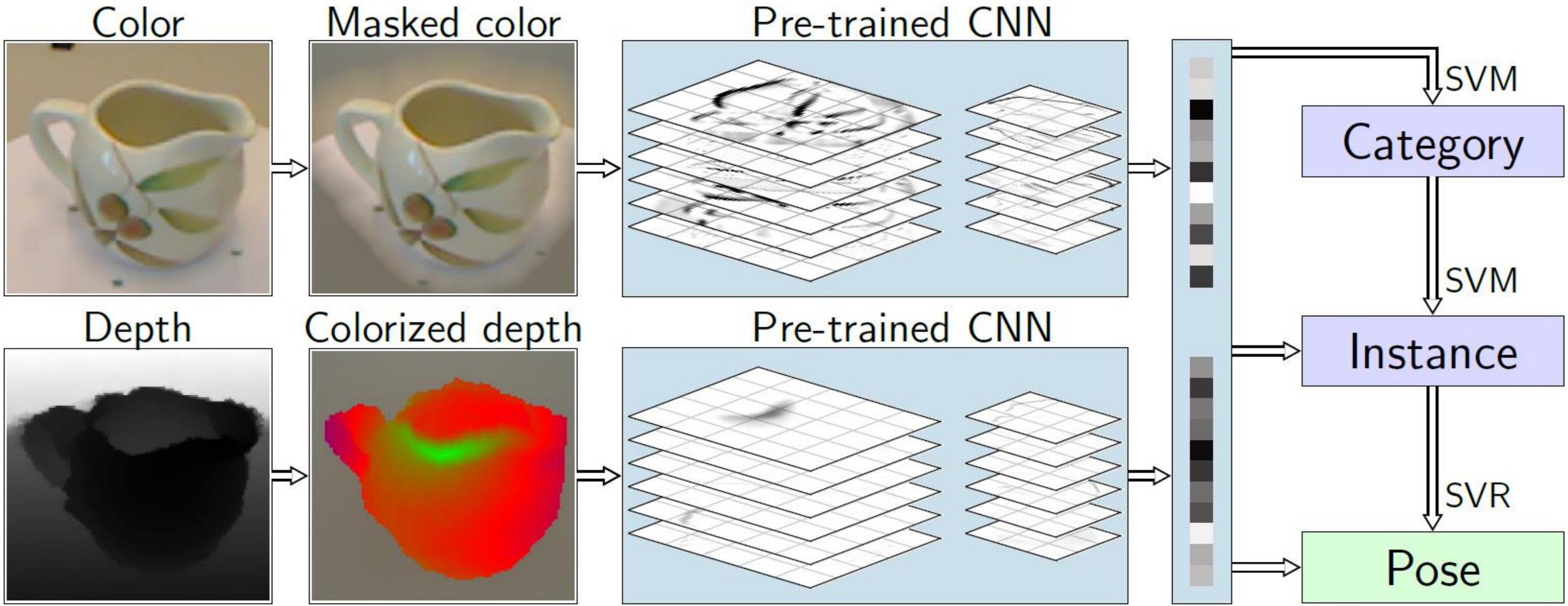
Geometric and Semantic Features for RGB-D Object-class Segmentation

- New **geometric** feature: distance from wall
- **Semantic** features pretrained from ImageNet
- Both help significantly



[Husain et al. RA-L 2017]

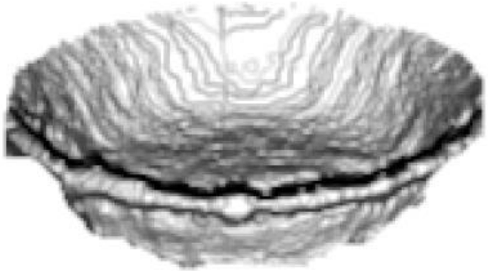
RGB-D Object Recognition and Pose Estimation

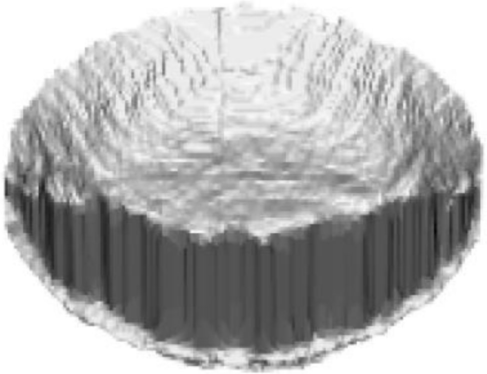


[Schwarz, Schulz, Behnke, ICRA2015]

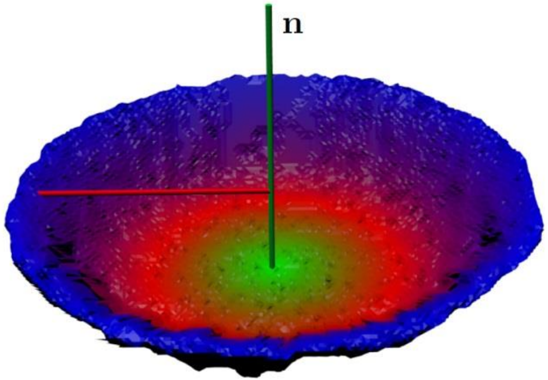
Canonical View, Colorization

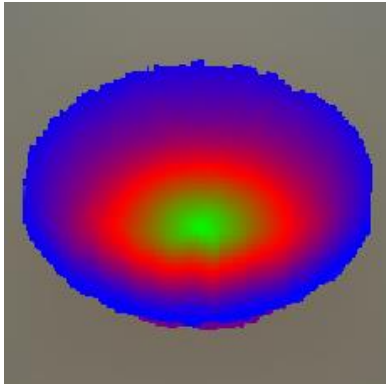
- Objects viewed from different elevation
- Render canonical view





- Colorization based on distance from center vertical

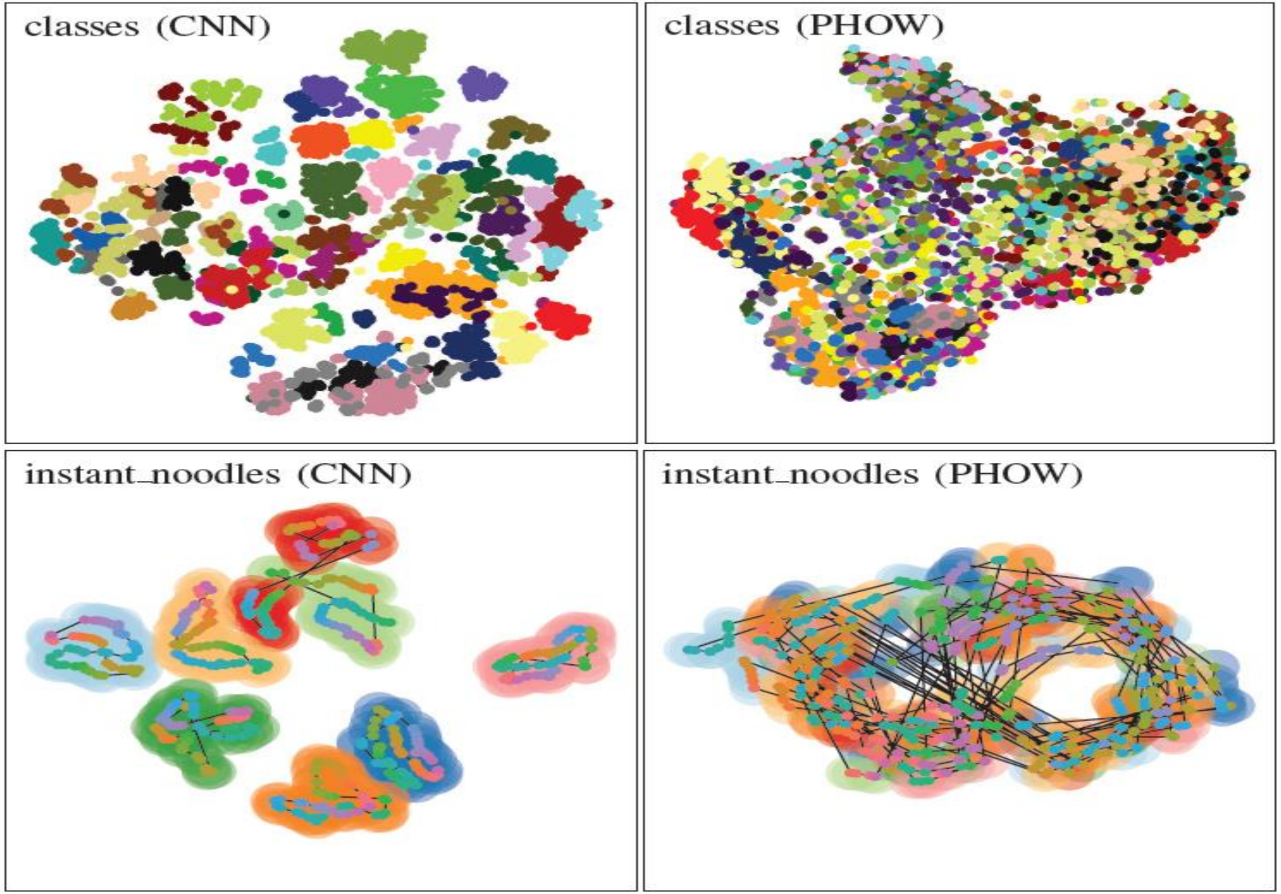




[Schwarz, Schulz, Behnke, ICRA2015]

Pretrained Features Disentangle Data

- t-SNE embedding



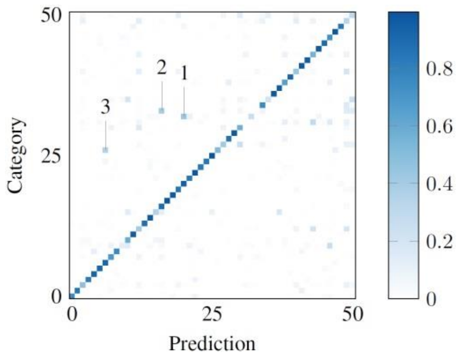
[Schwarz, Schulz, Behnke ICRA2015]

Recognition Accuracy

- Improved both category and instance recognition

Method	Category Accuracy (%) RGB	Category Accuracy (%) RGB-D	Instance Accuracy (%) RGB	Instance Accuracy (%) RGB-D
Lai *et al.* [1]	74.3 ± 3.3	81.9 ± 2.8	59.3	73.9
Bo *et al.* [2]	82.4 ± 3.1	87.5 ± 2.9	**92.1**	92.8
PHOW[3]	80.2 ± 1.8	—	62.8	—
Ours	$\mathbf{83.1 \pm 2.0}$	88.3 ± 1.5	92.0	**94.1**
Ours	$\mathbf{83.1 \pm 2.0}$	$\mathbf{89.4 \pm 1.3}$	92.0	**94.1**

- Confusion:



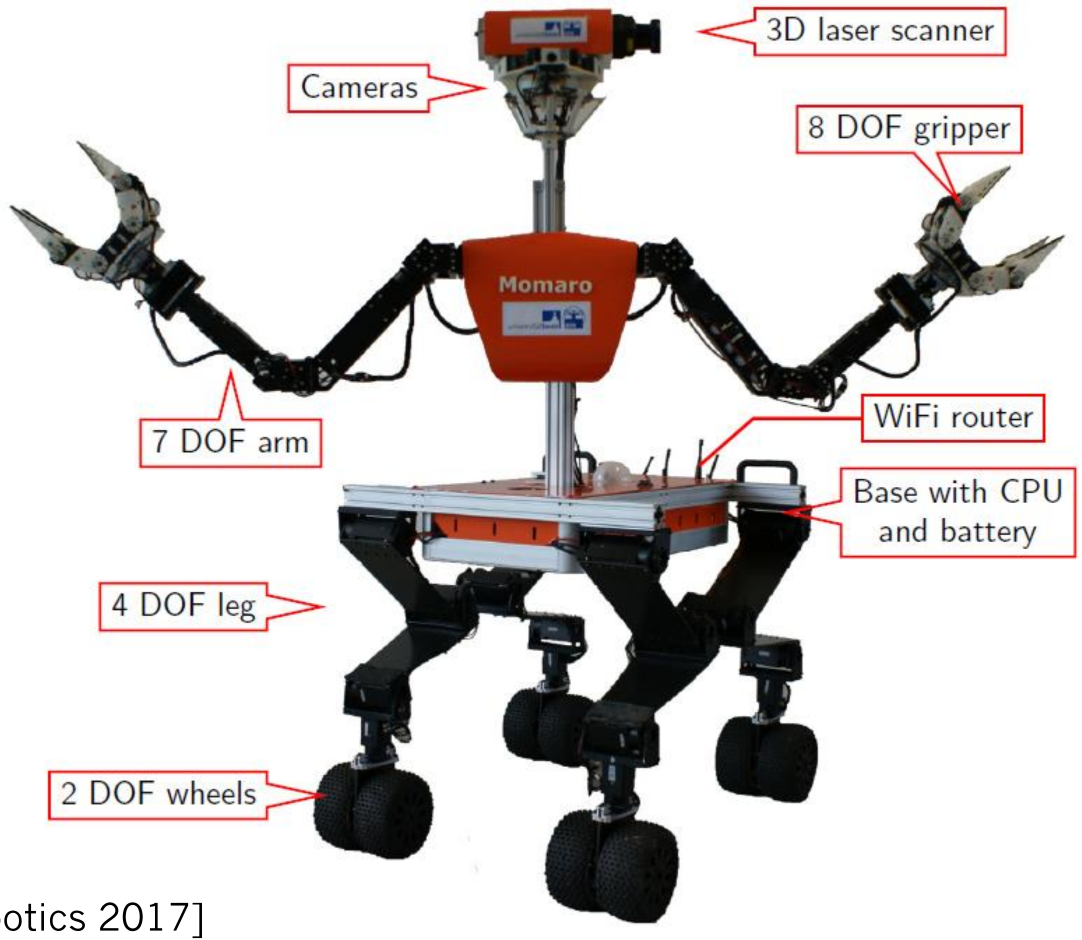
1: pitcher / coffe mug

2: peach / sponge

[Schwarz, Schulz, Behnke, ICRA2015]

Mobile Manipulation Robot Momaro

- Four compliant legs ending in pairs of steerable wheels
- Anthropomorphic upper body
- Sensor head
  - 3D laser scanner
  - IMU, cameras



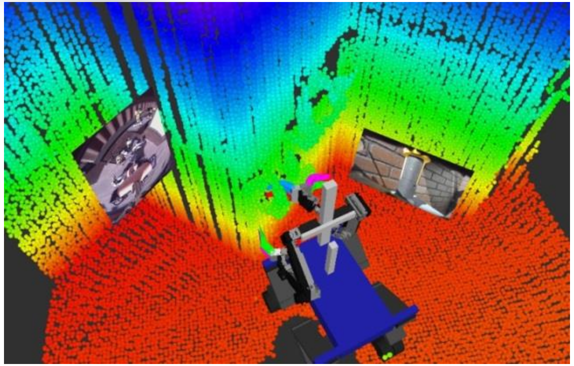
[Schwarz et al. Journal of Field Robotics 2017]

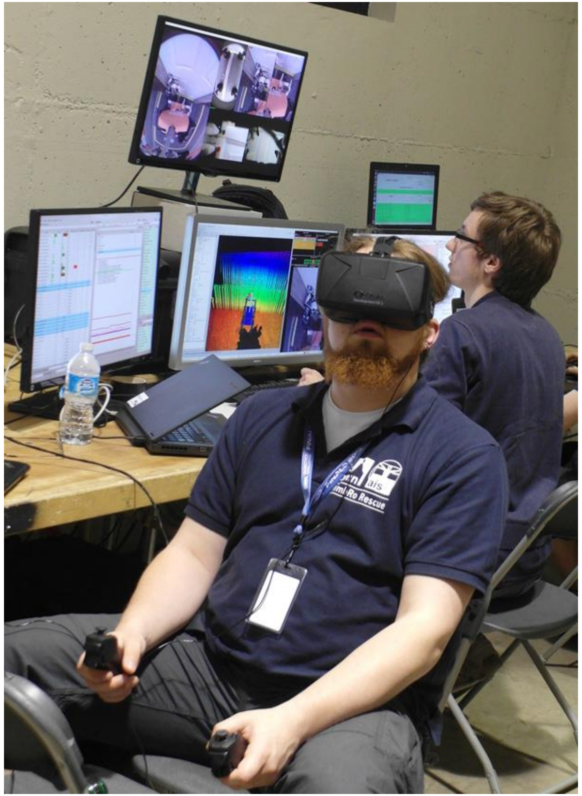
FAIRPLEX
FAIRPLEX
FAIRPLEX
POLARIS
23:15:03 05/06/2015 UTC
4x


CAUTION
POLARIS
23:16:59 05/06/2015 UTC
4x

Manipulation Operator Interface

- 3D head-mounted display
- 3D environment model + images
- 6D magnetic tracker





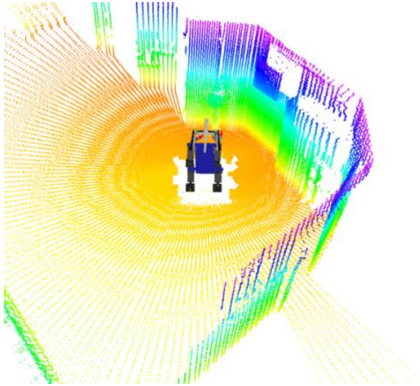
[Rodehutskors et al., Humanoids 2015]


FAIRPLEX
Momaro
RANGER
XP
23:20:32 05/06/2015 UTC
4x

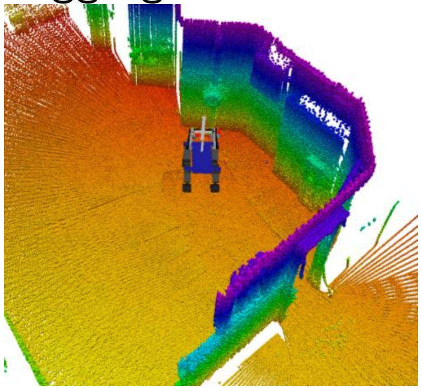
Local Multiresolution Surfel Map

- Registration and aggregation of 3D laser scans
- Local multi-resolution grid
- Surfel in grid cells

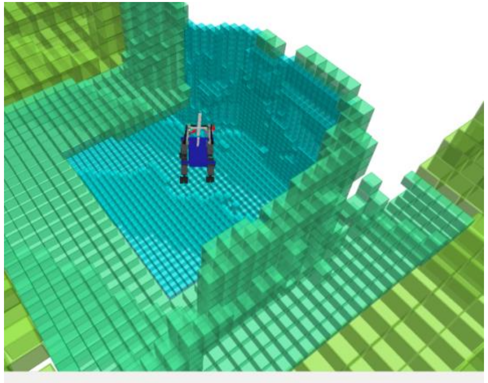
3D scan



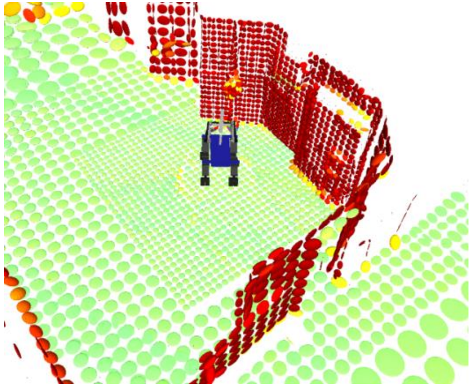
Aggregated scans



Multiresolution grid



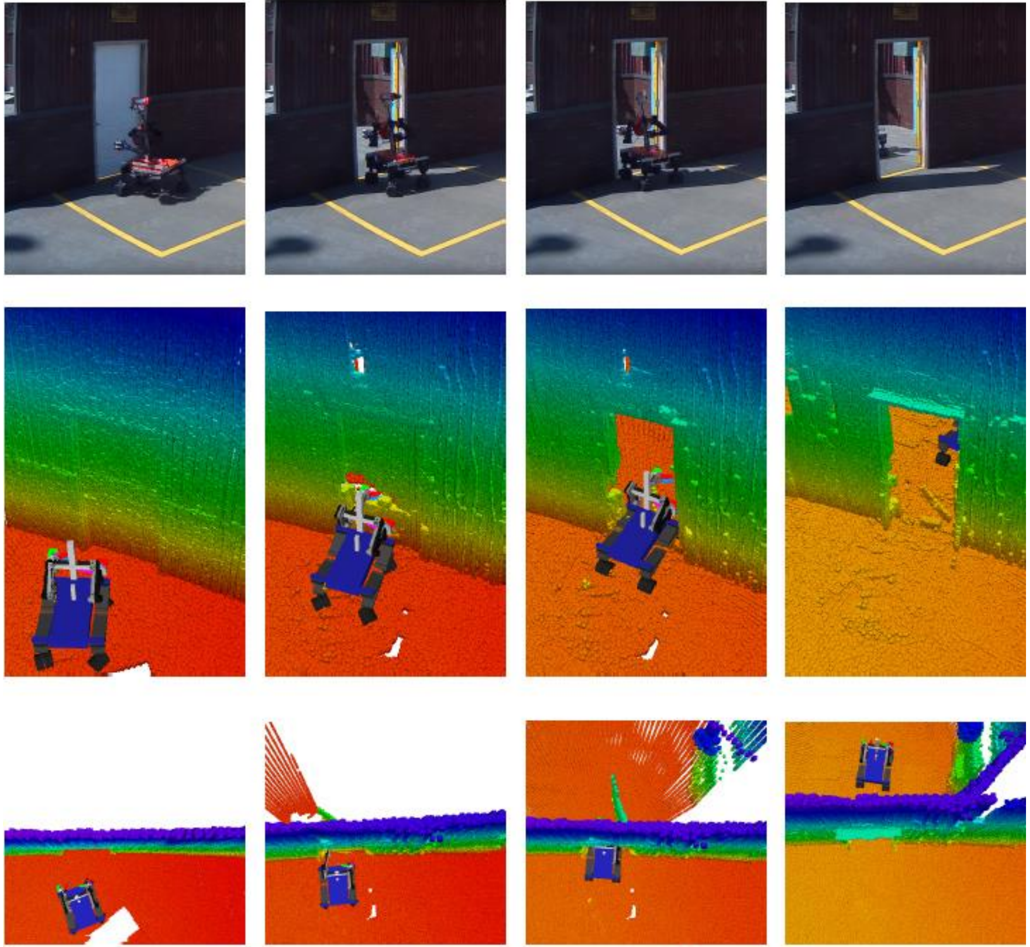
Surfels



[Droeschel et al., Robotics and Autonomous Systems 2017]

Filtering Dynamic Objects

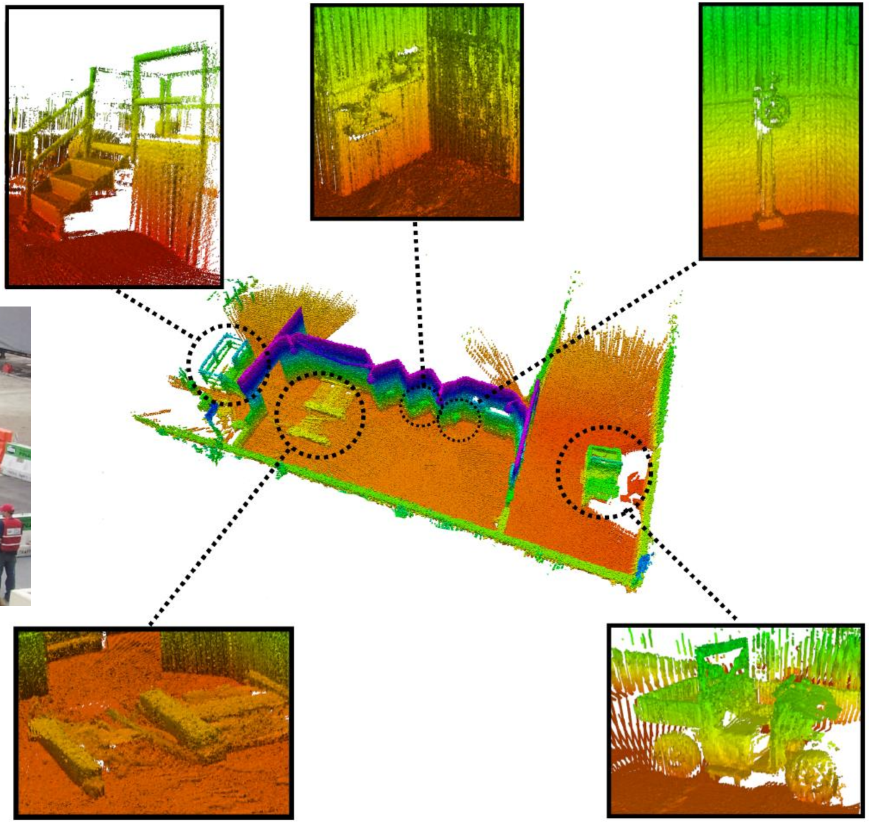
- Maintain occupancy in each cell
- Remove measurements of empty cells



1 scan (5 s) | 2 scans (10 s) | 5 scans (25 s)

Allocentric 3D Mapping

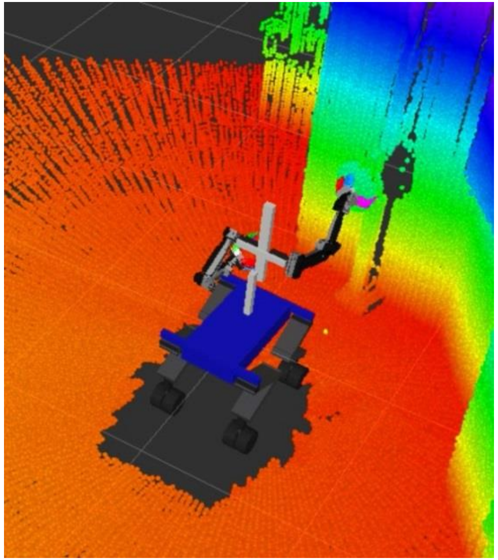
- Registration of egocentric maps by graph optimization



[Droeschel et al., Robotics and Autonomous Systems 2017]

Valve Turning Interface

- Align wheel model with 3D points using interactive marker



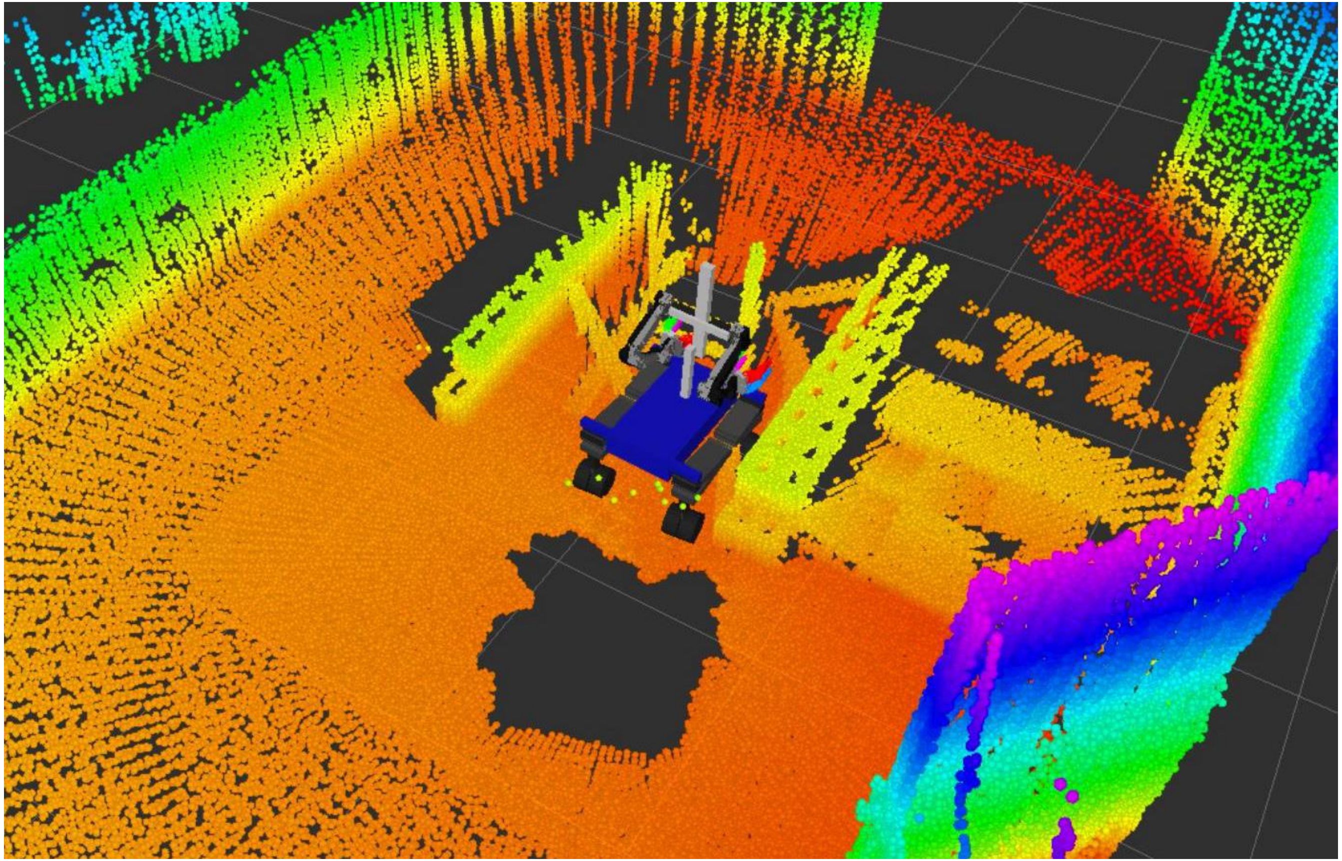
[Schwarz et al. Journal of Field Robotics 2017]

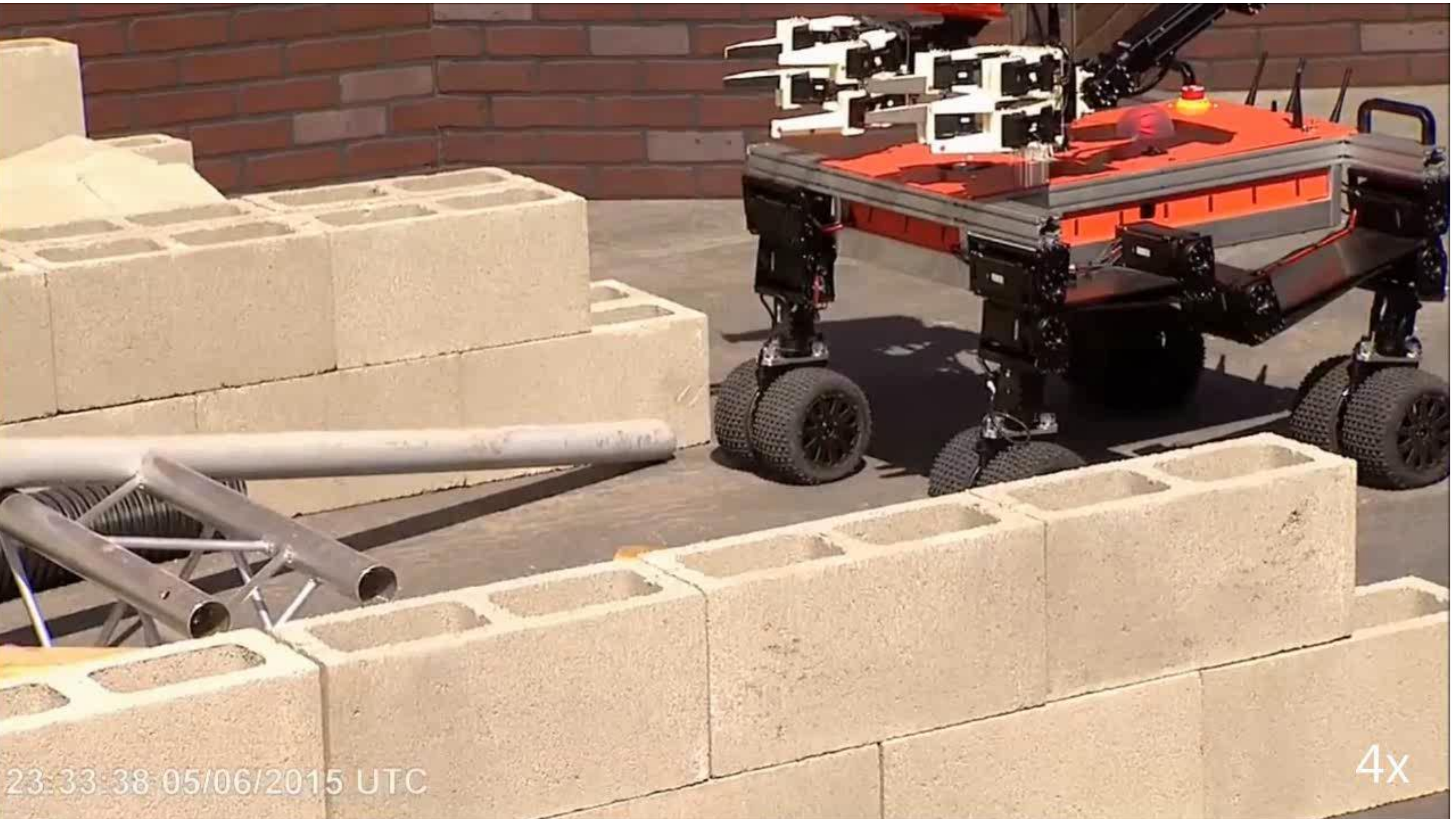

23:25:56 05/06/2015 UTC
4x


23:28:21 05/06/2015 UTC
4x


4X
02:23:20 07/06/2015 UTC

Debris Tasks




23:33:38 05/06/2015 UTC
4x

CHALLENGE
DARPA
23:36:46 05/06/2015 UTC
4x

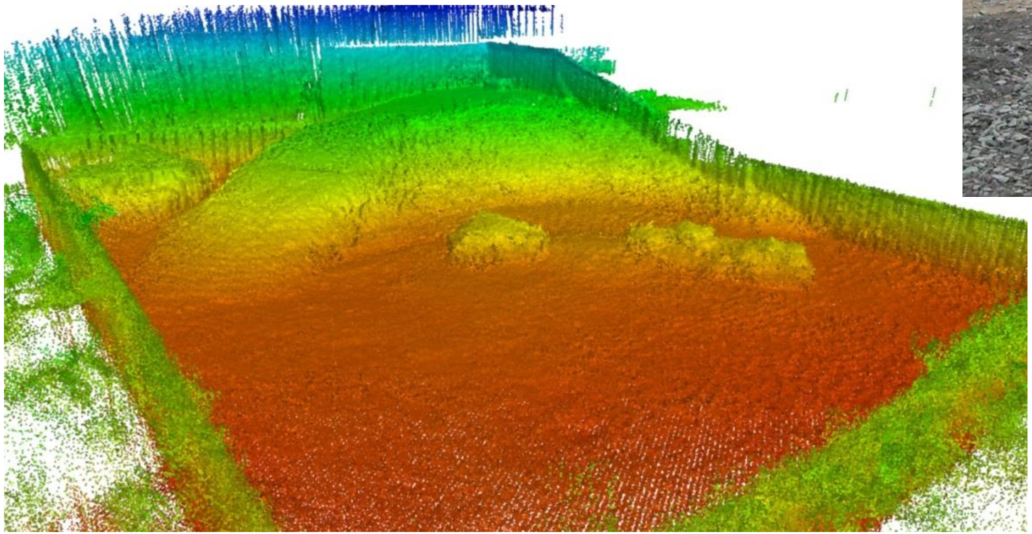
Team NimbRo Rescue

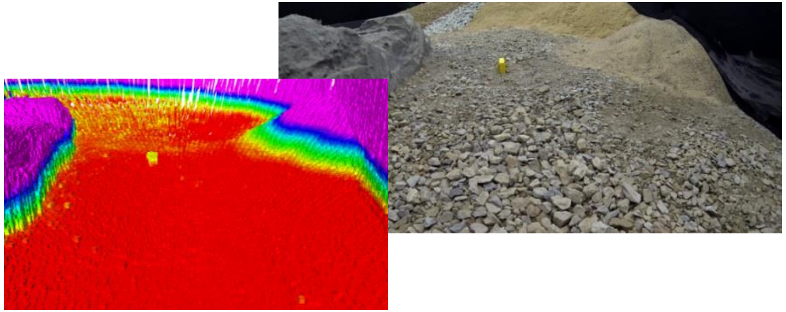
Best European Team (4th place overall), solved seven of eight tasks in 34 minutes

DLR SpaceBot Cup 2015

- Mobile manipulation in rough terrain

[Schwarz et al., Frontiers on Robotics and AI 2016]

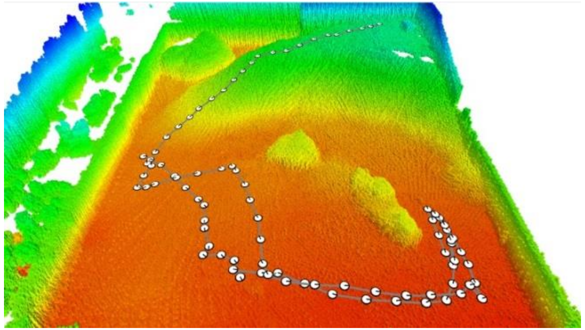




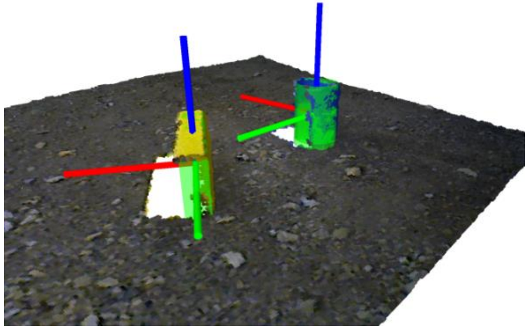

8X

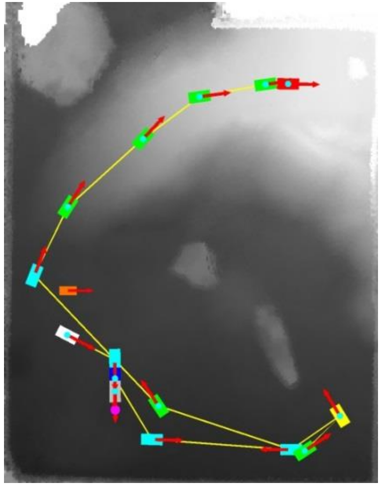
Autonomous Mission Execution

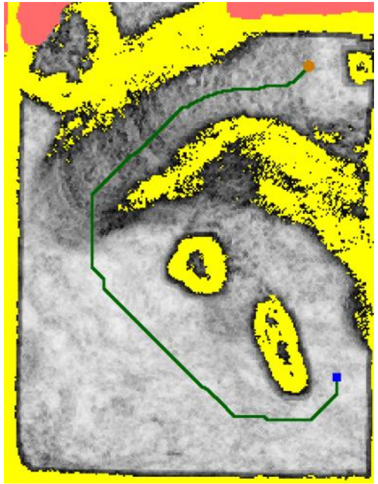
- 3D mapping, localization, mission and navigation planning

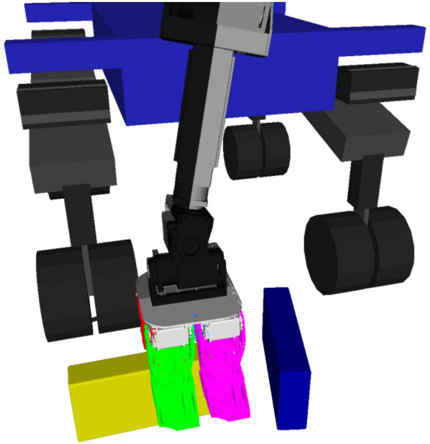


- 3D object perception and grasping







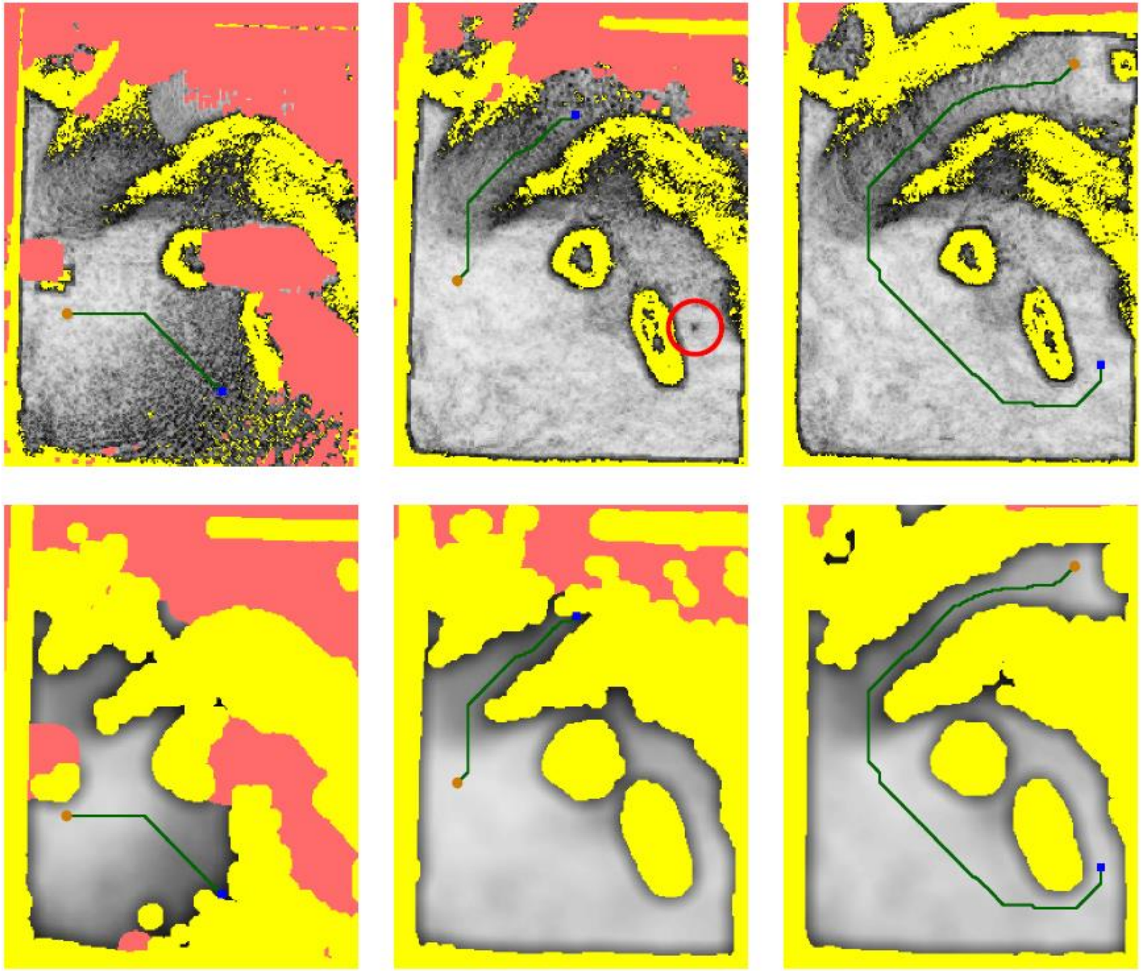


[Schwarz et al. Frontiers 2016]

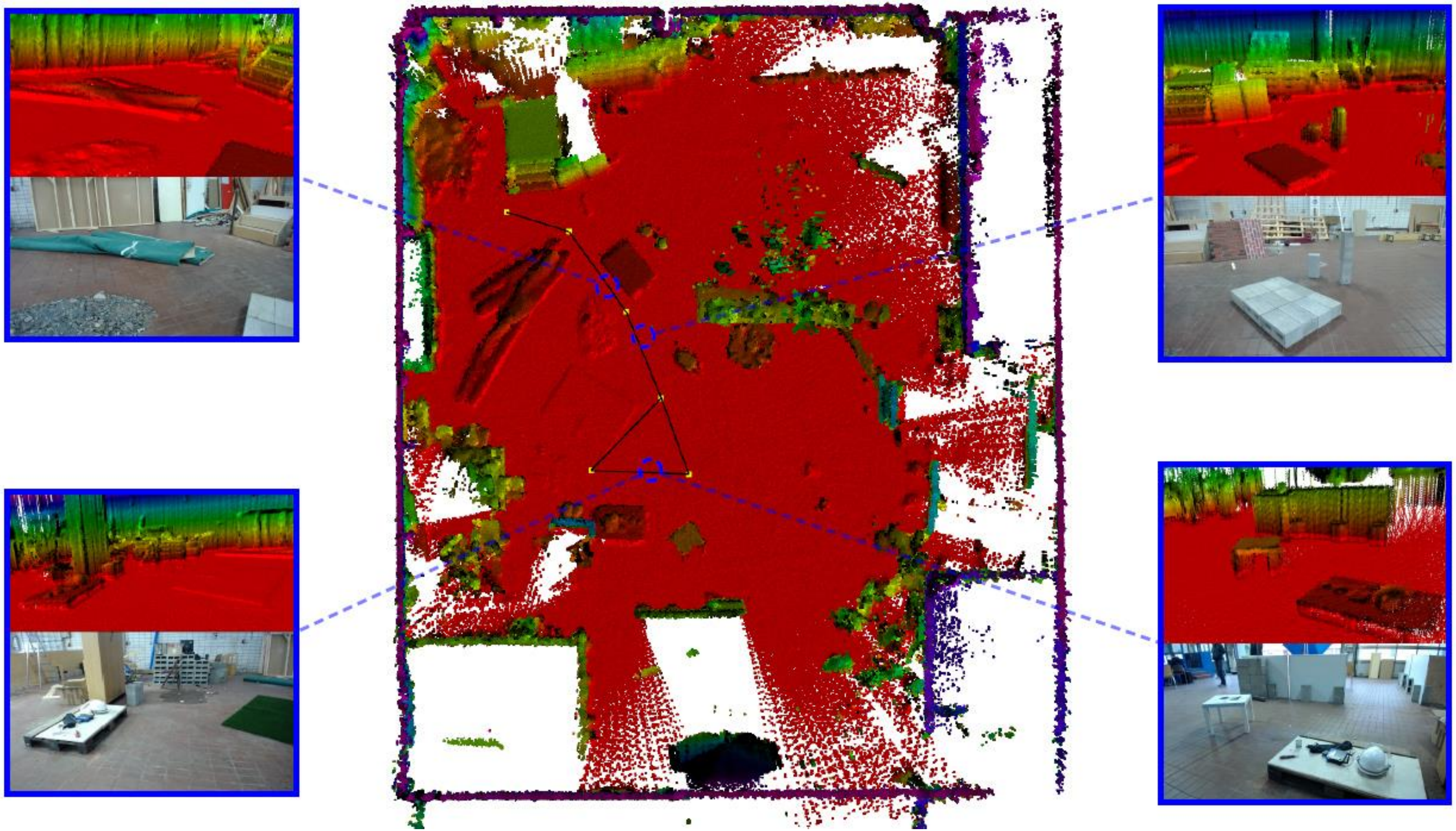
Navigation Planning

- Costs from local height differences
- A* path planning

[Schwarz et al., Frontiers in Robotics and AI 2016]



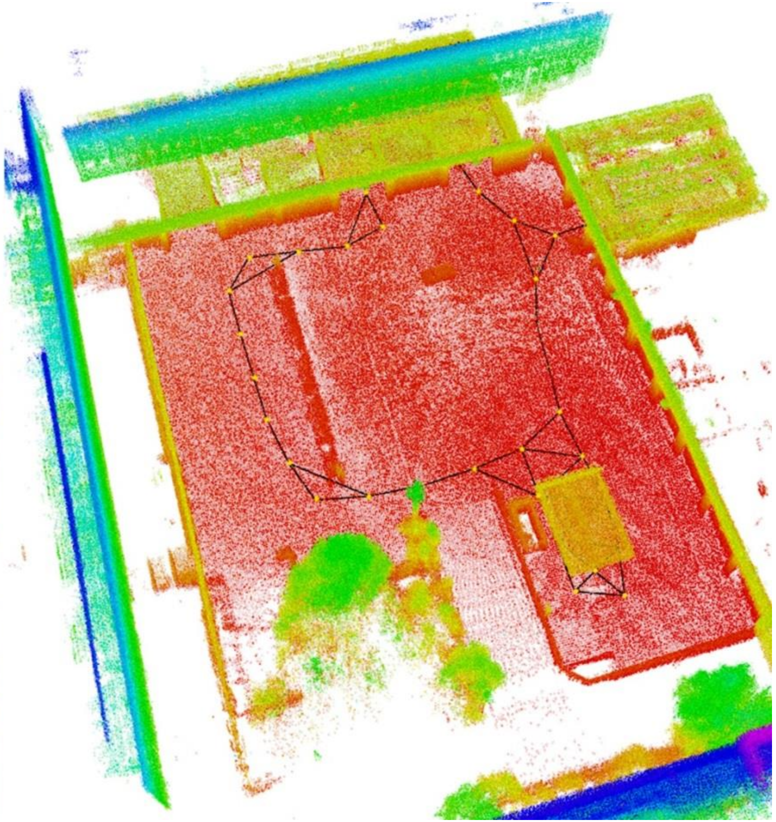
3D Map



Improved Sensor Head

- Continuously rotating Velodyne Puck VLP-16
  - 300,000 3D points/s
  - 100 m range
  - Spherical field of view
- Three wide-angle color cameras (total FoV 210×103°)
- Kinect V2 RGB-D camera on pan-tilt unit

3D Map of Indoor+Outdoor Scene

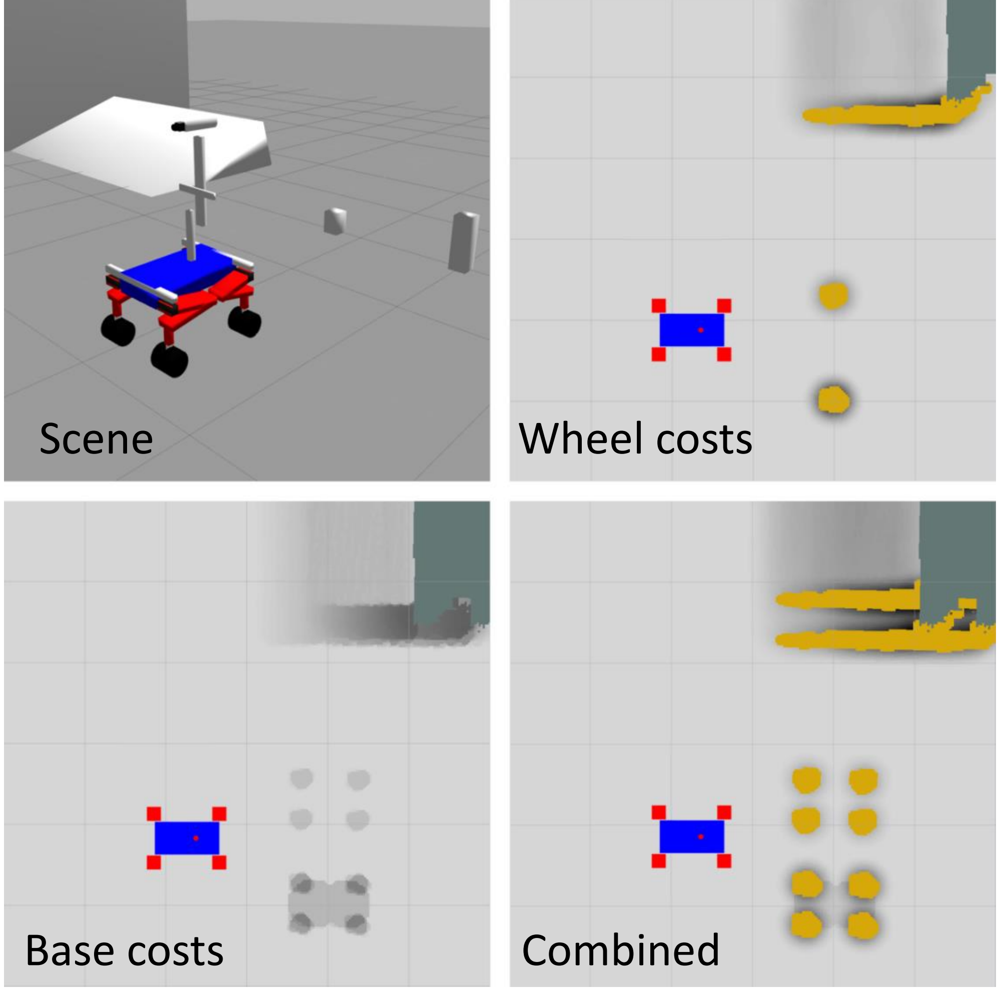


[Droeschel et al., Robotics and Autonomous Systems 2017]

UNIVERSITÄT BONN AIS

Considering Robot Footprint

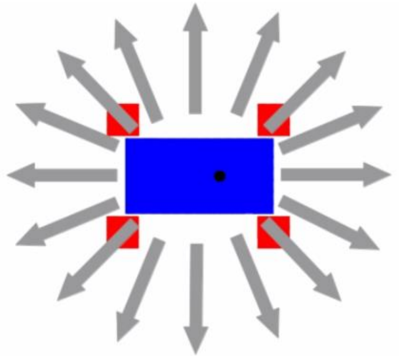
- Costs for individual wheel pairs from height differences
- Base costs
- Non-linear combination yields 3D (x, y, θ) cost map



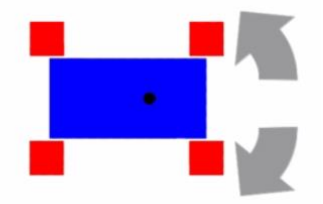
[Klamt and Behnke, IROS 2017]

3D Driving Planning (x, y, θ): A*

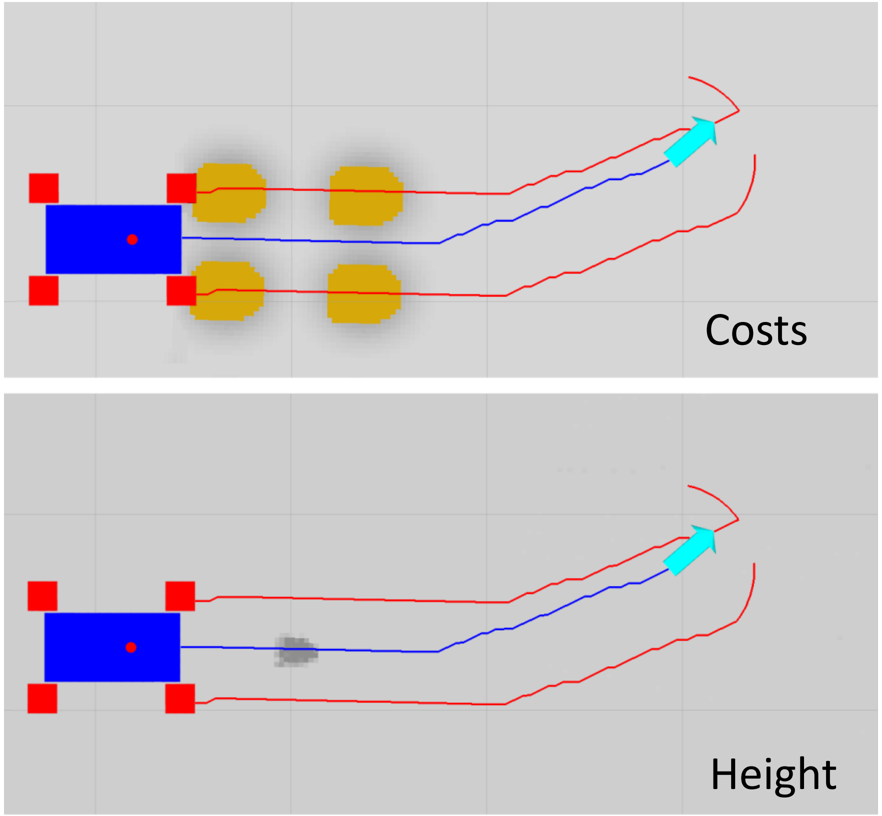
- 16 driving directions



- Orientation changes



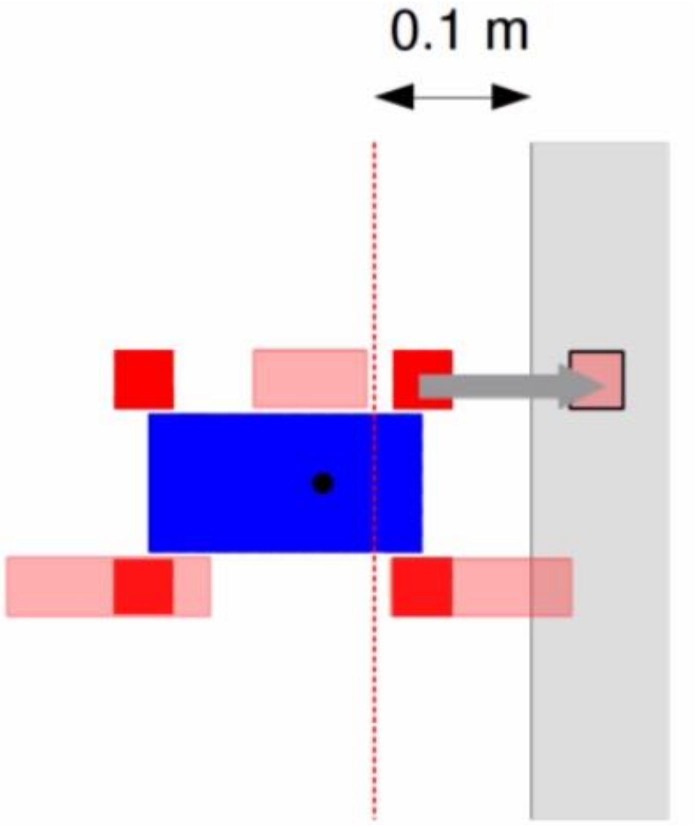
=> Obstacle between wheels



[Klamt and Behnke, IROS 2017]

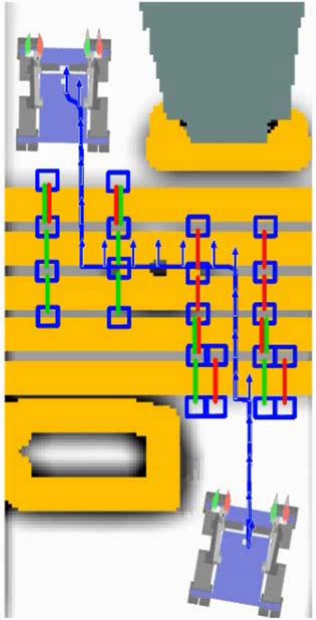
Making Steps

- If not drivable obstacle in front of a wheel
- Step landing must be drivable
- Support leg positions must be drivable



[Klamt and Behnke: IROS 2017]

Planning for Challenging Scenarios



[Klamt and Behnke: IROS 2017]

Centauro Robot

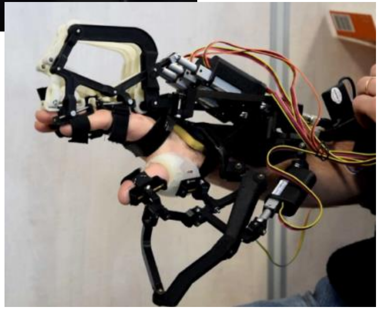
- Serial elastic actuators
- 42 main DoFs
- Schunk hand
- 3D laser
- RGB-D camera
- Color cameras
- Two GPU PCs

[Tsagarakis et al., IIT 2017]

Main Operator Telepresence Interface

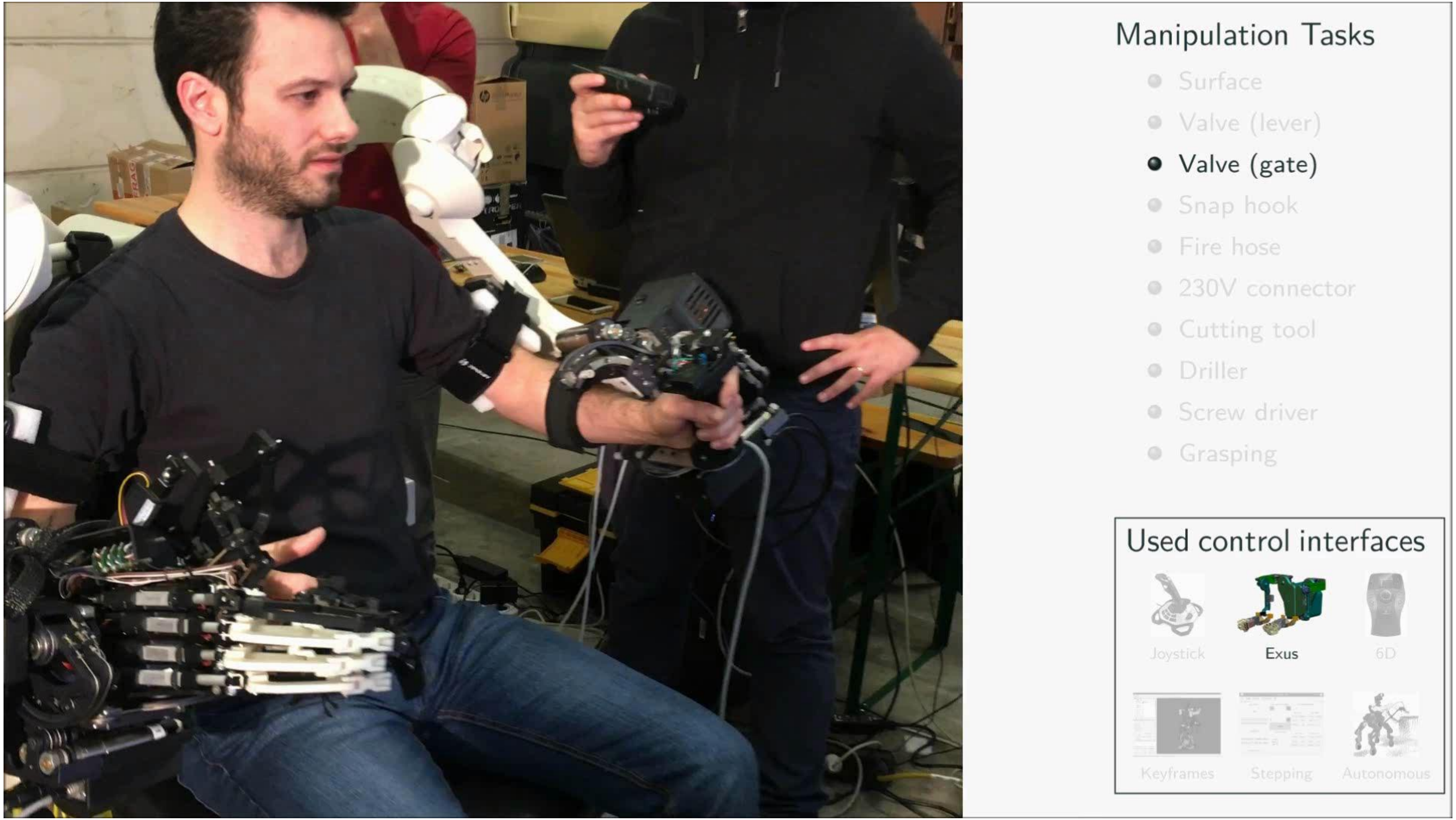
- Tendon-driven dual-arm exoskeleton
- Active wrist with differential tendon transmission
- Underactuated hand exoskeleton
- Head-mounted display
- Foot pedals





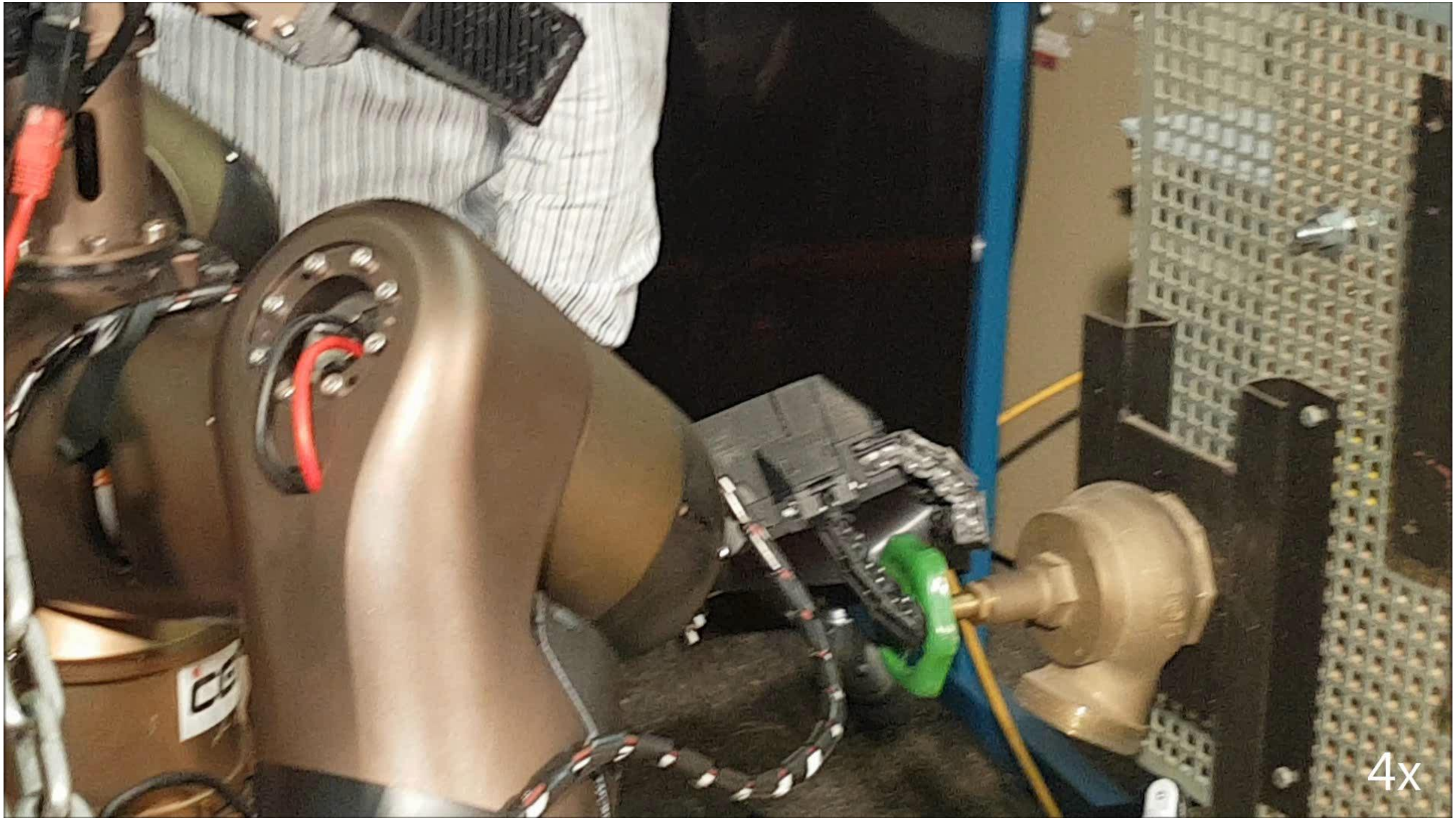

[Frisoli et al., SSSA 2017]

Main Operator Control

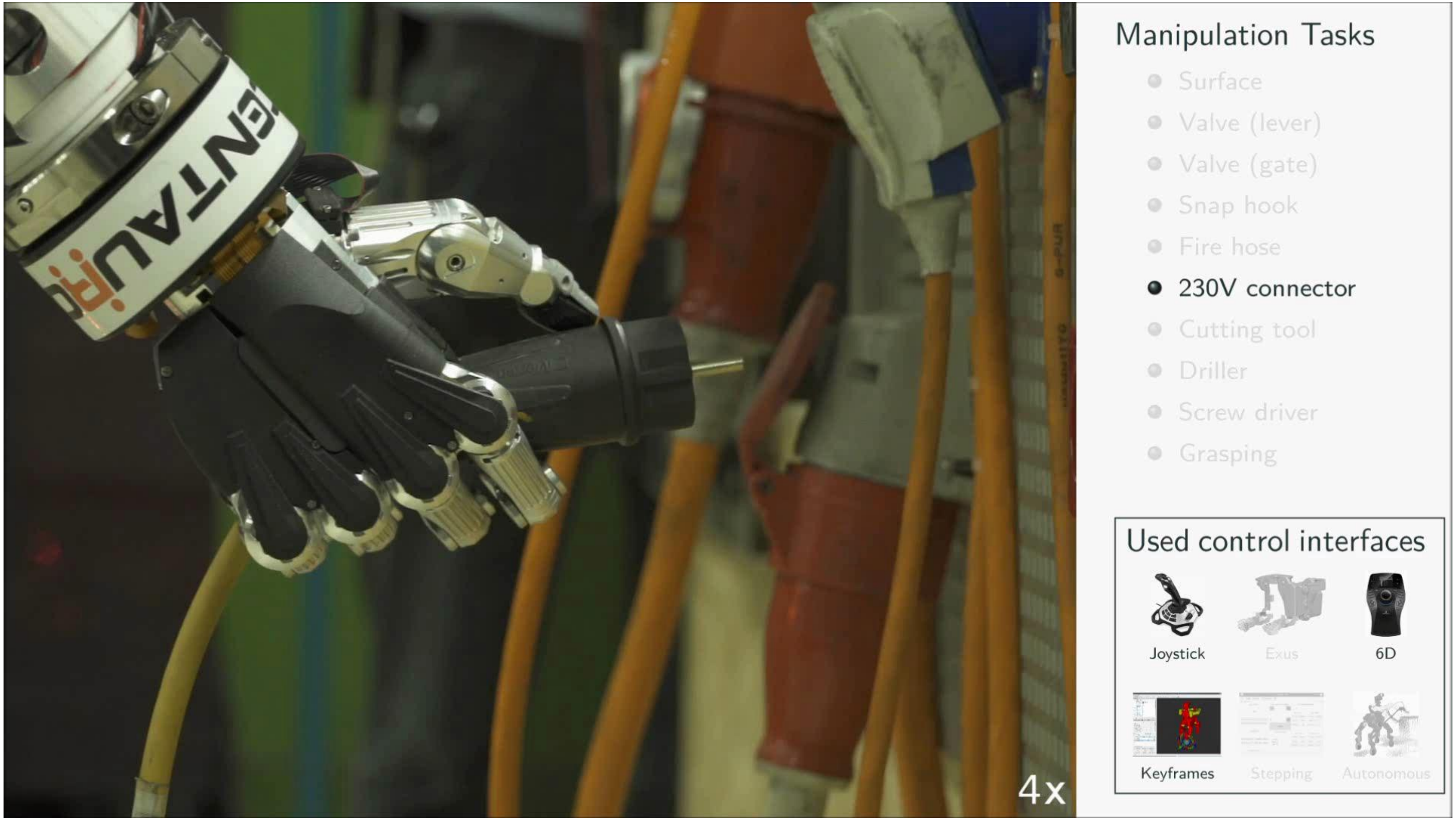


UNIVERSITÄT BONN AIS

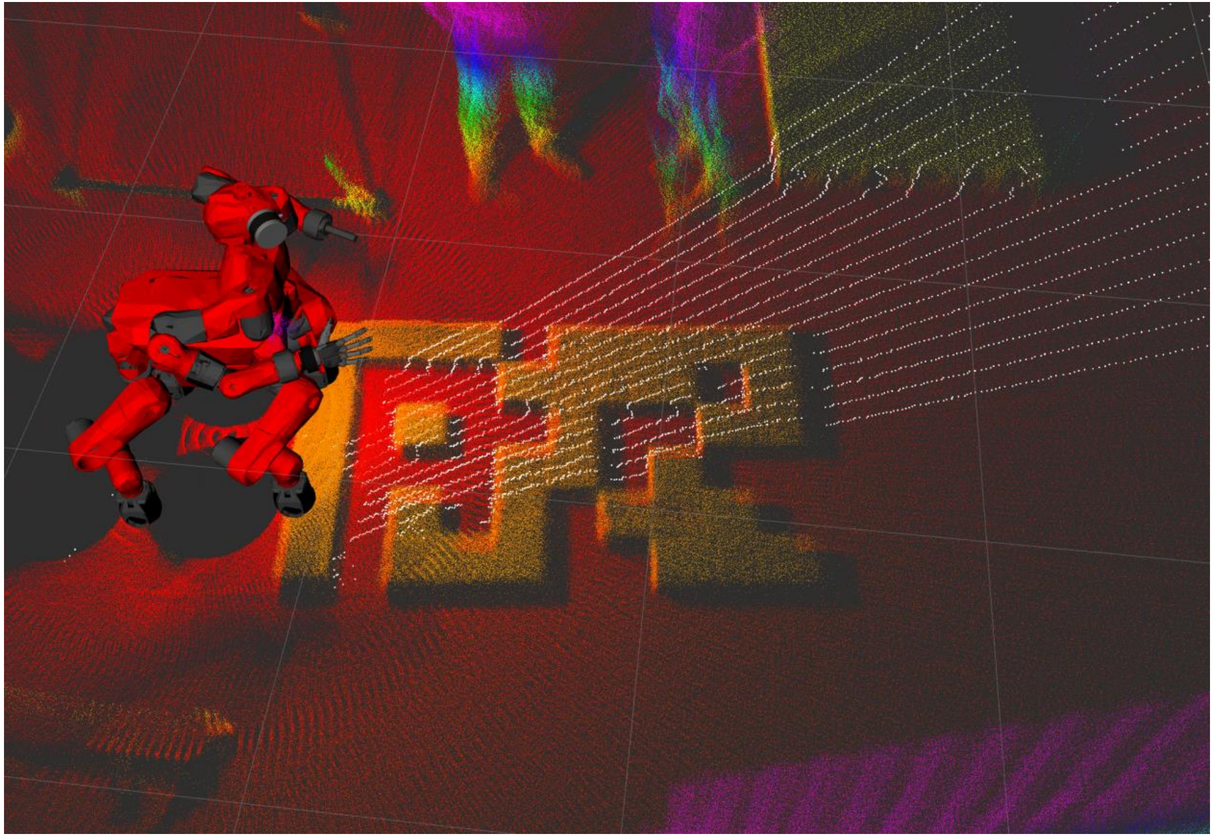
Turning a Valve



Connecting a Plug



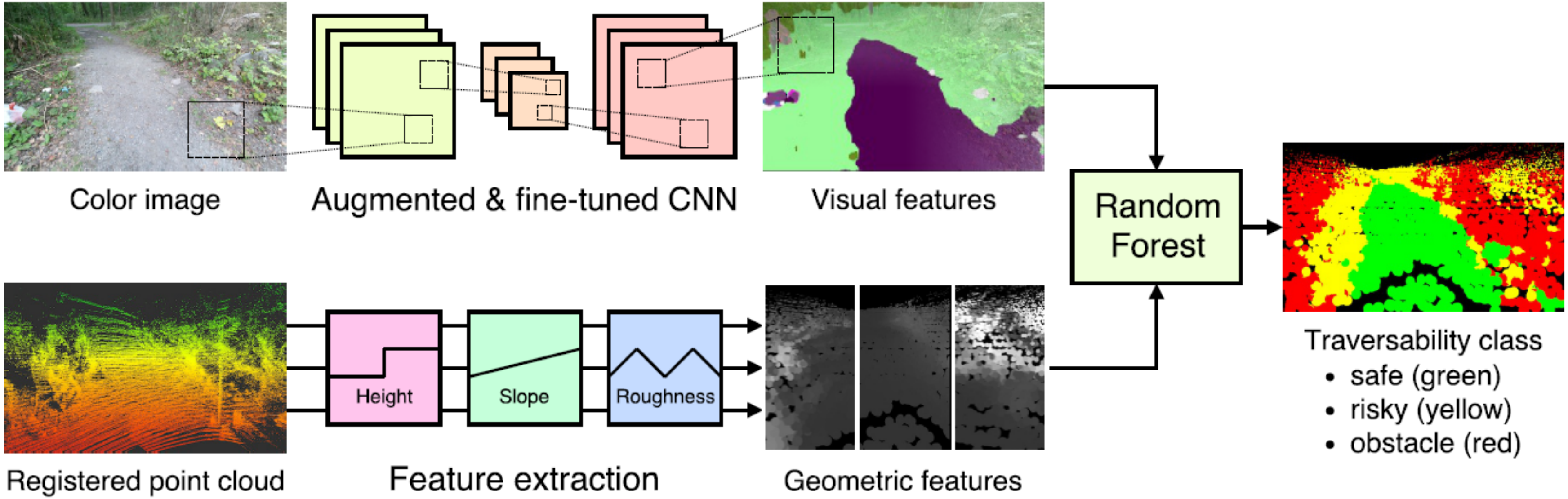
3D Mapping and Localization



Walking over a Step Field

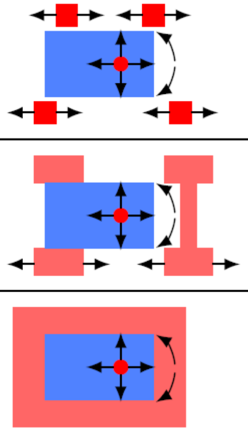
UNIVERSITÄT BONN AIS

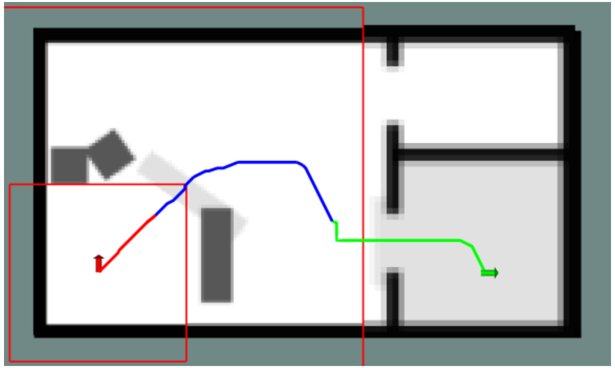
Terrain Classification

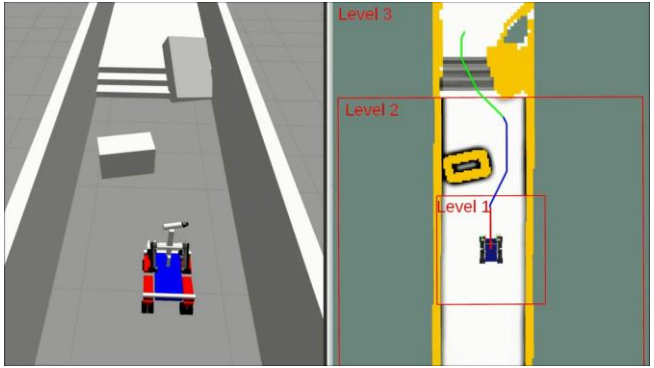


[Schilling et al., IROS 2017]

Hybrid Driving-Stepping Locomotion Planning: Abstraction

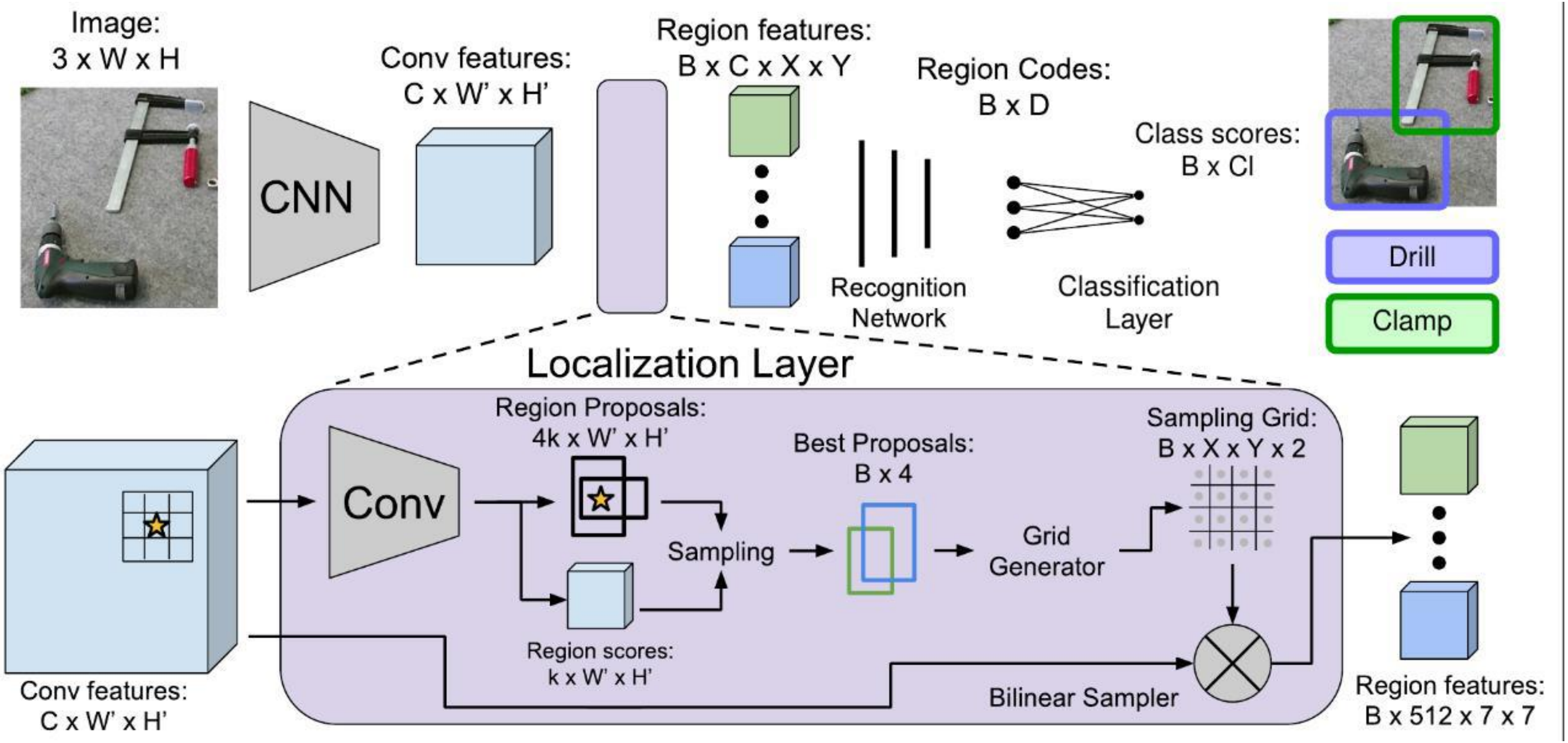
Level	Map Resolution	Map Features	Robot Representation	Action Semantics
1	• 2.5 cm • 64 orient.	• Height		• Individual Foot Actions
2	• 5.0 cm • 32 orient.	• Height • Height Difference		• Foot Pair Actions
3	• 10 cm • 16 orient.	• Height • Height Difference • Terrain Class		• Whole Robot Actions





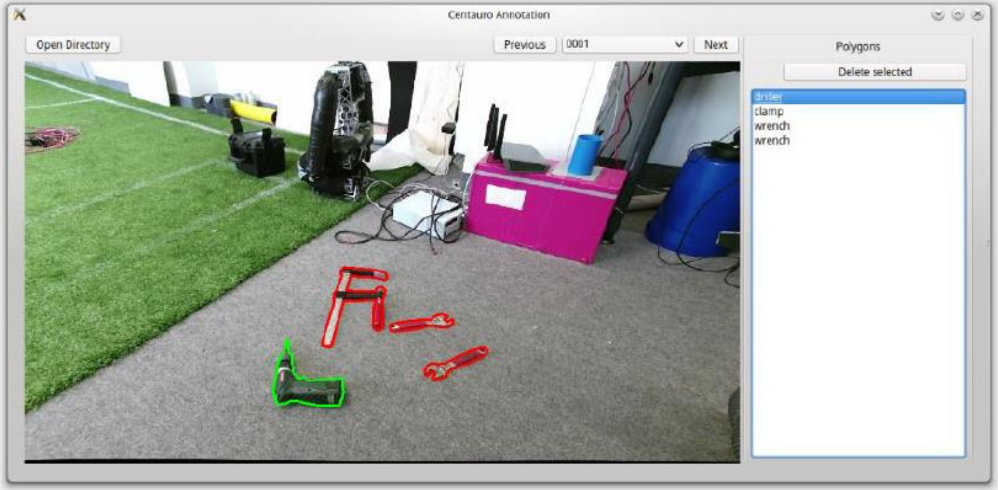
[Klamt and Behnke, IROS 2017, ICRA 2018]

Deep Learning Object Detection



[Johnson et al. 2015]

CENTAURO Workspace Perception Data Set



129 frames, 6 object classes

https://www.centauro-project.eu/data_multimedia/tools_data

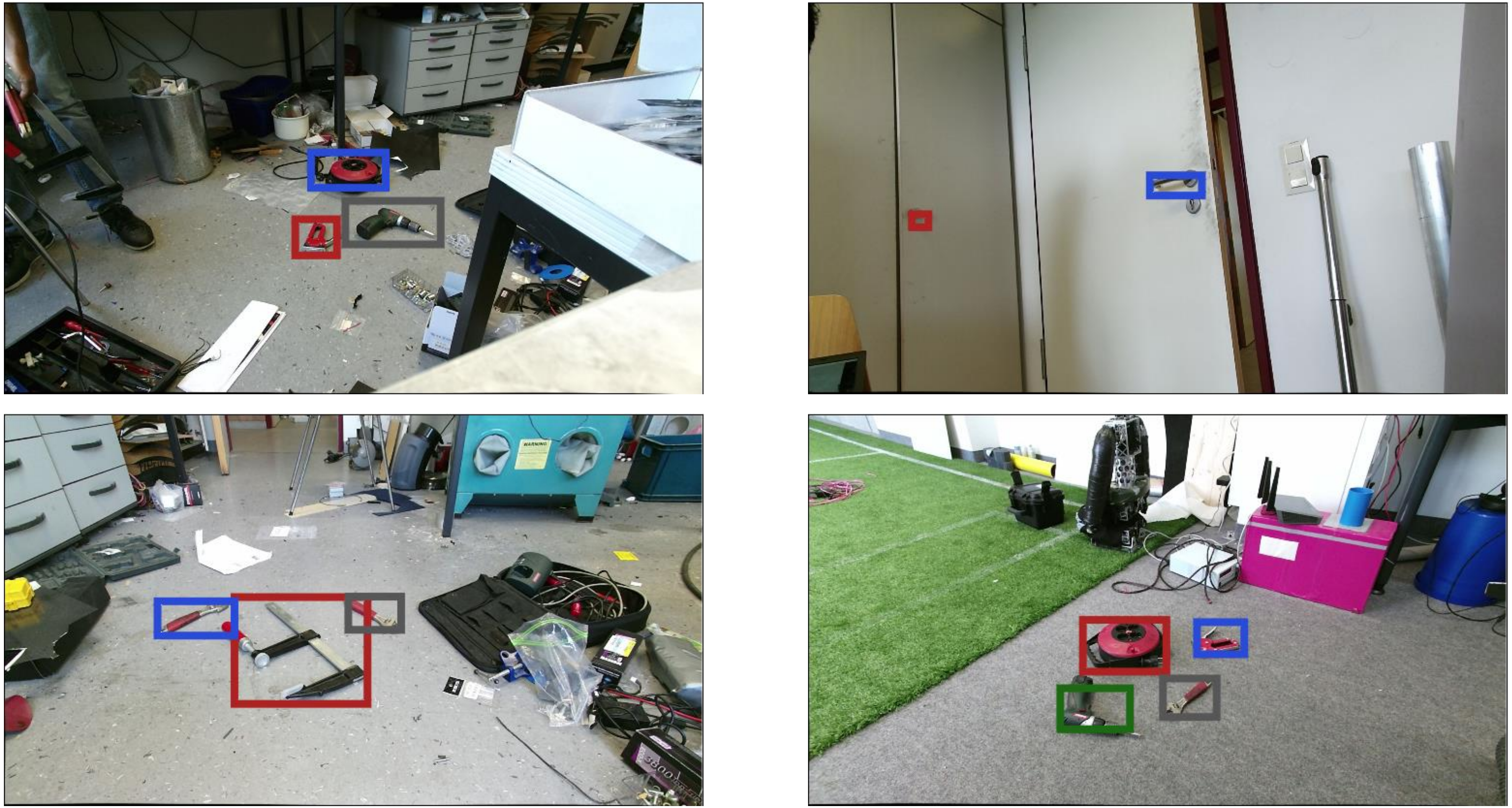
Tool Detection Results

[Schwarz et al. IJRR 2017]

extension_box stapler driller clamp [background]

Resolution	Clamp	Door handle	Driller	Extension	Stapler	Wrench	Mean
	AP / F1	AP / F1	AP / F1	AP / F1	AP / F1	AP / F1	AP / F1
720×507	0.881/0.783	0.522/0.554	0.986/0.875	1.000/0.938	0.960/0.814	0.656/0.661	0.834/0.771
1080×760	0.926/0.829	0.867/0.632	0.972/0.893	1.000/0.950	0.992/0.892	0.927/0.848	0.947/0.841
1470×1035	0.913/0.814	0.974/0.745	1.000/0.915	1.000/0.952	0.999/0.909	0.949/0.860	0.973/0.866

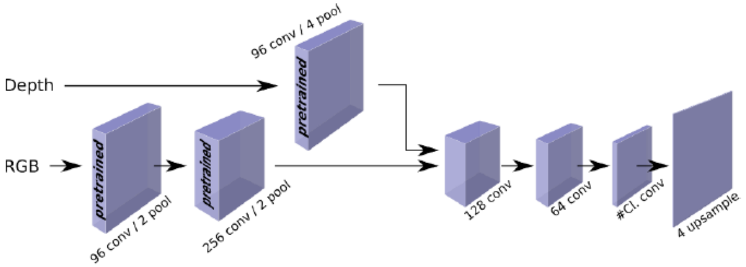
Tools Detection Examples



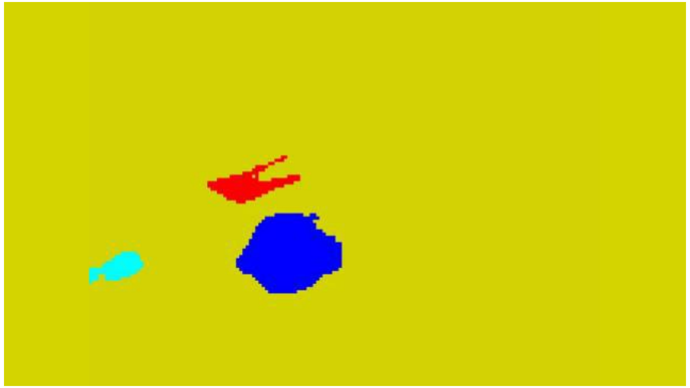
[Schwarz et al. IJRR 2017]

Semantic Segmentation

- Deep CNN



[Husain et al. RA-L 2016]

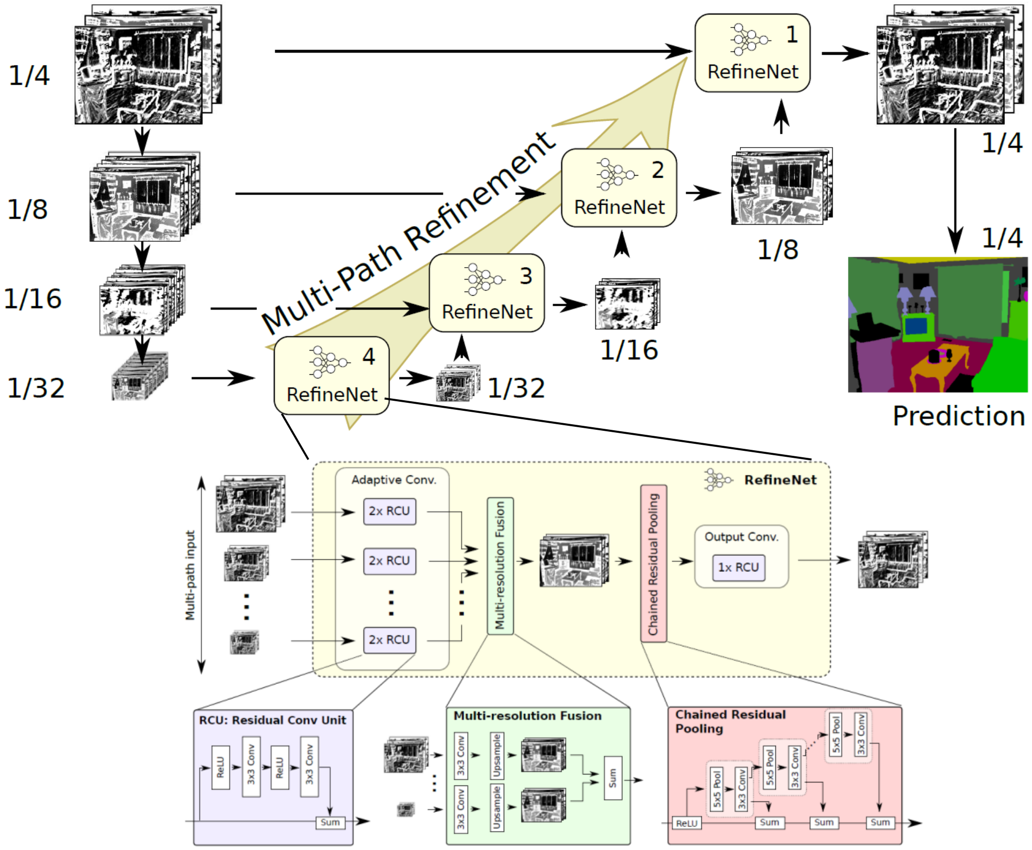


Pixel-wise accuracy:

Clamp	Door handle	Driller	Extension	Stapler	Wrench	Background	Mean
0.727	0.751	0.769	0.889	0.775	0.734	0.992	0.805

RefineNet for Semantic Segmentation

- Scene represented as feature hierarchy
- Corse-to-fine semantic segmentation
- Combine higher-level features with missing details



[Lin et al. CVPR 2017]

The Data Problem

- Deep Learning in robotics (still) suffers from shortage of available examples
- We address this problem in two ways:

1. **Generating data**:
   Automatic data capture,
   online mesh databases,
   scene synthesis

2. **Improving generalization**:
   Object-centered models,
   deformable registration,
   transfer learning,
   semi-supervised learning

Object Capture and Scene Rendering

- Turntable + DLSR camera

[Schwarz et al. ICRA 2018]

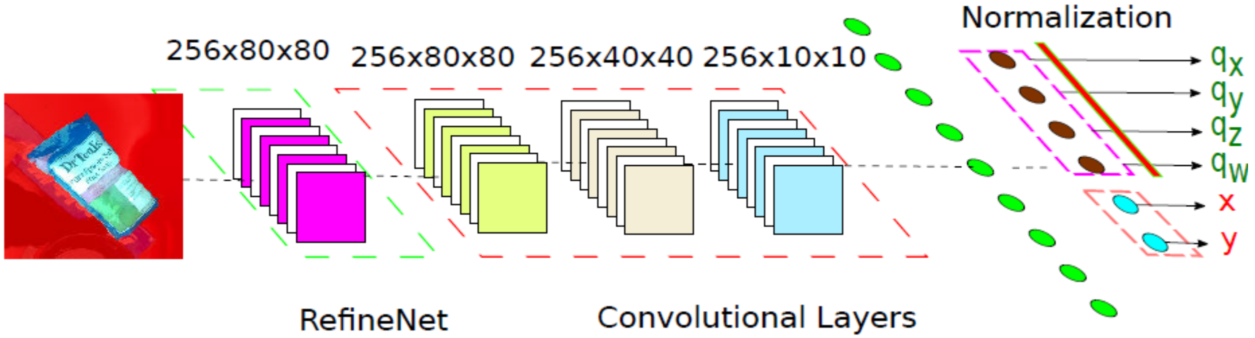
- Rendered scenes

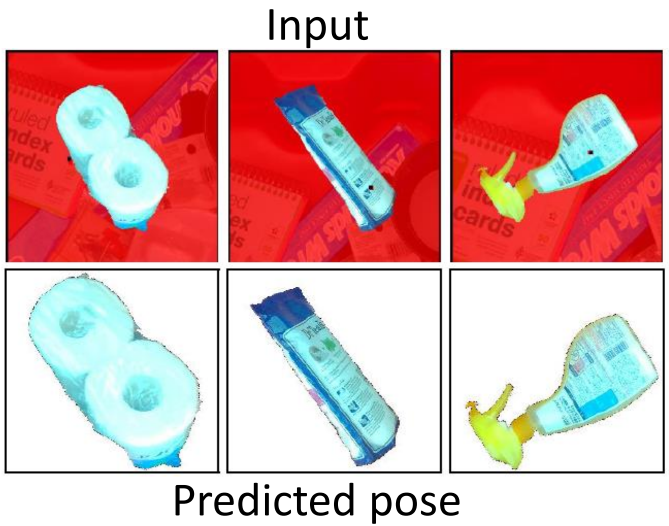
Semantic Segmentation Example

[Schwarz et al. ICRA 2018]

Object Pose Estimation

- Cut out individual segments
- Use upper layer of RefineNet as input
- Predict pose coordinates





[Schwarz et al. ICRA 2018, Periyasamy et al. IROS 2018]

From Turntable Captures to Textured Meshes

Fused & textured result

Transfer of Manipulation Skills

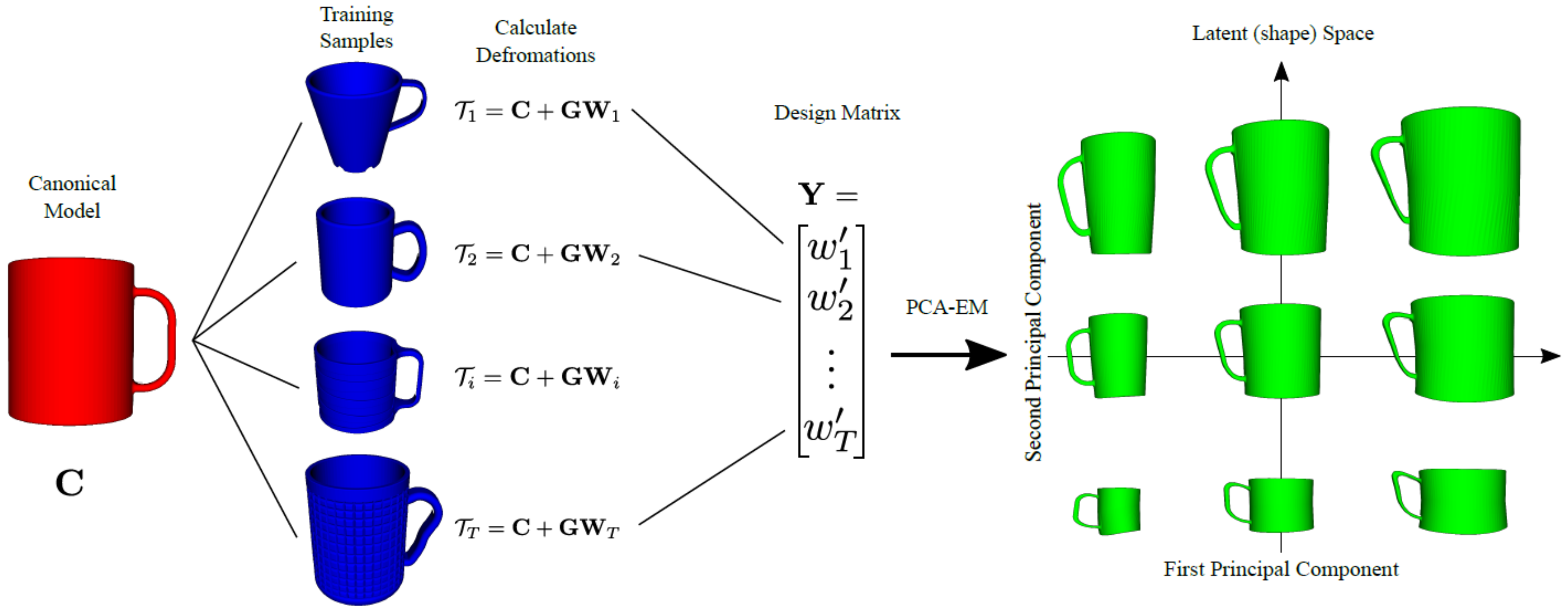
- Objects belonging to the same **category** can be handled in a very similar manner.

Transfer of Manipulation Skills

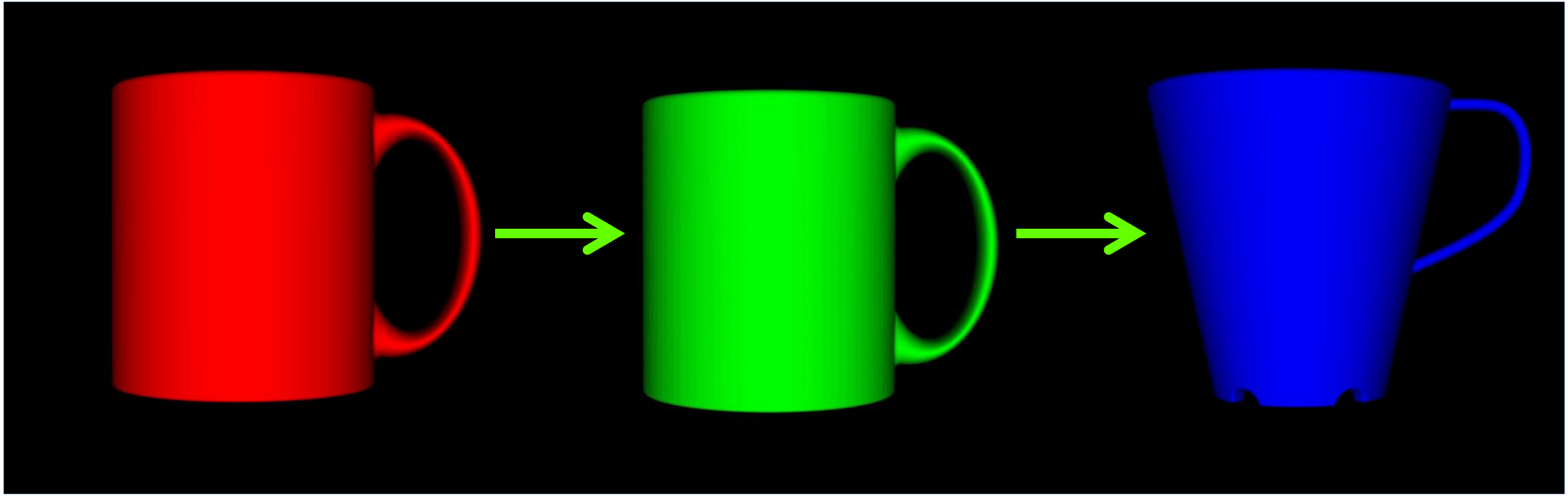
Knowledge Transfer

Learning a Latent Shape Space

- Non-rigid registration of instances and canonical model
- Principal component analysis of deformations



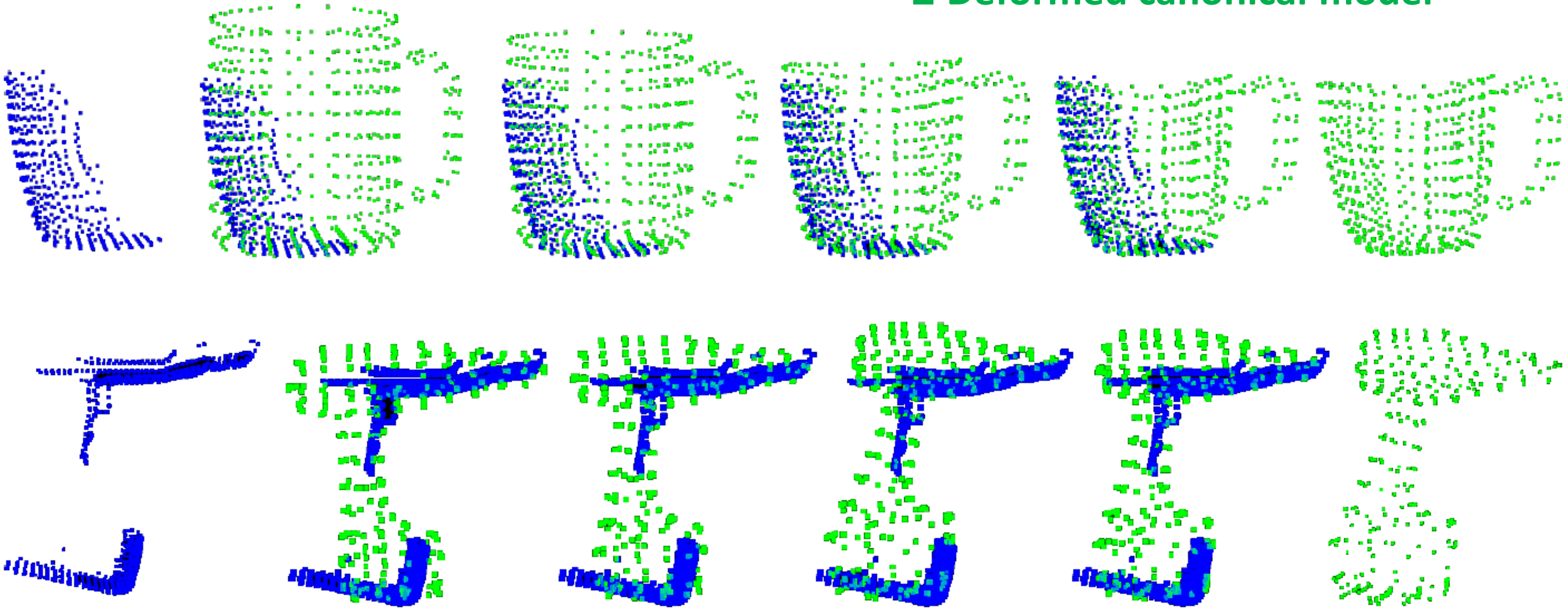
Interpolation in Shape Space



[Rodriguez and Behnke ICRA 2018]

Shape-aware Non-rigid Registration

■ Partial view of novel instance
■ Deformed canonical model

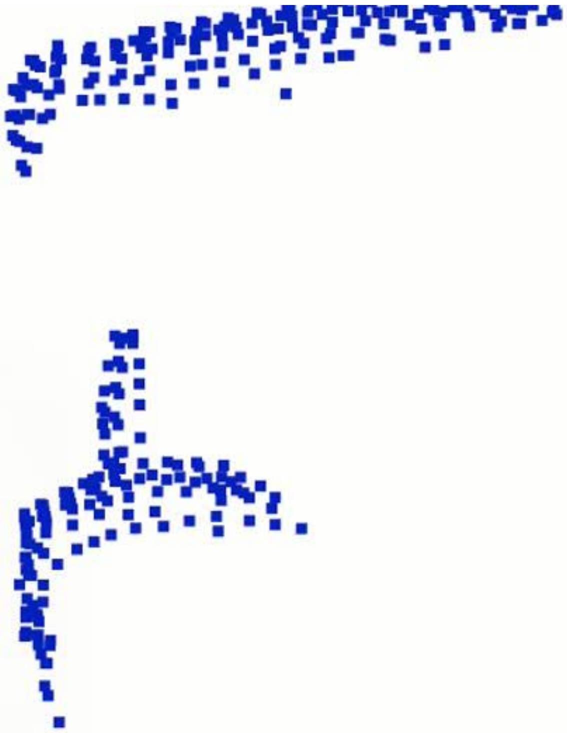


[Rodriguez and Behnke ICRA 2018]

Shape-aware Registration for Grasp Transfer

- Full point cloud

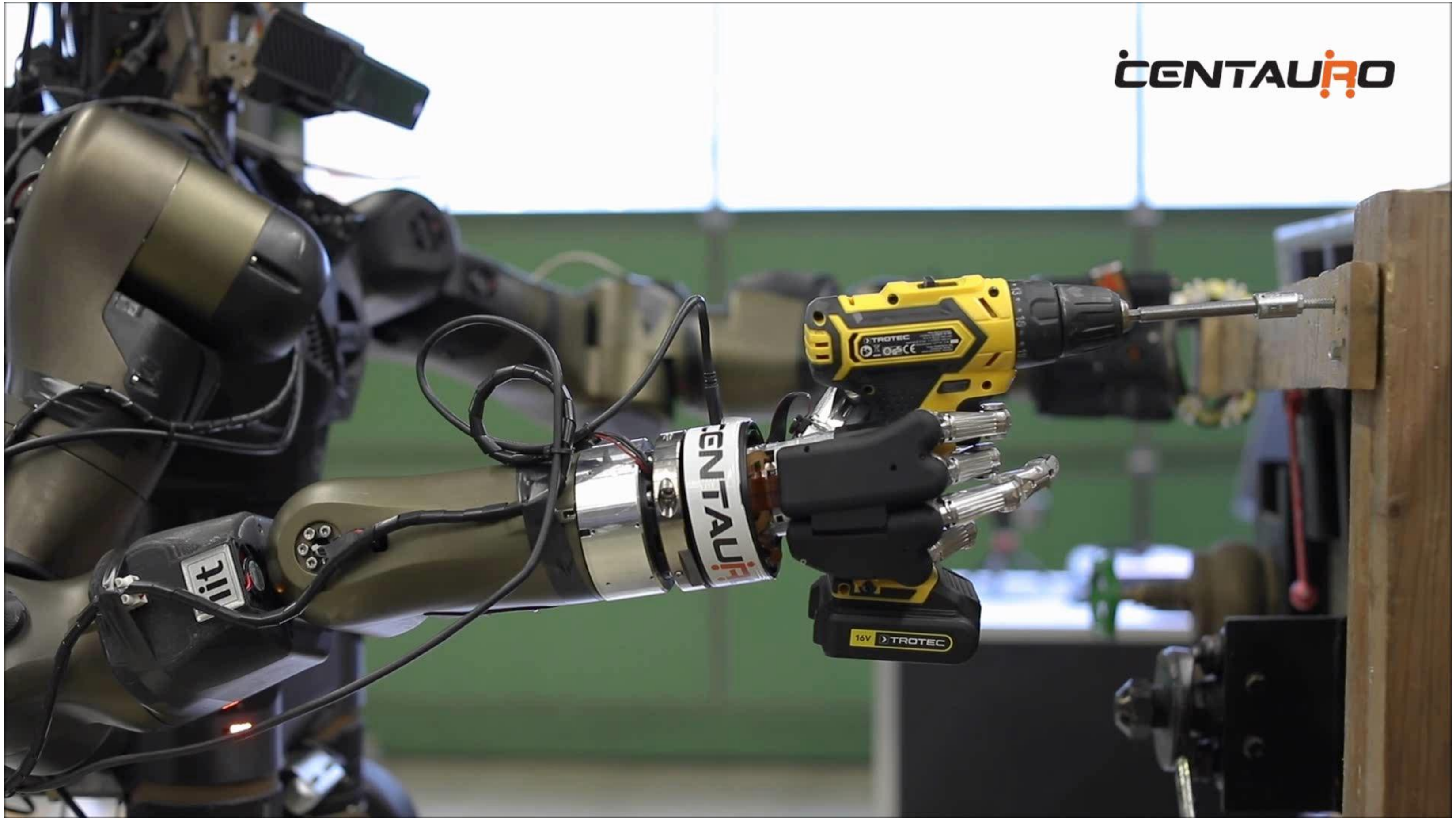
- Partial view



Grasping an Unknown Power Drill

Fastening a Screw

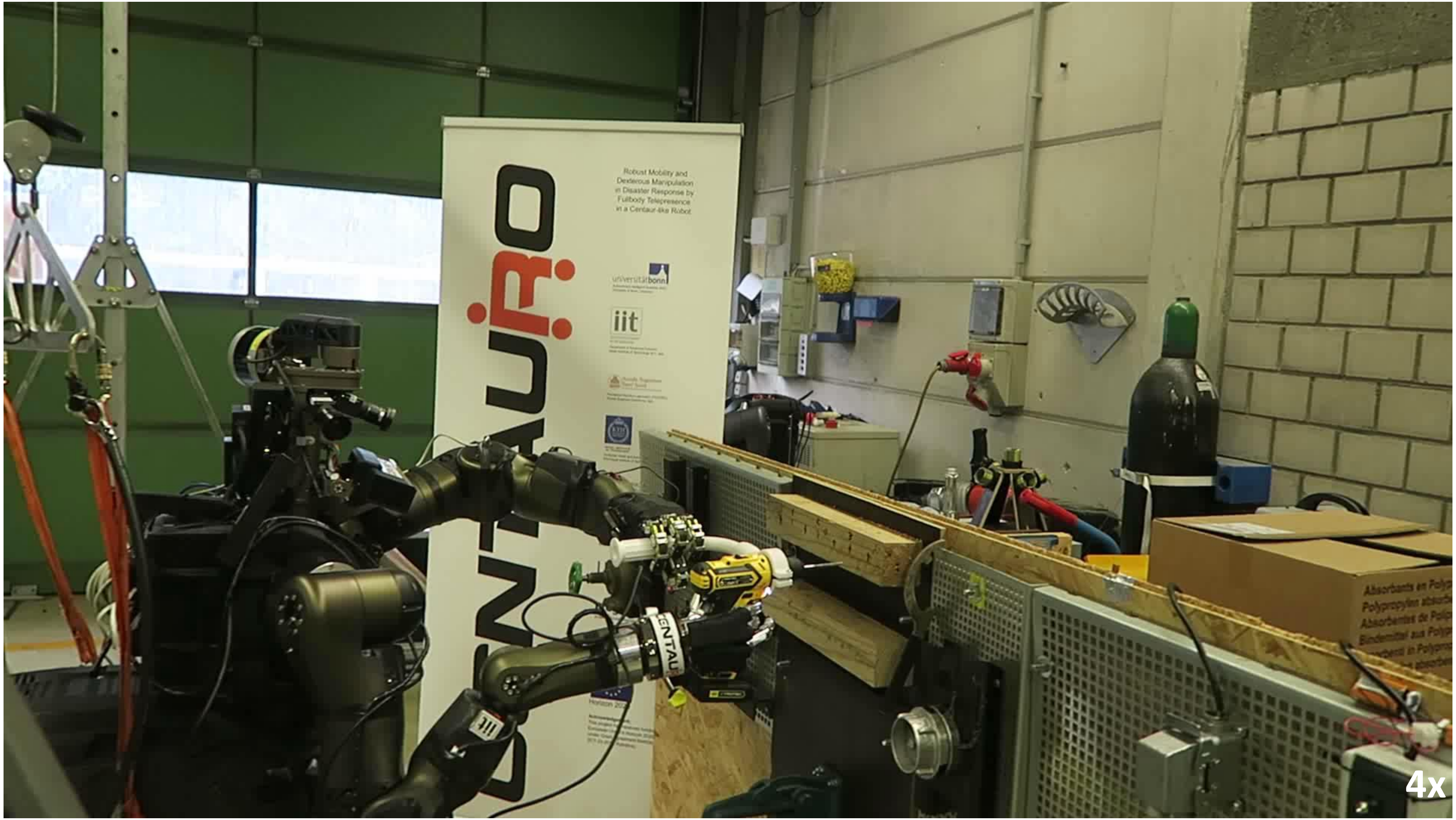
Bimanual Fastening Task



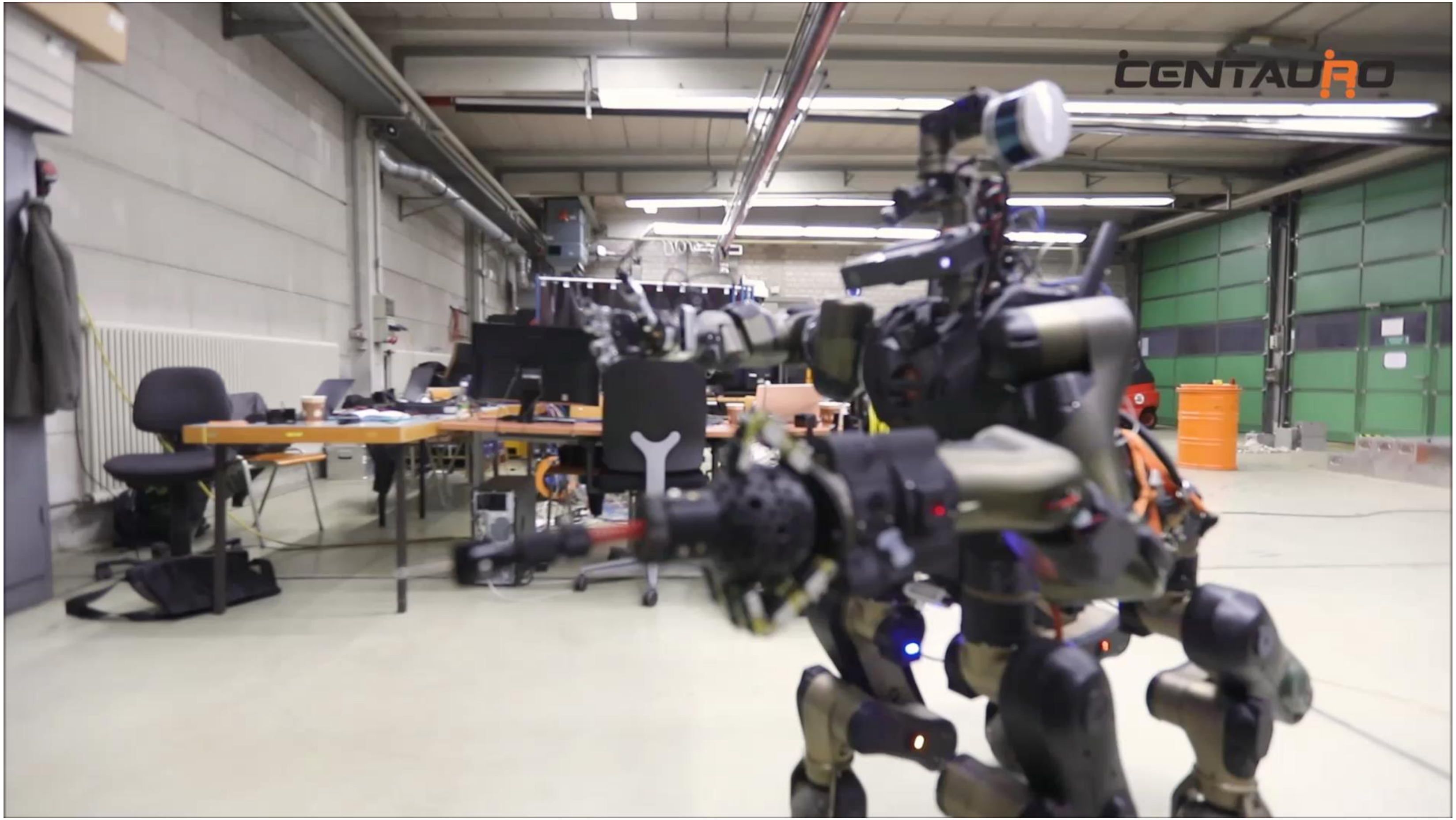
Bimanual Grasping

UNIVERSITÄT BONN AIS

Bimanual Drilling



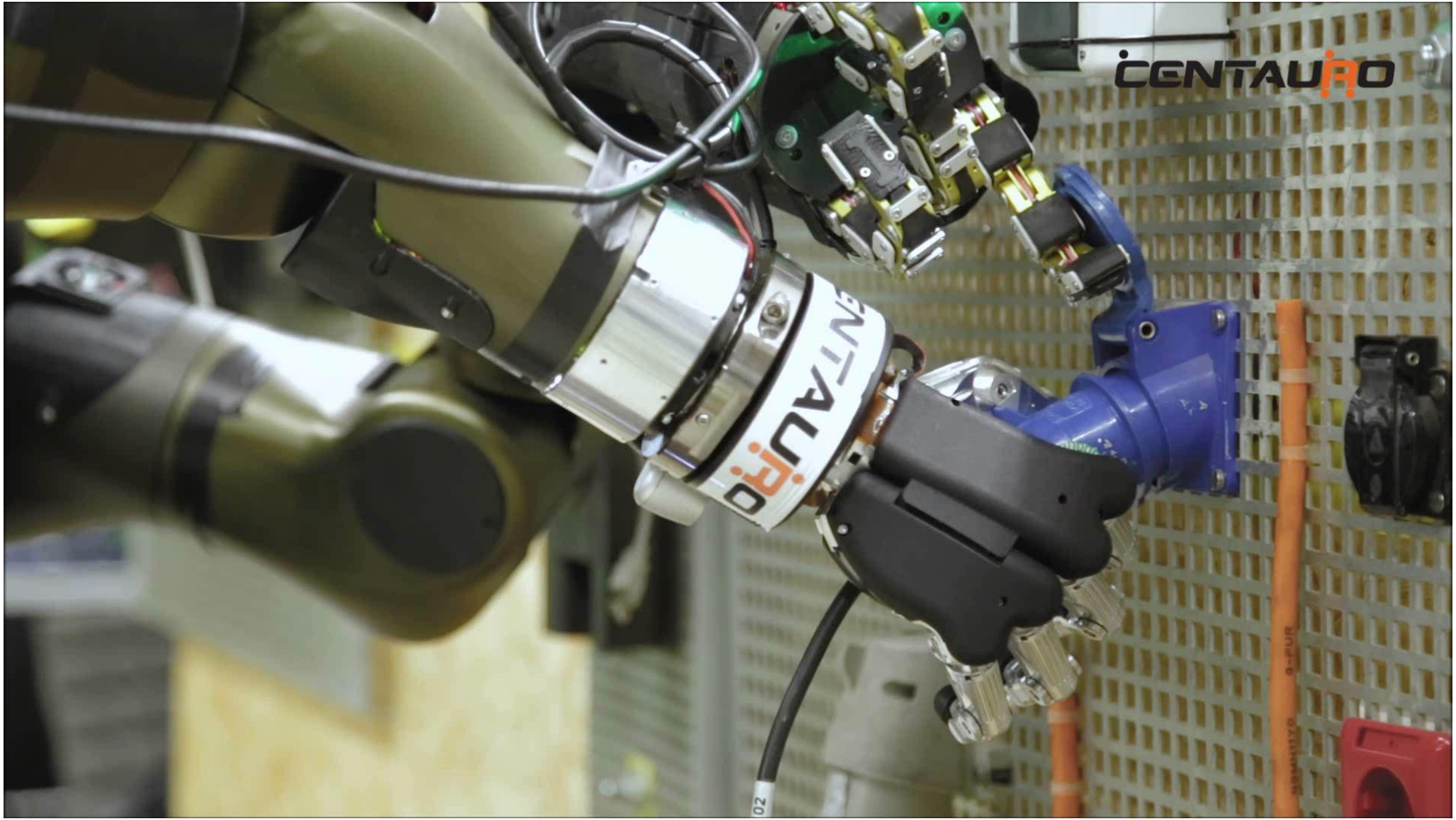
Opening a Door with a Key



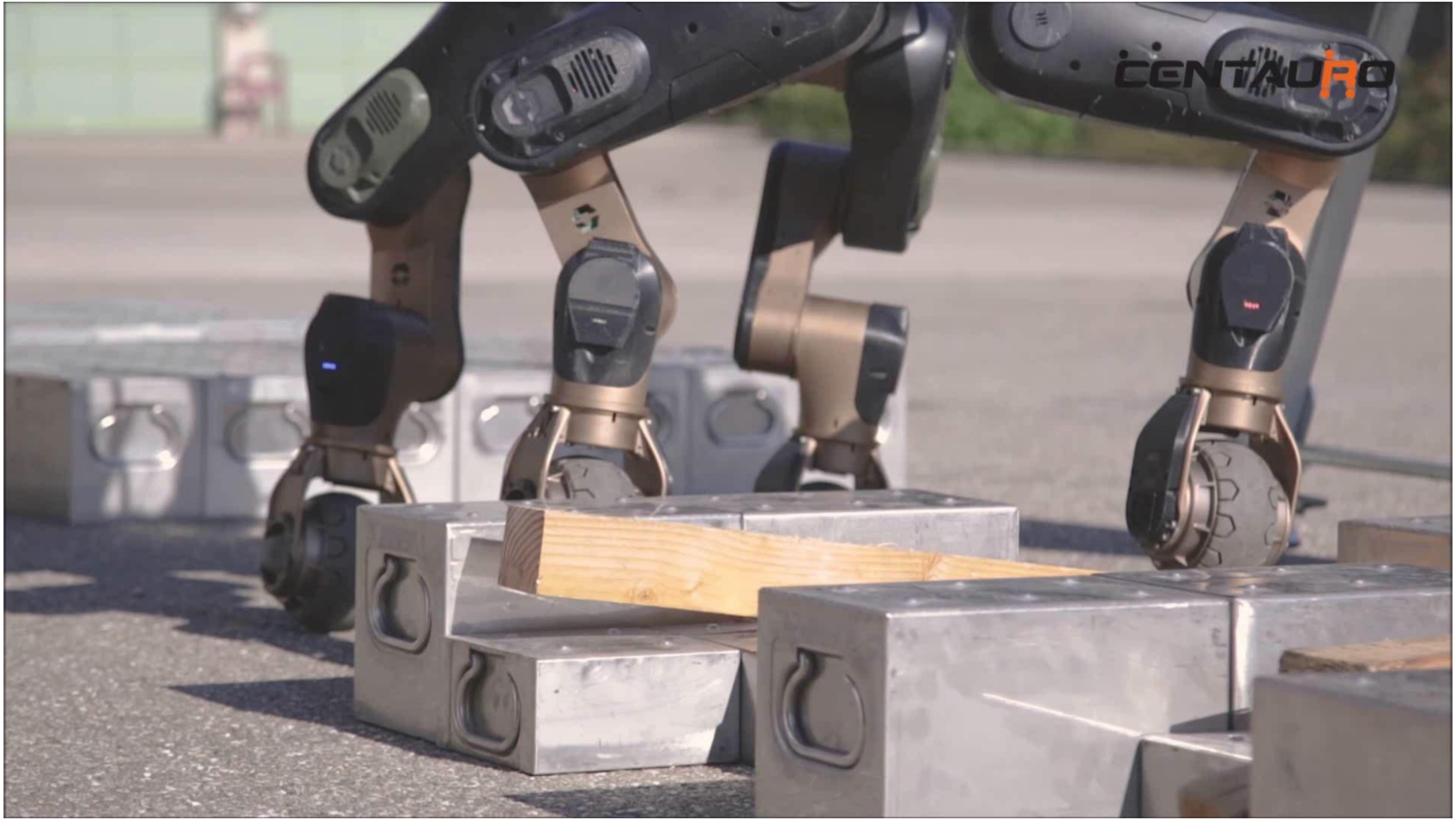
Closing a Shackle

UNIVERSITÄT BONN AIS

Bimanual Plug Tasks



Step Field with Debris

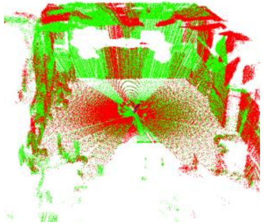


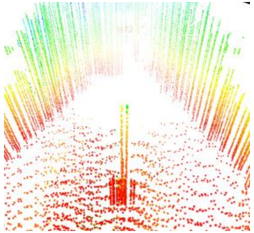
Autonomous Navigation

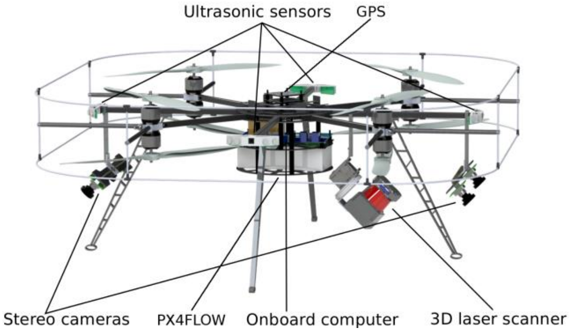
Autonomous Flight Near Obstacles

Multimodal obstacle detection

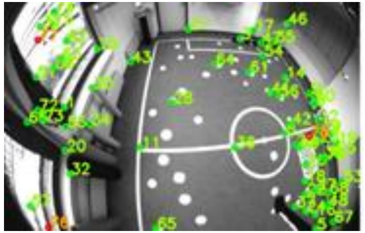
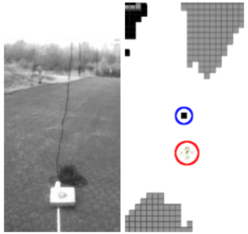
- 3D laser scanner

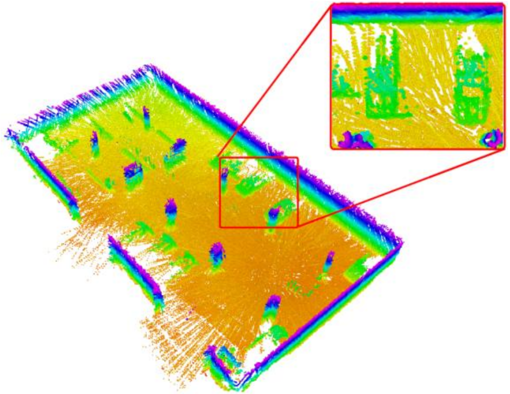







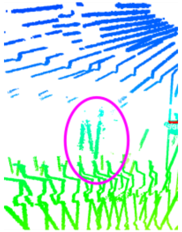
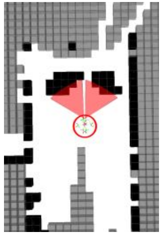
- Stereo cameras



- Ultrasound



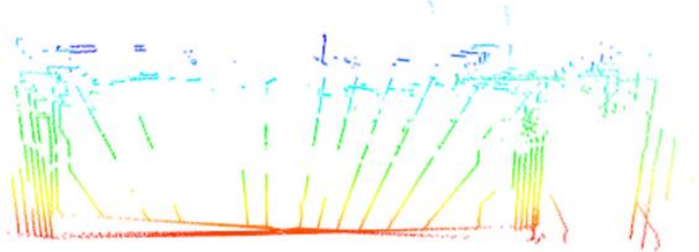



[Droeschel et al.: Journal of Field Robotics, 2015]

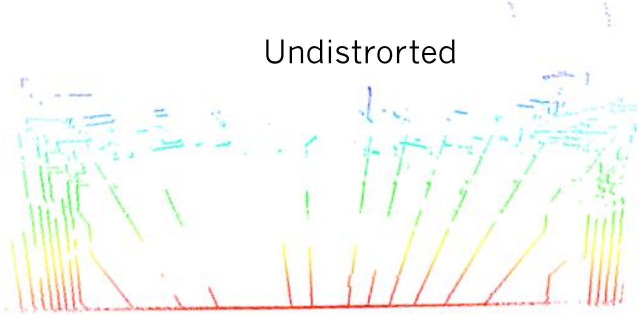
Egocentric Laser-based 3D Mapping

- Motion compensation

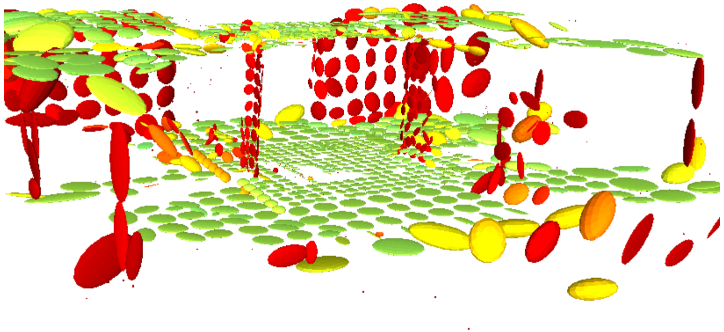
Distorted

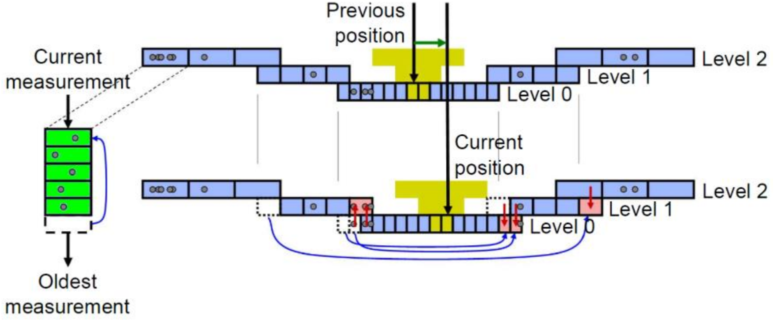


Undistrorted



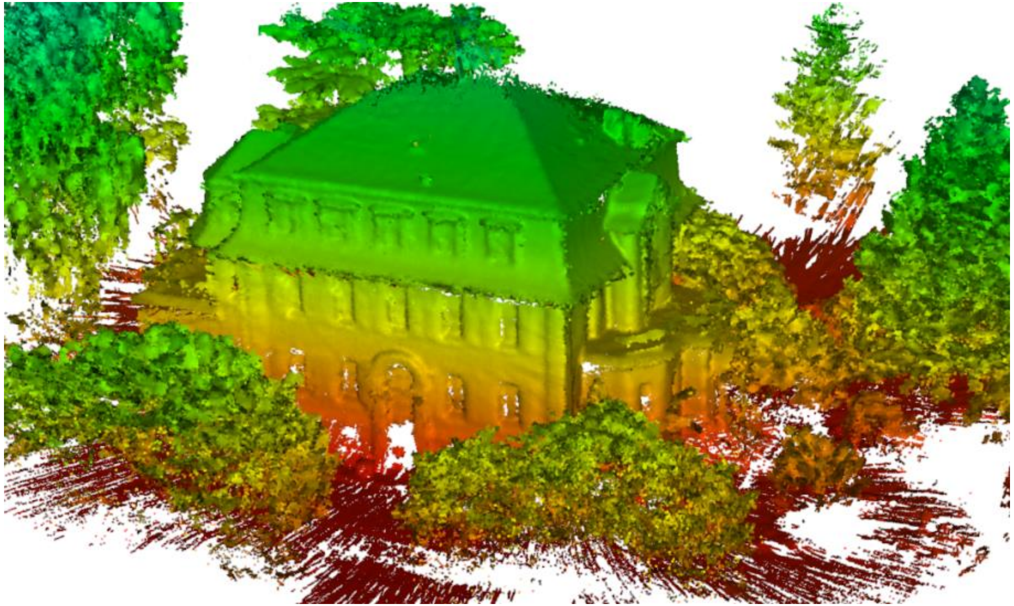
- Local multiresolution surfel maps





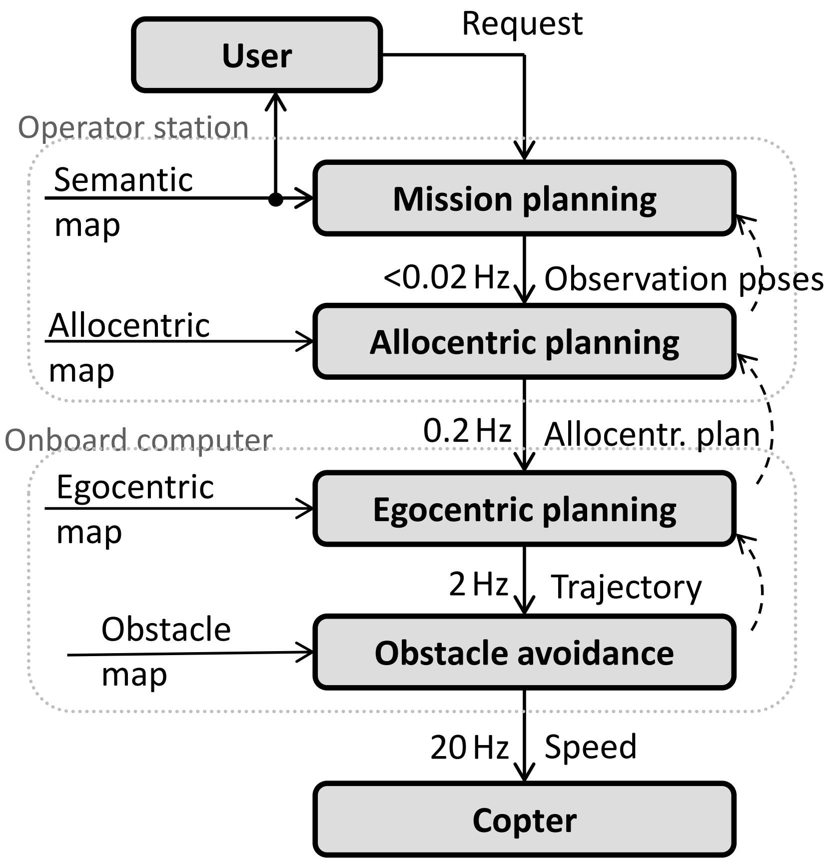
Allocentric 3D Map

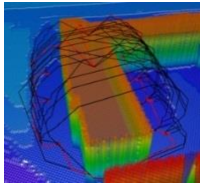
- Registration of egocentric maps
- Global optimization of registration error by GraphSLAM



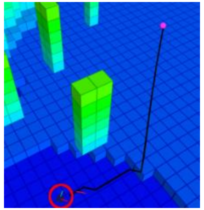
[Droeschel et al. JFR 2016]

Hierarchical Navigation

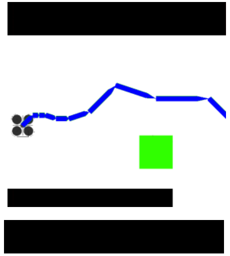




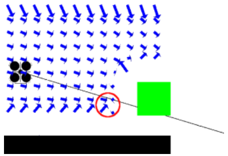
Mission plan



Allocentric planning



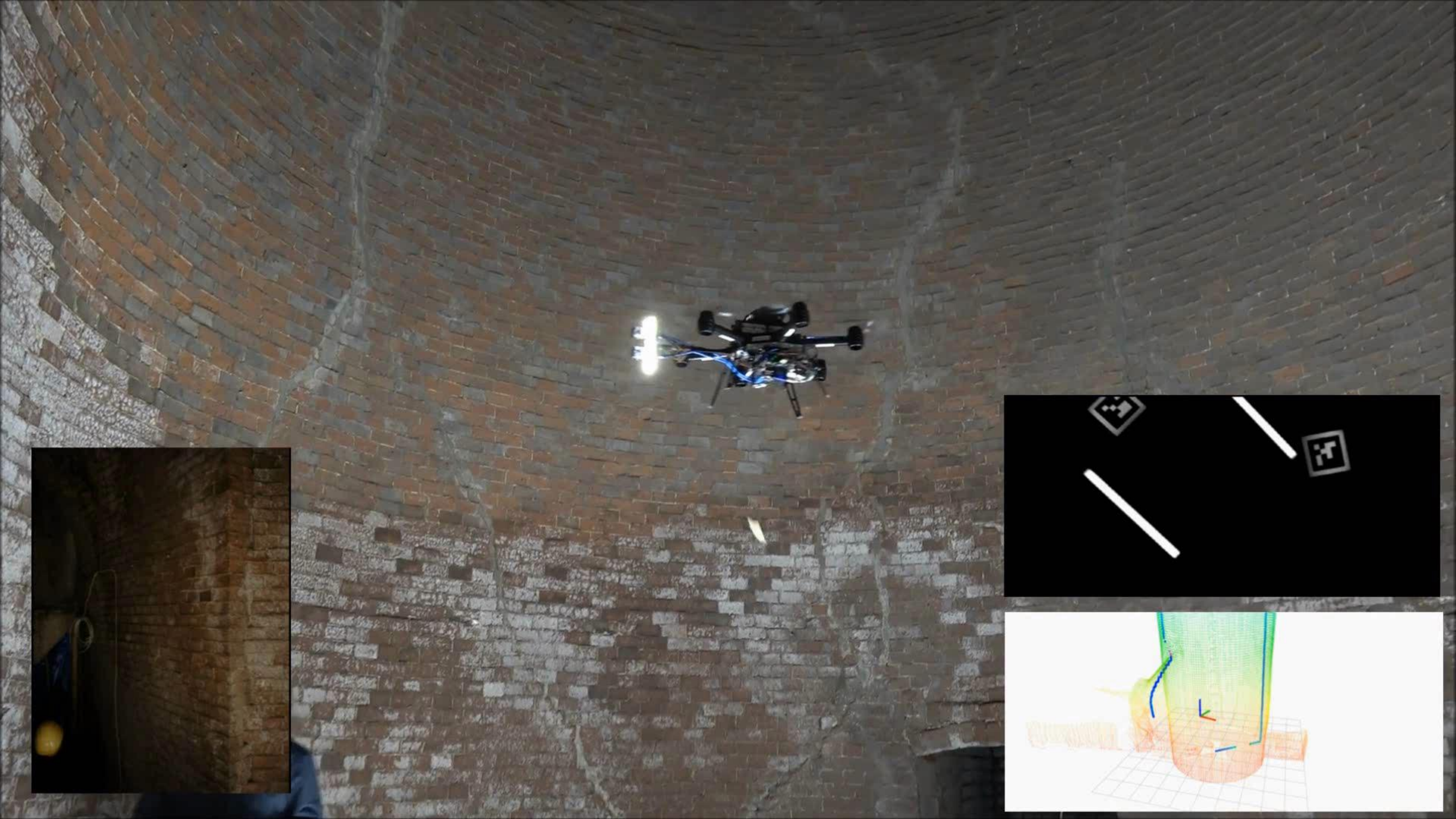
Egocentric planning



Obstacle avoidance

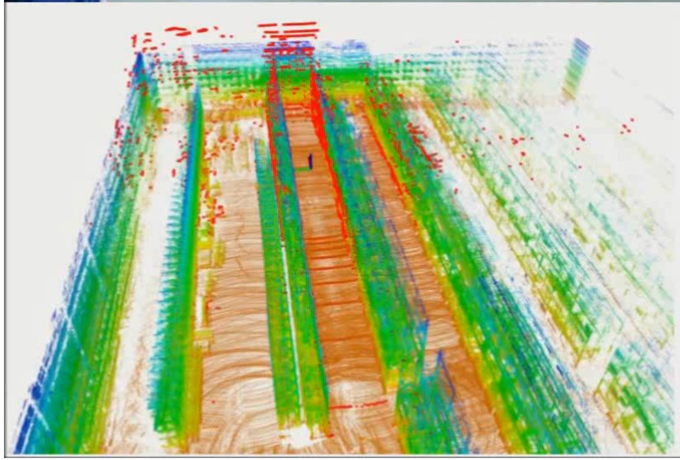
Mapping on Demand

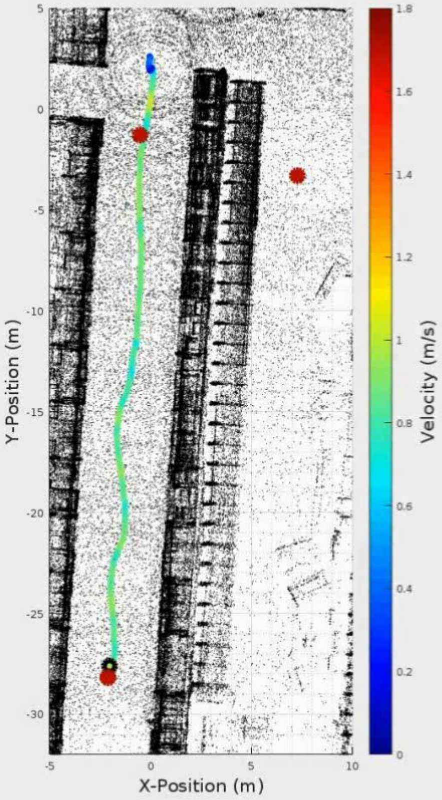
Autonomous Flight to Planned View Poses



DJI Matrice 600 with Velodyne Puck & Cameras

InventAIRy: Autonomous Navigation in a Warehouse

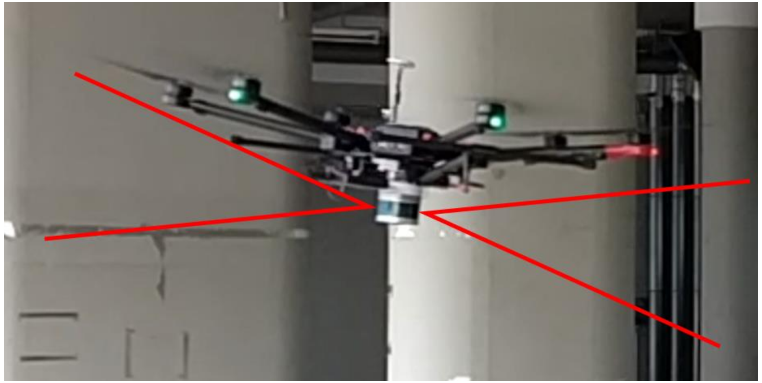




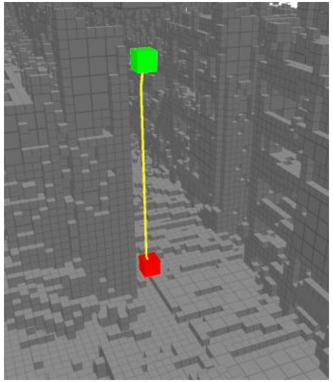
InventAIRy: Detected Tags in Shelf

Navigation Planning with Visibility Constraints

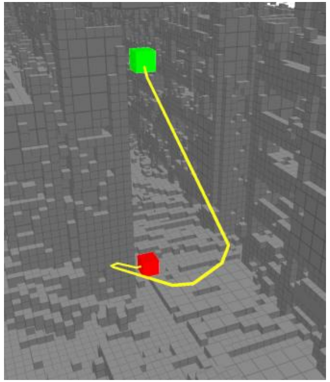
- Velodyne Puck has limited vertical field-of-view (30°)
- Must be considered in navigation planning
- Only fly in directions that can be measured



Lidar field-of-view



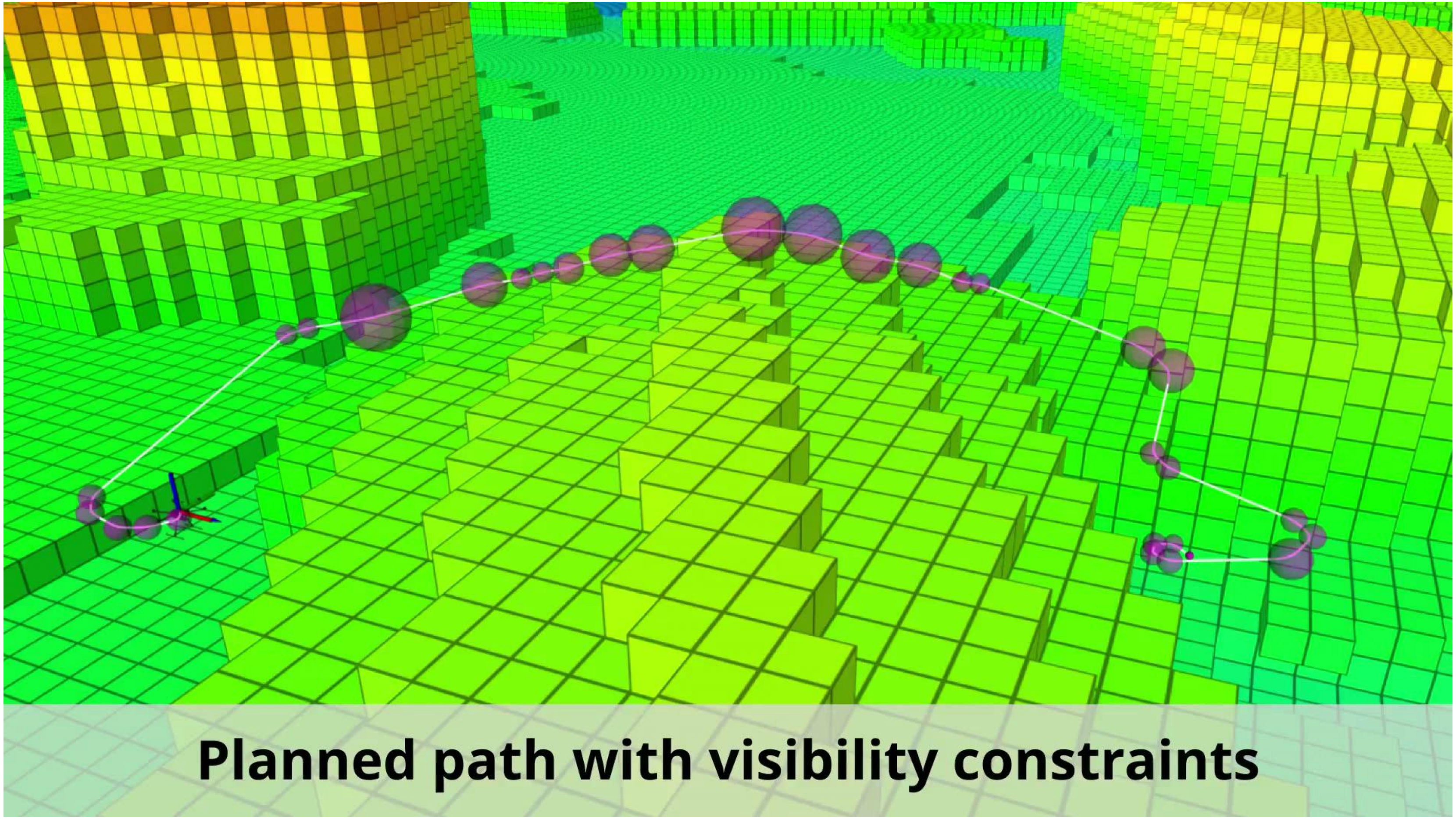
Fastest path



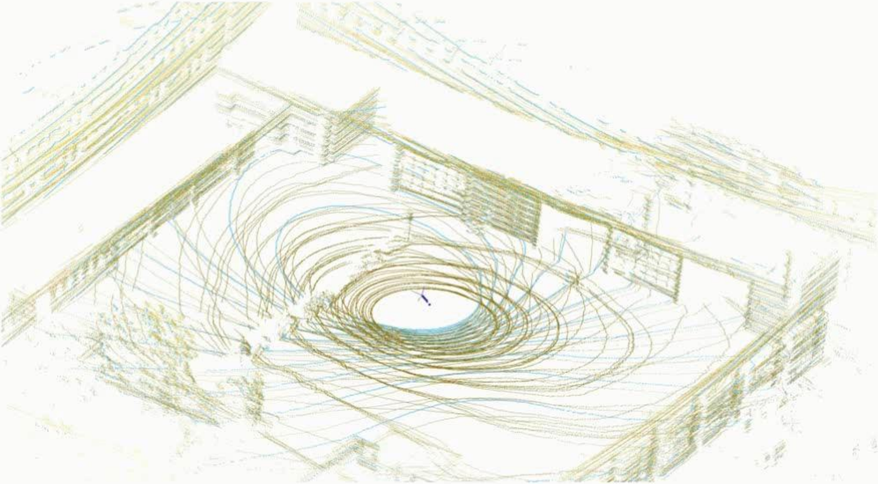
Safe path

[Nieuwenhuisen and Behnke, ICRA 2019]

Navigation Planning with Visibility Constraints



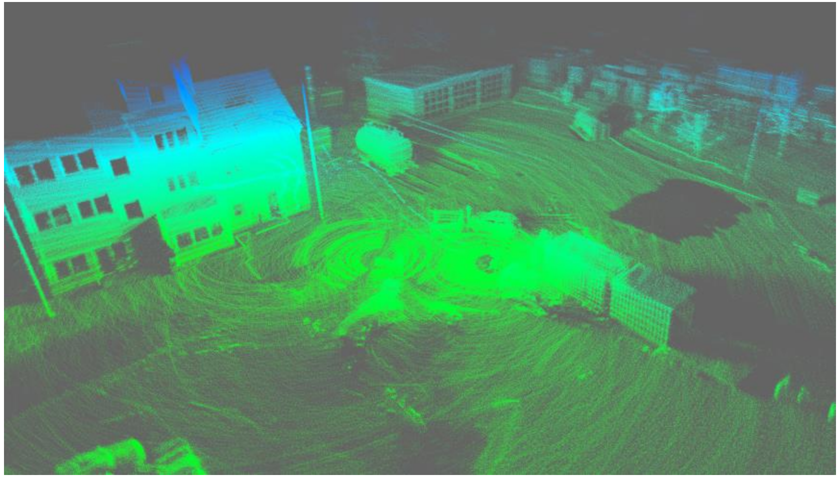
Lidar-based SLAM from MAV



[Droeschel & Behnke, ICRA 2018]

Supporting Fire Fighters (A-DRZ)

- Added thermal camera
- Flight at Brandhaus Dortmund

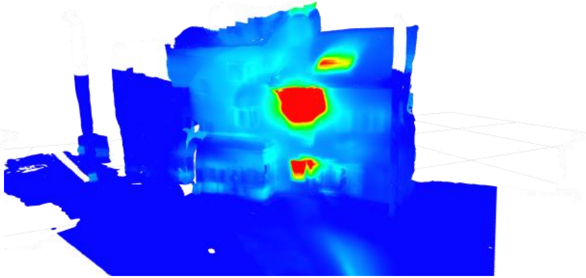


Mesh-based 3D Modeling + Textures

- Model 3D geometry with mesh
- Appearance and temperature as high-resolution texture

Mesh geometry

RGB texture

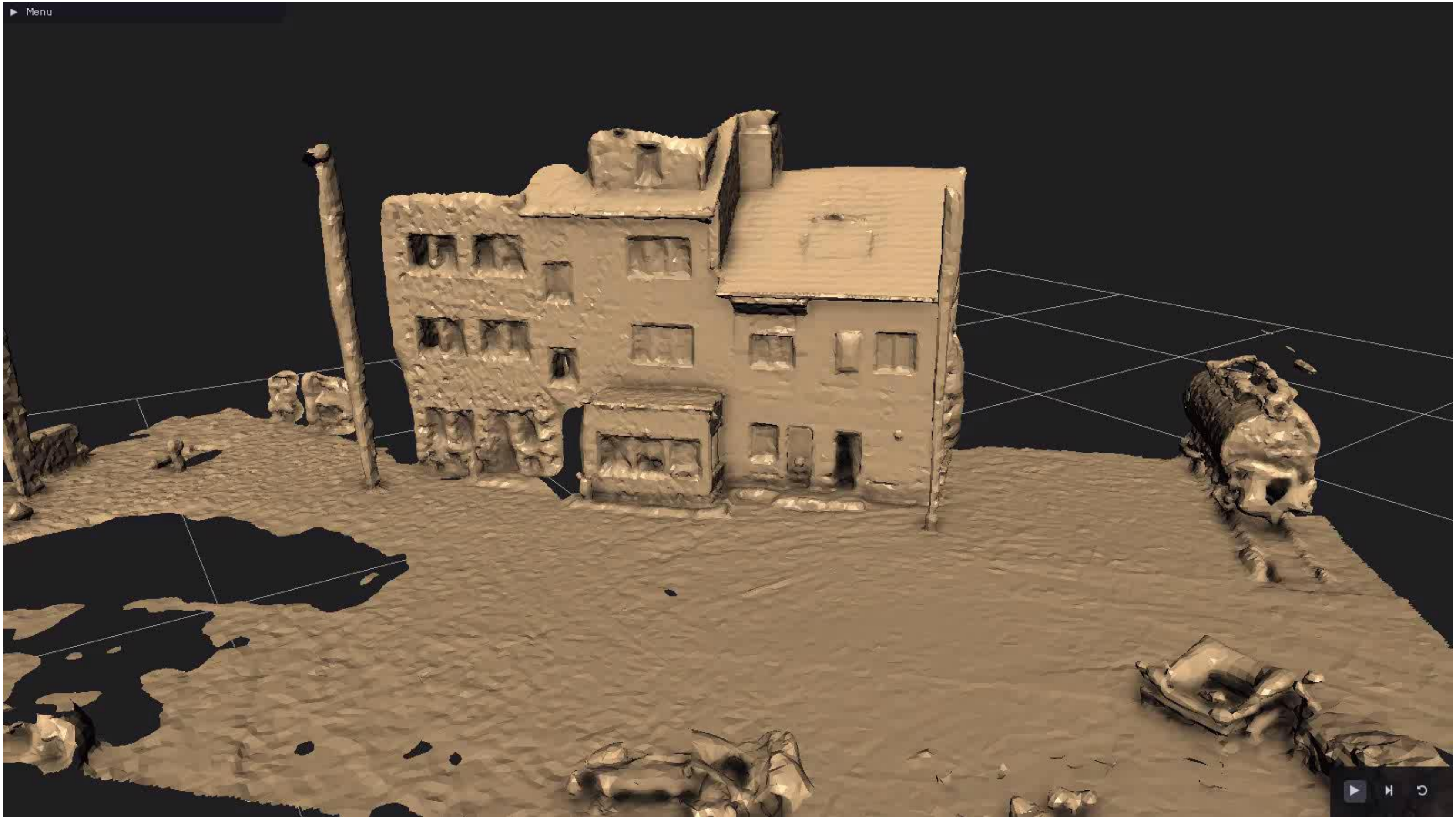


Thermal texture

- Mapping from 3D mesh to 2D texture

Texture mapping

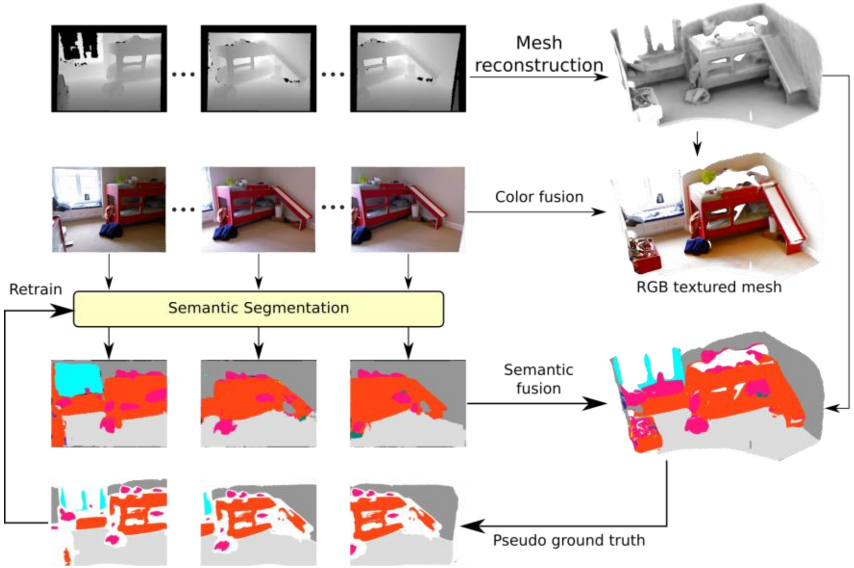
Modeling the Brandhaus Dortmund



UNIVERSITÄT BONN AIS

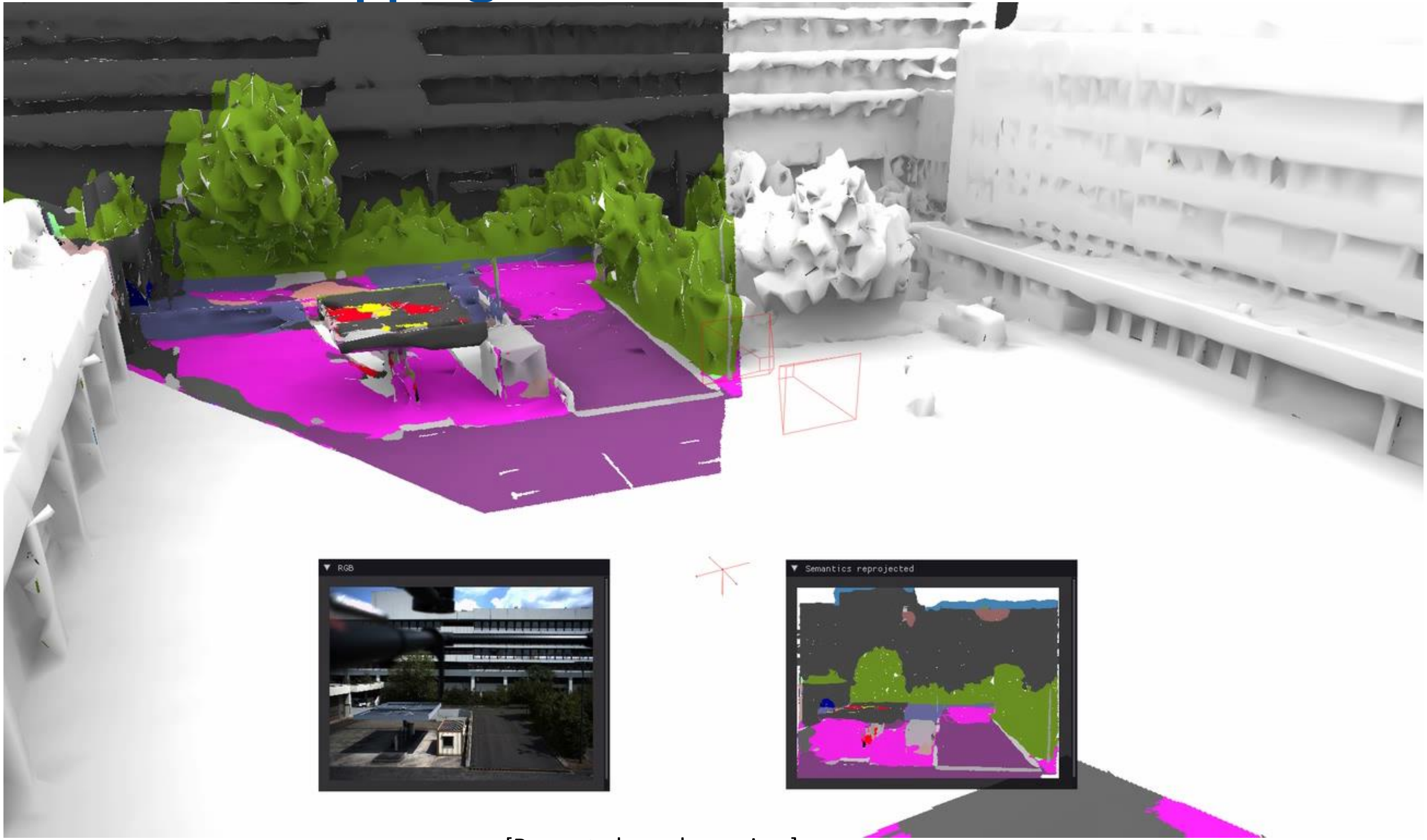
3D Semantic Mapping

- Image-based semantic categorization, trained with Mapillary data set
- 3D fusion in semantic texture
- Backprojection of labels to other views



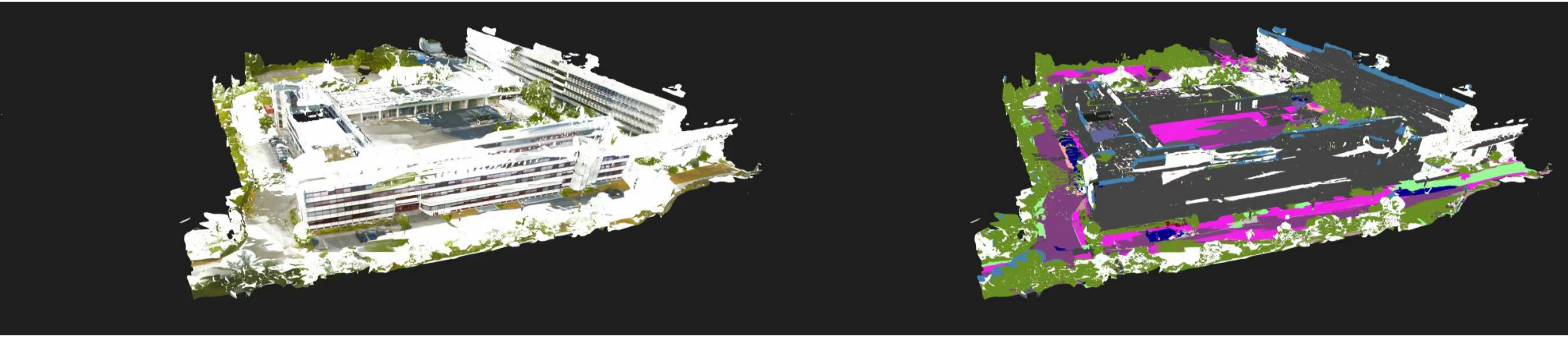
[Rosu et al., under review]

3D Semantic Mapping



[Rosu et al., under review]

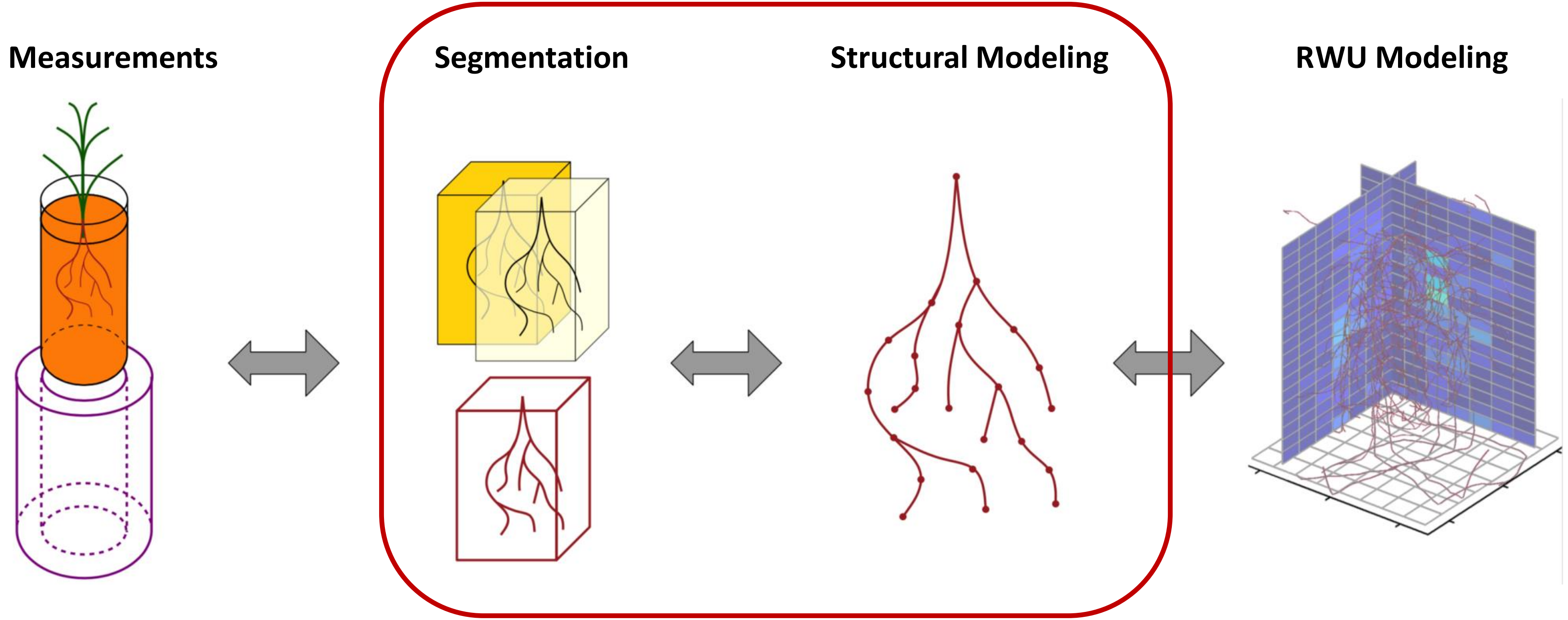
3D Semantic Map



[Rosu et al., under review]

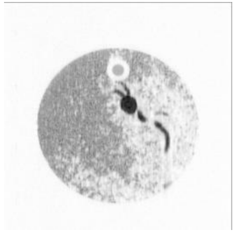
Reconstruction of Plant Roots from MRI

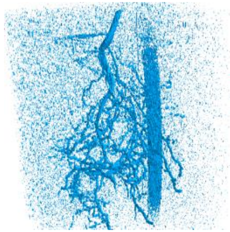
- DFG project with Andrea Schnepf (FZJ)

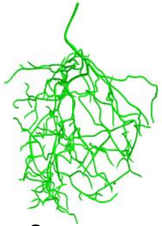


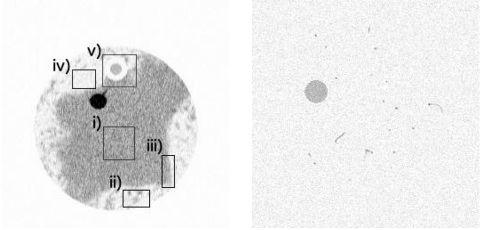
Learning Root vs. Soil Segmentation and Superresolution

- Input: MRI, manual root structure reconstructions
- Desired output: Increased MRI contrast & resolution
- Issues: Few data, reconstructions not well aligned
- Generate synthetic MRI training data
  - Geometric transformations
  - Various noise
- Learn segmentation & superresolution with Deep NN
- Apply to real MRI

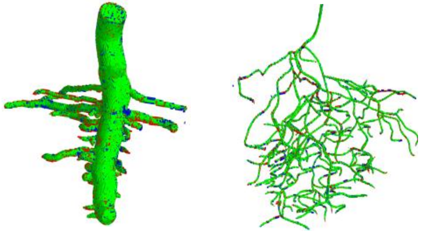

Real MRI slice

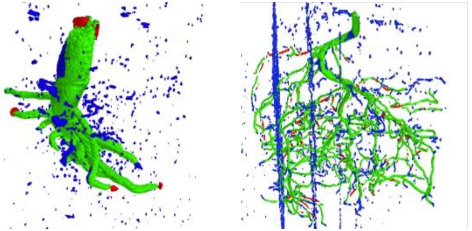

Real MRI 3D


Structure



Synthetic MRI slices

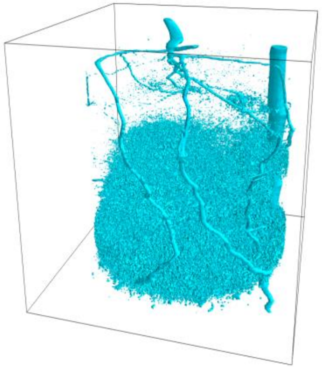

Output for synthetic MRI


Output for real MRI

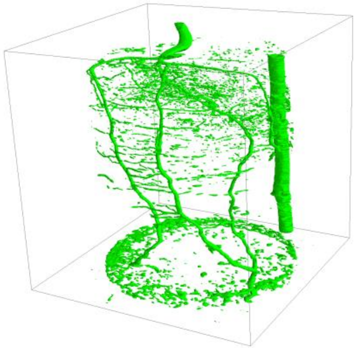
[Uzman et al. ESANN 2019]

Using Learned Model for Root Structure Reconstruction

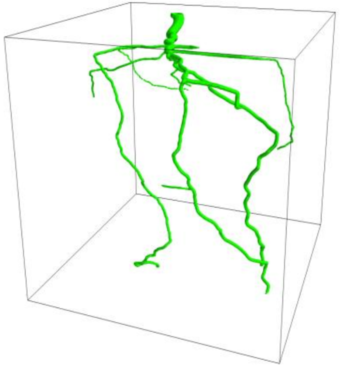
- DAP17



MRI

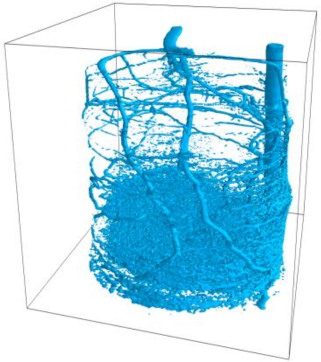


Model output

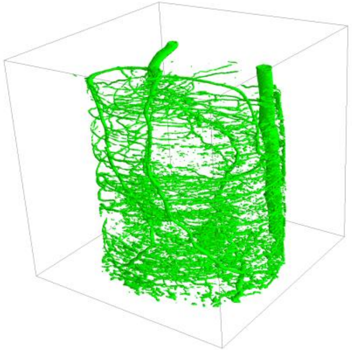


Manual reconstruction

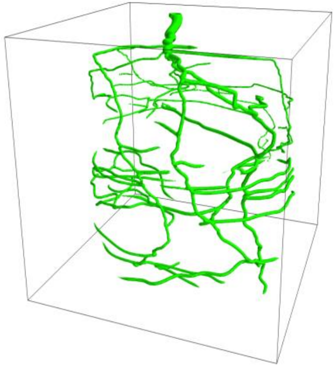
- DAP24



MRI



Model output



Manual reconstruction

Conclusions

- Developed capable robotic systems for challenging scenarios
  - Domestic service
  - Disaster response
  - Aerial inspection
- Autonomy for navigation and manipulation tasks
  - 3D semantic mapping
  - Navigation and manipulation planning
- Use as a tool in PhenoRob, e.g. in
  - CP1: 4D phenotyping of individual plants
  - CP4: Intervention
- Challenges include
  - Correspondences despite growth & deformations
  - Small and big data