Perception, Planning, and Learning for Cognitive Robots

Sven Behnke

University of Bonn
Computer Science Institute VI
Autonomous Intelligent Systems

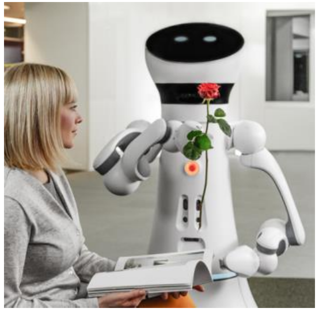
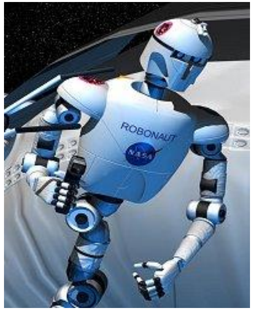
Many New Application Areas for Robots

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

Need more cognitive abilities!






Autonomous Intelligent Systems

UNIVERSITÄT BONN AIS

- Established 2008
- Research in Cognitive Robotics and Computer Vision

Some of our Cognitive Robots

- Equipped with numerous sensors and actuators
- Complex demonstration scenarios

Soccer

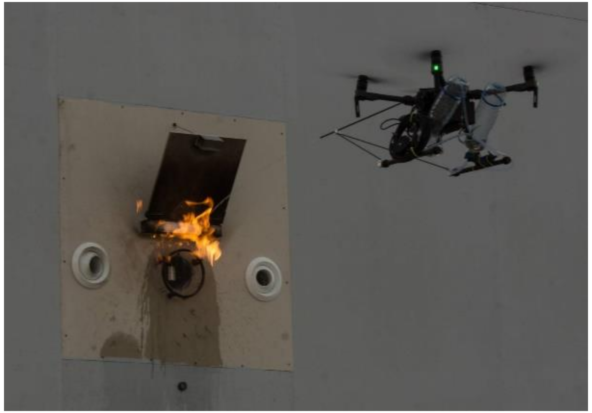
Domestic service

Mobile manipulation

Bin picking

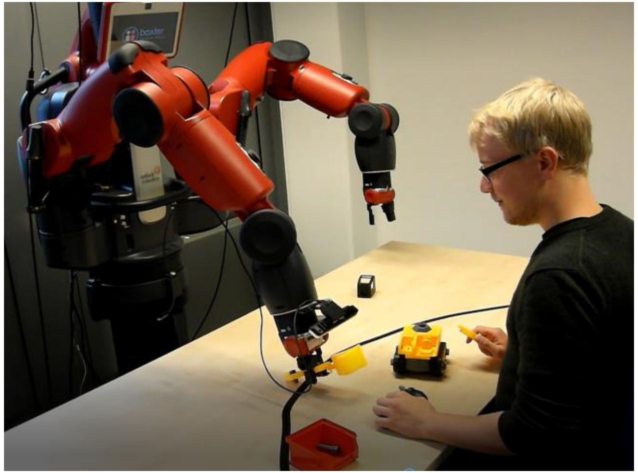
Aerial inspection

Some more of our Cognitive Robots

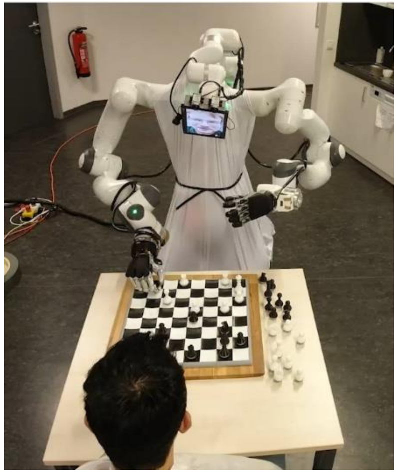


Rescue

Phenotyping

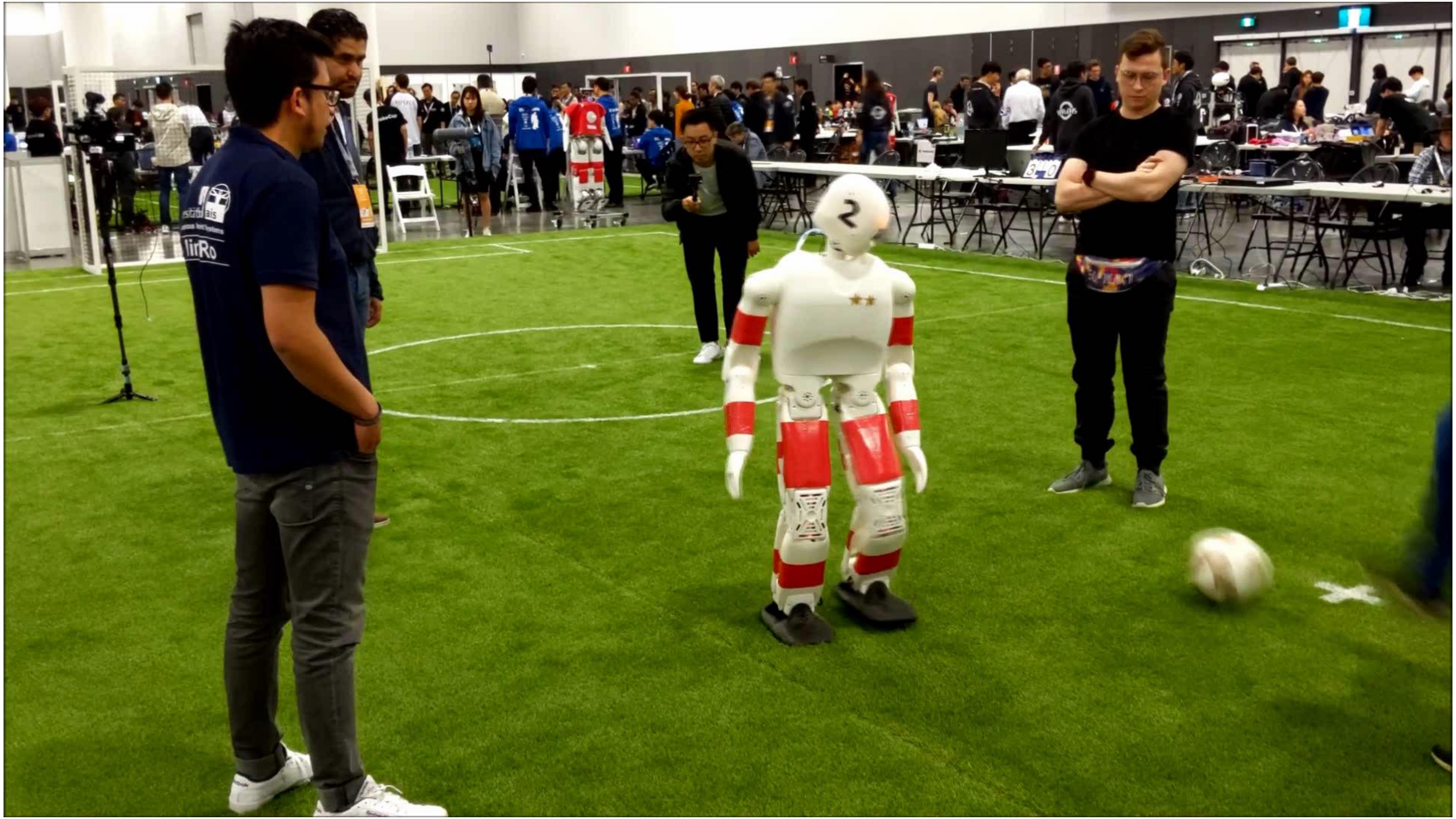


Human-robot collaboration



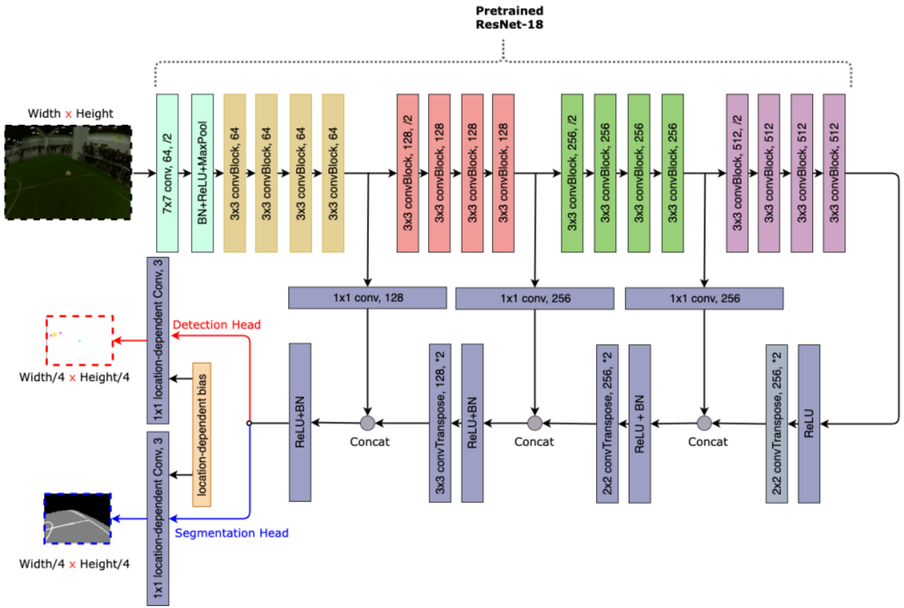
Telepresence

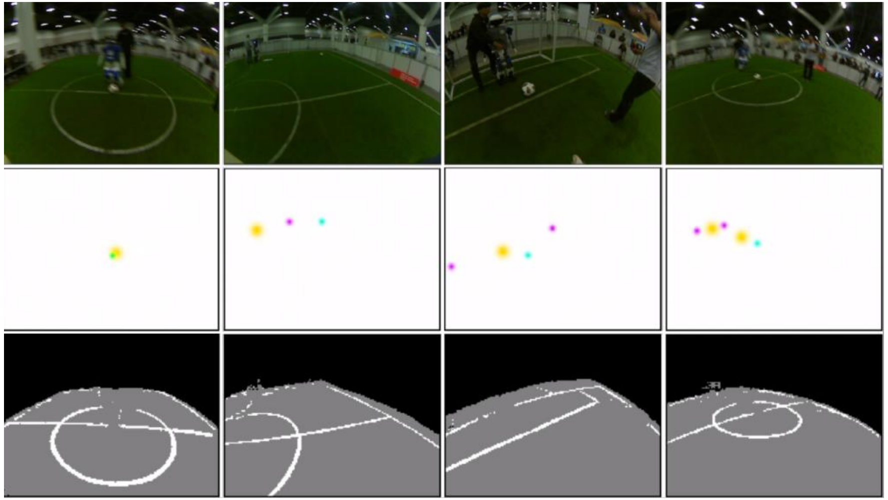
RoboCup 2019 in Sydney



Transfer Learning for Visual Perception

- Encoder-decoder network
- Two outputs
  - Object detection
  - Semantic segmentation
- Location-dependent bias



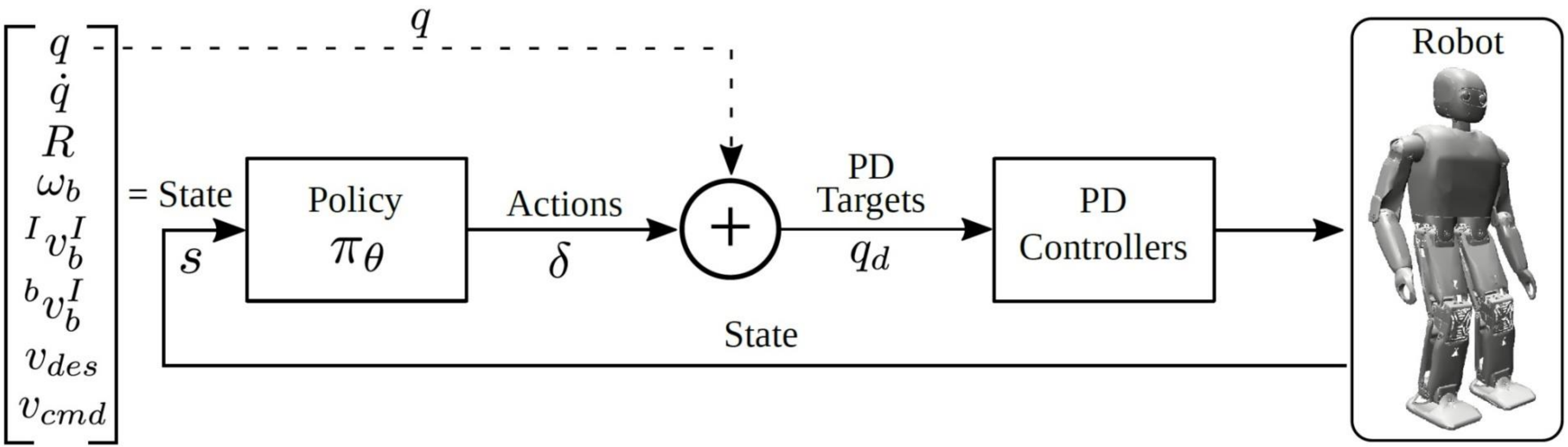


- Detects objects that are hard to recognize for humans
- Robust to lighting changes

[Rodriguez et al. 2019]

Learning Omnidirectional Gait from Scratch

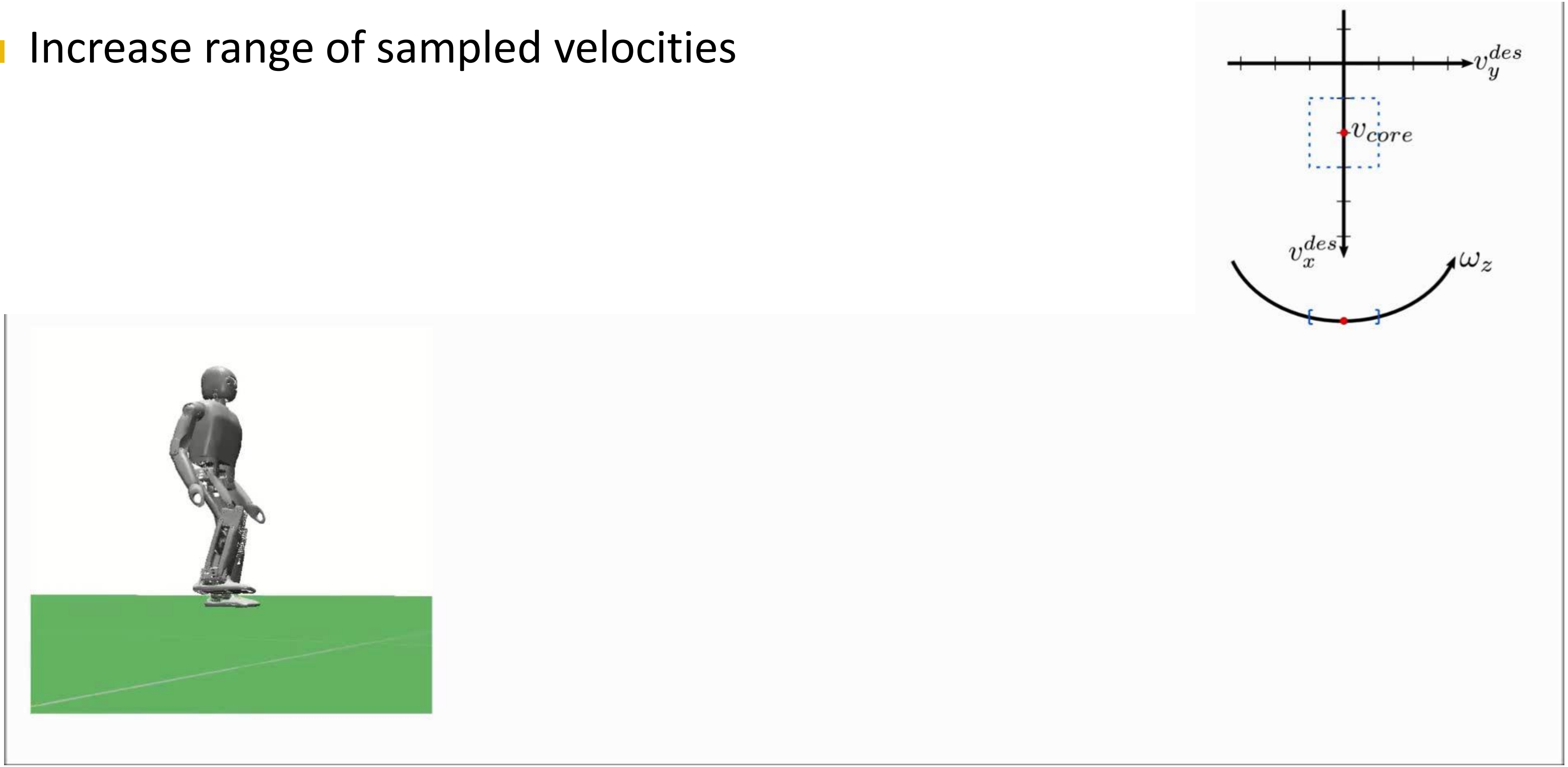
- State includes joint positions and velocities, robot orientation, robot speed
- Actions are increments of joint positions
- Simple reward structure
  - Velocity tracking
  - Pose regularization
  - Not falling



[Rodriguez and Behnke, ICRA 2021]

Learning Curriculum

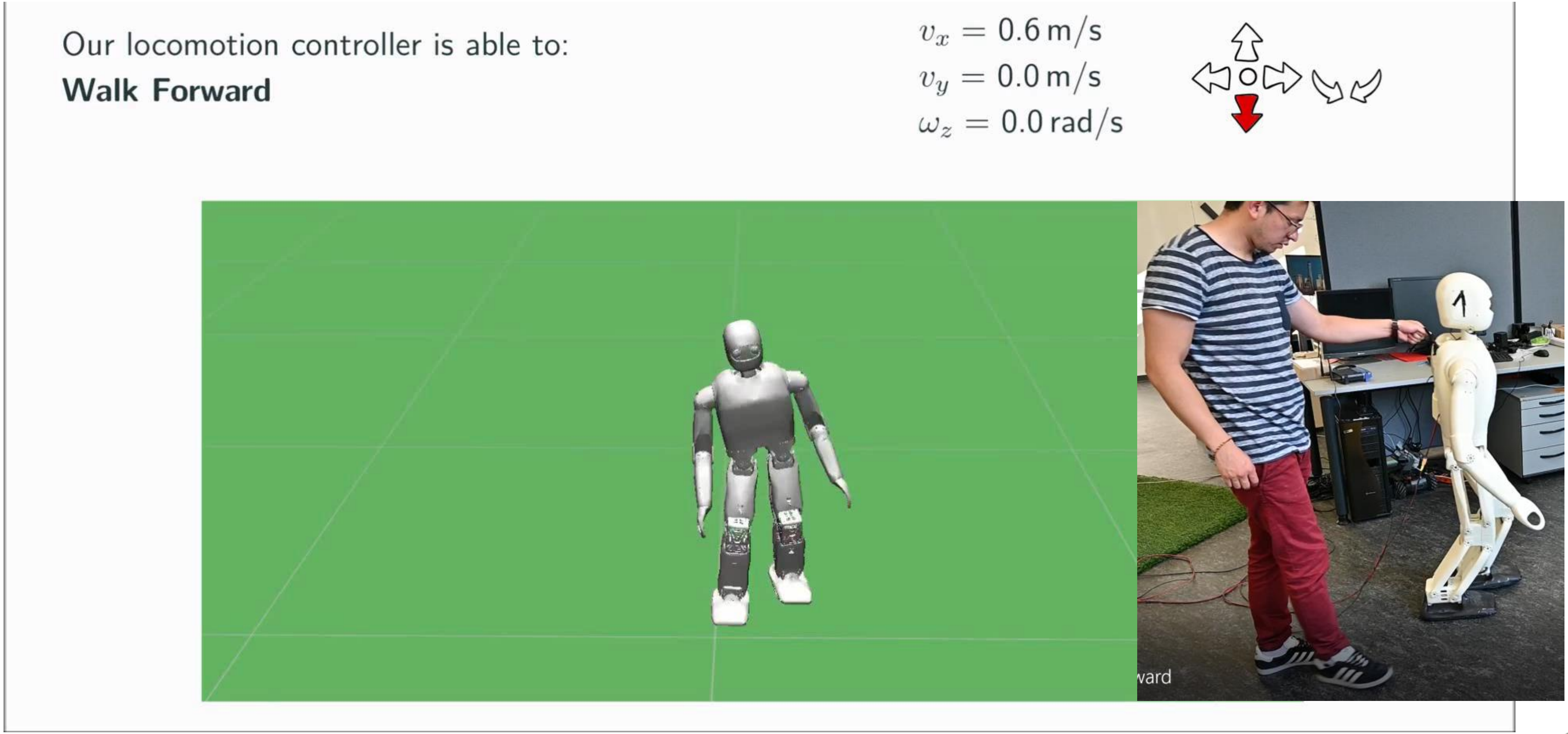
- Start with small velocities
- Increase range of sampled velocities



[Rodriguez and Behnke, ICRA 2021]

Learned Omnidirectional Gait

- Target velocity can be changed continuously

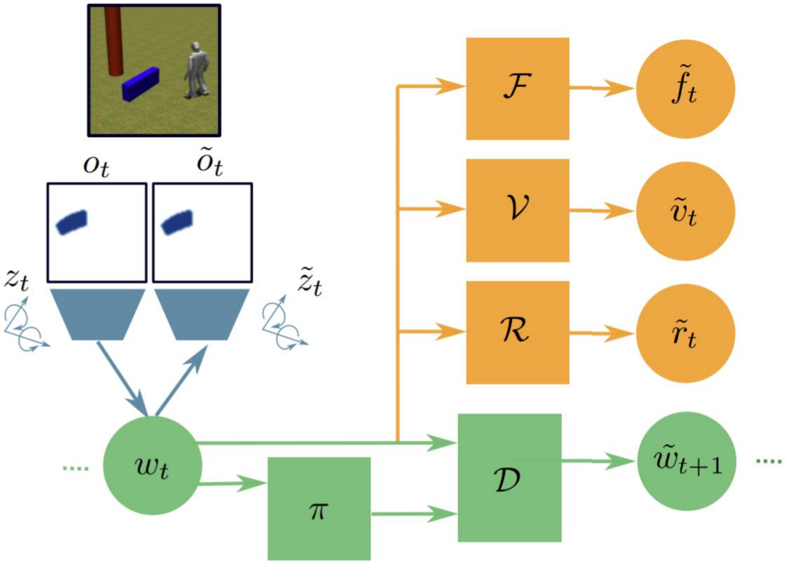


[Rodriguez and Behnke, ICRA 2021]

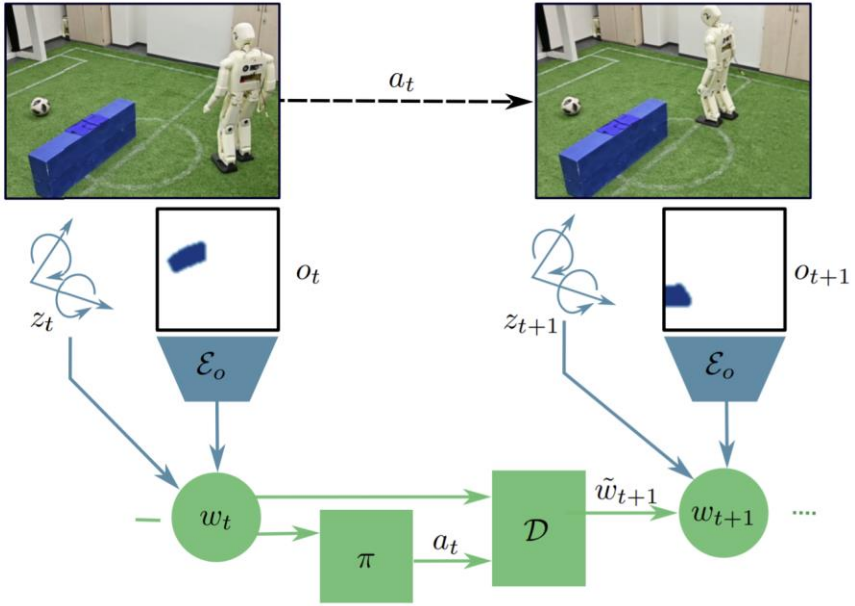
Mapless Humanoid Navigation

- Visual (RGB images) and nonvisual observations to learn a control policy and an environment dynamics model, extends Dreamer [Hafner et al. ICLR 2020]
- Anticipate terminal states of success and failure

Training



Inference



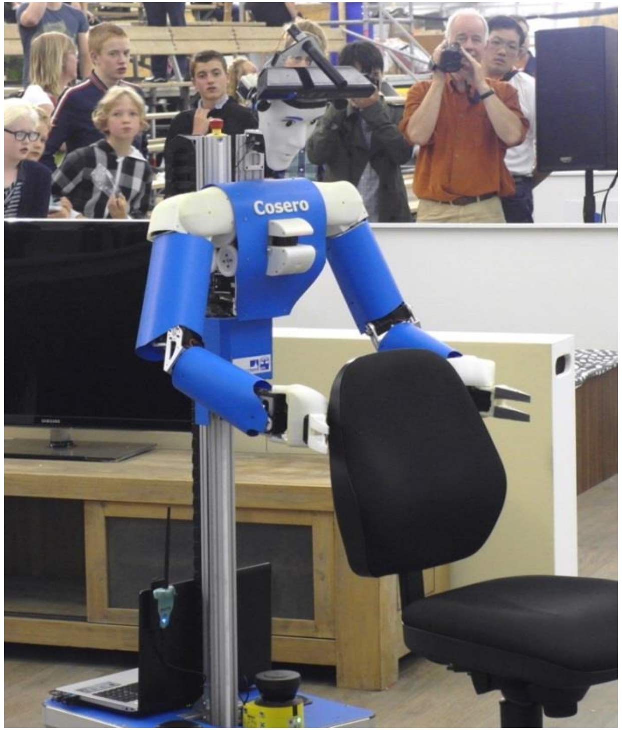
[Brandenburger et al. IROS 2021]

Mapless Humanoid Navigation

[Brandenburger et al. IROS 2021]

Our Domestic Service Robots

Dynamaid



Cosero

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

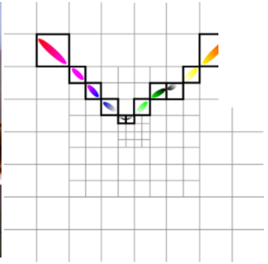
[Stückler et al.: Frontiers in Robotics and AI 2016]

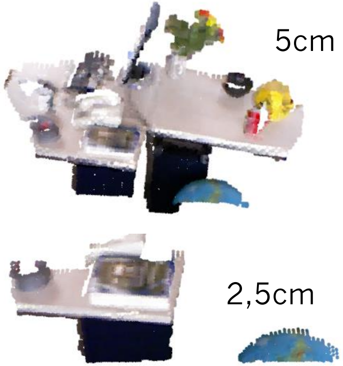
Cognitive Service Robot Cosero

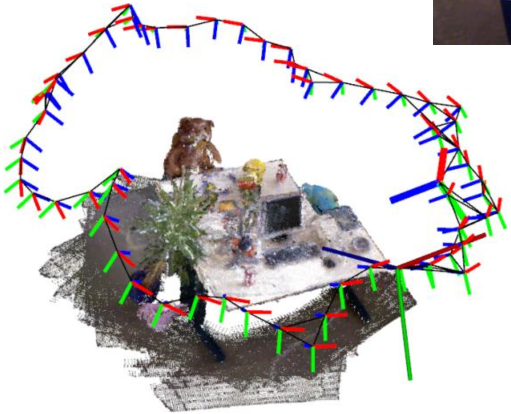
3D Mapping by RGB-D SLAM

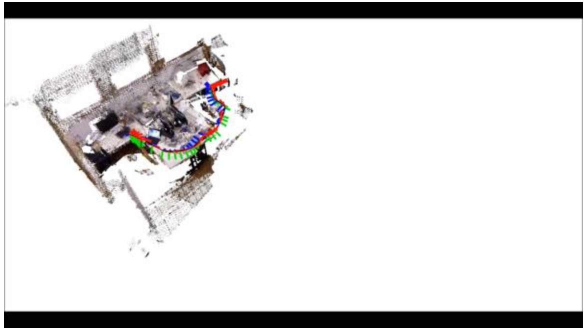
[Stückler, Behnke:
Journal of Visual Communication
and Image Representation 2013]

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

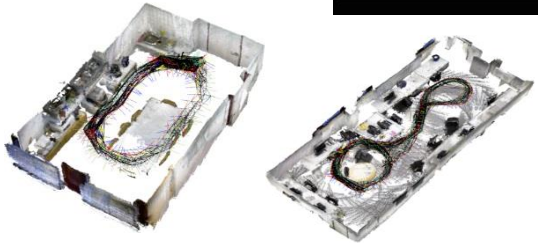








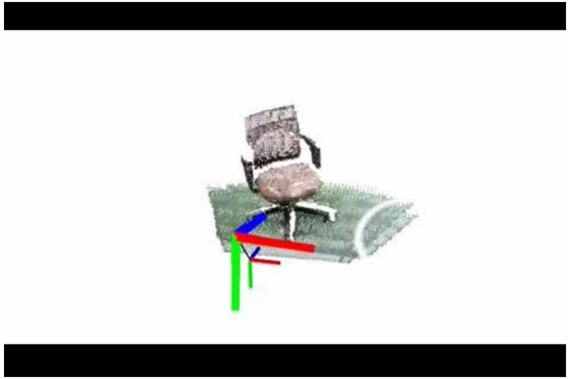
- Multi-camera SLAM



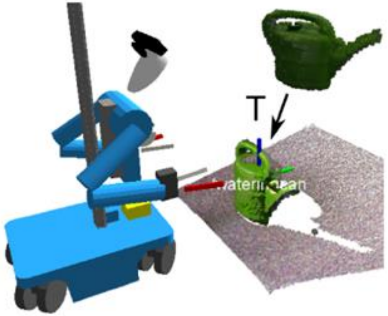
[Stoucken]

Learning and Tracking Object Models

- Modeling of objects by RGB-D-SLAM

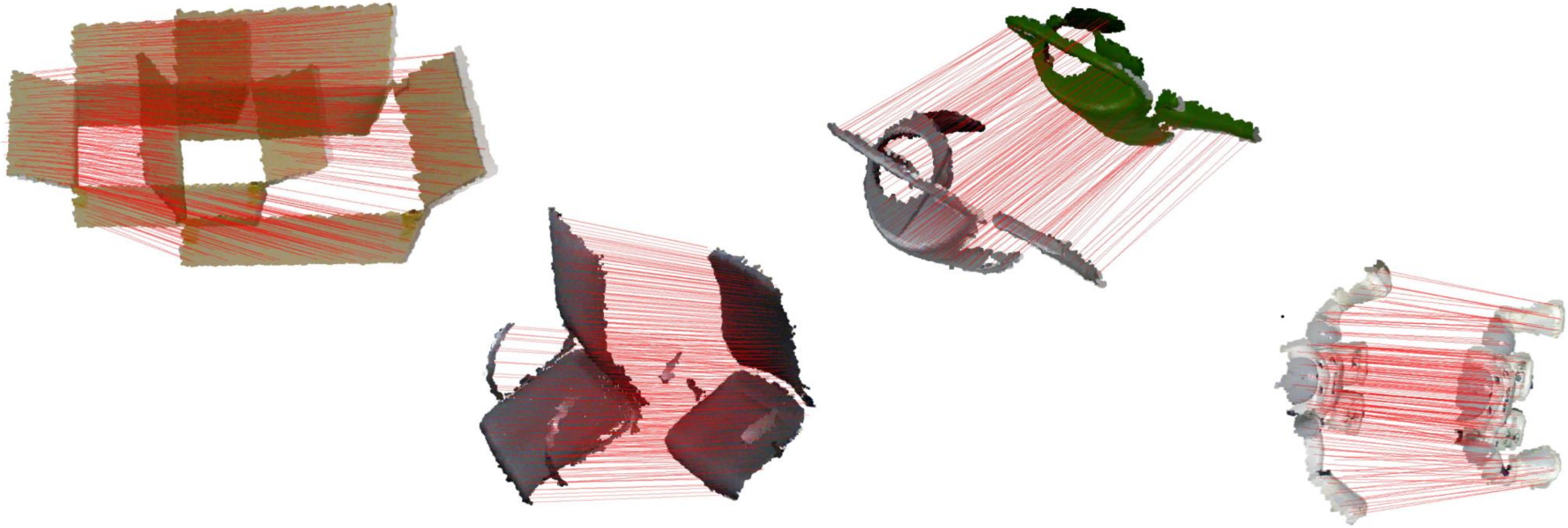



- Real-time registration with current RGB-D frame

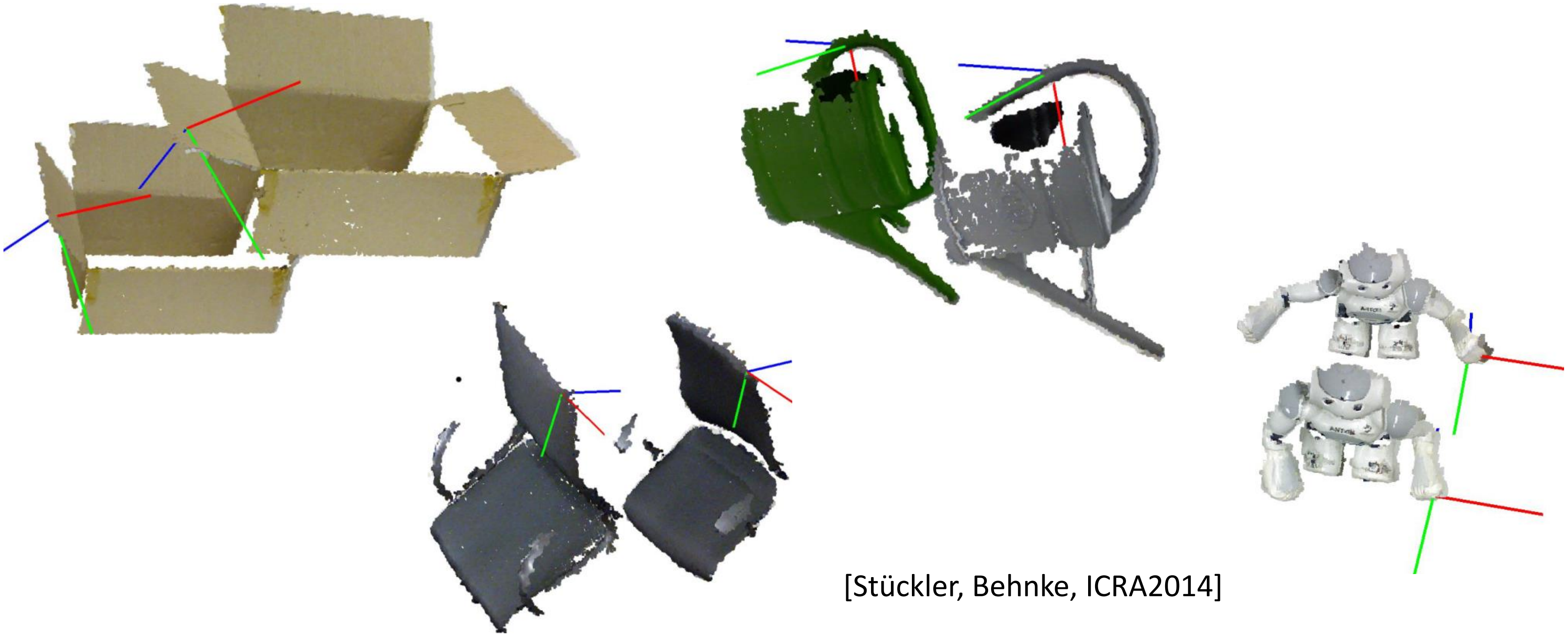
Deformable RGB-D-Registration

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



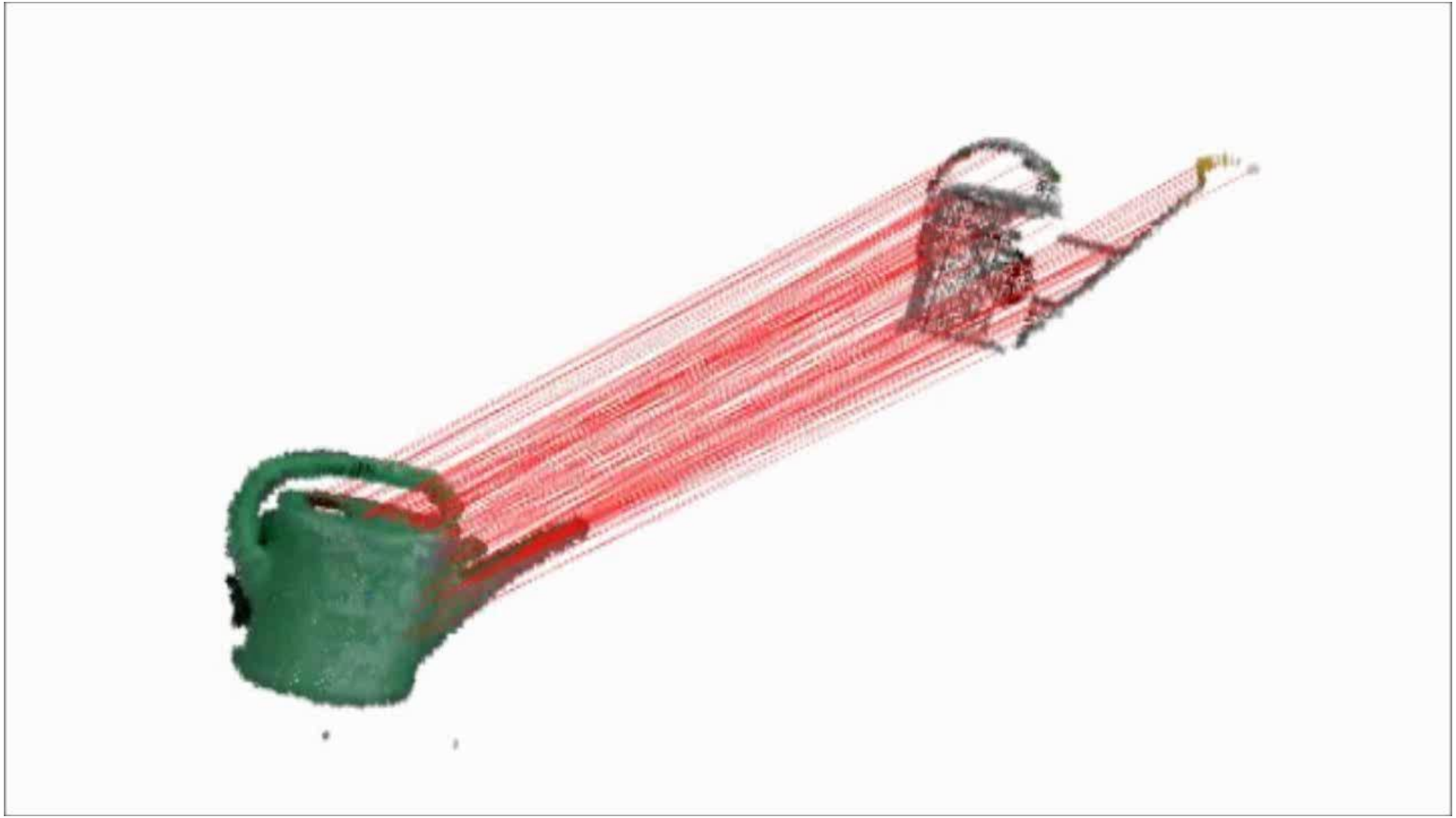
Transformation of Poses on Object

- Derived from the deformation field



[Stückler, Behnke, ICRA2014]

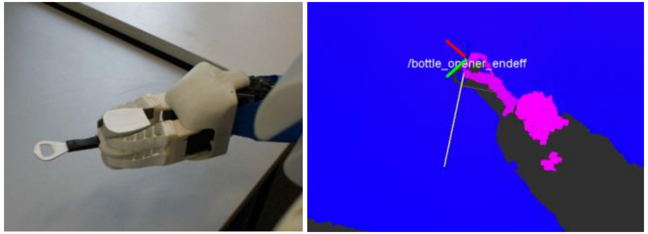
Grasp & Motion Skill Transfer



[Stückler,
Behnke,
ICRA2014]

Tool use: Bottle Opener

- Tool tip perception

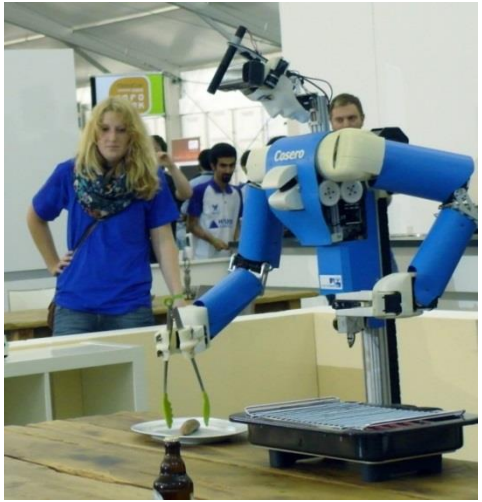


- Extension of arm kinematics
- Perception of crown cap
- Motion adaptation

[Stückler, Behnke, Humanoids 2014]

Picking Sausage, Bimanual Transport

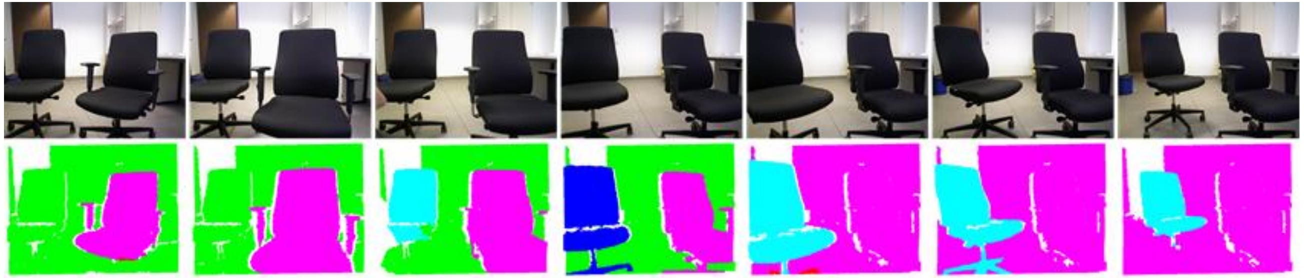
- Perception of tool tip and sausage
- Alignment with main axis of sausage



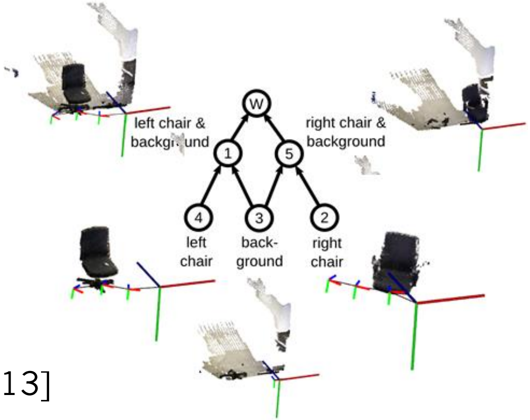
- Our team NimbRo won the RoboCup@Home League in three consecutive years

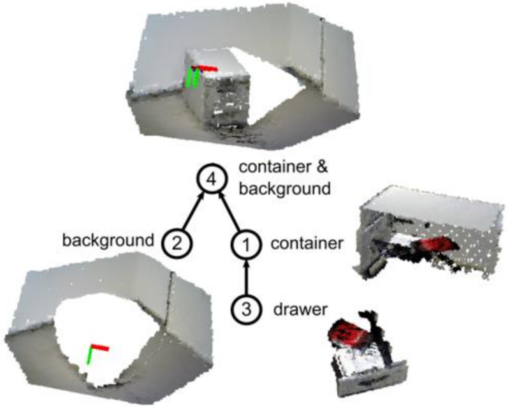
Hierarchical Object Discovery trough Motion Segmentation

- Simultaneous object modeling and motion segmentation



- Inference of a segment hierarchy

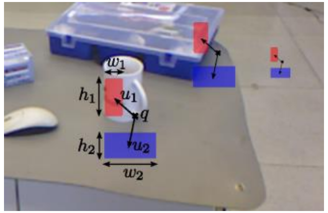


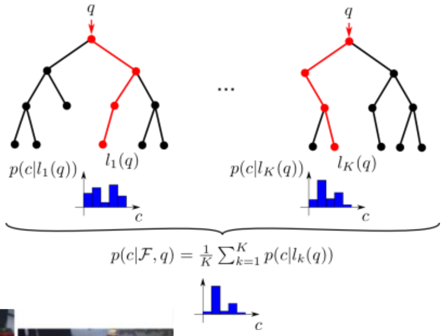


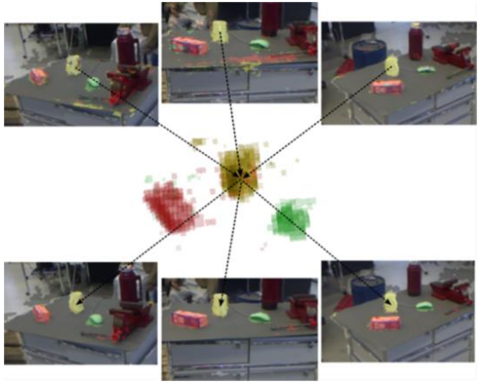
[Stückler, Behnke: IJCAI 2013]

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]

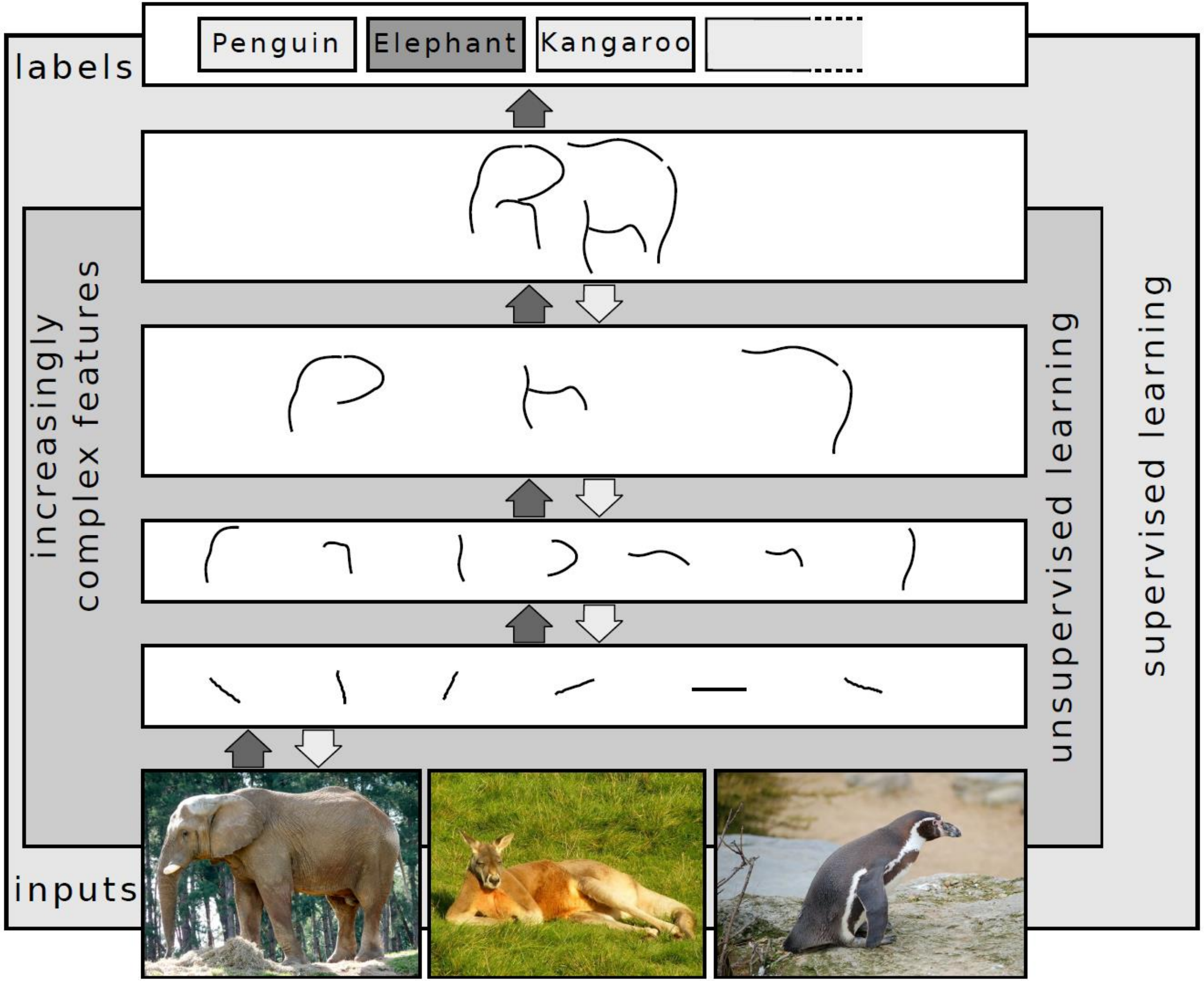
Ground truth

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



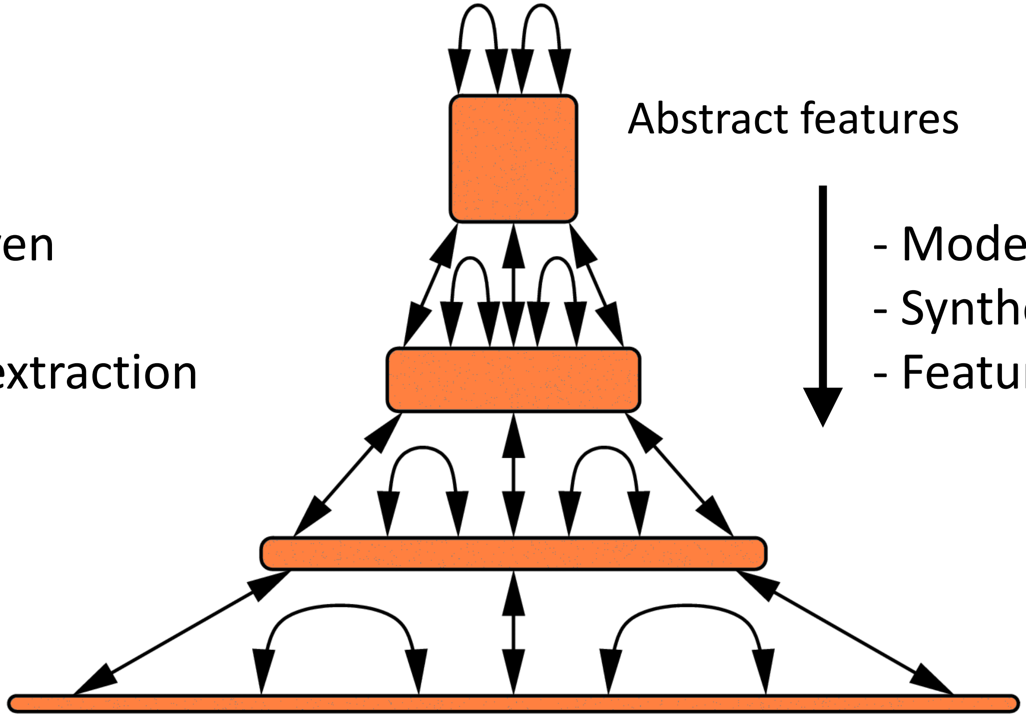
[Schulz; Behnke, KI 2012]

Neural Abstraction Pyramid

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

Abstract features

- Data-driven
- Analysis
- Feature extraction



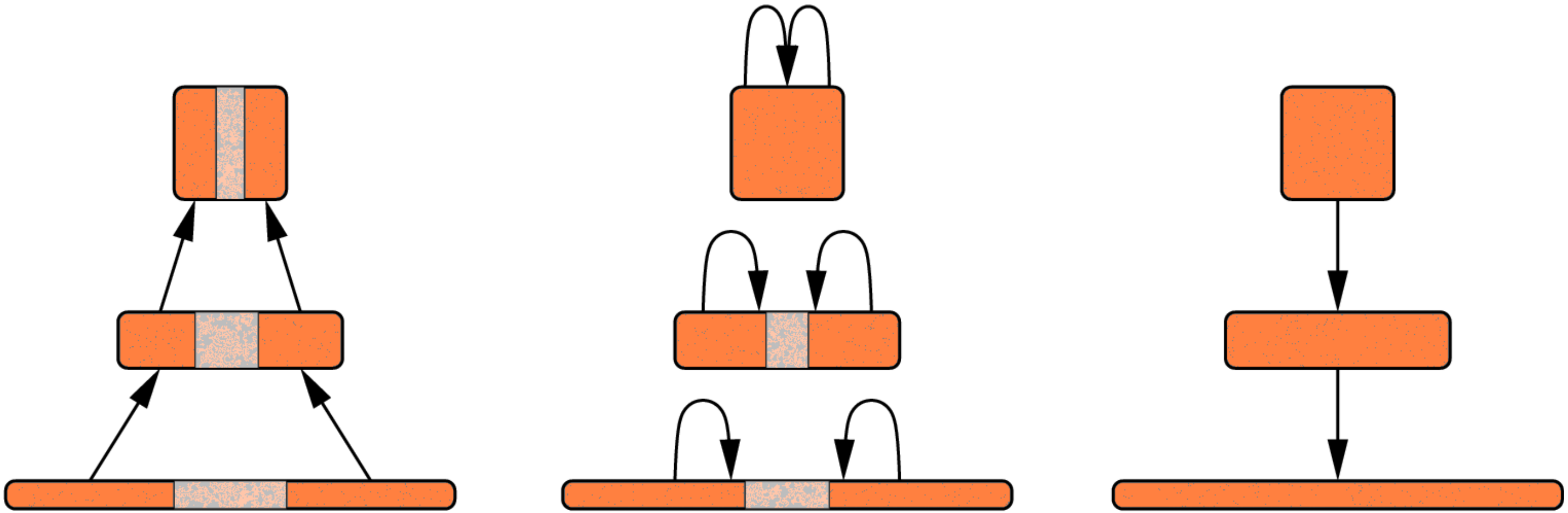
- Model-driven
- Synthesis
- Feature expansion

Signals

- Grouping - Competition - Completion

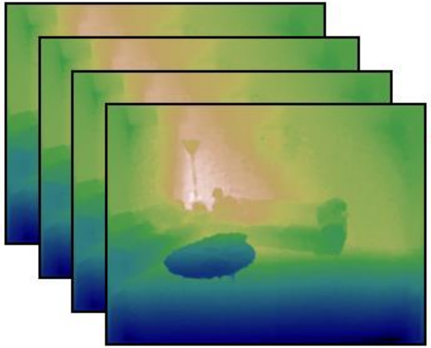
Iterative Image Interpretation

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

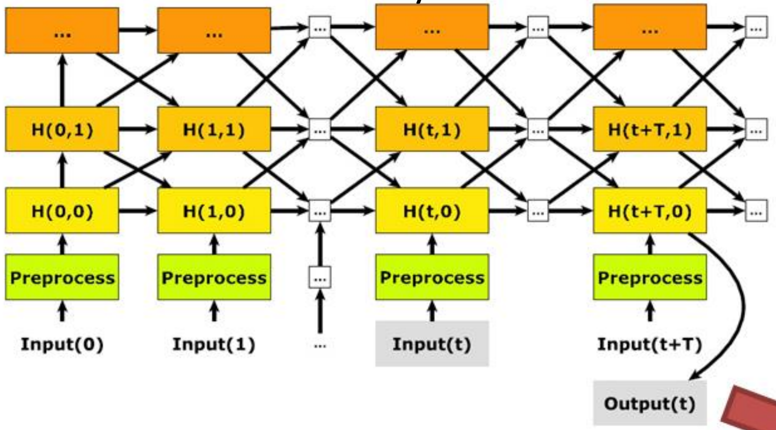


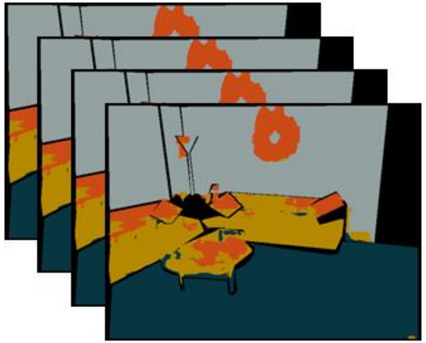
Neural Abstraction Pyramid for Object-class Segmentation of RGB-D Video

- Recursive computation is efficient for temporal integration



Neural Abstraction Pyramid





[Pavel, Schulz, Behnke, Neural Networks 2017]

The Data Problem

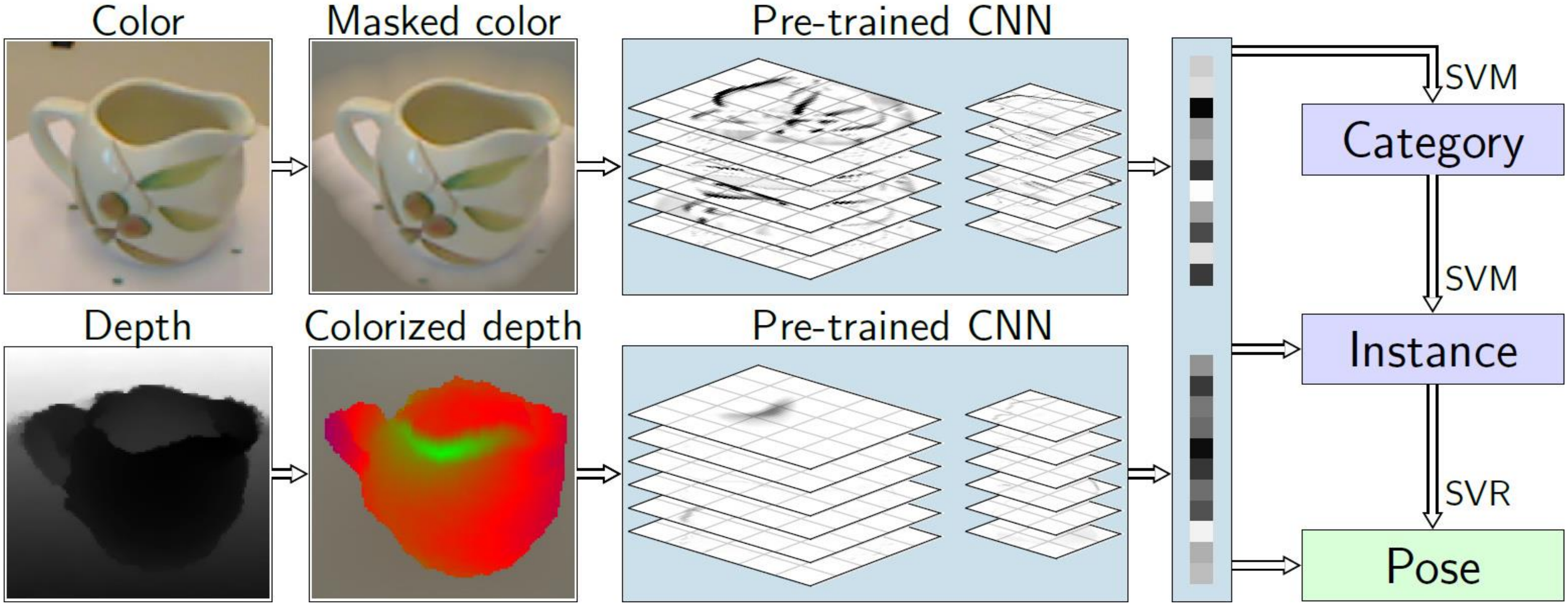
- Deep Learning in robotics (still) suffers from shortage of available examples
- We address this problem in two ways:

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

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

RGB-D Object Recognition and Pose Estimation

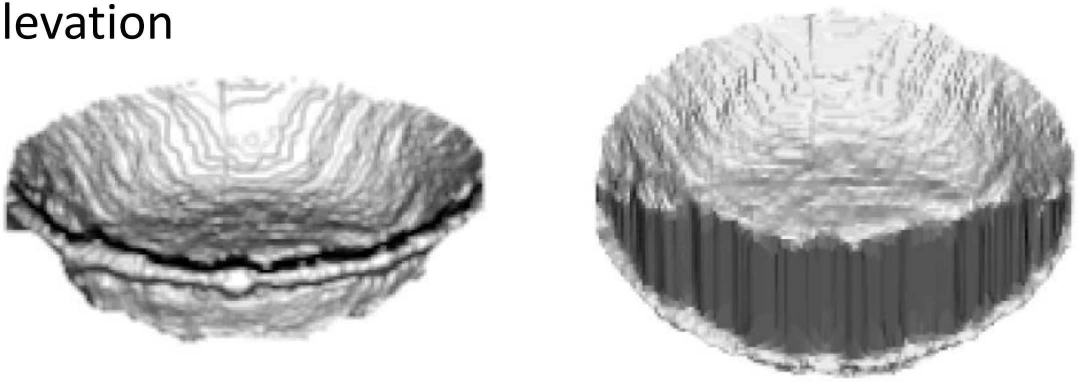
- Transfer learning from large-scale data sets



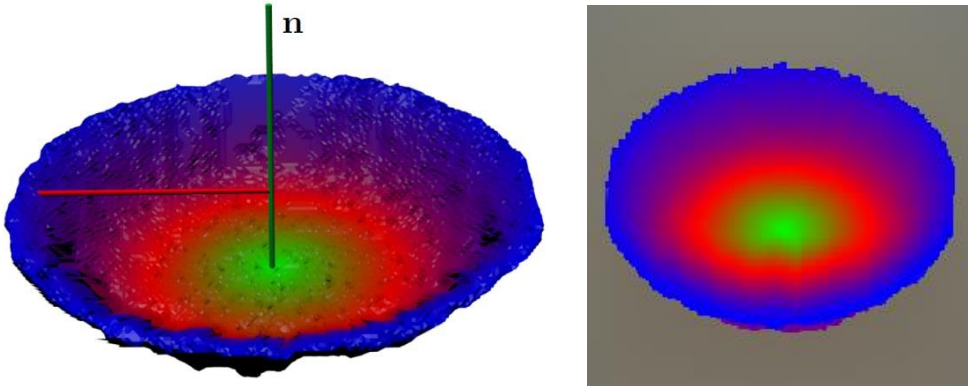
[Schwarz, Schulz, Behnke, ICRA2015]

Canonical View, Colorization

- Objects viewed from different elevation
- Render canonical view



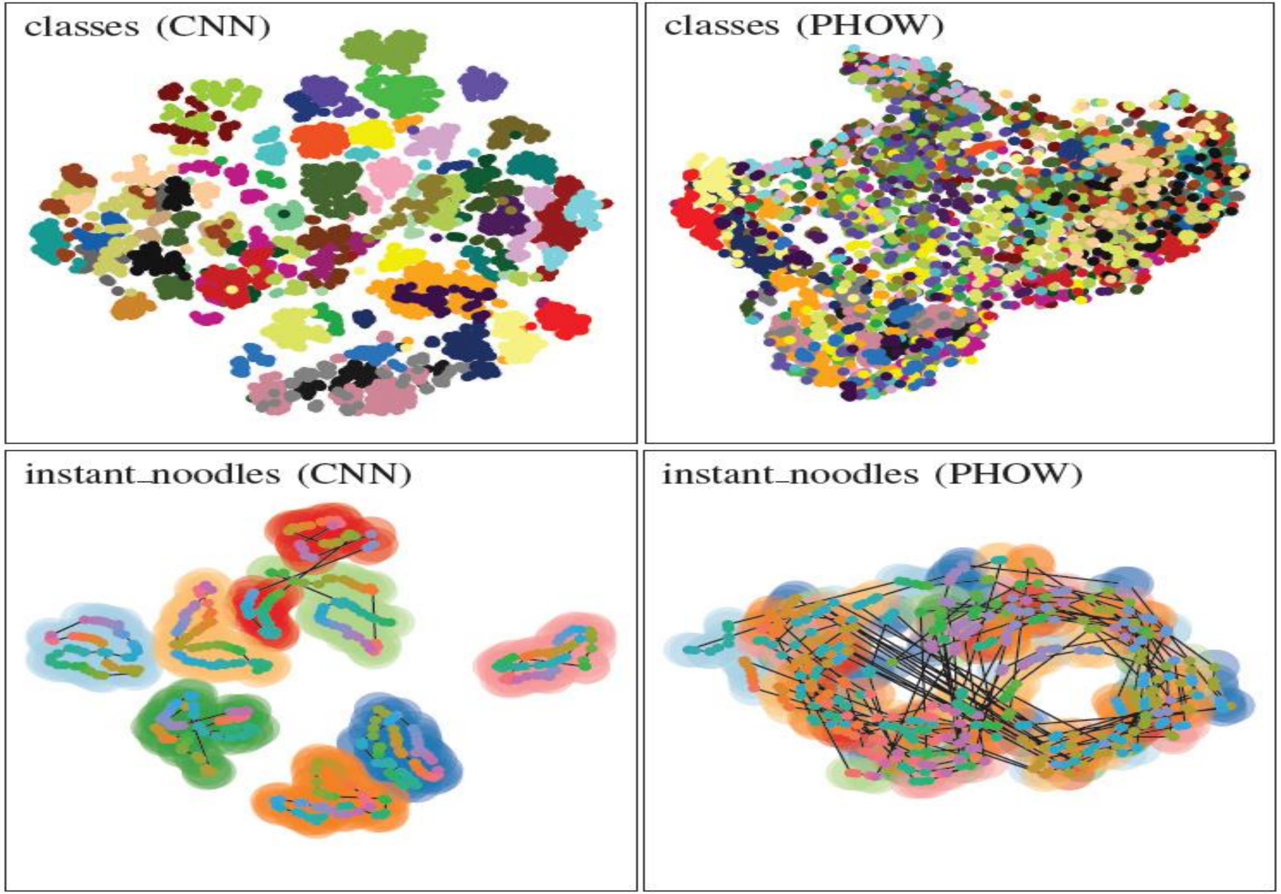
- Colorization based on distance from center vertical



[Schwarz, Schulz, Behnke, ICRA2015]

Pretrained Features Disentangle Data

- t-SNE embedding



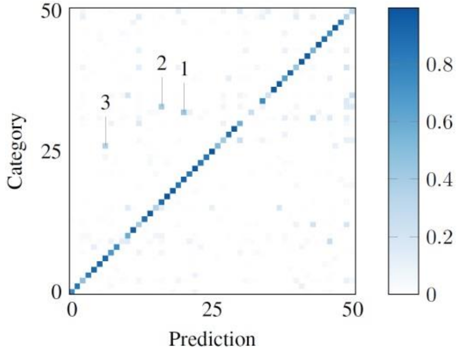
[Schwarz, Schulz, Behnke ICRA2015]

Recognition Accuracy

- Improved both category and instance recognition

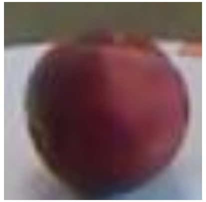
Method	Category Accuracy (%) RGB	Category Accuracy (%) RGB-D	Instance Accuracy (%) RGB	Instance Accuracy (%) RGB-D
Lai *et al.* [1]	74.3 ± 3.3	81.9 ± 2.8	59.3	73.9
Bo *et al.* [2]	82.4 ± 3.1	87.5 ± 2.9	**92.1**	92.8
PHOW[3]	80.2 ± 1.8	—	62.8	—
Ours	$\mathbf{83.1 \pm 2.0}$	88.3 ± 1.5	92.0	**94.1**
Ours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

- Confusion:



1: pitcher / coffe mug

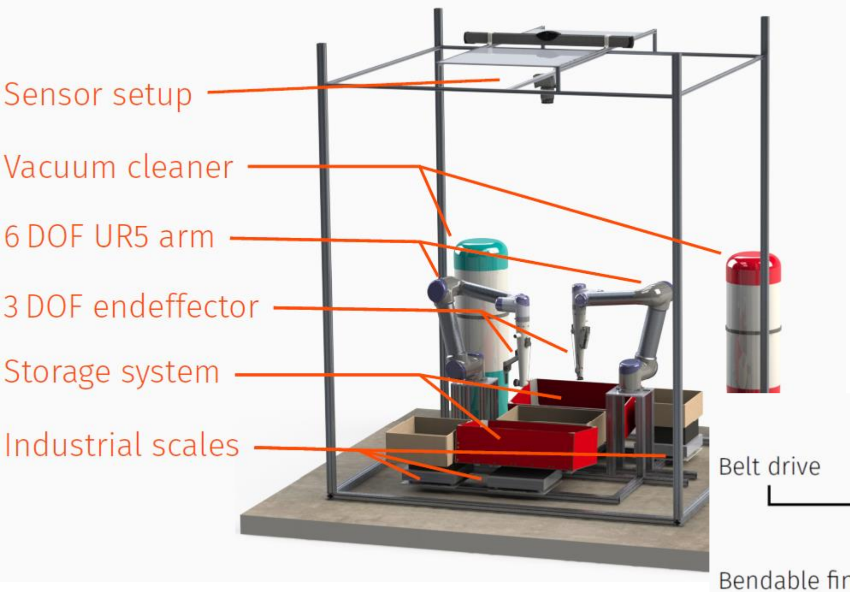
2: peach / sponge

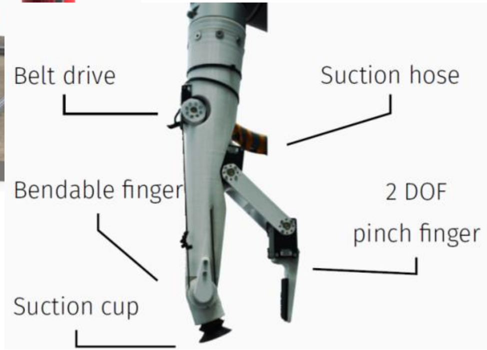


[Schwarz, Schulz, Behnke, ICRA2015]

Amazon Robotics Challenge

- Storing and picking of items
- Dual-arm robotic system





[Schwarz et al. ICRA 2018]

[Amazon]

Object Capture and Scene Rendering

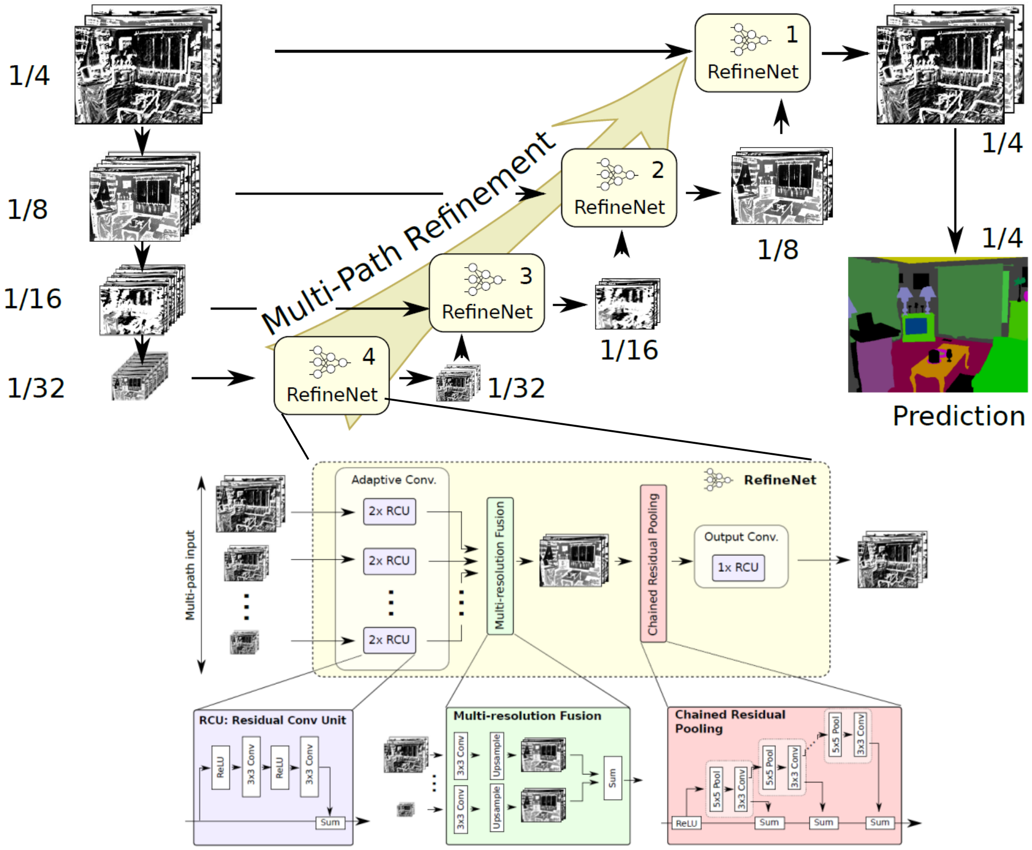
- Turntable + DLSR camera

- Insertion in complex annotated scenes

[Schwarz et al. ICRA 2018]

RefineNet for Semantic Segmentation

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

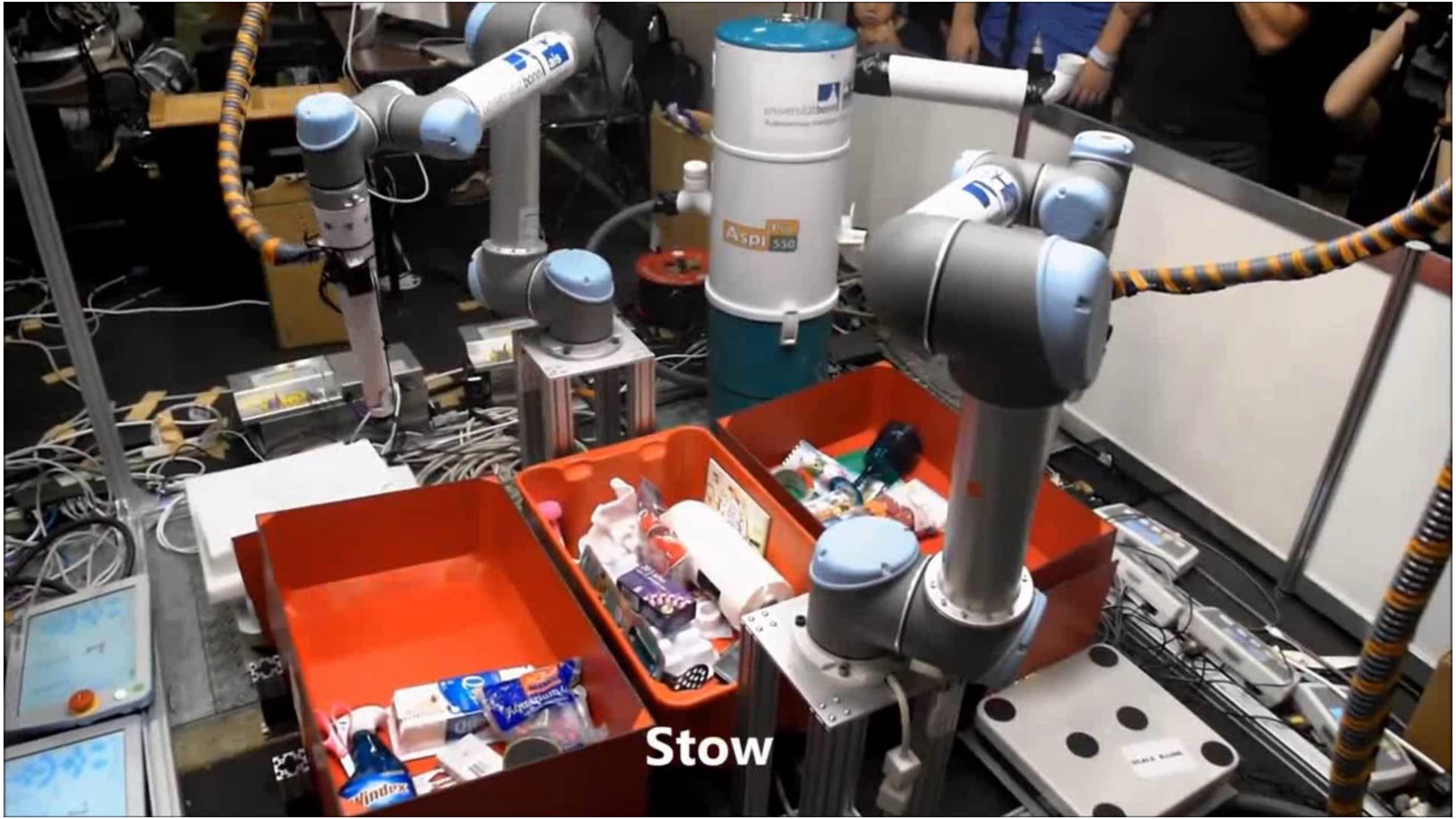


[Lin et al. CVPR 2017]

Semantic Segmentation Example

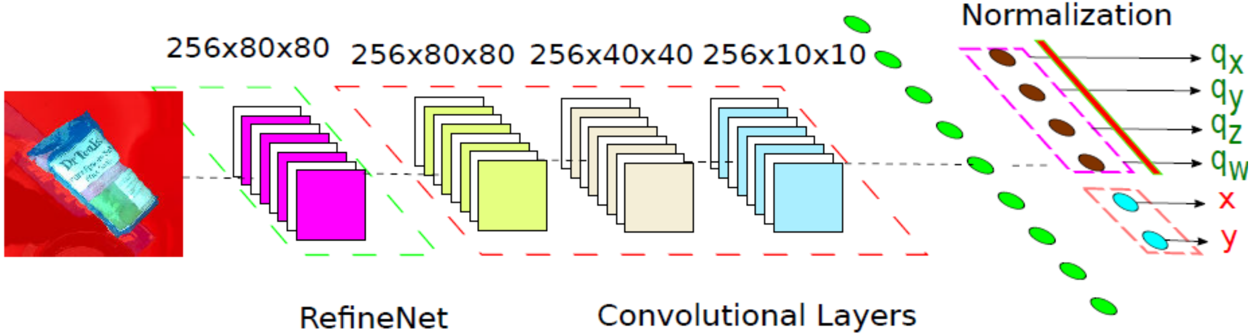
[Schwarz et al. ICRA 2018]

Amazon Robotics Challenge 2017



Object Pose Estimation

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



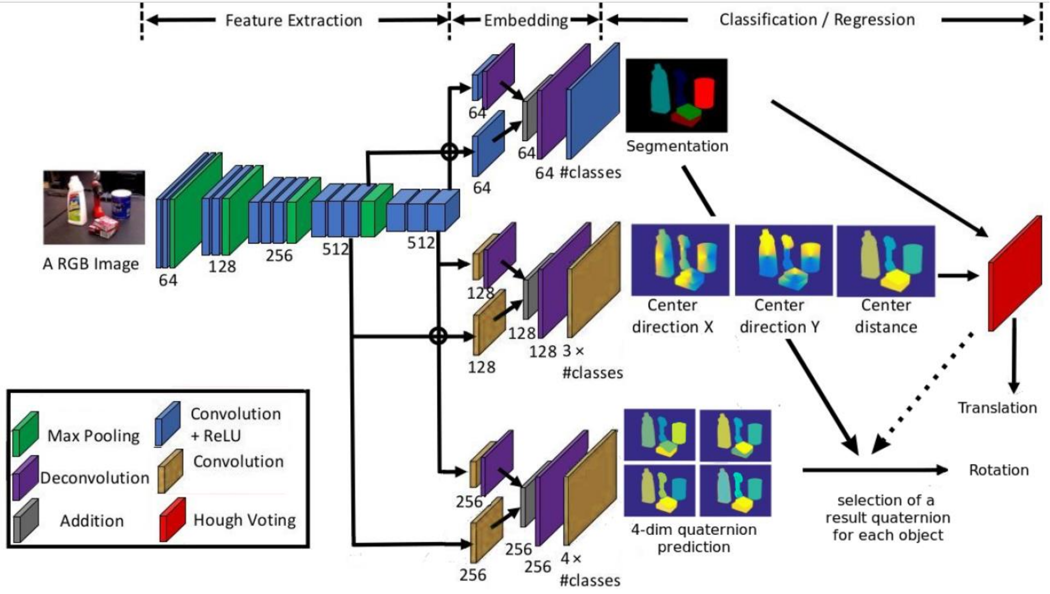
Input

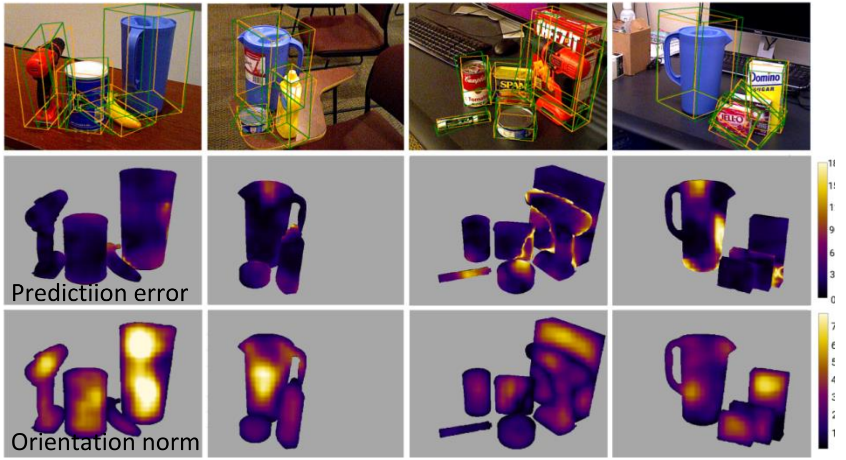
Predicted pose

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

Dense Convolutional 6D Object Pose Estimation

- Extension of PoseCNN [Xiang et al. RSS 2018]
- Dense prediction of object center and orientation, without cutting out

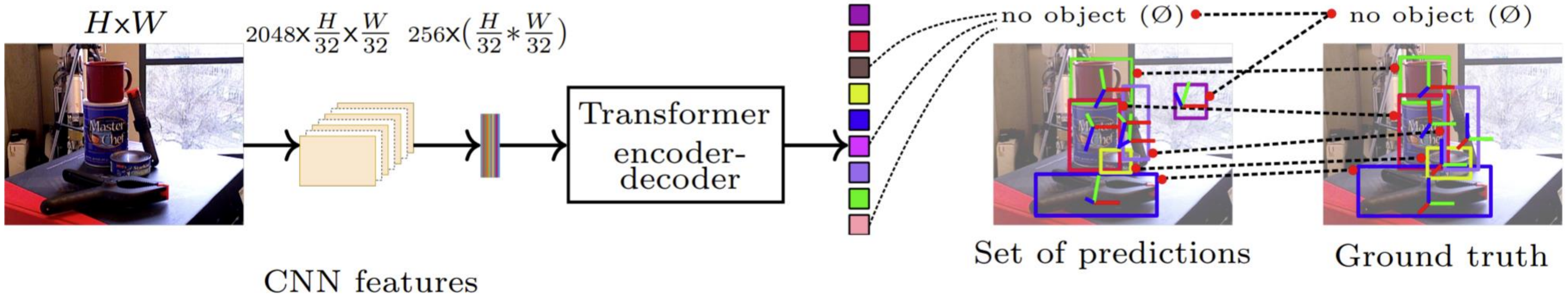




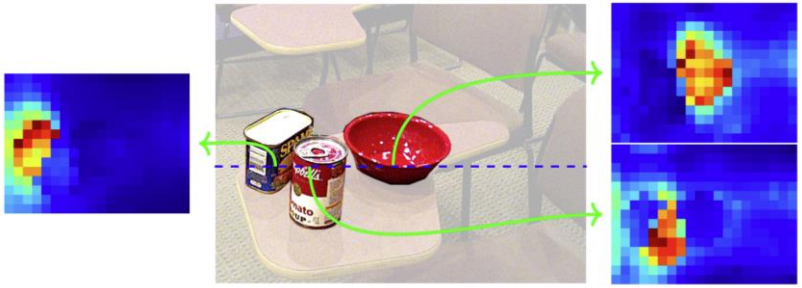
[Capellen et al., VISAPP 2020]

T6D-Direct: Transformers for Multi-Object 6D Pose Direct Regression

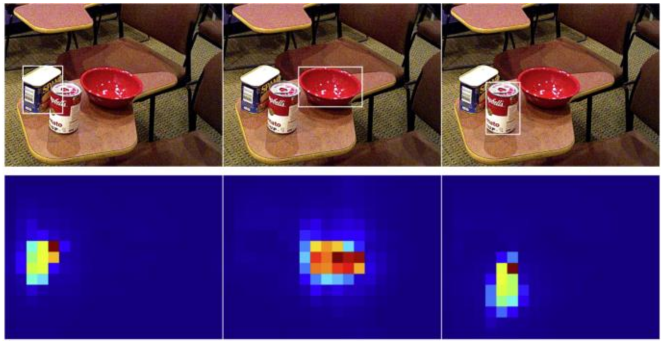
- Extends DETR: End-to-end object detection with transformers [Carion et al. ECCV 2020]
- End-to-end differentiable pipeline for 6D object pose estimation



Encoder self-attention

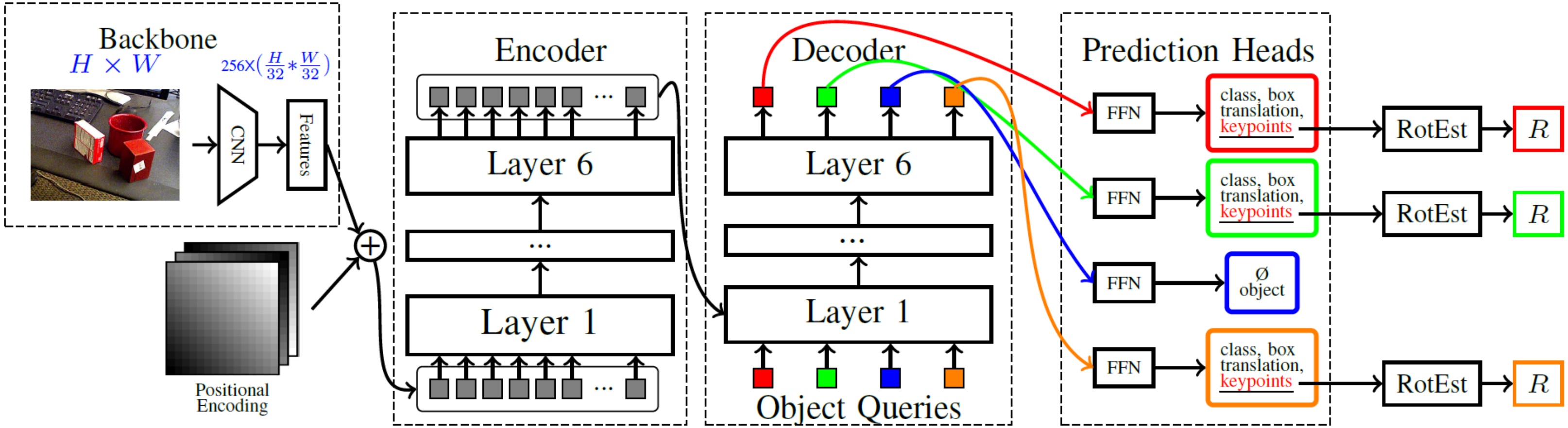


Object detections and decoder attention



[Amini et al. GCPR 2021]

Multi-Object 6D Pose Estimation using Keypoint Regression



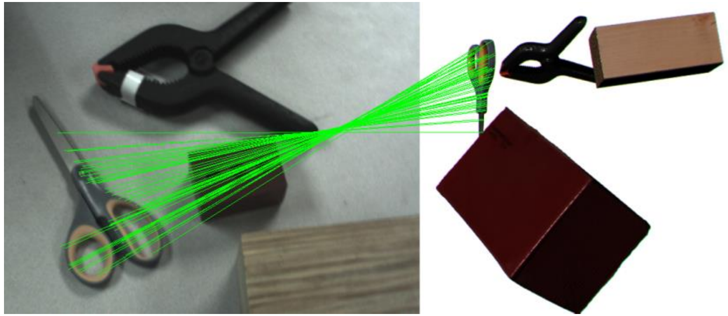
[Amini et al. under review]

From Turntable Captures to Textured Meshes

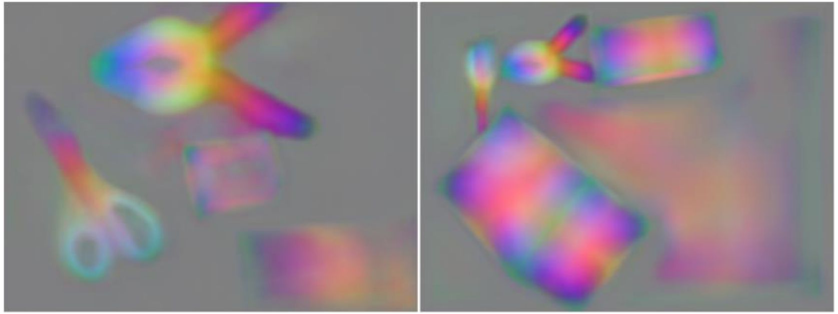
Fused & textured result

Self-Supervised Surface Descriptor Learning

- Feature descriptor should be constant under different transformations, viewing angles, and environmental effects such as lighting changes
- Descriptor should be unique to facilitate matching across different frames or representations
- Learn dense features using a contrastive loss



Known correspondences



Learned features

[Periyasamy, Schwarz, Behnke Humanoids 2019]

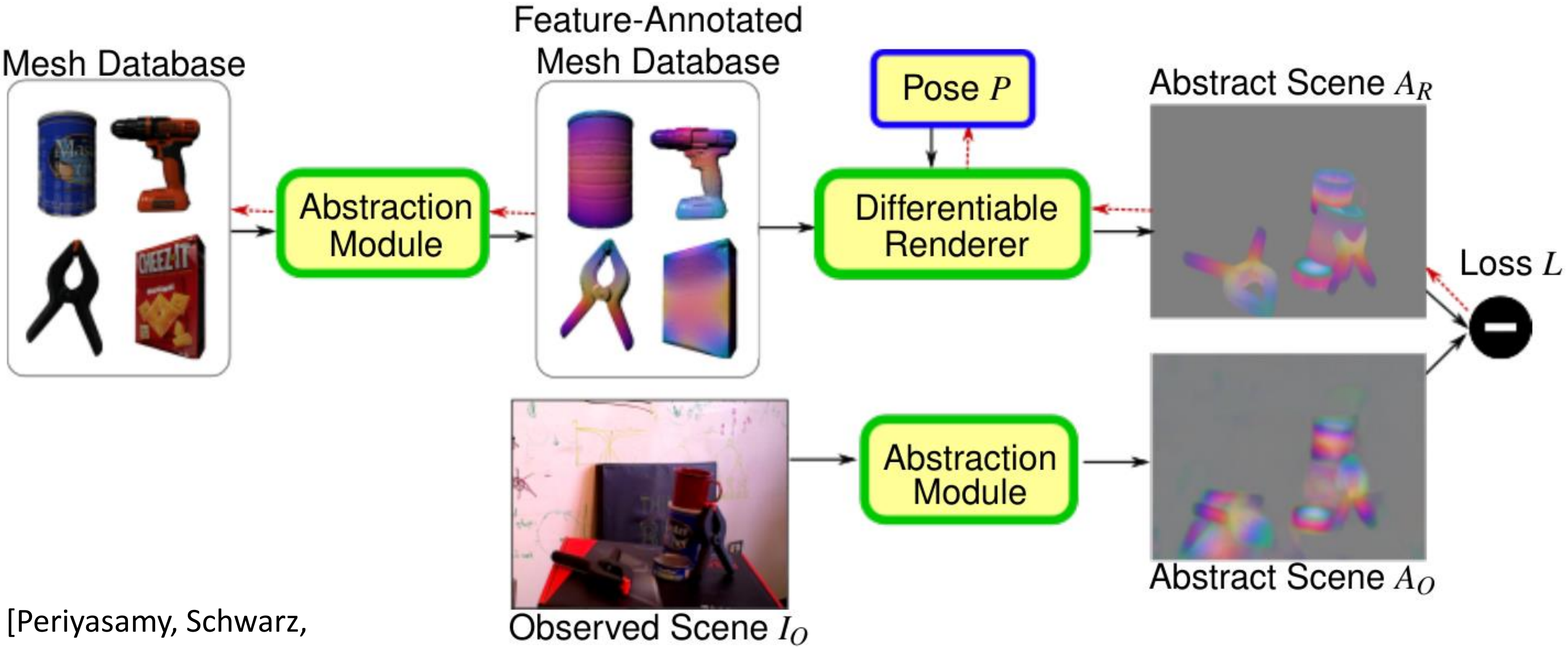
Descriptors as Texture on Object Surfaces

- Learned feature channels used as textures for 3D object models
- Used for 6D object pose estimation

[Periyasamy, Schwarz, Behnke Humanoids 2019]

Abstract Object Registration

- Compare rendered and actual scene in feature space
- Adapt model pose by gradient descent



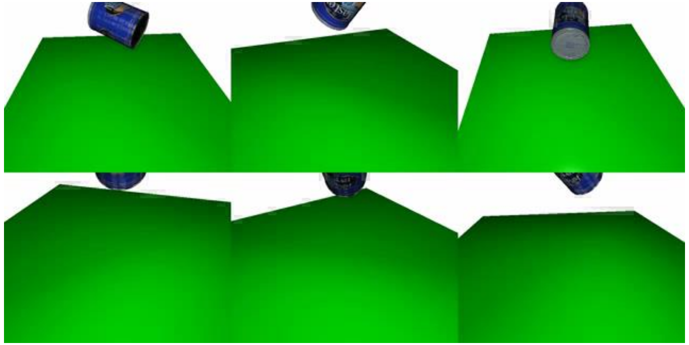
[Periyasamy, Schwarz, Behnke Humanoids 2019]

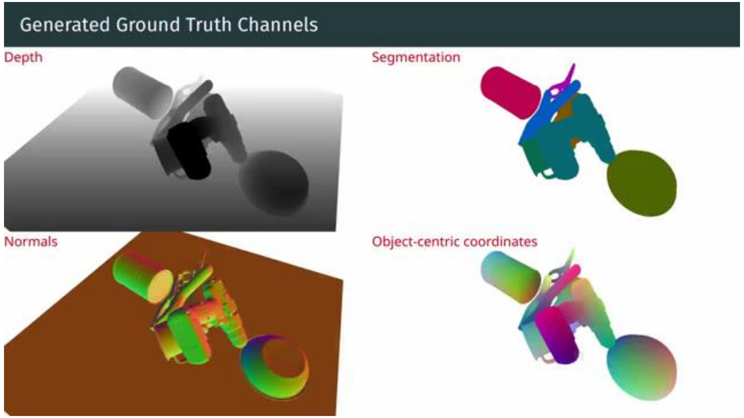
Registration Examples

[Periyasamy, Schwarz, Behnke Humanoids 2019]

Learning from Synthetic Scenes

- Cluttered arrangements from 3D meshes
- Photorealistic scenes with randomized material and lighting including ground truth
- For online learning & render-and-compare
- Semantic segmentation on YCB Video Dataset
  - Close to real-data accuracy
  - Improves segmentation of real data





[Schwarz and Behnke, ICRA 2020]

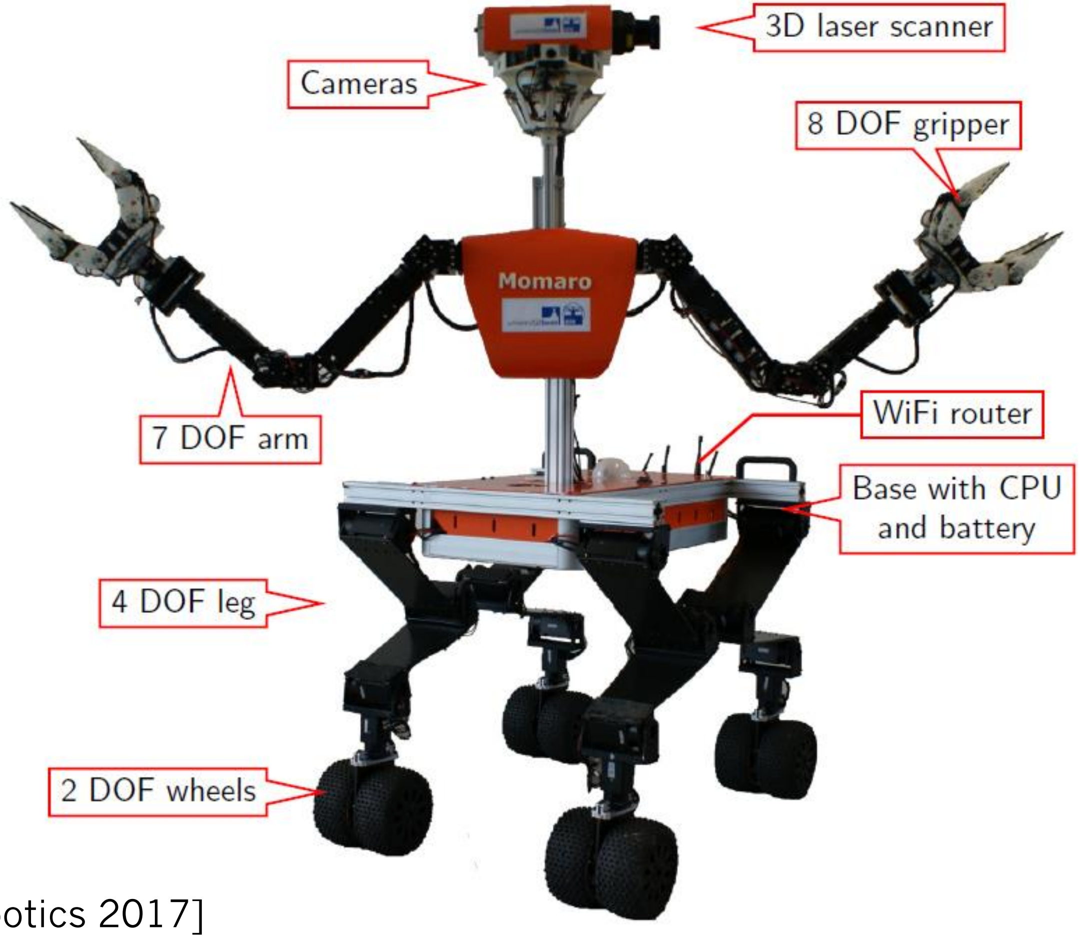
SynPick: A Dataset for Dynamic Bin Picking Scene Understanding

- Object arrangement and manipulation simulation using NVIDIA PhysX
- Untargeted and targeted picking actions, as well as random moving actions

[Periyasamy et al. CASE 2021]

Mobile Manipulation Robot Momaro

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

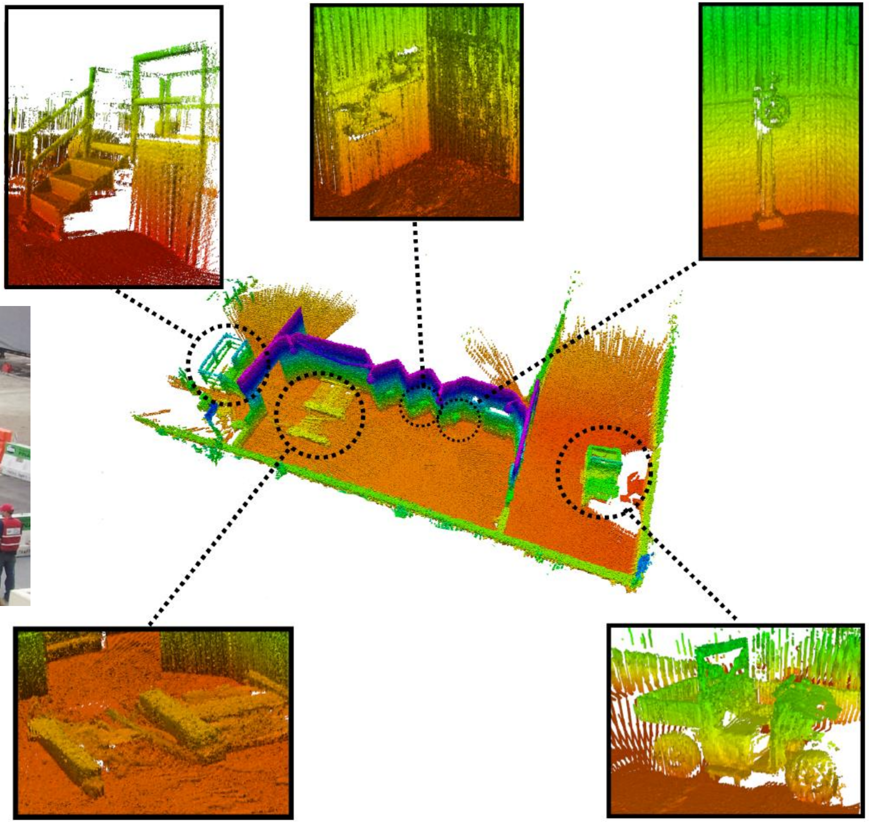


[Schwarz et al. Journal of Field Robotics 2017]

DARPA Robotics Challenge

Allocentric 3D Mapping

- Registration of egocentric maps by graph optimization

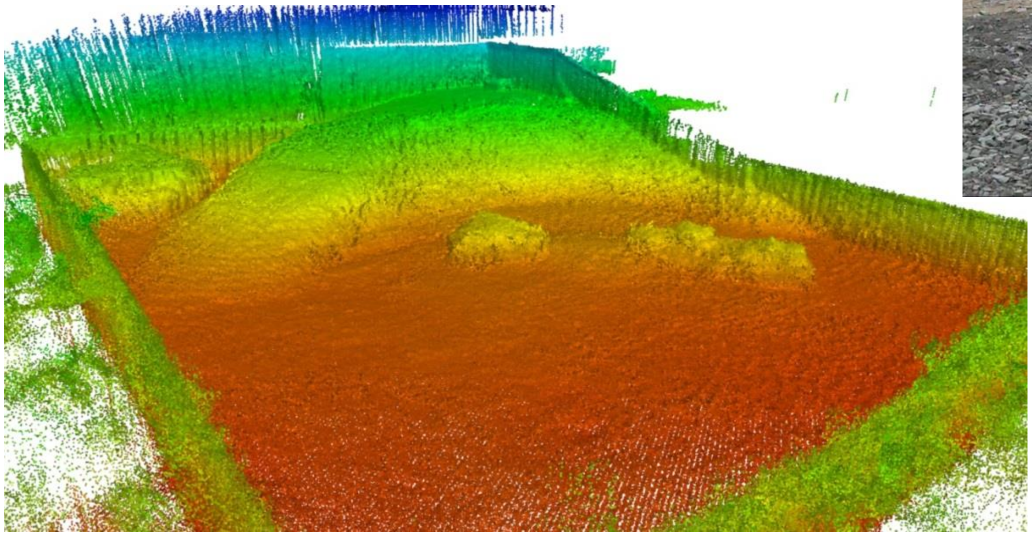


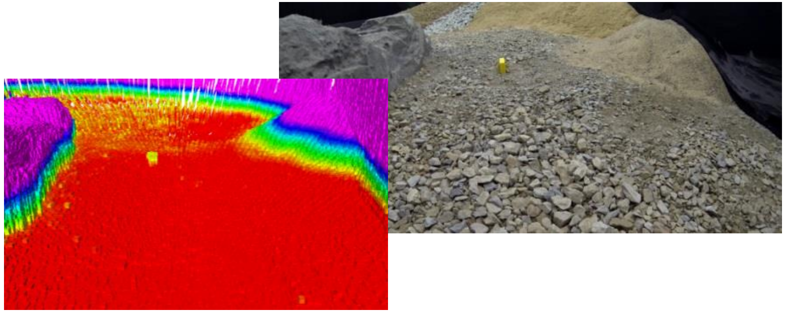
[Droeschel et al., Robotics and Autonomous Systems 2017]

DLR SpaceBot Cup 2015

- Mobile manipulation in rough terrain

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

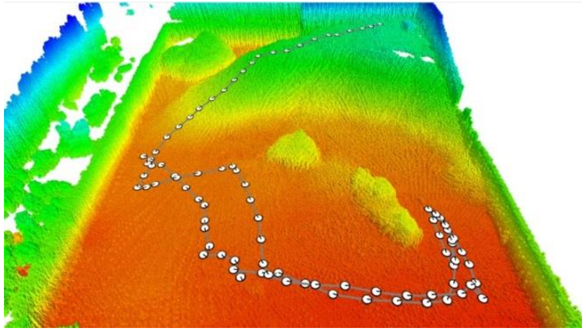




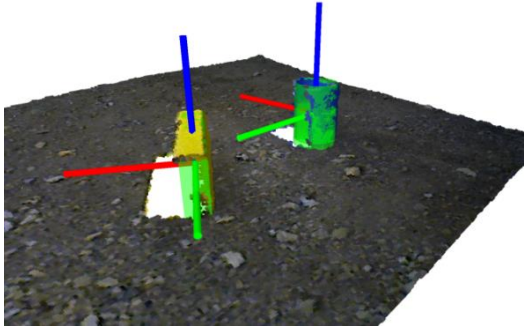

8X

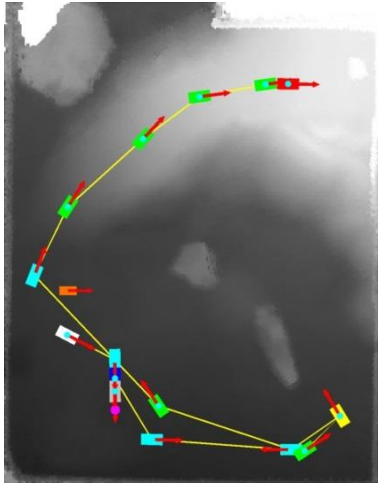
Autonomous Mission Execution

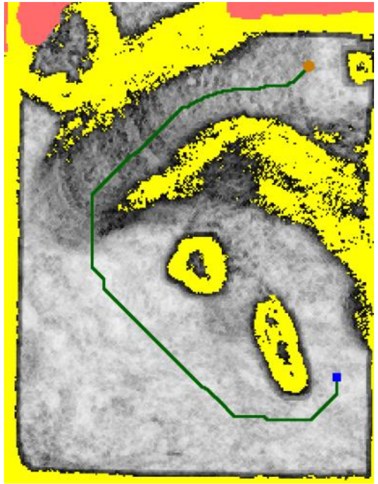
- 3D mapping, localization, mission and navigation planning

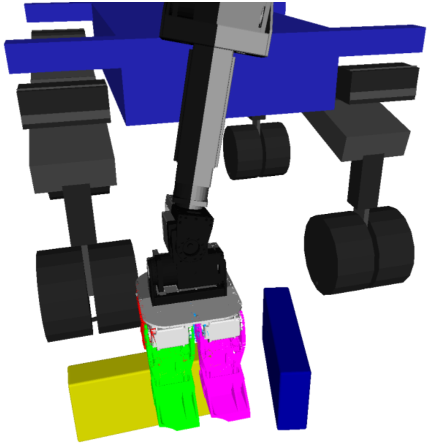


- 3D object perception and grasping





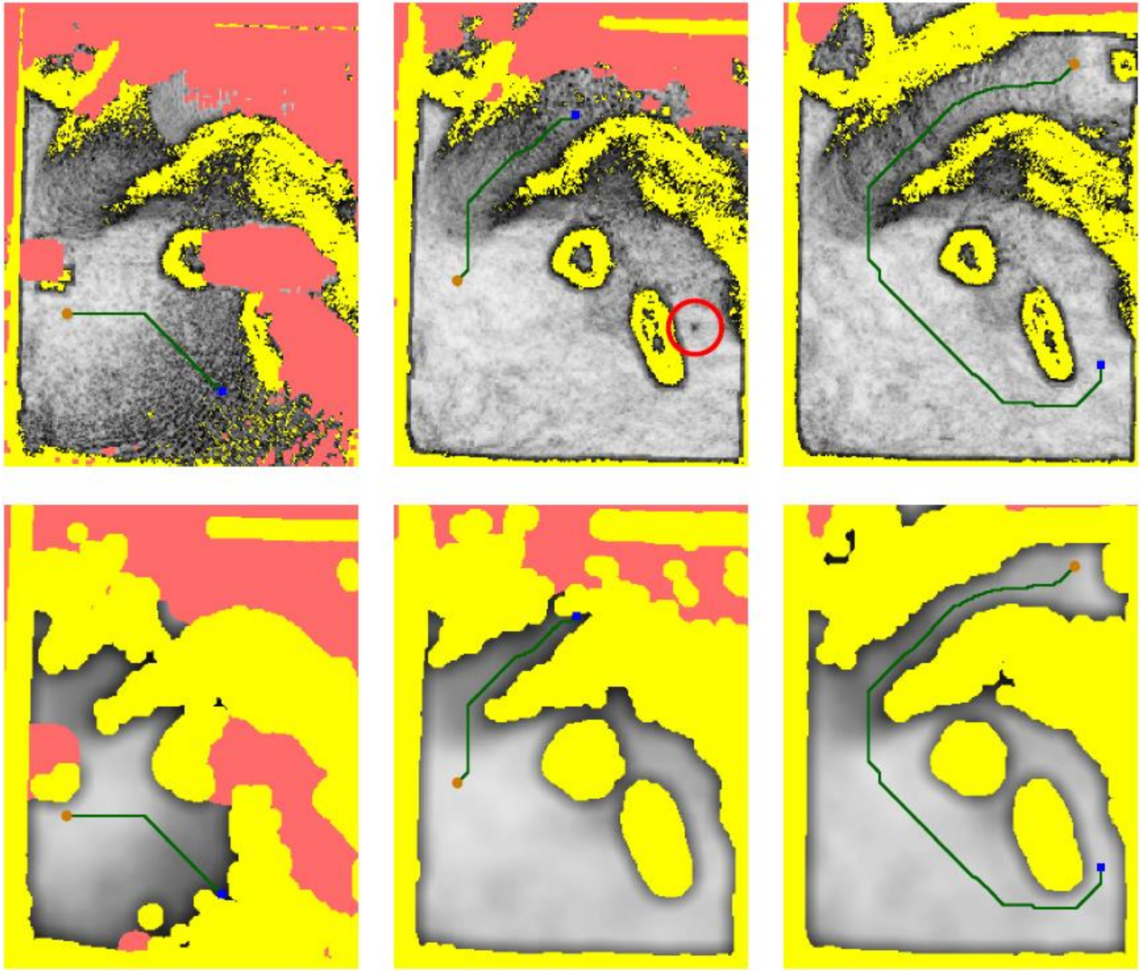




[Schwarz et al. Frontiers 2016]

Navigation Planning

- Costs from local height differences
- A* path planning

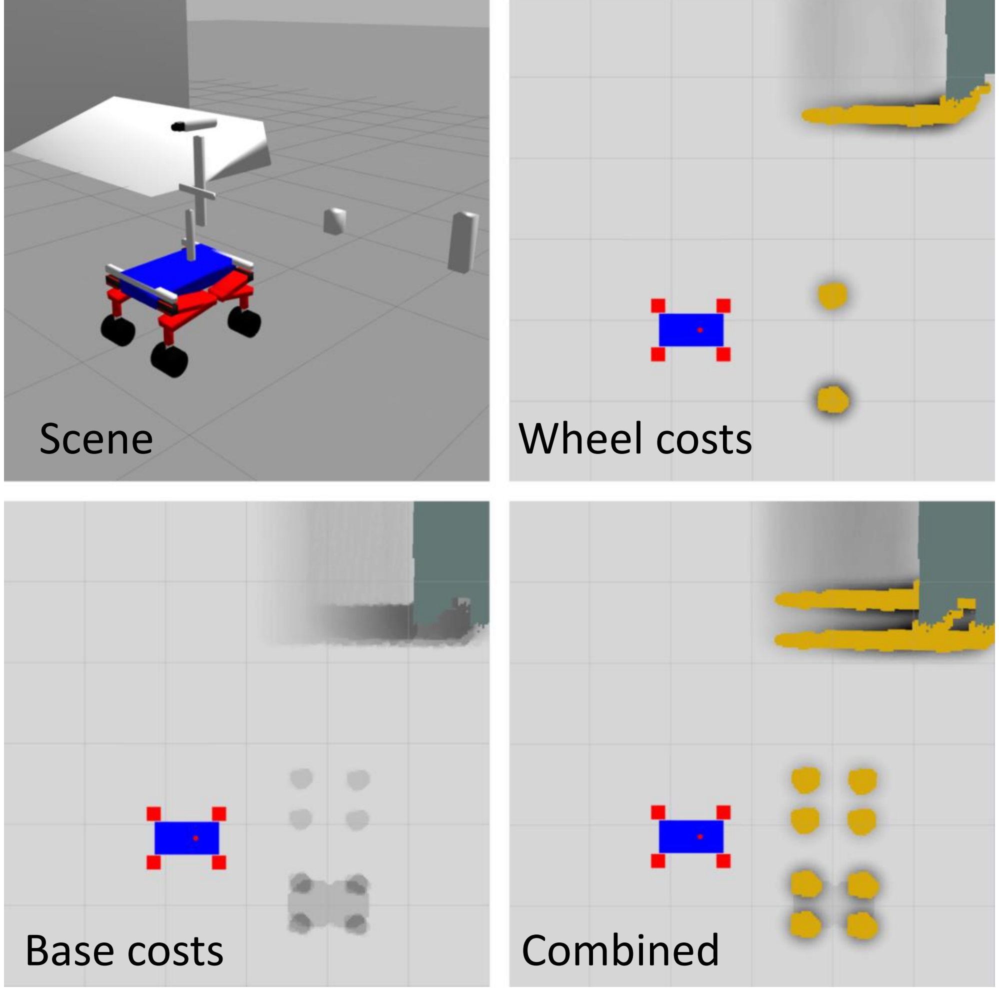


[Schwarz et al., Frontiers in Robotics and AI 2016]

Considering Robot Footprint

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

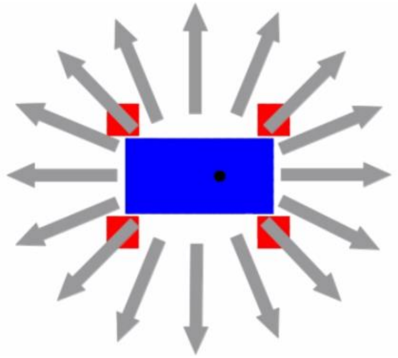
[Klamt and Behnke, IROS 2017]



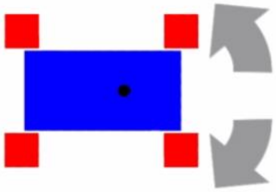
UNIVERSITÄT BONN AIS

3D Driving Planning (x, y, θ): A*

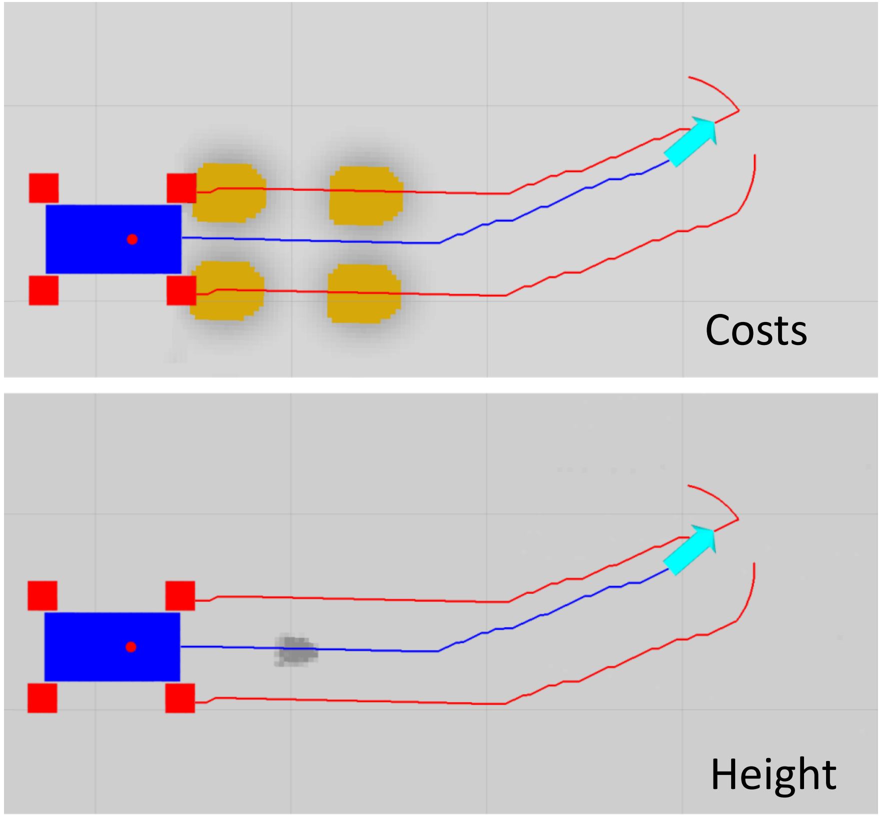
- 16 driving directions



- Orientation changes



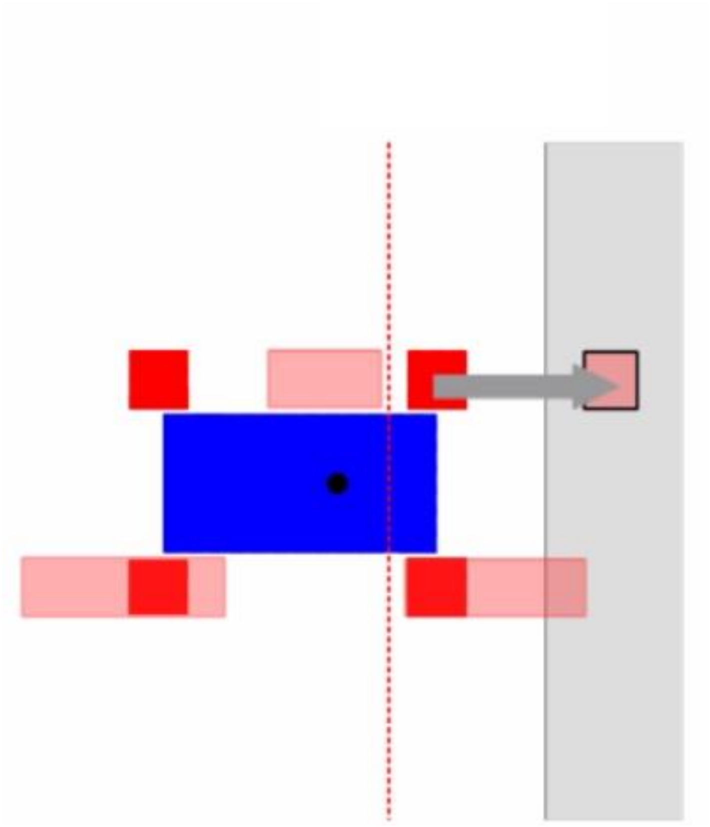
=> Obstacle between wheels



[Klamt and Behnke, IROS 2017]

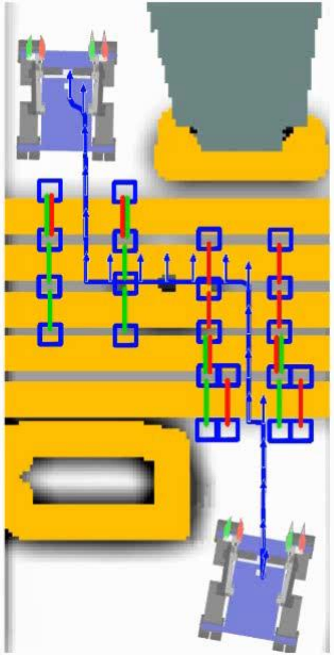
Making Steps

- If non-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 a Challenging Scenario



[Klamt and Behnke: IROS 2017]

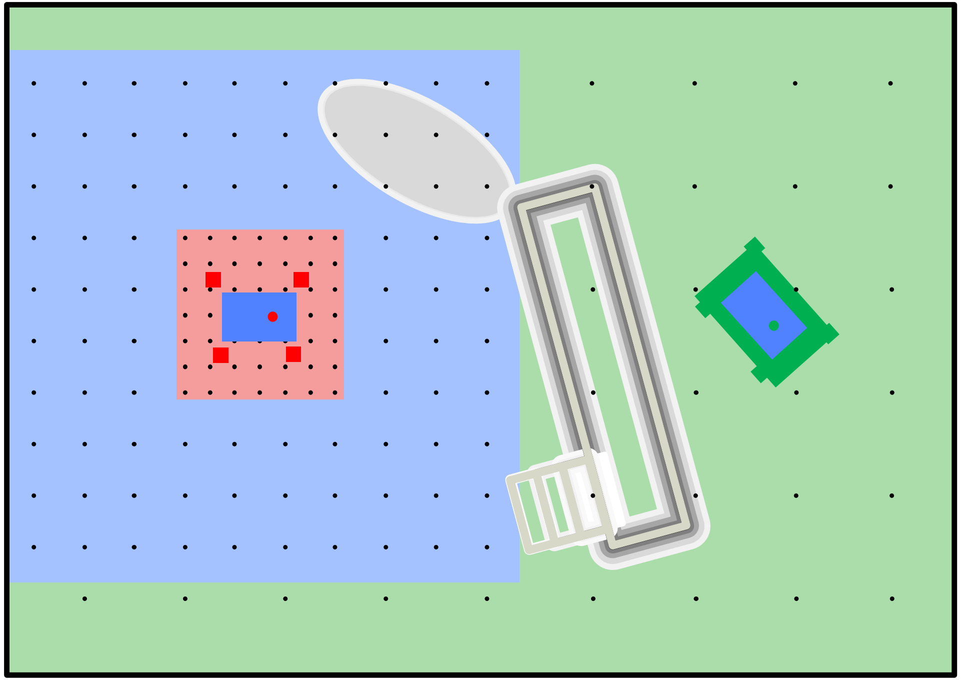
Centauro Robot

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

[Tsagarakis et al., IIT 2017]

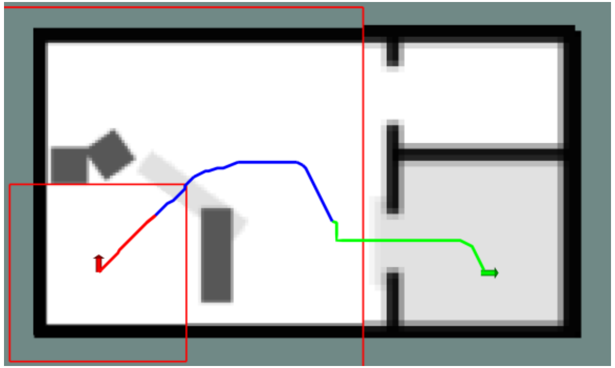
Hybrid Driving-Stepping Locomotion Planning: Abstraction

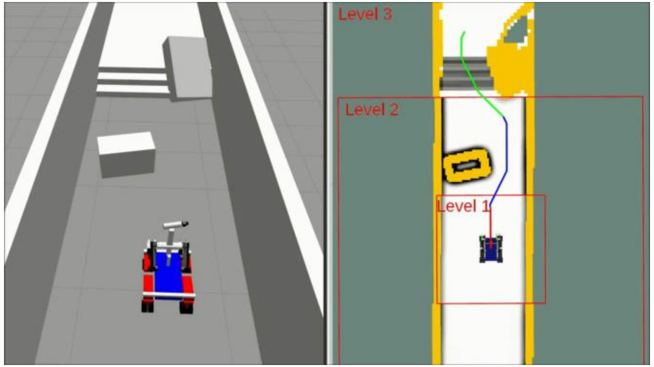
- Planning in the here and now
- Far-away details are abstracted away



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

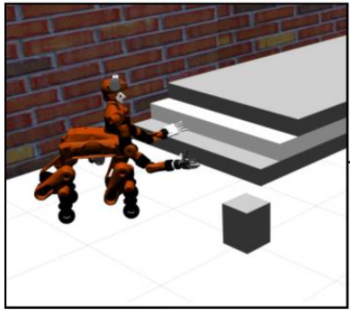




[Klamt and Behnke, IROS 2017, ICRA 2018]

Learning Cost Functions of Abstract Representations

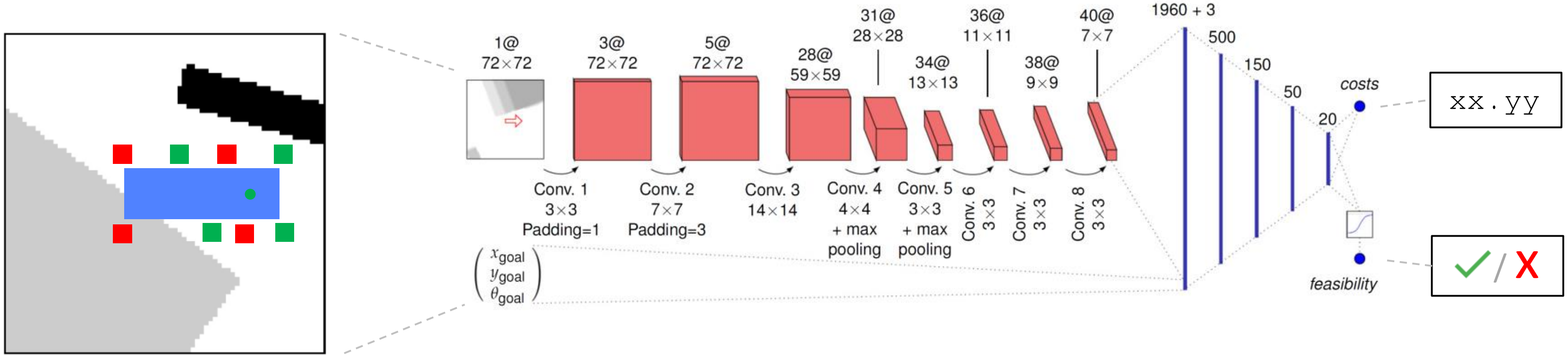
Planning problem



[Klamt and Behnke, ICRA 2019]

Abstraction CNN

- Predict feasibility and costs of local detailed planning



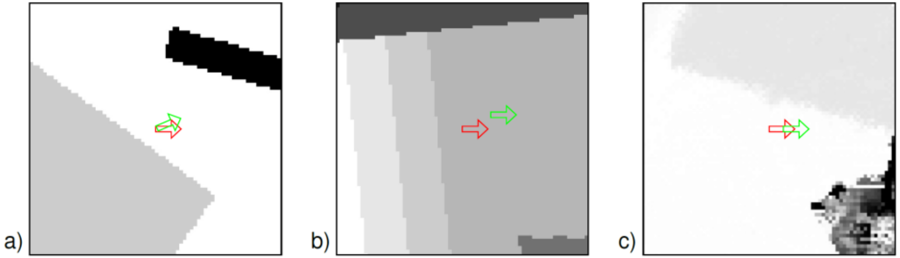
Training data

- generated with random obstacles, walls, staircases
- *costs* and *feasibility* from detailed A*-planner
- ~250.000 tasks

[Klamt and Behnke, ICRA 2019]

Learned Cost Function: Abstraction Quality

- CNN predicts feasibility and costs better than manually tuned geometric heuristics

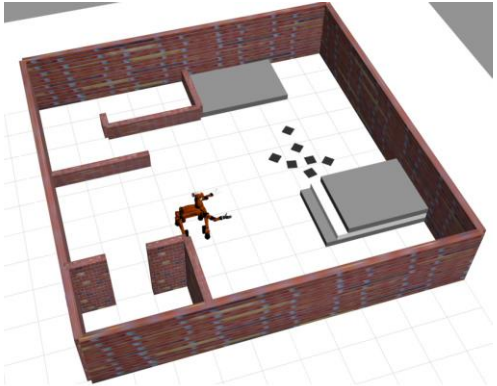


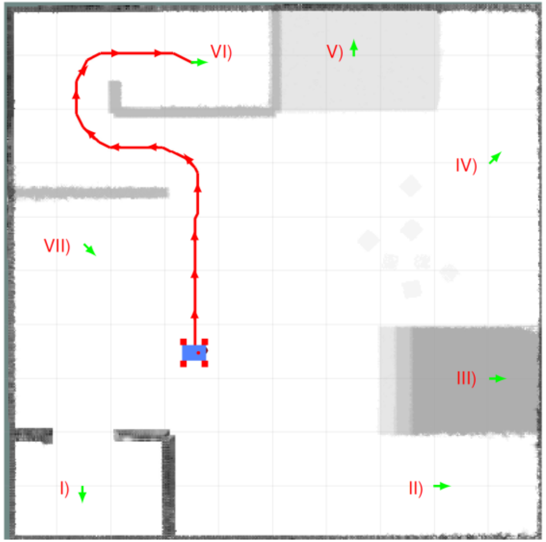
	random	*simulated*	*real*
feasibility correct, man.tuned	79.27%	65.35%	69.77%
Error($\mathcal{C}_{\text{a,man.tuned}}$)	0.057	0.021	0.103
feasibility correct, CNN	95.04%	96.69%	92.62%
Error($\mathcal{C}_{\text{a,CNN}}$)	0.027	0.013	0.081

[Klamt and Behnke, ICRA 2019]

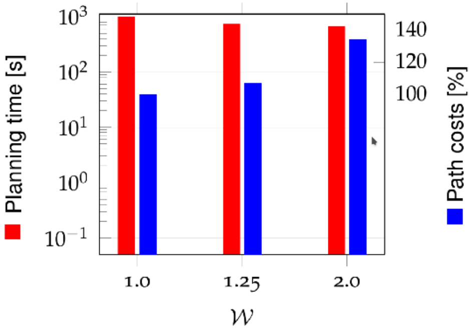
Experiments – Planning Performance

- Learned heuristics accelerates planning, without increasing path costs much

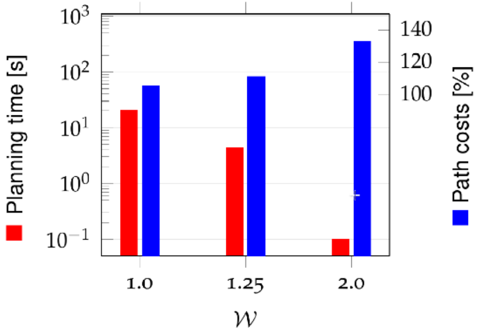




Geometric heuristic



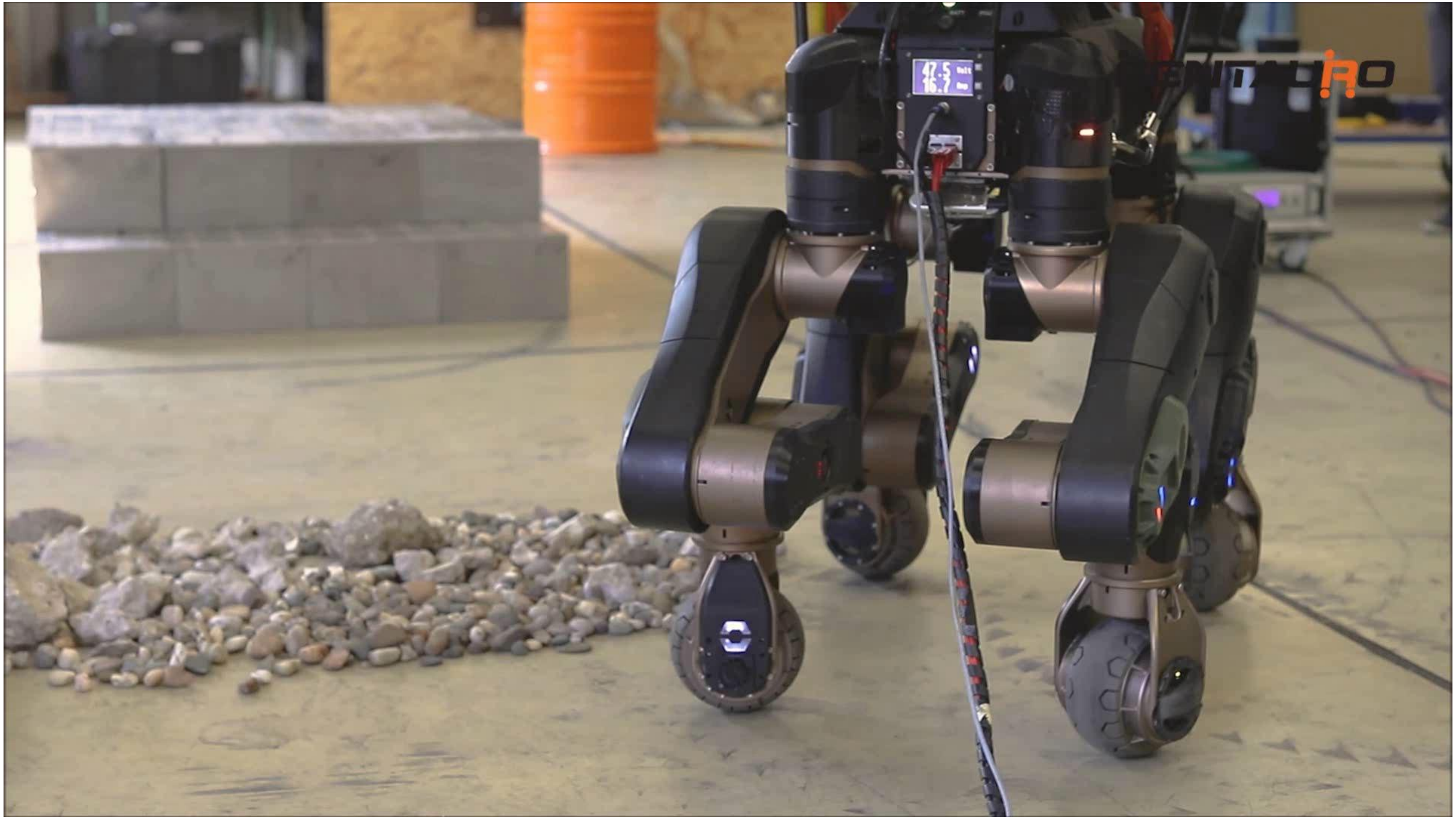
Abstract representation heuristic



Heuristic preprocessing: 239 sec

[Klamt and Behnke, ICRA 2019]

CENTAURO Evaluation @ KHG: Locomotion Tasks



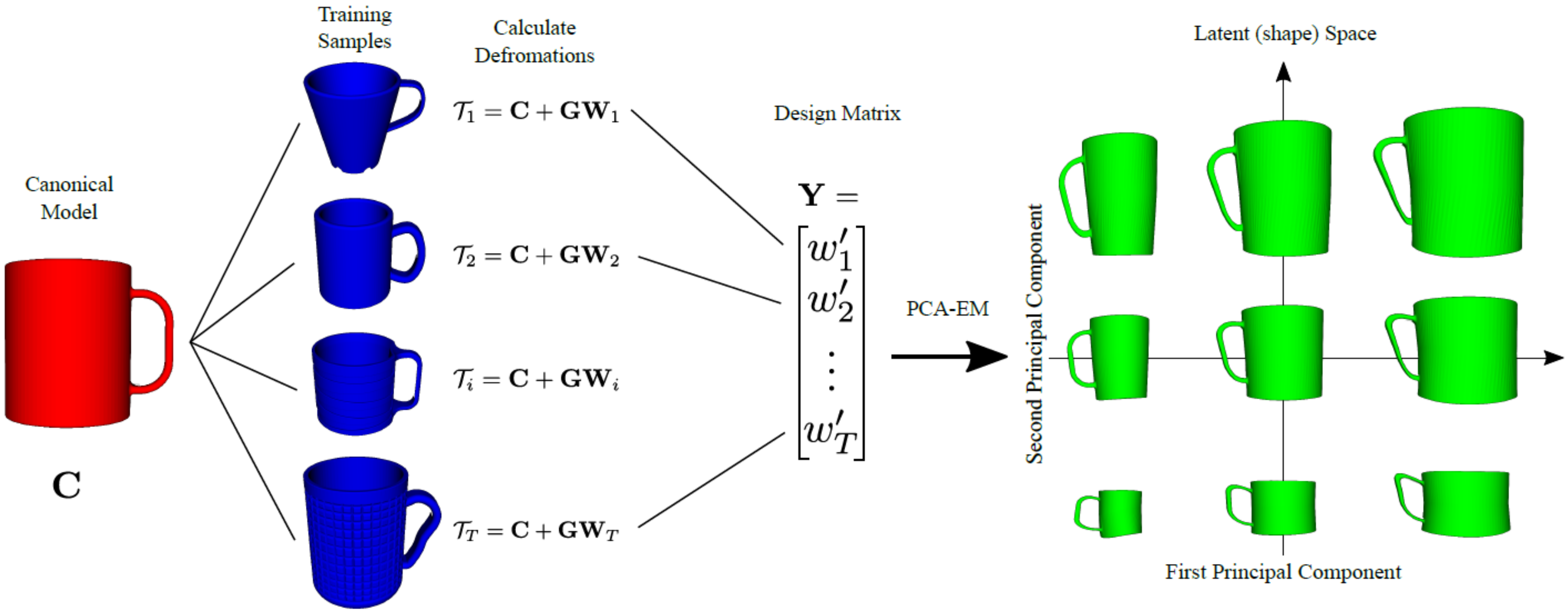
[Klamt et al. RAM 2019]

Transfer of Manipulation Skills

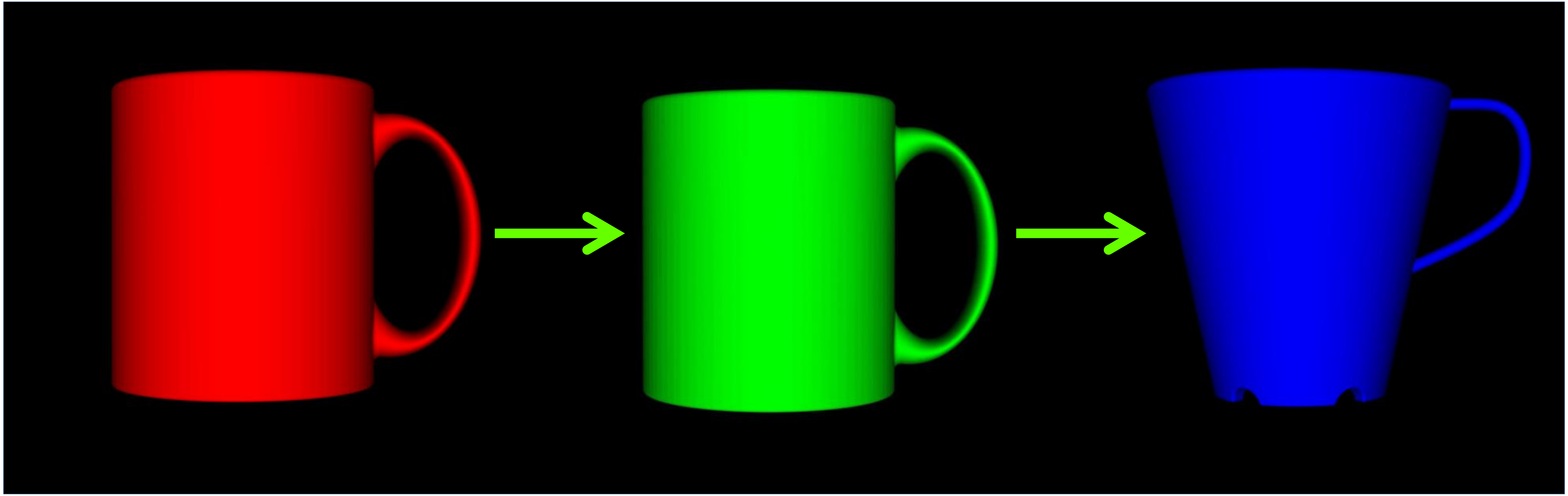
Knowledge Transfer

Learning a Latent Shape Space

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

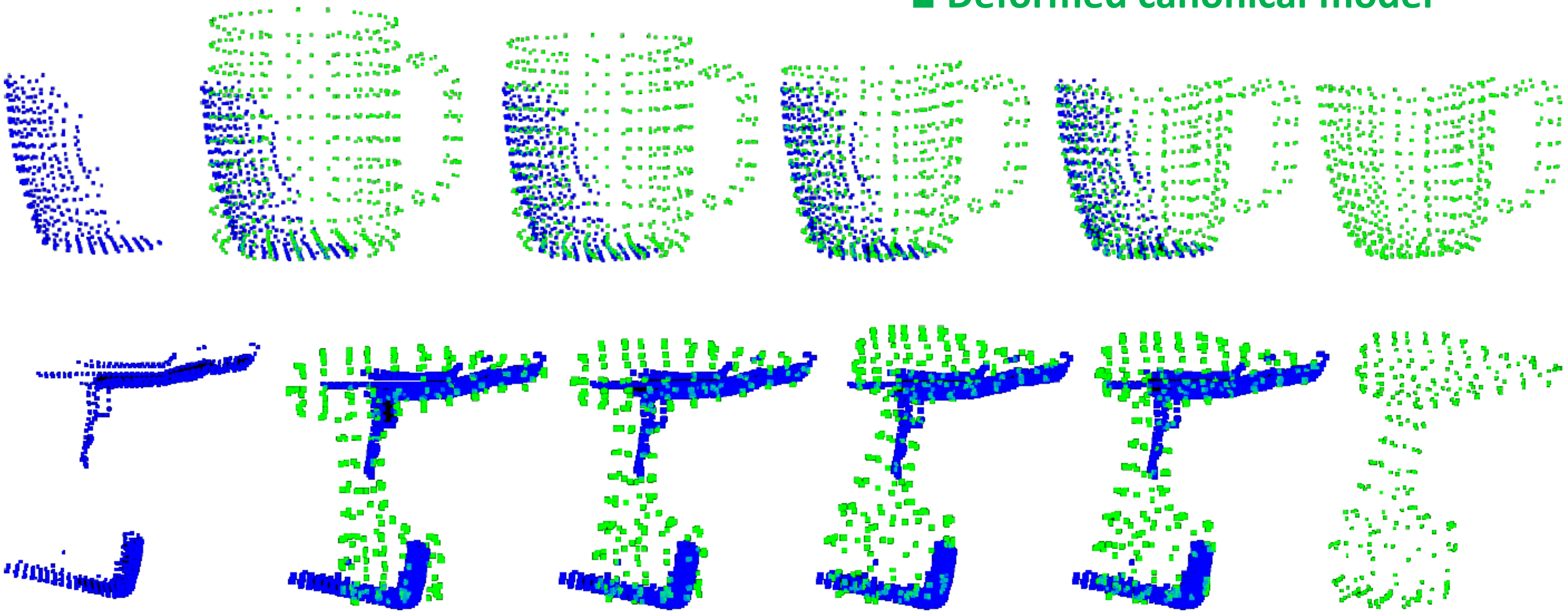


Interpolation in Shape Space



[Rodriguez and Behnke ICRA 2018]

Shape-aware Non-rigid Registration

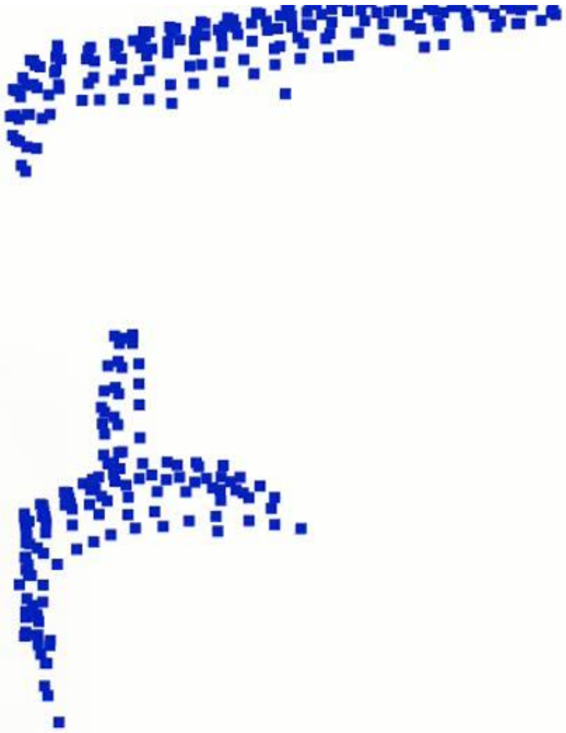


[Rodriguez and Behnke ICRA 2018]

Shape-aware Registration for Grasp Transfer

- Full point cloud

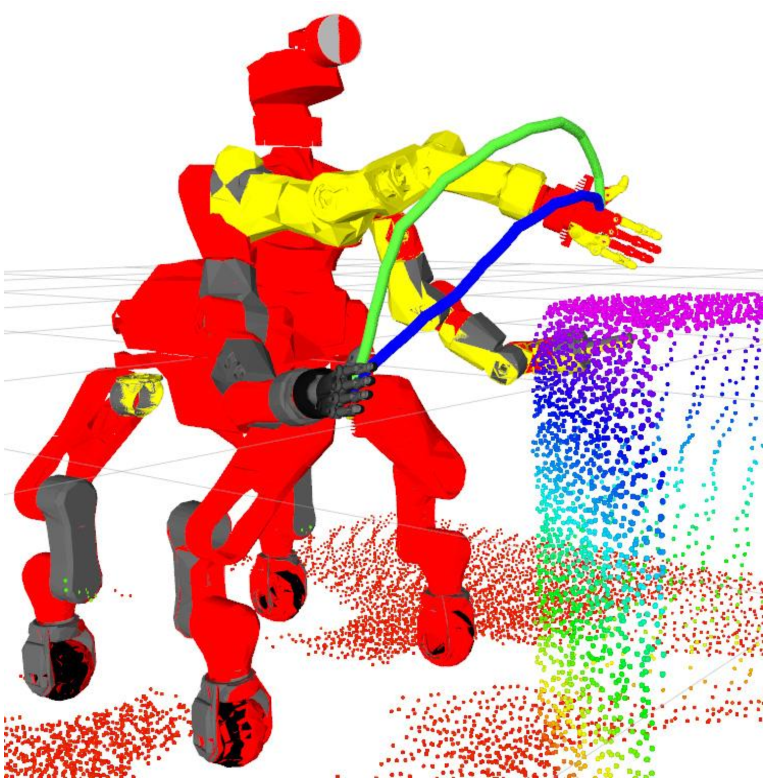
- Partial view



Collision-aware Motion Generation

Constrained Trajectory Optimization:

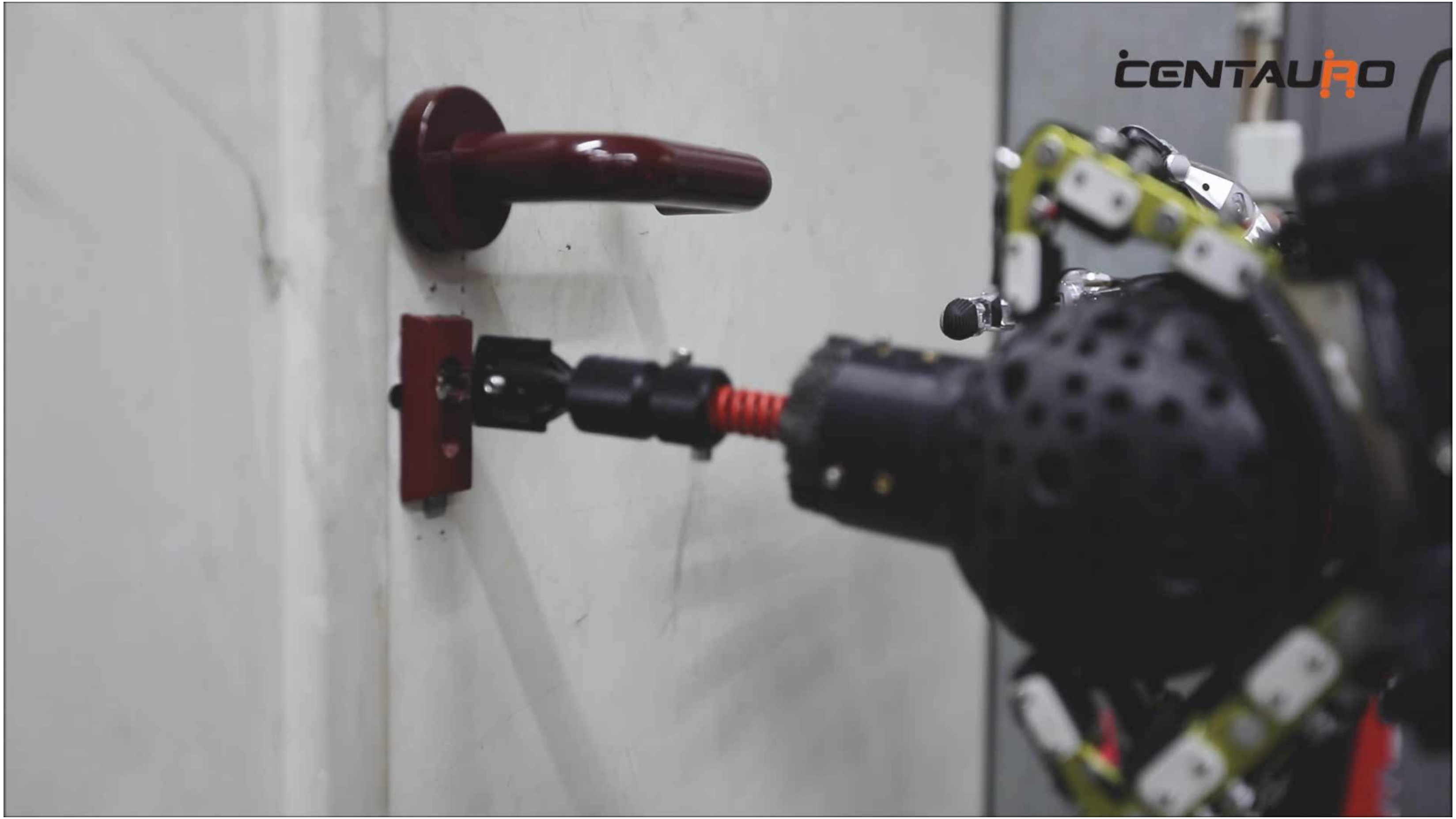
- Collision avoidance
- Joint limits
- Time minimization
- Torque optimization



[Pavlichenko et al., IROS 2017]

Grasping an Unknown Power Drill and Fastening Screws

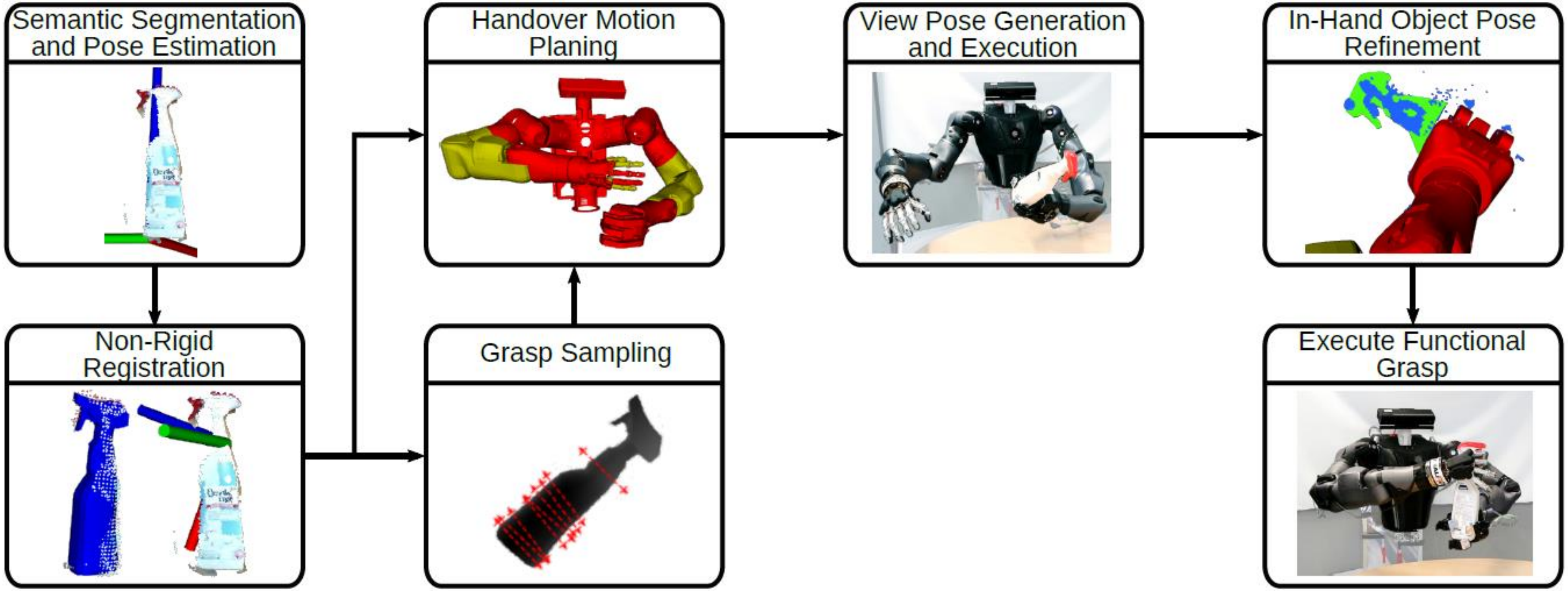
CENTAURO: Complex Manipulation Tasks



[Klamt et al. RAM 2019]

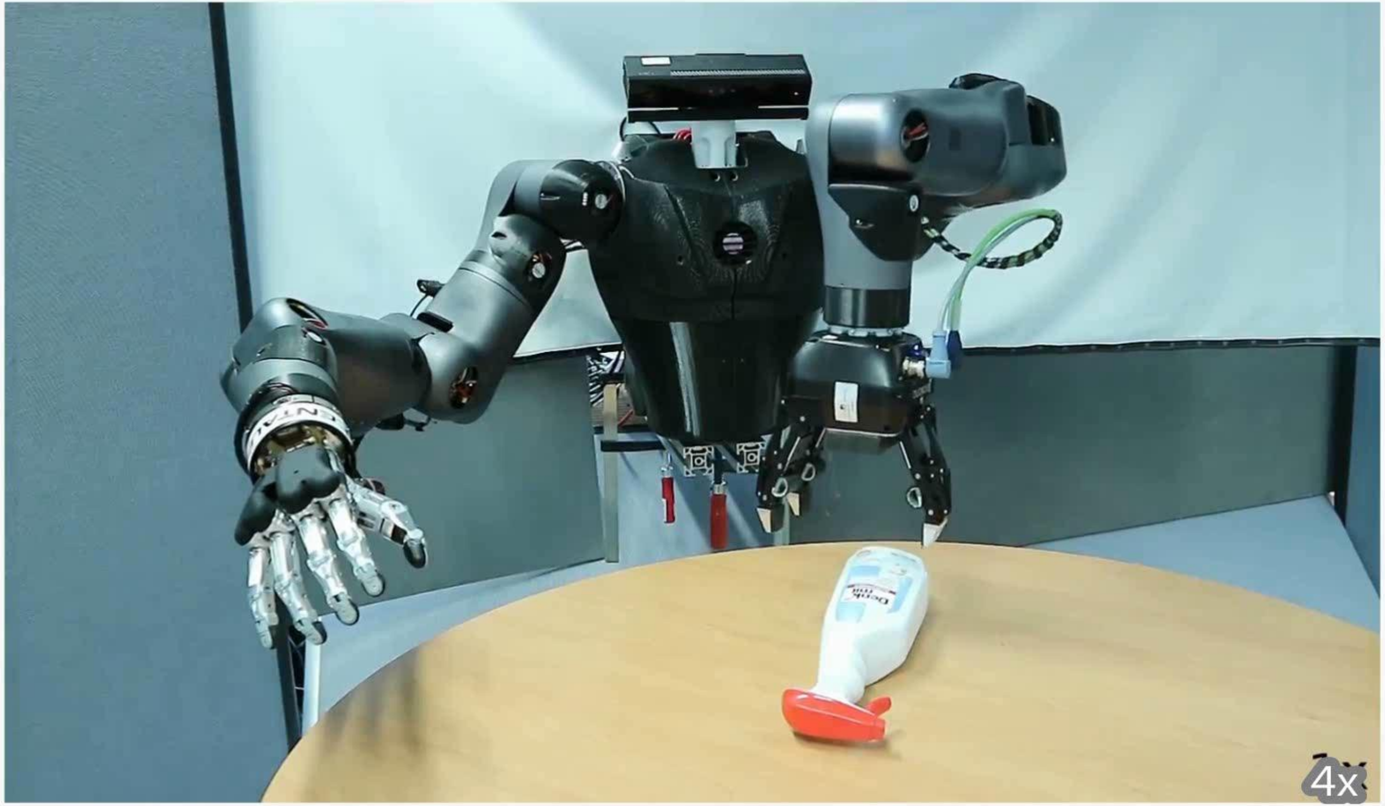
Regrasping for Functional Grasp

- Direct functional grasps not always feasible
- Pick up object with support hand, such that it can be grasped in a functional way



[Pavlichenko et al. Humanoids 2019]

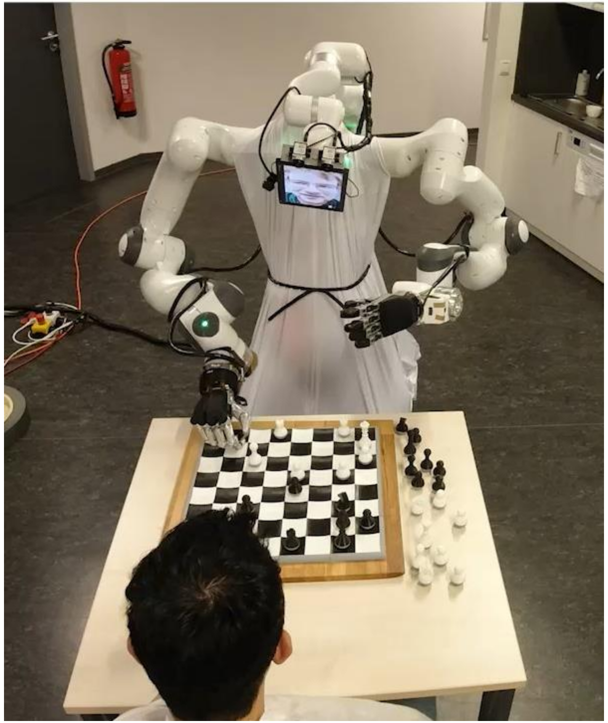
Regrasping Experiments

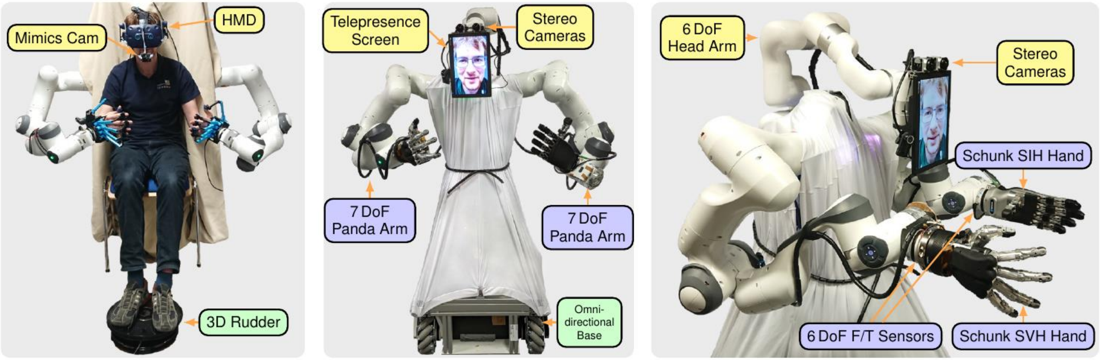


[Pavlichenko et al. Humanoids 2019]

NimbRo Avatar

ANA AVATAR XPRIZE





(a) Operator Station (b) Avatar Robot (c) Avatar Upper Body

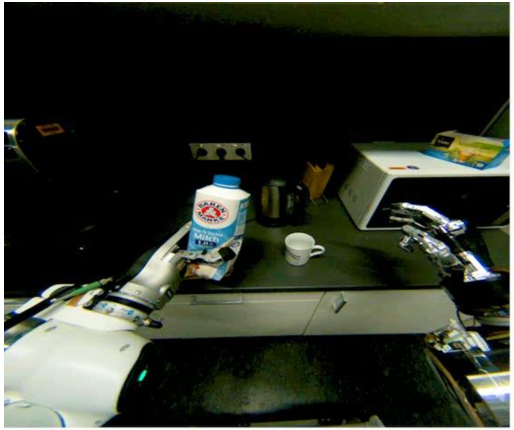
- Two-armed avatar robot designed for teleoperation with immersive visualization & force feedback
- Connected to operator station with HMD, exoskeleton and locomotion interface

[Schwarz et al. IROS 2021]

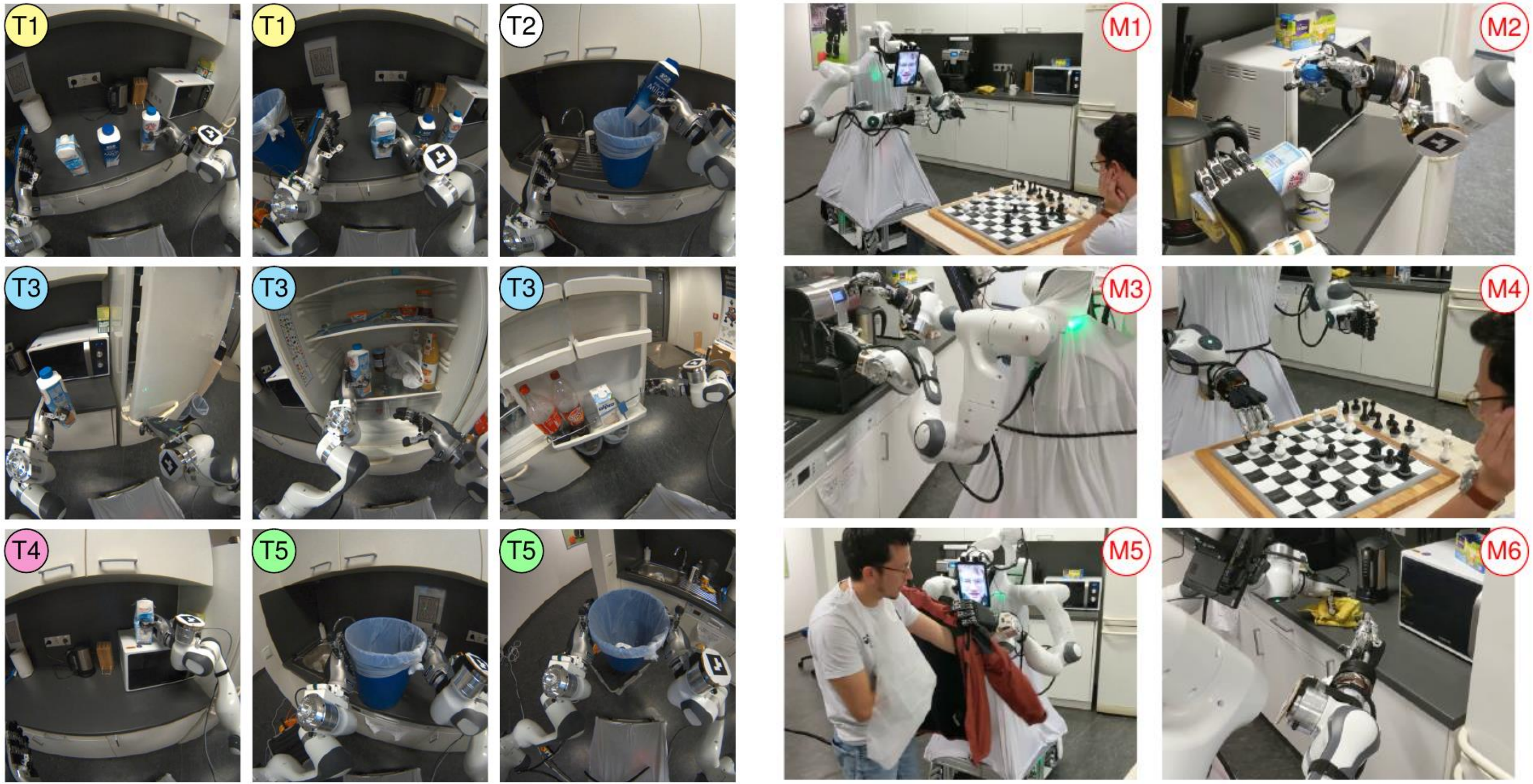
Team NimbRo Semifinal Submission

ANA AVATAR XPRIZE

[Schwarz et al. IROS 2021]

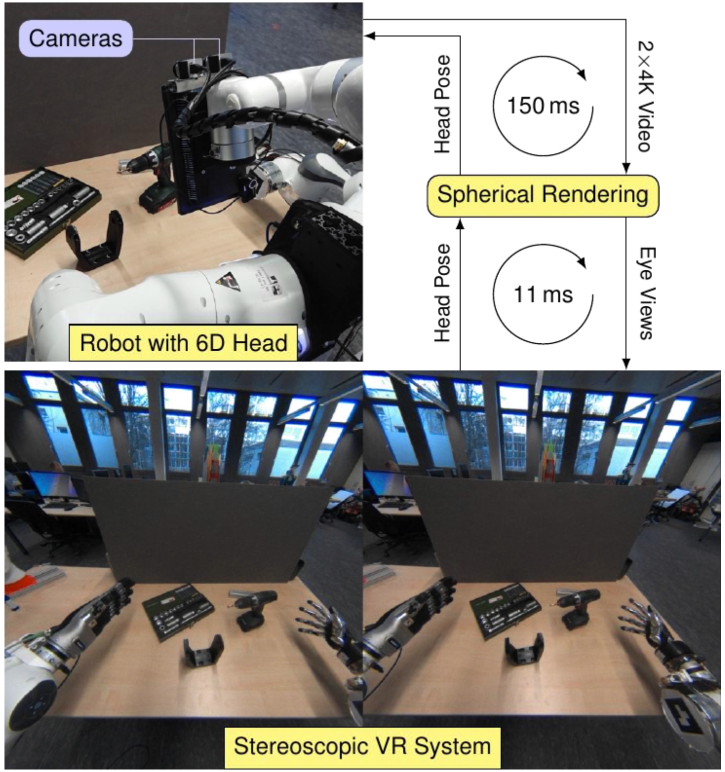


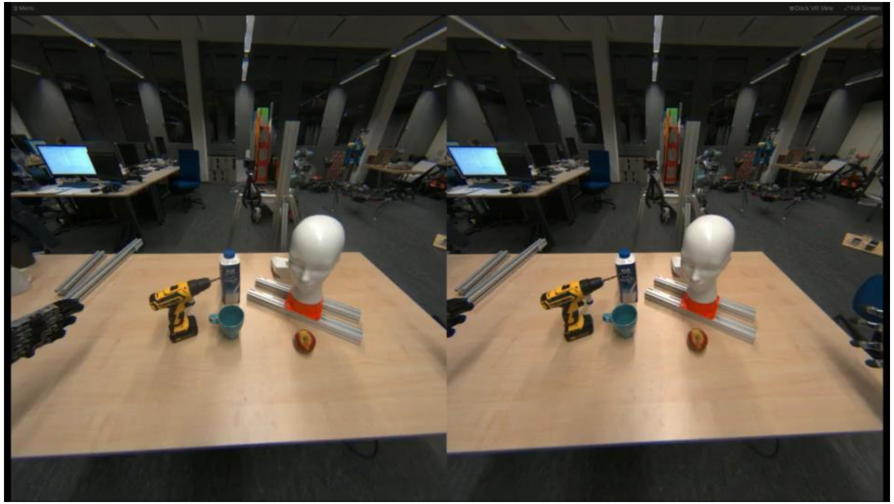
NimbRo Avatar – Applications



[Schwarz et al. IROS 2021]

NimbRo Avatar – VR Visualization System

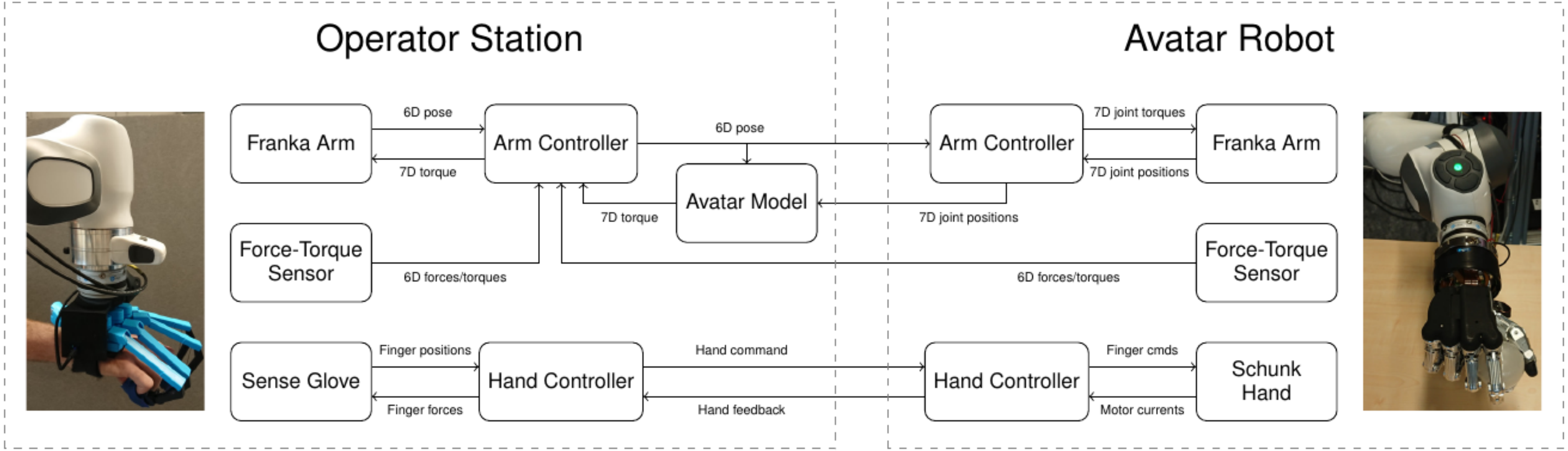




- 4K stereo video stream
- 6D head arm allows full head movement
- Spherical rendering technique hides movement latencies

[Schwarz and Behnke Humanoids 2021]

NimbRo Avatar – Haptic Manipulation

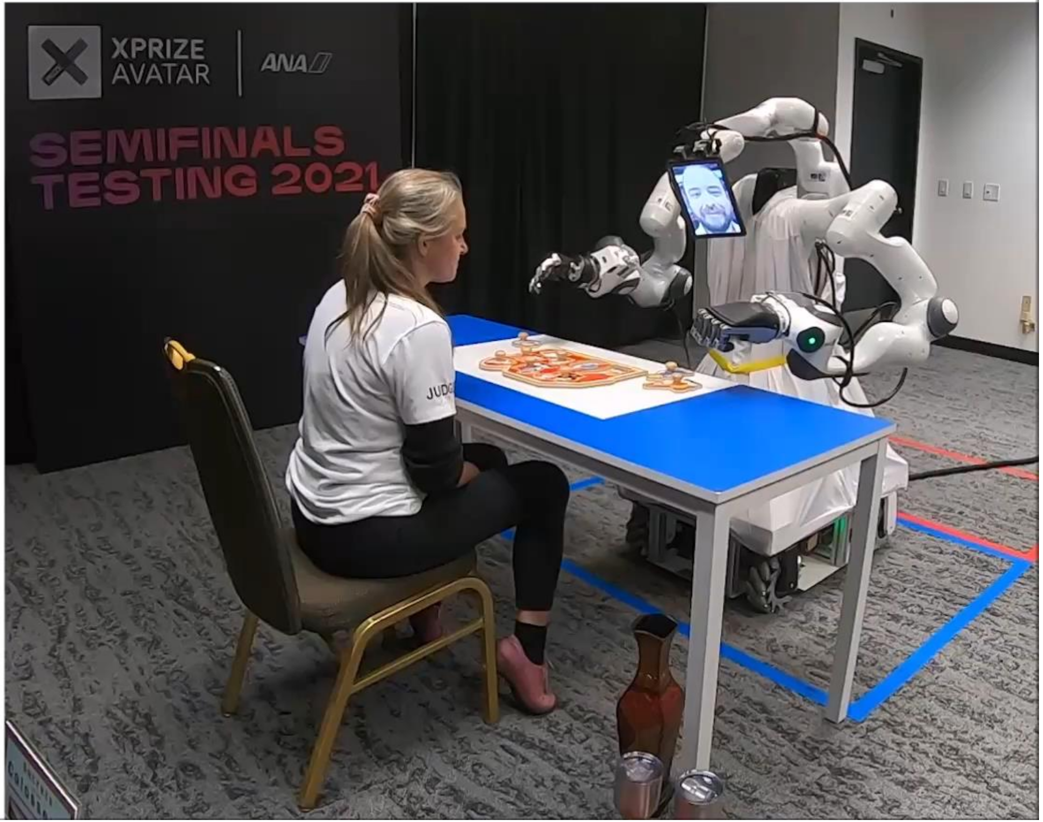


- Arm exoskeleton (Franka Emika Panda), F/T sensor (OnRobot HEX), hand exoskeleton (SenseGlove)
- Avatar side: Arm + F/T sensor + Schunk SVH / SIH hand
- Method provides force feedback for wrist & fingers
- Avatar limit avoidance using predictive model to reduce latencies

[Lenz and Behnke ECMR 2021]

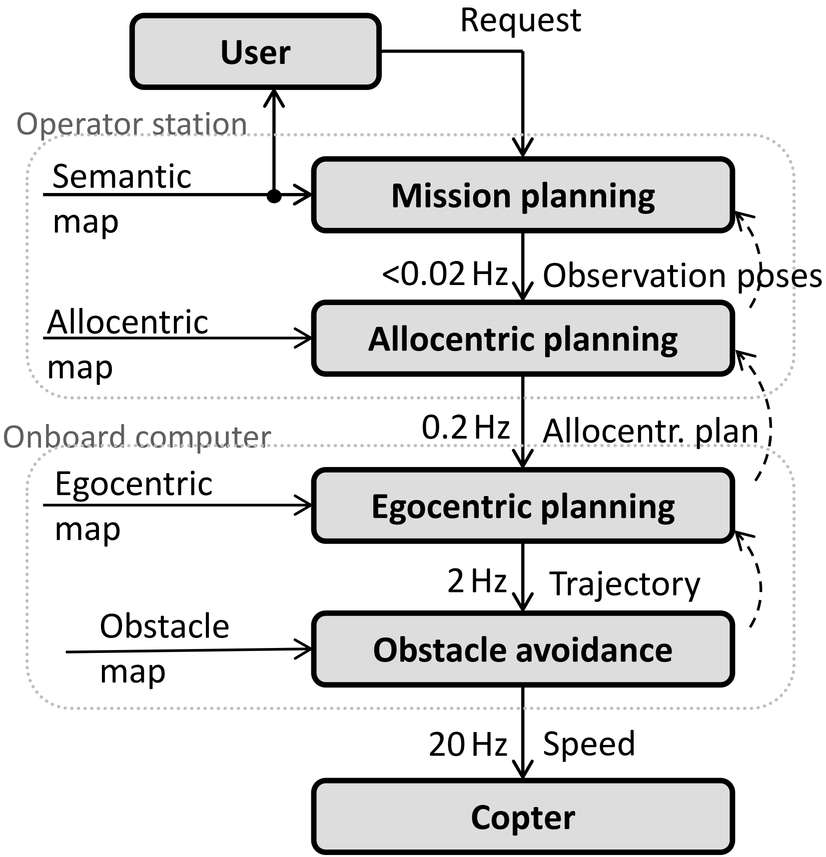
NimbRo Avatar

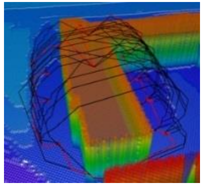
Avatar XPRIZE Semifinals



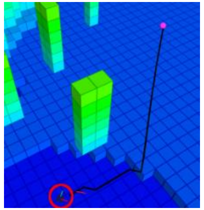
[Schwarz et al. IROS 2021]

Micro Aerial Vehicles: Hierarchical Navigation

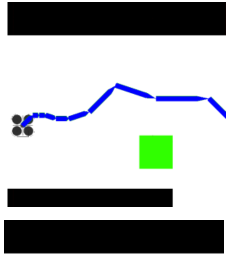




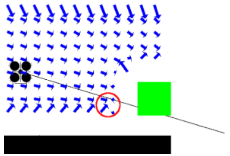
Mission plan



Allocentric planning



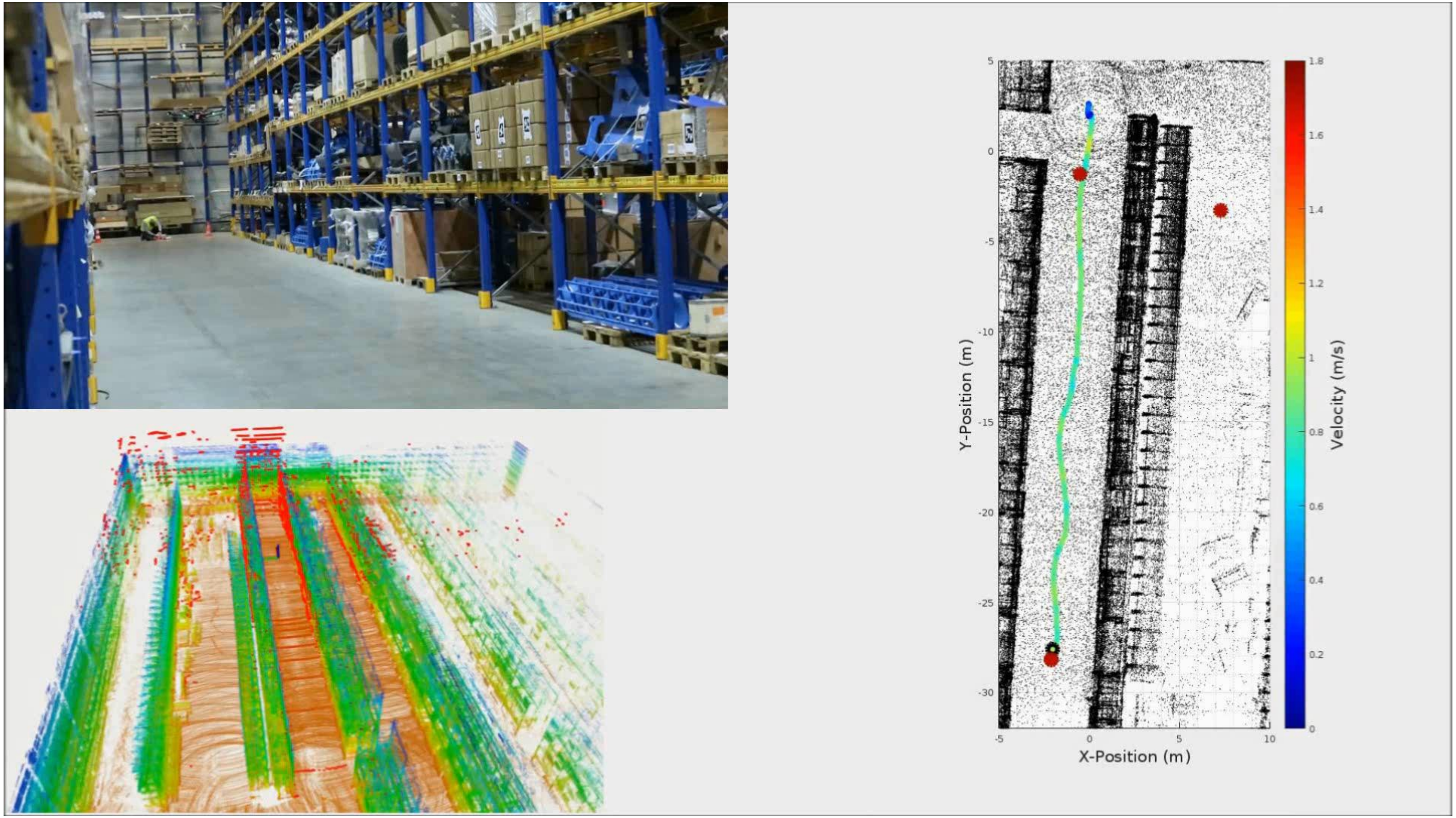
Egocentric planning



Obstacle avoidance

[Droeschel et al. JFR 2016]

InventAIRy: Autonomous Navigation in a Warehouse



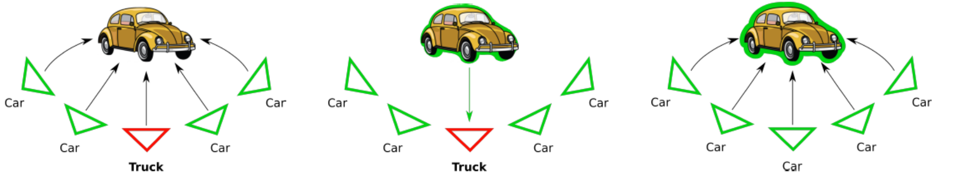
[Beul et al. RA-L 2018]

InventAIRy: Detected Tags in Shelf

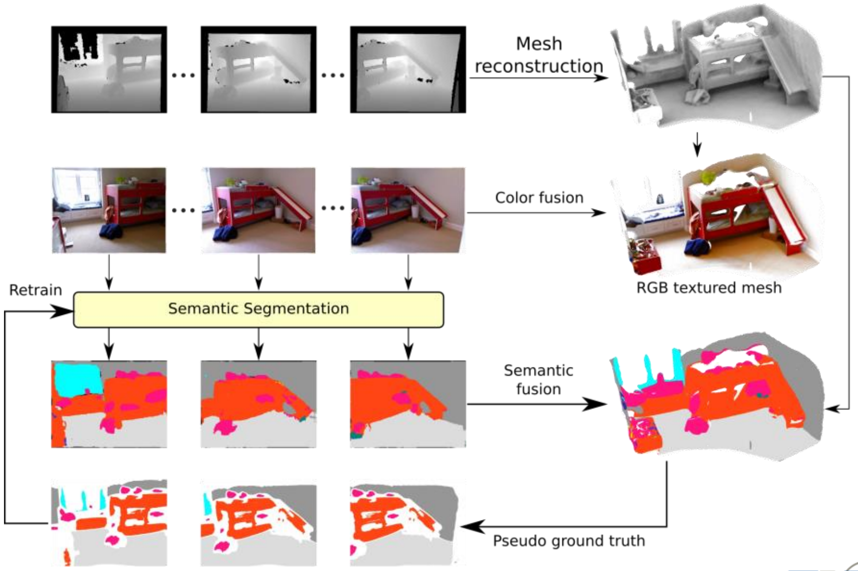
[Beul et al. RA-L 2018]

Label Propagation for 3D Semantic Mapping

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

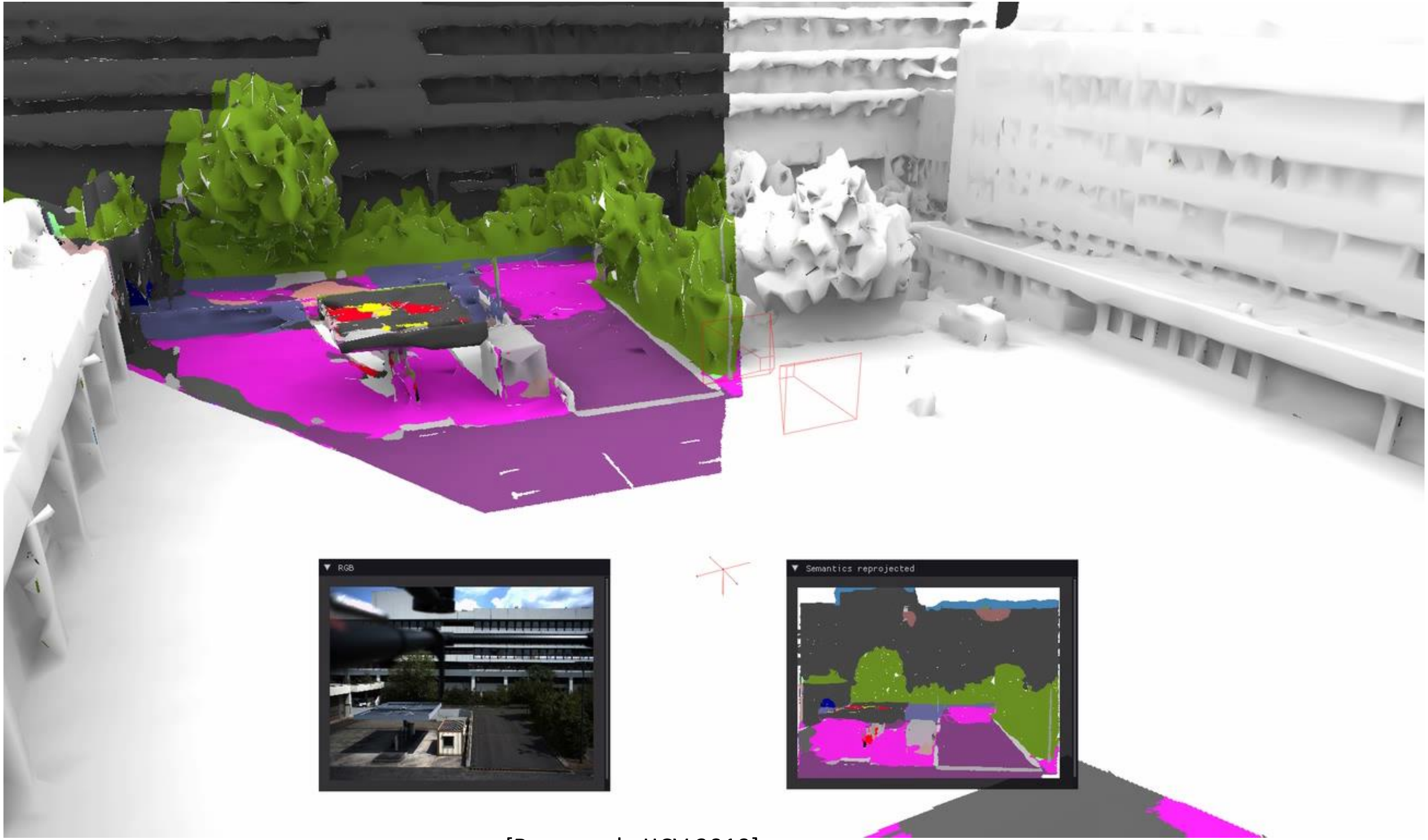


initial fusion feedback improved interpretation



[Rosu et al., IJCV 2019]

3D Semantic Mapping



[Rosu et al., IJCV 2019]

German Rescue Robotics Center

DEUTSCHES RETTUNGSROBOTIK ZENTRUM

Initial demonstrator

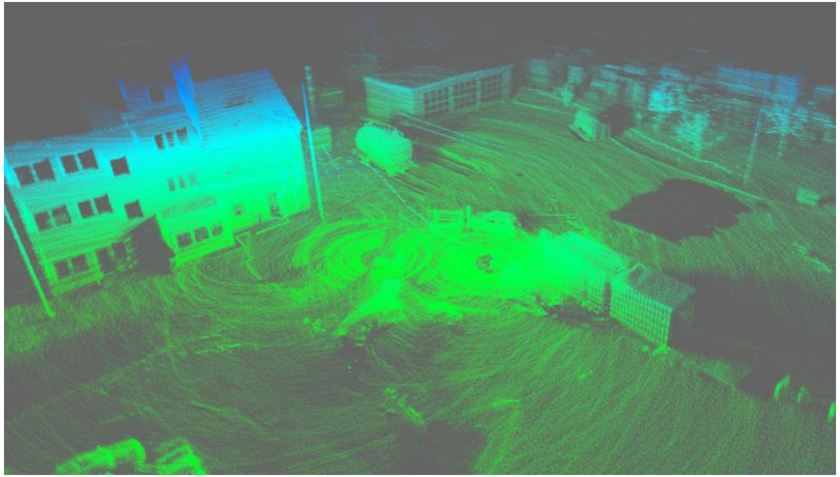
- Basis: DJI Matrice 600 Pro
- Sensors: Velodyne VLP 16, FLIR Boson, 2x FLIR BlackFly S
- Tiltable sensor head

Current demonstrator

- Basis: DJI Matrice 210 v2
- Sensors: Ouster OS-0, FLIR AGX, 2× Intel RealSense D455
- IP43 water resistance

Supporting Fire Fighters (A-DRZ)

- Added thermal camera
- Flight at Brandhaus Dortmund



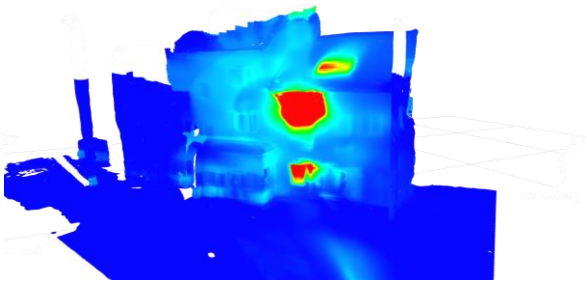
[Rosu et al. SSRR 2019]

Mesh-based 3D Modeling + Textures

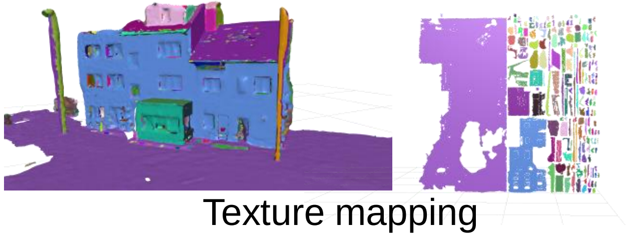
- Model 3D geometry with mesh
- Appearance and temperature as high-resolution texture


Mesh geometry


RGB texture

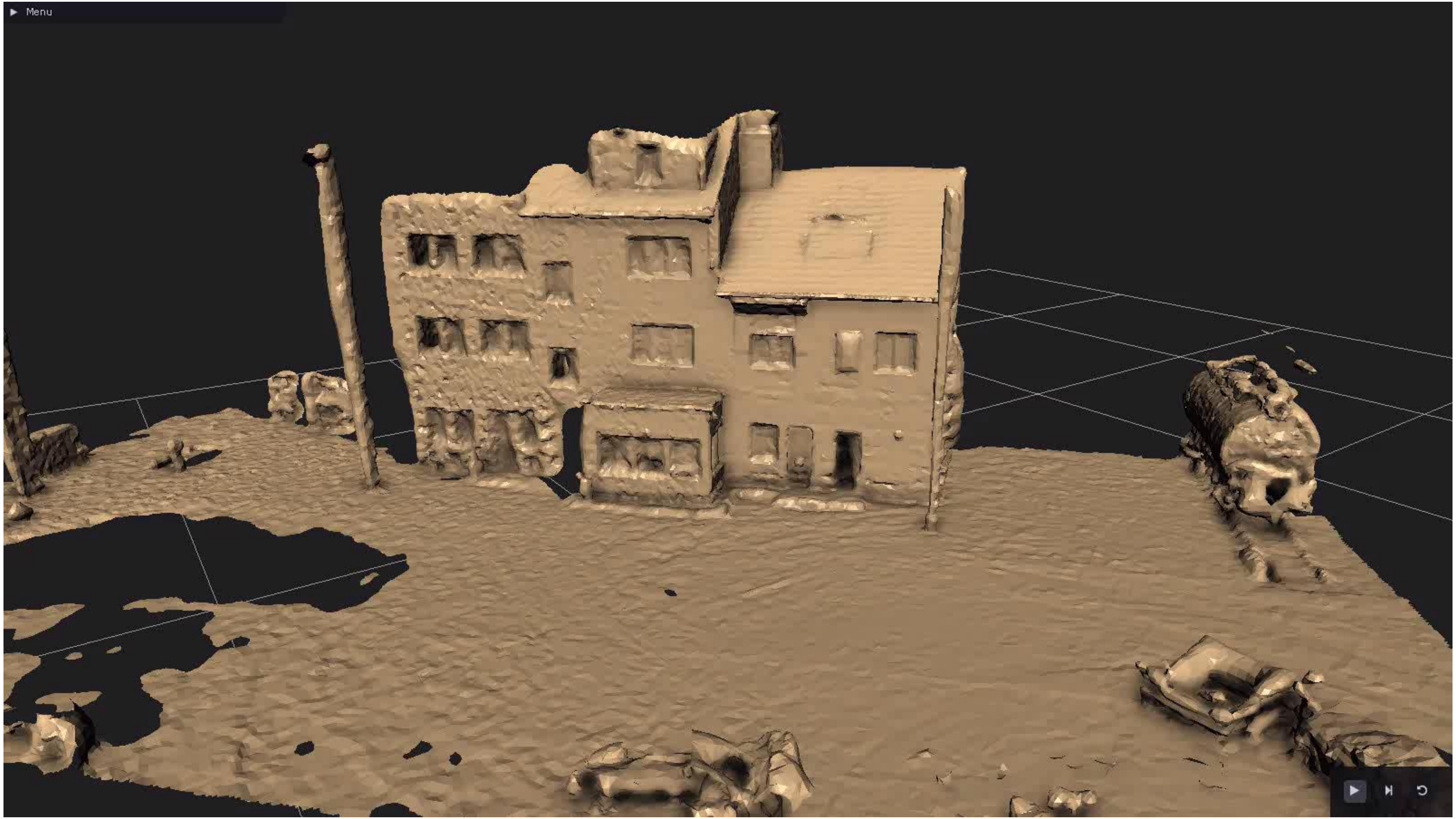

Thermal texture

- Mapping from 3D mesh to 2D texture


Texture mapping

[Rosu et al. SSRR 2019]

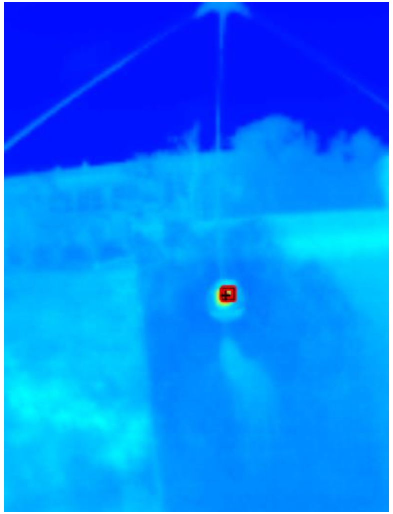
Modeling the Brandhaus Dortmund

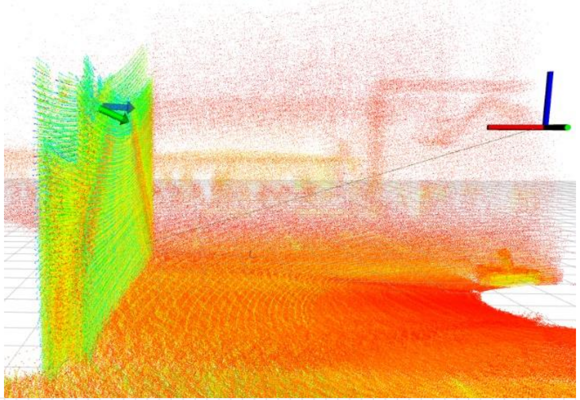


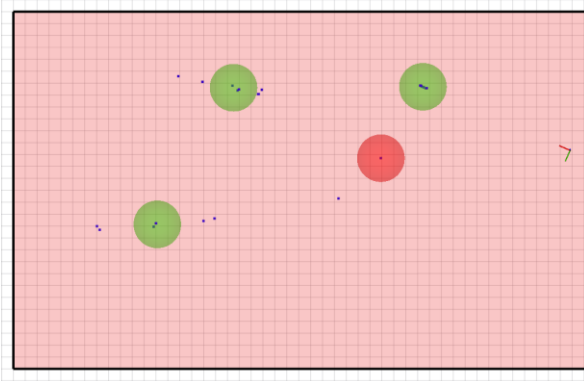
[Rosu et al. SSRR 2019]

Multi-hypothesis Tracking of Fire Detections

- Aggregation of egocentric fire detections to filtered allocentric fire hypotheses
- Integration of 2D detections (direction vector) by ray-casting and of 3D detections





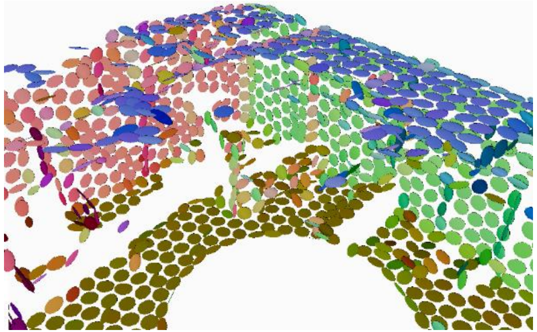


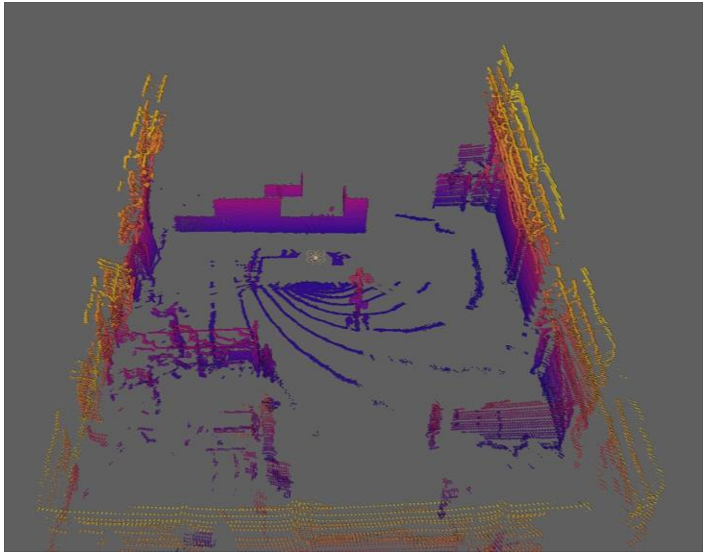
[Quenzel et al. ICUAS 2021]

Real-time LiDAR Odometry with Continuous-time Trajectory Optimization

- Simultaneous registration of multiple multiresolution surfel maps using Gaussian mixture models and temporally continuous B-spline
- Accelerated by sparse permutohedral voxel grids and adaptive choice of resolution
- Real-time onboard processing 16-20 Hz
- Open-Source
  https://github.com/AIS-Bonn/lidar_mars_registration

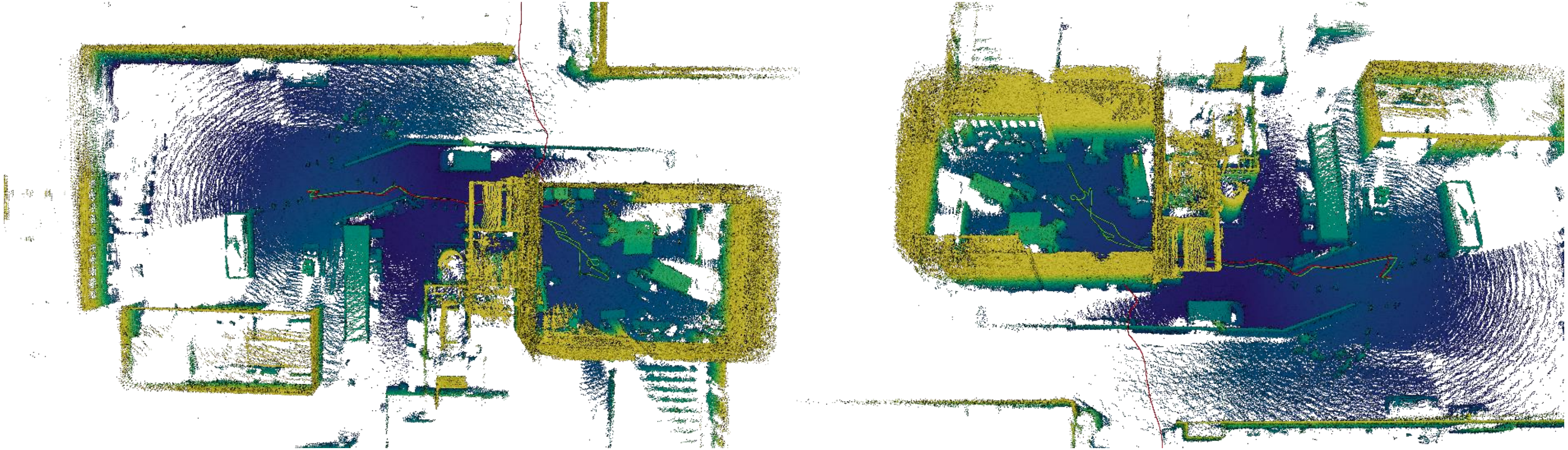
[Quenzel and Behnke, IROS 2021]





LiDAR Odometry

Minimax-Viking fire house

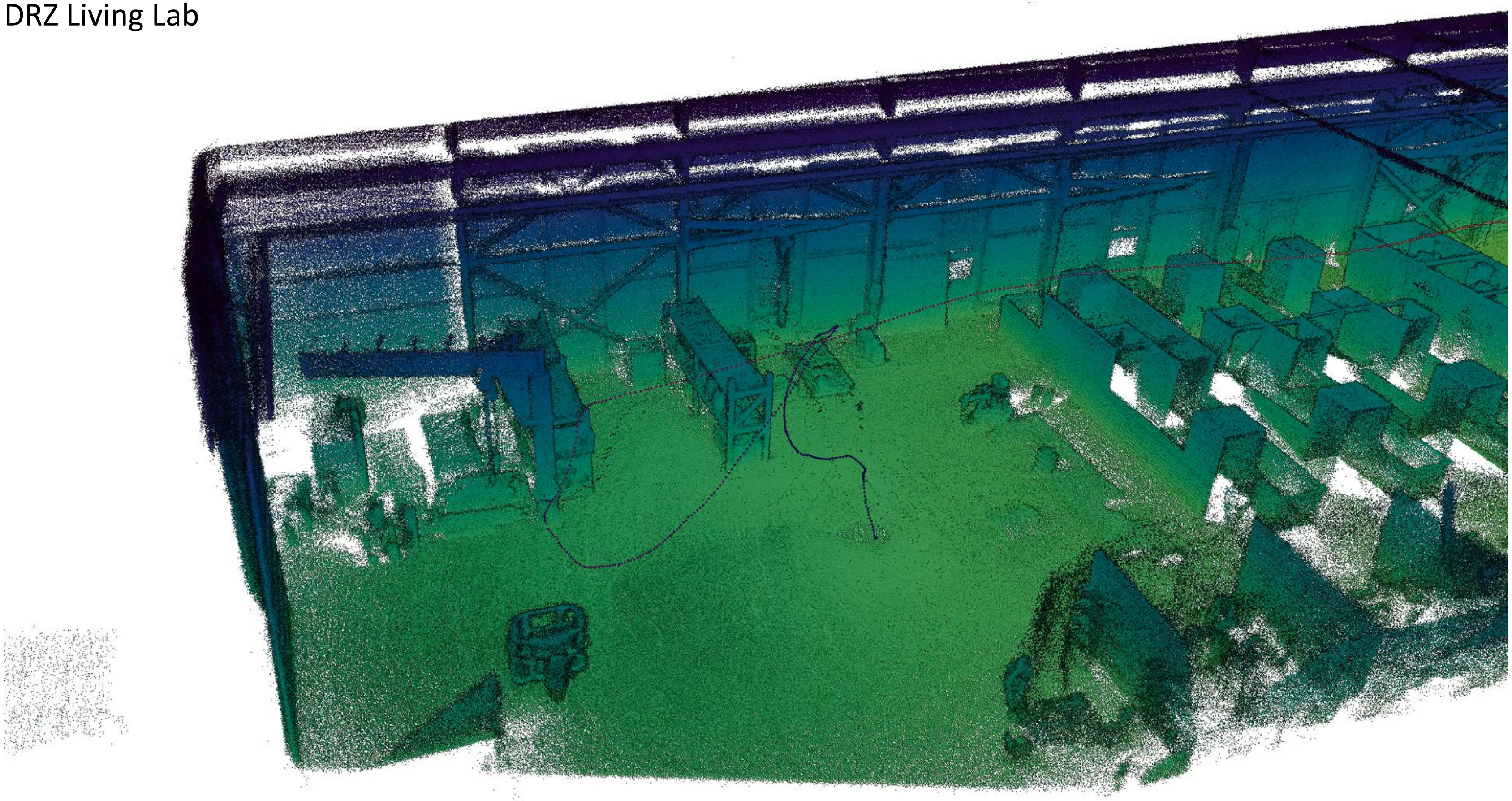


- Sliding window keyframe approach for drift reduction
- Scan fusion and moving the local map on the surfel level

[Quenzel and Behnke, IROS 2021]

3D LiDAR Mapping

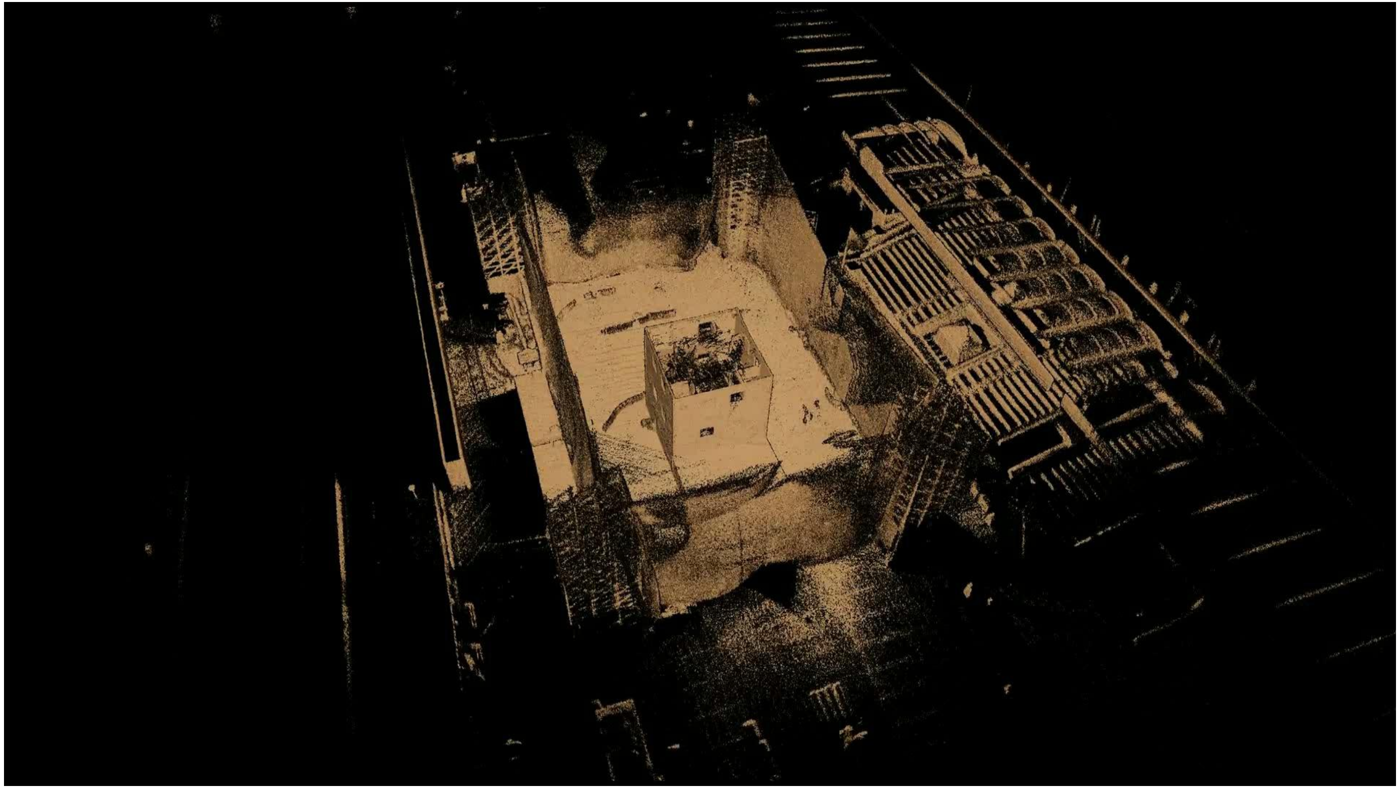
DRZ Living Lab



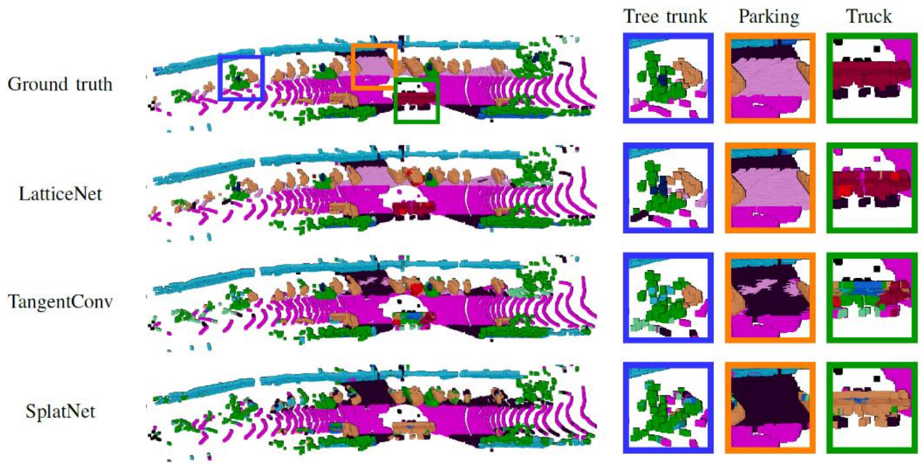
3D LiDAR Mapping

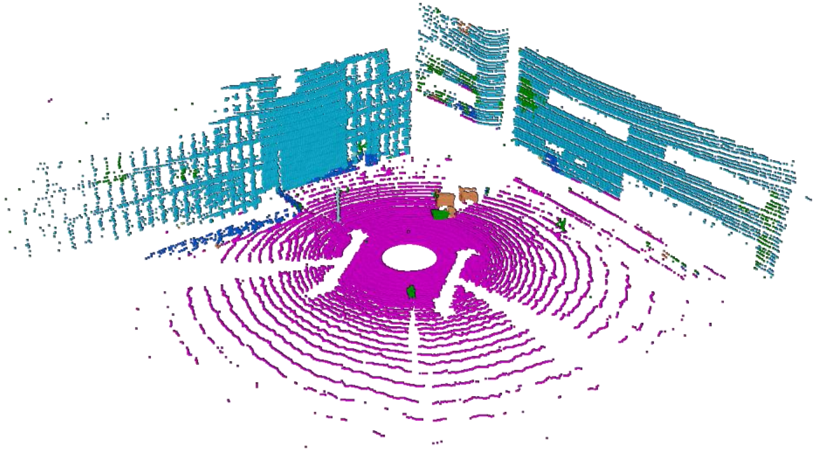
- Local mapping with position prior
- GPS offset correction for improved localization
- Dedicated outdoor and indoor maps with seamless localization switching

MBZIRC 2020



Semantic Perception: LiDAR Segmentation



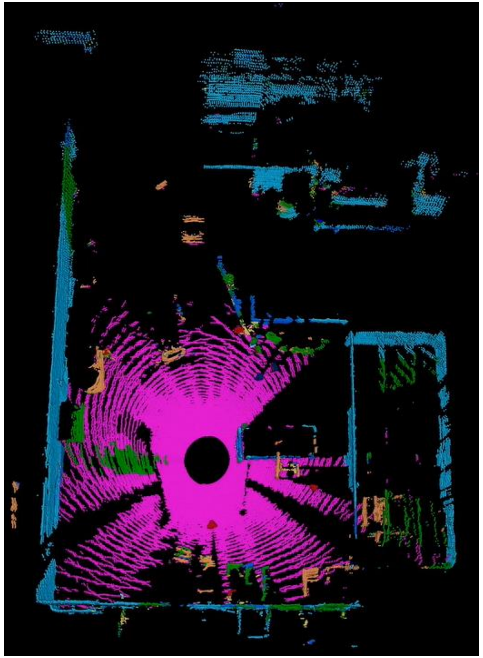


- LatticeNet segmentation of 3D point clouds based on sparse permutohedral grid
- Hierarchical information aggregation through U-Net architecture
- LatticeNet is real-time capable and achieves excellent results in benchmarks

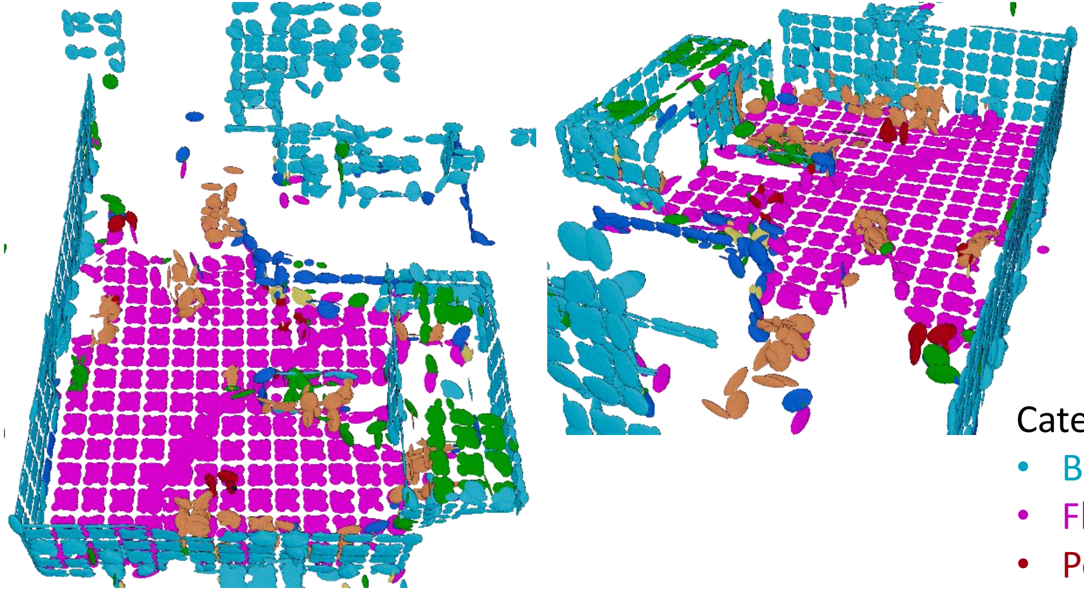
[Rosu et al., RSS 2020]

Semantic Fusion: 3D LiDAR Mapping

Minimax-Viking fire house



Segmented point cloud



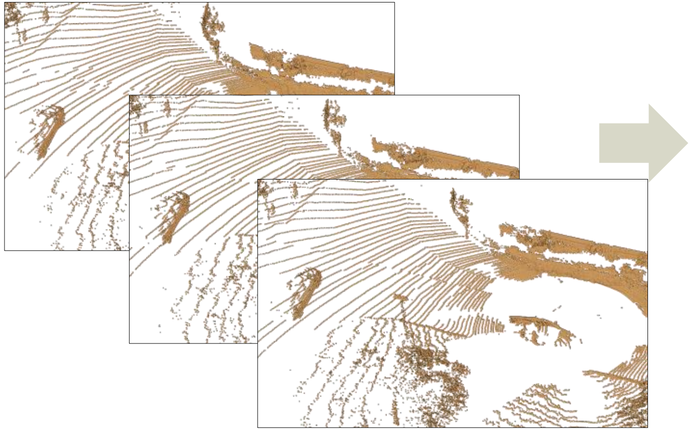
Semantic multiresolution surfel map

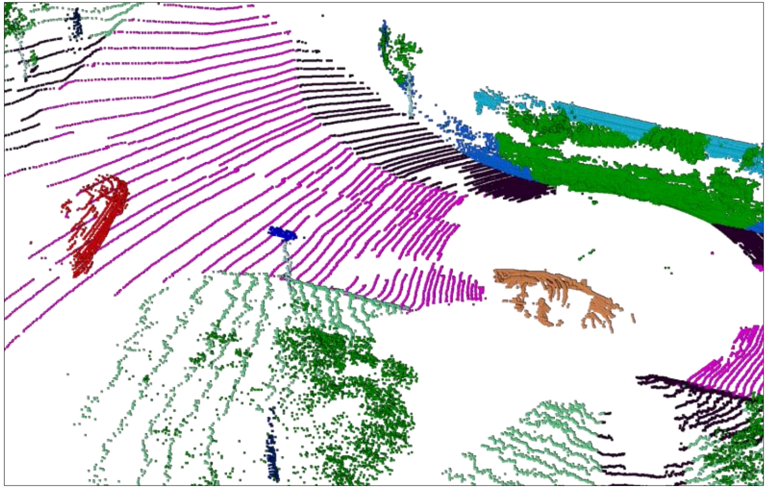
Categories:

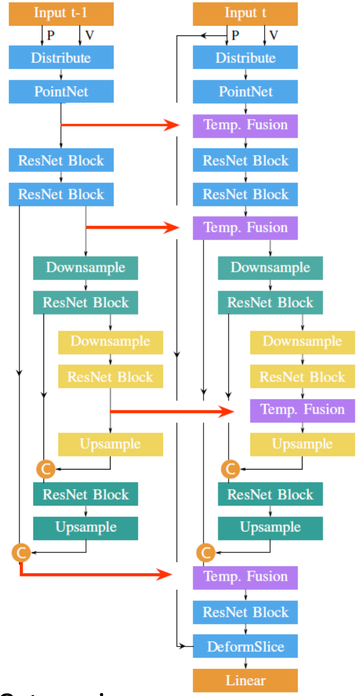
- Building
- Floor
- Persons
- Vehicles
- Fence
- Vegetation

Semantic Fusion: Temporal LatticeNet

- Semanic segmentation of sequences of 3D point clouds
- Integration of recurrent connections
- Trained on three scans of SemanticKITTI
- Distinguishing moving from parking vehicles







Categories:
- Street
- Moving Vehicle
- Parking Vehicle
- Vegetation

[Rosu et al. Autonomous Robots 2021]

Semantic Perception: Camera-based Segmentation + Detection

RGB image

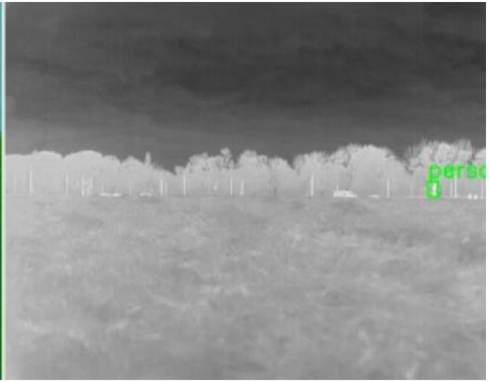
Semantic segmentation with overlaid detections at the DRZ integration sprint in Bad Oldesloe, Germany

- Pixel-wise semantic segmentation and object detection with Google Edge TPU
- Detection of e.g. buildings, vegetation etc. (DeepLab v3 CNN with MobileNet v3 Backbone)

Semantic Perception: Detection of Persons and Vehicles

RGB image

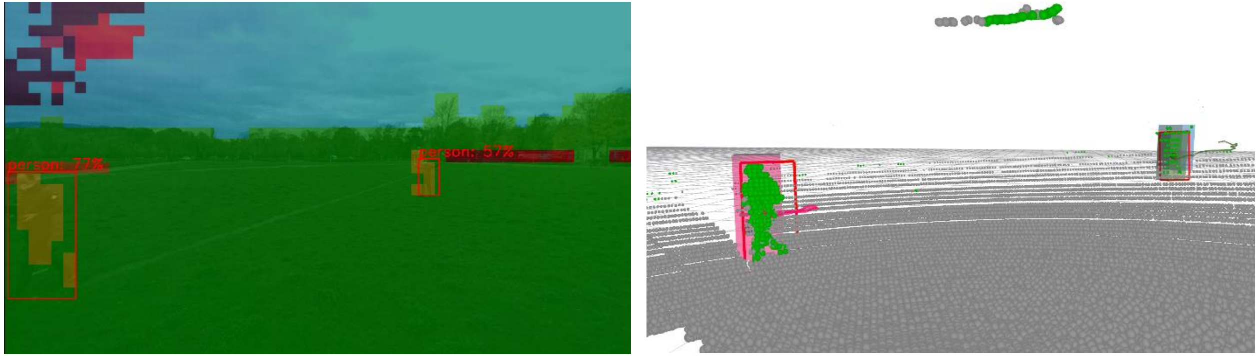
Semantic segmentation



Person detection in thermal images

- Detection of persons and vehicles in color and thermal images (SSD with MobileNet v3 backbone)
- Runs on board computer with approx. 5 fps

Multi-hypothesis Tracker for Dynamic Objects



- Multi-hypothesis tracker for combining detected objects from image and LiDAR
- Segmentation of LiDAR scan into foreground and background with subsequent grouping of foreground segments of adjacent scan lines and person detection
- 2D image detections + depth camera to derive a 3D detection hypothesis
- Movement of individual instances can be predicted

[Razlaw et al., ICRA 2019]

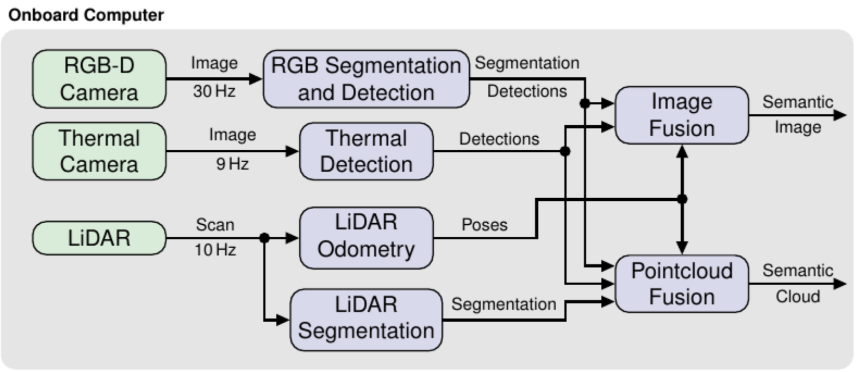
Semantic Perception: Synthesis of Training Data

- Identification of relevant object categories with DRZ partners IFR, FwDo and DFKI
- Review of available data sets
- Generation of synthetic training data with physics-based renderer EasyPBR

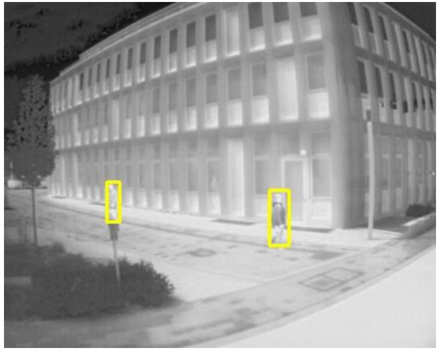
[Rosu and Behnke, GRAPP 2021]

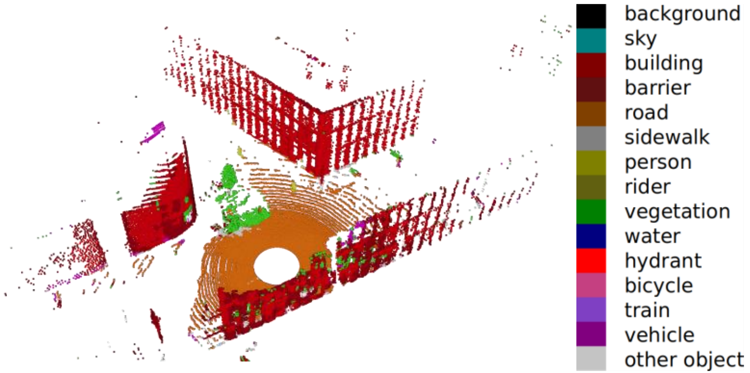
Onboard Multimodal Semantic Fusion

- Real-time semantic Segmentation and Object detection (≈9Hz) with EdgeTPU / iGPU
  - SalsaNext for LiDAR
  - DeepLabv3 for RGB images
  - SSD MobileDet for Thermal/RGB



- Late-Fusion for
  - Point cloud
  - Image segmentation

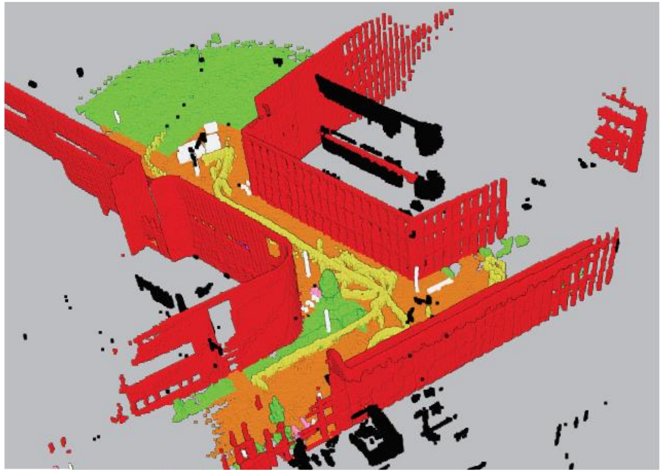


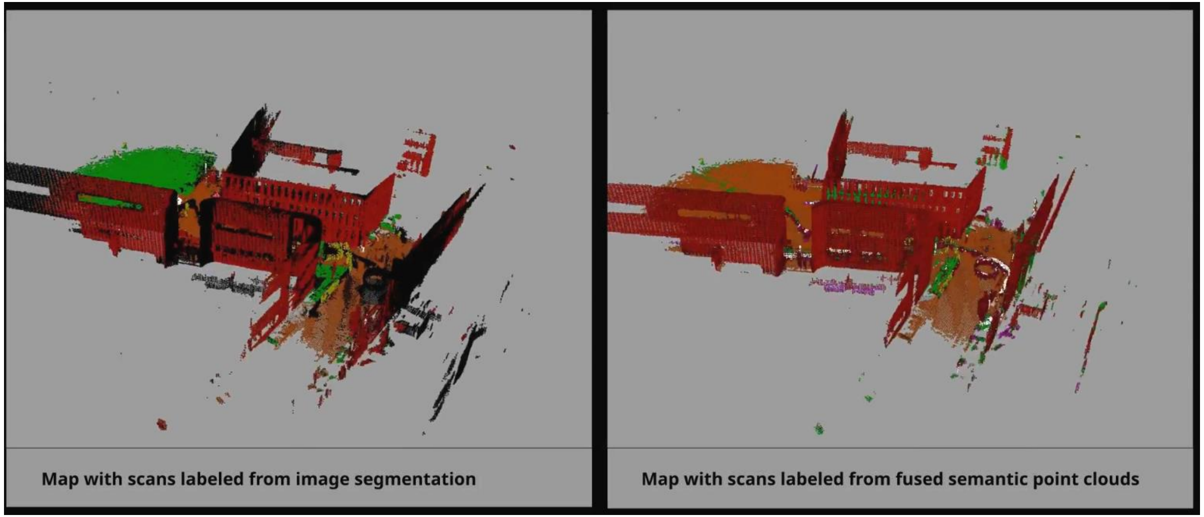


[Bultmann et al. ECMR 2021]

Onboard Multimodal Semantic Fusion

- Bayesian fusion of class probabilities in sparse voxel grid

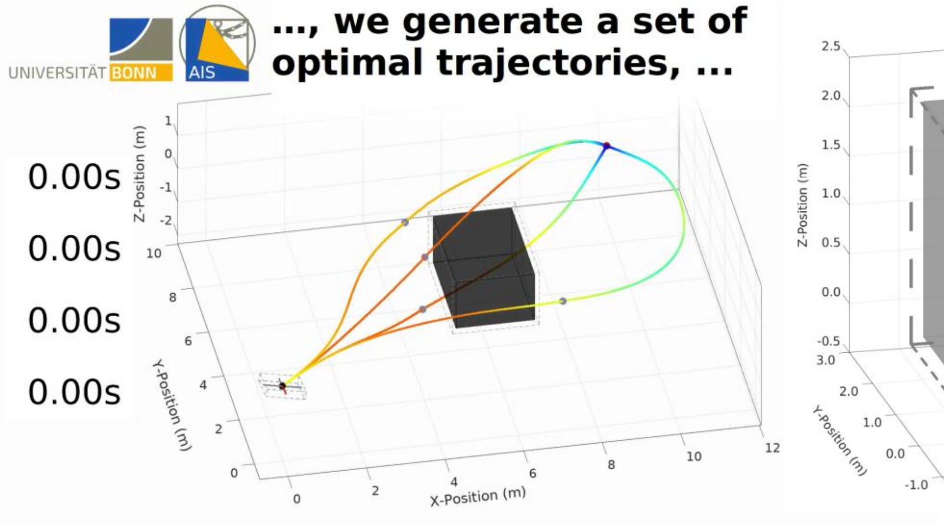


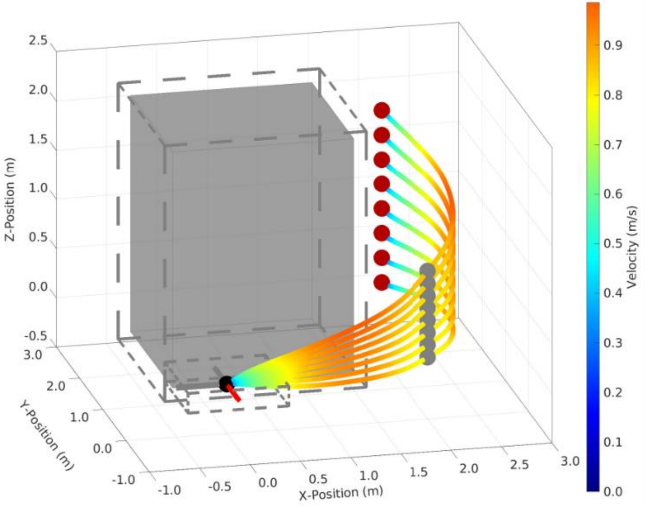


[Bultmann et al. ECMR 2021]

Optimal Obstacle Avoidance Trajectories

- Fast avoidance of immediately perceived obstacles (persons, birds, copters, ...)
- Modeling of dynamic obstacles with assumption of constant speed



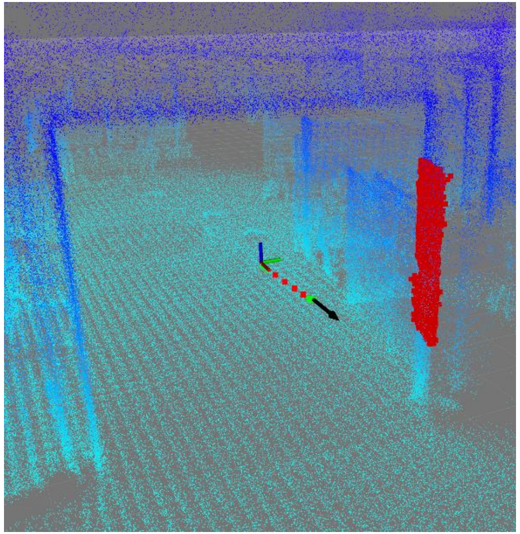


Optimale Ausweichtrajektorien um statische Hindernisse

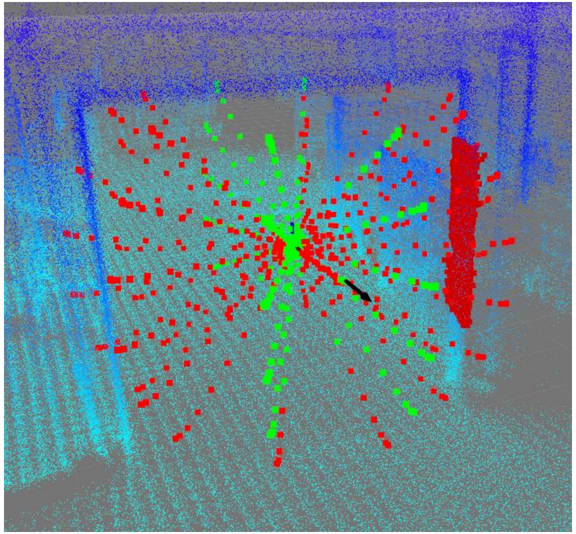
[Beul and Behnke, IROS 2019]

LiDAR-based Obstacle Avoidance

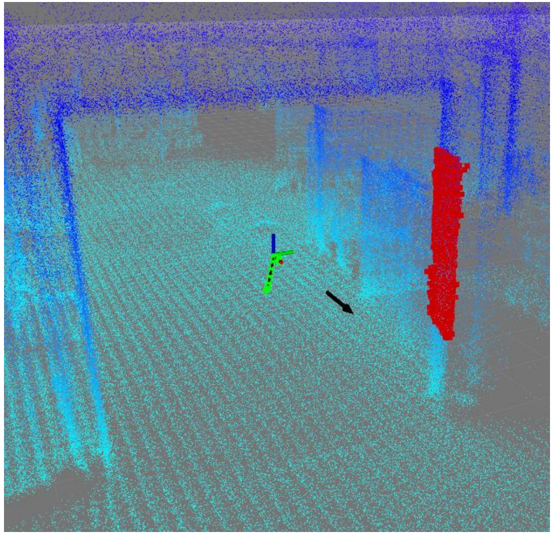
- Fast analytical collision check with 3D point cloud
- Planning of alternative trajectories if original trajectory causes collision
- Selection and execution of a collision-free alternative trajectory



Collision check



Generation of alternative trajectories

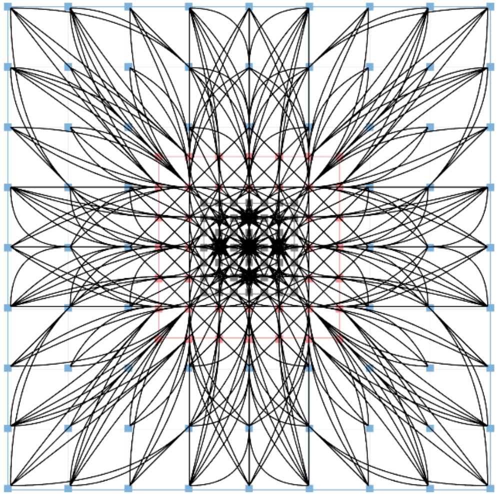


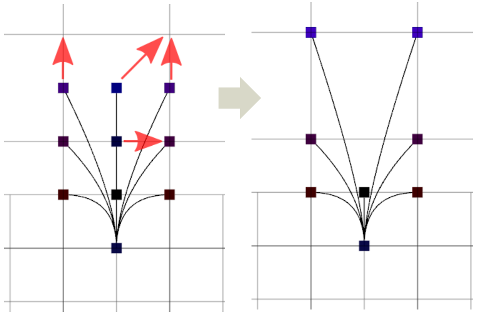
Selection based on distance to target and previous trajectory

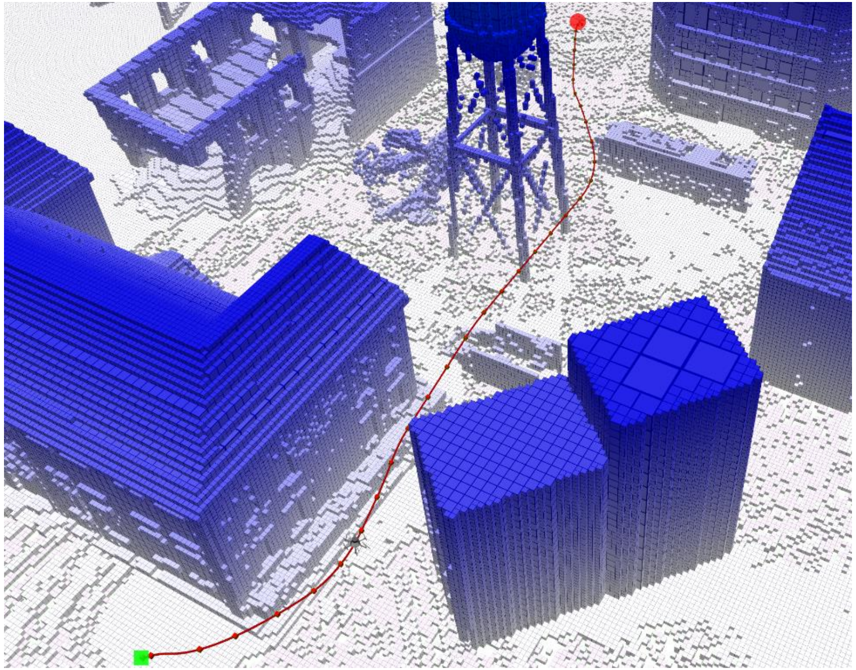
[Beul and Behnke, SSRR 2020]

Dynamic 3D Navigation Planning

- Positions and velocities in sparse local multiresolution grid
- Adaptation of movement primitives to grid
- Optimization of flight time and control costs
- 1 Hz replanning



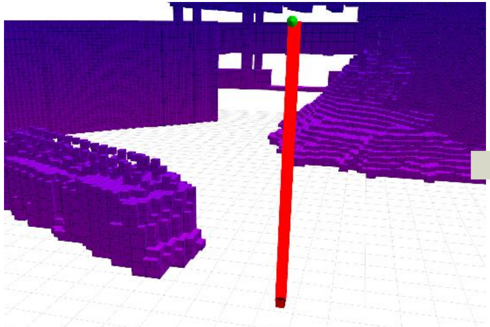




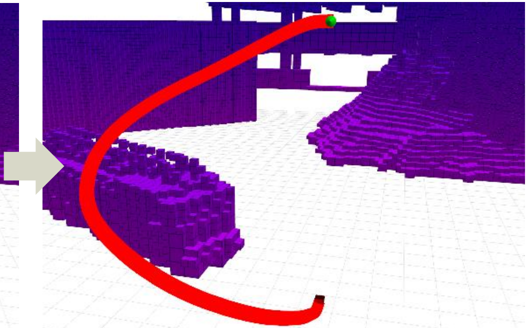
[Schleich and Behnke, ICRA 2021]

Planning with Visibility Constraints

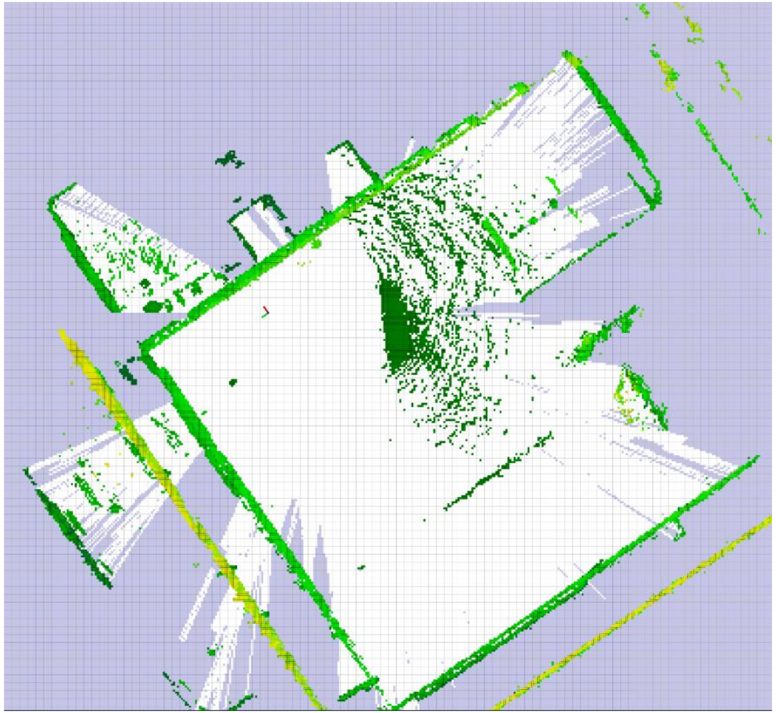
- Extra costs for flight through unmapped volumes
- Consideration of sensor frustum:
  - Coupling of vertical and horizontal motion
  - Preferred forward flight with limited rotational speed



Initial trajectory



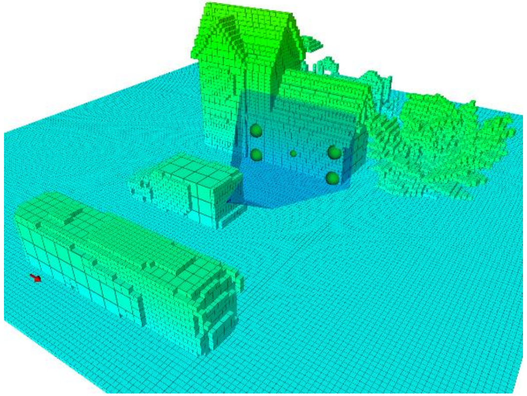
Optimized trajectory



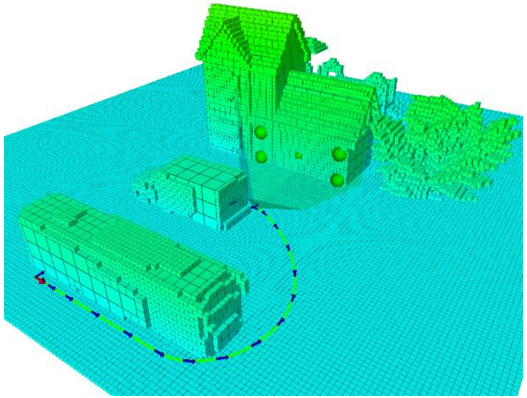
Obstacle map

Observation Pose Planning

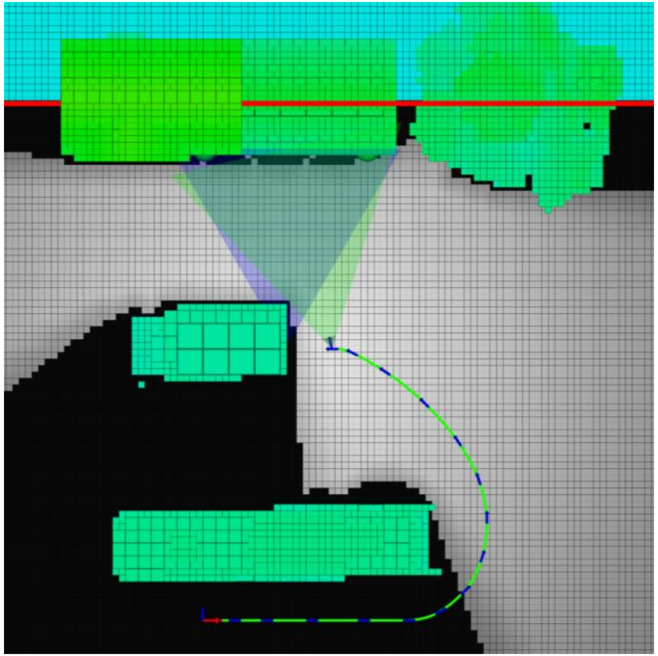
- Planning of observation poses with line of sight to the target object despite occlusions
- Target objects are defined by position, line of sight and distance
- Optimization of observation poses with regard to visibility quality and accessibility



Initial observation pose



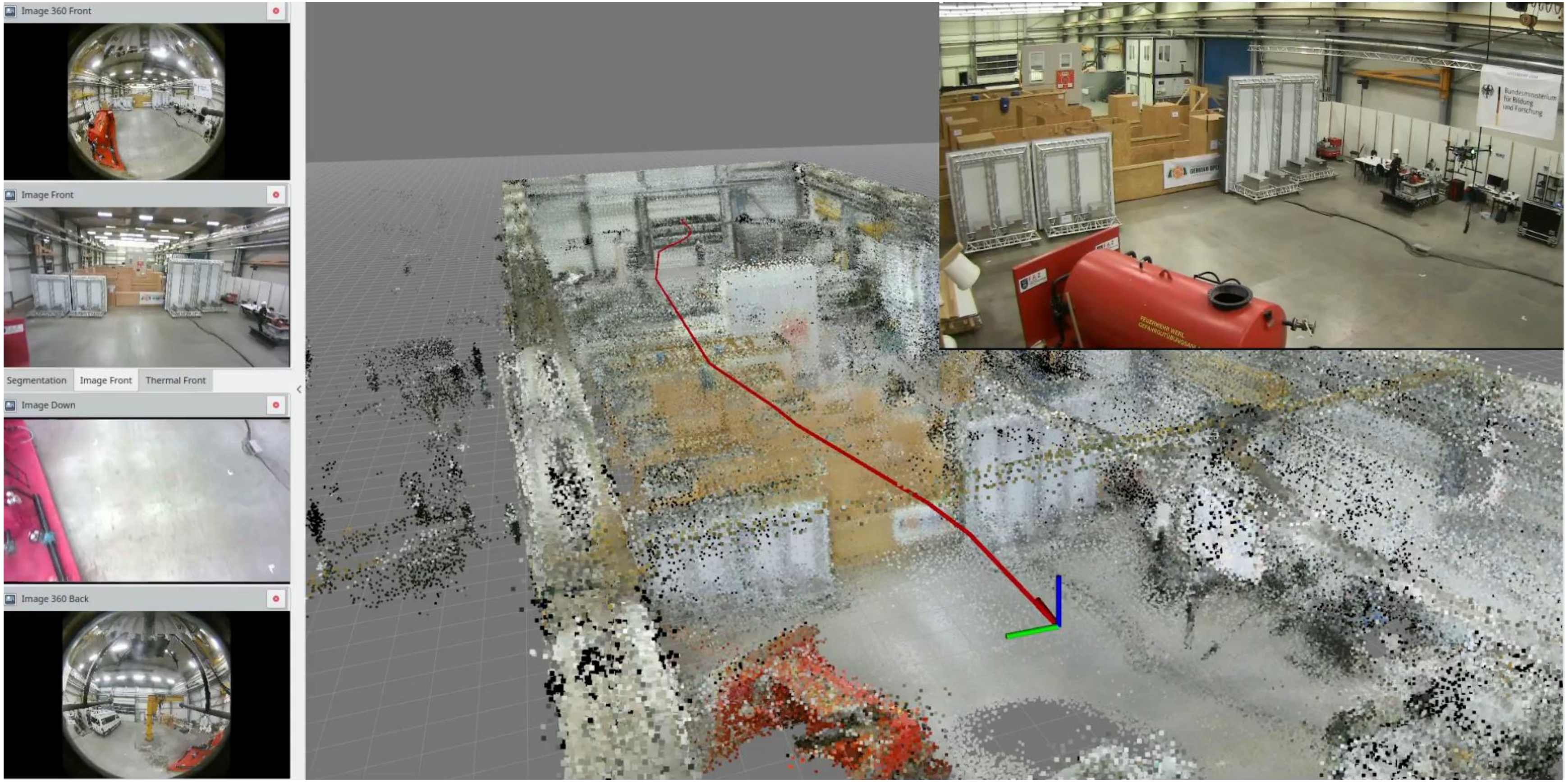
Optimized path



Top-down view

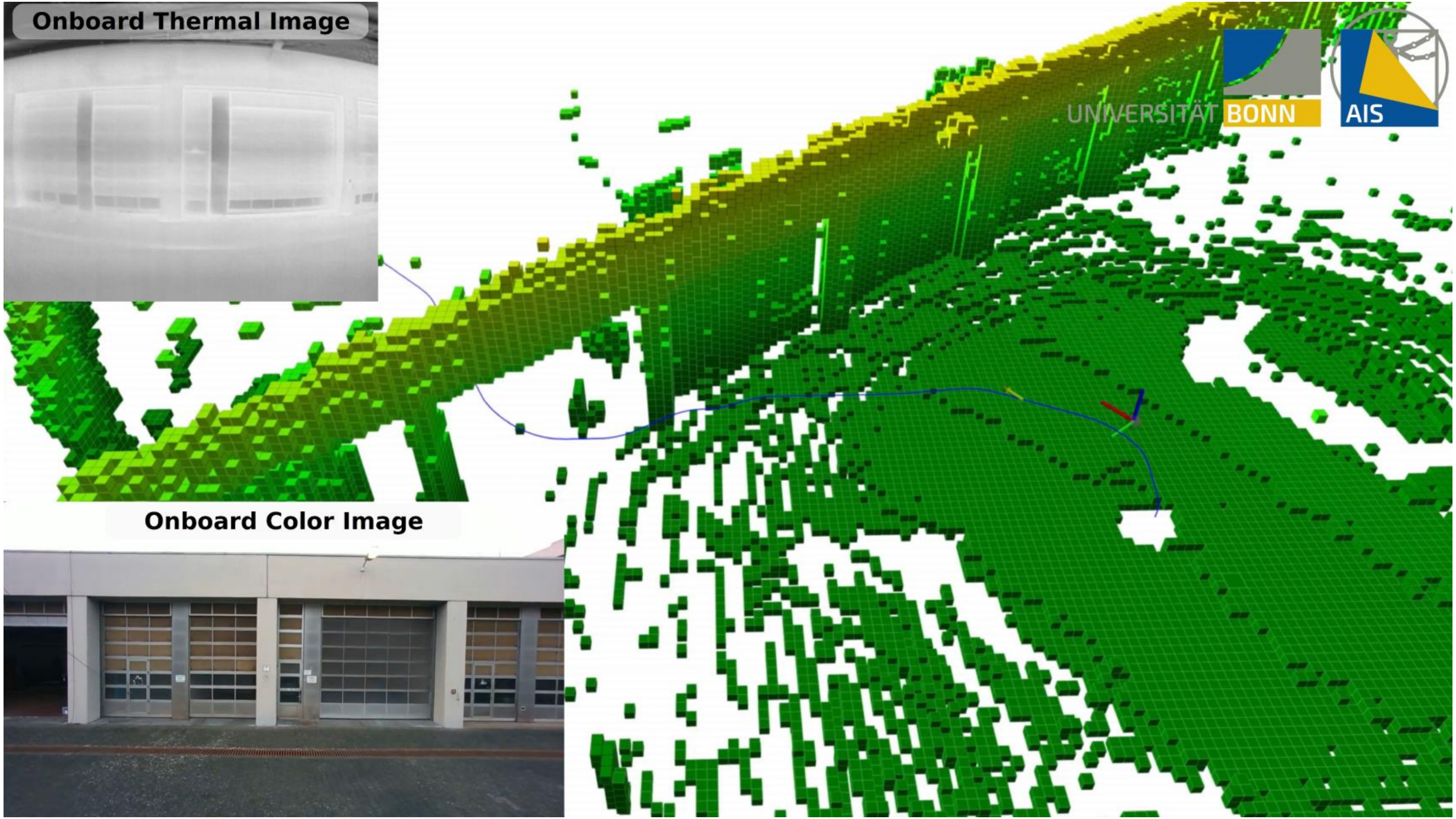
UNIVERSITÄT BONN AIS

Autonomous Flight without GNSS



DRZ Dortmund

Autonomous Flight without GNSS for Disaster Examination



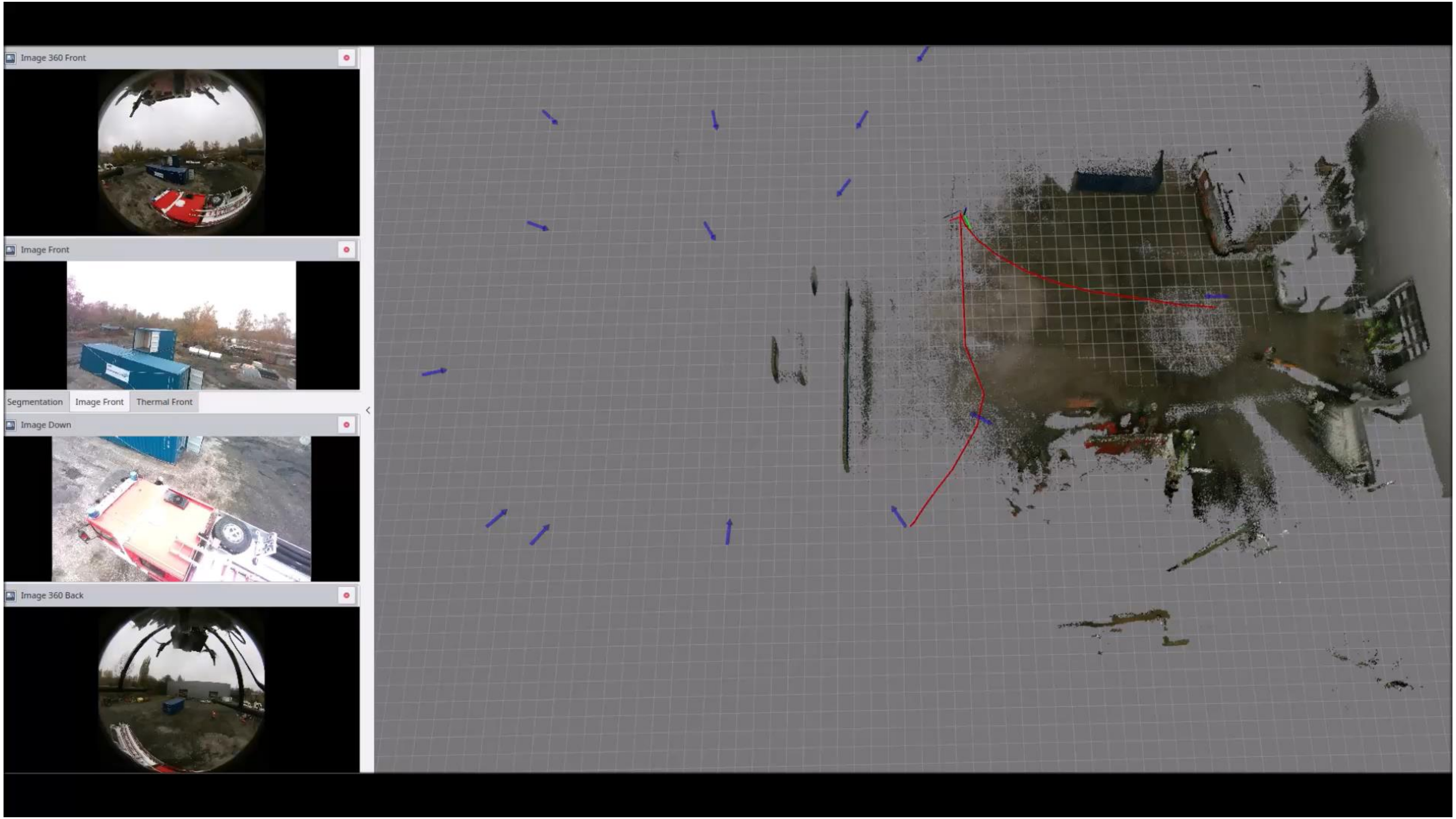
[Schleich et al., ICUAS 2021]

Exploration

- Definition of target area w.r.t. sattelite images or steet
- Simple exploration patterns (spirals, meanders, …)
- Collision check
- TSP to determine segment sequence
- Continous replanning

Campus Poppelsdorf

Autonomous Exploration



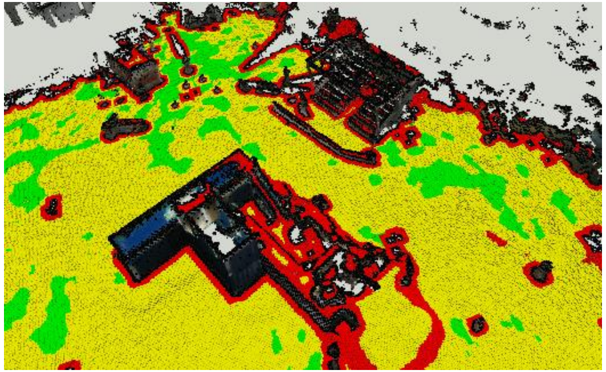
DRZ Dortmund

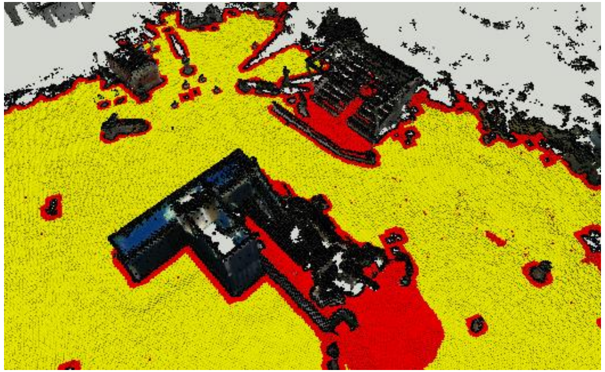
Terrain Classification for Traversability

- Based on voxel-filtered aggragated point cloud
- Terrain classification based on local height differences in the robot ground robot footprints
- Categories: drivable, walkable, unpassable
- Reachability analysis


Aggregated colored point cloud


Local height differences


Terrain category


Reachability

[Schleich et al., ICUAS 2021]

Conclusions

- Developed capable robotic systems for challenging scenarios
  - Humanoid soccer
  - Domestic service
  - Bin picking
  - Disaster response
  - Aerial robots
- Challenges include
  - 4D semantic perception
  - High-dimensional motion planning
- Promising approaches
  - Prior knowledge (inductive bias)
  - Shared experience (fleet learning)
  - Shared autonomy (human-robot)
  - Instrumented environments