

# Umgebungswahrnehmung und Verhaltensplanung für Kognitive Roboter

**Sven Behnke**

Universität Bonn

Institut für Informatik VI –

Intelligente Systeme und Robotik

Autonome Intelligente Systeme



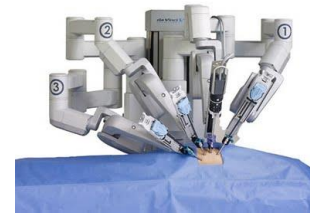
# Industrieroboter

- Kräftig, präzise, ausdauernd
- Hohe Produktivität
- Kontrollierte Umgebung erforderlich
- Unflexibel



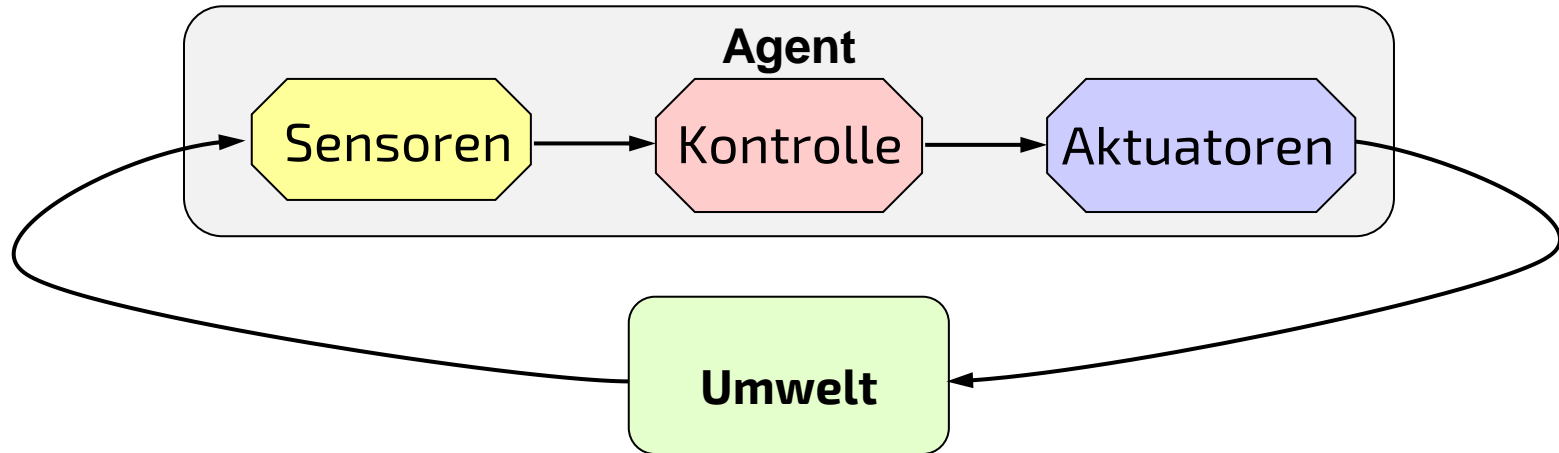
# Neue Anwendungsgebiete für Roboter

- Autonomes Fahren
- Logistik
- Landwirtschaft
- Kollaborative Produktion
- Alltagsassistentz
- Weltraum, Suche&Rettung
- Medizin, Pflege
- Unterhaltung, Spielzeuge
- **Brauchen mehr Kognition!**



# Teilprobleme

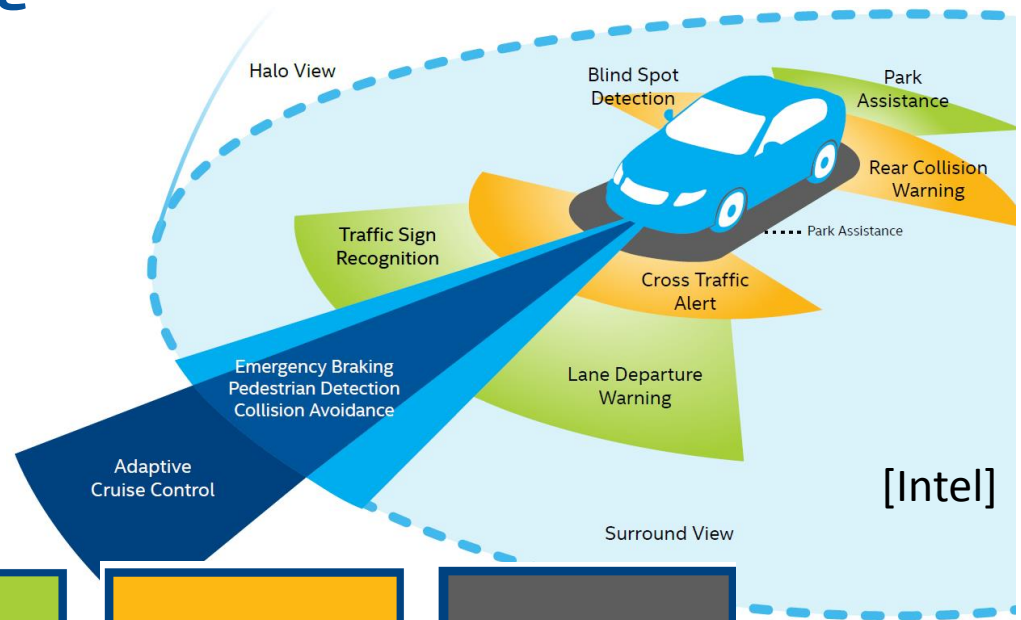
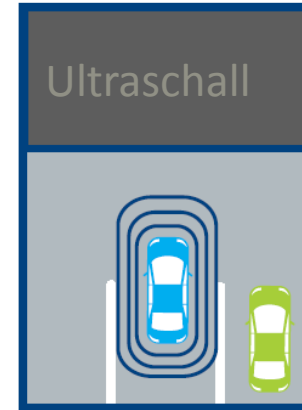
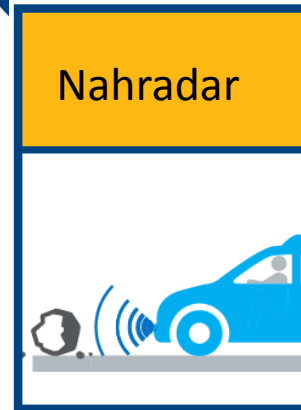
- Wahrnehmung der Umwelt
- Verhaltensplanung
- Beeinflussung der Umwelt



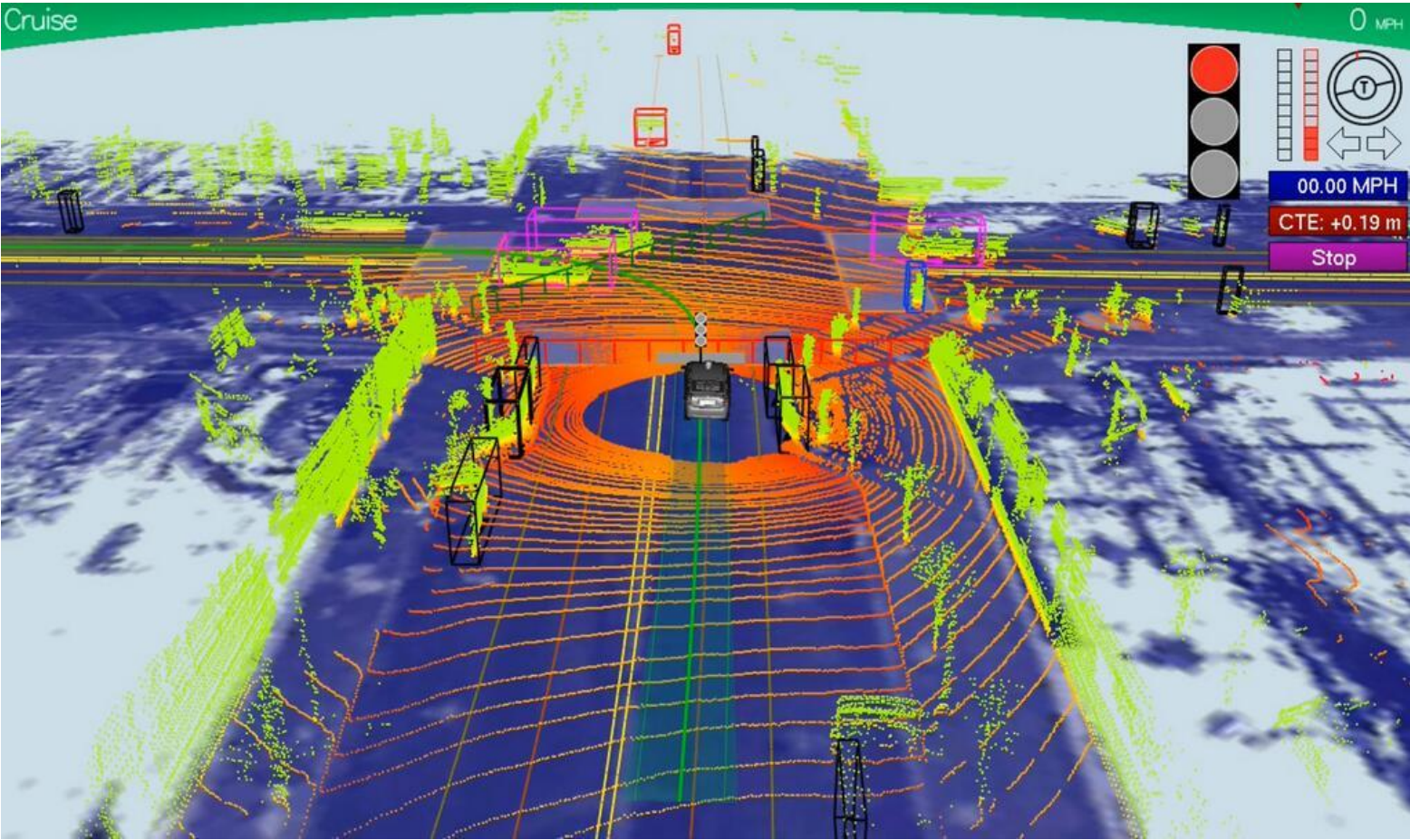
# Selbstfahrende Automobile



# Sensoren für Automobile

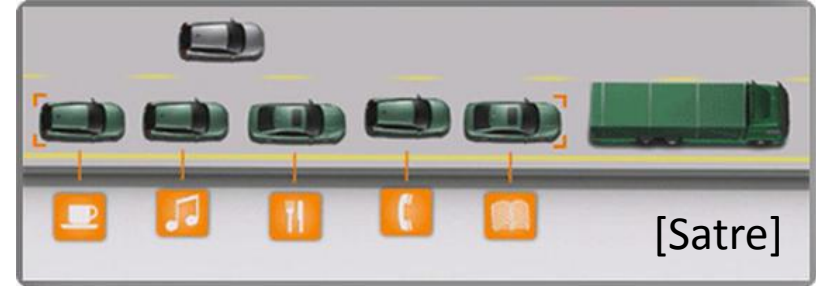


# Umgebungssicht des Google-Autos



# Möglicher Nutzen Selbstfahrender Automobile

- Höher Sicherheit
- Weniger Energieverbrauch
- Bessere Ausnutzung des Straßenraums
- Fahrer kann Zeit anders nutzen
- Preiswertere Taxis





# Einige unserer Kognitiven Roboter

- Ausgestattet mit zahlreichen Sensoren und Gelenken
- Demonstration in komplexen Szenarien



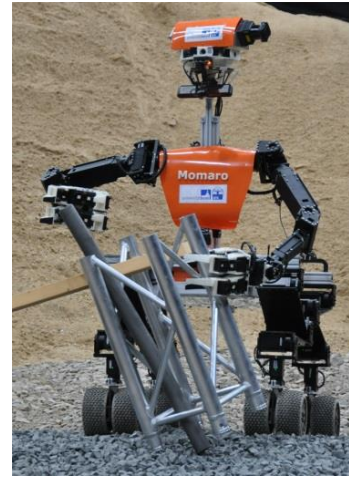
Fußball



Serviceaufgaben



Griff in die Kiste



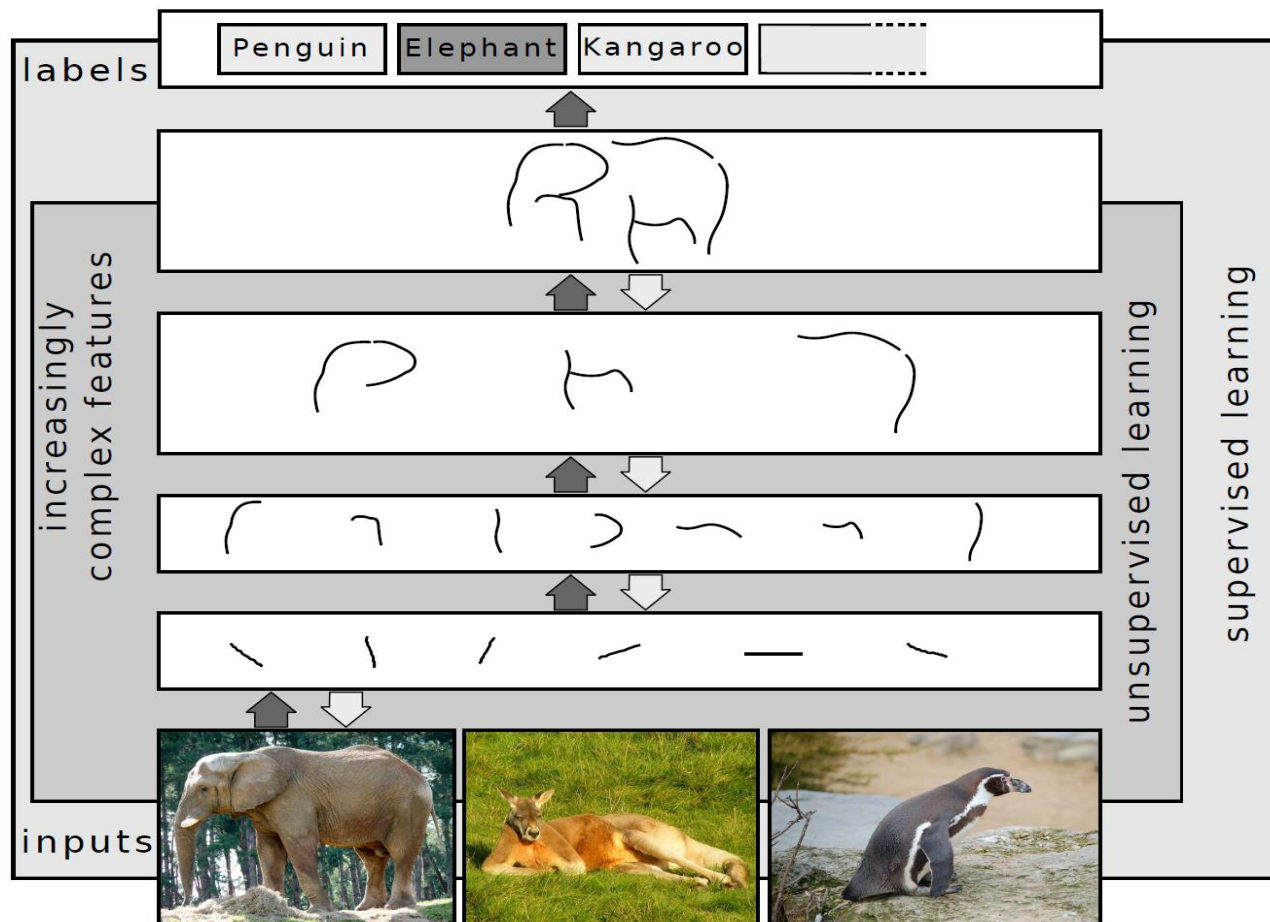
Mobile Manipulation



Inventur

# Deep Learning

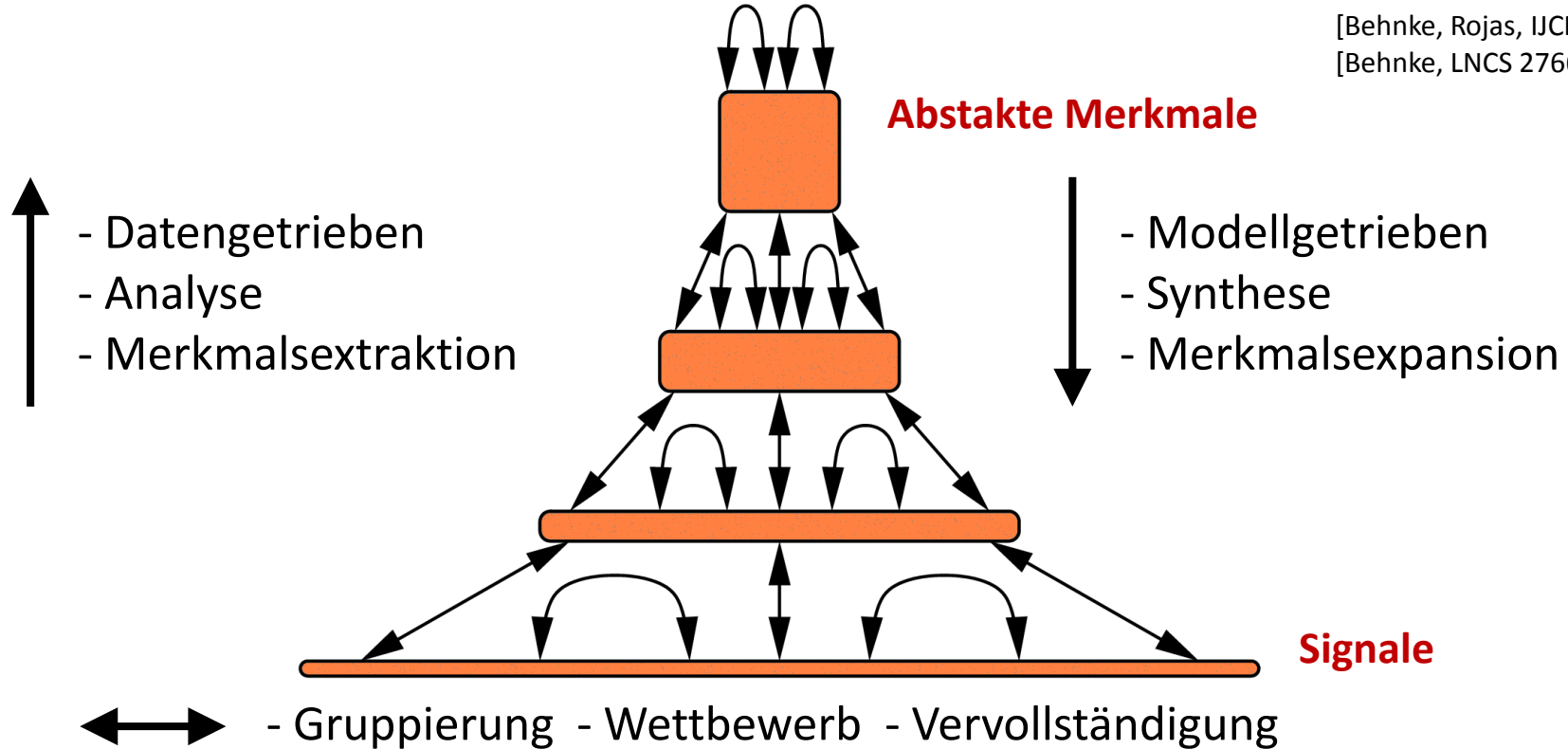
- Lernen immer abstrakterer Repräsentationen



[Schulz;  
Behnke,  
KI 2012]

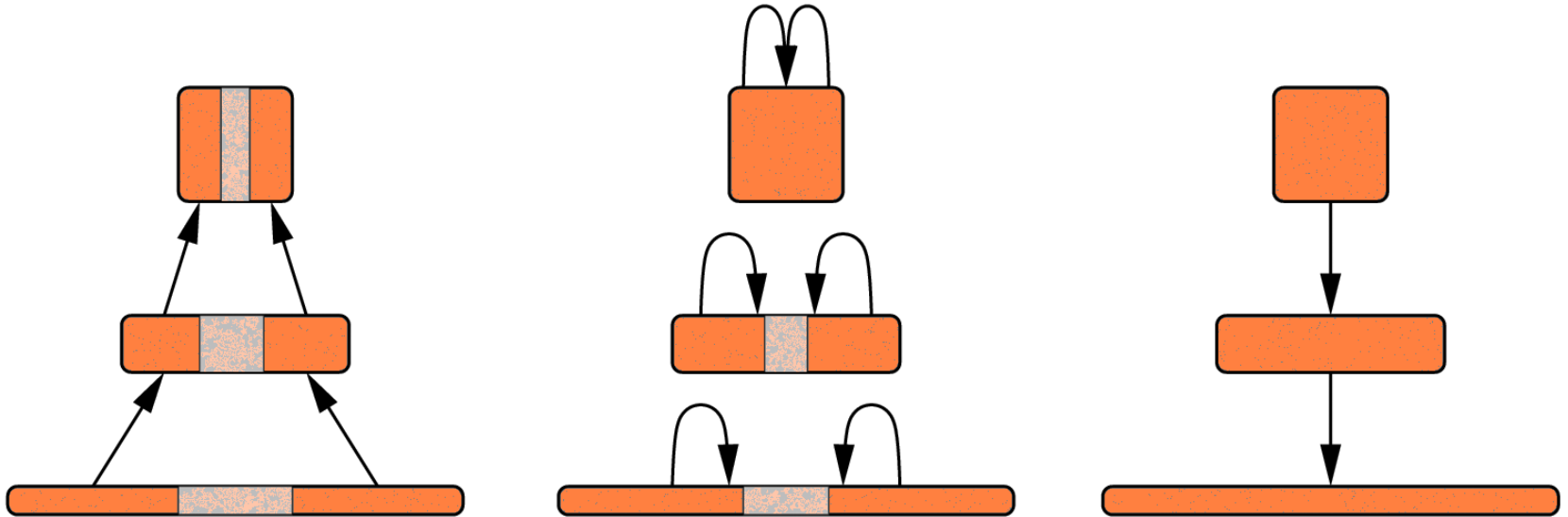
# Neuronale Abstraktionspyramide

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



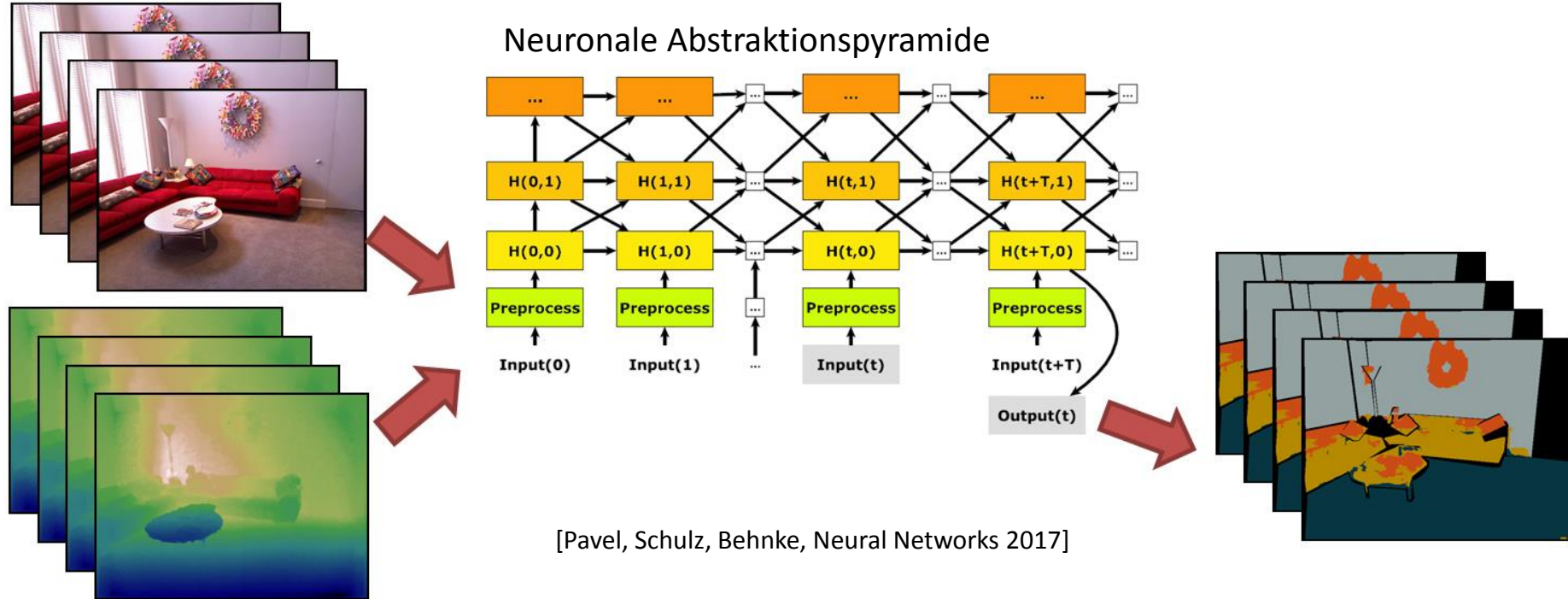
# Iterative Interpretation

- Interpretiere einfache Bereiche zuerst
- Nutze Teilinterpretationen als Kontext um lokale Mehrdeutigkeiten aufzulösen



# Neuronale Abstraktionspyramide zur Semantischen Segmentierung von RGB-D-Video

- Rekursive Berechnung effizient für zeitliche Integration

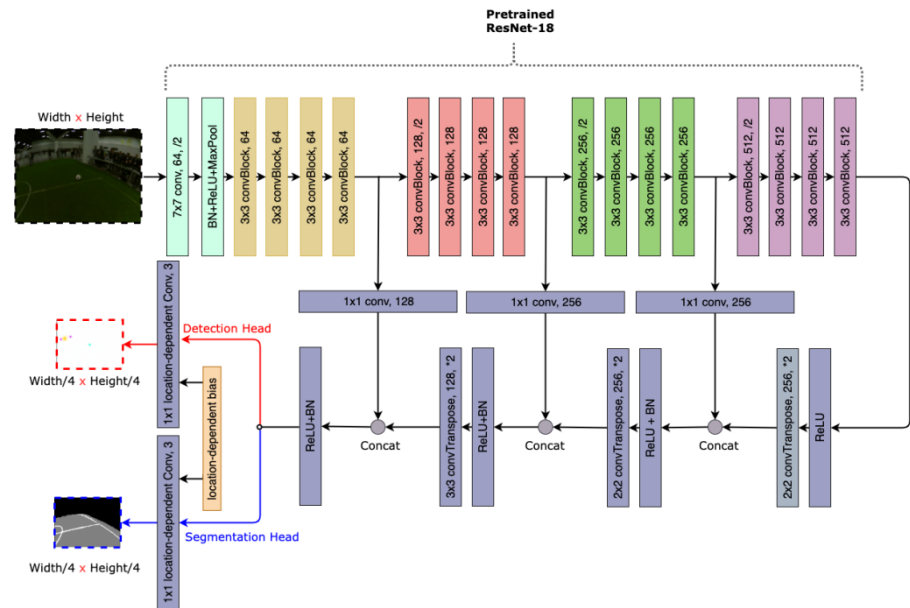
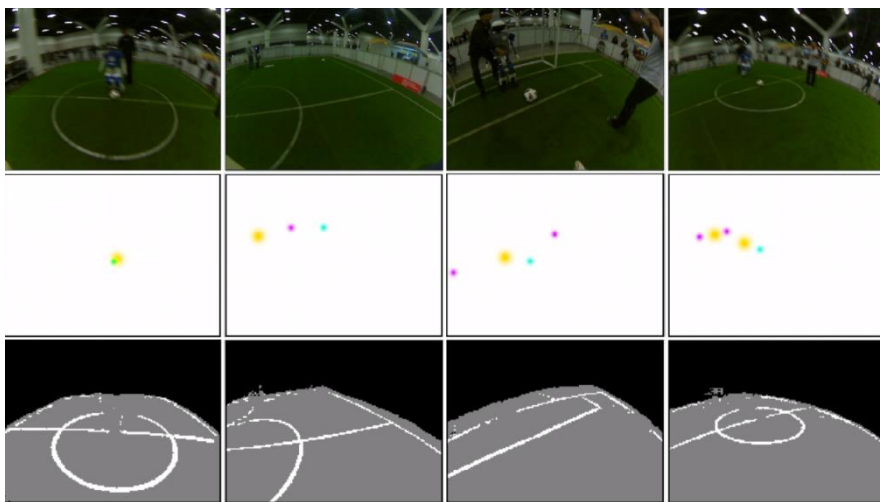


# RoboCup 2019 in Sydney



# Visuelle Wahrnehmung

- Encoder-Decoder-Netzwerk
- Zwei Ausgaben:
  - Objektdetektion
  - Semantische Segmentierung
- Ortsabhängiges Bias



- Detektiert Objekte die für Menschen schwer auszumachen sind
- Robust gegenüber Beleuchtungsänderungen

# Das Daten-Problem

- Deep-Learning Methoden benötigen viele Trainingsdaten

- Lösungsansätze:

## 1. Generieren von Daten:

Automatische Datenerfassung,  
Online-Datenmengen,  
Synthese von Szenen



## 2. Dateneffiziente Verfahren:

Objektzentrierte Modelle,  
Deformierbare Registrierung,  
Transfer-Lernen,  
Halbüberwachte Verfahren





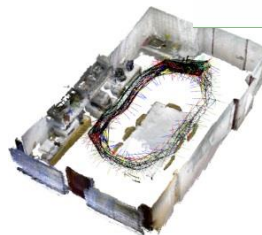
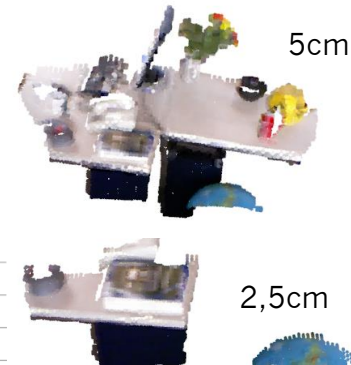
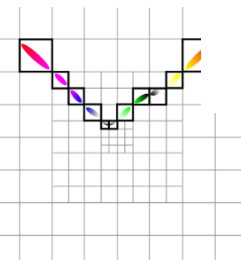
# Kognitiver Serviceroboter Cosero



# 3D-Kartierung durch RGB-D SLAM

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

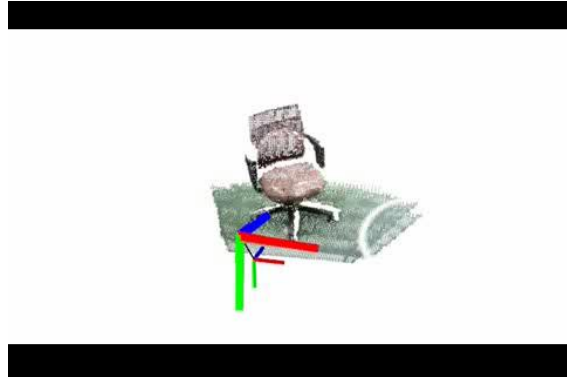
- Modellierung von Form und Farbverteilung in Voxeln
- Lokale Multiresolution
- Effiziente Registrierung mit CPU
- Globale Optimierung
- Multikamera-SLAM



[Stoucken]

# Lernen und Verfolgen von Objektmodellen

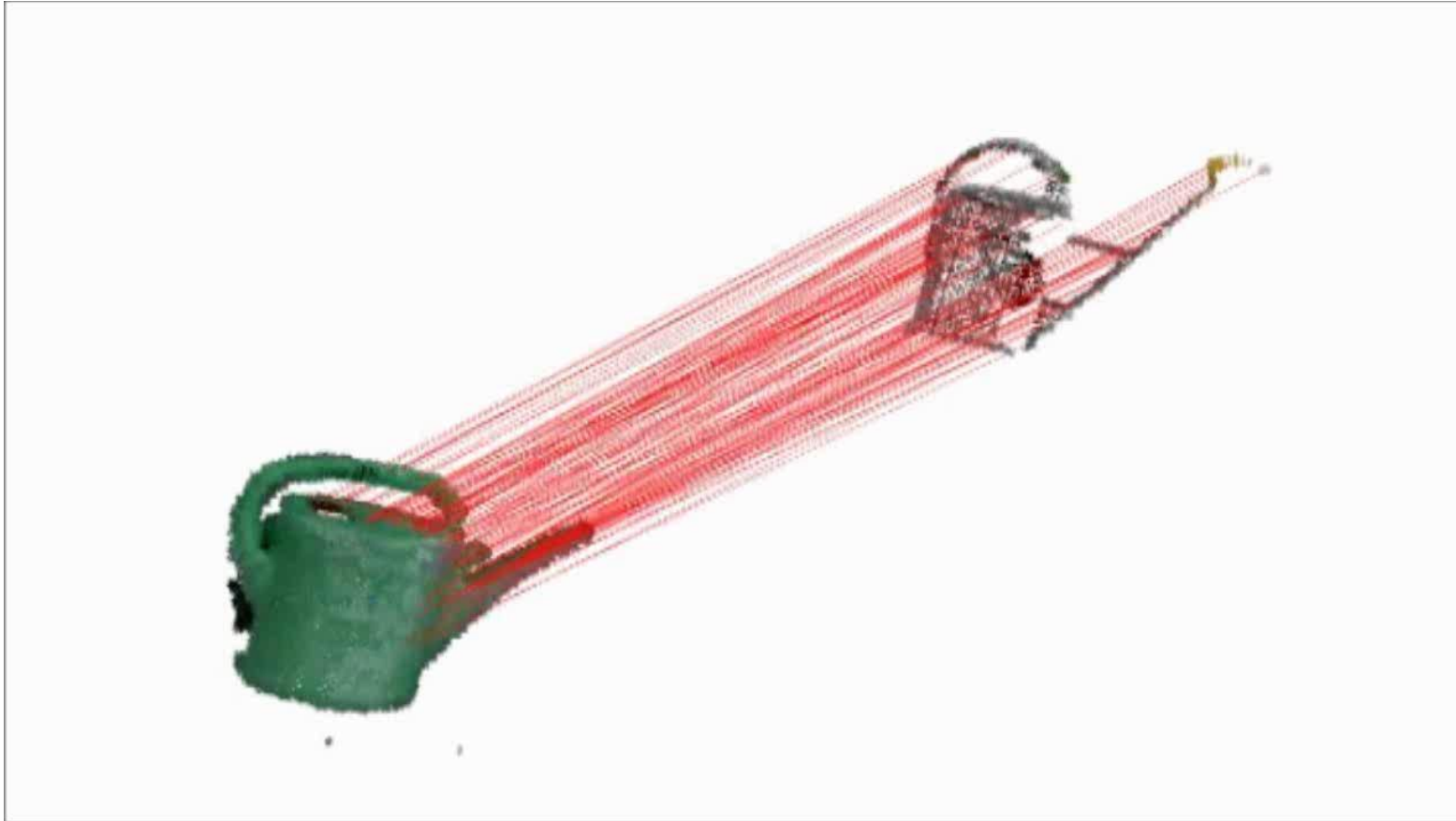
- Modellierung von Objekten durch RGB-D-SLAM



- Echtzeit-Registrierung mit aktuellem RGB-D-Frame



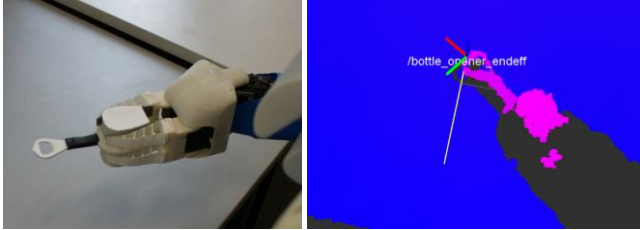
# Transfer von Handhabungswissen



[Stückler,  
Behnke,  
ICRA2014]

# Werkzeuggebrauch: Flaschenöffner

- Wahrnehmung der Werkzeugspitze



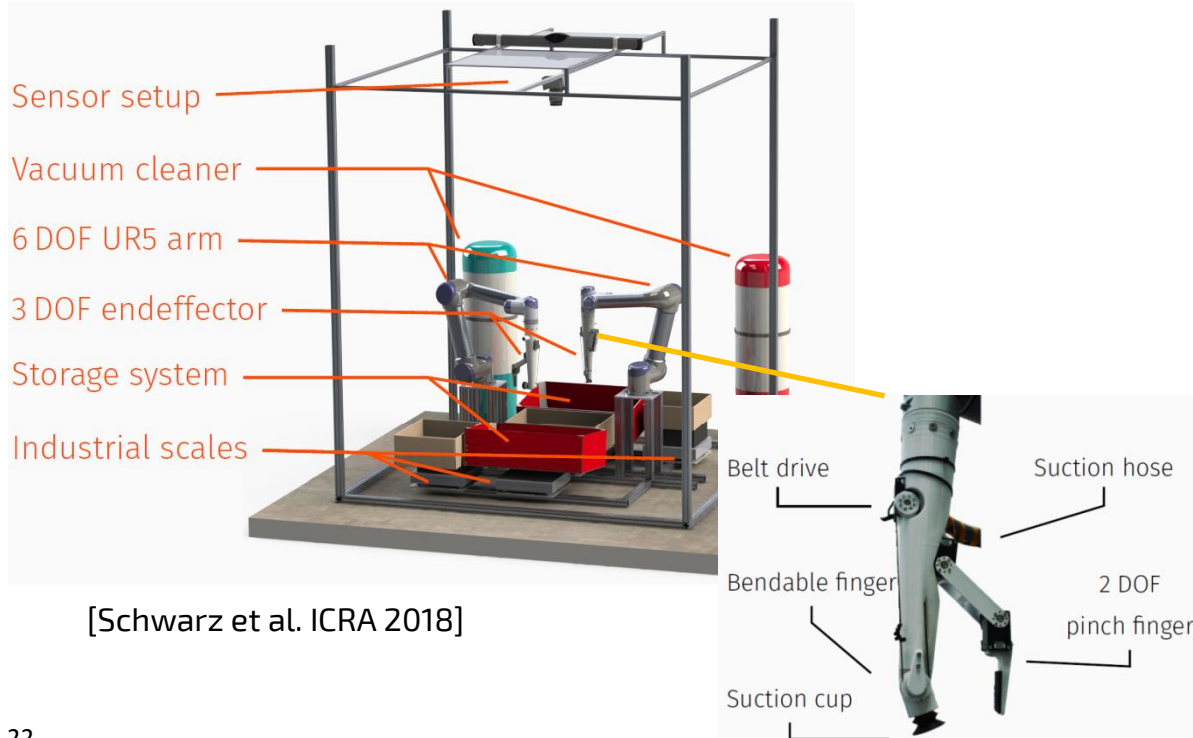
- Erweiterung der Armkinematik
- Wahrnehmung des Kronkorkens
- Anpassung der Bewegungsprimitive



[Stückler, Behnke, Humanoids 2014]

# Amazon Robotics Challenge

- Einlagern und Kommissionierung von Waren
- Zweiarmiges Robotersystem



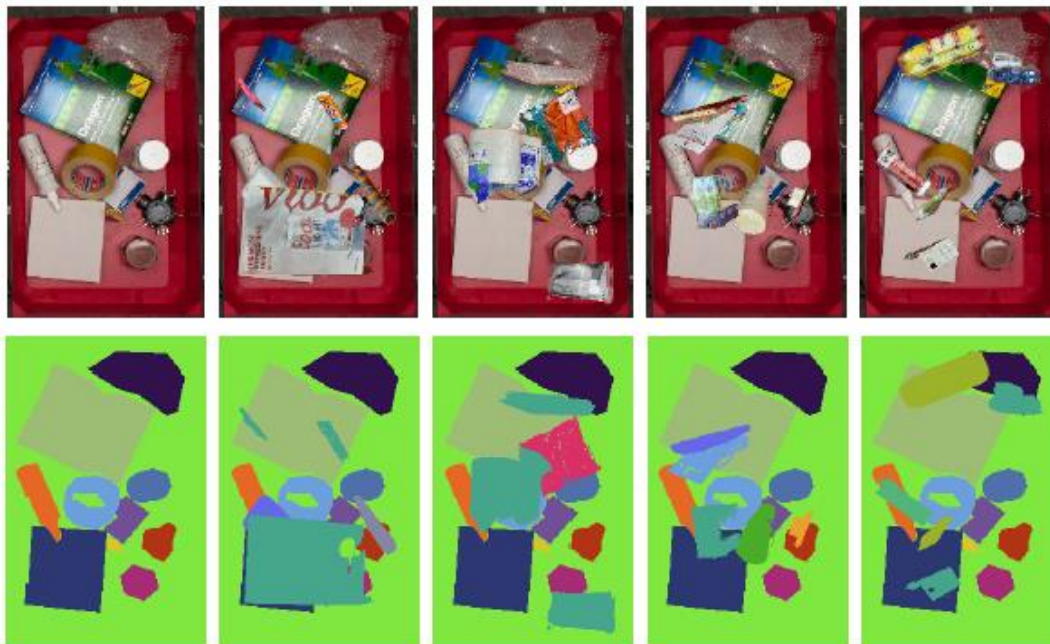
[Amazon]

# Objekterfassung und Szenengenerierung

## ■ Drehteller + DLSR



## ■ Einfügen in komplexe Szenen



# Szeneninterpretation und Greifplanung

- Semantische Segmentierung mit RefineNet [Lin et al. CVPR 2016]
- Saugpopen auf ebenen Flächen im Zentrum der Segmente



bronze\_wire\_cup  
conf: 0.749401

irish\_spring\_soap  
conf: 0.811500

playing\_cards  
conf: 0.813761

w\_aquarium\_gravel  
conf: 0.891001

crayons  
conf: 0.422604

reynolds\_wrap  
conf: 0.836467

paper\_towels  
conf: 0.903645

white\_facecloth  
conf: 0.895212

hand\_weight  
conf: 0.928119

robots\_everywhere  
conf: 0.930464



mouse\_traps  
conf: 0.921731

windex  
conf: 0.861246

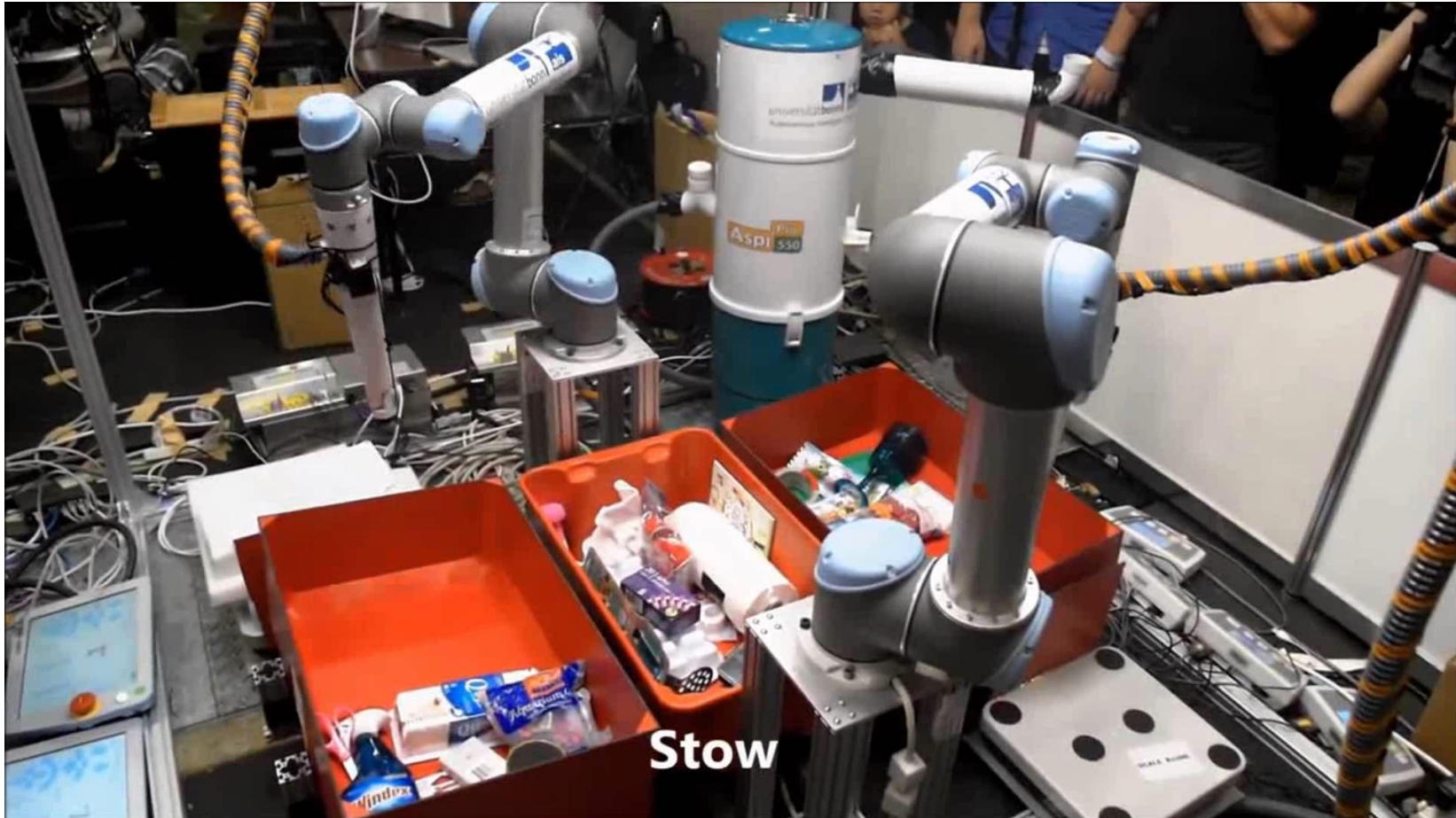
q-tips\_500  
conf: 0.475015

fiskars\_scissors  
conf: 0.831069

ice\_cube\_tray  
conf: 0.976856

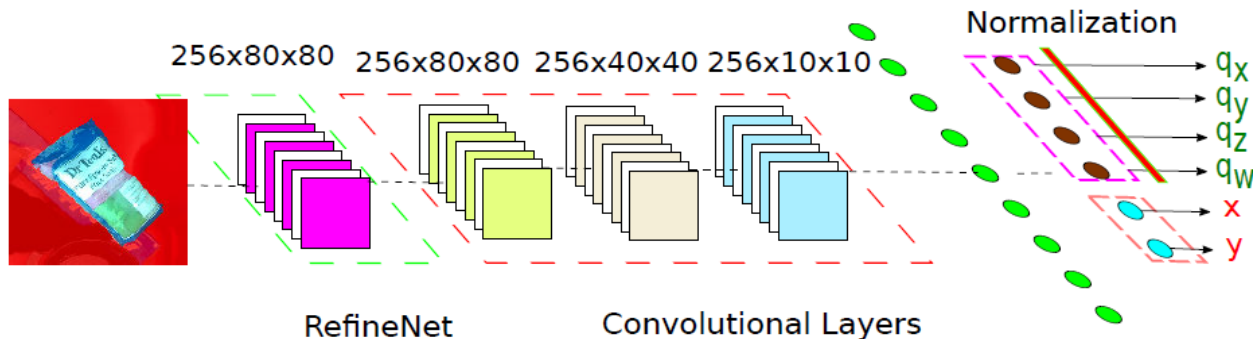


# Amazon Robotics Challenge 2017 Finale

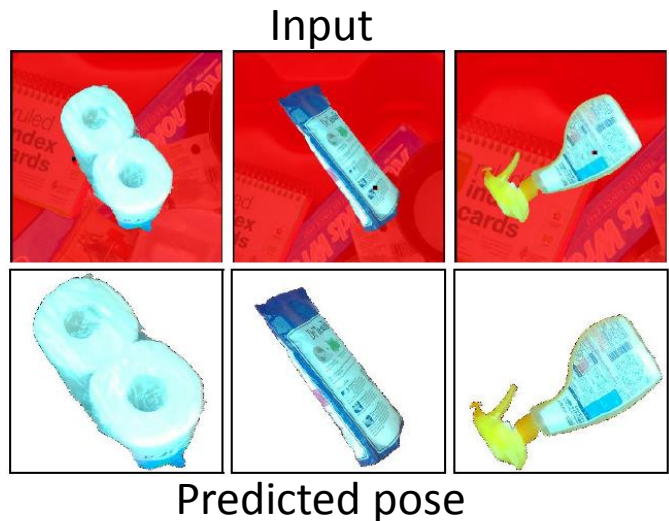


# Schätzung der 6D-Objektpose

- Ausschneiden einzelner Objekt-hypothesen (Segmente)

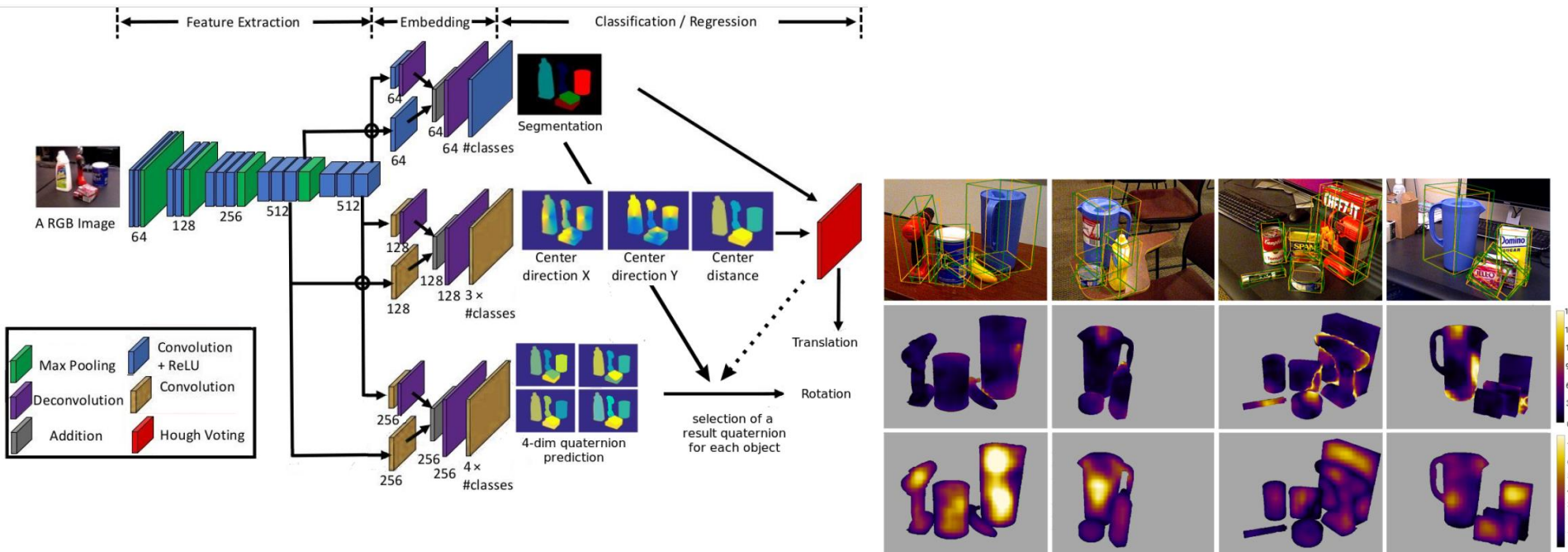


- Merkmale der obersten Schicht von RefineNet
- Vorhersage von Posenparametern
- Registrierung von Objektmodellen



# Dichte Konvolutionale 6D-Posenschätzung

- Erweiterung von PoseCNN [Xiang et al. RSS 2018]
- Dichte Vorhersage von Objektzentren und Orientierungen, ohne Ausschneiden



[Capellen et al, submitted 2019]

# Generieren von Trainingsdaten: Aggregieren der Objektansichten zu texturierten 3D-Objektmodellen



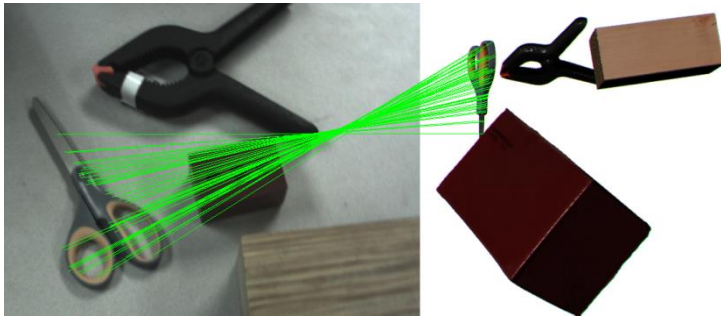
Ansichten



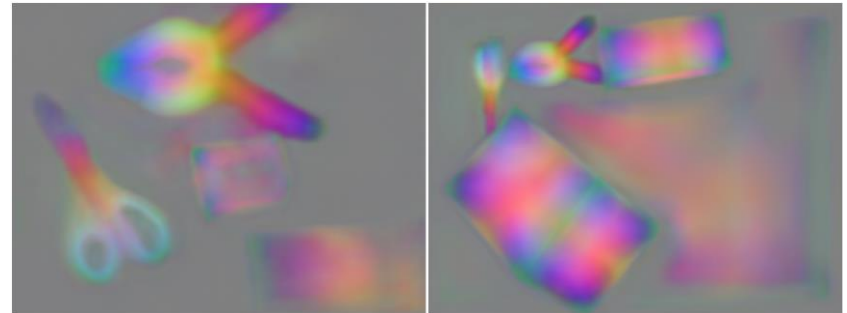
Texturierte  
3D-Modelle

# Selbstüberwachtes Lernen dichter Merkmale

- Oberflächenmerkmale sollen konstant bleiben, auch wenn sich Sichtwinkel, Beleuchtung, etc. ändern
- Merkmale sollen eindeutig sein, um Korrespondenzen zwischen verschiedenen Ansichten herzustellen
- Lernen dichter Merkmale aus Korrespondenzen mit Kontrast-Fehlerfunktion [Schmidt et al. 2016]

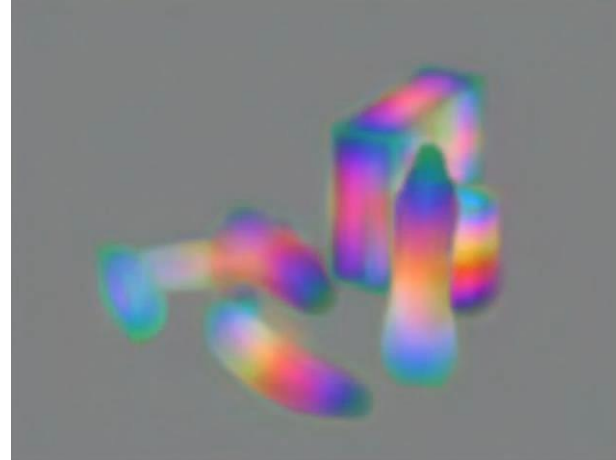


Bekannte Korrespondenzen



Gelernte Merkmale

# Gelernte Szenenabstraktion



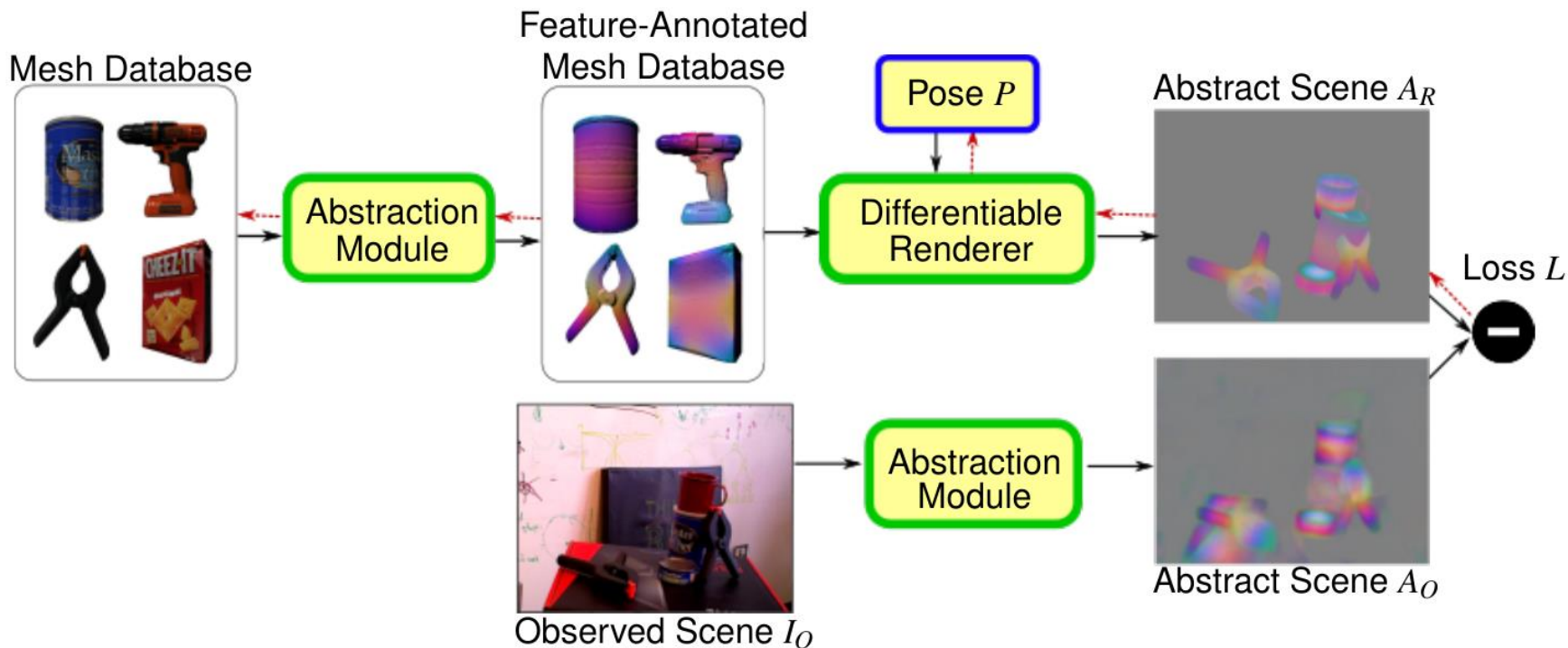
# Merkmalsdeskriptoren als Texturen für 3D-Objektmodelle

- Gelernte Merkmalskanäle als “Farben” für Textur in 3D-Objektmodellen
- Nutzung für 6D-Posenschätzung



# Abstrakte Objektregistrierung

- Vergleiche gerenderte und zu interpretierende Szene in gelerntem Merkmalsraum
- Passe Modellpose durch Gradientenabstieg an





# Beispiel-Szene



# Evaluation auf YCB-Video-Datenmenge

- Konsistente Verbesserung im Vergleich zu PoseCNN (Xiang et al. 2018)
- Monokular (keine Tiefe erforderlich)



PoseCNN



Optimierte Schätzung

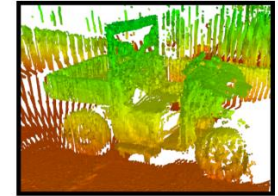
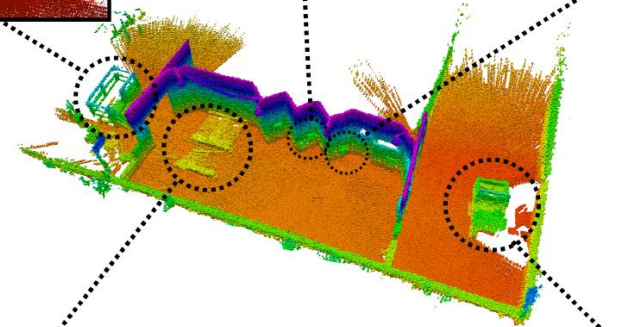
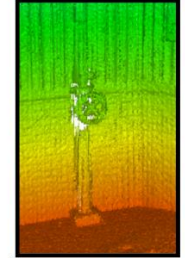
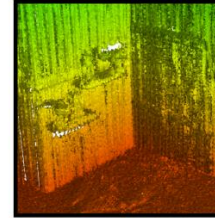
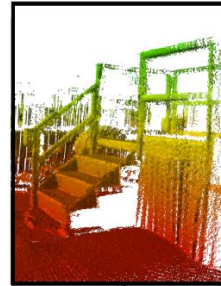
Object	PoseCNN [5]		PoseCNN refined (ours)	
	ADD	ADD-S	ADD( $\Delta$ )	ADD-S( $\Delta$ )
master_chef_can	50.2	83.9	63.3(+13.1)	91.7( +7.8)
cracker_box	53.1	76.9	65.3(+12.2)	81.7( +4.9)
sugar_box	68.4	84.2	85.3(+16.9)	92.0( +7.8)
tomato_soup_can	66.2	81.0	59.4( -6.8)	79.9( -1.1)
mustard_bottle	81.0	90.4	86.5( +5.5)	92.3( +1.9)
tuna_fish_can	70.7	88.0	81.1(+10.4)	94.3( +6.3)
pudding_box	62.7	79.1	71.1( +8.4)	83.1( +4.1)
gelatin_box	75.2	87.2	81.5( +6.3)	89.1( +1.9)
potted_meat_can	59.5	78.5	63.7( +4.2)	80.3( +1.8)
banana	72.3	86.0	82.1( +9.8)	91.8( +5.8)
pitcher_base	53.3	77.0	85.1(+31.8)	92.7(+15.7)
bleach_cleanser	50.3	71.6	65.0(+14.7)	80.4( +8.9)
bowl	3.3	69.6	6.5( +3.1)	75.5( +5.9)
mug	58.5	78.2	65.9( +7.4)	84.0( +5.9)
power_drill	55.3	72.7	73.7(+18.4)	85.9(+13.2)
wood_block	26.6	64.3	45.5(+18.9)	73.3( +9.0)
scissors	35.8	56.9	40.0( +4.1)	58.6( +1.7)
large_marker	58.3	71.7	63.9( +5.6)	77.3( +5.6)
large_clamp	24.6	50.2	37.0(+12.4)	65.1(+15.0)
extra_large_clamp	16.1	44.1	25.4( +9.3)	63.7(+19.6)
foam_brick	40.2	88.0	43.3( +3.1)	90.8( +2.8)
ALL	53.7	75.8	62.8( +9.1)	82.4( +6.6)

# DARPA Robotics Challenge



# Allozentrische 3D-Kartierung

- Registrierung egozentrischer Karten durch Graphoptimierung

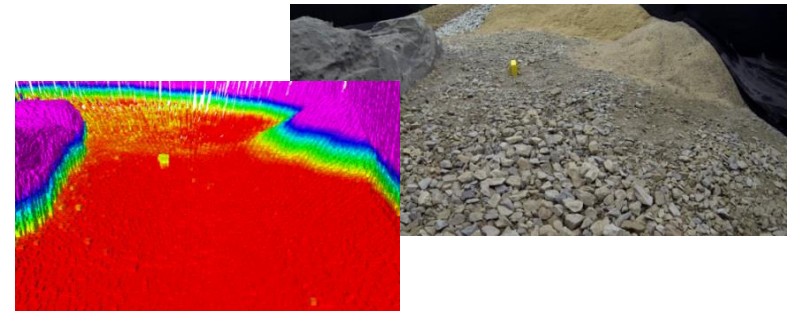
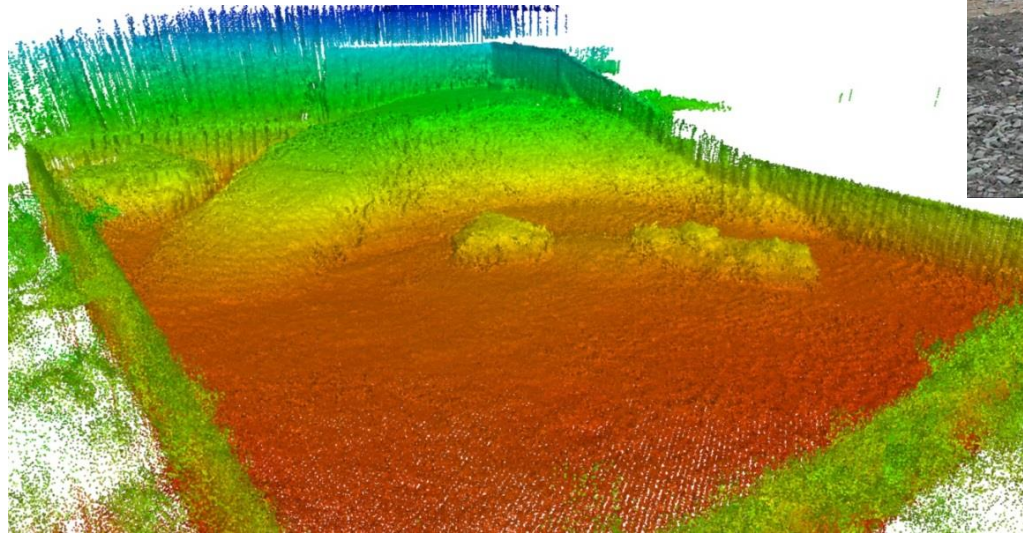


[Droeschel et al., Robotics and Autonomous Systems 2017]

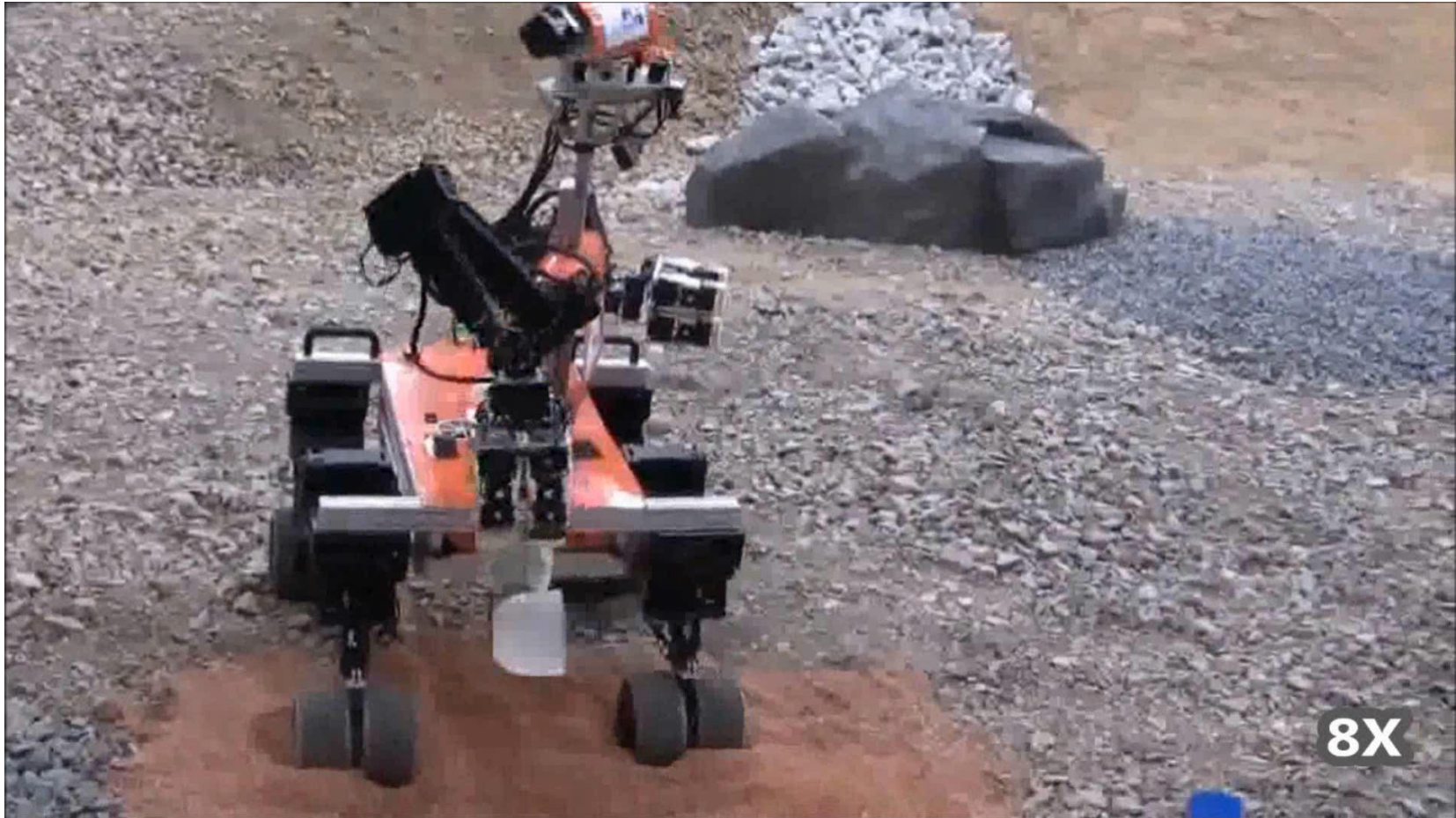
# DLR SpaceBot Cup 2015

## ■ Mobile Manipulation im Gelände

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

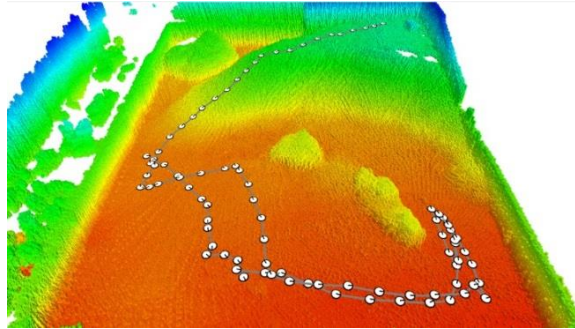


# DLR SpaceBot Cup 2015

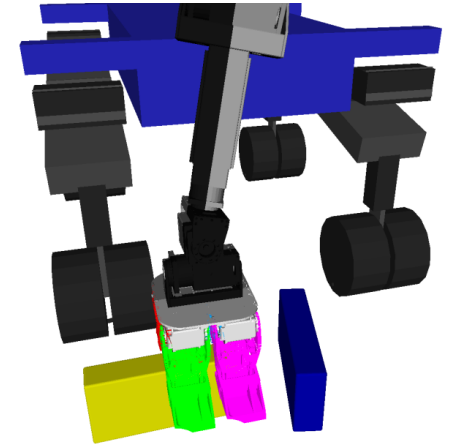
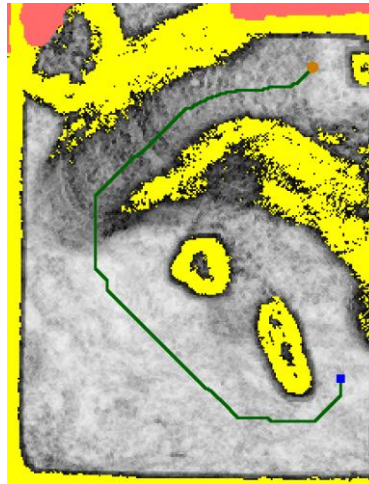
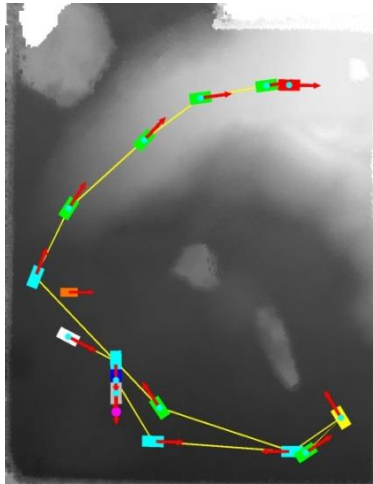
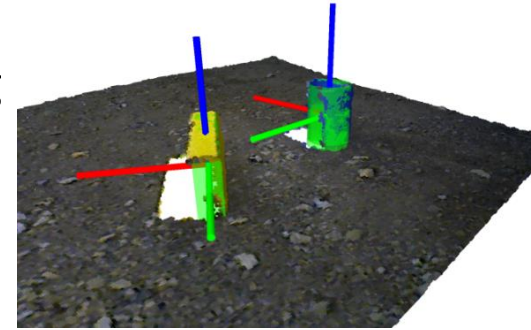


# Autonome Missionsausführung

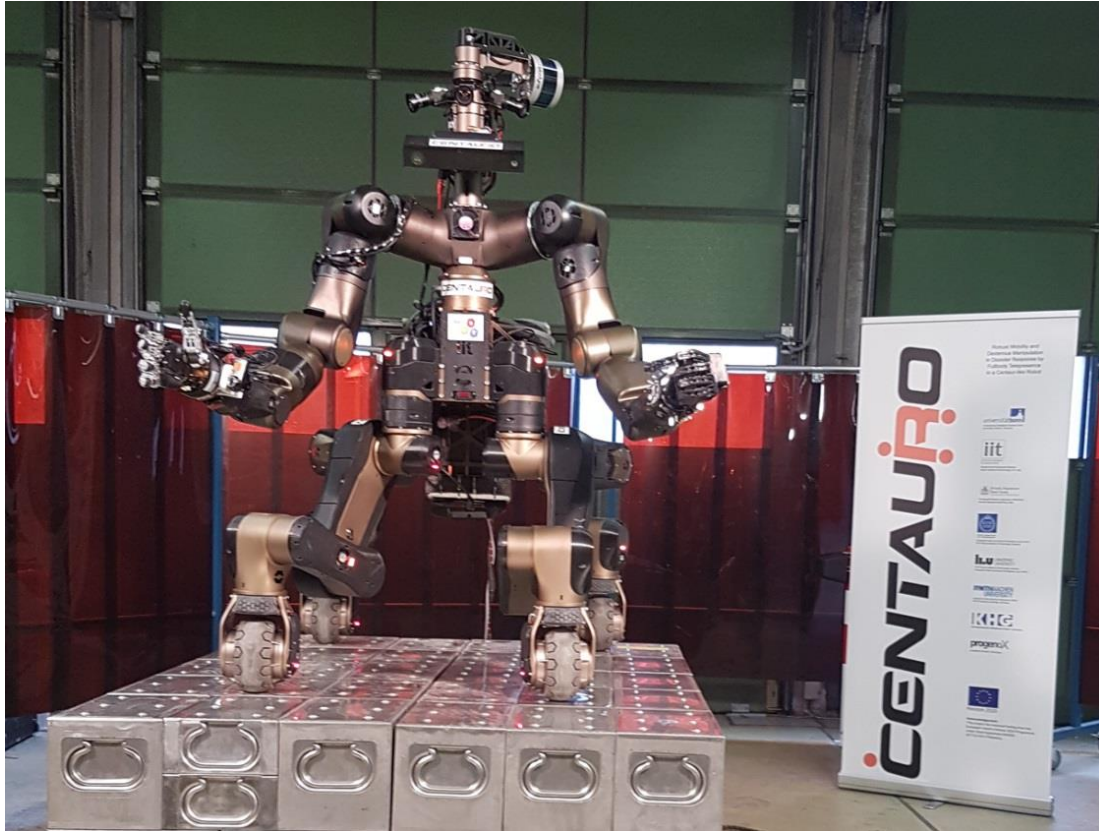
- 3D-Kartierung, Lokalisierung, Missions- und Navigationsplanung



- 3D-Objekt-wahrnehmung und Handhabung



# Centauro-Roboter



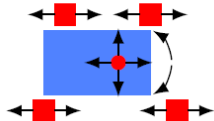
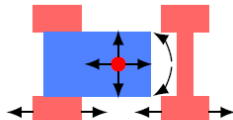
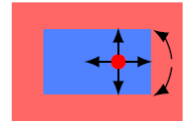
# CENTAURO

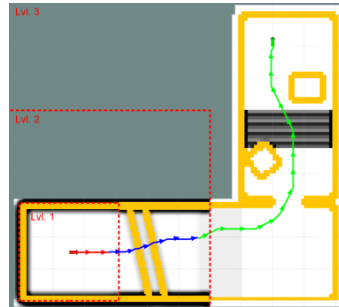
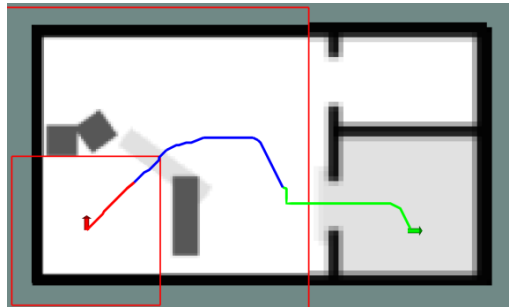
- Seriell-elastische Aktuatoren
- 42 große Gelenke
- Schunk-Hand
- 3D-Laserscanner
- RGB-D-Kamera
- Farbkameras
- Zwei GPU-PCs

[Tsagarakis et al., IIT 2017]



# Planen auf mehreren Abstraktionsstufen

Level	Map Resolution	Map Features	Robot Representation	Action Semantics
1	<ul style="list-style-type: none"> <li>• 2.5 cm</li> <li>• 64 orient.</li> </ul>	<ul style="list-style-type: none"> <li>• Height</li> </ul>		<ul style="list-style-type: none"> <li>• Individual Foot Actions</li> </ul>
2	<ul style="list-style-type: none"> <li>• 5.0 cm</li> <li>• 32 orient.</li> </ul>	<ul style="list-style-type: none"> <li>• Height</li> <li>• Height Difference</li> </ul>		<ul style="list-style-type: none"> <li>• Foot Pair Actions</li> </ul>
3	<ul style="list-style-type: none"> <li>• 10 cm</li> <li>• 16 orient.</li> </ul>	<ul style="list-style-type: none"> <li>• Height</li> <li>• Height Difference</li> <li>• Terrain Class</li> </ul>		<ul style="list-style-type: none"> <li>• Whole Robot Actions</li> </ul>



[Klamt and Behnke,  
IROS 2017, ICRA 2018]

# Evaluation @ KHG: Fortbewegungsaufgaben



# Transfer von Handhabungsfertigkeiten



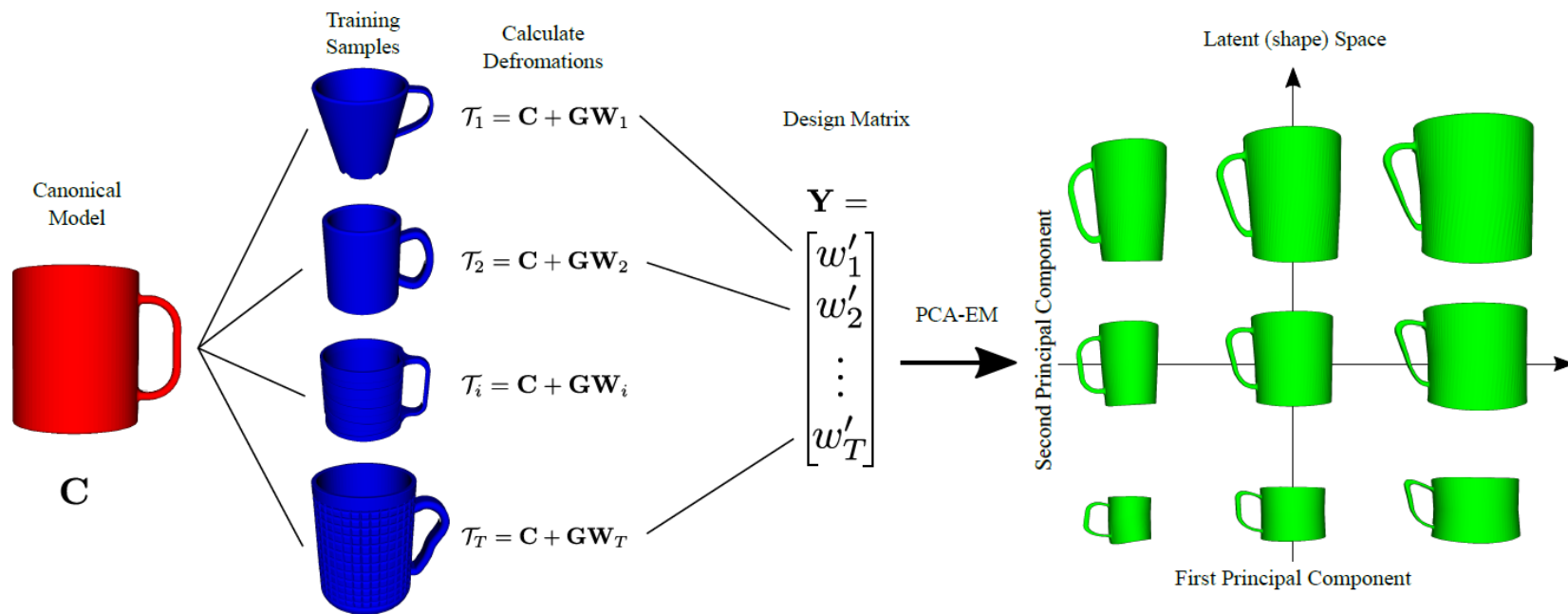
Wissens-  
transfer



[Rodriguez et al. ICRA 2018]

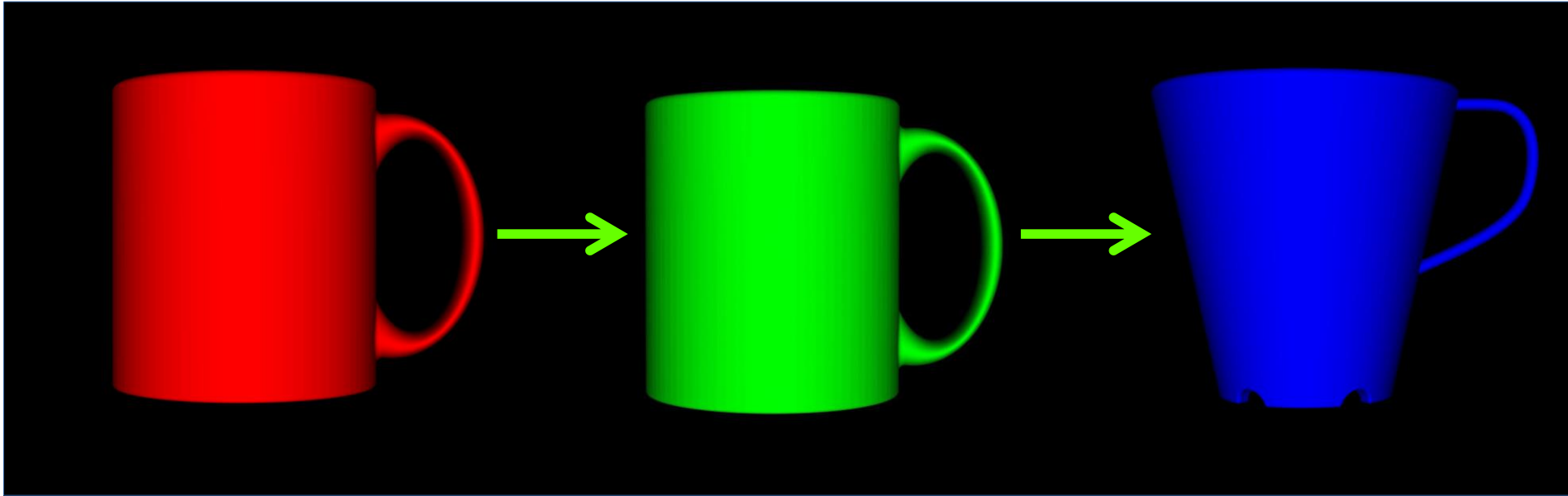
# Lernen von eines Formraums

- Deformierbare Registrierung von Instanz und kanonischem Modell
- Hauptkomponentenanalyse der Verformungen



[Rodriguez et al. ICRA 2018]

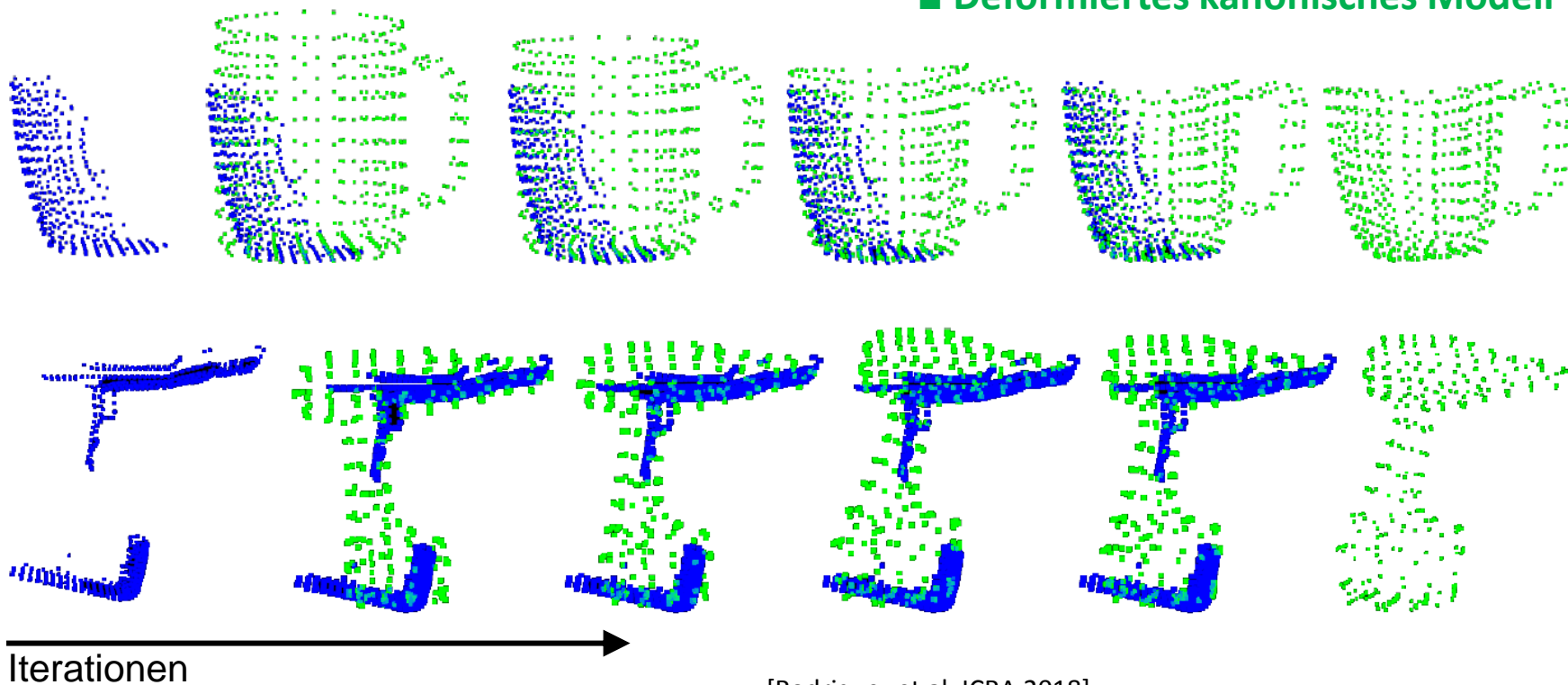
# Interpolation in Formraum



[Rodriguez et al. ICRA 2018]

# Deformierbare Registrierung der gelernten Form

- Teilansicht unbekannter Instanz
- Deformiertes kanonisches Modell



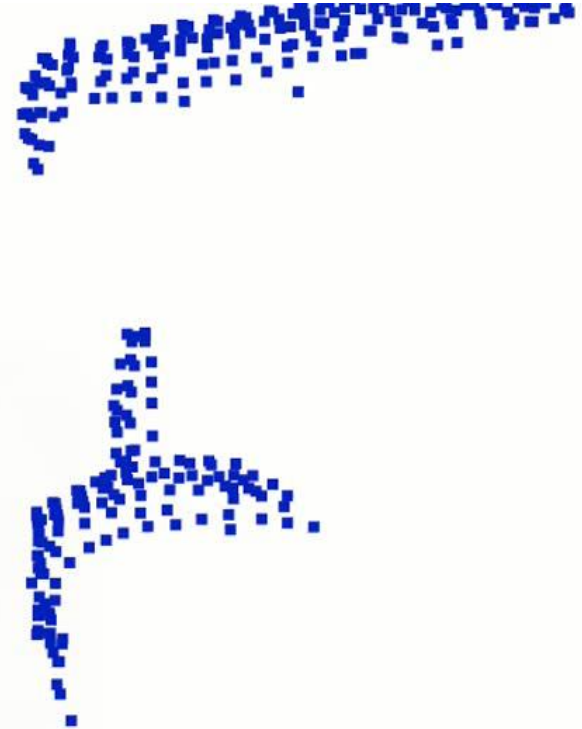
[Rodriguez et al. ICRA 2018]

# Transfer von Handhabungswissen

■ Vollständige Punktwolke



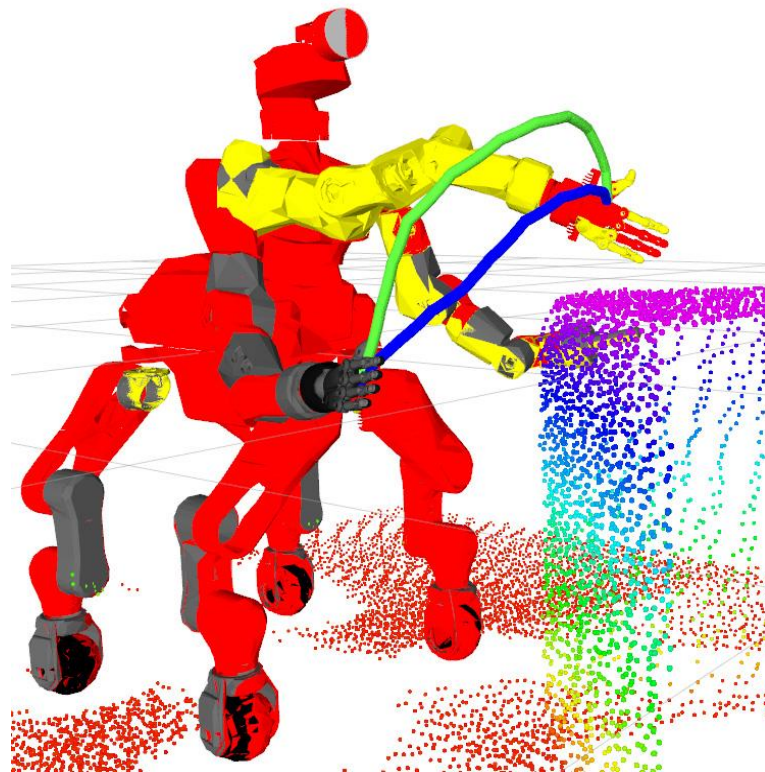
■ Teilansicht



# Armbewegungsoptimierung

Trajektorienoptimierung mit Nebenbedingungen:

- Hindernisvermeidung
- Gelenklimits
- Zeitminimierung
- Drehmoment-Minimierung



[Pavlichenko et al., IROS 2017]



# Greifen und Benutzen eines unbekannten Akkuschraubers

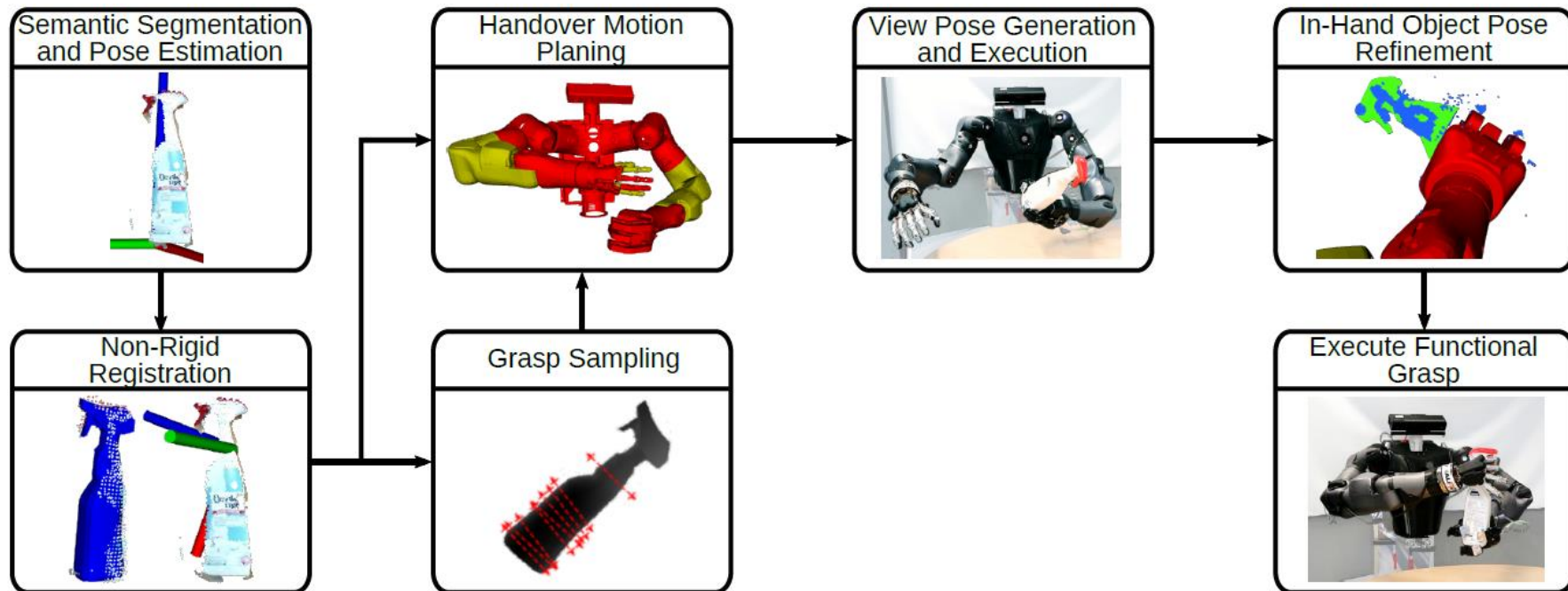


# Evaluation @ KHG: Handhabungsaufgaben



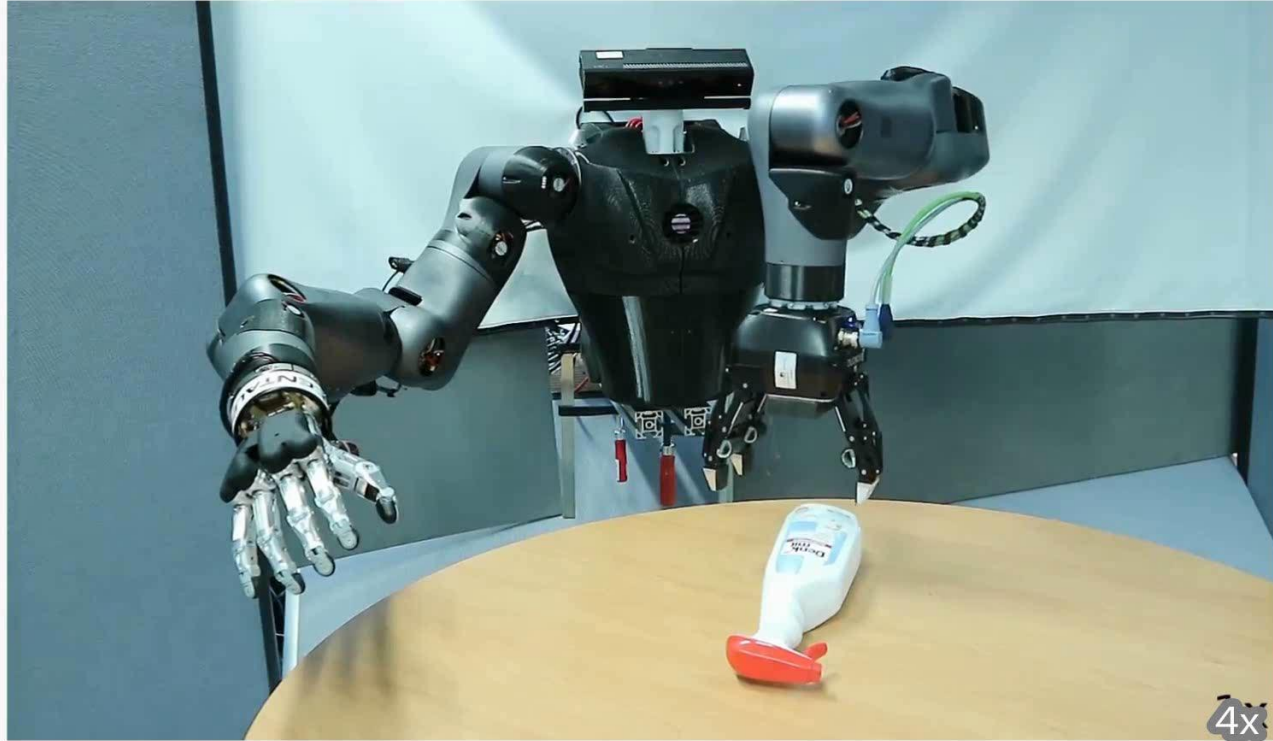
# Vorbereitung eines Funktionalen Griffs

- Direkte funktionale Griffe sind oft nicht möglich
- Aufheben des Werkzeugs so, dass andere Hand funktional greifen kann



# Vorbereitung eines Funktionalen Griffs

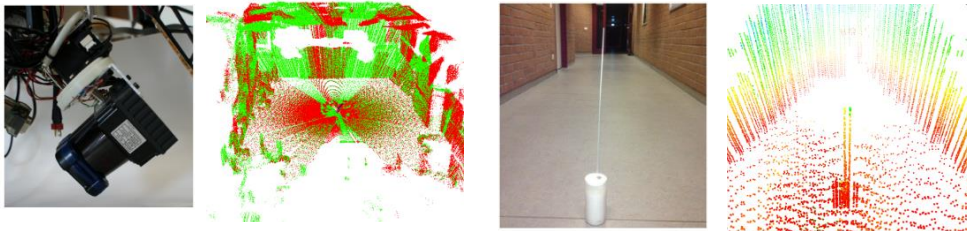
## Robot Experiments



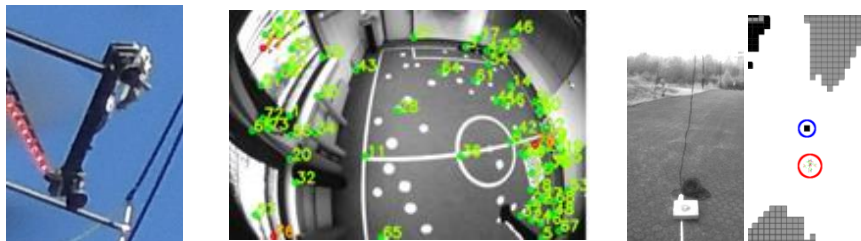
# Autonomer Flug in Hindernisnähe

## ■ Multimodale Hinderniserkennung

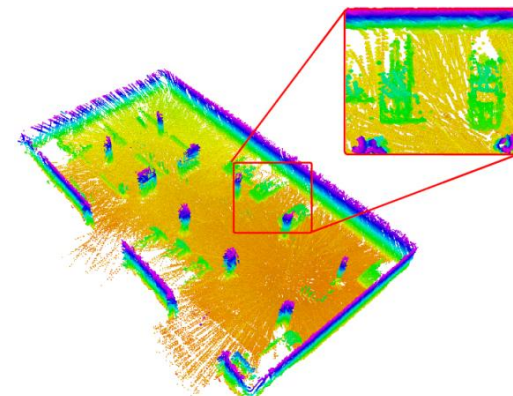
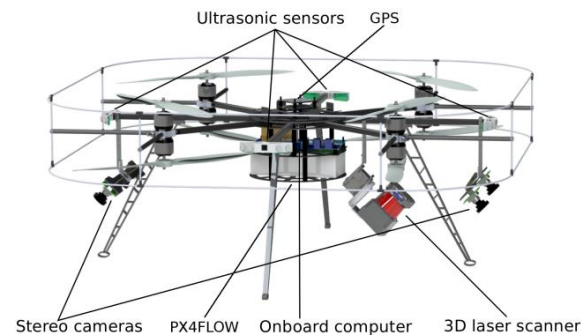
### ● 3D-Laserscanner



### ● Stereokameras



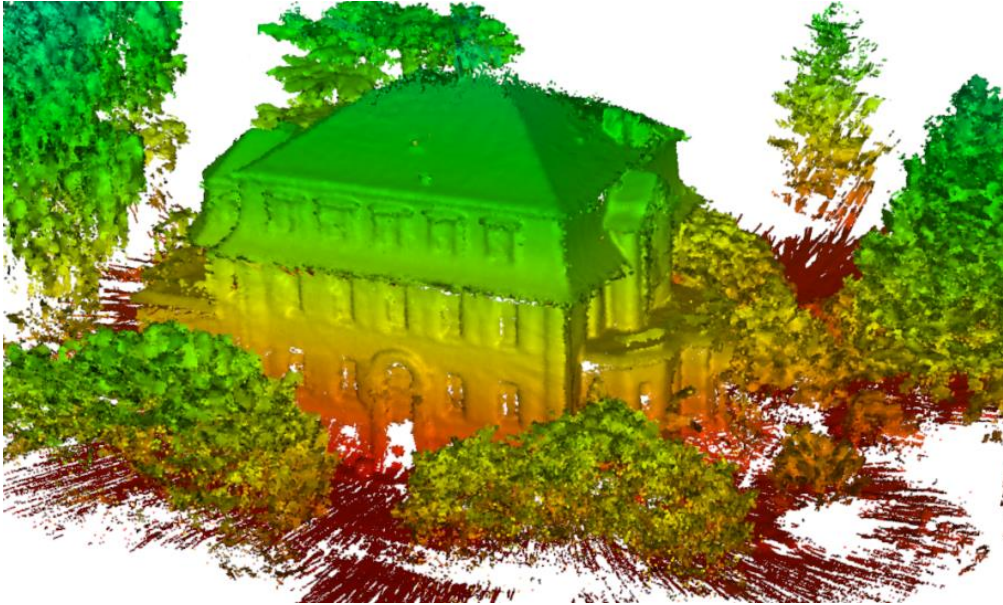
### ● Ultraschall



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

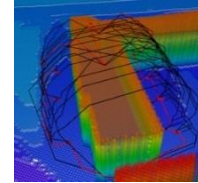
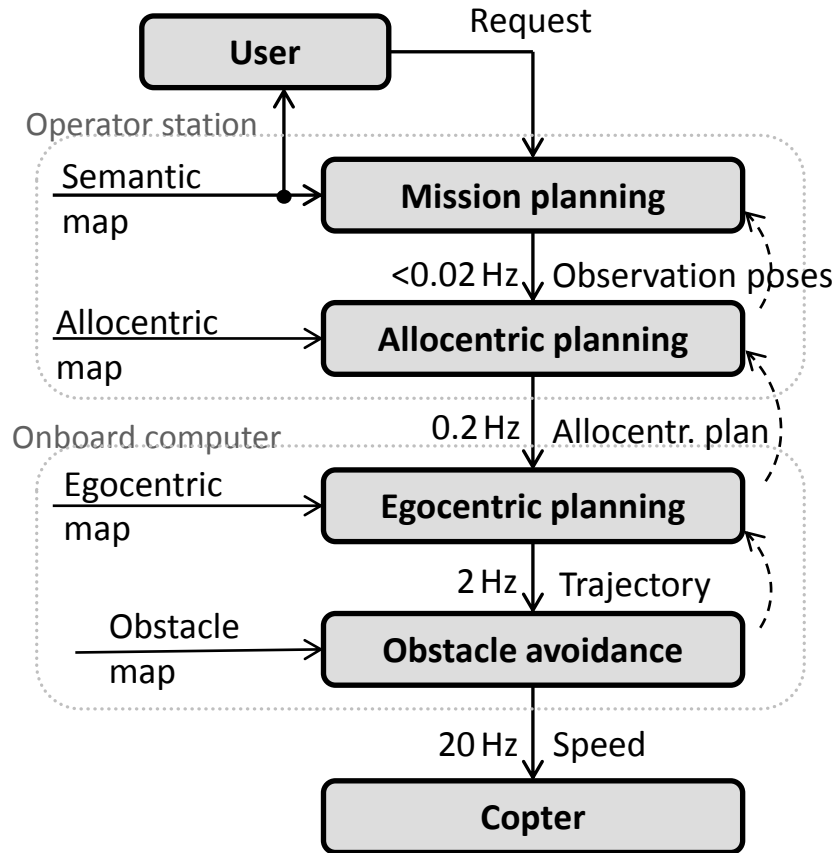
# 3D-Kartierung

- Registrierung von Laser-Distanzmessungen

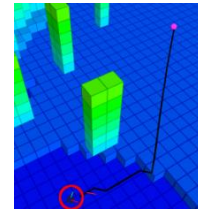


[Droeschel et al. JFR 2016]

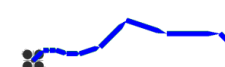
# Hierarchische Navigation



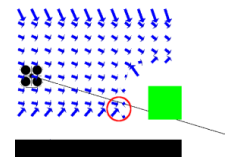
Missionsplanung



Allozentrische Planung



Egozentrische Planung



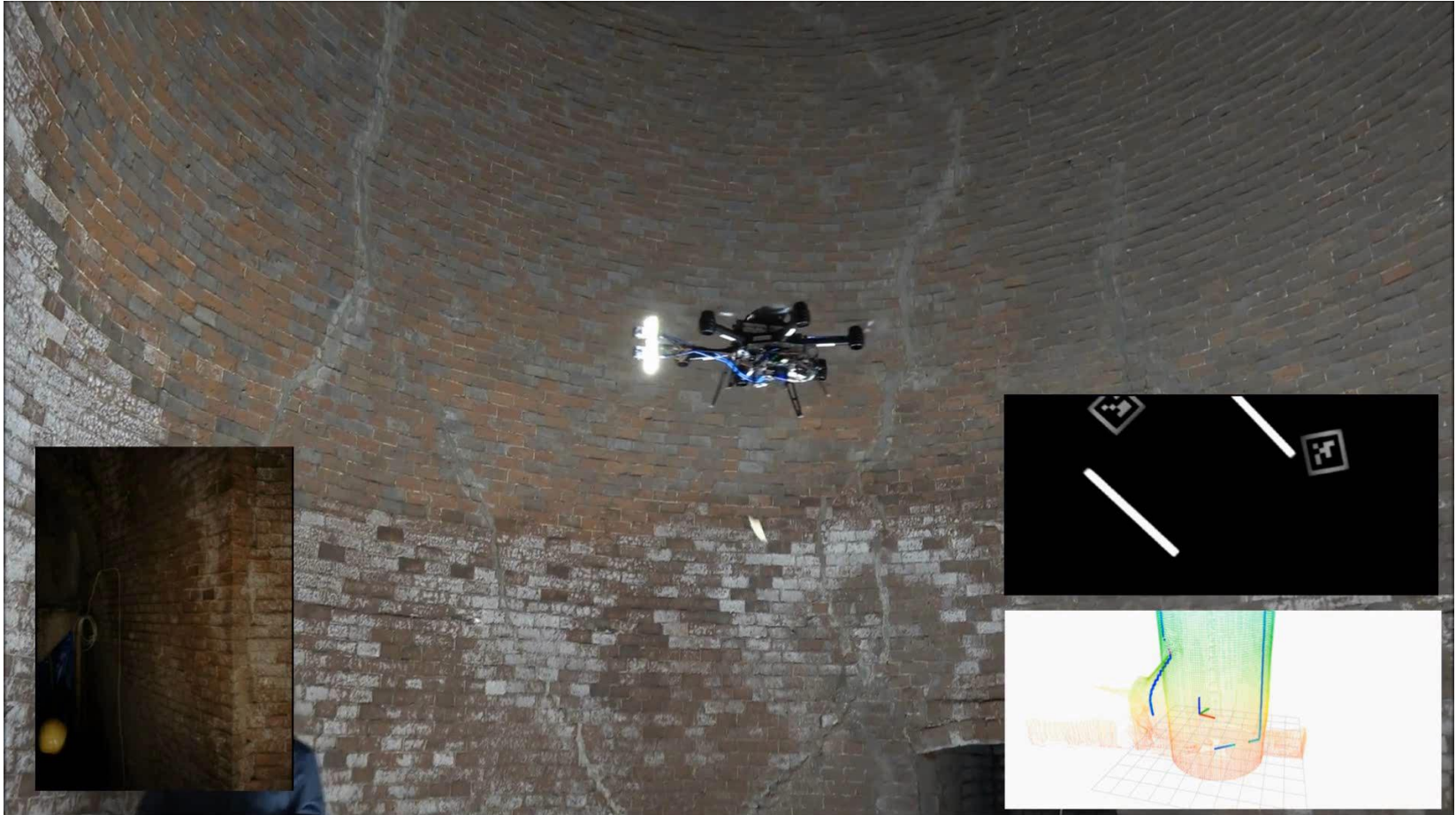
Hindernisvermeidung

## Mapping on Demand

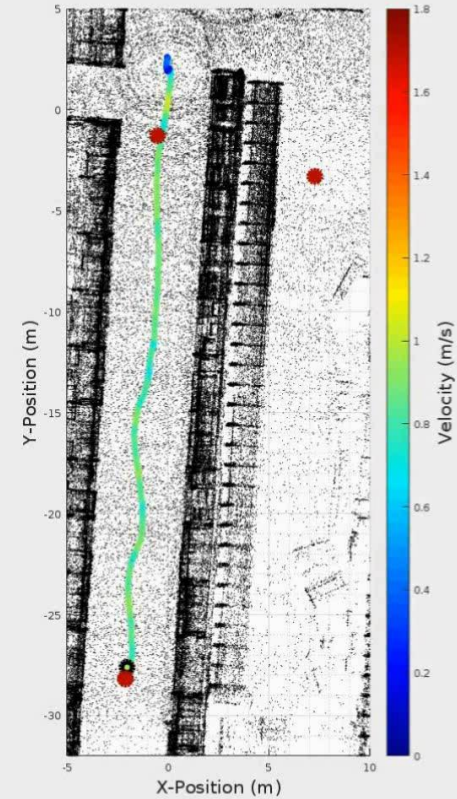
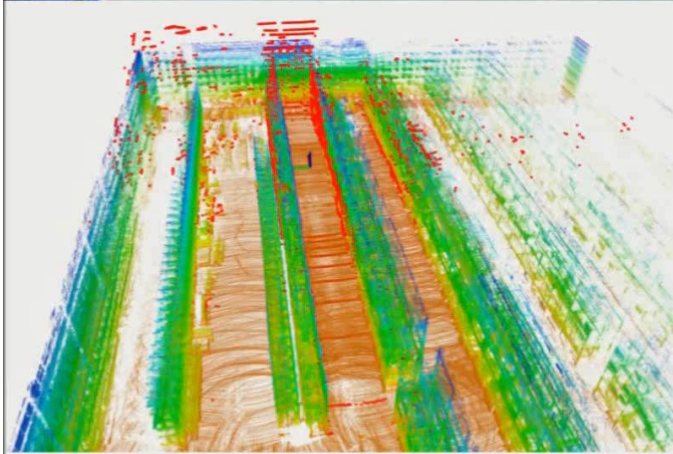
Autonomous Flight to Planned View Poses



# Schornsteininspektion



# InventAIRy: Autonome Navigation im Warenlager

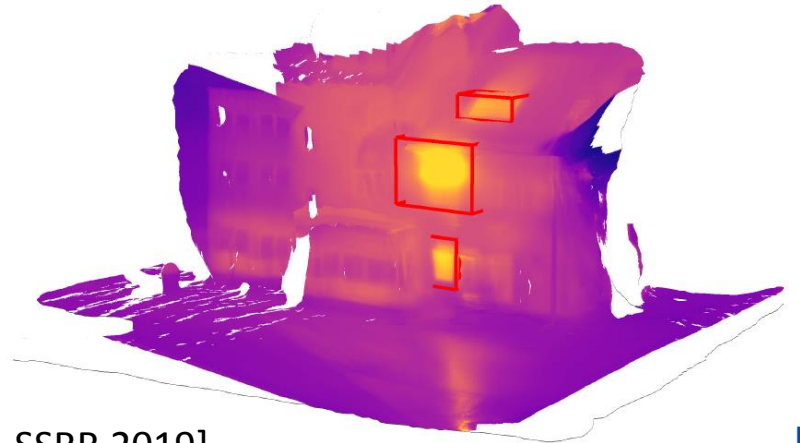
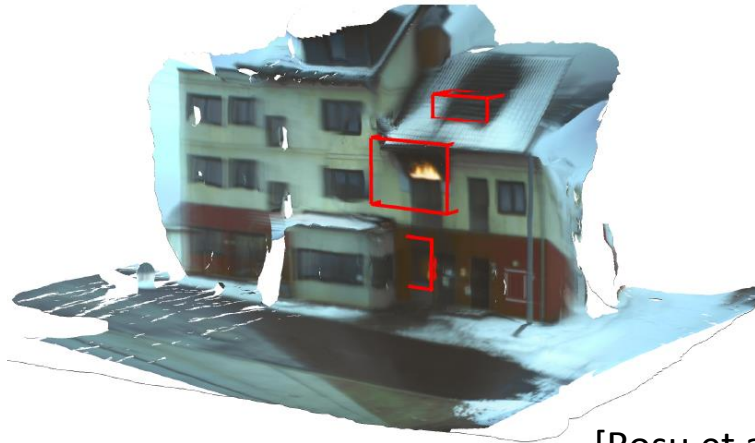
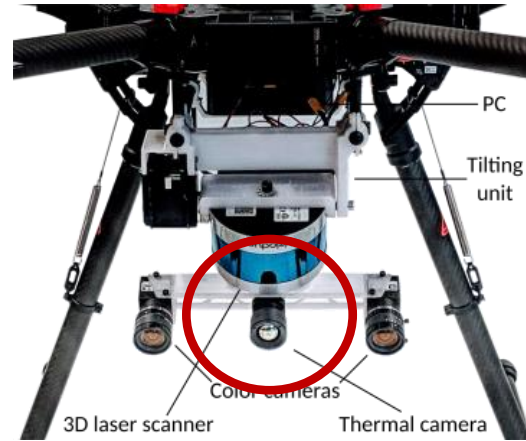


# InventAIRy: Detektierte Etiketten im Regal



# Unterstützung von Einsatzkräften (A-DRZ)

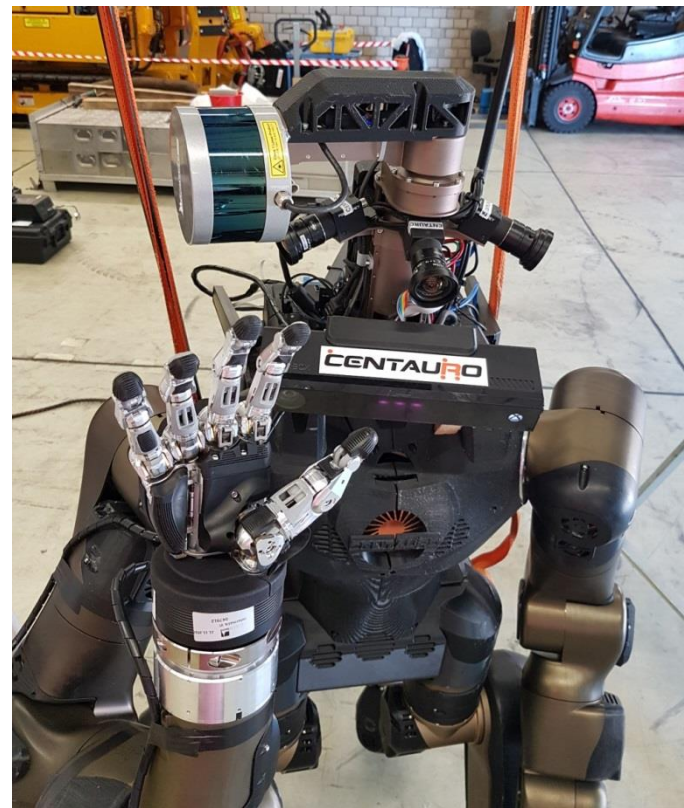
- Thermalkamera zugefügt
- Flug am Brandhaus in Dortmund
- Detektion von Brandherden und Glutnestern



[Rosu et al, SSRR 2019]

# Zusammenfassung

- Beispiele für kognitive Roboter in komplexen Szenarien
  - Humanoide Fußballroboter
  - Serviceaufgaben im Haushalt
  - Griff in die Kiste
  - Menschenfeindliche Umgebungen
  - Flugroboter
- Herausforderungen beinhalten
  - Semantische 4D-Kartierung, Vorhersagen
  - Robuste hochdimensionale Bewegungsplanung
- Mögliche Lösungsansätze
  - Zusammenführen von Erfahrungen vieler Roboter
  - Kombination von menschlicher Intelligenz und Autonomie
  - Instrumentierte Umgebungen

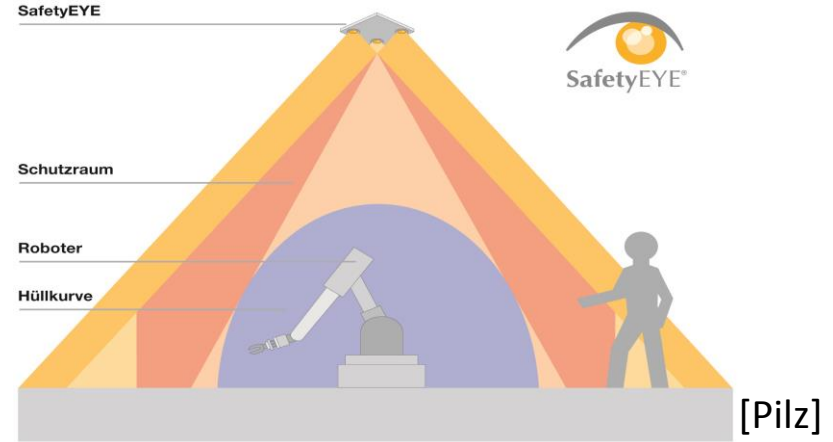
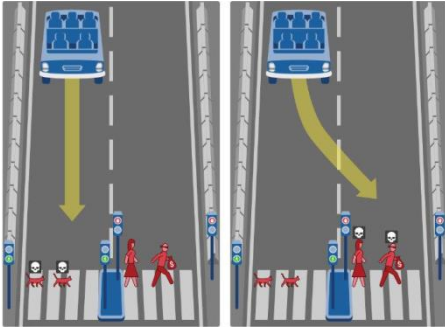


# Ausblick

- Kognitive Roboter haben hohes Anwendungspotential
- Werden Auswirkungen auf alle Lebensbereiche haben
- Gesellschaftliche Herausforderungen z.B.
  - Sicherheit,
  - Haftungsfragen,
  - Auswirkungen auf die Arbeitswelt,
  - Privatsphäre und Datenschutz

# Sicherheit

- Trennung von Mensch und Robotern aufgehoben
- Sicherheit schwer oder gar nicht zu garantieren
- Sicherheitsnormen müssen angepasst werden
- Entscheidungsdilemmata



[NTSB]

# Haftungsfragen

- Wer haftet?
  - Hersteller des selbst fahrenden Autos?
  - Anbieter der Steuersoftware?
  - Fahrzeughalter?
  - Fahrer?
- Das Wiener Übereinkommen über den Straßenverkehr von 1968 schreibt vor, dass jedes in Bewegung befindliche Fahrzeug einen Fahrer haben und dieser das Fahrzeug auch beherrschen muss.
  - Ist der Autopilot der Fahrer?



# Privatsphäre und Datenschutz

- Roboter dringen in die privatesten Bereiche der Nutzer ein
- Müssen sehr viel wissen, um ihre Dienste zu erbringen
- Maßnahmen zum Datenschutz und zur informationellen Selbstbestimmung erforderlich
- Prinzip der sparsamen Datengewinnung
- Verlust an Privatheit muss gegen den Nutzen aufgewogen werden

# Auswirkungen auf die Arbeitswelt

- Automatisierbare Tätigkeiten werden verdrängt



[CGP Grey  
<https://youtu.be/7Pq-S557XQU>]

- Müssen durch andere Tätigkeiten bzw. Umverteilung von Arbeit und Einkommen ersetzt werden

# Neue Anwendungsgebiete für Roboter

- Autonomes Fahren
- Logistik
- Landwirtschaft
- Kollaborative Produktion
- Alltagsassistentz
- Weltraum, Suche&Rettung
- Medizin, Pflege
- Unterhaltung, Spielzeuge

