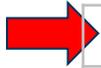


Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Abschlussworkshop mit dem Beirat, der TAB und dem TMWWDG
am 24.11.2021

Agenda



9:00	Prof. Groß	Begrüßung der Teilnehmer und Eröffnung
9:15	Dr. Müller	Vorstellung des realisierten Szenarios und der Systemarchitektur
9:30	jeweilige Partner	Vorstellung der Forschungsergebnisse zu den methodischen Teilaspekten des Szenarios <ul style="list-style-type: none"> - Personenwahrnehmung (TUI NIKR) - Real Time Localization System (HSM) - Multimodale 3D Sensorik für Hand/Objektsegmentierung (TUI QBV) - Objektposenschätzung (GFE) - Greifposen und Bewegungssteuerung (TUI NIKR) - Nutzerzentrierte Navigation (TUI NIKR)
10:45		Diskussion zu den methodischen Beiträgen
10:55		kurze Pause
11:00	Dr. Müller Yan Zhang	Vorstellung der Demonstratoren, Demonstration der Roboter im Labor
11:30	Prof. Notni	Freie Diskussion der Ergebnisse mit dem Beirat
11:45	Prof. Groß	Zusammenfassung der wissenschaftlichen Ergebnisse
11:55	Dr. Djambova	Feedback des Projektträgers

Vorstellung - SONARO Forschergruppe

TU Ilmenau,
FG Neuroinformatik und
Kognitive Robotik
TUI-NIKR



Prof. Groß



Dr. Müller



TU Ilmenau,
FG Qualitätssicherung und
Industrielle Bildverarbeitung
TUI-QBV



Prof. Notni



M. Sc. Zhang



Hochschule
Schmalkalden,
FG Eingebettete
Diagnosesysteme **HSM**



Prof. Wenzel



M. Sc. Schneider

Gesellschaft für
Fertigungstechnik und
Entwicklung Schmalkalden
e.V. **GFE**



Dr. Welzel



Dr. Garten

Vorstellung des Projektträgers



Thüringer Aufbaubank

Thüringer Aufbaubank (TAB)

- ✓ Dr. Djambova (fachlich inhaltliche Projektbetreuung)
- ✓ Frau Trommer (finanztechnische Projektbetreuung)
 - Dr. Halama (Abteilungsleiter Transferförderung)

TMWWDG

- Frau Haelbich

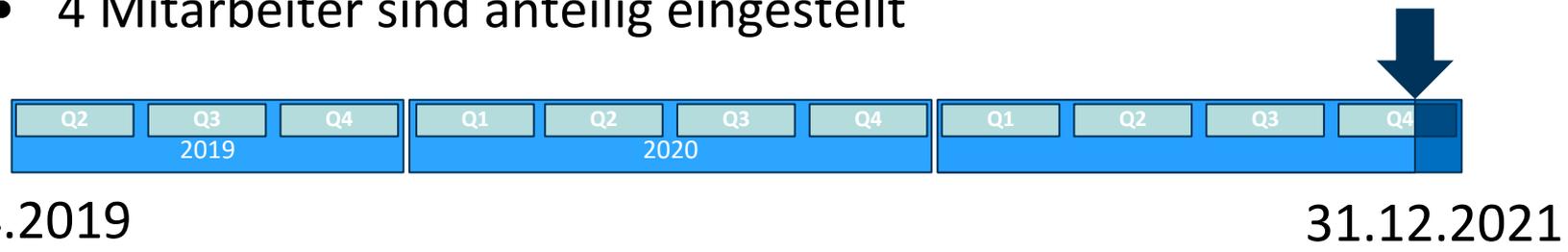
SONARO Unternehmensbeirat + Gäste

- ✓ • Dr. Trabert, Metralabs GmbH, Ilmenau
- ✓ • Herr Waldmann, Betriebsleiter HENKEL+ROTH GmbH, Ilmenau
- ✓ • Herr Hofmann, Cluster für Fertigungstechnik & Metallbearbeitung
- ✓ • Herr Strüning, SCS Robotik UG, Schmalkalden
- ✓ • Herr Richter, Honda Research Institute Europe
 - Dr. Behling, Götting KG, Lehrte
 - Vision & Control GmbH, Suhl
- ✓ • Herr Schäfer, Hörisch Präzision GmbH
- ✓ • Herr Welzel, Herr Lahmann
- ✓ • Herr Patschger (ThZM), Dr. Debes
- ✓ • Teile des E4SM-Teams (Carl-Zeiss-Förderprojekt)

Eckdaten der Forschergruppe



- ThZM als Antragsteller
- Koordinator: TU-Ilmenau FG NI&KR (Prof. Groß)
- Fördersumme: 699.020 EUR
- 4 Mitarbeiter sind anteilig eingestellt



Zielstellung des Projekts SONARO

Was meinen wir mit **smarte Objektübernahme und –übergabe**?

- Bei der **Übernahme** Gegenstände beliebiger Form aus der Hand eines Menschen greifen
- Dabei sicher etwaige Kollision mit der Hand vermeiden
- Greifpunkte am Objekt ergonomisch sinnvoll (taskabhängig!) wählen
- Bei der **Übergabe** die Gegenstände in die Hand des Nutzers legen
- Dabei gängige Normen für die Übergabe von teils gefährlichen Gegenständen einhalten (z.B.: Messer mit dem Griff nach vorn reichen)
- Dazu ggf. den Gegenstand mit dem Roboter umgreifen (wenn zweiarmig)

Wesentliche Entwicklungsschwerpunkte:

- Sensorik und Methodik für robuste Segmentierung von Hand und Objekt
- Sensorik für robuste Indoor-Lokalisation (anwendbar für Personen, Objekte, Roboter)
- Methodik für Steuerung des Greifprozesses in dynamischen Umgebungen

Angestrebte Innovationen in SONARO

Aspekt	Ausgangssituation im Jahre 2019	Zielstellung von SONARO
Objektübernahme	Mensch muss sich statisch verhalten und kooperativ sein und sich an Übernahmepose des Roboters anpassen (→ in die Hand legen)	Situationsabhängige freie Übernahme der Objekte durch den Roboter auch bei Bewegungen des Menschen
Objektübergabe	Mensch muss kooperativ sein und sich an die Übergabepose des Roboters anpassen (→ aus der Hand nehmen) Roboter übergibt Gegenstand so wie übernommen	Freie Übergabe an den Menschen – auch bei Bewegungen (dynamische Anpassung) Berücksichtigung der Objektaffordanzen , d.h. Unterscheidung zwischen Griffposen für Übergabe und Übernahme ; Realisierung eines Umgreifens zwecks korrekter Übergabe
Sicherheit	Gemeinsamer Interaktionsbereich für Mensch und Roboter ist sicher durch bauliche Einschränkungen der Roboter (max. Geschwindigkeit/Kraft)	Robuste Wahrnehmung der menschlichen Körperteile im Interaktionsbereich durch spezielle Sensorik und Berücksichtigung bei der Bewegungsausführung
Objektwissen	Modellwissen zu den Objekten vorab erforderlich (z.B. als 3D Mesh)	Übergabe/Übernahme auch für unbekannte Objekte möglich
Personenerfassung / Aktionsraum	Begrenzter Aktionsradius – Interaktionspartner muss im Sichtbereich sein	Erweiterter Aktionsradius - Interaktionspartner kann sich außerhalb des Sichtbereiches aufhalten

Weiterer Ablauf

- Vorstellung der Systemarchitektur als Grundlage für die Demonstratoren
- Gliederung der methodischen Beiträge der Projektpartner anhand der Systemarchitektur

- **Notation:**

Studentische Abschlussarbeiten

BA Amanda Robillard: *Untersuchung von 3D-Skelett-basierten Verfahren zur Aktionserkennung.* TU Ilmenau

Internationale Konferenzbeiträge

Müller, St., Stephan, B., Gross, H.-M. *MDP-based Motion Planning* in: Europ. Conf. on Mobile Robotics (ECMR)

Journalpublikationen

Zhang, Y., Müller, St., Stephan, B., Gross, H.-M., Notni, G. *Point Cloud Segmentation* in: Sensors 2021, 21(16), 5676

Zitierte Fremdliteratur

Li, Yi, et al. *DeepIM: Deep iterative matching for 6d pose estimation.* Proceedings of (ECCV). 2018.

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Beiratstreffen am 24.11.2021

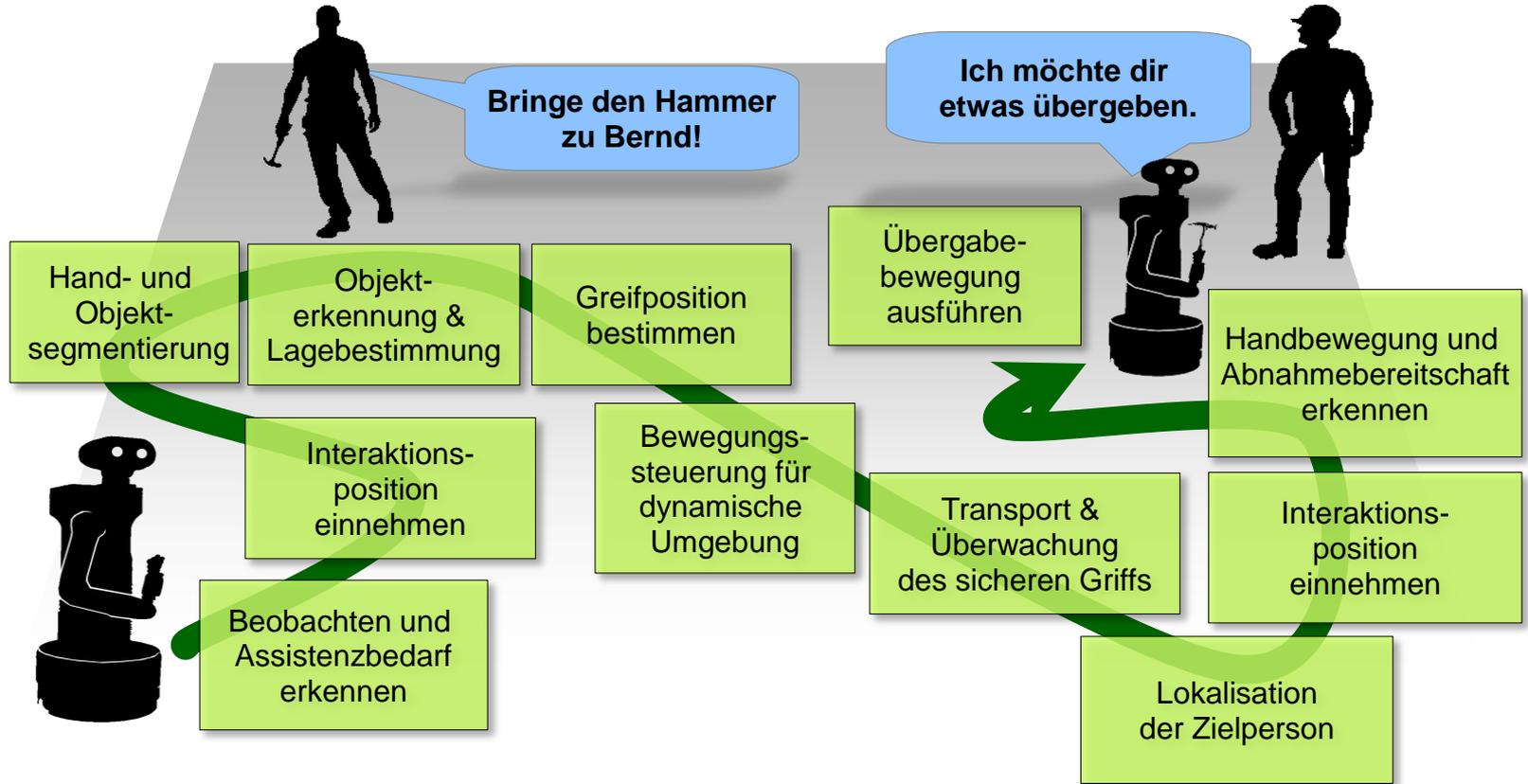
Vorstellung des realisierten Szenarios und der Systemarchitektur

Dr. Steffen Müller

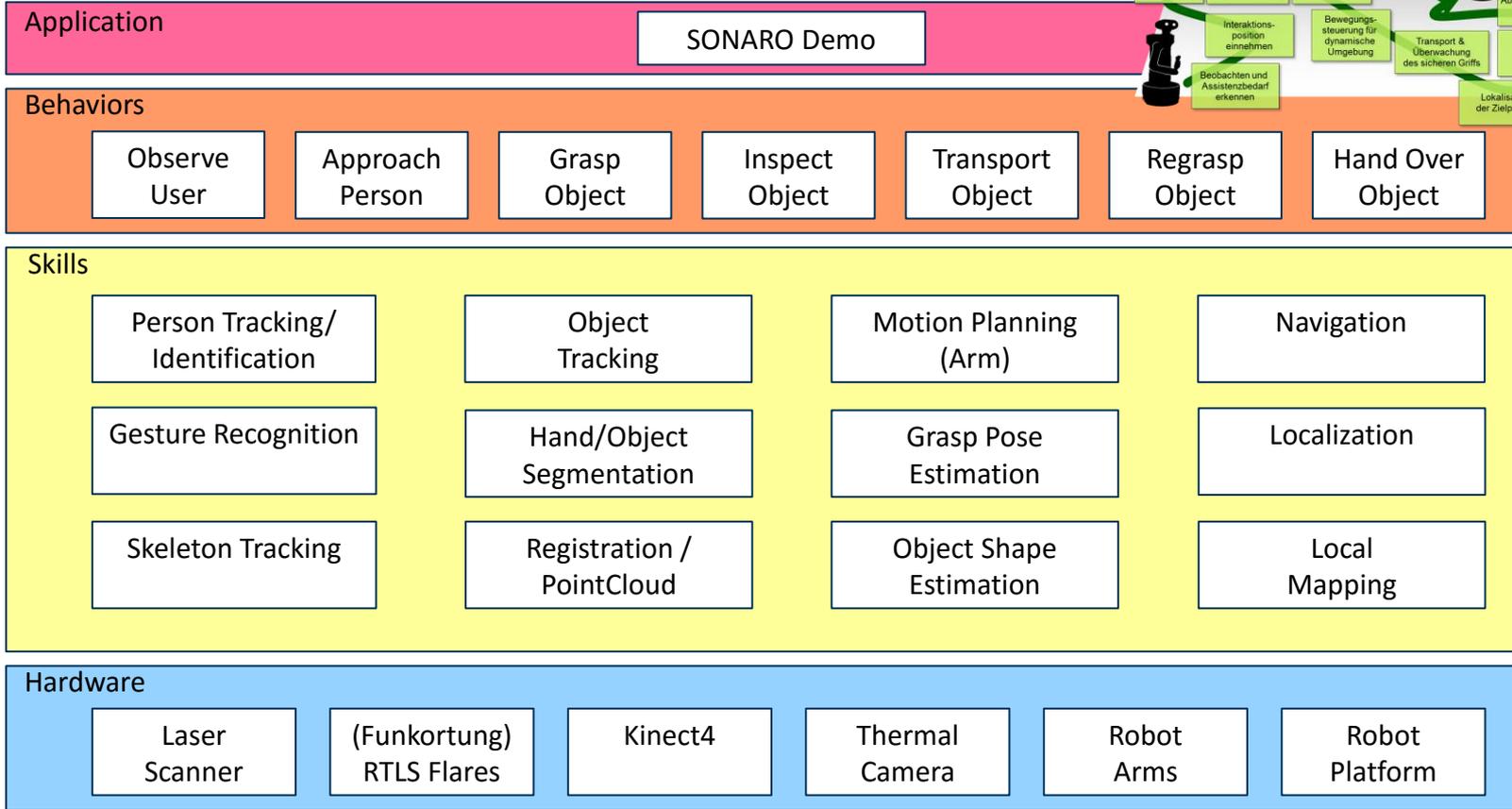
www.sonaro-projekt.de



Ablauf einer Objektübernahme/-übergabe



Systemarchitektur

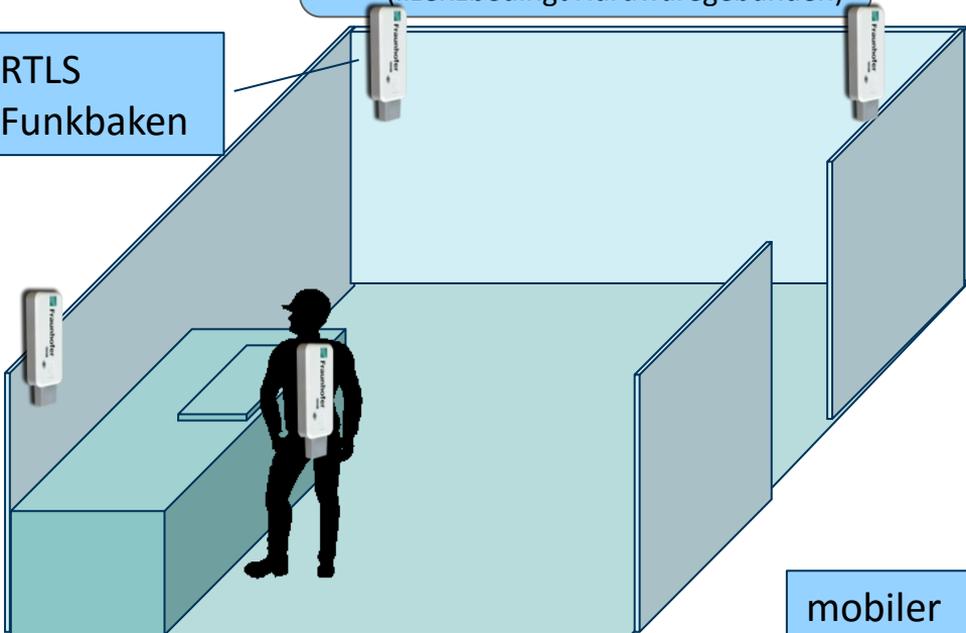


Konkrete Hardware Architektur

PC

- Real Time Localization System (RTLS) Empfänger
- Halcon Bibliothek (lizenzbedingt Hardwaregebunden)

RTLS Funkbaken



Thermokamera

RGB-D Kamera

Laptop mit GPU

- Kinect4 Tiefenkamera
- Skeletttracking

Nvidia Jetson Xavier GPU

- Thermokamera
- Hand-/Objektsegmentierung

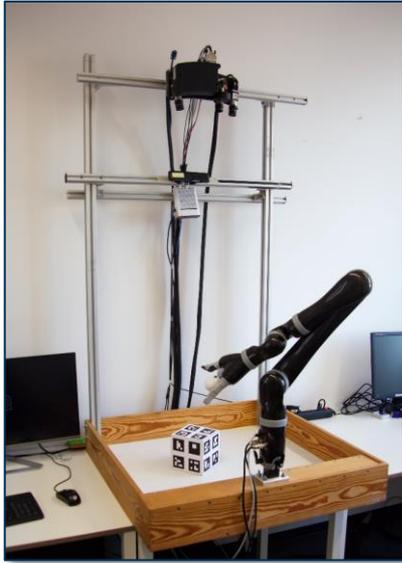
On Board PC

- Navigation,
- Bewegungssteuerung,
- Personenwahrnehmung,
- Greifen
- Ablaufsteuerung

mobiler
Roboter



Robotertechnik für die Umsetzung des Szenarios



Greiftestplatz

- Evaluation der Greifposenbestimmung
- Kameravergleich



Tiago

- Hauptentwicklungsplattform für SONARO
- Umsetzung des Demonstrationsszenarios



Scitos X3

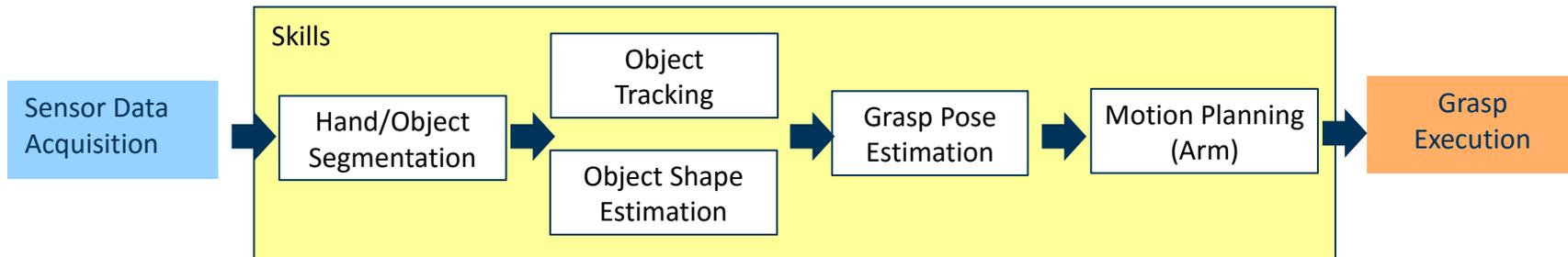
- Plattform mit zwei Armen
- Untersuchungen zum Umgreifen von Gegenständen



RB Kairos

- Untersuchungen zur Integration des multimodalen Kamerasystems in den Roboter

Untersuchte Methodenvielfalt für das Greifen



• RGB-D

Kameras:

- Kinect4Azure
- AstraOrbbec
- ASUS Xtion
- IOF Kamera

• Thermokamera

- I3system TE-Q1

• Heuristisch:

- Background Subtraction
- Planefitting

• 2D Farbbildsegmentierung

- PointRend
- YolactEdge

• 3D PointCloud Segmentierung

- XYZ RGB
- XYZ RGB Thermo

• Klassisch:

- ICP + Kalman-Filter Tracking
- Keypoint+ ICP

• Machine Learning:

- Pixel2Mesh
- Mesh RCNN
- ObMan (AtlasNet)
- DeepIM

• Analytisch:

- Online
- Approximativ
- Affordances

• Modellfrei

- Heuristisch

• Machine Learning:

- GG-CNN

• Dynamisch:

- ArmPilot

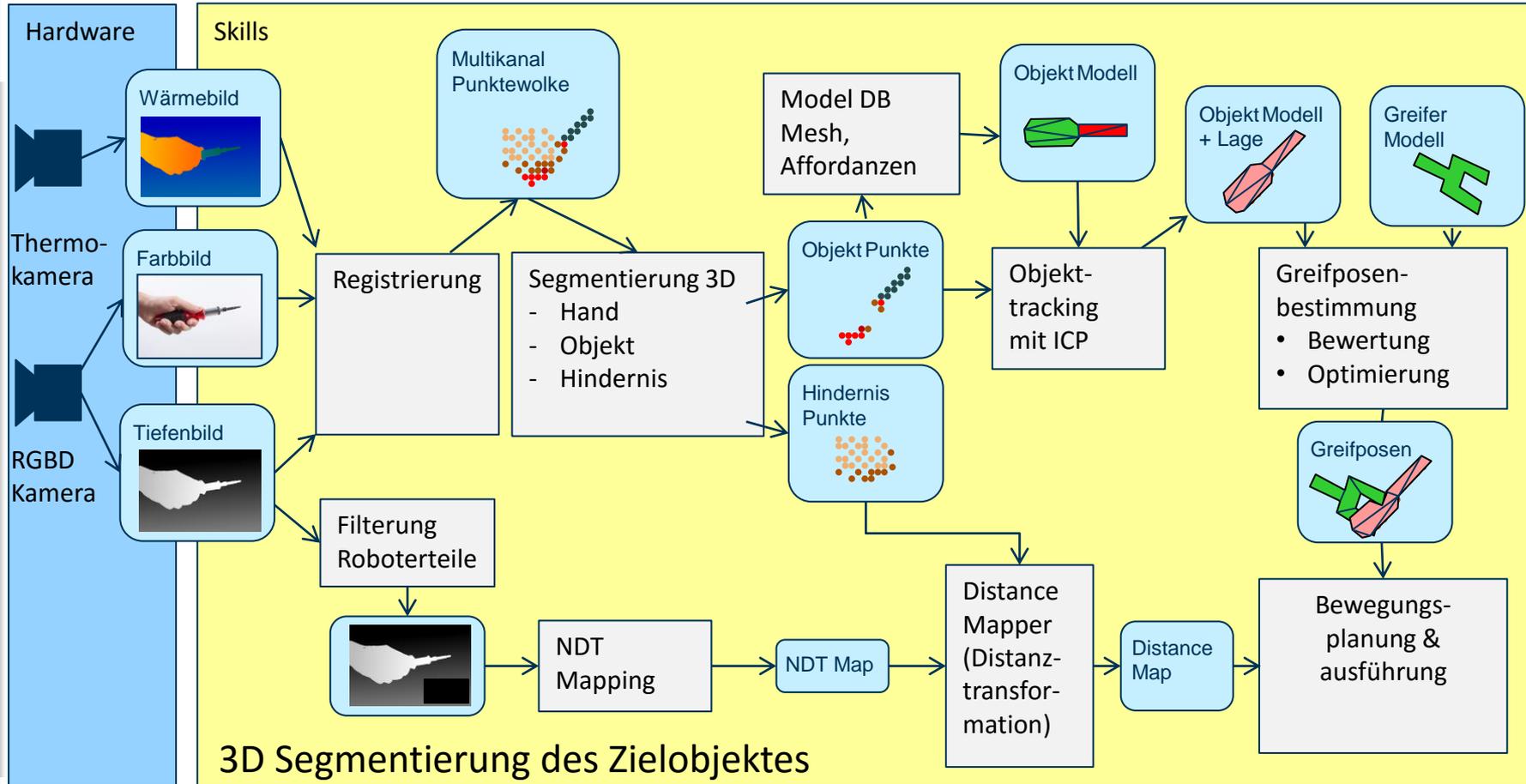
• Statisch:

- ROS MoveIt

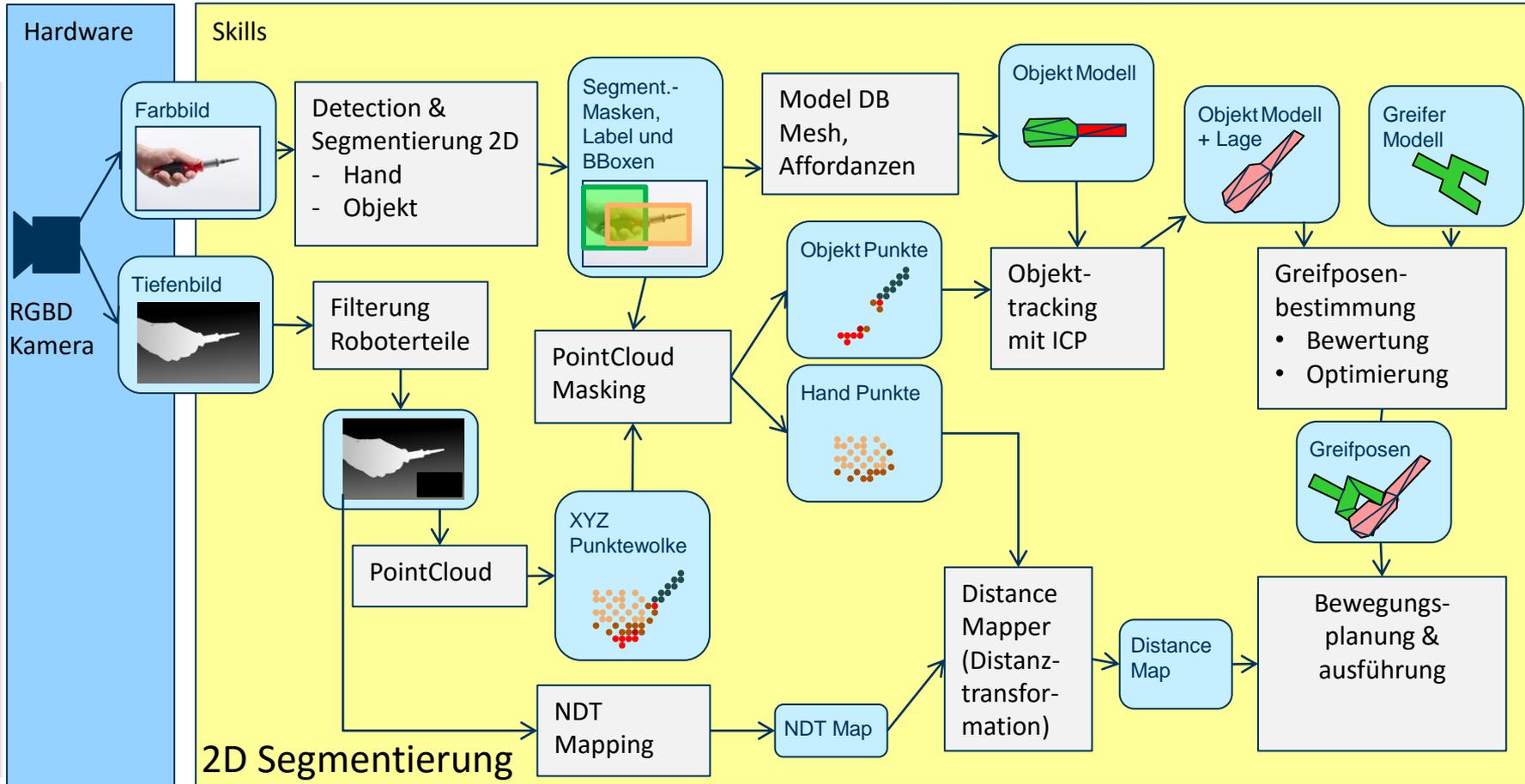
• Hardware:

- GraspBox (Kinova Arm)
- Tiago (5 Finger Hand, Parallel-Greifer)
- Scitos X3 (Kinova Arm)
- UR5 Arm

Umgesetzte Varianten der Greif-Pipeline (1)



Umgesetzte Varianten der Greif-Pipeline (2)



Forscherguppe SONARO

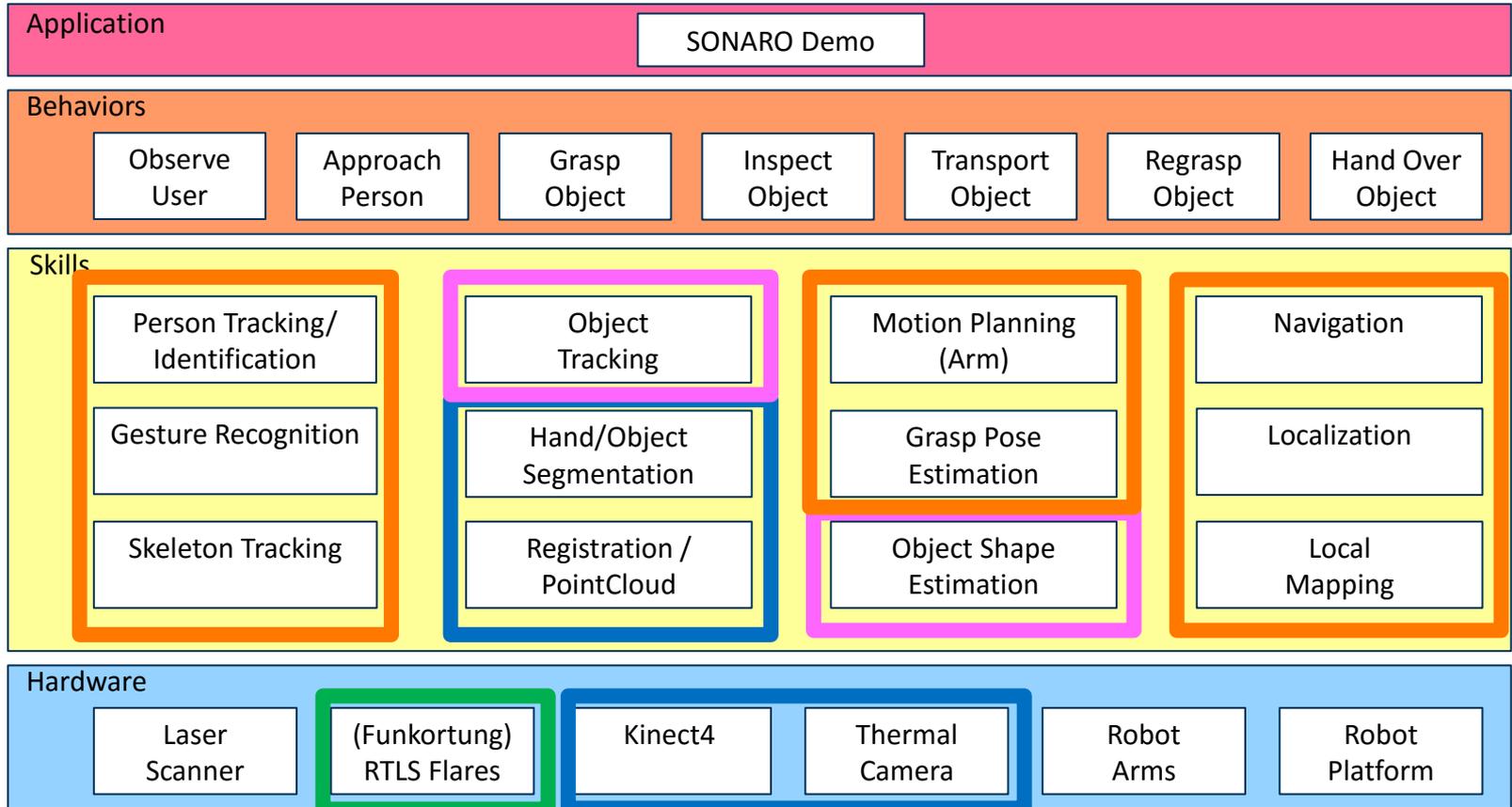
Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Beiratstreffen am 24.11.2021

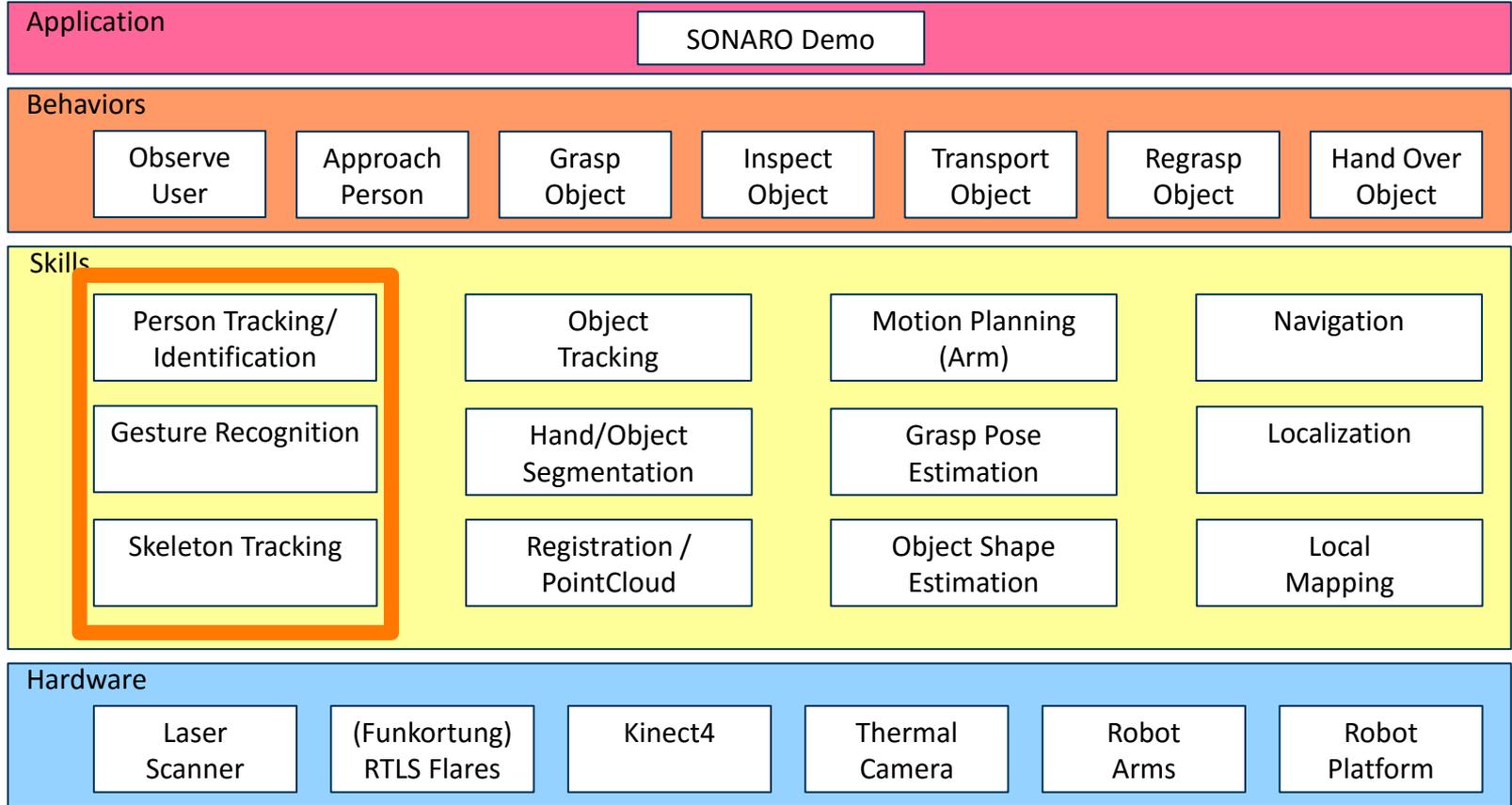
**Vorstellung der Forschungsergebnisse zu den methodischen
Teilaspekten des Szenarios**

www.sonaro-projekt.de





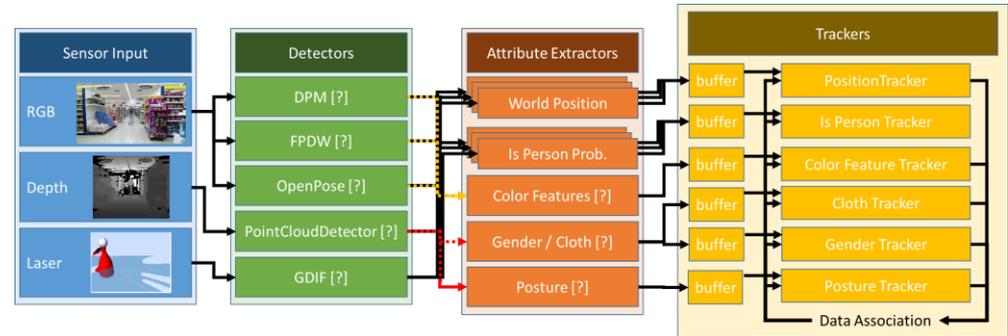
Personenwahrnehmung für die Interaktion (TUI NIKR)



Personenwahrnehmung für die Interaktion (1)

Ausgangssituation:

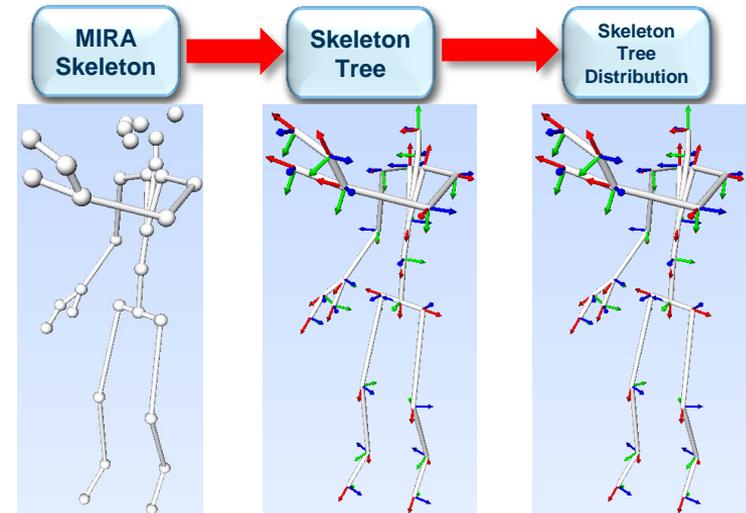
- Framework für multimodales Personentracking
- Verschiedene Detektoren können auf unterschiedlichen Sensordaten verschiedene Eigenschaften von Personen im Roboterumfeld erfassen
- Position der Person (nur Schwerpunkt) wird mittels Kalman-Filter geschätzt
- Wiedererkennung über Gesichtsfeatures oder Ansichts-basiert möglich (erfordert GPU Ressourcen / Deep Neural Networks)



Müller, St., Wengefeld, T., Trinh, T. Q., Aganian, D., Eisenbach, M., Gross, H.-M. *A Multi-Modal Person Perception Framework for Socially Interactive Mobile Service Robots*. Sensors, vol. 20 (2020) 3, 722, 18 pages

Entwicklungen im Projekt:

- Erweiterung des Trackingframeworks um Skelettdaten für jede getrackte Person
- Den Gelenkpunkten werden dabei auch Orientierungen hinzugefügt (benötigt für Übergabeposition)
- RTLS Tags (Funkortungssystem von HSM) als Beobachtungsquelle für Personenpositionen integriert
- Personenidentifikation mittels RTLS (keine aufwändige Gesichtserkennung auf dem Roboter PC)



Personenwahrnehmung für die Interaktion (3)

Intenterkennung:

- Zur Initiierung der robotischen Unterstützung sollen Zwischenschritte im Montageprozess und direkte Kommandos an den Roboter erkannt werden
- Aktivitäten und kommunikative Gesten → **Activity recognition**
- **State of the Art:**
 - Activity Recognition in Videosequenzen
 - **Activity Recognition auf Skelettdaten**

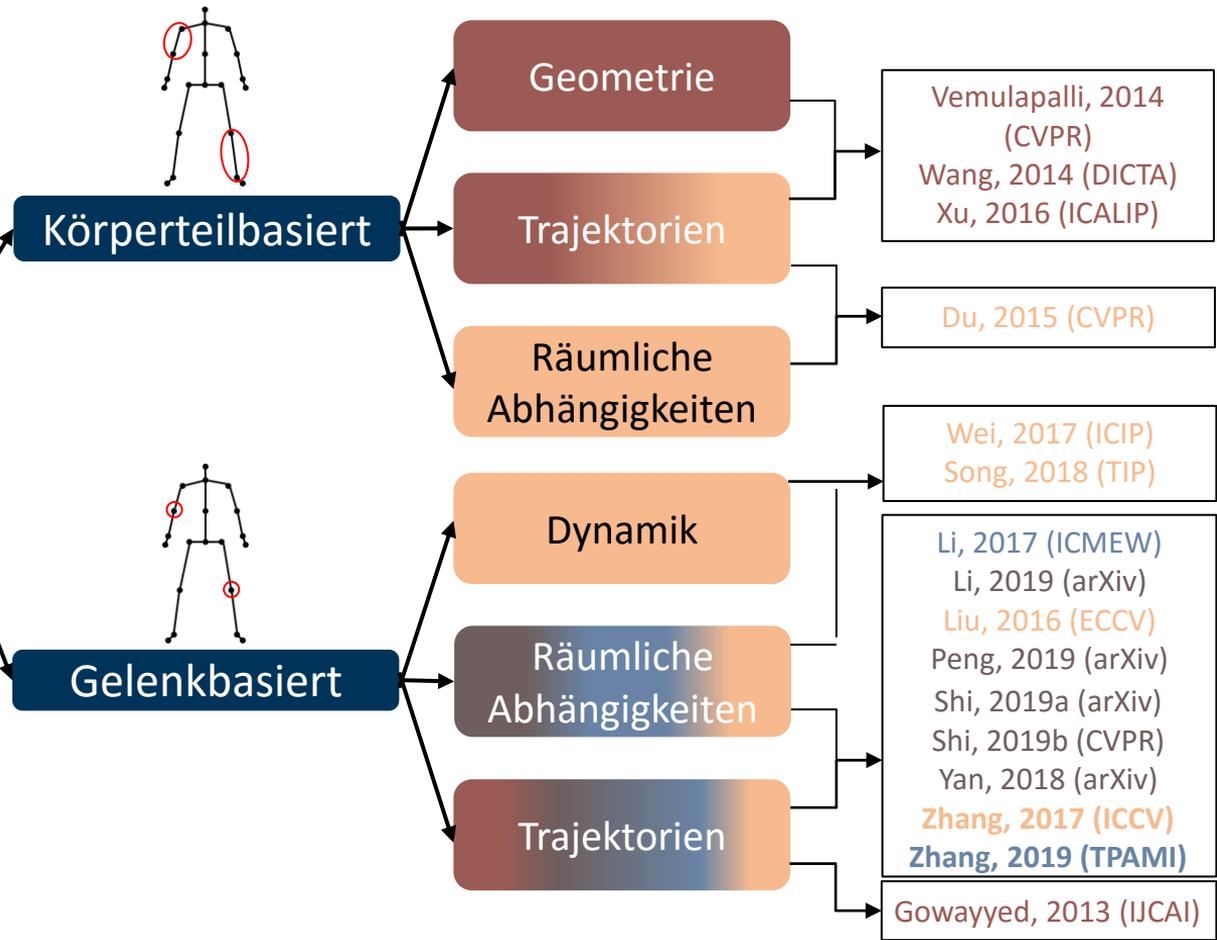
Personenwahrnehmung für die Interaktion (4)

State of the Art:

BA Amanda Robillard:
Untersuchung von 3D-Skelett-basierten Verfahren zur Aktionserkennung. TU Ilmenau 2020

Aktionserkennung

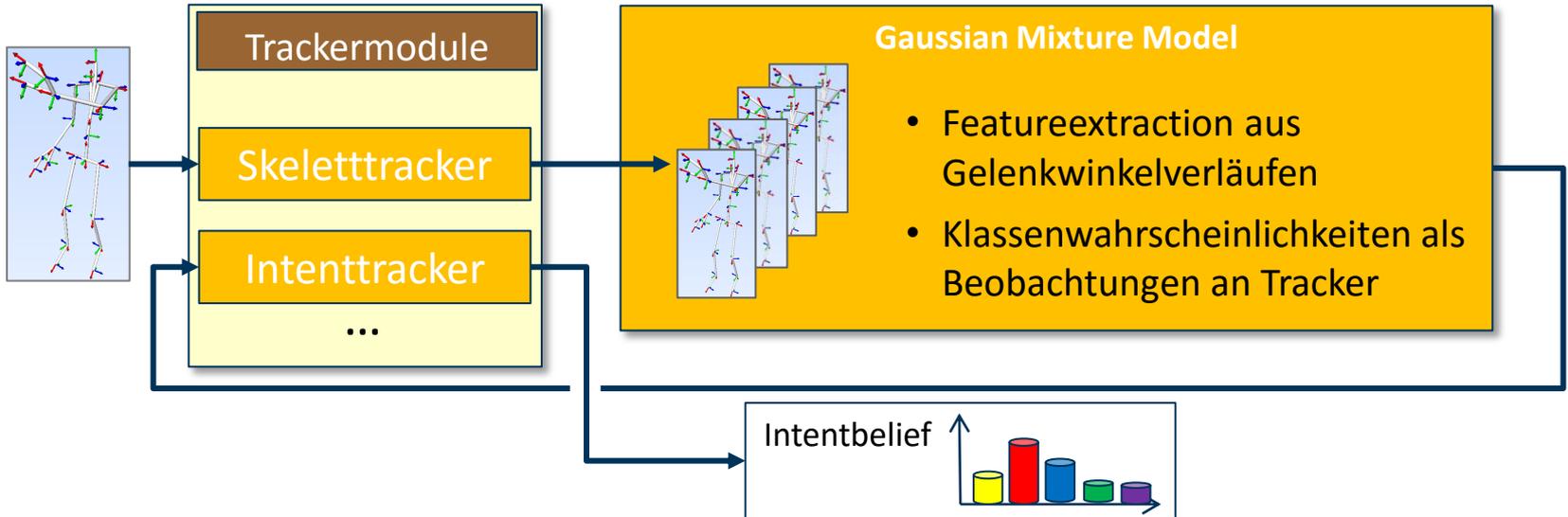
- Klassische
- RNN-basierte
- CNN-basierte
- GNN-basierte



Personenwahrnehmung für die Interaktion (5)

Umsetzung der Intenterkennung in SONARO:

- Pragmatische Lösung (Ressourcensparend / kein Deep Learning)
- Aus zeitlichen Verläufen der Skelettgelenkwinkel werden statistische Features extrahiert (automatische Featureselection)
- Klassifikation mittels Gaussian Mixture Model → zeitliche Integration im Tracker mittels Bayes-Filter



Konnten die Erwartungen erfüllt werden?

Aspekt	Erwartungen zu Projektbeginn	Erreichter Stand
Personenwahrnehmung	Neue Deep Learning Verfahren für Personendetektion sollen integriert werden	Yolo Detektor ist integriert, wird aber nicht genutzt (Ressourcenproblem), Skelettschätzung des Kinect 4 SDK integriert (auch GPU)
Personenwiedererkennung	Zielpersonen sollen ansichtsbasiert erkannt werden	RTLS System löst das Wiedererkennungsproblem im Demonstrator
Intenterkennung	Prozessüberwachung und Modell sollte zur Vorhersage des Assistenzbedarfs dienen	Direkte nonverbale Aufforderung an den Roboter über Gestenerkennung realisiert

Lokalisation / Identifikation der Zielpersonen (HSM)

Application

SONARO Demo

Behaviors

Observe
User

Approach
Person

Grasp
Object

Inspect
Object

Transport
Object

Regrasp
Object

Hand Over
Object

Skills

Person Tracking/
Identification ✓

Object
Tracking

Motion Planning
(Arm)

Navigation

Gesture Recognition ✓

Hand/Object
Segmentation

Grasp Pose
Estimation

Localization

Skeleton Tracking ✓

Registration /
PointCloud

Object Shape
Estimation

Local
Mapping

Hardware

Laser
Scanner

(Funkortung)
RTLS Flares

Kinect4

Thermal
Camera

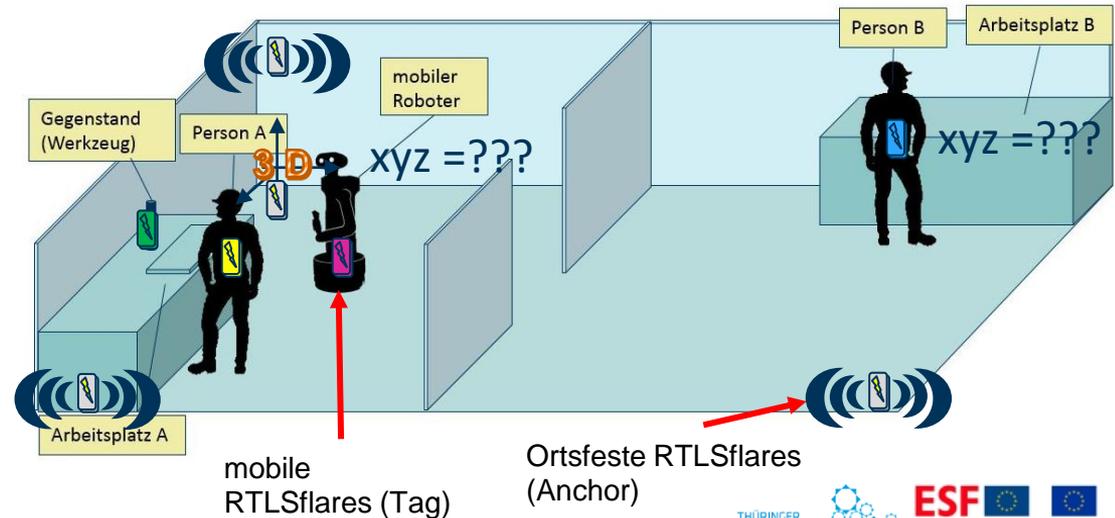
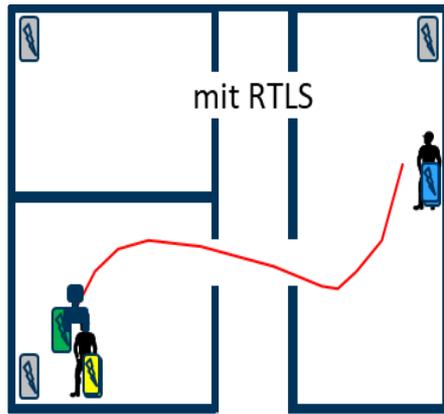
Robot
Arms

Robot
Platform

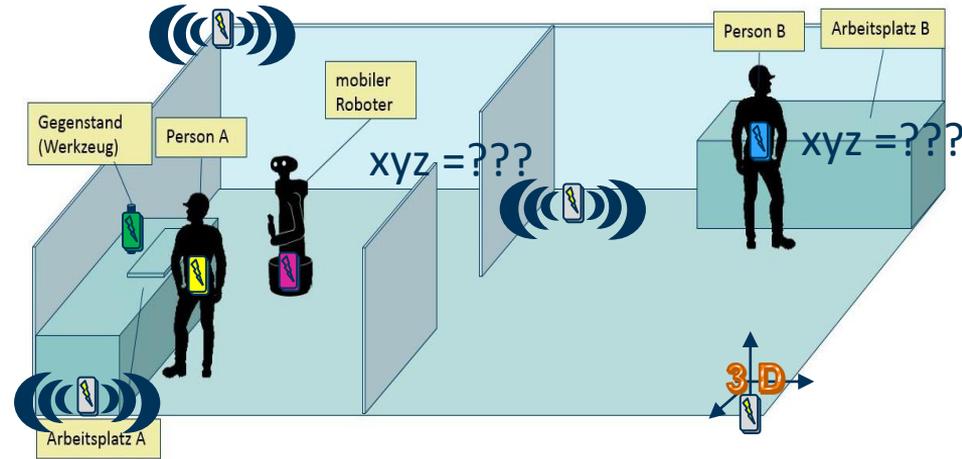
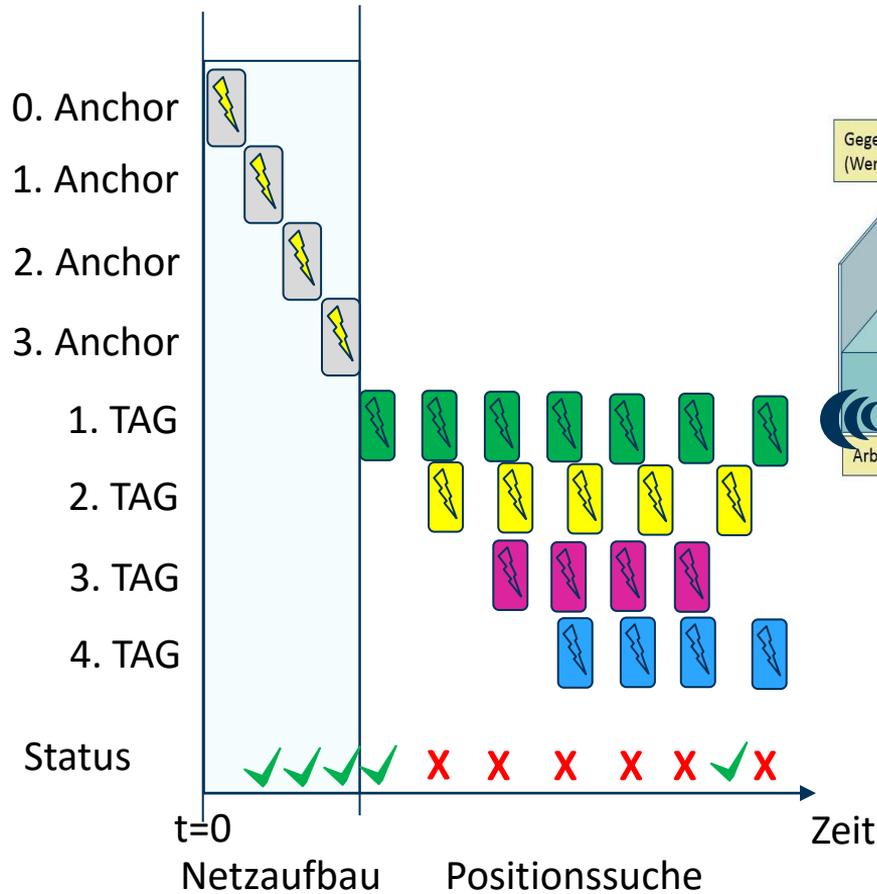
Anforderungsanalyse für das Funkortungssystem

Anwendungsmöglichkeiten des Real Time Localisation System (RTLS):

- Unterschiedliche Räume mit einem Funklokalisationsnetzwerk ausgestattet
- Mindestens 4 ortsfeste RTLS-Flares bilden ein Referenzkoordinatensystem in dem sich alle Teilnehmer orten können (insbesondere auch mobile RTLS-Flares)
- Im SONARO Szenario soll dem Roboter somit die Position von Zielpersonen und –objekten auch außerhalb seines Sichtbereiches übermittelt werden.

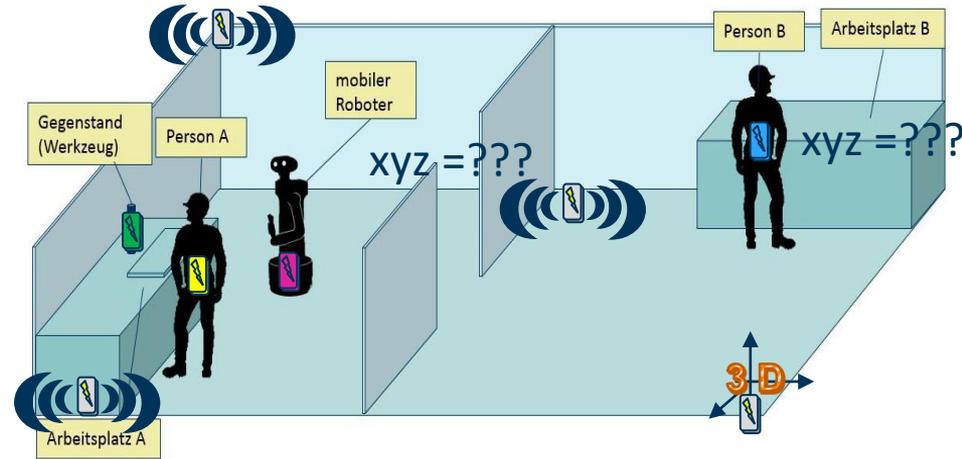
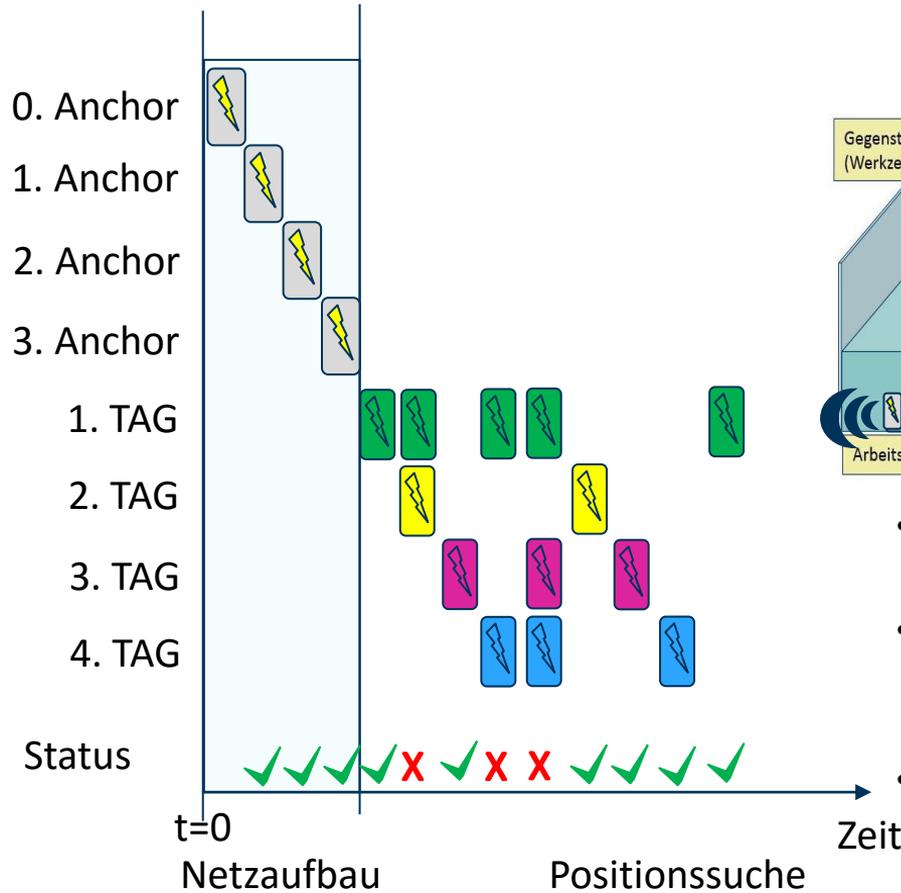


6. Ausgangssituation (Luftschnittstelle)



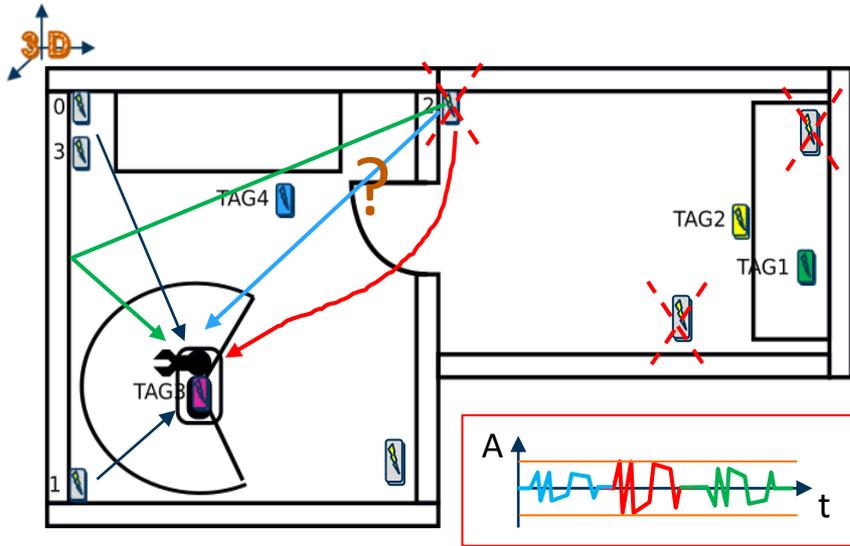
- Positionsbestimmung nicht in der erforderlichen Aktualisierungsrate realisierbar
- **X** Kollisionen der Funkpakete verschlechtern die Aktualisierungsrate aller Teilnehmer
- Ziel: Verbesserung der Aktualisierungsrate / Fehlererkennung

7. Optimierung der Luftschnittstelle (Skalierbarkeit)



- Verwendung eines Kollisionsvermeidungsalgorithmus (CSMA-CA)
- Zufalls-Synchronisation für die Entflechtung der Kommunikationspakete (random back off timer)
- Verbesserung der Aktualisierungsrate zur Positionsbestimmung

Untersuchung von Störungen bei der Positionsbestimmung (1)



Erkennung von gestörten
Entfernungsmessungen mit
Verwendung neuer Signalparameter

z.B.:

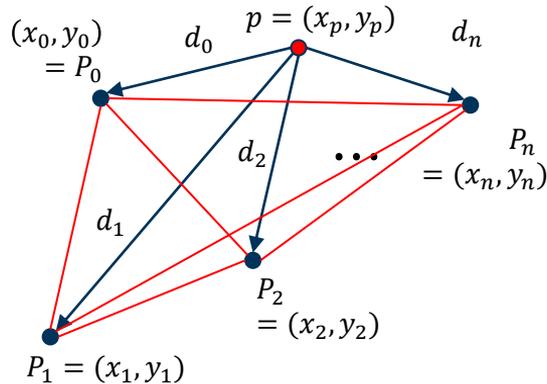
- First Path Power,
- Received Power,
- Channel Standard & Maximum Noise

Projektseminar Tony Schneider, Nicolas Schmitt:
*Bestimmung materialabhängiger Messfehler
beim Einsatz von RTLS Lokalisierungssystem,
Hochschule Schmalkalden, 2019*

**Projektseminar Stephan Hintz, Adrian Herrman,
Nicolas Schmitt:** *Messfehlerbestimmung im 3D-
Raum von UWB Signalen zur Indoor-Navigation
mit RTLS , Hochschule Schmalkalden, 2020*

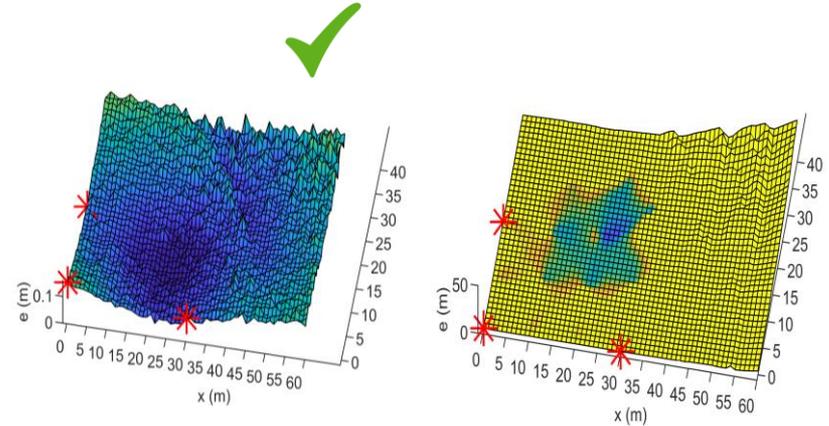
Untersuchung von Störungen bei der Positionsbestimmung (2)

Untersuchung weiterer Fehlerminimierungsmethoden zur Positionsschätzung im Sensornetzwerk:



Least Squares Schätzung

Taylorreihen-Entwicklung

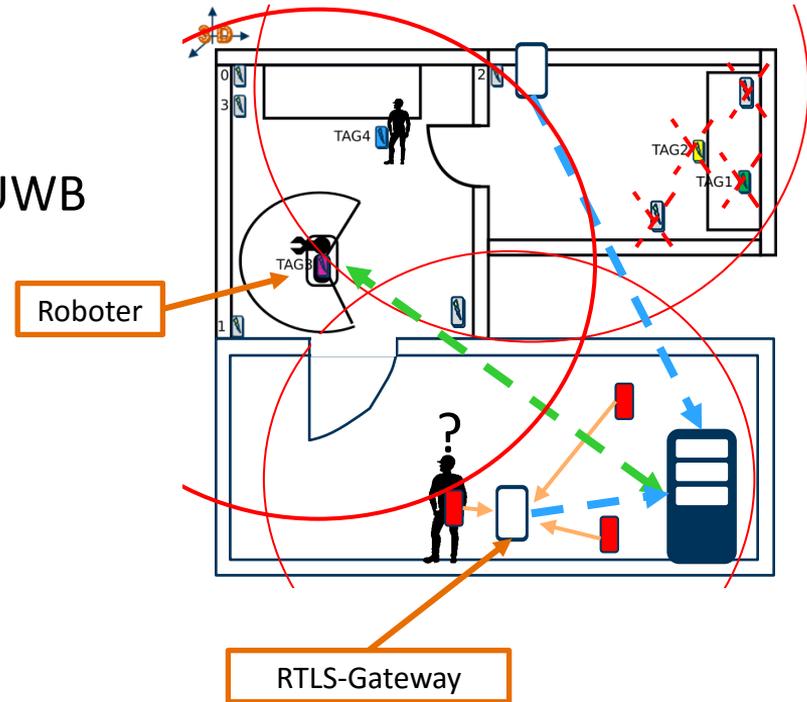


N. Fränzel, N. Greifzu, M. Schneider, A. Wenzel: *Bootstrapping a Wireless Localization System*, IEEE: The 4th International Conference on Advanced Communication Technologies and Networking, CommNet 2021, ISBN: 978-6654-0306-1, 12/2021

N. Fränzel, N. Greifzu, A. Wenzel: *Robuste Lokalisierung in drahtlosen Sensornetzwerken*, IEEE: The 4th International Conference on Advanced Communication Technologies and Networking, CommNet 2021, ISBN: 978-1-6654-0306-1, 12/2021

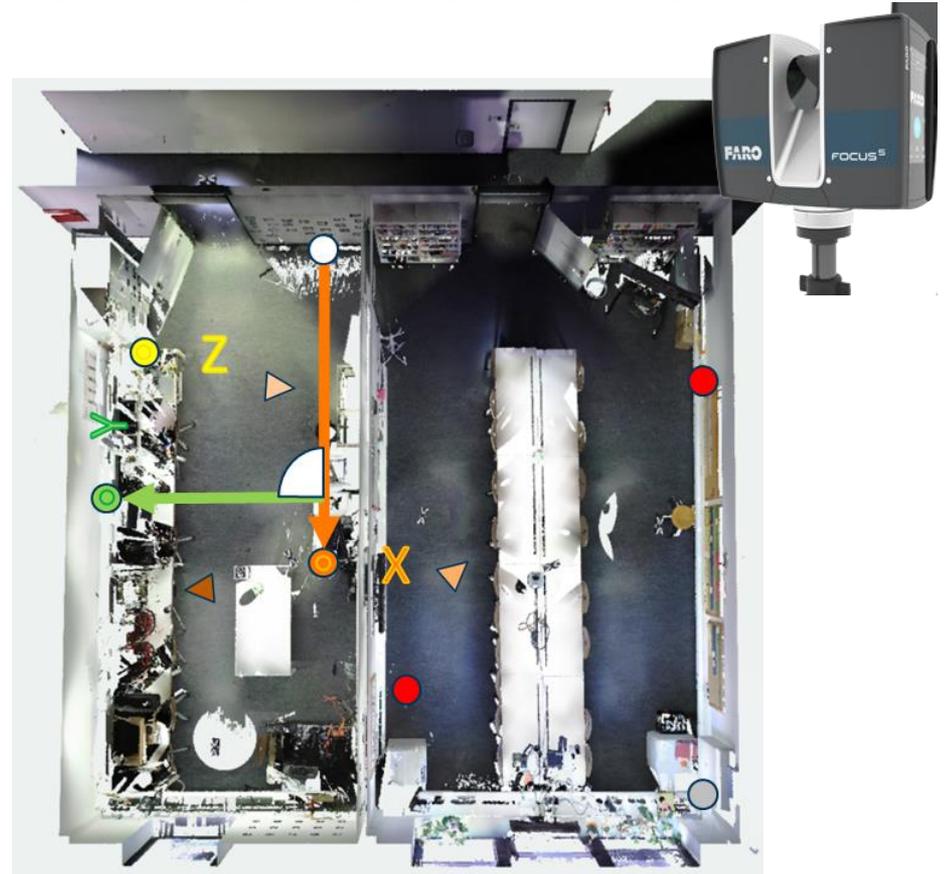
Vergrößerung der Funkortungsreichweite

- Netzwerkerweiterung mit RTLS-Gateway
- Neue Hardware mit:
 - PoE, WLAN, USB, SD-Speicher, UWB



Integration des RTLS im SONARO Demonstrator

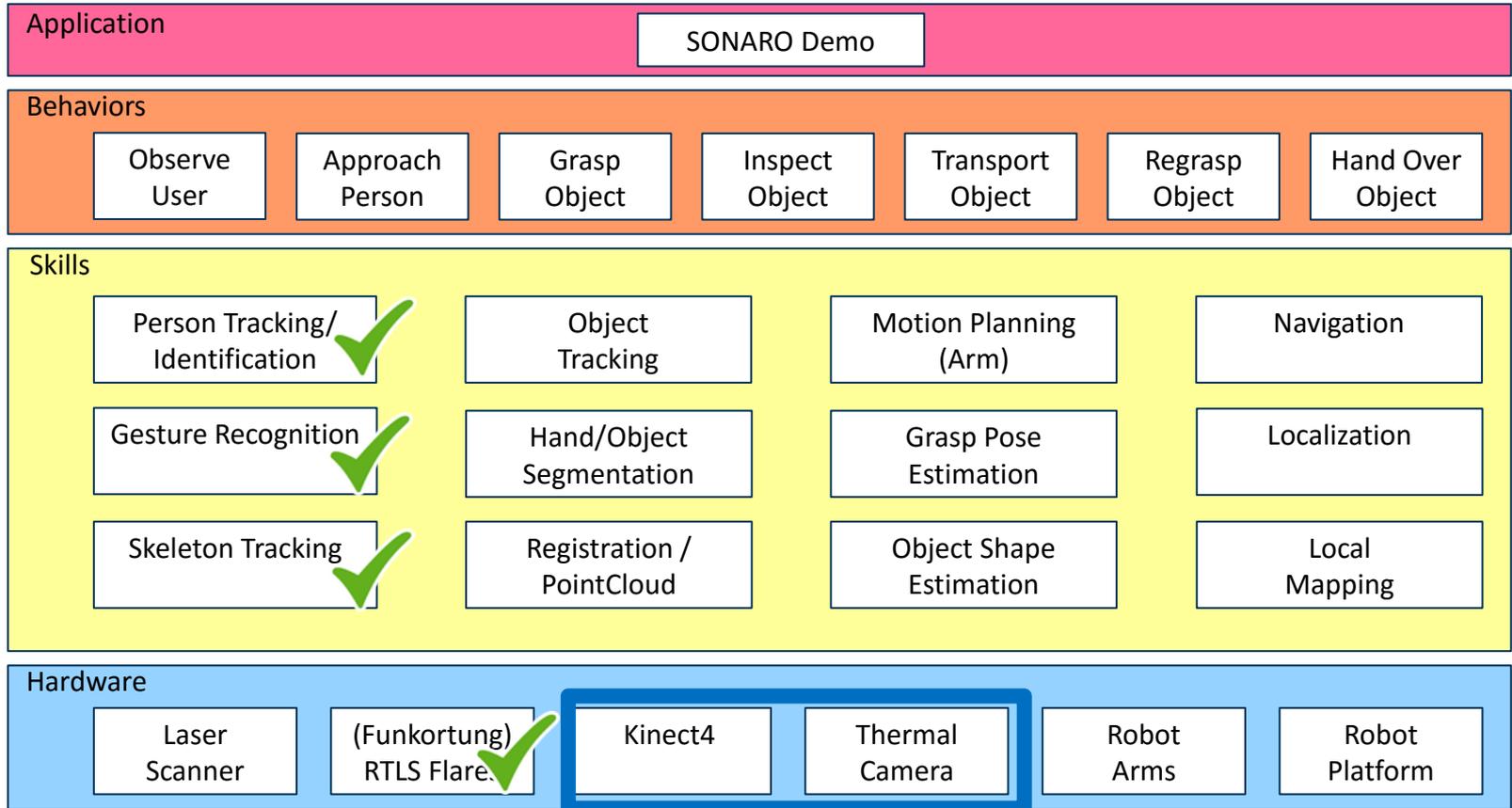
- Entwicklung von Softwaremodulen für die Robotik-Middlewareframeworks **MIRA** und **ROS**
- Anbindung an das Trackingframework von TUI NI&KR
- Installation und Einmessung des RTLS-Netzwerkes im Labor an der TU Ilmenau



Konnten die Erwartungen erfüllt werden?

Aspekt	Erwartungen zu Projektbeginn	Erreichter Stand
Lokalisationsgenauigkeit	Störeinflüsse bei der Signalausbreitung sollten untersucht und im System berücksichtigt werden	Störeinflüsse in verschiedenen Signalparametern identifiziert und Positionsgenauigkeit durch Fehlerminimierung im Netzwerk verbessert.
Updaterate	Das Netzwerk sollte skalierbar in der Anzahl der Teilnehmer werden	Kollisionsproblem durch CSMA-CA Verfahren gelöst
Reichweite	Der Roboter soll die Positionen aller getagten Personen und Objekte auch außerhalb seiner unmittelbaren Empfangsreichweite erhalten	Entwicklung eines WLAN basierten RTLS-Gateways zur Übertragung der Lokalisationsdaten

Auswahl und Evaluation der multimodalen 3D Sensorik (TUI QBV)



Auswahl und Evaluation der multimodalen 3D Sensorik (TUI QBV) (1)

Ausgangspunkt

- Verschiedene RGB-D Sensoren verfügbar
- Verschiedene 2D-Kamera verfügbar:
RGB, NIR, SWIR, FIR sowie **multispektral**
- Farbbasierte 2D- / 3D-Menscherkennung

Fragestellung

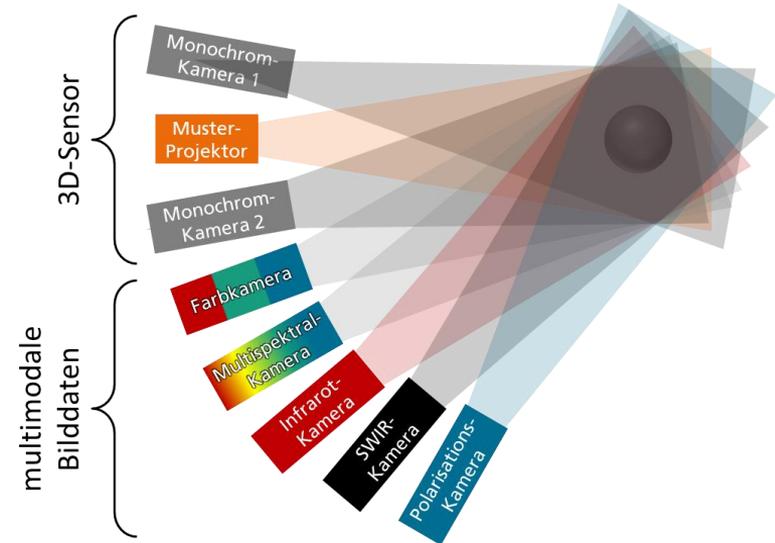
- In welchem Spektralbereich ist eine sichere und saubere Segmentierung der menschlichen Hand und der gehaltenen Objekte möglich?



Auswahl und Evaluation der multimodalen 3D Sensorik (TUI QBV) (2)

Entwicklungen im Projekt

- Ansatz: Verwendung von multimodalen Bilddaten, um sicherer und sauberer menschliche Hand und Objekte an der Hand zu segmentieren
- Kandidaten von Bildmodalitäten: RGB, 3D, SWIR, NIR, FIR (Thermo) oder Polarisation)
- Aufgrund der Thermostabilität des menschlichen Körpers wurde FIR gewählt, um multimodalen Bilddaten zu erhalten.
- RGB + 3D + FIR (Thermo) (**XYZ-RGB-T multimodale 3D-Bilddaten**)



Aufbau eines multimodalen Sensorsystems (TUI QBV) (1)

Kalibrierung:

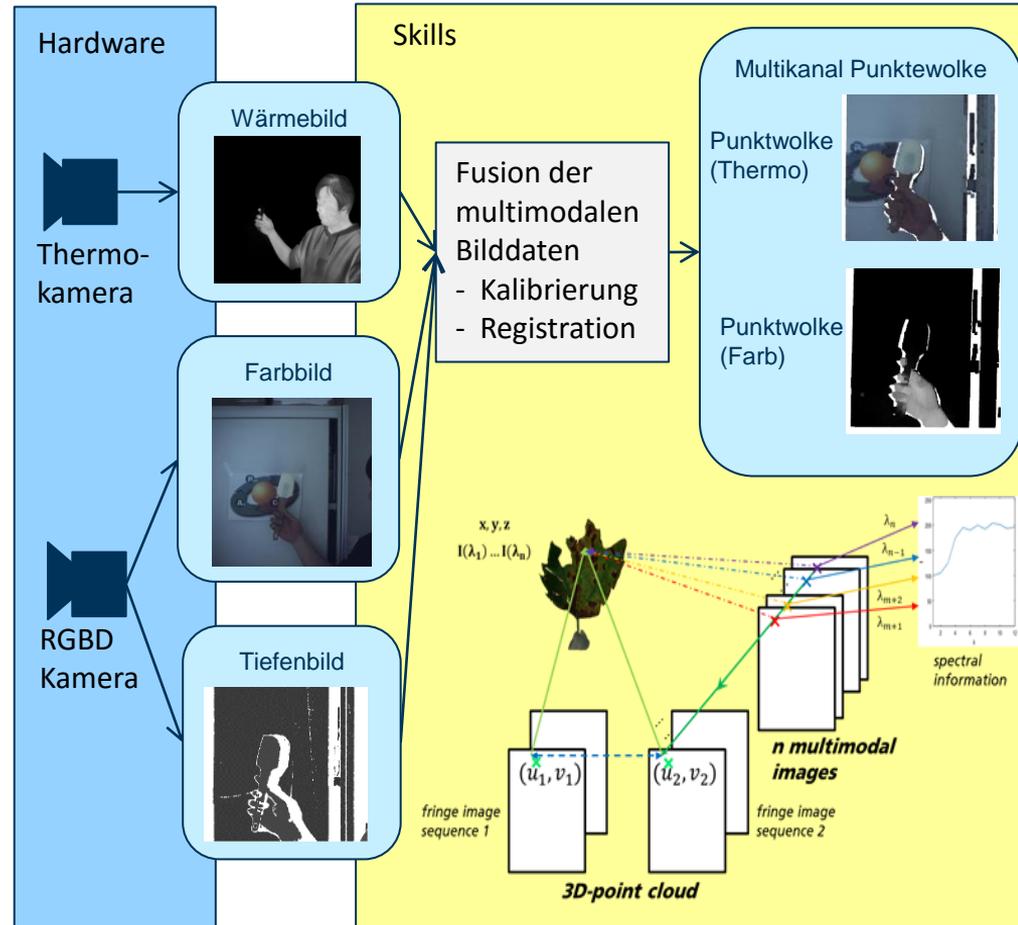
- Intrinsische Parameter von jedem Sensor und untereinander
- Extrinsische Parameter müssen kalibriert werden.

Registrierung:

- Mittels Intrinsischer- und Extrinsischer Parameter kann die sogenannte multimodale Punktwolke (XYZ-RGB-T) ermittelt werden.

Rosenberger M., Zhang C., Zhang Y., Notni G.

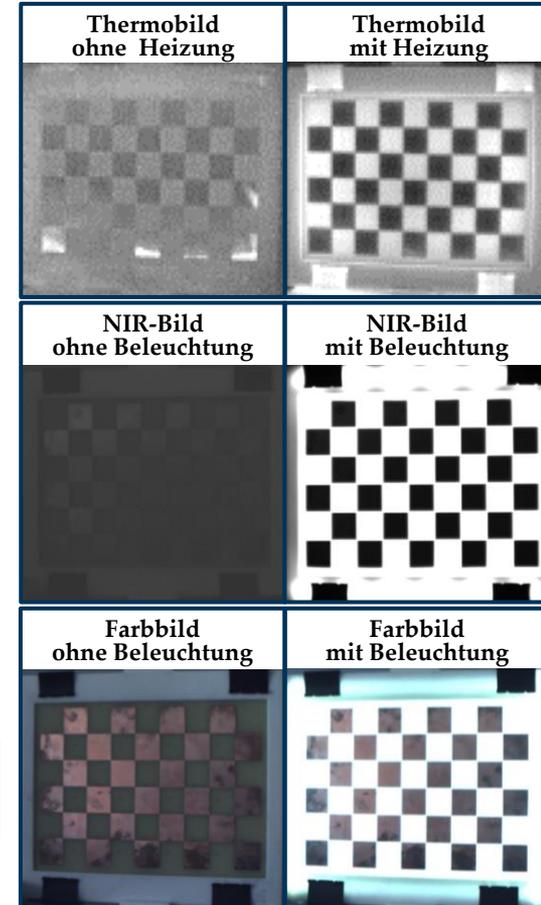
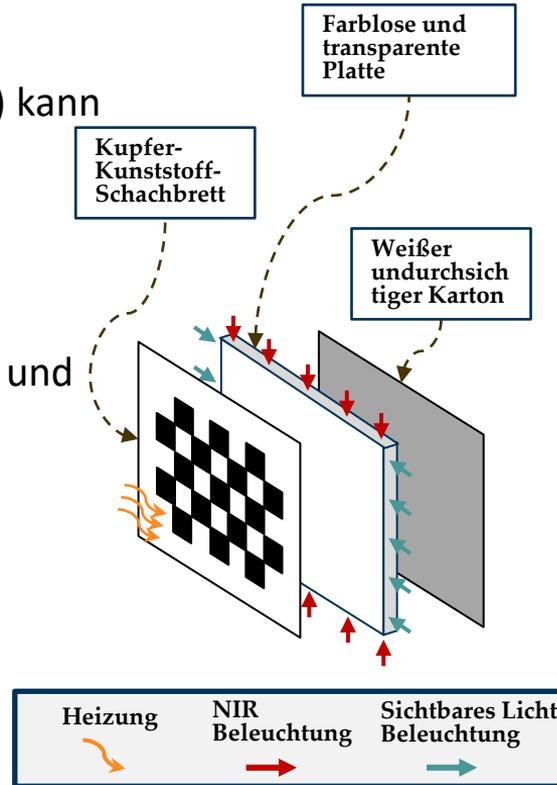
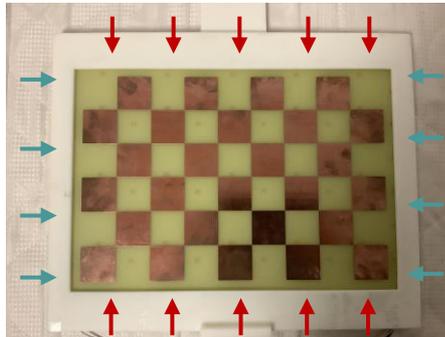
3D high-resolution multimodal imaging system for real-time applications; in: SPIE Dimensional Optical Metrology and Inspection for Practical Applications IX. München, 2020



Aufbau eines multimodalen Sensorsystems (TUI QBV) (2)

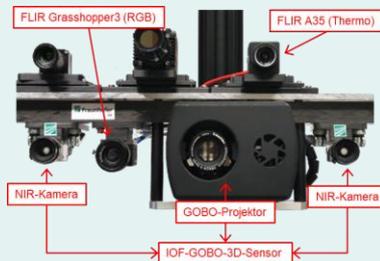
Kalibrierung eines Sensorsystems mit einer Thermokamera:

- Problem:** Ein gedrucktes Kalibrierungstarget (Schachbrett) kann nicht von der Thermokamera aufgenommen werden.
- Lösung:** Ein spezifisches Kalibrierungstarget aus Kupfer-Kunststoff mit Eigenbeleuchtung und Erwärmung hergestellt.



Stationäres Sensorsystem

- IOF-GOBO-3D-Sensor (3D)
- FLIR A35 (Thermo)
- FLIR Grasshopper3 (RGB)



Mobiles Sensorsystem (auf TIAGO-Roboter)

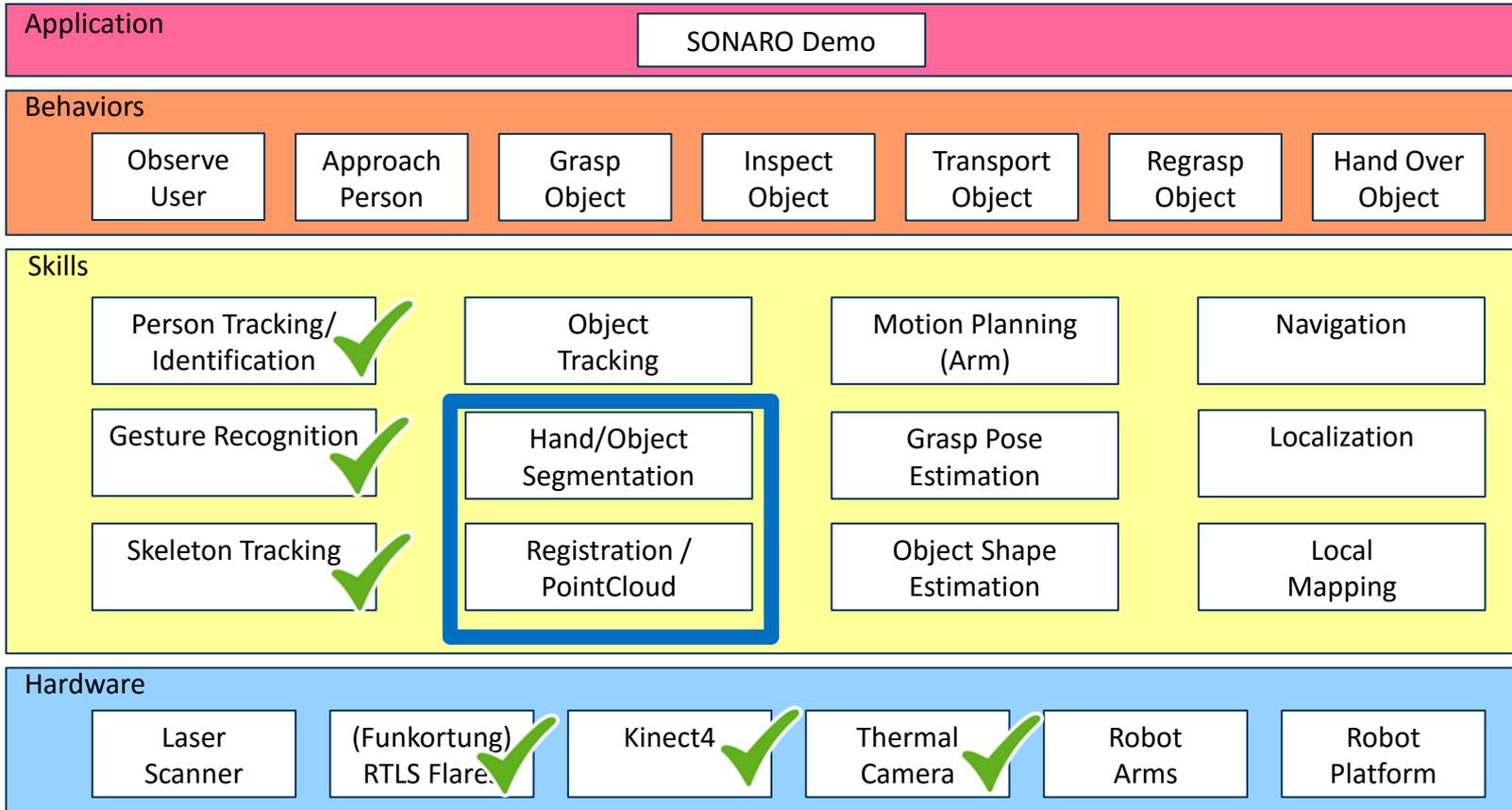
- Kinect 4 Azure (RGB-D)
- I3Systems TE-Q1 (Thermo)
- Astra Orbbec S (RGB-D)



	IOF-GOBO-3D-Sensor	FLIR A35 (Thermo-kamera)	FLIR Grasshopper 3 (Farbkamera)
Resolution	1024 x 1024	320 x 256	2048 x 2048
Frequency	36 Hz	60 Hz	90 Hz
Field of view	48° x 44°	63° x 50°	50° x 50°
Mean depth error	0,3 mm	-	-

	Kinect 4 Azure	I3System TE-Q1	Astra Orbbec (Astra S)
Resolution	1024 x 1024	384 x 288	640 x 480
Frequency	15 Hz	< 9 Hz	30 Hz
Field of view	120° x 120°	56° x 41°	60° x 49°
Mean depth error	> 1 mm	-	> 1 mm

3D Datenverarbeitung / Objektsegmentierung (TUI QBV)

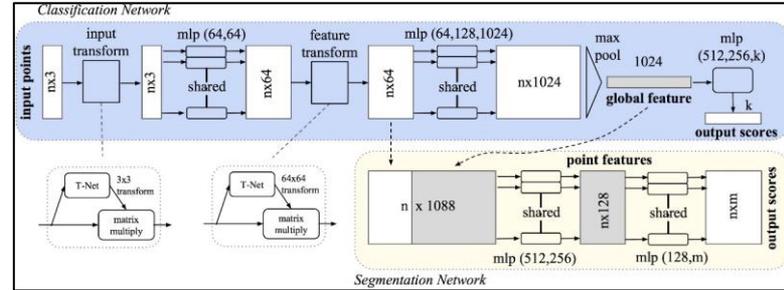


Ausgangspunkt

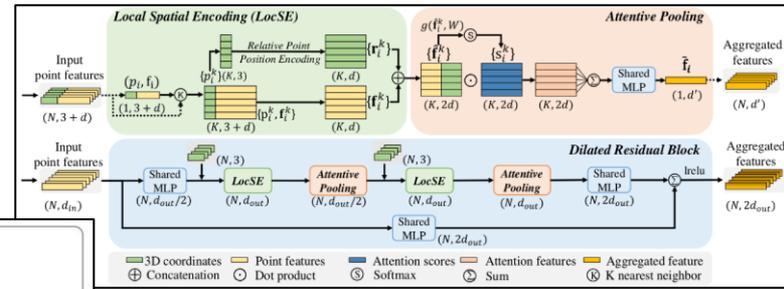
- 3D-Punktwolkensegmentierung:
 - PointNet, PointNet++ und RandLANet
- 2D-Bildsegmentierung:
 - MaskRCNN, PointRend und YolactEdge

Fragestellung

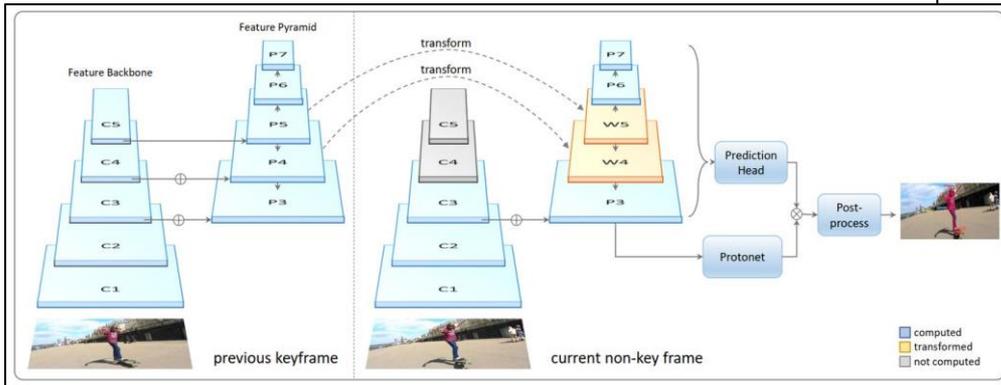
- Sind die Netzwerke für multimodale Daten geeignet?



PointNet



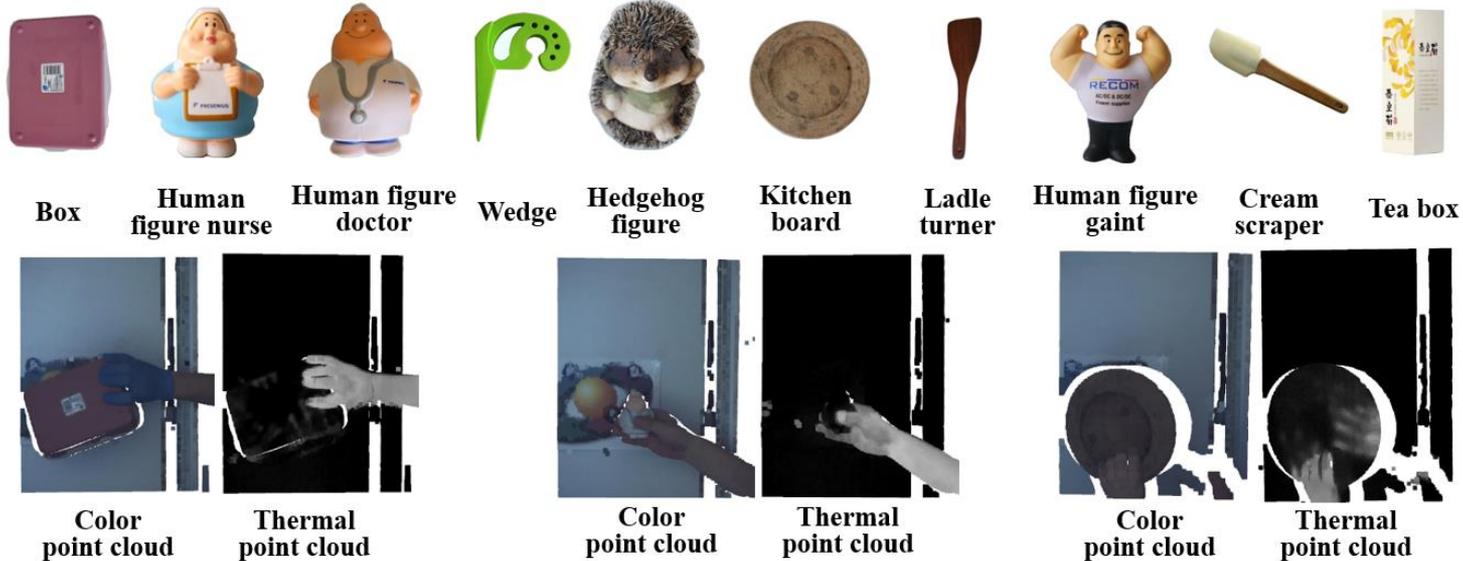
RandLANet



YolactEdge

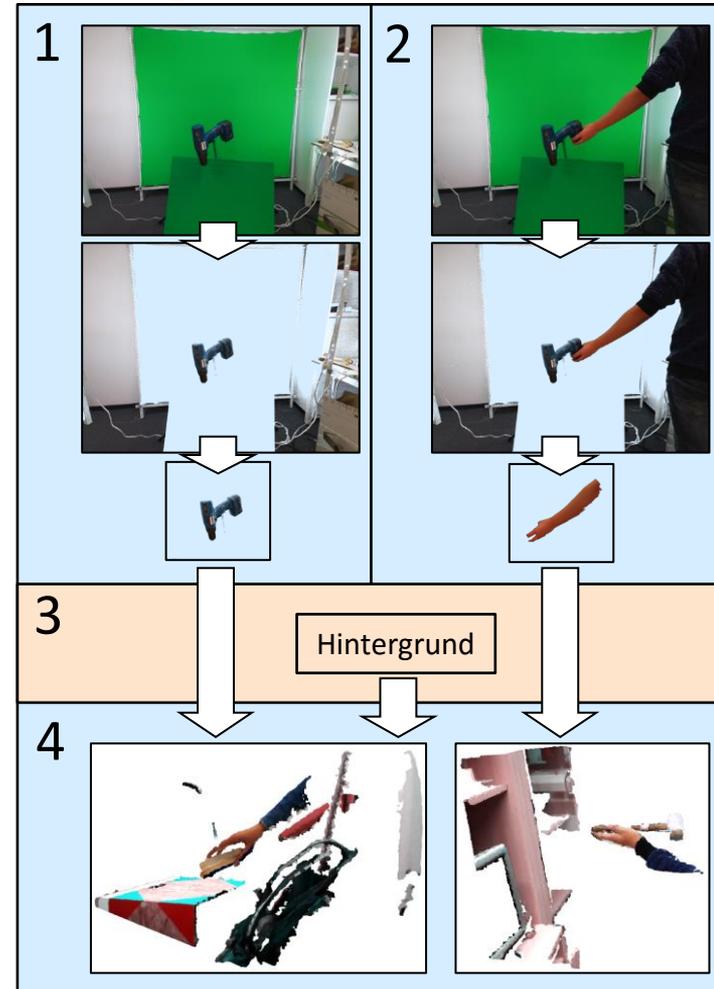
Multimodaler Datensatz 1 (10-Objekte-Datensatz)

- Aufgenommen mit **stationärem Sensorsystem**
- **12 Klassen:** 10 Objekte, Hand und Hindernis
- Zur Evaluation der Leistung der multimodalen (XYZ-RGB-T) Punktwolkensegmentierung



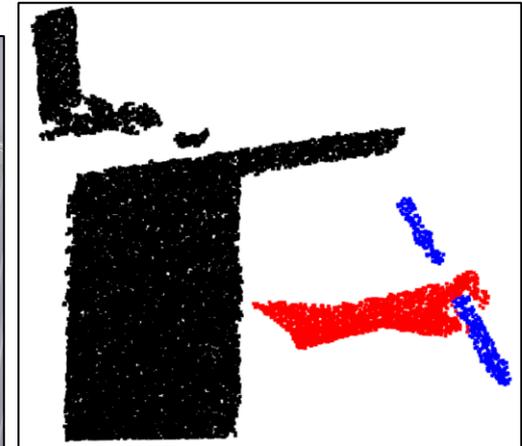
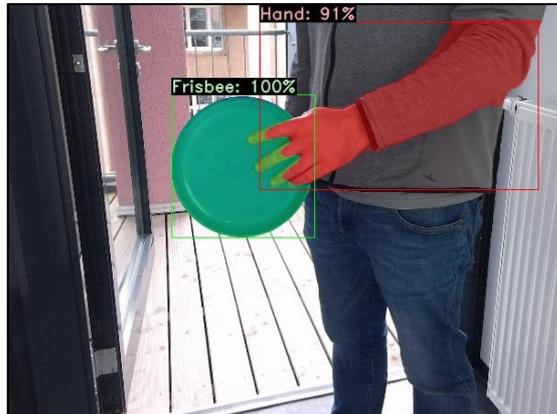
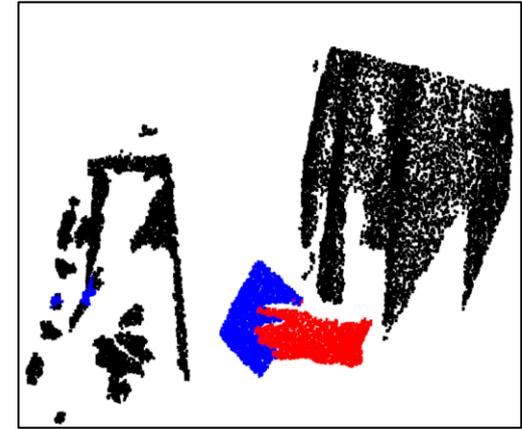
Multimodaler Datensatz 2 (3-Klassen-Datensatz)

- Aufgenommen mit **mobilem Sensorsystem**
- **Nur 3 Klassen:** Hand, alle Objekte an der Hand und Hindernis
- **Vollautomatische Annotation:**
 1. Punktwolken von Objekten in festen Posen werden mit Hilfe von einem grünen Hintergrund segmentiert.
 2. Handpunktwolken werden durch Halten des festen Objekts mit der Hand in verschiedenen Posen gewonnen.
 3. Hintergrundpunktwolke werden aufgenommen.
 4. Datensatz wird mit zufälligem Vorder- und Hintergrund erzeugt.
- Zur Evaluation der Leistung der Punktwolkensegmentierung **für neue oder unbekannte Objekte** (nicht vorhanden im Datensatz)



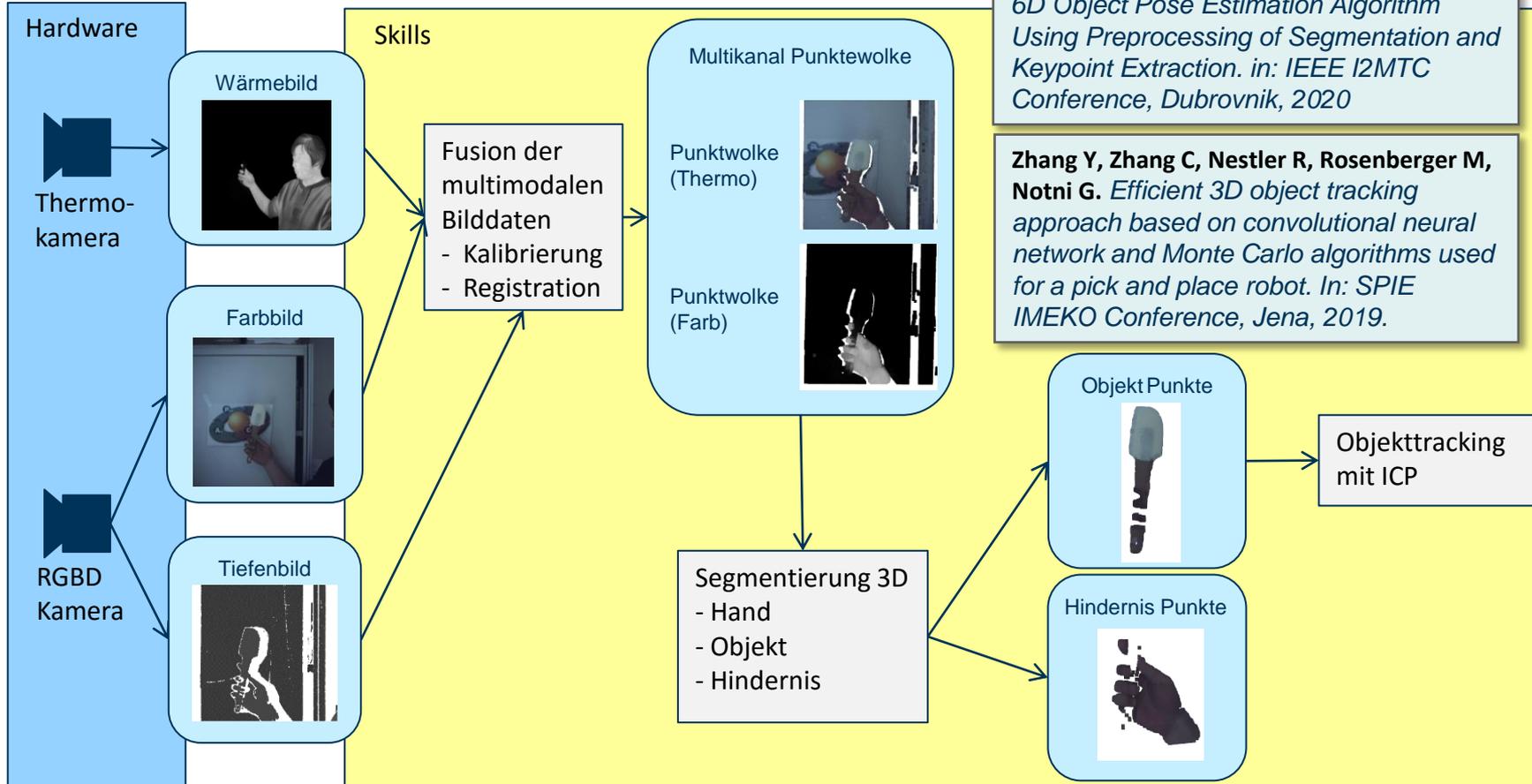
Entwicklungen im Projekt

- 3D-Punktwolkensegmentierung (semantische Segmentierung) mittels multimodaler Punktwolke zur sicheren und sauberer Hand- & Objektsegmentierung
- 2D-Bildsegmentierung (Instanzsegmentierung) mittels Farbbild
- Segmentierung für **unbekannte** Objekte



3D Hand- /Objektsegmentierung (1)

• 3D-Punktwolkensegmentierung (XYZ-RGB-T)



Zhang Y, Zhang C, Rosenberger M, Notni G. *6D Object Pose Estimation Algorithm Using Preprocessing of Segmentation and Keypoint Extraction*. in: *IEEE I2MTC Conference, Dubrovnik, 2020*

Zhang Y, Zhang C, Nestler R, Rosenberger M, Notni G. *Efficient 3D object tracking approach based on convolutional neural network and Monte Carlo algorithms used for a pick and place robot*. In: *SPIE IMEKO Conference, Jena, 2019*.

3D Hand- /Objektsegmentierung (2)

- **Versuch 1: Evaluation der Segmentierung mit multimodaler Punktwolke (XYZ-RGB-T)**
 - PointNet, PointNet++ und RandLANet Segmentierung untersucht
 - Ergebnis mit stationärem Sensorsystem (Datensatz 1: 10-Objekte-Datensatz)
 - RandLANet-Segmentierung erreicht höchstes IoU
 - Für XYZ-T ist die Segmentierung von unterschiedlichen Objekten schwierig
 - XYZ-RGB-T hat eine deutliche bessere Segmentierung im Vergleich zu allen anderen, insbesondere im Interaktionsbereich zwischen Hand und Objekt IoU Hand.

IoU:
Intersection
over Union

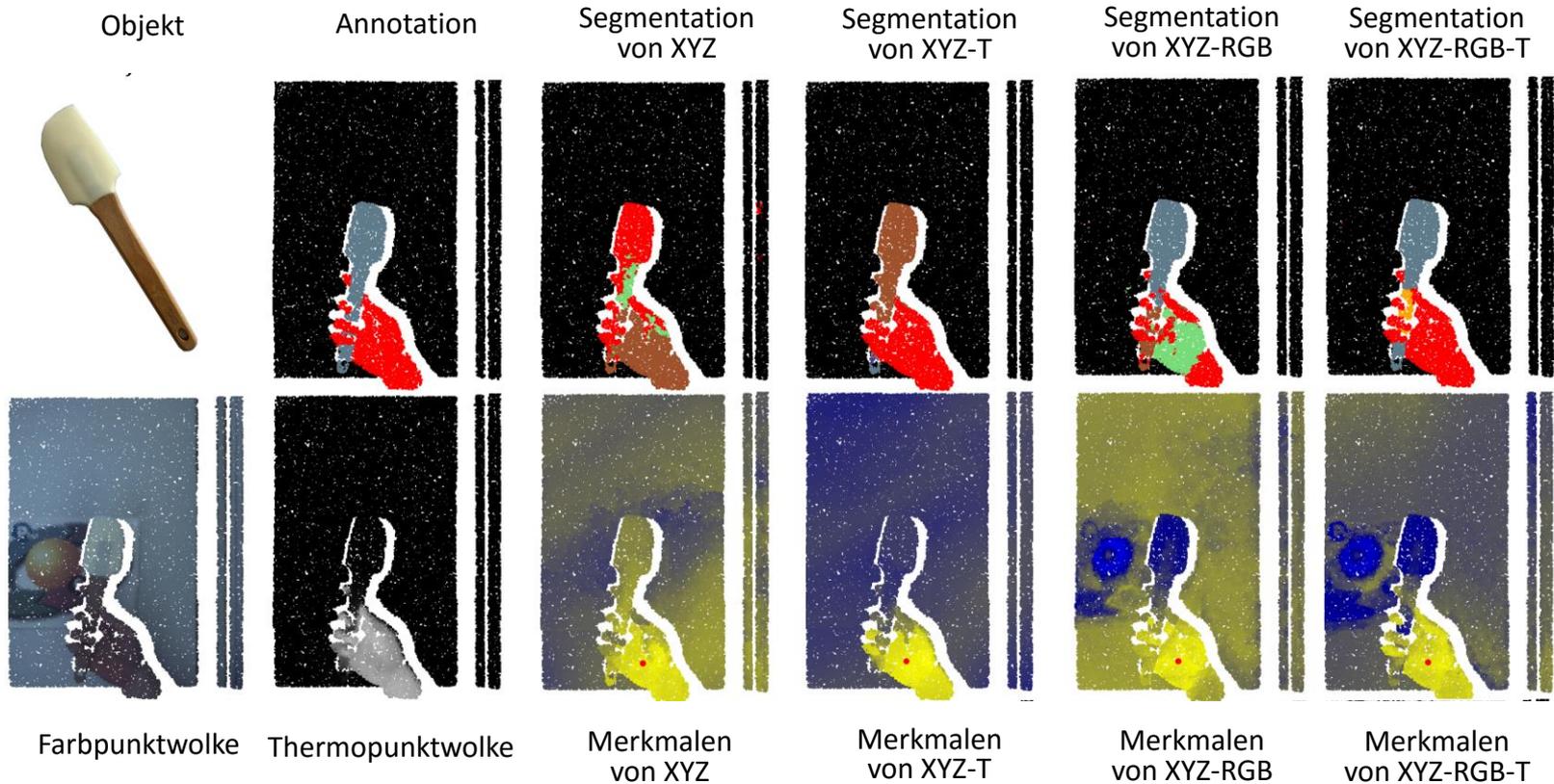
mIoU:
mean
Intersection
over Union

		mIoU insgesamt	IoU Hand	mIoU Objekte
RandLANet	XYZ-RGB-T	82,8 %	92,6 %	81,9 %
	XYZ-RGB	77,3 %	88,9 %	76,2 %
	XYZ-T	35,7 %	84,1 %	30,9 %
	XYZ	35,7 %	76,7 %	31,6 %
PointNet++	XYZ-RGB-T	55,0 %	79,8 %	52,5 %
PointNet	XYZ-RGB-T	45,9 %	79,9 %	42,5 %

Zhang, Y., Müller, S., Stephan, B., Gross, H. M., & Notni, G. *Point Cloud Hand-Object Segmentation Using Multimodal Imaging with Thermal and Color Data for Safe Robotic Object Handover*; in: *Sensors*, 2021, 21(16): 5676.

3D Hand- /Objektsegmentierung (3)

- Visualisierung der RandLANet-Segmentierung (10-Objekte-Datensatz)



Similar

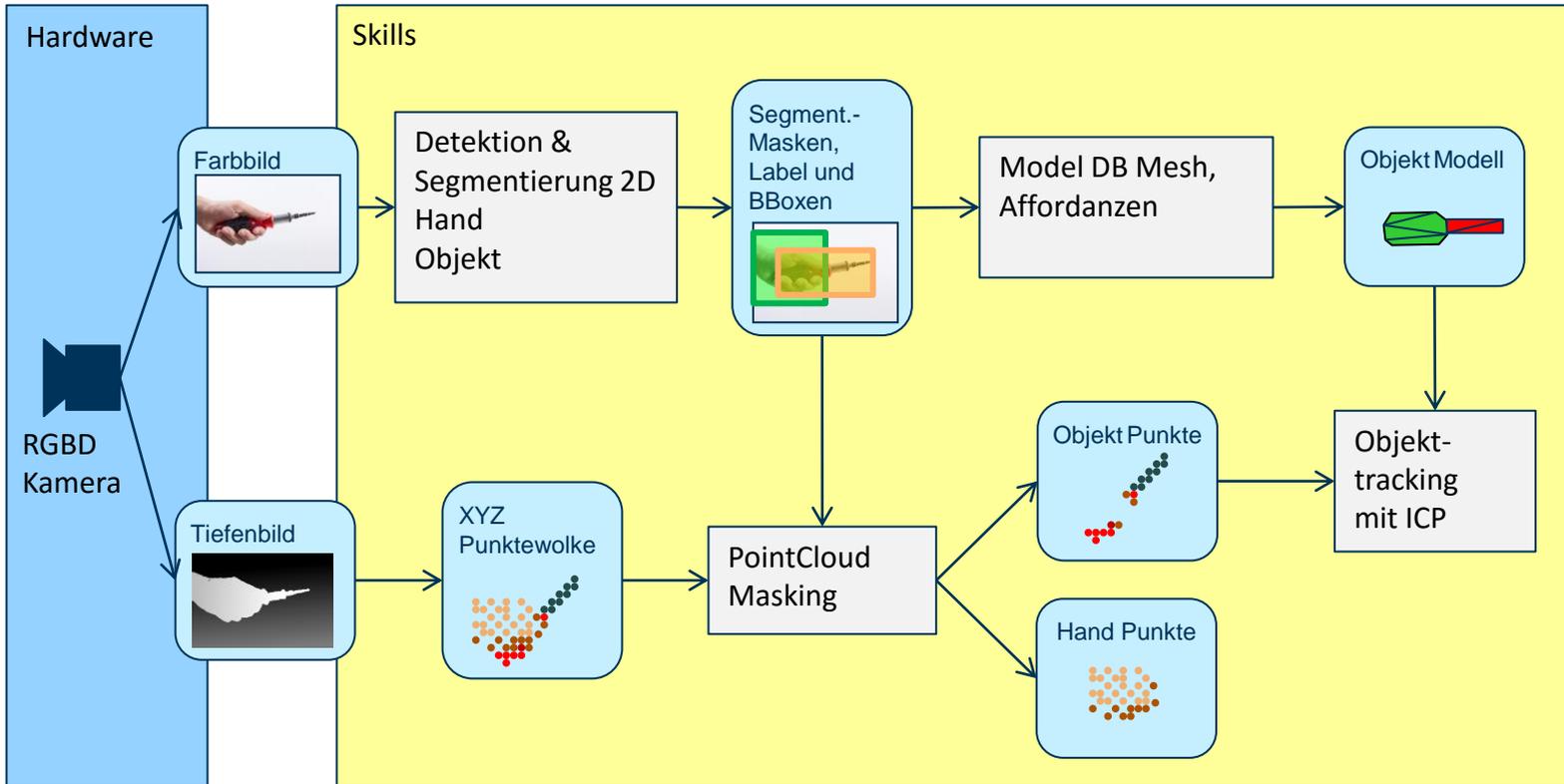
Not similar

3D Hand- /Objektsegmentierung (4)

- **Versuch 2: Hand- & Objektsegmentierung für unbekannte Objekte**
 - RandLANet-Segmentierung
 - Trainiert mit Datensatz 2 (3-Klassen-Datensatz)
 - Vergleich zum Versuch 1:
 - Segmentierung für nur drei Klassen (Hindernis, Hand, alle Objekte an der Hand)
 - Für neues Objekt ist neues Training **nicht** notwendig.
 - Segmentierung für **bekante** und **unbekante** Objekte
 - **Ergebnis:**
 - Saubere und stabile Vordergrundsegmentierung (Hand/Objekt)
 - Mittels Vorausschnitt der Punktwolke wurden für die Segmentierung 11 FPS auf Nvidia-Jetson erreicht
 - Es gibt manchmal False-Positive-Punkte von Objekte im Hintergrund

2D Hand- /Objektsegmentierung

- 2D-Bildsegmentierung & Abbilden in 3D-Punktwolke

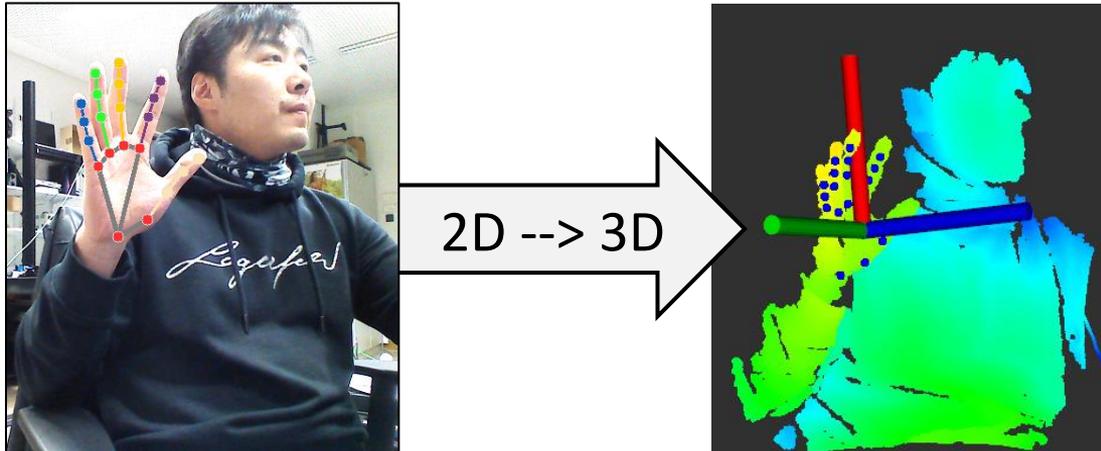


2D Hand- /Objektsegmentierung

- **Versuch 3: 2D-Bildsegmentierung & Abbilden in 3D-Punktwolke**
 - YolactEdge-Segmentierung (Farbbild)
 - Trainiert mit Datensatz 2 (3-Klassen-Datensatz)
 - Workflow:
 1. Instanzsegmentierung im Farbbild
 2. 2D-Maske generieren
 3. Punktwolkensegmentierung mittels 2D-Maske
 - **Ergebnis:**
 - Stabile Hand- /Objektsegmentierung
 - Wegen Instanzsegmentierung gibt es keine False-Positive-Punkte im Hintergrund
 - Die Segmentierung erreicht 3 FPS auf Nvidia-Jetson
 - Das Modell wird im Demo 2 verwendet.

• **Erwartete Bilder von Benedict!**

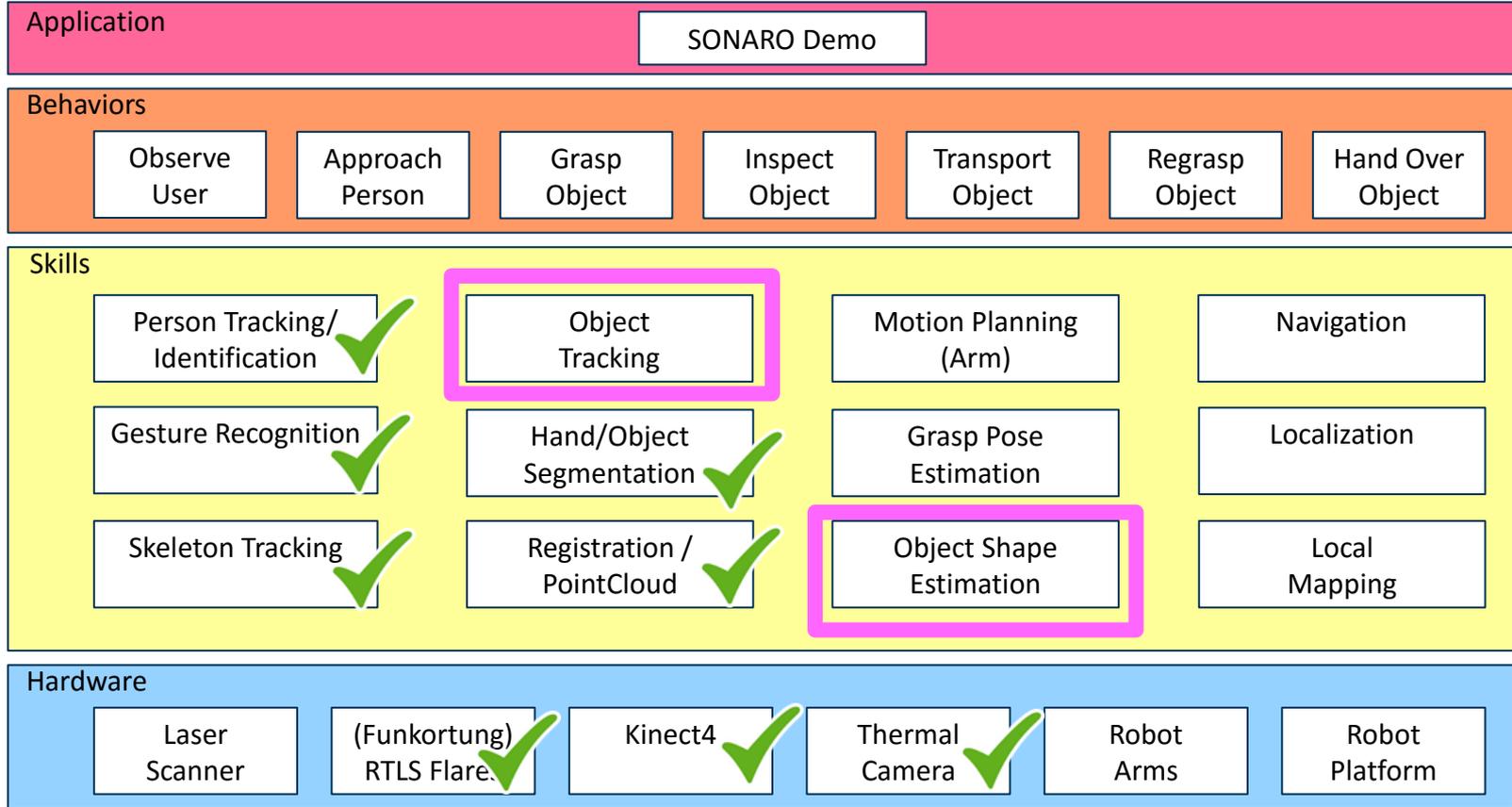
- Handkeypoints Detektion in 2D-Farbbild (Mediapipe Bibliothek)
- Konvertierung: von 2D-Keypoints zu 3D-Keypoints mittels Tiefenbild
- 3D-Handposeerkennung als eine Zielpose zur Objektübergabe des Roboters
- Nutzung der 3D-Handposeerkennung zur Vorauswahl der Punktwolke für eine effizientere Hand- /Objektsegmentierung



Konnten die Erwartungen erfüllt werden?

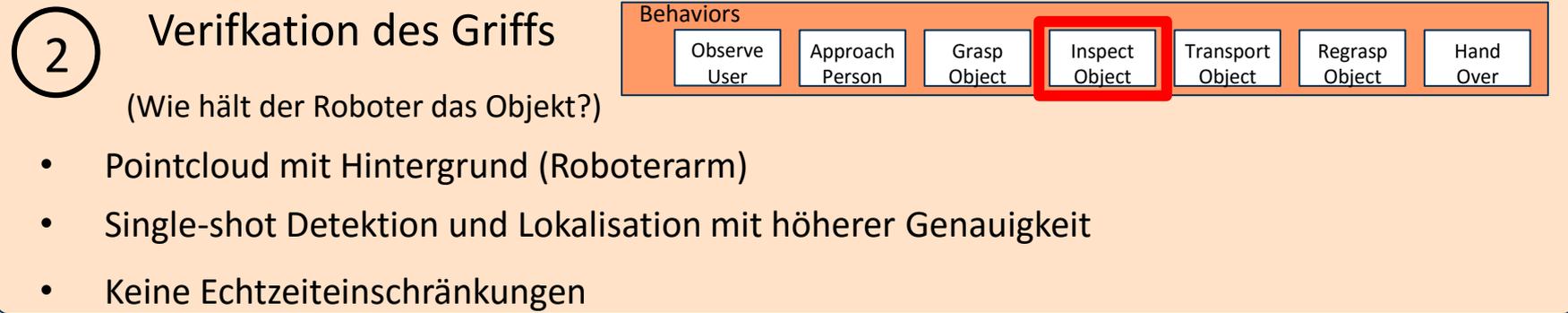
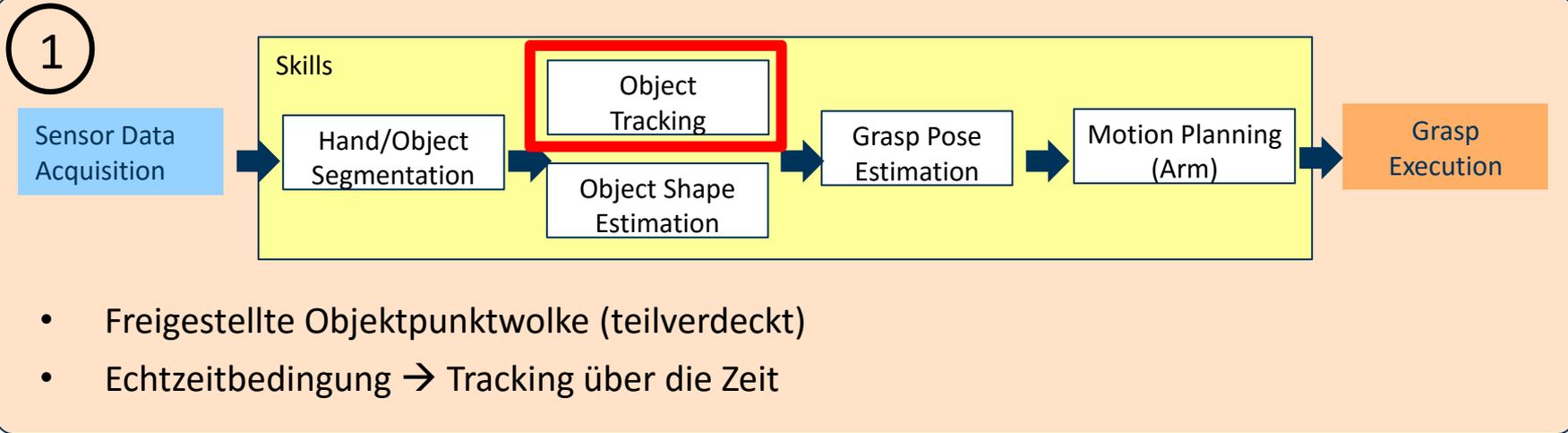
Aspekt	Erwartungen zu Projektbeginn	Erreichter Stand
Multimodale 3D-Sensorik	Sensorsystem zur Erfassung multimodaler Punktwolken sollte entwickelt werden	Kameras mit optimalem Spektrum (Thermo) wurde ausgewählt Kalibrierung und Registrierung wurde realisiert
Multimodale Tiefensensorik	Sensorsystem zur Erfassung multimodaler Punktwolken sollte entwickelt werden	Kameras mit optimalem Spektrum (thermo) wurde ausgewählt, Kalibrierung und Registrierung wurde realisiert und multimodale Punktwolken wurden in die Softwareframeworks integriert
Segmentierung Hand/Objekt	Robuste, echtzeitfähige Segmentierung auf Basis der multimodalen 3D Daten sollte entwickelt werden	Sowohl 2D, als auch 3D Segmentierungsnetzwerke wurde auf eigenen Daten evaluiert und integriert. Ergebnisse zeigen Überlegenheit von XYZ RGB+T gegenüber XYZ RGB

Objektmodellierung / -lageschätzung (GFE)

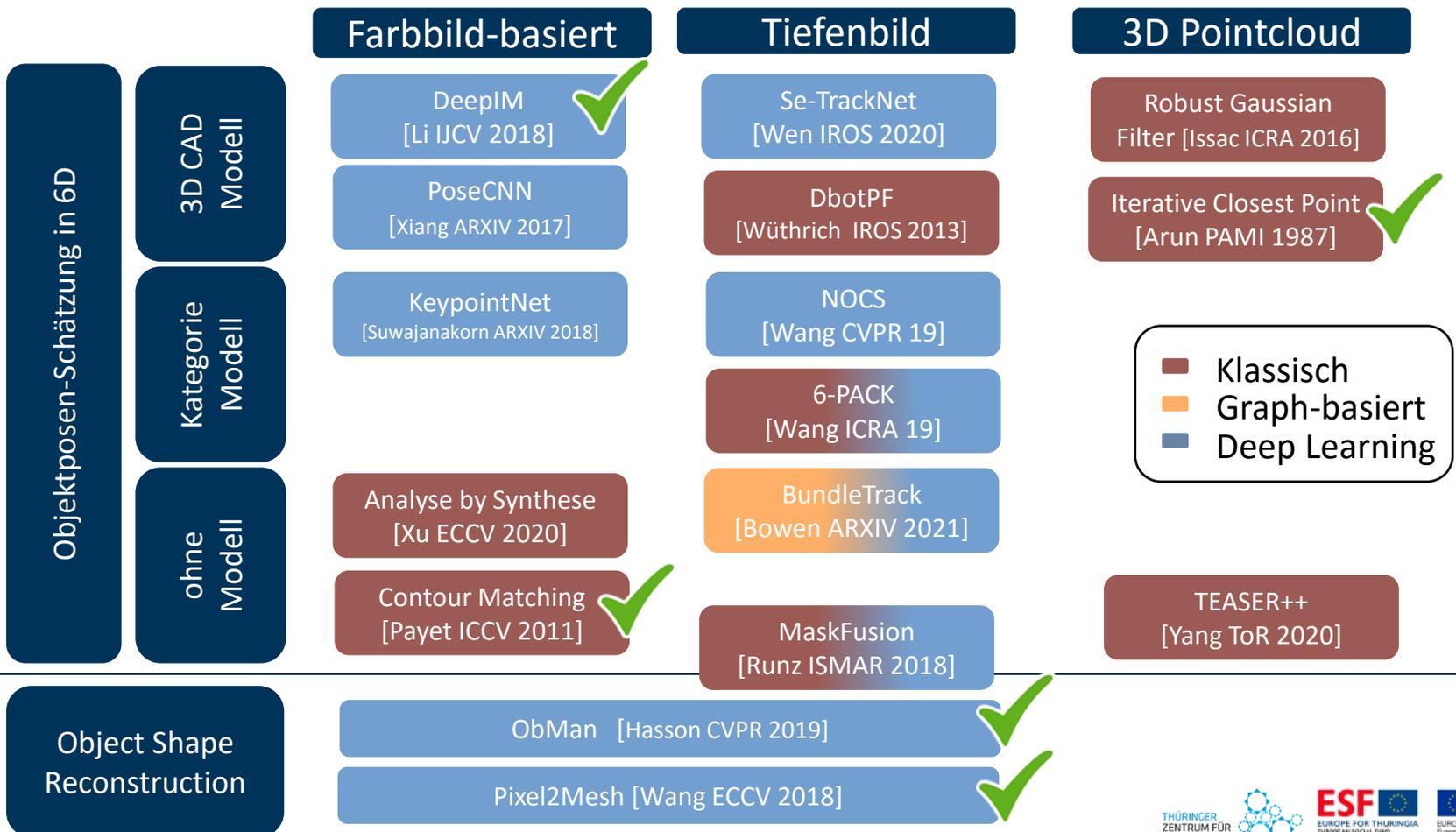


Objektposenschätzung in der Greifpipeline

Zwei Stellen im Ablauf mit unterschiedlichen Anforderungen:



State of the Art Objektposenschätzung



Legend:

- Klassisch
- Graph-basiert
- Deep Learning

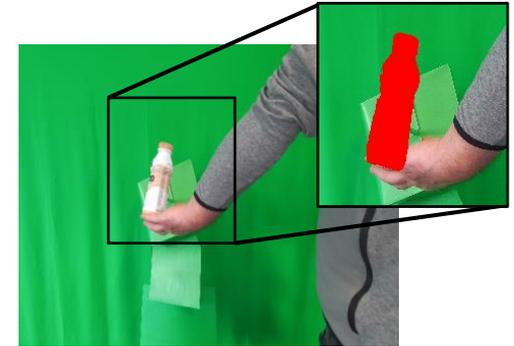
Untersuchte Verfahren zur modellbasierten Objekt-detektion (1)



Es wurden auf Basis der industriellen Bildverarbeitungs-bibliothek HALCON mehrere Ansätze entwickelt und evaluiert:

1. Form- und farbbasierte Matching-Verfahren auf Basis eines 2D-Referenzbildes und des aktuellen RGB-Bildes der Roboterumgebung
 - Nicht robust genug bei anderen Posen und starken Verdeckungen des Objektes

2. Key-Point-basiertes 3D-Matching auf Basis der 3D Modelle des SONARO HandObject Datensatzes
 - Konvergenz sowie Rechenzeit stark abhängig von Struktur der aktuellen Bildszene
 - Nicht vertretbare Rechenzeiten



Untersuchte Verfahren zur modellbasierten Objekt-detektion (2)

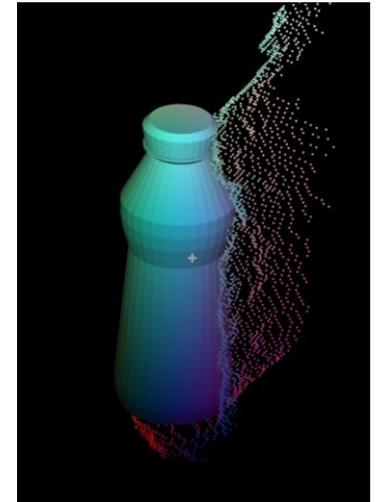


Es wurden auf Basis der industriellen Bildverarbeitungs-bibliothek HALCON mehrere Ansätze entwickelt und evaluiert:

3. Point Cloud Registration
3D Feature-basierte Initialschätzung +
Verfeinerung der Pose mittels Iterative Closest Point (ICP)
 - **Nutzung im Demonstrator für
Verifikation des gegriffenen Objekts** 2
 - **Verfahren braucht sehr lange abhängig von der
Punktwolkengröße > 1 sec**

4. Nutzung der Bildverarbeitungsbibliothek **PCL** → **Iterative Closest Point Algorithmus** zur Registrierung der bereits segmentierten Objektteilpunkt看ke gegen eine aus dem 3D Meshmodell gesampelten Punkt看ke

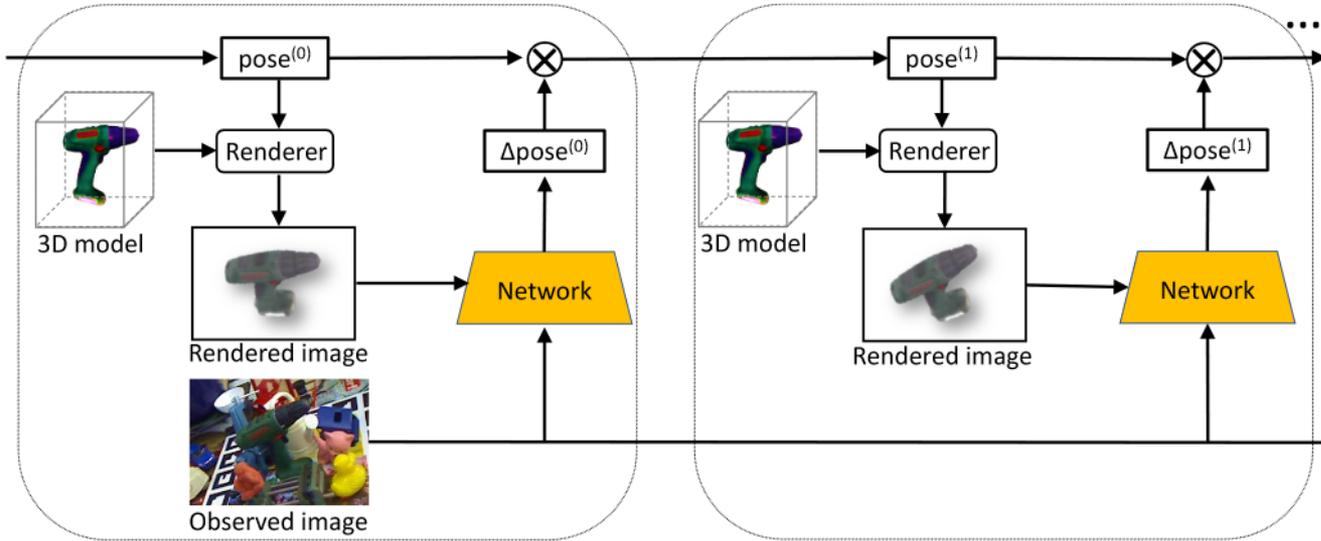
- Iterative Closest Point braucht eine einigermaßen passende Startpose
 - mehrfaches zufälliges Sampeln der Initialpose
 - Tracking der Pose über die Zeit (Kalman Filter)
- **Nutzung des ICP im Demonstrator für Livetracking der zu greifenden Objekte** 1
- **Problem:** Bei kleinen Punkt看kenausschnitten sind Posen nicht eindeutig → springendes Tracking
- **Fazit:** Verfahren sollten auch die reichhaltigeren Farbbildinformationen nutzen



Angepasstes 3D Modell in Punkt看ke von Kinect Azure (mit Artefakten)

Untersuchte Verfahren zur Objekt-Lageschätzung

Li, Yi, et al. *DeepIM: Deep iterative matching for 6d pose estimation*. Proceedings of the European Conference on Computer Vision (ECCV). 2018.

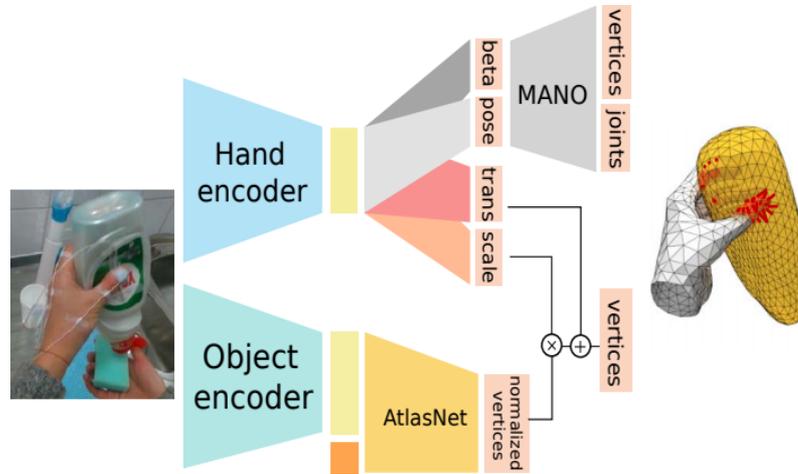


- Rendering eines Objektmodells und Vergleich mit Inputbild
- Neural Network prädiziert Posenänderung zur Minimierung des Bildfehlers
- Iterative Anpassung ausgehend von Startpose
- **Problem:** benötigt texturierte 3D Modelle und gute Startposition

Object Shape Reconstruction (1)

Objektposenschätzung und Modellauswahl wäre überflüssig, wenn Objektform im aktuellen Kamerabild geschätzt werden könnte.

Vielversprechende Deep Learning Verfahren in der aktuellen Literatur,



Hasson, Yana et al.: *Learning joint reconstruction of hands and manipulated objects*, CVPR, 2019

- Farbbild – basiert
- Für Video Activity Recognition gedacht
- **Gute Ergebnisse nur auf bekannten Datensätzen**

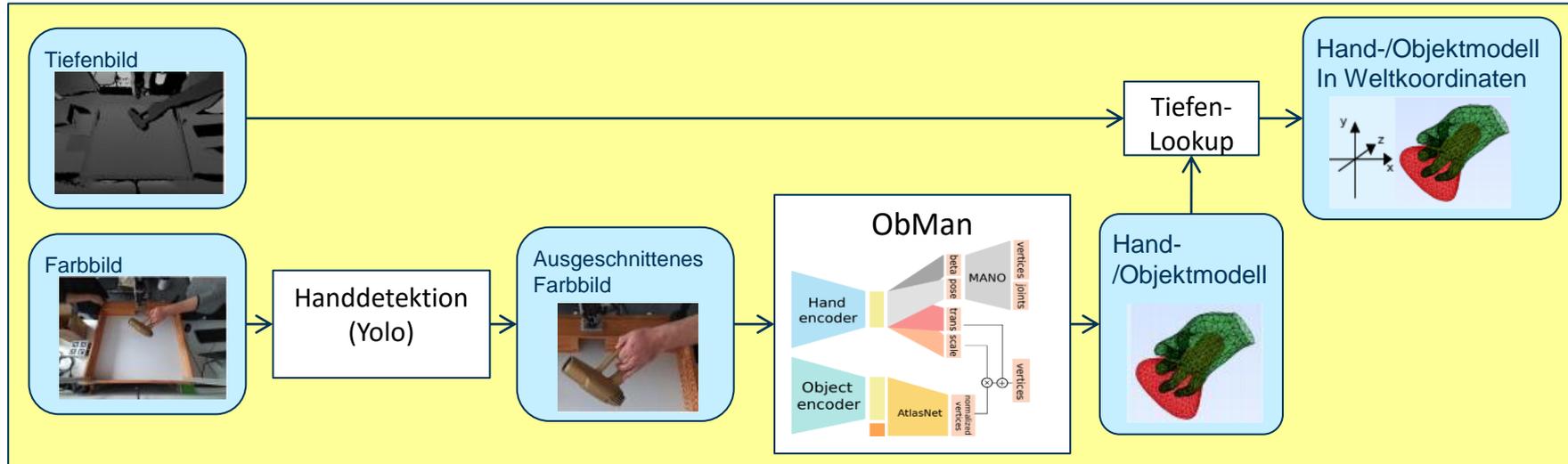
Object Shape Reconstruction (2)

MA Patrick Stäblein: *Integration eines Objekt-/Handtracking-Ansatzes in die reale Roboterapplikation, TU-Ilmenau 2021*

Masterarbeit Patrick Stäblein (zusammen mit TUI NI&KR)

→ Integration und Evaluation des **ObMan** Verfahrens in eine Pipeline für den Test auf dem Roboter

- **Hauptproblem:** Objektzentrierte Koordinaten → 3D Weltkoordinaten



- **Fazit: Formschätzung mit Fehlern >10cm viel zu ungenau!**

Object Shape Reconstruction (3)

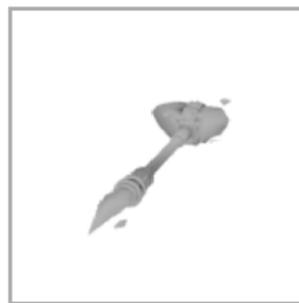
MA Markus Paschke: *Deep-Learning-Methoden zur Erstellung von 3D-Modellen*, TU-Ilmenau, 2020

Masterarbeit Markus Paschke (TUI NI&KR)

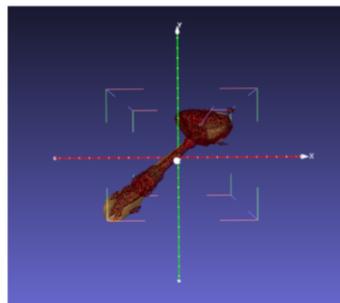
- Recherche und Vergleich mehrerer 3D Rekonstruktionsverfahren

Octree Generating Network, Pixel2Mesh, Point Set Generation Network

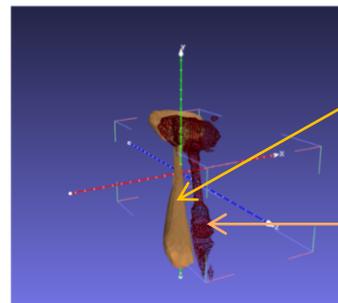
- Genauere Untersuchung von Pixel2Mesh → Neutraining auch auf Tiefenbildern



(a) Eingabebild



(b) Blick aus Richtung Z-Achse



(c) Andere Perspektive

rekonstruiertes
3D Meshmodell

Ground Truth
Punktwolke

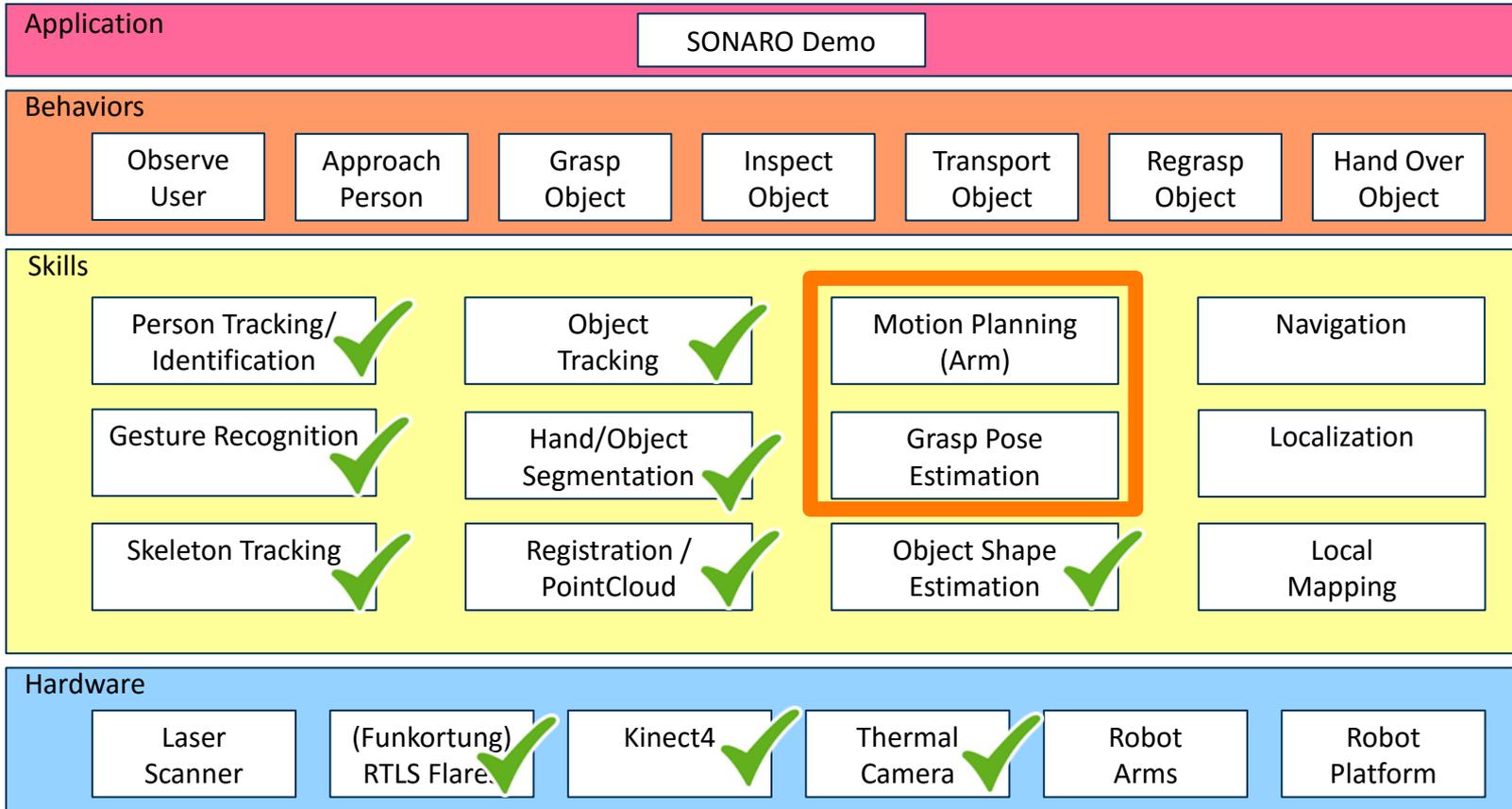
• Probleme:

- Objektform wird in Objektkoordinaten geschätzt -> keine Weltanbindung
- Freigestellte Objektbilder benötigt
- Formschätzung nur in Projektionsrichtung gut; **allgemein zu ungenau!**

Konnten die Erwartungen erfüllt werden?

Aspekt	Erwartungen zu Projektbeginn	Erreichter Stand
Objekttracking	Bekannte Objekte sollen im Livedatenstrom detektiert und ihre Position getrackt werden	Bekannte Objektmodell können mit der Punktwolke registriert werden. Trackingqualität hängt stark vom Verdeckungsgrad der Objekte ab und kann teilweise sehr stark springen
Objektmodellschätzung	Für unbekannte Objekte soll ein 3D Modell aus den vom Roboter aus aufgenommenen Sensordaten erstellt werden	Erprobte State of the Art Verfahren liefern keine hinreichend genauen Modell, um ein Greifen zu ermöglichen (weiterer FuE-Bedarf!)

Greifposenschätzung

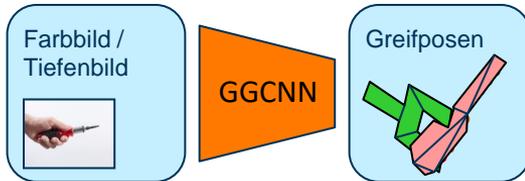


State of the Art Greifposenschätzung



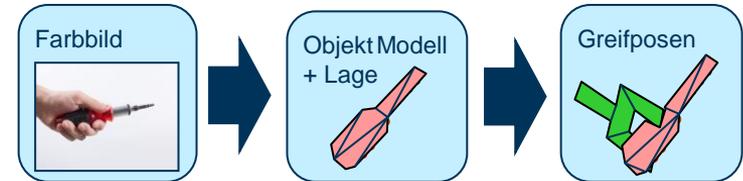
Ausgangspunkt:

Modellfreie Verfahren



- Direkte Abbildung von Bild/Tiefendaten oder Punktwolken auf die Greiferposition
- Keine Möglichkeit Zusatzwissen zum Objekt und dessen Eigenschaften einzubringen

Modell-basierte Verfahren



- Simulationsbasierte analytische Bewertung von Greifposen am 3D Modell
- Kontrollierbare Zwischenergebnisse

D. Morrison, P. Corke, and J. Leitner, *Closing the loop for robotic grasping: A real-time, generative grasp synthesis approach*, CoRR, vol. abs/1804.05172, 2018. [Online]. Available: <http://arxiv.org/abs/1804.05172>

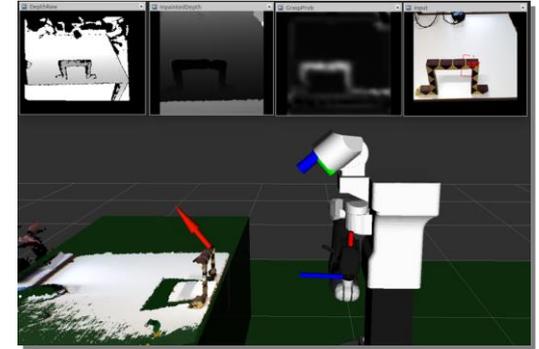
Greifposenschätzung (1)

Arbeiten zur **modellfreien** Greifposenschätzung in SONARO:

Untersuchung der Nutzbarkeit von GGCNN

D. Morrison, P. Corke, and J. Leitner, *Closing the loop for robotic grasping: A real-time, generative grasp synthesis approach*, CoRR, vol. abs/1804.05172, 2018. [Online]. Available: <http://arxiv.org/abs/1804.05172>

Leider nur bedingt zu gebrauchen da für Parallelgreifer in Top-Down Anordnung konzipiert

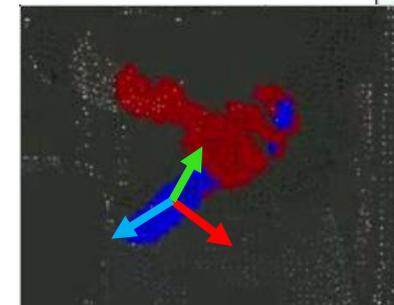


Heuristische Samplingmethode basierend auf segmentierten Punktwolken

Multikriterielle Bewertung:

- Handabstand
- Abstand zur aktuellen Roboterposition
- Abstand antipodaler Kontaktpunkte

TUI QBV



Greifposenschätzung (2)

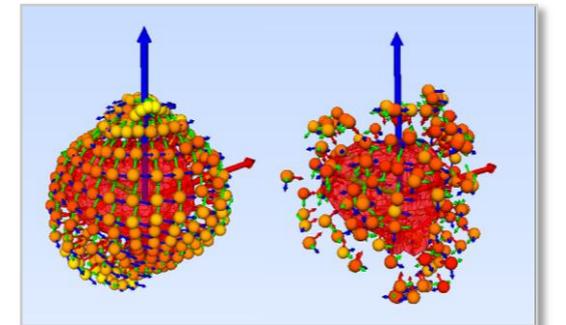
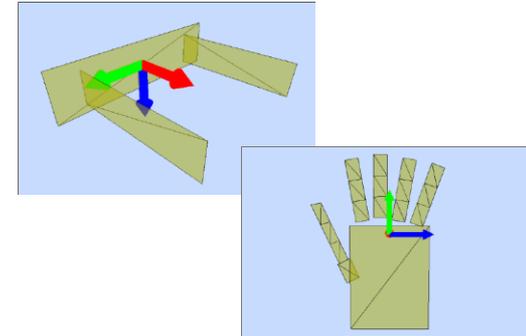
Eigene Arbeiten zur **modellbasierten** Greifposenschätzung:

Masterarbeit D. Rink *Greifposenbestimmung für mobile robotische Manipulatoren.* TU Ilmenau, 2019.

- Umsetzung einer analytischen Greifposenbestimmung mittels **Particle Swarm Optimization** für verschiedene Greifergeometrien
- Onlinefähig (optimiert kontinuierlich eine Menge von möglichen Greifposen relativ zum Objektmodell)

Bachelorarbeit Wei Dai *Approximation der Greifgütefunktion mit neuronalen Netzen.* TU Ilmenau, 2020

- Versuch der Beschleunigung einer Vorhersage von dichten Greifposenwolken mittels **Convolutional Neural Networks**
- Training mit den analytisch bestimmten Posen



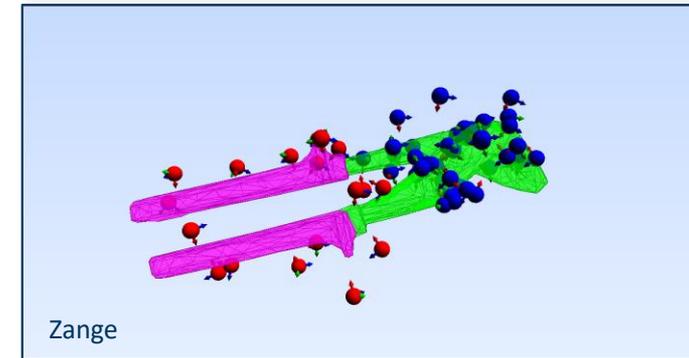
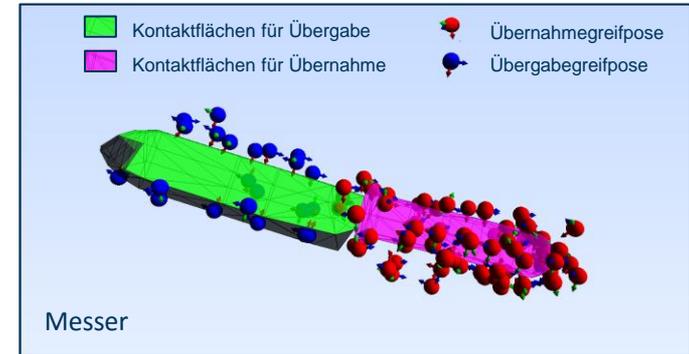
Netzoutput

analytisch bestimmte Trainingsdaten

Greifposenschätzung (3)

Eigene Arbeiten zur Greifposenschätzung in SONARO:

- **Affordanzen** der Objektteile für unterschiedliche Aufgaben
 - **Give, Take, Use**
- Integration der Affordanzlabel in die Analytische Greifposenschätzung liefert verschiedene Wolken an Greifposen
- Komplexere Greifer ermöglichen **unterschiedliche Griffarten**
 - **sphärischer Griff, Zylindergriff, Pinzettengriff**
- Analytischer Greifposenplaner berücksichtigt Griffarten und Label am Objekt



Greifposenschätzung (3)

In der Anwendung:

Verschiedene Greifposen mit unterschiedlichen Greifermodellen werden benötigt

1. Übernahme von Objekten:

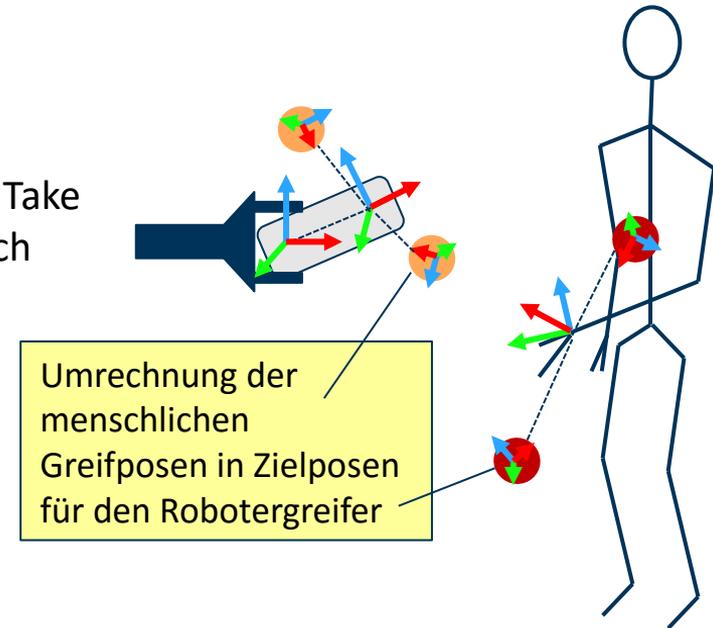
- Greifermodell des Roboters, Affordanz: Give

2. Umgreifen:

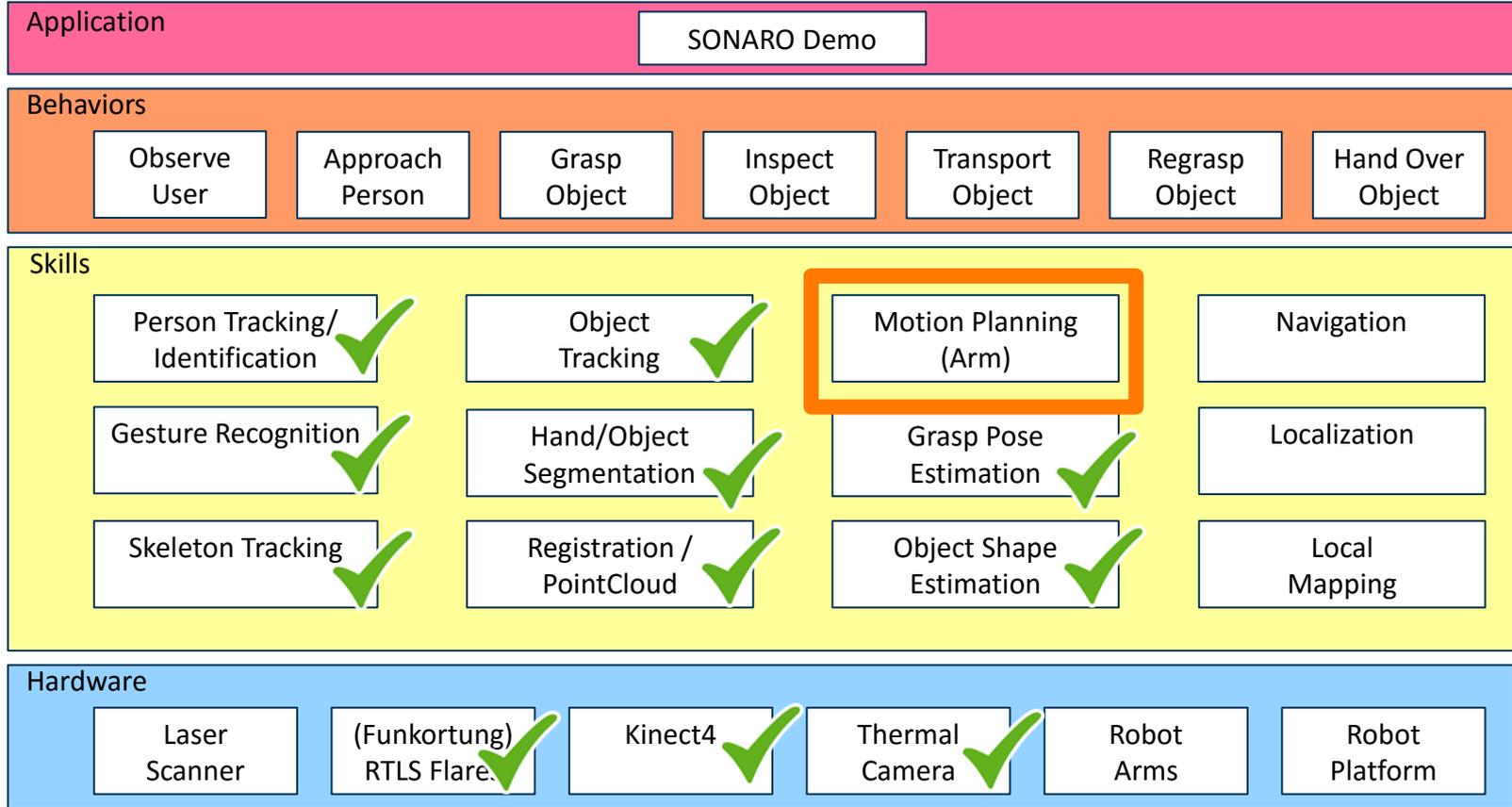
- Handmodell Mensch (links, rechts), Affordanz: Take (für die Entscheidung, ob diese Positionen durch Robotergreifer belegt sind)
- Greifermodell des Roboters (links, rechts), Affordanz: Give

3. Übergabe:

- Handmodell Mensch (links, rechts), Affordanz: Take



Dynamische Bewegungssteuerung für die Roboterarme



Bewegungssteuerung für die Roboterarme (1)

Ausgangspunkt



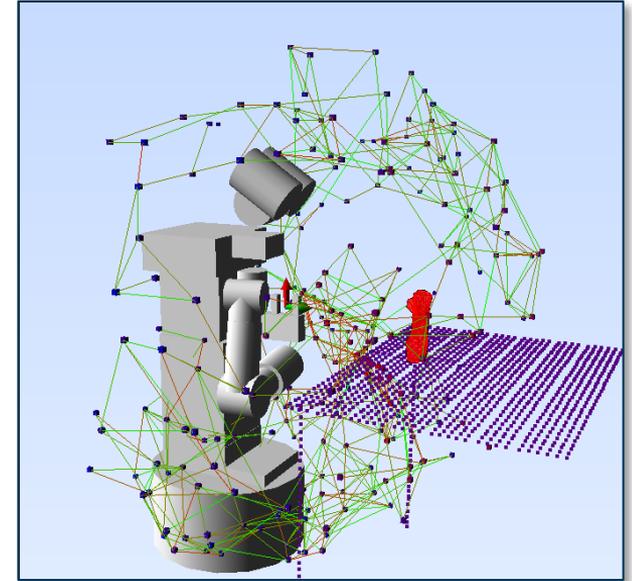
- Quasistandard ROS MoveIt
- Nur statische Trajektorienplanung und anschließend Ausführung
- Bei dynamischen Szenen Unterbrechung mit Neuplanung

→ Eigenentwicklung einer dynamischen Bewegungsteuerung erforderlich mit Integration in das genutzte Softwareframework MIRA

Bewegungssteuerung für die Roboterarme (3)

Entwicklung im Rahmen von SONARO:

- MDP (Markov Decision Process) basierter Planer
- **Kontinuierlich aktualisierter Roadmap Graph ermöglicht die Berücksichtigung von dynamischen Hindernissituationen und Zielen!**
- MDP ermöglicht Definition von nicht über einen Zielpunkt definierten Tasks (**keine inverse Kinematik notwendig**)
 - Z.B.: „Minimiere den Grundriss mit gegriffenem Objekt“,
 - oder „Erreiche eine von n möglichen Greifposen“
- **Publikation zur ECMR 2021:**



Müller, St., Stephan, B., Gross, H.-M.

MDP-based Motion Planning for Grasping in Dynamic Szenarios.

in: Europ. Conf. on Mobile Robotics (ECMR), Bonn, 2021.

Bewegungssteuerung für die Roboterarme (3)

Konkrete Beiträge des Papers:

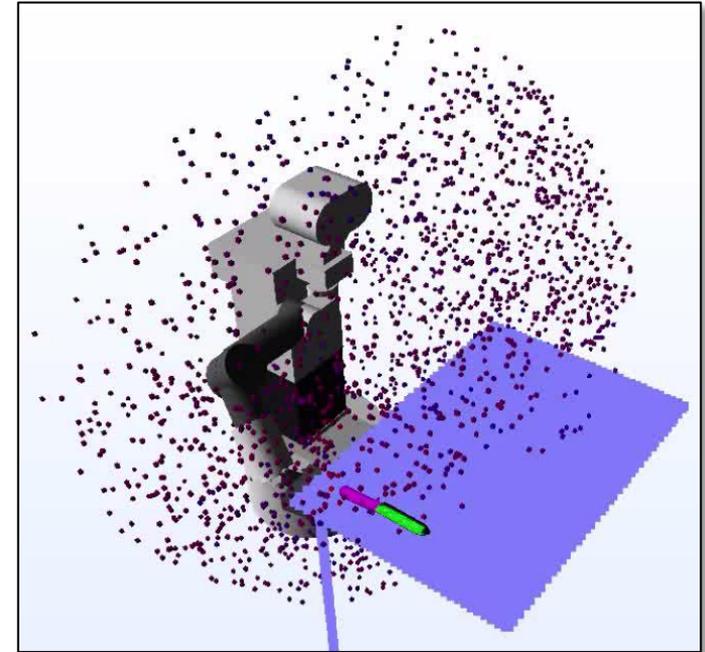
Müller, St., Stephan, B., Gross, H.-M. *MDP-based Motion Planning for Grasping in Dynamic Szenarios. In Proc. of ECMR, 2021.*

- MDP Formulierung des Ziels anstatt eine feste Zielposition (Dijkstra)
- Strategie für die Erstellung des Roadmap Graphs → weniger Kanten → effizienter

Ergebnisse:

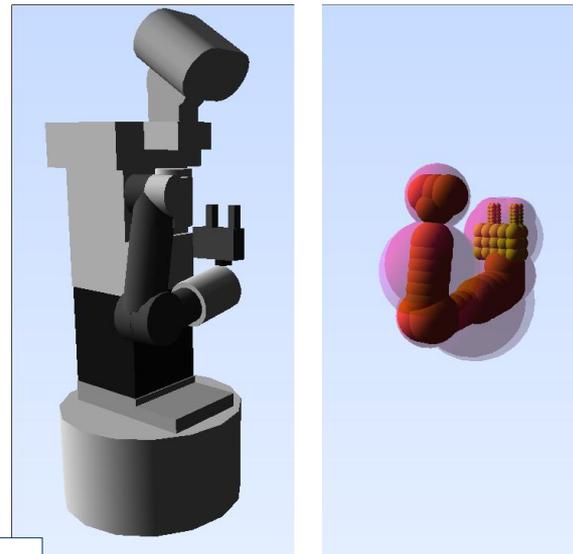
- Versuch:
 - Roboter greift aus Ruhelage nach bewegtem Objekt
 - Erfolg, wenn innerhalb 30sec das Objekt in der Bewegung gegriffen wurde

	success rate	Average duration
Dijkstra	59%	15.1 sec
MDP	78%	13.8 sec



Kollisionsberechnungen für die Bewegungssteuerung

- Effiziente Kollisionsberechnung ist Grundlage für die Bewegungsplanung
- Approximation der beweglichen Roboterteile mittels Kugelhierarchie
- Darstellung der Hinderniskonfiguration über eine Distanztransformierte Voxelkarte
- Abstandsberechnung durch Lookup der Kugelmittelpunkte

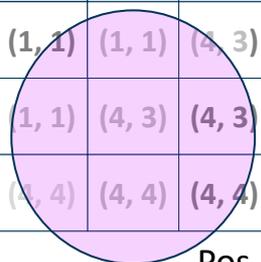


(∞, ∞)	(∞, ∞)	(∞, ∞)	(∞, ∞)	(4, 0)
(∞, ∞)	(1, 1)	(∞, ∞)	(∞, ∞)	(∞, ∞)
(∞, ∞)				
(∞, ∞)	(∞, ∞)	(∞, ∞)	(∞, ∞)	(4, 3)
(∞, ∞)	(∞, ∞)	(∞, ∞)	(∞, ∞)	(4, 4)



Distanz-
trans-
formation

(1, 1)	(1, 1)	(1, 1)	(4, 0)	(4, 0)
(1, 1)	(1, 1)	(1, 1)	(4, 0)	(4, 0)
(1,1)	(1, 1)	(1, 1)	(4, 3)	(4, 3)
(1, 1)	(1, 1)	(4, 3)	(4, 3)	(4, 3)
(1, 1)	(4, 4)	(4, 4)	(4, 4)	(4, 4)



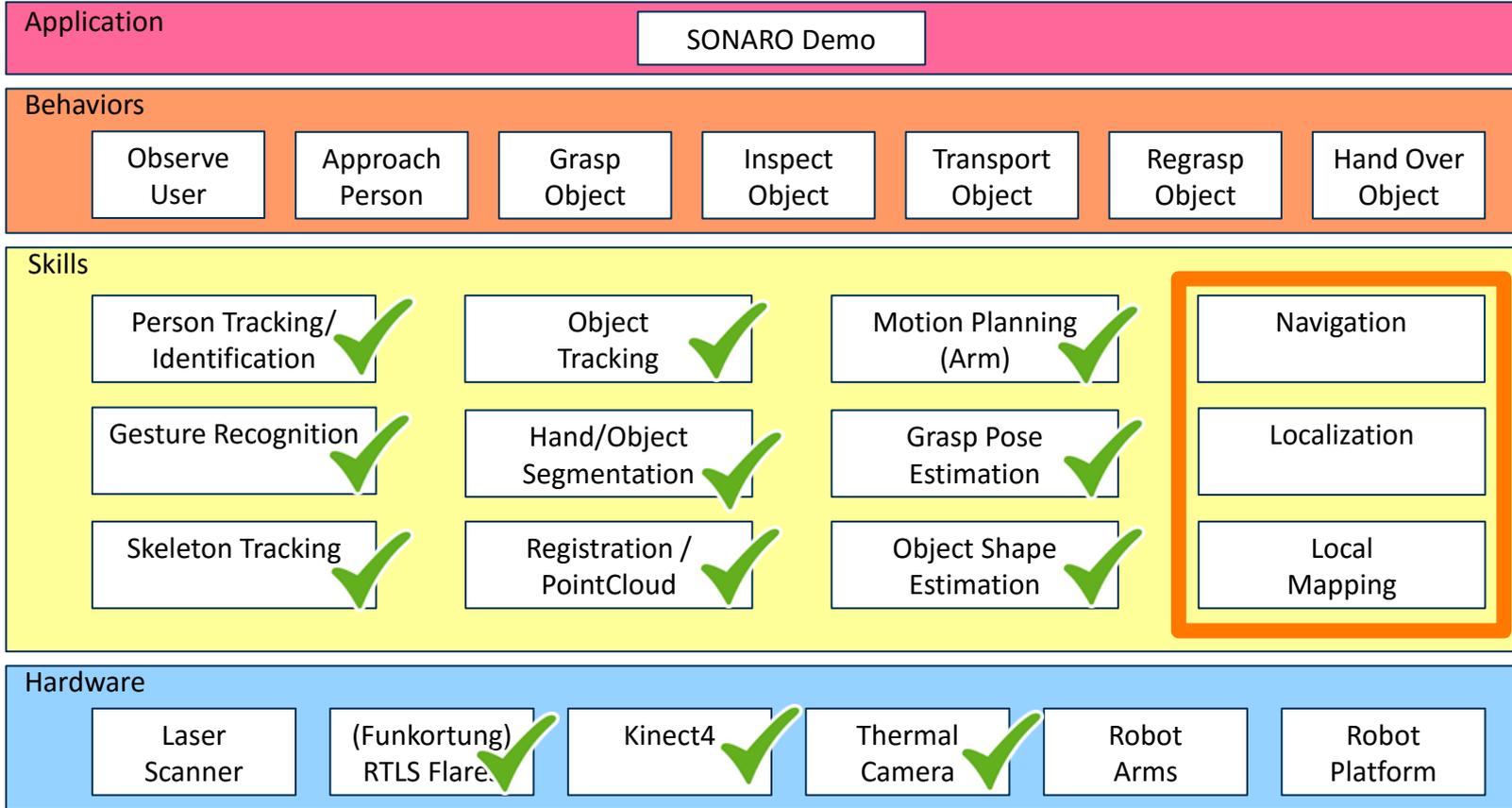
Pos.:(2,3) Radius r=1.5

Abstand:
 $a = |(2,3) - (4,3)| = 2$
 $a = 2 < r = 1.5$
 \rightarrow keine Kollision

Konnten die Erwartungen erfüllt werden?

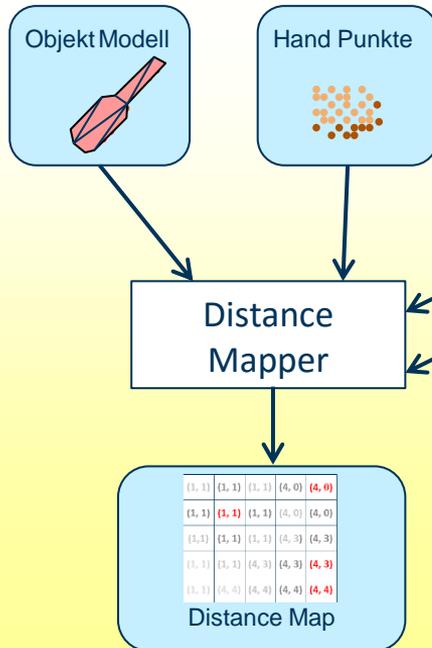
Aspekt	Erwartungen zu Projektbeginn	Erreichter Stand
Bewegungssteuerung des Roboterarmes	Bewegliche Ziele sollten online angefahren werden können. Menschliche Körperteile im Aktionsbereich stellen besonders kritische zu vermeidende Hindernisse dar, auf die dynamisch reagiert werden muss.	Der neu entwickelte Bewegungsplaner kann auf dynamische Hindernisse und Zieländerungen während der Bewegung reagieren. Dabei wird nicht nur lokal ausgewichen, sondern kontinuierlich global geplant.
Greifposenbestimmung	Für beliebige 3D Geometrien sollen für die verschiedenen Greiferarten sinnvolle Angriffspunkte bestimmt werden.	Verfahren zur analytischen Greifposenbestimmung wurde implementiert und berücksichtigt auch Affordanzlabel für Objektteile. Modellfreie Verfahren wurden erprobt.

Roboternavigation

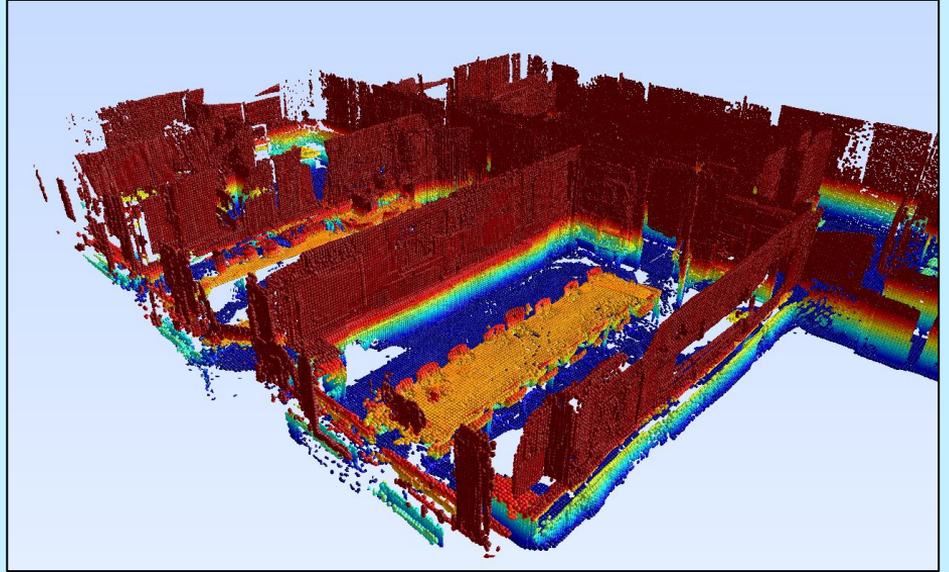


Umgebungsmodellierung für Bewegungssteuerung

- **Normal Distribution Transformed Map** (NDT-Map) integriert Messungen der Tiefenkameras → persistent auch dort wo aktuell nicht hingeschaut wird
- Statische + dynamische NDT-Map



Statische NDT Map der Laborumgebung



dynamische NDT Map

Lokalisation

State of the Art:

- SLAM (Simultaneous Localization and Mapping)
- Kamerabasiert (RTAB-Map)

Ausgangspunkt:

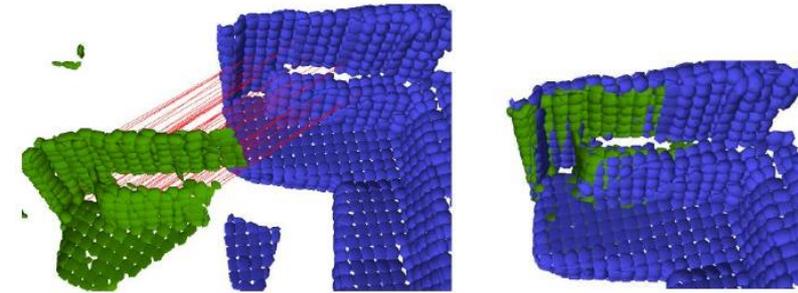
- Laser-basierte Monte Carlo Lokalisation

Entwicklung während SONARO:

- RTAB-Map wurde am Fachgebiet in anderen Projekten eingesetzt und erprobt
 → **keine bessere Genauigkeit, großer Ressourcenbedarf**
- Tiago Roboter hat viel Schlupf an den Rädern
 → **schlechte Odometrie**
- Nutzung der NDT Monte Carlo Localization (Vergleich der lokalen mit der globalen NDT Karte)

Methode	Zellgröße Karte	Zielstreuung
Laser AMCL	5 cm	22 mm
NDT AMCL	5 cm	43 mm
NDT AMCL	10 cm	44 mm
RTAB-Map +Laser	5 cm	51 mm
RTAB-Map (nur Bild)	5 cm	65 mm

Erreichte Genauigkeiten Bei Zielfahrten



Beispiel NDT Map Matching

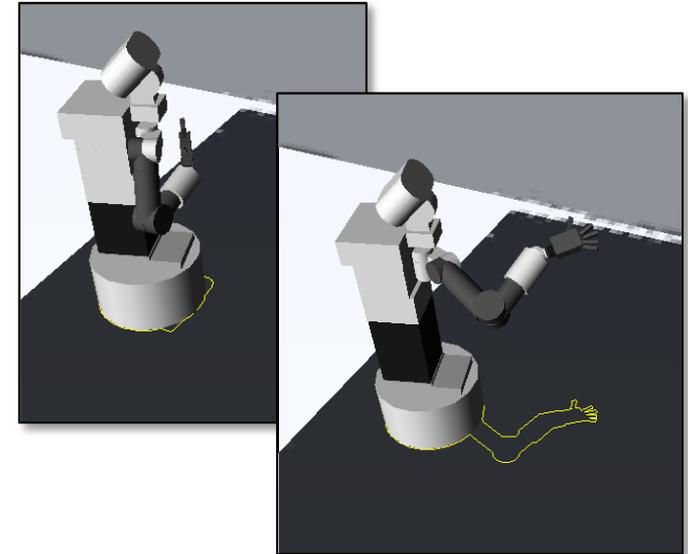
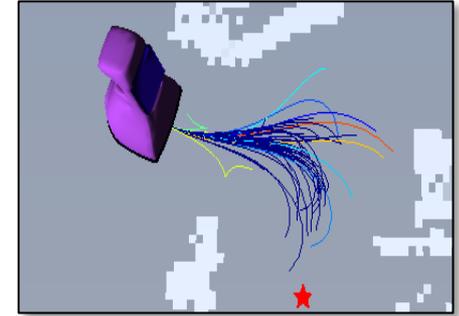
Roboternavigation

Ausgangspunkt:

- Evolutionary Motion Planner (EMP)
- Bewertung von evolutionär optimierten hypothetischen Bewegungstrajektorien bezüglich verschiedener Zielfunktionen
 - Zielannäherung
 - Hindernisabstand
 - Fahrtrichtung

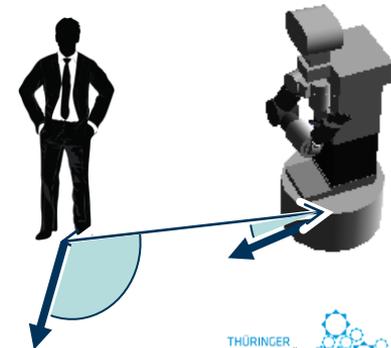
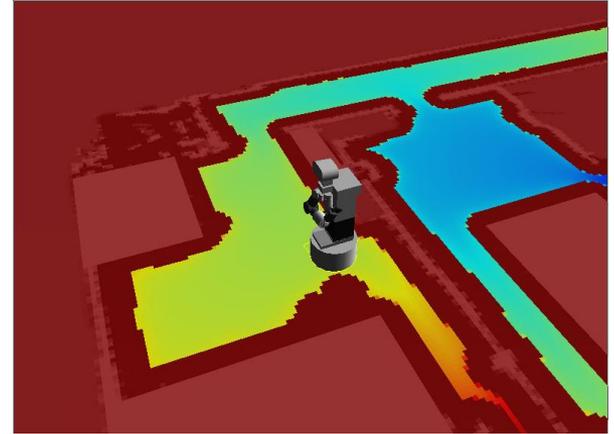
Erweiterungen im Rahmen von SONARO:

- Berücksichtigung des veränderlichen Robotergrundrisses bei ausgestrecktem Arm
- Modellierung der Hardwareregler in den neuen Roboterplattformen (PID-Regler)



Heranfahren an eine Zielperson

1. Optimierungsverfahren für eine geeigneten Zielposition (online)
 2. Navigation mit EMP zu beweglichem Ziel
- Bewertung der lokalen Umgebung um die Zielperson bzgl.:
 - **Erreichbarkeit für den Roboter**
 - **Abstand zur Person (Sollabstand 1,2m)**
 - **Orientierung Roboter zur Person**
 - **Orientierung Person zum Roboter**
 - Maximumsauswahl liefert Roboterzielposition



Konnten die Erwartungen erfüllt werden?

Aspekt	Erwartungen zu Projektbeginn	Erreichter Stand
Umgebungsmodellierung	Neue multimodale 3D-Sensorik sollte zur 3D Kartierung der lokalen Roboterumgebung genutzt werden	Statische und dynamische lokale 3D Karte werden mit den neuen Sensordaten gespeist.
Nutzerzentrierte Navigation	Die am Fachgebiet NI&KR genutzten Methoden sollten für die neuen Roboterplattformen und Einsatzszenarien angepasst werden	Neue Lokalisationsverfahren wurden erprobt und die Navigationsverfahren funktionieren auch auf den neuen Roboterplattformen.

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Beiratstreffen am 24.11.2021
Vorstellung der Demonstratoren

www.sonaro-projekt.de



Was wird demonstriert?

1. Modellfreies Objektgreifen mit stationärer Roboterinstallation



2. Transportszenario mit Tiago Roboter

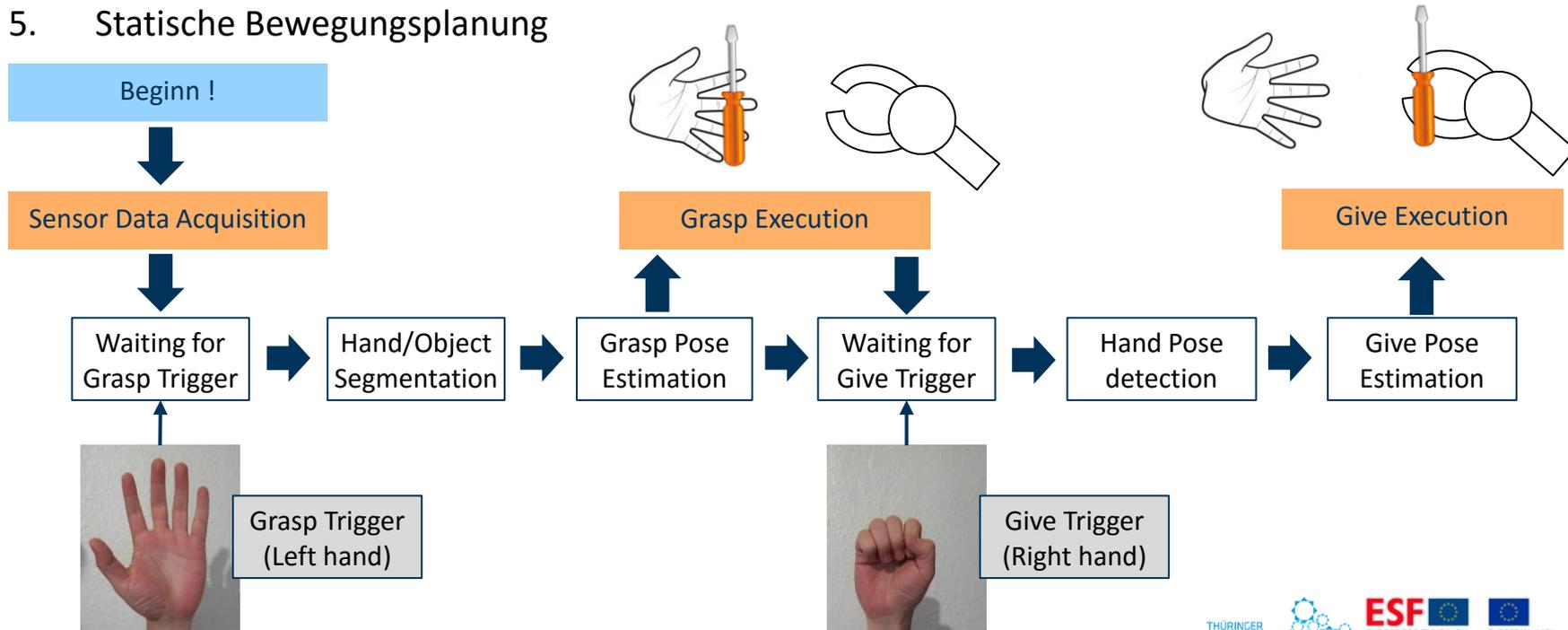


Application		SONARO Demo					
Behaviors							Object
Observe User	Approach Person	Grasp Object	Inspect Object	Transport Object	Regrasp Object	Hand Over	
Skills							
Person Tracking/ Identification ✓	Object Tracking ✓	Motion Planning (Arm) ✓	Navigation ✓				
Gesture Recognition ✓	Hand/Object Segmentation ✓	Grasp Pose Estimation ✓	Localization ✓				
Skeleton Tracking ✓	Registration / PointCloud ✓	Object Shape Estimation ✓	Local Mapping ✓				
Hardware							
Laser Scanner	RTLS Flores ✓	Kinect4 ✓	Thermal Camera ✓	Robot Arms	Robot Platform		

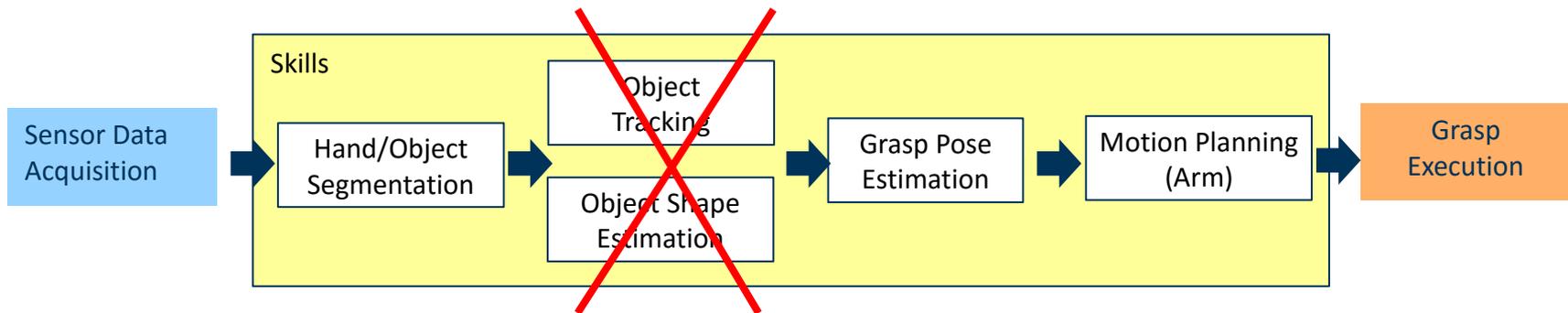
3. Umgreifen von Objekten zwischen Übernahme und Übergabe

Demonstration unbekanntes Objektgreifen

1. Greifen von bekannten und unbekannten Objekten mit stationärer Roboterinstallation.
2. Die zu greifende Objekte werden von der Hand in verschiedenen Körperstellungen gehalten.
3. Mittels Handgesten wird der Übernahmeprozess getriggert.
4. Modellfreie Greifposenbestimmung
5. Statische Bewegungsplanung



Konkrete Realisierung der Übernahme/Übergabe Demo



• RGB-D

Kameras:

- Kinect4Azure
- **AstraOrbbec**
- ASUS Xtion
- IOF Kamera

• Thermokamera

- **I3system TE-Q1**

• Heuristisch

• 2D Farbbild-segmentierung

- PointRend
- YolactEdge

• 3D PointCloud Segmentierung

- XYZ RGB
- **XYZ RGB Thermo (RandLANet)**

• Klassisch:

- ICP + Kalman-Filter Tracking
- Keypoint+ ICP

• Machine Learning:

- Pixel2Mesh
- Mesh RCNN
- ObMan (AtlasNet)
- DeepIM

• Analytisch:

- Online
- Approximativ
- Affordances

• Modellfrei

- **Centroid**

• Machine Learning:

- GG-CNN

• Dynamisch:

- ArmPilot

• Statisch:

- **ROS MoveIt**

• Hardware:

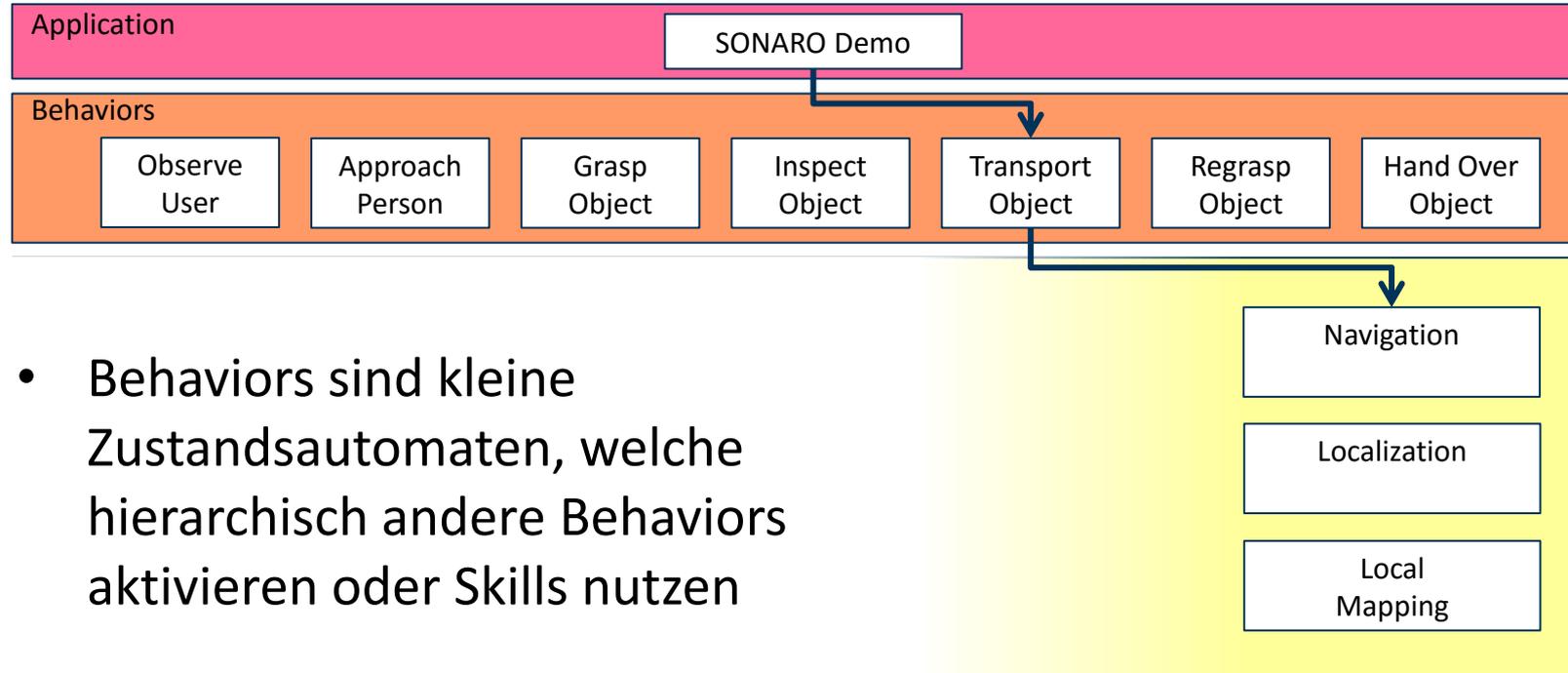
- GraspBox (Kinova Arm)
- Tiago (5 Finger Hand, Parallel-Greifer)
- Scitos X3 (Kinova Arm)
- **UR5 Arm**

Demonstration unbekanntes Objektgreifen

Kamerasystem
(RGB-D-T)

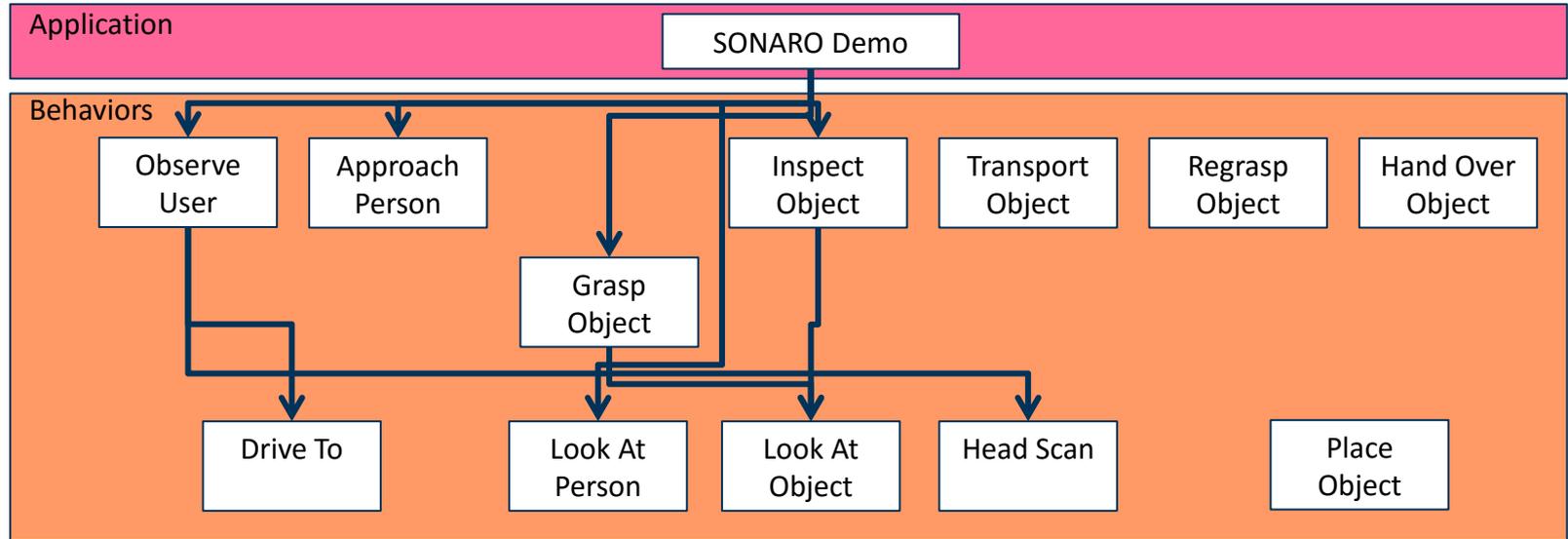


Behavior Architektur



- Behaviors sind kleine Zustandsautomaten, welche hierarchisch andere Behaviors aktivieren oder Skills nutzen

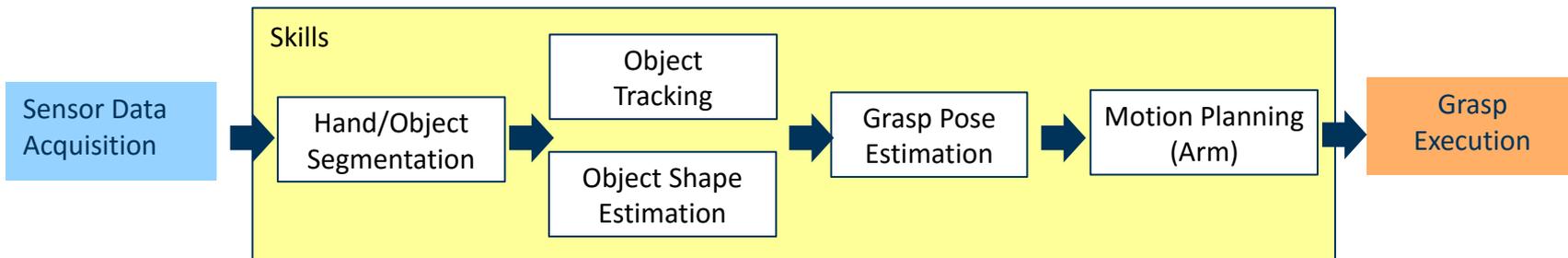
Ablauf der Demo aus Sicht des Roboters



1. Navigation zur Beobachtungsposition
2. Erstellung einer statischen Hinderniskarte
3. Warten auf Intenterkennung (Triggergeste)
4. Heranfahren an den Nutzer
5. Objekt greifen
6. Griff verifizieren (Objektlage checken)

7. Verstauen des Objekts zum Transport
8. Navigation zu Zielperson
9. Warten auf Hand in Reichweite
10. Übergabe in die detektierte Hand
11. Rückzug zur Beobachtungspose

Konkrete Realisierung der Übernahme/Übergabe Demo



• RGB-D

Kameras:

- **Kinect4Azure**
- **AstraOrbbec**
- ASUS Xtion
- IOF Kamera

• Thermokamera

- **I3system TE-Q1**

• Heuristisch

• 2D Farbbildsegmentierung

- PointRend
- YolactEdge

• 3D PointCloud Segmentierung

- XYZ RGB
- **XYZ RGB Thermo (RandLANet)**

• Klassisch:

- **ICP + Kalman-Filter Tracking**
- **Keypoint+ ICP**

• Machine Learning:

- Pixel2Mesh
- Mesh RCNN
- ObMan (AtlasNet)
- DeepIM

• Analytisch:

- **Online**
- Approximativ
- Affordances

• Modellfrei

- Centroid

• Machine Learning:

- GG-CNN

• Dynamisch:

- **ArmPilot**
- **Statisch:**
- ROS MoveIt

• Hardware:

- GraspBox (Kinova Arm)
- **Tiago (5 Finger Hand, Parallel-Greifer)**
- Scitos X3 (Kinova Arm)
- UR5 Arm

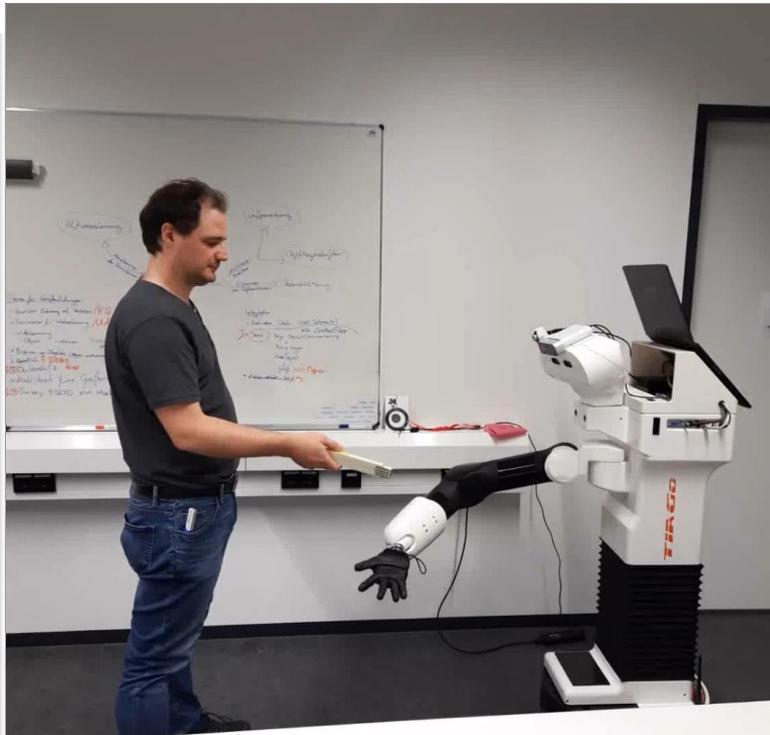
Demonstrator Übernahme/Übergabe

Gesamtablauf ist auf Tiago implementiert, aber:

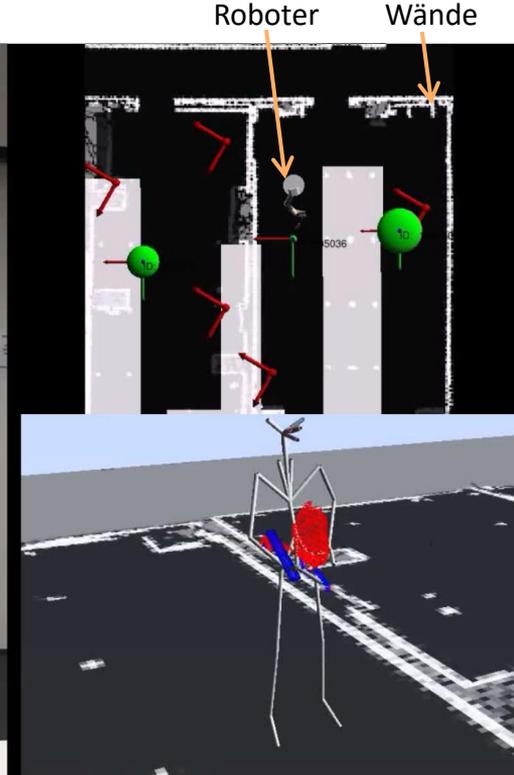
- Sehr komplexer Aufbau aus vielen Teilmodulen macht den Ablauf fehleranfällig!
- Schwerfällige Roboterbewegungen (langsam) stehen einem effizienten Praxiseinsatz entgegen (kooperativer Nutzer mit viel Geduld erforderlich)
- **Limitierung der greifbaren Objekte durch**
 - a) Wahrnehmbarkeit mit den Tiefenkameras (kein Metall oder Schwarz, genügend sichtbare Fläche mit 3d Features)
 - b) 5-Finger-Hand von Tiago kann nur stabförmige Objekte robust greifen (zu wenig sensorisches Feedback und nur 3DoF)



Demo Übernahme, Transport, Übergabe



ungeschnittene Ansicht von außen



Draufsicht auf die Testumgebung mit getrackten Personen

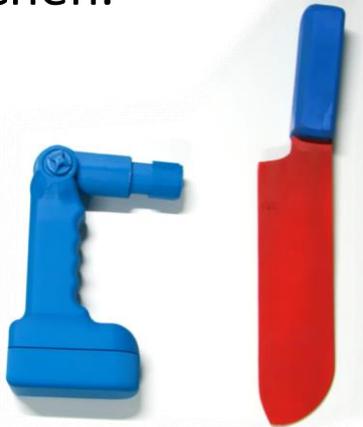
-  RTLS-Anchor
-  Personenhypothesen
Größe = Ortsunsicherheit

3D Darstellung der Skelettschätzung und segmentierten Punktwolke

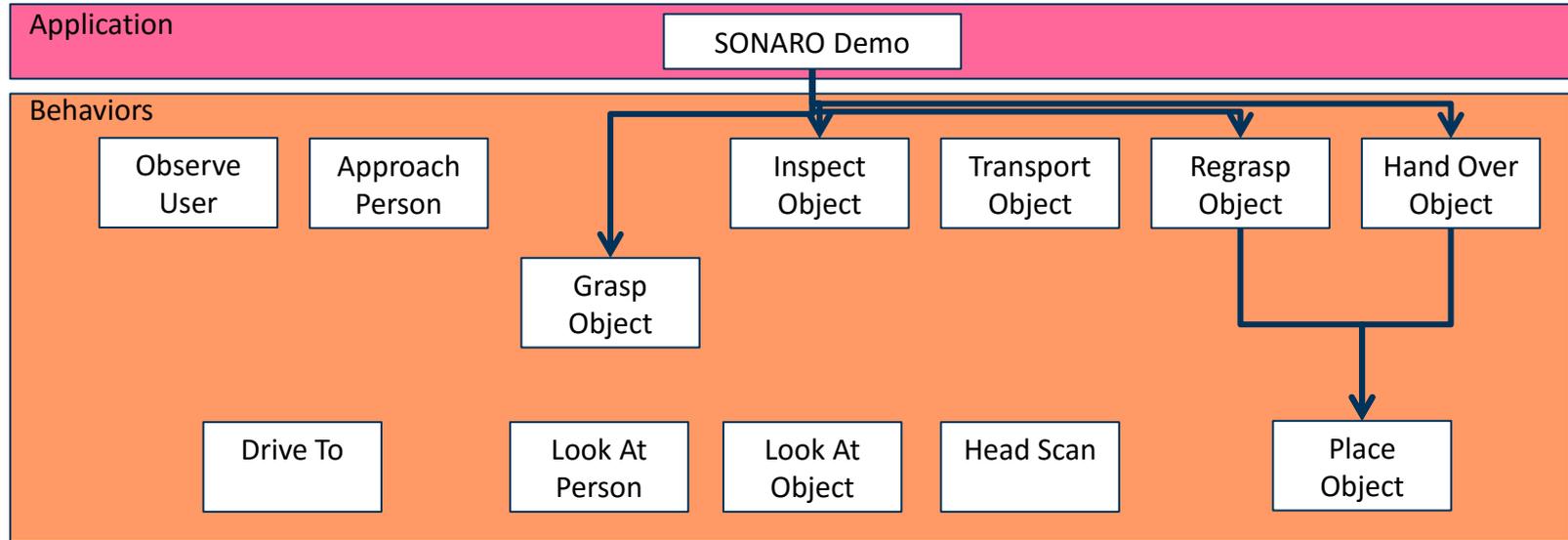
-  Segmentierte Handpunkte
-  Segmentierte Objektpunkte
-  Objektmodell
Positionsschätzung

Demonstration ergonomisches Umgreifen

- Wenn der Roboter die Griffe an Objekten mit seinem Greifer belegt, kann er umgreifen, um dem Menschen den Gegenstand mit dem Griff voran zu reichen.

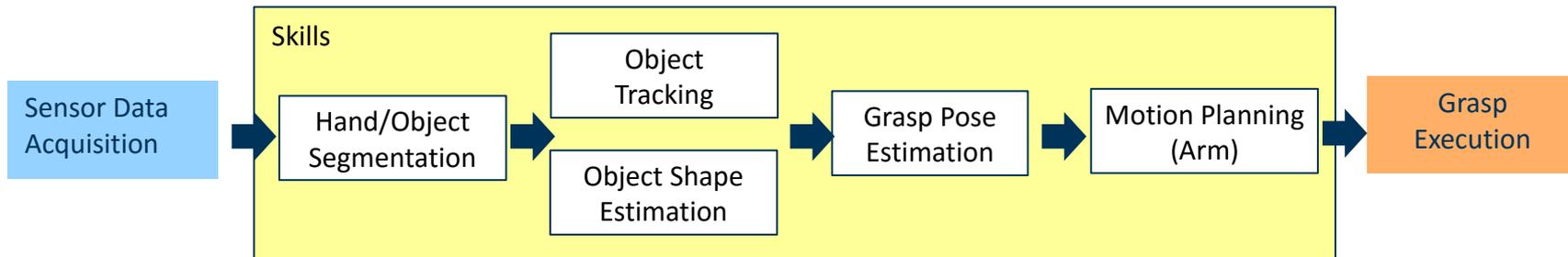


Demonstration ergonomisches Umgreifen



1. Objekt wird übernommen
2. Überprüfung, ob Übernahmegriffe für menschliche Hand blockiert sind
3. Umgreifen
4. Überprüfen des Griffs
5. Übergabe an Mensch

Konkrete Realisierung der Demo ergonomisches Umgreifen



• RGB-D

Kameras:

- Kinect4Azure
- AstraOrbbec
- **ASUS Xtion**
- IOF Kamera

• Thermokamera

- I3system TE-Q1

• Heuristisch

• 2D Farbbild-segmentierung

- PointRend
- **YolactEdge**

• 3D PointCloud Segmentierung

- XYZ RGB
- XYZ RGB Thermo (RandLANet)

• Klassisch:

- **ICP + Kalman-Filter Tracking**
- **Keypoint+ ICP**

• Machine Learning:

- Pixel2Mesh
- Mesh RCNN
- ObMan (AtlasNet)
- DeepIM

• Analytisch:

- **Online**
- Approximativ
- Affordances

• Modellfrei

- Centroid

• Machine Learning:

- GG-CNN

• Dynamisch:

- **ArmPilot**

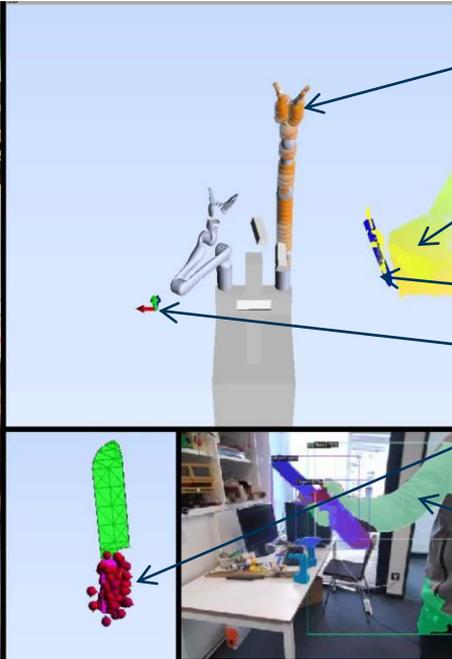
• Statisch:

- ROS MoveIt

• Hardware:

- GraspBox (Kinova Arm)
- Tiago (5 Finger Hand, Parallel-Greifer)
- **Scitos X3 (Kinova Arme)**
- UR5 Arm

Demonstration ergonomisches Umgreifen



- Robotermodell mit Hindernisgeometrie
- Punktwolke der Tiefenkamera
- Objektmodell
- Übergabepose (Hand)
- Greifposen für menschliche Hand
- 2D Segmentierung von Hand und Objekt

Konnten die Erwartungen erfüllt werden?



Aspekt	Angestrebter Idealzustand	Erreichter Stand
Objektmodellierung	Unbekannte Objekte können während des Vorhaltens in ihrer 3D Form erfasst werden	Objektmodelle müssen vorher bekannt sein, Objektklassifikation aus bekannter Menge ist möglich
Lageschätzung des Objekts	Dynamische Bewegungen sollen stabil nachverfolgt werden können	Segmentierung der Objektpunkte + ICP basiertes Tracking sind Fehleranfällig → Feedback für Nutzer über Trackingzustand erforderlich
Roboterbewegungen	Dynamischer, ergonomischer Bewegungsablauf	Bewegungen sind eher träge und langwierig. <ul style="list-style-type: none">• Beschränkungen der Hardware• Datenverarbeitungsgeschwindigkeit / Wahrnehmung
Praxistauglichkeit	Uneingewiesene Nutzer sollten sicher mit dem System interagieren können	Kooperativer Nutzer erforderlich

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Beiratstreffen am 24.11.2021

Zusammenfassung der wissenschaftlichen Ergebnisse

Managementreport

Prof. Groß

Studentische Arbeiten im Rahmen von SONARO

- **TUI-NIKR:**

- **MA Patrick Stäblein:** *Integration eines Objekt-/Handtracking-Ansatzes in die reale Roboterapplikation, TU-Ilmenau 2021*
- **MA Max Kogel:** *Intenterkennung für die Objektübergabe anhand von Skellettracking, TU-Ilmenau 2021*
- **MA Hendrik Köhler:** *Hands on Deep Learning 3D Handposen Schätzung, TU-Ilmenau, 2021*
- **MA Tim-Justin Aldinger:** *Bewegungssteuerung für einen Roboter mit Greifer, TU-Ilmenau, 2020*
- **MA Markus Paschke:** *Deep-Learning-Methoden zur Erstellung von 3D-Modellen, TU-Ilmenau, 2020*
- **BA Hung Le Huy:** *Automatische Kalibrierung der Kameras einer mobilen Roboterplattform, TU-Ilmenau 2020*
- **BA Julius Lerm:** *2D zu 3D Transformation für bildbasierte Skeletttracker, TU-Ilmenau, 2020*
- **BA Amanda Robillard:** *Untersuchung von 3D-Skelett-basierten Verfahren zur Aktionserkennung, TU-Ilmenau, 2020*
- **BA Wei Dai:** *Approximation von Greifposenbewertungsfunktionen mittels Neuronaler Netze, TU-Ilmenau, 2020*
- **Projektseminar Toan Phan Bao, Lukas Vollmar:** *Aufbau einer Testumgebung für robotisches Greifen, TU-Ilmenau, 2021*

Studentische Arbeiten im Rahmen von SONARO

- **TUI-QBV:**

- **MA Jinggang Yan:** *Untersuchung zur Nutzbarkeit des Convolutional-Neural-Networks für die echtzeitfähige Objektdetektion auf Basis von Farb- und 3D-Bilddaten.* TU-Ilmenau, 2019
- **MA Zaizhuang Miao:** *Untersuchung zur Beschleunigung der Partikel-Filter basierend Objekttrackingsverfahren.* TU-Ilmenau, 2019
- **Projektarbeit Yujian Yuan, Xiaojiang Han und Jinxin Zhu:** *Untersuchung zum Verfahren der Kalibrierung und der Registrierung von 3D-Kamera und Thermokamera.* TU-Ilmenau, 2020
- **Projektarbeit Yang Li, Xiao Lin und Jingyu Wang:** *Die Entwicklung der effizienten teilweise automatischen Annotationsapplikation für die multimodale Punktwolke.* TU-Ilmenau, 2021

- **HSM:**

- **Projektseminar Tony Schneider, Nicolas Schmitt:** *Bestimmung materialabhängiger Messfehler beim Einsatz von RTLS Lokalisierungssystem,* Hochschule Schmalkalden, 2019
- **Projektseminar Stephan Hintz, Adrian Herrman, Nicolas Schmitt:** *Messfehlerbestimmung im 3D-Raum von UWB Signalen zur Indoor-Navigation mit RTLS ,* Hochschule Schmalkalden, 2020
- **Projektseminar Tim Pfeffer:** *Erarbeitung eines SCITOS/Mira-Labordemonstrators ,* Hochschule Schmalkalden, 2021

Wissenschaftliche Publikationen

Internationale
wissenschaftl.
Konferenz-
beiträge:

Wengefeld, T., Müller, St., Lewandowski, B., Gross, H.-M. *A Multi Modal People Tracker for Real Time Human Robot Interaction.* in: IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN), New Delhi, India, 8 pages, IEEE 2019

Zhang Y, Zhang C, Rosenberger M, Notni G. *6D Object Pose Estimation Algorithm Using Preprocessing of Segmentation and Keypoint Extraction.* in: IEEE I2MTC Conference, Dubrovnik, 2020

Zhang Y, Zhang C, Nestler R, Rosenberger M, Notni G. *Efficient 3D object tracking approach based on convolutional neural network and Monte Carlo algorithms used for a pick and place robot.* In: SPIE IMEKO Conference, Jena, 2019.

N. Fränzel, N. Greifzu, M. Schneider, A. Wenzel: *Bootstrapping a Wireless Localization System,* IEEE: The 4th International Conference on Advanced Communication Technologies and Networking, CommNet 2021, ISBN: 978-6654-0306-1, 12/2021

Rosenberger, Maik, et al. *3D high-resolution multimodal imaging system for real-time applications.* Dimensional Optical Metrology and Inspection for Practical Applications IX. Vol. 11397. International Society for Optics and Photonics, 2020.

N. Fränzel, N. Greifzu, A. Wenzel: *Robuste Lokalisierung in drahtlosen Sensornetzwerken,* IEEE: The 4th International Conference on Advanced Communication Technologies and Networking, CommNet 2021, ISBN: 978-1-6654-0306-1, 12/2021

Müller, St., Stephan, B., Gross, H.-M. *MDP-based Motion Planning for Grasping in Dynamic Szenarios.* in: Europ. Conf. on Mobile Robotics (ECMR), Bonn, Germany, 2021

Wissenschaftliche Publikationen

Journalpublikationen:

Zhang, Y., et al. *Efficient 3D object tracking approach based on convolutional neural network and Monte Carlo algorithms used for a pick and place robot.* Photonics and Education in Measurement Science 2019. Vol. 11144. International Society for Optics and Photonics, 2019.

Müller, St., Wengefeld, T., Trinh, T. Q., Aganian, D., Eisenbach, M., Gross, H.-M. *A Multi-Modal Person Perception Framework for Socially Interactive Mobile Service Robots.* Sensors, vol. 20 (2020) 3, 722, 18 pages

Zhang, Y., Müller, St., Stephan, B., Gross, H.-M., Notni, G. *Point Cloud Hand-Object Segmentation Using Multimodal Imaging with Thermal and Color Data for Safe Robotic Object Handover.* in: Sensors 2021, 21(16), 5676

Fazit zu wissenschaftlichen Ergebnissen

- **17** Studentische Abschlussarbeiten
(7 **Masterarbeiten**, 4 **Bachelorarbeiten**, 6 **Projektarbeiten**)
- **7** internationale Konferenzbeiträge
- **3** internationale Journalpublikationen
- **3** Präsentationen auf Fachtagungen

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Beiratstreffen am 24.11.2021

Diskussion der Ergebnisse mit dem Beirat

www.sonaro-projekt.de



Diskussionspunkte

- Industrielle Nutzbarkeit von Teilaspekten des Projektes – wo und wozu?
➔ RTLS-Flares, Planer eher nicht, RGB-D-T-Sensorik
- Synergie ThZM (Innovationszentrum) + Forschergruppe
- Genauigkeit und Aufwand der RTLS-Kalibrierung?
- Behandlung und Integration der semantischen Informationen in den 3D-Modellen? (Integration von Common Sense Wissen)
- Dank für umfangreiche Arbeit in den verschiedenen Teilteams der FoGru
- ...