

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die nutzerzentrierte mobile Assistenzrobotik

5. Workshop / Beiratstreffen
am 20.04.2021

www.sonaro-projekt.de



Agenda

9:00	Prof. Groß	Begrüßung der Teilnehmer und Einführung zu den Eckdaten des Projektes im 4 Halbjahr
9:10	Dr. Müller	Einführung des Demonstratorszenarios und der Verarbeitungspipeline für den Zugriff
9:25	M.Sc. Zhang	Fortschrittsbericht zu den Arbeitspaketen von TUI-QBV
9:45	Dr. Garten	Fortschrittsbericht zu den Arbeitspaketen von GFE
10:05	Dr. Müller	Fortschrittsbericht zu den Arbeitspaketen von TUI-NIKR
10:25	M.Sc. Schneider	Fortschrittsbericht zu den Arbeitspaketen von HSM
10:45	Dr. Müller	Demonstration des Roboters im Labor
11:05	Prof. Notni	Frei Diskussion der Ergebnisse mit dem Beirat
11:25	Dr. Djambova	Feedback des Projektträgers



Vorstellung - SONARO Forschergruppe

TU Ilmenau,
FG Neuroinformatik und
Kognitive Robotik
TUI-NIKR



Prof. Groß



Dr. Müller



TECHNISCHE UNIVERSITÄT
ILMENAU

TU Ilmenau,
FG Qualitätssicherung und
Industrielle Bildverarbeitung
TUI-QBV



Prof. Notni



M. Sc. Zhang



TECHNISCHE UNIVERSITÄT
ILMENAU



HOCHSCHULE
SCHMALKALDEN
UNIVERSITY OF APPLIED SCIENCES



Hochschule
Schmalkalden,
FG Eingebettete
Diagnosesysteme **HSM**



Prof. Wenzel



M. Sc. Schneider

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



Dr. Welzel



Dr. Garten

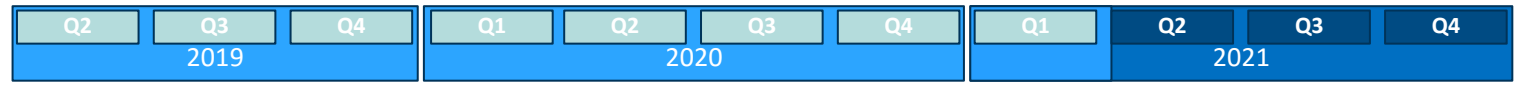
SONARO Unternehmensbeirat

- ✓ Dr. Trabert, Metralabs GmbH, Ilmenau
- ✓ Herr Waldmann, Betriebsleiter HENKEL+ROTH GmbH, Ilmenau
- ✓ Herr Hofmann, Cluster für Fertigungstechnik & Metallbearbeitung
 - Herr Sommer, SCS Robotik UG, Schmalkalden
- ✓ Herr Richter, Honda Research Institute Europe
 - Dr. Behling, Götting KG, Lehrte
 - Vision & Control GmbH, Suhl
 - Hörisch Präzision GmbH

Eckdaten der Forschergruppe



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

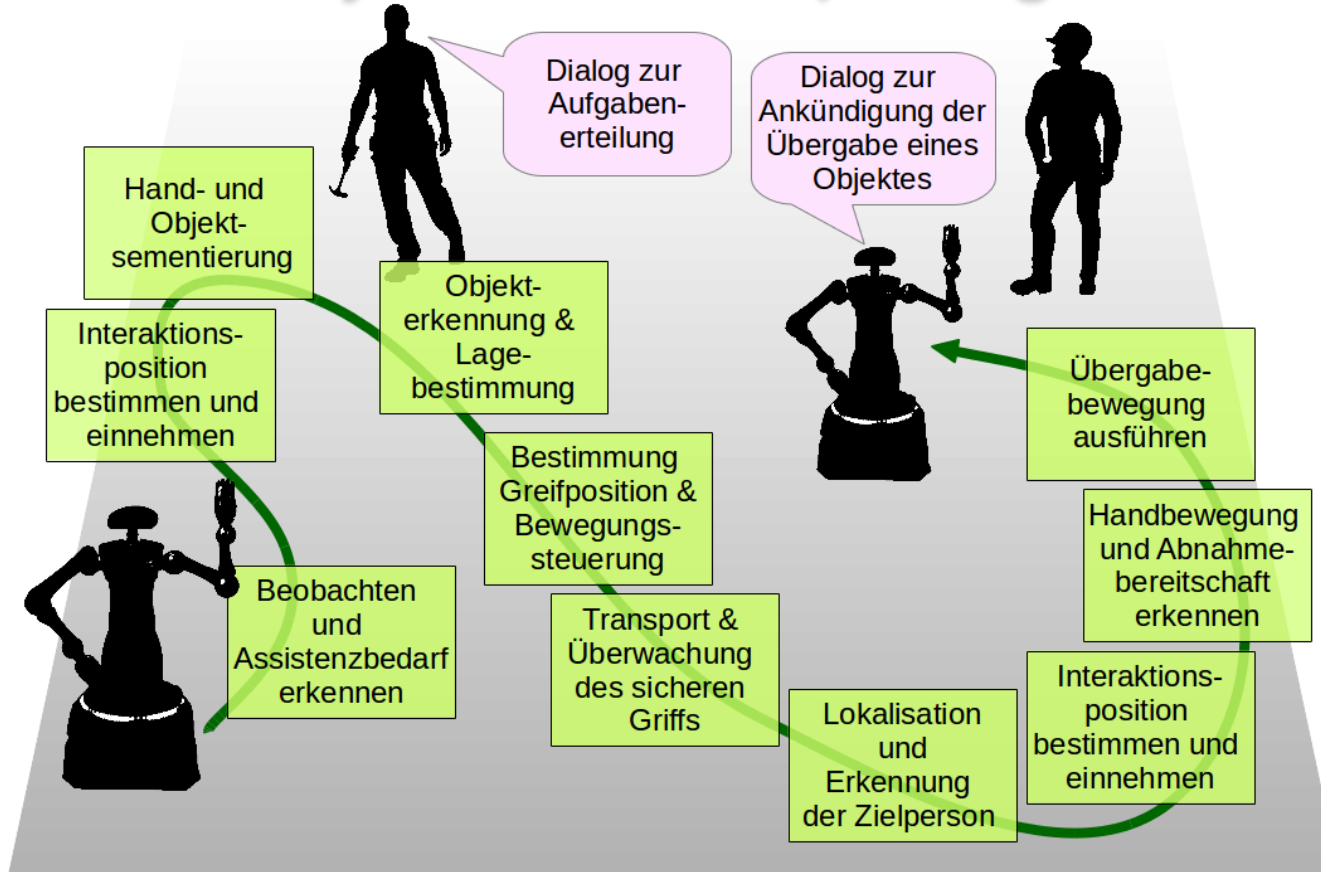


1.4.2019



31.12.2021

Ablauf einer Objektübernahme/-übergabe



Präzisierung der Grobplanung/Zielfunktionalitäten für die kommenden zwei Jahre

Milestones	Monat	Beschreibung
MS1 	01/2020	Roboter betriebsbereit, kann Personen wahrnehmen und autonom navigieren
MS2 	07/2020	Mobile RTLS-Flares sind funktionsfähig und in Demonstrator integriert
MS3 	07/2020	Multispektraler 3D-Sensor ist betriebsbereit und auf Roboter integriert
MS4 	01/2021	Präsentierte Objekte können von Hand getrennt, klassifiziert und im Raum getrackt werden
MS5	10/2021	Roboter kann im vollständigen Ablauf Gegenstände von Person A greifen und bei B wieder übergeben
MS6	12/2021	Die Leistungsfähigkeit des Systems ist evaluiert

Studentische Arbeiten im Wintersemester 2020/21

- **TUI-NIKR:**



- **MA Patrick Stäblein:** *Deep-Learning-Methoden zur Erstellung von, 3D-Modellen, TU-Ilmenau 2021*

- **MA Max Kogel:** *Intenterkennung für die Objektübergabe anhand von Skelettracking, TU-Ilmenau 2021*



- **Projektseminar Lucas Vollmar, Toan Phan-Bao:** *Aufbau einer Testumgebung für robotisches Greifen, TU Ilmenau 2021*

- **TUI-QBV:**



- **Projektarbeit Yang Li, Xiao Lin und Jingyu Wang:** *Untersuchung des Bewegungsplanens des Roboterarms (UR-10) mit der Hilfe von multimodalen 3D-Sensoren. TU-Ilmenau 2021*

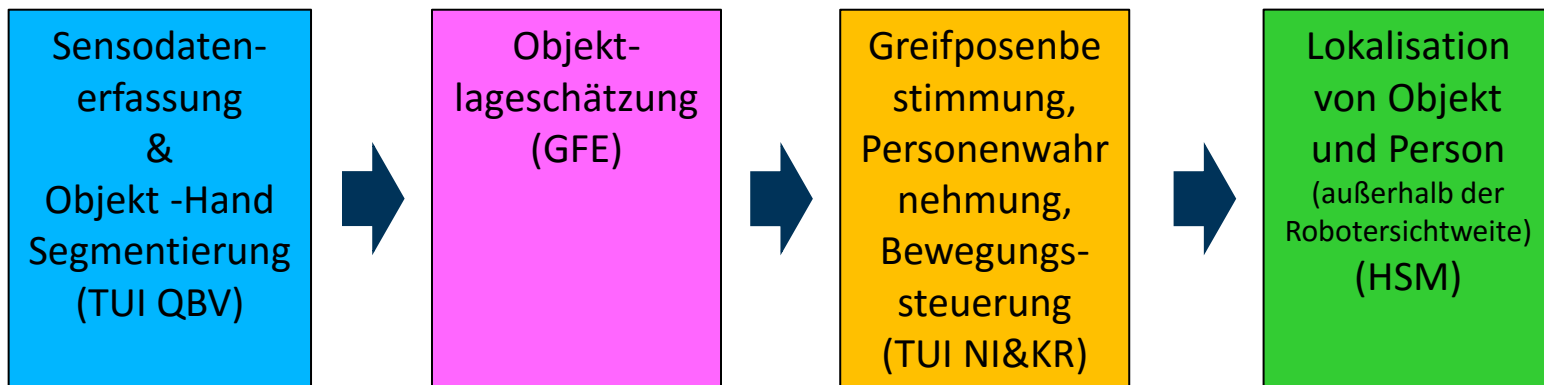
- **Bachelorarbeit Hake, Cornelius;** *Anwendungsorientierte Charakterisierung und Beurteilung des Intel RealSense Lidars L515 nach VDI 2634, TU-Ilmenau 2021*

- **HSM:**

- **Projektarbeit Sasi Kumar Selvaraj:** *Implementation for RTLS localisation in ROS and rviz (abgebrochen)*
- **Projektarbeit Stephan Hintz, Adrian Herrman, Nicolas Schmitt:** *Messfehlerbestimmung von UWB Signalen zur Indoor-Navigation (kurz vor Abschluss)*

Ablauf im Folgenden

- Überblick zu den Realisierungsmöglichkeiten des Zugreifens mit dem Roboter
- Anschließend Bearbeitungsschwerpunkte gemäß des Datenflusses/Ablaufs



Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die nutzerzentrierte mobile Assistenzrobotik

Einführung des Demonstratorszenarios
und der Verarbeitungspipeline für das Zugreifen

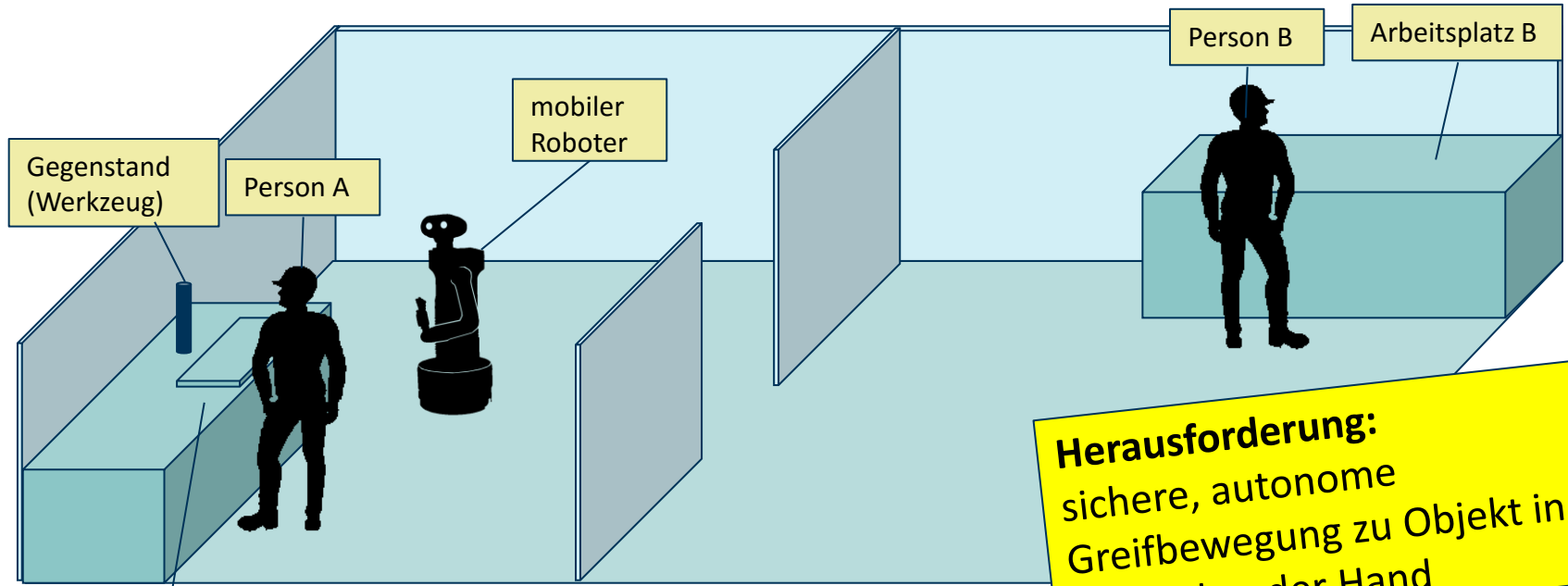
Dr. Steffen Müller

www.sonaro-projekt.de



Demonstrationsszenario: Einsatz in der Produktion

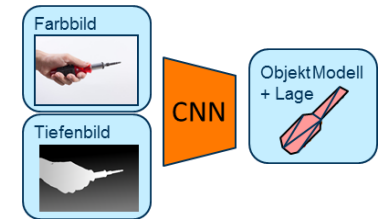
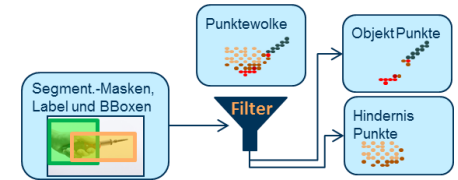
- Zwei oder mehr Montageplätze an denen Werker Montagearbeiten ausführen
- **Bedarf:** Werkzeug/Werkstück wird durch den Roboter von Person A übernommen und zu Person B gebracht und übergeben



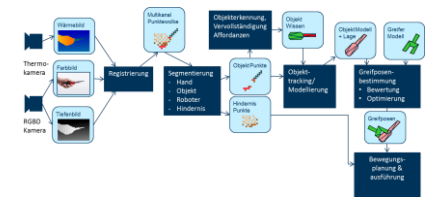
Herausforderung:
sichere, autonome
Greifbewegung zu Objekt in
darreichender Hand

Designalternativen für die Umsetzung des robotischen Greifens

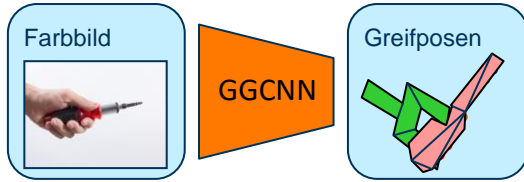
- **Modellfrei vs. Modell-basiert**
- **Segmentierung im 3D vs. Segmentierung in 2D**
- **Modellschätzung vs. Bekanntes Objekt**



Die verschiedenen Verfahren wurden (gemäß der Spezifika der versch. Projektpartner) in unterschiedliche Pipelines integriert.



Modellfrei



Vorteile:

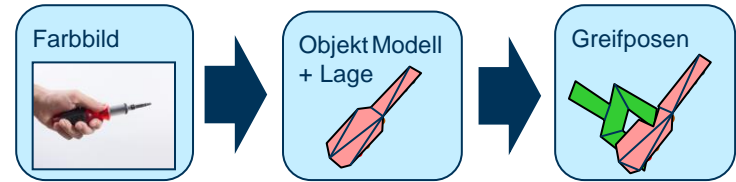
- Keine fehleranfälligen Zwischenergebnisse
- Unabhängig vom Objekt
- Auch mit unbekanntem Objekten

Nachteile:

- Keine Trennung von Hand und Objekt
- Keine Kontrolle über das Ziel
- Keine Berücksichtigung von Task und Affordances
- Keine Berücksichtigung verdeckter Objektteile
- Aufwändiges Training für neue HW und neue Aufgabe

vs.

Modell-basiert



Vorteile:

- Mehr Kontrolle über Zielobjekt und Taskspezifische Greifaufgaben
- Explizite Trennung von Hand und Objekt
- HW unabhängig

Nachteile:

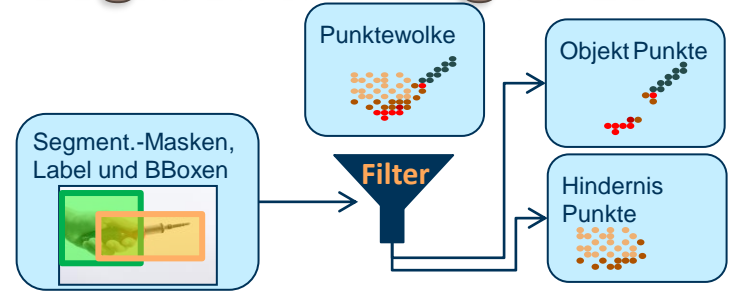
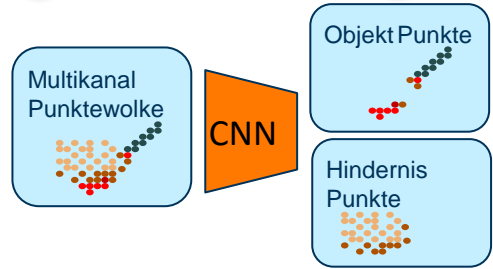
- Mehr Points of Failure in der Kette
- Komplexere Architektur

Untersuchte Ansätze:
 GG-CNN (modellfrei) nicht zufriedenstellend
 → Konzentration auf Modellbasierte Pipeline

Segmentierung im 3D

vs.

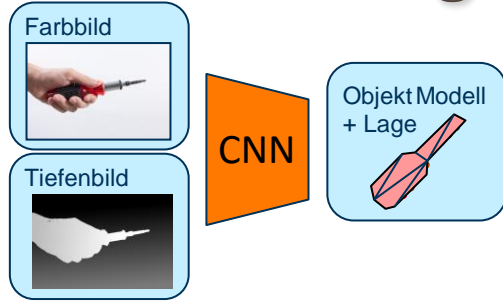
Segmentierung in 2D



Ziel: Trennung von Hand- und Objektpunktwolke (Hindernis und Ziel)

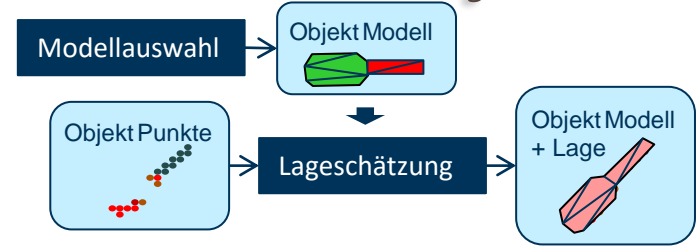
- Beide Varianten wurden untersucht (3D Segmentierung → TUI QBV), (2D Segmentierung → TUI NI&KR)
- 2D Semantische Segmentierung mittels PointRend → trainiert auf eigenem Datensatz von in Händen gehaltenen Objekten
- Sehr vielversprechende und performante Lösung (10FPS auf GPU)
- Objekterkennung fällt mit ab → Auswahl des geeigneten Objektmodells für Tracking und Greifposenermittlung

Modellschätzung



vs.

Bekanntes Objekt



Ziel: Form und Lage des Objektes im Raum schätzen als Grundlage für Greifposenbestimmung

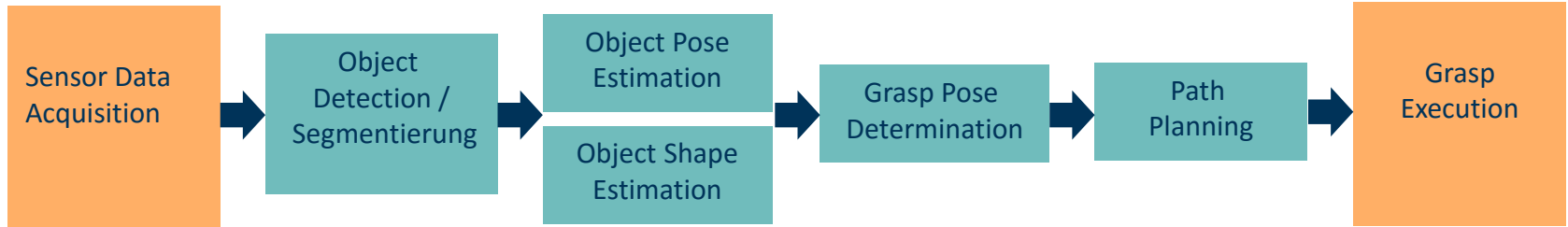
- Drei Machine-Learning-basierte Verfahren zur Modellschätzung wurden untersucht (ObMan [1] , Mesh RCNN [2], Pixel2Mesh [3])
- Alle haben Probleme bei der Generalisierung auf ungesehene Objekte (Training auf künstlichen Daten gerendert aus dem ShapeNet Datensatz)
- Genauigkeit der Objektform und der Position in Kamerakoordinaten ist bislang unzureichend für Greifversuche!
- → Implementierung der Lageschätzung für bekannte Objektmodelle (parallel zu Arbeiten von GFE wurde von TUI-NIKR ein Tracker auf Basis des ICP Algorithmus entwickelt)

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

[2] Georgia Gkioxari, Jitendra Malik, Justin Johnson: Mesh R-CNN, ICCV 2019

[3] Wang, Nanyang, et al. "Pixel2mesh: Generating 3d mesh models from single rgb images." Proceedings of the European Conference on Computer Vision (ECCV). 2018.

Generischer Ablauf des Greifens



• RGB-D

Kameras:

- Kinect4Azure
- AstraOrbbec
- ASUS Xtion
- IOF Kamera

• Thermokamera

- I3system TE-Q1

• Heuristisch:

- Background Subtraction
- Planefitting

• 2D Farbbildsegmentierung

- PointRend
- Contour Matching

• PointCloud Segmentierung

- RGB-D
- RGB-D+Thermo

• Klassisch:

- ICP + Kalman-Filter Tracking
- Keypoint+ ICP

• Machine Learning:

- Pixel2Mesh
- Mesh RCNN
- ObMan (AtlasNet)

• Analytisch:

- Online
- Approximativ
- Affordances

• Machine Learning:

- GG-CNN

• Dynamisch:

- ArmPilot
- Grasp-Objective

• Statisch:

- ROS MoveIt

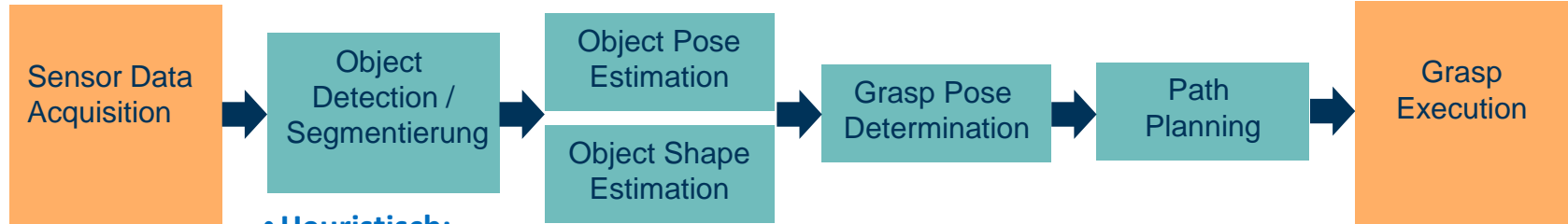
• Hardware:

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

Untersuchte Alternativen für die einzelnen Teilschritte

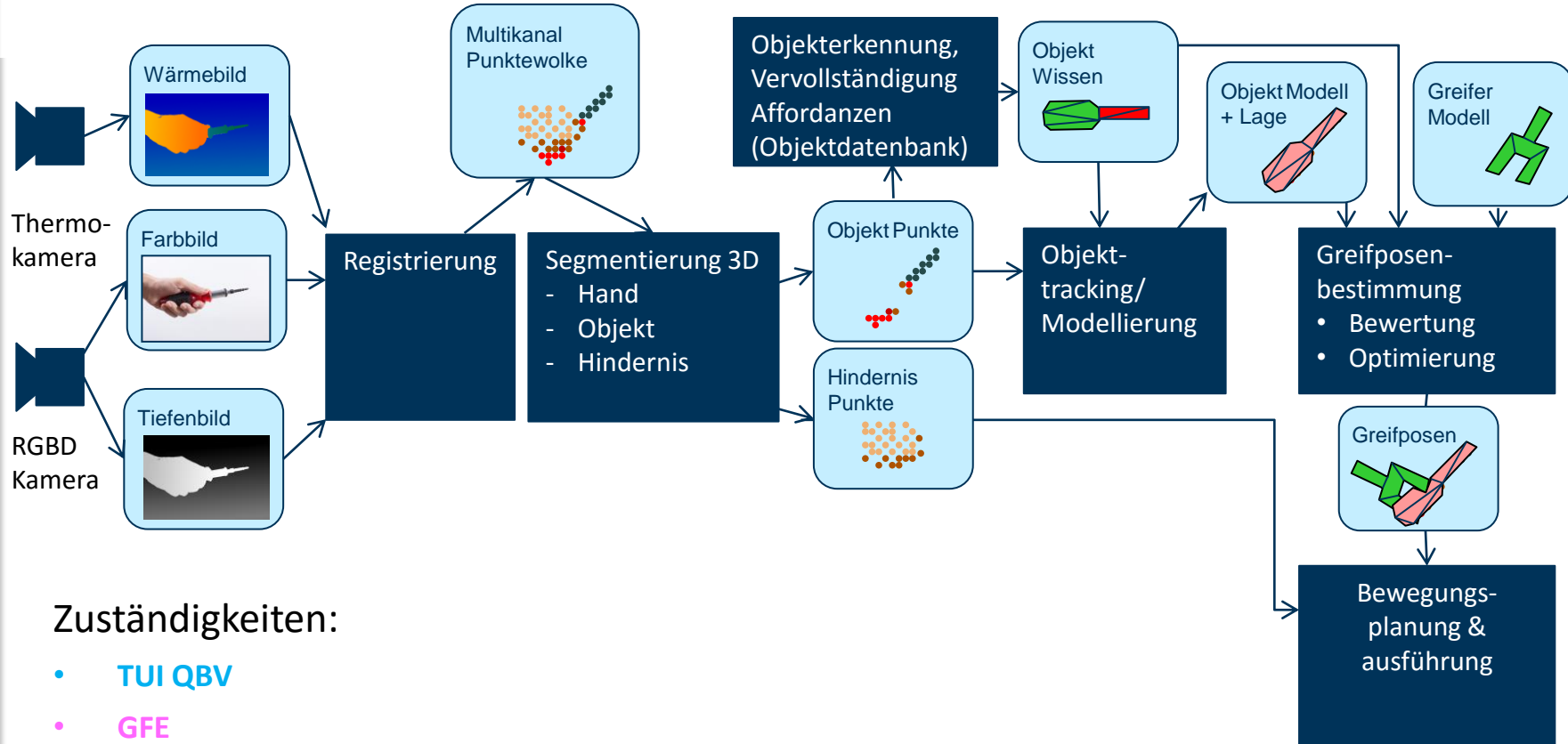
Ursprünglich geplante Umsetzung der Greif-Pipeline

- Viele methodische Auswahlmöglichkeiten in den einzelnen Zwischenschritten
→ Evolution einer praktikablen Lösung



- **Heuristisch:**
 - **RGB-D Kameras:**
 - Background Subtraction
 - Plane fitting
 - Kinect4Azure
 - AstraOrbbec
 - ASUS Xtion
 - **IOF Kamera**
 - **Thermokamera**
 - **I3system TE-Q1**
 - **2D Farbbild-segmentierung**
 - PointRend
 - **PointCloud Segmentierung**
 - RGB-D
 - **RGB-D+Thermo**
- **Klassisch:**
 - **ICP + Kalman-Filter Tracking**
 - Keypoint+ ICP
- **Machine Learning:**
 - Pixel2Mesh
 - Mesh RCNN
 - ObMan (AtlasNet)
- **Analytisch:**
 - **Online**
 - Approximativ
 - Affordances
- **Machine Learning:**
 - GG-CNN
- **Dynamisch:**
 - **ArmPilot**
 - **Grasp-Objective**
- **Statisch:**
 - ROS MoveIt
- **Hardware:**
 - GraspBox (Kinova Arm)
 - **Tiago (5 Finger Hand, Parallel-Greifer)**
 - **Scitos X3 (Kinova Arm)**

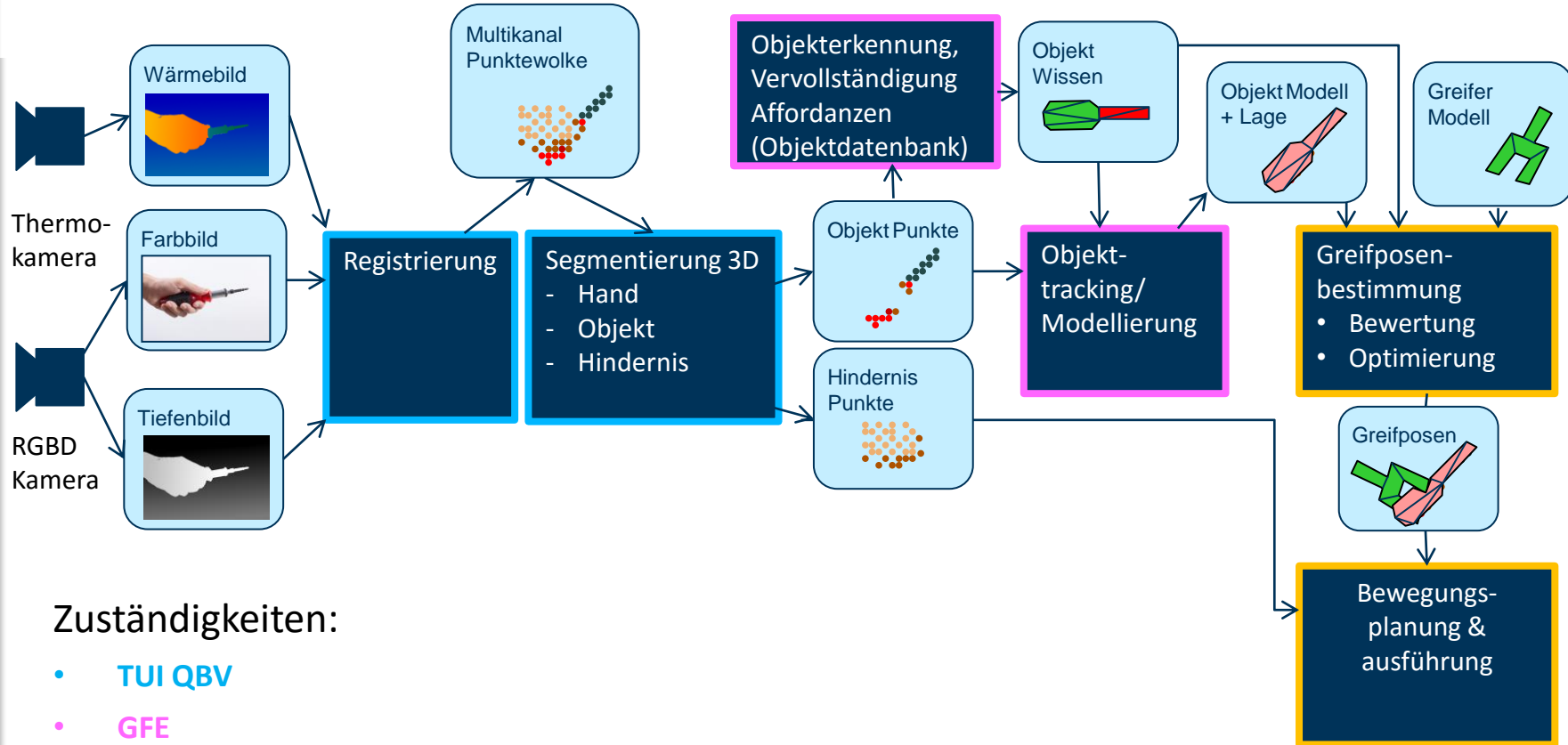
Ursprünglich geplante Umsetzung der Greif-Pipeline (a)



Zuständigkeiten:

- TUI QBV
- GFE
- TUI NI&KR

Ursprünglich geplante Umsetzung der Greif-Pipeline (a)

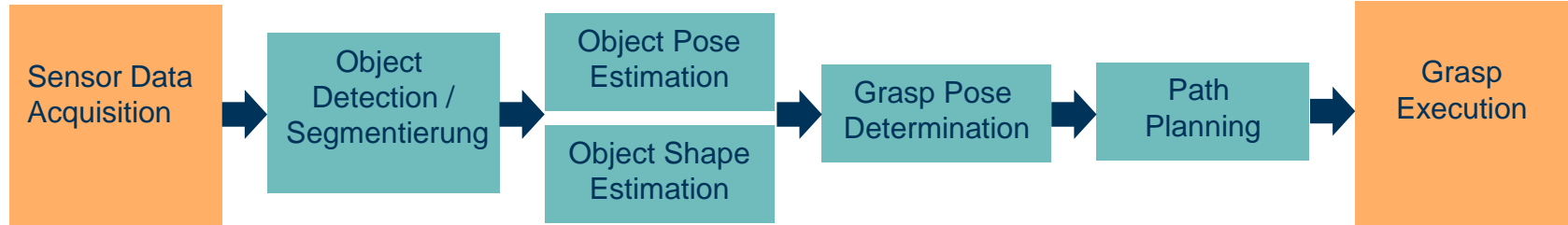


Zuständigkeiten:

- TUI QBV
- GFE
- TUI NI&KR

Integration von SoTA Deep Learning Methoden

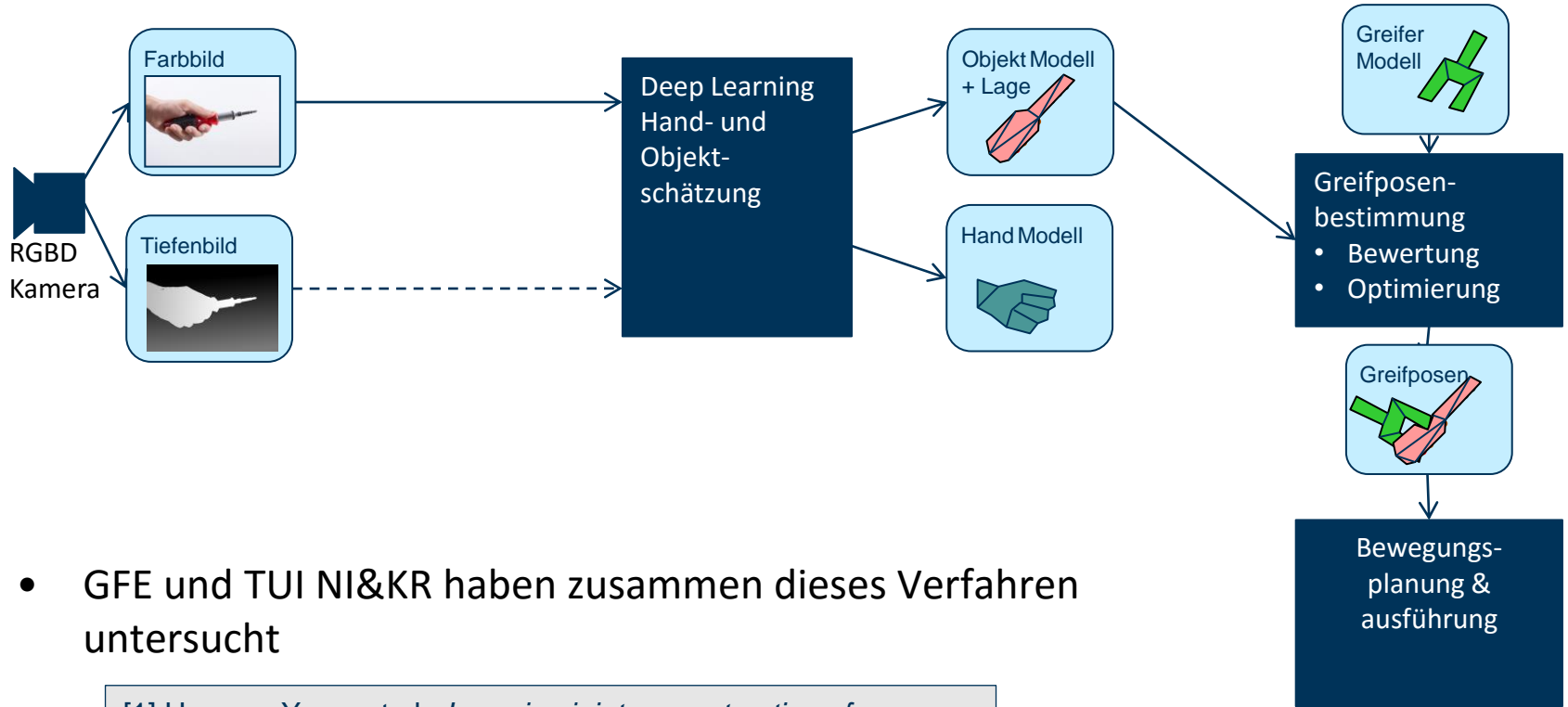
- Neue Machine-Learning Methoden führten zu einer alternativen Pipeline



- **Heuristisch:**
 - Background Subtraction
 - Planefitting
- **RGB-D Kameras:**
 - Kinect4Azure
 - AstraOrbbec
 - ASUS Xtion
 - IOF Kamera
- **Thermokamera**
 - I3system TE-Q1
- **2D Farbbild-segmentierung**
 - PointRend
 - Contour Matching
- **PointCloud Segmentierung**
 - RGB-D
 - RGB-D+Thermo
- **Klassisch:**
 - ICP + Kalman-Filter Tracking
 - Keypoint+ ICP
- **Machine Learning:**
 - **Pixel2Mesh**
 - **Mesh RCNN**
 - **ObMan (AtlasNet)**
- **Analytisch:**
 - Online
 - **Approximativ**
 - Affordances
- **Machine Learning:**
 - GG-CNN
- **Dynamisch:**
 - ArmPilot
 - Grasp-Objective
- **Statisch:**
 - ROS MoveIt
- **Hardware:**
 - GraspBox (Kinova Arm)
 - Tiago (5 Finger Hand, Parallel-Greifer)
 - Scitos X3 (Kinova Arm)

Nicht bis zur Integration auf Roboter gelangt

Datenverarbeitungspipeline (b) Nutzung des ObMan Verfahrens

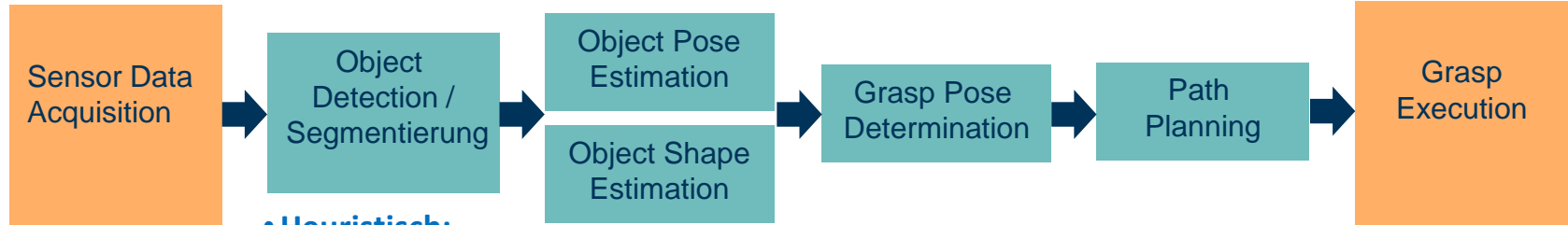


- GFE und TUI NI&KR haben zusammen dieses Verfahren untersucht

[1] Hasson, Yana, et al. : *Learning joint reconstruction of hands and manipulated objects*. In: CVPR. (2019)

Integration SoTA 2D Segmentierungsverfahren (c)

- Leistungsfähige 2D Segmentierungsverfahren sollten auch erprobt werden
- Möglichkeiten der Objekterkennung integriert



• RGB-D

Kameras:

- Kinect4Azure
- AstraOrbbec
- ASUS Xtion
- IOF Kamera

• Thermokamera

- I3system TE-Q1

• Heuristisch:

- Background Subtraction
- Planefitting
- 2D Farbbild-segmentierung
- PointRend
- Contour Matching
- PointCloud Segmentierung
- RGB-D
- RGB-D+Thermo

• Klassisch:

- ICP + Kalman-Filter Tracking
- Keypoint+ ICP

• Machine Learning:

- Pixel2Mesh
- Mesh RCNN
- ObMan (AtlasNet)

• Analytisch:

- Online
- Approximativ
- Affordances

• Machine Learning:

- GG-CNN

• Dynamisch:

- ArmPilot
- Grasp-Objective

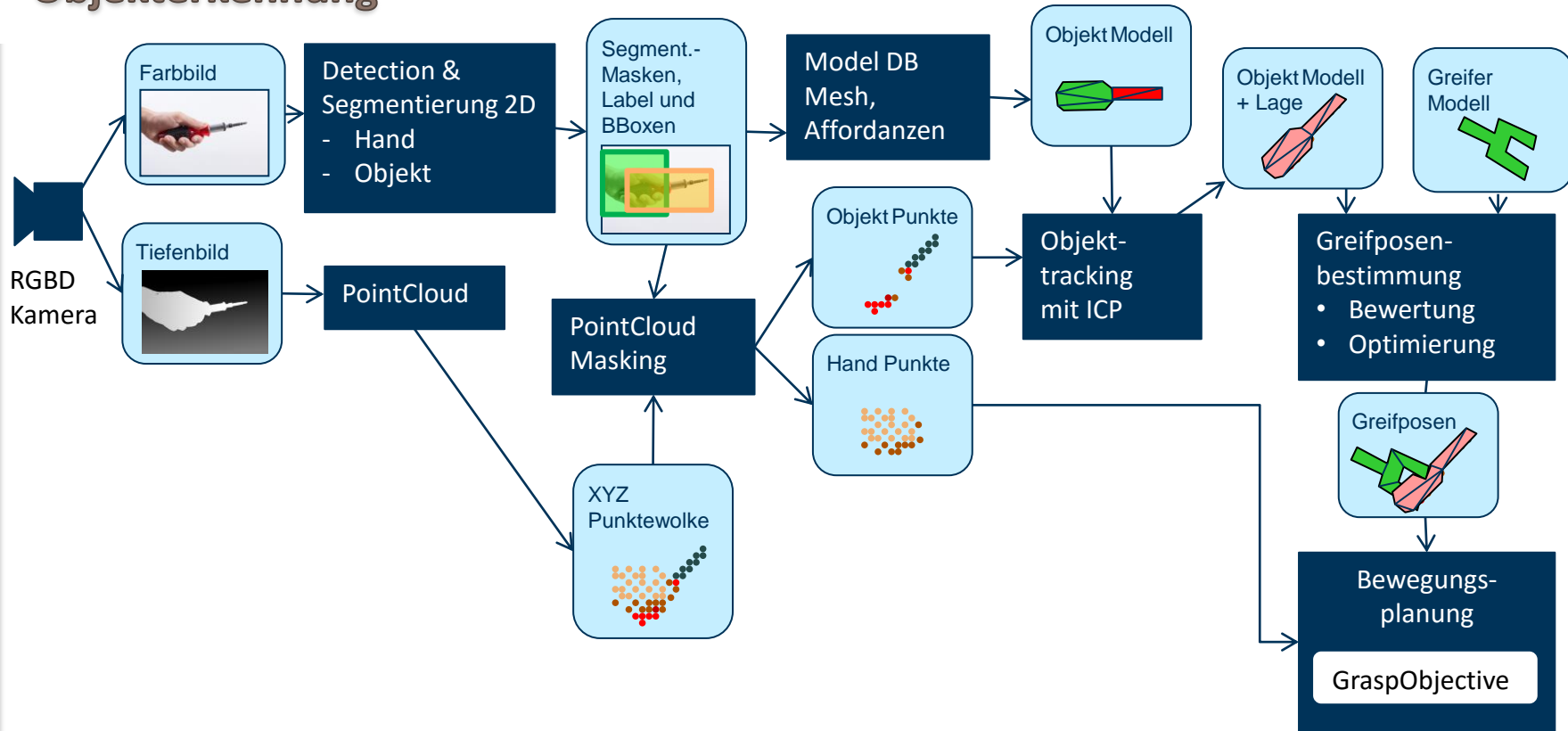
• Statisch:

- ROS MoveIt

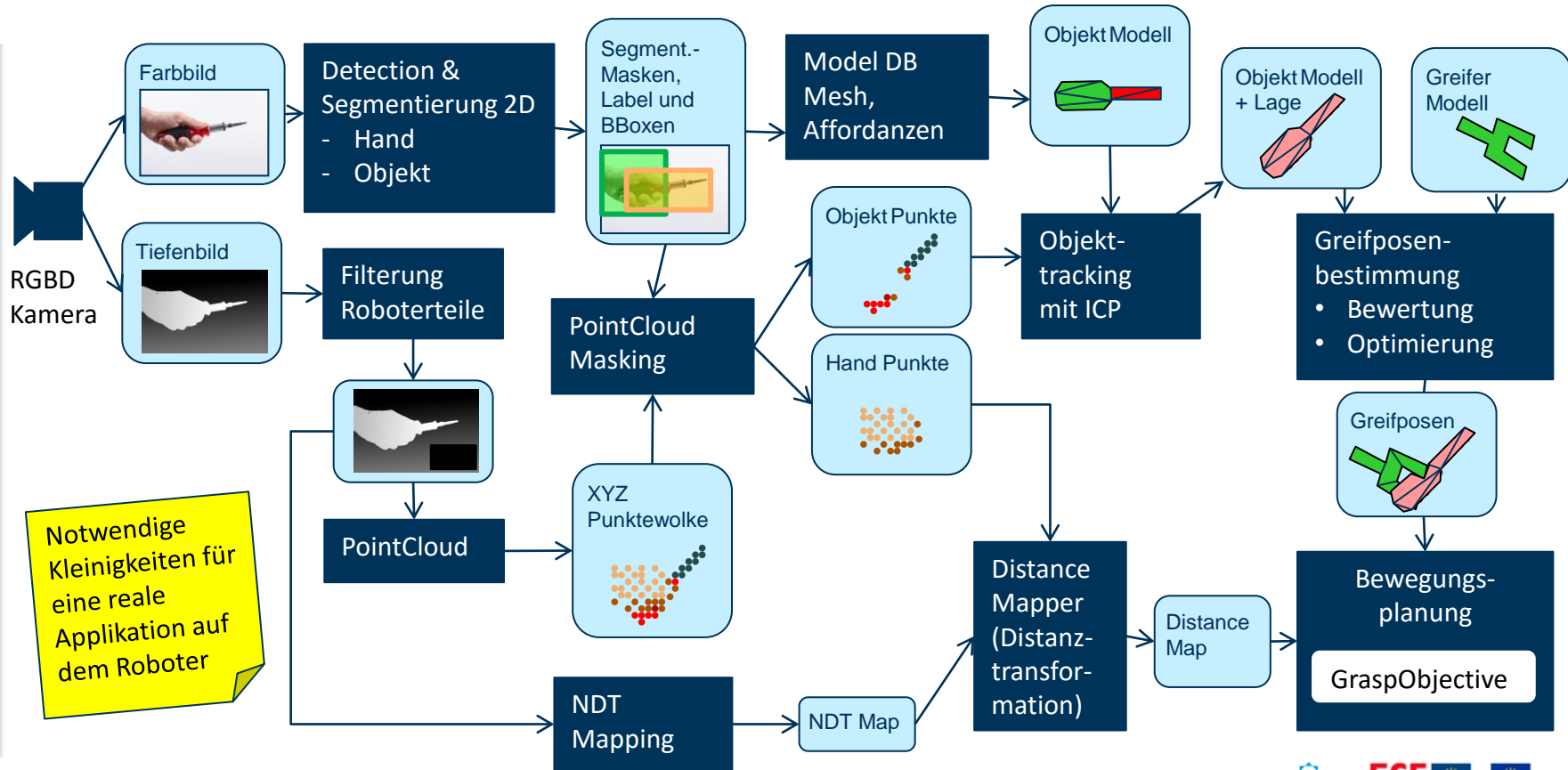
• Hardware:

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

Datenverarbeitungspipeline (c) 2D Segmentierung mit Objekterkennung



Datenverarbeitungspipeline (c) integriert auf Robotern



Notwendige Kleinigkeiten für eine reale Applikation auf dem Roboter

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Beiratstreffen am 20.04.2021

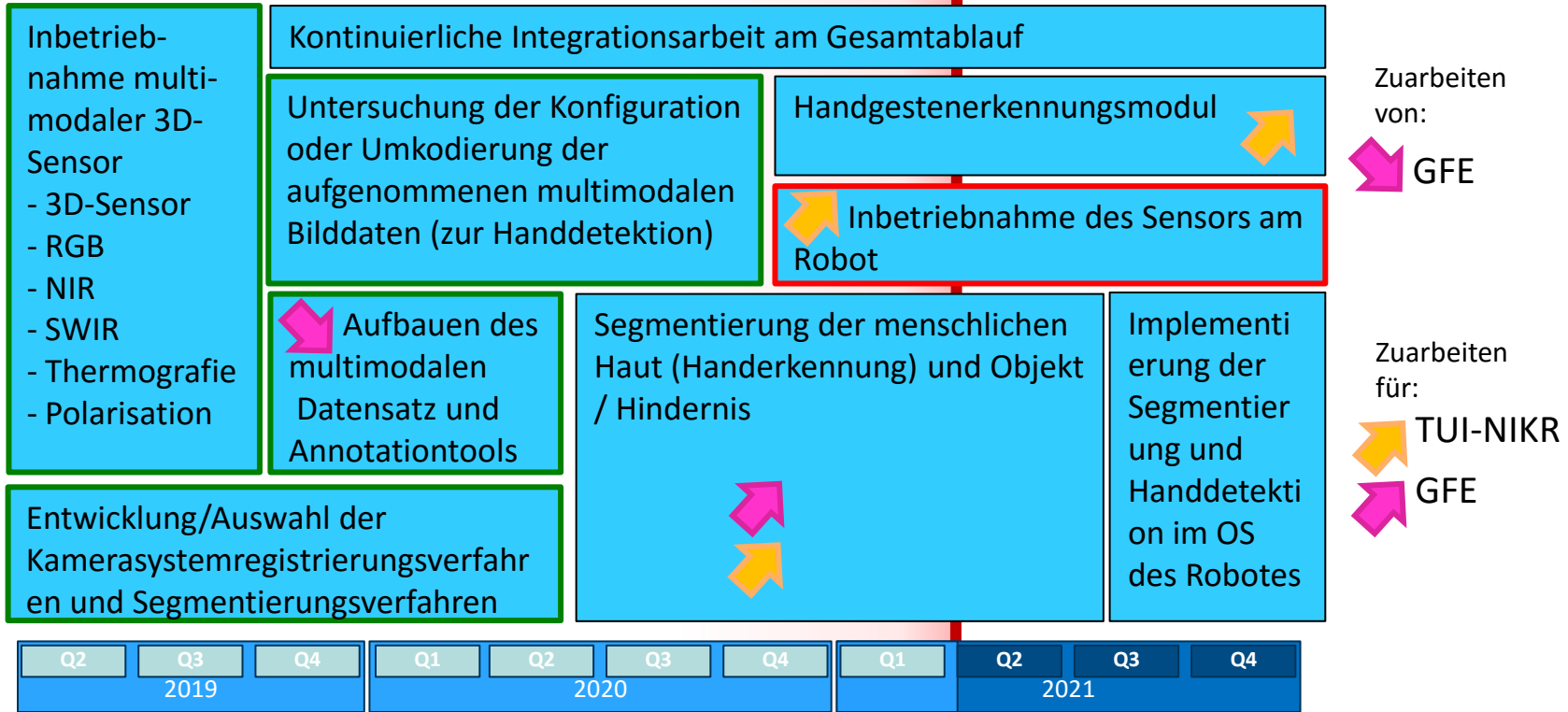
Vorstellung der Arbeiten von TU-Ilmenau QBV

M.Sc. Yan Zhang

www.sonaro-projekt.de

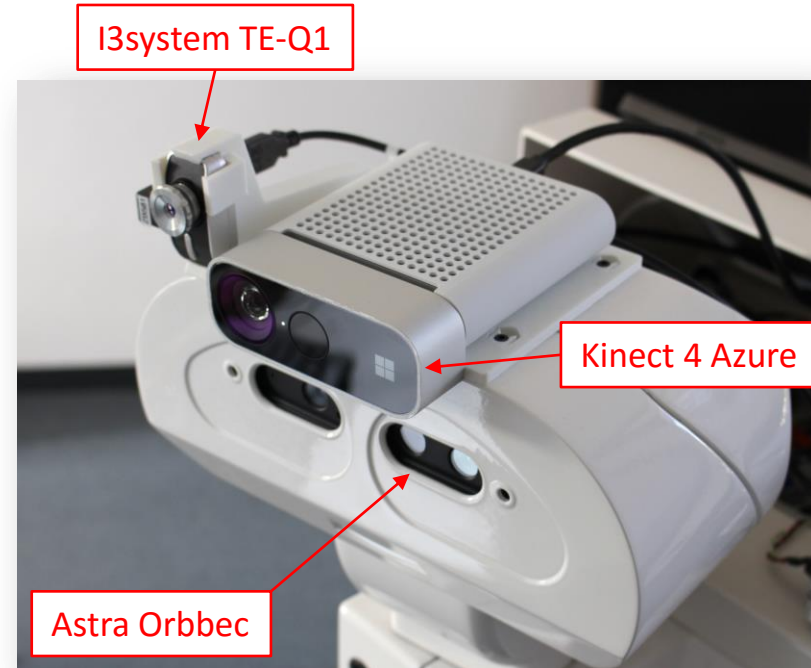


Arbeitsplan für TUI-QBV über die Projektlaufzeit



Integration des multimodalen 3D-Sensorsystems auf dem Tiago-Roboter

- **Stand zum letzten Workshop:**
 - Roboter wurde ausgestattet mit:
 - Kinect 4 Azure -> X, Y, Z, R, G, B
 - I3Systems TE-Q1 Thermal Camera (384x288, 8-14 μ m) -> T
 - Astra Orbbec (ASTRA S) -> X, Y, Z, R, G, B, B
 - Registrierung aller Bilder für die Berechnung einer multimodalen Punktwolke (X, Y, Z; R, G, B; T)



Datensatz für Multimodal Point Cloud Segmentierung

- **Erreichter Stand**

- Datensatz zur **Evaluation** der 3D Segmentierungsverfahren aus SONARO HandObject Datensatz generiert (X, Y, Z; R, G, B; T)
 - Kombination von zufälligen Hintergründen und Hand-Objekt Punktwolken
 - Train-Datensatz: 10024
 - Validation-Datensatz: 1726
 - Test-Datensatz: 1136 (Bekannte Objekte)
 - Test-Datensatz: 1129 (Unbekannte Objekte)



Vergleich zwischen 3D-Sensoren

Sensor Data Acquisition

• RGB-D

Kameras:

– Kinect4Azure

– AstraOrbbec

– ASUS Xtion

– IOF Kamera

• Thermokamera

– I3system TE-Q1

• Vergleichsexperimente zwischen drei 3D-Sensoren:

Motivation: Es gibt viele alternative 3D-Sensoren. Die Vor- und Nachteile sowie Eigenschaften der Sensoren sollen mittels dieser Experimente zusammengefasst werden. Vorschlag zum Verwendungsbereich oder -zweck jedes Sensors wird gegeben.

- **Experiment 1:** Evaluation der durchschnittlichen Punktpositionsabweichung der aufgenommenen Punktwolke für Objekte aus unterschiedlichen Materialien
- **Experiment 2:** Evaluation der Punktwolkeaufnahme für geneigte Flächen

Segmentierung

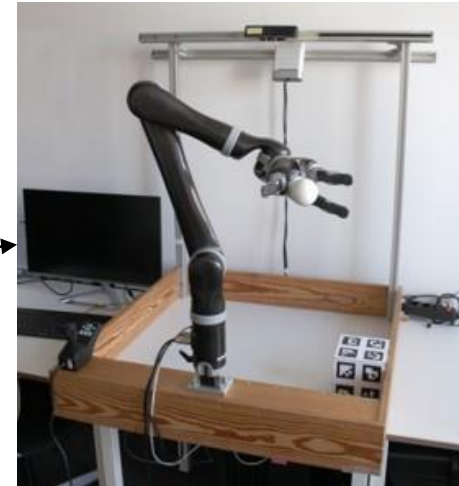
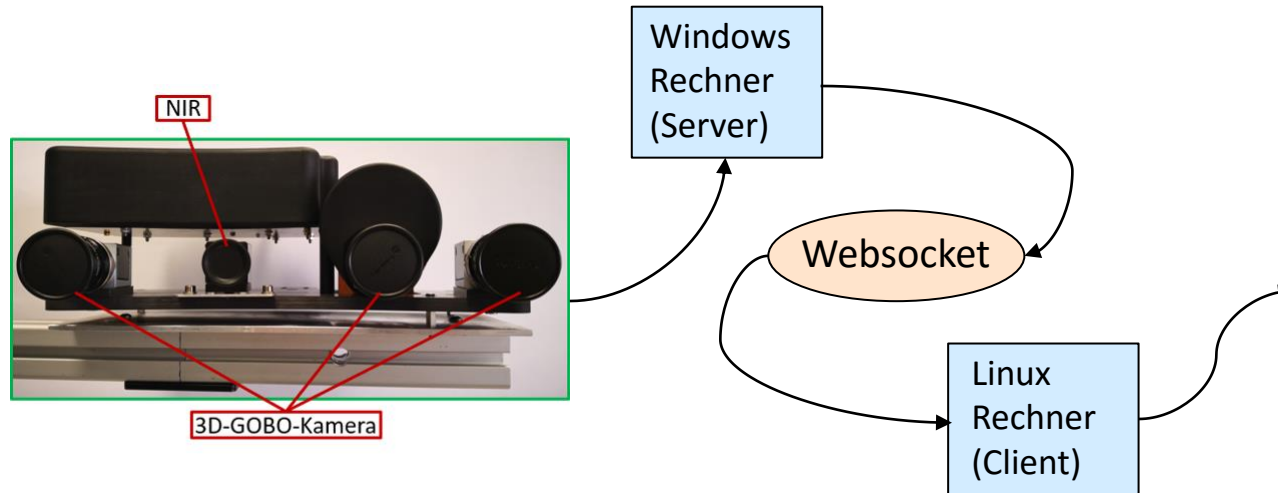
– RGB-D

– RGB-D+Thermo

– ObMan
(AtlasNet)

Vergleich von 3D-Sensoren Testplattform (Graspbox)

- **IOF-GOBO-M, ASUS-Xtion** und **Kinect-Azure** wurden an der Testplattform montiert und kalibriert.
- Die Punktwolken Daten von IOF-GOBO-M werden mittels WebSocket von einem Windows-PC an den Linux-PC übertragen.
- Vergleichsexperimente zwischen den drei 3D-Sensoren wurden durchgeführt.



Vergleich zwischen IOF-Kamera, ASUS und Kinect

- Technische Daten von IOF-GOBO-M, ASUS-Xtion und Kinect-Azure

	IOF-GOBO-M	ASUS-Xtion	Kinect-Azure
Pixelzahl pro 3D-Bild	1024x1024 (variabel)	640x480	1024x1024
3D-Datenrate	36 fps	30 fps	15 fps
Öffnungswinkel *	H: 48°, V: 44°	H: 116°, V: 90°	H: 120°, V: 120°
Nominale Messgenauigkeit	0,1mm	ca. 2 mm [1]	1,5 mm [2]
Messfeld	0,4 – 1,0 m (variabel)	0,8 – 3,5 m	0,25 – 2,21 m

* H: Horizontal, V: Vertikal

[1] Rauscher, G., Dube, D., & Zell, A. (2016). A comparison of 3d sensors for wheeled mobile robots. In Intelligent Autonomous Systems 13 (pp. 29-41). Springer, Cham.

[2] Tölgyessy, M., Dekan, M., Chovanec, L., & Hubinský, P. (2021). Evaluation of the Azure Kinect and Its Comparison to Kinect V1 and Kinect V2. Sensors, 21(2), 413.

Experiment 1

- **Durchschnittliche Punktpositionsabweichung der aufgenommenen Punktwolke für Objekt aus unterschiedlichen Materialien**



Lackierte Holzplatte



Buchcover



Schneidbrett (Kunststoff)



Rückseitige bedruckte Glasplatte



Alufolie



Papier



Stoff (Textil)



Schaumstoff



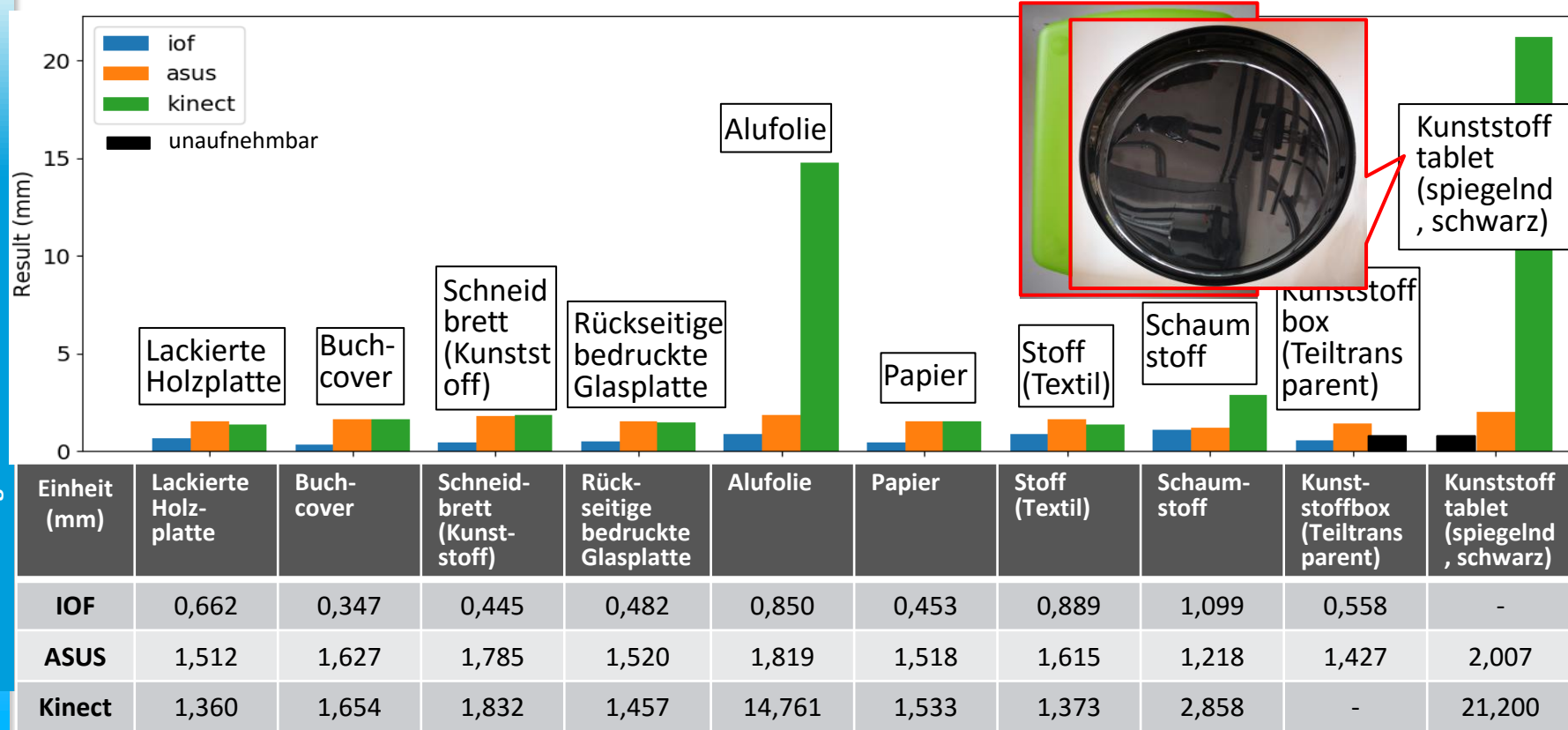
Kunststoffbox (Teiltransparent)



Kunststofftablet (spiegelnd, schwarz)

Experiment 1

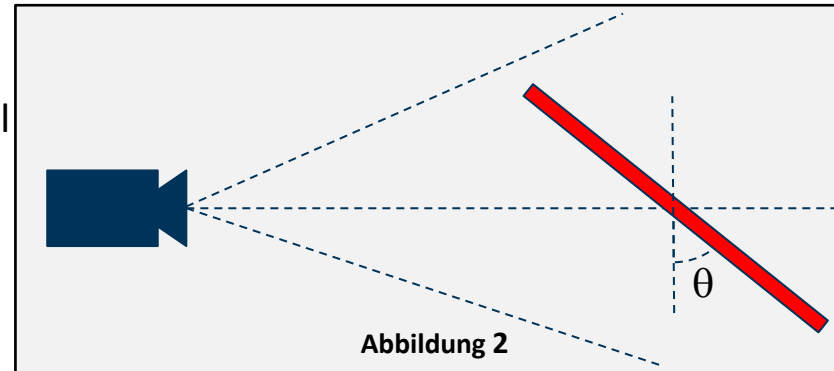
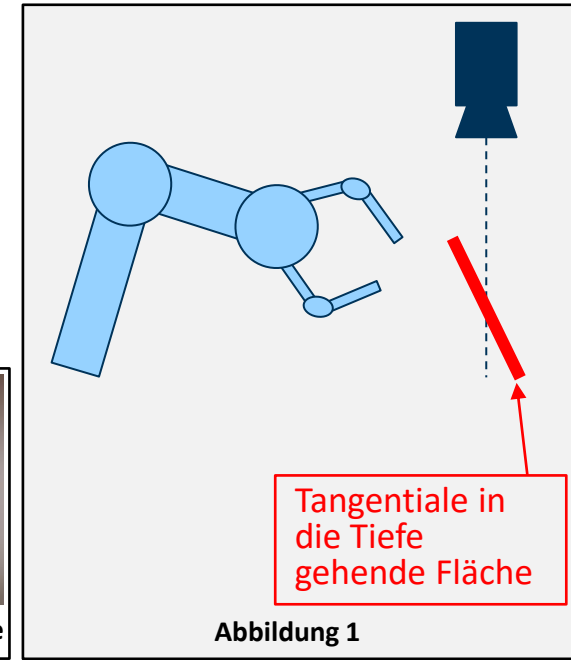
- Durchschnittliche Punktpositionsabweichung der aufgenommenen Punktwolken für Objekte aus unterschiedlichen Materialien



Experiment 2

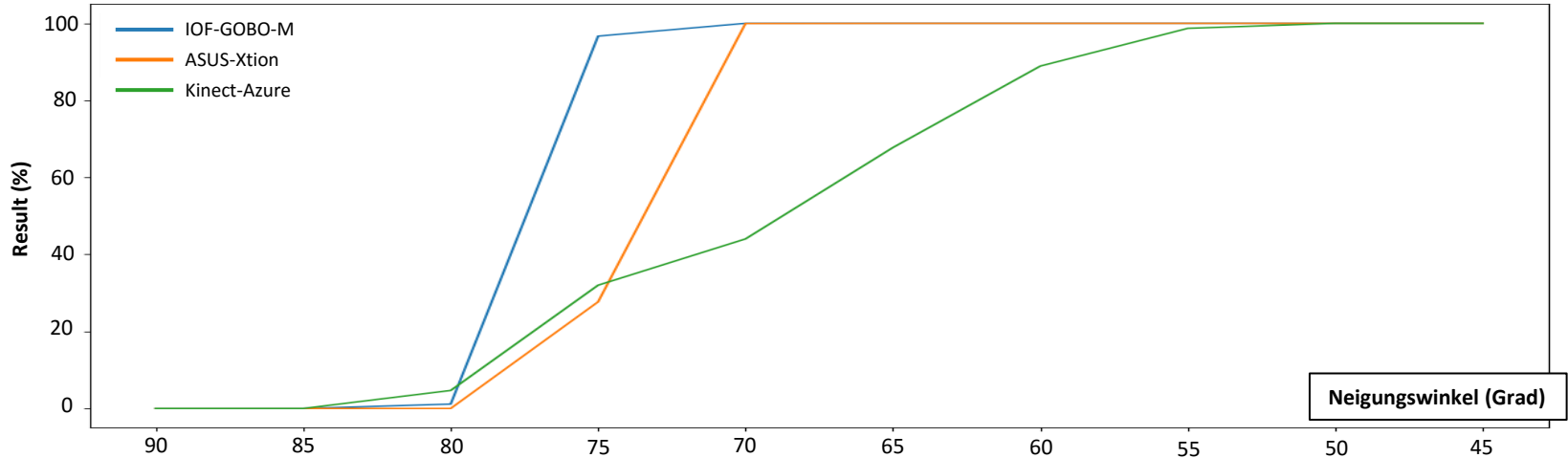
• Punktwolkenaufnahme geneigter Flächen

- Motivation:** Für Singleview-3D-Sensor, gibt es immer ein Problem, dass geneigte Fläche (tangente in die Tiefe gehende Fläche) des Objekts schwierig von 3D-Sensor aufgenommen werden kann. (Siehe Abbildung 1)
- Ziel des Experiments:** Evaluation der Punktwolkenaufnahme der Sensoren für die tangente Fläche.
- Eine Platte (Lackierte Holzplatte) (siehe Abbildung 2) wird bei unterschiedlichen Neigungswinkeln θ von 90° bis 45° gemessen.
- Auswertungskriterium:** Quotient von Anzahl der aufgenommenen Punkte und der theoretischen Max-Punktzahl.
- $$Result = \frac{Ist-Anzahl}{Soll-Anzahl}$$



Experiment 2

- Punktwolkenaufnahme für unterschiedliche Neigungswinkel (Objekt: Lackierte Holzplatte)



	90°	85°	80°	75°	70°	65°	60°	55°	50°	45°
IOF	0%	0%	1,13%	96,68%	100%	100%	100%	100%	100%	100%
ASUS	0%	0%	0%	27,68%	99,95%	100%	100%	100%	100%	100%
Kinect	0%	0%	4,67%	32%	44,02%	67,74%	88,9%	96,9%	100%	100%

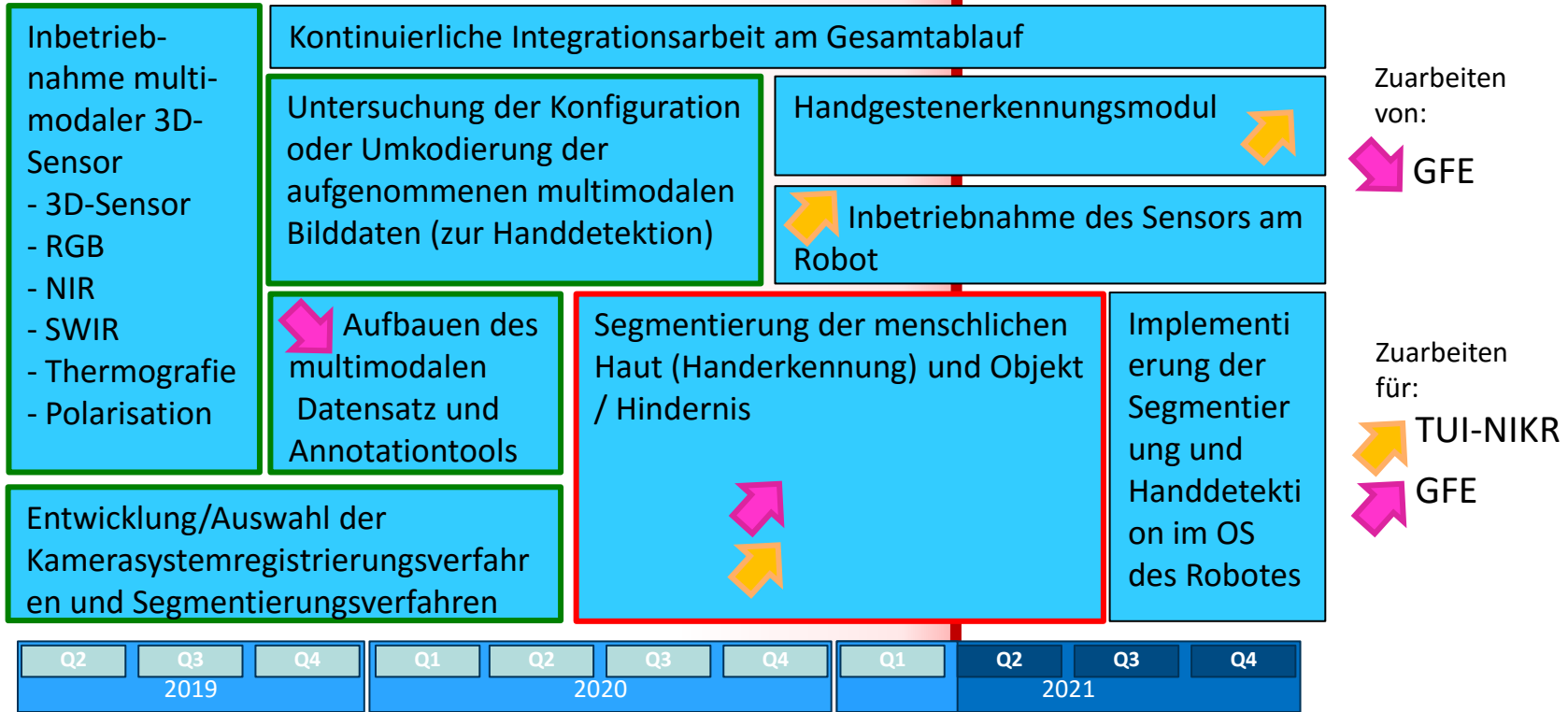
Vergleich zwischen IOF-Kamera, ASUS und Kinect

	Stereo Camera		Time-of-Flight Camera
	IOF-GOBO-M	ASUS-Xtion	Kinect-Azure
Vorteile	<ul style="list-style-type: none"> - Hohe Ortauflösung & Bildauflösung - Hohe Genauigkeit - Vollständige Erfassung bei größtem Neigungswinkel von Flächen kann erreicht werden (bis 75°) 	<ul style="list-style-type: none"> - Dunkle und spiegelnde Objekte können robuster aufgenommen werden. - Opensource SDK (OpenNI unterstützt) 	<ul style="list-style-type: none"> - Großer Öffnungswinkel & Arbeitsbereich - Leistungsstarkes SDK (z.B: Menschliches Skelett erkennen) - Hohe Bildauflösung
Nachteile	<ul style="list-style-type: none"> - Kleiner Öffnungswinkel & Arbeitsbereich - Hohe Leistungsaufnahme - SDK unterstützt nur Windows 	<ul style="list-style-type: none"> - Niedrige Bildauflösung - Niedrige Ortauflösung - Niedrige Genauigkeit 	<ul style="list-style-type: none"> - Niedrige Genauigkeit - Problem für teiltransparentes Objekt - Unscharfe Kanten - Problem bei geneigten Flächen - Vollständige Erfassung nur bis zu einem Neigungswinkel von bis zu 50° möglich

Vergleich zwischen IOF-Kamera, ASUS und Kinect

- **Zusammenfassung:**
 - **IOF-GOBO-M:**
 - IOF-3D-Sensor hat sehr gute Genauigkeit aber kleineren Arbeitsbereich und ist somit für stationäre Roboter mit Feinarbeit besonders geeignet.
Für ein mobiles System wegen des großen Strombedarfs nicht geeignet.
 - **Kinect-Azure:**
 - Kinect-Azure hat großen Arbeitsbereich und Öffnungswinkel, aber niedrige Genauigkeit. Deshalb ist Kinect-Azure geeignet für grobe Umgebungserkennung oder -verständnis einer große Szene.
 - **ASUS-Xtion:**
 - Im Gegensatz zu den anderen zwei Kameras können viele dunkle und spiegelnde Objekte bei mittlerer Genauigkeit aufgenommen werden. Deswegen ist ASUS-Xtion geeignet, an einem mobilen Roboter eingesetzt zu werden, um einfache Aufgaben zu erledigen.

Arbeitsplan für TUI-QBV über die Projektlaufzeit



Motivation Segmentierung mit multimodalen 3D-Daten



• RGB-D Kameras:

- Kinect4Azure
- AstraOrbbec
- ASUS Xtion
- IOF Kamera

• Thermokamera

- I3system TE-Q1

• Heuristisch:

- Background Subtraction
- Planefitting
- 2D Farbbildsegmentierung
- PointRend
- Contour Matching

• PointCloud Segmentierung

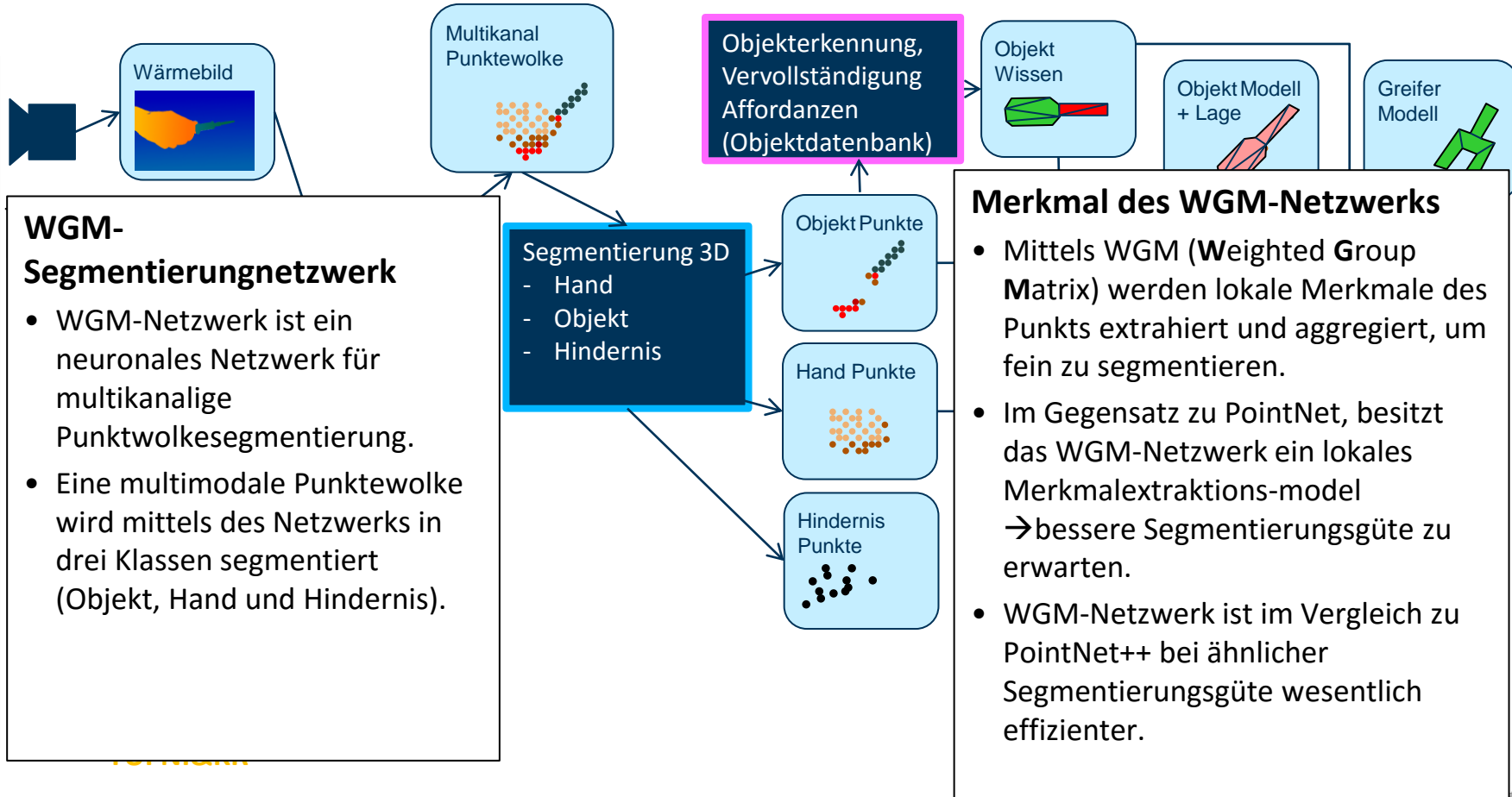
- RGB-D

RGB-D+Thermo

• Segmentierung mit multimodalen Daten

- **Motivation:** Für sichere Hand-Objekt-Segmentierung, gibt es das Problem, dass eine saubere Segmentierung des überlappenden Bereich zwischen Hand und Objekt schwierig ist. Menschen haben eine Temperatur von 37 °C und Objekte weichen davon ab (oft Raumtemperatur ca. 20°C). Deswegen soll eine multimodale Punktwolke (XYZ RGB T) mit Thermobild zur sicheren Hand-Objekt-Segmentierung genutzt werden.

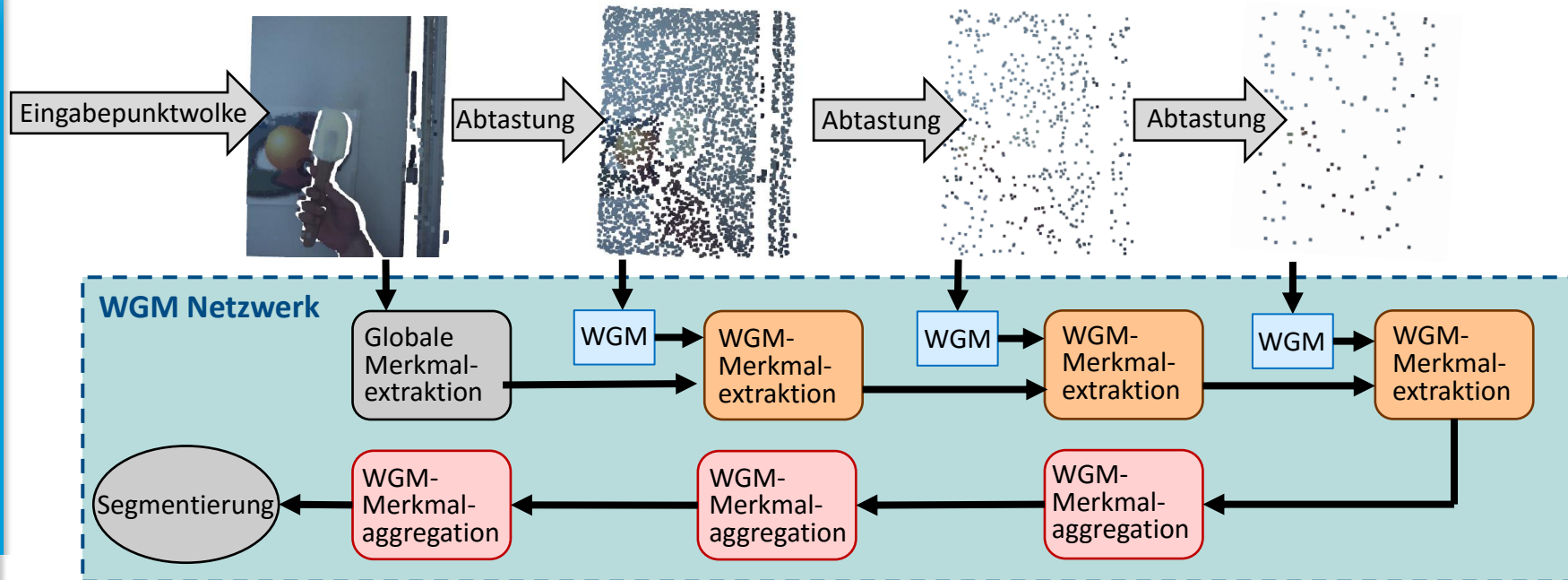
Ansatz zur Segmentierung: WGM-Netzwerk



Segmentierung: WGM-Netzwerk

- **Stand zum letzten Workshop:**

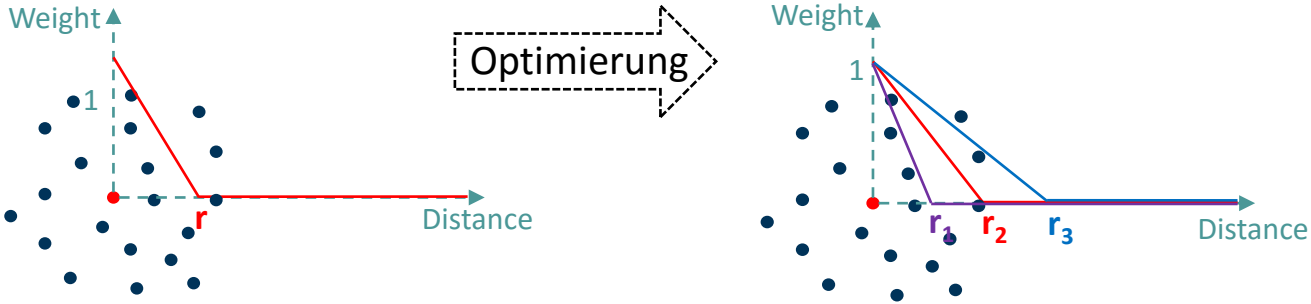
- Eingabepunktwolke wird homogen in drei verschiedenen Auflösungen abgetastet
- Segmentierung wird mit der Hilfe von **WGM-Merkmalextraktion** und **WGM-Merkmalaggregation** in jeder abgetasteten Punktwolke erledigt.



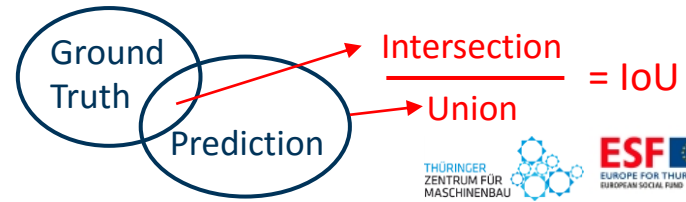
Segmentierung: WGM-Netzwerk

- Erreichter Stand:

- Optimierung der Segmentierungsnetzwerk.
 - Multisearchradius-Mechanismus zur lokale Merkmalextraktion

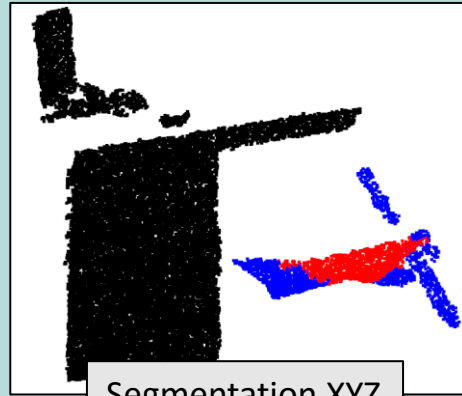
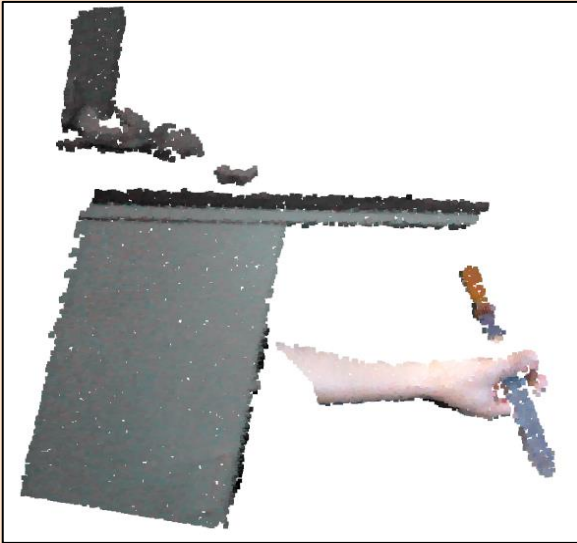


- Evaluation basiert auf eigenem domänenspezifischen Datensatz
- **Trainingphase:** Downsampling der Punktwolken von Datensatz auf gleiche Größe (10k, 7) für Trainingphase.
- **Testingphase:** In Inference- oder Testingphase wird die Punktwolke mit originalen Größe eingegeben.
- **Auswertungskriterium:** mean IoU (Quotient von Intersection und Union.)

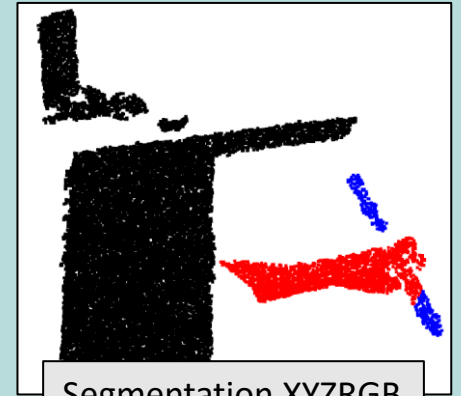


Segmentierung: WGM-Netzwerk

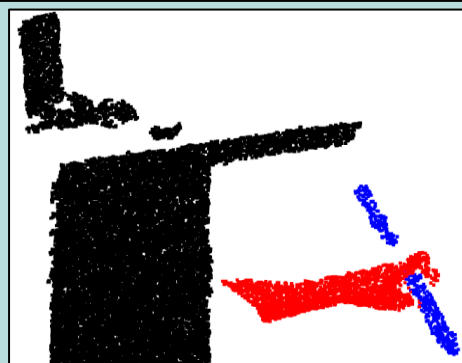
- Vergleichen zwischen XYZRGB-T, XYZRGB und XYZ (Test-Datensatz bekanntes Objekt: Feile)



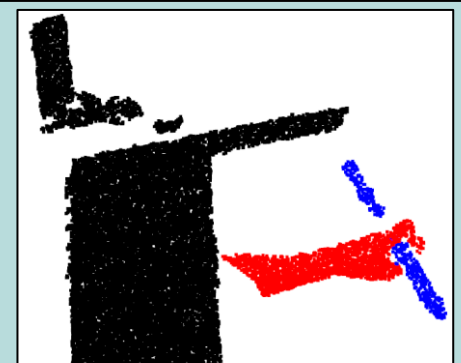
Segmentation XYZ



Segmentation XYZRGB



Segmentation XYZRGB-T



Ground-Truth

	Testdatensatz (Bekanntes Objekt)
XYZRGB-T	94,45%
XYZRGB	86,67%
XYZ	66,70%

Mean IoU: mean Intersection over Union

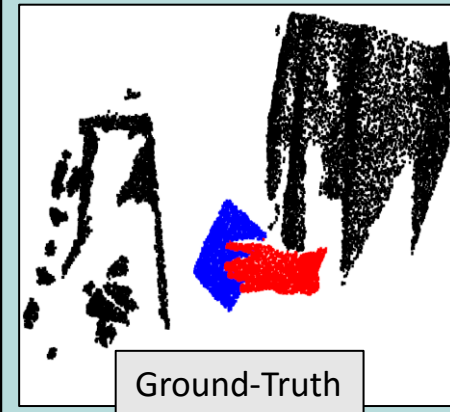
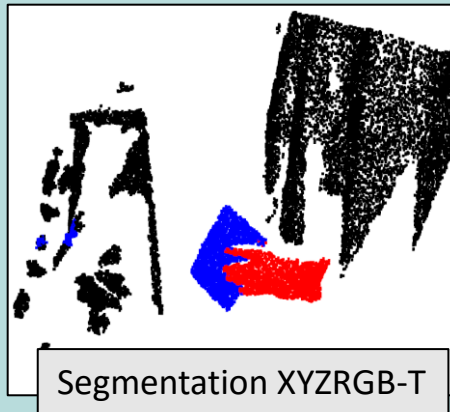
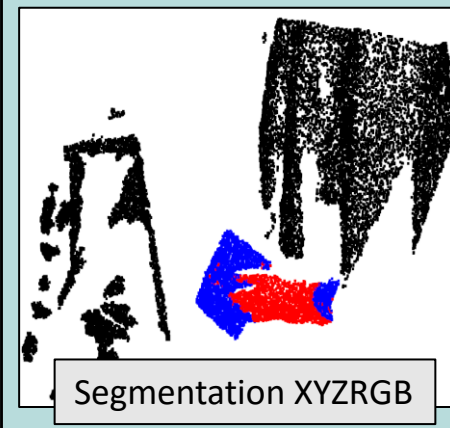
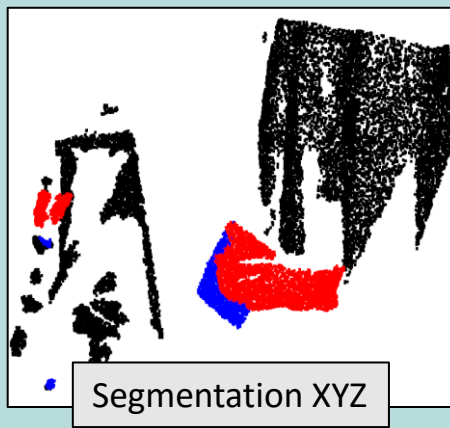
Segmentierung: WGM-Netzwerk

- Vergleichen zwischen XYZRGB-T, XYZRGB und XYZ (Test-Datensatz unbekanntes Objekt: CD-Box)

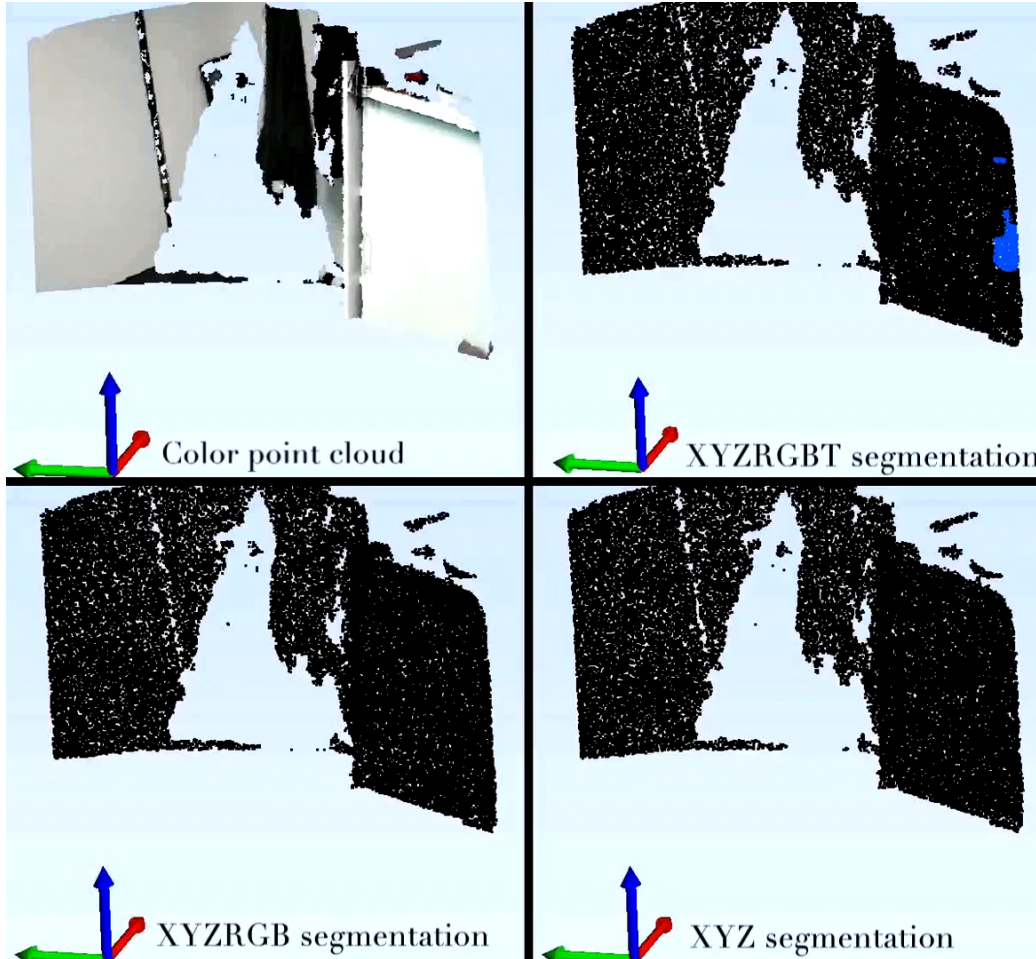


Testdatensatz (Unbekanntes Objekt)	
XYZRGB-T	94,27%
XYZRGB	85,62%
XYZ	58,39%

Mean IoU: mean Intersection over Union



Videobeispiel Segmentierung (unbekante Objekte)



**OBJECT:
Bottle**

Red points:
Hand

Blue points:
Object

Black points:
Background

Segmentierung auf multimodalen 3D-Daten

- **Zusammenfassung:**

- Die Erhöhung der Segmentierungsleistung mit Hilfe der multimodalen 3D-Daten (XYZRGB-T) ist deutlich.
- Gute Generalisierungsfähigkeit auf unbekannte Objekte
- Abfall der IoU zwischen bekannten Testobjekten und Unbekannten:
 - **XYZRGB-T**: fast keine Reduzierung (c.a. 0,2%)
 - **XYZRGB**: kleine Reduzierung (c.a. 1%)
 - **XYZ**: große Reduzierung (c.a. 10%)

	Testdatensatz (Bekannte Objekte)	Testdatensatz (Unbekannte Objekte)
XYZRGB-T	94,45%	94,27%
XYZRGB	86,67%	85,62%
XYZ	66,70%	58,39%

Mean IoU: mean Intersection over Union

- Sowohl für bekannte Objekte als auch unbekannte Objekten haben multimodale 3D-Daten immer die beste Segmentierungsleistung.
- Multimodalen 3D-Daten hat bessere Robustheit bei unterschiedlichem Umgebungslicht.

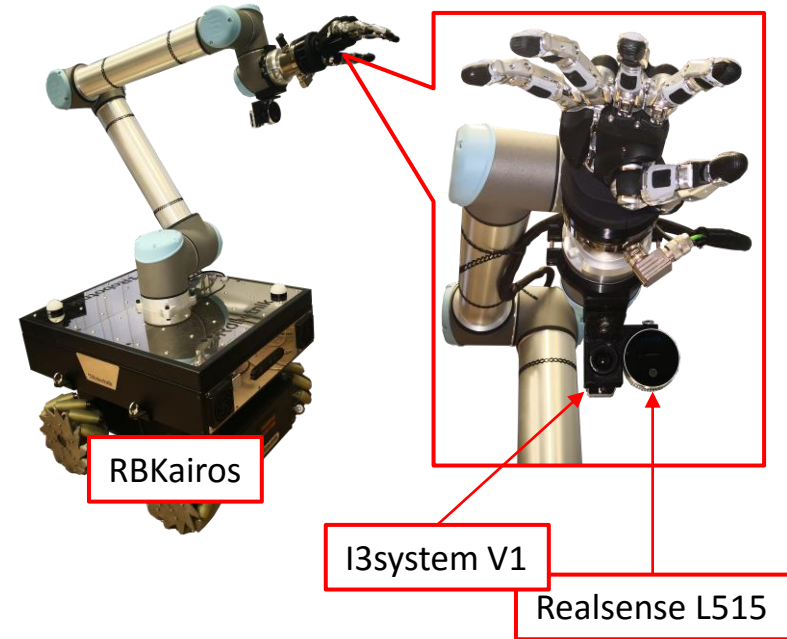
Zusammenfassung

- **An der Testplattform (Graspbox) wurde der Vergleich zwischen drei 3D-Sensoren (Kinect-Azure, ASUS-Xtion und IOF-GOBO-M) durchgeführt.**
- **Analyse zum Anwendungsbereich der drei 3D-Sensoren im Bereich von Robotik durchgeführt.**
- **Optimierung des WGM-Segmentierungsverfahrens.**
- **Evaluation des Segmentierungsverfahrens wurde durchgeführt. Der Vorteil der Verwendung der multimodalen 3D-Daten konnte gezeigt werden.**

Kameraposition am Roboter

- **Statisch extern:**
 - **Bsp. Tiago:** Sensorik am Kopf
 - Multimodales 3D-Sensorsystem:
 - I3system Q1 (Thermo)
 - Orbbec Astra (RGB-D)
 - Nachteil: Verdeckungen durch Roboterarm

- **Mitbewegt am Roboterarm:**
 - **Bsp. RBKairos:** Sensorik an Greifhand
 - Multimodales 3D-Sensorsystem:
 - I3system V1 (Thermo);
 - Realsense L515 (RGB-D)
 - Nachteil: kleiner Wahrnehmungsbereich im Nahbereich



Ausblick auf das letzte Projekthalbjahr

Projektziele	Offene Punkte (todos)	Zusätzliche erwartete Ergebnisse bis Projektende
Robuste Trennung von Objekt und Hand	<ul style="list-style-type: none"> • Verbesserung der Segmentierungsleistung 	<ul style="list-style-type: none"> • Erweiterung der Untersuchung zur Nutzung multimodaler Daten auf 2D Segmentierung (NIKR, QBV)
Konzeption und Evaluation multimodaler Tiefensensoren		<ul style="list-style-type: none"> • Vergleich von externer zu mitbewegter Sensorik im Greifer
Integration auf Roboterplattform für Demonstrator	<ul style="list-style-type: none"> • Integration der 3D Segmentierung in Demonstrator-Pipeline auf Tiago 	<ul style="list-style-type: none"> • Integration auf RBKairos



Forscherguppe SONARO

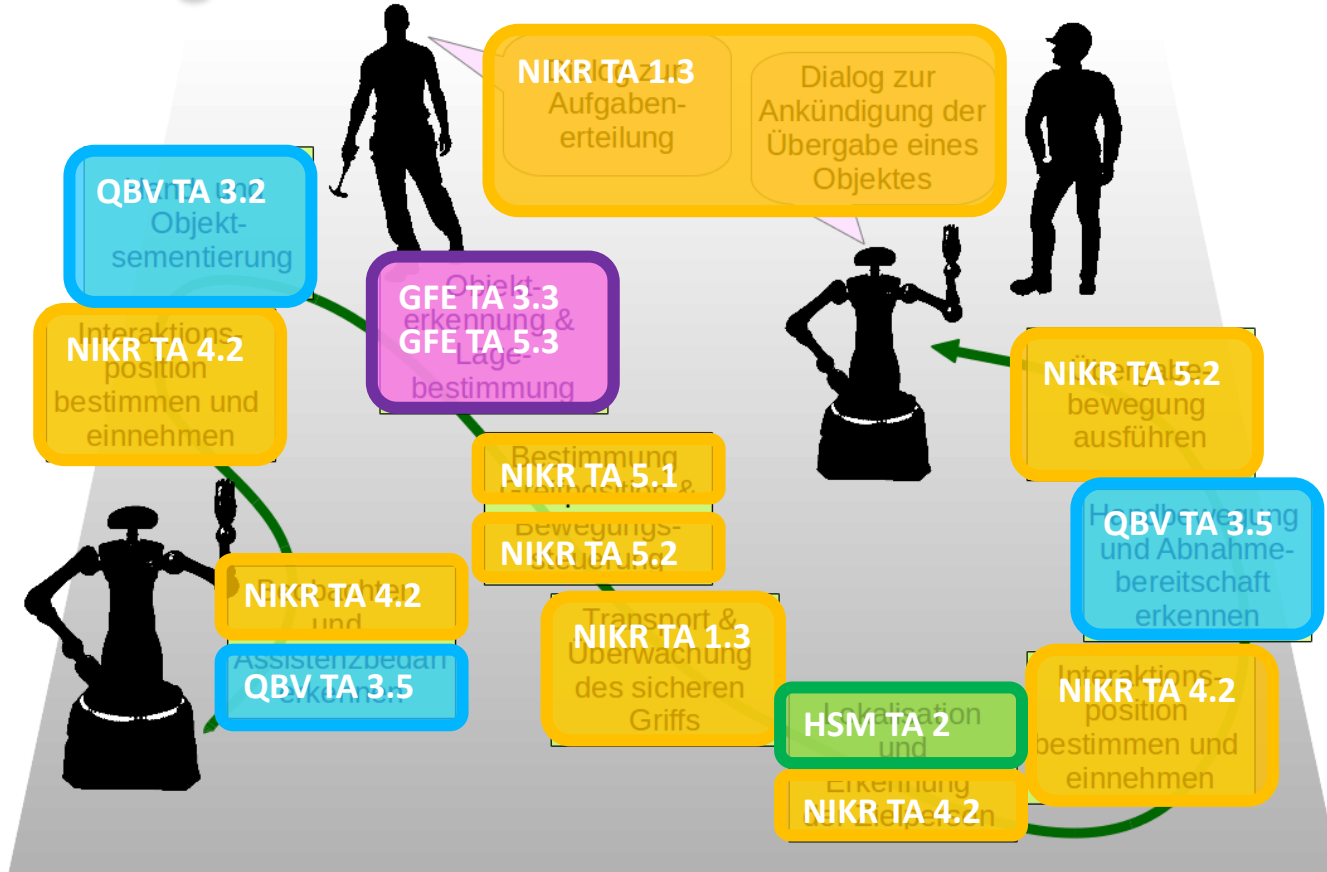
Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Beiratstreffen am 20.04.2021

Vorstellung der Arbeiten von GFE Schmalkalden e. V.

Dr. Daniel Garten

Einordnung der GFE in SONARO



Arbeitsplan für GFE über die Projektlaufzeit

TA 1.1 Spezifikation

TA 1.4 Kontinuierliche Integrationsarbeit am Gesamtablauf und Test

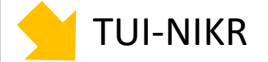


TA 3.3 Anlernen von Objekten durch Präsentation / haptische Exploration

- Anlegen Objektdatenbank
- Entwicklung der Benutzeroberfläche
- Umsetzung eines Online-Trainingsmechanismus für die lebenslange Erweiterung der Wissensbasis des initialen Klassifikators



Zuarbeiten von:

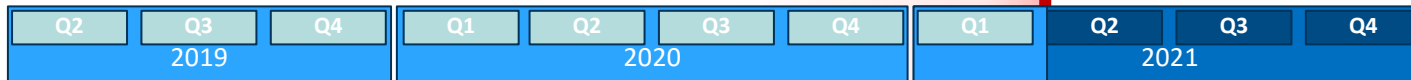
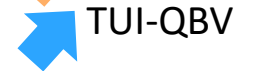


TA 5.3 Objekttracking,-klassifikation und Lageerkennung

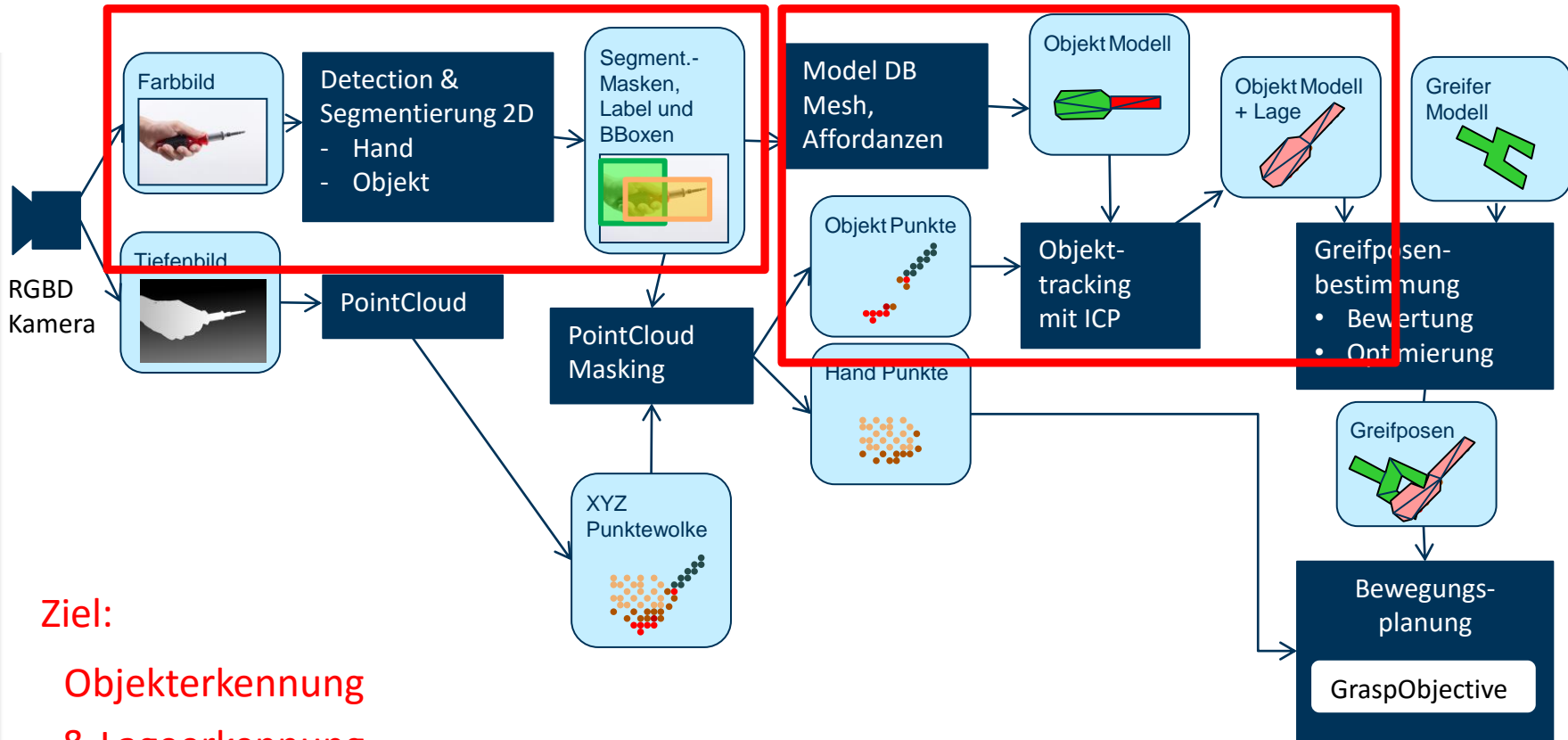
- Untersuchung und Auswahl geeigneter Klassifikatoren
- Erstellung eines parametrischen Modells für den zu greifenden Gegenstand



Zuarbeiten für:



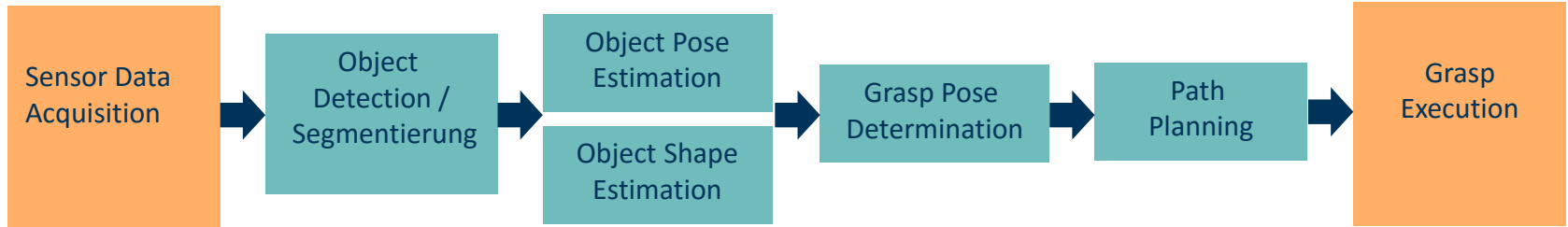
Datenverarbeitungspipeline



Ziel:

Objekterkennung
& Lageerkennung
für spätere Griffposenbestimmung

Datenverarbeitungspipeline mit verwendeten Verfahren



• RGB-D

Kameras:

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

• Thermokamera

- I3system TE-Q1

• Heuristisch:

- Background Subtraction
- Planefitting

• 2D Farbbild-segmentierung

- PointRend

- Contour Matching

• PointCloud Segmentierung

- **RGB-D**
- RGB-D+Thermo

• Klassisch:

- ICP + Kalman-Filter Tracking

- Keypoint+ ICP

• Machine Learning:

- Pixel2Mesh
- Mesh RCNN
- ObMan (AtlasNet)

• Analytisch:

- Online
- Approximativ
- Affordances

• Machine Learning:

- GG-CNN

• Dynamisch:

- ArmPilot
- Grasp-Objective

• Statisch:

- ROS MoveIt

• Hardware:

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

Untersuchte Verfahren zur modellbasierten Objektdetektion



Es wurden auf Basis der industriellen Bildverarbeitungsbibliothek HALCON mehrere Ansätze entwickelt und evaluiert

1. Form- und farbbasierte Matching-Verfahren auf Basis eines 2D-Referenzbildes und des aktuellen RGB-Bildes der Roboterumgebung
2. Key-Point-basiertes 3D-Matching auf Basis der 3D Modelle des SONARO HandObject Datensatzes
3. Kombiniertes Verfahren aus Deep-Learning und ICP (Iterative-Closest-Point) auf Basis eines vorher aufgenommenen 3D-Datensatzes der zu detektierenden Objekte sowie des aktuellen RGB- und Tiefenbildes der Roboterumgebung

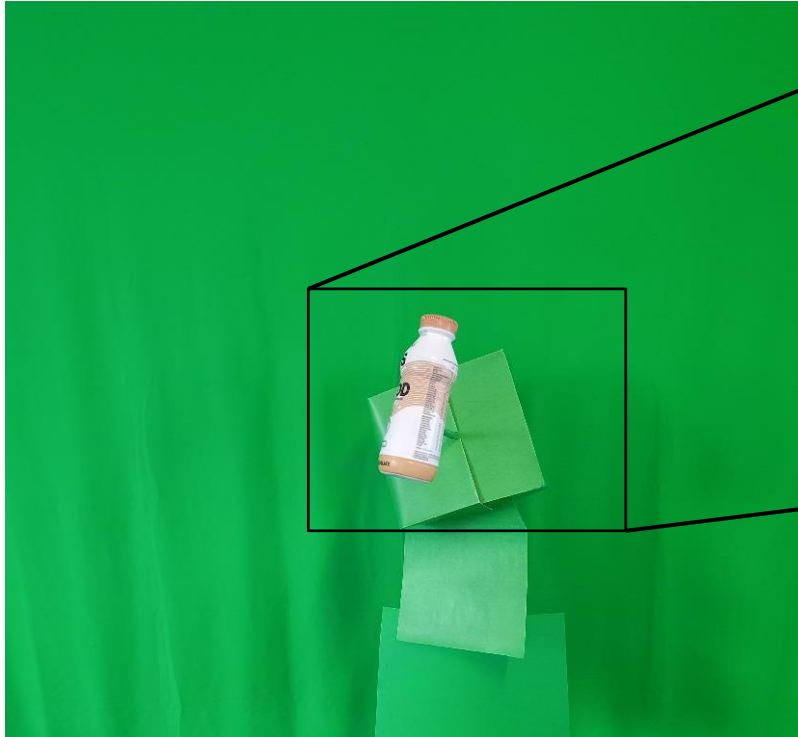
Ansatz zur konturbasierten Detektion



1. Nutzung von Referenzobjekten (aufgenommen vor stabilem Hintergrund - Greenscreen) als Template für 2-D-Matching (kontur- und texturbasiert)
2. Suche des Referenzobjektes im RGB-Bild der Astra-Orbbec oder Kinect-Kamera (Ergebnis: Bildregion des Objektes inklusive verdeckter Bereiche)
3. Lookup der Tiefendaten des Objektes
→ Region of Interest (RoI) in Point Cloud
4. Registrierung der Tiefendaten des Objektes (RoI) mit dem zugehörigen 3D-Modell aus Datensatz

Beispiel Milchflasche – Referenz-Template generieren

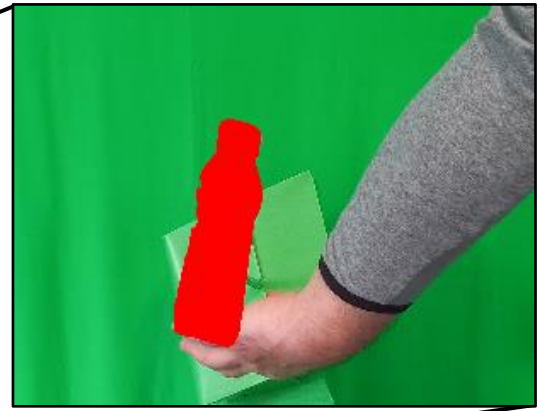
Referenzbild



Segmentierung und
Konturdetektion Referenz-Template

Beispiel Milchflasche – Referenz-Template suchen

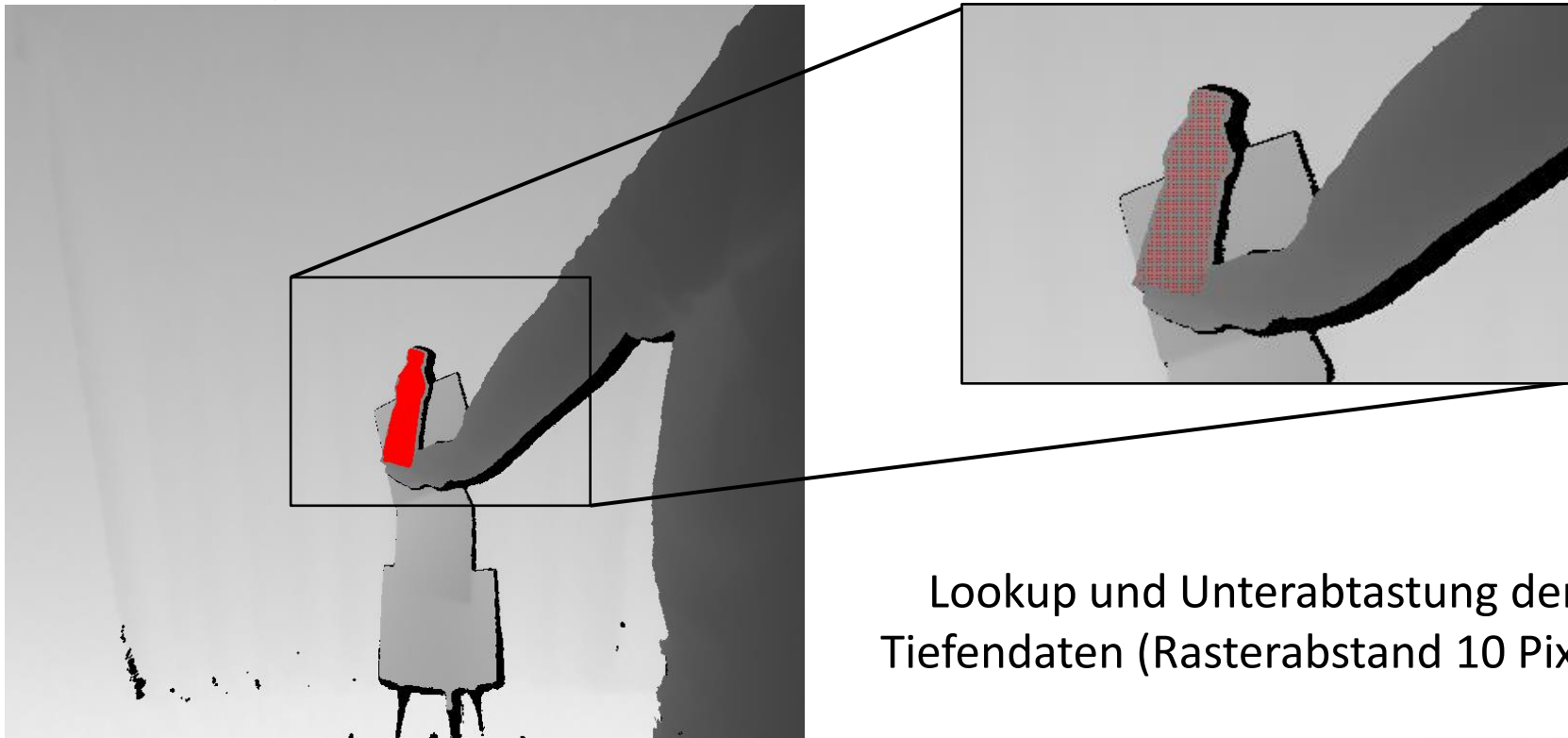
Suchbild



Detektion des Referenztemplates
auf Basis von Kontur-Matching

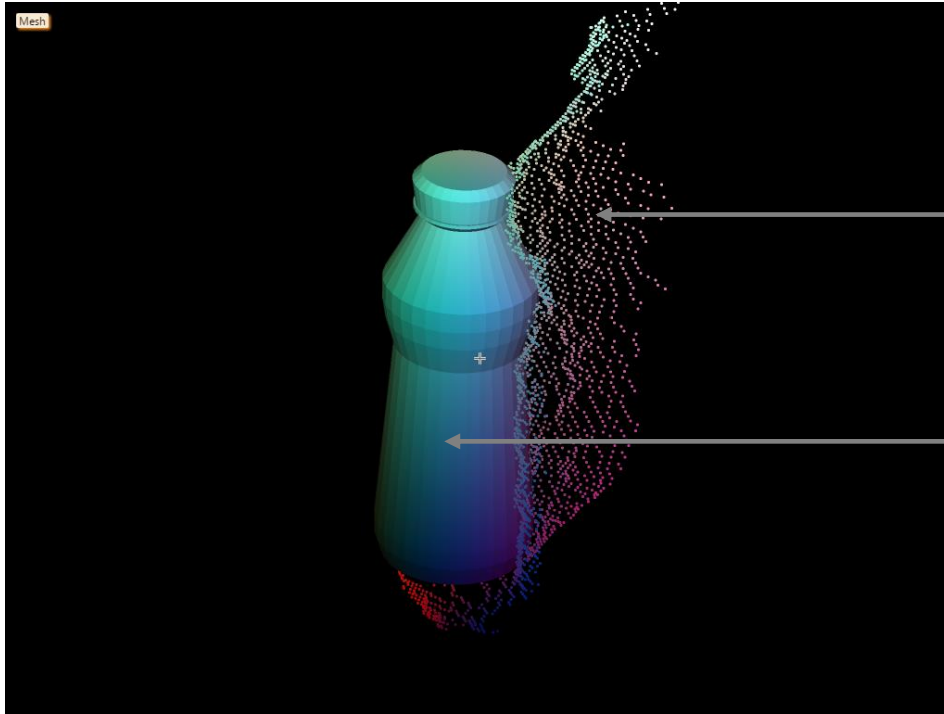
Beispiel Milchflasche – Tiefen-Lookup

Tiefen-Lookup



Lookup und Unterabtastung der Tiefendaten (Rasterabstand 10 Pixel)

Beispiel Milchflasche – Tiefen-Lookup



Punkte aus Tiefen-Lookup

3-D-Modell des zu suchenden Gegenstandes

Fehlerbetrachtung: Bereiche der Hand sowie Ungenauigkeiten in den Datenskalierungen führen zu Fehlern

Ergebnisse zum Konturbasierenden Matching

Nachteile:

1. Konturbasiertes Matching erfordert Ansichten aus mehreren Perspektiven
2. Da ebenfalls Innenkonturen verarbeitet werden und insbesondere auch Aufdrucke (Beispiel: Beschriftung auf der Milchflasche) berücksichtigt werden wird Nachteil 1 noch verstärkt
3. Nur ca. 25 % Verdeckung des zu detektierenden Objektes durch die Hand ist handhabbar
4. Invarianz gegenüber affinen Transformationen liegt vor, jedoch nicht gegenüber perspektivischen

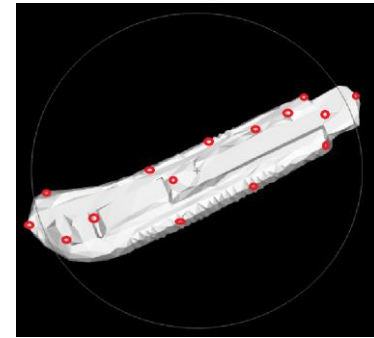
Fazit:

Für Realweltanwendung ist dieses Verfahren nicht robust und effizient genug

Alternativer Detektionsansatz: Key-Point-basiertes 3-D-Matching

Grundgedanke:

Bestimmung von markanten Punkten (Key-Points) am 3-D-Referenzmodell und Einpassung dieser in die aufgenommenen 3-D-Daten (siehe Beispiel: Cutter-Messer)

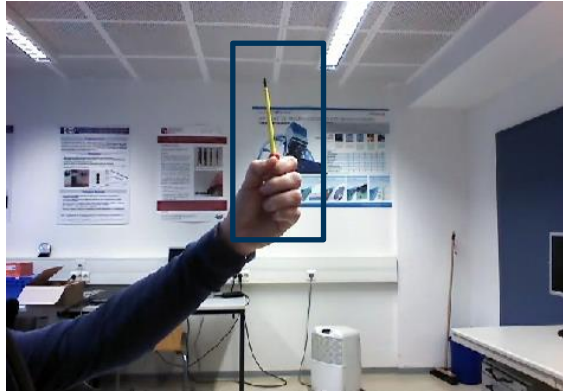


Ergebnis der Untersuchung:

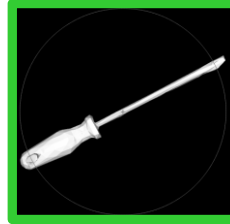
1. Konvergenz sowie Rechenzeit stark Abhängig von Struktur der aktuellen Bildszene der Roboterumgebung
2. Suchbereich teilweise sehr groß (Erfassung des Bereiches von 1 bis 3 m vor dem Sensor bei Azure Kinect)
3. Verkleinerung des Suchbereiches führt zu Artefakten an den Rändern
4. Sehr lange Rechenzeiten (teilweise mehrere Minuten für eine Bildszene)

⇒ **Insbesondere aufgrund der hohen Rechenzeit ist das Verfahren nicht praktikabel für den Anwendungsbereich**

Kombiniertes Verfahren (Deep-Learning + ICP)



Objektdetektion und
Bestimmung ROI



Referenzmodell
suchen



Referenzmodell
einpassen

Kombiniertes Verfahren (Deep-Learning + ICP)

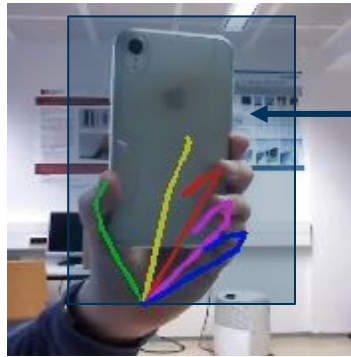
Grundgedanke: Bisher untersuchte Deep-Learning-Verfahren ermöglichen eine Lokalisation des Gegenstandes im Raum, welche jedoch zu ungenau für die Griffpunktbestimmung ist.
=> Verfeinerung der Objektkoordinaten durch ICP-basierte Nachverarbeitung

Verfahrensablauf:

1. Detektion des Gegenstandes auf Basis von Deep-Learning (YOLO Detektor) im RGB-Bild
2. Tiefen-Lookup und Berechnung der 3-D-Koordinaten mittels der intrinsischen Kameraparameter => initiale Punktmenge Gegenstand
3. Einpassung eines vorher aufgenommenen 3-D-Referenzobjektes in die initiale Punktmenge und Verwendung von deren Pose zur Griffpunktbestimmung

Kombiniertes Verfahren (Deep-Learning + ICP)

1. Nutzung der Ergebnisse der Objektsegmentierung des ungenauen aber schnellen Deep-Learning-Detektors als ROI



ROI im RGB-Bild

2. Ausschneiden eines rechteckigen Bereiches mit entsprechendem Toleranzfenster um das gefundene Objekt aus den 3-D-Daten
3. Durchführung des 3-D-Matchings auf den ausgeschnittenen Daten in diesem Bereich

Verwendete Datenstrukturen

Eingabedaten:

- 3-D-Daten
- 3-D-Modell der Referenzgegenstände
- 3-D-Daten der Szene

Ausgabedaten:

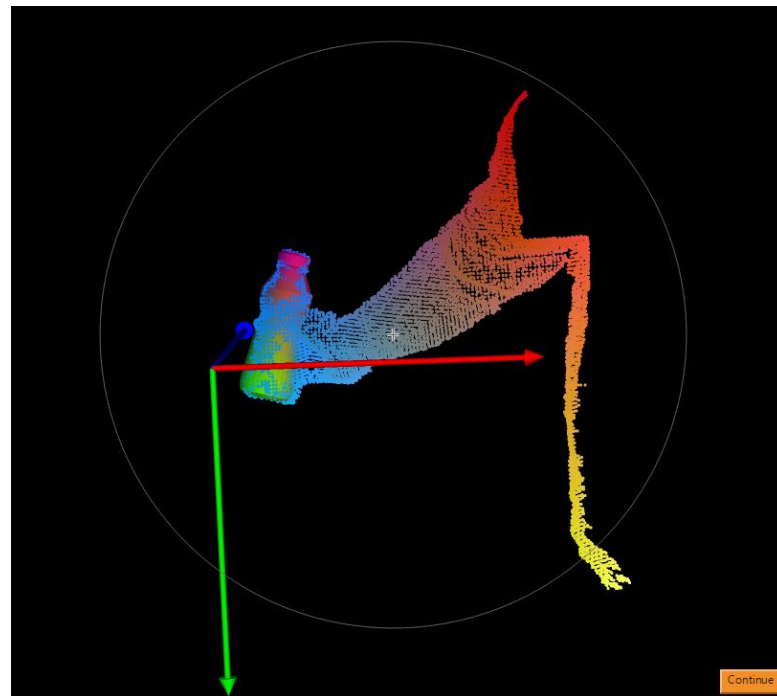
1. Transformiertes 3-D-Modell der Referenzgegenstände
2. Optional: Affine Transformation zwischen 3-D-Referenzmodell und Ist-Daten der aktuellen Szene

Verwendeter Sensor zur Aufnahme der RGB- und Tiefendaten:

Microsoft Azure Kinect

3-D-Matching

3-D-Registrierung nach Parameteroptimierung: Eingepasstes 3-D-Referenzmodell

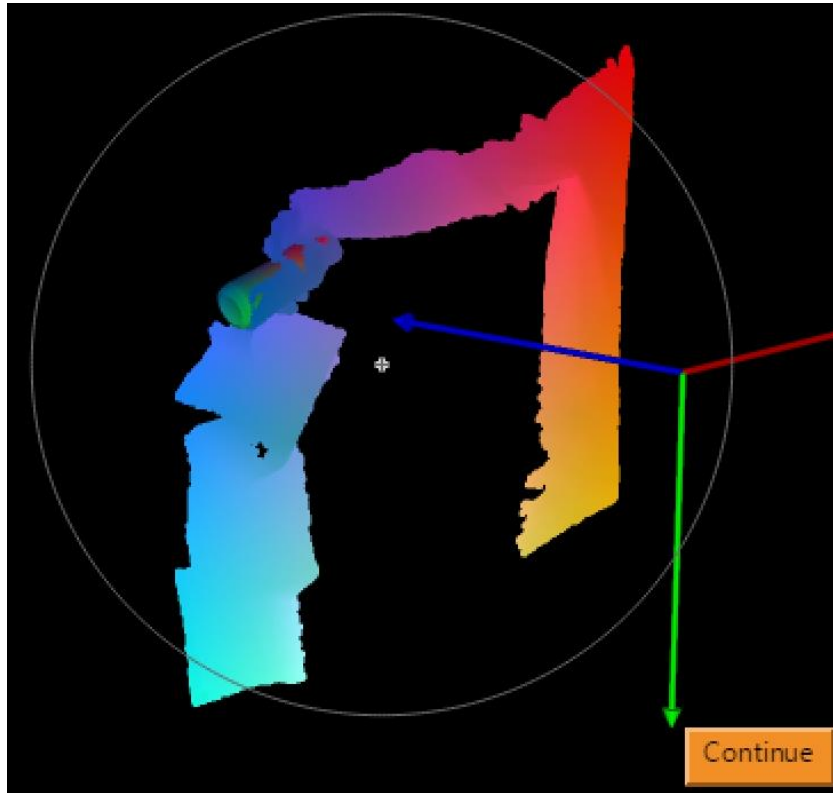


unterabgetastet auf Punktabstand max. 0.005

3-D-Matching

3-D-Szene mit eingepasstem 3-D-Modell

RGB-Bild der Szene



Ergebnisse auf Referenzdatensatz mit 32 Objekten

1. 3D-Matching liefert eine korrekte Einpassung in 94 % der Fälle auf dem Testobjekt Milchflasche (80 Samples)
2. Keypoint-basiertes Matching liefert gute Ergebnisse bei nicht vorliegender Grobausrichtung
3. ICP konvergiert nur bei hinreichender Grobausrichtung und macht Probleme mit verdeckten Bereichen
4. Objekte mit eindeutiger und charakteristischer Form wie Akkuschrauber, Frisbee oder Kunststoffflasche werden mit hoher Sicherheit gefunden
5. Objekte mit Stil wie Drahtbürste, Feile oder Hammer werden aufgrund geringer Verdeckungen und eindeutiger Geometrie gut gefunden

Zusammenfassung

Stand zum letzten Workshop:

Deep-Learning-basiertes Verfahren zur Schätzung des Objekt-Meshes auf Basis von RGB- und Tiefendaten erprobt.

Erreichter Stand:

Konzeption eines kombinierten Verfahrens aus Deep-Learning + ICP zur Steigerung der Genauigkeit der Posenschätzung für eine optimale nachgelagerte Griffplanung

Weiteres Vorgehen:

Rechenzeitoptimierung der ICP-Komponente des Verfahrens

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Einführung des Demonstratorszenarios
und der Verarbeitungspipeline für das Zugreifen

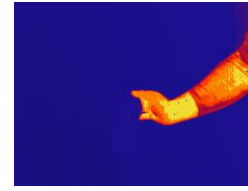
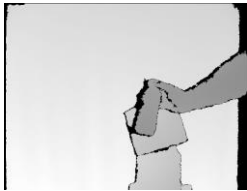
Dr. Steffen Müller

www.sonaro-projekt.de



SONARO HandObject Datensatz

- **Ziel:**
 - Datensatz mit Farb, Thermo und Tiefenbildern von Objekten und haltenden Händen
 - 3D Position und Mesh der Objekte
 - Segmentierte Punktwolke (Hand vs. Objekt)
 - Objekte aus Kategorie Werkzeuge und greifbare Gegenstände



Objektmesh
+ Pose im
Raum



Referenz
ohne Hand

Astra Orbbec S RGB und Depth , Kinect4 RGB und Depth, Thermo

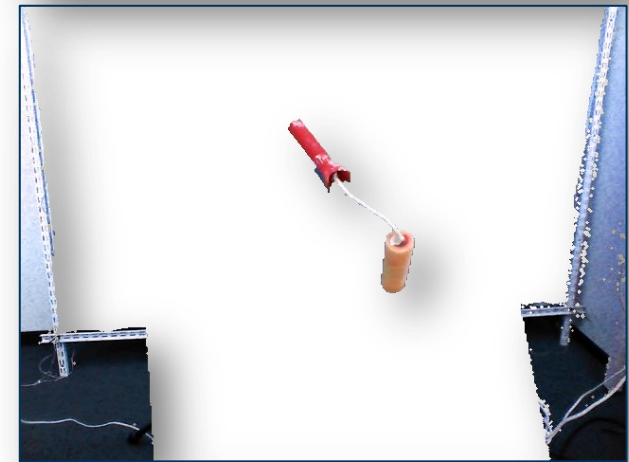
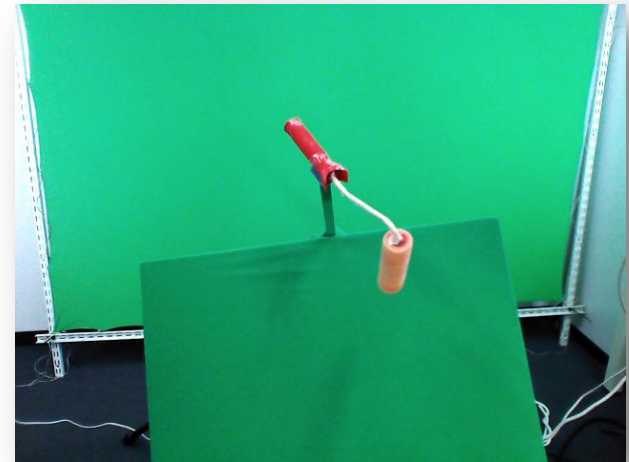
Datensatzaufnahmesetup

Erreichter Stand:

- Objektmodelle mit 3D Scanner erfasst oder CAD Modelliert
- Objekte fest positioniert und dann Hand zu bekannter Objektpose geführt
- 49 Objekte a ca. 120 Bilder → mehr als 300GB
- → Green Screen Segmentierung ermöglicht automatische Labelung im 2D Bild
- Objektmodell und registrierte Lage ermöglicht automatische Segmentierung der Punktwolke in Hand und Objekt



Ausblick:

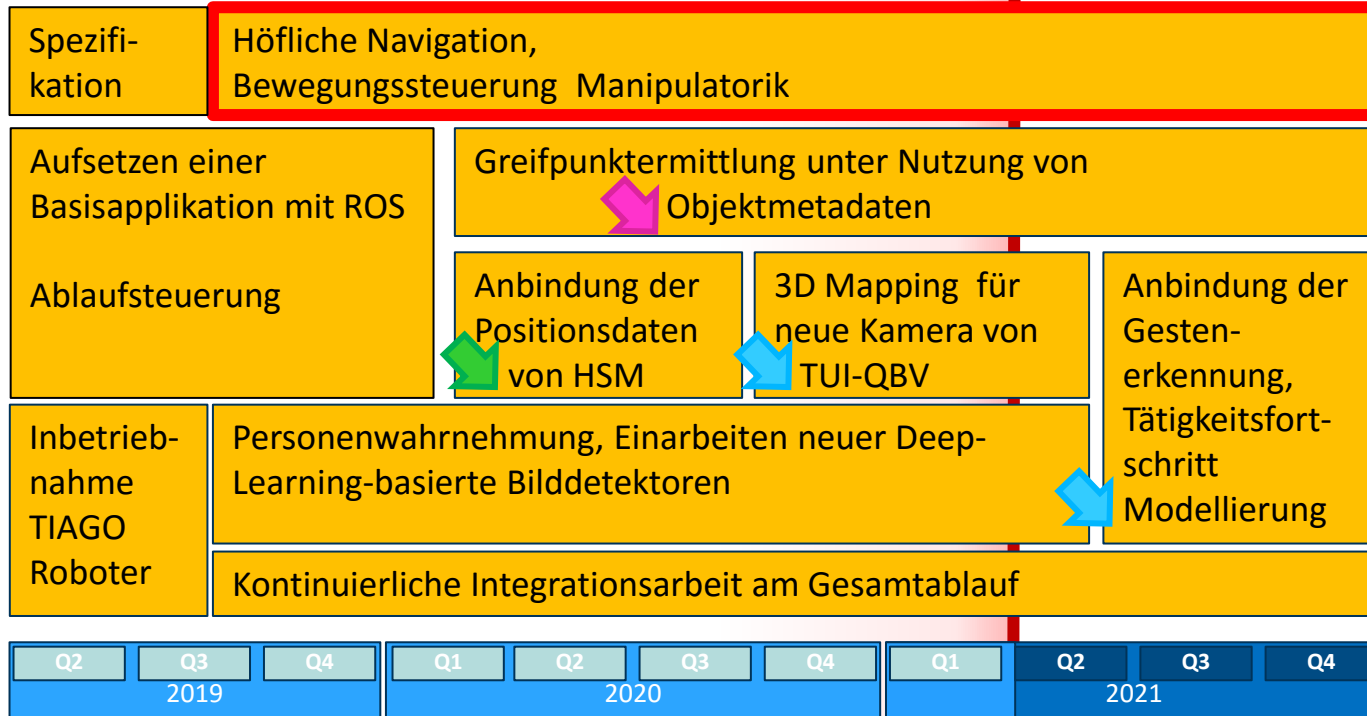
- Datensatz soll publiziert werden



Arbeitsplan für TUI-NIKR über die Projektlaufzeit

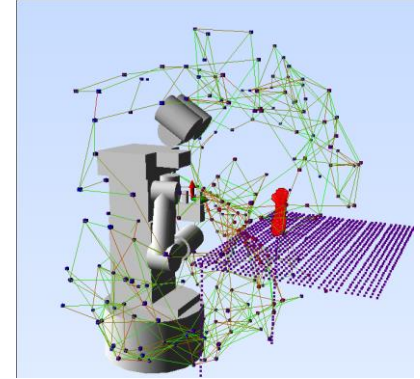
Zuarbeiten von:

-  HSM
-  TUI-QBV
-  GFE



Arbeitspaket Navigation & Bewegungssteuerung

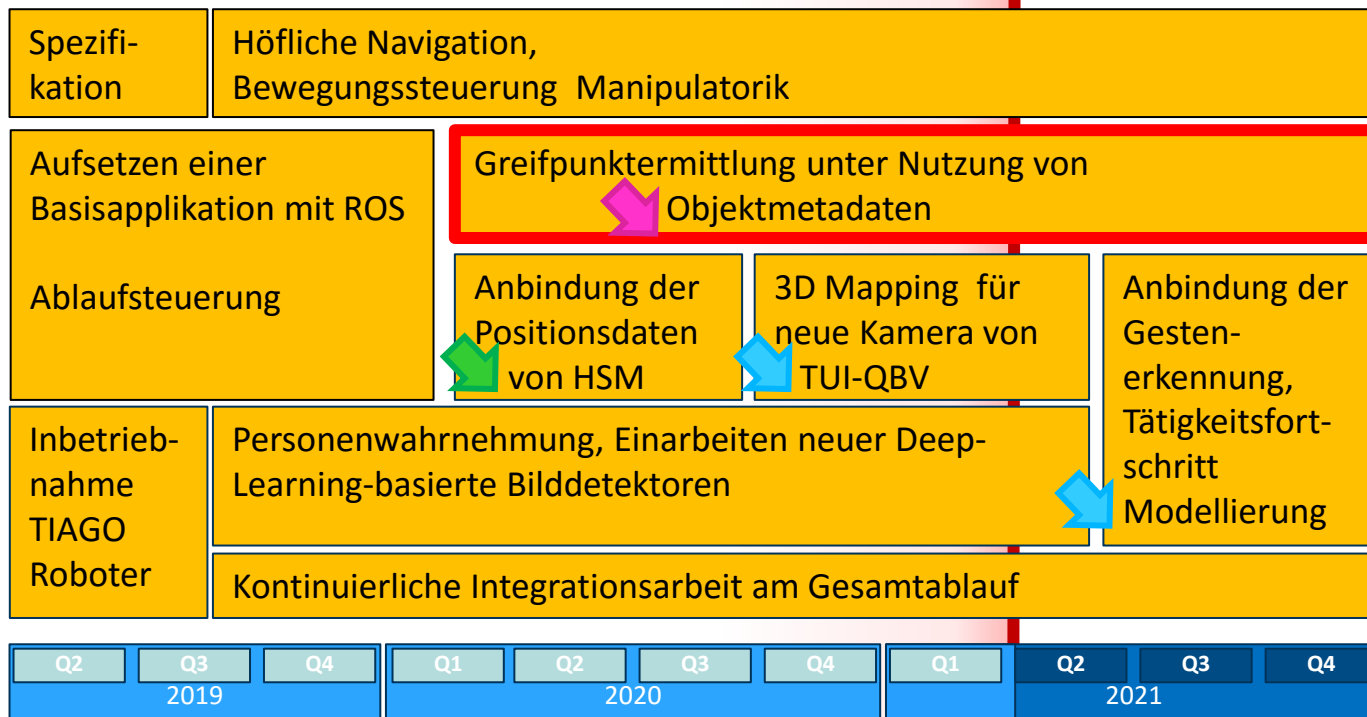
- **Stand zum letzten Workshop:**
 - dynamischer Bewegungsplaner einsatzbereit
 - Zielfunktion für den Greifprozess noch suboptimal
- **Erreichter Stand:**
 - Neuentwickelte Greifzielfunktion
 - keine Zustandstrennung mehr zwischen Annäherung und Zugriff
 - Kraftmessung als Trigger für Zugriff (weil Mensch das Objekt nicht still hält bis der Roboter angekommen ist)
 - Verbesserte Trajektorienplanung → glattere Bewegungen
 - Untersuchung zur Integration externer Bibliotheken für Kollisionsberechnung führte nicht zum Ziel (abgebrochene studentische Arbeit)
- **Weiteres Vorgehen:**
 - Implementierung Übergabeverhalten



Arbeitsplan für TUI-NIKR über die Projektlaufzeit

Zuarbeiten von:

- HSM
- TUI-QBV
- GFE



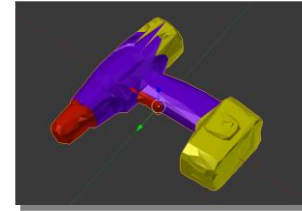
Greifposenermittlung

- **Stand zum letzten Workshop:**

- Analytischer Greifposenplaner (Eigenentwicklung)
- Approximation mittels CNN → Beschleunigung (kann nicht alle Kriterien berücksichtigen)

- **Erreichter Stand:**

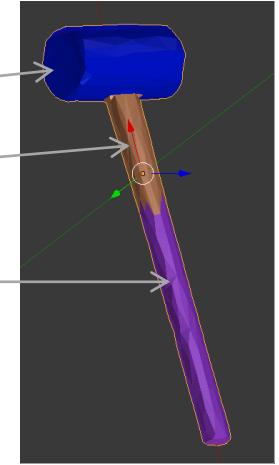
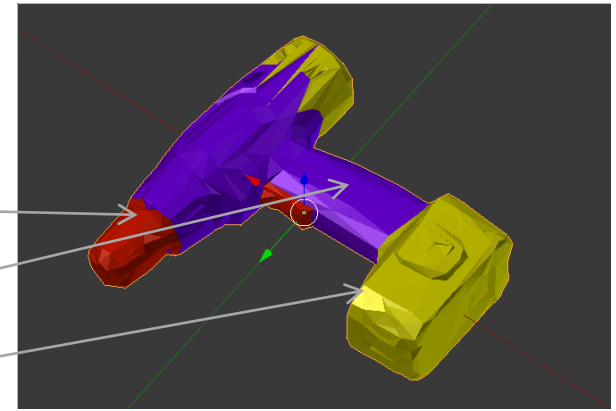
- Berücksichtigung von verschiedenen Griffarten (Pinzettengriff, Zylindergriff, Sphärischer Griff)
- Berücksichtigung von Taskspezifischen Affordanzen (unterschiedliche Präferenzen für Übernahme, Übergabe und Nutzung)
- Abschluss Projektseminar zur Evaluation der Greiferfolge mittels GraspBox



Projektseminar Lucas Vollmar, Toan Phan-Bao: Aufbau einer Testumgebung für robotisches Greifen, TU Ilmenau 2021

Affordanzen und Griffarten

- Objekte des Datensatzes wurden manuell gelabelt
 - Verbotene (gefährliche Angriffsflächen)
 - Mögliche Griffart:
 - Pinzettengriff
 - Stick
 - Zylinder
 - Sphärisch
 - Taskspezifische Affordanzen:
 - Give
 - Take
 - Use
- Berücksichtigung der Informationen im analytischen Greifposenplaner als zusätzliche Bewertungsfunktionen
- **Ausblick:** Schätzung der Labels mittels Machine-Learning



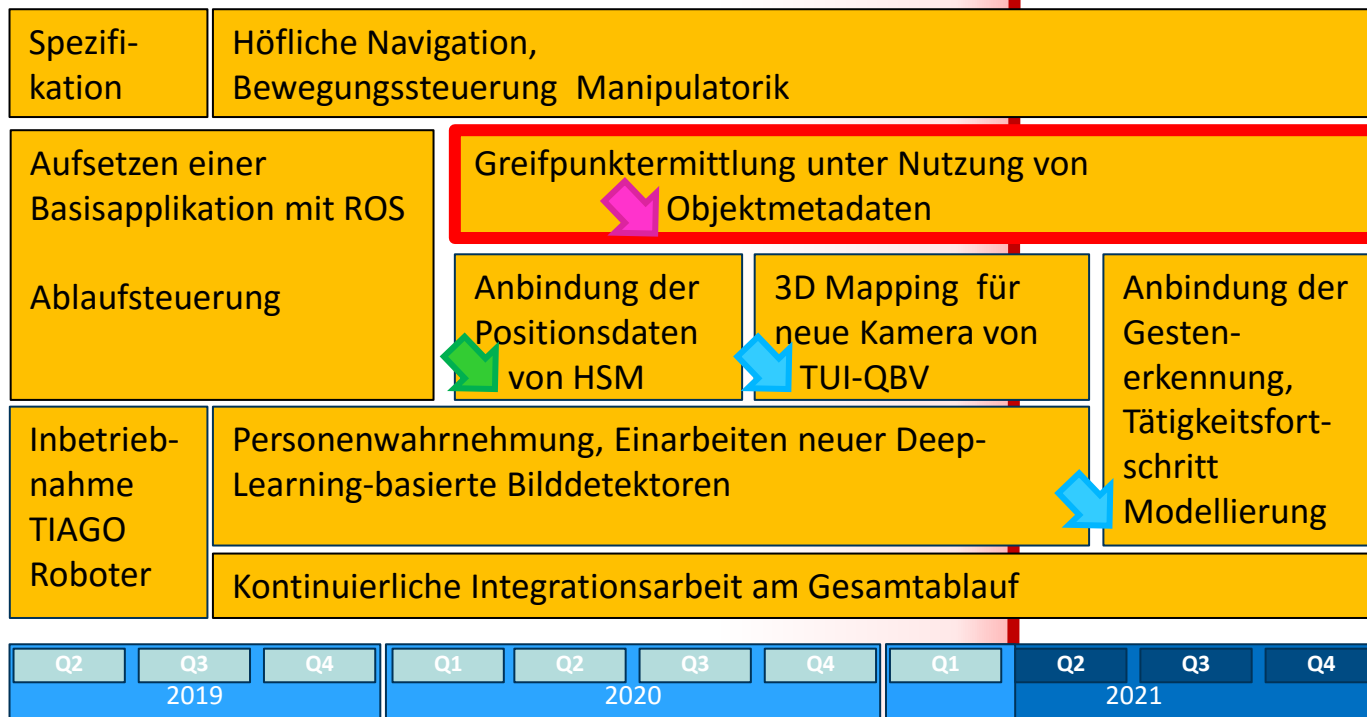
Taskabhängige Greifposenbestimmung

demonstration of task dependent online grasp pose
detection for three finger Kinova gripper

Arbeitsplan für TUI-NIKR über die Projektlaufzeit

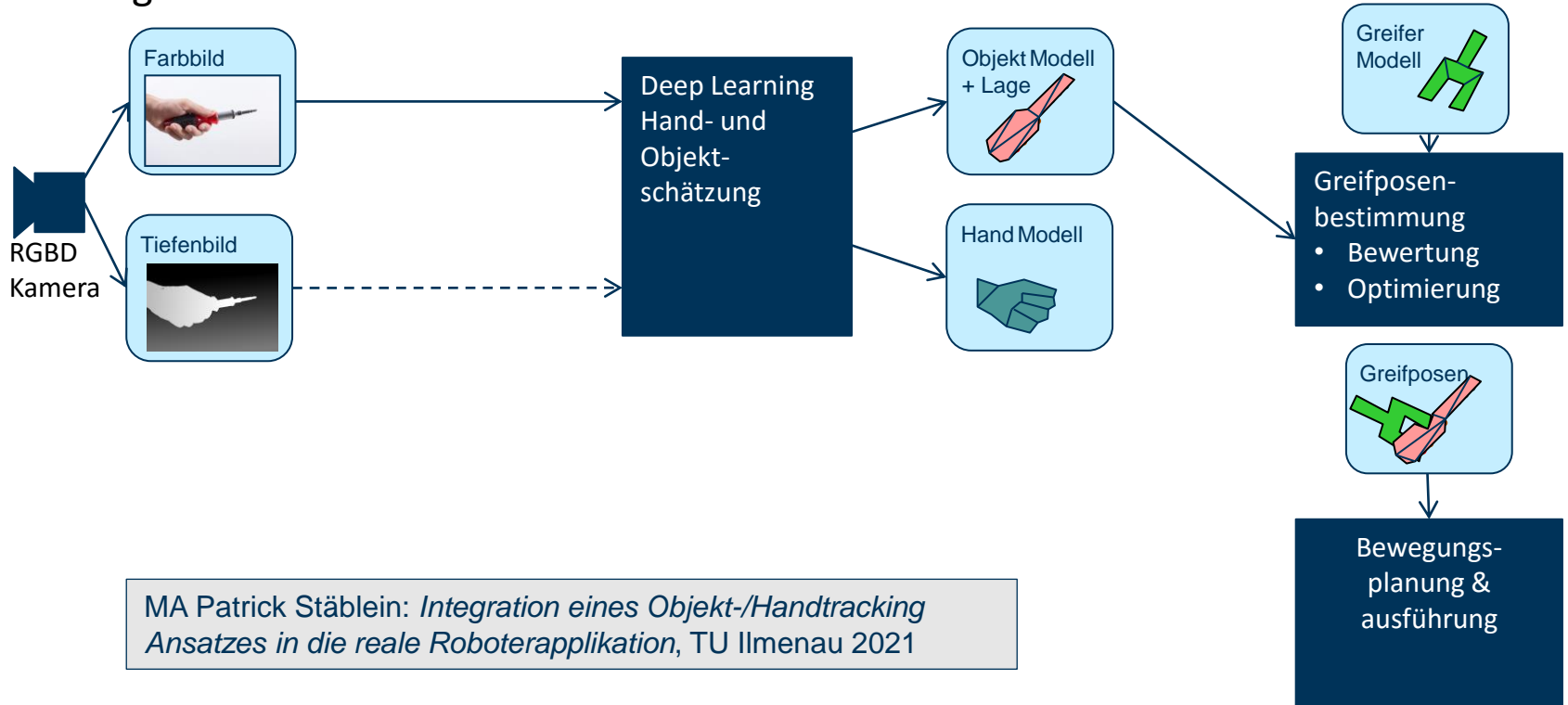
Zuarbeiten von:

- HSM
- TUI-QBV
- GFE



Untersuchung der DeepLearning Verfahren für Objektschätzung

- Abgeschlossene Masterarbeit



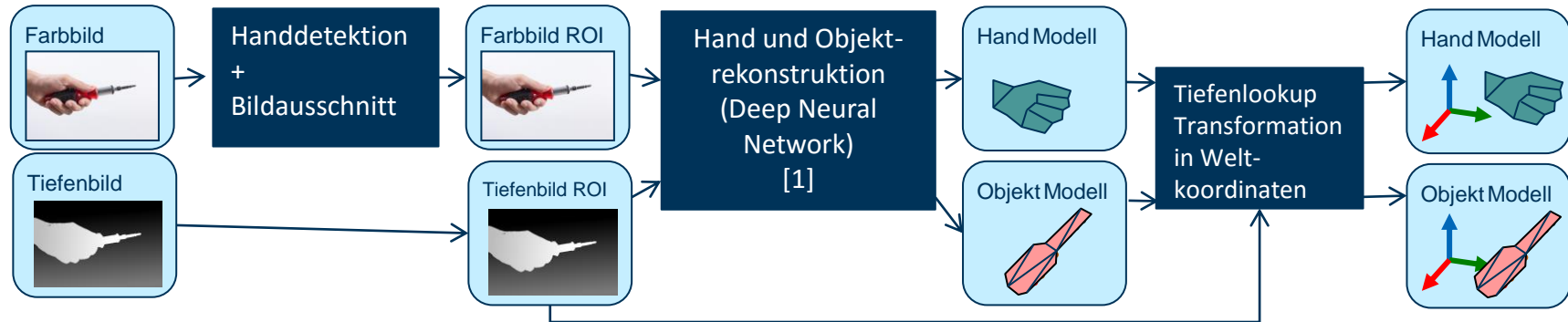
MA Patrick Stäblein: *Integration eines Objekt-/Handtracking Ansatzes in die reale Roboterapplikation*, TU Ilmenau 2021

DeepLearning Verfahren für Objektschätzung

- Entwurf einer Pipeline für die Integration des Verfahrens [1] in die Roboteranwendung

[1] Hasson, Yana, et al. : *Learning joint reconstruction of hands and manipulated objects*. In: CVPR. (2019)

- Ursprünglich nur für Aktivitätserkennung in YouTube Videos gedacht → andere Ansprüche an Positionsgenauigkeit
- Vorgeschaltete Handdetektion mittels YOLO Detektor
- Anschließend Lookup der Entfernung im Tiefenbild für Anbindung an Weltkoordinaten

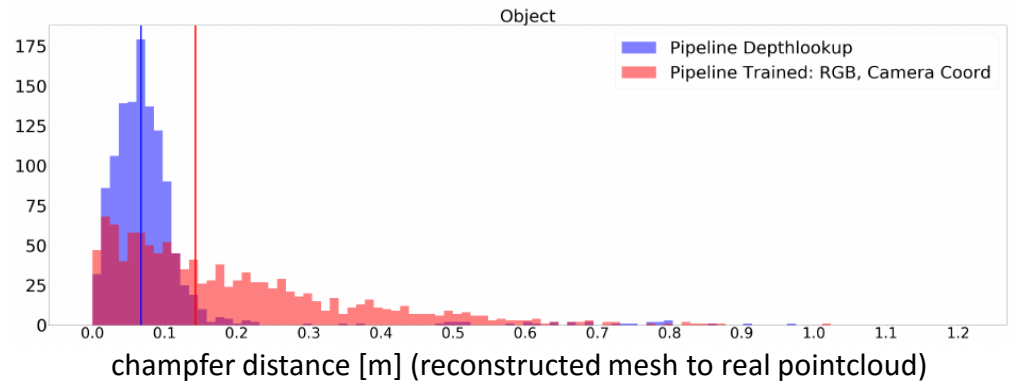
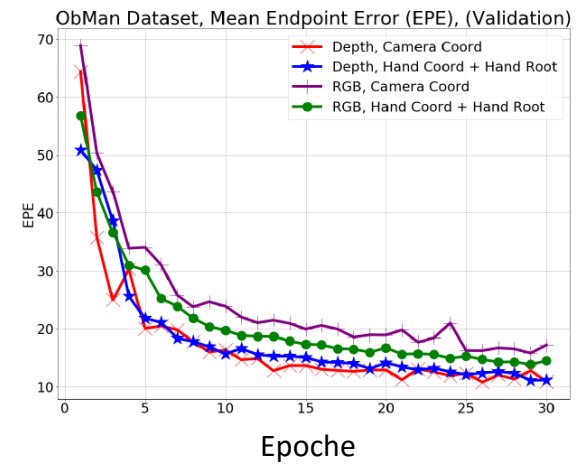


DeepLearning Verfahren für Objektschätzung

EndPointError (EPE):
Rekonstruktionsfehler
des Handskeletts

- **Ergebnisse:**
 - Training auf Tiefenbildern vs. Training auf RGB Bildern
 - Kodierung in Weltkoordinaten vs. Handrelativposition
 - Training auf Weltkoordinaten vs. Lookup in Tiefenbildern

Variante	EPE
Depth, Camera Coord	10.80
Depth, Hand Coord + Hand Root	11.13
RGB, Camera Coord	15.86
RGB, Hand Coord + Hand Root	13.91
HASSON et al., 2019	11.6

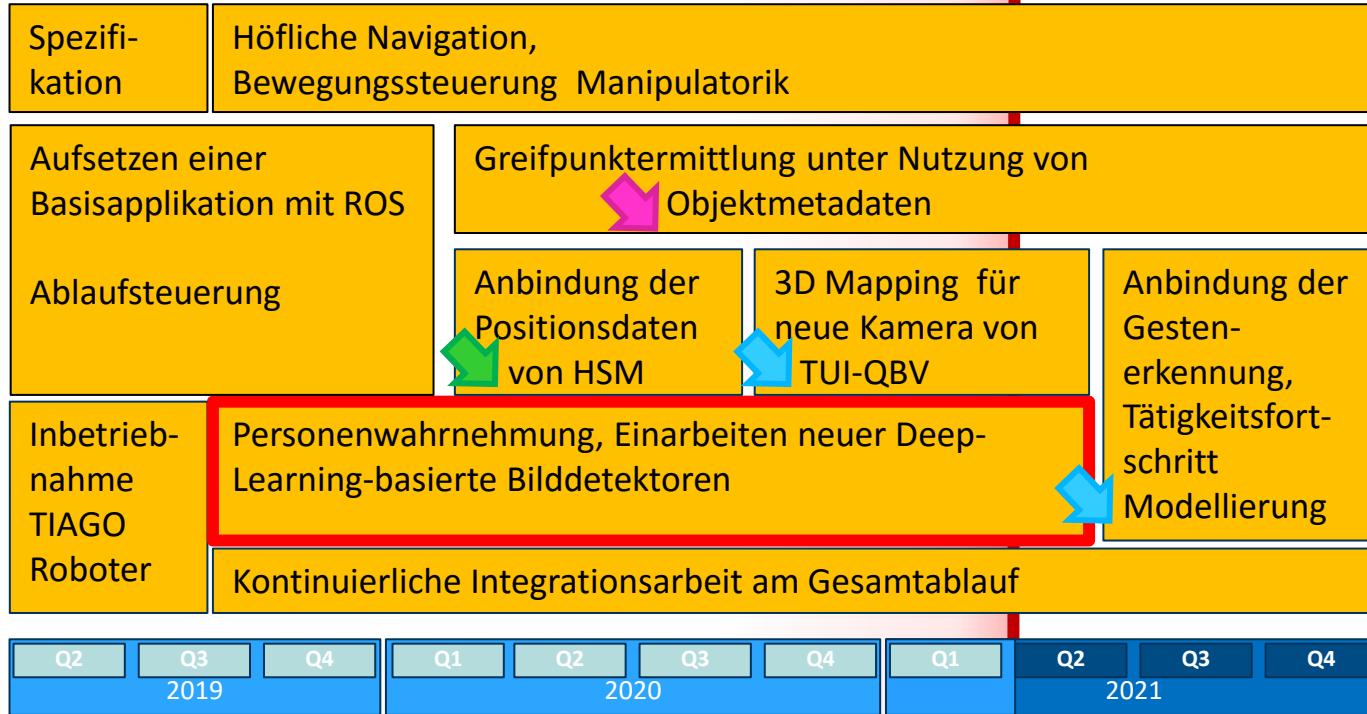


- **Fazit:**
Für Greifaufgaben ist die Objektschätzung zu ungenau.
(insbesondere bei unbekanntem Objekten)

Arbeitsplan für TUI-NIKR über die Projektlaufzeit

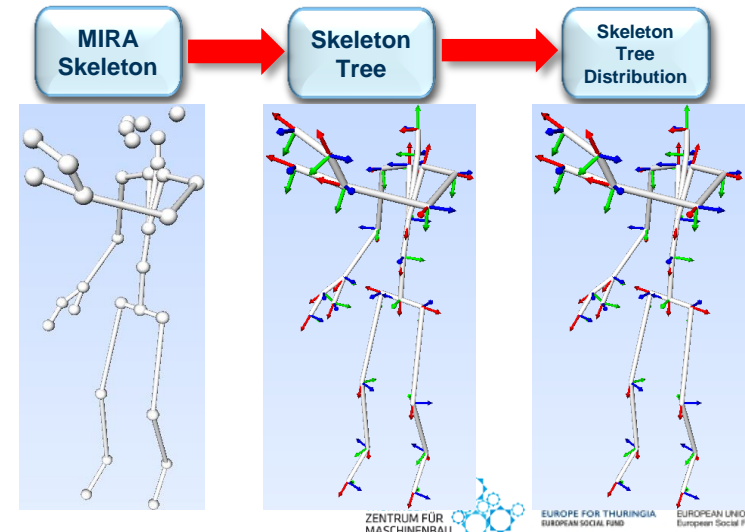
Zuarbeiten von:

-  HSM
-  TUI-QBV
-  GFE



Intenterkennung mittels Skelettracking

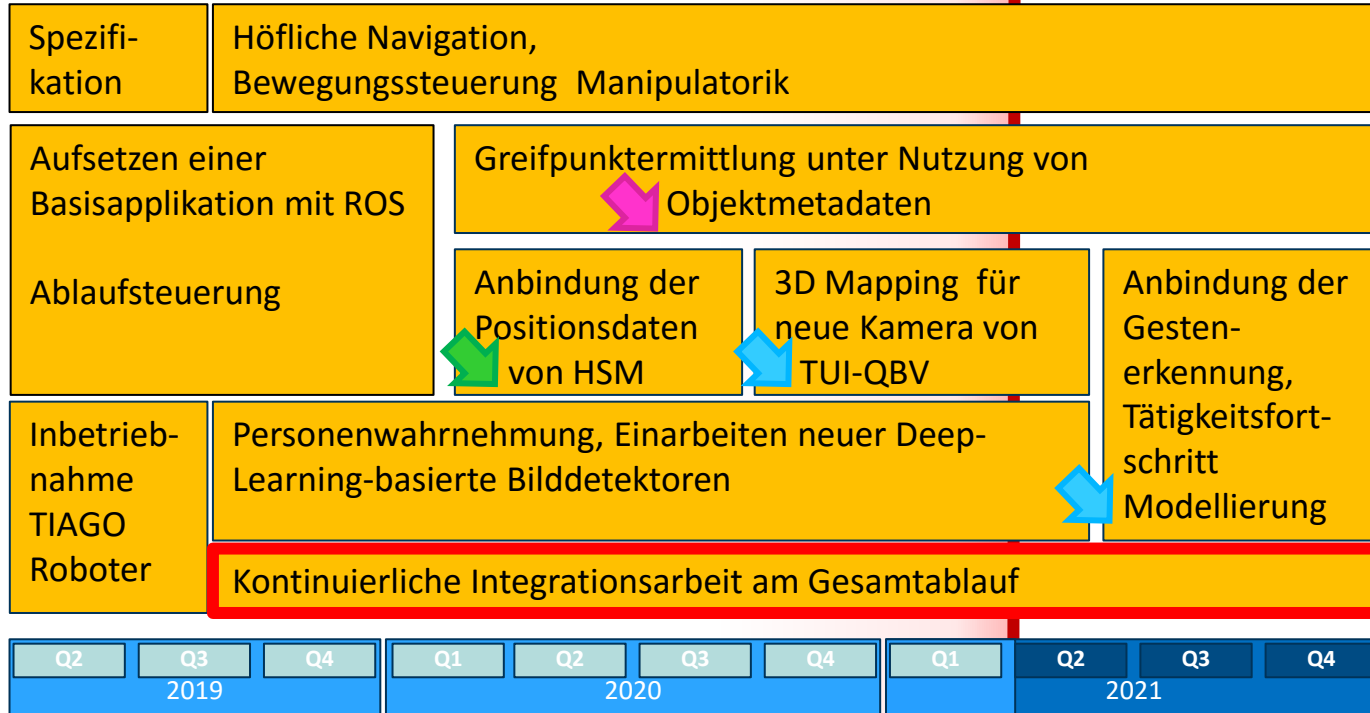
- **Masterarbeit kurz vor Abschluss.**
- Integration der Skelettschätzung von Kinect4 Azure in das multimodale Trackingframework (ermöglicht Fussion von mehreren Ansichten und persistente Hypothesen auch bei kurzzeitigen Verdeckungen)
- Ableiten von Gelenkwinkelfeatures aus unterschiedlichen Zeitfenstern
- Klassifikation der Intentgesten (z.B. „Komm her“, „Nimm Gegenstand“, „Gib mir was“, Geh weg“,...) mittels **Gaussian Mixture Models**
- **Ausblick:**
 - Nutzung der Skelettdaten zur Einschränkung der Objektposition für Übernahme
 - Nutzung der Intenterkennung als Trigger für die Übergabe/Übernahme
 - Nutzung der Handposition als Übergabeposition



Arbeitsplan für TUI-NIKR über die Projektlaufzeit

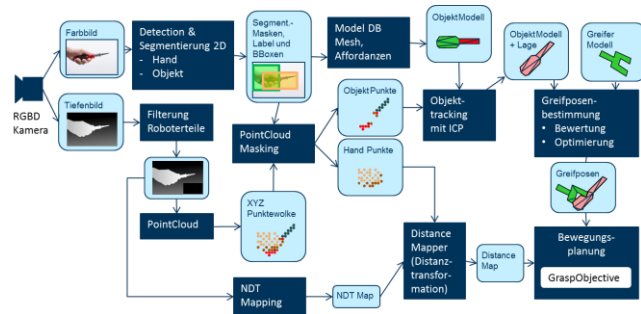
Zuarbeiten von:

-  HSM
-  TUI-QBV
-  GFE




Hardwareplattformen für die Demonstration der Übergabe

- Mobile Roboter:
 - Tiago
 - ScitosX3
- Stationärer Teststand:
 - GraspBox



Pipeline (c) wurde auf den Plattformen integriert → Demo

Ausblick auf das letzte Projekthalbjahr

Projektziele	Offene Punkte (todos)	Zusätzliche erwartete Ergebnisse bis Projektende
Demonstration der Objektübergabe- / übernahme	<ul style="list-style-type: none"> • Entwicklung der Übergabestrategie • Integration der Abläufe auf dem Roboter • Intenterkennung als Trigger 	<ul style="list-style-type: none"> • Umgreifen der übernommenen Objekte zwecks ergonomischer Übergabe 
Greifprozess	<ul style="list-style-type: none"> • Optimierung des Trackingprozesses • Auswertung der Greiferfolge aus der haltenden Hand • Auswertung und Benchmark der Bewegungssteuerung 	<ul style="list-style-type: none"> • Affordanzschätzung für unbekannte Objekte mittels Machine Learning
Navigation	<ul style="list-style-type: none"> • Heranfahen an die Zielperson 	

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die
nutzerzentrierte mobile Assistenzrobotik

Beiratstreffen am 20.04.2021

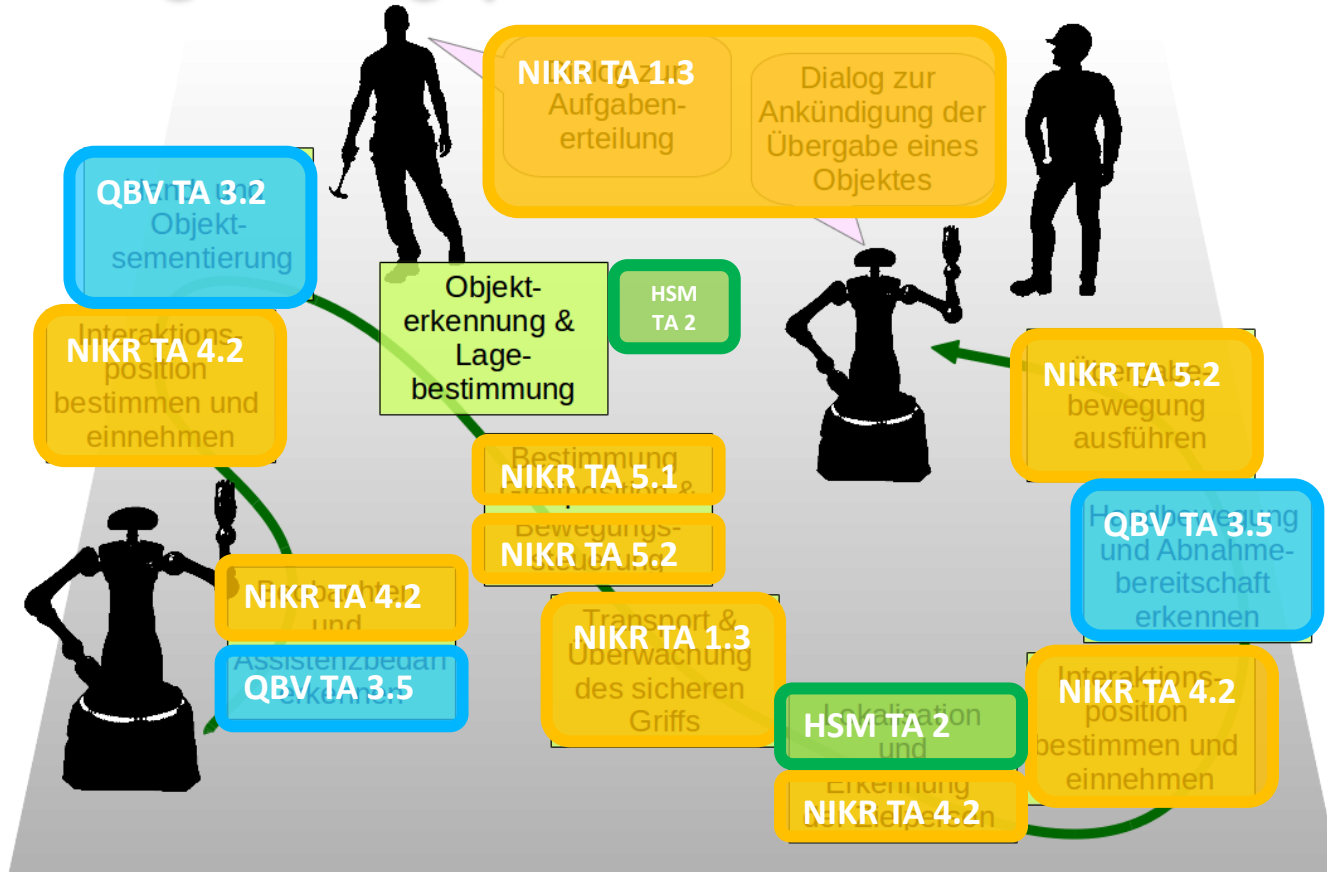
Vorstellung der Arbeiten von HSM FG EDS

M.Sc. Manuel Schneider

www.sonaro-projekt.de



Abdeckung durch geplante AP in SONARO

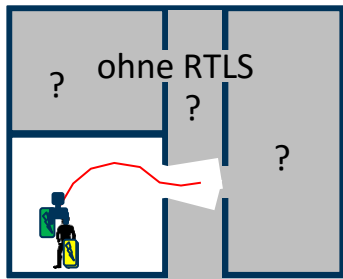
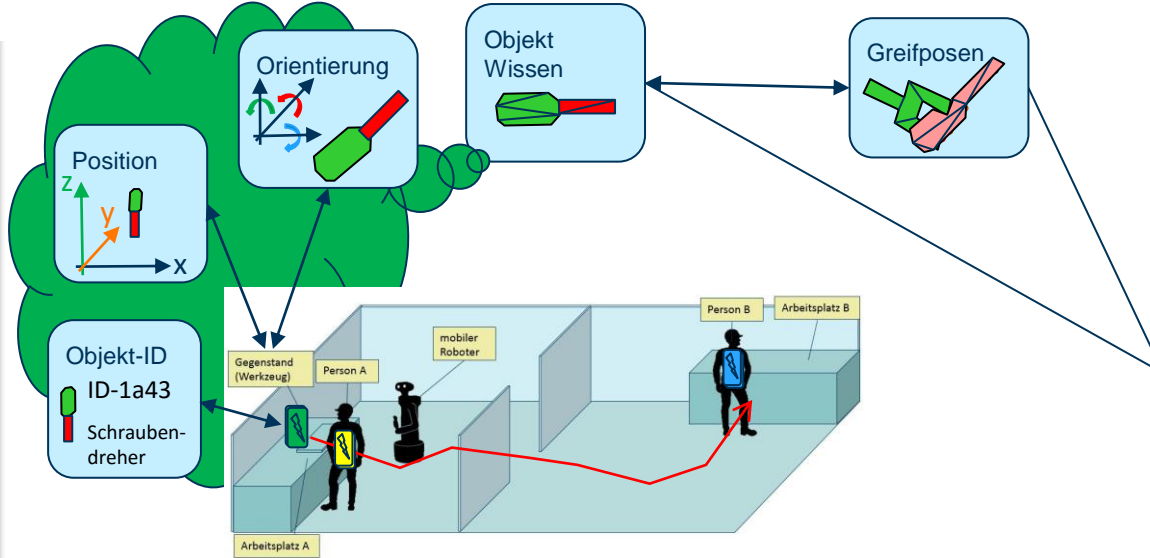


Ortung von Personen / Gegenstände mit RTLS

TIAGo



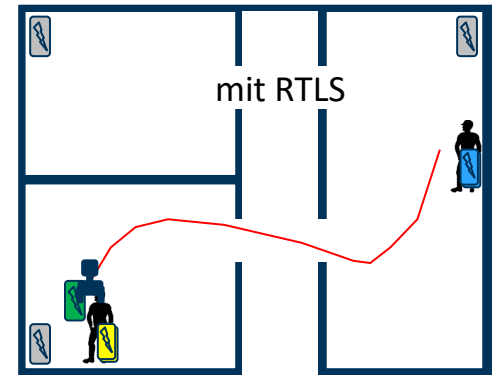
Softwareplattform



RTLSflares (Funk-Ortungssystem)



Erweiterung der Umgebungswahrnehmung



Weitere zeitliche Planung

Anforderungsanalyse, Spezifikation und Schnittstellendefinition

Anbindung an die genutzten Frameworks



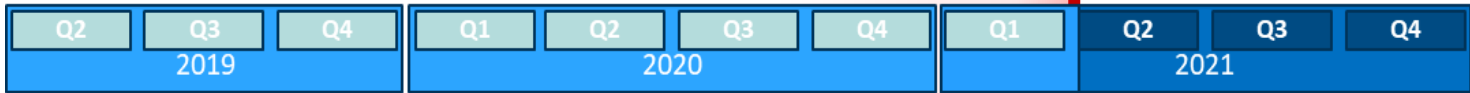
Messszenarien mit Hinderniserkennung
Datenaufzeichnung, Entwicklung von Algorithmen zur Optimierung der Positionsschätzung

Optimierung Hinderniserkennung





Hardwareoptimierung

Datenaufzeichnung und Prüfen der neuen RTLS – Funktionen

Zuarbeiten für:
TUI-NIKR





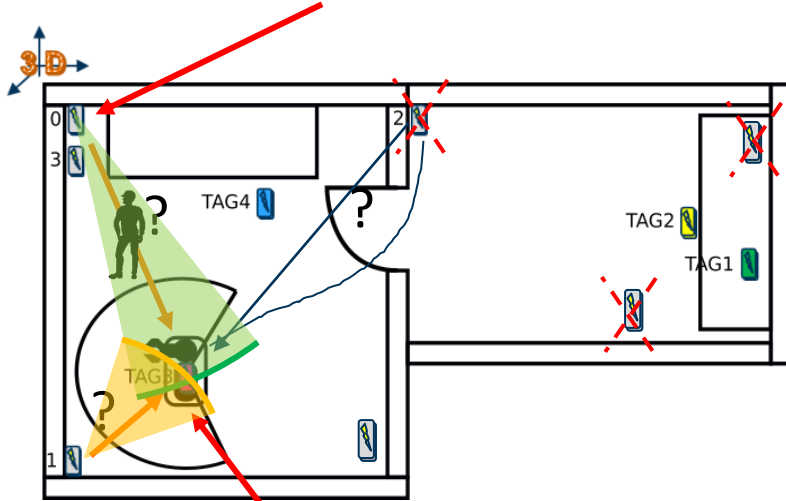
-  RTLS Anker (feste)
-  Personenortung
-  Bewegungstracker der Person (UWM / Laser)
-  Objektortung

Probleme:

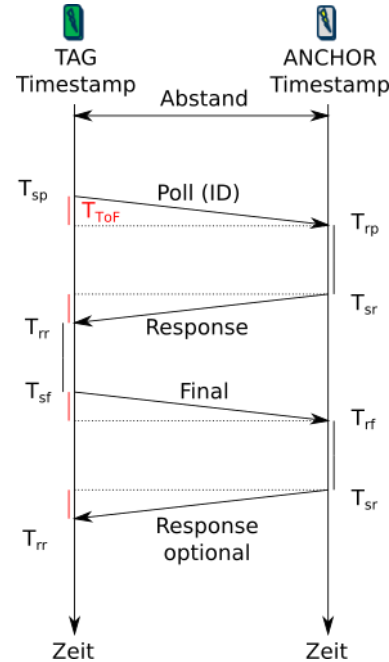
- Verbindung Roboter zum zu trackenden Flare kann gestört / zu weit weg sein.

Grundlagen RTLS Ortung / Erkennung einer gestörten Messung

Ortsfeste RTLSflares (Anker)



mobile RTLSflares (Tag)

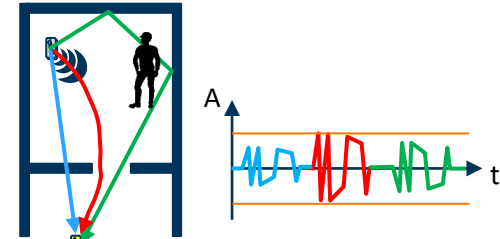


$$\text{Abstand} = \text{ToF} * c$$

ToF = Time of Flight

c = Lichtgeschwindigkeit

Fehlerursachen:



Wann ist eine Verbindung gestörte?

- Dämpfung der Signalstärke bei Hindernissen
- Entfernung, Reflexion

Optimierung der Firmware / Informationen vom Anker

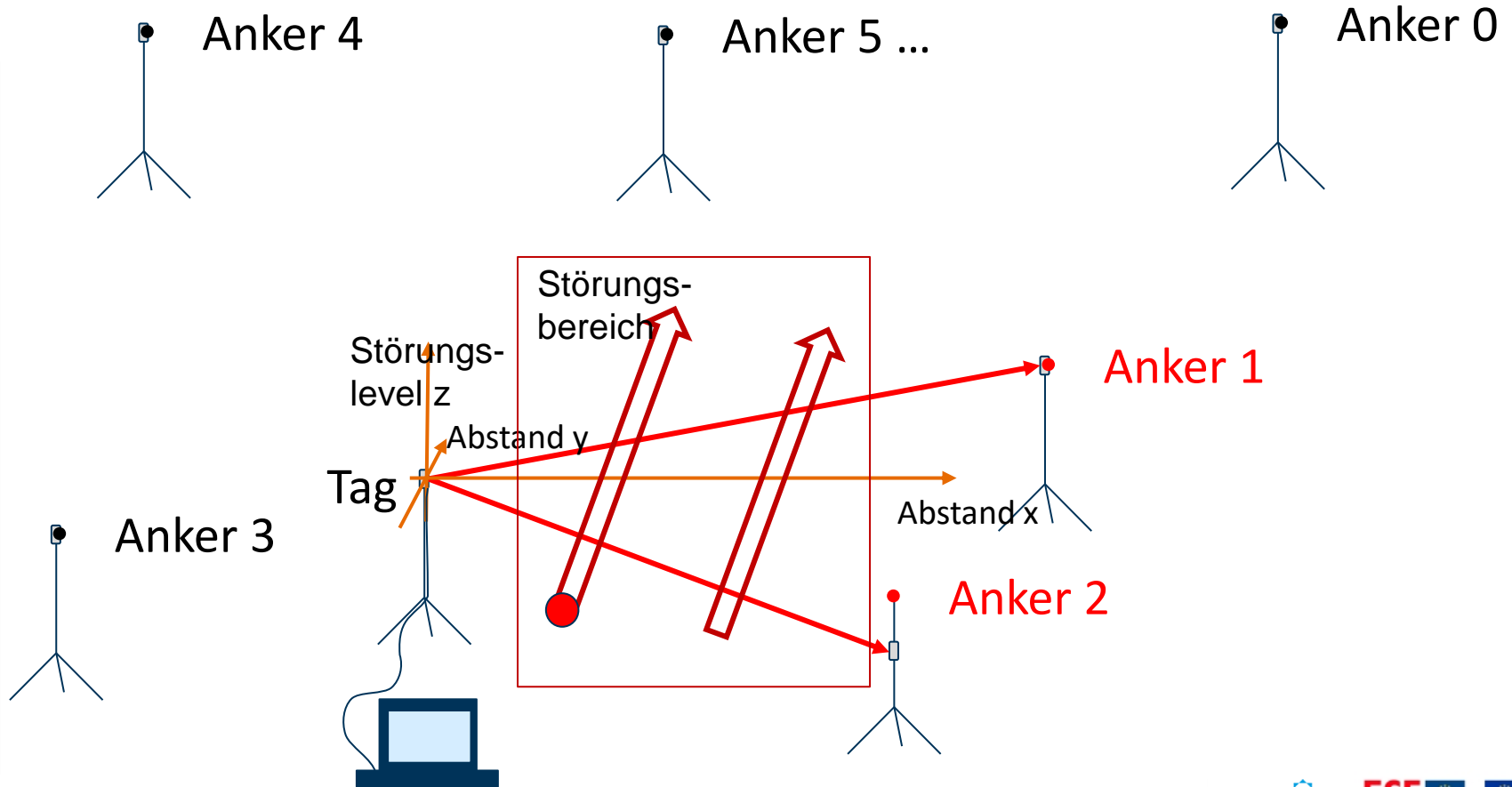
"al": [Anchorliste als JSON Array mit bis zu 31 Anchorobjekten
{	Anchorobjekt für jeden bekannten Anchor
"id16": "0x4e1a",	UDID des Anchors als JSON String
"pos": {	Positionsobjekt des Anchors
"x": -34.16,	
"y": -2.38,	
"z": -1.48	
},	
"dof": 0,	Freiheitsgrad des Anchors, 0 bis 3 für die Basiszelle, 4 alle anderen
"pr": 0.998,	künstliches Hilfsmaß „Positionrate“ für die Güte der Positionsbestimmung des Anchors
"ldis": 22.87,	aktueller Abstand zum Anchor in m
"adis": 22.55,	Mittelwert des Abstands zum Anchor seit Reset in m
"var": 0.0712,	„Varianz“ des Abstands zum Anchor
"sqn": 5,	Sequenznummer, bezieht sich auf die Reihenfolge, in der die Anchor abgefragt wurden
"nom": 433,	Anzahl der Abstandsmessungen seit Positionsreset
"blk": 579591,	Zeitstempel des letzten „life sign“
"fpp": -32.54,	First Path Power (siehe nächstes Kapitel)
"rxp": -30.17,	Received Power (siehe nächstes Kapitel)
"csn": 104.00,	Channel Standard Noise
"cmn": 2652.00,	Channel Maximum Noise
"toc": 0,	Time out counter, wird erhöht, wenn Abstandsmessung fehlgeschlagen ist
"ra": 1,	Messanchor, 1 bedeutet, Daten sind valide und werden zur Positionsbestimmung verwendet

Berechnung im Anker

Berechnung im Tag

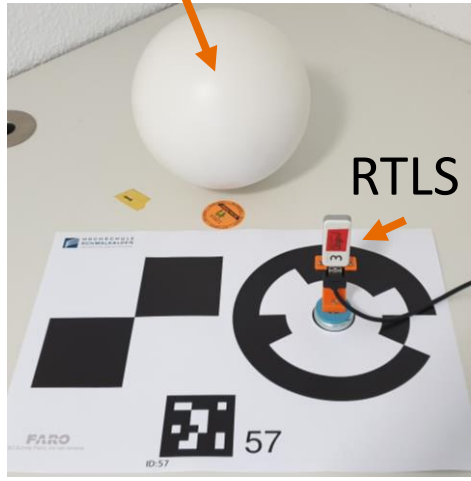
Neue Merkmale zur Erkennung einer gestörten Verbindung

Schematische Darstellung Teilergebnis der Untersuchung



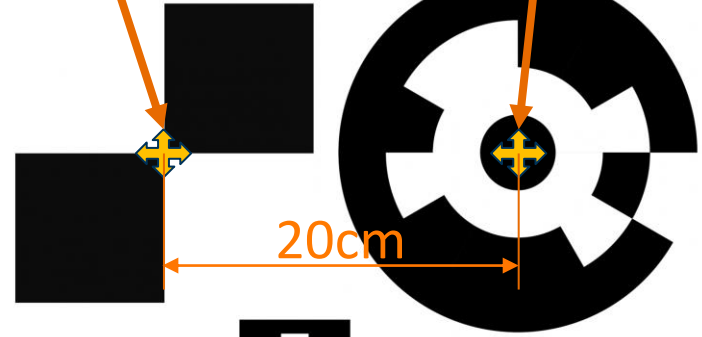
Neue FARO Messmarken zur Positionsbestimmung der RTLS Anker

Faro Messkugel

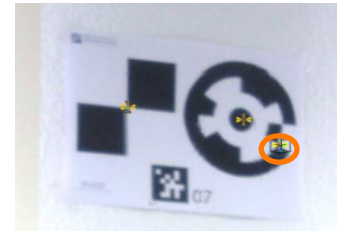


Faro Schachbrett
Position

Radialmarkierung



Messfehler

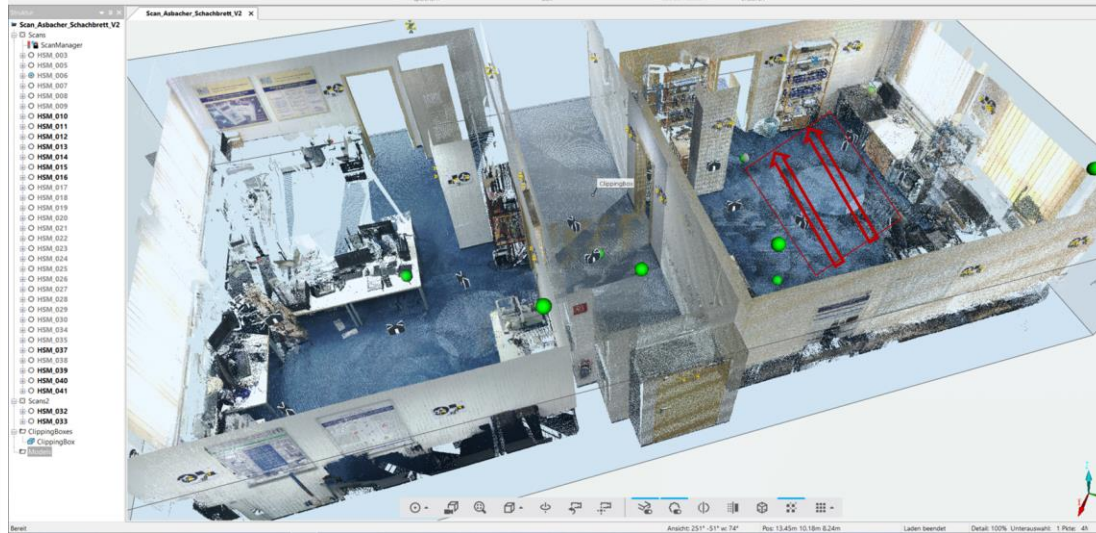


Aruco Marker

Position und ID
z.B. 001408

Indirekte Position über
Kamerabild und ID

Ergebnisdarstellung der Untersuchung von 3D Scans

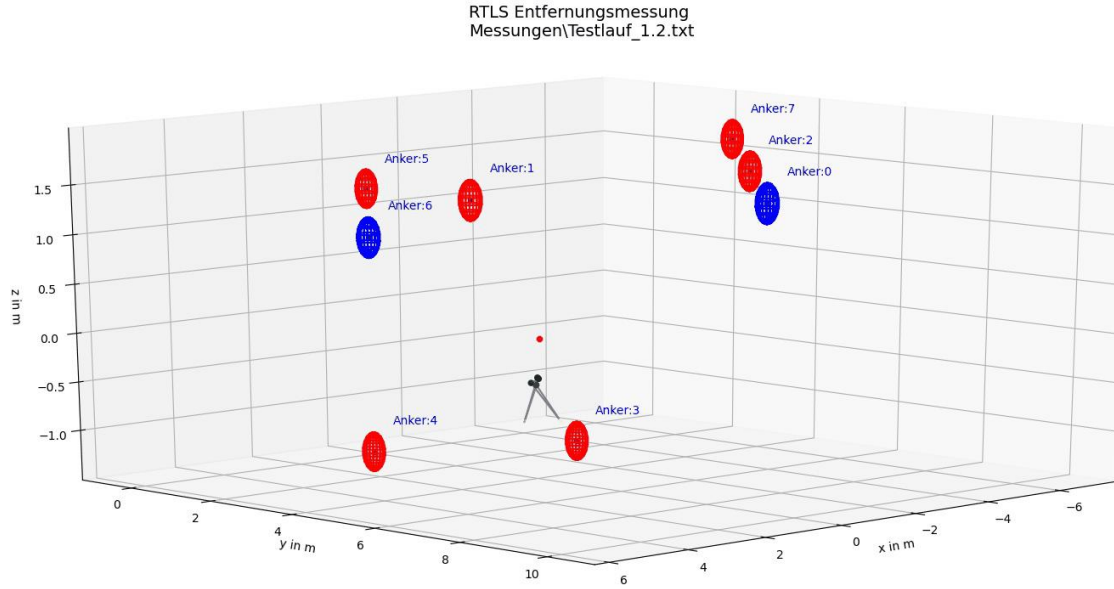


- 3D Vermessung der Versuchsumgebung
- Genaue Bestimmung der Abstände zwischen den Ankern und dem Tag (auch durch Wände)
- Bildet die Bezugsgröße für die Fehlerbestimmung zum RTLS Ortungssystem

Messung	1.	2.	3.	4.
Max. Abstandsfehler (mm)	5,0	4,3	4,5	1,0
Durchschnittlicher Abstandsfehler (mm)	1,1	2,7	2,3	0,8
Max. Horizontaler Fehler (mm)	4,8	3,9	4,0	0,6
Max. Vertikaler Fehler (mm)	0,8	2,5	2,1	0,3

Nur mit Faro Messkugel

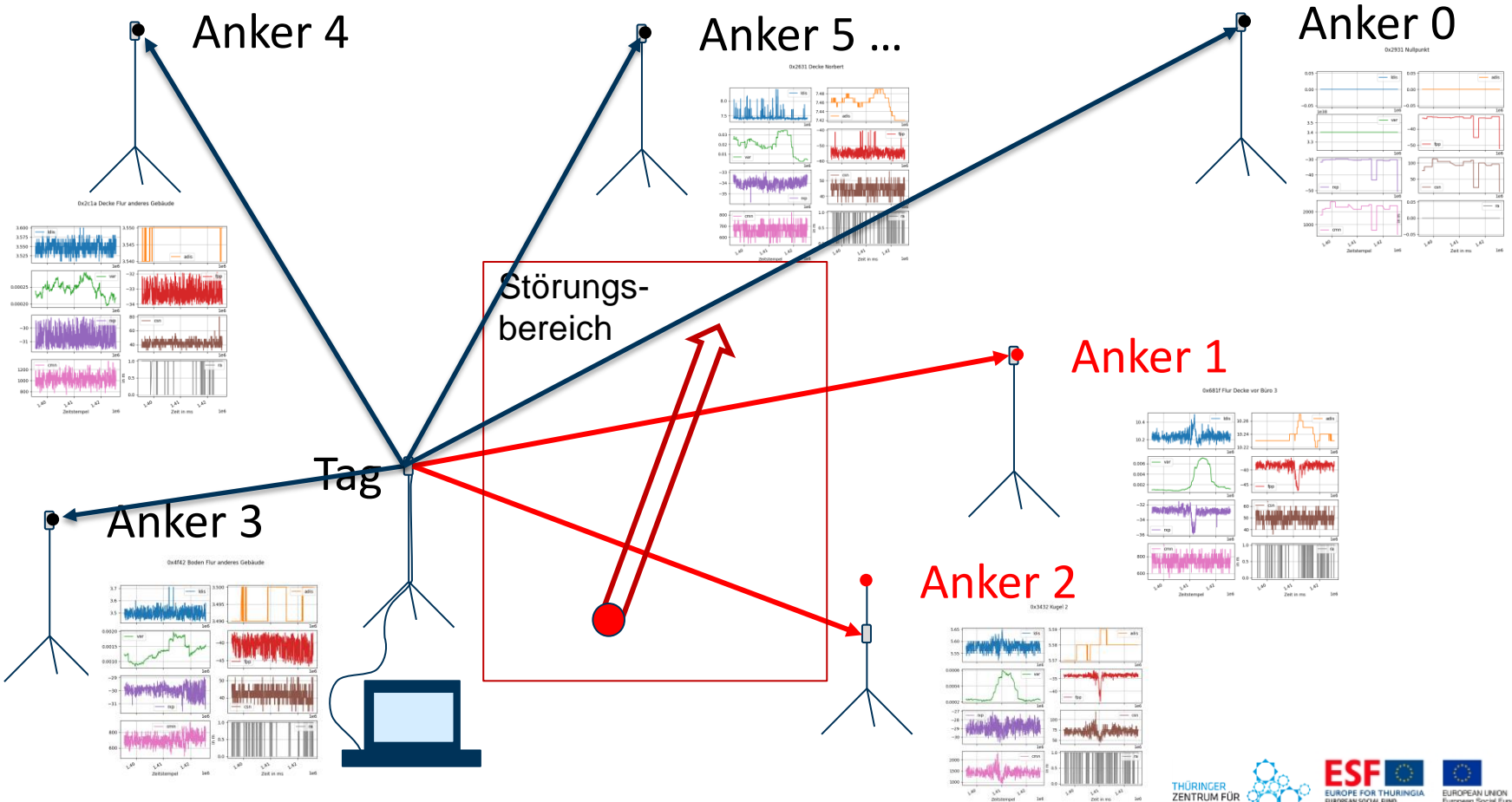
RTLS Untersuchung



- RTLS POS Path
- REAL_POS
- RTLS POS xyz
- 0x2931 Anker:0
- 0x6b1f Anker:1
- 0x3432 Anker:2
- 0x3632 Anker:3
- 0x4f42 Anker:4
- 0x2631 Anker:5
- 0x2c1a Anker:6
- 0x681f Anker:7

- Aktive Verbindung zur Positionsbestimmung
- Inaktive Verbindung zur Positionsbestimmung

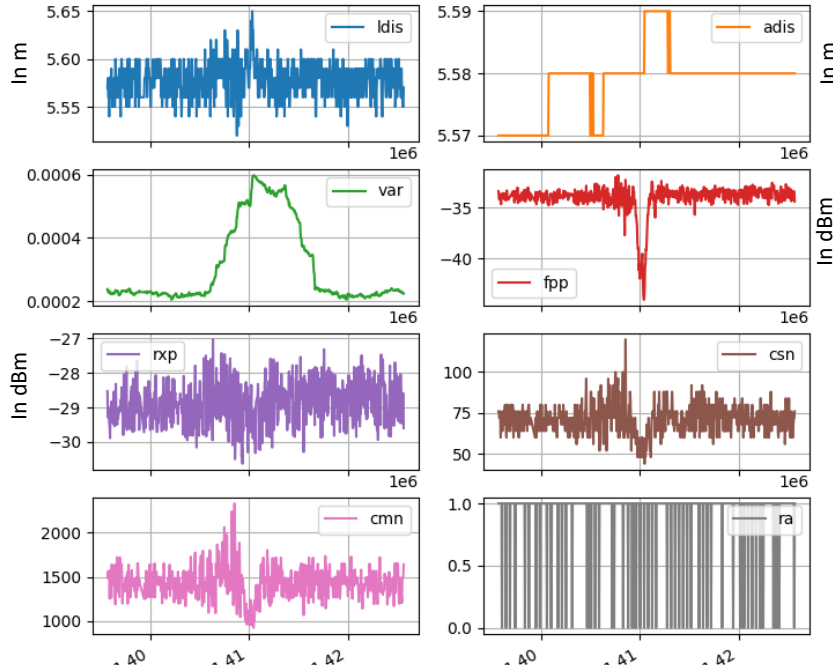
Schematische Darstellung Teilergebnis der Untersuchung



Detaillierte Darstellung der gestörten Verbindungen

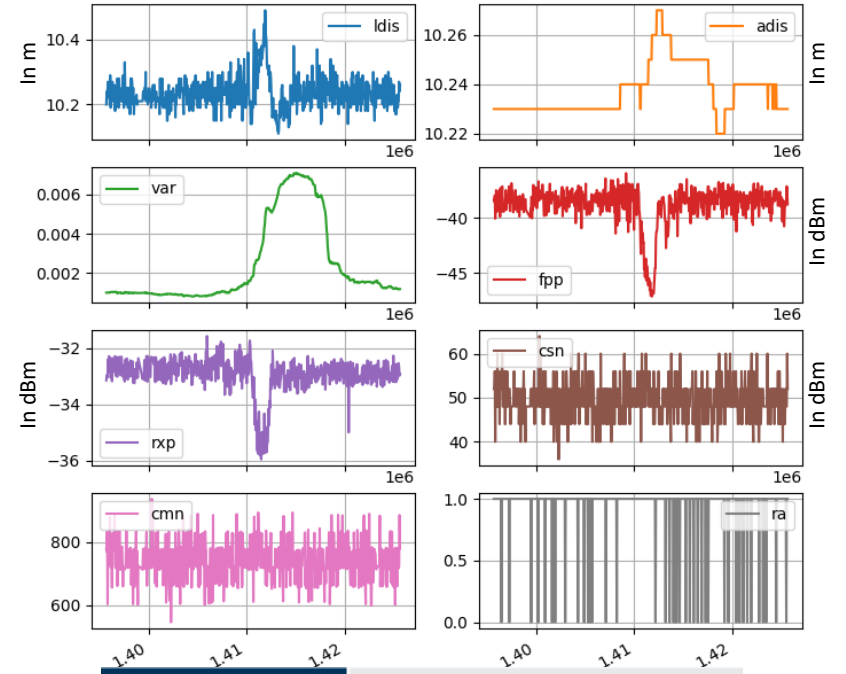
0x3432 Kugel 2

0x681f Flur Decke vor Büro 3



"ldis": 22.87,	aktueller Abstand zum Anchor in m
"adis": 22.55,	Mittelwert des Abstands zum Anchor seit Reset in m
"var": 0.0712,	„Varianz“ des Abstands zum Anchor

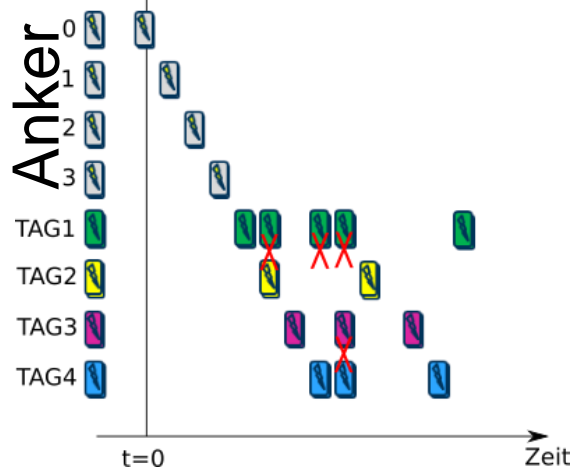
"ra": 1,	Messanchor, 1 bedeutet, Daten sind valide und werden zur Positionsbestimmung verwendet
----------	--



"fpp": -32.54,	First Path Power (siehe nächstes Kapitel)
"rxp": -30.17,	Received Power (siehe nächstes Kapitel)
"csn": 104.00,	Channel Standard Noise
"cmn": 2652.00,	Channel Maximum Noise

Verbesserung der RTLS Hardware

Routing-Problem der RTLS Informationen in großen RTLS-Netzwerken



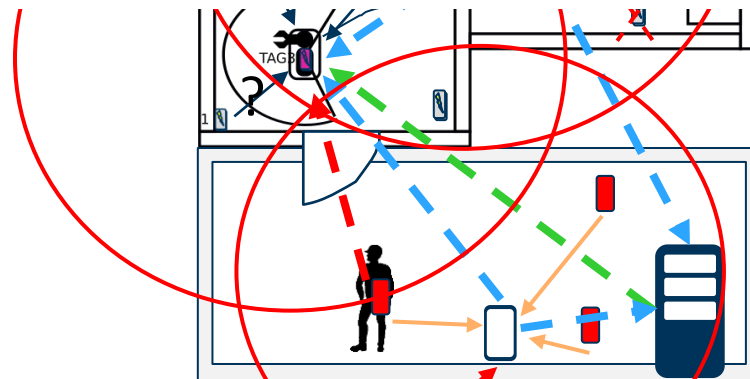
CSMA/CA
Carrier Sense Multiple Access/Collision
Avoidance

- ermöglicht die stabile
Kommunikation zwischen

Lösung: Hardware

RTL Erweiterung mit
Anbindung an ein WLAN-
Infrastruktur

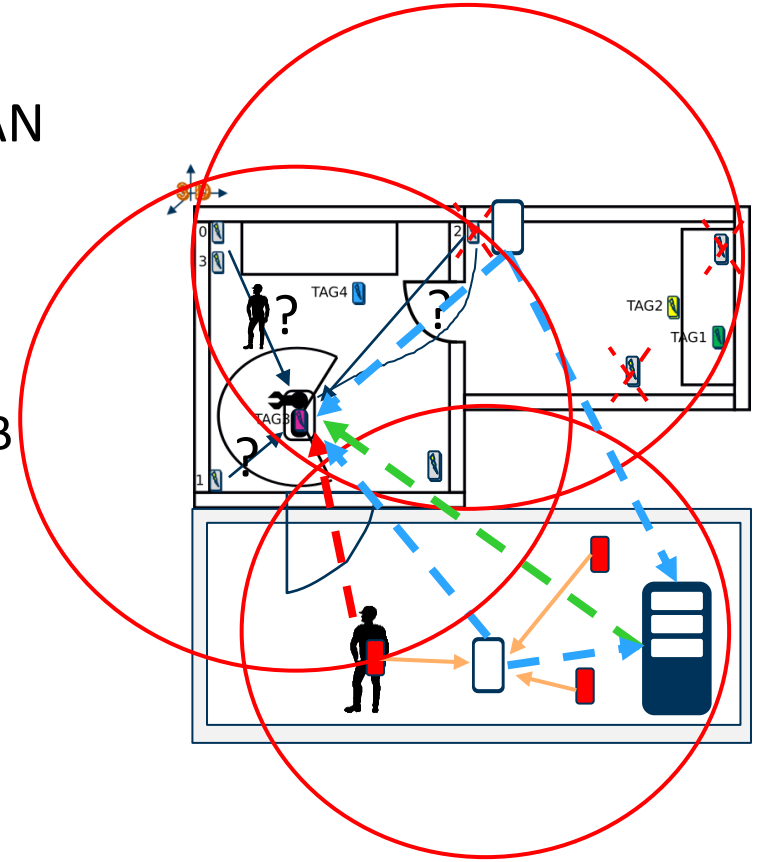
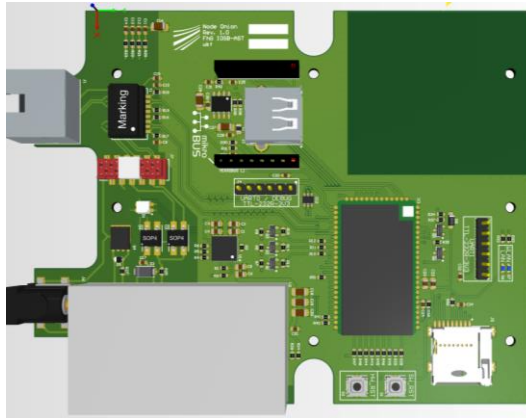
S
(U
WB
)



RTLS-ext
(UWB +

Optimierung der Hardware

- Erweiterung der Hardware mit WLAN
 - Neuen Schaltplan / Layout
 - Programmierung der Firmware
 - Tests der neuen Hardware
 - PoE, WLAN, USB, SD-Speicher, UWB



Veröffentlichungen / Studentische Arbeiten

- **Veröffentlichung:**
 - N. Fränzel, N. Greifzu, M. Schneider, A. Wenzel: Robuste Lokalisierung in drahtlosen Sensornetzwerken, Tag der Forschung, HSM-Print, 2020
 - Simulativer Vergleich bei der Positionsberechnung mittels Least Squares Schätzung und Taylorreihen-Entwicklung
 - Unter NLOS Bedingungen ist Least Squares Schätzung robuster
- **Projektarbeiten:**
 - Sasi Kumar Selvaraj: Implementation for RTLS localisation in ROS and rviz (abgebrochen)
 - Stephan Hintz, Adrian Herrman, Nicolas Schmitt: Messfehlerbestimmung von UWB Signalen zur Indoor-Navigation (fast fertig)

Ausblick auf das letzte Halbjahr

Projektziele	Offene Punkte (todos)	Zusätzliche erwartete Ergebnisse bis Projektende
Objekt- / Personenortung Reichweite bis 20m im Gebäude	<ul style="list-style-type: none"> • Genaue Auswertung der erreichbaren Genauigkeit im Versuchsaufbau 	<ul style="list-style-type: none"> • Reichweitenerweiterung über WLAN Infrastruktur im Gebäude
Firmware-optimierung zur Verbesserung der Positionsbestimmung	<ul style="list-style-type: none"> • Auswertung der Ergebnisse der Hardware-Erweiterung 	<ul style="list-style-type: none"> • Neue Signalparameter zur Qualitätsbestimmung der Verbindung • Untersuchung neuer Algorithmen zur Detektion von Signalstörungen
Anbindung an die Roboterplattform	<ul style="list-style-type: none"> • Aufbau der finalen Demonstrationsumgebung 	<ul style="list-style-type: none"> • Alle sichtbaren Teilnehmer werden dem Roboter dargestellt

Forscherguppe SONARO

Smarte Objektübernahme und –übergabe für die nutzerzentrierte mobile Assistenzrobotik

Demonstration des Standes zum Milestone MS4

Präsentierte Objekte können von Hand getrennt, klassifiziert und im Raum getrackt werden

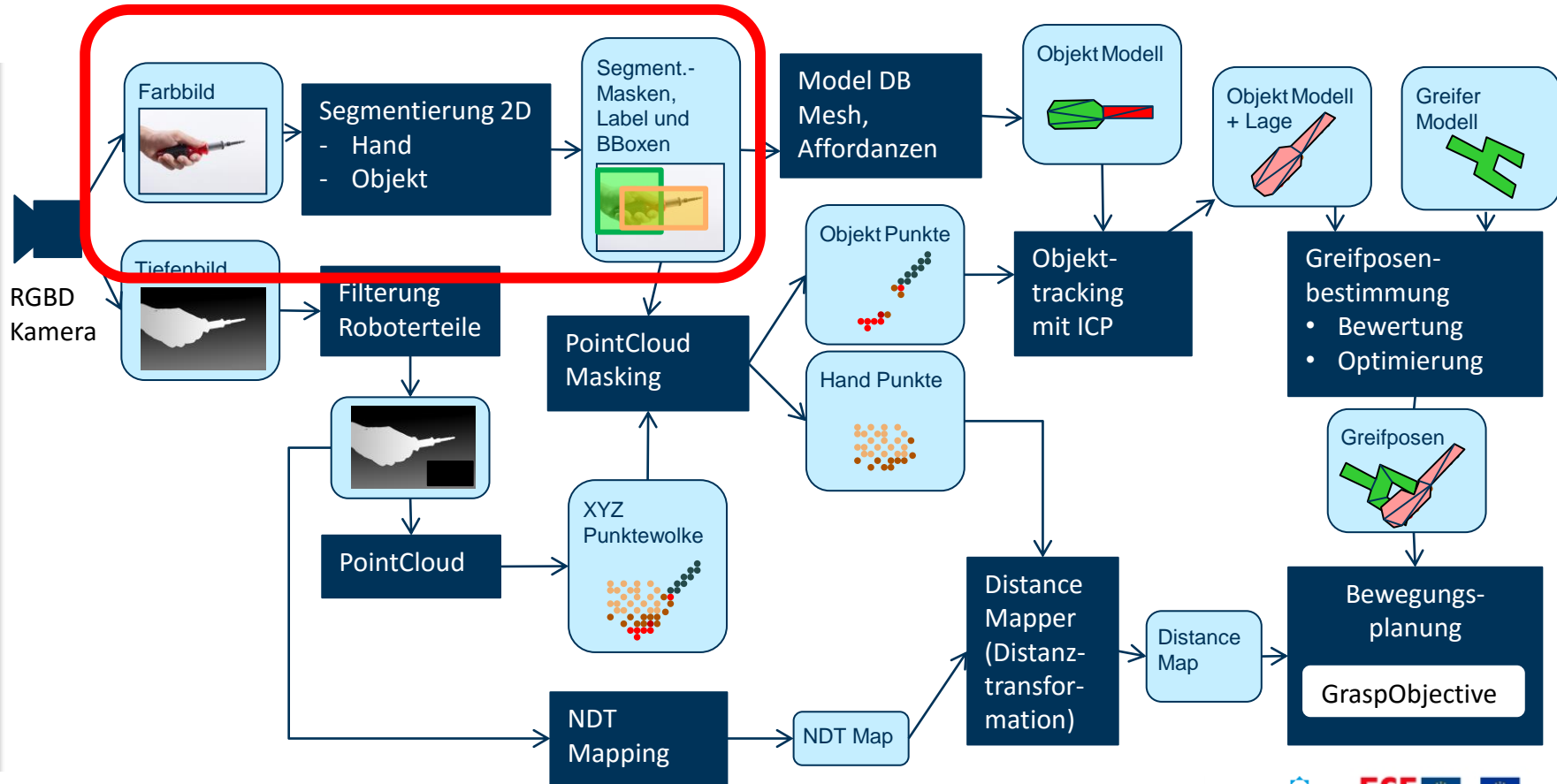
www.sonaro-projekt.de



Präzisierung der Grobplanung/Zielfunktionalitäten für die kommenden zwei Jahre

Milestones	Monat	Beschreibung
MS1 	01/2020	Roboter betriebsbereit, kann Personen wahrnehmen und autonom navigieren
MS2 	07/2020	Mobile RTLS-Flares sind funktionsfähig und in Demonstrator integriert
MS3 	07/2020	Multispektraler 3D-Sensor ist betriebsbereit und auf Roboter integriert
MS4 	01/2021	Präsentierte Objekte können von Hand getrennt, klassifiziert und im Raum getrackt werden
MS5	10/2021	Roboter kann im vollständigen Ablauf Gegenstände von Person A greifen und bei B wieder übergeben
MS6	12/2021	Die Leistungsfähigkeit des Systems ist evaluiert

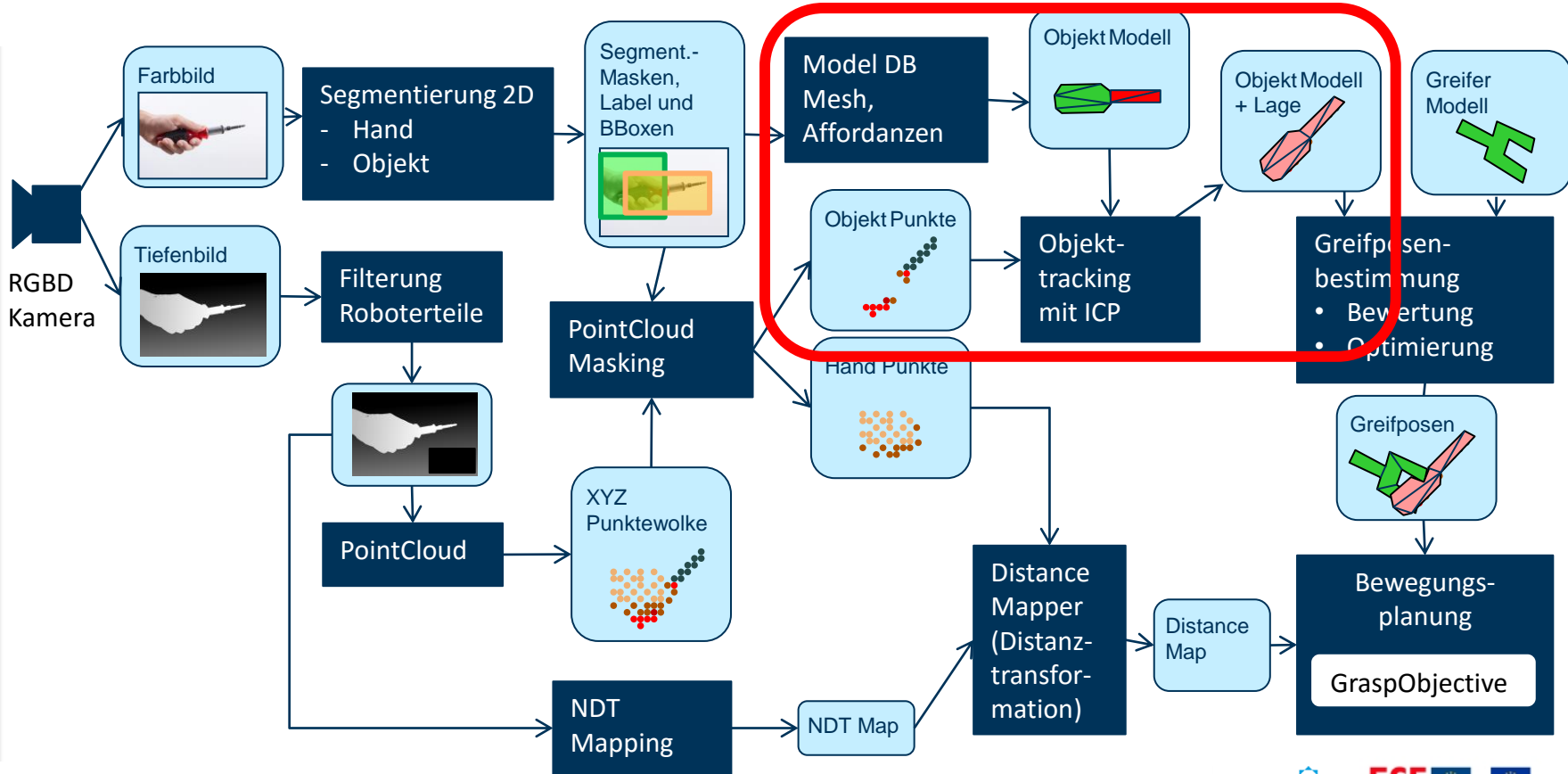
Datenverarbeitungspipeline



PointRend trainiert auf SONARO Datensatz (17 Objekte)

Segmentation of presented objects and holding hand
in 2D RGB video using the PointRend Network
trained on 17 objects of the SONARO HandObjects dataset

Datenverarbeitungspipeline



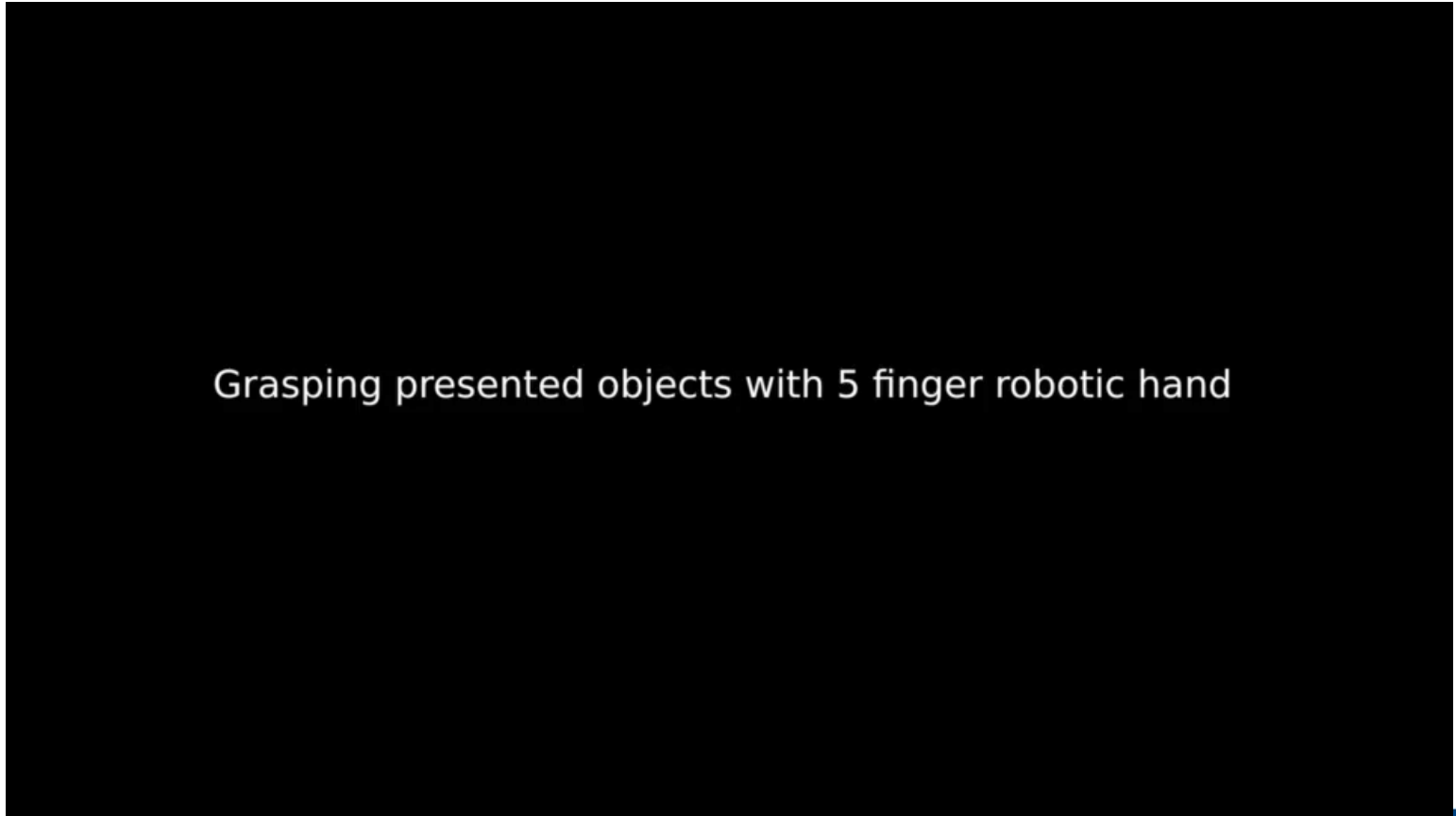
ICP Tracker im Livebetrieb

Tracking of known objects in 3d point cloud
Kamera: Asus Xtion pro live
Semantic segmentation: PointRend
(trained on objects held by human hand)
Matching: ICP
Tracking: Kalman Filter

Zugriff mit dem Tiago roboter (5-Finger Hand)

- Nur geeignet für stabförmige Gegenstände
- Stabile Greiftrajektorie hängt stark von einem robusten Tracking ab
- Probleme durch Verdeckungen mit der Roboterhand → Tracking springt und damit auch die Greifposen
 - In Zukunft Objekttracking auch an Skeletttracking binden (robuster gegen Teilverdeckung)
 - Sichtbarkeit der Zielobjekte bei Bahnplanung berücksichtigen (Roboterarm von der Seite annähern und nicht in Sichtrichtung)

Zugriff mit dem Tiago roboter (5-Finger Hand)



Grasping presented objects with 5 finger robotic hand

Agenda



9:00	Prof. Groß	Begrüßung der Teilnehmer und Einführung zu den Eckdaten des Projektes im 4 Halbjahr
9:10	Dr. Müller	Einführung des Demonstratorszenarios und der Verarbeitungspipeline für den Zugriff
9:25	M.Sc. Zhang	Fortschrittsbericht zu den Arbeitspaketen von TUI-QBV
9:45	Dr. Garten	Fortschrittsbericht zu den Arbeitspaketen von GFE
10:05	Dr. Müller	Fortschrittsbericht zu den Arbeitspaketen von TUI-NIKR
10:25	M.Sc. Schneider	Fortschrittsbericht zu den Arbeitspaketen von HSM
10:45	Dr. Müller	Demonstration des Roboters im Labor
11:05	Prof. Notni	Frei Diskussion der Ergebnisse mit dem Beirat
11:25	Dr. Djambova	Feedback des Projektträgers

Diskussionspunkte

- Erfolg der Förderpolitik - alle geförderten Roboter und Sensoren des ThZM Ilmenau Teils sind in SONARO als Forschungstechnik im Einsatz (Notni, ThZM)
- sehr gut aufgesetztes Projekt; bezüglich der Robustheit im algorithmischen Bereich hat das Projekt den State-of-Art erreicht; Robustheit vor Geschwindigkeit; Geschwindigkeit mit den vorhandenen Robotern nicht möglich (Richter, HRI)
- Rückfragen zur Bauform und Daten der eingesetzten Thermokameras (Richter, HRI)
- Rückfrage zur Lichtempfindlichkeit bei den gezeigten Greifprozessen? (Waldmann, Henkel & Roth)
 - Invarianzeigenschaften der Machine-Learning verfahren hängt stark von den Trainingsdaten ab. Mit großem Trainingsdatensatz sollten Beleuchtungsschwankungen kein Problem mehr darstellen
- Problematik der Erkennung von Händen in Thermobildern bei Einsatz von Handschuhen?
 - Thermo ist nicht der einzige Kanal für die Klassifikation, sondern lediglich unterstützend -> Trainingsdatensatz umfasst auch Handschuhaufnahmen
- Letztliche Nutzung der untersuchten Kameras aus der Studie des FG QBV? (TAB)
 - Rapide weiterentwicklung der Kameratechnik → Einsatz auch anderer verfügbarer Sensoren
 - Untersuchte Kamera ASUS Xtion steht stellvertretend für die Gruppe der Sensoren, welche mit Stereo & Musterprojektion arbeiten, Solche Kameras sind auch auf den Roboterplattformen integriert und werden genutzt.