

MIXED REALITY OHNE HOLOLENS?

Entwicklung eines Prototypen für SLAM



Agenda

- Einleitung
- Zielsetzung
- Methodik
- Stand der Technik
- Konzeption
- Test
- Ergebnisse
- Auswertung & Evaluation
- Fazit & Ausblick

RV-Continuum



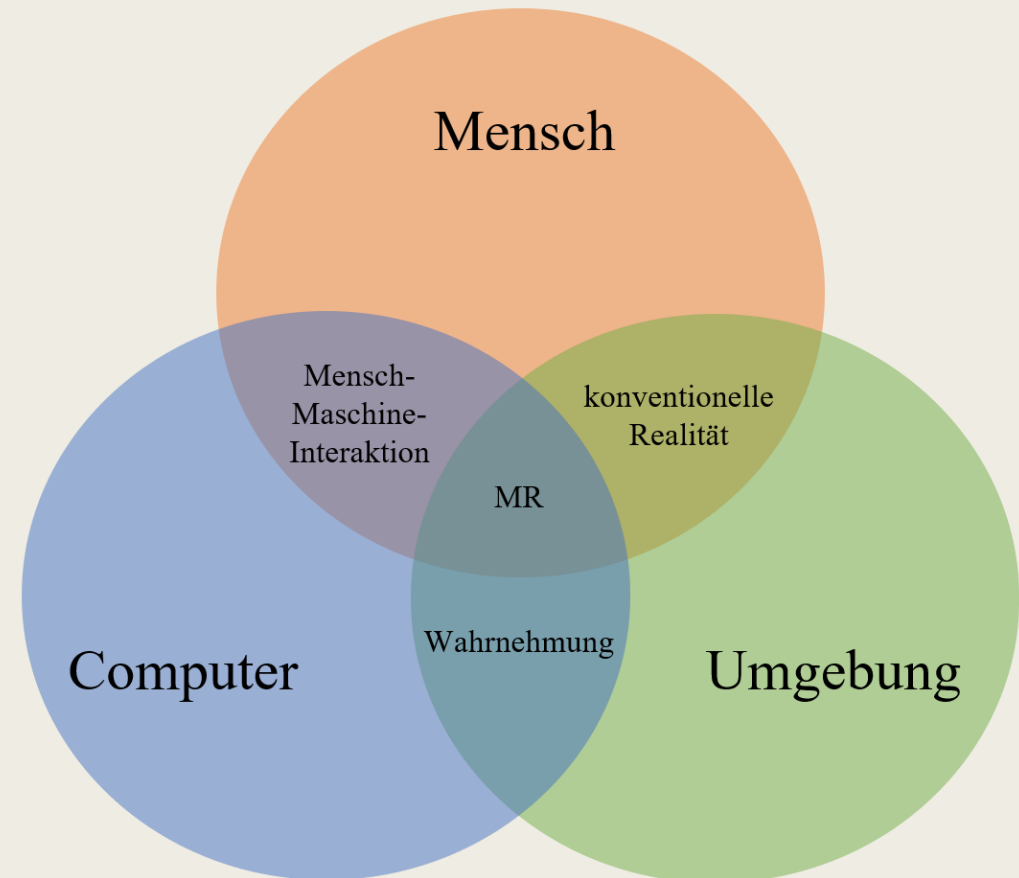
(Dörner et al., 2013)

- Beschreibt das Spektrum von der physischen Realität, bis zur vollständig virtuellen Umgebung
 - *Wichtige Definitionen: Augmented Reality und Augmented Virtuality*
 - *Die Natur des Spektrums lässt weitere Interpretationen der MR zu*

Windows Mixed Reality

- „Real“ Mixed Reality (Microsoft)
- Spatial Context-aware
- HoloLens als Input-/Output-Device
- Markerless Tracking
- Environment Understanding
- Spatial Mapping
- Pose-Estimation

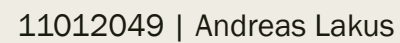
-> Selbstlokalisierung in unbekannten Raum



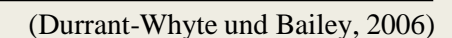
(Bray et al., 2021)

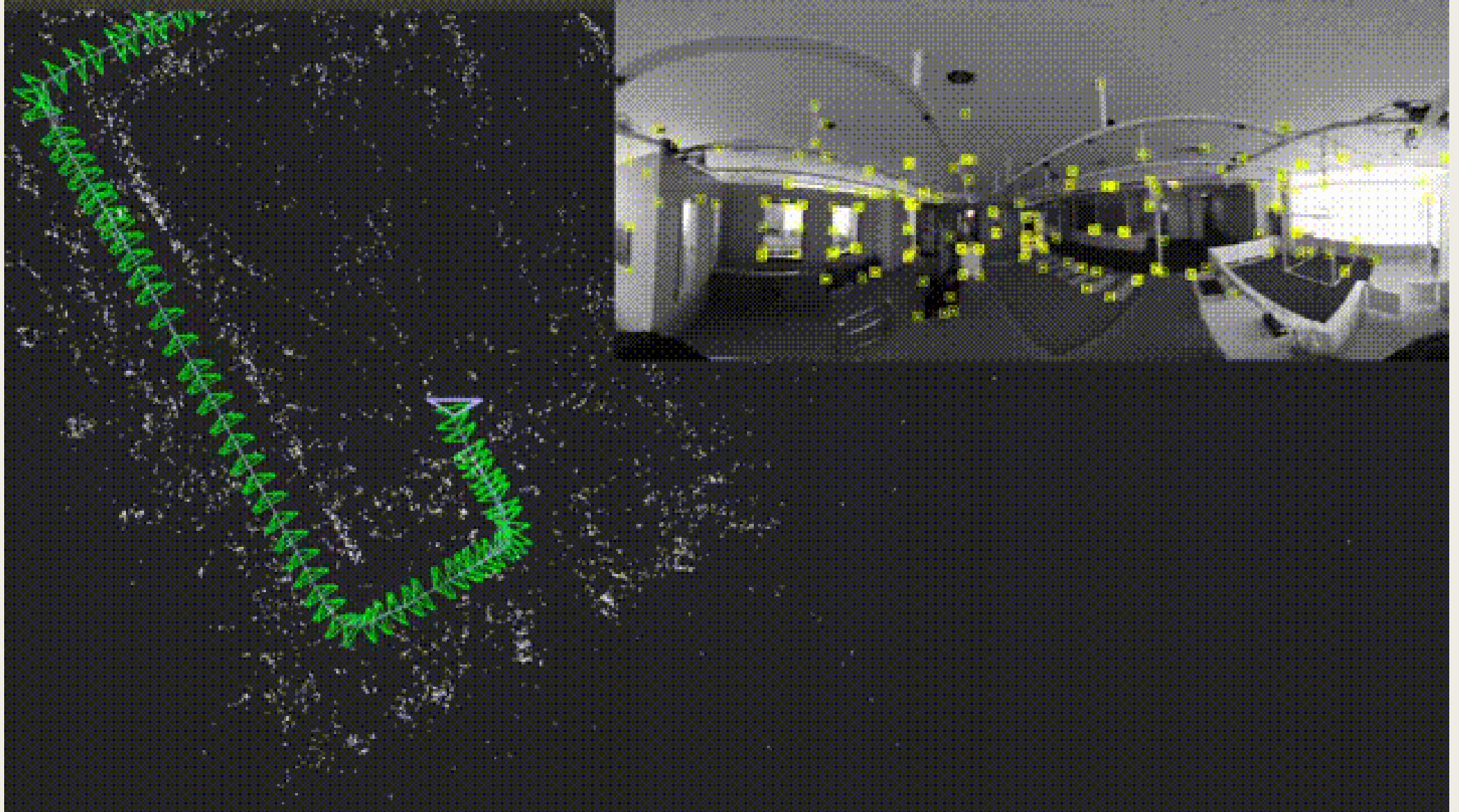
5

29.09.2021



- Local Pose-Estimation zeitgleich zur globalen Kartierung pro Zeiteinheit
- Wahre Trajektorie im Kontext MR schwer messbar
- Optische Messung immer fehlerbehaftet
- Optimierung bzw. Filterung der Messungen essentiell
- Roboter durch Mensch mit HMD ersetzbar





Quelle: <https://github.com/xdspacelab/openslam>

Probleme

HoloLens und Windows Mixed Reality

- Nicht Open-Source
- Durch Patente geschützt
 - Daher konkreter Algorithmus nicht einsehbar
 - Eingeschränkter Funktionsumfang in „Research-Mode“
 - Nur voll einsatzfähig in Verbindung mit Windows Mixed Reality

(Evans et al., 2017)

These

MR müsste in Verbindung mit SLAM, in Kombination mit Kameras realisierbar sein

Hypothese

Microsoft löst das SLAM-Problem mittels der HoloLens, insbesondere mit Ansätzen aus dem State of the Art bzgl. visual SLAM (vSLAM)

Zielsetzung

Entwicklung eines Prototypen zur Lösung des SLAM-Problems für MR-Anwendungen

Einschränkungen:

- Keine Entwicklung einer konkreten MR-App
- Nutzbar mit der Unity Engine für nachfolgende Forschung und Entwicklung
- Eine Kamera (kein viSLAM)
- Reine Machbarkeit
- Open Source bzw. Anpassbar

Methodik

- State of the Art vSLAM durch Literaturanalyse
- Konzeption
 - *Definition Einsatzgebiet des Algorithmus*
 - *Anforderungsergebnung*
 - *Identifikation einer Lösungsstrategie für vSLAM*
 - *Konzeption des Systemaufbaus*
- Realisierung
 - *Umsetzung des Systementwurfs*
- *Test ORB-SLAM-2-Native*
 - *TUM-Datasets zur Evaluation von ORB-SLAM-2*
- *Live-Test des AR-Demos*
- *ORB-SLAM-2-Native und ROS*
 - *Entfernungstest*
 - *Rekonstruktionstest*
 - *Live-Mapping*
- *Implementierung in der Unity Engine*

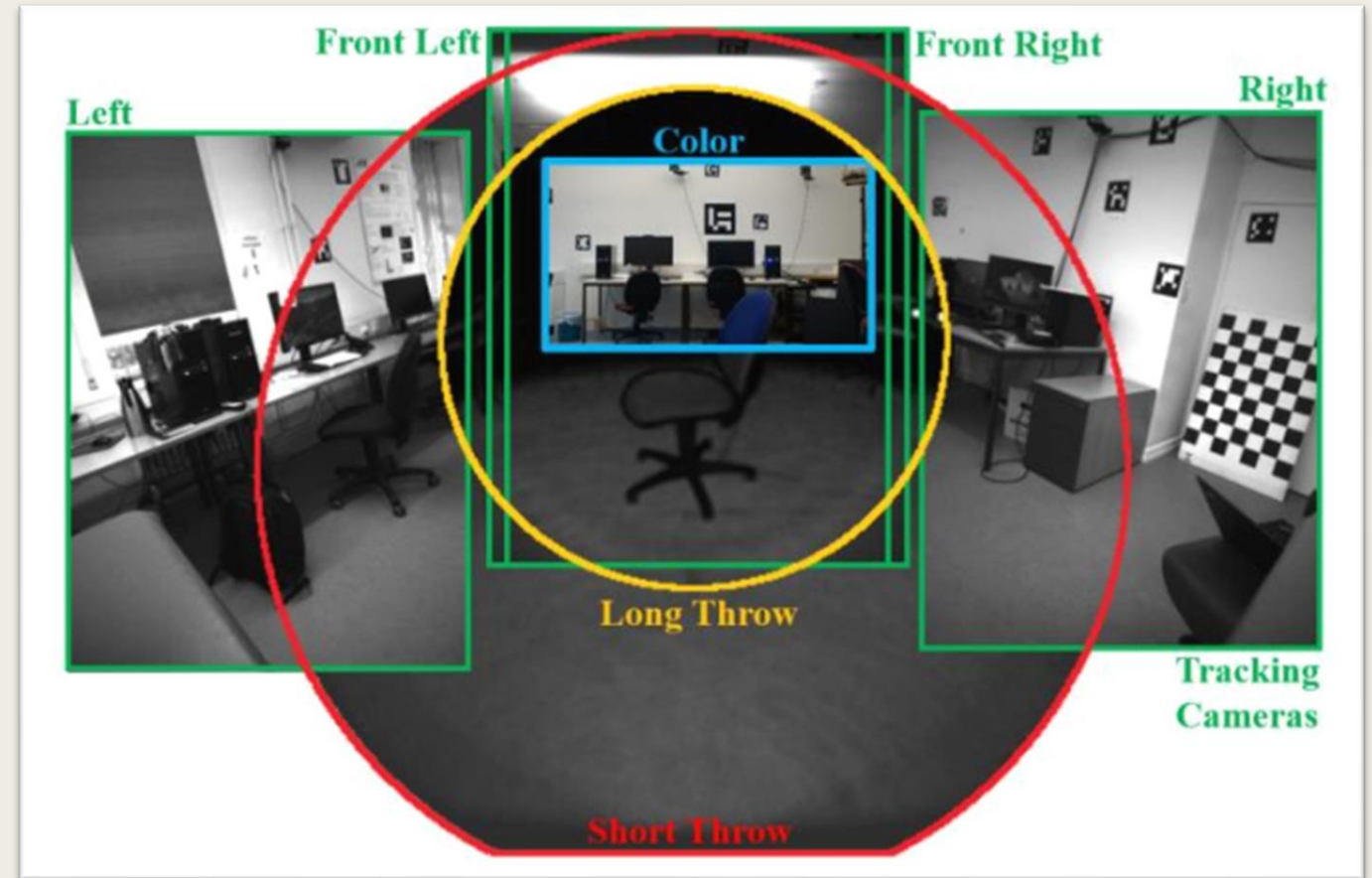
HoloLens



(Khoshelham, Tran und Acharya, 2019)

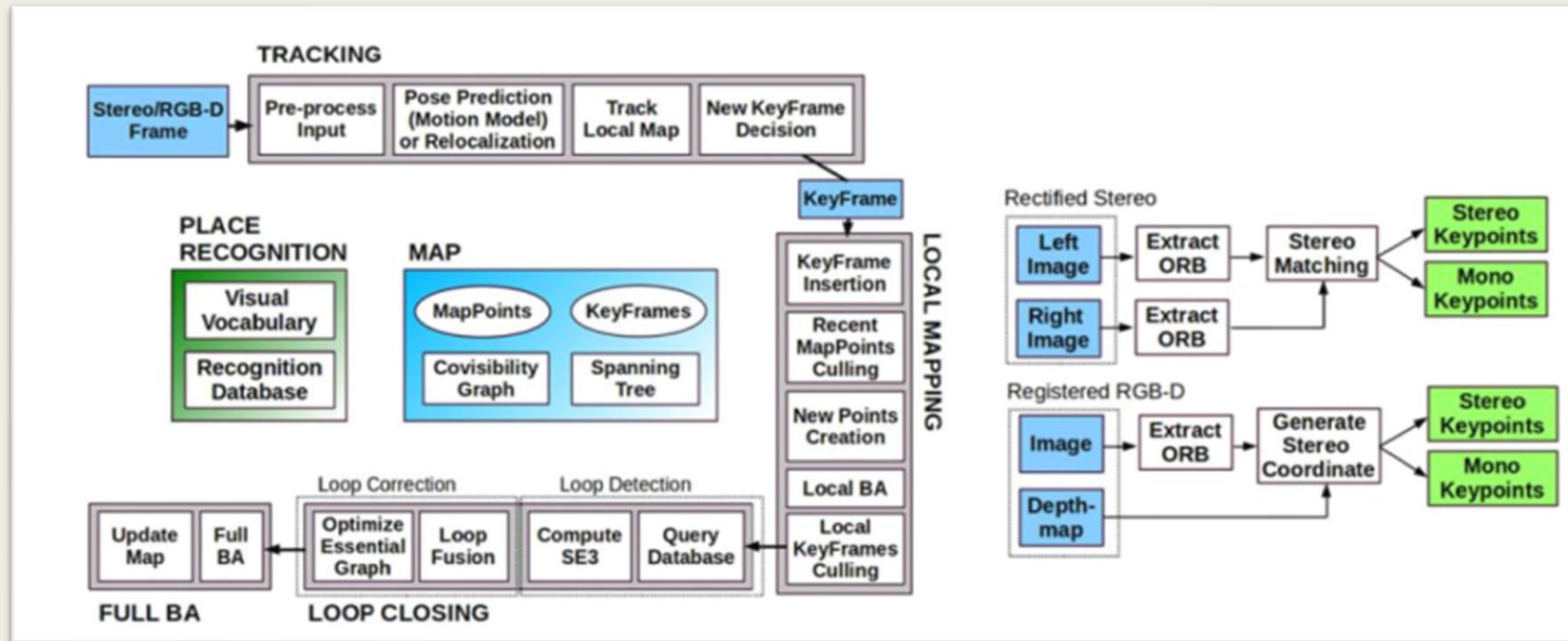
HoloLens

- Applikationen
- Tracking Cameras (RGB)
- kurze Entfernung
- mittlere Entfernung



(Hübner et al., 2020)

ORB-SLAM 2



(Mur-Artal und Tardos, 2017)

- Loop Closing
- Akkommodabel
- Robust (Innen, Außen)
- Echtzeit
- Open Source
- State of the Art
- Hohe Verbreitung
- Kompatibilität

Anforderungen: Funktional

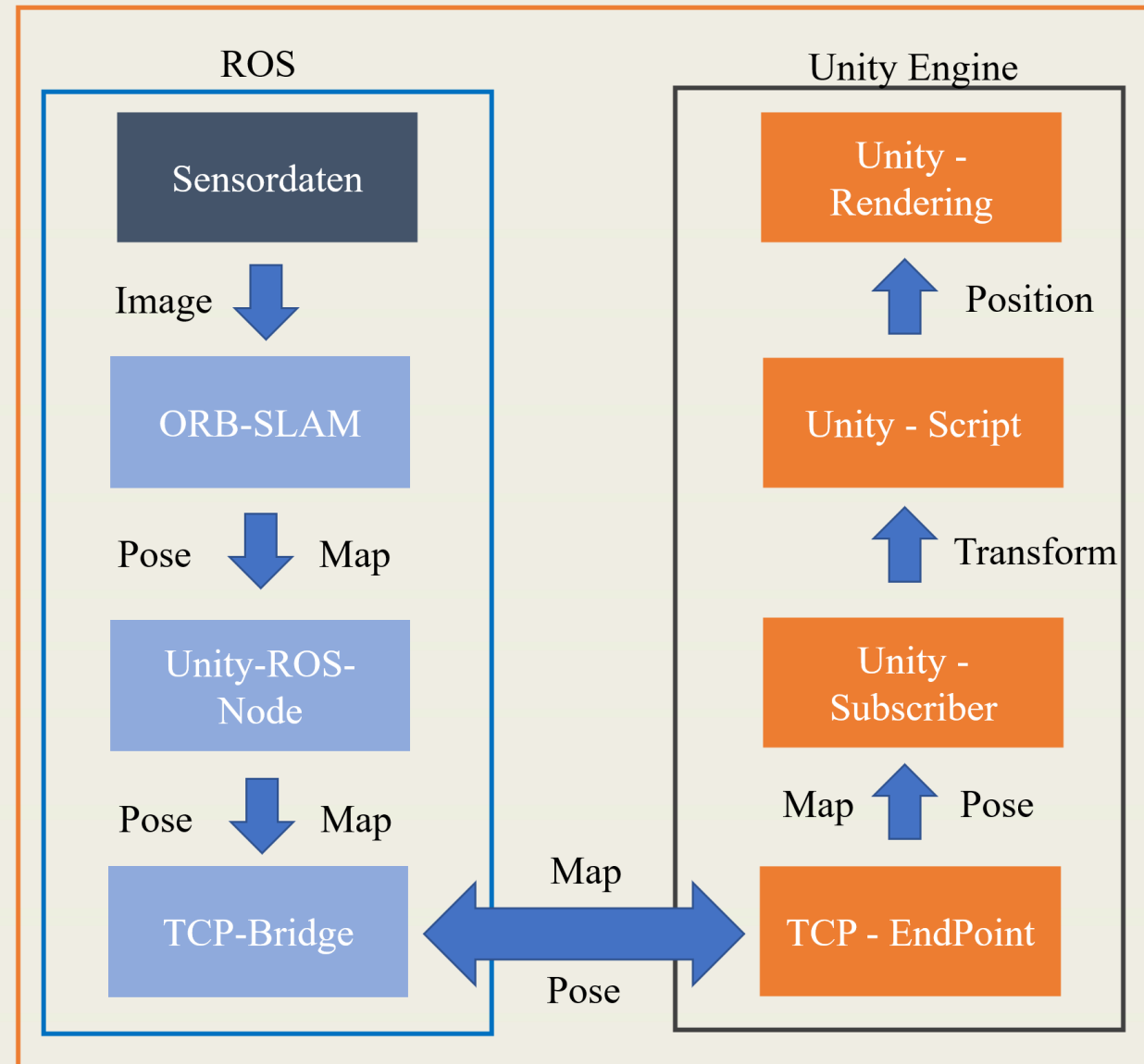
- *Auswertung von Bilddaten durch den SLAM-Algorithmus in Echtzeit*
- *Ausgabe von Positionsinformationen zur Raumgeometrie und Pose der Kamera*
- *Anbindung an die Unity Engine mit schneller Datenübertragung für Karten- und Positionsdaten*
- *Rendering in Unity Engine mit den Daten des SLAM-Algorithmus*
- *Echtzeit-Tracking natürlicher Merkmale*

Anforderungen: Nicht-funktional

- *Echtzeitfähigkeit*
- *Nachvollziehbare Dokumentation*
- *Modularer aufgebaut des SLAM-Algorithmus*
- *Anpassbar hinsichtlich Weiterentwicklung*
- *Entwicklung einer MR-Anwendung soll möglich sein*
- *Robuste Datenverarbeitung und Fehlertoleranz zur Vermeidung von Systemabstürzen*

Systementwurf

- Ubuntu 20.04
 - Plattform
- ROS Noetic
 - SLAM
 - Schnittstelle
- Unity Engine
 - Pose-Data
 - Rendering



- AR-Demo von ORB-SLAM-2
- Rotationstest
- Entfernungstest
- Rekonstruktion
- Live-Mapping

Diagram illustrating a 2D environment for robot navigation. The environment is defined by a horizontal axis (distance in cm) and a vertical axis (distance in cm).

Key elements and coordinates (approximate):

- Figur 2:** Red circle at (30cm, 30cm).
- Figur 1:** Green circle at (10cm, 10cm).
- Cube:** Gray square at (100cm, 150cm).
- Buch:** Green arrow pointing right at (150cm, 30cm).
- Pferd:** Brown horse icon at (100cm, 10cm).
- Initial Position of the Camera:** Blue arrow points to the origin (0,0).
- Angles:** A red arc indicates angles from -90° to 90° relative to the horizontal axis.

■ TUM-RPY

- TUM-XYZ

-
- The chart, titled "Fehlgeschlagene Initialisierungen", displays the number of failed initializations for two methods across five tests. The y-axis represents the count, ranging from 0 to 8. The x-axis lists the tests from Test 1 to Test 5. TUM-RPY (blue line) starts at 1 for Test 1, rises to 2 for Test 2, peaks at 7 for Test 3, drops to 2 for Test 4, and reaches 0 for Test 5. TUM-XYZ (red line) starts at 4 for Test 1, drops to 2 for Test 2, and remains at 2 for Tests 3, 4, and 5.
- | Test | TUM-RPY | TUM-XYZ |
|--------|---------|---------|
| Test 1 | 1 | 4 |
| Test 2 | 2 | 2 |
| Test 3 | 7 | 2 |
| Test 4 | 2 | 2 |
| Test 5 | 0 | 2 |

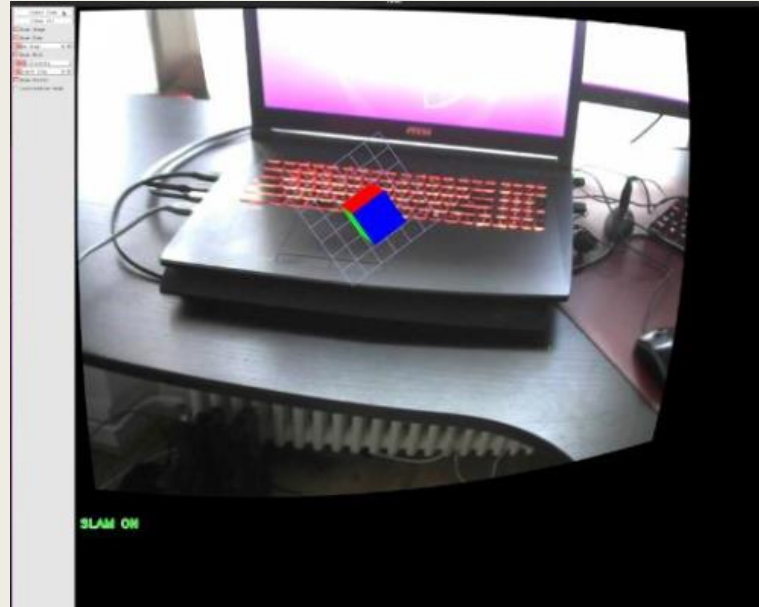
Nr.	Failed Initializations	Map-Points	Median tracking time (s)	Mean tracking time (s)	RMSE (cm)
1	80%	135	0,0188573	0,0211068	5,1052
2	66%	104	0,0187696	0,0209025	9,4834
3	66%	100	0,0183865	0,0204396	1,0138
4	66%	101	0,0200443	0,0219075	0,8699
5	66%	100	0,0186890	0,0207981	4,7984
Mean RMSE					4,25414
Median RMSE					4,7984

- 23

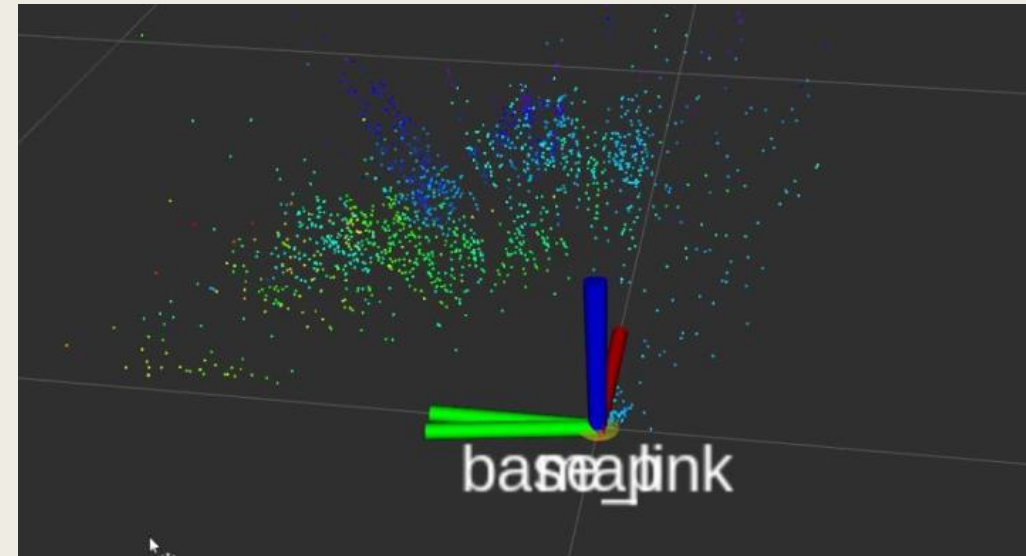
Nr.	Failed Initializations	Map-Points	Median tracking time (s)	Mean tracking time (s)	RMSE (cm)
1	50%	107	0,0153205	0,0159898	5,3008
2	66%	105	0,0164173	0,0181199	5,5013
3	87%	101	0,016211	0,0195156	3,9010
4	87%	115	0,0166927	0,0186322	5,1062
5	0%	121	0,0168280	0,0181311	3,6771
Mean RMSE					4,69728
Median RMSE					5,1062

- 24

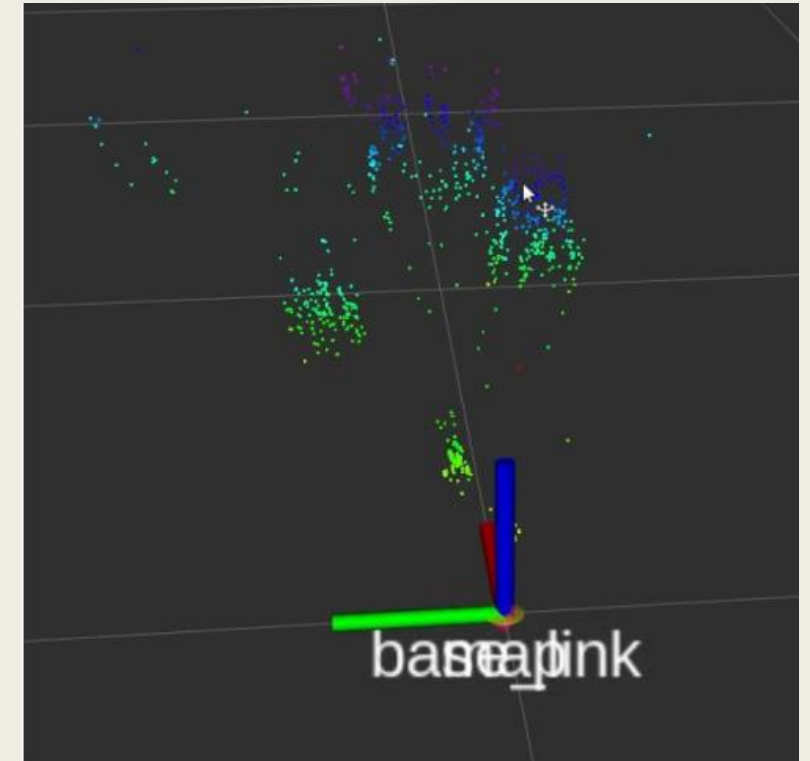
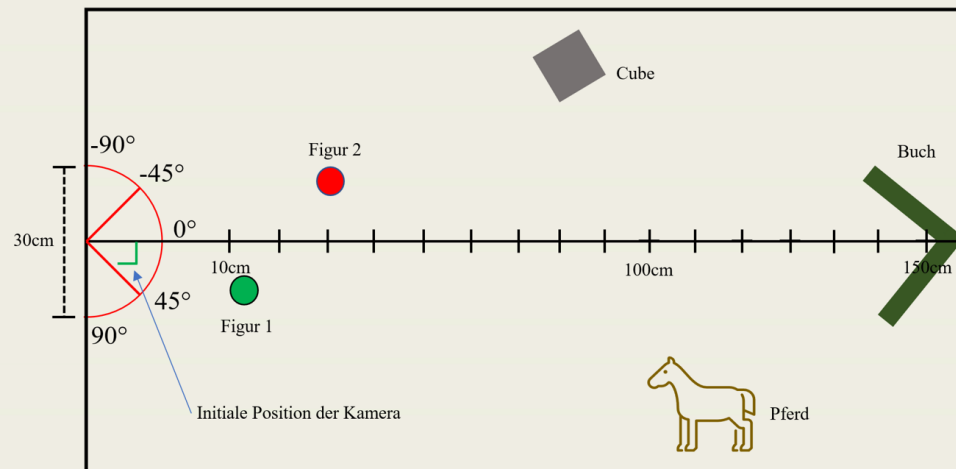
AR-Demo



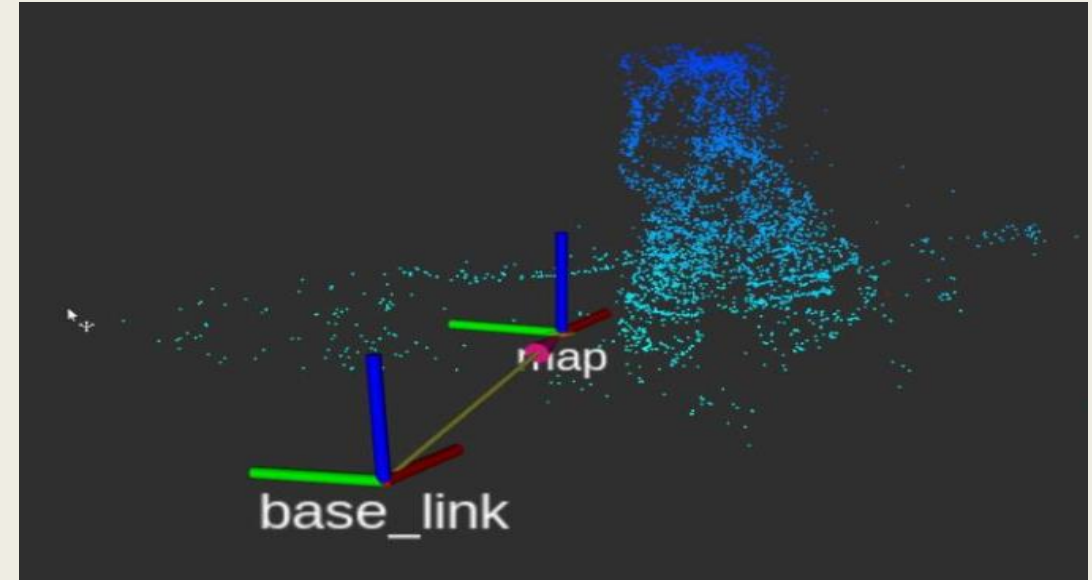
- Markerless Tracking konnte demonstriert werden
- Verankerung und 3D-Registratur konnte nachgewiesen werden
- Orientierung des virtuellen Objekts schwer planar realisierbar
- Plane-Detection durch RANSAC nicht ausreichend implementiert, da Demo



- 26

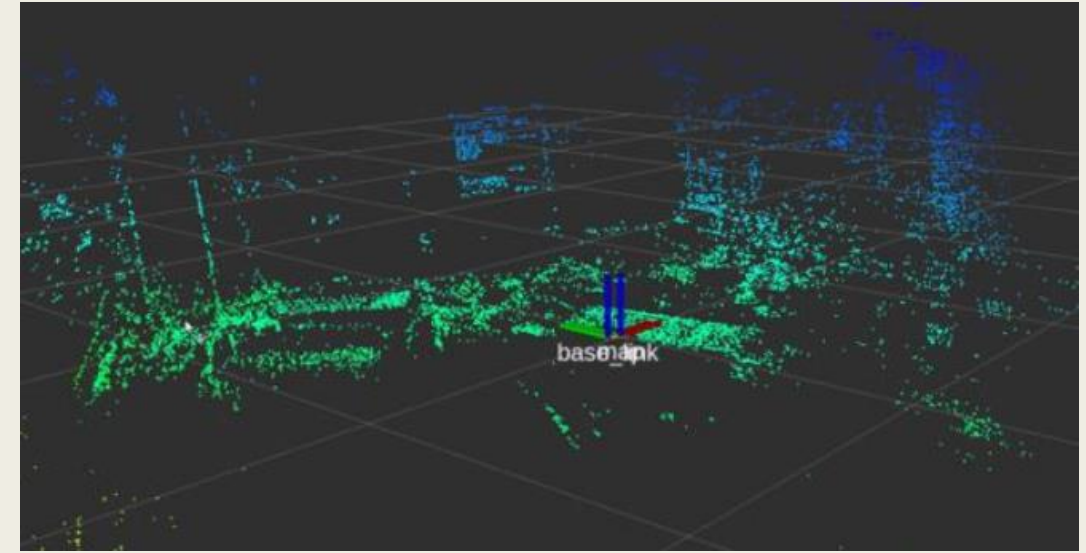


A 3D visualization of a point cloud. The point cloud is composed of many small dots, colored red and black. A green bounding box is drawn around a cluster of points on the left, and a red bounding box is drawn around a cluster of points on the right. The green bounding box is slightly tilted, while the red bounding box is more rectangular. The points are scattered throughout the space, with a higher density in the two main clusters.



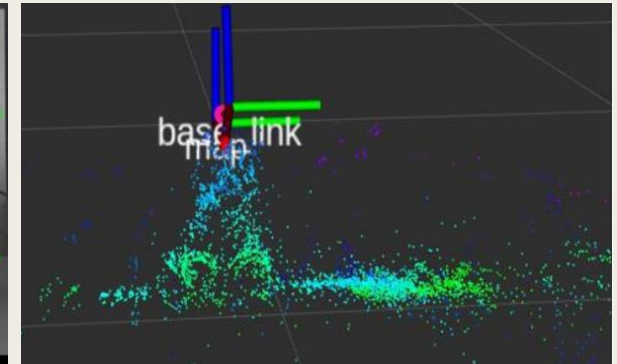
- ORB-SLAM-2-Native rekonstruiert das Objekt
- ORB-SLAM-2-ROS ermöglicht klarere Strukturen

A 3D point cloud visualization of a scene, likely a chair, with a green mesh overlay and a blue wireframe structure. The point cloud is composed of numerous black dots representing individual data points. The green mesh is a semi-transparent surface that follows the general shape of the chair. The blue wireframe structure is a skeletal representation of the chair, showing the underlying geometry. The background is white, and the overall image is a screenshot from a presentation or software interface.



- ORB-SLAM-2-Native rekonstruiert ähnliches Ergebnis wie ORB-SLAM-2-ROS

-



Vergleich TUM-Dataset

- Höhere RMSE als Vergleichsdaten
 - *Implementierung*
 - *Kalibrierung*
 - *Versionierung*
- RPY generell schlecht
 - *Präziser als Liu & Zhang*
 - *Zu wenige Daten um aussagekräftig zu sein*

Autor	Dataset	RMSE (cm)
Liu, Zhang et al. 2016	fr1_xyz	1,05
	fr1_rpy	5,53
Liu, Yulei et al. 2020	fr1_xyz	1,06
	fr1_rpy	-
Mur-Artal et al. 2015	fr1_xyz	0,90
	fr1_rpy	-
Implementierung	fr1_xyz	4,25414
	fr1_rpy	4,69728

(Liu, Zhang und Bao, 2016, Liu et al., 2020, Mur-Artal und Tardos, 2017)

Fazit & Ausblick

- Mixed Reality ohne HoloLens ist realisierbar
 - *SLAM im Kontext MR wurde demonstriert und durch Tests validiert*
- ORB-SLAM-2 stellt nicht die optimale Lösung dar, da die Initialisierung, respektive rotierende Bewegungen schwer realisierbar sind bzw. abhängig von der Umgebung sind
 - *Stereosensoren in Kombination mit IMU bzw. Sensorfusion*
- Alternative zu ROS
 - *Direkte Integration von SLAM-Algorithmus als statische/dynamische Bibliothek als sog. native Plugin für die Unity Engine inklusive Rendering Pipeline für AR*
- Planare Projektion nur schwer realisierbar durch RANSAC
 - *AI-Object Detection*
 - *RANSAC neu implementieren anhand State of the Art*
- Alternativen reiner vSLAM-methoden für MR kaum vorhanden, werden jedoch erwartet

Quellen

■ F3:

- MILGRAM, P., H. TAKEMURA, A. UTSUMI und F. KISHINO, 1995. *Augmented reality: a class of displays on the reality-virtuality continuum*. In: H. DAS, Hg. *Telemanipulator and Telepresence Technologies*. Bellingham Washington: SPIE, S. 282-292
- DÖRNER, R., W. BROLL, P. GRIMM und B. JUNG, 2013. *Virtual und Augmented Reality (VR / AR)*. Berlin, Heidelberg: Springer Berlin Heidelberg

■ F4:

- BRAY, B., 2021. *Was ist Mixed Reality? - Mixed Reality* [online]. 1. Juli 2021 [Zugriff am: 21. Januar 2021]. Verfügbar unter: <https://docs.microsoft.com/de-de/windows/mixed-reality/discover/mixed-reality>

■ F7:

- DURRANT-WHYTE, H. und T. BAILEY, 2006. *Simultaneous Localization and Mapping: part I* [online]. *IEEE Robotics & Automation Magazine*, **13**(2), 99-110. ISSN 1070-9932. Verfügbar unter: doi:10.1109/MRA.2006.1638022

■ F9:

- EVANS, G., J. MILLER, M. IGLESIAS PENA, A. MACALLISTER und E. WINER, 2017. *Evaluating the Microsoft HoloLens through an Augmented Reality assembly application*. In: J.N. SANDERS-REED und J.J. ARTHUR, Hg. *Degraded Environments: Sensing, Processing, and Display 2017*: SPIE, 101970V.

■ F13:

- KHOSHELHAM, K., H. TRAN und D. ACHARYA, 2019. *Indoor mapping exewear: geometric evaluation of Spatial Mapping capability of HoloLens* [online]. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **XLII-2/W13**, 805-810. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Verfügbar unter: doi:10.5194/isprs-archives-XLII-2-W13-805-2019

■ F14:

- HÜBNER, P., K. CLINTWORTH, Q. LIU, M. WEINMANN und S. WURSTHORN, 2020. *Evaluation of HoloLens tracking and depth sensing for indoor mapping applications* [online]. *Sensors (Basel, Switzerland)*, **20**(4). *Sensors (Basel, Switzerland)*. Verfügbar unter: doi:10.3390/s20041021

Quellen

■ F15:

- MUR-ARTAL, R. und J.D. TARDOS, 2017. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras [online]. *IEEE Transactions on Robotics*, **33**(5), 1255-1262. ISSN 1552-3098. Verfügbar unter: doi:10.1109/TRO.2017.2705103

■ F19:

- STURM, J., 2021. *Computer Vision Group - Submission form for automatic evaluation of RGB-D SLAM results* [online] [Zugriff am: 13. September 2021]. Verfügbar unter: https://vision.in.tum.de/data/datasets/rgbd-dataset/online_evaluation
- STURM, J., N. ENGELHARD, F. ENDRES, W. BURGARD und D. CREMERS, 2012. A benchmark for the evaluation of RGB-D SLAM systems. In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*: IEEE, S. 573-580.

■ F31:

- LIU, H., G. ZHANG und H. BAO, 2016. Robust keyframe-based Monocular SLAM for Augmented Reality. In: *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*: IEEE, S. 1-10
- LIU, J., Y. XIE, S. GU und X. CHEN, 2020. A SLAM-based mobile Augmented Reality tracking registration algorithm [online]. *International Journal of Pattern Recognition and Artificial Intelligence*, **34**(01), 2054005. ISSN 0218-0014. Verfügbar unter: doi:10.1142/S0218001420540051
- MUR-ARTAL, R. und J.D. TARDOS, 2017. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras [online]. *IEEE Transactions on Robotics*, **33**(5), 1255-1262. ISSN 1552-3098. Verfügbar unter: doi:10.1109/TRO.2017.2705103

Kolloquium zur Thesis

MIXED REALITY OHNE HOLOLENS?

Entwicklung eines Prototypen für SLAM