

Proseminar Objektposenschätzung

Robert Jeutter

3. November 2021

Zur Interaktion mit seiner Umwelt muss ein Roboter die Lage der Objekte in seiner Umgebung erkennen. Der Artikel schafft einen Überblick über aktuelle Verfahren mit Fokus auf Verfahren die keine Objektmodelle benötigen.

1 Motivation

Roboter in Industrie und Assistenz treffen häufig auf nicht vorhersehbare oder vorab programmierbare Herausforderungen. Damit Die Schätzung der 6D-Position bekannter Objekte ist wichtig für die Interaktion von Robotern mit der realen Welt wichtig. Das Problem ist aufgrund der Vielfalt der Objekte sowie der Komplexität einer Szene, die durch Unordnung und Verdeckungen zwischen den Objekten verursacht wird, eine Herausforderung.

2 Anforderungen

Objektmodelle, schablonenbasierte Methoden und merkmalsbasierte Methoden

Hardware-Ausstattung: 2D/3D Kamera, RGB-D, positionsveränderung der Kamera

Verarbeitung: Komplexität, Geschwindigkeit, Genauigkeit

Verkettete Verarbeitung: verarbeitung eines Bilderstroms statt einzelner Bilder für sich

3 Verschiedene Verfahren

3.1 BundleTrack[WB21]

3.2 DeepIM[Li+18]

3.3 MaskFusion[RBA18]

3.4 Neural Analysis-by-Synthesis[Che+20]

3.5 6-PACK[Wan+19]

3.6 PoseCNN[Xia+17]

neues Convolutional Neural Network für die 6D-Objektposenschätzung. PoseCNN schätzt die 3D-Verschiebung eines Objekts, indem es sein Zentrum im Bild lokalisiert und seinen Abstand zur Kamera vorhersagt. Die 3D-Rotation des Objekts wird durch Regression auf eine Quaternion-Darstellung geschätzt. Dabei führt man eine neue Verlustfunktion ein, die es PoseCNN ermöglicht, symmetrische Objekte zu behandeln.

3.7 Robust Gaussian Filter[Iss+16]

4 Vergleich verschiedener Verfahren

5 Fazit

Glossar

Convolutional Neural Network Besitzt pro Convolutional Layer mehrere Filterkerne, sodass Schichten an Feature Maps entstehen, die jeweils die gleiche Eingabe bekommen, jedoch aufgrund unterschiedlicher Gewichtsmatrizen unterschiedliche Features extrahieren.. 1

Quaternion-Darstellung Darstellung V einer Gruppe G , die einen G -invarianten Homomorphismus $J : V \rightarrow V$ besitzt, der antilinear ist und $J^2 = -Id$ erfüllt.. 1

Literatur

- [Che+20] Xu Chen u. a. *Category Level Object Pose Estimation via Neural Analysis-by-Synthesis*. Aufgerufen 27.10.2021. 2020. arXiv: [2008.08145](https://arxiv.org/abs/2008.08145). URL: arxiv.org/abs/2008.08145.
- [Iss+16] Jan Issac u. a. „Depth-based object tracking using a Robust Gaussian Filter“. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)* (März 2016). Aufgerufen 27.10.2021. DOI: [10.1109/icra.2016.7487184](https://doi.org/10.1109/icra.2016.7487184). URL: dx.doi.org/10.1109/ICRA.2016.7487184.
- [Li+18] Yi Li u. a. „DeepIM: Deep Iterative Matching for 6D Pose Estimation“. In: *International Journal of Computer Vision* 128.3 (Nov. 2018). Aufgerufen 16.10.2021, S. 657–678. ISSN: 1573-1405. DOI: [10.1007/s11263-019-01250-9](https://doi.org/10.1007/s11263-019-01250-9). URL: arxiv.org/abs/1804.00175.
- [RBA18] Martin Rünz, Maud Buffier und Lourdes Agapito. *MaskFusion: Real-Time Recognition, Tracking and Reconstruction of Multiple Moving Objects*. Aufgerufen 27.10.2021. 2018. arXiv: [1804.09194](https://arxiv.org/abs/1804.09194). URL: arxiv.org/abs/1804.09194.
- [Wan+19] Chen Wang u. a. *6-PACK: Category-level 6D Pose Tracker with Anchor-Based Keypoints*. Aufgerufen 27.10.2021. 2019. arXiv: [1910.10750](https://arxiv.org/abs/1910.10750). URL: arxiv.org/abs/1910.10750.
- [WB21] Bowen Wen und Kostas Bekris. *BundleTrack: 6D Pose Tracking for Novel Objects without Instance or Category-Level 3D Models*. Website. Aufgerufen 23.10.2021. 2021. arXiv: [2108.00516](https://arxiv.org/abs/2108.00516). URL: arxiv.org/abs/2108.00516.
- [Xia+17] Yu Xiang u. a. *PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes*. Website. Aufgerufen 16.10.2021. 2017. arXiv: [1711.00199](https://arxiv.org/abs/1711.00199). URL: arxiv.org/abs/1711.00199.