Bei diesem Beitrag handelt es sich um einen wissenschaftlich begutachteten und freigegebenen Fachaufsatz („reviewed paper“).

doi.org/10.37544/1436-4980-2023-03-6

Entwicklung eines Objekterkennungssystems für die Erzeugung und Pflege digitaler Umgebungsmodelle

Flexible Objektdetektion in 3D-Lidar-Scans

M. Willenbrink, D. Karelina, I. Effenberger

Eine wachsende Zahl von Unternehmen setzt auf Digitalisierung, um ihre Prozesse zu optimieren und angesichts zunehmender Unsicherheiten resilienter zu werden. Die Erzeugung und Pflege der dazu eingesetzten digitalen Modelle ist aber mit hohem Aufwand verbunden. Eine flexible KI-basierte Objekterkennung, die nur geringe Ansprüche an die Datenbasis stellt und dem Anwender außerdem Korrekturmöglichkeiten bietet, kann als Schlüsseltechnologie helfen, die bestehenden Hindernisse zu überwinden und die Transformation zu effizienteren Prozessen zu beschleunigen.

STICHWÖRTER

Automatisierung, Digitalisierung, Künstliche Intelligenz (KI)

Flexible object detection in 3D lidar scans – Development of an object recognition system for creating and updating digital models

More and more companies focus on digitization to optimize their processes and become more resilient in the light of increasing uncertainties. However, generating and updating digital models used for this purpose involves a significant amount of effort. Flexible AI-based object detection places fewer demands on the database while offering additional user correction options. It is the key to overcoming the barriers and accelerating the transformation towards more efficient processes.

1 Motivation

Fachkräftemangel, stockende Lieferketten, kürzere Produktlebenszyklen – dies sind die Herausforderungen, mit denen sich Unternehmen zunehmend konfrontiert sehen. In einer Studie zum Stand der Digitalisierung wurden von Pricewaterhouse-Coopers verschiedene Maßnahmen zur Stärkung der Resilienz im Betrieb vorgestellt [1]. Ein digitales Rückgrat wird darin als Grundvoraussetzung für eine zukunftsgerichtete Transformation identifiziert. Mit digitalen Zwillingen kann die physische Realität in digitalen Modellen abgebildet werden und durch die erweiterten Steuerungs- und Simulationsmöglichkeiten ein präziseres Management erfolgen.

Auch wenn die Notwendigkeit einer effektiven Digitalisierung erkannt worden ist, stockt in vielen Unternehmen die Umsetzung. Dies liegt nicht zuletzt am erheblichen Aufwand, der bei der Erstellung und Pflege der digitalen Datengrundlage betrieben werden muss. Allein die initiale Bestandsdigitalisierung kann Monate in Anspruch nehmen. Dabei kann heute mit Lidar-Scannern in wenigen Minuten eine auf wenige Millimeter genaue Erfassung der Umgebung stattfinden [2]. Die so entstandenen Punktwolken und Panoramabilder können anschließend in digitale geometrische Modelle überführt werden.

Gerade dieser letzte Schritt ist noch immer zeit- und arbeitsaufwendig und erfolgt größtenteils manuell [3]. Unterstützende Tools beschränken sich häufig auf die automatische Erkennung einfacher Elemente wie Wände [4]. Sollen aber weitere Entitäten in der Punktwolke, beispielsweise Werkzeugmaschinen, digitalisiert werden, müssen diese händisch in das Modell eingepflegt werden. Um diesen Prozess zu beschleunigen und die Anwender bei der repetitiven Arbeit zu unterstützen, wird ein flexibles Verfahren zur automatischen Objekterkennung in industriellen Umgebungen vorgestellt. Durch die Identifikation und exakte Lokalisierung von Objekten in der aufgenommenen Szene können in Zukunft digitale geometrische Modelle mit geringerem Zeit- und Personalaufwand erstellt und in höherer Regelmäßigkeit aktualisiert werden.

2 Stand der Technik

Mit Lidar-Scannern aufgenommene Punktwolken lassen sich nutzen, um geometrische Modelle für digitale Zwillinge zu erzeugen. Der Transfer der unstrukturierten Daten zu attribuierten digitalen Modellen erfordert jedoch erheblichen Arbeitsaufwand [3]. Softwareprodukte wie „Autodesk ReCap“, „Leica Cyclone“ oder „Faro Scene“ bieten verschiedene Werkzeuge zur automatischen Registrierung von Teilpunktwolken, zum Post-Processing und zur (semi-)automatischen Objekterkennung. Diese Objekterkennungsverfahren bieten Möglichkeiten zur Erkennung von Strukturelementen wie Boden, Wänden und Decken [4]. Zunehmend werden auch kleinere Objekte in den Erkennungskatalog aufgenommen, wie etwa Rohre, Lüftungsschächte oder andere Elemente häufig verbauter Gebäudetechnik [5]. Technisch setzen diese Verfahren meist auf klassische Methoden, wie „Ransac“ oder „Region Growing“, um die geometrisch einfachen Objekte (Quader oder Zylinder) zu detektieren [6–8].

In der Forschung gibt es inzwischen deutlich leistungsfähigere Ansätze zur Objektdetektion, die primär mit Machine-Learning-

https://doi.org/10.37544/1436-4980-2023-03-6 https://www.inlibra.com/de/page/agb - Open Access - CC BY

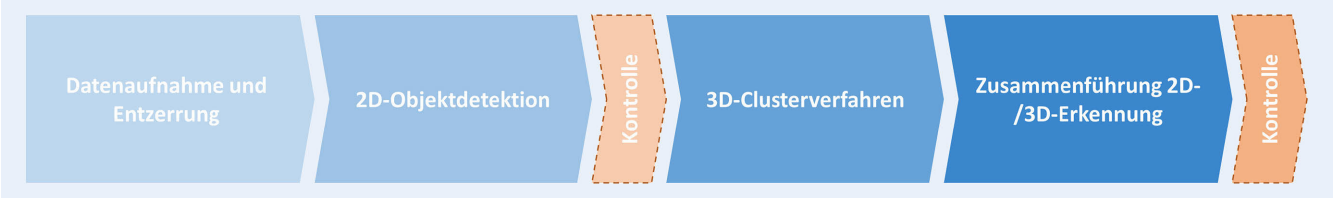

Bild 1. Verfahrensschritte einer flexiblen Objektdetektion mit multimodalen Daten. *Grafik: Fraunhofer IPA*

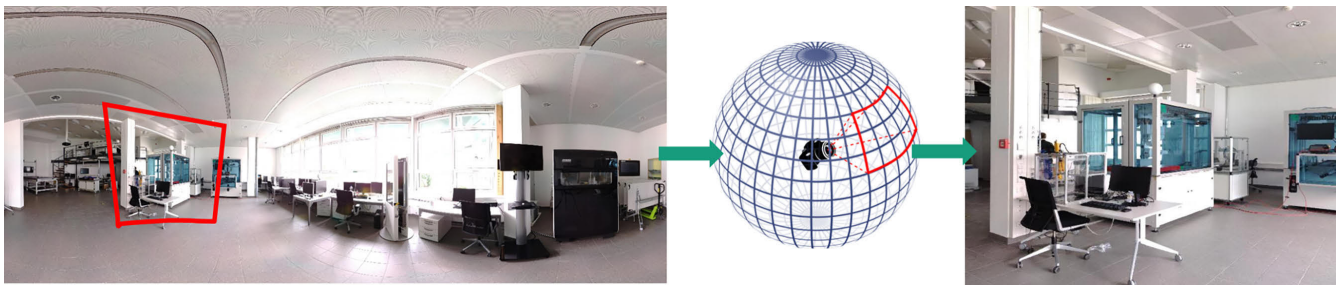

Bild 2. Entzerrung von 360°-Panoramabildern in Kugelprojektion zu Kacheln gewöhnlicher Rektangularprojektion. *Foto: Fraunhofer IPA*

Verfahren umgesetzt werden. Für Punktwolken existiert etwa „PointNet“ und darauf aufbauende Varianten wie „PointNet++“ oder „PointPillars“ [9–11]. Diese leiten auf unterschiedlichen Skalen Features aus Gruppierungen von räumlich benachbarten Punkten ab und prädizieren für jeden Punkt die zugehörige Klasse. Die intrinsische Unsicherheit der Detektion in industriellen Szenen konnte durch eine Reformulierung mit statistischen Bausteinen quantifiziert und eine Aussage über die Zuverlässigkeit der Erkennungen gewonnen werden [12]. Few-Shot-Learning macht durch Eintrainieren universeller Features die Objektdetektion effizienter und reduziert die Abhängigkeit von großen Trainingsdatensätzen [13].

Diese neuen Möglichkeiten werden von Startups (zum Beispiel Riiico GmbH, aurivus GmbH) von der Forschung in die Industrie getragen. Häufig sind die Verfahren aber für den Einsatz im breiten Markt noch nicht robust genug und für spezialisierte Anwendungsfälle fehlen in der Regel passende Trainingsdaten. Nicht zuletzt deshalb beschränkt sich das marktverfügbare Angebot in der Detektion nach wie vor auf wenige, einfach zu detektierende Elemente und bietet kaum Möglichkeiten, Anpassungen für individuelle Objektklassen vorzunehmen.

Um Anwender bei der Erstellung der digitalen Modelle aus Lidar-Daten effektiv zu unterstützen, ist ein Prozess nötig, der eine echte 3D-Erkennung auch für neue Klassen liefert und dem Anwender die Möglichkeit zur Korrektur einräumt. Im Folgenden wird ein solches flexibles KI-basiertes Verfahren vorgestellt, das nur von einer geringen Menge an Trainingsdaten abhängt.

3 Flexible Objekterkennung in multimodalen Daten

Das gewählte Vorgehen setzt auf multimodale Daten aus Punktwolken und 2D-Bildern und gliedert sich in die vier Schritte in **Bild 1**, die durch zwei optionale Kontrollschritte des Anwenders ergänzt werden.

Durch die Auftrennung der Erkennung in eine 2D-Detektion und eine anschließende Lokalisierung der Objekte im 3D-Raum ergeben sich gegenüber Ende-zu-Ende-Verfahren mehrere große Vorteile. Wenn neue Objektklassen eintrainiert werden sollen, kann der dafür nötige Datensatz in 2D annotiert werden, wodurch die deutlich aufwendigere Annotation von Objekten in 3D entfällt. Des Weiteren bietet die Trennung in zwei Teilschritte flexible Korrekturmöglichkeiten: Sowohl nach Schritt 2, als auch in Schritt 4 kann der Anwender die Zwischenergebnisse überprüfen und bei Bedarf anpassen.

3.1 Datenaufnahme und Entzerrung

Da 360°-Panoramabilder mit ihren starken Verzerrungen von üblichen 2D-Detektoren nicht direkt verarbeitet werden können, müssen diese zunächst, wie in **Bild 2** dargestellt, von der Kugelprojektion in eine Rektangularprojektion transformiert werden [14]. Als Ergebnis dieses Schritts steht eine große Zahl sich teilweise überlappender, entzerrter Kacheln zur Verfügung.

3.2 2D-Objekterkennung

Auf jede der erzeugten Kacheln wird anschließend ein 2D-Objektdetektor angewendet, der die Position und Klasse der vorhandenen Objekte prädiziert. Im aktuellen Entwicklungsstadium sind zwei verschiedene 2D-Detektoren implementiert: „Yolov5“ [15, 16] und ein Few-Shot-Ansatz. Yolov5 ist ein etablierter Ansatz für schnelle 2D-Erkennungen, der auf öffentlichen Datensätzen vortrainiert wurde. Experimentell wird zusätzlich mit einem Few-Shot-Ansatz gearbeitet, der weniger abhängig von vorhandenen Datensätzen ist und anhand weniger Beispiele die Erkennung neuer Klassen lernt. Durch den flexiblen Aufbau des Verfahrens sind beim Wechsel des Detektors keine Anpassungen nötig – auch andere Detektoren können als Drop-In-Replacement genutzt werden.

Beim Einsatz des Yolov5-Detektors muss für eine genaue Erkennung das Modell auf einer großen Menge annotierter Daten trainiert werden. Für einige Objektklassen existieren dafür frei verfügbare Datensätze. Sind spezielle Objekte von Interesse, wie sie im Industrieumfeld oft auftreten, muss für diese ein Datensatz mit einer hinreichend großen Anzahl an annotierten Objekten erstellt werden. Zur Reduktion des Aufwands für die Hinzunahme neuer Objektklassen ist ein Few-Shot-Verfahren von Vorteil.

https://doi.org/10.37544/1436-4980-2023-03-6 https://www.inlibra.com/de/page/agb - Open Access - CC BY

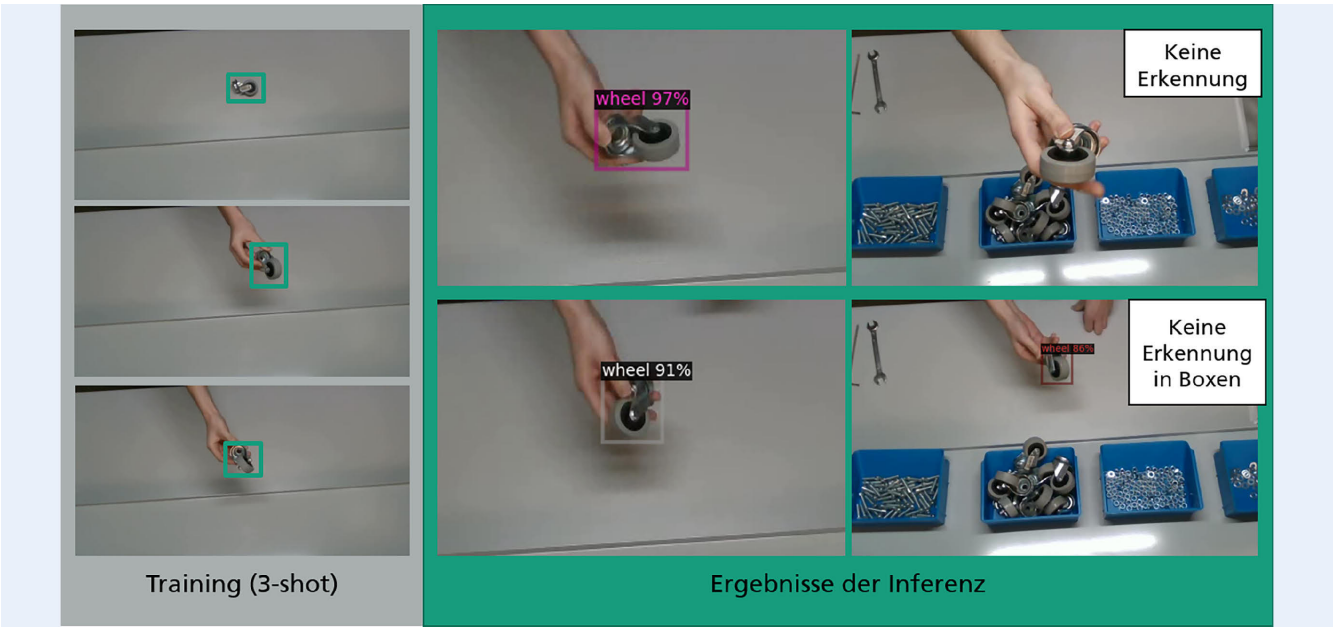

Bild 3. Training und Ergebnisse des „DeFRCN"-Detektors. *Foto: Fraunhofer IPA*

Few-Shot-Detektoren wurden zu dem Zweck entwickelt, aus einer sehr geringen Menge an annotierten Daten die Erkennung neuer Objekte lernen zu können. Ein Ansatz, der laut Übersichtsstudie von *Köhler et al.* [17] eines der besten Ergebnisse liefert, ist das Decoupled-Faster-RCNN–Verfahren (DeFRCN) von *Qiao et al.* [18]. Erste Experimente zur Objekterkennung mit dem DeFRCN-Ansatz ergaben, dass diese Methode im Fall einer geringen Variation des Objektes vielversprechend ist. In **Bild 3** ist das Ergebnis für das Training mit nur drei Beispielen von Rollen exemplarisch dargestellt.

In den beiden linken Ergebnisbildern wichen die Position, Orientierung und Umgebung nur leicht von den drei Trainingsbildern ab und es konnten die Rollen mit hoher Zuverlässigkeit erkannt werden. In den beiden rechten Ergebnisbildern sind Fehlerfälle zu sehen, die durch eine zu stark vom Trainingsdatensatz abweichende Umgebung und Positionierung verursacht wurden.

Je nach verwendetem Objektdetektor können einem Anwender unterschiedliche Korrekturmöglichkeiten angeboten werden. Liegen mehrere Detektionen vor, von denen diejenigen mit einer ausreichend hohen Konfidenz im weiteren Verfahren berücksichtigt werden, kann der Nutzer proaktiv eine der Detektionen mit geringer Konfidenz zusätzlich selektieren oder Detektionen hoher Konfidenz abwählen. Alternativ können durch händisches Einzeichnen von 2D-Bounding-Boxes zusätzliche Detektionen hinzugefügt werden.

3.3 3D-Clustering

Die Positionskoordinaten der 2D-Detektionen werden für die 3D-Lokalisierung in die Punktwolke projiziert. Anhand der Kameraposition kann ein Teilbereich der Punktwolke extrahiert werden, der in unmittelbarer Umgebung des in 2D gefundenen Objekts liegt. Durch diese Maßnahme lässt sich der Suchbereich für die nachfolgenden Verfahrensschritte erheblich verkleinern.

Mit „Efficient Ransac" [19], einem intelligenten Clustering-Verfahren, wird die Teilszene anschließend, wie in **Bild 4** dargestellt, in geometrische Grundkörper zerlegt.

Die Prämisse für jeden nachfolgenden Gruppierungsschritt ist, dass das gesuchte Objekt aus einer Kombination dieser geometrischen Primitive gewonnen werden kann. Um eine passende Gruppierung zu finden, werden die Primitive hierarchisch geclustert und jeder Zwischenschritt als potenzieller Objektkandidat gespeichert. Das Verfahren terminiert, sobald die letzten Objekte zur gesamten Teilszene gruppiert werden. Das Ergebnis ist eine große Zahl an 3D-Bounding-Boxes, die jeweils eine mögliche Position eines Objektes darstellen. Die Information, um welches Objekt es sich handelt, ist hierbei noch nicht enthalten.

3.4 Zusammenführung 2D- und 3D-Erkennung

Die Auswahl des passenden Objektkandidaten und die Zuordnung zu einer Klasse gelingt durch die erneute Betrachtung der ursprünglichen 2D-Detektion und einer Datenbank mit Vorwissen zum gesuchten Objekt. Durch die Rückprojektion der Objektkandidaten von 3D in 2D kann ein Abgleich der erwarteten Detektion mit der tatsächlich stattgefundenen erfolgen. Zusätzlich erlaubt die Berücksichtigung des Vorwissens eine weitere Kontrolle, sodass Objektkandidaten mit nicht plausibler Größe oder Textur verworfen werden können. Im Ergebnis wird derjenige Objektkandidat als erkanntes Objekt ausgewählt, der die höchste Übereinstimmung mit der 2D-Detektion und dem Vorwissen aufweist. Wie auch schon in der 2D-Erkennung kann der Anwender bei Bedarf korrigierend eingreifen und einen beliebigen anderen Objektkandidaten auswählen. Auch ein direktes Eingeben der 3D-Bounding-Box ist möglich, was jedoch mit zusätzlichem Aufwand verbunden ist.

3.5 Ergebnisse

Das Verfahren wurde an drei Szenen in den Laboren des Fraunhofer IPA in Stuttgart getestet. Um statistisch aussagekräftige Ergebnisse zu erhalten, wurde eine Einschränkung auf die Objektklassen „Bildschirm" und „Stuhl" vorgenommen, da diese in den Szenen zahlreich vorhanden sind. Außerdem konnte so ein

https://doi.org/10.37544/1436-4980-2023-03-6 https://www.inlibra.com/de/page/agb - Open Access -

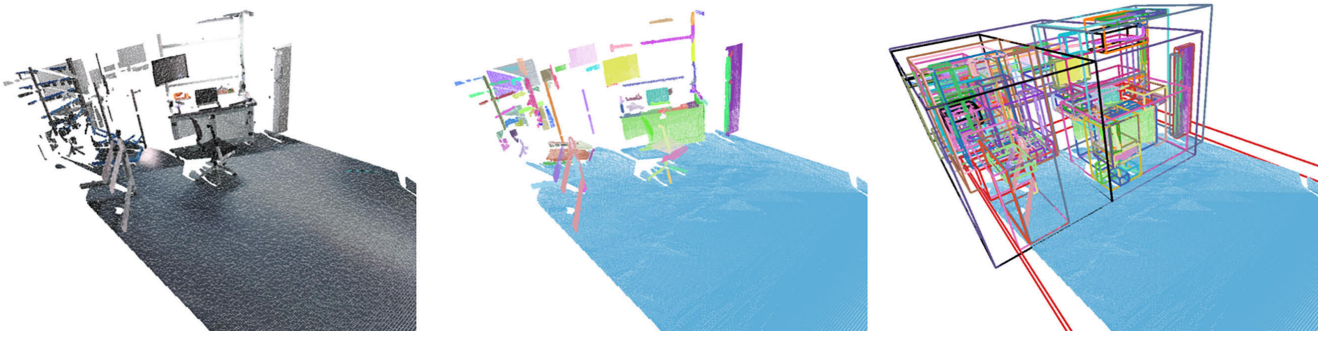

Bild 4. Clustering einer Szene zu geometrischen Primitiven und anschließende Gruppierung zu Objektkandidaten. *Foto: Fraunhofer IPA*

Tabelle. Auswertung zur Detektionsgenauigkeit von Stühlen und Bildschirmen in den Räumlichkeiten des Fraunhofer.

SCAN-ID	CLASS	TOTAL	TP	FP	FN	ACC	PREC
150645	Chairs	7	7	0	0	1,00	1,00
150645	Screens	20	11	2	9	0,55	0,85
150716	Chairs	15	13	1	2	0,87	0,93
150616	Chairs	8	8	0	0	1,00	1,00
150616	Screens	14	10	1	4	0,71	0,91
ALL	ALL	64	49	4	15	0,77	0,92

Bild 5. 3D-Annotationen nach Abschluss des Verfahrens. Bildschirme sind mit grüner Farbe markiert, Stühle mit roter Farbe. *Foto: Fraunhofer IPA*

Anwendungsfall zur Erfassung von Arbeitsplätzen im digitalen Zwilling simuliert werden. Die Ergebnisse der Auswertung sind in der **Tabelle** aufgeführt und zeigen für die beiden genannten Klassen die Gesamtzahl der Objekte in der jeweiligen Szene (TOTAL), die korrekt erkannten Instanzen (True Positive TP), Fehldetektionen (False Positive FP) und vorhandene Objekte, die nicht erkannt wurden (False Negative FN). Aus diesen Werten lässt sich die Detektionsgenauigkeit (ACC=TP/TOTAL) und die Precision (PREC=TP/(TP+FP)) ableiten.

Für die Praxis deutet die Detektionsgenauigkeit (ACC) darauf hin, dass etwa Dreiviertel der getesteten Objekte automatisch erkannt werden können und somit nur für die verbleibenden knapp 25 % ein Eingriff durch den Anwender stattfinden muss – mit proportionaler Zeitersparnis gegenüber der vollständigen Modellierung von Hand. Die hohe Precision lässt den Schluss zu, dass es sich bei den identifizierten Objekten in den meisten Fällen auch tatsächlich um diese Objekte gehandelt hat, es also wenige Fehldetektionen gab. Der häufigste Fehlerfall ist das Nicht-Erkennen von tatsächlich vorhandenen Objekten, die folglich händisch nachgetragen werden müssen. Exemplarisch sind für Scan-ID 150645 in **Bild 5** die 3D-Bounding-Boxes dargestellt.

Die Erkennung von Bildschirmen funktionierte dabei etwas schlechter als die von Bürostühlen. Dies kann durch die kleineren Abmessungen der Bildschirme und somit geringeren Abtastraten erklärt werden. Auch existieren an den in den Laboren stehenden Aufbauten Scheiben und Fenster, die in ihrem Erscheinungsbild an Bildschirme erinnern. Nicht zuletzt hat die Lidar-Sensorik Probleme stark spiegelnde Objekte zu erfassen. So wird ein Bildschirm in 2D häufig erkannt, kann in 3D aber einem Objektkandidaten nicht korrekt zugeordnet werden. Positiv fällt auf, dass auch teilverdeckte Objekte erkannt werden konnten.

4 Anwendungsszenarien

Der vielversprechendste Einsatz für das vorgestellte Verfahren liegt im Bereich der initialen Erzeugung und Pflege digitaler geometrischer Modelle, wie sie insbesondere als Baustein für digitale Zwillinge benötigt werden. Sollen in einer Produktionsanlage beispielsweise Restrukturierungen vorgenommen werden, ist zur vorbereitenden Planung eine akkurate Simulation unerlässlich. Die Basis dafür sind digitale Modelle, die neben der räumlichen Anordnung auch die Verflechtungen und Interaktionen verschiedenster Elemente untereinander, wie etwa Warenflüsse oder Tragfähigkeiten, abbilden. Auch bei der Aktualisierung bereits existierender Modelle ist oft ein erneutes Vermessen des Bestands nötig. Mit einer intelligenten Objekterkennung kann ein Großteil dieser Arbeit automatisiert werden. Bis die Algorithmen zuverlässig genug sind (vergleiche ungelöste Probleme beim autonomen Fahren), kann der Anwender nicht ersetzt werden. Eine optimale Unterstützung durch intelligente Vorschläge wirkt aber entlastend und führt zu Zeit- und Kostenersparnissen.

https://doi.org/10.37544/1436-4980-2023-03-6 https://www.inlibra.com/de/page/agb - Open Access - CC BY

Ein zweiter, großer Anwendungsfall ist das Bauwesen. Digitale Zwillinge für den gesamten Lebenszyklus von Bauwerken, sogenanntes Building Information Modelling (BIM), wird heute vom Entwurf über die Planung, den tatsächlichen Bau, den späteren Betrieb und am Ende des Lebenszyklus für den Rückbau eingesetzt. Für Bestandsgebäude, für die bislang keine oder keine aktuelle digitale Dokumentation vorliegt, kann auch nachträglich eine digitale Bestandsdokumentation erzeugt werden, welche dieselben Vorteile für die Bewirtschaftung bestehender Gebäude erlaubt. Diese As-Built-Dokumentation des tatsächlichen Zustands in Form digitaler Modelle ist nach wie vor meist Handarbeit. Durch diesen Aufwand wird die Aktualisierung der digitalen Modelle nicht so häufig durchgeführt, wie es nötig wäre, und die BIM-Modelle unterscheiden sich über die Zeit zunehmend vom tatsächlichen Bauwerk, sodass kein aktuelles Modell mehr verfügbar ist.

Um die Architekten, Bauingenieure und Handwerker bei der Pflege der BIM-Modelle zu unterstützen, eine häufigere Aktualisierung zu ermöglichen und die Bestandsdokumentation zu vereinfachen, kann das oben vorgestellte Verfahren zur automatischen Identifikation relevanter Objekte und Strukturen in diesen Szenen eingesetzt werden.

5 Zusammenfassung und Ausblick

Motiviert durch den aktuellen Bedarf an digitalen Zwillingen und der zunehmenden Verbreitung von Lidar-Technologie wurde ein flexibles Verfahren zur Digitalisierung einer Umgebung entwickelt. Die Flexibilität zeichnet sich durch die Kombination der Objektdetektion in 2D mit der Lokalisierung in 3D aus. Das aufwendige Annotieren von Daten in 3D kann so vermieden werden. Mit Few-Shot-Learning wurde zudem ein Verfahren vorgestellt, das den Bedarf an Trainingsdaten deutlich reduziert und auf Basis einiger weniger Annotationen bereits vielversprechende Ergebnisse liefert. Durch den mehrstufigen Prozess wird gegenüber Ende-zu-Ende-Verfahren eine höhere Transparenz erreicht, die es dem Anwender ermöglicht, korrigierend einzugreifen.

Diese Eigenschaften machen das vorgestellte 3D-Objekterkennungsverfahren interessant für verschiedenste Anwendungsfälle in der Industrie, in denen digitale Modelle der Umgebung dringend benötigt werden und zudem mit vertretbarem Aufwand erstellt werden können.

Literatur

[1] PricewaterhouseCoopers GmbH: PwC Digital Factory Transformation Survey 2022. Stand: 2022. Internet: www.pwc.de/de/content/0f96ea9c-992c-4ba7–8c4d-b4637cf81d9f/pwc-digital-factory-transformation-survey-2022.pdf. Zugriff am 01.03.2023

[2] Chen, Y.; Tang, J.; Jiang, C. et al.: The accuracy comparison of three simultaneous localization and mapping (SLAM)-based indoor mapping technologies. Sensors 18 (2018) 10, pp. 3228

[3] Borrmann, A.; König, M.; Koch, C. et al.: Building Information Modeling – Technology Foundations and Industry Practice. Cham: Springer International Publishing 2018

[4] Undet: How to improve modeling speed using point clouds in Revit by 2x times?. Stand: 2022. Internet: www.undet.com/how-to-improve-modeling-speed-using-point-clouds-in-revit-by-2x-times/. Zugriff am 01.03.2023

[5] BI-Medien: Reales BIM-konform erfassen – Marktüberblick 3D-Laserscanner. Stand 2022. Internet: bi-medien.de/fachzeitschriften/baumagazin/hochbau/marktueberblick-3-d-laserscanner-reales-bim-konform-erfassen-b12232. Zugriff am 01.03.2023

[6] Fayez, T.-K.; Landes, T.; Gussenmeyer, P.: Hough-transform and extended ransac algorithms for automatic detection of 3d building roof planes from lidar data. ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, Espoo/Finnland, 2007, pp. 407–412

[7] Son, H.; Kim, C.; Turkan, Y.: Scan-to-BIM-An Overview of the Current State of the Art and a Look Ahead. Proceedings of the International Symposium on Automation and Robotics in Construction 32 (2015), #1

[8] Zeiser, R.; Ullmann, F.; Neuhäuser, T. et al.: System concept for semi-automated generation of layouts for simulation models based on point clouds. Simulation in Produktion und Logistik 2021 (2021), pp. 443

[9] Qi, C. R.; Su, H.; Mo, K et al.: Pointnet: Deep learning on point sets for 3d classification and segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition, CVPR 2017, New York/USA, 2017, pp. 652–660

[10] Qi, C. R.; Yi, L.; Su, H et al.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Proceedings of the 31st International Conference on Neural Information Processing Systems NIPS'17, 2017, pp. 5105–5114

[11] Lang, A. H.; Vora, S.; Caesar, H. et al.: Pointpillars: Fast encoders for object detection from point clouds. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 12697–12705

[12] Petschnigg, C.; Pilz, J.: Uncertainty aware deep point based neural network for 3D object classification. Data Science–Analytics and Applications: Proceedings of the 3rd International Data Science Conference, iDSC2020, 2021, pp. 66–73

[13] Zhao, N.; Chua, T. S.; Lee, G. H.: Few-shot 3d point cloud semantic segmentation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 8873–8882

[14] Viu, J. G.: Object recognition on panoramic images. Bachelorarbeit, Universität Zaragoza, 2019

[15] Ultralytics: yolov5. Stand 2023. Internet: github.com/ultralytics/yolov5. Zugriff am 01.03.2023

[16] Redmon, J.; Divvala, S.; Girshick, R. et al.: You only look once: Unified, real-time object detection. Proceedings of the IEEE conference on computer vision and pattern recognition 2016, pp. 779–788

[17] Köhler, M.; Eisenbach, M.; Gross, H. M.: Few-Shot Object Detection: A Survey. ACM Computing SurveysVolume 54 (2022) 11s, # 242, pp 1–37, doi.org/10.1145/3519022

[18] Qiao, L.; Zhao, Y.; Li, Z. et al.: Defrcn: Decoupled faster r-cnn for few-shot object detection. Proceedings of the IEEE/CVF International Conference on Computer Vision 2021, pp. 8681–8690

[19] Schnabel, R.; Wahl, R.; Klein, R.: Efficient RANSAC for point-cloud shape detection. Computer graphics forum, 2007, pp. 214–226, doi.org/10.1111/j.1467–8659.2007.01016.x

Dipl.-Ing. **Matthias Willenbrink**
Foto: Autor

Dascha Karelina, M. Sc.

Dr.-Ing. **Ira Effenberger**

Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA
Nobelstr. 12, 70569 Stuttgart
Tel. +49 711 / 970-1225
matthias.willenbrink@ipa.fraunhofer.de
www.ipa.fraunhofer.de

LIZENZ

Dieser Fachaufsatz steht unter der Lizenz Creative Commons Namensnennung 4.0 International (CC BY 4.0)

https://doi.org/10.37544/1436-4980-2023-03-6 https://www.inlibra.com/de/page/agb - Open Access -