

Untersuchung von Neural Radiance Fields zur Modellrekonstruktion in der Fabrikplanung

Neural Radiance Fields in der Fabrikplanung

P. Ruediger-Flore, M. Klar, M. Hussong, J. Mertes, L. Yi, M. Glatt, P. Kölsch, J. C. Aurich

Neural Radiance Fields (NeRF) bieten eine kostengünstige und effiziente Alternative im Vergleich zu herkömmlichen Verfahren, um 3D-Modelle aus realen Objekten zu generieren. Dementsprechend bieten NeRF große Potenziale zur Nutzung in diversen Anwendungsfällen, wie der Fabrikplanung. In diesem Beitrag wird die NeRF-Technologie an einem Beispiel aus der Fabrikplanung angewendet und daran aktuelle Herausforderungen sowie Möglichkeiten zur Nutzung der Technologie diskutiert.

STICHWÖRTER

Bildverarbeitung, Fabrikplanung, Digitalisierung

On the applicability of Neural Radiance Fields for virtual model reconstruction in factory planning

Compared to conventional methods, Neural Radiance Fields (NeRF) offer a cost-effective and efficient alternative for generating 3D models from real objects. Therefore, NeRF has great potential for use in diverse applications, such as factory planning. Accordingly, this paper applies the NeRF technology to an example use case from factory planning and discusses current challenges and possibilities for using the technology.

1 Motivation

Neural Radiance Fields (NeRF) sind eine vielversprechende Technologie zur Generierung realistischer 3D-Modelle aus 2D-Bildern [1]. Auch andere Formen von Eingangsdaten wie etwa Videos [2] oder Punktwolken [3] sind für NeRF geeignet. Somit ermöglichen NeRF die Erstellung von 3D-Modellen mit hoher räumlicher Auflösung bei gleichzeitig hoher Flexibilität in Bezug auf die Eingangsdaten. Im Vergleich zu herkömmlichen 3D-Modellierungstechniken, wie zum Beispiel der Verwendung von LiDAR-Sensoren (Light Detection and Ranging) oder 3D-Laserscannern, bieten NeRF eine potenziell kostengünstige, schnelle und effiziente Alternative.

NeRF haben das Potenzial in einer Vielzahl von Anwendungen eingesetzt zu werden, wie der Gesichtserkennung, der Bildverarbeitung, der Rekonstruktion von Denkmälern oder Gebäuden, der Robotik oder auch der Fabrikplanung [4]. In der Fabrikplanung bieten NeRF die Möglichkeit, realistische 3D-Modelle von Fabriken zu generieren. Dadurch können Fabrikplaner virtuelle Umgebungen erstellen und testen, bevor physische Anlagen realisiert werden. Die Folge sind Kosteneinsparungen und die Vermeidung von Fehlern, die sonst zu teuren Nacharbeiten und Verzögerungen führen können [5]. Durch die einfache Erstellung von 3D-Modellen lassen sich verschiedene Planungsvarianten und -konfigurationen im Vergleich zu den genannten Methoden schneller analysieren, da lediglich Bilder der geplanten Fabrikflächen als Eingabedaten benötigt werden [6]. Ein weiterer möglicher Nutzen des Einsatzes von NeRF in der Fabrikplanung liegt in der Erstellung realistischer Simulationen von Produktionsprozessen. NeRF können dazu beitragen, Ineffizienzen in der aktuel-

len Ausgestaltung von Produktionssystemen, etwa häufige Staus im Materialfluss, visualisierbar und analysierbar zu machen. Dies führt dazu, dass gezielte Umplanungsprozesse in der Produktion durchgeführt werden können, die zu einer Steigerung der Effizienz und Qualität der Produktion führen [7].

Vor dem Hintergrund dieser Potenziale wird in diesem Beitrag die NeRF-Technologie für den Anwendungsfall der Fabrikplanung einer näheren Betrachtung unterzogen. Das nächste Kapitel beschreibt das Modell der NeRF und die Vorgehensweise bei der Implementierung. Anschließend wird dieses Verfahren auf einen Anwendungsfall aus dem Bereich der Fabrikplanung angewendet. Die Ergebnisse und die Diskussion der Herausforderungen und Möglichkeiten für den Einsatz von NeRF sind Gegenstand von Kapitel 3. Kapitel 4 fasst die Ergebnisse kurz zusammen und gibt einen Ausblick.

2 Mathematische Modellierung und Vorgehen zur Anwendung von NeRF

Die NeRF repräsentieren eine Szene in einem vollständig verbundenen künstlichen neuronalen Netzwerk (englisch: artificial neural network, ANN) F_θ . Das Training des ANN kann etwa aus vorhandenen Bildern einer statischen Szene mit bekannten Kamerapositionen erfolgen. Dabei transformiert das ANN eine gegebene Position (x, y, z) und Blickrichtung (θ, ϕ) in der trainierten Szene in eine Dichte σ und eine Lichtfarbe c . Wenn Lichtstrahlen in der Szene von der Kameraposition aus anhand der Vektorbeschreibung $r(t)$ verfolgt werden, können die Volumendichte- und Farbinformationen entlang der Strahlen mit NeRF approximiert und mit Methoden des Volume-Rendering zu einem

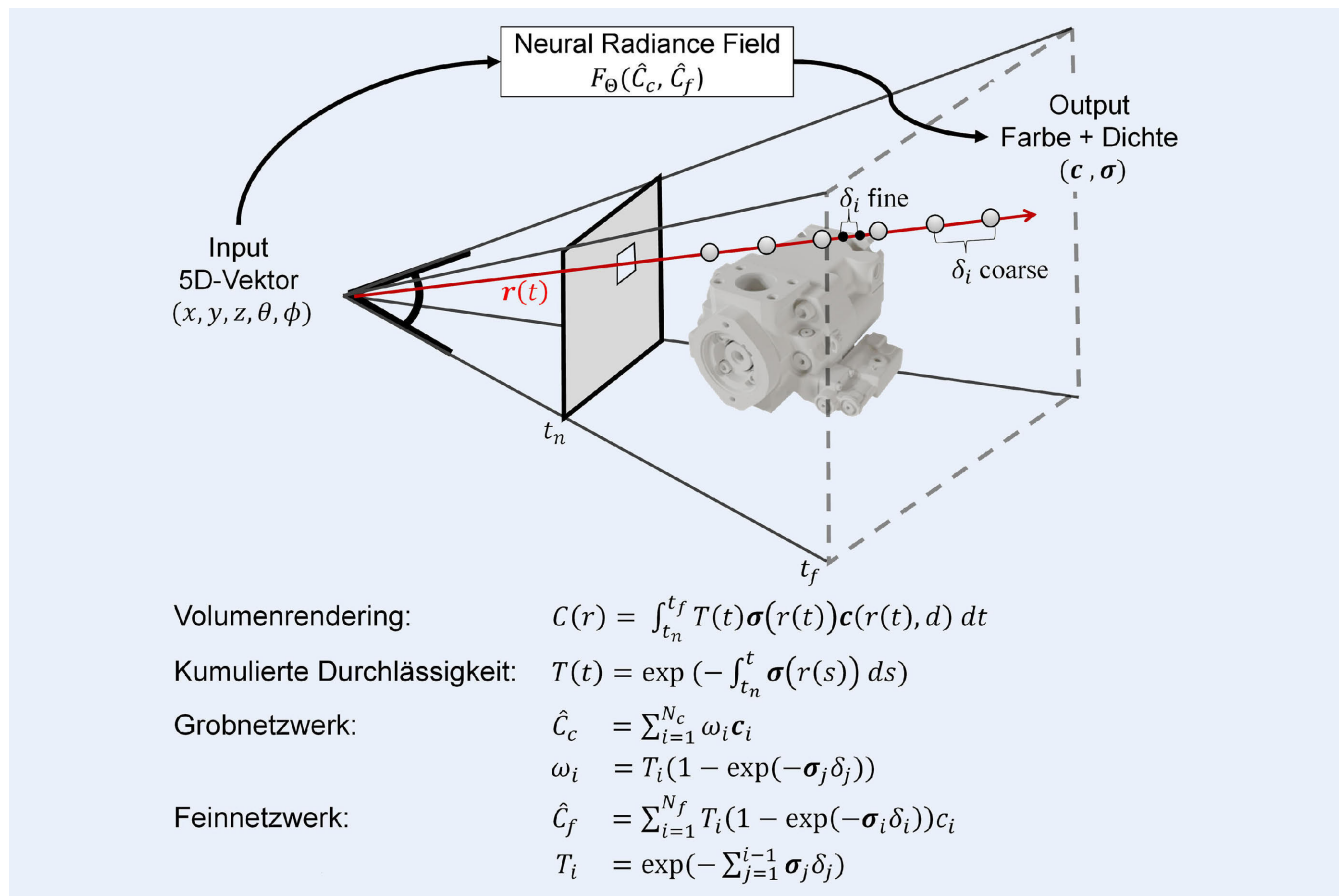


Bild 1. Allgemeines Konzept von NeRF (Neural Radiance Fields). Grafik: RPTU Kaiserslautern

2D-Bild kombiniert werden. Die Dichte- und Farbinformationen werden getrennt vom Netz bestimmt. Die Dichte an einem Punkt wird nur als Funktion der drei kartesischen Positionskoordinaten über acht ANN-Schichten mit je 256 Neuronen berechnet. Das Ergebnis dieser Faltung ist die Dichte und ein weiterer Merkmalsvektor. Dieser Vektor ist mit der Blickrichtung verknüpft und dient als Input für die Berechnung des Farbwertes, welche über eine weitere ANN-Schicht erfolgt.

Durch eine blickrichtungsabhängige Berechnung der Farbe in einem Punkt können Reflexionen durch die NeRF abgebildet werden. Da die Blickrichtungen keinen Einfluss auf die Dichte eines Objekts haben, sind sie für die Berechnung nicht nötig. Neue Blickrichtungen werden durch Strahlen synthetisiert, die von der gewünschten Position in die Szene gesendet werden. Um feine Details mit NeRF darstellen zu können, wird jede einzelne Einflussgröße (drei Ortskoordinaten (x, y, z) sowie die drei Komponenten des Richtungsvektors d) durch Sinus- und Kosinustransformationen in einen höher-dimensionalen Raum transformiert. Der Informationsgehalt der eindimensionalen Einflussgrößen wird so deutlich verbessert und gewährleistet das Lernen auch feiner Details innerhalb der Szene. Die Farb- und Dichteinformationen der Strahlen werden von NeRF ermittelt, sodass die resultierende Farbe jedes Strahls bestimmt werden kann. Diese Informationen werden durch zwei Netze, ein „grobes“ und ein „feines“, begrenzt und innerhalb eines Intervalls $[t_n, t_f]$ ausgewertet, siehe Bild 1. Eine genaue mathematische Herleitung der Funktionsweise dieser Netze ist in [8] zu finden.

Die allgemeine Vorgehensweise bei der Verwendung von NeRF zur Erstellung eines virtuellen 3D-Modells in Form einer vernetzten STL (Stereolithografie)-Datei besteht aus sechs Schritten. Zuerst werden Bilder der Szene aufgenommen, sodass das Zielobjekt von allen Seiten vollständig erfasst wird. Für einen benutzerfreundlicheren Ansatz kann ein Video des Objekts verwendet werden. Bei einer flüssigen Videoaufnahme können mehrere Bilder mit zeitdiskreter Abtastung extrahiert werden. Dies ermöglicht auch eine aufwandsarme Generierung von Trainings- und Testdaten, da nur die Abtastgröße verändert werden muss.

Im zweiten Schritt werden die Parameter für die Kamera kalibrierung und den 5D-Eingangsvektor bestimmt. Dies wird durch den Einsatz von Structure-from-Motion-Methoden [9] zur Strukturierung aus der Bewegung erreicht. Innerhalb dieser Pipeline werden Schritte bis zur Bildregistrierung betrachtet. Hier werden neue Bilder zum aktuellen Modell registriert, indem das Perspective-n-Point (PnP)-Problem [10] unter Verwendung von Merkmalskorrespondenzen zu triangulierten Punkten in bereits registrierten Bildern gelöst wird. Das Ergebnis ist die Ausgabe der Kamerapositionen und Blickwinkel in Form von Tensoren.

Von allen Verfahrensschritten ist dies der rechenintensivste und zeitaufwendigste, da er mit zunehmender Bildanzahl und Eingabeauflösung schlecht skalierbar ist. Mit geeigneter Handhabungstechnik können Kamerapositionen und Blickwinkel direkt gesteuert und erfasst werden, sodass dieser Schritt übersprungen werden kann. Dabei ist zu beachten, dass bereits eine gute Näherung ausreicht und somit die Handhabung mit einfachen Mitteln realisiert werden kann.

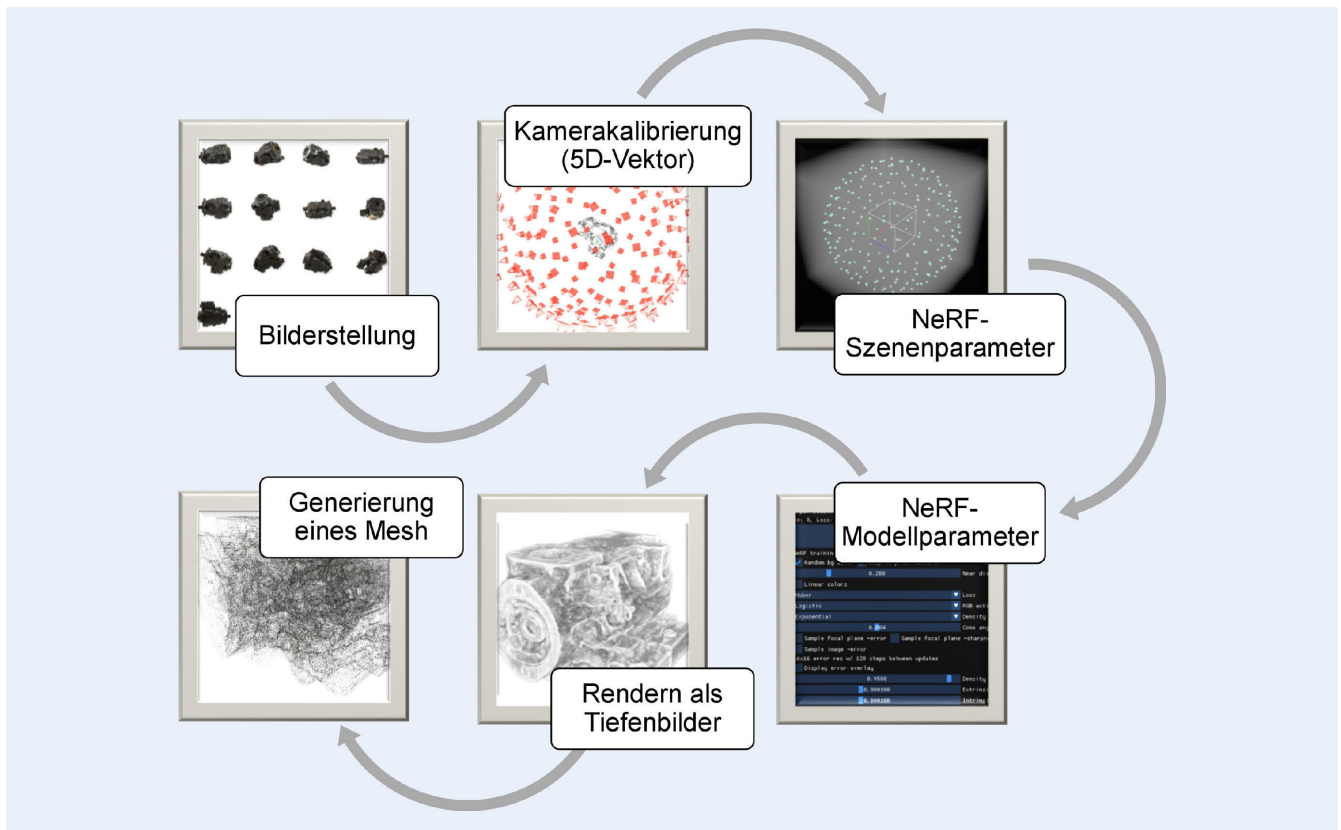


Bild 2. Allgemeines Vorgehen zur Anwendung von NeRF. Grafik: RPTU Kaiserslautern

Im dritten Schritt wird die Parametrisierung der NeRF bestimmt. Dazu muss sichergestellt werden, dass sich das Zielobjekt innerhalb der Grenzen der 3D-Szene befindet. Außerdem muss die Strahllänge so eingestellt werden, dass sie durch alle relevanten Objekte der Szene verläuft.

Im vierten Schritt wird das eigentliche Modell trainiert, wobei verschiedene modellspezifische Parameter wie zum Beispiel die Verlustfunktion oder die Lernrate eingestellt werden können. Eine vollständige Liste der Parameter findet sich in [8]. Im untersuchten Anwendungsszenario der Fabrikplanung werden die Standardparameter der NeRF verwendet.

Im fünften Schritt kann das Modell entweder zur Erzeugung beliebiger neuer Sichten verwendet werden, eine Punktwolke durch strukturelles Sampling neuer Sichten im Modell erzeugt werden oder nur die Volumendichte ausgewertet werden.

Im sechsten und letzten Schritt können Nachbearbeitungswerkzeuge für Punktwolken, wie etwa der Marching-Cubes-Algorithmus, zur Vernetzung der Punktwolken und zur Reparatur der resultierenden 3D-Modelle entsprechend der weiteren Anwendungsanforderungen eingesetzt werden.

Die allgemeinen Schritte zur Anwendung von NeRF sind in **Bild 2** nochmals zusammengefasst.

3 Ergebnisse und Diskussion

Der Fabrikplanungsprozess profitiert von 3D-Modellen, die den aktuellen Zustand der Fabrik abbilden, wie in Kapitel 1 beschrieben. Im Anwendungsszenario werden daher die Möglichkeiten von NeRF zur digitalen und flexiblen Layoutabbildung für Fabriken untersucht. Dazu wird ein bestehendes Produktionsge-

bäude als exemplarisches Anwendungsszenario verwendet. Dieses ist in **Bild 3** dargestellt.

In **Bild 3** sind im Szenario verschiedene Werkzeugmaschinen, Tische, Lagerregale, ein Brückenkran, Netzwerkinfrastruktur sowie Medienver- und -entsorgungsinfrastruktur abgebildet. Zu den unterschiedlichen Objektmerkmalen gehören feine und grobe Strukturen, matte und reflektierende Oberflächen sowie helle und schlecht beleuchtete Bereiche. Somit ist eine Untersuchung der Fähigkeiten von NeRF in einer für die Industrie relevanten Umgebung möglich.

Die für das Training verwendeten Bilder werden mit einer Smartphone-Kamera aufgenommen und mit der Structure-from-Motion-Methode von [9] vorverarbeitet, um die intrinsischen und extrinsischen Kameraparameter zu approximieren. Zudem wird sichergestellt, dass die Aufnahmeorte gleichmäßig über die gesamte Anlage verteilt sind, indem das aufgenommene Video in Intervallen von 3 bis 5 Bildern pro Sekunde gleichmäßig abgetastet wird. Die **Tabelle** zeigt die Parameter des Modells, die für das Anwendungsszenario definiert wurden.

Die Ergebnisse zeigen, dass die Qualität des Modells stark vom Versuchsaufbau und den Parametern des Algorithmus abhängt. Vor allem die Kartierung von Fabriken ist eine Herausforderung, da die Fabrik nicht schnell von allen Seiten umrundet werden kann. Stattdessen werden Bilder aus dem Inneren des modellierten Objekts aufgenommen, was den Trainingsprozess von NeRF vor neue Schwierigkeiten stellt, da der primäre Fokus des Algorithmus auf der Konstruktion eines einzelnen zentrierten Objekts liegt. Außerdem beeinflussen die in der Tabelle aufgeführten Parameter des Algorithmus die Qualität der Ergebnisse.

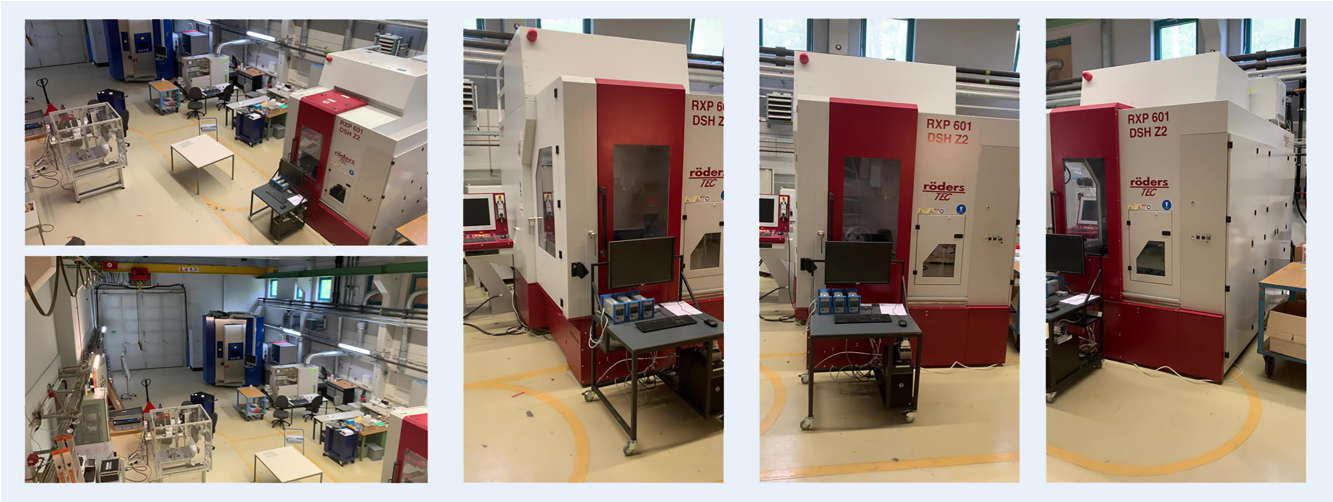


Bild 3. Exemplarische Bildaufnahmen einer Fabrikhalle für das Training des NeRF. Foto: RPTU Kaiserslautern

Tabelle. Experimentelles Design für die Anwendung von NeRF für die Fabriklayoutplanung.

Experiment	1	2	3	4
Szene	Halle Vogelperspektive			Einzelne Maschine
Einheitswürfel	klein	groß	klein	klein
Bildauflösung	1920 x 1080	1920 x 1080	3840 x 2160	1920 x 1080



Bild 4. Ergebnisse zur Rekonstruktion einer Fabrikhalle mithilfe von NeRF. Foto: RPTU Kaiserslautern

Die Trainingsbilder der ersten drei Modelle in **Bild 4** wurden auf einer Plattform oben in der Fabrikhalle aufgenommen. Der einzige Unterschied zwischen erstem und zweitem Modell ist die Größe des Einheitswürfels, ein Hyperparameter, der das Zentrum der zu rekonstruierenden Szene beschreibt und vor dem Training definiert werden muss.

Es ist zu beobachten, dass eine Vergrößerung des Einheitswürfels zu einer Verschlechterung der Modellgüte führt, da die Komplexität des Fokuspunktes für die Rekonstruktion zunimmt. Der Einfluss ist in Bild 4 zu sehen, in dem die Unterschiede in der Modellqualität visualisiert sind. Zudem werden gut beleuchtete Kanten im Modell besser dargestellt als ebene Bereiche oder Bereiche mit geringer Lichtintensität. Das dritte Modell verwendet die gleichen Perspektiven und Trainingsparameter wie das erste Modell. Die Bildauflösung wurde jedoch auf 3840 x 2160 erhöht. Das Rekonstruktionsziel des Modells ist identisch mit der Eingangsauflösung, daher führt eine höhere Auflösung ab einem bestimmten Punkt zu einer deutlichen Qualitätsverschlechterung, wenn die Modellparameter den Detailgrad nicht mehr erfassen können. Die Fräsmaschine steht im Mittelpunkt des letzten

Modells. Wie zu sehen ist, werden die Konturen auf der linken Seite der Fräsmaschine mit hoher Qualität rekonstruiert. Auf der rechten Seite der Maschine verbleiben jedoch einige Artefakte.

Für den erfolgreichen Einsatz von NeRF zur Rekonstruktion von Fabriklayouts ist eine mehrstufige Sequenz mit zunehmendem Detaillierungsgrad vorteilhaft. Im ersten Schritt wird der allgemeine Bereich des Layouts visualisiert. Modelle einzelner Entitäten (zum Beispiel ein Fräsmaschinenmodell) können dann verwendet werden, um Teile des Gesamtmodells zu ersetzen, die eine geringere Qualität aufweisen. Die Frage, wie einzelne Modellteile ersetzt werden können, bleibt jedoch offen.

Zusammenfassend lässt sich sagen, dass die Auswahl der geeigneten Kombinationen von Modell- und Szenenparametern mit zunehmender Komplexität des zu betrachtenden Objekts nicht mehr gelöst werden kann, indem alle Kombinationen ausprobiert werden (curse of dimensionality). Zudem darf die Vorverarbeitung der Trainingsdaten für die Blickwinkelschätzung nicht vernachlässigt werden. Denn die genaue Berechnung der Kameraposition hat mit zunehmender Komplexität einen größeren Einfluss auf die Ergebnisqualität. Für größere Fabrik-

hallen erscheint es vielversprechend, Drohnen zur Bildaufnahme einzusetzen, was die Blickwinkelschätzung in Systemen mit hoher räumlicher Dynamik vor weitere Herausforderungen stellt.

4 Zusammenfassung und Ausblick

Effiziente und einfache Möglichkeiten für die Erstellung von virtuellen 3D-Modellen sind für die Planung von Fabriken von immer größerer Bedeutung. Daher wurde das Potenzial von NeRF für die Fabrikplanung untersucht und bewertet. Auf der Grundlage dieser Anwendungsszenarien können wir die Schlussfolgerung ziehen, dass NeRF ein hohes Potenzial für die Lösung von Problemen im Zusammenhang mit Rekonstruktionsszenarien in einer Fabrik hat, wofür die Fotogrammetrie bisher zu teuer oder zu ungenau in der Berechnung oder Bilderfassung war.

Um jedoch in der Fabrikplanung eingesetzt zu werden, besteht weiterer Forschungsbedarf zur Modellskalierung. In der derzeitigen Struktur ist ein Modell kaum in der Lage, eine komplette Maschinenhalle in ausreichender Qualität zu rekonstruieren. Die Ergebnisse der Rekonstruktion einzelner Maschinen zeigen aber bereits heute ein großes Potenzial für Planungsaufgaben. Entlang des gesamten Fabrikplanungsworkflows, angefangen von der Definition der Sichten und der Parametrisierung des Modells über die Segmentierung und Vernetzung der Objekte bis hin zur Dimensionierung und Zuordnung des vorhandenen Know-hows, gibt es noch viele ungelöste Herausforderungen.

Zukünftige Forschungsaufgaben sollten die Integration der NeRF-Technologie in den Fabrikplanungsprozess näher untersuchen, beispielsweise die Verwendung anwendungsspezifischer NeRF-Eingabeparameter im Kontext der Fabriklayoutplanung. Das Verständnis dieser Parameter könnte bei komplexen Szenen, vor allem bei Fabriklayouts, den Nachbearbeitungsaufwand für die Vernetzung der resultierenden Punktwolken und die Objektsegmentierung deutlich reduzieren.

FÖRDERHINWEIS

Diese Forschung wurde gefördert durch die Europäische Union (EU) im Rahmen des Europäischen Fonds für regionale Entwicklung (EFRE) und des Commercial Vehicle Cluster Südwest (P1-SZ2-3 Project: Nutzung von Künstlicher Intelligenz in der Nutzfahrzeugproduktion).

Literatur

- [1] Zhang, K.; Riegler, G.; Snively, N. et al.: NeRF++: Analyzing and Improving Neural Radiance Fields. arXiv:2010.07492v2 [cs.CV] (2020), doi.org/10.48550/arXiv.2010.07492
- [2] Gao, X.; Zhong, C.; Xiang, J. et al.: Reconstructing Personalized Semantic Facial NeRF Models from Monocular Video. ACM Transactions on Graphics 41 (2022) 6, pp. 1–12
- [3] Zimny, D.; Trzciński, T.; Spurek, P.: Points2NeRF: Generating Neural Radiance Fields from 3D point cloud. arXiv:2206.01290 [cs.CV] (2022), doi.org/10.48550/arXiv.2206.01290
- [4] Malamas, E. N.; Petrakis, E. G.; Zervakis, M. et al.: A survey on industrial vision systems, applications and tools. Image and Vision Computing 21 (2003) 2, pp. 171–188
- [5] Denkena, B.; Dittrich, M.-A.; Stobrawa, S.: Automated generation of a digital twin using scan and object detection for data acquisition. Simulation in Produktion und Logistik (2019), pp. 49–60
- [6] Gong, L.; Berglund, J.; Fast-Berglund, Å. et al.: Development of virtual reality support to factory layout planning. International Journal on Interactive Design and Manufacturing 13 (2019) 3, pp. 935–945
- [7] Sommer, M.; Seiffert, K.: Scan methods and tools for reconstruction of built environments as basis for digital twins. In: Stjepandić, J., Sommer, M., Denkena, B. (eds): DigiTwin: An Approach for Production Process Optimization in a Built Environment. Springer Series in Advanced Manufacturing. Cham: Springer. 2022, pp. 51–77
- [8] Mildenhall, B.; Srinivasan, P. P.; Tancik, M. et al.: NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. Communications of the ACM 65 (2022) 1, pp. 99–106
- [9] Schonberger, J. L.; Frahm, J.-M.: Structure-from-Motion Revisited. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 4104–4113
- [10] Fischler, M.; Bolles, R.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24 (1981) 6, pp. 381–395



Dr.-Ing. **Patrick Ruediger-Flore**
Foto: RPTU Kaiserslautern

Matthias Klar, M.Sc.

Marco Hussong, M.Sc.

Jan Mertes, M.Sc.

Dr.-Ing. **Li Yi**

Moritz Glatt, M.Sc.

Dr.-Ing. **Patrick Kölsch**

Prof. Dr.-Ing. **Jan C. Aurich**

Lehrstuhl für Fertigungstechnik
und Betriebsorganisation [FBK]
Rheinland-Pfälzische Technische Universität RPTU
Gottlieb-Daimler-Str. 42, 67663 Kaiserslautern
Tel. +49 631 / 205-2618
patrick.ruediger@mv.uni-kl.de
www.mv.uni-kl.de/fbk/

LIZENZ



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