DGK Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 899

Simon Burkard

Geodatenbasierte Lokalisierung für Outdoor Augmented Reality

München 2023

Bayerische Akademie der Wissenschaften

ISSN 0065-5325 ISBN 978-3-7696-5311-3

Diese Arbeit ist gleichzeitig veröffentlicht in:
RosDoc – Online Publikationen der Universität Rostock
https://doi.org/10.18453/rosdok_id00004305, Rostock 2023

DGK Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 899

Geodatenbasierte Lokalisierung für Outdoor Augmented Reality

An der Agrar- und Umweltwissenschaftlichen Fakultät
der Universität Rostock
zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing.)
vorgelegte Dissertation

von

Simon Burkard, M. Sc.

Berlin

München 2023

Bayerische Akademie der Wissenschaften

ISSN 0065-5325 ISBN 978-3-7696-5311-3

Diese Arbeit ist gleichzeitig veröffentlicht in:
RosDoc – Online Publikationen der Universität Rostock
https://doi.org/10.18453/rosdok_id00004305, Rostock 2023

Adresse der DGK:

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften (DGK)

Alfons-Goppel-Straße 11 ● D – 80 539 München

Telefon +49 – 331 – 288 1685 ● E-Mail post@dgk.badw.de

http://www.dgk.badw.de

Prüfungskommission:

Vorsitzender: Prof. Dr. Henning Bombeck (Universität Rostock)

Referent: Sen.-Prof. Dr. Ralf Bill (Universität Rostock)

Korreferenten: Prof. Dr. Frank Fuchs-Kittowski (HTW Berlin)
Prof. Dr. Jörg Blankenbach (RWTH Aachen)

Tag der mündlichen Prüfung: 27.03.2023

© 2023 Bayerische Akademie der Wissenschaften, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen

ISSN 0065-5325 ISBN 978-3-7696-5311-3

Danksagung

Ohne die Unterstützung vieler Menschen wäre die Umsetzung dieser Arbeit nicht möglich gewesen. Ich möchte mich daher bei allen herzlich bedanken, die mich in der Vergangenheit bei meiner Forschungstätigkeit unterstützt haben. Mein Dank gilt dabei den Projektträgern und Mittelgebern sowie allen Projektpartnern der Forschungsprojekte *VGI4HWM*, *mARGo* und *AR4WIND* sowie allen Kolleg:innen und Student:innen, die bei meiner Forschung in den vergangenen Jahren mitgewirkt und mir bei der Entwicklung dieser Arbeit geholfen haben.

Ich möchte mich an dieser Stelle auch bei Herrn Prof. Ralf Bill bedanken, der als Doktorvater diese Arbeit begleitet und durch wertvolles Feedback zum Gelingen der Arbeit beigetragen hat. Ein ganz besonderer Dank gilt Herrn Prof. Frank Fuchs-Kittowski, der mich nicht nur zur Dissertation motiviert, sondern mich auch mehrere Jahre lang an der HTW Berlin in meiner wissenschaftlichen Tätigkeit und bei zahlreichen Publikationen gefördert und unterstützt hat. Außerdem möchte ich mich ganz herzlich bei Dr. Rahul Swaminathan bedanken. Als Mentor und Kollege bei den Telekom Innovation Laboratories hat er mir bereits während meiner Studienzeit geholfen, die komplexen Herausforderungen der Computer Vision zu erarbeiten und schon früh meine Begeisterung für das Thema der Augmented Reality geweckt.

Zuletzt gilt ein großer Dank meiner Familie. Anni, Nuri und Fiete, ohne euch und eure Unterstützung wäre diese Arbeit nicht möglich gewesen!

Zusammenfassung

Die Technologie der Augmented Reality (AR) ermöglicht neue Arten der Visualisierung, indem virtuelle Informationen positionsgetreu in der Kameraansicht der realen Umgebung auf mobilen Endgeräten eingeblendet werden. Auch Vor-Ort-Visualisierungen im Freien können damit einfacher und verständlicher gestaltet werden, z.B. im Kontext von Bauplanungen durch eine Darstellung von geplanten Bauwerken im realen Landschaftsbild. Eine große Herausforderung bei der AR-Visualisierung von Geodaten (GeoAR) besteht in der präzisen globalen Lokalisierung (Registrierung) des AR-Systems, d.h. bei der Bestimmung von Position und Orientierung (Pose) des Endgeräts in Bezug auf ein geografisches Referenzsystem. Die Genauigkeit der üblicherweise in mobilen Geräten integrierten Sensoren zur Bestimmung der globalen Geräte-Pose ist für realitätsnahe Visualisierungen meist nicht ausreichend und die Nutzung alternativer Registrierungsmethoden daher notwendig.

In dieser Arbeit wird ein flexibles manuelles Registrierungsverfahren für Outdoor AR-Anwendungen vorgestellt, welches existierende 3D-Geodatenmodelle (Oberflächenmodelle, Geländemodelle und Stadtmodelle) als virtuelle Repräsentation der Außenumgebung in der AR-Umgebung visualisiert. Durch Interaktion der Nutzer:innen ist anschließend ein Angleichen der virtuellen Umgebung an die reale Welt und somit eine globale Registrierung des AR-Systems möglich.

Die Arbeit diskutiert zunächst die besonderen Herausforderungen der Technologie bei der AR-Visualisierung von Geodaten und präsentiert die Stärken und Schwächen existierender Arbeiten im Bereich der AR-Registrierung. Anschließend werden die Hauptkomponenten der entwickelten Methode präsentiert. Dazu zählen insbesondere eine Geodatenverarbeitungskette zur Umwandlung der eingesetzten Geodaten in geeignete AR-fähige Formate sowie die technische und mathematische Beschreibung der entworfenen Methode. Im Rahmen einer praxisnahen Evaluation werden schließlich die Genauigkeit und Bedienbarkeit sowie das Anwendungspotenzial des Verfahrens diskutiert und Handlungsempfehlungen für einen idealen Praxiseinsatz vorgeschlagen.

Der Praxiseinsatz verdeutlicht, dass das manuelle Registrierungsverfahren gut dazu geeignet ist, die globale Pose eines mobiles AR-System in unterschiedlichen Umgebungen auf einfache und intuitive Art und Weise zu bestimmen, so dass Outdoor AR-Szenarien mit hohem Realitätsgrad realisiert werden können.

Abstract

Augmented reality (AR) is a technology that enables new types of visualization by superimposing virtual information in the camera view of the real environment on mobile devices. On-site outdoor visualizations can thus be made simpler and more comprehensible, e.g. in the context of construction planning by realistically representing planned structures in the real landscape.

A major challenge in AR visualizations of geospatial data (GeoAR) is the precise global localization (registration) of the AR system, i.e., determining the position and orientation (pose) of the end device with respect to a geographic reference system. The accuracy of the localization sensors integrated in mobile devices is usually insufficient for realistic visualizations. The use of alternative registration methods is therefore necessary.

In this work, a flexible manual registration method for outdoor AR applications is presented. It visualizes existing 3D geospatial models (3D surface models, 3D terrain models, and 3D city models) as a virtual representation of the outdoor environment in the AR view. Through user interaction, an alignment of the virtual environment with the real world and thus a global registration of the AR system can be achieved in a fast and efficient way.

The work first discusses the particular challenges of the AR technology for visualizations of geospatial data and presents the strengths and weaknesses of existing research in the field of AR registration. Then, the main components of the developed method are outlined. These include a geospatial data processing pipeline for converting geospatial data into suitable AR-enabled formats, as well as the technical and mathematical description of the designed method. Finally, the accuracy and usability as well as the actual potential of the method are discussed in the context of an user-centered evaluation. Recommendations for an ideal use of application in practice are proposed eventually.

The field tests illustrate that the developed user-aided registration method is well suited to determine the global pose of a mobile AR system in different environments in a simple and intuitive way, thus enabling outdoor AR scenarios with a high degree of realism.

Inhaltsverzeichnis

1 Einleitung **1**

1.1 Motivation 1
1.2 Problemstellung 3
1.3 Zielstellung 3
1.4 Forschungsfragen 5
1.5 Struktur der Arbeit 6

2 Stand der Forschung und Entwicklung zu Augmented Reality **9**

2.1 Begriffsdefinition und Ausprägungen von AR-Systemen 10
2.2 Technologie der Mobile Augmented Reality 13
2.2.1 AR-Visualisierungen als fortlaufender Kreislauf 13
2.2.2 Tracking, Kalibrierung und Registrierung 14
2.2.3 Mobile AR-Sensoren 17
2.3 Geodatenbasierte AR 22
2.3.1 Begriffsdefinition und Abgrenzung 22
2.3.2 Herausforderungen 23
2.4 Geodatenbasierte AR-Anwendungen 26
2.5 Mathematische Grundlagen der Projektion von 3D-Inhalten in den 2D-Raum 28
2.6 Zusammenfassung 33

3 Stand der Forschung und Entwicklung zur AR-Registrierung **35**

3.1 Lokale Registrierung von AR-Systemen 36
3.1.1 Registrierung anhand bekannter 2D-Referenzmarker 37
3.1.2 Registrierung mittels Visual-SLAM-Technologie 38
3.1.3 Registrierung mittels Visual-Inertial-Odometry 40
3.2 Globale Registrierung von AR-Systemen 41
3.2.1 Pose-Bestimmung mittels GNSS- und IMU-Sensoren 41
3.2.2 Automatische bildbasierte globale Registrierung 43
3.2.3 Nutzergesteuerte globale Registrierung 46
3.3 Zusammenfassung und Forschungslücke 48

4 Nutzergesteuerte geodatenbasierte AR-Registrierung 51
4.1 System-Architektur und -Komponenten . . . 52
4.2 Geodatenverarbeitungskette . . . 56
4.2.1 Digitale 3D-Geodaten . . . 56
4.2.2 3D-Geodatenverarbeitung . . . 62
4.3 Geodaten-CMS . . . 70
4.4 Globales Registrierungs- und Tracking-System . . . 73
4.4.1 Koordinatensystem-Transformationen . . . 74
4.4.2 Kalibrierung der globalen Position und Orientierung . . . 78
4.5 Zusammenfassung . . . 84

5 Evaluation 87
5.1 Hypothesen zur Evaluation . . . 88
5.2 Mobile Evaluations-Anwendung . . . 89
5.2.1 Anforderungen . . . 89
5.2.2 User Flow . . . 92
5.2.3 Grafische Benutzeroberfläche . . . 94
5.2.4 Implementierung . . . 96
5.3 Testmethode . . . 100
5.3.1 Evaluationsmethoden und Bewertungskriterien . . . 101
5.3.2 Evaluationsablauf . . . 106
5.3.3 Test-Standorte . . . 108
5.4 Ergebnisse . . . 111
5.5 Diskussion und Zusammenfassung . . . 122
5.5.1 Diskussion der Ergebnisse . . . 122
5.5.2 Handlungsempfehlungen . . . 130

6 Zusammenfassung und Ausblick 133
6.1 Zusammenfassung der Ergebnisse . . . 133
6.2 Ausblick . . . 136

Literaturverzeichnis 143

Abbildungsverzeichnis 155

Abkürzungsverzeichnis 157

Eidesstattliche Erklärung 159

Lebenslauf 161

Kapitel 1

Einleitung

1.1 Motivation

Bei der mobilen Augmented Reality (AR) wird die menschliche Wahrnehmung der unmittelbaren realen Umgebung durch virtuelle Informationen überlagert und somit erweitert [11]. Dies ermöglicht neue Arten der Visualisierung und neuartige Wege, mit digitalen Informationen zu interagieren. Die AR-Technologie wird heutzutage primär dazu eingesetzt, um virtuelle Inhalte in der unmittelbaren Umgebung im Nahbereich der Nutzer:innen – in der Regel in Innenräumen – darzustellen (*indoor*), z.B. zur Inneneinrichtung von Wohnräumen [111]. Die AR-Technologie besitzt jedoch auch ein großes Potenzial, mobile Anwendungen zur Visualisierung im Freien (*outdoor*) und in einer großräumigen Umgebung (*large-scale environment*) einfacher und effizienter zu gestalten, indem virtuelle Informationen bzw. Modelle (z.B. Gebäudedaten und -modelle, geplante Windkraftanlagen o.ä.) positionsgetreu in der Live-Kameraansicht der realen Umgebung visualisiert werden [25, 73].

Auf diese Weise können vielfältige Outdoor AR-Szenarien mit hohem Realitätsgrad entstehen, bei denen virtuelle Daten in die naturräumlichen Gegebenheiten eingepasst und somit im realen Kontext besser verstanden, diskutiert und dokumentiert werden können. Derartige Vor-Ort-Visualisierungen können Arbeitsprozesse in verschiedenen Bereichen optimieren, beispielsweise im Rahmen von Bauplanungen (z.B. 3D-Darstellung eines geplanten Gebäudes) oder bei Navigationsaufgaben (z.B. realitätsgetreue Einblendung von Routen oder Points of Interests). Auch im Kontext von Beteiligungsprozessen bei geplanten Neubauprojekten kann mit Hilfe der AR-Technologie eine realitätsnahe Vor-Ort-Visualisierung erfolgen, um die tatsächlichen Auswirkungen von geplanten Bauwerken (z.B. Brücken, Stromtrassen, Windparks etc.)

auf das reale Landschaftsbild von verschiedenen Perspektiven aus den verschiedenen Interessensgruppen (z.B. Anwohnern, Behörden) zu kommunizieren [26, 68].

Gründe für eine bisher nur zögerlich erfolgte Einführung und Etablierung von erfolgreichen AR-Anwendungen im Außenbereich liegen in den komplexen technischen Herausforderungen, die notwendig sind, um AR-Darstellungen korrekt und benutzerfreundlich realisieren zu können [8, 63]. Die Grundvoraussetzung für die Akzeptanz von AR-Darstellungen und somit für eine erfolgreiche Verbreitung von AR-Anwendungen liegt vor allem in einer möglichst genauen und realitätsnahen Visualisierung der virtuellen Inhalte im Kamerabild. Für eine überzeugende AR-Visualisierung sind dabei mehrere Rahmenbedingungen notwendig: Zum einen müssen die darzustellenden virtuellen Inhalte in möglichst hochwertiger Qualität bereitgestellt werden, z.B. als hochaufgelöstes texturiertes 3D-Modell mit geeigneten Materialeigenschaften. Zum anderen ist eine präzise Kenntnis der 3D-Umgebung (z.B. zur korrekten Wiedergabe von Verdeckungen) sowie eine präzise Lokalisierung des mobilen Endgeräts (Smartphone, Tablet) in der realen Welt notwendig, d.h. eine genaue Bestimmung von Position und Orientierung des AR-Systems (Geräte-Pose). Dieser Prozess, virtuelle und reale Welt korrekt anzugleichen, d.h. virtuelle AR-Objekte an den realen Objekten auszurichten, so dass die Illusion einer kohärenten Umgebung steht, wird als Registrierung (AR-Registrierung) bezeichnet [126, S.179ff].

Zum Zweck einer genauen Registrierung (Lokalisierung) des mobilen AR-Systems innerhalb der physischen Welt müssen Position (Translation) und Orientierung (Rotation) des mobilen Endgeräts fortwährend in sechs Freiheitsgraden bestimmt werden [126, S.86]. Für kleinräumige AR-Szenarien (z.B. in Innenräumen) kann die Bestimmung dieser Geräte-Pose in der Regel in Bezug auf ein lokales Koordinatensystem ohne Bezug zu einem geografischen Referenzsystem erfolgen. Die Bestimmung der genauen Position und Ausrichtung des Geräts erfolgt dabei im Nahbereich fortwährend durch inzwischen etablierte und zuverlässige bildbasierte Tracking-Verfahren [126, S.105ff]. Infolgedessen sind robuste AR-Tracking-Tools für AR-Anwendungen in kleinräumigen Umgebungen entstanden, die mittlerweile auch von großen kommerziellen Unternehmen wie Google oder Apple bereitgestellt werden.

Zur Darstellung digitaler georeferenzierter Informationen, z.B. Geodaten aus Geo-Informationsystemen (GIS), innerhalb mobiler Outdoor AR-Anwendungen ist es jedoch notwendig, das AR-System nicht nur relativ zu einem beliebigen lokalen Bezugssystem, sondern relativ zu einem globalen Geo-Koordinatensystem zu registrieren (globale Registrierung) [126, S.92f]. Derartige Outdoor AR-Anwendungen können als GeoAR-Anwendungen bezeichnet werden.

1.2 Problemstellung

Eine solche globale Registrierung eines handelsüblichen mobilen AR-Systems ist nach wie vor eine große Herausforderung, insbesondere aufgrund der Ungenauigkeiten der gängigen Sensoren in den mobilen Endgeräten zur Bestimmung der globalen Geräte-Pose [8]. Neben der Notwendigkeit einer präzisen Pose-Bestimmung besteht außerdem eine weitere wichtige Voraussetzung für realitätsnahe AR-Anwendungen in der korrekten digitalen dreidimensionalen Erfassung der virtuellen Umgebung. Nur durch Kenntnis der 3D-Struktur der Umgebung kann ein AR-System mögliche Verdeckungen der darzustellenden virtuellen Inhalte durch reale physische Objekte wiedergeben und somit realistische Darstellungen erzeugen. Solche Verdeckungen können bei Outdoor AR-Anwendungen beispielsweise durch Vegetation, Gelände oder Gebäude entstehen.

Für die Lösung dieser Herausforderungen bei GeoAR-Anwendungen existieren einige erfolgversprechende Ansätze, die auf vollautomatischen Bilderkennungsverfahren zuvor registrierter, großflächiger 3D-Umgebungen basieren [126, S.164ff]. Diese sind jedoch zeitaufwändig, fehleranfällig und nicht auf beliebige großflächige Gebiete übertragbar, da sie große Datenbanken mit zuvor aufgenommenen georeferenzierten Bildern bzw. 3D-Punktwolken als Referenzmodelle benötigen. Die Entwicklung von alternativen flexibleren Registrierungsverfahren für GeoAR-Anwendungen mit einfacherer Handhabung ist daher notwendig.

Manuelle AR-Registrierungsmethoden können unter Umständen eine solche Alternative bieten. Die Grundidee besteht dabei darin, dass durch aktive Mithilfe der Nutzer:innen bei Ausrichtung und Positionierung des AR-Systems ein flexibles und robustes Registrierungsverfahren entstehen kann, das in vielen Situationen vollautomatischen Erkennungsverfahren überlegen ist. Die wenigen existierenden manuellen Verfahren können jedoch globale Position und Ausrichtung des Systems nicht in vollständigem Umfang bzw. nicht mit hoher Präzision bestimmen oder sind nur unter bestimmten Rahmenbedingungen und in eingeschränkten Umgebungen (z.B. in bergiger Umgebung) nutzbar [48, 72, 132].

1.3 Zielstellung

Hauptziel dieser Forschungsarbeit ist daher die Entwicklung und Evaluation eines nutzergesteuerten Verfahrens, das vollumfänglich zur robusten und effizienten globalen Registrierung von mobilen Outdoor AR-Systemen geeignet ist. Dazu soll eine manuelle bzw. halbautomatische Registrierungsmethode vorgestellt werden, die existierende 3D-

Geodatenmodelle (3D-Oberflächenmodelle, 3D-Geländemodelle und 3D-Stadtmodelle) während des Registrierungsprozesses als Referenzobjekte miteinbezieht und durch Benutzerinteraktion ein Angleichen (Kalibrierung) der virtuellen AR-Umgebung mit der realen Welt ermöglicht.

Derartige 3D-Geodatenmodelle können die Außenumgebung der Nutzer:innen in unterschiedlicher Art und Weise virtuell beschreiben und stehen bereits in verschiedenen Ausprägungen in Fachanwendungen oder über Geo-Webdienste zur Verfügung. Neben gewöhnlichen Offline-Visualisierungen von 3D-Geodatenmodellen innerhalb von Geo-Informationssystemen bieten derartige Geodaten als Hilfsmittel zur Registrierung von mobilen AR-Systemen einen vollkommen neuartigen Einsatzzweck. Ein Verwendung komplexer und aufwändiger bildbasierter globaler Registrierungsverfahren ist dann nicht mehr notwendig.

Die Grundidee einer solchen nutzergesteuerten Registrierung mithilfe von 3D-Geodatenmodellen ist in Abbildung 1.1 skizziert und lässt sich wie folgt beschreiben: Verschiedenartige 3D-Geodaten werden aufbereitet und in der Live-AR-Ansicht des mobilen Endgeräts als virtuelle Repräsentation der Umgebung dargestellt. Basierend auf einer initialen Bestimmung der Geräte-Pose mittels GNSS (*Global Navigation Satellite System*) zur ungefähren Positionsbestimmung und IMU (*Inertial Measurement Unit*) zur Schätzung der Geräteorientierung kann bereits eine erste grobe, aber unpräzise und

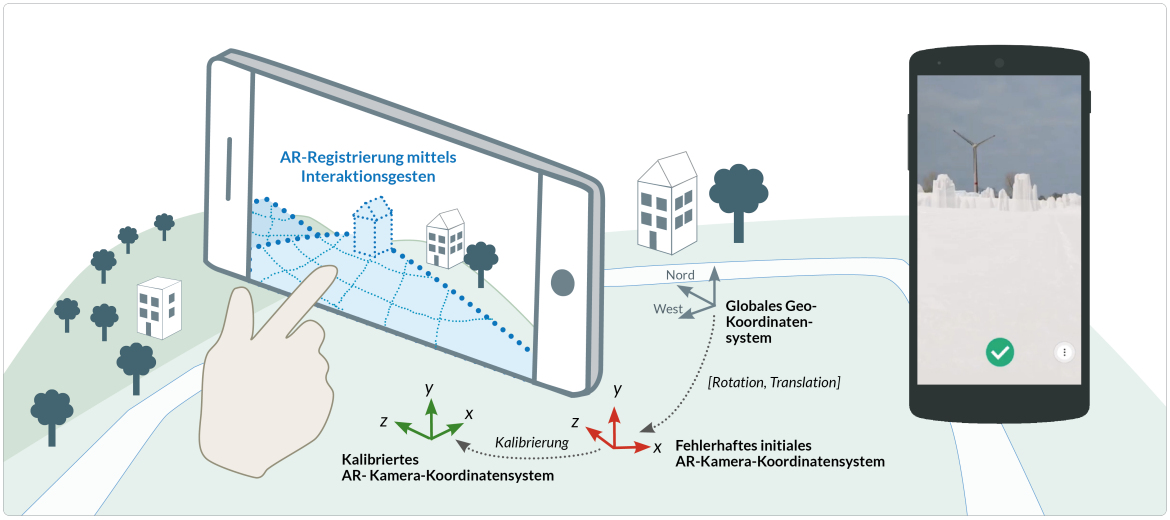


Abbildung 1.1 Grundidee der nutzergesteuerten globalen Registrierungsmethode: Digitale 3D-Geodatenmodelle (z.B. 3D-Oberflächenmodelle) werden als visuelle Repräsentation der aktuellen Umgebung in die AR-Ansicht integriert und mit grober, jedoch fehlerhafter AR-Projektion dargestellt. Durch Benutzerinteraktion können die Modelle so verschoben und der realen Umgebung angepasst werden, dass eine Korrektur von globaler Position und Ausrichtung des AR-Systems möglich ist (Kalibrierung der Geräte-Pose).

fehlerhafte AR-Darstellung der 3D-Geodatenmodelle im Kamerabild erfolgen. Durch unterschiedliche Formen von benutzergesteuerter Interaktion können diese anfangs fehlerhaft platzierten Geodatenmodelle schließlich auf dem Bildschirm so ausgerichtet und verschoben werden, dass die virtuellen Umgebungsmodelle mit der tatsächlichen Abbildung der physischen Realität im Kamerabild möglichst passend übereinstimmen. Dadurch ist eine Korrektur der globalen Geräteposition und -rotation und somit eine Kalibrierung des AR-Systems möglich. Durch diesen Ansatz, die Nutzer:innen selbst im Registrierungsprozess miteinzubeziehen, kann ein neuartiges und unkompliziertes, aber gleichzeitig sehr effektives und robustes Verfahren entstehen, das in vielfältigen GeoAR-Anwendungsszenarien zum Einsatz kommen kann, um eine realitätsnahe AR-Visualisierung von georeferenzierten AR-Inhalten zu ermöglichen. Gleichzeitig eröffnet sich durch diese Grundidee ein neuartiger und praxisnaher Einsatzzweck für die mobile Visualisierung von dreidimensionalen Geodaten.

1.4 Forschungsfragen

Diese Arbeit behandelt also das Thema der Registrierung von mobilen AR-Systemen für geodatenbasierte AR-Szenarien. Der wissenschaftliche Kernbeitrag der Arbeit liegt dabei in der Lösung der Forschungsfrage, wie – unter Einbeziehung von 3D-Geodatenmodellen – eine genaue und nutzerfreundliche AR-Registrierung ermöglicht werden kann, bei der die globale Position und Orientierung eines mobilen Geräts in Bezug auf ein geografisches Referenzkoordinatensystem präzise in sechs Freiheitsgraden bestimmt wird. Im Gesamtkontext der Arbeit sollen dabei die folgenden Forschungsfragen im Detail betrachtet und gelöst werden:

- **Herausforderungen von geodatenbasierter Augmented Reality**:
  Herausarbeiten von Besonderheiten und Herausforderungen der AR-Technologie zur Darstellung von Geodaten im Außenbereich (Kapitel 2.3: Geodatenbasierte AR) sowie der Schwierigkeiten der Registrierung von GeoAR-Systemen, insbesondere im Vergleich zur lokalen AR-Registrierung im Nahbereich (Kapitel 3: Stand der Forschung und Entwicklung zur AR-Registrierung).

- **Entwicklung eines GeoAR-Registrierungsverfahrens**:
  Entwicklung eines neuartigen Verfahrens, um durch Nutzerinteraktionen und mithilfe von 3D-Geodatenmodellen die globale Geräteposition und -ausrichtung in bis zu sechs Freiheitsgraden mit metrischer Skalierung zu korrigieren (Kapitel 4: Nutzergesteuerte geodatenbasierte AR-Registrierung).

- **Evaluation des entwickelten Registrierungsverfahrens**:
  Praxisnahe Evaluation von Genauigkeit und Bedienbarkeit des Verfahrens in unterschiedlichen Umgebungen zur Untersuchung des Anwendungspotenzials und zum Ableiten von konkreten Handlungsempfehlungen für den Praxiseinsatz der entwickelten Methode (Kapitel 5: Evaluation).

1.5 Struktur der Arbeit

In **Kapitel 2** der Arbeit sollen zunächst wissenschaftlich-technische Grundlagen der AR- bzw. GeoAR-Technologie erläutert werden. Ziel dieses Kapitels ist das Herausarbeiten von Basiswissen, das zur Realisierung von AR-Anwendungen im Allgemeinen und von GeoAR-Anwendungen im Speziellen notwendig ist. Ausgehend von der Fragestellung, wie die technischen Herausforderungen bei mobilen AR-Anwendungen (z.B. Bestimmung von Geräte-Pose) in der Regel gelöst werden und welche mobilen Sensoren dazu üblicherweise verwendet werden, erfolgt anschließend eine besondere Betrachtung der geodatenbasierten AR-Technologie. Dabei sollen die besonderen Herausforderungen und technischen Möglichkeiten der AR-Technologie bei der Visualisierung von Geodaten im Außenbereich präsentiert und die Wichtigkeit von präzisen Registrierungsverfahren für geodatenbasierte AR-Anwendungen herausgestellt werden. Auch beispielhafte GeoAR-Anwendungsszenarien werden in diesem Zusammenhang vorgestellt. Abschließend erfolgt in diesem Grundlagenkapitel eine mathematische Beschreibung des Projektionsverfahrens von 3D-Inhalten in den 2D-Raum. Dieser Abbildungsvorgang bildet die Grundlage für die immersive Wahrnehmung von AR-Visualisierungen und damit auch für das Verständnis der entwickelten Registrierungsmethode.

Nachdem in Kapitel 2 bereits die hohe Bedeutung einer genauen Registrierung von mobilen AR-Systemen für realitätsnahe AR-Visualisierungen verdeutlicht wird, sollen in **Kapitel 3** konkrete aktuelle Entwicklungs- und Forschungsarbeiten im Bereich der AR-Registrierung sowie deren Stärken und Schwächen im Praxiseinsatz vorgestellt und erläutert werden. Ausgehend von der aktuellen Forschung zur Bestimmung der Geräte-Pose mobiler AR-Systeme im Nahbereich werden schließlich existierende globale Registrierungsverfahren für Outdoor AR-Anwendungen und existierende Ansätze für benutzergesteuerte Registrierungsmethoden vorgestellt. Die Unterschiede und Limitierungen der jeweiligen existierenden Arbeiten werden dabei erläutert. Ein Herausarbeiten der Forschungslücke im Bereich der geodatenbasierten AR-Registrierung betont schließlich den Innovationsgehalt der hier entwickelten Methode.

In **Kapitel 4** wird das entwickelte geodatenbasierte Registrierungsverfahren im Detail vorgestellt. Dazu werden zunächst notwendige System-Komponenten des Verfahrens präsentiert, unter anderem eine Geodatenverarbeitungskette zur Prozessierung und Umwandlung von 3D-Geodaten in AR-geeignete Formate. In diesem Kontext werden die Spezifikationen der eingesetzten Geodatenmodelle und die einzelnen Verarbeitungsschritte zur Generierung von kleinflächigen 3D-Kacheln beschrieben, die zur effizienten Speicherung und Visualisierung der virtuellen 3D-Landschaft innerhalb des AR-Clients genutzt werden. Anschließend erfolgt die Erläuterung der Funktionsweise der entwickelten Methode. Dazu werden die Koordinaten- und Bezugssysteme, die im gesamten Registrierungsprozess involviert sind, beschrieben und die Interaktionsformen zur Kalibrierung von globaler Geräteposition und -orientierung mathematisch herausgearbeitet.

In **Kapitel 5** wird die praxisnahe Evaluation der entwickleten Methode hinsichtlich Genauigkeit und Bedienbarkeit an verschiedenen Standorten vorgestellt. Die Evaluation erfolgt auf Basis einer prototypischen Anwendung zur AR-Visualisierung von Windenergieanlagen im Landschaftsbild. Dazu wird zunächst das fachliche und technische Konzept dieser Evaluationsanwendung präsentiert sowie die Vorgehensweise und Methodik bei der Durchführung der Evaluation erläutert. Die anschließend vorgestellten und diskutierten Evaluationsergebnisse der Vor-Ort-Tests heben die Vorteile, aber auch die Beschränkungen und Limitierungen des Verfahrens hervor. Daraus abgeleitete konkrete Handlungsempfehlungen für einen optimalen Praxiseinsatz der AR-Registrierungsmethode schließen das Kapitel ab.

Kapitel 6 fasst die Inhalte und Ergebnisse der Arbeit zusammen und stellt die Potenziale sowie mögliche Ansätze für Optimierungen und Weiterentwicklungen des geodatenbasierten Registrierungsverfahrens vor.

Kapitel 2

Stand der Forschung und Entwicklung zu Augmented Reality

Für das Verständnis der entwickelten geodatenbasierten AR-Registrierungsmethode ist es notwendig, die Funktionsweise der Augmented Reality Technologie im Allgemeinen, sowie die Besonderheiten und Herausforderungen der geodatenbasierten AR-Technologie im Speziellen zu verstehen. Ziel der Arbeit soll dabei nicht sein, sämtliche Facetten und Ausprägungen der AR-Technologie aufzuzeigen. Stattdessen sollen nur diejenigen wesentlichen Aspekte der Technologie betrachtet werden, die auch im weiteren Verlauf für die Entwicklung der geodatenbasierten AR-Registrierungsmethode von grundsätzlicher Bedeutung sind. Detaillierte Ausführungen zur AR-Technologie sind in weiterführender Literatur zu finden, beispielsweise im Werk von Schmalstieg und Hollerer [126].

Nach einer kurzen Begriffsdefinition der AR-Technologie und einer kurzen Vorstellung üblicher Ausprägungen von AR-Systemen (Kapitel 2.1) wird im folgenden Kapitel zunächst erläutert, wie AR-Visualisierungen auf mobilen Endgeräten technisch erzeugt werden und welche Arbeits- und Verarbeitungsschritte für AR-Projektionen notwendig sind (Kapitel 2.2). Dabei werden die Konzepte des AR-Trackings und der AR-Registrierung genauer beleuchtet und die dafür in der Regel verwendeten mobilen AR-Sensoren im Detail beschrieben. Anschließend erfolgt eine gesonderte Betrachtung der geodatenbasierten Augmented Reality (GeoAR) und ein Herausarbeiten der besonderen Herausforderungen bei der AR-basierten Visualisierung von Geodaten im Außenbereich (Kapitel 2.3). Mögliche GeoAR-Szenarien werden anschließend in Kapitel 2.4 vorgestellt. Dort wird eine Auswahl an existierenden geodatenbasierten AR-Anwendungen präsentiert wird.

Die AR-Technologie bietet den Vorteil, dass virtuelle 3D-Inhalte realitätsgetreu aus verschiedenen Positionen und Blickwinkeln und somit mit einem hohen Grad an

Immersion wahrgenommen werden können. Der immersive Effekt entsteht durch die Bewegung des AR-Systems in Verbindung mit einer realitätsgetreuen und perspektivisch korrekten Projektion der virtuellen 3D-Inhalte in einen 2D-Raum, nämlich den Bildschirm des AR-Endgeräts. Das Grundlagenkapitel endet daher mit einem Blick auf die Erzeugung von mobilen AR-Visualisierungen aus der Perspektive der Computer Vision und der mathematischen Beschreibung (Kapitel 2.5) der perspektivischen Projektion von 3D-Punkten in den 2D-Raum. Dies bildet schließlich die Grundlage für das Verständnis der entwickelten Methode zur manuellen Registrierung von globaler Geräte-Position und -Ausrichtung.

2.1 Begriffsdefinition und Ausprägungen von AR-Systemen

Bei der *Virtual Reality* (Virtuelle Realität; VR) tauchen die Nutzer:innen vollständig in eine synthetische Umgebung ein und können die reale Welt gar nicht mehr sehen. Im Gegensatz dazu können Nutzer:innen bei der **Augmented Reality** (Erweiterte Realität) die reale Umgebung sehen, wobei virtuelle (computer-generierte) Inhalte oder Objekte die reale Welt überlagern bzw. mit dieser verschmolzen sind [11]. Der Übergang von Realität zur vollständig virtuellen Umgebung und damit die Abgrenzung zwischen Augmented Reality und Virtual Reality kann durch das *Reality-Virtuality Continuum* beschrieben werden, welches von Milgram et al. eingeführt wurde [97].

Nach der Definition von Azuma wird bei der Augmented Reality die Wahrnehmung der realen Welt so angereichert, dass die virtuellen Informationen als Teil der realen räumlichen Umgebung wahrgenommen werden und die Nutzer:innen im Idealfall annehmen würden, dass die virtuellen und realen Objekte im selben Raum koexistieren [9]. Ein AR-System wird dabei so definiert, dass es die folgenden drei Eigenschaften besitzen muss:

- Ein AR-System kombiniert reale und virtuelle Objekte innerhalb einer realen Umgebung miteinander.
- Ein AR-System läuft interaktiv und in Echtzeit.
- Ein AR-System registriert reale und virtuelle Objekte miteinander im gleichen 3D-Raum, d.h. es richtet diese in gleicher Richtung und Position aus.

Es ist bei dieser Definition zu beachten, dass der AR-Begriff dabei weder auf ein bestimmtes Ausgabegerät (z.B. mobiles Display), noch auf ein bestimmtes Medium

limitiert wird. Damit kann das Konzept der Augmented Reality theoretisch auf beliebige Sinne (z.B. Hören, Geruch oder Berührung) übertragen werden. In den meisten Fällen - und auch in der folgenden Arbeit - soll jedoch der AR-Begriff allein auf die visuelle Wahrnehmung beschränkt sein.

Bei der **Mobilen Augmented Reality** wird dieses Konzept der Erweiterten Realität auf den mobilen Kontext übertragen, d.h. auf Umgebungen außerhalb von sorgfältig überwachten Arbeitsbereichen und auf mobile Ausgabegeräte, die sich bewegen können und nicht stationär befestigt sind [57]. Mobile AR-Systeme haben insbesondere durch die stark gestiegene Verfügbarkeit und Rechenleistung mobiler Endgeräte an Bedeutung gewonnen und ermöglichen AR-Anwendungen, die sich an den unmittelbaren Ortskontext der Nutzer:innen anpassen können [9], d.h. je nach aktueller Umgebung eine angepasste Situationswahrnehmung ermöglichen und ortskontextbezogene Informationen anzeigen können (*Location-based Services*). Der spezielle Begriff der *Mobile Augmented Reality* ist demzufolge erst durch die breite Verfügbarkeit mobiler tragbarer Endgeräte (Smartphones, Tablets) geprägt worden [12]. Heutzutage werden jedoch bei der überwiegenden Zahl der AR-Anwendungen mobile AR-Systeme eingesetzt. Auch in der folgenden Arbeit sind die Begriffe der *Augmented Reality* bzw. *Outdoor Augmented Reality* implizit stets im Kontext der mobilen AR zu sehen.

In den vergangenen Jahrzehnten sind im Zuge fortschreitender Entwicklungen verschiedene – mobile und stationäre – AR-Systeme in unterschiedlichen Ausprägungen entstanden [9]. Es existieren dabei verschiedene Ansätze und Taxonomien, diese AR-Systeme in unterschiedlichen Kategorien systematisch zu sammeln und zu ordnen, beispielsweise hinsichtlich des AR-Displays (z.B. Projektor, Head Mounted Display oder Handheld-Display), hinsichtlich der Interaktionsform (z.B. Steuerung mittels externem Controller oder Touch-Gesten-Steuerung) oder hinsichtlich des Anwendungsgebietes (z.B. Bildung, Medizin, Tourismus, Gaming etc.) [9, 73, 21, 60]. Im Folgenden sollen zur besseren Einordnung der in dieser Arbeit vorgestellten Technologie typische Ausprägungen von AR-Systemen hinsichtlich Endgerät, Interaktion sowie Kommunikation und Datenbereitstellung kurz vorgestellt werden:

AR-Endgeräte Mobile AR-Anwendungen lassen sich auf unterschiedlichen Endgeräte nutzen. Üblich ist hierbei beispielsweise eine Unterscheidung zwischen videobasierten AR-Überlagerungen (*Video-See-Through*) und optischen AR-Projektionen (*Optical-See-Through*). Bei dem videobasierten Ansatz erzeugt zunächst eine Kamera ein Videobild der realen Welt und stellt dieses auf einem Bildschirm dar (z.B. Kamerabilder auf einem Smartphone- oder Tablet-Display). Darauf werden dann virtuelle AR-Inhalte ein-

geblendet. Die Videokamera kann dabei entweder im AR-Ausgabegerät selbst integriert sein (Smartphone, Tablet) oder die Bilder werden von einem separaten (mobilen) Gerät gesendet (z.B. Live-Bilder einer AR-Drohne). Bei den Optical-See-Through-Geräten werden die virtuellen Inhalte hingegen direkt auf transparente Displays projiziert, z.B. auf die Brillengläser von AR-Brillen (*Head-Mounted Displays*) [36, S.315ff].

Interaktion Je nach Ausprägung und Wahl des AR-Endgeräts unterscheiden sich auch die Möglichkeiten, mit der virtuellen AR-Welt zu interagieren. Während bei videobasierten AR-Systemen (z.B. Smartphones) meist mit Hilfe von Touch-Gesten auf dem Bildschirm mit virtuellen AR-Inhalten interagiert wird, sind bei anderen Endgeräten auch alternative Interaktionsformen üblich. Ein Anvisieren von virtuellen Objekten wird dann beispielsweise durch externe Controller, deren Pose im Raum ebenfalls verfolgt werden kann, ermöglicht oder durch das Erkennen und Verfolgen von Handbewegungen und Handgesten oder der Blickrichtung des Auges (*Eye-Tracking*) [36, S.117ff].

Kommunikation und Daten-Bereitstellung Je nach Art des AR-Systems werden auch unterschiedliche Architekturen und Kommunikationsformen zwischen den einzelnen Systemkomponenten eingesetzt. Während in vielen Fällen alle zur AR-Visualisierung notwendigen Arbeitsschritte (z.B. Videoaufnahme, Datenbereitstellung, AR-Registrierung, Darstellung und Ausgabe) auf demselben Gerät stattfinden können (z.B. Standalone-AR-Anwendung auf dem Smartphone), ist in anderen Fällen eine Verteilung der Verarbeitungsschritte auf mehrere Geräte vorgesehen. So existieren beispielsweise serverbasierte AR-Architekturen, bei denen rechenintensive Operationen an Server ausgelagert werden. Dann werden z.B. im Kontext einer globalen Registrierung einzelne Kamerabilder an einen Server geschickt, der das Kamerabild auf Basis großer Bilddatenbanken analysiert und nur die berechnete Geräte-Pose an das AR-Endgerät zurücksendet [146]. Üblich ist zudem die dynamische Bereitstellung der darzustellenden AR-Inhalte (z.B. 3D-Modelle) über Web-Server und eine Client-Server-Architektur [126, S.379ff].

Auch wenn heutzutage für mobile AR-Szenarien mit intuitiver Bedienung oftmals auch AR-Brillen eingesetzt werden (siehe auch Kapitel 6), sollen im Kontext dieser Arbeit die folgenden Betrachtungen jedoch auf übliche mobile videobasierte Endgeräte, d.h. **Smartphones und Tablets**, beschränkt werden.

2.2 Technologie der Mobile Augmented Reality

Zum besseren Verständnis der Herausforderungen der mobilen AR-Technologie soll im Folgenden die technische Funktionsweise zur korrekten Platzierung virtueller AR-Inhalte im Kamerabild zur Realisierung mobiler AR-Szenarien genauer erläutert werden. Insbesondere die dabei eingesetzten mobilen Sensoren zur Pose-Bestimmung auf handelsüblichen mobilen Endgeräten sollen dabei genauer beschrieben werden.

2.2.1 AR-Visualisierungen als fortlaufender Kreislauf

Die Definition des AR-Begriffs macht deutlich, dass bei einem mobilen AR-System immer eine Interaktivität zwischen Nutzer:innen und System notwendig ist [126, S.2-4]. Das heißt, die Nutzer:innen kontrollieren und steuern die AR-Ansicht, beispielsweise durch Bewegungen der Kamera, um die Szene aus verschiedenen Blickwinkeln zu betrachten. Das AR-System wiederum muss in der Lage sein, diesen Input der Nutzer:innen zu verarbeiten, indem die Bewegungen der mobilen Kamera fortlaufend bestimmt werden, um den aktuellen Standpunkt und die aktuelle Blickrichtung der Kamera – also die **Kamera-Pose** – innerhalb der realen Welt zu kennen. Basierend auf diesen Kenntnissen kann das System wiederum eine aktualisierte AR-Darstellung virtueller Inhalte erzeugen, bei der die virtuellen Inhalte registriert (verankert) an den Objekten der realen Welt bleiben. Es ergibt sich somit ein fortlaufender Kreislauf der AR-Visualisierung, der in Abbildung 2.1 skizziert ist:

Die Nutzer:innen des AR-Systems führen aktiv durch eine Bewegung der mobilen Kamera eine Veränderung der Kamera-Ansicht herbei. Mithilfe des aktualisierten Kamerabildes – sowie oftmals mit Berücksichtigung anderer mobiler Sensoren – kann das System nun eine neue Kamera-Position und -Ausrichtung bestimmen (*Tracking*). Auf Basis dieser aktualisierten Kamera-Pose kann eine erneute Registrierung der virtuellen Inhalte in der realen Welt erfolgen. Die virtuellen AR-Inhalte werden dabei in der physischen Welt verankert. Die Tracking- und Registrierungskomponente benötigt für diese Zwecke Informationen sowohl über die reale Welt als auch über die virtuelle Welt. Diese Informationen können in der Regel aus einer lokalen Datenbank oder einem Online-Service abgerufen werden. Das Modell der realen Welt dient dabei als Referenz, um die Kamera-Pose in der tatsächlichen Welt zu bestimmen. Es kann sich dabei beispielsweise um eine Karte aus 3D-Punkten der realen Welt (3D-Punktwolke, 3D-Karte) handeln oder um Bild-Informationen eines Markers (Referenzbild), der in der realen Welt ausliegt. Eine erfolgreiche Registrierung des AR-Systems ermöglicht schließlich wieder eine aktualisierte Visualisierung der Inhalte im Kamerabild, also

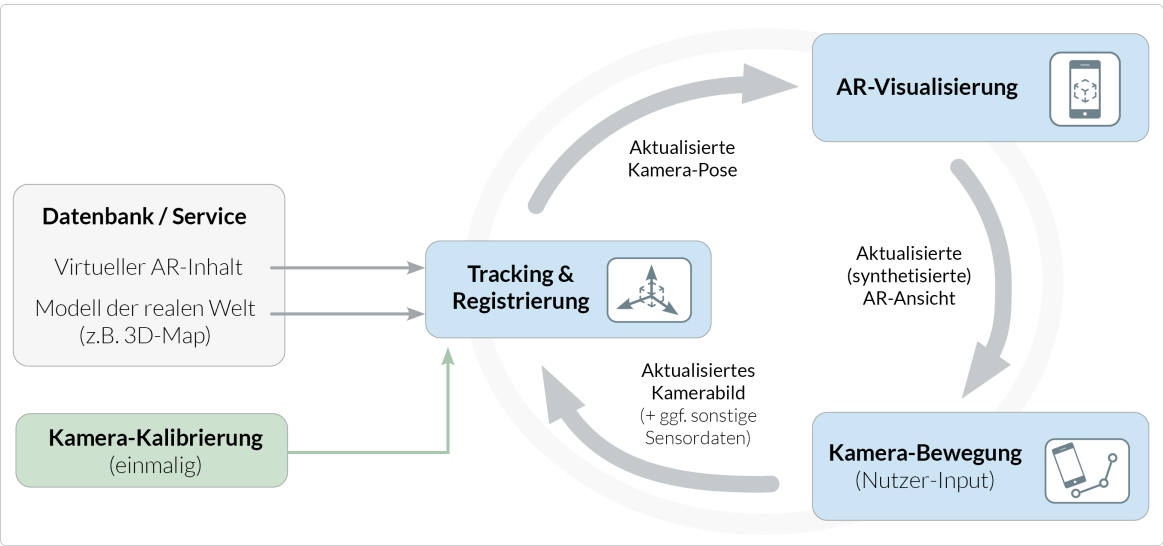


Abbildung 2.1 Kreislauf der AR-Visualisierung: Nutzer:innen steuern und bewegen die AR-Ansicht. Das AR-System erkennt die Kamera-Bewegung und bestimmt eine neue Kamera-Pose zur Registrierung der virtuellen Inhalte in der realen Welt. Dies ermöglicht eine aktualisierte AR-Darstellung (Synthese) virtueller Inhalte im Kamerabild. (Eigene Darstellung in Anlehnung an [126, S.4])

eine aktualisierte Bildsynthese der virtuellen Inhalte (Rendering) auf dem mobilen AR-Display.

2.2.2 Tracking, Kalibrierung und Registrierung

Im Kontext von AR-Systemen existieren drei wesentliche Begrifflichkeiten, die im Zusammenhang mit der korrekten Ausrichtung virtueller AR-Inhalte innerhalb der realen Welt stehen: *Tracking*, *Registrierung* und *Kalibrierung*. Die Bedeutung dieser Begriffe wurde in vorherigen Kapiteln bereits grob beschrieben. Im Folgenden sollen diese nun konkret definiert und abgegrenzt werden:

Tracking

Der Begriff *Tracking* wird genutzt, um das dynamische Erfassen der Kamera-Pose mittels Sensoren zu beschreiben. Um virtuelle AR-Inhalte positionsgetreu im dreidimensionalen Raum visualisieren zu können, muss zumindest die relative Pose des AR-Displays bekannt sein – also die Position und Orientierung der AR-Kamera in Bezug auf die realen Objekte. Eine vollständige Kamera-Pose ist dabei in sechs Dimensionen zu bestimmen: Die Position im 3D-Raum entlang der x-, y- und z-Achse sowie die Orientierung (Rotation) der Kamera um diese drei Achsen, z.B. definiert durch die

Euler-Winkel *Roll*, *Pitch* und *Yaw*. Man spricht im Kontext dieser Dimensionen auch von sechs **Freiheitsgraden** (*Degrees of Freedom; DOF*). Da AR-Systeme in Echtzeit funktionieren sollen, muss das Tracking zur Erfassung der Kamera-Pose fortlaufend erfolgen [126, S.86]. Die Realisierung des Trackings erfolgt auf mobilen Endgeräten in der Regel durch Verwendung der im Gerät integrierten mobilen Sensoren, beispielsweise optische Sensoren (Kamera) oder Orientierungssensoren einer inertialen Messeinheit (siehe Kapitel 2.2.3).

Kalibrierung

Der Prozess der *Kalibrierung* beschreibt üblicherweise das Angleichen von Messgrößen, die mit zwei unterschiedlichen Geräten – einem Referenzgerät und dem zu kalibrierenden Gerät – erfasst wurden. Das Ziel der Kalibrierung besteht darin, Parameter zu finden, um das zu kalibrierende Gerät so einzustellen, dass es brauchbare Messwerte liefert, die im Idealfall mit den Referenzwerten (dem *Normal*) übereinstimmen [126, S.87]. In Bezug auf die AR-Technologie bedeutet das insbesondere, dass die Komponenten des AR-Systems, die das Tracking ermöglichen, kalibriert werden müssen. Während das Tracking fortlaufend erfolgen muss, wird die Kalibrierung in der Regel nur einmalig ausgeführt, beispielsweise vor dem Einsatz des AR-Visualisierungssystems.

Im Kontext dieser Arbeit wird der Begriff der Kalibrierung in zwei Zusammenhängen verwendet: Zum einen beschreibt die Kalibrierung das einmalige Bestimmen von intrinsischen Kamera-Parametern (z.B. Kamera-Brennweite), die im Rahmen der perspektivischen Projektion von 3D-Inhalten auf das 2D-Kamerabild benötigt werden (Kapitel 2.5). Zum anderen kann auch der Prozess zur Bestimmung der realen Kamera-Pose, also die Bestimmung von Translation und Rotation der Kamera, als Kalibrierung (*extrinsic calibration*) bezeichnet werden [137, S.693]. Dementsprechend wird im Folgenden auch die entwickelte Methode (Kapitel 4) zum Auffinden von idealer globaler Translation und Rotation als **geodatenbasierte Kalibrierungsmethode** bezeichnet. Auch hier findet ein Angleichen von Messparametern (nämlich die ungenaue initiale globale Kamera-Pose) des Messsystems (das mobile AR-System) an eine Referenz-Messgröße (die tatsächliche globale Kamera-Pose) statt. Dieser Vorgang ist ebenfalls nicht fortlaufend, sondern in der Regel nur einmalig beim Start der AR-Anwendung durchzuführen.

Registrierung

Der Begriff der *Registrierung* bezieht sich auf das Angleichen der Koordinatensysteme der virtuellen und realen Objekte, d.h. das korrekte geometrische Ausrichten virtueller

Inhalte an der realen Welt, so dass die Illusion einer kohärenten Umgebung entsteht [126, S.179ff]. Eine möglichst akkurate Registrierung der virtuellen AR-Inhalte in der Wahrnehmung der realen Welt ist daher ein primäres Ziel eines AR-Systems. Der Begriff der Registrierung wird in der Literatur oftmals auch synonym mit dem Prozess der **Lokalisierung** (*localization*) verwendet (z.B. [6]). In der folgenden Arbeit soll jedoch für den Vorgang der Angleichung der virtuellen Welt durch Bestimmung der Geräte-Pose ausschließlich der Begriff der Registrierung genutzt werden.

Bei Bewegung der Kamera ist ein präzises *Tracking* der Kamera-Pose also Voraussetzung für eine akkurate **dynamische Registrierung**. Das Erreichen einer **statischen Registrierung** (bei Nicht-Bewegung der Kamera) erfordert wiederum eine erfolgreiche *Kalibrierung* des AR-Systems, um die Koordinatensysteme der virtuellen und realen Objekte in Einklang zu bringen [126, S.87].

Im Kontext der Registrierung kann außerdem zwischen einer lokalen Registrierung und einer globalen Registrierung unterschieden werden (siehe Abbildung 2.2): Bei der **lokalen Registrierung** erfolgt die Bestimmung der Geräte-Pose in Bezug auf ein beliebiges lokales Bezugssystem ohne Georeferenzierung (lokales Weltkoordinatensystem). Sowohl das AR-System als auch die virtuellen AR-Inhalte werden bezogen auf dieses lokale Bezugssystem registriert. In der Regel liegt der Ursprung dieses lokalen Bezugssystems im Nahbereich in der unmittelbaren Anwendungsumgebung (z.B. Ursprung einer initial erstellten 3D-Karte oder Mittelpunkt eines Referenzmarkers). Sofern eine Positionierung der AR-Inhalte in Bezug auf dieses lokale Bezugssystem im Rahmen des AR-Szenarios möglich ist (z.B. AR-Darstellung eines virtuellen Möbelstücks im Wohnzimmer), ist eine lokale Registrierung des AR-Systems ausreichend. Eine **globale Registrierung** ist dann notwendig, wenn darzustellende Inhalte in Bezug auf ein geografisches Koordinatensystem (Geo-Koordinatensystem) definiert sind, also an einer bestimmten Geo-Position in der AR-Ansicht visualisiert werden sollen (Geodaten). In diesem Fall muss auch die Bestimmung der Pose des AR-Systems durch eine globale Registrierung in Bezug auf ebendieses Geo-Koordinatensystem erfolgen.

Auch in dem hier entwickelten Verfahren findet im Kontext der globalen Pose-Bestimmung auch eine Angleichung der Koordinatensysteme zwischen den virtuellen Inhalten (virtuelle 3D-Landschaftsmodelle) und den realen Objekten (tatsächliche Landschaft) statt. Das Verfahren kann daher allgemein auch als **geodatenbasiertes Registrierungsverfahren** bezeichnet werden. Die virtuellen Landschaftsmodelle sind dabei in Bezug auf ein Geo-Koordinatensystem definiert, d.h. es findet eine globale Registrierung statt.

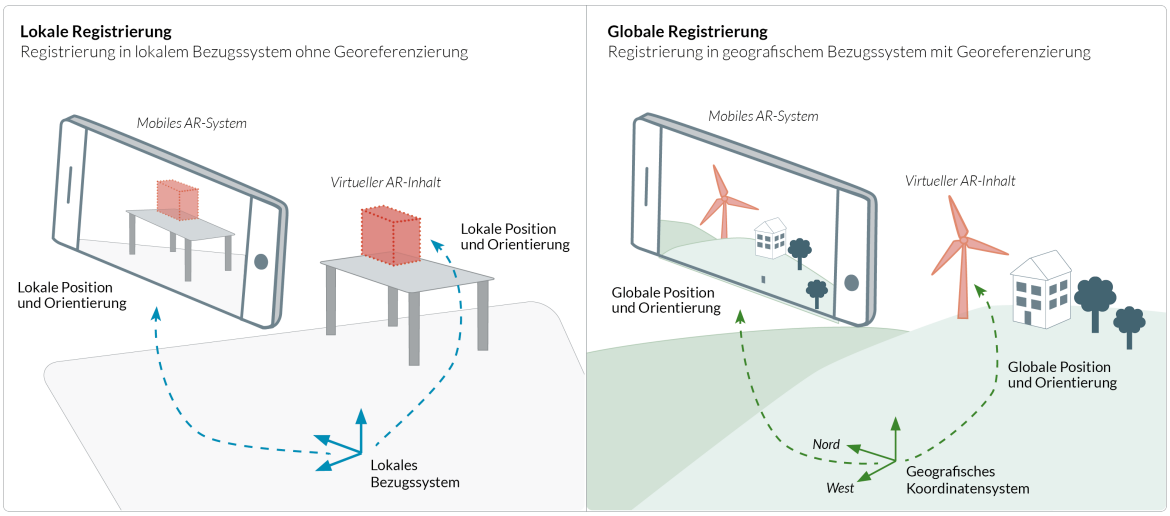


Abbildung 2.2 Bei der lokalen Registrierung werden das AR-System sowie die virtuellen AR-Inhalte in Bezug auf ein lokales Bezugssystem ohne Georeferenzierung definiert (links). Bei der globalen Registrierung sind virtuelle AR-Inhalte in Bezug auf ein globales geografisches Bezugssystem definiert (rechts). Eine Georeferenzierung der Geräte-Pose des mobilen AR-Systems ist dann notwendig.

2.2.3 Mobile AR-Sensoren

In mobilen AR-Systemen wird die Erfassung (*Sensing*) der Umgebung zur Bestimmung der Geräte-Pose innerhalb der Trackingkomponente in der Regel auf dem mobilen Endgerät selbst durchgeführt. Übliche mobile AR-taugliche Endgeräte (Smartphones, Tablets) sind dazu mit einer Reihe mobiler Sensoren ausgestattet (siehe Abbildung 2.3), welche in unterschiedlichem Maße für das Tracking geeignet sind. Die im Kontext des AR-Tracking bzw. der AR-Registrierung üblicherweise eingesetzten mobilen Sensoren sollen im Folgenden genauer vorgestellt werden.

Inertial Measurement Unit (IMU)

Eine inertiale Messeinheit (*Inertial Measurement Unit; IMU*) kombiniert mehrere Inertialsensoren, die in Kombination ihrer Messwerte ein geeignetes Maß für Rotationsbewegungen des Endgeräts und die Orientierung des Mobilgeräts im Raum liefern. Die IMU-Einheit besteht in der Regel aus einem Magnetometer, einem Accelerometer und einem Gyroskop.

Das **Magnetometer** misst die Richtung des Erdmagnetfeldes in allen drei Achsen. Damit kann die Orientierung des Geräts in Bezug auf den magnetischen Nordpol bestimmt werden (digitaler Kompass) [126, S.102]. In der Praxis sind diese Messwerte

jedoch in der Regel unzuverlässig und unpräzise und können um bis zu 30 Grad oder mehr abweichen [123].

Ein **Gyroskop** ist ein Gerät zur Messung der Winkelgeschwindigkeit mithilfe eines sich auf einer Drehachse frei drehenden Kreisels [126, S.102f]. Wird das umgebende System rotiert bzw. gekippt, behält der Kreisel seine Lage und die Geschwindigkeit der Drehbewegung kann gemessen werden. In mobilen AR-Systemen werden Gyroskope zur Messung der Drehgeschwindigkeit in allen drei Achsen als elektromechanisches Bauteil eingesetzt.

Ein **Accelerometer** ist ein Beschleunigungssensor. Der Sensor misst dabei die Trägheitskraft, die bei einer Bewegung des Sensors auf ein federgelagertes Testgewicht wirkt [126, S.103f]. In mobilen AR-Systemen werden ebenfalls Beschleunigungssensoren eingesetzt, welche die Beschleunigung in allen drei Achsen messen können. Auf diese Weise ist es nicht nur möglich, die Richtung der Gerätebeschleunigung im Raum zu messen, sobald das Gerät bewegt wird, sondern auch die Richtung der konstant wirkenden Erdanziehungskraft, wenn das Gerät nicht bewegt wird.

Bei der Bestimmung der Geräte-Pose mobiler AR-Systeme können die genannten IMU-Sensoren zur Erfassung von Rotationsbewegungen bzw. zur Bestimmung der globalen Orientierung des Systems eingesetzt werden. Eine geeignete Kombination der Sensordaten aller Inertialsensoren ist dabei empfehlenswert (*Sensor Fusion*) [51]. Hierfür ist eine aufwändige Kalibrierung der Sensoren sowie eine Verarbeitung und Kombination der Sensor-Rohdaten notwendig [77]. Inzwischen werden auch von den Betriebssystemen mobiler Endgeräte robuste softwareseitige Implementierungen bereitgestellt, welche durch Nutzung aller im Gerät verfügbaren Inertial-Sensoren geeignete kombinierte Rotationsvektoren zur Verfügung stellen, um Rotationsbewegungen des Endgeräts fortwährend zu beschreiben [4]. Die Problematik der fehlerhaften globalen Orientierungsbestimmung aufgrund unpräziser Magnetometer-Messwerte bleibt dabei jedoch bestehen.

GNSS-Empfänger

Das globale Navigationssatellitensystem (*Global Navigation Satellite System; GNSS*) ermöglicht eine globale Positionsbestimmung mithilfe von elektromagnetischen Signalen, die von Navigationssatelliten verschickt und von GNSS-Empfängern empfangen werden können. Neben dem US-amerikanischen Satellitensystem GPS (*Global Positioning System*) existieren mittlerweile auch weitere Systeme, die von GNSS-Empfängern in handelsüblichen mobilen Endgeräten empfangbar sind. Dazu zählen das chinesische System *Beidou*, das russische System *GLONASS* und das europäische System *Galileo*

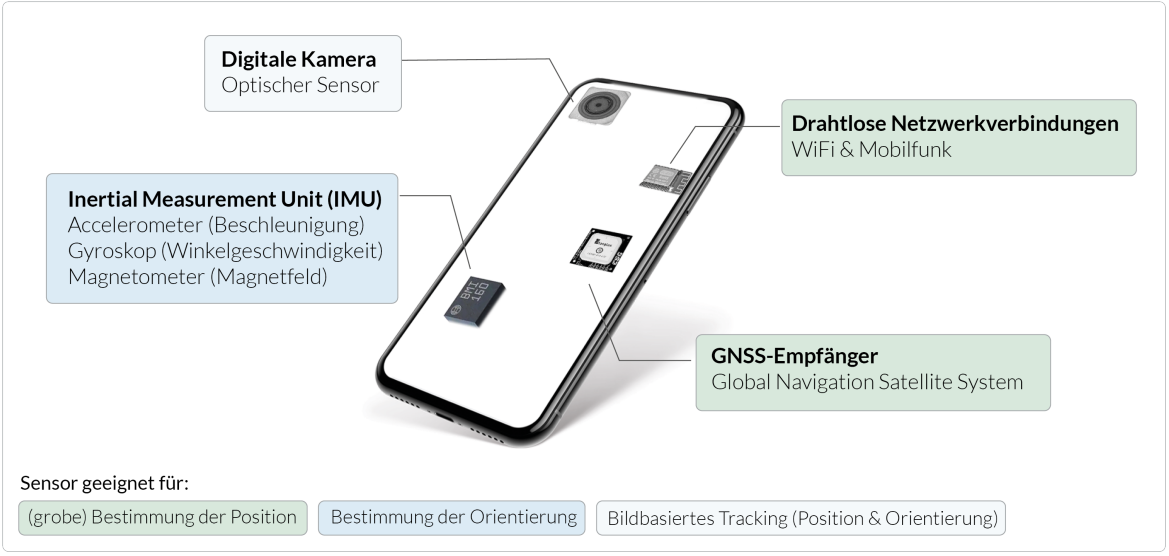


Abbildung 2.3 Mobile Sensoren in handelsüblichen Smartphones und Tablets, die für das Tracking in AR-Systemen geeignet sind: Eine inertiale Messeinheit (Inertial Measurement Unit) für Orientierungsmessungen, ein GNSS-Empfänger und drahtlose Netzwerkverbindungen für eine grobe globale Positionsbestimmung sowie ein optischer Kamera-Sensor für bildbasiertes Tracking.

[138]. Zur Positionsbestimmung werden dazu im GNSS-Empfänger die Laufzeiten der von den Satelliten ausgesendeten Signale gemessen. Mit Kenntnis der Satellitenstandorte kann damit die aktuelle globale Position des Empfängers auf der Erdoberfläche berechnet werden [126, S.99]. Theoretisch werden hierfür die Signale von mindestens vier Satelliten benötigt. Eine verlässliche Positionsbestimmung ist in der Regel zudem nur im Freien möglich.

Die Genauigkeit der Positionsbestimmung hängt dabei von mehreren Faktoren ab, beispielsweise den vorherrschenden Wetterverhältnissen, der Anzahl der aktuell sichtbaren Satelliten und der Signalqualität aufgrund der aktuellen Anwendungsumgebung (z.B. fehlerhafte Signale in urbaner Umgebung aufgrund von Reflexionen an Hausfassaden). In der Praxis sind Ungenauigkeiten bei der Positionsbestimmung von 5-10m in urbanem Gebiet durchaus üblich [113]. Die Bestimmung der geografischen Höhe ist besonders fehlerhaft und oftmals um ein Vielfaches ungenauer als die Bestimmung der geografischen Länge oder Breite [138].

Durch Korrektursignale kann die Genauigkeit der GNSS-basierten Positionsbestimmung deutlich verbessert werden, z.B. durch RTK-Technologie (*Real Time Kinematic, RTK-GNSS*). Hierfür werden jedoch zusätzlich große Empfangsantennen benötigt, sodass derartige System in üblichen mobilen Endgeräten als mobile AR-Systeme keine

Verwendung finden [126, S.100]. In der folgenden Arbeit sollen diese hochpräzisen GNSS-Systeme daher auch nicht weiter berücksichtigt werden.

Drahtlose Netzwerke

Eine mobile Positionsbestimmung kann auch durch Analyse der Signalstärken drahtloser Netzwerk-Infrastruktur erfolgen. Dazu können beispielsweise Signale aus Mobilfunknetzwerken oder aus WiFi- oder Bluetooth-Netzen genutzt werden [99]. Die grundsätzliche Idee dabei besteht darin, dass die beobachteten Signalstärken existierender Netzwerke an möglichst vielen Orten präzise vorab erfasst und gespeichert werden und auf Basis dieser Signalinformationen im Anschluss die Positionen mobiler Empfangsgeräte approximiert werden können (*Fingerprinting*) [126, S.101]. Auf diese Weise ist in urbanen Gegenden mit hoher WiFi-Signaldichte eine präzisere Positionsbestimmung möglich als durch rein GNSS-basierte Verfahren [120].

Die Betriebssysteme handelsüblicher mobiler Endgeräte stellen inzwischen GNSS-basierte Positionierungsmethoden zur Verfügung, die nicht nur reine GNSS-Signale verarbeiten, sondern diese – sofern verfügbar – auch mit Signalinformationen aus WiFi- und Mobilfunknetzen kombinieren [50]. Ist im Folgenden der Arbeit also von GNSS-basierten Positionierungsverfahren die Rede, so soll dies implizit auch die Möglichkeit einschließen, dass die reine GNSS-Positionsbestimmung möglicherweise durch Signale drahtloser Funknetzwerke optimiert wird.

Optische Kamera

Die bisher genannten mobilen AR-Sensoren ermöglichen eine grobe Positionsbestimmung (GNSS-Empfänger; Drahtlose Netwerke) bzw. eine – mehr oder weniger genaue – Bestimmung der Geräteorientierung im 3D-Raum (IMU-Sensoren). Die Genauigkeit dieser Sensoren reicht jedoch nicht für hochwertige AR-Szenarien aus, die robuste und positionsgetreue Tracking-Mechanismen benötigen. Für diese Zwecke sind bildbasierte Tracking-Verfahren mithilfe von optischen Kamera-Sensoren unabdingbar. Derartige digitale Bildsensoren können einfallende Lichtintensitäten mit zugehörigen Farbinformationen aufnehmen und als zweidimensionale pixelbasierte Kamerabilder digitalisieren [126, S.105]. Diese Bilder können anschließend mithilfe von komplexen Verfahren der *Computer Vision* analysiert werden, z.B. um Kamerabewegungen präzise zu erfassen.

Es existieren verschiedene Ansätze, um mittels bildbasierten Verfahren die Pose der Kamera zu bestimmen (siehe Abbildung 2.4): Im Falle des **Model-based Tracking** ist ein bekanntes Referenzmodell oder Referenzbild in der Umgebung verfügbar, z.B. ein Bildmarker oder ein vorab gescanntes und registriertes 3D-Modell eines Objekts [126,

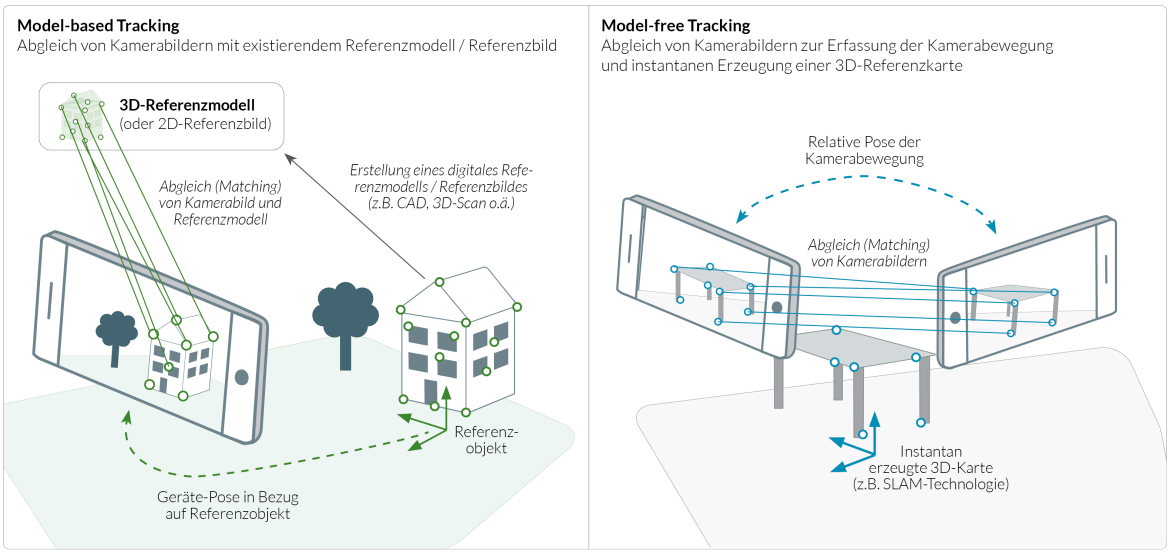


Abbildung 2.4 Bestimmung der Kamera-Pose mittels bildbasierter Verfahren: Falls ein Referenzbild oder Referenzmodell der Umgebung existiert, können Kamerabilder mit diesem Modell abgeglichen werden, um die Pose der Kamera in Bezug auf das Referenzobjekt zu bestimmen (Model-based Tracking). Falls keine Referenz vorhanden ist, können aufeinander folgende Kamerabilder per Feature Tracking abgeglichen werden, um die Bewegung der Kamera zu erfassen und instantan eine 3D-Referenzkarte zu erzeugen (Model-free Tracking).

S.106ff]. Die Aufgabe des bildbasierten Trackingverfahrens besteht in diesem Fall primär darin, aktuelle Kamerabilder mit dem vorab bekannten Referenzmodell bzw. -bild abzugleichen, um auf diese Weise die Position und Ausrichtung der Kamera in Bezug auf dieses Referenzobjekt zu bestimmen (siehe auch Kapitel 3.1.1). Im Falle des **Model-free Tracking** ist zu Beginn des Trackingvorgangs keine Referenz vorhanden [126, S.106ff]. Das bildbasierte Trackingverfahren kann dann versuchen, ein temporäres Modell der Umgebung instantan selbst zu erzeugen (siehe z.B. SLAM-Technologie in Kapitel 3.1.2). Hierzu werden zunächst aufeinander folgende Kamerabilder miteinander abgeglichen, um die Bewegung der Kamera zwischen diesen Bildern zu bestimmen. Auf dieser Basis können anschließend Referenzpunkte im sichtbaren 3D-Raum grob geschätzt werden. Es muss dabei beachtet werden, dass derartige bildbasierte Verfahren ohne vorab verfügbare Referenzobjekte nur eine Registrierung im Nahbereich ermöglichen. Sind markante 3D-Punkte zu weit entfernt, erzeugen Kamerabewegungen nur geringfügige Änderungen der Projektion der Ansicht im Kamerabild. Unterschiede zwischen aufeinander folgenden Kamerabildern sind dann durch die Bildanalyseverfahren nicht mehr erkennbar.

Der Abgleich von 2D-Kamerabildern sowohl mit 3D-Referenzmodellen als auch mit anderen 2D-Bildern erfolgt auf Basis des **Feature Tracking** [126, S.138ff]: Zunächst werden in den Kamerabildern markante Bildpunkte (*Keypoints*, *Interest Points*) de-

tektiert (z.B. Kantenpunkte an Hell-/Dunkel-Übergängen) und zu jedem Punkt ein sog. Deskriptor (*Descriptor*) erzeugt. Dabei handelt es sich um eine Datenstruktur, die den Bildpunkt möglichst eindeutig beschreibt, beispielsweise indem Helligkeitsinformationen zu umliegenden Pixeln gespeichert werden. Diese Deskriptoren können schließlich mit Deskriptoren aus anderen Kamerabildern oder aus Referenzmodellen verglichen werden, um auf diese Weise korrespondierende Bildpunkte zu finden (*Descriptor Matching*). Durch Lösen von mathematischen Optimierungsproblemen kann schließlich auf Basis dieser korrespondierenden Punkte die Bewegung der Kamera zwischen zwei Kamerabildern oder zwischen einem Kamerabild und einem Referenzobjekt geschätzt werden. Idealerweise ist für diesen Prozess das Auffinden von mehreren Hundert Korrespondenzpunkten in jedem Kamerabild erforderlich.

Neben diesen üblichen bildbasierten Tracking-Verfahren unter Verwendung einer einzelnen optischen Kamera existieren auch weitere optische Tracking-Methoden, bei denen z.B. Stereo-Kameras oder Laser-Sensoren zum Einsatz kommen. Letztere bieten eine schnelle Methode zur Abstandsmessung und damit zur Erfassung von geometrischen Strukturen in der unmittelbaren Umgebung des AR-Systems. Zwar werden derartige Laser-Sensoren (*Lidar; Light Detection and Ranging*) inzwischen auch in ausgewählten handelsüblichen mobilen Endgeräten integriert, sie werden jedoch nicht primär zur Erfassung der Geräte-Pose genutzt [104]. Im Kontext der vorliegenden Arbeit sollen daher nur monokulare Verfahren auf Basis einer einzelnen gewöhnlichen optischen Kamera weiter betrachtet werden.

2.3 Geodatenbasierte AR

Während in einem Großteil mobiler AR-Szenarien virtuelle Inhalte ohne Bezug auf eine geografische Position visualisiert werden, sind zahlreiche AR-Szenarien vorstellbar, bei denen georeferenzierte Inhalte (Geodaten) in einer AR-Ansicht visualisiert werden sollen. Im folgenden Kapitel soll diese Technologie der geodatenbasierten Augmented Reality als Spezialfall der Augmented Reality genauer vorgestellt werden. Dazu sollen die besonderen Herausforderungen der GeoAR-Technologie beschrieben sowie existierende GeoAR-Anwendungsszenarien kurz präsentiert werden.

2.3.1 Begriffsdefinition und Abgrenzung

Sind AR-Szenarien primär für die Nutzung im Freien (*Outdoor*) gedacht, so kann diese Technologie als **Outdoor Augmented Reality** bezeichnet werden [10]. Derartige

AR-Entwicklungen im Outdoor-Bereich implizieren heutzutage üblicherweise auch eine Nutzung mobiler AR-Systeme [73, 126]. Der Begriff *Outdoor Augmented Reality* soll daher im Kontext dieser Arbeit auch stets mit der Bezeichnung *Mobile Outdoor Augmented Reality* gleichgesetzt werden.

Eine klare Abgrenzung von *Outdoor* AR-Szenarien zu *Indoor* AR-Anwendungen oder AR-Szenarien für den *Nahbereich* ist dabei nicht immer möglich. Viele AR-Anwendungen können freilich sowohl innerhalb von Gebäuden als auch im Freien genutzt werden. Beispielsweise könnte eine AR-App zur Visualisierung virtueller Möbelstücke gleichermaßen zur Platzierung eines virtuellen Regals im Wohnzimmer und zur Platzierung eines virtuellen Liegestuhls im Garten genutzt werden [111]. Zudem sind viele gängige mobile AR-Anwendungen nur für den Einsatz im Nahbereich konzipiert, während Outdoor AR-Szenarien oftmals großflächigere Areale berücksichtigen müssen, was zu zusätzlichen Herausforderungen führt (siehe Kapitel 2.3.2) [126, S.164ff]. Natürlich existieren aber auch Outdoor AR-Szenarien, die im Rahmen der reinen AR-Visualisierung nur sehr kleinflächige Areale berücksichtigen müssen, beispielsweise das AR-Game *Pokemon Go*, bei dem virtuelle AR-Inhalte zwar im Freien und an Orten weltweit darstellbar sind, die AR-Darstellung selbst jedoch nur in unmittelbarer Umgebung direkt vor den Anwender:innen erfolgt [110].

Ein besonderer Fall der Outdoor Augmented Reality liegt dann vor, wenn georeferenzierte Daten AR-basiert dargestellt werden sollen und eine globale Registrierung des AR-Systems in Bezug auf ein geografisches Referenzsystem notwendig ist (siehe Kapitel 2.2.2). Ein solcher Fall wird in dieser Arbeit als **geodatenbasierte Augmented Reality** (*GeoAR, Geospatial Augmented Reality)* bezeichnet.

2.3.2 Herausforderungen

Outdoor AR-Anwendungen im Allgemeinen sowie GeoAR-Anwendungen im Speziellen müssen – im Vergleich zu Anwendungen, die für den Einsatz in Innenräumen im Nahbereich konzipiert wurden – oftmals zusätzliche Rahmenbedingungen berücksichtigen und besondere Herausforderungen bewältigen, was die technische Umsetzung derartiger Anwendungen nochmals erschwert [63, 126, 117, 8]. Abbildung 2.5 fasst Herausforderungen zusammen, die üblicherweise bei Outdoor AR- bzw. GeoAR-Szenarien auftreten:

- **Globale Registrierung**: Für GeoAR-Anwendungen muss das AR-System global in Bezug auf ein geografisches Koordinatensystem registriert werden. Die in mobilen Endgeräten integrierten Sensoren sind dabei nicht ausreichend, um eine präzise Pose-Bestimmung zu realisieren (siehe Kapitel 2.2). Existierende Verfahren

für eine präzise globale Registrierung sind daher fehlerintensiv bzw. nur sehr aufwändig umzusetzen (siehe Kapitel 3.2). Diese technische Herausforderung gilt nur für GeoAR-Anwendung. Für übrige AR-Anwendungen ist eine georeferenzierte Registrierung nicht notwendig.

- **Anwendungsgebiet (Ausdehnung)**: Die meisten AR-Szenarien müssen in der Regel nur in einem kleinräumigen Gebiet in der unmittelbaren Umgebung funktionieren (z.B. innerhalb eines Raums). Outdoor AR-Szenarien sind hingegen oftmals so konzipiert, dass sie in einer großflächigen Umgebung AR-Visualisierungen ermöglichen sollen, z.B. innerhalb der gesamten Stadt (*large-scale AR*). Für modellbasierte Tracking- und Registrierungsverfahren ist dann ein umfangreiches großflächiges Referenzmodell notwendig, welches in der Regel nur mit großem Zeit- und Kostenaufwand und umfangreicher IT-Infrastruktur bereitgestellt werden kann (siehe Kapitel 3.2.2).

- **Umgebungscharakteristik**: Umgebungen im Freien beinhalten oftmals Störfaktoren, die eine robuste AR-Visualisierung erschweren. Dazu zählen beispielsweise Flächen, deren Texturen für bildbasierte Trackingverfahren problematisch sind, z.B. stark homogene oder monotone Flächen (z.B. Gras, Straßen, Fenster, Zaun). Zudem können äußere Umwelteinflüsse die Trackingverfahren beeinträchtigen, z.B. sich bewegende Objekte (z.B. Fahrzeuge, Personen) oder saisonal bedingte Veränderungen der Vegetation (z.B. Baum mit und ohne Blätter). Existierende Referenzmodelle können durch derartige zeitliche Veränderungen unbrauchbar

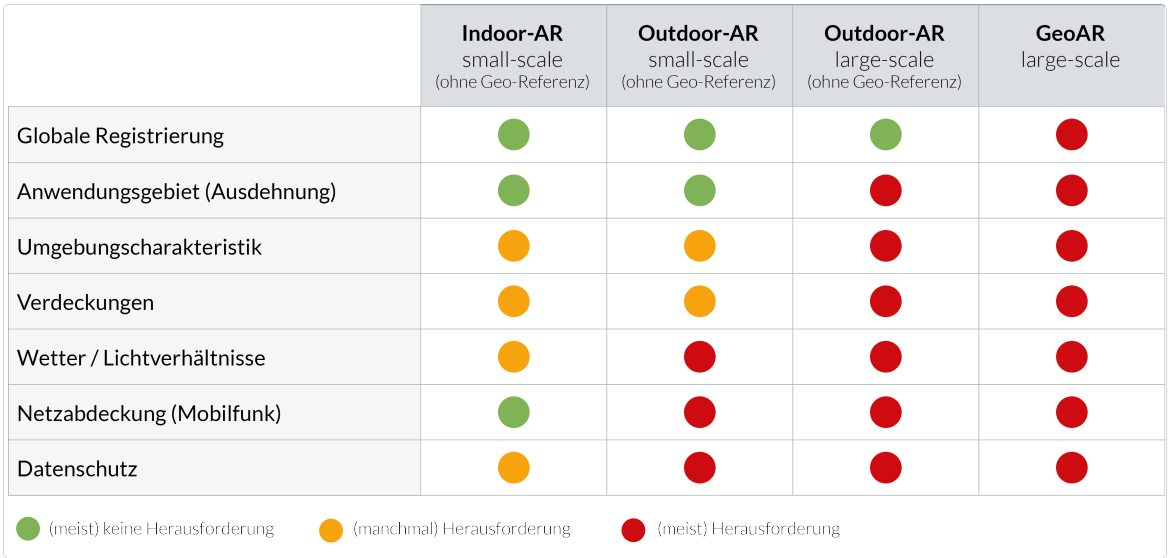


Abbildung 2.5 Herausforderungen von GeoAR-Szenarien (eigene Bewertung)

werden. Indoor-AR-Anwendungen müssen solche Einflüsse und Störfaktoren in der Regel nicht (Vegetationsveränderungen) oder nur sehr eingeschränkt (bewegende Objekte, monotone Flächen) bewältigen.

- **Verdeckungen**: Realitätsnahe AR-Visualisierungen sollten in der Lage sein, Verdeckungen der virtuellen AR-Inhalte durch reale Objekte (Personen, Gegenstände, Gelände, Vegetation etc.) zu berücksichtigen. Bei AR-Visualisierungen, die im Nahbereich stattfinden, ist die Berücksichtigung von Verdeckungen möglich, indem mithilfe der beim Tracking erstellten 3D-Punkt-Karte die Umgebung detailgetreu rekonstruiert wird. Derartige Mechanismen sind in neuartigen AR-Tracking-Methoden im Nahbereich inzwischen üblich [141]. Um übliche Verdeckungen in großflächigen Szenarien (z.B. durch Häuser oder Gelände) bei Outdoor AR-Szenarien zu berücksichtigen, ist die genaue Kenntnis der gesamten 3D-Landschaft notwendig, beispielsweise durch digitale 3D-Oberflächenmodelle der Umgebung. Die Erzeugung und Bereitstellung derartiger Modelle ist aufwändig (siehe auch Kapitel 3.2.2).

- **Wetter und Lichtverhältnisse**: Externe Wettereinflüsse können bei Outdoor AR-Anwendungen nicht nur die Qualität der bildbasierten Trackingverfahren beeinflussen, wenn sich durch Niederschlag oder Wind das Aussehen der Umgebung verändert, sondern auch unmittelbar die Bedienbarkeit der mobilen Anwendung behindern, z.B. wenn bei Regen oder bei kalten Außentemperaturen die Bedienung der Anwendungsoberfläche per Touch-Gesten erschwert wird. Starke Sonneneinstrahlung kann zudem die Sichtbarkeit des Bildschirms beeinträchtigen. Auch die Beleuchtung der virtuellen AR-Szene muss für realitätsnahe Darstellungen optimal an wechselnde Lichtverhältnisse im Freien angepasst werden. Indoor-AR-Anwendungen sollten zwar auch unterschiedliche Lichtverhältnisse berücksichtigen können, werden aber in der Regel durch externe Wettereinflüsse weniger stark beeinflusst.

- **Netzabdeckung**: AR-Anwendungen setzen oft einen verfügbaren Onlinezugang voraus, beispielsweise um virtuelle Inhalte oder Referenzmodelle für die AR-Registrierung vor Ort per Online-Service nachladen zu können. Bei Outdoor AR-Anwendungen kann das instantane Nachladen von digitalen Zusatzdaten oftmals an der fehlenden Verfügbarkeit schneller Mobildatennetze scheitern, insbesondere in ländlichen Gegenden. Bei Indoor-AR-Anwendung ist die Verfügbarkeit einer schnellen Internetverbindung in der Regel weniger problematisch.

- **Datenschutz (Persönlichkeitsrechte)**: Werden AR-Anwendungen mit handelsüblichen mobilen Endgeräten genutzt, muss die Kamera der Geräte zur Darstellung der virtuellen Inhalte auf die Umgebung gerichtet werden. Wenn sich der Einsatz im öffentlichen Raum abspielt und unbeteiligte Personen per Kamera erfasst werden, kann dies sowohl für die Nutzer:innen der AR-Anwendung als auch für die zufällig im Kamerabild auftauchenden Personen als unangenehm bzw. störend wahrgenommen werden [119]. Dies wird dadurch verstärkt, dass die unbeteiligt per Kamera erfassten Personen nicht wissen, zu welchem Zweck die Kameraaufnahmen erfolgen bzw. inwieweit die Aufnahmen dauerhaft gespeichert werden. Indoor-AR-Anwendungen spielen sich hingegen meist eher im halböffentlichen bzw. privaten Raum ab.

Es wird deutlich, dass Outdoor AR-Anwendungen – und insbesondere GeoAR-Szenarien – technisch schwieriger umzusetzen sind als gewöhnliche AR-Szenarien, die nur in Innenräumen in der unmittelbaren Umgebung stattfinden. Dennoch sollte es ein wichtiges Ziel zur weiteren Etablierung der mobilen AR-Technologie sein, geeignete AR-Szenarien auf einfache Art und Weise *überall*, sowohl in Gebäuden als auch im Freien, nutzbar zu machen. Es ist also sehr erstrebenswert, im Kontext von künftigen AR-Entwicklungen insbesondere die Etablierung und Erforschung verbesserter Outdoor AR- und GeoAR-Technologien in den Vordergrund zu rücken [126, S.415f]. In diesem Kontext wird auch die Notwendigkeit des Forschungsziels dieser Arbeit nochmals verdeutlicht, nämlich die Entwicklung einer flexibel einsetzbaren und leicht bedienbaren Registrierungsmethode für GeoAR-Szenarien.

2.4 Geodatenbasierte AR-Anwendungen

Aufgrund der genannten Herausforderungen sind bislang nur wenige geodatenbasierten AR-Anwendungen auf dem Markt verfügbar und praxistauglich einsetzbar. Viele dieser existierenden GeoAR-Anwendungen können entweder keine qualitativ hochwertigen AR-Visualisierungen anbieten (z.B. [148, 24]) oder sind nur in eingeschränktem Umfang oder in Rahmen von Forschungsarbeiten nutzbar (z.B. [1, 153]. Im folgenden Kapitel soll eine Auswahl an existierenden GeoAR-Anwendungen kurz präsentiert werden. Einige der genannten Anwendungen sind in Abbildung 2.6 dargestellt.

Im Zuge der Verfügbarkeit mobiler Endgeräte für eine breite Nutzerschicht und der Entstehung von ersten mobilen AR-Anwendungen sind einige **AR-Plattformen** (sog. *AR-Browser*) entstanden, die auf Basis einfacher GNSS- und IMU-gestützter Registrierung (siehe Kapitel 3.2.1) punktförmige Geodaten im Kamerabild einblenden

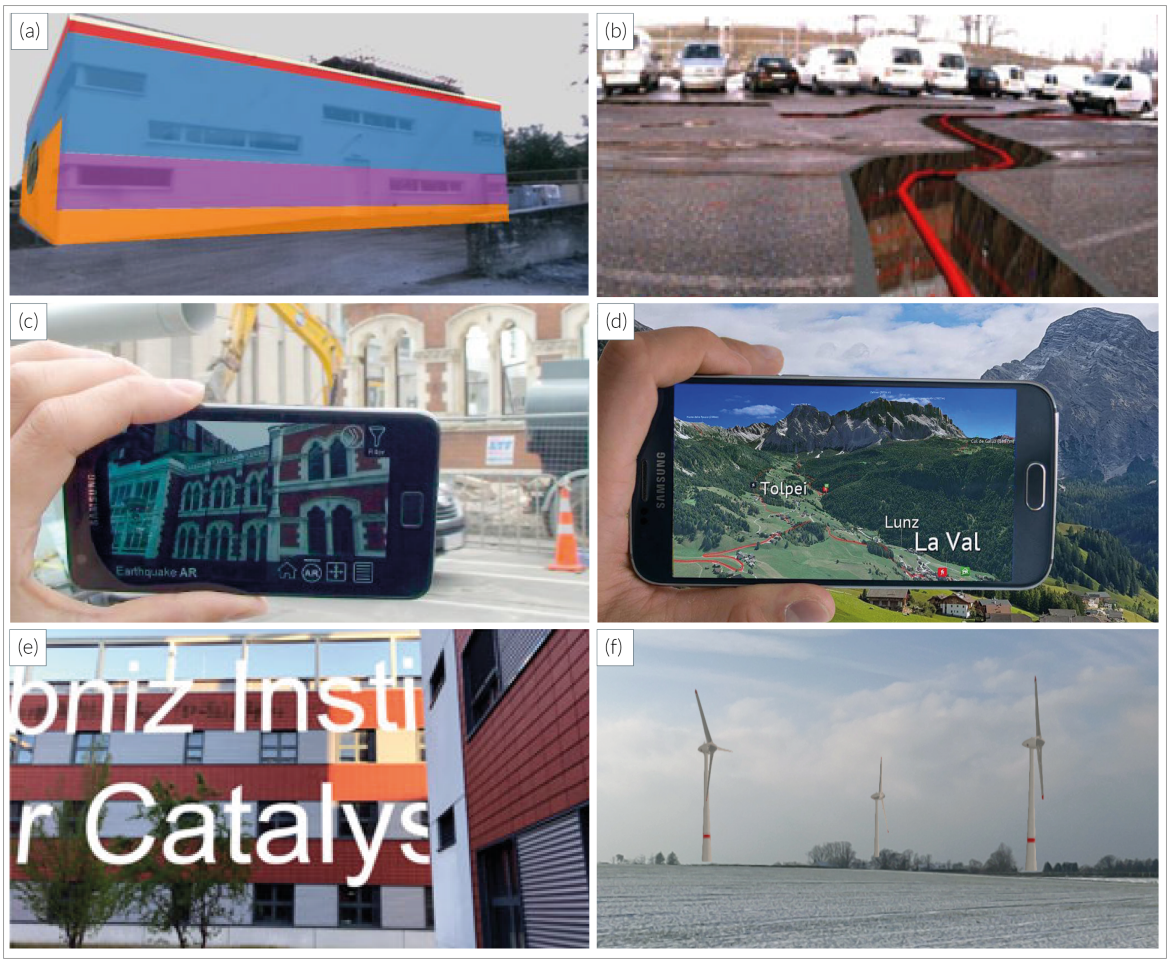


Abbildung 2.6 (a) AR-basierte Visualisierung von GIS-Daten zur Bauüberwachung [153] sowie (b) zur Darstellung von unterirdischen Leitungen [124]. (c) AR-Visualisierung von durch Erdbeben zerstörten Gebäuden [82]. (d) AR-basierte Einbettung von großflächigen Geländemodellen zur Wanderroutenplanung [1]. (e) Integration von 3D-Stadtmodellen zur Berücksichtigung von Verdeckungen bei Outdoor AR-Anwendungen [116]. (f) AR-Visualisierung von geplanten Windenergieanlagen im Landschaftsbild [26].

konnten. Die dargestellten Geodaten konnten dabei beliebige Anwendungsbereiche abdecken (z.B. nahegelegene Sehenswürdigkeiten, Restaurants oder Berggipfel). Zu diesen AR-Browsern zählen beispielsweise *Wikitude* [148] und *Layar* [24].

Daneben entstanden einige **prototypische Fach-Anwendungen** zur Visualisierung von Geodaten, die der Unterstützung der Feld- und Geländearbeit von Fachexperten dienen, beispielsweise Anwendungen zur AR-Darstellung von unterirdischen Leitungen [124], zur Darstellung von Sensordaten im Umweltbereich (z.B. dem zeitlichen Verlauf von Schneehöhen) [145], zur Aufnahme von Bohrlöchern als Markierungspunkte [84], zur Unterstützung von planerischen Aufgaben im Rahmen der Flurneuordnung [78]

oder zur Dokumentation und Bauüberwachung auf Baustellen [153]. Andere Outdoor AR-Anwendungen versuchen, historische Gebäude an realer geografischer Position mittels AR-Darstellung wiederaufleben zu lassen, beispielsweise eine Anwendung von Lee et al. zur AR-Darstellung von Gebäuden, die durch Erdbeben zerstört wurden [82].

Die AR-bezogene **Einbettung von Geodatenmodellen** – wie in dieser Arbeit zum Zwecke der Registrierung praktiziert – ist ebenfalls bei einigen existierenden Anwendungen möglich. Eine derartige Einbettung von großflächigen digitalen Gelände-, Oberflächen-, oder Stadtmodellen erfolgt dabei jedoch allein zum Zweck der AR-Visualisierung, ohne die Geodatenmodelle auch zur Kalibrierung des AR-Systems zu nutzen. In diesem Kontext sind beispielsweise die Anwendung 3D RealityMaps [1] zu nennen, die eine AR-Darstellung von texturierten 3D-Geländekarten zur Routenplanung ermöglicht, oder die Arbeiten von Postert et al. und Kasperi et al., um mittels AR-basierter Integration von 3D-Stadtmodellen [116] oder 3D-Gebäudemodellen [66] Verdeckungen in Outdoor AR-Szenarien realistisch zu berücksichtigen.

Weitere GeoAR-Anwendungsszenarien nutzen AR-Visualisierungen zur realitätsnahen Vor-Ort-Darstellung der **visuellen Auswirkungen geplanter Bauwerke** auf das Landschaftsbild und die Landschaftsästhetik. Derartige Anwendungen können beispielsweise im Rahmen von Bürgerbeteiligungsverfahren eingesetzt werden, um die Transparenz innerhalb von Bauplanungsprozessen zu erhöhen und die Akzeptanz der Bauprojekte bei Anwohnern zu steigern. In diesem Kontext existieren beispielsweise Demonstrator-Anwendungen zur AR-basierten Darstellung von Windenergieanlagen im Landschaftsbild [26, 68].

2.5 Mathematische Grundlagen der Projektion von 3D-Inhalten in den 2D-Raum

Im Gegensatz zu klassischen 2D-Abbildungen (z.B. übliche Video- oder Fotoaufnahmen) bietet die AR-Technologie den Vorteil, dass virtuelle 3D-Inhalte durch einen hohen Grad an Immersion wahrgenommen werden können. Durch Bewegungen des mobilen AR-Gerätes im Raum können Nutzer:innen virtuelle 3D-Objekte aus verschiedenen Positionen und Blickwinkeln betrachten und damit auch dreidimensional erleben und wahrnehmen. Die AR-Visualisierungen selbst finden jedoch trotzdem auf einem 2D-Bildschirm des Smartphones oder Tablets statt. Der immersive Effekt entsteht nur durch die Bewegung des AR-Systems in Verbindung mit einer realitätsgetreuen und perspektivisch korrekten **Projektion der 3D-Inhalte in den 2D-Raum**, also den Bildschirm des Endgeräts.

Bei dieser perspektivischen Projektion findet also ein Abbildungsvorgang von 3D-Punkten aus einem metrischen 3D-Raum (Weltkoordinatensystem) auf 2D-Punkte des pixelbasierten Kamerabildes statt. Die mathematische Definition dieser AR-Projektion soll im Folgenden kurz beschrieben werden. Dadurch soll ein besseres Verständnis vermittelt werden für die später mathematisch beschriebenen Zusammenhänge des entwickelten GeoAR-Kalibrierungsverfahrens (siehe Kapitel 4.4.1).

Pinhole-Kameramodell

Die perspektivische 3D-2D-Projektion erfolgt in der *Computer Vision* auf Basis eines idealisierten Lochkameramodells (*Pinhole Camera Model*) zur Beschreibung der Bilderzeugung einer realen physischen Digitalkamera [137, S.48ff] [126, S.124ff]. Das Prinzip dieser Projektion ist in Abbildung 2.7 skizziert.

In diesem Modell wird ein 3D-Punkt $P = [X_W, Y_W, Z_W]$, der im Weltkoordinatensystem O_W definiert ist, auf einen 2D-Punkt $q = [q_u, q_v]$ im Pixelkoordinatensystem O_{UV} projiziert. Der Punkt q befindet sich dabei auf einer Bildebene π. Alle Projektionslinien treffen sich in einem gemeinsamen Punkt, dem optischen Zentrum der Kamera (*Center of Projection*). Dieser Punkt stellt den Ursprung des Kamerakoordinatensystems O_{Cam} dar. Die Projektion von O_{Cam} auf die Bildebene π wird als *Principle Point* $c = [c_u, c_v]$ bezeichnet. Eine gedachte Linie vom Projektionszentrum O_{Cam} durch den Punkt c wird *optische Achse* genannt, wobei der Abstand zwischen O_{Cam} und π als Brennweite f definiert ist.

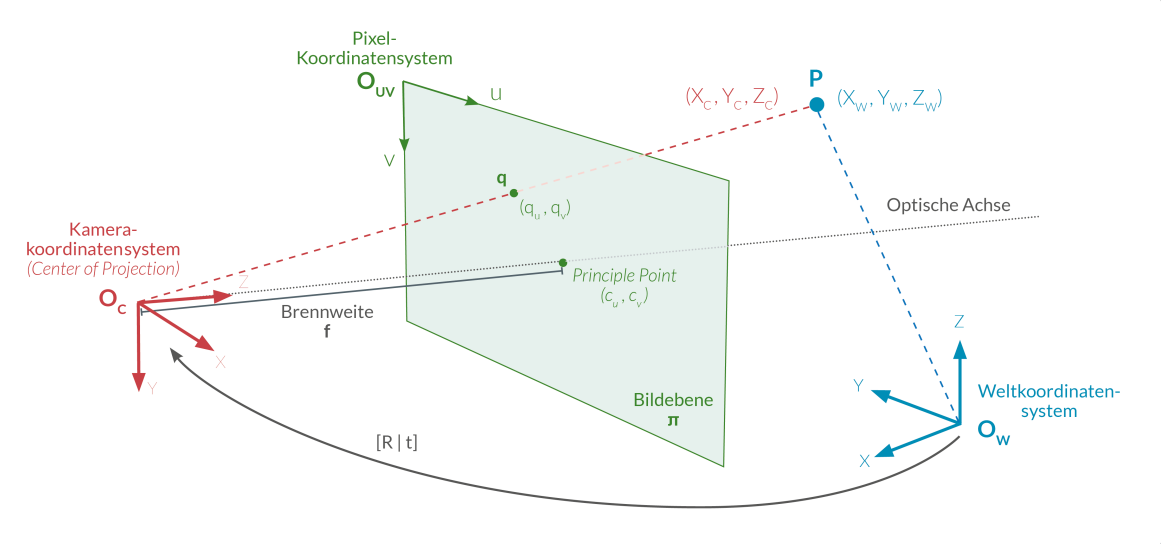


Abbildung 2.7 Perspektivische Projektion von 3D-Punkten auf 2D-Bildpunkte auf Basis des Pinhole-Kameramodells.

Extrinsische Kameraparameter

Das Kamerakoordinatensystem O_{Cam} und das Weltkoordinatensystem O_W sind durch eine Rotationsmatrix R und einen Translationsvektor t miteinander verknüpft. Eine Rotation im 3D-Raum kann durch drei Parameter – beispielsweise durch drei Eulerwinkel *Yaw*, *Pitch* und *Roll* – erfasst werden. Die 3x3-Matrix R besitzt dementsprechend drei Freiheitsgrade. Gemeinsam mit dem Translationsvektor t (3-DOF) kann also die Beziehung der beiden Koordinatensystem mit sechs Freiheitsgraden beschrieben werden. Der 3D-Punkt P mit den Koordinaten $P_W = [X_W, Y_W, Z_W]$ im Weltkoordinatensystem kann dementsprechend unter Beachtung der Rotation und Translation als $P_C = [X_C, Y_C, Z_C]$ im Kamerakoordinatensystem wie folgt ausgedrückt werden:

$$\begin{pmatrix} X_C \\ Y_C \\ Z_C \end{pmatrix} = R \begin{pmatrix} X_W \\ Y_W \\ Z_W \end{pmatrix} + t = [R|t] \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix} \tag{2.1}$$

Die Rotation R und Translation t werden dabei oftmals als **extrinsische Parameter** und die 3x4-Matrix $[R|t]$ als extrinsische Matrix bezeichnet [137, S.51].

Intrinsische Kameraparameter

Die perspektivische Projektion des 3D-Punktes $P_C = [X_C, Y_C, Z_C]$ auf einen 2D-Punkt q im Pixelkoordinatensystem O_{UV} kann schließlich mit **intrinsischen Kameraparametern** und einer intrinsischen Matrix K mit fünf Parametern beschrieben werden:

$$K = \begin{pmatrix} f_x & \gamma & c_u \\ 0 & f_y & c_v \\ 0 & 0 & 1 \end{pmatrix} \tag{2.2}$$

Die Parameter $f_x = s_x f$ und $f_y = s_y f$ definieren die Brennweite, skaliert mit der Größe der einzelnen Pixel s_x und s_y entlang der horizontalen und vertikalen Achsen des Pixelkoordinatensystems. In den meisten Fällen kann jedoch davon ausgegangen werden, dass das durch die digitale Kamera bereitgestellte Bild bereits quadratische Pixel aufweist, sodass $f_x = f_y$ angenommen werden kann [126, S.126].

Im definierten Kameramodell stimmt der *Principle Point c* nicht mit dem Ursprung des Pixelkoordinatensystems O_{UV} überein. Die Verschiebung von c in Bezug auf O_{UV} wird daher über die Parameter c_u und c_y beschrieben. Der Parameter γ beschreibt schließlich einen *skew factor*, falls die u- und v-Achse des Pixelkoordinatensytems nicht

senkrecht zueinander stehen. Dies ist bei üblichen digitalen Kameras jedoch in der Regel nicht der Fall, sodass $\gamma = 0$ angenommen werden kann.

Um die Pixel-Koordinaten q_u, q_v der Projektion des Punktes P_C auf die Bildebene zu erhalten, wird P_C also zunächst mit der intrinsischen Kameramatrix K multipliziert und der resultierende Punkt schließlich durch die dritte Komponente Z_C dividiert. Dadurch entsteht der Punkt $q = [q_c, q_v, 1]$ in homogenen Bildkoordinaten:

$$\begin{pmatrix} q'_u \\ q'_v \\ Z_C \end{pmatrix} = K \begin{pmatrix} X_C \\ Y_C \\ Z_C \end{pmatrix} \tag{2.3}$$

$$\begin{pmatrix} q_u \\ q_v \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{q'_u}{Z_C} \\ \frac{q'_v}{Z_C} \\ 1 \end{pmatrix} \tag{2.4}$$

Durch die Division verliert der Punkt seine Tiefeninformation. Lediglich die Richtung des Projektionsstrahls kann aus dem 3D-Bildpunkt q mithilfe der inversen Matrix K^{-1} rekonstruiert werden. Eine Rekonstruktion der Entfernung des ursprünglichen 3D-Punktes P aus dem Bildpunkt q entlang dieses Projektionsstrahls ist nicht mehr eindeutig möglich.

Die Bestimmung der intrinsischen Kameraparameter wird in der Regel nur einmalig vorgenommen (Kamera-Kalibrierung). Anschließend kann die Matrix K als fix betrachtet werden, zumindest so lange keine Änderungen an der Brennweite (z.B. durch Zoomen) vorgenommen werden. Im Rahmen der Entwicklung von Anwendungen auf mobilen Endgeräten mit integrierter Kamera können die intrinsischen Parameter meist durch die Betriebssysteme der Mobilgeräte bereitgestellt werden [3].

Kameramatrix für AR-Projektionen

Die gesamte perspektivische Projektion kann schließlich über eine 3x4-Matrix M, der sogenannten **Kameramatrix**, ausgedrückt werden. Diese Matrix integriert sowohl die extrinsischen Parameter $[R|t]$ als auch die intrinsischen Parameter K:

$$M = K[R|t] \tag{2.5}$$

Die Anwendung der Kameramatrix M auf einen Punkt P_W im Weltkoordinatensystem projiziert diesen also auf einen eindeutigen Bildpunkt q auf der Bildebene:

$$\begin{pmatrix} q_u \\ q_v \\ 1 \end{pmatrix} \sim M \begin{pmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{pmatrix} \tag{2.6}$$

Der Punkt q korrespondiert hingegen mit unendlich vielen anderen 3D-Punkten entlang des Projektionsstrahls. Die Kameramatrix definiert also – ausgehend von einem bekannten 2D-Bildpunkt – einen 3D-Punkt P_W nicht eindeutig, sondern nur mit einem Skalierungsfaktor (*"up to scale"*).

Zusammenfassend lässt sich festhalten, dass mithilfe der Kameramatrix eine perspektivische Projektion des 3D-Punktes P_W aus dem Weltkoordinatensystem auf den 2D-Bildpunkt q im Pixelkoordinatensystem auf Basis der folgenden beiden Schritte realisiert wird:

1. Der 3D-Punkt P_W wird mithilfe der **extrinsischen Parameter** (Rotation und Translation) vom Weltkoordinatensystem in das Kamerakoordinatensystem überführt.

2. Der neue Punkt P_C wird aus dem Kamerakoordinatensystem mithilfe von **intrinsischen Parametern** (Brennweite, Principle Point) auf den 2D-Bildpunkt q perspektivisch projiziert.

Mithilfe dieser Kameramatrix können schließlich AR-Visualisierungen realisiert werden. Das heißt, virtuelle Modelle, die aus 3D-Punkten mit bekannten Koordinaten bestehen (z.B. Vertices eines 3D-Modells), werden fortlaufend mithilfe der Kameramatrix auf Basis der aktuellen Kameraparameter auf das Kamerabild projiziert. Bei Anwendung der Projektionsgleichung innerhalb mobiler AR-Anwendungen besteht die Herausforderung nun darin, die Rotation und Translation des Kamerakoordinatensystems, die sich ständig durch Kamerabewegungen ändert, zu bestimmen (Tracking; siehe Kapitel 2.2.2). Dabei ist zu unterscheiden, ob die zu projizierenden 3D-Punkte in einem lokalen oder globalen Weltkoordinatensystem definiert sind und demzufolge eine Bestimmung von lokaler oder globaler Rotation und Translation des Kamerakoordinatensystems – also des AR-Systems – notwendig ist (siehe Kapitel 2.2.2). Aktuelle Lösungsansätze für eine lokale und globale AR-Registrierung werden ausführlich in Kapitel 3 vorgestellt.

2.6 Zusammenfassung

Bei der mobilen Augmented Reality werden virtuelle Inhalte in Echtzeit im Kamerabild eines mobilen Endgeräts eingeblendet. Die unmittelbare reale Umgebung wird somit erweitert. Zur Realisierung derartiger AR-Visualisierungen ist eine Registrierung der virtuellen Inhalte in der realen Welt notwendig. Dies erfordert ein fortlaufendes Erfassen (Tracking) der aktuellen Kamera-Pose, d.h. der Position und Ausrichtung der Kamera in Bezug auf ein Referenzsystem in der realen Welt.

Das AR-Tracking kann mithilfe unterschiedlicher mobiler Sensoren erfolgen. Die Orientierung des AR-Systems kann mit einer inertialen Messeinheit (Inertial Measurement Unit) erfasst werden, GNSS-Empfänger und drahtlose Netzwerkverbindungen können für eine grobe globale Positionsbestimmung eingesetzt werden, und ein optischer Kamera-Sensor ist für bildbasiertes Tracking geeignet. Insbesondere mittels bildbasierter Trackingverfahren – idealerweise in Kombination mit IMU-Sensoren (Accelerometer, Gyroskop, Magnetometer) – ist eine sehr genaue Registrierung virtueller Inhalte in Bezug auf ein lokales Bezugssystem ohne Georeferenzierung möglich, entweder mit Unterstützung vorhandener Referenzobjekte (Model-based Tracking) oder ohne vorhandenes Referenzobjekt per Analyse fortlaufender Kamerabilder mittels Feature Tracking (Model-free Tracking). Eine globale Registrierung in Bezug auf ein geografisches Bezugssystem mit Georeferenzierung ist jedoch schwierig und anspruchsvoll, insbesondere aufgrund der Ungenauigkeit globaler Orientierungs- und Positionierungssensoren.

Sind AR-Szenarien primär für die Nutzung im Freien gedacht, kann dies als Outdoor Augmented Reality bezeichnet werden. Geodatenbasierte Augmented Reality (GeoAR) liegt dann vor, wenn georeferenzierte Daten als AR-Visualisierung dargestellt werden. Im Vergleich zu AR-Szenarien, die für den Einsatz in Innenräumen im Nahbereich konzipiert wurden, müssen bei Outdoor AR-Anwendungen und insbesondere bei GeoAR-Anwendungen in der Regel zusätzliche Herausforderungen bewältigt werden, was die technische Umsetzung derartiger Anwendungen nochmals erschwert. Neben der genannten Problematik der globalen Registrierung bei GeoAR-Szenarien können beispielsweise auch wechselnde Wetter- und Lichtverhältnisse, bewegliche oder veränderliche Objekte in der Außenumgebung als Störfaktoren oder Datenschutzproblematiken im öffentlichen Raum für zusätzliche Schwierigkeiten sorgen. Aufgrund der großen Herausforderungen sind bislang nur wenige GeoAR-Anwendungen auf dem Markt verfügbar und in der Praxis einsetzbar. Die existierenden Anwendungen in diesem Bereich zeigen zwar großes Potenzial für vielfältige Anwendungsmöglichkeiten der AR-Technologie durch Integration von Geodaten, sind jedoch bislang meist nur in eingeschränktem Umfang oder minderwertiger Darstellungsqualität nutzbar.

Die AR-Technologie bietet den Vorteil, dass virtuelle 3D-Inhalte realitätsgetreu aus verschiedenen Positionen und Blickwinkeln und somit mit einem hohen Grad an Immersion wahrgenommen werden können. Das Grundlagenkapitel endet daher mit einem Blick auf die Erzeugung von mobilen AR-Visualisierungen aus der Perspektive der Computer Vision und der mathematischen Beschreibung (Kapitel 2.5) der perspektivischen Projektion von 3D-Punkten in den 2D-Raum. Dies bildet schließlich die Grundlage für das Verständnis der entwickelten Methode zur manuellen Registrierung von globaler Geräte-Position und -Ausrichtung.

Der immersive Effekt von AR-Visualisierungen entsteht durch die Bewegung des AR-Systems im Raum in Verbindung mit einer realitätsgetreuen und perspektivisch korrekten Projektion der virtuellen 3D-Inhalte in einen 2D-Raum, nämlich den Bildschirm des AR-Endgeräts. Bei dieser Projektion auf Basis des Lochkameramodells geht die Tiefeninformation der 3D-Punkte verloren. Die Projektion kann mathematisch mithilfe einer Kameramatrix beschrieben werden, die aus extrinsischen Parametern (Rotation, Translation) und intrinsischen Parametern (z.B. Brennweite) besteht. Während die intrinsischen Parameter als bekannt und unveränderbar betrachtet werden können, besteht die Aufgabe des AR-Trackings darin, die sich durch Kamerabewegungen ständig verändernden extrinsischen Parameter fortlaufend zu bestimmen.

Kapitel 3

Stand der Forschung und Entwicklung zur AR-Registrierung

Im vorherigen Kapitel wurde bereits verdeutlicht, dass eine genaue Registrierung des mobilen AR-Systems in seiner dreidimensionalen physischen Umgebung eine der wichtigsten Voraussetzungen für die Umsetzung einer realitätsnahen AR-Anwendung ist. Nur durch Kenntnis von Position und Orientierung der mobilen Kamera innerhalb des Referenzsystems ist eine Projektion der virtuellen AR-Inhalte an korrekter Position auf dem AR-Display möglich (siehe Kapitel 2.5).

In dieser Problemstellung der korrekten Registrierung besteht somit eine wesentliche technische Herausforderung bei der Entwicklung mobiler AR-Systeme, für deren Lösung in der Forschung in den vergangenen Jahren – je nach Anwendungsfall und Rahmenbedingungen – unterschiedliche Verfahren und Vorgehensweisen entstanden sind. Ein Überblick über mögliche AR-Registrierungsverfahren und deren grobe Funktionsweise wurde bereits in den vorherigen Kapiteln bei der allgemeinen Vorstellung der AR-Technologie und der dabei eingesetzten mobilen AR-Sensoren präsentiert (siehe Kapitel 2.2). Konkrete aktuelle Entwicklungen und Forschungsarbeiten zu praxistauglichen AR-Registrierungsverfahren sowie deren Stärken und Limitierungen sollen im folgenden Kapitel nun genauer erläutert werden. Die Registrierungsverfahren, die dabei betrachtet werden sollen, sind in Abbildung 3.1 zusammenfassend skizziert.

Für die Registrierung eines AR-Systems in einem lokalen Referenzsystem im Nahbereich existieren zahlreiche etablierte bildbasierte Methoden, die sowohl mit als auch ohne bekannte Referenzmarker funktionieren. Derartige lokale AR-Registrierungsmethoden sollen zunächst in Kapitel 3.1 beschrieben werden. Für geodatenbasierte AR-Szenarien werden jedoch globale Registrierungsmethoden zur Pose-Bestimmung des mobilen Endgeräts in einem globalen Geo-Referenzsystem benötigt. Existierende Ansätze für

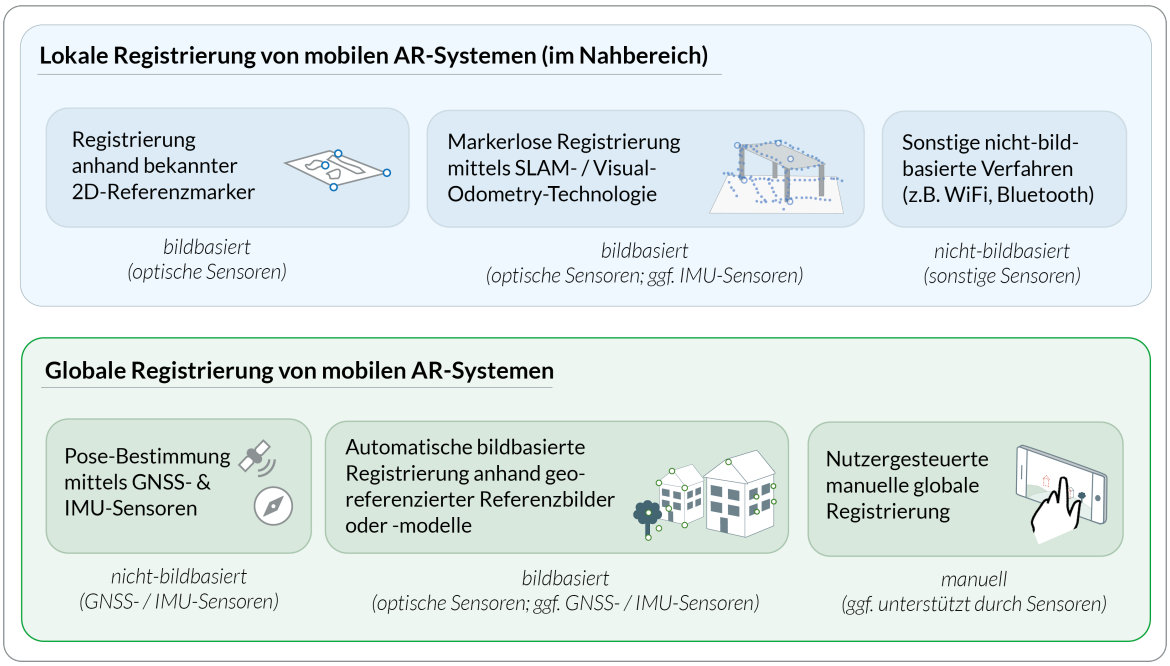


Abbildung 3.1 Übersicht über Registrierungsverfahren für mobile AR-Anwendungen

derartige Registrierungsmethoden in der aktuellen Forschung werden in Kapitel 3.2 vorgestellt. Dazu zählen insbesondere nicht-bildbasierte Verfahren mittels GNSS- und IMU-Sensoren sowie automatische bildbasierte Registrierungsverfahren anhand georeferenzierter Referenzbilder oder -modelle. Das in dieser Arbeit entwickelte manuelle Registrierungsverfahren stellt einen Spezialfall der globalen Registrierung von AR-Systemen dar. Derartige nutzergesteuerte globale Registrierungsmethoden wurden in der Forschung bisher kaum untersucht und entwickelt. Die wenigen existierenden Arbeiten in diesem Bereich werden in Kapitel 3.2.3 präsentiert. Die in diesem Kontext existierende Forschungslücke und die daraus resultierenden neuartigen Aspekte des entwickelten geodatenbasierten Registrierungsverfahrens werden abschließend in Kapitel 3.3 zusammengefasst.

3.1 Lokale Registrierung von AR-Systemen

Bei den meisten AR-Szenarien zur Visualisierung von virtuellen Inhalten im Nahbereich kann eine schnelle und robuste lokale Registrierung des AR-Systems mittels bildbasierter Tracking-Verfahren erfolgen. Zwar existiert grundsätzlich auch die Möglichkeit einer groben Positionsbestimmung in Innenräumen mittels nicht-bildbasierter Verfahren, z.B. durch Analyse der Signalstärken von WiFi- oder Bluetooth-Signalen [99]. Derartige

Methoden kommen bei praxistauglichen AR-Szenarien jedoch aufgrund der geringen Präzision eher selten zum Einsatz und sollen daher im Folgenden nicht näher betrachtet werden. Die grundsätzliche Funktionsweise der bildbasierten Verfahren besteht darin, dass in einem Kamerabild markante Bildpunkte (*image features*) detektiert werden und mit Bildpunkten in anderen Kamera-Frames oder mit Bildpunkten aus bekannten zweidimensionalen Referenzbildern oder dreidimensionalen Referenzkarten verglichen werden, um Position und Orientierung der Kamera innerhalb des gemeinsamen Referenzsystems zu bestimmen (siehe Kapitel 2.2). Auf diese Weise ist eine fortlaufende Bestimmung der Geräte-Pose in sechs Freiheitsgraden möglich.

3.1.1 Registrierung anhand bekannter 2D-Referenzmarker

Erste AR-Anwendungen verwendeten in diesem Kontext oft künstliche zweidimensionale Marker. In der Regel handelt es sich dabei um kontrastreiche schwarz-weiße Strukturen, welche auch in unterschiedlichen Lichtverhältnissen gut detektierbar sind, beispielsweise übliche QR-Codes oder spezielle Marker der AR-Tracking-Bibliothek *ARToolKit* oder *ARTag* [44] (siehe Abbildung 3.2(a)). Auch natürliche zweidimensionale Referenzbilder (z.B. Poster oder Farbbilder in Zeitschriften) können verwendet werden, um Position und Orientierung der Kamera in vorab bekannter Umgebung zu bestimmen und somit virtuelle Objekte oder Informationen positionsgetreu in Bezug auf diese Referenzbilder zu positionieren. Auf diese Weise können beispielsweise statische Werbeplakate durch virtuelle Animationen oder Videos mittels AR-Visualisierungen „zum Leben erweckt“ oder Produktabbildungen beim Blättern in Werbebroschüren durch virtuelle Zusatzinhalte „augmentiert“ und erweitert werden, z.B. beim Betrachten eines

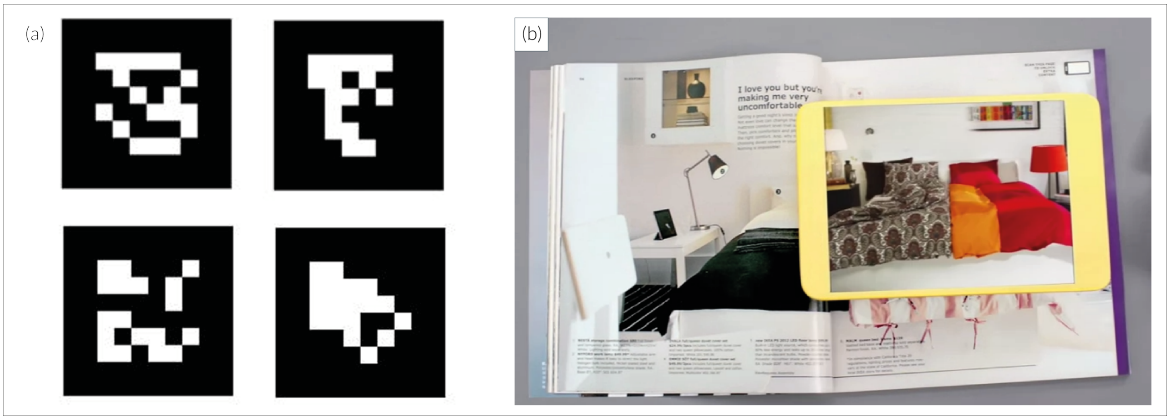


Abbildung 3.2 (a) Registrierung im Nahbereich mittels künstlicher AR-Marker, z.B. mit der Tracking-Bibliothek ARTag [44], oder (b) mittels natürlicher zweidimensionaler Referenzbilder, z.B. den Seiten eines IKEA-Katalogs [111].

IKEA-Katalogs durch Einblenden von zusätzlichen Ansichten der abgebildeten Möbel auf einzelnen Katalogseiten [111] (siehe Abbildung 3.2(b)).

Technische Grundlage für diese Augmentierungen bildet das Detektieren, Extrahieren und Vergleichen von Bilddeskriptoren im Live-Kamerabild und im vorab bekannten Referenzbild. Die in unterschiedlichen Arbeiten und Anwendungen verwendeten Bilddeskriptoren (z.B. SURF, SIFT, ORB) unterscheiden sich dabei vor allem hinsichtlich Robustheit, Präzision sowie Komplexität, und somit auch hinsichtlich Rechenaufwand und Geschwindigkeit bei den notwendigen Vergleichs- und Erkennungsoperationen [64]. Derartige bildbasierte Registrierungsmethoden für AR-Anwendungen anhand bekannter Referenzbilder ermöglichen zwar präzise und schnelle AR-Erlebnisse, funktionieren jedoch nur in vorab bekannter Umgebung, da Art und Aussehen des Referenzbildes dem System vorab bekannt sein müssen. Zudem ist dem AR-System nur die eine Ebene im Raum bekannt, auf der sich auch das Referenzbild befindet (z.B. Wand oder Katalogseite). Eine Kenntnis der restlichen dreidimensionalen räumlichen Umgebung, um beispielsweise Hindernisse und Verdeckungen im Anwendungsszenario berücksichtigen zu können, ist auf diese Weise nicht möglich.

3.1.2 Registrierung mittels Visual-SLAM-Technologie

Neue Möglichkeiten der mobilen AR-Visualisierung in unbekannten Umgebungen entstehen durch Einsatz der SLAM-Technologie (*Simultaneous Localization and Mapping*). Grundidee dieser Technik ist, dass ein mobiles Endgerät eine dreidimensionale Karte der Umgebung (3D-Punktwolke) erstellt und gleichzeitig seine aktuelle Position innerhalb dieser Karte bestimmt. Mit zunehmender Bewegung der mobilen Kamera kann die Karte erweitert und verbessert und die Position der Kamera innerhalb dieser Karte dadurch präzisiert werden (siehe Abbildung 3.3). Bei dem aus der Robotik entstammenden Verfahren kommen dabei oft auch Stereo-Kameras [94] oder RGB-D Kameras mit Tiefenerkennung [105] zum Einsatz. Für mobile AR-Anwendungen auf gewöhnlichen mobilen Endgeräten ist hingegen besonders die monokulare visuelle SLAM-Technologie von Bedeutung, bei der nur eine einzige optische Kamera eingesetzt wird.

Die initiale Erstellung und Erweiterung dieser 3D-Karte erfolgt bei monokularem SLAM auf Basis von *Triangulation* [137, S.345ff]: Bewegt sich die Kamera, so sind in aufeinander folgenden Kamerabildern dieselben 3D-Punkte aus unterschiedlichen Blickwinkeln im Kamerabild abgebildet. Können nun die Positionen von markanten 3D-Punkten in den aufeinander folgenden zweidimensionalen Kamerabildern gefunden werden (z.B. auf Basis von *Image Feature Tracking*), können mithilfe von mathematischen Optimierungsverfahren sowohl die relative Bewegung der Kamera als auch

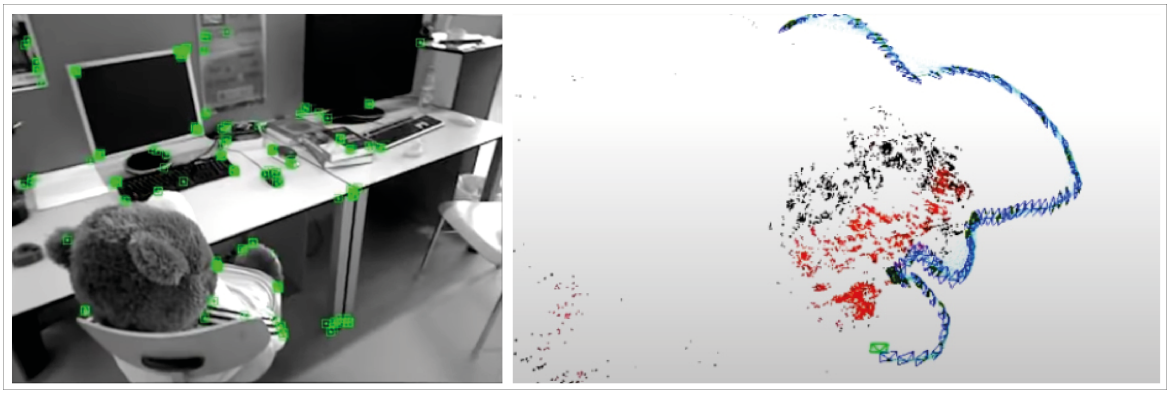


Abbildung 3.3 Registrierung in unbekannter Umgebung mittels monokularer SLAM-Technologie, z.B. ORB-SLAM [102]. In den Kamerabildern werden markante Feature-Punkte der Umgebung erfasst (links; grüne Punkte) und in einer 3D-Punktwolken-Karte gespeichert (rechts; rote Punkte). Mit zunehmender Kamerabewegung kann die Karte erweitert und die Kamerapositionen innerhalb dieser Karte fortlaufend erfasst werden (rechts; blauer Pfad).

die 3D-Positionen dieser Punkte geschätzt werden. Auf diese Weise ist eine initiale Schätzung einer 3D-Punkt-Karte sowie mit zunehmender Kamera-Bewegung eine Erweiterung und Optimierung dieser 3D-Punktwolke möglich. Dabei sollte jedoch beachtet werden, dass derartige Verfahren nur dann robust funktionieren, wenn die sichtbaren 3D-Punkte der Umgebung nicht zu weit entfernt sind. Ansonsten werden durch die Kamerabewegungen nur geringfügige Änderungen der Bildprojektion erzeugt (geringe Parallaxe), sodass Unterschiede zwischen aufeinander folgenden Kamerabildern nicht mehr erkennbar sind.

Ein SLAM-Verfahren wurde 2009 von Klein und Murray erstmals erfolgreich im Kontext von mobilen AR-Anwendungen eingesetzt [75]. Mit rasch zunehmender Rechenleistung mobiler Endgeräte und durch Weiterentwicklungen von monokularen SLAM-Verfahren, z.B. DTAM [106] und LSD-SLAM [40] sowie in jünger Vergangenheit insbesondere ORB-SLAM [102], entstanden im Laufe der letzten Jahre etablierte bildbasierte SLAM-Tracking-Systeme. Diese für mobile Endgeräte optimierten SLAM-Varianten integrieren weitere Mechanismen (z.B. *Loop Closure*, *Bundle Adjustment*), um den Ungenauigkeiten der ursprünglichen SLAM-Varianten (z.B. Drift), entgegenzuwirken und somit auch bei länger andauernden Kamerabewegungen eine präzise 6-DOF-Pose-Bestimmung mobiler AR-Geräte im Nahbereich in zuvor unbekannter Umgebung zu ermöglichen.

Ein großer Nachteil dieser monokularen visuellen SLAM-Verfahren besteht darin, dass zwar Position und Orientierung einer Kamera relativ zu einem Referenzkoordinatensystem präzise bestimmt werden können, jedoch ohne metrische Skalierung (*scale*

ambiguity) [139]: Kamerabewegungen und damit auch die Entfernung der Kamera zu den erkannten 3D-Punkten im Raum können lediglich mit relativer Skalierung zueinander, nicht jedoch in absoluter metrischer Skalierung (z.B. in cm) erfasst werden. Für viele AR-Anwendungen ist dies irrelevant, da eine relative Skalierung der virtuellen Objekte im Raum ausreicht. Für andere AR-Szenarien ist jedoch eine metrische Skalierung der AR-Inhalte zwingend notwendig (z.B. AR-basierte Entfernungsmessung oder AR-Anwendung zur größengetreuen Platzierung von bestimmten Möbelstücken im Raum). Zur Lösung dieses Problems ist die Integration einer bekannten (oder geschätzten) Referenzgröße notwendig. Es existieren daher einige Ansätze für SLAM-Verfahren mit metrischer Skalierung, die zu diesem Zweck die geschätzte Höhe der mobilen Kamera über dem Boden [74] einbeziehen oder davon ausgehen, dass die Dimension eines bestimmten 3D-Objektes am Startpunkt des Trackingvorgangs bekannt ist [34].

3.1.3 Registrierung mittels Visual-Inertial-Odometry

Zur Bestimmung der Geräte-Pose eines AR-Systems mit metrischer Skalierung hat sich in den letzten Jahren außerdem die Technologie der *Visual Inertial Odometry (VIO)* etabliert. Der Begriff der *Visual Odometry* bezeichnet in erster Linie die inkrementelle bildbasierte Bestimmung der Kamerabewegung von Frame zu Frame. Die Problemstellung der Visual Odometry, also die fortlaufende Schätzung der sich bewegenden Kamera, kann damit als Teilaspekt der SLAM-Technologie verstanden werden [150]. Durch geeignete Kombination der optischen Sensordaten der Visual Odometry-Verfahren mit Sensordaten von inertialen Messeinheiten (Gyroskop, Accelerometer, Magnetometer) entstehen schließlich Visual-Inertial-Odometry-Systeme (VIO-Systeme), welche die Kamerabewegung nicht nur mit relativer, sondern ebenfalls mit metrischer Skalierung fortlaufend erfassen können [39, 86] (siehe Abbildung 3.4). In Kombination mit den genannten Eigenschaften und Optimierungen der monokularen SLAM-Verfahren spricht man in diesem Kontext auch von Visual-Inertial-SLAM-Verfahren [103].

Derartige bildbasierte Verfahren sind aktuell als die fortschrittlichsten Verfahren zu betrachten, um die Pose mobiler AR-Endgeräte im Nahbereich fortlaufend zu bestimmen. Durch die Bereitstellung dieser VIO-Verfahren im Rahmen von AR-Entwicklerwerkzeugen – hier sind in erster Linie Googles *ARCore* (Release 2018) und Apples *ARKit* (Release 2017) zu nennen – sind inzwischen etablierte und robuste Tools für eine breite Gruppe von Anwendungsentwicklern vorhanden, um AR-Systeme präzise im Nahbereich zu registrieren [80].

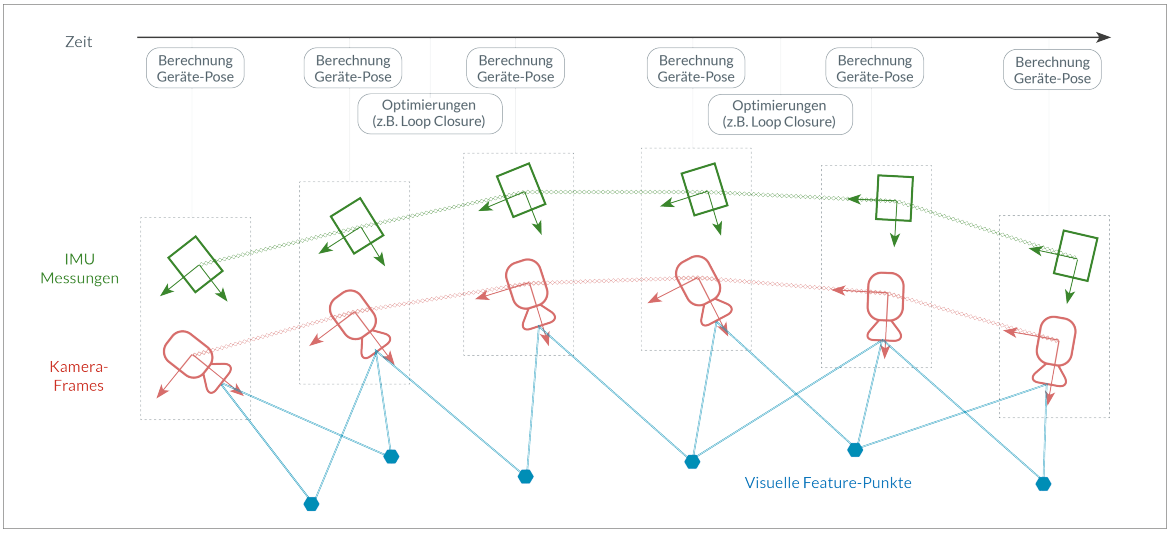


Abbildung 3.4 Optische Tracking-Verfahren können durch das Erkennen von markanten Feature-Punkten (blau) in fortlaufenden Kamera-Frames die Bewegung der Kamera (rot) kontinierlich erfassen und verfolgen (Visual Odometry). In Kombination mit Messungen von inertial IMU-Sensoren (grün) entstehen Visual-Inertial-Odometry-System, welche eine fortlaufende Erfassung der Geräte-Pose mit absoluter metrischer Skalierung ermöglichen.

3.2 Globale Registrierung von AR-Systemen

Diese bisher genannten Verfahren sind für eine Registrierung mobiler Systeme im Nahbereich in Bezug auf ein lokales Koordinatensystem geeignet, welches jedoch willkürlich in Bezug auf ein globales geografisches Referenzsystem ausgerichtet ist. Für eine globale Registrierung von mobilen AR-Systemen sind daher andere Verfahren notwendig.

3.2.1 Pose-Bestimmung mittels GNSS- und IMU-Sensoren

Oftmals werden zum Zweck der globalen Registrierung von Mobilgeräten integrierte GNSS-Sensoren sowie Sensoren der inertialen Messeinheit (IMU) genutzt. Zwar ist der Einsatz von GNSS-Empfängern in der Regel in Innenräumen nicht problemlos möglich, für die globale Registrierung in GeoAR-Anwendungen stellt dies jedoch eher keine Einschränkung dar, da derartige Anwendungen ohnehin im Normalfall für den Außeneinsatz gedacht sind. GNSS- und IMU-Sensoren sind heutzutage in nahezu allen mobilen Endgeräten verbaut, ermöglichen jedoch nur eine grobe und in der Regel fehlerbehaftete Bestimmung von globaler Position und Orientierung (siehe Kapitel 2.2.3). Frühe mobile Anwendungen im Bereich der geodatenbasierten Augmented Reality verwendeten zur globalen Registrierung ausschließlich diese Möglichkeiten, bei-

Abbildung 3.5 (a) In mobilen Endgeräten integrierte GNSS- und IMU-Sensoren ermöglichen eine grobe, jedoch fehlerbehaftete Bestimmung von globaler Geräte-Pose. Einige Outdoor AR-Anwendungen nutzen diese Registrierungsmethode, beispielsweise zur Darstellung von AR-Markern nahegelegener Immobilien [149]. (b) Der Einsatz zusätzlicher externer Sensoren erlaubt eine globale Registrierung mit deutlich höherer Präzision, beispielsweise das Produkt *Trimble SiteVision*, das z.B. zu AR-basierten Darstellungen während Bauausführungen eingesetzt werden kann [140].

spielsweise die AR-Anwendung *SkyView* [131] zur Visualisierung von Sternbildern am Nachthimmel oder die Anwendung *ImmobiliAR* [149], mit der nahegelegene Immobilien als Points-Of-Interest-Marker in einer AR-Ansicht dargestellt werden können (siehe Abbildung 3.5(a)).

Ein Einsatz ausschließlich dieser Sensoren führt jedoch in der Regel zu mehreren Metern Abweichung bei der globalen Positionsbestimmung sowie mehreren Grad Abweichung bei der Bestimmung der globalen Orientierung, was in einer deutlich fehlerhaften AR-Projektion der virtuellen Inhalte im Kamerabild resultiert [113, 95, 124]. Eine derart fehlerhafte Platzierung von AR-Inhalten auf dem Bildschirm ist für die meisten GeoAR-Anwendungen sehr unbefriedigend.

Die Genauigkeit der georäumlichen Positionierung kann durch den Einsatz von zusätzlicher externer Hardware in Form von externen IMU-Sensoren und RTK-GNSS-Empfängern mit externer Antenne deutlich erhöht werden [124]. Zusätzlich zur gewöhnlichen Vorgehensweise der GNSS-Positionierung werden dabei Korrektursignale von ortsfesten Referenzstationen empfangen und verarbeitet (siehe auch Kapitel 2.2.3). Im Idealfall ist so eine hochpräzise, zentimetergenaue globale Positionsbestimmung als Basis für realistische Outdoor AR-Szenarien möglich [134]. In marktreifer Ausführung wird dieses Verfahren beispielsweise durch das Produkt Trimble SiteVision genutzt, welches AR-Visualisierungen im Außenbereich auch für großflächige Szenarien ermöglicht, z.B. für die Darstellung von geplanten Gebäuden oder Verkehrsanlagen

aus größerer Entfernung [140] (siehe Abbildung 3.5(b)). Die Notwendigkeit derartiger kostenintensiver externer Sensoren und Antennen mag diese Registrierungsmethode zwar für Fachanwender:innen praktikabel machen, jedoch nicht für die gewöhnliche Nutzung mobiler GeoAR-Anwendungen im Alltag.

3.2.2 Automatische bildbasierte globale Registrierung

In den vergangenen Jahren sind mit bildbasierten globalen Registrierungsmethoden – auch im mobilen Kontext – robuste Alternativen zu den universell einsetzbaren, jedoch ungenauen GNSS- und IMU-basierten Registrierungsverfahren entstanden. Die Grundidee dieser bildbasierten Verfahren besteht darin, dass Live-Bilder der Kamera vor Ort mit ähnlichen, vorab aufgenommenen georeferenzierten Bildern aus einer Bilder-Datenbank [7] oder mit einer vorab generierten georeferenzierten 3D-Punktwolke [87, 121, 136, 146] verglichen werden. Meist werden lokale Kamerabewegungen dabei auf herkömmliche Art und Weise mittels SLAM-Methoden in einem lokalen Bezugssystem registriert (siehe Kapitel 3.1.2). Eine Lokalisierung innerhalb der global referenzierten 3D-Punktwolken-Karte erfolgt dann nicht mit allen Kamerabildern, sondern in regelmäßigen Abständen nur mit ausgewählten Kamerabildern (*Keyframes*). Die dadurch berechnete globale Position und Orientierung der mobilen Kamera kann dadurch regelmäßig aktualisiert werden. Manche Arbeiten ermöglichen auch eine fortwährende dynamische Erweiterung der globalen 3D-Punktwolken-Karte, indem Teile der lokalen SLAM-Karte sukzessive in die globale Karte integriert werden [136]. Weitere Optimierungen können erreicht werden, indem die lokalen bildbasierten Trackingverfahren und die globale Registrierung mit den in Mobilgeräten integrierten GNSS- und IMU-Sensoren kombiniert werden [65, 135, 96].

Auf diesen Ansätzen und Forschungsarbeiten basierend sind in jüngster Zeit verschiedenartige proprietäre Anwendungen und Produkte von kommerziellen Anbietern entstanden, um eine bildbasierte globale Registrierung mobiler Endgeräte – in bestimmten ausgewählten Einsatzgebieten – zu ermöglichen: *Google Maps Live View* ermöglicht innerhalb bestimmter städtischer Gebiete eine präzise bildbasierte Pose-Bestimmung und AR-basierte Fußgängernavigation, basierend auf zuvor aufwändig erstellter StreetView-Bildaufnahmen [98] (siehe Abbildung 3.6). Apple realisiert auf sehr ähnliche Art und Weise ebenfalls in ausgewählten Städten eine AR-basierte Navigationsmöglichkeit, basierend auf der *Location Anchors*-Funktionalität des ARKit-SDK [5]. Das US-Unternehmen Niantic veröffentlichte vor kurzem ebenfalls eine AR-Plattform *Lightship*, die in ausgewählten Gebieten eine bildbasierte globale Positionierung und damit eine Platzierung von individuellen AR-Inhalten erlaubt [107]. Ein ähnliches

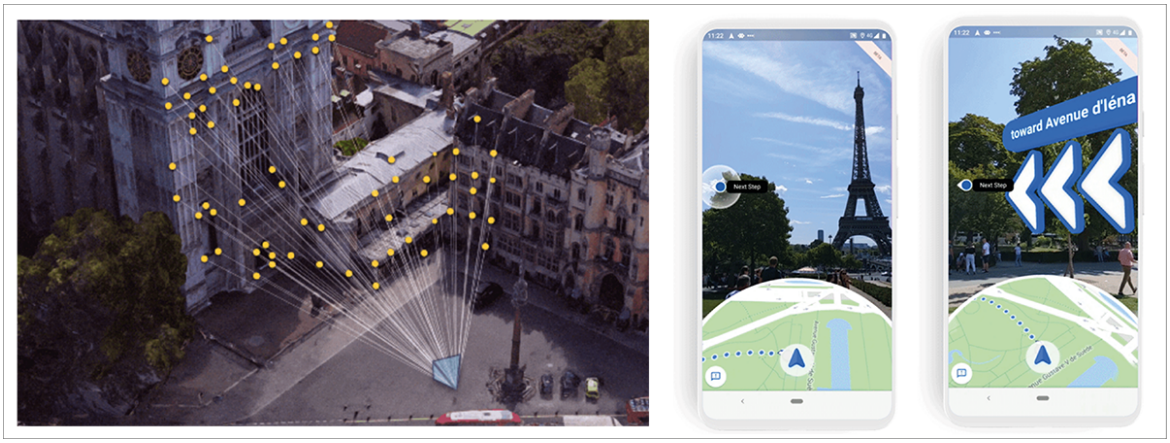


Abbildung 3.6 Auf Basis von georeferenzierten 3D-Punktwolken (links) ermöglicht Google Maps Live View in ausgewählten Städten eine automatische bildbasierte Registrierung mobiler Geräte, z.B. zur AR-basierten Fußgängernavigation (rechts) [98].

Anwendungsszenario wird vom Unternehmen Snap mit der Funktion *Local Lenses* bereitgestellt, mit der an bestimmten Standorten kollaborative AR-Interaktionen mit tatsächlichen physischen Objekten vor Ort ermöglicht werden (z.B. Virtuelle Street-Art-Malerei an Hausfassaden) [16].

Andere bildbasierte Registrierungsverfahren nutzen existierende Geodaten zur Unterstützung der globalen Positionierung: Arth et al. entwickelten ein System für eine globale Registrierung durch Einbindung von groben Gebäudemodellen aus OpenStreetMaps [6, 88] (siehe Abbildung 3.7(a)). Die Gebäudemodelle werden dabei nicht nur für die Bestimmung von georäumlicher Position und Orientierung genutzt, sondern ermöglichen auch eine absolute metrische Skalierung des SLAM-basierten lokalen Tracking-Systems. Lothe et al. präsentierten ein ähnliches System für eine globale automatische Registrierung von Fahrzeugen durch Nutzung eines groben 3D-Stadtmodells [89]. Das digitale Stadtmodell wird dabei zusätzlich verwendet, um anhand geometrischer Randbedingungen sich anhäufende Fehler innerhalb der SLAM-Karte zu korrigieren. Neben digitalen Stadt- und Gebäudemodellen existieren auch einige Ansätze von globalen bildbasierten Registrierungsverfahren auf Basis von digitalen Geländemodellen (DGM). Mithilfe von Bildanalyseverfahren wird dabei versucht, automatisch eine Horizontlinie (Horizontsilhouette) aus den Kamerabildern zu extrahieren, um diese schließlich mit virtuellen Horizontsilhouetten der Geländemodelle zu vergleichen und daran auszurichten [43, 13] (siehe Abbildung 3.7(b)). Weitere ähnlich geartete bildbasierte globale Registrierungsansätze sind in einer Arbeit von Brejcha und Čadík zusammengefasst [23].

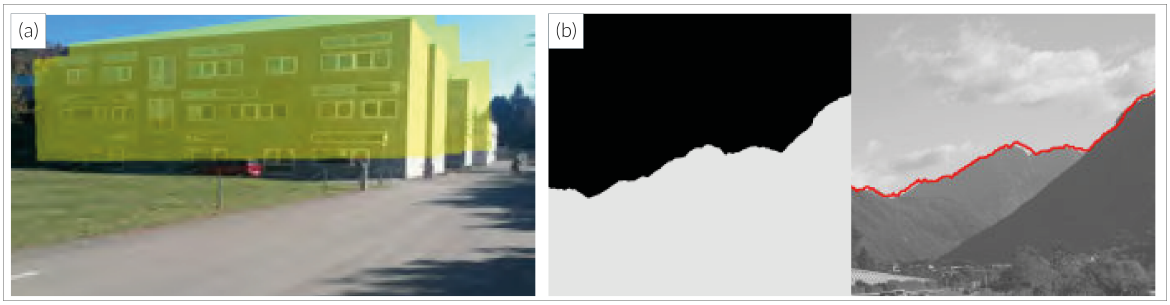


Abbildung 3.7 (a) In einer Arbeit von Arth et al. werden Gebäudemodelle aus OpenStreetMap-Datensätzen für eine automatische bildbasierte globale Registrierung mobiler AR-Systeme genutzt [6]. (b) Baatz et al. vergleichen virtuelle Horizont-Silhouetten aus digitalen Höhenmodellen (links) mit extrahierten Horizontlinien aus realen Kamerabildern (rechts), um die globale Position und Orientierung mobiler Endgeräte zu bestimmen [13].

All diese Ansätze setzen auf bildbasierte Vergleichs- und Tracking-Verfahren und besitzen daher – neben dem großen Vorteil der vollautomatischen, nicht-manuellen und dadurch meist schnellen Registrierung – auch mehrere Nachteile und Einschränkungen:

- **Umfangreiche Referenz-Bilddaten**: Die meisten der genannten Verfahren sind abhängig von vorab generierten und georeferenzierten Bilddatenbanken oder 3D-Punktewolken-Karten in hoher Auflösung und Qualität. Die Generierung und Speicherung derartiger Datensätze ist kostenintensiv und aufwändig, besonders wenn großflächige Gebiete abgedeckt werden sollen. Die Bereitstellung einer solchen Datengrundlage ist flächendeckend daher nur durch sehr große Organisationen oder Unternehmen zu leisten.

- **Datenschutz**: Zur Funktionsfähigkeit der automatischen bildbasierten Methoden müssen die zugrundeliegenden Referenzbilddaten bzw. Punktwolken in der Regel in recht hoher Auflösung erstellt und bereitgestellt werden. Dabei ist zu berücksichtigen, dass personenbezogene Daten (Privatgrundstücke, KFZ-Kennzeichen o.ä.) anonymisiert und etwaige persönliche Schutzrechte nicht verletzt werden. Eine derartige nachträgliche Anonymisierung der Bilddaten ist aufwändig und unter Umständen nur durch manuelle Bearbeitung möglich.

- **Beschränkte Umgebungscharakteristik**: Einige der genannten Verfahren sind angewiesen auf bestimmte Umgebungsvoraussetzungen, z.B. eine freie Fernsicht in bergiger Umgebung für einen erfolgreichen automatischen Abgleich der Horizontlinien [43, 13], oder andere bestimmte Umgebungscharakteristiken, z.B. das Vorhandensein von planaren Hausfassaden [6, 88].

- **Unzureichende Bildqualität und Störfaktoren**: Bildbasierte Trackingverfahren sind fehleranfällig für unzureichende oder uneinheitliche Beleuchtung der Umgebung, also bei sehr heller oder dunkler Umgebung oder wenn sich die aktuellen Lichtverhältnisse deutlich unterscheiden von den Referenzbilddaten. Zudem sind derartige Bilderkennungsmethoden oftmals nur sehr eingeschränkt funktionsfähig in dynamischen nicht-statischen Szenen (z.B. sich bewegender Verkehr oder Personen) oder in Szenen mit sich wiederholender und sehr homogenen Strukturen (z.B. weiße Hauswand, Gras etc.).

Diese Einschränkungen führen dazu, dass automatische bildbasierte globale Registrierungsverfahren zwar unter bestimmten Rahmenbedingungen (Licht, Wetter, Art der Umgebung) und bei Einsatz von großen Ressourcen (Datengrundlage, Infrastruktur) eine schnelle, nutzerfreundliche und präzise AR-Registrierung ermöglichen können, jedoch in bestimmten Fällen auch nur eine sehr eingeschränkte Nutzbarkeit aufweisen.

3.2.3 Nutzergesteuerte globale Registrierung

Im Gegensatz zu solchen automatischen bildbasierten Registrierungsmethoden lassen sich nur wenige Ansätze finden, die eine AR-basierte benutzergesteuerte Interaktion für eine präzise globale Registrierung nutzen. Die Grundidee dabei besteht darin, dass durch aktive Mithilfe und Einbeziehen der Kenntnisse und Fähigkeiten der Nutzer:innen bei der Ausrichtung und Positionierung des AR-Systems ein flexibler und unkomplizierter Registrierungsprozess entstehen kann, der in gewissen Fällen den automatisierten Prozessen hinsichtlich Robustheit und Fehleranfälligkeit überlegen sein kann. Analog zu den bildbasierten Methoden sind als Grundlage für manuell unterstützte globale Registrierungsverfahren ebenfalls global referenzierte Referenzdaten notwendig. Diese können jedoch von sehr unterschiedlicher Art und Qualität sein: punktförmige Geodaten können ebenso eingesetzt werden wie komplexe dreidimensionale Geodatenmodelle in unterschiedlicher Auflösung.

In der Arbeit von Fukuda et al. erfolgt eine manuelle kartenbasierte Positionsbestimmung zur Präzisierung einer initialen GNSS-basierten Positionierung für Outdoor-AR-Anwendungen [45]. Dort werden zudem digitale Geländemodelle eingesetzt, um an der präzisierten 2D-Kartenposition eine Kalibrierung des Höhenwertes durchzuführen. Eine Kalibrierung und Korrektur der globalen Orientierung erfolgt dabei nicht. Kilimann et al. implementierten eine Methode, um virtuelle AR-Marker an manuell definierten, gut sichtbaren Orientierungspunkten in der realen Landschaft (z.B. Kirchtürme, Straßenkreuzungen) auszurichten, um damit einen einzigen Freiheitsgrad

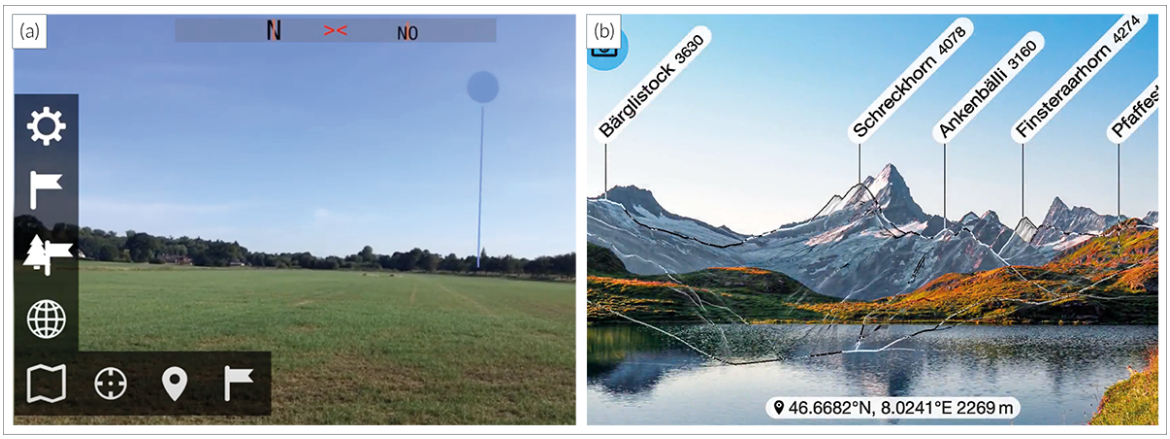


Abbildung 3.8 (a) In einer Arbeit von Kilimann et al. kann die Kameraorientierung anhand markanter punktförmiger Orientierungspunkte im Kamerabild manuell korrigiert werden [72]. (b) In der kommerziellen Anwendung PeakFinder AR ist eine manuelle Kalibrierung der Kameraorientierung anhand umliegender, virtuell eingeblendeten Bergkonturen möglich [132].

der globalen Kameraausrichtung zu korrigieren (Rotation um die vertikale Achse) [72] (siehe Abbildung 3.8(a)). Eine präzise globale Positionierung erfolgt in diesem System ebenfalls manuell über eine zweidimensionale Kartenansicht.

Huang et al. präsentierten einen manuellen Registrierungsmechanismus für Outdoor AR-Systeme, der speziell für dreidimensionale Wireframe-Modelle von einzelnen Gebäuden ausgelegt ist [58]. Die Interaktion mit dem System erfolgt dabei nicht direkt auf dem Bildschirm des AR-Systems, sondern durch Interaktion auf einem speziellen mobilen Controller. Es existieren weitere ähnliche Ansätze, die auf dem Einsatz spezieller Hardware oder Zusatz-Sensoren beruhen, beispielsweise statischen Panorama-Videoaufnahmen [56] oder speziellen Handschuhen und Laser-Scannern zur manuellen Steuerung und globalen Ausrichtung von virtuellen Modellen [33].

Thematisch verwandt zu der in dieser Arbeit entwickelten Registrierungsmethode ist eine Arbeit von Gazcón et al., die eine manuelle nutzergesteuerter AR-Registrierung zur AR-basierten Unterstützung der Geländearbeit von Geologen vorschlägt [48]. Bevor virtuelle georeferenzierte Punkte oder Ebenen zur Unterstützung der geologischen Feldarbeit als AR-Markierungen in der Szene erzeugt werden können, kann durch Verschieben eines virtuellen Geländemodells das mobile Endgerät an der realen Landschaftsstruktur korrekt ausgerichtet werden. Außerdem existiert die kommerziell erhältliche mobile Anwendung *PeakFinderAR* [132], welche Berggipfel in der Umgebung per verschiebbarer und korrigierbarer AR-Ansicht darstellen kann (siehe Abbildung 3.8(b)). Durch manuelles Verschieben der virtuell gezeichneten Bergkonturen auf dem AR-

Display ist eine Kalibrierung der globalen Geräteausrichtung in zwei Freiheitsgraden möglich. Die in diesen beiden Systemen eingesetzten AR-basierten nutzergesteuerten Registrierungsverfahren basieren jedoch lediglich auf sehr grob aufgelösten digitalen Geländemodellen (DGM) und sind daher nicht überall einsetzbar, sondern auf gebirgige Umgebungen mit markanten Landschaften beschränkt. Die dort eingesetzten Interaktionsarten wurden zudem nur sehr grob bzw. gar nicht wissenschaftlich hinsichtlich Genauigkeit sowie Bedienbarkeit untersucht und evaluiert. Die in dieser Arbeit entwickelte Methode ist hingegen nicht auf bestimmte Umgebungen und grob aufgelöste Geländemodelle beschränkt, sondern integriert explizit verschiedene Arten von Geodaten – insbesondere auch texturierte und nicht-texturierte 3D-Stadtmodelle und hochaufgelöste Oberflächenmodelle – und stellt damit eine robuste Registrierungsmethode dar, die in unterschiedlichen Umgebungen flexibel einsetzbar ist. Zudem ist bei der hier vorgestellten Methode nicht nur eine Kalibrierung der globalen Geräteausrichtung, sondern auch der globalen Geräteposition möglich.

3.3 Zusammenfassung und Forschungslücke

Mit bildbasierten SLAM- und VIO-Technologien existieren heutzutage robuste und etablierte Verfahren zur präzisen Registrierung mobiler AR-Systeme im Nahbereich (siehe Kapitel 3.1). Auch für die Problemstellung der globalen Registrierung, die für GeoAR-Anwendungen notwendig ist, existieren bildbasierte Verfahren, die präzise und vielversprechend sind, jedoch aufgrund hoher Komplexität und hohem Ressourcenaufwand in vielen Situationen nicht ohne Weiteres einsetzbar sind (siehe Kapitel 3.2.2). Nutzergesteuerte globale Registrierungsverfahren stellen in diesem Kontext eine flexible und robuste Alternative dar. Die aktive Mithilfe der Nutzer:innen und die spielerische Interaktion bei der Durchführung einer globalen Registrierung – z.B. durch interaktives Verschieben von dreidimensionalen Umgebungsmodellen in der AR-Ansicht – kann dabei helfen, den Registrierungsprozess nachvollziehbar zu machen, während der dafür notwendige Zeitaufwand vertretbar bleibt. Die AR-basierte Integration von digitalen Geodatenmodellen als virtuelle Repräsentation der Nutzerumgebung wird zwar bereits in einigen Anwendungen praktiziert, jedoch existieren nur sehr wenige Arbeiten, die derartige Geodaten ebenfalls zum Zwecke der AR-basierten globalen Registrierung einsetzen (siehe Kapitel 3.2.3). Die wenigen präsentierten Verfahren nutzen dabei nur punktförmige Geodaten als Referenzobjekte [72], nutzen 3D-Geodatenmodelle nur zur automatischen Höhenkalibrierung [45] oder sind auf den Einsatz von groben Geländemodellen beschränkt und daher nur in eingeschränkten Umgebungen, nämlich in bergigem

Umfeld, nutzbar [48, 132]. Auch die Genauigkeit und Bedienbarkeit der Verfahren in unterschiedlichen Umgebungen wurde in diesen Arbeiten nicht genauer untersucht. Im Bereich der manuellen globalen Registrierungsmethoden mittels Einbindung von 3D-Geodaten herrscht daher noch großer Forschungsbedarf.

Die vorliegende Arbeit soll an dieser Stelle ansetzen und dabei helfen, die Forschungslücke in diesem Bereich weiter zu schließen. Dazu sollen insbesondere die folgenden neuartigen Aspekte betrachtet werden:

- **Vollständige Kalibrierung der 6-DOF-Geräte-Pose sowie deren mathematische Beschreibung**: Existierende Arbeiten präsentieren zwar Mechanismen, um einzelne Freiheitsgrade (z.B. Geräterotation oder geografische Höhe) mithilfe von 3D-Geodatenmodellen zu korrigieren, setzen die eingesetzten Interaktionsformen jedoch gar nicht mathematisch in einen Zusammenhang mit Koordinatensystem-Transformationen zur Kalibrierung der globalen Geräte-Pose. In der vorliegenden Arbeit sollen unterschiedliche Interaktionsformen (2D-Bildschirmgesten) zur vollständigen Kalibrierung der 6-DOF Geräteposition und -ausrichtung unter Zuhilfenahme von virtuellen Geodaten nicht nur entwickelt und evaluiert, sondern auch mathematisch beschrieben werden.

- **3D-Geodatenmodelle für unterschiedliche Umgebungen**: Die existierenden Arbeiten zu diesem Thema beschränken sich auf eine bestimmte Art von Geodaten (3D-Geländemodelle) und sind daher nur in bestimmten sehr markanten Landschaftsstrukturen einsetzbar. In der vorliegenden Arbeit sollen zusätzlich digitale 3D-Stadtmodelle sowie 3D-Oberflächenmodelle in verschiedenen Ausprägungen (z.B. mit und ohne Farbtextur) integriert werden, um das Registrierungsverfahren in beliebigen Umgebungen mit unterschiedlicher Charakteristik (flache Umgebung, urbane Umgebung, bergige Umgebung) zu testen und zu vergleichen.

- **Genauigkeit und Anwendungspotenzial**: Eine ausführliche Evaluation hinsichtlich Genauigkeit und Nutzerfreundlichkeit ist bislang bei keiner anderen Arbeit zu diesem Thema verfügbar. Im Rahmen dieser Arbeit soll das entwickelte Verfahren praxisnah in unterschiedlicher Ausführung und an unterschiedlichen Orten getestet werden, um Vor- und Nachteile sowie generelle Rückschlüsse auf das Anwendungspotenzial des Verfahrens zu erörtern und zu diskutieren.

Kapitel 4

Nutzergesteuerte geodatenbasierte AR-Registrierung

Im folgenden Kapitel soll das entwickelte geodatenbasierte Registrierungsverfahren vorgestellt werden. Die Grundidee der entwickelten Methode besteht darin, durch AR-basierte Einblendung verschiedenartiger 3D-Geomodelle die räumliche Umgebung um die Nutzer:innen herum virtuell zu repräsentieren, um durch dieses visuelle Hilfsmittel eine manuelle Bestimmung von globaler Geräteposition und -orientierung zu ermöglichen. Das nutzergesteuerte Verfahren zur Kalibrierung der globalen Kamera-Pose ist dabei eingebettet in ein Gesamtsystem, welches alle Prozessierungsschritte, die zum operativen Einsatz der geodatenbasierten Registrierung notwendig sind, integriert – von der Vorverarbeitung und Bereitstellung AR-geeigneter 3D-Geodatenmodelle bis hin zur AR-basierten Visualisierung der Geodatenmodelle auf dem AR-Client zum Zwecke der manuellen Kalibrierung.

Nach einem Überblick über alle notwendigen System-Komponenten (Kapitel 4.1) werden zunächst die zur Registrierung eingesetzten Geodaten genauer beschrieben sowie die Verarbeitungsschritte vorgestellt, die zur Prozessierung und Umwandlung der Geodaten in AR-geeignete 3D-Formate notwendig sind (Kapitel 4.2). Die Schnittstellen und Abläufe zur Speicherung und Verwaltung der generierten 3D-Modelle werden danach in Kapitel 4.3 präsentiert. Anschließend erfolgt eine mathematisch-technische Beschreibung der entwickelten Registrierungsmethode. Dazu werden die involvierten Koordinatensysteme (Kapitel 4.4.1) und die Interaktionsformen zur Kalibrierung von globaler Kameraposition und -orientierung (Kapitel 4.4.2) mathematisch hergeleitet und erläutert.

4.1 System-Architektur und -Komponenten

Zur Realisierung der geodatenbasierten Registrierungsmethode wurde ein Gesamtsystem entwickelt, welches aus zwei Hauptkomponenten besteht: Eine Komponente zur Geodatenverarbeitung (Geodatenverarbeitungskette) konvertiert gängige 3D-Geodaten (GIS-Daten) in kleinflächige AR-fähige 3D-Modelle, welche effizient gespeichert, einem Client zur Verfügung gestellt und dort performant visualisiert werden können. Die zweite Hauptkomponente entspricht einem Echtzeit-Registrierungs- und Tracking-System, welches clientseitig implementiert wird, und sowohl eine präzise globale nutzergesteuerte Geräteregistrierung als auch ein robustes automatisches lokales Geräte-Tracking ermöglicht. Als Bindeglied zwischen diesen beiden Hauptkomponenten existiert ein Geodaten-Content-Management-System (Geodaten-CMS) zur serverseitigen Speicherung und Verwaltung der 3D-Geodatenmodelle. Die grundsätzliche Funktionsweise der genannten System-Komponenten und deren Beziehungen zueinander ist in Abbildung 4.1 skizziert:

Geodatenverarbeitungskette Die 3D-Geodaten, die für die nutzergesteuerte Kalibrierung des AR-Systems eingesetzt werden sollen, liegen in der Regel nicht in Formen und Formaten vor, die eine problemlose serverbasierte Bereitstellung sowie AR-basierte Integration innerhalb des mobilen Clients erlauben. Eine Offline-Komponente wurde daher entworfen, um 3D-Geodaten aus üblichen GIS-Formaten anzugleichen und so in kleinflächige 3D-Kacheln umzuwandeln, dass diese clientseitig effizient als AR-Visualisierung integriert werden können. Hierfür sind mehrere Umwandlungsschritte notwendig (siehe Kapitel 4.2.2):

- **Koordinatensystem-Konvertierung**: Zunächst werden die eingesetzten Daten uniform in ein lokales metrisches Geo-Koordinatensystem konvertiert.
- **Kachel-Generierung**: Anschließend werden die Quelldaten für ein effizientes Datenhandling in kleinflächige quadratische Kacheln mit vorab definierten Abmessungen zerlegt (Tiling).
- **Reduktion und Komprimierung**: Zur Reduktion der Zieldateigröße ist zudem unter Umständen auch eine Verringerung der räumlichen Auflösung oder eine stärkere Kompression etwaiger Modell-Texturen angebracht.
- **3D-Mesh-Generierung**: Sofern die Quelldateien anfangs nur in textbasierter Form als 3D-Punktwolke vorliegen, erfolgt auch die Generierung eines 3D-Oberflächennetzes (3D-Mesh), um ein renderfähiges 3D-Modell zu erhalten.

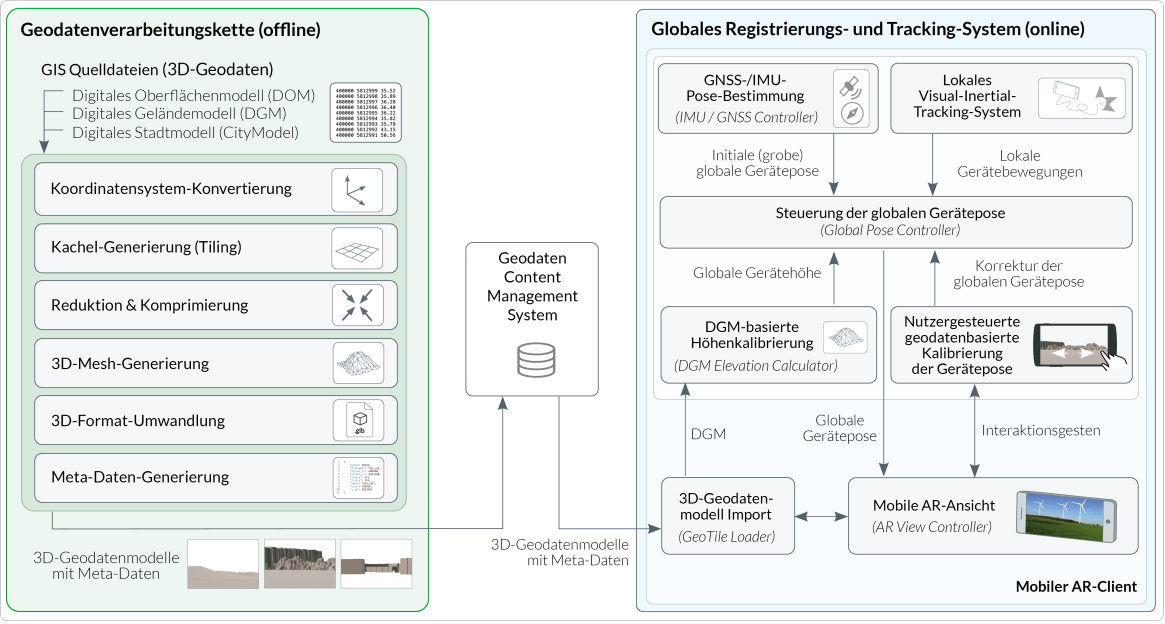


Abbildung 4.1 System-Architektur der geodatenbasierten AR-Registrierungsmethode.

- **3D-Format-Umwandlung**: Die 3D-Kacheln werden schließlich in ein binäres 3D-Dateiformat umgewandelt, sodass diese effizient gespeichert, verteilt und clientseitig geladen werden können.
- **Meta-Daten-Generierung**: In einem abschließenden Schritt werden für jede 3D-Kachel alle relevanten Meta-Daten (geografische Position, Abmessungen, verwendetes Koordinatensystem u.a.) in einem einheitlichen Datenformat zusammengestellt, um ein effektives Datei-Management der Modelldateien zu ermöglichen.

Diese einzelnen Verarbeitungsschritte der Geodatenverarbeitungskette werden in Kapitel 4.2.2 im Detail erläutert.

Geodaten-Content-Management-System Die primäre Aufgabe dieser serverbasierten Komponente ist die Speicherung der generierten 3D-Kacheln auf Basis der bei der Generierung gesammelten Meta-Daten sowie die Bereitstellung der 3D-Geodatenmodelle nach Anfrage durch einen Client über definierte Schnittstellen (siehe Kapitel 4.3).

Globales Registrierungs- und Tracking-System Innerhalb des AR-Clients wird zur Laufzeit ein System zur globalen Registrierung und zum lokalen Tracking des mobilen Endgeräts ausgeführt:

- **GNSS-/IMU-Pose-Bestimmung**: Die integrierten GNSS- und IMU-Sensoren werden zunächst für eine ungefähre initiale Schätzung von globaler Kameraposition und -orientierung genutzt.
- **3D-Geodatenmodell-Import**: 3D-Kacheln zur Repräsentation der aktuellen Umgebung werden vom Client geladen (3D-Geodatenmodell-Import) und basierend auf dieser initialen Schätzung mit ihrer tatsächlichen metrischen Skalierung als AR-Visualisierung in der mobilen AR-Ansicht dargestellt.
- **DGM-basierte Höhenkalibrierung**: Digitale Geländemodelle können genutzt werden, um eine automatische Höhenkalibrierung der GNSS-basierten Positionsschätzung durchzuführen, also eine Kalibrierung der vertikalen globalen Kameraposition.
- **Nutzergesteuerte geodatenbasierte Kalibrierung**: Diese initiale Schätzung der globalen Kamera-Pose erzeugt eine grobe – in der Regel jedoch deutlich fehlerhafte – virtuelle 3D-Projektion der Umgebung in der AR-Ansicht. Mittels Interaktionsgesten kann nun die nutzergesteuerte geodatenbasierte Kalibrierung erfolgen (Kapitel 4.4.2): Die Nutzer:innen bewegen und verschieben dabei die projizierte virtuelle 3D-Umgebung auf dem Bildschirm so, dass sie mit der realen Sicht der tatsächlichen Umgebung möglichst übereinstimmt. Die globale Kamera-Pose wird dadurch kalibriert.
- **Lokales Visual-Inertial-Tracking-System**: Sowohl während als auch im Anschluss an den manuellen Kalibrierungsprozess läuft im Hintergrund ständig ein lokales Visual-Inertial-Tracking-System (siehe Kapitel 3.1.3), um fortlaufende Kamerabewegungen im lokalen Raum zu detektieren. Dadurch ist durchwegs – auch nach Abschluss der benutzergesteuerten Kalibrierung – eine stabile AR-Projektion von georeferenzierten AR-Inhalten ohne Ruckeln oder Drift während der gesamten AR-Session möglich.

Der hier beschriebene grobe Ablauf der globalen nutzergesteuerten Registrierungsmethode ist in Abbildung 4.2 in Form eines Sequenzdiagramms dargestellt: Bevor die Nutzer:innen eine manuelle Kalibrierung der Geräte-Pose durchführen, werden zunächst durch einen *Global Pose Controller* eine grob geschätzte globale Orientierung und Position von einem *IMU-Controller* und einem *GNSS-Controller* abgefragt und der mobilen AR-Komponente (*AR View Controller*) bereitgestellt. Auf dieser Basis können entsprechende Geodatenmodelle (*GeoTiles*) bei einem (*GeoTile Loader*) angefragt, geladen und zur weiteren Verarbeitung zur Verfügung gestellt werden.

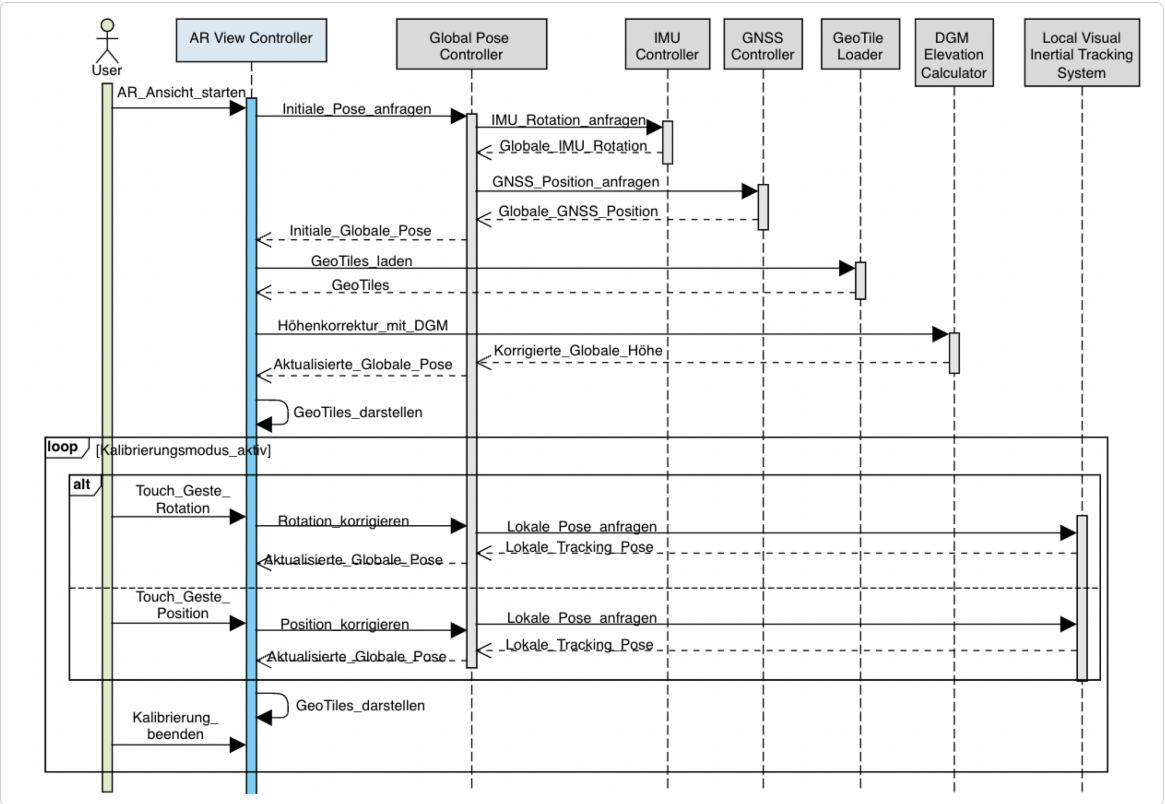


Abbildung 4.2 Sequenzdiagramm des globalen Registrierungs- und Tracking-Systems.

Die Gerätehöhe an der grob geschätzten globalen Position kann anschließend mithilfe eines digitalen Geländemodells präzisiert werden. Dazu wird in einem *DGM Elevation Calculator* der vertikale Positionswert an der geschätzten GNSS-basierten 2D-Geo-Koordinate aus einem geladenen digitalen Geländemodell extrahiert und somit die globale Höhe kalibriert. Hierbei wird berücksichtigt, dass sich das mobile AR-System nicht direkt auf Geländehöhe, sondern in den Händen der Nutzer:innen befindet. Auf die bestimmte Geländehöhe wird daher pauschal eine weitere Höhenkalibrierung von 1,5 m addiert.

Die geladenen GeoTiles können schließlich auf Basis der initial geschätzten globalen Orientierung und Position in der AR-Ansicht dargestellt werden. Die manuelle Kalibrierung der Geräte-Pose erfolgt daraufhin separat entweder über mobile Touch-Gesten zur Steuerung der Rotation oder durch Touch-Gesten zur Kalibrierung der Position. Die Nutzer:innen initiieren die Interaktionsgesten jeweils über die AR-Ansicht und der *Global Pose Controller* übersetzt die Touch-Gesten auf dem Bildschirm mittels mathematischer Projektionsalgorithmen in eine entsprechende Korrektur der globalen Pose (siehe Kapitel 4.4.2). In jedem Fall werden die manuellen Korrekturen mit

einer aktuellen lokalen Geräte-Pose kombiniert, welche kleinere – unvermeidbare – Kamerabewegungen im Nahbereich beschreibt. Diese Bewegungen werden durch ein *Local Visual Inertial Tracking System* ständig erfasst und mit der manuell kalibrierten globalen Kamera-Pose verrechnet. Auf diese Weise entsteht die finale globale Pose. Auf Basis dieser globalen Pose können die Geodatenmodelle neu gerendert und dargestellt werden. Die Beendigung des Registrierungsverfahrens erfolgt schließlich manuell durch die Nutzer:innen.

Die drei grob beschriebenen Komponenten des Gesamtsystems (Geodatenverarbeitungskette, Geodaten-CMS und Registrierungs- und Tracking-System) sollen im Folgenden nun im Detail vorgestellt werden.

4.2 Geodatenverarbeitungskette

Das Ziel der Geodatenverarbeitungskette besteht darin, 3D-Geodaten, die in diversen Ursprungsformaten vorliegen, in ein einheitliches AR-fähiges Format zu konvertieren. Primäres Ziel dabei ist die Umwandlung der Ursprungsdaten in kleinflächige 3D-Kacheln mit möglichst geringer Dateigröße ohne merkliche Qualitätsverluste. Im Zuge der Kachel-Generierung wird dabei ein eigenes Verfahren und Austauschformat eingesetzt. Dieses ist zwar an den 3D-Tiles-Standard des *Open Geospatial Consortium* [109] angelehnt (z.B. hinsichtlich glTF-Dateiformat), baut jedoch auf individuell entworfenen Verarbeitungsschritten und Datenmodellen auf, die den Spezifikationen der eingesetzten 3D-Geodaten (siehe Kapitel 4.2.1) optimal Rechnung tragen können und nur die notwendigen Anforderungen berücksichtigen.

Im Folgenden sollen zunächst die im Kontext des entwickelten Registrierungsverfahrens eingesetzten 3D-Geodatenmodellarten vorgestellt und deren Eigenschaften und jeweiligen Eignungen und Spezifikationen beschrieben werden (Kapitel 4.2.1). Anschließend werden die einzelnen Verarbeitungsschritte zur Umwandlung der 3D-GIS-Rohdaten in AR-fähige 3D-Kacheln im Detail beschrieben (Kapitel 4.2.2).

4.2.1 Digitale 3D-Geodaten

Die Bedeutung dreidimensionaler digitaler Geodaten zur Beschreibung von realen Landschaften hat in den letzten Jahren bei der Nutzung von Geo-Informationssystemen deutlich zugenommen. Waren in klassischen GIS-Anwendungen oft zweidimensionale Geodaten (z.B. topografische Karten, Orthofotos o.ä.) ausreichend, existieren mittlerweile viele Anwendungsszenarien, bei denen die Integration von dreidimensionalen

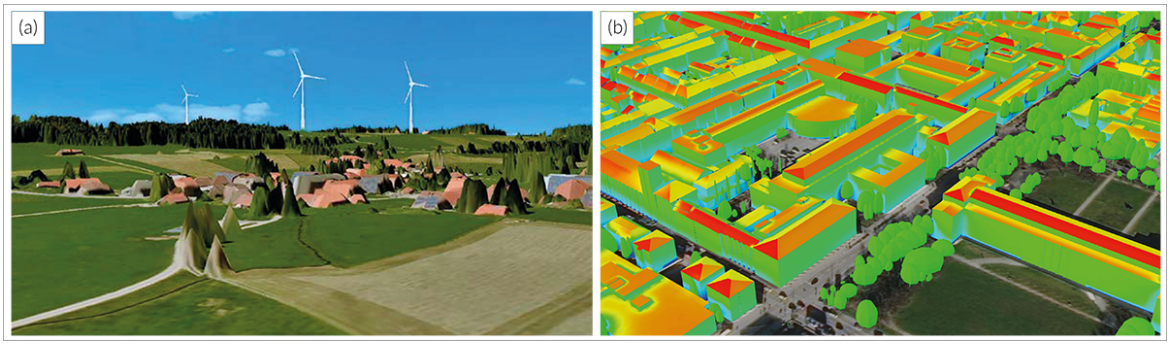


Abbildung 4.3 Mögliche GIS-Anwendungsszenarien durch Integration digitaler 3D-Geodatenmodelle. (a) Nutzung von digitalen texturierten Oberflächenmodellen zur Sichtbarkeitsanalyse von geplanten Windenergieanlagen im Energie-Atlas Bayern [15]. (b) Nutzung von digitalen 3D-Stadtmodellen bei Potenzialanalyse von Dach-Photovoltaikanlagen [17].

GIS-Daten einen deutlichen Mehrwert bietet (z.B. Verschattungsanalyse, Standortplanung, Infrastrukturbau etc.; siehe Abbildung 4.3) [93]. Aufgrund dieses hohen Anwendungspotenzials haben in jüngster Vergangenheit die Vermessungsämter in vielen europäischen Ländern intensiv daran gearbeitet, derartige 3D-Daten zur Modellierung der realen Landschaft in vielfältigen Ausprägungen zu generieren und bereitzustellen [133]. Im Folgenden sollen diejenigen 3D-Geodaten, die in dieser Arbeit als Hilfsmittel zur Gerätekalibrierung eingesetzt werden, genauer vorgestellt und hinsichtlich ihrer Eigenschaften beschrieben werden.

Begriffsdefinitionen

Es existieren verschiedene Arten, um die reale Landschaft der physischen Welt digital als dreidimensionales Modell zu beschreiben und zu repräsentieren. Je nach Anwendungsfall sind dabei verschiedenartige Modellierungen der Umgebung von Bedeutung. Die genauen Definitionen der einzelnen Modellbegriffe unterscheiden sich dabei in der Literatur im Detail [127]. Nach Jäger und Heipke beschreibt ein **Digitales Geländemodell (DGM)** dabei immer das Relief der Geländeoberfläche durch räumliche Koordinaten einer repräsentativen Menge an Geländepunkten [62]. Ein **Digitales Oberflächenmodell (DOM)** beinhaltet im Gegensatz zum DGM hingegen immer die Höhen von auf der Oberfläche befindlichen natürlichen und künstlichen Objekte (z.B. Bauwerke und Vegetation) statt die Höhen des darunter liegenden Reliefs. Der Begriff des Digitalen Höhenmodells kann schließlich als Oberbegriff von DGM und DOM verstanden werden [127].

In urbaner Gegend stellen **3D-Stadtmodelle** eine weitere Möglichkeit zur digitalen dreidimensionalen Repräsentation der realen Umgebung dar. In solchen Stadtmodellen

werden in der Regel nicht nur die 3D-Geometrien, sondern auch semantischen Eigenschaften urbaner Objekte (z.B. Gebäudeart, Adresse etc.) in einheitlicher Weise und mit unterschiedlichem Detailgrad (*Level of Detail*; *LoD*) erfasst [52].

Für die vorliegende Arbeit sollen ebendiese drei Modellarten (DGM, DOM und Stadtmodell) berücksichtigt werden, um mit deren Hilfe eine visuelle Repräsentation der Nutzerumgebung zu erschaffen und somit eine globale manuelle Gerätekalibrierung zu ermöglichen.

Datengenerierung

Zur Erfassung und Generierung dieser 3D-Geodatenmodelle existieren unterschiedliche Vorgehensweisen. Ältere digitale Geländemodelle basierten oft noch auf einer Analog-Digital-Wandlung von analogen Höhenlinien (Isolinien) als Datengrundlage [38]. Digitale Oberflächenmodelle wurden oftmals durch bildbasierte stereoskopische Auswertung (Photogrammetrie) generiert. Aus Stereobildpaaren werden dazu korrespondierende Bildbereiche gefunden, wobei die Korrelation der beiden Bilder Rückschlüsse auf die Höhe der Oberfläche an dieser Stelle zulässt [38].

Inzwischen bilden in der Regel 3D-Punktwolken aus flugzeugbasierten Laserscans (*Airborne Laserscanning, ALS*) die primäre Datengrundlage für DOM, DGM und 3D-Stadtmodelle. Auf Flugzeugen installierte Laserscanner senden dabei in hoher Frequenz Signale aus, welche an Oberflächen reflektiert werden. Durch präzise Bestimmung der Flugrouten mittels D-GNSS kann durch Registrierung der ersten Signalreflexion (first pulse) an der äußersten Geländeoberfläche ein digitales Oberflächenmodell erfasst werden. Die Erfassung der letzten Signalreflexion auf dem Gelände kombiniert mit einer Nachverarbeitung mit unterschiedlichen Klassifikationsebenen (Bauwerke, Brücken o.ä.) liefert die notwendigen Informationen zur Generierung eines hochaufgelösten digitalen Geländemodells [62, 67].

Die 3D-Punktdaten der Laserscans bilden üblicherweise auch die Grundlage für die Erzeugung von digitalen 3D-Stadtmodellen. Zur präzisen Modellierung der einzelnen Gebäudegeometrien werden dabei noch zusätzlich Gebäudeinformationen (z.B. Gebäudegrundrisse) aus digitalen Liegenschaftskatastern integriert [62]. Außerdem werden bei der Beschreibung virtueller 3D-Stadtmodelle oftmals auch nicht nur die reinen geometrischen Strukturen der Gebäude, sondern auch zusätzliche Semantiken und Beziehungen der modellierten Objekte zueinander miteinbezogen (z.B. im CityGML-Standard) [90].

Datenquellen und Datenverfügbarkeit

Die amtliche Bereitstellung von topographischen und kartographischen Geobasisinformationen liegt in Deutschland klassischerweise in der Zuständigkeit der Länder. In diesem Zusammenhang existiert seit etwa 1990 beispielsweise deutschlandweit das Amtliche Topographisch-Kartographische Informationssystem (ATKIS), in dessen Rahmen diverse digitale Geobasisdaten, u.a. digitale Geländemodelle in unterschiedlicher räumlicher Auflösung, durch das amtliche Vermessungs- und Geoinformationswesen der Öffentlichkeit bereitgestellt werden [53]. Im Zuge technischer Weiterentwicklungen und der Verfügbarkeit neuerer digitaler Vermessungs- und Verarbeitungstechnologien (Photogrammetrie, Lidar) konnten die Landesvermessungsämter mittlerweile ihre Produktpaletten stetig erweitern und bieten inzwischen auch komplexere Geodaten wie (bildbasierte) digitale Oberflächenmodelle oder Stadt- und Gebäudemodelle zur öffentlichen Nutzung an. Durch die föderale Struktur der Vermessungsverwaltung in Deutschland unterscheiden sich jedoch insbesondere diese höherwertigen 3D-Geodatenmodelle (DOM, Stadtmodelle) je nach Bundesland hinsichtlich der genauen Spezifikationen (z.B. Aktualität, räumliche Auflösung) sowie hinsichtlich der Nutzungskosten. Während für die Bereitstellung derartiger Geodatenmodelle in vielen Bundesländern noch flächenabhängige Nutzungsentgelte zu zahlen sind, bieten mittlerweile immer mehr Landesvermessungsämter diese 3D-Geodaten inzwischen im Rahmen von OpenData-Initiativen entgeltfrei zur Nutzung an (z.B. Berlin, Brandenburg oder Thüringen).

Spezifikationen der verwendeten 3D-Geodaten

Die 3D-Geodaten werden von den Landesvermessungsämtern in unterschiedlichen Ausführungen und Quelldateiformaten zur Verfügung gestellt. Abbildung 4.4 zeigt beispielhaft einige Auszüge aus den textbasierten Quelldateien. Abbildung 4.5 gibt einen detaillierten Überblick über die Spezifikationen der hier im Rahmen dieser Arbeit verwendeten 3D-Geodaten. Die eingesetzte Geodatenverarbeitungskette wurde dazu so entworfen, dass sie Geodaten mit diesen Daten-Spezifikationen verarbeiten kann, um Zieldateien mit den in der Abbildung genannten Eigenschaften zu erzeugen.

Das **Digitale Geländemodell** eignet sich als Kalibrierungsmodell lediglich in bergigen Umgebungen bzw. an Standorten mit markanten Geländeerhebungen in Sichtweite. Eine großflächige Modellierung der Umgebung ist hierfür notwendig und eine relative geringe räumliche Auflösung des Modellgitters ausreichend. Da in diesem Fall zudem nur texturlose Modellkacheln eingesetzt werden, entstehen im Falle des DGM finale 3D-Kacheln mit recht geringer Dateigröße. Das DGM ist zudem am ehesten

Abbildung 4.4 Auszüge aus Quelldateien der 3D-Geodatenmodelle. (a) Digitales (untexturiertes) Oberflächenmodelle als .xyz-Datei (ASCII-Format). (b) Digitales 3D-Stadtmodell im CityGML-Format. (c) Texturiertes Oberflächenmodell als 3D-Modell im textbasierten .obj-Format.

zur Schätzung der globalen Höhe des AR-Systems an einer gegebener zweidimensionalen Geo-Koordinate geeignet, da davon auszugehen ist, dass sich das AR-System in der Regel auf der Oberfläche des Geländereliefs befindet und nicht auf dem obersten Punkt der sichtbaren Oberfläche (Baumwipfel, Gebäudedach etc.).

Das **Digitale Oberflächenmodell** mit Modellierung sämtlicher Umgebungsoberflächen (Gebäude, Vegetation) weist eine deutlich höhere Komplexität auf und eignet sich in kleinflächiger Ausdehnung als Kalibrierungsmodell sowohl in ländlichen als auch urbanen Standorten mit markanten Gebäuden oder markanter Vegetation in unmittelbarer Umgebung und Sichtweite. Zu diesem Zweck ist eine möglichst hohe räumliche Modellauflösung von Vorteil, was jedoch – je nach Charakteristik der Umgebung – zu großen Dateigrößen führen kann. Für einen effizienten und flexiblen Einsatz ist daher die Erzeugung von Kacheln mit relativ geringer Ausdehnung angebracht. Um die Nutzerumgebung noch realitätsnaher repräsentieren zu können, sollen für diesen Modelltyp zudem auch texturierte 3D-Kacheln eingesetzt werden können. Die Integration der zusätzlichen Texturen als Bilddateien führt zu einer noch größeren Dateigröße, weshalb für diesen Fall nur sehr kleinflächige Kachelgrößen generiert werden sollen, um als Hilfsmittel bei der Gerätekalibrierung nur die unmittelbare tatsächliche Nutzerumgebung virtuell in der AR-Ansicht abzubilden.

In urbanen Einsatzorten bzw. an Standorten mit markanten Gebäuden in Sichtweite kann ein digitales **3D-Stadtmodell** als Kalibrierungsmodell zum Einsatz kommen. Zu diesem Zweck werden LoD2-Gebäudemodelle genutzt, welche – im Gegensatz zu den Blockmodellen mit einheitlichen Flachdächern der LoD1-Qualitätsstufe – auch unterschiedliche Dachformen abbilden können. Auch für diesen Modelltyp sollen zusätzlich

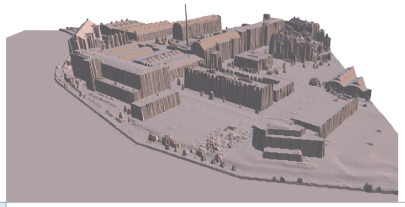
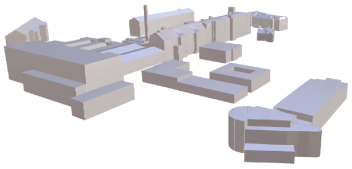
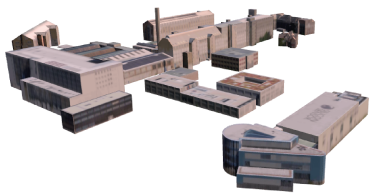
		Digitales Geländemodell (DGM)	
Kurzbeschreibung		3D-Modellierung des Erdoberflächenreliefs ohne Vegetation und Bebauung	
Eignung (Kalibrierung)		Geeignet als Kalibrierungsmodell an Einsatzorten in bergiger Umgebung bzw. mit markanten Bergen in Sichtweite; nicht geeignet in flachen Umgebungen ohne markante Erhebungen in Sichtweite; Auch geeignet zur intialen Schätzung der globalen Höhe des AR-Systems an gegebener 2D-Geokoordinate.	
Quelldatei	Dateiformat: Kachelgröße (Dateigröße):	ASCII (.xyz) 5,0 x 5,0 km (~8 MB)	
Zieldatei	Räumliche Auflösung: Kachelgröße (Dateigröße):	10m 5,0 x 5,0 km (~ 2 MB)	
		Digitales Oberflächenmodell (DOM)	
		ohne Textur (dom)	mit Textur (dom_rgb)
			
Kurzbeschreibung		3D-Modellierung der Erdoberfläche inklusive der darauf befindlichen Objekte (Vegetation, Bebauung)	
Eignung (Kalibrierung)		Geeignet als Kalibrierungsmodell an Einsatzorten mit markanten Gebäuden oder markanter Vegetation (z.B. Bäume) in direkter Umgebung; nicht geeignet zur Abdeckung sehr großer Gebietsareale aufgrund der hohen Komplexität und Datengröße des 3D-Modells	
Quelldatei	Dateiformat: Kachelgröße (Dateigröße):	ASCII (.xyz) 2,0 x 2,0 km (~100 MB)	OBJ (+ Texturen im .jpg-Format) 0,2 x 0,2 km (~150 - 200 MB)
Zieldatei	Räumliche Auflösung: Kachelgröße (Dateigröße):	1m; 2m; 4m 0,5 x 0,5 km (~ 0,5 - 7 MB)* ** je nach Komplexität der Landschaft & räumlicher Auflösung der Zieldatei*	~ 1m 0,2 x 0,2 km (~ 2 - 5 MB)* ** je nach Komplexität der Landschaft*
		Digitales 3D-Stadtmodell (CityModel)	
		ohne Textur (cityModel)	mit Textur (cityModel_rgb)
			
Kurzbeschreibung		3D-Modellierung von urbanen Objekten (Gebäuden) ohne Geländerelief	
Eignung (Kalibrierung)		Geeignet als Kalibrierungsmodell an urbanen Einsatzorten mit Gebäuden in unmittelbarer Nähe oder mit markanten Gebäuden (z.B. Kirchtürme, Hochhaus) in größerer Entfernung.	
Quelldatei	Dateiformat: Kachelgröße (Dateigröße):	CityGML (.gml) 1,0 x 1,0 km (~0,2 - 10 MB) *	OBJ (+ Texturen im .jpg-Format) 1,0 x 1,0 km (~ 50 - 100 MB) *
Zieldatei	Räumliche Auflösung: Kachelgröße (Dateigröße):	LoD2 1,0 x 1, 0 km (~ 0,1 - 2 MB) * ** je nach urbanem Charakter der Umgebung (Anzahl Gebäude)*	LoD2 1,0 x 1,0 km (~ 2 - 15 MB) * ** je nach urbanem Charakter der Umgebung (Anzahl Gebäude)*

Abbildung 4.5 Spezifikationen der eingesetzten 3D-Geodatenmodelle sowie angestrebte Eigenschaften der Zieldateien.

zu texturlosen 3D-Modellen auch texturierte 3D-Kacheln erzeugt und genutzt werden können. Da sich auch weit entfernte markante Gebäude oftmals als Referenzobjekt während der Kalibrierung eignen, ist bei der Nutzung von 3D-Stadtmodellen in vielen Fällen eine großflächige virtuelle Abbildung der Umgebung und somit auch die Wahl relativ großflächiger Kachelgrößen sinnvoll. Die Dateigröße der final generierten Kacheln ist dabei stark abhängig von der Charakteristik der Umgebung (Anzahl der Gebäude).

Im Falle der eingesetzten texturlosen 3D-Geodatenmodelle werden Quelldateien in textbasierten Dateiformaten genutzt. Bei Gelände- und Oberflächenmodellen handelt es sich dabei um die textbasierte Beschreibung eines 3D-Punktegitters (x,y,z) im ASCII-Format. Bei 3D-Stadtmodellen werden Dateien in standardisiertem CityGML-Schema als Ursprungsdateien genutzt. Die geometrische Struktur der Gebäude wird dabei durch eine Beschreibung von Gebäudefassaden und -dächern in Form einzelner Polygone wiedergegeben [22]. Bei den texturierten Geodatenmodellen werden Quelldateien im offenen Wavefront OBJ-Dateiformat verwendet, in Kombination mit Materialdateien (.mtl) sowie Bildtexturen im jpeg-Bildformat.

Die eingesetzten Geodaten verwenden zudem in allen Fällen ein metrisches UTM-Koordinatensystem (*Universal Transverse Mercator*) als globales Bezugssystem. Je nach Bundesland handelt es sich dabei entweder um die Zone 32N (westliche Länder) oder 33N (östliche Länder) des ETRS89-Systems (EPSG 25832 bzw. EPSG 25833).

4.2.2 3D-Geodatenverarbeitung

Für eine geeignete Integration der 3D-Geodatenmodelle innerhalb des AR-basierten Kalibrierungsverfahrens ist eine Umwandlung dieser Basisdaten aus den genannten Ursprungsspezifikationen in renderfähige 3D-Kacheln mit geringerer Dateigröße notwendig. Die hierfür realisierten Verarbeitungsschritte sind in Abbildung 4.6 skizziert.

Lokale Koordinatensystem-Anpassung

In einem ersten Schritt erfolgt für die Eingangsdaten in der Regel zunächst eine lokale UTM-Koordinatensystem-Anpassung. Dabei erfolgt keine Transformation der Eingangsdaten in ein anderes Georeferenzsystem mit anderer Maßstabsskalierung oder anderer 2D-Kartenprojektion. Stattdessen wird hier lediglich eine zweidimensionale Verschiebung der Eingangsdaten vorgenommen, von einem standardisierten globalen UTM-Koordinatensystem O_{UTM} (EPSG 35833 oder 35832) in ein lokales metrisches Bezugssystem O_{Geo} mit dem Koordinatenursprung an einer geografischen Position in der jeweiligen Einsatzregion der Kalibrierungsdaten (siehe Abbildung 4.7). Diese

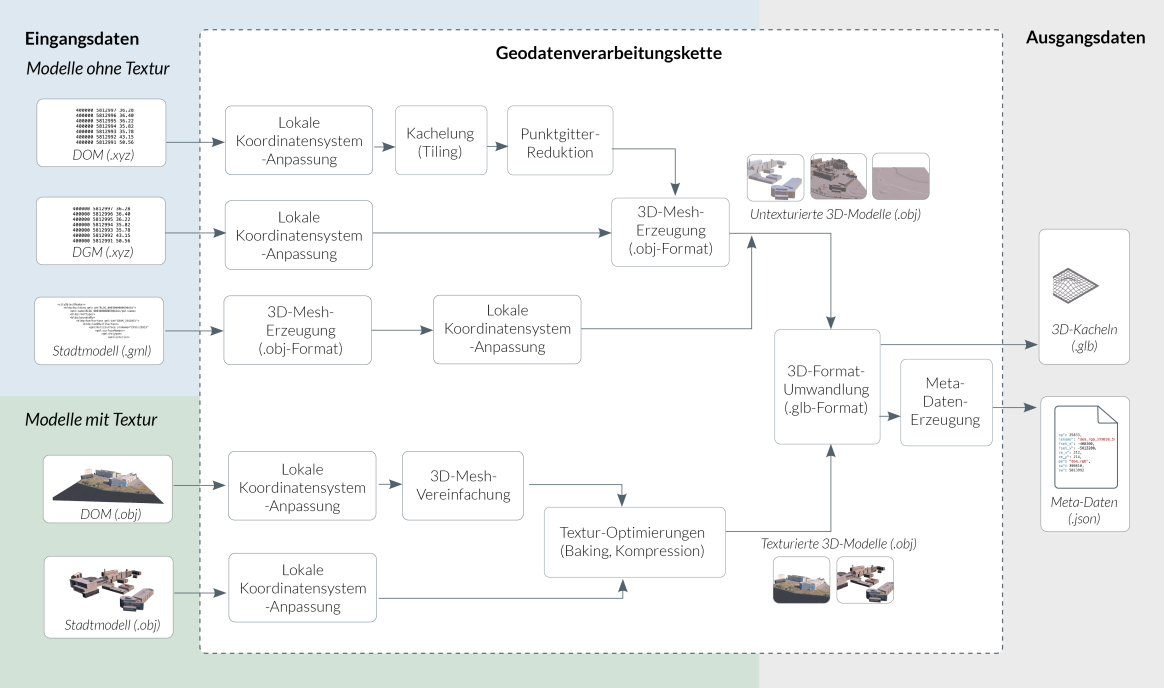


Abbildung 4.6 Verarbeitungsschritte der Geodatenverarbeitungskette zur Umwandlung von Geodatenmodellen aus Quellformaten in AR-fähige uniforme 3D-Kacheln.

statische zweidimensionale Verschiebung t_{offset} der Eingangsdaten ist notwendig, um numerische Instabilitäten bei Anwendung der mathematischen Projektionsalgorithmen (siehe Kapitel 4.4.1) zu verhindern. Außerdem realisieren manche 3D-Dateiformate die Speicherung der 3D-Vertices nur als 32-bit Gleitkommazahlen mit einfacher Genauigkeit (*single precision floating point number*), was nur eine Auflösung von maximal 7-stelligen Dezimalzahlen erlaubt [125]. Die eingesetzten globalen UTM-Koordinaten werden hingegen in ihrer ursprünglichen Form durch 7- bis 8-stellige Dezimalzahlen repräsentiert (z.B. eine UTM-Koordinate $pos_{XY} = [399830,5; 5813090,0]$ in Berlin). Durch die zweidimensionale Verschiebung können die Koordinatenwerte in der Regel auf Dezimalzahlen mit maximal 5 Stellen reduziert werden.

Die lokale Koordinatensystem-Anpassung wird für alle Geodatenmodelle realisiert, sobald diese entweder im .xyz-Format oder .obj-Format vorliegen, was in der Regel gleich zu Beginn der Verarbeitungskette der Fall ist. Lediglich im Falle der Stadtmodelle ist zunächst eine 3D-Mesh-Erzeugung ins .obj-Format aus dem .gml-Quellformat notwendig. Sowohl .xyz- als auch .obj-Formate können problemlos als maschinenlesbare Textdateien geöffnet, analysiert und verändert werden. Innerhalb der Textdatei können somit alle UTM-Koordinaten per individuellem Skript identifiziert und durch neue Koordinaten unter Berücksichtigung der gewünschten zweidimensiona-

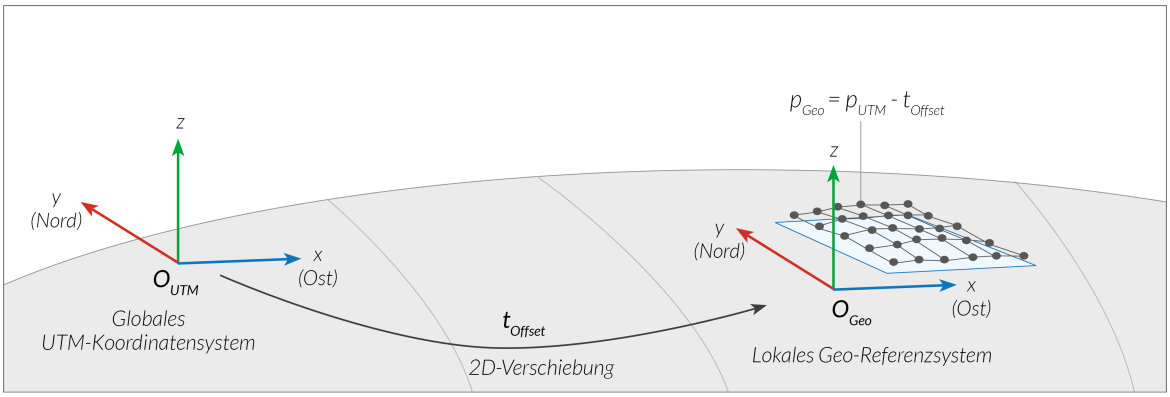


Abbildung 4.7 Die lokale Koordinatensystem-Anpassung führt zu einer zweidimensionalen Verschiebung t_{offset} der Koordinaten aus dem Ursprungs-UTM-Koordinatensystem O_{UTM} in ein lokales metrisches Bezugssystem O_{Geo}. Die Verschiebung reduziert den Wertebereich der Koordinaten und vermeidet damit numerische Instabilitäten und Inkompatibilitäten.

len Verschiebung ersetzt werden. Die aktualisierte Textdatei kann schließlich wieder in gleichem Ursprungsformat gespeichert werden. Da alle eingesetzten Basisdaten in UTM-Koordinatensystemen als Ausgangssystem vorliegen, erfolgt dieser Verarbeitungsschritt für alle Eingangsdaten gleichermaßen.

Kachelung

Die Bereitstellung der Eingangsdaten erfolgt in der Regel bereits in Form von Kacheln mit einer festgelegten fixen Kachelgröße. Die Größe der Kacheln ist jedoch in vielen Fällen nicht geeignet für eine effiziente mobile 3D-Visualisierung. Dies betrifft insbesondere die hochaufgelösten 3D-Oberflächenmodelle, welche als großflächige Kacheln letztendlich 3D-Modelle mit sehr großer Dateigröße (z.B. Dateigrößen von über 100 MB für 2x2 km Kacheln) hervorrufen würden. Für ein effizientes und flexibleres Handling der Kalibrierungsmodelle werden diese großflächigen Modelle daher in Kacheln mit kleinerer Ausdehnung und geringerer Dateigröße geteilt (siehe Abbildung 4.8(a)).

Die Realisierung dieses Verarbeitungsschrittes erfolgt im Falle der digitalen Oberflächenmodelle, indem für jede Eingangsdatei zunächst auf Basis der gewünschten Zielkachelgröße die gewünschte Anzahl neuer Kacheln sowie deren jeweilige geografischer Position und Ausdehnung bestimmt werden. Für jede neue Kachel kann nun eine leere Datei angelegt werden. Anschließend können die Eingangsmodelle in textbasierter Form geöffnet und zeilenweise für jeden Gitterpunkt auf Basis dessen Position entschieden werden, in welche der neu angelegten Kacheln der Punkt geschrieben werden muss. Am Ende enthalten die neuen Kachel-Dateien alle 3D-Gitterpunkte, die im Abdeckungsgebiet der Kachel liegen.

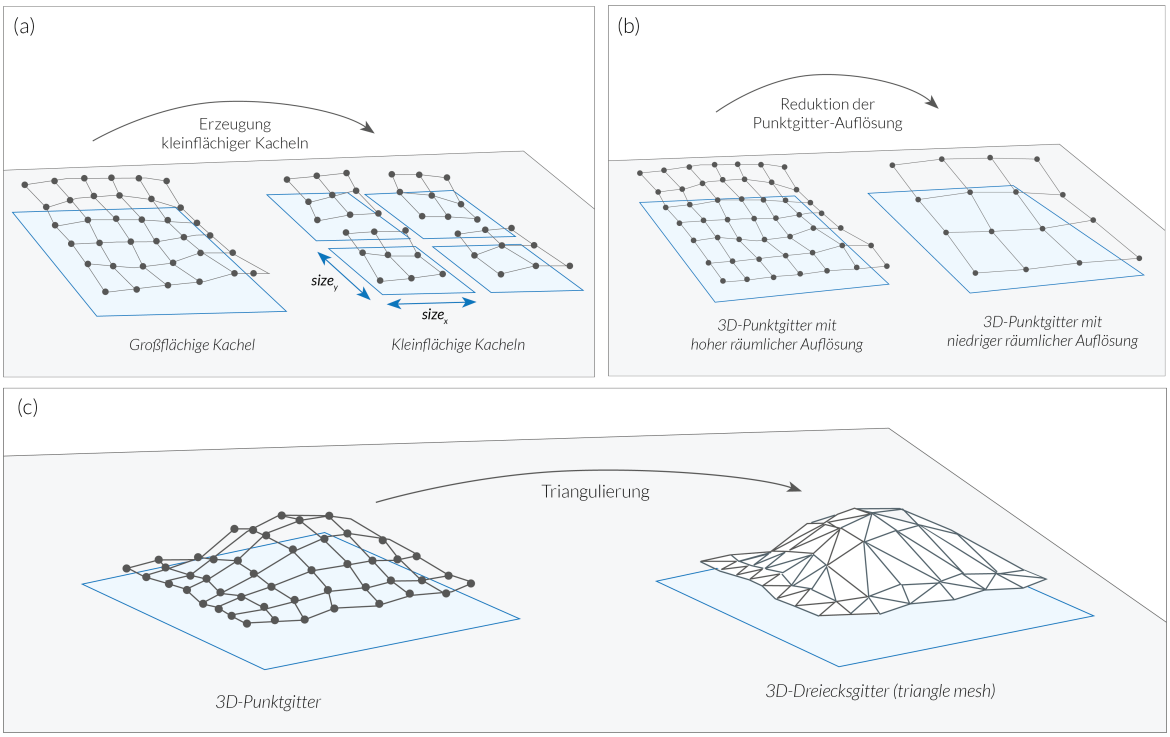


Abbildung 4.8 (a) Großflächige Kacheln werden in mehrere Kacheln mit kleiner Ausdehnung aufgesplittet (Kachelung). (b) Eine Reduktion der 3D-Punktgitterstruktur führt zu einer geringen räumlichen Auflösung und damit geringen Dateigröße der Kacheln. (c) Die Modelloberflächen werden mittels Triangulierung durch Dreiecksnetze approximiert (Meshing).

Punktgitter-Reduktion

Nicht nur die Flächenausdehnung der Eingangskacheln, sondern auch die Gitterstruktur der 3D-Punkte der Oberflächenmodelle entspricht in der ursprünglichen Form nicht vollständig der gewünschten räumlichen Auflösung. Mit dem Ziel, die Dateigröße der Kalibrierungsmodelle zum Zwecke einer effizienten mobilen AR-Visualisierung möglichst gering zu halten, werden aus den hochaufgelösten 3D-Oberflächenmodellen (1 m räumliche Auflösung) auch Modelle mit reduzierter Gitter-Auflösung (2 m, 4 m) generiert (siehe Abbildung 4.8(b)). Auf diese Weise können je nach Einsatzzweck Oberflächenmodelle mit unterschiedlicher Auflösung – und damit Genauigkeit – geladen und dargestellt werden. So können beispielsweise zur virtuellen Modellierung der Umgebung in größerer Entfernung Kacheln mit geringer Auflösung und in unmittelbarer Umgebung Kacheln mit höherer Auflösung eingesetzt werden.

Die Reduktion der räumlichen Auflösung der digitalen Oberflächenmodelle erfolgt ebenfalls durch eine textbasierte Verarbeitung der erzeugten Kacheln, und zwar indem je nach gewünschtem Grad der Reduktion nur bestimmte Gitterpunkte in eine neue

Datei übernommen und übrige Gitterpunkte eliminiert werden. Soll beispielsweise eine Gitterauflösung von 2 m erreicht werden, kann einfach jeder zweite Gitterpunkt aus der Quelldatei (1 m Auflösung) gelöscht bzw. bei der Generierung einer neuen Datei übersprungen werden. Dies ist möglich, da die Gitterpunkte in der Datei in nach räumlicher Position sortierter Form vorliegen.

3D-Mesh-Erzeugung

Damit dreidimensionale Oberflächen mithilfe einer üblichen 3D-Render-Engine darstellbar sind, müssen diese als Dreiecksgitter (*triangle mesh*) vorliegen [91]. Die texturlosen Oberflächen werden jedoch ursprünglich entweder als 3D-Punktegitter (.xyz-Format) oder im Falle des 3D-Stadtmodells durch Polygonflächen (CityGML-Format) beschrieben. Die auf diese Weise vorgegebenen Oberflächen müssen daher zunächst durch eine Menge kleiner zusammenhängender Dreiecke (Dreiecksnetz) approximiert werden (*meshing*). In dieser Arbeit wird dazu mithilfe der Delaunay-Triangulierung die Oberfläche als *Triangulated Irregular Network (TIN)* approximiert [46] (siehe Abbildung 4.8(c)). Im Gegensatz zum *Triangulated Regular Network*, bei dem aus gegebenen Rasterpunkten ein gleichförmiges Dreiecksgitter gebildet wird, wird bei der TIN-Erzeugung die Tatsache ausgenutzt, dass natürliche Landschaften oftmals aus sich langsam verändernden Bereichen bestehen, welche mit einer geringeren Anzahl an Dreiecken approximiert werden können. Dies führt zu einer geringeren Anzahl an zu speichernden Polygonen und damit zu einer kleineren Dateigröße des resultierenden 3D-Modells. Als Zielformat zur Speicherung der generierten Meshes wird das Wavefront OBJ-Format genutzt [29].

3D-Mesh-Vereinfachung

Die eingesetzten texturierten 3D-Modelle liegen in ihrer Ausgangsform bereits als renderfähige Meshes mit Texturdateien vor. Im Falle der texturierten Oberflächenmodelle liegen jedoch Oberflächennetze vor, die eine sehr hohe Anzahl an Dreiecken (*triangles*) besitzen und damit viel dichter modelliert sind als für die beabsichtigte Anwendung erforderlich. Dies ist auf die Art der Erfassungsmethode der texturierten Oberflächenmodelle durch flugzeugbasierte Laserscans zurückzuführen, welche 3D-Punktgitternetze mit sehr hoher räumlicher Auflösung erzeugen. Eine Integration derartig komplexer Modelle mit mehreren Millionen Triangles wäre problematisch auf mobilen Endgeräten mit eingeschränkter Rechenkapazität. Mithilfe geeigneter Algorithmen kann die Oberflächenstruktur iterativ so vereinfacht werden, dass Materialeigenschaften (z.B. Textur-Koordinaten) sowie die grundsätzliche Oberflächenstruktur erhalten bleiben, während die Anzahl an Polygonen und damit die Komplexität der Mesh-Struktur

deutlich reduziert wird – beispielsweise durch den Algorithmus von Garland und Heckbert [47]. Beim Einsatz dieser Vereinfachungsmethode mit texturierten Oberflächenmodellen des Landes Berlin im Rahmen der Evaluation dieser Arbeit (siehe auch Kapitel 5) zeigte sich, dass durch deren Einsatz die Zahl der Polygone innerhalb der Oberflächenstruktur um etwa das 10-fache verringert werden kann, während die Qualität der Oberflächenmodellierung dennoch ausreichend bleibt für den Einsatzzweck als Kalibrierungsmodell. Die Oberflächenstruktur wirkt durch die Vereinfachung an planaren Flächen mitunter sogar deutlich verbessert und geglättet – im Gegensatz zu den oftmals deformierten Flächen im Original-Mesh (siehe Abbildung 4.9(a)). Die Dateigröße des Oberflächenmodells wird dadurch ebenfalls annähernd um den gleichen Faktor reduziert.

Im Falle der texturierten 3D-Stadtmodelle ist eine zusätzliche Mesh-Vereinfachung nicht notwendig, da diese Mesh-Strukturen unmittelbar auf den planaren LoD2-Gebäudemodellen basieren und keine unnötigen zu dichten Netzstrukturen aufweisen.

Textur-Optimierung

Für eine Reduktion des Speicherbedarfs der 3D-Modelle und zum Zwecke eines performanteren Einsatzes der texturierten Modelle innerhalb der AR-Anwendung sind weitere Optimierungsschritte in Bezug auf die Texturierung der Modelle sinnvoll. Dazu zählt zunächst eine Vereinfachung der Modell-Texturierung durch eine Zusammenfassung mehrerer einzelner Texturbilder innerhalb einer 3D-Kachel zu einer einzigen großen Texturdatei mittels ***texture baking***. Dieser Schritt ist insbesondere bei der Verarbeitung der originalen 3D-Stadtmodelle angebracht. Diese Modelle bestehen unter Umständen aus mehreren tausend einzelner Texturen, sofern eine einzelne Textur für jedes Gebäude bzw. jede Hausfassade o.ä. angelegt wurde. Der Einsatz derartiger Modelle würde zu sehr langen Ladezeiten und einer schlechten Darstellungsperformance innerhalb der Client-Anwendung führen. Durch die Generierung einer einzigen neuen Textur und Übertragen aller Texturinformationen aus dem Originalmodell in die neue Textur kann dieses Problem behoben werden.

Ein weiterer wesentlicher Verarbeitungsschritt zur Reduktion des Speicherbedarfs der 3D-Modelle kann durch **höhere Bildkompressionsraten** der Texturbilder erreicht werden. Die verwendeten Bilddateien zur Texturierung der 3D-Modelle besitzen – je nach Modelltyp – einen mehr oder weniger großen Anteil an der Gesamtdateigröße des Modells: Im Falle der texturierten Oberflächenmodelle mit aufwändig modellierter geometrischer Struktur und dementsprechend hohem Speicherbedarf zur Beschreibung der reinen Oberflächenstruktur machen die Texturdateien lediglich einen Anteil von

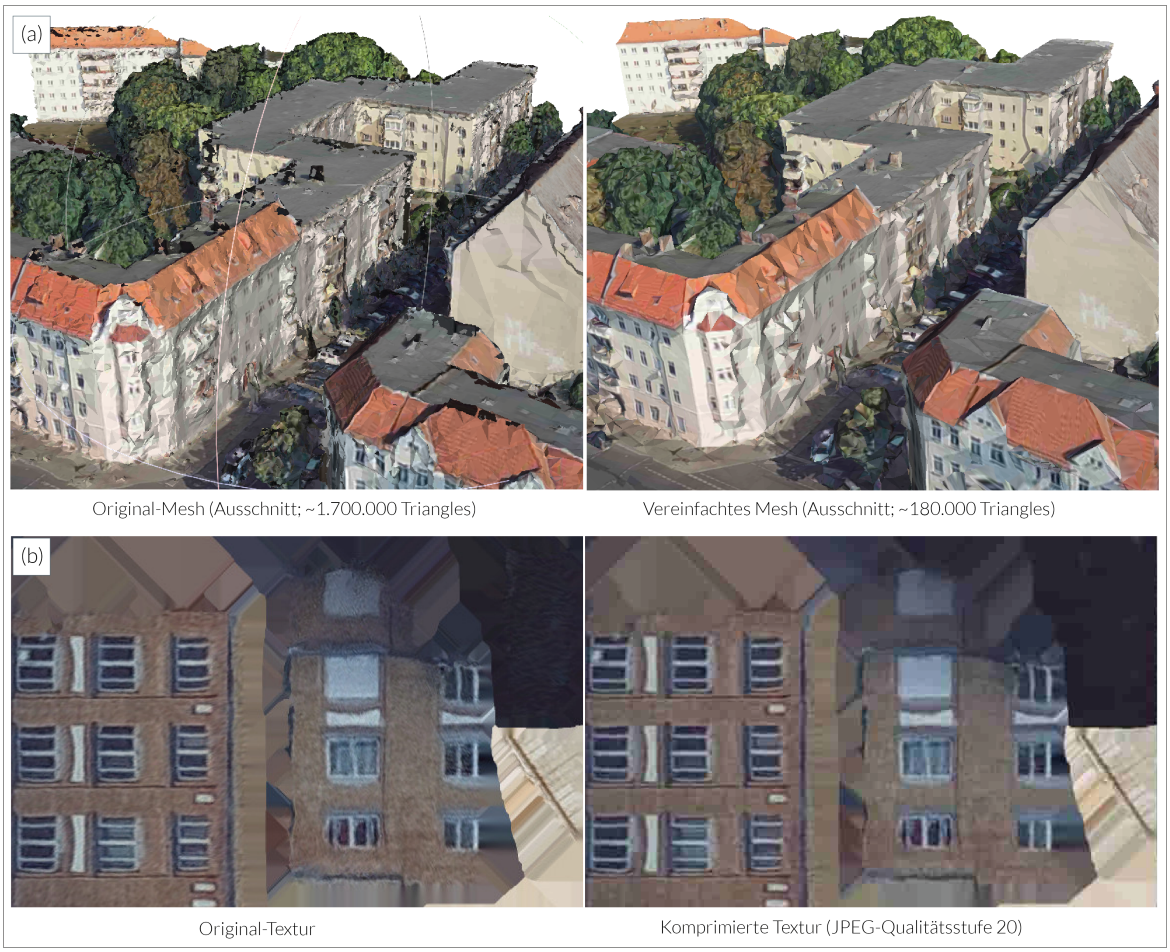


Abbildung 4.9 (a) Durch Vereinfachung der Oberflächenstruktur kann die Polygonanzahl des Original-Oberflächenmodells deutlich reduziert werden, während die Struktur der Oberfläche grundsätzlich erhalten bleibt. (b) Eine hohe jpeg-Kompression der Original-Texturen führt zu einer deutlichen Reduktion der Dateigröße bei kaum sichtbarem Qualitätsverlust.

etwa 10-15% der Gesamtgröße aus – bezogen auf die Originaldatei. Im Falle des 3D-Stadtmodells, das mit einer deutlich geringeren Polygonanzahl und somit geringerem Speicherbedarf modelliert werden kann, ist etwa 95% der Gesamtgröße einer Kachel auf die Bilddateien zur Texturierung zurückzuführen. Dennoch kann in jedem Fall durch eine stärkere Kompression dieser Bilddateien die Gesamtdateigröße jeder Kachel merklich reduziert werden, ohne dabei deutlich sichtbare Qualitätseinbußen hinzunehmen (siehe Abbildung 4.9(b)). Die Reduktion wird in dieser Arbeit erreicht durch das erneute Abspeichern jeder Bildtextur als jpeg-Bilddatei mit Kompressionsfaktor bzw. Qualitätsstufe 20.

3D-Format-Umwandlung

Die finalen 3D-Modelle sollen möglichst in einem 3D-Dateiformat ausgegeben werden, welches für eine performante Nutzung innerhalb einer mobilen AR-Anwendung gut geeignet ist. Das glTF-Dateiformat bietet für diese Einsatzzwecke gute Voraussetzungen. Eine erste Spezifikation dieses von der *Khronos Group* entwickelten quelloffenen Grafikformats wurde 2014 vorgestellt [118]. Seitdem ist durch einige Erweiterungen und Weiterentwicklungen mit der Spezifikation glTF 2.0 [71] ein modernes Dateiformat entstanden, welches durch eine sehr kleine Dateigröße eine effiziente Übertragung der 3D-Modelle ermöglicht, während der Verarbeitungsaufwand zur Laufzeit (Entpacken und Laden der 3D-Modelle) dennoch – im Vergleich zu anderen gängigen 3D-Formaten – gering bleibt [83]. So hat sich dieses 3D-Format mittlerweile als effizientes Dateiformat etabliert, um auch großflächige 3D-Stadt- oder -Landschaftsmodelle in mobilen Anwendungen nutzbar zu machen [125]. In dieser Arbeit wird die binäre Form des glTF-Formats (.glb) genutzt, die in die Spezifikation von glTF 2.0 integriert wurde [70]. In einem letzten Arbeitsschritt werden daher die zuvor erzeugten und im .obj-Format vorliegenden – texturierten und untexturierten – 3D-Modelle in das binäre .glb-Format umgewandelt. Verwendete Texturdateien werden in diesem Format nicht mehr als externe Bilddateien referenziert, sondern direkt in die .glb-Datei integriert. Für jede final entstandene Kachel liegt auf diese Weise nur eine einzige Datei zur Wiedergabe des 3D-Modells vor.

Meta-Daten-Erzeugung

Neben diesen 3D-Kacheln im binären .glb-Format werden als finale Ausgabedateien der Geodatenverarbeitungskette zusätzlich Meta-Dateien zur Beschreibung der einzelnen

```
1  {
2      "filename": "cityModel_rgb_398000_5813000.glb",
3      "type": "cityModel_rgb",
4      "epsg": 25833,
5      "offset_x": -400300,
6      "offset_y": -5813200,
7      "size_x": 1000.0,
8      "size_y": 1000.0,
9      "x_sw": 398000.0,
10     "y_sw": 5813000.0
11 }
```

Abbildung 4.10 Beschreibung der Meta-Daten einer 3D-Kachel (GeoTile) im JSON-Format am Beispiel einer Kachel eines Stadtmodells

3D-Kacheln generiert. Der Zweck dieser Dateien besteht darin, notwendige Zusatzinformationen zu den generierten 3D-Mesh-Kacheln zu speichern, um diese auf einfache und schnelle Art und Weise zu organisieren, zu verwalten und für Client-Anwendungen in geeigneter Form auffindbar zu machen. Zu jeder einzelnen 3D-Kachel wird dazu eine Datei im JSON-Format (*JavaScript Object Notation*) angelegt (siehe Abbildung 4.10). Neben einer eindeutigen Kachel-ID und der Referenz auf den Dateinamen der zugehörigen 3D-Kachel werden in diesem JSON-Format die georäumliche Position, die räumliche Ausdehnung und Auflösung, das genutzte geografische Bezugssystem sowie die Art des Geodatenmodells (z.B. DGM, DOM, Stadtmodell) gespeichert. Eine genauere Beschreibung des Datenmanagements der eingesetzten Kalibrierungsdaten erfolgt in folgendem Kapitel 4.3.

4.3 Geodaten-CMS

Die geodatenbasierte Kalibrierungsmethode wurde so entworfen, dass sie einen Einsatz von Geodatenmodellen unterschiedlicher Art mit jeweils unterschiedlichen Eigenschaften realisieren kann. So sollen – je nach Einsatzzweck bzw. Charakteristik der aktuellen Umgebung – Geodatenmodelle unterschiedlicher Art und in flexiblem Umfang geladen und dargestellt werden können. Für eine derartig flexible Datenintegration ist ein geeignetes Datenhandling und -management notwendig. Um Geodatenmodelle sowohl persistent speichern als auch flexibel nach Bedarf dem mobilen AR-Client zur Verfügung stellen zu können, bietet sich der Einsatz einer Server-Client-Architektur an. Die hierfür entworfenen Schnittstellen und Abläufe sollen in diesem Kapitel vorgestellt werden.

Die Beschreibung der einzelnen Geodatenmodelle erfolgt auf Basis der Meta-Daten, die am Ende der Verarbeitungskette für jede einzelne 3D-Kachel (*GeoTile*) im JSON-Format erzeugt werden (siehe Abbildung 4.10). Für die Datenhaltung der Kacheln wird dabei nicht auf ein standardisiertes Geodaten-Austauschformat, z.B. GeoJSON [61], zurückgegriffen. Stattdessen wird für die selbst generierten Kacheln ein eigenes einfaches Datenmodell entworfen, welches deren individuelle Beschaffenheit und Eigenschaften berücksichtigt. Da die Kacheln immer in rechteckiger Form erzeugt werden sowie immer parallel zu den nach Osten bzw. Norden verlaufenden x- und y-Achsen des UTM-Koordinatensystems verlaufen, kann die geografische Position und Ausdehnung jeder Kachel auf sehr einfache Art und Weise eindeutig und kompakter beschrieben werden: Hierfür wird lediglich die vertikale und horizontale Ausdehnung in Metern sowie die geografische Position des südwestlichen Randpunktes der Kachel gespeichert. Die erfassten Meta-Daten für eine Kachel sind im Einzelnen:

filename (string)
Verweis auf den eindeutigen Dateinamen der zugehörigen generierten 3D-Kachel im glb-Format

type (enum <dgm, dom, dom_rgb, citymodel, citymodel_rgb>)
Art des Geodatenmodells: Geländemodell (dgm), untexturiertes/texturiertes Oberflächenmodell (dom/dom_rgb) oder untexturiertes/texturiertes Stadtmodell (cityModel/cityModel_rgb)

epsg (int)
EPSG-Code des verwendeten UTM-Koordinatenreferenzsystems

offset_x (int)
X-Koordinate der lokalen zweidimensionalen Koordinatensystem-Anpassung t_{offset}

offset_y (int)
Y-Koordinate der lokalen zweidimensionalen Koordinatensystem-Anpassung t_{offset}

size_x (float)
Ausdehnung der Kachel in Metern entlang der x-Achse des UTM-Systems

size_y (float)
Ausdehnung der Kachel in Metern entlang der y-Achse des UTM-Systems

resolution (int)
Räumliche Auflösung (z.B. 1 m, 2 m) der ursprünglichen 3D-Gitterpunkte der Kachel in Metern (nur bei dom, dom_rgb und dgm)

x_sw (float)
X-Koordinatenwert des südwestlichen Randpunktes der Kachel

y_sw (float)
Y-Koordinatenwert des südwestlichen Randpunktes der Kachel

Die Geodatenverarbeitungskette erzeugt für jede Kachel zwei Dateien (siehe Kapitel 4.2.2): Eine binäre 3D-Grafikdatei (.glb), welche das 3D-Geodatenmodell als 3D-Mesh beinhaltet, sowie eine JSON-Datei zur Beschreibung der Eigenschaften der Kachel. Ein geeignetes *Content Management System* (*CMS*) kann schließlich genutzt werden, um diese Daten zu speichern, zu verarbeiten sowie einer mobilen AR-Client-Anwendung bei Bedarf zur Verfügung zu stellen (siehe Abbildung 4.11):

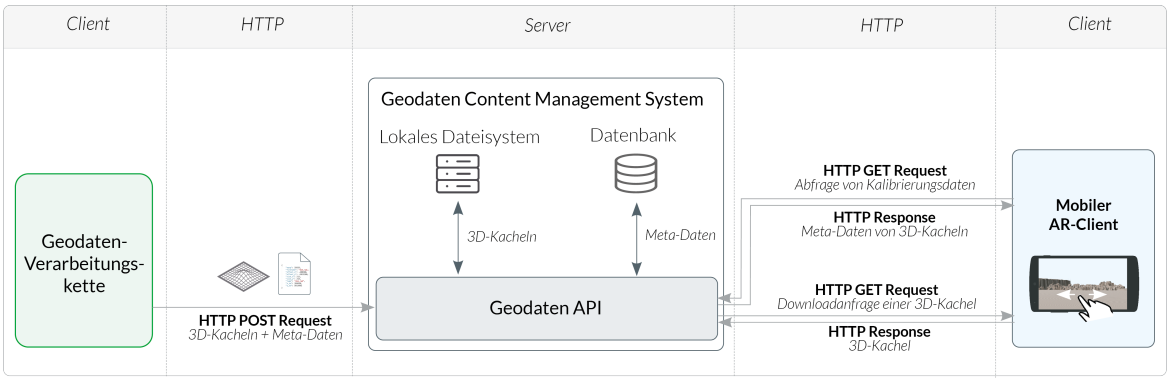


Abbildung 4.11 Client-Server-Architektur zur Speicherung und Abfrage der 3D-Kacheln als Kalibrierungsdateien

- Über einen *HTTP-POST-Request* können die binären 3D-Kacheln gemeinsam mit ihren JSON-basierten Meta-Daten clientseitig von der Komponente der Geodatenerarbeitungskette an eine entsprechende Server-Schnittstelle (*Application Programming Interface; API*) des CMS gesendet werden.
- Im Geodaten-CMS können die Daten serverseitig gespeichert werden. Die binäre 3D-Grafikdatei wird auf einem lokalen Dateisystem abgelegt, während die Meta-Daten der Kachel in einer Datenbank gespeichert werden.
- Ein mobiler AR-Client kann die benötigten Kacheln bei Bedarf über die API des Geodaten-CMS laden. Dazu initiiert der Client zunächst per *HTTP-GET-Request* eine Abfrage der Meta-Daten verfügbarer Kacheln. Auf Basis dieser Meta-Daten kann der Client schließlich Anfragen zum Download einzelner 3D-Kacheln an das Geodaten-CMS richten und die 3D-Modelle einzelner Kacheln downloaden.

Das genaue Vorgehen zum clientseitigen Abrufen der GeoTiles zum Zwecke der Kalibrierung ist in Abbildung 4.12 skizziert: Zunächst initiieren die Nutzer:innen clientseitig eine Abfrage an die Server-API, um Informationen über verfügbare Kalibrierungsmodelle an der aktuellen Position zu erhalten. Der Client kann dabei zusätzlich die gewünschte Modellart sowie den Such-Radius bestimmen. Mittels einer Datenbankabfrage werden schließlich serverseitig alle GeoTiles des passenden Modelltyps abgefragt und die Meta-Daten derjenigen Kacheln an den Client zurückgeschickt, die sich innerhalb des gewünschten Umkreises befinden.

Der Bezug der 3D-Modelle selbst erfolgt in einem zweiten separaten Schritt: Anhand der erhaltenen Meta-Daten der verfügbaren Kacheln kann der Client für jede einzelne gewünschte Kachel eine Download-Abfrage starten. Dazu wird der jeweilige Kachel-Dateiname (*filename*), der jede Kachel als ID eindeutig beschreibt, als Request an

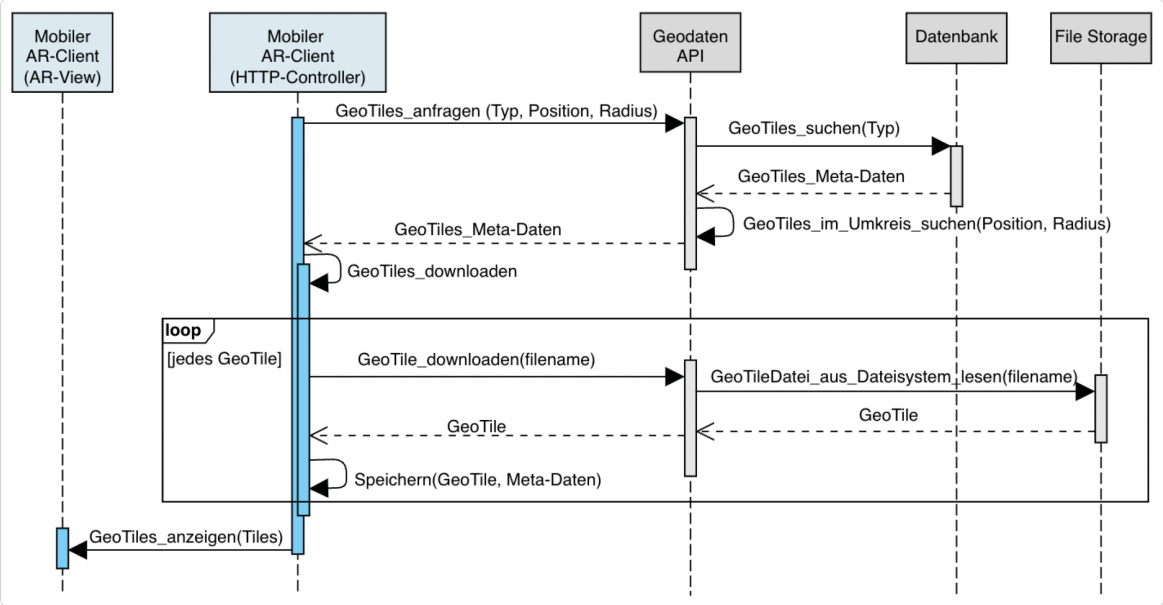


Abbildung 4.12 Sequenzdiagramm zur Abfrage von Kalibrierungsdaten (GeoTiles) durch den mobilen AR-Client

den Server geschickt. Dort kann die entsprechende binäre 3D-Grafikdatei aus dem lokalen Dateisystem gelesen und als *HTTP-Response* an den Client zurückgeschickt werden. Da eine einzelne Kachel mitunter eine Dateigröße von mehreren MB besitzt (siehe Kapitel 4.2.1), kann dieser Download-Vorgang je nach Umfang der angefragten Kalibrierungsdaten und verfügbarer Bandbreite einige Zeit in Anspruch nehmen. Es ist daher sinnvoll, jede Kachel anschließend – gemeinsam mit den zugehörigen JSON-basierten Meta-Daten – auch lokal auf dem mobilen Client abzuspeichern, sodass der Download-Prozess für jede Kachel nur einmalig notwendig ist. Sobald der Download-Vorgang für alle Kacheln beendet ist, können die bezogenen GeoTiles schließlich innerhalb des Clients an die AR-Ansicht übergeben und dort visualisiert werden.

4.4 Globales Registrierungs- und Tracking-System

Eine initiale Darstellung der geladenen Kalibrierungsmodelle in der mobilen AR-Ansicht kann auf Basis der globalen Geräte-Pose erfolgen, die durch IMU- und GNSS-Sensoren erfasst wird. Das Ziel der geodatenbasierten Registrierungsmethode besteht anschließend darin, diese initiale, jedoch sehr unpräzise globale Pose manuell durch Interaktionen zu korrigieren und somit das AR-System zu kalibrieren. Parallel dazu erfolgt eine ständige automatische Berücksichtigung von kleineren lokalen Gerätebewegungen

durch Integration eines lokalen bildbasierten Tracking-Systems. Bevor die einzelnen Interaktionsformen zur Kalibrierung der globalen Position und Orientierung im Detail beschrieben werden, sollen im Folgenden zunächst alle Koordinaten- und Bezugssysteme, die im Gesamtprozess involviert sind, genauer erläutert werden.

4.4.1 Koordinatensystem-Transformationen

Die im gesamten Registrierungsprozesse beteiligten Koordinatensystemen sowie deren Bezug zueinander ist in Abbildung 4.13 skizziert. Dabei werden die folgenden fünf Koordinatensysteme definiert:

1. **Globales UTM-Koordinatensystem O_{UTM}:** Globales geografisches Bezugssystem, in dem die verwendeten Geodaten in ihrer ursprünglichen Form zur Verfügung gestellt werden.

2. **Lokales Geo-Referenzsystem O_{Geo}:** Lokales geografisches Bezugssystem, das sich durch eine zweidimensionale Verschiebung des globalen UTM-Systems in das Anwendungsgebiet der Nutzung ergibt, d.h. der Ursprung des Systems liegt nahe der geografischen Position der Nutzer:innen. Die im Registrierungsprozess eingesetzten Geodaten sind nach der Konvertierung durch die Geodatenverarbeitungskette immer in Bezug auf dieses Referenzsystem definiert.

3. **Lokales VI-Tracking-Koordinatensystem O_{VI}:** Lokales metrisches Bezugssystem, in dem die lokalen Bewegungen des mobilen Endgeräts (Kamera) ständig durch bildbasiertes Visual-Inertial-Tracking erfasst werden.

4. **Kamera-Koordinatensystem O_{Cam} / $O_{CamCalib}$:** Koordinatensystem des mobilen Endgeräts. Im Zuge der manuellen Registrierung wird die Pose des Kamera-Koordinatensystems O_{Cam} hin zu einem kalibrierten Kamera-Koordinatensystem $O_{CamCalib}$ mit präziser globaler Pose korrigiert.

5. **Pixel-Koordinatensystem O_{UV}:** Pixel-basiertes zweidimensionales Koordinatensystem des Kamerabildes auf dem AR-Display.

Alle beteiligten Koordinatensysteme – mit Ausnahme des Pixel-Koordinatensystems – sind als metrische Bezugssysteme mit dem gleichen metrischen Raum (Einheit Meter) definiert. Das **globale UTM-System** O_{UTM} ist ein rechtsdrehendes Koordinatensystem, bei dem die x-Achse immer in Richtung des geografischen Ostens, die y-Achse immer nach Norden und die z-Achse in Richtung des Himmels zeigt. Die Verschiebung des Bezugssystems der eingesetzten Geodaten von dem globalen UTM-System

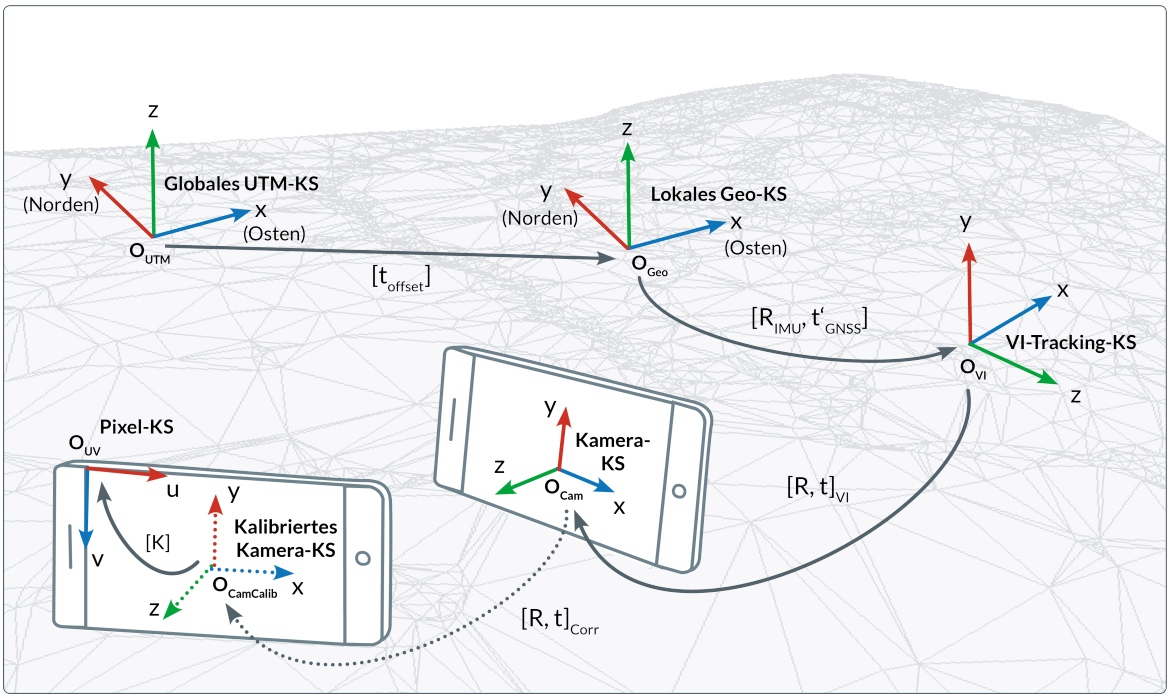


Abbildung 4.13 Koordinatensysteme der globalen Registrierungsmethode: Georeferenzierte Daten sind in Bezug auf ein lokales Geo-Referenzsystem definiert. Ein VI-Tracking-System dient als Bezugssystem aller lokaler Kamerabewegungen. Die nutzergesteuerte globale Registrierung kann als Korrektur der Kamera-Pose in Bezug auf das VI-System betrachtet werden.

in ein **lokales geografisches Referenzsystem** O_{Geo} in die geografische Umgebung der Ziel-Anwendung erfolgt zur Vermeidung von Inkompatibilitäten und numerischen Instabilitäten bei der weiteren Verarbeitung und Darstellung der 3D-Geodatenmodelle (siehe Kapitel 4.2.2). Ziel ist es also, dass sich der Ursprung dieses lokalen Geo-Referenzsystems in der Nähe der geografischen Position der Anwender:innen befindet. Auch dieses lokale geografische Bezugssystem ist gleichermaßen immer nach Osten bzw. Norden ausgerichtet. Da diese Koordinatensystem-Transformation bereits im Rahmen der Daten-Generierung innerhalb der Geodatenverarbeitungskette erfolgt, sind alle verwendeten georeferenzierten Daten immer in Bezug auf ebendieses System definiert, wenn sie innerhalb der mobilen AR-Anwendung eingesetzt werden.

Ein beliebiger dreidimensionaler georeferenzierter Punkt $P_{Geo} = (X_{Geo}, Y_{Geo}, Z_{Geo})^T$ des lokalen geografischen Bezugssystems O_{Geo} kann daher in Bezug auf das globale

UTM-Koordinatensystem O_{UTM} definiert werden als:

$$P_{Geo} = P_{UTM} + t_{offset} = \begin{pmatrix} X_{UTM} \\ Y_{UTM} \\ Z_{UTM} \end{pmatrix} + \begin{pmatrix} x_{offset} \\ y_{offset} \\ 0 \end{pmatrix} \tag{4.1}$$

Dieser Bezug ermöglicht eine einfache Umrechnung von lokalen Koordinaten der verwendeten Geodatenmodelle (z.B. Koordinaten der Mesh-Vertices) zu globalen UTM-Koordinaten.

Das **lokale Tracking-Koordinatensystem** O_{VI} wird durch eine globale Position und Orientierung auf Basis der initialen GNSS- und IMU-Pose-Bestimmung definiert. Das VI-Tracking-System ist ebenfalls als rechtsdrehendes Bezugssystem definiert, jedoch so, dass die y-Achse in Richtung Himmel zeigt. Das heißt, die x-Achse des Systems zeigt ungefähr in Richtung Osten und die z-Achse ungefähr in Richtung Süden. Das System O_{VI} kann initialisiert werden, sobald Sensordaten der GNSS- und IMU-Sensorik verfügbar sind. Angesichts der Ungenauigkeiten dieser Registrierungsmethode ist dieses Bezugssystem jedoch nur unpräzise global ausgerichtet. Nach initialer Schätzung von Position und Ausrichtung durch die GNSS- und IMU-Sensordaten erfolgt zunächst außerdem eine Präzision der vertikalen Positionskomponente des Systems O_{VI} (Gerätehöhe) auf Basis eines digitalen Geländemodells an der geschätzten Geo-Position.

Die initiale GNSS- und IMU-basierte Pose-Bestimmung kann durch eine Rotationsmatrix $R_{IMU} = R_{Init}$ der inertialen Messeinheit und eine translatorische Komponente t_{GNSS} des GNSS-basierten Positionierungssensors beschrieben werden. Die Positionierung erfolgt bei der GNSS-Sensorik bezogen auf das globale Referenzsystem O_{UTM}. Die GNSS-basierte Position des Systems O_{VI} auf das lokale geografische Koordinatensystem O_{Geo} kann daher – bezogen auf die globale Position – zunächst mit einer Translations-Komponente t'_{GNSS} beschrieben werden:

$$t'_{GNSS} = \begin{pmatrix} X'_{GNSS} \\ Y'_{GNSS} \\ Z'_{GNSS} \end{pmatrix} = t_{GNSS} - t_{offset} = \begin{pmatrix} X_{GNSS} \\ Y_{GNSS} \\ Z_{GNSS} \end{pmatrix} - \begin{pmatrix} x_{offset} \\ y_{offset} \\ 0 \end{pmatrix} \tag{4.2}$$

Eine Funktion $DGMElevation(x_{UTM}, y_{UTM})$ soll nun als Operation definiert werden, die an gegebener zweidimensionaler UTM-Position $p_{UTM} = (x_{UTM}, y_{UTM})$ die Gerätehöhe auf Basis des Geländereliefs eines digitalen Geländemodells schätzt. Die Funktion soll dabei auch berücksichtigen, dass das mobile Gerät nicht unmittelbar auf Geländehöhe positioniert ist, sondern in den Händen gehalten wird. Hierfür wird ein pauschaler Versatz von 1,5 m oberhalb der Geländehöhe angenommen. Der vertikale

Wert Z'_{GNSS} der Translations-Komponente t'_{GNSS} kann nun mithilfe dieser Funktion durch einen präziseren Wert Z_{DGM} aktualisiert werden:

$$Z'_{GNSS} = Z_{DGM} = DGMElevation(X_{GNSS}, Y_{GNSS}) \tag{4.3}$$

Ein Translationsvektor t_{Init} kann schließlich auf dieser Basis die initiale globale Translation des Trackingsystems O_{VI} – bezogen auf das lokale geografische Koordinatensystem O_{Geo} – beschreiben:

$$t_{Init} = \begin{pmatrix} X'_{GNSS} \\ Y'_{GNSS} \\ Z_{DGM} \end{pmatrix} = \begin{pmatrix} X_{GNSS} - x_{offset} \\ Y_{GNSS} - y_{offset} \\ Z_{DGM} \end{pmatrix} \tag{4.4}$$

Für jeden Punkt P_{VI}, der bezogen auf das lokale Tracking-Koordinatensystem definiert ist, gilt damit:

$$\begin{pmatrix} P_{VI} \\ 1 \end{pmatrix} = \begin{pmatrix} R_{IMU} & t_{Init} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} P_{Geo} \\ 1 \end{pmatrix} \tag{4.5}$$

Das bildbasierte Tracking-System erfasst lokale Bewegungen der sich bewegenden Kamera mittels Feature-Tracking innerhalb fortlaufender Kameraframes in Kombination mit Messwerten der inertialen IMU-Sensorik (siehe Kapitel 3.1.3). Dieser Visual-Inertial-Tracking-Algorithmus bestimmt somit fortlaufend die relative Position t_{VI} und relative Orientierung R_{VI} der sich bewegenden Kamera O_{Cam} in Bezug auf das Koordinatensystem O_{VI}. Ein beliebiger georäumlicher Punkt P_{Geo}, der im geografischen Koordinatensystem O_{Geo} global definiert ist, kann daher im sich bewegenden **nicht-kalibriertem Kamera-Koordinatensystem** O_{Cam} ausgedrückt werden durch:

$$\begin{pmatrix} P_{Cam} \\ 1 \end{pmatrix} = \begin{pmatrix} R_{VI} & t_{VI} \\ 0 & 1 \end{pmatrix} \left[\begin{pmatrix} R_{IMU} & t_{Init} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} P_{Geo} \\ 1 \end{pmatrix} \right] \tag{4.6}$$

Das primäre Ziel des nutzergesteuerten interaktiven Registrierungsverfahrens ist eine anschließende Kalibrierung dieser initialen, jedoch ungenauen globalen Kamera-Pose. Diese Kalibrierung kann dabei als Translations- und Rotationskorrektur t_{Corr} und R_{Corr} der Kamera-Pose in Bezug auf das Tracking-Koordinatensystem O_{VI} gesehen werden.

Das durch diese Kalibrierung entstehende Bezugssystem soll als **kalibriertes Kamera-Koordinatensystem** $O_{CamCalib}$ definiert werden. Mit gegebenen intrinsischen Kamera-Parametern K können schließlich beliebige 3D-Punkte aus diesem

Bezugssystem in das zweidimensionale **Pixel-Koordinatensystem** O_{UV} des Kamerabildes überführt werden. Ein 3D-Punkt $P_{Cam} = (X_{Cam}, Y_{Cam}, Z_{Cam})^T$, der im initialen nicht-kalibriertem Kamerasystem O_{Cam} definiert ist, kann somit auf folgende Weise auf einen 2D-Bildpunkt $p_i = (u, v)$ in das Pixel-Koordinatensystem des kalibrierten Kamera-System $O_{CamCalib}$ projiziert werden:

$$\lambda \begin{pmatrix} p_i \\ 1 \end{pmatrix} = K \begin{pmatrix} R_{Corr} & t_{Corr} \end{pmatrix} \begin{pmatrix} P_{Cam} \\ 1 \end{pmatrix} \tag{4.7}$$

Auf Basis der hier formulierten Definitionen kann somit jeder beliebige georeferenzierte Punkt – und somit auch jedes beliebige eingesetzte 3D-Geodatenmodell – aus dem Geo-Koordinatensystem O_{Geo} bzw. dem globalen UTM-Koordinatensystem O_{UTM} eindeutig auf einen Bildpunkt in der kalibrierten Kameraansicht der mobilen AR-Ansicht projiziert werden. Auf ebendiese Art und Weise kann schließlich die AR-Visualisierung georeferenzierter Daten in der entwickelten Anwendung realisiert werden.

Die Aufgabe des manuellen Registrierungsverfahrens kann in diesem Kontext darin verstanden werden, mithilfe von Bildschirm-Interaktionen durch die Nutzer:innen entsprechende Rotations- und Translations-Korrekturmatrizen R_{Corr} und t_{Corr} zu formen, welche eine kalibrierte Abbildungsprojektion von georeferenzierten 3D-Punkten auf pixelgenaue Bildpunkte im Kamerabild ermöglichen.

4.4.2 Kalibrierung der globalen Position und Orientierung

Zu Beginn des manuellen Registrierungsprozesses können die eingebetteten virtuellen 3D-Geodatenmodelle nur mit ungefährer Ausrichtung und Skalierung dargestellt werden. Der Grad der Ungenauigkeit der Visualisierung hängt primär von der Genauigkeit der initialen GNSS-/IMU-Pose-Bestimmung ab. Üblicherweise führt dies zu einer horizontalen Verschiebung der AR-Visualisierung der Geodatenmodelle im Kamera-Blickfeld um etwa 10-20 Grad – im Vergleich zur Darstellung der realen Landschaft [95]. Detaillierte Ergebnisse hinsichtlich initialer Genauigkeit sind auch in der Praxis-Evaluation in Kapitel 5 beschrieben.

Das allgemeine Ziel des Registrierungsprozesses ist nun eine möglichst genaue Korrektur dieser inkorrekten virtuellen Repräsentation der Landschaft (Kalibrierung). Zu diesem Zweck stehen verschiedene Touch-Gesten als Interaktionsmethoden zur Verfügung, um die virtuellen Modelle direkt auf dem Bildschirm mit der Ansicht der realen Welt anzugleichen und somit die globale Geräte-Pose zu kalibrieren (siehe Abbildung 4.14):

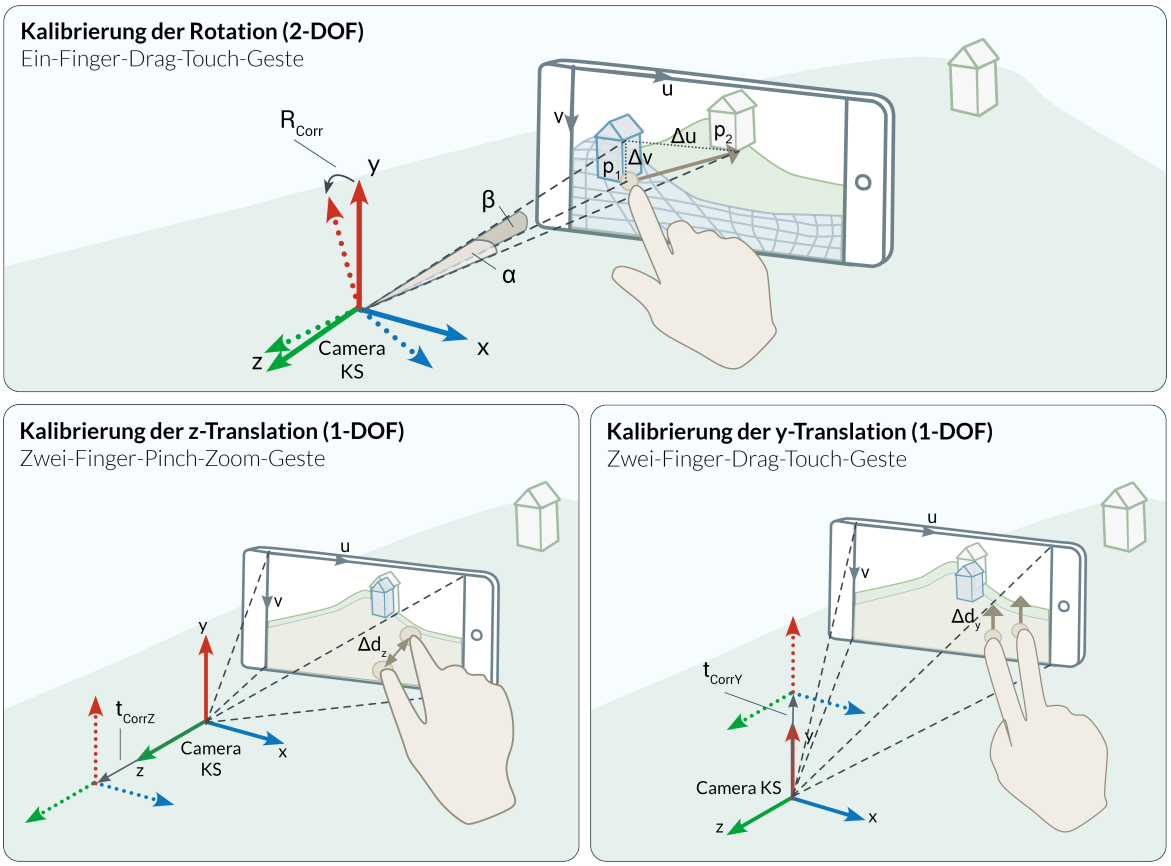


Abbildung 4.14 Interaktionsgesten zur Kalibrierung der globalen Rotation und Translation

- **Kalibrierung der Rotation (2-DOF):** Mithilfe einer *Ein-Finger-Drag-Geste* kann die Orientierung der Kamera um die x- und y-Achse des Kamera-Koordinatensystems kalibriert werden. In der resultierenden AR-Darstellung entspricht dies einer *Drehung* der Geodatenmodelle um die aktuelle Position herum nach links und rechts bzw. nach oben und unten.

- **Kalibrierung der z-Translation (1-DOF):** Mithilfe einer *Zwei-Finger-Pinch-Zoom-Geste* kann die Translation der Kamera entlang der z-Achse des Kamera-Koordinatensystems kalibriert werden. In der resultierenden AR-Darstellung entspricht dies einer *Skalierung* der Geodatenmodelle – also einer Vergrößerung oder Verkleinerung der dargestellten Modelle.

- **Kalibrierung der y-Translation (1-DOF):** Mithilfe einer *Zwei-Finger-Drag-Geste* nach oben oder unten kann die Kamera-Translation entlang der y-Achse des Kamera-Systems kalibriert werden. In der resultierenden AR-Darstellung entspricht dies einer *Verschiebung* der Geodatenmodelle nach oben und unten.

Im Folgenden soll nun beschrieben werden, auf welche Weise durch diese zweidimensionalen Bewegungsgesten auf der Oberfläche des mobilen AR-Bildschirms eine Kalibrierung der globalen Geräte-Pose erreicht wird.

Interaktionsformen zur Kalibrierung der Kamera-Rotation

Bei einer Kalibrierung der Rotation kann die virtuelle Umgebung um die y-Achse und die x-Achse des aktuellen Kamera-Koordinatensystems rotiert werden, indem die Nutzer:innen einen Finger in einer ununterbrochenen Geste über die Bildschirmoberfläche bewegen (Drag-Geste). Der Startpunkt dieser Geste soll dabei als $p_1 = (u_1, v_1)$ und der Endpunkt der Bewegung als $p_2 = (u_2, v_2)$ bezeichnet werden (siehe Abbildung 4.14). Die dabei mit dem Finger zurückgelegte Strecke auf dem Bildschirm in Pixeln kann durch $\Delta u = u_2 - u_1$ entlang der horizontalen u-Achse und $\Delta v = v_2 - v_1$ entlang der vertikalen v-Achse des Pixel-Koordinatensystems definiert werden.

Die durch diese Drag-Geste zurückgelegten Distanzen können in korrespondierende Rotationswinkel α und β umgewandelt werden. Mit Kenntnis der intrinsischen Kamera-Parameter K, die sich aus der Kamera-Brennweite f und dem Kamera-Hauptpunkt $p_c = (u_c, v_c)$ zusammensetzt (siehe Kapitel 2.5), können dazu zunächst zwei Richtungsvektoren $\vec{P_1}$ und $\vec{P_2}$ definiert werden, welche die Richtung vom Ursprung des Kamera-Koordinatensystems O_{Cam} hin zu den Punkten p_1 und p_2 auf dem Kamerabild im 3D-Raum beschreiben. Es wird hierbei vereinfacht angenommen, dass keine Verzeichnung der Kameralinse existiert. Für diese Richtungsvektoren gilt dann:

$$\begin{aligned} \vec{P_1} = K^{-1} p_1 &= \begin{pmatrix} \frac{1}{f} & 0 & -\frac{c_u}{f} \\ 0 & \frac{1}{f} & -\frac{c_v}{f} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{c_u - u_1}{f} \\ \frac{c_v - v_1}{f} \\ 1 \end{pmatrix} \\ \vec{P_2} = K^{-1} p_2 &= \begin{pmatrix} \frac{1}{f} & 0 & -\frac{c_u}{f} \\ 0 & \frac{1}{f} & -\frac{c_v}{f} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{c_u - u_2}{f} \\ \frac{c_v - v_2}{f} \\ 1 \end{pmatrix} \end{aligned} \tag{4.8}$$

Eine Aufspaltung dieser dreidimensionalen Richtungsvektoren jeweils auf einzelne Richtungsvektoren innerhalb der zweidimensionalen horizontalen x-z-Ebene bzw. der vertikalen y-z-Ebene ermöglicht schließlich eine Beschreibung von horizontalen und vertikalen Rotationswinkeln. Für die horizontale Rotation können – normiert auf die Brennweite f – die beiden Vektoren $\vec{r_{u,1}}$ und $\vec{r_{u,2}}$ gebildet werden:

$$\vec{r_{u,1}} = f \begin{pmatrix} \frac{c_u - u_1}{f} \\ 1 \end{pmatrix} = \begin{pmatrix} c_u - u_1 \\ f \end{pmatrix} \qquad \vec{r_{u,2}} = f \begin{pmatrix} \frac{c_u - u_2}{f} \\ 1 \end{pmatrix} = \begin{pmatrix} c_u - u_2 \\ f \end{pmatrix} \tag{4.9}$$

Analog dazu werden die Richtungsvektoren $\vec{r_{v,1}}$ und $\vec{r_{v,2}}$ innerhalb der vertikalen y-z-Ebene definiert:

$$\vec{r_{v,1}} = f \begin{pmatrix} \frac{c_v - v_1}{f} \\ 1 \end{pmatrix} = \begin{pmatrix} c_v - v_1 \\ f \end{pmatrix} \qquad \vec{r_{v,2}} = f \begin{pmatrix} \frac{c_v - v_2}{f} \\ 1 \end{pmatrix} = \begin{pmatrix} c_v - v_2 \\ f \end{pmatrix} \tag{4.10}$$

Der Rotationswinkel α beschreibt schließlich den Winkel zwischen den beiden Richtungsvektoren $\vec{r_{u,1}}$ und $\vec{r_{u,2}}$ und damit die horizontale Rotationskorrektur durch die Touch-Geste der Nutzer:innen in horizontaler Richtung. Der vertikale Rotationswinkel β ist analog dazu als Winkel zwischen den beiden Richtungsvektoren $\vec{r_{v,1}}$ und $\vec{r_{v,2}}$ und somit als vertikale Rotationskorrektur zu betrachten. Ein von zwei Vektoren eingeschlossener Winkel kann über den Arkuskosinus aus der Division des Skalarprodukts der beiden Vektoren durch das Produkt der Vektorlängen berechnet werden. Für die beiden Winkel ergibt sich somit:

$$\begin{aligned} \alpha &= \arccos\left(\frac{\vec{r_{u,1}} \cdot \vec{r_{u,2}}}{\|\vec{r_{u,1}}\| \|\vec{r_{u,2}}\|}\right) \\ \beta &= \arccos\left(\frac{\vec{r_{v,1}} \cdot \vec{r_{v,2}}}{\|\vec{r_{v,1}}\| \|\vec{r_{v,2}}\|}\right) \end{aligned} \tag{4.11}$$

Mithilfe dieser beiden Rotationswinkel kann schließlich eine Rotationsmatrix R_{Corr} gebildet werden, welche die durch die Nutzer:innen durchgeführte korrektive Kamerarotation um die x- und y-Achse beschreibt:

$$\begin{aligned} R_{Corr} &= R_Y(\alpha) R_X(\beta) \\ &= \begin{bmatrix} \cos(\alpha) & 0 & \sin(\alpha) \\ 0 & 1 & 0 \\ -\sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{bmatrix} \end{aligned} \tag{4.12}$$

Eine manuelle Kalibrierung der Kamera-Rotation um die z-Achse wird durch die beschriebene Drag-Touch-Geste nicht realisiert. Eine Kalibrierung der Rotation um diese Achse würde eine zusätzliche Interaktionsgeste erfordern und den Registrierungsprozess damit komplexer gestalten und erschweren. Durch eine relativ präzise und stabile Bestimmung des Gravitationsvektors durch die IMU-Sensorik – insbesondere im ruhigen Zustand – erscheint eine Korrektur dieser dritten Kamera-Achse zudem nicht zwingend notwendig. Dies wird ebenfalls durch die Praxistests (siehe Kapitel 5) bestätigt.

Interaktionsformen zur Kalibrierung der Kamera-Translation

Eine Kalibrierung der globalen Kamera-Position im 3D-Raum kann über zwei verschiedene Touch-Gesten erfolgen: Eine *Zwei-Finger-Pinch-Zoom-Geste* zur Kalibrierung der Kamera-Translation entlang der z-Achse in Kamera-Blickrichtung sowie eine *Zwei-Finger-Drag-Geste* nach oben oder unten zur Kalibrierung der Kamera-Translation entlang der vertikalen y-Achse des Kamera-Koordinatensystems (Höhenkalibrierung).

Eine **Translation der Kamera entlang der z-Achse** des Kamera-Systems O_{Cam} resultiert visuell in einer Vergrößerung oder Verkleinerung der dargestellten Geodatenmodelle in der AR-Ansicht. Zu diesem Zweck wenden die Nutzer:innen auf dem Bildschirm eine Pinch-Zoom-Geste mit zwei Fingern an, d.h. sie bewegen zwei Finger – den Daumen und Zeigefinger – aufeinander zu, um die virtuelle Umgebung zu verkleinern und somit virtuell wegzubewegen. Alternativ können die zwei Finger voneinander wegbewegt werden, um die virtuelle Landschaft zu vergrößern und näherzubringen. Die Länge der Zoom-Geste Δd_z auf der Bildschirmoberfläche in Pixeln kann schließlich mit einem Skalierungsfaktor s_z skaliert werden, um ein Maß für die tatsächliche metrische Verschiebung der Kamera-Position entlang der z-Achse im Metern zu bestimmen. Ein aus dieser Interaktionsgeste resultierender Translationsvektor t_{CorrZ} zur Beschreibung der manuellen Positionsverschiebung des Kamera-Ursprungs $O_{CorrCam}$ in Bezug auf den unkalibrierten Kamera-Ursprung O_{Cam} kann daher wie folgt definiert werden:

$$t_{CorrZ} = \begin{pmatrix} 0 \\ 0 \\ s_z \cdot \Delta d_z \end{pmatrix} \tag{4.13}$$

Der Wert des Skalierungsfaktors s_z ist dabei experimentell zu bestimmen und kann unter Umständen auch variabel gewählt werden, beispielsweise je nach Charakteristik der Umgebung und erwarteter Genauigkeit der initialen GNSS-basierten Positionsbestimmung.

Es ist bei dieser Form der Positionsverschiebung zu betonen, dass die Kalibrierung der globalen Kamera-Position immer nur entlang der Achse der aktuellen Kamera-Blickrichtung erfolgt. Soll eine globale Positionskalibrierung in andere Richtungen erfolgen, muss dazu die Kamera zunächst entsprechend gedreht werden, um dann in geänderter Blickrichtung eine erneute Positionskalibrierung entlang einer aktualisierten Kamera-Blickrichtung (z-Achse) durchzuführen. Soll also beispielsweise die globale Position nach links (oder rechts) kalibriert werden, muss die Kamera zunächst entsprechend nach links (oder rechts) rotiert werden, bevor schließlich in dieser geänderten Blickrichtung eine Positionskalibrierung per Zoom-Geste auf dem Bildschirm erfolgt.

Ein derartiges Vorgehen ermöglicht zwar theoretisch eine flexible Kalibrierung der globalen Kamera-Position in allen drei Freiheitsgraden bezogen auf das globale Referenzsystem. Sofern eine umfangreiche Positionskalibrierung entlang verschiedener Achsen durchgeführt werden soll, resultiert dies jedoch zwangsläufig auch in einer aufwändigen und wenig intuitiven Bedienung. Dies betrifft insbesondere eine etwaige Kalibrierung der Kamera-Position entlang der vertikalen y-Achse. Hierzu müsste die Kamera-Blickrichtung zunächst nach unten direkt auf den Boden (oder senkrecht nach oben gen Himmel) gerichtet werden, um eine Kalibrierung entlang ebendieser vertikalen Achse der aktuellen Kamera-Blickrichtung zu realisieren.

Um eine derartige unnatürliche und aufwändige Bedienung zu vermeiden, wurde eine dritte Interaktionsgeste entworfen, um eine separate **Kalibrierung der Kamera-Position entlang der vertikalen y-Achse** zu ermöglichen. Dies entspricht einer *Höhenkalibrierung* der Kamera-Position und resultiert in einer Verschiebung der dargestellten Geodatenmodelle in der AR-Ansicht nach oben oder unten. Die Kalibrierung wird über eine Drag-Geste in vertikaler Richtung nach oben (oder unten) mit zwei Fingern auf dem Bildschirm durchgeführt. Die Geste beginnt, sobald zwei beliebige Finger gleichzeitig den Bildschirm berühren und beide Finger eine Bewegung entlang der v-Achse des Pixelkoordinatensystems starten. Sie endet, sobald einer der beiden Finger den Bildschirm nicht mehr berührt. Die Länge dieser Geste Δd_y kann als Mittelwert der vertikalen Verschiebungskomponenten beider Finger-Gesten auf dem Bildschirm in Pixeln definiert werden. Analog zur Translationskalibrierung in z-Richtung kann auch hier die tatsächliche Positionsverschiebung in Metern über einen Skalierungsfaktor s_y gesteuert werden. Ein Translationsvektor t_{CorrY} zur Beschreibung der Positionsverschiebung des Kamera-Ursprungs $O_{CorrCam}$ in Bezug auf den unkalibrierten Kamera-Ursprung O_{Cam} entlang der vertikalen y-Achse kann schließlich wie folgt definiert werden:

$$t_{CorrY} = \begin{pmatrix} 0 \\ s_y \cdot \Delta d_y \\ 0 \end{pmatrix} \tag{4.14}$$

Die gesamte nutzergesteuerte Translationskorrektur t_{Corr} der globalen Kamera-Pose ergibt sich schließlich aus der Überlagerung der beiden manuellen Korrekturen t_{CorrY} und t_{CorrZ}:

$$t_{Corr} = t_{CorrY} + t_{CorrZ} = \begin{pmatrix} 0 \\ s_y \cdot \Delta d_y \\ s_z \cdot \Delta d_z \end{pmatrix} \tag{4.15}$$

Auf eine weitere Interaktionsgeste zur vereinfachten Kalibrierung der Kamera-Position entlang der x-Achse des Kamera-Systems wird bewusst verzichtet. Die Integration einer vierten Kalibrierungsgeste würde den gesamten Registrierungsprozess noch komplexer gestalten und zudem auf programmiertechnischer Ebene die Unterscheidbarkeit und Erkennung der einzelnen Gesten weiter erschweren.

4.5 Zusammenfassung

In diesem Kapitel wurde das entworfene geodatenbasierte manuelle Registrierungsverfahren im Detail vorgestellt. Durch Einblendung von 3D-Geodatenmodellen in der AR-Ansicht entsteht ein visuelles Hilfsmittel zur virtuellen Repräsentation der aktuellen Umgebung. Mittels Interaktionsgesten auf dem Bildschirm kann diese virtuelle Umgebung durch die Nutzer:innen so verschoben und der realen Ansicht angepasst werden, dass eine Kalibrierung von globaler Kameraposition und -orientierung erreicht wird.

Zur effektiven Integration unterschiedlicher Geodatenmodelle in die mobile AR-Anwendung konnte eine Geodatenverarbeitungskette entworfen werden, welche eine automatische Umwandlung von digitalen Geländemodellen, Oberflächenmodellen und 3D-Stadtmodellen in AR-fähige Formate ermöglicht. Im Rahmen mehrerer Umwandlungsschritte werden dabei Eingangsdaten in Form von texturlosen und texturierten 3D-Geodaten angepasst, hinsichtlich ihrer Dateigröße optimiert und als kleinflächige und renderfähige 3D-Kacheln in ein geeignetes 3D-Dateiformat umgewandelt, welches effektiv über eine server-basierte Dateimanagement-Komponente (Geodaten-CMS) einem AR-Client bereitgestellt und dort geladen werden kann. Die Berücksichtigung verschiedenartiger 3D-Geodatenmodelle erlaubt dabei je nach Charakteristik der aktuelle umgebenden Landschaft eine geeignete virtuelle Abbildung der Umgebung, sei es die Abbildung von markanten Gebäudestrukturen als Referenzobjekte in der unmittelbaren urbanen Umgebung durch kleinflächige digitale 3D-Stadtmodelle, oder die virtuelle Abbildung von markanten Bergketten in weiter Entfernung in bergiger Umgebung durch großflächige digitale Geländemodelle.

Der manuelle Registrierungsprozess zur Kalibrierung der Kamera-Pose ist dabei eingebettet in ein Gesamtsystem für die globale Registrierung mit Echtzeit-Tracking. Eine

grobe initiale Pose-Bestimmung erfolgt zunächst auf Basis der GNSS- und IMU-Sensoren sowie durch eine automatische Höhenkalibrierung mithilfe digitaler Geländemodelle. Diese erste Schätzung der globalen Kamera-Pose definiert den Ursprung eines lokalen bildbasierten Tracking-Systems, welches im Hintergrund während des Kalibrierungsprozesses ständig alle lokalen Kamera-Bewegungen vor Ort registriert. Der manuelle Prozess zur Kalibrierung der globalen Kamera-Pose mittels Verschiebungen und Interaktionsgesten auf dem Bildschirm des Endgeräts kann dabei als eine Translations- und Rotations-Korrektur der Kamera-Pose in Bezug auf das lokale Tracking-System verstanden werden. Für die Kalibrierung selbst stehen insgesamt drei verschiedene Interaktionsgesten zur Verfügung: eine einfache Verschiebungsgeste zur Kalibrierung der Kamera-Rotation, eine Zoom-Geste zur Kalibrierung der Kamera-Translation entlang der aktuellen Blickrichtung sowie eine Verschiebungsgeste mit zwei Fingern zur nachträglichen manuellen Kalibrierung der globalen Gerätehöhe.

Nach dieser detaillierten Vorstellung der Funktionsweise der entwickelten Registrierungsmethode soll im folgenden Kapitel schließlich die Genauigkeit und das Anwendungspotenzial der Methode im Rahmen einer Praxis-Evaluation untersucht werden.

Kapitel 5

Evaluation

Das Anwendungspotenzial der entwickelten AR-Registrierungsmethode hängt maßgeblich von der Genauigkeit und der Bedienbarkeit des Verfahrens im Praxiseinsatz ab. Diese Thematik soll im folgenden Kapitel genauer betrachtet und beantwortet werden.

Die Evaluation des entwickelten geodatenbasierten Verfahrens erfolgt im Rahmen einer umfangreichen Feldstudie mit mehreren Testpersonen. Bevor die Ergebnisse dieser Praxistests präsentiert und diskutiert werden, sollen zunächst das Testmaterial und die Testmethode genauer vorgestellt werden. Die Entwicklung von Testmaterial und -methode erfolgt dazu auf Basis von vorab definierten Hypothesen hinsichtlich Genauigkeit und Bedienbarkeit des Kalibrierungsverfahrens. Diese Hypothesen zur Evaluation werden zunächst in Kapitel 5.1 vorgestellt. Als Testmaterial dient eine zum Zwecke der Evaluation entwickelte mobile AR-Anwendung, welche das vorgestellte nutzergesteuerte Registrierungsverfahren implementiert und eine Bewertung des Anwendungspotenzials der Methode ermöglicht. Die Anforderungen an diese mobile Test-Anwendung, deren grafische Benutzeroberfläche sowie Details zur Implementierung der App werden in Kapitel 5.2 präsentiert. Die genaue Vorgehensweise der Evaluation wird anschließend in Kapitel 5.3 erläutert. In diesem Kontext werden neben dem Evaluationsablauf auch die genutzten Test-Standorte und deren Charakteristiken sowie die eingesetzten Evaluationsinstrumente vorgestellt.

Anschließend erfolgt eine Präsentation der Evaluationsergebnisse, die sowohl die erreichte Präzision hinsichtlich globaler Position und Ausrichtung des Endgeräts zeigen als auch die subjektiv durch die Testnutzer bewertete Bedienbarkeit des Verfahrens (Kapitel 5.4). Eine abschließende Zusammenfassung der Ergebnisse bewertet schließlich die evaluierten Werte und verdeutlicht die Vorteile, aber auch die Beschränkungen und Limitierungen des Registrierungsverfahrens (Kapitel 5.5). Auch die Richtigkeit der aufgestellten Hypothesen kann in dieser Diskussion abschließend geprüft werden.

5.1 Hypothesen zur Evaluation

Die Entwicklung einer geeigneten Evaluations-Anwendung sowie eines geeigneten Evaluationskonzepts erfolgt auf Basis von vorab definierten Hypothesen hinsichtlich Genauigkeit und Bedienbarkeit des entwickelten Registrierungsverfahrens. Jede dieser Hypothesen ermöglicht Rückschlüsse darauf, wie der Ablauf der Evaluation bzw. die mobile Evaluations-App gestaltet werden sollten. Die Hypothesen sowie die resultierenden Maßnahmen hinsichtlich der Gestaltung der Evaluation sind:

H1 *Die Registrierungsmethode ermöglicht eine präzise Bestimmung der Geräte-Pose in unterschiedlichen Umgebungscharakteristiken mit geringem Zeitaufwand.*

Maßnahme: Die erreichte Genauigkeit der Geräte-Pose (Ausrichtung, Position) sowie der zur Kalibrierung benötigte Zeitaufwand sollen objektiv beurteilt werden, und zwar an verschiedenen Standorten mit unterschiedlicher Umgebungscharakteristik. Die objektive Beurteilung der Genauigkeit kann durch nachträgliche Analyse erfolgen, indem wesentliche Parameter während der Nutzung aufgezeichnet werden sowie georeferenzierte Validierungsobjekte nach erfolgter Kalibrierung in der AR-Ansicht eingeblendet werden.

H2 *Die Registrierungsmethode ist an allen Standorten intuitiv bedienbar. Die Bedienbarkeit ist jedoch abhängig vom verwendetem Geodatenmodell und der Umgebung.*

Maßnahme: Die Bedienbarkeit soll an unterschiedlichen Standorten und bei Nutzung unterschiedlicher 3D-Geodatenmodelle bewertet werden, und zwar mittels Befragung der testenden Nutzer per Fragebogen.

H3 *Mit der geodatenbasierten Registrierungsmethode ist eine Kalibrierung der globalen Position schwieriger realisierbar als eine Kalibrierung der globalen Orientierung.*

Maßnahme: Eine Evaluation der Kalibrierung der globalen Ausrichtung mittels virtuellem Geodatenmodell soll auch unabhängig von einer Positionskalibrierung ermöglicht werden. Dazu soll alternativ auch eine manuelle Positionsbestimmung über eine Kartenansicht angeboten werden. Bedienbarkeit und Genauigkeit der geodatenbasierten AR-Kalibrierung sollen sowohl mit als auch ohne Vorab-Positionsbestimmung per Kartenansicht abgefragt werden, um Rückschlüsse auf Vor- und Nachteile der Positionierungsmethoden zu ermöglichen.

H4 *Die Bedienung der geodatenbasierten Registrierungsmethode fällt auf einem Tablet mit großem Bildschirm leichter als auf einem Smartphone.*

Maßnahme: Die Evaluation soll sowohl auf einem Tablet als auch auf einem Smartphone durchgeführt werden. Die Bedienbarkeit auf den jeweiligen Geräten soll per Befragung an die Testperson beurteilt werden.

Ziel der Evaluation soll nun eine Prüfung sein, ob und in welchem Maße diese Hypothesen im Praxiseinsatz korrekt sind. Diese Prüfung findet in der abschließenden Diskussion der Evaluationsergebnisse statt (siehe Kapitel 5.5). Zunächst sollen in den folgenden Kapiteln die auf Basis der Hypothesen entwickelte Evaluations-Anwendung sowie der detaillierte Evaluationsablauf vorgestellt werden.

5.2 Mobile Evaluations-Anwendung

Die zum Zweck der Evaluation entwickelte mobile Anwendung integriert das entwickelte geodatenbasierte Registrierungsverfahren und berücksichtigt weitere Rahmenbedingungen, die auf Basis der oben definierten Evaluations-Hypothesen innerhalb der Anwendung realisiert werden müssen. Die konkreten Anforderungen an die mobile Anwendung, die einzelnen Schritte zur Nutzung der Anwendung (*User Flow*), die daraus resultierende grafische Benutzeroberfläche sowie Details zur Implementierung werden im Folgenden präsentiert.

5.2.1 Anforderungen

Die wesentlichen Anforderungen an die mobile Evaluations-Anwendung ergeben sich vor allem aus den oben genannten Hypothesen hinsichtlich Genauigkeit und Bedienbarkeit des Registrierungsverfahrens. Abbildung 5.1 fasst die Anforderungen zusammen. Im Einzelnen lauten diese:

A1: Art der Geodatenmodelle Die Anwendung soll so gestaltet sein, dass alle 3D-Geodatenmodelle, die im Rahmen der Geodatenverarbeitungskette mit den genannten Spezifikationen (siehe Kapitel 4.2.1) verarbeitet, integriert und geladen werden können. Das heißt, digitale Geländemodelle (DGM), texturierte und untexturierte digitale Oberflächenmodelle (DOM, DOM_RGB) sowie texturierte und untexturierte 3D-Stadtmodelle (CityModel, CityModel_RGB) sollen im Rahmen der Evaluation getestet werden können.

A1		**Art der Geodatenmodelle:** Integration von fünf verschiedenen Arten von 3D-Geodatenmodellen
A2		**Wahl der Geodatenmodelle:** Kartenbasierte Anzeige der verfügbaren und manuelle Wahl der zu ladenden Geodatenmodelle
A3		**Abdeckung:** Integration der Geodatenmodelle innerhalb von vorab definierten Test-Standorten ausreichend
A4		**Manuelle Kalibrierung:** Geodatenbasierte nutzergesteuerte Kalibrierung basierend auf der entwickelten Methode
A5		**Positionsbestimmung:** Möglichkeit einer alternativen kartenbasierten Positionsbestimmung
A6		**Endgeräte:** Realisierung der mobilen Anwendung auf handelsüblichen Smartphones und Tablets
A7		**Daten-Logging**: Persistentes Speichern aller Interaktionen und internen Berechnungen bzgl. Position und Ausrichtung zur nachträglichen Auswertung
A8		**Validierungsobjekte:** AR-Visualisierung von virtuellen georeferenzierten Windenergieanlagen als Validierungsobjekte nach erfolgter Kalibrierung, um Korrektheit der Ausrichtung nachträglich zu beurteilen

Abbildung 5.1 Anforderungen an die mobilen Evaluations-Anwendung

A2: Wahl der Geodatenmodelle Über eine separate Kartenansicht sollen alle verfügbaren Kacheln der Geodatenmodelle dargestellt werden. Diese Darstellung soll es zudem ermöglichen, einzelne Kacheln manuell an- und abzuwählen. Nur gewählte Kacheln sollen dann bei der AR-basierten Kalibrierung berücksichtigt und in die AR-Ansicht geladen werden können. Über eine weitere Auswahlmöglichkeit (z.B. Checkbox-Menü) soll dann die Sichtbarkeit der fünf Geodatenmodellarten gesteuert werden. Auf diese Weise kann – je nach Standort und gewünschter Evaluationsart – händisch gesteuert werden, welche Kacheln und welche Modellart im jeweiligen Testdurchlauf dargestellt werden sollen. Um die Vor- und Nachteile der einzelnen Geodatenmodellarten bestmöglich beurteilen zu können, soll bei der Evaluation die Nutzbarkeit der einzelnen Modellarten isoliert betrachtet werden, nicht die Kombination und Überblendung mehrerer Modellarten gleichzeitig. Es sollen also zwar beliebig viele Kacheln dargestellt werden können, jedoch immer nur von einer einzigen Modellart gleichzeitig.

A3: Räumliche Abdeckung Die Evaluation soll an vorab definierten Standorten stattfinden. Es müssen daher nur Geodatenmodelle prozessiert und in die Anwendung geladen werden, die sich im sichtbaren Umfeld dieser Test-Standorte befinden.

A4: Manuelle Kalibrierung Das manuelle geodatenbasierte Kalibrierungsverfahren selbst soll so implementiert und realisiert werden, wie zuvor in Kapitel 4.4 beschrieben und in Abbildung 4.2 dargestellt wurde. Das heißt, basierend auf einer initialen GNSS- und IMU-gestützten Schätzung von Position und Ausrichtung und automatischen DGM-basierten Höhenschätzung erfolgt eine erste AR-Visualisierung der jeweils gewählten 3D-Kacheln. Daraufhin ist eine manuelle Kalibrierung der AR-Darstellung mit den genannten Interaktionsgesten möglich.

A5: Kartenbasierte Positionsbestimmung Zusätzlich soll die Test-Anwendung noch eine alternative Möglichkeit bieten, die eigene zweidimensionale geografische Position über eine Kartenansicht (Satellitenbild) zu bestimmen. Diese Anforderung basiert auf der Hypothese, dass die Kalibrierung der globalen Position mithilfe der visualisierten 3D-Geodatenmodelle schwieriger durchzuführen ist als eine Kalibrierung der globalen Ausrichtung mithilfe der Modelle. Erfolgt eine präzise Positionsbestimmung zunächst alternativ über eine Kartenansicht, kann eine initiale AR-Projektion der 3D-Geodatenmodelle auf Basis einer manuell präzisierten Position forciert werden und die Bedienbarkeit der globalen Rotationskalibrierung auch isoliert auf Basis dieser präziseren Positionierung betrachtet werden.

A6: Endgeräte Die Evaluation soll mithilfe von handelsüblichen Smartphones und Tablets durchgeführt werden. Damit ist eine Beurteilung der Bedienbarkeit sowohl auf mobilen Endgeräten mit großem als auch kleinerem Bildschirm möglich.

A7: Daten-Logging Die mobile Evaluations-App soll alle wesentlichen Parameter, die während der Kalibrierung berechnet und verändert werden, auf persistente Art und Weise speichern, beispielsweise in Form einer lokal auf dem Endgerät gespeicherten Textdatei. Auf diese Weise ist eine nachträgliche objektive Analyse von erreichter Genauigkeit und benötigtem Zeitaufwand möglich. Zu den relevanten Parametern zählen u.a. die geografische Position und globale Ausrichtung der virtuellen AR-Umgebung, die aktuell sichtbare Geodatenmodellart sowie Angaben zum Zeitpunkt von Beginn und Ende des Kalibrierungsvorgangs.

A8: Virtuelles Validierungsobjekt Eine nachträgliche Prüfung der Genauigkeit der kalibrierten globalen Ausrichtung ist möglich, indem die kalibrierte AR-Ansicht nach erfolgter Kalibrierung mit einer ideal kalibrierten Ansicht (*Ground Truth*) verglichen wird. Dies ist im Nachhinein möglich, indem nach jedem Kalibrierungsvorgang

3D-Modelle (Validierungsobjekte) an einer bestimmten geografischen Position positioniert und in der AR-Ansicht sichtbar gemacht werden. Diese Bildaufnahme kann im Nachgang dann händisch mit einer idealen AR-Projektion, einer Ground-Truth-Aufnahme, verglichen werden (siehe Kapitel 5.3.1 für eine detailliertere Beschreibung der Vorgehensweise). In der Evaluations-Anwendung sollen als Validierungsobjekte nach jedem Kalibrierungsvorgang an jedem Standort zwei virtuelle Windenergieanlagen visualisiert werden.

5.2.2 User Flow

Die mobile Evaluations-Anwendung wird sowohl von den *Testpersonen* selbst zum Durchführen von Testdurchläufen der Kalibrierungsmethode als auch von einem *Evaluationsleiter* zur Administration der Geodatenmodelle genutzt. Die Benutzerführung der mobilen Anwendung (User Flow) wird im Folgenden kurz erläutert.

Testdurchlauf (Testperson)

Primäres Ziel der mobilen Evaluations-App aus Sicht der Testperson besteht darin, am aktuellen Test-Standort die manuelle geodatenbasierte Kalibrierung mit einer bestimmten Geodatenmodellart einmalig und mit möglichst hoher Genauigkeit durchzuführen. Nach erfolgter Kalibrierung kann die Anwendung wieder beendet und ein neuer Kalibrierungsvorgang mit geänderten Parametern (z.B. veränderter Geodatenmodellart) gestartet werden. Abbildung 5.2 skizziert die einzelnen Schritte dieses Programmablaufs.

Nach Start der Anwendung wählt die Testperson zunächst die gewünschten Einstellungen zum folgenden Testdurchlauf. Dazu wird aus den fünf verfügbaren Geodatenmodellarten die gewünschte zu ladende Modellart gewählt sowie der aktuelle Test-Standort manuell aus einer vorgegebenen Liste selektiert, sodass die korrekten, für den jeweiligen Standort vorgesehenen Validierungsobjekte nach erfolgter Kalibrierung anzeigt werden können. Zudem legt die Testperson fest, ob eine alternative kartenbasierte Positionsbestimmung vorab erfolgen soll.

Anschließend wird der Kalibrierungsvorgang manuell gestartet. Dazu wird entweder direkt die AR-Ansicht mit den gewählten virtuellen Geodatenmodellen geladen, in der die nutzergesteuerte Kalibrierung erfolgen kann, oder – falls eine kartenbasierte Positionsbestimmung aktiviert wurde – zunächst eine Kartenansicht eingeblendet, mit der Möglichkeit, die eigene Position auf dieser zu präzisieren. Nach manuellem Angleichen der virtuellen Landschaft per Interaktionsgesten entscheiden die Nutzer:innen

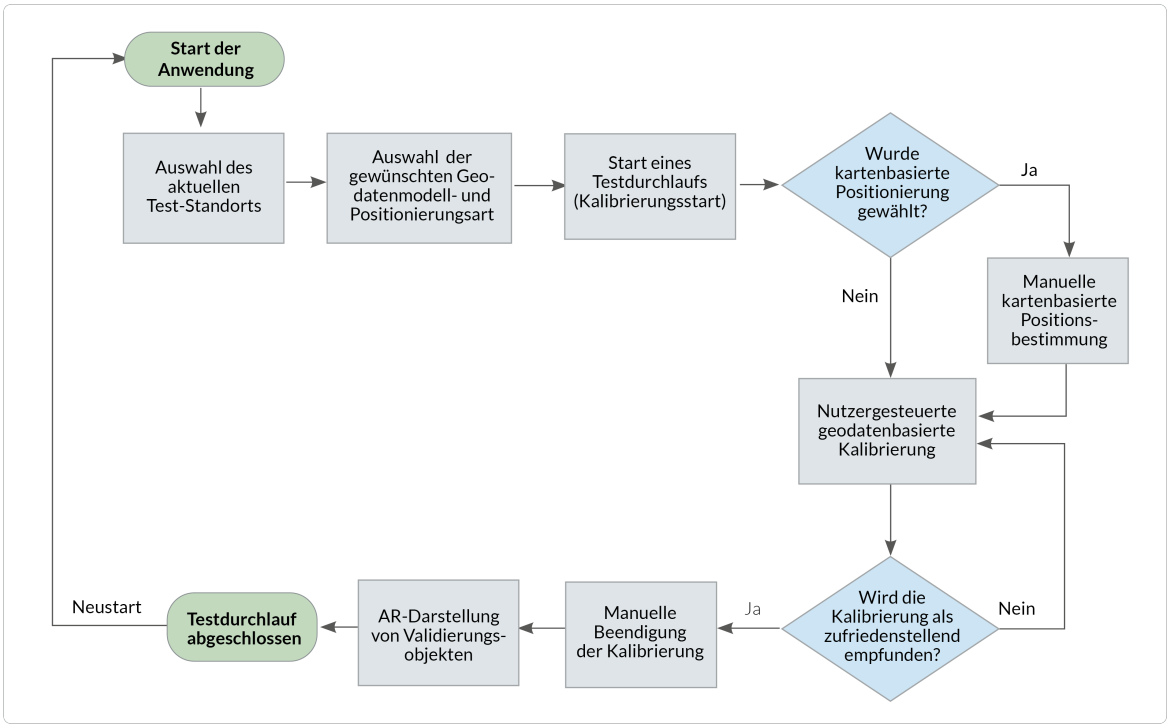


Abbildung 5.2 Benutzerführung (User Flow) der mobilen Evaluations-Anwendung zur Durchführung eines Testdurchlaufs aus Sicht der Testperson.

schließlich selbst, wann die Angleichung der virtuellen Modelle an die reale Ansicht als zufriedenstellend empfunden wird und der Kalibrierungsprozess beendet werden soll. Nach Beendigung der Kalibrierung werden in der AR-Ansicht die für den jeweiligen Test-Standort vorab definierten Validierungsobjekte (zwei virtuelle Windenergieanlagen) an ihrer geografischen Position eingeblendet. Diese können dann nachträglich für die Beurteilung der Kalibrierungsqualität genutzt werden (siehe Kapitel 5.3.1). Die zur Kalibrierung genutzten Geodatenmodelle werden gleichzeitig ausgeblendet, d.h. sie sind nicht mehr sichtbar, tragen jedoch zur Verdeckung der virtuellen Validierungsmodelle bei. Eine Verdeckung der in der AR-Ansicht dargestellten Windenergieanlagen durch Gebäude oder Vegetation wird so also berücksichtigt, was zu einer – je nach verwendeter Geodatenmodellart und Qualität der Kalibrierung – mehr oder weniger realitätsnahen AR-Darstellung führt.

Administration (Evaluationsleiter)

Der Leiter der Evaluation nutzt die mobile Anwendung primär zur Administration der einsetzbaren 3D-Geodatenmodelle im Vorfeld der Evaluation an jedem Test-Standort. Dazu werden über eine Kartenansicht alle verfügbaren Kacheln der 3D-Geodatenmodelle

angezeigt. Der Administrator kann schließlich für jede Geodatenmodellart einzelne Kacheln manuell an- und abwählen. Nur gewählte Kacheln werden dann beim Kalibrierungsvorgang berücksichtigt und in die AR-Ansicht geladen. Dieses Vorgehen wird bei jedem Test-Standort durch den Evaluationsleiter vorab durchgeführt, sodass bei allen Testdurchläufen am Standort die gleichen Modelle in gleicher und geeigneter räumlicher Abdeckung zur Verfügung stehen.

Außerdem definiert der Evaluationsleiter vor Beginn der Evaluation die Test-Standorte mit den geografischen Positionen der jeweils zu visualisierenden Validierungsobjekte (Windenergieanlagen). Die Festlegung dieser Standorte geschieht jedoch nicht innerhalb der mobilen Anwendung, sondern separat über Textdateien, die beim Start in die Anwendung geladen werden.

5.2.3 Grafische Benutzeroberfläche

Die grafische Benutzeroberfläche (*Graphical User Interface; GUI*) ergibt sich aus den Anforderungen und dem eben beschriebenen User Flow der mobilen Anwendung. Die Anwendung besteht primär aus fünf verschiedenen Oberflächen (*Screens*). Diese sind in Abbildung 5.3 dargestellt:

S1: Start-Screen Nach dem Start der Anwendung erscheint ein Bildschirm zur Konfiguration des Testdurchlaufs. Über ein Checkbox-Menü kann die zu ladende Geodatenmodellart gewählt sowie eine Option zur Vorab-Positionsbestimmung per Karte an- und abgewählt werden. Außerdem werden in einer Listenansicht alle verfügbaren Test-Standorte angezeigt. Diese basieren auf vorab durch den Evaluationsleiter erstellten Textdateien, in welchen die geografischen Positionen der jeweiligen Validierungsobjekte definiert werden. Nach Selektion des aktuellen Test-Standortes in der Liste ist ein Start des Kalibrierungsvorgangs über einen *Laden*-Button möglich.

S2: Geodatenmodell-Administration Über einen Konfigurationsbutton auf dem Start-Screen kann die Ansicht zur Verwaltung der verfügbaren Geodatenmodelle geöffnet werden. In dieser Kartenansicht kann der Evaluationsleiter alle verfügbaren Kacheln der Geodatenmodelle einsehen sowie durch Selektion einzelner Kacheln diese für den Einsatz im Kalibrierungsprozess aktivieren. Über Buttons am oberen Bildschirmrand ist ein Wechsel zwischen den fünf integrierten Geodatenmodellarten möglich. Eine Listenansicht im unteren Bildschirmbereich gibt zudem einen Überblick über die Gesamtzahl der jeweils verfügbaren und gewählten Kacheln.

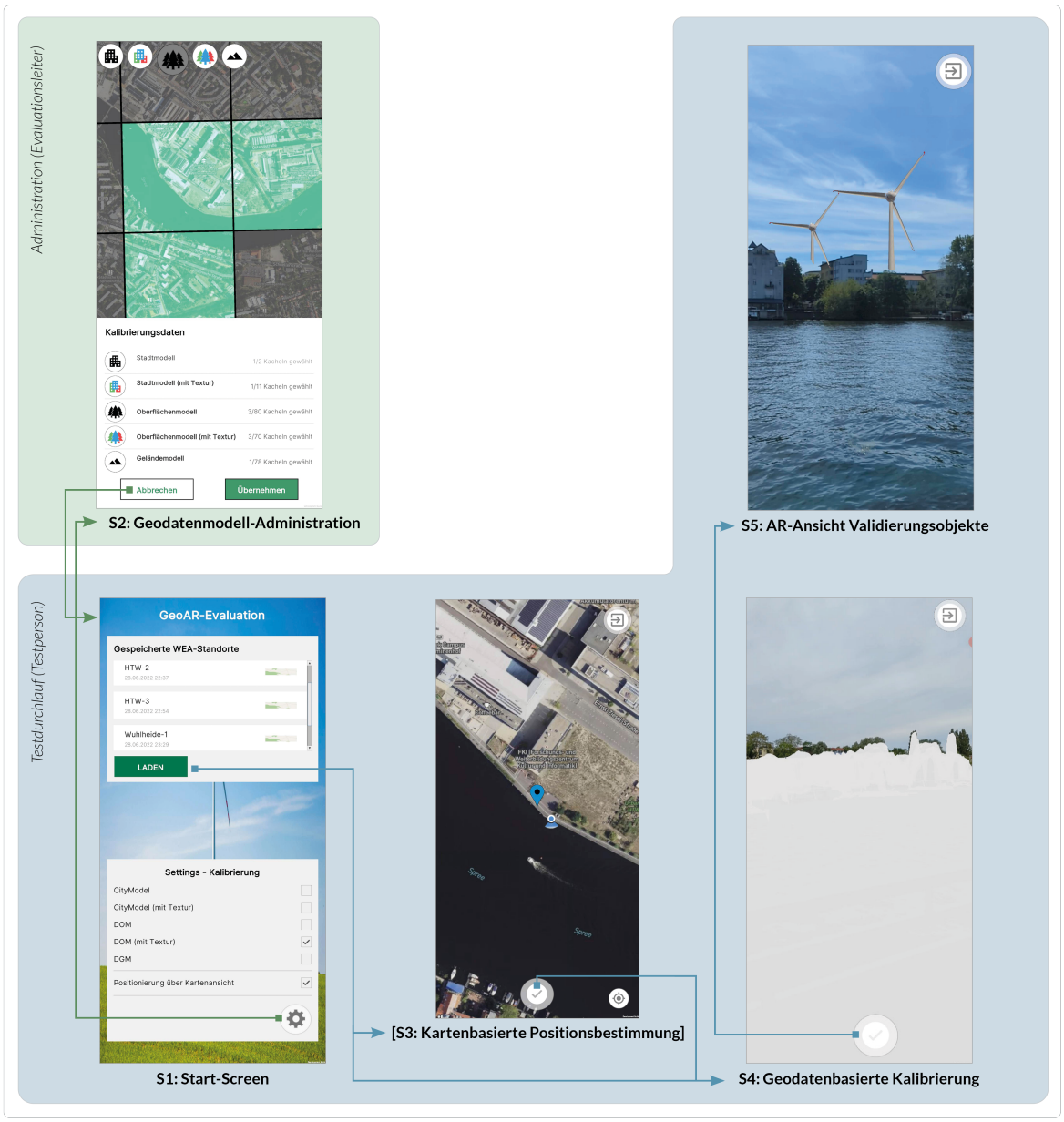


Abbildung 5.3 Grafische Benutzeroberfläche der mobilen Evaluations-Anwendung

S3: Kartenbasierte Positionsbestimmung Wurde eine manuelle kartenbasierte Vorab-Positionsbestimmung aktiviert, so öffnet sich vor dem Öffnen der AR-basierten Kalibrierungsansicht zunächst eine Kartenansicht mit Luftbildaufnahmen (Satellitenbild). Die Karte zeigt per Icon den aktuellen Nutzerstandort basierend auf der initialen GNSS-basierten Positionsschätzung. Per Touch-Gesten kann die Kartenansicht verschoben und skaliert werden, um den tatsächlichen aktuellen Standort auf einen stets

in der Bildschirmmitte platzierten Marker zu verschieben und zu präzisieren. Über einen *Bestätigen*-Button wird die aktualisierte Position bestätigt.

S4: Geodatenbasierte Kalibrierung Die Hauptansicht der Kalibrierungsmethode zeigt eine AR-Visualisierung der gewählten 3D-Geodatenmodelle mit ihrer initialen Projektion. Das heißt, die selektierten Kacheln der gewünschten Geodatenmodellart werden basierend auf der initialen Schätzung der globalen Gerätepose auf dem Kamerabild eingeblendet und können nun über Touch-Gesten verschoben und der realen Umgebung angeglichen werden (siehe Kapitel 4.4.2). Über einen *Bestätigen*-Button kann die durchgeführte Kalibrierung durch die Testperson bestätigt werden. Ein *Exit*-Button ermöglicht ein sofortiges Beenden der Test-Anwendung, z.B. wenn der Testperson eine Kalibrierung unmöglich erscheint und der Test-Durchlauf abgebrochen werden soll.

S5: AR-Ansicht Validierungsobjekte Nach Bestätigung der Kalibrierung werden die Geodatenmodelle in der AR-Ansicht ausgeblendet und die am Standort definierten Validierungsobjekte an ihrer geografischen Position in der AR-Ansicht eingeblendet. Je nach Art und Qualität der durchgeführten Kalibrierung wirkt diese AR-Visualisierung mehr oder weniger realitätsnah. In jedem Fall kann die Visualisierung der Validierungsobjekte für eine nachträgliche Analyse der Registrierungsgenauigkeit verwendet werden (siehe Kapitel 5.3.1). Die Testperson hat einen Testdurchlauf nun vollständig absolviert. Die Anwendung kann nun über einen *Exit*-Button geschlossen und ein neuer Testdurchlauf gestartet werden.

5.2.4 Implementierung

Die mobile Evaluations-Anwendung wurde basierend auf der konzipierten System-Architektur der geodatenbasierten Registrierungsmethode (siehe Kapitel 4.1) und unter Berücksichtigung der eben beschriebenen Anforderungen und Rahmenbedingungen implementiert. Wesentliche Aspekte bezüglich der Implementierung der beiden Hauptkomponenten des Gesamt-Systems, d.h. der Geodatenverarbeitungskette und des mobilen AR-Clients, werden im Folgenden beschrieben. Eine Implementierung des entworfenen Geodaten-CMS erfolgte zwar im Rahmen des Forschungsprojekts *mARGo* [115], diese serverseitige Implementierung wurde jedoch im Rahmen der Evaluation nicht verwendet. Stattdessen wurden die von der Geodatenverarbeitungskomponente erzeugten Kacheln direkt auf dem Dateisystem der mobilen Testgeräte gespeichert und von dort geladen, um einen möglichst stabilen und zuverlässigen Ablauf der Tests zu gewährleisten.

Geodatenverarbeitungskette

Die Realisierung der einzelnen Verarbeitungsschritte der Geodatenverarbeitungskette (siehe Kapitel 4.2.2) erfolgte auf Basis von individuell erstellten Python-Scripts sowie unter Verwendung von mehreren Open-Source-Tools. Die für jede integrierte Geodatenmodellart notwendigen Verarbeitungsschritte sind in Abbildung 4.6 dargestellt. Da sich die Verarbeitungsschritte für jede Geodatenmodellart unterscheiden, wurde für jede Geodatenmodellart ein individuelles Script entworfen, welches nacheinander alle jeweils für eine Quelldatei notwendigen Verarbeitungsschritte abarbeitet und somit eine automatische Verarbeitung beliebig vieler Eingangsdaten einer bestimmten Modellart ermöglicht (siehe Abbildung 5.4(a)).

Die lokale **Koordinatensystem-Anpassung** wird realisiert, indem die textbasierten .xyz- oder .obj-Quelldateien über das Python-Script zeilenweise eingelesen und die Koordinaten aller 3D-Gitterpunkte bzw. Vertices entsprechend der gewünschten Verschiebungskomponente aktualisiert und neu abgespeichert werden. Ein ähnliches Vorgehen erfolgt bei den Verarbeitungsschritten der **Kachelung** und der **Punktgitter-Reduktion**, welche bei der Verarbeitung der untexturierten digitalen Oberflächenmodelle durchgeführt werden. Auch diese Verarbeitungsschritte können – auf Basis der in Kapitel 4.2.2 beschriebenen Vorgehensweise – durch eine zeilenweise Verarbeitung der .xyz-Dateien innerhalb des Python-Scipts ohne den Einsatz externer Tools realisiert werden. Aus einer großflächigen DOM-Quelldatei entstehen auf diese Weise mehrere kleinflächige .xyz-Dateien mit reduzierter räumlicher Auflösung.

Für die **3D-Mesh-Erzeugung** zur Generierung von Dreiecksnetzen im .obj-Format mittels Triangulation wird auf externe Tools zurückgegriffen. Eine Umwandlung von .xyz-Punktgitter-Dateien in .obj-Formate erfolgt mithilfe des Tools *TIN Terrain* [54]. Dieses Programm nutzt dabei unter anderem wesentliche Funktionalitäten der *GDAL*-Bibliothek [147], einer Open-Source-Bibliothek zur Verarbeitung räumlicher Geodaten. Das Tool *TIN Terrain* kann als *Docker*-Container [35] genutzt werden. Dieser wird im Zuge der Verarbeitungskette über das Python-basierte Verarbeitungsskript mit entsprechenden Parametern als externer Prozess gestartet. Eine .obj-Mesh-Erzeugung der CityGML-basierten 3D-Stadtmodelle erfolgt mithilfe des Tools *cjio* [59]. Damit werden die .gml-Dateien der Stadtmodelle zunächst in das artverwandte *CityJSON*-Format [81] konvertiert, um daraus dann Oberflächennetze im .obj-Format zu generieren.

Auch bei den Verarbeitungsschritten zur Optimierung der texturierten Geodatenmodelle wird auf externe Tools zurückgegriffen, die dann über das Python-basierte Verarbeitungsskript als externe Prozesse gestartet werden. Eine **3D-Mesh-Vereinfachung** der texturierten Oberflächenmodelle wird mithilfe des Open-Source-Tools *MeshLab*

[30] bzw. der darauf basierenden Python-Bibliothek *PyMeshLab* realisiert. Diese Bibliothek integriert einen Filter *Quadric Edge Collapse Decimation*, der einen von Garland und Heckbert entwickelten Algorithmus implementiert [47], um die Komplexität der Mesh-Struktur auf effektive Art zu vereinfachen, indem die Anzahl der Triangles im Oberflächennetz iterativ reduziert wird. Das Tool *MeshLab* kann ebenfalls zur **Textur-Optimierung** eingesetzt werden, nämlich durch das Kombinieren von Texturinformationen aus mehreren Einzeltexturen in eine einzige neue Texturdatei (*texture baking*). Diese Übertragung von Texturinformationen von einem Mesh (mit sehr vielen Texturbildern) auf ein anderes Mesh (mit einem einzigen Texturbild) kann durch die Anwendung des Filters *Transfer: Vertex Attributes to Texture* umgesetzt werden. Durch diese Optimierung kann die Ladezeit der 3D-Modelle deutlich reduziert werden. Ein weiterer Verarbeitungsschritt zur Optimierung der Modell-Texturen besteht in der Erzeugung neuer Texturbilder mit deutlich höheren Bildkompressionsraten. Dieser Verarbeitungsschritt wird mithilfe des Open-Source-JPEG-Encoders *MozJPEG* [101] implementiert. Damit kann die Dateigröße der JPEG-basierten Texturbilder der eingesetzten texturierten 3D-Geodatenmodelle um etwa das 10-fache reduziert werden bei kaum merklichem Qualitätsverlust.

Die finale **3D-Format-Umwandlung** aller .obj-Modelle in .glb-Dateien wird innerhalb der Geodatenverarbeitungskette schließlich mithilfe des Tools *obj2gltf* [28] realisiert, das als Teil der umfangreichen *Cesium-Plattform* [27] für webbasierte 3D-Visualisierungen von Geodatenmodellen zur Verfügung gestellt wird. Die abschließende **JSON-Meta-Daten-Erzeugung** kann schließlich wieder ohne den Einsatz externer Tools unmittelbar mittels des integrierten Python-JSON-Encoders innerhalb des Python-basierten Verarbeitungsskripts erfolgen.

Die im Rahmen der Evaluation verwendeten **3D-Geodatenmodelle** stammen aus offiziellen, frei verfügbaren Quellen der Berliner und Brandenburger Landesbehörden. Die für die Test-Standorte genutzten digitalen Oberflächenmodelle und 3D-Stadtmodelle wurden dazu mit den in Kapitel 4.2.1 genannten Spezifikationen über das *Berliner 3D-Downloadportal* [19] (siehe Abbildung 5.4(b)) sowie über das *Berliner Open-Data-Portal* [129] kostenlos bezogen. Digitale Geländemodelle für alle Test-Standorte sind über den *LGB-Geobroker* [85] erhältlich.

Mobiler AR-Client

Die Implementierung der mobilen **AR-Evaluations-Anwendung** wurde mithilfe der Spiele-Entwicklungsumgebung *Unity* [144] für Android-basierte Endgeräte durchgeführt (siehe Abbildung 5.4(c)). Wesentliche AR-bezogene Funktionalitäten werden dabei

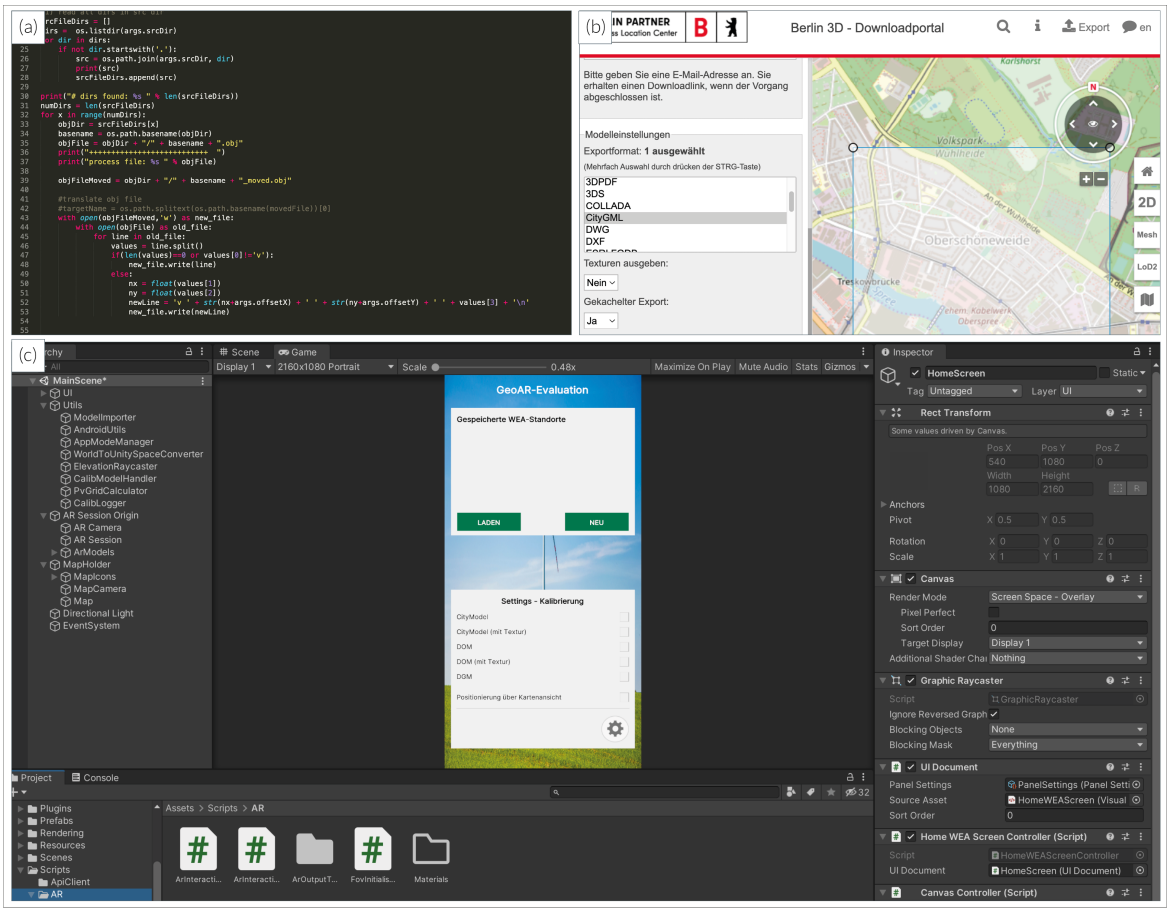


Abbildung 5.4 (a) Über Python-Skripte werden die einzelnen Schritte der Geodatenverarbeitungskette gesteuert und automatisiert. (b) Die eingesetzten 3D-Geodatenmodelle wurden unter anderem über ein offizielles Berliner 3D-Downloadportal bezogen [19]. (c) Die Implementierung der mobilen Evaluations-Anwendung wurde mithilfe der Unity-Entwicklungsumgebung realisiert [144].

über das Framework *Unity AR Foundation* [143] realisiert. Dazu zählt insbesondere die Funktionalität zum Erfassen von lokalen Kamerabewegungen mittels Visual-Inertial-Tracking-Methoden, die für Android-basierte Mobilgeräte durch die *Google ARCore API* [49] bereitgestellt werden.

Die Umsetzung der entworfenen **grafischen Benutzeroberflächen** erfolgte primär mittels des in der Unity-Entwicklungsumgebung integrierten *Unity UI Toolkit*. Die Oberflächen wurden dabei so realisiert, dass diese sowohl im Querformat (Landscape-Modus) für Tablet-Endgeräte als auch im Hochformat (Portrait-Modus) für Smartphones darstellbar und bedienbar sind. Alle zweidimensionalen Kartenansichten innerhalb der Unity-Anwendung basieren auf dem *Mapbox Maps SDK for Unity* [92], welches auch die Darstellung von hochaufgelösten Satellitenbildern ermöglicht.

Der **3D-Modell-Import** der von der Verarbeitungskette erzeugten .glb-Modelle wurde mithilfe des Unity-Plugins *GLTFUtility* [130] realisiert. Um die texturlosen Geodatenmodelle sichtbar zu machen, wurde diesen nach dem Import ein hellgraues, leicht transparentes Material zugewiesen. Die Interaktionsformen der **nutzergesteuerten geodatenbasierten Kalibrierung** wurden schließlich über individuelle Algorithmen zur Erkennung der entworfenen Touch-Gesten implementiert. Analog zu den in Kapitel 4.4.1 dargestellten mathematischen Zusammenhängen erzeugen diese Interaktionen dann eine entsprechende Korrektur von Position und Ausrichtung der virtuellen Unity-AR-Kamera in Bezug auf ein Referenzsystem in der realen Welt.

Die **initiale GNSS- und IMU-basierte Schätzung** der globalen Gerätepose nutzt die durch die Unity-Umgebung bzw. das Android-Betriebssystem bereitgestellten internen Lokalisierungsfunktionen. Zur Bereitstellung einer initialen globalen Rotationsschätzung wurde zudem ein eigenes Android-Plugin für Unity erstellt, welches den durch das Android-Betriebssystem erzeugten *Rotation-Vector-Sensor* in die Unity-Umgebung integrieren kann, der durch software-gestützte Kombination alle verfügbaren IMU-Sensoren für eine robuste Rotationsbestimmung nutzt. Eine initiale **DGM-basierte Höhenkalibrierung** erfolgt anschließend mittels *Raycasting*, d.h. innerhalb der Unity-Umgebung wird an der per GNSS geschätzten Position ein virtueller Strahl in vertikale Richtung auf die importierten digitalen Geländemodelle als Kollisionsobjekte geschickt. Der Kollisionspunkt mit diesen Modellen entspricht dann der geschätzten Höhe an dieser geografischen Position. Die manuelle Kalibrierung der Gerätehöhe sowie Geräteposition und -orientierung kann dann anschließend mittels Interaktionsgesten auf Basis dieser initialen Schätzung der Gerätepose erfolgen.

5.3 Testmethode

Mithilfe der entwickelten mobilen Evaluationsanwendung konnte schließlich eine Evaluation des Registrierungsverfahrens in einer Feldstudie durchgeführt werden. Im Folgenden soll die Methodik, die bei der Evaluation eingesetzt wurde, vorgestellt werden. Dazu sollen zunächst die eingesetzten Evaluationsmethoden und Bewertungskriterien erläutert und anschließend der Ablauf der Evaluation sowie die individuellen Charakteristiken der einzelnen Test-Standorte beschrieben werden.

5.3.1 Evaluationsmethoden und Bewertungskriterien

Die Bedeutung der wissenschaftlichen Evaluation von AR-Anwendungen – insbesondere hinsichtlich der Nutzererfahrung und Nutzerfreundlichkeit – hat in den vergangenen Jahren zugenommen, u.a. weil AR-Anwendungen in zunehmend ausgereifter und marktreifer Form realisiert werden [73]. Die dazu eingesetzten Evaluierungsmethoden unterscheiden sich jedoch und sind in der Regel abhängig von der Fragestellung [37]. Meist wird zur AR-Evaluation eine Mischung aus **qualitativen** (z.B. Umfrage, Fragebögen) und **quantitativen** Methoden (z.B. Messung von Zeit oder Genauigkeit) genutzt [73]. In Anlehnung an die von Dünser und Billinghurst vorgeschlagenen Methoden für eine nutzerorientierte Evaluation [37] sollen im Kontext der hier durchgeführten Tests die folgenden drei Evaluationsmethoden eingesetzt werden:

- **Objektive Bewertungen** (*objective measurements*): Diese Bewertungen basieren auf quantitativen und reproduzierbaren Beobachtungen und Messwerten. Auf diese Weise sollen die Genauigkeit (Geräteposition und -orientierung) sowie der Zeitaufwand der entwickelten Kalibrierungsmethode gemessen werden.
- **Subjektive Bewertungen** (*subjective measurements*): Diese Bewertungen basieren auf subjektiven Beurteilungen der teilnehmenden Testpersonen und werden z.B. mithilfe von Fragebögen ermittelt. Auf diese Weise soll in dieser Arbeit die Bedienbarkeit des entwickelten Registrierungsverfahrens beurteilt werden.
- **Qualitative Analyse** (*qualitative analysis*): Bei einer qualitativen Analyse werden Daten durch strukturierte Beobachtungen oder durch Interviews oder Fragerunden gesammelt. Diese Evaluationsmethode wurde eingesetzt, um durch allgemeine Beobachtungen während der Tests durch den Evaluationsleiter sowie einer abschließenden Fragerunde an die teilnehmenden Personen zusätzliche Erkenntnisse hinsichtlich der Bedienbarkeit zu sammeln.

Die im Rahmen dieser Evaluationsmethoden genutzten Bewertungskriterien sollen im Folgenden im Detail erläutert werden.

Objektive Bewertungen (Genauigkeit und Zeitaufwand)

Im Rahmen der Evaluation sollen durch objektive Bewertungen die Genauigkeit der Kalibrierungsmethode, d.h. die Genauigkeit der globalen Kameraausrichtung und -position nach erfolgter Kalibrierung, sowie der für die Kalibrierung benötigte Zeitaufwand objektiv gemessen werden. In wissenschaftlichen Evaluationen von mobilen Registrierungsmethoden erfolgt die Bestimmung der Genauigkeit der gemessenen Geräte-Pose

in der Regel durch Vergleich mit der tatsächlichen idealen Geräte-Pose (*Ground Truth*), die oftmals vorab auf separatem Weg (z.B. händisch) bestimmt wird [117, 122].Auch in der folgenden Evaluation soll eine manuelle Bestimmung der tatsächlich angenommenen Geräte-Pose erfolgen.

Als Evaluationsinstrument zur Bestimmung der Abweichungen von gemessener und tatsächlicher Pose dient die digitale Aufzeichnung der Interaktionen und internen Berechnungen innerhalb der Anwendung auf dem Testgerät. Die Aufzeichnung erfolgt unsichtbar im Hintergrund durch Speichern wesentlicher Parameter als Textdatei (Daten-Logging) sowie durch Aufnahme des gesamten Bildschirmes als Videodatei (Screen-Capturing).

Fehler der Geräteorientierung ($\Delta\alpha, \Delta\beta$) Die tatsächliche Abweichung der kalibrierten globalen Geräteorientierung von der tatsächlichen globalen Geräterotation kann nicht in Echtzeit während der Testdurchführung bestimmt werden. Zwar blicken alle Testpersonen bei Kalibrierungsstart grob in die gleiche Richtung, die genaue tatsächliche reale Orientierung des Testgeräts ist jedoch bei Start des Testdurchgangs unbekannt bzw. nicht in jedem Durchgang genau identisch. Daher erfolgt eine Fehlerbestimmung der kalibrierten Geräteorientierung im Nachgang auf Basis der aufgenommenen Bildschirmvideos und mithilfe der visualisierten Validierungsobjekte. Dazu wird für jeden Testdurchlauf aus Videoaufnahmen ein Einzelframe mit den nach der Kalibrierung sichtbaren Windenergieanlagen exportiert und mit einer Aufnahme möglichst ideal platzierter Windenergieanlagen (*Ground Truth*) verglichen (siehe Abbildung 5.5). Die Generierung der Ground-Truth-Aufnahmen für jeden Standort erfolgte im Nachgang händisch durch den Evaluationsleiter ebenfalls mithilfe der Evalations-App. Hierzu wurde die globale Geräteposition bei der Kalibrierung jeweils auf die festgelegte globale Position der Test-Standorte (siehe Abbildung 5.9) fixiert und anschließend eine sehr sorgfältige manuelle Bestimmung der globalen Geräteorientierung anhand von texturierten Oberflächenmodellen durchgeführt. Daraus resultierte eine nahezu ideale Angleichung der 3D-Geodatenmodelle an die reale Umgebung. Die damit erzeugten AR-Visualisierungen der Validierungsobjekte sollen somit als möglichst ideale Ground-Truth-Aufnahmen betrachtet werden.

Mithilfe einer Bildbearbeitungssoftware können anschließend die Abweichungen der platzierten Windenergieanlagen zwischen Testaufnahme und Ground-Truth-Aufnahme pixelgenau (am Mittelpunkt der Rotoren) händisch in vertikaler und horizontaler Richtung bestimmt werden. Mit Kenntnis der Bildschirmbreite und -höhe w, h in Pixeln sowie des horizontalen und vertikalen Sichtfeldes (*Field of View*) θ_h, θ_v des

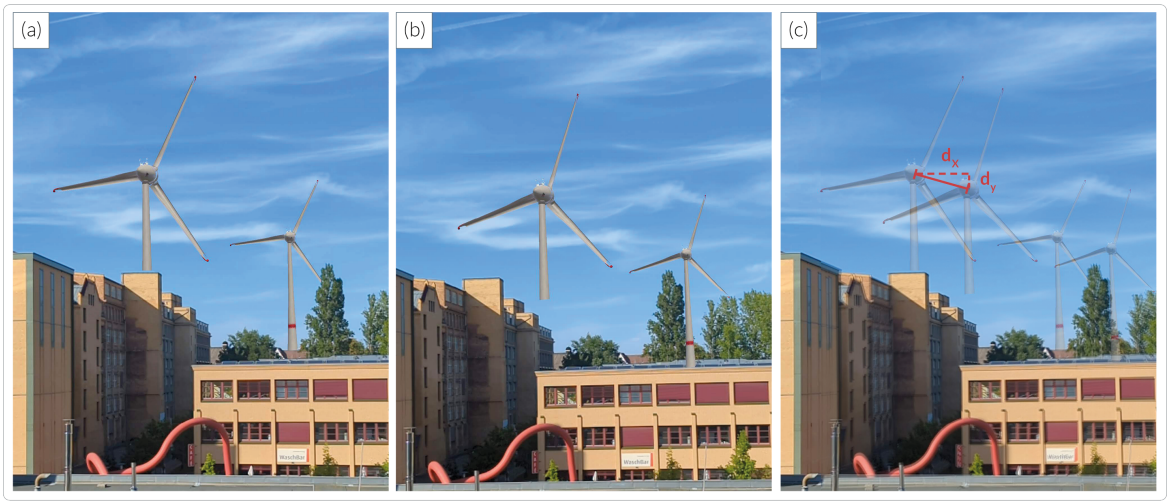


Abbildung 5.5 Manuelle Bestimmung der Abweichung der Geräteorientierung per Bildvergleich. Die erstellte Ground-Truth-Aufnahme mit möglichst idealer Kalibrierung (a) wird mit einer von Testpersonen erstellten Kalibrierung (b) verglichen. Durch Überlagerung beider Aufnahmen können die Abweichungen der Validierungsobjekte händisch in Pixeln gemessen werden (c). Die in der dargestellten Aufnahme gemessene Abweichung der kalibrierten Geräteorientierung beträgt etwa 3,5 Grad in horizontaler Richtung und etwa 1 Grad in vertikaler Orientierung (Bildausschnitt ist vergrößert dargestellt).

Testgeräts kann dann auf Basis der gemessenen Abweichungen d_x, d_y die Abweichung von kalibrierter und realer Geräteorientierung als Winkel in horizontaler Richtung ($\Delta\alpha$) und vertikaler Richtung ($\Delta\beta$) vereinfacht wie folgt bestimmt werden:

$$\Delta\alpha = \theta_h \frac{d_x}{w} \qquad \Delta\beta = \theta_v \frac{d_y}{h} \tag{5.1}$$

Die Abweichung der initialen IMU-basierten globalen Geräteorientierung von der realen Geräteausrichtung kann auf einfache Weise im Anschluss abgeschätzt werden, da der Grad der manuell im Laufe des Kalibrierungsvorgangs erfolgten Korrektur der Ausrichtung per Daten-Logging aufgezeichnet wird.

Fehler der Geräteposition (Δp) Im Gegensatz zur Fehlerabweichung der Geräteorientierung kann die Abweichung der kalibrierten globalen Geräteposition von der realen geografischen Position automatisiert auf Basis der per Daten-Logging aufgezeichneten Positionswerte ausgelesen werden. Die Bestimmung der realen idealen zweidimensionalen Geräteposition $p_{UTM} = [x_{UTM}, y_{UTM}]$ in UTM-Koordinaten basiert auf manueller, sehr sorgfälter kartenbasierter Bestimmung bei der Festlegung der genauen Test-Standorte (siehe Kapitel 5.3.3). Der Fehler der globalen Geräteposition

Δp ist dann einfach die euklidische Distanz zwischen der während der Tests kalibrierten geografischen Position der virtuellen AR-Umgebung $p_{AR} = [x_{AR}, y_{AR}]$ und der realen Position p_{UTM}:

$$\Delta p = \sqrt{(x_{UTM} - x_{AR})^2 + (y_{UTM} - y_{AR})^2} \tag{5.2}$$

Es ist dabei festzuhalten, dass die reale globale Geräteposition in den Testdurchläufen nicht immer zentimetergenau mit der vorab per Kartenansicht bestimmten Position der Test-Standorte übereinstimmt. Zwar wurde während der Tests darauf geachtet, dass alle Testpersonen den gewählten Test-Standort möglichst genau einhalten, von leichten Abweichungen aufgrund von leichten Bewegungen, unterschiedlichen Armlängen o.ä. muss jedoch ausgegangen werden. Eine spätere hochpräzise Analyse der Messergebnisse mit Genauigkeiten im Zentimeter-Bereich ist daher nicht sinnvoll.

Fehler der Gerätehöhe (Δh) Die Bestimmung der Abweichung der kalibrierten globalen Gerätehöhe von der realen globalen Höhe soll separat neben der Fehlerbestimmung der UTM-basierten 2D-Geräteposition und ebenfalls automatisiert mithilfe der per Daten-Logging aufgezeichneten Kalibrierungswerte erfolgen. Die Bestimmung der realen globalen Gerätehöhe erfolgt für jeden Teststandort manuell auf Basis einer interaktiven Abfrage der Geländehöhe per Online-Tool (*ATKIS DGM*) [20] auf Grundlage eines hochaufgelösten Geländemodells an der genauen Position des Test-Standortes. Hierbei wird auch der pauschal geschätzte Höhenversatz des Testgerätes über der Geländeoberfläche während der Testdurchführung (1,5 Meter) berücksichtigt. Der Test-Standort 2 befindet sich zudem nicht direkt auf Geländehöhe, sondern auf Höhe des zweiten Obergeschosses eines Gebäudes. Die reale globale Gerätehöhe an diesem Standort wurde in der Außenansicht des Gebäudes und mithilfe eines Zollstocks als Referenzhöhe abgeschätzt. Die geschätzte Abweichung Δh der globalen Gerätehöhe ist dann die absolute Differenz zwischen der im Rahmen der Kalibrierung bestimmten Höhe und der im Nachgang bestimmten real geschätzten globalen Gerätehöhe.

Zeitaufwand (t) Die Bestimmung des Zeitaufwandes t, der zur Durchführung der Kalibrierung benötigt wird, erfolgt automatisiert auf Basis der per Daten-Logging aufgezeichneten Zeitstempel bei Start und bei Beendigung des Kalibrierungsvorgangs. Falls eine Vorab-Positionsbestimmung per Kartenansicht erfolgt, wird die dafür benötigte Zeit ebenfalls zum Zeitaufwand der Kalibrierung gezählt.

Subjektive Bewertungen (Bedienbarkeit)

Im Gegensatz zu den objektiven Bewertungskriterien wird die subjektive Bewertung der Registrierungmethode durch die persönliche Wertung der Testperson beeinflusst. Bei der Evaluation soll durch eine subjektive Bewertungensmethode die Bedienbarkeit der Methode an den Test-Standorten bei Verwendung unterschiedlicher Geodatenmodellarten abgefragt werden, und zwar mithilfe eines Fragebogens.

Nach dem Absolvieren von Kalibrierungsvorgängen mit allen am Standort verfügbaren Geodatenmodellarten wird den Teilnehmenden ein Fragebogen ausgehändigt, um die Schwierigkeit der Kalibrierung bei Nutzung der verschiedenen Modellarten am Standort subjektiv abzufragen. Dabei sind jeweils fünf verschiedene Antwortmöglichkeiten möglich (siehe Abbildung 5.6). Zudem muss von der Testperson angegeben werden, mit welchem Testgerät getestet wurde und ob eine Vorab-Positionsbestimmung per Kartenansicht genutzt wurde. So sind bei der Auswertung Rückschlüsse möglich, inwieweit diese Parameter die Schwierigkeit der Bedienung beeinflussen.

Qualitative Analyse

Durch qualitative Beobachtungen durch den Evaluationsleiter während der Testdurchläufe lassen sich weitere wesentliche Aussagen über die Schwierigkeiten bei der Nutzung der Registrierungsmethode ableiten. Auf diese Weise kann beispielsweise von außen

Fragebogen

Evaluation
Geodatenbasierte Lokalisierung für Outdoor Augmented Reality

Genutztes Endgerät: ☐ Smartphone ☐ Tablet **Standort:** _____

Genutzte Positionierungsart: ☐ per Karte ☐ per AR-Ansicht

Wie schwierig (intuitiv) habe ich die **AR-Kalibrierung** (Anpassung der Landschaftsmodelle im Kamerabild) empfunden?

	sehr schwierig (sehr unintuitiv)	schwierig (unintuitiv)	mittelschwierig	einfach (intuitiv)	sehr einfach (sehr intuitiv)
DOM					
DOM (farbig)					
CityModel					
CityModel (farbig)					
DGM nur Standort 5 (Fhagen)					

Abbildung 5.6 Fragebogen zur Evaluation der subjektiv empfundenen Bedienbarkeit der AR-Kalibrierung bei Nutzung verschiedener Geodatenmodellarten. Der Fragebogen wird von jeder Testperson nach Testen aller am Standort verfügbaren Geodatenmodellarten ausgefüllt.

erfasst werden, welche Interaktionsgesten durch die Testpersonen primär genutzt wurden oder ob eine Testperson die initiale Darstellung des virtuellen Landschaftsmodells problemlos oder nur mit Schwierigkeiten auf die reale Umgebung übertragen kann.

Weitere wesentliche Aussagen lassen sich aus einer abschließenden Fragerunde an die Testpersonen ableiten. Diese Feedback-Runden fanden am Ende der Test-Sessions statt und umfassten unter anderem die folgenden Fragen:

- Was waren mögliche Gründe, weshalb eine Kalibrierung als schwierig wahrgenommen wurde (z.B. schwieriges Auffinden von realen Referenzpunkten; verwirrendes, uneindeutiges virtuelles Landschaftsmodell)?
- War die Nutzung der Kalibrierung mit dem Smartphone oder Tablet einfacher?
- Waren die Touch-Gesten zur Modellangleichung intuitiv oder verwirrend?
- Wie wurde die manuelle Positionsbestimmung per Karte empfunden?
- An welchen Standorten war die Kalibrierung sehr einfach und wo sehr schwierig?

5.3.2 Evaluationsablauf

Die Evaluation fand im Rahmen von drei Test-Sessions an zwei Tagen im Juni bzw. Juli 2022 bei klarem und sonnigen Wetter statt. An den einzelnen Test-Sessions nahmen unterschiedliche Personengruppen mit jeweils drei bis vier Testpersonen teil (siehe Abbildung 5.7). Insgesamt waren an der Feldstudie neben dem Evaluationsleiter elf Testpersonen (neun männlich; zwei weiblich) im Alter von 25 bis 36 Jahren beteiligt. Darunter befanden sich einige Studierende und Absolventen der Umweltinformatik, aber auch einige Personen in Berufen ohne Informatikbezug. Alle Testpersonen waren grundsätzlich mit der Bedienung von Smartphones und Tablets vertraut. Zwei Testpersonen kannten die Kalibrierungsmethode bereits, zumindest in Grundzügen. Den übrigen teilnehmenden Personen war die Testanwendung völlig unbekannt. Als Testgeräte standen ein Tablet (Samsung Tab S7 Plus; 12,4 Zoll) sowie ein Smartphone (Google Pixel 6 Pro; 6,7 Zoll) mit Android-Betriebssystem und vorinstallierter Evaluations-Anwendung zur Verfügung.

In jeder Test-Session wurden Test-Standorte mit der jeweils teilnehmenden Testgruppe nacheinander besucht. In der ersten und zweiten Test-Session konnten aufgrund von Zeitbeschränkungen der Teilnehmer:innen nicht alle vier Test-Standorte besucht werden (siehe Abbildung 5.7). Die Tests fanden dann nur an den Test-Standorten auf dem Campus der HTW Berlin und deren Umgebung statt (siehe auch Kapitel 5.3.3). Bei der dritten Test-Session wurden Test an allen Standorten durchgeführt. Dabei

	Test-Session #1	Test-Session #2	Test-Session #3
Datum	29.06.2022	29.06.2022	30.07.2022
Testpersonen	4 männlich (Alter: 25-31J)	2 männlich, 1 weiblich (Alter: 25-30J)	3 männlich, 1 weiblich (Alter: 34-36J)
Testdurchläufe	44	24	60
Standorte	1 → 2 → 3	1 → 2	4 → 3 → 2 → 1

Abbildung 5.7 Durchgeführte Test-Sessions zur Evaluation der Kalibrierungsmethode

fand die Abfolge der Standorte in umgekehrter Reihenfolge statt, um mögliche Gewöhnungseffekte mit zunehmender Nutzung der Kalibrierungsmethode an nachfolgenden Standorten möglichst auszugleichen. Vor Beginn einer Test-Session fand zudem stets eine kurze Einweisung durch den Evaluationsleiter statt, in der die Grundidee und die Bedienung der Anwendung erläutert wurden.

Der grundsätzliche Ablauf der Evaluation an jedem Standort ist in Abbildung 5.8 skizziert: Jede Testperson führt am Standort mit jeder jeweils am Standort zur Verfügung stehenden Geodatenmodellart einen Testdurchlauf der Kalibrierungsmethode (siehe Abbildung 5.2) durch. Die Tests erfolgen für alle Teilnehmenden von derselben Position aus und mit möglichst derselben, vom Evaluationsleiter vorgegebenen Blickrichtung. So wird sichergestellt, dass am Ende einer erfolgreichen Kalibrierung die voreingestellten Validierungsobjekte stets im Blickfeld sind und dass alle Testpersonen von Beginn an dieselben Orientierungspunkte in der realen Landschaft einsehen. Das dabei verwendete Testgerät sowie die einzusetzende Positionierungsart (mit oder ohne Positionsbestimmung per Karte) wird vom Evaluationsleiter bestimmt, wobei darauf geachtet wurde, dass diese für eine Testperson im Laufe der Evaluation variieren. Nach entsprechender Konfiguration der Evaluationsanwendung durch den Evaluationsleiter startet die Testperson dann einen Testdurchlauf. Das heißt, sie nimmt die vorgesehene Position ein und absolviert den Registrierungsvorgang mit einer bestimmten Geodatenmodellart. Der Evaluationsleiter achtet dabei darauf, dass die vorgesehene Standort-Position möglichst genau eingehalten wird. Er verfolgt außerdem die Durchführung der Kalibrierung durch die Testperson und notiert seine Beobachtungen.

Nach erfolgter Kalibrierung beginnt die Testpersonen einen neuen Testdurchlauf mit geänderter Geodatenmodellart. Wurden alle verfügbaren Modellarten von der Person getestet, wird ihr ein Fragebogen ausgehändigt, auf dem die Schwierigkeit der Kalibrierung bei Verwendung der jeweiligen Modellarten subjektiv bewertet wird (siehe Abbildung 5.6). Wurden am Standort Testdurchläufe von allen Testpersonen absolviert,

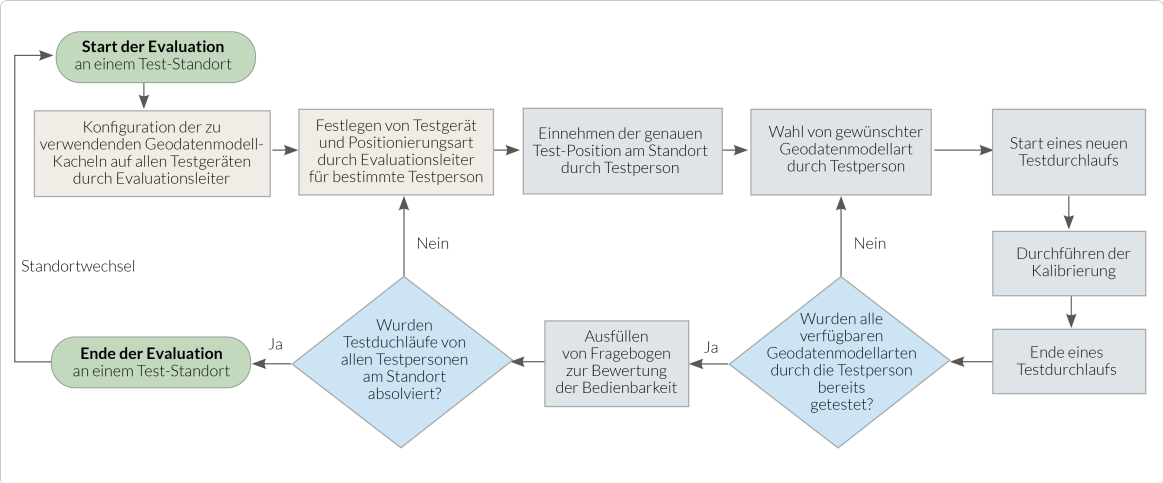


Abbildung 5.8 Ablauf der Evaluation der Kalibrierungsmethode durch eine Testperson an einem Test-Standort.

wird ein neuer Standort aufgesucht. Am Ende jeder Test-Session findet schließlich die abschließende Feedback-Runde statt, in der die Testpersonen die Erfahrungen über alle besuchten Test-Standorte hinweg reflektieren können.

5.3.3 Test-Standorte

Die Evaluation fand an vier verschiedenen Test-Standorten in Berlin statt. Die Standorte wurden so gewählt, dass unterschiedliche Umgebungen mit sowohl städtischer als auch naturräumlicher Charakteristik abgedeckt werden (siehe Abbildung 5.9). So kann getestet werden, inwieweit die verschiedenen Geodatenmodellarten für die Nutzung in unterschiedlichen Umgebungen geeignet sind. Die Nutzung von **digitalen Oberflächenmodellen**, welche sowohl Vegetation als auch Gebäude abbilden können, ist an jedem Test-Standort möglich. Der Einsatz von **digitalen 3D-Stadtmodellen** ist möglich, sobald Gebäude beliebiger Art (Wohnhäuser, Industriegebäude, Kirchtürme etc.) im Blickfeld sichtbar sind. Dies ist bei drei der vier gewählten Test-Standorte möglich. **Digitale Geländemodelle** können hingegen nur dann sinnvoll als Kalibrierungsmodelle verwendet werden, wenn deutliche Geländeerhebungen (Hügel, Berge) im Umkreis erkennbar sind. Dies ist nur beim Test-Standort 4 der Fall. Die Charakteristik der vier Test-Standorte lässt sich wie folgt beschreiben:

Standort 1 (HTW1): Der Standort befindet sich in urbaner Umgebung am Geländer eines Fußwegs unmittelbar am Spree-Ufer auf dem Campus der HTW Berlin. Die Tests erfolgen mit Blickrichtung auf die gegenüberliegende Uferseite. Dort sind

Standort 1 (HTW-1)

Ort	Direkt an einem Geländer am Spreeufer auf dem Campus der HTW Berlin in Berlin-Oberschöneweide
Position & Höhe	Position (UTM 33N): 399787,0 / 5812627.7 Höhe: 33,5 m
Eingesetzte Geodatenmodellarten	✓ CityModel ✓ CityModel_RGB ✓ DOM ✓ DOM_RGB ✗ DGM
Charakteristik	Urbane Umgebung; freies Blickfeld bis zur gegenüberliegenden Flussseite in etwa 150 Entfernung mit nebeneinander liegenden Häusern und vereinzelten Baumgruppen mit markanten Baumkronen.

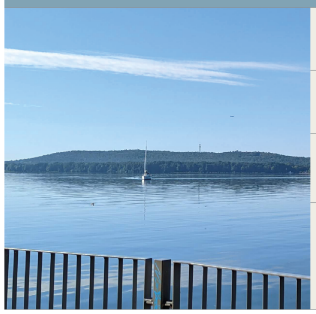
Standort 2 (HTW-2)

Ort	An einem geöffneten Fenster eines Büroraums im 2. Obergeschoss eines Gebäudes auf dem Campus der HTW Berlin in Berlin-Oberschöneweide
Position & Höhe	Position (UTM 33N): 399833,6 / 5812739,8 Höhe: 44,3 m (Geländehöhe: 34,8 m)
Eingesetzte Geodatenmodellarten	✓ CityModel ✓ CityModel_RGB ✓ DOM ✓ DOM_RGB ✗ DGM
Charakteristik	Urbane Umgebung; diverse Gebäude in unterschiedlichen Ausrichtungen und Gebäudehöhen sowie einzelne hohe Bäume sichtbar; Standort nicht auf Geländehöhe, sondern etwa 10 Meter über dem Boden.

Standort 3 (Wuhlheide)

Ort	Mitten auf einer großen Wiese innerhalb des städtischen Waldgebiets der Wuhlheide in Berlin-Oberschöneweide
Position & Höhe	Position (UTM 33N): 399760,0 / 5814476,5 Höhe: 34,6 m
Eingesetzte Geodatenmodellarten	✓ CityModel ✓ CityModel_RGB ✓ DOM ✓ DOM_RGB ✗ DGM
Charakteristik	Ländliche, naturräumliche Gegend; freies Blickfeld auf große Waldwiese mit dichtem Wald in etwa 180 Meter Entfernung; einzelne markante Bäume und Sträucher in etwa 40 Meter sichtbar; in großer Entfernung lässt sich eine Kirchturmspitze oberhalb der Wald-Silhouette erahnen.

Standort 4 (Friedrichshagen)



Ort	Direkt an einem Geländer am Ufer des großen Müggelsees in Berlin-Friedrichshagen
Position & Höhe	Position (UTM 33N): 406646,7 / 5811357,3 Höhe: 33,3 m
Eingesetzte Geodatenmodellarten	✗ CityModel ✗ CityModel_RGB ✓ DOM ✓ DOM_RGB ✓ DGM
Charakteristik	Ländliche Gegend; freier Blick über den großen Müggelsee; im Hintergrund in ca. 3km Entfernung die Silhouette der Müggelberge (bis zu 115m hoch) deutlich erkennbar; in etwa 80m Entfernung einige markante Baumstrukturen am Rande des Blickfeldes; Gebäude nicht erkennbar.

Abbildung 5.9 Übersicht über alle vier genutzten Test-Standorte und deren jeweiliger Umgebungscharakteristik

Baumreihen sowie bunte Häuser mit markanten Dachformen deutlich sichtbar. Eine Kalibrierung mithilfe von Stadt- und Oberflächenmodellen anhand dieser Baum- und Häuserstrukturen als Referenzobjekte erscheint problemlos möglich.

Standort 2 (HTW2): Der Standort befindet sich an einem großen geöffneten Fenster eines Büroraums im zweiten Obergeschoss eines Gebäudes auf dem Campus der HTW Berlin, ebenfalls in urbaner Umgebung. Im Blickfeld befinden sich mehrere Gebäude in unterschiedlicher Ausrichtung sowie einige hohe auffällige Bäume. Der Standort ist zwar nur über Innenräume zugänglich, der Blick zeigt jedoch ausschließlich auf Außenbereiche. Eine Kalibrierung mithilfe von Stadt- sowie Oberflächenmodellen erscheint daher gut realisierbar zu sein. Auch zuverlässiger GNSS-Empfang ist gegeben. Der Ort wurde vor allem deshalb als Test-Standort gewählt, weil er sich nicht auf Geländeniveau, sondern knapp 10 Meter über der Geländeoberfläche befindet. Eine manuelle Höhenbestimmung durch die Testpersonen ist daher hier in jedem Fall notwendig.

Standort 3 (Wuhlheide): Der Standort befindet sich in naturnaher Umgebung auf einer großen Wiese innerhalb eines Waldgebietes. Die Tests erfolgen mit freiem Blick über die Wiese auf einen etwa 200 Meter entfernten Wald. Die Silhouette des Waldes am Horizont weist keine auffälligen Strukturen auf. Eine Kalibrierung nur anhand dieser Waldsilhouette könnte daher möglicherweise schwerfallen. Leicht abseits befinden sich in kurzer Entfernung weitere kleine Bäume. Außerdem ist bei genauem Hinsehen die Spitze eines Kirchturmes hinter dem Wald erkennbar. Neben der Nutzung von Oberflächenmodellen ist daher – trotz des naturnahen Charakters – auch die Nutzung von Gebäudemodellen möglich. Der Standort wurde ausgewählt, um die Kalibrierungsmethode auch in freier, naturnaher und relativ gleichförmiger Umgebung (homogene Wiese und Waldsilhouette) beurteilen zu können.

Standort 4 (Friedrichshagen): Der Standort befindet sich an einem Geländer unmittelbar am Rande des großen Müggelsees mit freiem Blick über den See. Im Hintergrund ist in etwa 3 Kilometer Entfernung die Hügelkette der Müggelberge gut erkennbar. Als einziger Standort ist daher an dieser Stelle auch eine Kalibrierung mithilfe eines digitalen Geländemodells möglich. In etwa 80 Meter Entfernung sind abseits außerdem einige markante Baumreihen am Ufer sichtbar. Gebäude sind nicht erkennbar. Ein Einsatz von 3D-Stadtmodellen ist daher am Standort nicht möglich.

5.4 Ergebnisse

Nach Auswertung der Testdurchläufe anhand der genannten Bewertungskriterien sollen im Folgenden die Ergebnisse der Evaluation präsentiert werden. Die Vorstellung der Ergebnisse berücksichtigt dabei stets sowohl die gemessenen Messwerte (objektive Bewertung) und Beurteilungen der Testpersonen (subjektive Bewertungen) als auch die Beobachtungen durch den Evaluationsleiter während der Testdurchführung (qualitative Analyse).

Es wurden insgesamt über alle Standorte betrachtet 128 Testdurchläufe mit unterschiedlichen Modellarten absolviert: jeweils 44 Tests an Standort 1 und Standort 2, 28 Testdurchläufe an Standort 3 und 12 Tests an Standort 4. Beispielhafte Bildschirmaufnahmen der Testdurchläufe an allen Test-Standorten während sowie bei Abschluss der Kalibrierung sind in Abbildung 5.10 dargestellt.

Von allen durchgeführten Testdurchläufen wurden 6 Kalibrierungsvorgänge nicht erfolgreich absolviert, die übrigen Kalibrierungen – also etwa 95 Prozent – wurden erfolgreich abgeschlossen. Bei den nicht erfolgreichen Tests war die Kalibrierung der Geräte-Pose nach Beendigung des Kalibrierungsprozesses derart misslungen, dass die Validierungsmodelle gar nicht sichtbar waren und damit keine vollständige Analyse der Testergebnisse möglich war. Der Grund lag in der Regel in einer deutlich inkorrekten initialen GNSS-basierten Standortbestimmung und damit einer sehr fehlerhaften Anfangsdarstellung der virtuellen Geodatenmodelle.

Genauigkeit der Geräteorientierung

Die im Durchschnitt erreichte Genauigkeit der kalibrierten vertikalen und horizontalen Geräteausrichtung an den einzelnen Standorten ist in Abbildung 5.11 dargestellt. Es zeigt sich, dass an allen Orten sowie mit allen eingesetzten Modellarten eine erfolgreiche Kalibrierung der Geräteorientierung möglich war.

Die **initiale horizontale Abweichung** $\Delta\alpha$ der globalen Orientierung war in nahezu allen Fälle erwartungsgemäß sehr fehlerhaft, was auf die Ungenauigkeit des integrierten Magnetometer-Sensors zurückzuführen ist (siehe auch Kapitel 2.2.3). In einzelnen Testdurchläufen konnte an Standort 3 und Standort 4 die Kalibrierung zwar mit einer relativ genauen initialen IMU-basierte Ausrichtung (etwa 5-10 Grad Abweichung) begonnen werden, im Mittel war jedoch eher eine horizontale Abweichung der initialen Blickrichtung von etwa 15-20 Grad üblich. Die anschließende Verschiebung der Modelle per One-Finger-Drag-Geste klappte schließlich in den meisten Fällen sehr

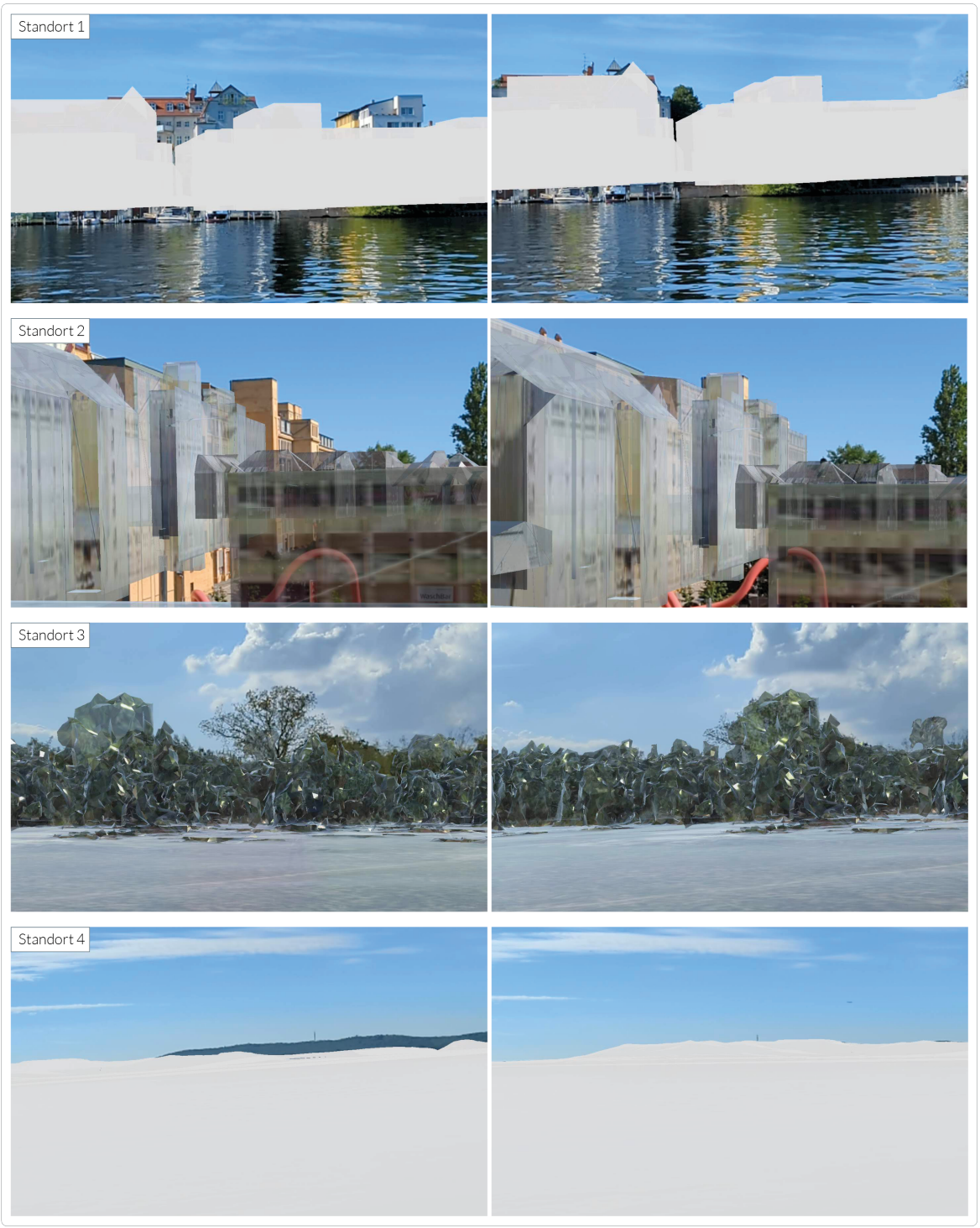


Abbildung 5.10 Beispielhafte Bildschirmaufnahmen während der Tests an den vier Test-Standorten, jeweils während der Kalibrierung (links) sowie bei Abschluss der Kalibrierung (rechts). Die Beispiel-Aufnahmen zeigen den Einsatz eines untexturierten Stadtmodells (City-Model) an Standort 1, eines texturierten Stadtmodells (CityModel_RGB) an Standort 2, eines texturierten Oberflächenmodells (DOM_RGB) an Standort 3 sowie eines Geländemodells (DGM) an Standort 4.

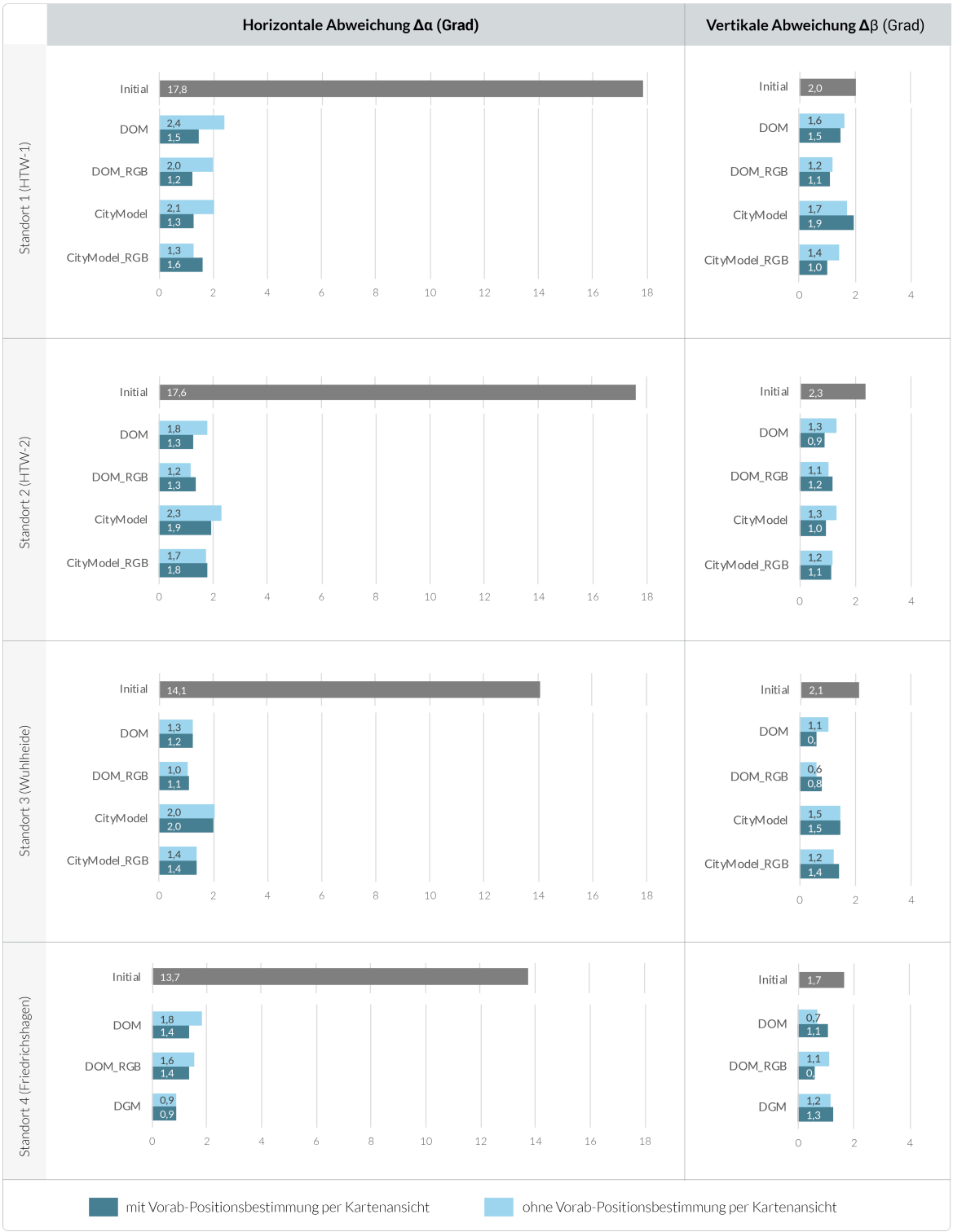


Abbildung 5.11 Gemessene durchschnittliche horizontale und vertikale Abweichung der kalibrierten globalen Geräterotation.

genau, sodass üblicherweise die horizontale Orientierung auf etwa 1,5 Grad genau kalibriert werden konnte.

Die Schätzung der **initialen vertikalen Orientierung** auf Basis der IMU-Sensoren konnte hingegen deutlich genauer erfolgen. Die vertikale Abweichung $\Delta\beta$ der virtuellen Umgebung zu Beginn der Tests betrug in der Regel kaum mehr als 2 Grad. Eine Verschiebung der Modelle in vertikaler Richtung ist daher meist kaum mehr notwendig. Die erreichte Genauigkeit der vertikalen Orientierung nach erfolgter Kalibrierung betrug im Mittel auch etwa 1,5 Grad.

Es zeigte sich, dass eine genaue Kalibrierung mit allen **Geodatenmodellarten** möglich ist, wobei es tendenziell beim Einsatz untexturierter Modellarten eher zu einzelnen fehlerhaften Ausrichtungen kommt, und der Einsatz texturierter Modelle, insbesondere des texturierten Oberflächenmodells, tendenziell eine etwas höhere Genauigkeit erreichen kann. Eine sehr hohe Genauigkeit konnte ebenfalls beim Einsatz des digitalen Geländemodells an Standort 4 erreicht werden (etwa 1 Grad Abweichung).

Eine genaue Kalibrierung der Geräteausrichtung ist ebenfalls an allen **Standorten** möglich. Dies liegt daran, dass an allen gewählten Orten durch die Testpersonen geeignete **Referenzobjekte** zur Ausrichtung der Blickrichtung gefunden werden konnten. Meist handelte es sich dabei um markante Silhouetten im Kamerabild durch Bäume (bei Oberflächenmodellen) oder Strukturen von Gebäudedächern (bei Stadtmodellen und Oberflächenmodellen). Als primäre Referenz zum Abgleich der Orientierung diente dabei stets die Horizontlinie, d.h. die Silhouette der sichtbaren Objekte am Horizont. Geeignete Punkte am Horizont zum Abgleich der Orientierung fanden sich in der Regel sehr schnell, lediglich in Ausnahmefällen, z.B. bei der Identifikation eines weit entfernten Kirchturms an Standort 3, musste das Modell zunächst mehrere Male hin- und hergeschoben werden, um mögliche Referenzpunkte zu finden und abzugleichen. Insgesamt betrachtet erfolgte die Kalibrierung der Geräteorientierung jedoch an allen Standorten ähnlich genau. Das heißt, sofern geeignete Referenzpunkte vorhanden waren – was bei allen Standorten der Fall war – ist kein signifikanter Vorteil eines einzelnen Standortes erkennbar.

Als wesentliche Beobachtung während der Testdurchläufe zeigte sich vor allem, dass die erreichte Genauigkeit bei der Kalibrierung der Geräteorientierung weniger von der Genauigkeit der initialen Orientierungsschätzung abhängt, sondern vielmehr davon, ob die **initiale Positionierung** des Geräts genau oder weniger genau erfolgte. Bei relativ genauer initialer Positionsschätzung werden die virtuellen Landschaftsmodelle bereits so in die AR-Ansicht projiziert, dass für eine gute Kalibrierung lediglich eine einfache Verschiebung des Modells zur Korrektur der Blickrichtung erfolgen muss. Bei ungenauer

initialer Positionierung erscheint die initiale AR-Projektion des Landschaftsmodells hingegen fehlerhaft und damit verwirrend. Den Testpersonen erscheint es dann – auch nach mehrmaligem Hin- und Herschieben – in vielen Fällen unmöglich, das Modell korrekt anzugleichen. Die Folge ist dann oftmals eine ungenauere Orientierungskalibrierung, sowohl in horizontaler als auch vertikaler Richtung. Demzufolge wird bei Einsatz der kartenbasierten Vorab-Positionsbestimmung tendenziell auch eine genauere Kalibrierung der Geräteorientierung erreicht, da diese Art der Positionsbestimmung in der Regel eine genauere Bestimmung der initialen Positionierung ermöglicht und somit auch eine genauere Kalibrierung der Geräteorientierung erreicht als bei initialer AR-Projektion allein auf Basis der ungenaueren, rein GNSS-basierten Positionsschätzung.

Genauigkeit der Geräteposition

Die im Durchschnitt an allen Standorten gemessene Abweichung der kalibrierten globalen Geräteposition ist in Abbildung 5.12 dargestellt. Die **initiale GNSS-basierte Positionsschätzung** erreicht dabei je nach Standort im Mittel Genauigkeiten von etwa 3-6 Meter, wobei auf der freien Wiese an Standort 3 die höchste (3,4 Meter) und an Standort 2 die geringste (5,3 Meter) initiale Genauigkeit verzeichnet wird.

Mithilfe der manuellen **Vorab-Positionsbestimmung per Kartenansicht** kann diese Schätzung über alle Standorte hinweg deutlich präzisiert werden, sodass auf diese Weise die Ungenauigkeiten der globalen Geräteposition im Idealfall sogar auf unter 1 Meter reduziert werden können. Im Mittel erreicht diese Methode der Positionsbestimmung Genauigkeiten von etwa 1-2 Meter. Die Testpersonen gaben in der Gesamtbetrachtung an, dass die kartenbasierte Bestimmungsmethode zwar mitunter zeitaufwändig sei, da anfangs erst Orientierungspunkte der realen unmittelbaren Umgebung (z.B. einzelne Bäume, Wege oder Sitzbänke) auf den Satellitenbildern der Karte wiedergefunden werden müssten, die Methode jedoch ansonsten über alle Standorte hinweg gut und intuitiv bedienbar sei.

Die Testergebnisse bestätigen auch, dass die manuelle kartenbasierte Positionsbestimmung – im Gegensatz zur reinen GNSS-basierten Schätzung – in jedem Testdurchlauf gute Ergebnisse lieferte und nie grob fehlerhaft war. Lediglich an Standort 3 inmitten einer freien Wiese erwies sich die kartenbasierte Positionsbestimmung als herausfordernder und ungenauer als an anderen Standorten, da eindeutige Orientierungspunkte in der unmittelbaren Umgebung auf den Satellitenbildern der Kartenansicht fehlten. An diesem Standort war daher die Genauigkeit der kalibrierten Geräteposition auf Basis der manuellen kartenbasierten Positionsbestimmung kaum höher als auf Basis der GNSS-basierten Positionsschätzung.

Abbildung 5.12 Gemessener durchschnittlicher Fehler der kalibrierten globalen Geräteposition und Gerätehöhe.

Die aus den Messergebnissen abgeleiteten Erkenntnisse decken sich mit den beschriebenen Ausführungen hinsichtlich der Genauigkeit der Geräteorientierung: Eine genau **initiale Positionsbestimmung** bleibt sehr entscheidend, sowohl für eine Kalibrierung von globaler Orientierung als auch bei einer etwaigen **Kalibrierung der Position durch Interaktionsgesten** in der AR-Ansicht. Wird die initiale AR-Projektion aufgrund ungenauer Anfangsposition deutlich fehlerhaft dargestellt, erscheint für die Testperson auch keine erkennbare Optimierung der globalen Position durch AR-basierte Interaktionsgesten möglich. Beobachtungen während der Testdurchführungen zeigten, dass eine Positionsoptimierung durch die Zwei-Finger-Pinch-Geste von Testpersonen zumeist nur dann erfolgreich durchgeführt wird, wenn ein anvisiertes Referenzobjekt nahe gelegen ist und eine einzige Korrekturgeste in Kamera-Blickrichtung ausreicht, um eine deutlich erkennbare Optimierung der virtuellen Projektion zu erreichen. Ein wiederholtes Drehen der Kamera und eine Kalibrierung der Position in geänderter Blickrichtung wird von der Testperson in der Regel nicht durchgeführt, vor allem, da bei Ansicht der (fehlerhaften) AR-Projektion der virtuellen Landschaft oftmals gar nicht ersichtlich erscheint, in welche Richtung die Position korrigiert werden müsste, um die Projektion zu optimieren. Während der Tests fand demzufolge eine tatsächliche Positionsoptimierung auf Basis dieser Interaktionsgeste nur in wenigen Fällen statt. Wurde keine kartenbasierte Positionsbestimmung durchgeführt, stimmte also die kalibrierte Geräteposition dann meist unmittelbar mit der (ungenauen) initialen GNSS-basierten Positionsbestimmung überein.

Genauigkeit der Gerätehöhe

In Bezug auf die Genauigkeit der globalen Gerätehöhe zeigt sich, dass an den genutzten Standorten die Ungenauigkeiten der GNSS-basierten Höhenschätzung mithilfe der DGM-basierten automatischen Höhenkalibrierung sehr genau bestimmt werden können (siehe Abbildung 5.12). Die auf diese Weise bestimmte **initiale Gerätehöhe** vor Beginn des Kalibrierungsvorgangs ist zumeist kaum von der vermuteten tatsächlichen globalen Gerätehöhe (Ground Truth) zu unterscheiden. Eine zusätzliche manuelle Kalibrierung der virtuellen Umgebung per Zwei-Finger-Drag-Geste ist daher in der Regel gar nicht mehr notwendig und wurde von den Testpersonen in der Regel auch nicht durchgeführt. Beobachtungen während der Tests zeigten, dass eine zusätzliche manuelle Höhenkalibrierung meist nur dann vorgenommen wurde, wenn die Testperson auch die Zwei-Finger-Pinch-Geste zur manuellen Positionskalibrierung anwendete.

Die Tests an Standort 2 bildeten davon jedoch eine Ausnahme. Dort befanden sich die Testpersonen nicht auf Geländehöhe, sondern etwa 10 Meter oberhalb der Erdober-

fläche. Hier zeigte sich, dass die Interaktionsgeste zur **manuellen Kalibrierung der Gerätehöhe** gut und recht genau funktionierte (etwa 1,2 m Abweichung), sofern sich die Person der Problematik der anfangs fehlerhaften Gerätehöhe bewusst war. Diese Problematik an Standort 2 wurde während der Evaluation von den Testpersonen gut verstanden. In Einzelfällen war jedoch ein kurzer Hinweis durch den Evaluationsleiter erforderlich. Eine Höhenkalibrierung zur Angleichung der virtuellen Landschaft wurde demzufolge in allen Tests durchgeführt. Zu beobachten war dabei, dass es teilweise für die Testperson schwierig war, festzustellen, wann genau die korrekte Höhe erreicht wurde. Ein mehrmaliges Hin- und Herbewegen des virtuellen Modells nach oben und unten zum Ausprobieren verschiedener Höhen und ein daraus resultierender größerer Zeitaufwand war dementsprechend die Konsequenz.

Zeitaufwand

Der im Mittel benötigte Zeitaufwand zur Durchführung des Kalibrierungsvorgangs ist in Abbildung 5.13 dargestellt. Meist dauerte die Kalibrierung etwa 25 Sekunden. Es zeigte sich, dass der Zeitaufwand in der Regel etwas höher war, wenn als zusätzlicher Schritt die kartenbasierte Vorab-Positionsbestimmung durchgeführt wurde, wobei dabei – verständlicherweise – vor allem die erstmalige kartenbasierte Positionsbestimmung an einem Standort durch eine Testperson zeitaufwändig war, während in den folgenden Positionsbestimmungen am selben Standort ein Lerneffekt eintrat, da wesentliche Orientierungspunkte bereits identifiziert und bekannt waren.

Ein deutlicher Unterschied des benötigten Zeitaufwandes an den verschiedenen **Standorten** war ansonsten nicht bemerkbar. Lediglich an Standort 2 war im Mittel ein höherer Zeitaufwand erkennbar. Dort war zusätzlich in jedem Fall eine – mitunter zeitwaufändige – manuelle Höhenkalibrierung notwendig.

Der gemessene Zeitaufwand war zudem kaum von der genutzten **Modellart** abhängig. Eine Ausnahme stellte hier der Einsatz digitaler Geländemodelle an Standort 4 dar. Hier erfolgte die Kalibrierung durch Angleichen des Geländereliefs der sichtbaren Hügelkette nicht nur sehr präzise, sondern auch sehr schnell. Die Nutzung von CityModel-Arten hingegen erfolgte tendenziell mit leicht höherem Zeitaufwand. Dort kann eine Kalibrierung – im Gegensatz zu digitalen Oberflächen- und Geländemodelle - nicht immer allein durch Angleichen der Horizontlinie erfolgen, sondern es müssen zunächst einzelne korrespondierende Gebäude identifiziert und die Korrektheit deren virtueller Projektionen jeweils überprüft und ggf. korrigiert werden. Dies ist mutmaßlich etwas zeitaufwändiger als ein Angleichen der virtuellen Landschaft allein anhand einer durchgängigen Horizontlinie.

Abbildung 5.13 Gemessener durchschnittlicher Zeitaufwand zur Durchführung der Kalibrierung sowie subjektiv per Fragebogen bewertete Bedienbarkeit der Kalibrierungsmethode.

Ein höherer Zeitaufwand der Kalibrierung war auch dann zu beobachten, wenn die Testperson das Endgerät stark rotiert hat, um das virtuelle Landschaftsmodell in verschiedenen **Blickfeldern** zu betrachten und anhand mehrerer, weit voneinander entfernter Referenzen abzugleichen. Dies resultierte oftmals darin, dass das Modell dann bei anderem Blickwinkel wieder von einer optimalen Angleichung abwich und erneut ausgerichtet werden musste. Testpersonen, die sich bei der Kalibrierung hingegen stets auf ein kleineres Sichtfeld konzentrierten, konnten den Vorgang deutlich schneller abschließen.

In den Beobachtungen während der Testdurchläufe war schließlich auch hier erkennbar, dass der benötigte Zeitaufwand oft von der Qualität der **initialen Positionsbestimmung** abhing: Wurde das virtuelle Landschaftsmodell aufgrund ungenauer initialer Positionierung deutlich fehlerhaft in die AR-Ansicht projiziert, führte diese verfälschte Darstellung oftmals dazu, dass die Testperson unter allen Umständen und durch Aneinanderreihung einer Vielzahl an Interaktionsgesten versuchte, das Modell der realen Umgebung anzupassen. Dies resultierte dann in diesen Einzelfällen in einem deutlich höheren Zeitaufwand von 40 Sekunden und länger.

Bedienbarkeit

Die subjektive, per Fragebogen ermittelte Beurteilung der Bedienbarkeit der Kalibrierungsmethode ist ebenfalls in Abbildung 5.13 dargestellt. Es wird deutlich, dass in der Gesamtbetrachtung über alle Standorte hinweg die Nutzung des **texturierten Oberflächenmodells** als die Modellart wahrgenommen wird, die am einfachsten zu nutzen ist. An Standort 4 mit freier Sicht auf eine markante Hügelkette wurde außerdem der Gebrauch des digitalen **Geländemodells** als sehr intuitiv bewertet. Dort wurde als sehr angenehm angemerkt, dass zur erfolgreichen Kalibrierung eine Fokussierung allein auf die im Gelände deutlich abgrenzbare Hügelkette ausreiche, während bei anderen Modellarten verschiedenartige Oberflächenstrukturen (Bäume, Sträucher, Häuser etc.) für mehr Komplexität und eine eher schwierigere Handhabung sorgten.

Die Verwendung der **3D-Stadtmodelle** wurde von Fall zu Fall teilweise sehr unterschiedlich beurteilt. Während in vielen Fällen einige Testpersonen mit dieser Modellart sehr gut zurechtkamen, wurden die Gebäudemodelle in anderen Fällen – mitunter auch am gleichen Standort – als wenig intuitiv bewertet. Insgesamt lässt sich in jedem Fall festhalten, dass diese Modellart schwieriger zu nutzen ist bzw. mehr Nutzungserfahrung benötigt als andere Modellarten, bei denen ein Abgleich der gesamten Horizontlinie zur Kalibrierung genutzt werden kann. Die Bewertungen zur Bedienbarkeit der einzelnen Modellarten korrespondieren somit größtenteils mit den

erreichten Genauigkeiten der Ausrichtungskalibrierung bei Nutzung der entsprechenden Modellarten.

Es zeigt sich in der Gesamtbetrachtung, dass die Nutzung von **texturierten Modellen** etwas positiver bewertet wurde als die Verwendung der jeweils nicht-texturierten Modellarten. In den meisten Fällen ist mit den texturierten Modellen eine Repräsentation der physischen Umgebung mit höherem Realitätsgrad und damit ein einfacheres Abgleichen von realen und virtuellen Referenzpunkten und -objekten möglich als mit den untexturierten, einfarbigen Modellvarianten. In wenigen Einzelfällen an Standort 2 wurde lediglich geäußert, dass die Nutzung des texturierten Oberflächenmodells nachteilig gegenüber der untexturierten Variante sei, weil die Qualität der Textur sehr schlecht und damit die Texturdarstellung eher störend und verwirrend sei. Auch an Standort 3 wurde die Nutzung des texturierten Oberflächenmodells nicht von allen als intuitiv bewertet. Als Grund wurde angegeben, dass eine Referenzierung des Modells in der realen Umgebung aufgrund der sehr homogenen grünen Textur von Wald und Wiese verwirrend und schwierig sei.

Eine schwierige Bedienung wurde in manchen Fällen ebenfalls bescheinigt, wenn Objekte, die als **Referenzobjekte** zum korrekten Angleichen genutzt wurden, nicht korrekt im virtuellen Modell wiedergegeben wurden und damit für Verwirrung sorgten (z.B. ein zu niedriger Turm im Oberflächenmodell an Standort 2). Über alle Standorte hinweg war ansonsten zu beobachten, dass die Modelle in fast allen Fällen nur auf Basis der sichtbaren Horizontlinie angeglichen wurden. Wenn der Horizont sehr markant und wenig homogen beschaffen war, resultierte dies tendenziell in einer schnellen und einfachen Bedienung (Standorte 1, 2 und 4). Bei einer sehr homogenen Horizontlinie (Standort 3) wurde die Bedienung hingegen tendenziell als etwas schwieriger beurteilt. In diesen Fällen mussten unter Umständen zusätzlich noch andere Referenzobjekte (z.B. nahegelegende Bäume und Sträucher) einbezogen werden.

Mobiles Endgerät

Im Zuge der Evaluation zeigte sich, dass die Bedienung der Kalibrierungsmethode bei Einsatz eines Tablets mit großem Bildschirm nicht leichter fällt als auf einem handelsüblichen Smartphone mit deutlich kleinerem Display. In den abschließenden Feedbackrunden äußerte stattdessen eine Mehrzahl der Testpersonen, dass die Bedienung der Test-Anwendung auf dem Tablet schwieriger sei. Das Testgerät müsse während der Tests die meiste Zeit aufrecht und fast auf Augenhöhe in Blickrichtung mit nur einer Hand gehalten werden, während die zweite Hand für die Durchführung der Kalibrierung benötigt werde. Aufgrund der Größe und des Gewichts sei dies bei

Tablet-Nutzung schwieriger als bei Verwendung des Smartphones, welches leichter mit nur einer Hand gehalten werden kann. Die Displaygröße des Smartphones (6,7 Zoll) wurde hingegen als ausreichend beurteilt, um die virtuellen Umgebungsmodelle in ausreichender Genauigkeit und Größe abzubilden.

5.5 Diskussion und Zusammenfassung

Die vorgestellten Ergebnisse der Praxis-Evaluation des entwickelten Registrierungsverfahrens lassen Möglichkeiten und Stärken, aber auch mögliche Limitierungen und Optimierungsmöglichkeiten des Verfahrens in Bezug auf das Anwendungspotenzial des Verfahrens im Praxiseinsatz erkennen. Im Folgenden sollen daher die Evaluationsergebnisse zusammenfassend diskutiert und die Richtigkeit der zuvor aufgestellten Evaluations-Hypothesen überprüft werden. Abschließend sollen daraus konkrete Empfehlungen zum Einsatz des Verfahrens abgeleitet werden.

5.5.1 Diskussion der Ergebnisse

Die folgende Diskussion der Messergebnisse und Beobachtungen der Evaluation soll zunächst aufzuzeigen, unter welchen Rahmenbedingungen die Entwicklungen in der Praxis am besten einsetzbar sind und welche Aspekte des Verfahrens gut funktionieren bzw. bestimmte Limitierungen aufweisen.

Genauigkeit der Registrierung

Die Evaluation konnte zeigen, dass mithilfe der entwickelten Methode bei geeigneten Rahmenbedingungen (z.B. geeignete Geodatenmodellart, gute initiale Positionierung) eine genaue Bestimmung der Geräte-Pose möglich ist. Sowohl die globale Geräteorientierung (etwa 1 Grad Abweichung) als auch die globale Geräteposition (etwa 1 Meter Abweichung) und Gerätehöhe (unter 1 Meter Abweichung) können dann unter guten Rahmenbedingungen recht genau bestimmt werden. Die Realisierung von vielfältigen Outdoor-AR-Anwendungen ist auf dieser Grundlage möglich. Es ist aber auch festzuhalten, dass bestimmte AR-Szenarien möglicherweise eine noch präzisere, zentimetergenaue Registrierung des AR-Systems benötigen. Dies wäre dann der Fall, wenn sich bereits leicht fehlerhafte AR-Projektionen auf dem Bildschirm als sehr problematisch erweisen (beispielsweise bei bestimmten AR-Szenarien zur Vermessung oder Bauüberwachung).

Initiale Geräteposition

Die Genauigkeit der initialen Schätzung der Geräteposition erweist sich als einer der wichtigsten Faktoren für ein einfaches und schnelles Gelingen des Kalibrierungsvorgangs. Bei einer genauen initialen Positionserfassung kann die AR-Projektion des virtuellen Landschaftsmodells zwar mit fehlerhafter Rotation, jedoch ohne sonstige Verzerrungen erfolgen, sodass durch einfache Verschiebungsgesten ein korrektes Angleichen des virtuellen Modells an die reale Landschaft möglich ist. Bei ungenauer initialer Positionserfassung hingegen wird die fehlerhafte und verzerrte initiale AR-Projektion der Geodatenmodelle oftmals als sehr verwirrend wahrgenommen. Die Interaktionsgesten zur Positionskorrektur, die dann zusätzlich für ein korrektes Angleichen des virtuellen Modells erforderlich sind, führen zwar in vielen Fällen zu einer Optimierung der globalen Position, werden aber als weniger intuitiv wahrgenommen, sind zeitaufwändig und gelingen meist nur solchen Nutzer:innen, die mit dem Kalibrierungsverfahren bereits sehr vertraut sind, d.h. die Methode bereits oftmals genutzt und die Funktionsweise und die Interaktionsgesten gut verinnerlicht haben (*erfahrene* Nutzer:innen).

Die hohe Bedeutung der Genauigkeit der initialen Geräteposition ist auch in Abbildung 5.14 deutlich erkennbar. Dort sind für den Test-Standort 1 die wahrgenommene Bedienbarkeit der Kalibrierung sowie die gemessene Genauigkeit der kalibrierten horizontalen Geräteorientierung in Bezug auf den initialen Positionierungsfehler dargestellt. Es wird deutlich, dass bei genauer initialer Positionserfassung die Bedienbarkeit der Kalibrierungsmethode als deutlich einfacher wahrgenommen wird und gleichzeitig auch die Geräteorientierung mit höherer Genauigkeit kalibriert werden kann.

Methoden der Positionsbestimmung

Eine manuelle Vorab-Positionsbestimmung per Kartenansicht ist daher in den meisten Fällen empfehlenswert. Diese Positionsbestimmung erwies sich in der Gesamtbetrachtung als genau, relativ einfach bedienbar und wenig zeitaufwändig. Probleme ergeben sich bei der kartenbasierten Positionsbestimmung nur in wenigen Fällen, nämlich dann, wenn eindeutig identifizierbare Orientierungspunkte in der unmittelbaren Umgebung fehlen (z.B. inmitten einer freien Wiese).

Die AR-basierte Positionskalibrierung mittels Zwei-Finger-Pinch-Geste ist wie erwähnt – im Vergleich zur kartenbasierten Positionsbestimmung – schwieriger zu bedienen und vor allem dann brauchbar, wenn eine einfache Korrekturgeste in Kamerablickrichtung genügt, was jedoch selten der Fall ist. Auch die Nutzung eines texturierten Landschaftsmodells ist in diesen Fällen sinnvoll, da dort besser sichtbar wird, in welchen

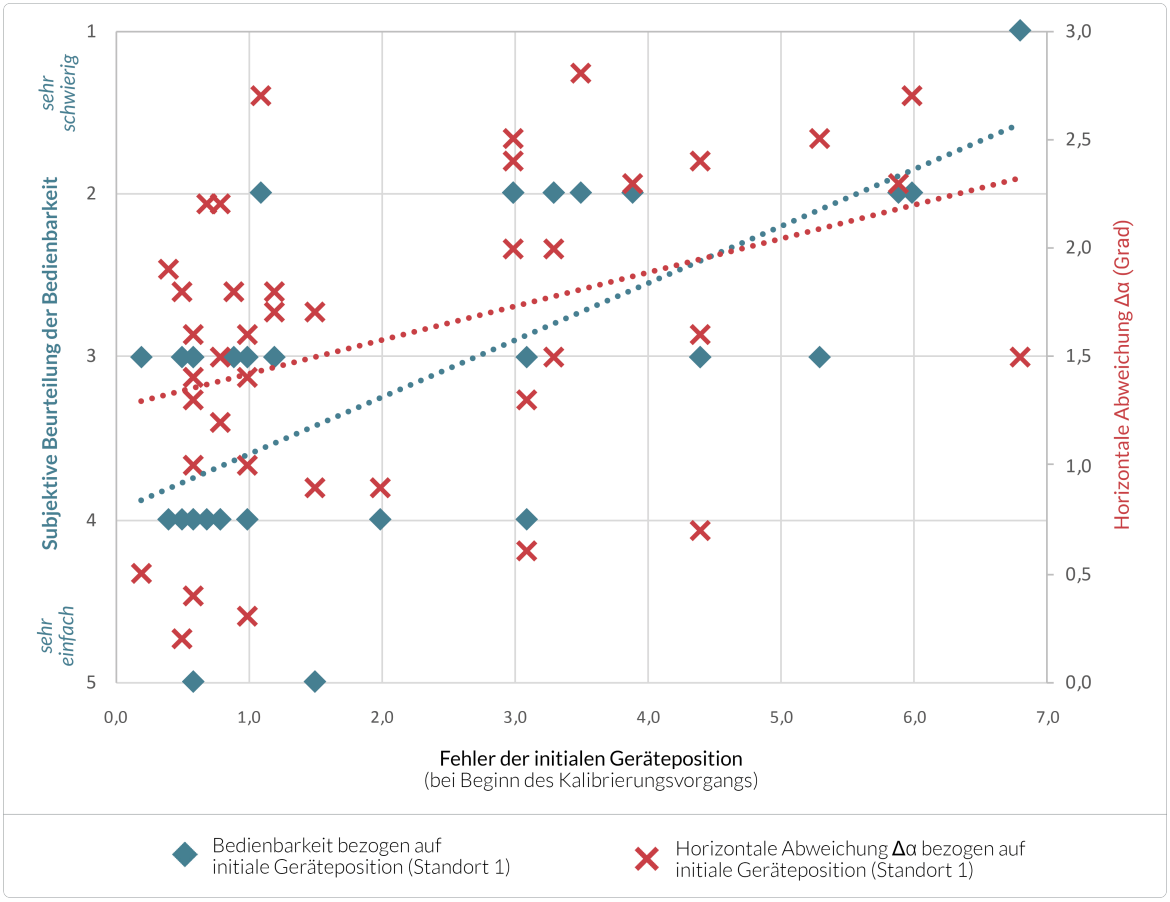


Abbildung 5.14 Bedienbarkeit und Genauigkeit der kalibrierten horizontalen Geräteorientierung bezogen auf die Genauigkeit der initialen Geräteposition an Test-Standort 1. Die Messwerte zeigen, dass mit zunehmender Ungenauigkeit der anfänglichen Positionsschätzung vor dem Start der Kalibrierung die Genauigkeit der kalibrierten Geräteorientierung abnimmt und die Bedienbarkeit der Kalibrierungsmethode als schwieriger bewertet wird.

Richtungen eine Korrektur der Position am ehesten notwendig wäre. Die Nutzung dieser Kalibrierungsgeste sollte jedoch eher erfahrenen Nutzer:innen vorbehalten sein, z.B. in Form eines Expertenmodus.

Es ist in diesem Kontext jedoch auch festzuhalten, dass in bestimmten Anwendungsfällen eine Positionskalibrierung bzw. -präzisierung gar zwingend nicht notwendig ist, um realitätsnahe AR-Visualisierungen bzw. eine gute Bedienbarkeit der Kalibrierungsmethode zu ermöglichen. Sind Referenzobjekte und darzustellende AR-Inhalte sehr weit entfernt, ist die Genauigkeit der GNSS-basierten Positionsschätzung in der Regel sogar ausreichend für korrekte und realistische AR-Projektionen. Dies kann beispielsweise in einem Anwendungsszenario der Fall sein, wenn virtuelle AR-Modelle (z.B. geplante Windenergieanlagen) in größerer Entfernung visualisiert werden sol-

len, sich keine Verdeckungsobjekte in der näheren Umgebung befinden, und auch die Referenzobjekte im virtuellen Modell weit entfernt sind (z.B. Kalibrierung mittels DGM an Test-Standort 4). Etwaige Positionsänderungen haben dann nur sehr geringe bzw. gänzlich vernachlässigbare Auswirkungen auf die AR-Projektionen und können in solchen Fällen sogar ganz entfallen. Ein derartiges Szenario bildet jedoch eher die Ausnahme. Üblicherweise ist eine manuelle Positionsbestimmung auf jeden Fall sinnvoll und notwendig für einen erfolgreichen Einsatz der Registrierungsmethode.

Standortwahl und Umgebungscharakteristik

Die Praxistests zeigten, dass die entwickelte Kalibrierungsmethode an verschiedenen Standorten mit unterschiedlicher Umgebungscharakteristik funktioniert. Voraussetzung ist lediglich, dass eindeutige Referenzobjekte in beliebiger Form in der realen Umgebung sichtbar und im Modell identifizierbar sind. Dabei kann es sich um Vegetation (Bäume, Sträucher), Gebäude oder sonstige markante Geländestrukturen handeln. Wesentlich für eine gute Bedienbarkeit ist dabei, dass die Silhouette einzelner Referenzobjekte am Horizont als auffällig wahrgenommen werden kann (z.B. auffällige Dachkontur eines Gebäudes, einzelne auffällige Baumspitze o.ä.) und die Horizontlinie des Modells nicht vollkommen monoton erscheint (wie z.B. bei entferntem Wald ohne auffällige Strukturen). Selbst sehr kleine, jedoch auffällige Objekte (z.B. kleine Kirchturmspitze hinter einem Wald) sind ausreichend. An den Teststandorten konnten stets eindeutig identifizierbare Referenzobjekte gefunden werden. Es ist also davon auszugehen, dass für den Einsatz der Kalibrierungsmethode verschiedenartige Orte genutzt werden können und im Allgemeinen daher keine wesentlichen Einschränkungen hinsichtlich der konkreten Standortwahl zu beachten sind.

Die Tests machten außerdem deutlich, dass die Kalibrierungsmethode auch an einem Standort genutzt werden kann, deren Höhe nicht mit der Geländehöhe übereinstimmt. Da die Notwendigkeit einer manuellen Höhenanpassung dann zwar den meisten, jedoch nicht allen Anwender:innen unmittelbar während des Kalibrierungsvorgangs bewusst ist, wäre möglicherweise ein zusätzlicher Hinweis bzw. eine Vorab-Abfrage sinnvoll, um über die bestehende Möglichkeit der Höhenanpassung zu informieren bzw. deren Bedarf abzufragen.

Es ist weiterhin festzuhalten, dass ein starkes und häufiges Rotieren des Endgeräts während der Kalibrierung problematisch sein kann. Insbesondere hat dies zur Folge, dass die Nutzer:innen versuchen, das virtuelle Modell in mehreren Blickrichtungen korrekt auszurichten. Ein ideales Angleichen der Horizontsilhouette in vielen Blickrichtungen ist jedoch nur bei optimaler Positionsbestimmung möglich. Bei leichten Abweichungen

der globalen Geräteposition resultiert ein etwaiges Angleichen des Modells in einer Blickrichtung in einer leicht fehlerhaften Modellangleichung in veränderter Blickrichtung. Ein ständiges Hin- und Herbewegen des Gerätes und wiederholtes Justieren des Landschaftsmodells in verschiedenen Richtungen durch die Nutzer:innen ist dann die Folge, ohne jedoch dabei die tatsächliche globale Geräte-Pose zu optimieren. Es ist daher empfehlenswert, sich bei der Kalibrierung möglichst auf eine feste Blickrichtung zu konzentrieren – idealerweise die Blickrichtung, in der auch die anschließend zu visualisierenden AR-Inhalte dargestellt werden sollen.

Leichte Positionsänderungen (Translationen) des Endgerätes während der Kalibrierung stellten hingegen kaum Probleme dar. Diese erzeugten kaum Änderungen in der AR-Projektion der Landschaftsmodelle und konnten zuverlässig durch das lokale VIO-System erfasst werden. Es ist also weniger problematisch, wenn sich die Nutzer:innen während der Kalibrierung einen Schritt zur Seite oder nach vorne bewegen.

Wahl der Geodatenmodellart

Die Evaluation machte deutlich, dass die entwickelte Kalibrierungsmethode mit allen fünf verwendeten Geodatenmodellarten funktionieren kann. Wie eben beschrieben ist die Voraussetzung für ein Gelingen dabei immer, dass die initiale Gerätepositionierung akzeptabel ist und eindeutig identifizierbare und auffällige Referenzobjekte am Standort und im Modell sichtbar sind. Ist eine markante Berg- oder Hügelkette in Sichtweite, ist der Einsatz eines digitalen Geländemodells zur Kalibrierung empfehlenswert. Damit ist eine schnelle, genaue und einfach bedienbare Registrierung möglich. In anderen Fällen wäre ein texturiertes Oberflächenmodell empfehlenswert, um eine Kalibrierung anhand nahegelegener Referenzen (Sträucher, Bäume, Häuser) durchzuführen. Falls auch diese Modellart nicht verfügbar ist oder deren Einsatz mit zu hohen Aufwänden verbunden wäre (z.B. aufgrund der höheren Dateigröße), kann jedoch auch bedenkenlos auf die übrigen Modellarten zurückgegriffen werden, ohne sehr große Einschränkungen hinsichtlich Bedienbarkeit und Genauigkeit in Kauf nehmen zu müssen.

Zu beachten bei der Integration bzw. Wahl der Geodatenmodelle ist außerdem, dass veraltete Modelle möglicherweise Probleme versuchen können, da real existierende Objekte dann im Modell nicht oder in geänderter Form modelliert sein können und in der Folge dann das Angleichen des Modells für Verwirrung sorgen kann. Eine Problematik hinsichtlich saisonaler Veränderung der Vegetation war bei den Tests nicht erkennbar. Dies lag jedoch auch daran, dass nicht nur die Tests im Sommer stattfanden, sondern auch die Aufnahmedaten der Oberflächenmodelle aus den Sommermonaten stammten.

Einweisung in die Bedienung der Methode

Die Testpersonen waren zum Großteil anfangs nicht mit der Bedienung der Kalibrierungsmethode vertraut. Während der Tests konnte jedoch beobachtet werden, dass die Teilnehmer:innen die Kalibrierung mit zunehmender Nutzung schneller und problemloser absolvieren konnten. Ein Gewöhnungs- und Trainingseffekt bei der Bedienung der Anwendung war also mitunter erkennbar.

Dabei sollte auch beachtet werden, dass alle Testpersonen zu Beginn eine ausführliche Einweisung in die Bedienung der Registrierungsmethode erhalten haben, bei der auch die Nutzung der einzelnen Interaktionsgesten im Detail gezeigt und nachvollzogen werden konnte. Auch während der Evaluationsdurchführung konnten etwaige Fragen der Teilnehmenden durch den Evaluationsleiter unmittelbar beantwortet werden. Bei einem Praxiseinsatz der Methode außerhalb des Evaluationssettings wäre eine derart intensive Betreuung und Schulung der Nutzer:innen nicht immer möglich. Es ist unklar, ob die Aufgabenstellung der Kalibrierung und die Interaktionsgesten intuitiv von allen Nutzer:innen auch ohne vorherige Einweisung verstanden werden. Alternative Mittel zur Einweisung in die Bedienung des Kalibrierungsverfahrens wären daher empfehlenswert (z.B. Hilfetexte oder Tutorials innerhalb der App).

Wahl des Endgeräts

Entgegen der ursprünglichen Annahme ist nach Aussagen der Testpersonen für eine gelungene Durchführung des Kalibrierungsvorgangs ein sehr großes Tablet-Display nicht zwingend erforderlich. Auch das Display eines großen Smartphones erscheint ausreichend und sogar vorteilhafter, da ein Smartphone problemlos einhändig vor dem Körper gehalten und mit einer zweiten Hand bedient werden kann. Diese Aussagen decken sich auch mit wissenschaftlichen Untersuchungen, nach denen Tablets üblicherweise mit zwei Händen gehalten werden und die Bedienung des Touchscreens dann nur mit beiden Damen geschieht [31, 41]. Dies führt zwangsläufig dazu, dass die hier entworfenen Interaktionsgesten, die den Einsatz von zwei Fingern einer Hand bzw. den Einsatz des Zeigefingers erfordern, auf Tablets herausfordernder und weniger intuitiv durchzuführen sind. Bei einer Tablet-Nutzung wäre demzufolge die Entwicklung von alternativen Interaktionsgesten, die allein per Daumen bedient werden können, zur Angleichung der virtuellen Modelle überlegenswert.

Sobald das AR-Szenario ein problemloses Halten des Geräts mit zwei Händen zulässt und keine Interaktion per Touchgeste notwendig ist (z.B. beim reinen Betrachten der dargestellten AR-Inhalten), kann sich der große Bildschirm eines Tablets im Vergleich

zum Smartphone dann jedoch durchaus auch im Kontext von GeoAR-Szenarien als vorteilhafter erweisen.

Berücksichtigung von Verdeckungen

Bei erfolgreicher Kalibrierung kann das global ausgerichtete Geodatenmodell dazu eingesetzt werden, Verdeckungen von darzustellenden AR-Inhalten durch reale Objekte zu berücksichtigen und korrekt wiederzugeben. Üblicherweise handelt es sich dabei um Verdeckungen, die durch reale Gebäude, Vegetation (Bäume, Sträucher) oder Gelände (Berge, Hügel) entstehen. Auf diese Weise können AR-Szenarien mit hohem Realitätsgrad realisiert werden, die ansonsten nur durch aufwändige alternative Verfahren (z.B.

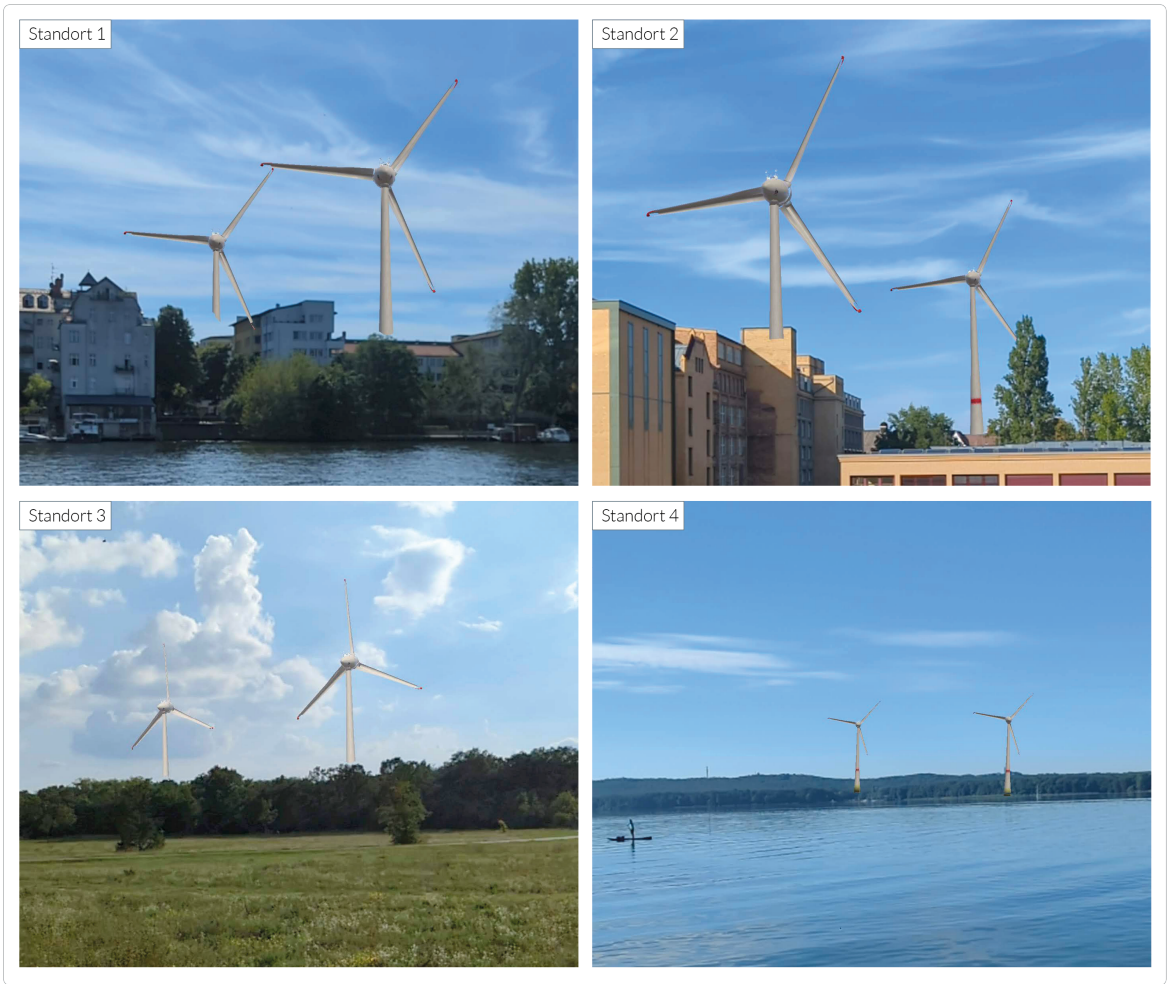


Abbildung 5.15 Beispielhafte AR-Visualisierungen von Windenergieanlagen nach Durchführung der Kalibrierung an den Test-Standorten. Die Berücksichtigung von Verdeckungen der virtuellen Windenergieanlagen durch Gebäude und Vegetation erfolgt je nach Genauigkeit der durchgeführten Kalibrierung mehr oder weniger realitätsnah.

manuelle Segmentierung der Verdeckungsobjekte) möglich wären. Abbildung 5.15 zeigt beispielhaft AR-Visualisierungen von Windenergienanlagen, die als Validierungsmodelle im Rahmen der Evaluation genutzt wurden. Es ist zu erkennen, dass je nach Qualität der durchgeführten Kalibrierung auch Verdeckungen der virtuellen Windenergieanlagen mehr oder weniger realitätsgetreu wiedergegeben werden können. Die Nutzung eines Oberflächenmodells ist dabei in den meisten Fällen notwendig, da ansonsten beispielsweise Verdeckungen durch Vegetation nicht berücksichtigt werden können.

Prüfung der Hypothesen zur Evaluation

Die Evaluationsergebnisse lassen erkennen, dass sich die meisten der zuvor definierten vier Hypothesen hinsichtlich Genauigkeit und Bedienbarkeit der Registrierungsmethode durch die Feldstudie bestätigten. Die Richtigkeit der einzelnen Hypothesen lässt sich abschließend wie folgt zusammenfassen:

H1 *Die Registrierungsmethode ermöglicht eine präzise Bestimmung der Geräte-Pose in unterschiedlichen Umgebungscharakteristiken mit geringem Zeitaufwand.*

Ergebnis: **Teilweise korrekt.** Sowohl die globale Orientierung des mobilen AR-Systems als auch die geografische Position und globale Gerätehöhe können mithilfe des entwickelten Registrierungsverfahrens kalibriert und präzisiert werden. Dies gelingt an verschiedenen Standorten mit unterschiedlicher Umgebungscharakteristik. Der dafür benötigte Zeitwand bleibt akzeptabel und beträgt meist etwa 20-30 Sekunden. Jedoch bleibt festzuhalten, dass eine Kalibrierung mit sehr hoher Genauigkeit (1 Grad Abweichung) mit der Methode insbesondere nur dann möglich ist, wenn bereits eine präzise initiale Positionsbestimmung vorab erfolgte.

H2 *Die Registrierungsmethode ist an allen Standorten intuitiv bedienbar. Die Bedienbarkeit ist jedoch abhängig vom verwendetem Geodatenmodell und der Umgebung.*

Ergebnis: **Teilweise korrekt.** Bei geeigneten Rahmenbedingungen (insb. genauer initialer Positionsbestimmung) ist die Methode an allen Standorten intuitiv bedienbar. Zwar zeigte sich, dass bei bestimmten Geodatenmodellarten (z.B. texturiertem Oberflächenmodell) oder bestimmten Umgebungen (z.B. markante Hügelkette in Sichtweite) die Kalibrierung tendenziell einfacher und nutzerfreundlicher gelingt, eine sehr deutliche Abhängigkeit von verwendeter Modellart und Umgebungscharakteristik konnte jedoch nicht festgestellt werden.

H3 *Mit der geodatenbasierten Registrierungsmethode ist eine Kalibrierung der globalen Position schwieriger realisierbar als eine Kalibrierung der globalen Orientierung.*

Ergebnis: **Korrekt.** Es zeigte sich, dass eine Kalibrierung der globalen Orientierung durch manuelles Verschieben der projizierten 3D-Geodatenmodelle einfach und intuitiv möglich ist, sofern die globale Geräteposition präzise bestimmt wurde, beispielsweise durch eine manuelle kartenbasierte Positionsbestimmung. Die Kalibrierung einer fehlerhaften globalen Position ist mithilfe der geodatenbasierten AR-Registrierungsmethode zwar auch möglich, jedoch herausfordernder, zeitaufwändiger und oftmals nur mit unbefriedigender Genauigkeit zu realisieren.

H4 *Die Bedienung der geodatenbasierten Registrierungsmethode fällt auf einem Tablet mit großem Bildschirm leichter als auf einem Smartphone.*

Ergebnis: **Falsch.** Die Evaluation zeigte zwar, dass die Bedienung der Registrierungsmethode auch auf einem Tablet mit großem Display gut gelingen kann, in der Gesamtbetrachtung wurde jedoch die Nutzung eines Smartphones als einfacher und nutzerfreundlicher bewertet, da dieses leichter einhändig gehalten und bedient werden kann.

Es kann also festgehalten werden, dass auf Basis der Evaluationsergebnisse drei der vier anfangs aufgestellten Hypthesen zur Genauigkeit und Bedienbarkeit der entwickelten Methode als korrekt bzw. größtenteils korrekt bestätigt werden konnten.

5.5.2 Handlungsempfehlungen

Auf Basis der diskutierten Ergebnisse lassen sich konkrete Handlungsempfehlungen für einen idealen Einsatz des Verfahrens in der Praxis ableiten (siehe Abbildung 5.16):

Integration von Hilfetexten und Tutorials Die Bedienung des Kalibrierungsverfahrens wird durchweg als intuitiv wahrgenommen, jedoch ist eine einmalige Einweisung in die Aufgabenstellung der Kalibrierung (also das Angleichen der virtuellen Modelle an die reale Umgebung) und die Nutzung der Interaktionsgesten empfehlenswert. Dies könnte in Form von kurzen Hilfetexten bzw. Bedienungsanweisungen auf dem Bildschirm erfolgen oder in Form eines Tutorials, bei dem alle notwendigen Bedienschritte auf dem Bildschirm nacheinander im Detail erklärt und durch die Nutzer:innen nachvollzogen werden können.

Abbildung 5.16 Handlungsempfehlungen zum Praxiseinsatz der Kalibrierungsmethode

Kartenbasierte Vorab-Positionsbestimmung Eine genaue initiale globale Gerätepositionierung ist von hoher Bedeutung für ein erfolgreichen Gelingen des Verfahrens. Eine manuelle Vorab-Positionsbestimmung per Kartenansicht ist daher sehr empfehlenswert, um eine hohe Genauigkeit und eine gute Bedienbarkeit des Registrierungsverfahrens zu erreichen. Auf eine Kalibrierung der globalen Geräteposition in der AR-Ansicht, die sich als schwieriger und weniger intuitiv herausstellte, kann dann verzichtet werden.

Flexible Standortwahl Das Verfahren ist in unterschiedlichen Umgebungen (naturnah, urban) gut einsetzbar. Voraussetzung ist lediglich, dass identifizierbare Referenzobjekte (Hausdächer, Baumsilhouetten, Hügel o.ä.) in der Umgebung sichtbar sind. Es kann davon ausgegangen werden, dass solche Referenzobjekte an nahezu allen denkbaren Standorten vorhanden sind. Es brauchen daher keine allgemeinen

Einschränkungen hinsichtlich der Standortwahl beachtet werden, so dass das Verfahren an beliebigen Orten flexibel eingesetzt werden kann.

Fokussierung auf ein Blickfeld Ein wiederholtes Hin- und Herbewegen des Endgerätes und ein häufiger Wechsel verschiedener Blickfelder kann problematisch sein. Es sollte daher während der Kalibrierung eine Fokussierung auf ein bestimmtes Blickfeld stattfinden. Entsprechende Hinweise innerhalb der Anwendung bei wiederholtem Verlassen des Blickfeldes wären denkbar, um die Nutzer:innen anzuleiten, sich auf ein bestimmtes Blickfeld zu konzentrieren.

Manuelle Höhenkalibrierung mit Vorab-Hinweis Das Verfahren ist gut geeignet für eine nachträgliche manuelle Kalibrierung der globalen Gerätehöhe, d.h. eine Nutzung an Orten oberhalb (oder auch unterhalb) der Geländehöhe (z.B. am Fenster, auf Balkonen etc.) ist ebenfalls gut möglich. Das Einblenden eines Hinweises oder eine Vorab-Abfrage innerhalb der Anwendung ist jedoch sinnvoll, um auf die zusätzliche Möglichkeit der Höhenanpassung hinzuweisen. Dies kann beispielsweise erfolgen, indem der Nutzer oder die Nutzerin vor Beginn der Kalibrierung selbst in Form einer simplen Abfrage händisch Auskunft gibt, ob die Nutzung auf Geländehöhe stattfindet oder nicht. Die Möglichkeit zur manuellen Höhenkalibrierung per Interaktionsgeste könnte dann nur in letzterem Fall aktiviert werden.

Texturiertes DOM und DGM als bevorzugte Kalibrierungsmodelle Das Verfahren kann mit Oberflächenmodellen, Geländemodellen oder Stadtmodellen genutzt werden. Bei markanten Berg- oder Hügelketten in Sichtweite sollten Geländemodelle (DGM) zur Kalibrierung genutzt werden. In allen anderen Fällen ist die Nutzung eines texturierten Oberflächenmodelles (DOM_RGB) die geeignetste Wahl.

Großes Smartphone als präferiertes mobiles Endgerät Ein Smartphone mit großem Display ist als Endgerät für die entworfene Kalibrierungsmethode empfehlenswert (Portrait-Modus). Die Nutzung eines Tablets (Landscape-Modus) ist ebenfalls möglich. Die Entwicklung alternativer Interaktiongesten zur Modellverschiebung, die dann primär mit Daumen bedient werden können, wäre dann überlegenswert.

DOM zur realitätsnahen Wiedergabe von Verdeckungen Die ausgerichteten Geodatenmodelle können nach erfolgter Kalibrierung für eine realitätsnahe Wiedergabe von Verdeckungen (z.B. durch Gelände, Vegetation oder Gebäude) in einem AR-Szenario genutzt werden. Hierfür ist der Einsatz von Oberflächenmodellen empfehlenswert.

Kapitel 6

Zusammenfassung und Ausblick

Das folgende abschließende Kapitel soll die wesentlichen Ergebnisse der Arbeit zusammenfassen und einen Ausblick auf Potenziale und mögliche Weiterentwicklungen des entwickelten geodatenbasierten Registrierungsverfahrens geben.

6.1 Zusammenfassung der Ergebnisse

Das Hauptziel dieser Arbeit lag in der Entwicklung und Evaluation einer nutzergesteuerten geodatenbasierten Registrierungsmethode zur genauen und robusten Registrierung von mobilen AR-Systemen in Bezug auf ein globales Geo-Referenzsystem. Eine derartige globale Registrierung dient als wesentliche Grundlage für die Realisierung von mobilen Anwendungen zur AR-Visualisierung von georeferenzierten Informationen im Kamerabild von mobilen Endgeräten (Smartphones, Tablets). Auf Basis einer derartigen AR-Registrierungsmethode können somit flexible Vor-Ort-Visualisierungen im Außenbereich erfolgen, beispielsweise für Navigationsaufgaben (Einblendung von Routen oder Wegpunkten) oder um die optischen Auswirkungen geplanter Bauwerke (z.B. Gebäude, Windenergieanlagen etc.) auf das reale Landschaftsbild von verschiedenen Standpunkten aus zu visualisieren.

Die wesentlichen Ergebnisse der Arbeit lassen sich auf Basis der drei zu Beginn der Arbeit aufgestellten Forschungsfragen (siehe Kapitel 1.4) zusammenfassen:

Herausforderungen von geodatenbasierter Augmented Reality

Es konnte gezeigt werden, dass bei der Realisierung von GeoAR-Anwendungen große Herausforderungen zu bewältigen sind. Neben erschwerten Rahmenbedingungen aufgrund des Einsatzes der Technologie im Außenbereich (z.B. wechselnde Wetter- und

Lichtverhältnisse, veränderliche Objekte und Vegetation als Störfaktoren) stellt vor allem eine **präzise globale Registrierung** des mobilen AR-Systems eine große technische Herausforderung dar. Existierende GeoAR-Anwendungen nutzen daher oftmals nur vereinfachte und ungenaue globale Registrierungsverfahren, die primär auf dem Einsatz von integrierten GNSS- und IMU-Sensoren basieren, jedoch nicht für praxistaugliche AR-Visualisierungen mit hohem Realitätsgrad geeignet sind. Automatische bildbasierte AR-Registrierungsmethoden stellen als Alternative zwar einen genaueren Ansatz dar, funktionieren jedoch nur unter bestimmten Rahmenbedingungen (z.B. an vorab festgelegten Standorten) oder benötigen eine hochaufgelöste 3D-Punktwolke der gesamten Umgebung als Referenzmodell, deren Erstellung mit hohem Aufwand verbunden wäre.

Zur Bewältigung der genannten Herausforderungen bieten sich **nutzergesteuerte Registrierungsverfahren** als weitere Methode an. Durch aktive Mithilfe der Nutzer:innen bei der korrekten Bestimmung von Position und Ausrichtung des Systems kann somit mehr Flexibilität und Robustheit erreicht werden. Bei der Analyse aktueller Forschungs- und Entwicklungsarbeiten konnte jedoch gezeigt werden, dass die wenigen existierenden manuellen globalen Registrierungsverfahren deutliche Limitierungen aufweisen und nur in bestimmten Gegenden (z.B. bergige Landschaft) funktionieren oder nur einzelne Freiheitsgrade der Geräte-Pose kalibrieren können. Der Innovationsgrad einer Registrierungsmethode, welche an beliebigen Standorten eingesetzt werden kann und die Pose eines AR-Systems vollumfänglich mit hoher Genauigkeit und mit intuitiver Bedienbarkeit bestimmen kann, konnte dadurch verdeutlicht werden.

Entwicklung eines geodatenbasierten AR-Registrierungsverfahrens

Als wesentliches Ergebnis der Arbeit wurde schließlich ein neuartiges Registrierungsverfahren entwickelt, welches durch aktive Mithilfe der Nutzer:innen die Position und Ausrichtung des AR-Systems in bis zu sechs Freiheitsgraden global registrieren kann. Das Verfahren nutzt dazu digitale 3D-Landschaftsmodelle als Hilfsmittel zur virtuellen Repräsentation der Umgebung und kann flexibel an unterschiedlichsten Standorten genutzt werden. Die Realisierung dieser Methode basiert auf zwei entwickelten Hauptkomponenten:

- Eine **Geodatenverarbeitungskette** konvertiert unterschiedliche Arten von 3D-Geodatenmodellen (texturierte sowie nicht-texturierte 3D-Oberflächen- und Stadtmodelle sowie digitale Geländemodelle) in kleinflächige 3D-Kacheln mit möglichst geringer Dateigröße und ohne merkliche Qualitätsverluste. Eine effiziente Speicherung und AR-basierte Integration innerhalb eines mobilen Clients ist dadurch

möglich. Zu diesem Zweck werden mehrere Umwandlungsschritte realisiert, unter anderem eine Koordinatensystem-Anpassung, eine 3D-Mesh-Generierung, eine Reduktion der räumlichen Modellauflösungen sowie eine Kompression etwaiger Modelltexturen.

- Ein **globales Registrierungs- und Tracking-System** wird zur Laufzeit innerhalb des mobilen AR-Clients ausgeführt. Eine grobe initiale Schätzung der globalen Kamera-Pose erfolgt zunächst mithilfe der GNSS- und IMU-Sensoren sowie durch eine automatische Höhenkalibrierung auf Basis digitaler Geländemodelle. Für den manuelle Prozess zur Kalibrierung der globalen Pose durch Angleichung der Modelle an die reale Umgebung stehen drei verschiedene Interaktionsgesten zur Verfügung: eine Verschiebungsgeste zur Kalibrierung der Kamera-Rotation, eine Zoom-Geste zur Kalibrierung der Kamera-Translation entlang der aktuellen Blickrichtung sowie eine Verschiebungsgeste mit zwei Fingern zur Höhenkalibrierung der Kamera. Im Hintergrund läuft ständig ein lokales Visual-Inertial-Tracking-System, um Kamerabewegungen im lokalen Raum zu detektieren und eine stabile und ruckelfreie AR-Visualisierung zu ermöglichen. Es konnte im Zuge der Arbeit herausgearbeitet und hergeleitet werden, dass der manuelle Kalibrierungsprozess mathematisch als Korrektur der Translation und Rotation der Kamera-Pose in Bezug auf das lokale Tracking-System beschrieben werden kann.

Evaluation des entwickelten Registrierungsverfahrens

Die Genauigkeit und Bedienbarkeit des Registrierungsverfahrens konnte schließlich im Rahmen einer praxisnahen Evaluation untersucht werden. Daraus konnten wesentliche **Erkenntnisse zum Anwendungspotenzial** der Methode gewonnen werden. Die Feldstudie mit mehreren Testpersonen verdeutlichte, dass das Verfahren gut geeignet ist, in Umgebungen mit unterschiedlicher Charakteristik (urban, naturnah) die globale Pose eines mobilen AR-Systems auf einfache und intuitive Art und Weise zu kalibrieren und zu präzisieren. Mit sehr hoher Genauigkeit ist dabei vor allem eine Kalibrierung der globalen Geräteausrichtung möglich (etwa 1 Grad Abweichung). Auch eine nachträgliche manuelle Höhenkorrektur ist gut und intuitiv realisierbar, wenn die Kalibrierung an Standorten oberhalb der Geländehöhe (z.B. am Fenster) durchgeführt wird. Alle eingesetzten Geodatenmodellarten sind dabei für den Registrierungsvorgang geeignet, wobei texturierte Oberflächenmodelle und – in hügeliger Umgebung – digitale Geländemodelle bevorzugt eingesetzt werden sollten. Der für die Kalibrierung

benötigte Zeitwand bleibt in der Regel ebenfalls akzeptabel (etwa 20-30 Sekunden). Außerdem wird bei der Durchführung der manuellen Kalibrierung die Verwendung eines Smartphones, das problemlos mit einer Hand gehalten werden kann, der Nutzung eines großen Tablets vorgezogen. Vor allem zeigte sich durch die Evaluation, dass Genauigkeit und Bedienbarkeit der Methode stark von der Genauigkeit der initialen Positionsschätzung abhängen. Eine solche genaue Positionierung kann manuell mithilfe einer 2D-Kartenansicht schnell und nutzerfreundlich vorab erfolgen, während sich eine nachträgliche AR-basierte Positionskalibrierung als schwieriger und herausfordernder gestaltet. Unter den genannten Bedingungen ist schließlich nicht nur eine schnelle und genaue Kalibrierung der globalen Geräteorientierung, sondern auch eine präzise Kalibrierung der globalen Geräteposition (Abweichung etwa 1 Meter) möglich.

Aus den genannten Erkenntnissen konnten abschließend konkrete **Handlungsempfehlungen für den Praxiseinsatz** des Verfahrens abgeleitet werden. Neben der Empfehlung zur Integration einer kartenbasierten Vorab-Positionsbestimmung als zusätzlicher Baustein des entwickelten Systems und der Empfehlung zur Nutzung von Smartphones als mobiles Endgerät erscheint auch die Integration von zusätzlichen Hilfetexten oder Tutorials sinnvoll, um die durchzuführende Kalibrierungsaufgabe und die Interaktionsgesten anschaulich zu erklären. Auch die Fokussierung auf ein bestimmtes Blickfeld während der Kalibrierung ohne wiederholtes Wechseln der Blickrichtung ist empfehlenswert.

Der Praxiseinsatz verdeutlichte zudem einen weiteren Vorteil der entwickelten Methode: Durch die Integration der 3D-Geodatenmodelle zur Repräsentation der Umgebung steht eine geeignete Möglichkeit zur Verfügung, um in der AR-Ansicht etwaige Verdeckungen darzustellender georeferenzierter Modelle durch reale Objekte (Gebäude, Gelände oder Vegetation) effektiv zu berücksichtigen. Hierfür ist besonders die Integration von digitalen Oberflächenmodellen empfehlenswert, um alle denkbaren Verdeckungen bestmöglich berücksichtigen und somit AR-Szenarien mit hohem Realitätsgrad realisieren zu können.

6.2 Ausblick

Die Grundidee der Arbeit, 3D-Geodatenmodelle innerhalb einer AR-Ansicht zu visualisieren und als Hilfsmittel zur Registrierung mobiler AR-Geräte einzusetzen, stellt ein neuartiges und innovatives Anwendungsszenario dar. Im Folgenden soll das perspektivisch hohe Potenzial von Anwendungen mit 3D-Geodaten-Integration betont sowie etwaige Limitierungen der darauf aufbauenden GeoAR-Registrierungsmethode

aufgezeigt werden. Außerdem werden daraus resultierende Ansätze zur Optimierung des entworfenen Registrierungsverfahrens vorgestellt. Etwaige Potenziale und Möglichkeiten zur Optimierung der entwickelten Methode sind in Abbildung 6.1 zusammengefasst.

Crowdsourcing von 3D-Geodaten

Die Verfügbarkeit von 3D-Geodatenmodellen, die auch zur Kalibrierung von AR-Systemen geeignet sind, stellt momentan noch eine entscheidende Hürde bei der praxistauglichen Entwicklung derartiger Systeme dar. Es ist jedoch erkennbar, dass sich die Aufnahme und Erzeugung von 3D-Geodaten in Zukunft immer einfacher und kostengünstiger gestalten wird [112]. Angetrieben von wachsender Verbreitung von künstlicher Intelligenz, mobiler Sensorik und Cloud-Computing-Technologien entsteht ein großes Potenzial für die einfache und schnelle Aufnahme, Verarbeitung und Visualisierung von 3D-Daten im georäumlichen Kontext [112]. Die Generierung dieser Daten erfolgt dabei zunehmend durch die Masse an Endnutzer:innen (*Crowdsourcing*) mittels kostengünstiger Bildgebungs- und Lidar-Systemen in selbstfahrenden Autos, Mini-Drohnen oder gewöhnlichen Smartphones [2]. Auf diese Weise entstehen zunehmend – auf kostengünstige Art und Weise – digitale und georeferenzierte 3D-Abbilder von realen physischen Umgebungen (*Digital Twins*). Es ist schließlich gut denkbar, dass derartige 3D-Modelle der georäumlichen Umgebung auch als Datengrundlage für geodatenbasierte AR-Registrierungsverfahren genutzt werden können. Notwendig für eine effektive Umsetzung dieser Vision wäre neben der fortschreitenden Entwicklung kostengünstiger präziser 3D-Sensorik (in Smartphones, Drohnen etc.) auch die Verfügbarkeit entsprechender Softwareplattformen zur kollaborative Erzeugung und zum Austausch von generierten 3D-Geodatenmodellen durch die Endnutzer:innen.

Verfügbarkeit von 3D-Geodaten als Open Data

Die im Kontext dieser Arbeit eingesetzten 3D-Geodatenmodelle entstammen aus offiziellen Datenquellen von Landesvermessungsämtern, die jedoch innerhalb von Deutschland nur für bestimmte Regionen kostenfrei zur Verfügung stehen (siehe auch Kapitel 4.2.1). Alternative Ansätze, um nutzergenerierte georeferenzierte 3D-Daten großflächig global verfügbar zu machen, existieren zwar, z.B. in Form einer *Open AR Cloud*, stecken aber noch im Anfangsstadium [108]. Zunehmend Unterstützung findet auch die grundsätzliche Idee, offizielle digitale Geobasisdaten öffentlich und kostenfrei als *Open Data* zur Verfügung zu stellen. So stuft beispielsweise die Open-Data-Strategie der Europäischen Kommission georäumlichen Daten als *high-value datasets* ein, welche kostenfrei der Allgemeinheit in digitaler Form zur Verfügung gestellt werden sollten [42]. Auch

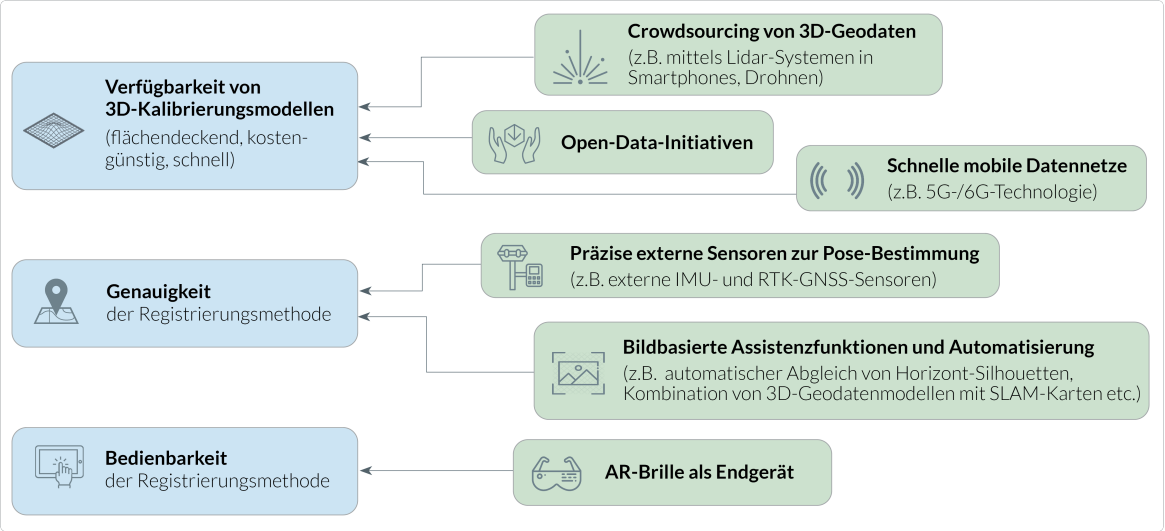


Abbildung 6.1 Ansätze zur Optimierung der entwickelten GeoAR-Registrierungsmethode

immer mehr Bundesländer sind diesem Trend in jüngster Zeit gefolgt und bieten 3D-Geobasisdaten kostenfrei zum Download an, z.B. Niedersachsen im Jahr 2021 [79] und Hessen im Jahr 2022 [55].

Es ist demzufolge davon auszugehen, dass die freie Verfügbarkeit von Geodatenmodellen, die für geodatenbasierte AR-Registrierungsverfahren geeignet sind, perspektivisch weiter steigen wird. Weitere politische Open-Data-Initiativen, klare rechtliche Regelungen hinsichtlich der Open-Data-Nutzung sowie zunehmender gesellschaftlicher Druck wäre hilfreich zur fortschreitenden kostenfreien Bereitstellung von derartigen Geobasisdaten.

Verfügbarkeit schneller mobiler Datennetze

Zur effektiven flächendeckenden Bereitstellung von 3D-Geodatenmodellen sind neben der kostenfreien Verfügbarkeit von 3D-Geodatenmodellen auch entsprechende Technologien notwendig, um die Modelldaten in großem Umfang auf schnelle und einfache Art und Weise in die mobilen AR-Systeme zu integrieren. Hierfür ist insbesondere eine Bereitstellung schneller drahtloser Breitbandzugänge (z.B. 5G-/6G-Technologie) erforderlich. Auch in Deutschland ist eine steigenden Verfügbarkeit schneller drahtloser Datennetze zu beobachten, dennoch sind weitere Anstrengungen notwendig für eine flächendeckende Verfügbarkeit derartiger Datennetze, insbesondere in ländlichen naturräumlichen Umgebungen [76].

Optimierung der Genauigkeit der entwickelten Methode

Die Praxis-Evaluation der entwickelten Methode zeigte Stärken, aber auch einige Beschränkungen und Limitierungen hinsichtlich Genauigkeit und Bedienbarkeit der entwickelten Methode. Neben der genannten Herausforderung der Verfügbarkeit geeigneter 3D-Kalibrierungsmodelle sind daher noch weitere Möglichkeiten denkbar, um die Nutzbarkeit der Registrierungsmethode zu optimieren.

Zur Verbesserung der Genauigkeit der Pose-Bestimmung wäre ein naheliegender Ansatz, die Genauigkeit der initialen GNSS-basierten Positionsschätzung zu verbessern. Dies könnte durch den Einsatz zusätzlicher **externer IMU- und RTK-GNSS-Sensorik** gelingen (siehe Kapitel 3.2.1). Dadurch wäre sogar eine zentimetergenaue globale Positionierung möglich, jedoch nur unter Zuhilfenahme externer Hardware mit zusätzlicher Antenne [124, 134, 140]. Für den Einsatz der geodatenbasierten AR-Registrierungsmethode in speziellen Fach-Anwendungen wäre eine Integration derartiger Zusatzhardware zum Zwecke einer genaueren Initial-Positionierung möglicherweise dennoch überlegenswert, trotz zusätzlicher Kosten und Einschränkungen hinsichtlich der Handhabung des Systems (z.B. höheres Gewicht, Zeitaufwand für RTK-GNSS-basierte Positionsbestimmung). Auch die initiale Schätzung der globalen Orientierung kann mittels hochwertiger externer IMU-Sensoren grundsätzlich verbessert werden, jedoch sind grobe Ungenauigkeit des Magnetometer-Sensors auch dann nicht auszuschließen, insbesondere in städtischer Umgebung [134]. Durch derartige externe Zusatzhardware könnte dann aber zumindest eine manuelle Positionsbestimmung bzw. -kalibrierung des AR-Systems entfallen. Nur die globale Orientierung des mobilen AR-Systems müsste dann durch einfaches Verschieben von 3D-Geodatenmodellen in der AR-Ansicht leicht korrigiert werden. Mit einer solchen zentimergenauer Positionsbestimmung wären somit möglicherweise auch hochpräzise AR-Visualisierungen möglich, beispielsweise für die Realisierung bestimmter AR-Szenarien in der Bauplanung, z.B. zur AR-Darstellung im Kontext des *Building Information Modeling* (*BIM*).

Für weitere Optimierungen und zur Unterstützung des manuellen Kalibrierungsprozesses wäre außerdem der Einsatz von **bildbasierten Assistenzfunktionen und Automatisierungen** denkbar. Durch Verfahren der Bildverarbeitung könnte dazu beispielsweise die **Horizontsilhouette** in den Kamerabildern robust erkannt und extrahiert werden (z.B. auf Basis von Kantenerkennung und grober manueller Vorselektion von Himmelsbereichen) und dann automatisiert mit der Kontur der virtuellen Horizontsilhouette der integrierten 3D-Landschaftsmodelle abgeglichen werden. Über eine Optimierungsfunktion könnte dann eine ideale globale Position bzw. Rotation berechnet und den Anwender:innen als Schätzung der initialen AR-Projektion vorge-

schlagen werden. Auf Basis dieser Schätzung wären dann unter Umständen nur noch sehr kleine manuelle Korrekturgesten notwendig. Die Machbarkeit derartiger automatischer bildbasierter Registrierungsverfahren mittels Abgleichen von realen Horizontlinien an virtuellen Gelände- oder Gebäudemodellen wurde in mehreren Forschungsarbeiten bereits gezeigt [32, 18, 14].

Neben dem automatisierten Abgleich von Horizontkonturen könnten auch weitere markante Konturen oder Oberflächen des sichtbaren Kamerabildes genutzt werden, um diese mit georeferenzierten Modelldaten automatisiert abzugleichen und somit den manuellen Kalibrierungsprozess zu unterstützen. Liu et al. nutzten beispielsweise die Oberflächen (Fassaden) von **georeferenzierten Gebäudemodellen**, um lokale 3D-SLAM-Karten global zu registrieren und damit die globale Pose des AR-Systems automatisiert zu kalibrieren [88]. Auch eine Optimierung mithilfe von bildbasierten **semantischen Registrierungsansätzen** (*semantic localization*) wäre denkbar. Mit Methoden der Künstlichen Intelligenz kann versucht werden, ein semantisches Verständnis der Kamerabilder zu entwickeln (z.B. Segementieren von Gebäuden, Bäumen, Straßen etc.), um dann durch Abgleichen dieser Bilder mit einer semantischen georeferenzierten Datenbank die globale Pose der Kamera zu schätzen [100, 128]. Möglicherweise könnte eine derartige Referenz-Datenbank auch durch eine automatische semantische Segmentierung von 3D-Geodatenmodellen erzeugt werden [114, 69]. Derartige bildbasierte Verfahren sind zwar an bestimmte Rahmenbedingungen gebunden (z.B. Vorhandensein von Gebäuden) und könnten nicht an beliebigen Standorten vollautomatisiert eingesetzt werden, jedoch könnten sie zumindest in bestimmten Situationen als Assistenzfunktionen einen manuellen Registrierungsvorgang unterstützen.

Denkbar wäre auch eine **vollumfängliche bildbasierte Registrierung** innerhalb von hochaufgelösten 3D-Geodatenmodellen, die z.B. vorab durch Mini-Drohnen mit Lidar-Sensoren generiert werden. Zhang et al. entwickelten beispielsweise ein derartiges System, bei dem aus Drohneaufnahmen 3D-Punktwolken erzeugt werden, die dann wiederum für eine automatische bildbasierte Registrierung mobiler AR-Systeme aus Bodenperspektive genutzt werden können [151, 152]. Es ist jedoch weitere intensive Forschungsarbeit zur Entwicklung von solchen Ansätzen notwendig, um aus üblichen texturierten, per ALS erzeugten 3D-Geodatenmodellen (wie z.B. dem hier eingesetzten texturierten digitalen Oberflächenmodell) 3D-Punktwolken so zu erzeugen, dass diese auch für automatische SLAM-basierte Registrierungsverfahren aus Bodenperspektive effektiv genutzt werden können. In jedem Fall ist in weiteren Optimierungsschritten ein Übergang von manuellen geodatenbasierten Registrierungsverfahren hin zu robusten, vollautomatischen Registrierungsmethoden denkbar und erstrebenswert.

Optimierung der Bedienbarkeit der entwickelten Methode

Derartige Automatisierungen könnten nicht nur die Genauigkeit, sondern auch die Bedienbarkeit der Registrierung optimieren. Die Bedienbarkeit der entwickelten manuellen Methode könnte zudem noch intuitiver gestaltet werden, indem nicht übliche Smartphones oder Tablets, sondern **AR-Brillen als Endgeräte** genutzt werden. Ein Halten des AR-Systems mit ein bzw. zwei Händen wäre dann nicht mehr notwendig, da die virtuellen Inhalt dann nicht mehr auf ein tragbares Touchscreen-Gerät, sondern auf die Brillengläser und damit direkt ins Sichtfeld der Nutzer:innen projiziiert werden. Zur Bedienung müssten dann alternative Interaktionsgesten implementiert werden, um die Angleichung der virtuellen 3D-Modelle an die reale Landschaft per Gestensteuerung in der Luft (*mid-air interaction*) zu realisieren. Mit einer breiten Verfügbarkeit geeigneter AR-Brillen wird jedoch erst in mehreren Jahren gerechnet [142].

Ergebnisse der Arbeit als Beitrag zur Etablierung der GeoAR-Technologie

Im Kontext künftiger AR-Forschung und -Entwicklung sollte es ein wesentliches Ziel sein, praxistaugliche AR-Szenarien nicht nur innerhalb begrenzter Räume, sondern großflächig und an beliebigen Orten, innerhalb von Gebäuden oder im Freien, zu ermöglichen. Diese Forschungsarbeit zeigte in diesem Kontext das große Potenzial von Outdoor AR-Anwendungen zur realitätsnahen Vor-Ort-Visualisierung von Geodaten und die Notwendigkeit zur Bereitstellung geeigneter Registrierungsmethoden mobiler AR-Systeme im Außenbereich. Eine solche flexible und nutzerfreundliche Registrierungsmethode wurde schließlich durch Integration von 3D-Geodatenmodellen erarbeitet und präsentiert sowie deren Anwendungspotenzial für einen Praxiseinsatz evaluiert.

Zu Beginn der Arbeit wurden die Limitierungen und Herausforderungen von vollautomatischen bildbasierten Registrierungsverfahren als Ausgangspunkt und Motivation zur Entwicklungen innovativer manueller Verfahren genannt. Dennoch ist es absehbar und auch erstrebenswert, Möglichkeiten zu entwickeln, um die Limitierungen derartiger automatisierter Verfahren zunehmend abzubauen. Denn ein grundsätzliches Ziel im Bereich der AR-Registrierung sollte es durchaus sein, möglichst einfache, schnelle, robuste – und somit eben auch automatisierte – Verfahren bereitzustellen. Das hier entwickelte manuelle Verfahren wäre dann möglicherweise obsolet. Diese Entwicklung kann daher mitunter eher als temporäre Übergangslösung betrachtet werden für den Zeitraum, in dem automatisierte Verfahren noch nicht mit hoher Robustheit und flächendeckend verfügbar sind. Eine parallele Weiterentwicklung der manuellen Registrierungsmethode in die Richtung hin zu automatisierten Verfahren, z.B. durch

bildbasierte Assistenzfunktionen (Abgleich von Horizontkonturen oder Gebäudefassaden, Generierung von 3D-SLAM-Karten aus Geodatenmodellen etc.), erscheint daher in jedem Fall erstrebenswert. Die hier erarbeitete Methode kann daher ebenfalls als ein Schritt zur zunehmenden Optimierung und Automatisierung der GeoAR-Registrierung verstanden werden und ist daher in jedem Fall auch als Beitrag für eine fortschreitende Etablierung und Verbesserung der geodatenbasierten AR-Technologie zu sehen.

Literaturverzeichnis

[1] 3D RealityMaps (2022). 3D RealityMaps - Kartografie, 3D-Vermessung und 3D-Visualisierung. https://www.realitymaps.de/kartenloesungen-apps/ (besucht am 27.09.2022).

[2] Abdullah, Q. (2021). Top Geospatial Trends to Watch in 2021. *Photogrammetric Engineering & Remote Sensing*, 87(3):149–151.

[3] Android Developers (2022a). CameraCharacteristics - LENS INTRINSIC CALIBRATION. https://developer.android.com/reference/android/hardware/camera2/CameraCharacteristics#LENS_INTRINSIC_CALIBRATION (besucht am 27.09.2022).

[4] Android Developers (2022b). Motion sensors. https://developer.android.com/guide/topics/sensors/sensors_motion (besucht am 27.09.2022).

[5] Apple Developer (2022). ARKit - Augmented Reality. https://developer.apple.com/augmented-reality/arkit/ (besucht am 27.09.2022).

[6] Arth, C., Pirchheim, C., Ventura, J., Schmalstieg, D., und Lepetit, V. (2015). Instant outdoor localization and slam initialization from 2.5 d maps. *IEEE transactions on visualization and computer graphics*, 21(11):1309–1318.

[7] Arth, C., Reitmayr, G., und Schmalstieg, D. (2012). Full 6DOF pose estimation from geo-located images. In *Asian Conference on Computer Vision*, Seiten 705–717. Springer.

[8] Arth, C. und Schmalstieg, D. (2011). Challenges of large-scale augmented reality on smartphones. *Graz University of Technology, Graz*, Seiten 1–4.

[9] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., und MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE computer graphics and applications*, 21(6):34–47.

[10] Azuma, R., Hoff, B., Neely, H., und Sarfaty, R. (1999). A motion-stabilized outdoor augmented reality system. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, Seiten 252–259. IEEE.

[11] Azuma, R. T. (1997). A survey of augmented reality. *Presence: teleoperators & virtual environments*, 6(4):355–385.

[12] Azuma, R. T., Billinghurst, M., und Klinker, G. (2011). Special Section on Mobile Augmented Reality. *Computers & Graphics*, 35(4):vii–viii.

[13] Baatz, G., Saurer, O., Köser, K., und Pollefeys, M. (2012). Large scale visual geo-localization of images in mountainous terrain. In *European conference on computer vision*, Seiten 517–530. Springer.

[14] Baboud, L., Čadík, M., Eisemann, E., und Seidel, H.-P. (2011). Automatic photo-to-terrain alignment for the annotation of mountain pictures. In *CVPR 2011*, Seiten 41–48. IEEE.

[15] Bayerisches Staatsministerium für Wirtschaft, Energie und Technologie (2017). Energie-Atlas Bayern - 3D-Analyse - Simulation von Windenergieanlagen in der Landschaft. https://www.bestellen.bayern.de/shoplink/07000233.htm (besucht am 30.04.2022).

[16] Bedingfield, W. (2022). Snapchat has turned London into an augmented reality experiment. https://www.wired.co.uk/article/snapchat-launches-local-lenses (besucht am 27.09.2022).

[17] Behnisch, M. und Kolbe, T. (2020). Solarpotenzialanalyse und Web-Visualisierung von 3D-Geomassendaten am Beispiel der Stadtregion München. https://katalog.rundertischgis.de/datahub_resource/geomassendaten-teilprojekttum (besucht am 30.04.2022).

[18] Behringer, R. (1999). Registration for outdoor augmented reality applications using computer vision techniques and hybrid sensors. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, Seiten 244–251. IEEE.

[19] Berlin Partner für Wirtschaft und Technologie GmbH (2022). Berlin 3D - Downloadportal. https://www.businesslocationcenter.de/downloadportal/ (besucht am 27.09.2022).

[20] Berliner Senatsverwaltung für Stadtentwicklung, Bauen und Wohnen (2022). Interaktive Abfrage der Geländehöhe - ATKIS DGM. https://www.stadtentwicklung.berlin.de/geoinformation/landesvermessung/atkis/de/dgm.shtml (besucht am 27.09.2022).

[21] Billinghurst, M., Clark, A., Lee, G., u a. (2015). A survey of augmented reality. *Foundations and Trends® in Human–Computer Interaction*, 8(2-3):73–272.

[22] Blut, C., Blut, T., und Blankenbach, J. (2019). CityGML goes mobile: application of large 3D CityGML models on smartphones. *International Journal of Digital Earth*, 12(1):25–42.

[23] Brejcha, J. und Čadík, M. (2017). State-of-the-art in visual geo-localization. *Pattern Analysis and Applications*, 20(3):613–637.

[24] Bruck, B. (2010). Layar-App lässt die Berliner Mauer wieder auferstehen. https://hauptstadtstudio.com/2010/05/07/layar-app-lasst-die-berliner-mauer-wieder-auferstehen/ (besucht am 27.09.2022).

[25] Burkard, S., Fuchs-Kittowski, F., Abecker, A., Haß, E., Heise, F., Miller, R., Runte, K., und Hosenfeld, F. (2021). Grundbegriffe, Anwendungen und Nutzungspotenziale von geodatenbasierter mobiler Augmented Reality im Umweltbereich. In *Umweltinformationssysteme–Wie verändert die Digitalisierung unsere Gesellschaft?*, Seiten 243–259. Springer.

[26] Burkard, S., Fuchs-Kittowski, F., Deharde, M., Poppel, M., und Schreiber, S. (2022). Eine mobile Augmented Reality-Anwendung für die Darstellung von geplanten Windenergieanlagen. In *Umweltinformationssysteme-Wie trägt die Digitalisierung zur Nachhaltigkeit bei?*, Seiten 21–41. Springer.

[27] Cesium GS Inc. (2022a). Cesium - The Platform for 3D Geospatial https://cesium.com/. (besucht am 27.09.2022).

[28] Cesium GS Inc. (2022b). OBJ2GLTF. https://github.com/CesiumGS/obj2gltf (besucht am 27.09.2022).

[29] Chakravorty, D. (2021). OBJ File Format – All You Need to Know. https://all3dp.com/1/obj-file-format-3d-printing-cad/ (besucht am 30.04.2022).

[30] Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., und Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. In *Eurographics Italian Chapter Conference*. The Eurographics Association.

[31] Clark, J. und Frost, B. (2015). *Designing for Touch*. A Book Apart : brief books for people who make websites. Book Apart, A.

[32] Coors, V., Huch, T., und Kretschmer, U. (2000). Matching buildings: Pose estimation in an urban environment. In *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*, Seiten 89–92. IEEE.

[33] Côté, S., Trudel, P., Desbiens, M., Giguère, M., und Snyder, R. (2013). Live mobile panoramic high accuracy augmented reality for engineering and construction. *Proceedings of the Construction Applications of Virtual Reality (CONVR), London, England*, Seiten 1–10.

[34] Davison, A. J., Reid, I. D., Molton, N. D., und Stasse, O. (2007). MonoSLAM: Real-time single camera SLAM. *IEEE transactions on pattern analysis and machine intelligence*, 29(6):1052–1067.

[35] Docker Inc. (2022). What is a container? https://www.docker.com/resources/what-container (besucht am 27.09.2022).

[36] Dörner, R., Broll, W., Grimm, P., und Jung, B. (2013). Virtual und Augmented Reality. *Grundlagen und Methoden der Virtuellen und Augmentierten Realität. Berlin und Heidelberg: Springer*.

[37] Dünser, A. und Billinghurst, M. (2011). Evaluating augmented reality systems. In *Handbook of augmented reality*, Seiten 289–307. Springer.

[38] Elias, B. (2012). 3D-Geobasisinformationen. *ZfV-Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, (zfv 5/2012).

[39] Engel, J., Koltun, V., und Cremers, D. (2017). Direct sparse odometry. *IEEE transactions on pattern analysis and machine intelligence*, 40(3):611–625.

[40] Engel, J., Schöps, T., und Cremers, D. (2014). LSD-SLAM: Large-Scale Direct Monocular SLAM. In Fleet, D., Pajdla, T., Schiele, B., und Tuytelaars, T., Herausgeber, *Computer Vision – ECCV 2014*, Seiten 834–849, Cham. Springer International Publishing.

[41] Esoldo, Christine (2022). The Fundamental Guide to Mobile Usability. https://www.toptal.com/designers/mobile-ui/fundamental-guide-mobile-usability (besucht am 28.08.2022).

[42] European Commission (2021). European legislation on open data. https://digital-strategy.ec.europa.eu/en/policies/legislation-open-data (besucht am 22.09.2022).

[43] Fedorov, R., Frajberg, D., und Fraternali, P. (2016). A framework for outdoor mobile augmented reality and its application to mountain peak detection. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, Seiten 281–301. Springer.

[44] Fiala, M. (2004). Artag, an improved marker system based on artoolkit. *National Research Council Canada, Publication Number: NRC*, 47419:2004.

[45] Fukuda, T., Zhang, T., und Yabuki, N. (2014). Improvement of registration accuracy of a handheld augmented reality system for urban landscape simulation. *Frontiers of Architectural Research*, 3(4):386–397.

[46] Garland, M. und Heckbert, P. S. (1995). Fast polygonal approximation of terrains and height fields.

[47] Garland, M. und Heckbert, P. S. (1998). Simplifying surfaces with color and texture using quadric error metrics. In *Proceedings Visualization'98 (Cat. No. 98CB36276)*, Seiten 263–269. IEEE.

[48] Gazcón, N. F., Nagel, J. M. T., Bjerg, E. A., und Castro, S. M. (2018). Fieldwork in Geosciences assisted by ARGeo: A mobile Augmented Reality system. *Computers & Geosciences*, 121:30–38.

[49] Google Developers (2022a). ARCore – Augmented-Reality-Erlebnisse schaffen, die die digitale und physische Welt nahtlos zusammenführen. https://unity.com/de/unity/features/arfoundation/ (besucht am 27.09.2022).

[50] Google Developers (2022b). Fused Location Provider API. https://developers.google.com/location-context/fused-location-provider (besucht am 27.09.2022).

[51] Gośliński, J., Nowicki, M., und Skrzypczyński, P. (2015). Performance comparison of EKF-based algorithms for orientation estimation on Android platform. *IEEE Sensors Journal*, 15(7):3781–3792.

[52] Gröger, G., Benner, J., Dörschlag, D., Drees, R., Gruber, U., Leinemann, K., und Löwner, M.-O. (2005). Das interoperable 3D-Stadtmodell der SIG 3D. *Zeitschrift für Vermessungswesen*, 130(6):343–353.

[53] Harbeck, R. (2005). Zur Situation der amtlichen topographischen Kartographie in Deutschland. *KN-Journal of Cartography and Geographic Information*, 55(6):297–307.

[54] HERE Europe B.V. (2018). TIN Terrain. https://github.com/heremaps/tin-terrain (besucht am 27.09.2022).

[55] Hessisches Ministerium für Wirtschaft, Energie, Verkehr und Wohnen (2022). Land gibt amtliche Geodaten frei. https://www.hessen.de/Presse/Land-gibt-amtliche-Geodaten-frei (besucht am 22.09.2022).

[56] Hoang, T. N. und Thomas, B. H. (2009). In-situ refinement techniques for outdoor geo-referenced models using mobile AR. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, Seiten 193–194. IEEE.

[57] Höllerer, T. und Feiner, S. (2004). Mobile augmented reality. *Telegeoinformatics: Location-based computing and services*, 21.

[58] Huang, W., Sun, M., und Li, S. (2016). A 3D GIS-based interactive registration mechanism for outdoor augmented reality system. *Expert Systems with Applications*, 55:48–58.

[59] Hugo Ledoux (2022). cjio (CityJSON/io). https://github.com/cityjson/cjio (besucht am 27.09.2022).

[60] Hugues, O., Fuchs, P., und Nannipieri, O. (2011). New augmented reality taxonomy: Technologies and features of augmented environment. *Handbook of augmented reality*, Seiten 47–63.

[61] Internet Engineering Task Force (IETF) (2016). The GeoJSON Format. https://datatracker.ietf.org/doc/html/rfc7946 (besucht am 27.09.2022).

[62] Jäger, E. und Heipke, C. (2014). Geotopographie und Photogrammetrie. *Das deutsche Vermessungs-und Geoinformationswesen. Wichmann*, Seiten 375–431.

[63] Kalkofen, D., Zollman, S., Schall, G., Reitmayr, G., und Schmalstieg, D. (2009). Adaptive visualization in outdoor AR displays. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality, ISMAR*, Band 9, Seite 37.

[64] Karami, E., Prasad, S., und Shehata, M. (2017). Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images. *arXiv preprint arXiv:1710.02726*.

[65] Karlekar, J., Zhou, S. Z., Lu, W., Loh, Z. C., Nakayama, Y., und Hii, D. (2010). Positioning, tracking and mapping for outdoor augmentation. In *2010 ieee international symposium on mixed and augmented reality*, Seiten 175–184. IEEE.

[66] Kasperi, J., Edwardsson, M. P., und Romero, M. (2017). Occlusion in outdoor Augmented Reality using geospatial building data. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, Seiten 1–10.

[67] Katzur, L. und Wedel, H. (2010). Landesflug Digitales Geländemodell. *Vermessung Brandenburg. Potsdam*, 15(1):53–62.

[68] Kauling, S., Taeger, S., Sondershaus, F., und Lensing, P. (2021). Frühzeitige Visualisierung in Planungsprozessen–ein Dilemma und mögliche Lösungsansätze durch flexible partizipative Ansätze am Beispiel der Windenergie. In *Handbuch Methoden Visueller Kommunikation in der Räumlichen Planung*, Seiten 79–99. Springer.

[69] Kazimi, B., Thiemann, F., und Sester, M. (2019). Object instance segmentation in digital terrain models. In *International Conference on Computer Analysis of Images and Patterns*, Seiten 488–495. Springer.

[70] KHRONOS (2021a). GLB File Format Specification. https://www.khronos.org/registry/glTF/specs/2.0/glTF-2.0.html#glb-file-format-specification (besucht am 27.09.2022).

[71] KHRONOS (2021b). glTF Specification. https://www.khronos.org/gltf (besucht am 27.09.2022).

[72] Kilimann, J.-E., Heitkamp, D., und Lensing, P. (2019). An Augmented Reality Application for Mobile Visualization of GIS-Referenced Landscape Planning Projects. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*, Seiten 1–5.

[73] Kim, K., Billinghurst, M., Bruder, G., Duh, H. B.-L., und Welch, G. F. (2018). Revisiting trends in augmented reality research: A review of the 2nd decade of ISMAR (2008–2017). *IEEE transactions on visualization and computer graphics*, 24(11):2947–2962.

[74] Kitt, B. M., Rehder, J., Chambers, A. D., Schonbein, M., Lategahn, H., und Singh, S. (2011). Monocular visual odometry using a planar road model to solve scale ambiguity.

[75] Klein, G. und Murray, D. (2009). Parallel tracking and mapping on a camera phone. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, Seiten 83–86. IEEE.

[76] Kretschmer, C. (2020). Schnelles Mobilfunknetz: 5G-Ausbau in großen Schritten. https://www.tagesschau.de/wirtschaft/technologie/5g-ausbau-in-deutschland-101.html (besucht am 22.09.2022).

[77] Kreuziger, U. (2017). *Entwicklung georeferenzierender Augmented Reality Systeme auf Mobilgeräten mit Mikrosensoren.* Dissertation, Otto-Friedrich-Universität Bamberg, Fakultät Geistes-und Kulturwissenschaften.

[78] Kreuziger, U. und Hehl, K. (2012). Entwicklung einer AR-Applikation für die Planung und Bodenordnung. *Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 1(137):37–45.

[79] Landesamt für Geoinformation und Landesvermessung Niedersachsen (2021). Ab sofort kostenfrei in Niedersachsen: Amtliche Orthophotos (Luftbilder), 3D-Gebäudemodelle und ein Landschaftsmodell. https://www.lgln.niedersachsen.de/startseite/wir_uber_uns_amp_organisation/presse_amp_broschuren/ab-sofort-kostenfrei-in-niedersachsen-amtliche-orthophotos-luftbilder-3d-gebaudemodelle-und-ein-landschaftsmodell-202028.html (besucht am 22.09.2022).

[80] Lanham, M. (2018). *Learn ARCore-Fundamentals of Google ARCore: Learn to build augmented reality apps for Android, Unity, and the web with Google ARCore 1.0*. Packt Publishing Ltd.

[81] Ledoux, H., Arroyo Ohori, K., Kumar, K., Dukai, B., Labetski, A., und Vitalis, S. (2019). CityJSON: A compact and easy-to-use encoding of the CityGML data model. *Open Geospatial Data, Software and Standards*, 4(1):1–12.

[82] Lee, G. A., Dünser, A., Kim, S., und Billinghurst, M. (2012). CityViewAR: A mobile outdoor AR application for city visualization. In *2012 IEEE International Symposium on Mixed and Augmented Reality-Arts, Media, and Humanities (ISMAR-AMH)*, Seiten 57–64. IEEE.

[83] Lee, G.-h., Choi, P.-h., Nam, J.-h., Han, H.-s., Lee, S.-h., und Kwon, S.-c. (2019). A Study on the Performance Comparison of 3D File Formats on the Web. *International journal of advanced smart convergence*, 8(1):65–74.

[84] Lee, S., Suh, J., und Park, H.-D. (2015). BoreholeAR: A mobile tablet application for effective borehole database visualization using an augmented reality technology. *Computers & Geosciences*, 76:41–49.

[85] LGB (Landesvermessung und Geobasisinformation Brandenburg) (2022). Geobroker – Der Internetshop der LGB. https://geobroker.geobasis-bb.de/ (besucht am 27.09.2022).

[86] Li, P., Qin, T., Hu, B., Zhu, F., und Shen, S. (2017). Monocular visual-inertial state estimation for mobile augmented reality. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Seiten 11–21. IEEE.

[87] Li, Y., Snavely, N., Huttenlocher, D., und Fua, P. (2012). Worldwide pose estimation using 3d point clouds. In *European conference on computer vision*, Seiten 15–29. Springer.

[88] Liu, R., Zhang, J., Chen, S., und Arth, C. (2019). Towards SLAM-based outdoor localization using poor GPS and 2.5 D building models. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Seiten 1–7. IEEE.

[89] Lothe, P., Bourgeois, S., Royer, E., Dhome, M., und Naudet, S. (2010). Real-Time Vehicle Global Localisation with a Single Camera in Dense Urban Areas: Exploitation of Coarse 3D City Models. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seiten 863–870.

[90] Löwner, M.-O., Casper, E., Becker, T., Benner, J., Gröger, G., Gruber, U., Häfele, K.-H., Kaden, R., und Schlüter, S. (2013). CityGML 2.0–ein Internationaler Standard für 3D-Stadtmodelle, Teil 2: CityGML in der Praxis. *Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, 2(2013):131–143.

[91] Luebke, D. und Humphreys, G. (2007). How gpus work. *Computer*, 40(2):96–100.

[92] Mapbox (2022). Maps SDK for Unity: 3D worlds, AR, POIs. https://www.mapbox.com/unity (besucht am 27.09.2022).

[93] Marx, C., Donaubauer, A., Fiutak, G., und Kolbe, T. H. (2017). Digitales Landschaftsmodell in 3D. *Geoinformationssysteme 2017–Beiträge zur 4. Münchner GI-Runde*.

[94] Mei, C., Sibley, G., Cummins, M., Newman, P. M., und Reid, I. (2009). A Constant-Time Efficient Stereo SLAM System. In *BMVC*, Seiten 1–11.

[95] Michel, T., Genevès, P., und Layaïda, N. (2018). A Method to Quantitatively Evaluate Geo Augmented Reality Applications. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Seiten 9–14. IEEE.

[96] Middelberg, S., Sattler, T., Untzelmann, O., und Kobbelt, L. (2014). Scalable 6-DOF Localization on Mobile Devices. In *Proceedings of the 13th European Conference on Computer Vision (ECCV'14)*.

[97] Milgram, P., Takemura, H., Utsumi, A., und Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, Band 2351, Seiten 282–292. Spie.

[98] Mirko, R. (2020). A new sense of direction with Live View. https://blog.google/products/maps/new-sense-direction-live-view/ (besucht am 27.09.2022).

[99] Mirowski, P., Ho, T. K., Yi, S., und MacDonald, M. (2013). SignalSLAM: Simultaneous localization and mapping with mixed WiFi, Bluetooth, LTE and magnetic signals. In *International Conference on Indoor Positioning and Indoor Navigation*, Seiten 1–10. IEEE.

[100] Mousavian, A., Košecká, J., und Lien, J.-M. (2015). Semantically guided location recognition for outdoors scenes. In *2015 IEEE international conference on robotics and automation (ICRA)*, Seiten 4882–4889. IEEE.

[101] Mozilla (2022). Mozilla JPEG Encoder Project. https://github.com/mozilla/mozjpeg (besucht am 27.09.2022).

[102] Mur-Artal, R., Montiel, J. M. M., und Tardos, J. D. (2015). ORB-SLAM: a versatile and accurate monocular SLAM system. *IEEE transactions on robotics*, 31(5):1147–1163.

[103] Mur-Artal, R. und Tardós, J. D. (2017). Visual-inertial monocular SLAM with map reuse. *IEEE Robotics and Automation Letters*, 2(2):796–803.

[104] Müller, P. (2022). Lidar-Sensor im iPhone: Was das ist und wofür. https://www.macwelt.de/news/Lidar-Sensor-im-iPhone-12-Pro-Was-das-ist-und-wofuer-10899937.html (besucht am 27.09.2022).

[105] Newcombe, R. A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A. J., Kohi, P., Shotton, J., Hodges, S., und Fitzgibbon, A. (2011a). Kinectfusion: Real-time dense surface mapping and tracking. In *2011 10th IEEE international symposium on mixed and augmented reality*, Seiten 127–136. IEEE.

[106] Newcombe, R. A., Lovegrove, S. J., und Davison, A. J. (2011b). DTAM: Dense tracking and mapping in real-time. In *2011 international conference on computer vision*, Seiten 2320–2327. IEEE.

[107] Niantic (2022). Lightship ARDK – Niantic Lightship. https://lightship.dev/products/ardk/ (besucht am 27.09.2022).

[108] Open AR Cloud Association (2022). Open AR Cloud - Building a better reality together. https://www.openarcloud.org/ (besucht am 22.09.2022).

[109] Open Geospatial Consortium (2019). OGC 3D Tiles Specification. http://docs.opengeospatial.org/cs/18-053r2/18-053r2.html (besucht am 27.09.2022).

[110] Paavilainen, J., Korhonen, H., Alha, K., Stenros, J., Koskinen, E., und Mayra, F. (2017). The Pokémon GO experience: A location-based augmented reality mobile game goes mainstream. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, Seiten 2493–2498.

[111] Paul Ridden (2013). IKEA catalog uses augmented reality to give a virtual preview of furniture in a room. https://newatlas.com/ikea-augmented-reality-catalog-app/28703/ (besucht am 27.09.2022).

[112] Pauli, F. (2022). The Rise of 3D Geospatial Data. https://www.geospatialworld.net/prime/prime-opinion/the-rise-of-3d-geospatial-data/ (besucht am 22.09.2022).

[113] Petovello, M. und Dabove, P. (2014). What are the actual performances of GNSS positioning using smartphone technology? *InsideGNSS*, 9:34–37.

[114] Politz, F. und Sester, M. (2018). Exploring ALS and DIM data for semantic segmentation using CNNs. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences; 42-1*, 42(1):347–354.

[115] Poppel, M. (2022). Geodaten-Service für GeoAR-Anwendungen. Bachelorarbeit, Hochschule für Technik und Wirtschaft Berlin.

[116] Postert, P., Berger, M., und Bill, R. (2021). Utilizing CityGML for AR-Labeling and Occlusion in Urban Spaces. In *Advances and New Trends in Environmental Informatics*, Seiten 209–223. Springer.

[117] Reitmayr, G. und Drummond, T. W. (2006). Going out: robust model-based tracking for outdoor augmented reality. In *2006 IEEE/ACM international symposium on mixed and augmented reality*, Seiten 109–118. IEEE.

[118] Robinet, F., Arnaud, R., Parisi, T., und Cozzi, P. (2014). gltf: Designing an open-standard runtime asset format. *GPU Pro*, 5:375–392.

[119] Roesner, F., Kohno, T., und Molnar, D. (2014). Security and privacy for augmented reality systems. *Communications of the ACM*, 57(4):88–96.

[120] Sapiezynski, P., Stopczynski, A., Gatej, R., und Lehmann, S. (2015). Tracking human mobility using wifi signals. *PloS one*, 10(7):e0130824.

[121] Sarlin, P.-E., Cadena, C., Siegwart, R., und Dymczyk, M. (2019). From coarse to fine: Robust hierarchical localization at large scale. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seiten 12716–12725.

[122] Sattler, T., Maddern, W., Toft, C., Torii, A., Hammarstrand, L., Stenborg, E., Safari, D., Okutomi, M., Pollefeys, M., Sivic, J., u a. (2018). Benchmarking 6dof outdoor visual localization in changing conditions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Seiten 8601–8610.

[123] Schall, G., Wagner, D., Reitmayr, G., Taichmann, E., Wieser, M., Schmalstieg, D., und Hofmann-Wellenhof, B. (2009). Global pose estimation using multi-sensor fusion for outdoor augmented reality. In *2009 8th ieee international symposium on mixed and augmented reality*, Seiten 153–162. IEEE.

[124] Schall, G., Zollmann, S., und Reitmayr, G. (2013). Smart Vidente: advances in mobile augmented reality for interactive visualization of underground infrastructure. *Personal and ubiquitous computing*, 17(7):1533–1549.

[125] Schilling, A., Bolling, J., und Nagel, C. (2016). Using glTF for streaming CityGML 3D city models. In *Proceedings of the 21st International Conference on Web3D Technology*, Seiten 109–116.

[126] Schmalstieg, D. und Hollerer, T. (2016). *Augmented reality: principles and practice*. Addison-Wesley Professional.

[127] Schmidt, F. (2002). *Hochgenaue digitale Geländemodelle: Untersuchungen zur Erstellung, Analyse und Anwendung in der Landwirtschaft*. Dissertation, Universität Rostock.

[128] Schönberger, J. L., Pollefeys, M., Geiger, A., und Sattler, T. (2018). Semantic visual localization. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Seiten 6896–6906.

[129] Senatsverwaltung für Wirtschaft, Energie und Betriebe (2022). Offene Daten Berlin. https://daten.berlin.de/ (besucht am 27.09.2022).

[130] Siccity (2022). GLTFUtility: Simple GLTF importer for Unity. https://github.com/Siccity/GLTFUtility (besucht am 27.09.2022).

[131] Skyview (2022). Skyview Lite - Apps on Google Play. https://play.google.com/store/apps/details?id=com.t11.skyviewfree (besucht am 27.09.2022).

[132] Soldati, F. (2020). PeakFinder AR. http://www.peakfinder.org/mobile (besucht am 30.04.2022).

[133] Stoter, J., Vallet, B., Lithen, T., Pla, M., Wozniak, P., Kellenberger, T., Streilein, A., Ilves, R., und Ledoux, H. (2016). State-of-the-art of 3D national mapping in 2016. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 41:653.

[134] Stranner, M., Arth, C., Schmalstieg, D., und Fleck, P. (2019). A high-precision localization device for outdoor augmented Reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Seiten 37–41. IEEE.

[135] Sweeney, C., Flynn, J., Nuernberger, B., Turk, M., und Höllerer, T. (2015). Efficient computation of absolute pose for gravity-aware augmented reality. In *2015 IEEE International Symposium on Mixed and Augmented Reality*, Seiten 19–24. IEEE.

[136] Sweeney, C., Höllerer, T., und Turk, M. (2013). Improved outdoor augmented reality through "Globalization". In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Seiten 1–4. IEEE.

[137] Szeliski, R. (2010). *Computer vision: algorithms and applications*. Springer Science & Business Media.

[138] Szot, T., Specht, C., Specht, M., und Dabrowski, P. S. (2019). Comparative analysis of positioning accuracy of Samsung Galaxy smartphones in stationary measurements. *PloS one*, 14(4):e0215562.

[139] Taketomi, T., Uchiyama, H., und Ikeda, S. (2017). Visual SLAM algorithms: a survey from 2010 to 2016. *IPSJ Transactions on Computer Vision and Applications*, 9(1):1–11.

[140] Trimble Inc. (2021). Trimble Sitevision - High accuracy augmented reality system. https://sitevision.trimble.com/ (besucht am 27.09.2022).

[141] Turner, E. (2022). Use the ARCore Depth API for immersive augmented reality experiencesr. https://codelabs.developers.google.com/codelabs/arcore-depth#0 (besucht am 27.09.2022).

[142] Ubrani, J. und Shirer, M. (2022). IDC Tracker Sees a Long Road Ahead for Augmented Reality Headsets. https://www.idc.com/getdoc.jsp?containerId=prUS49692622 (besucht am 27.09.2022).

[143] Unity Technologies (2022a). AR Foundation-Framework von Unity – Plattformübergreifende Augmented-Reality-Entwicklungssoftware. https://unity.com/de/unity/features/arfoundation/ (besucht am 27.09.2022).

[144] Unity Technologies (2022b). Unity Real-Time Development Platform. https://unity.com/ (besucht am 27.09.2022).

[145] Veas, E., Grasset, R., Ferencik, I., Grünewald, T., und Schmalstieg, D. (2013). Mobile augmented reality for environmental monitoring. *Personal and ubiquitous computing*, 17(7):1515–1531.

[146] Ventura, J., Arth, C., Reitmayr, G., und Schmalstieg, D. (2014). Global localization from monocular slam on a mobile phone. *IEEE transactions on visualization and computer graphics*, 20(4):531–539.

[147] Warmerdam, Frank and Rouault, Evan (2022). GDAL - Geospatial Data Abstraction Library. https://gdal.org/index.html (besucht am 27.09.2022).

[148] Wikitude (2022a). Location-based Augmented Reality. https://www.wikitude.com/geo-augmented-reality/ (besucht am 27.09.2022).

[149] Wikitude (2022b). Real Estate agency is listing properties with AR/VR app. https://www.wikitude.com/showcase/real-estate-agency-is-listing-properties-with-ar-app/ (besucht am 27.09.2022).

[150] Yousif, K., Bab-Hadiashar, A., und Hoseinnezhad, R. (2015). An overview to visual odometry and visual SLAM: Applications to mobile robotics. *Intelligent Industrial Systems*, 1(4):289–311.

[151] Zhang, W., Shao, J., Jin, S., Luo, L., Ge, J., Peng, X., und Zhou, G. (2021). Automated marker-free registration of multisource forest point clouds using a coarse-to-global adjustment strategy. *Forests*, 12(3):269.

[152] Zhang, Y., Yue, P., Zhang, G., Guan, T., Lv, M., und Zhong, D. (2019). Augmented reality mapping of rock mass discontinuities and rockfall susceptibility based on unmanned aerial vehicle photogrammetry. *Remote Sensing*, 11(11):1311.

[153] Zollmann, S., Hoppe, C., Kluckner, S., Poglitsch, C., Bischof, H., und Reitmayr, G. (2014). Augmented reality for construction site monitoring and documentation. *Proceedings of the IEEE*, 102(2):137–154.

Abbildungsverzeichnis

1.1 Grundidee der nutzergesteuerten Registrierungsmethode 4

2.1 Kreislauf der AR-Visualisierung 14
2.2 Lokale und globale Registrierung von mobilen AR-Systemen 17
2.3 Mobile Sensoren für AR-Tracking in Smartphones und Tablets 19
2.4 Bestimmung der Kamera-Pose mittels bildbasierter Verfahren 21
2.5 Herausforderungen von GeoAR-Szenarien 24
2.6 Beispielhafte geodatenbasierte AR-Anwendungen 27
2.7 Pinhole-Kameramodell für perspektivische 3D-2D-Projektion 29

3.1 Übersicht über Registrierungsverfahren für mobile AR-Anwendungen . 36
3.2 Registrierung mittels 2D-Referenzmarker 37
3.3 Registrierung mittels Visual-SLAM-Technologie 39
3.4 Registrierung mittels Visual-Inertial-Odometry 41
3.5 Globale Registrierung mittels GNSS- und IMU-Sensoren 42
3.6 Automatische bildbasierte globale Registrierung mit Google Live View . 44
3.7 Bildbasierte globale Registrierungsverfahren auf Basis von Geodaten . . 45
3.8 Existierende nutzergesteuerte globale Registrierungsverfahren 47

4.1 System-Architektur der geodatenbasierten AR-Registrierungsmethode . 53
4.2 Sequenzdiagramm des globalen Registrierungs- und Tracking-Systems . 55
4.3 Mögliche GIS-Anwendungsszenarien mit integrierten 3D-Geodatenmodellen 57
4.4 Auszüge aus Quelldateien der 3D-Geodatenmodelle 60
4.5 Spezifikationen der eingesetzten 3D-Geodatenmodelle 61
4.6 Verarbeitungsschritte der Geodatenverarbeitungskette 63
4.7 Lokale Geo-Koordinatensystem-Anpassung 64
4.8 Verarbeitungsschritte für texturlose Geodatenmodelle 65
4.9 Mesh-Vereinfachung und Optimierung der Modell-Texturen 68
4.10 Meta-Daten einer 3D-Kachel (GeoTile) im JSON-Format 69
4.11 Client-Server-Architektur zur Speicherung und Abfrage der 3D-Kacheln 72

4.12 Sequenzdiagramm zur clientseitigen Abfrage von Kalibrierungsdaten . . 73
4.13 Koordinatensysteme der entwickelten Registrierungsmethode 75
4.14 Interaktionsgesten zur Kalibrierung von Rotation und Translation . . . 79

5.1 Anforderungen an die mobilen Evaluations-Anwendung 90
5.2 Benutzerführung (User Flow) der mobilen Evaluations-Anwendung . . . 93
5.3 Grafische Benutzeroberfläche der Evaluations-Anwendung 95
5.4 Implementierung von Geodatenverarbeitungskette und AR-App 99
5.5 Manuelle Bestimmung des Orientierungsfehlers per Bildvergleich 103
5.6 Fragebogen zur Evaluation der Bedienbarkeit der Kalibrierung 105
5.7 Test-Sessions zur Evaluation der Kalibrierungsmethode 107
5.8 Ablauf der Evaluation an einem Test-Standort 108
5.9 Übersicht über alle vier genutzten Test-Standorte 109
5.10 Bildschirmaufnahmen während der Kalibrierung 112
5.11 Gemessene Abweichung der Geräterotation 113
5.12 Gemessener Fehler der globalen Geräteposition und Gerätehöhe 116
5.13 Zeitaufwand und Bedienbarkeit der Kalibrierungsmethode 119
5.14 Bedienbarkeit und Genauigkeit bezogen auf initiale Geräteposition . . . 124
5.15 AR-Visualisierung von Windenergieanlagen nach der Kalibrierung . . . 128
5.16 Handlungsempfehlungen zum Praxiseinsatz der Kalibrierungsmethode . 131

6.1 Ansätze zur Optimierung der entwickelten Registrierungsmethode . . . 138

Hinweis zu Quellenangaben bei Abbildungen: Sofern nicht anderweitig in der Abbildungsbeschreibung vermerkt, handelt es sich bei allen Abbildungen um eigene Darstellungen, die vollständig selbst erstellt wurden.

Abkürzungsverzeichnis

ALS Airborne Laserscanning
API Application Programming Interface
AR Augmented Reality
ATKIS Amtlich Topographisch-Kartographisches Informationssystem
BIM Building Information Modeling
CMS Content Management System
DGM Digitales Geländemodell
DOF Degrees of Freedom
DOM Digitales Oberflächenmodell
GeoAR Geodatenbasierte Augmented Reality
GIS Geo-Informationsystem
GNSS Global Navigation Satellite System
GPS Global Positioning System
GUI Graphical User Interface
IMU Inertial Measurement Unit
JSON JavaScript Object Notation
Lidar Light Detection and Ranging
LoD Level of Detail
RTK Real Time Kinematic
SLAM Simultaneous Localization and Mapping
TIN Triangulated Irregular Network
UTM Universal Transverse Mercator
VIO Visual Inertial Odometry
VR Virtual Reality

Eidesstattliche Erklärung

Hiermit erkläre ich durch eigenhändige Unterschrift, die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet zu haben. Die aus den Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Dissertation ist in dieser Form noch keiner anderen Prüfungsbehörde vorgelegt worden.

Ort, Datum	Unterschrift des Doktoranden (Simon Burkard)

Lebenslauf

Simon Burkard
geboren am 15. September 1986 in Coburg

s.burkard@htw-berlin.de 0049 175 521 9768

Simon Burkard ist seit 2016 als **Wissenschaftlicher Mitarbeiter an der HTW Berlin** tätig. In seiner Forschungstätigkeit befasst er sich mit Themen der **Outdoor Augmented Reality**, des **Mobile Sensing** sowie der **3D-Visualisierung von Geodaten** und entwickelt dabei neuartige Lokalisierungs- und Visualisierungsmethoden auf Basis von Geodaten. Sein Masterstudium der Technischen Informatik schloss er mit Auszeichnung im Jahr 2015 an der Technischen Universität Berlin ab.

Ausbildung und Akademischer Werdegang

Seit 02/2016	**Wissenschaftlicher Mitarbeiter, HTW Berlin** Entwicklung innovativer bildbasierter Messverfahren und neuartiger mobiler Visualisierungsmöglichkeiten für Geo- und Umweltdaten mit Hilfe mobiler Augmented Reality im Rahmen von Drittmittel-Forschungsprojekten
04/2010 - 09/2015	**M.Sc. Technische Informatik, Technische Universität Berlin** Ingenieurstechnischer Studiengang mit Inhalten aus der Elektrotechnik und Informatik; Spezialisierung in Bildverarbeitung und Augmented Reality
12/2009 - 04/2014	**Studentischer Mitarbeiter, Telekom Innovation Laboratories, Berlin** Softwareentwicklungen in den Bereichen Bildanalyse, Augmented Reality, Sensor Fusion und Vision-Based Navigation für mobile Endgeräte
04/2006 - 09/2009	**B.Sc. Internationale Medieninformatik, HTW Berlin** Interdisziplinäres Studium im Bereich Visual Computing & Webtechnologie
09/1997 - 06/2006	**Gymnasium Casimirianum, Coburg** Abitur mit Abschlussprüfungen in Mathematik, Wirtschaft und Physik

Forschungsprojekte

Seit 08/2021	**AR4WIND - Unterstützung von Bürgerbeteiligungsprozessen durch AR-Visualisierung geplanter Windenergieanlagen** Forschungsschwerpunkte: Forschung und Entwicklung im Bereich der GeoAR-Technologie (nutzerfreundliche und robuste Lokalisierung und Visualisierung)
12/2018 - 07/2021	**mARGo - Augmented Reality für Geoobjekte: Neue Technologien und Nutzungsparadigmen für die mobile Geodatenverarbeitung** Forschungsschwerpunkte: Erforschung von neuartigen Lokalisierungsverfahren für realitätsnahe Outdoor Augmented Reality
02/2016 - 11/2018	**VGI4HWM - Entwicklung eines Hochwasserprognosesystems mit Mobile Sensing** Forschungsschwerpunkt: Erforschung und Entwicklung von bildbasierten Messverfahren zur Erfassung des Wasserstandes mit mobilen Endgeräten